

Using naturalistic driving data to identify driving style based on longitudinal driving operation conditions

Nengchao Lyu and Yugang Wang

Intelligent Transportation Systems Research Center, Wuhan University of Technology, Wuhan, China

Chaozhong Wu

Intelligent Transportation Systems Research Center, Wuhan University of Technology, Wuhan, China and National Engineering Research Center for Water Transport Safety, Wuhan University of Technology, Wuhan, China, and

Lingfeng Peng and Alieu Freddie Thomas

Intelligent Transportation Systems Research Center, Wuhan University of Technology, Wuhan, China

Abstract

Purpose – An individual's driving style significantly affects overall traffic safety. However, driving style is difficult to identify due to temporal and spatial differences and scene heterogeneity of driving behavior data. As such, the study of real-time driving-style identification methods is of great significance for formulating personalized driving strategies, improving traffic safety and reducing fuel consumption. This study aims to establish a driving style recognition framework based on longitudinal driving operation conditions (DOCs) using a machine learning model and natural driving data collected by a vehicle equipped with an advanced driving assistance system (ADAS).

Design/methodology/approach – Specifically, a driving style recognition framework based on longitudinal DOCs was established. To train the model, a real-world driving experiment was conducted. First, the driving styles of 44 drivers were preliminarily identified through natural driving data and video data; drivers were categorized through a subjective evaluation as conservative, moderate or aggressive. Then, based on the ADAS driving data, a criterion for extracting longitudinal DOCs was developed. Third, taking the ADAS data from 47 Kms of the two test expressways as the research object, six DOCs were calibrated and the characteristic data sets of the different DOCs were extracted and constructed. Finally, four machine learning classification (MLC) models were used to classify and predict driving style based on the natural driving data.

Findings – The results showed that six longitudinal DOCs were calibrated according to the proposed calibration criterion. Cautious drivers undertook the largest proportion of the free cruise condition (FCC), while aggressive drivers primarily undertook the FCC, following steady condition and relative approximation condition. Compared with cautious and moderate drivers, aggressive drivers adopted a smaller time headway (THW) and distance headway (DHW). THW, time-to-collision (TTC) and DHW showed highly significant differences in driving style identification, while longitudinal acceleration (LA) showed no significant difference in driving style identification. Speed and TTC showed no significant difference between moderate and aggressive drivers. In consideration of the cross-validation results and model prediction results, the overall hierarchical prediction performance ranking of the four studied machine learning models under the current sample data set was extreme gradient boosting > multi-layer perceptron > logistic regression > support vector machine.

Originality/value – The contribution of this research is to propose a criterion and solution for using longitudinal driving behavior data to label longitudinal DOCs and rapidly identify driving styles based on those DOCs and MLC models. This study provides a reference for real-time online driving style identification in vehicles equipped with onboard data acquisition equipment, such as ADAS.

Keywords Machine learning, Advanced driver assistant systems, Driver behaviors and assistance, Sensor data processing

Paper type Research paper

1. Introduction

Driving style can be defined as an individual's habitual manner of driving (Elander *et al.*, 1993; Lajunen and Özkan, 2011; Sagberg *et al.*, 2015) (i.e. a person's preference of velocity distribution), which is formed over time as that person accumulates driving experience (Suzdaleva and Nagy., 2018).

The current issue and full text archive of this journal is available on Emerald Insight at: <https://www.emerald.com/insight/2399-9802.htm>



Journal of Intelligent and Connected Vehicles
5/1 (2022) 17–35
Emerald Publishing Limited [ISSN 2399-9802]
[DOI 10.1108/JICV-07-2021-0008]

© Nengchao Lyu, Yugang Wang, Chaozhong Wu, Lingfeng Peng and Alieu Freddie Thomas. Published in *Journal of Intelligent and Connected Vehicles*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence maybe seen at <http://creativecommons.org/licences/by/4.0/legalcode>

This research was funded by the National Nature Science Foundation of China (No.52072290), Hubei Province Science Fund for Distinguished Young Scholars (No.2020CFA081) and the Fundamental Research Funds for the Central Universities (No.191044003, No. 2020-YB-028).

Received 5 July 2021

Revised 2 October 2021

Accepted 5 November 2021

Studies also explicitly describe the importance of acceleration behavior as a key indicator of driving style because individuals have different preferences for speed (Müller *et al.*, 2013; Reiser, 2008). To differentiate between driving skill and driving style (Elander *et al.*, 1993; Taubman-Ben-Ari *et al.*, 2004), “skill” is defined as the driver’s ability to maintain control of the vehicle and adapt to complex traffic conditions, and driving skill is expected to improve with practice or training. On the other hand, “style” is defined as the manner in which a driver chooses to drive or habitually drives (i.e. his/her choice of driving speed and headway).

A number of studies have shown that driving style has a significant impact on traffic safety (Evans, 1996), vehicle dynamics control (Plöchl *et al.*, 2007) and the economic and ecological efficiency of driving (Mensing *et al.*, 2014). However, driving style information cannot be directly measured nor detected. Existing studies have categorized driving behavior into driving maneuvers (e.g. following, hard braking, lane changing, etc.) (Bellem *et al.*, 2016). These studies estimate driving style in terms of the durations or frequencies of individual maneuver states. However, driving style is easily affected by and fluctuates with the road traffic environment. Additionally, relatively static and singular driving data does not fully reflect the true driving style. On the other hand, one of the main factors affecting the identification of driving style is the real-time ability and effectiveness of data acquisition. Therefore, how to effectively use driving data to comprehensively and quantitatively analyze driving style has become a new field to be further explored (Qi *et al.*, 2019).

In recent years, advanced driving assistance systems (ADASs) have significantly progressed, opening novel horizons in reducing traffic accidents (Rezaei *et al.*, 2021). Specifically, with the rapid development of in-vehicle information systems and collision warning systems, a large amount of natural driving data can be acquired through these types of ADASs (Bao *et al.*, 2020; Orlovska *et al.*, 2020). In response to the great need of driving style identification for traffic safety and fuel economy, naturalistic data collection is becoming ever more feasible as the penetration rate of ADASs increases in vehicles and on roadways around the world.

Therefore, to explore the influence of different driving behavior data on driving style identification and realize the rapid and efficient detection of driving style, this study obtained a large amount of naturalistic driving data through an ADAS-equipped vehicle and proposes a solution framework for rapid detection of driving style based on the driver’s longitudinal driving operation conditions (DOCs). The proposed framework calibrates the driver’s DOCs through naturalistic driving data and rapidly detects driving style through a machine learning model according to the driving behavior parameter characteristics of different DOCs. To achieve the main goal of this research, 44 subjects participated in naturalistic driving experiments and data from the driver characteristics, vehicle motion attitude and micro driving operation was collected. The framework for rapid identification of driving styles proposed in this research may be applied in intelligent connected and vehicle-road cooperative scenarios, providing a reference for real-time and efficient identification of driving style to help drivers make real-time driving decisions.

2. Literature review

In recent years, to discover and present driving style information in a scientific method, many models have been developed that assess driving style from different aspects. Since its publication, the multidimensional driving style inventory (MDSI) (Taubman-Ben-Ari *et al.*, 2004) has been the subject of research around the world. It defines an individual’s driving style as a driving-specific factor that can contribute to both crashes and traffic violations directly and in terms of more general socio-demographic and personal factors. The MDSI can increase driver awareness of his/her own and others’ driving styles and be used to identify baseline driving styles prior to the implementation of road safety interventions as well as inform post-intervention assessments (Taubman-Ben-Ari *et al.*, 2016). To determine whether the MDSI is consistent with actual driving behavior, Van Huysduynen *et al.* (2018) conducted a simulation experiment with 88 participants. The objective data retrieved from the simulator was compared with the scores obtained from questionnaire data. The analysis showed that there is a moderate correlation between self-reported driving style and driving behavior in the simulator. This suggests that MDSI can be used as a diagnostic tool to identify typical driving behaviors of individuals in driving simulators. Ishibashi *et al.* (2007) developed a driving style questionnaire (DSQ) to extract key indicators from self-reports and calibrate different driving styles. However, the DSQ focuses more on preferences for driving behavior, which is limited by sample characteristics and structural validity (van Huysduynen *et al.*, 2018). In other words, the DSQ cannot fully describe an objective condition.

There have been many studies that classify driving style based on actual vehicle operating parameters, such as naturalistic driving and field operational tests (FOTs). For instance, Toledo *et al.* (2008) developed pattern recognition algorithms to identify more than 20 maneuvers (such as lane change and sudden braking) using naturalistic driving data on different roads; this information was collected by onboard data loggers. On this basis, drivers were divided into three categories combined with the weighted maneuvering frequency. The results showed that this method effectively predicts driving style. Wang *et al.* (2015) extracted emergency braking maneuver features from naturalistic driving data. On this basis, a classification regression tree model was established to estimate driving style, and drivers were divided into three risk groups according to nine rules. Xu *et al.* (2015) used naturalistic driving data from American highways and adopted a neural network (NN) model to divide driver styles into three types. In a simulated scenario, Baer *et al.* (2011) rated five driving styles: aggressive, anxious, economical, sensitive and calm.

Judging from the literature described above, it can be observed that driving style classification methods and standards are not uniform. That said, previous studies have found that in naturalistic driving, drivers generically categorized as high-risk drive faster, exhibit shorter time headways (THWs), brake harder and change lanes more frequently than low-risk drivers (Sagberg *et al.*, 2015; Xiong *et al.*, 2012). It was also found from field operation tests that low-risk drivers engage in fewer risky maneuvers (Simons-Morton *et al.*, 2015; Kusano *et al.*, 2015).

While the aforementioned studies did identify differences between driving styles, they did not establish evaluation models

to estimate driving style through different driving maneuvers. In contrast, Guo and Fang (2013) classified drivers into three risk groups by a K-means clustering method according to the maneuvers detected from naturalistic driving data on different roads in the USA; the authors established a logistic model to predict driving style, which showed that the frequency of emergency braking events was a valid indicator of high-risk drivers. Li et al. (2017) proposed a new method to identify driving style according to the transition patterns between maneuvering states. Driving behavior in highway traffic was divided into 12 maneuvering states. A conditional likelihood maximization method was used to extract typical maneuverability transfer patterns, which represented driving styles from 144 probabilities and the selected features were classified by a random forest algorithm. The results showed that the transitions concerning five maneuver states – free driving, approach, near following, constrained left lane changes and constrained right lane changes – can reliably classify driving style. Suzdaleva and Nagy (2018) proposed an online driving style detection model based on both a normal component and classification component mixed recursive Bayesian estimation. Seven driving styles associated with fuel economy were identified using an online estimation algorithm. That algorithm can also be used to model and predict fuel consumption, speed, throttle pedal position and gear selection. Lu et al. (2021) tried to understand the influence of different driving styles (such as cautious, normal and aggressive) on key variables (such as speed) in traffic flow theory and revealed the influence on network efficiency. The characteristics of different driving styles were extracted from high-dimensional data clustering classes and transformed into different vehicle-following models, which were simulated in a SUMO traffic simulator.

The key to the modeling and analysis of driving style is the extraction of driving maneuver features. Driving maneuvers are mainly divided into longitudinal or lateral. Longitudinal maneuvers include free driving, approaching, following, opening and emergency braking. Longitudinal maneuvers are classified according to the value of the THW, longitudinal acceleration (LA) and the perception of changes in the outward size of the vehicle ahead (Toledo et al., 2007). More specifically, the THW and LA are commonly used to describe the following maneuvers (Kondoh et al., 2008). Further, when rapid deceleration is not occurring, a 3-s THW or less is considered to be car-following (Kusano et al., 2015; Transportation Research Board, The Highway Capacity Manual, 2010). In other words, If the THW of the front and rear vehicles exceeds 3.0 s, it is considered a free drive operation.

This study focuses on longitudinal driving behavior and simplifies the impact of lateral driving behavior. The scope of this study was based on an urban expressway with high traffic flow and speed and the influence of acceleration and deceleration during the process of vehicle following was considered. According to the literature summary and the understanding and analysis of naturalistic driving data, this study took a 6.0-s THW as one of the criteria to indicate car-following. The longitudinal driving data was extracted from naturalistic driving data to classify different driving conditions, and different machine learning models were selected to construct driving style classification models; the accuracy of the various models was then compared to find the best fit.

3. Data collection

3.1 Test equipment and test route

To obtain real and reliable driving data, in this study, an automatic GAC Trumpchi passenger car equipped with FOT data acquisition equipment, as shown in Figure 1, was used to perform FOTs on various types of roads in Wuhan, China.

The multi-functional road test vehicle platform is shown in Figure 1 and the data types and parameter descriptions collected by each experimental equipment are shown in Table 1.

The installation of all instruments and equipment did not hinder normal driving, such that the driver could maintain a naturalistic driving state. The sampling frequency of the in-vehicle devices was 20~100 Hz and the sampling interval of all devices was set to 0.1 s. The naturalistic driving data was obtained in real time through the onboard laptop and the driving video data was continuously stored in the memory card.

As shown in Figure 2, the experimental route consisted of four sections. Detailed information for each section is provided in Table 2.

As can be observed from Table 1, Section 2 was a highway with dispersed traffic volume. During the FOT drives, the traffic flow on this section was low and the traffic density was sparse, such that the experimental vehicle was in a free-driving state for a long time. As can also be observed, Section 4 was an arterial with congested traffic volume. During the FOT drives, this section of the road had a high traffic flow and density, such that the experimental vehicle was in a car-following state for a long time. Therefore, in both Sections 2 and 4, the motion posture of the experimental vehicle was relatively stable; drivers did not make any significant operations that would make driving style identifiable.

On the contrary, Sections 1 and 3 were both expressways with moderate traffic volume, and the road parameters were similar. During the FOT drives, the traffic flow was moderate and traffic density was balanced, such that the experimental vehicle made a variety of motion postures, and the driver's operating characteristics were significantly different, rendering driving style easily identifiable. Therefore, 47 Kms of Sections 1 and 3 were selected as the expressway test bed from which to observe the naturalistic driving data.

3.2 Participants

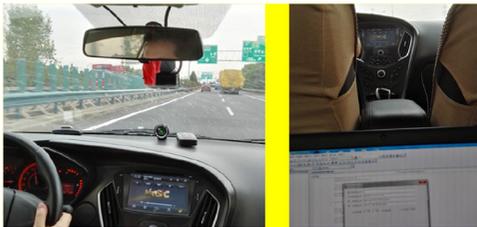
This study mainly focused on model and data analysis. The experiment was outdoor naturalistic driving, the experimental

Figure 1 Multi-functional road test vehicle platform



Table 1 Original data collected by naturalistic driving experimental platform

Data acquisition equipment	Data type	Parameter
OBD-II	Vehicle operation and kinematics data	Speed, accelerator pedal opening, braking pressure, steering wheel angle, steering wheel angular speed
Mobileye M630 INS RT2500	Position information Vehicle's movement and longitude and latitude information	Distance from right and left lane line, THW Lateral acceleration, longitudinal acceleration, longitude, latitude, yaw rate
IBEO LUX-4 LiDAR	Forward target and road edge information	Lateral and longitudinal distance of forward target, lateral and longitudinal relative velocity of forward target
MOVON camera	Video information	Driving video

Figure 2 Naturalistic driving experiment data acquisition equipment

road environment was good, the traffic volume was moderate and the weather was sunny. During the whole process of the experiment, an experimental assistant was arranged to monitor the risk factors and explain the experimental requirements in real time. The research plan was discussed with the research group, and all participants were informed of the experimental requirements and impacts.

Sample size selection is critical to obtaining sufficient experimental data. If the sample size is too small, the reliability of the results will be reduced and if the sample size is too large, resources will be wasted. For this study, the correct sample size was calculated based on expected variance, target confidence and error margin according to reference (Zhao *et al.*, 2020) as follows:

$$N = Z^2 \sigma^2 / E^2 \quad (1)$$

where N is the sample size; Z is the standard normal distribution statistic; σ is the standard deviation; E is the maximum error.

Generally, a significance level of 10% is chosen to reflect the 90% confidence level of the unknown parameter. In this study, when the confidence level was 90%, $Z = 1.25$, σ was 0.25~0.5 (Chow, 2007) and $E = 10\%$. Therefore, the minimum sample size required for calculation ranged from 10 to 39.

For this study, a total of 44 participants were recruited (female = 19; male = 25). The participants' age ranged from 22 to 55 years old (mean = 32.8, SD = 8.2). Their driving experience ranged from 2 to 18 years (mean = 6.9) and their total lifetime driving mileage ranged from 400 to 400,000 Kms (mean = 110,000). The distribution of gender, age and experience of the sample was consistent with the distribution of the general driving population in China.

3.4 Test process

In this study, naturalistic driving data was collected using a single test vehicle and a continuous measurement method. Each subject drove the test vehicle one time along the test road during a weekday. To avoid traffic flow disturbance caused by peak periods, the test was run between 09:00 to 16:00 (outside of rush hour). Each test provided subjects with route guidance only and did not interfere with their daily driving habits so as to keep the subjects in a naturalistic driving state. The test data was preprocessed to facilitate statistical analysis.

3.5 Data processing

The raw data collected by the natural driving experimental platform and the other methods is shown in Table 2. Because the original data collected by the onboard sensor inevitably experienced defects, such as missing frames, discontinuity and jump, it was necessary to clean and preprocess the original data to ensure quality. Therefore, this study used cubic spline interpolation to supplement the lost frames, filtered the noise and corrected the jump data based on the Savitzky-Golay filter and finally obtained accurate vehicle motion attitude.

The data collected in this study included driver attributes, operation parameters and road characteristics, as shown in Table 3. Driver attributes included driver ID, age and gender. Operation parameters included speed, LA, THW, time to

Table 2 Detailed information for each section

Section	Road type	Speed limit (km/h)	Lanes in each direction	Length (km)	Traffic volume
1	Expressway	70	3	13	Moderate
2	Highway	100–120	3–4	45	Dispersed
3	Expressway	80	3–4	34	Moderate
4	Arterial	40–60	2–3	12	Congested

Figure 3 Test route

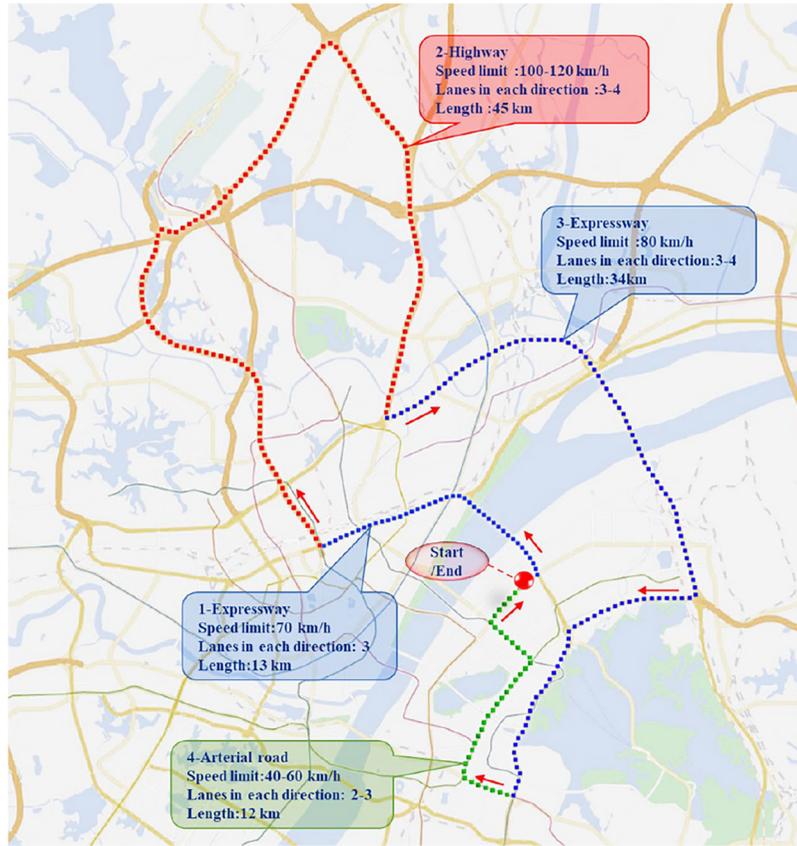


Table 3 Data collection

Data structure	Variable
Driver attribute data	ID, age, gender,
Driving operation data	Speed, THW, TTC, DHW, longitudinal acceleration
Road data	Type, length

collision (TTC) and distance headway (DHW). Road characteristics included road type and length.

3.6 Subjective driving style evaluation

As the DSQ uses subjective responses for driving style calibration, the analysis results are not only limited by sample characteristics and structural validity, but the data focuses more on driving behavior preferences and cannot fully describe a true objective driving condition. In this study, using the three-point scale method (Li et al., 2017), three drivers with rich driving experience (the actual driving mileage per person was more than 60,000 Kms and the driving experience per person was more than eight years) were selected as the scoring experts. Driving style was scored according to the video data based on three points, namely, 1 indicated a conservative driving style, 2 indicated a moderate driving style and 3 indicated an aggressive driving style. The scoring rules were set as follows:

Score

$$= \begin{cases} E_A \text{ or } E_B \text{ or } E_C, & | \text{ if } E_A = E_B = E_C \\ E_A \text{ or } E_B, & | \text{ if } E_A = E_B \neq E_C, |E_A - E_C| \leq 1 \\ E_A \text{ or } E_C, & | \text{ if } E_A = E_C \neq E_B, |E_A - E_B| \leq 1 \\ E_B \text{ or } E_C, & | \text{ if } E_B = E_C \neq E_A, |E_A - E_B| \leq 1 \\ \text{rerate}, & | \text{ otherwise} \end{cases} \quad (2)$$

where E_A is the scoring value of the first expert, E_B is the scoring value of the second expert and E_C is the scoring value of the third expert.

The results from the DSQ are shown in Figure 4. In total, 16 drivers were scored as cautious, 22 drivers were scored as moderate and 6 drivers were scored as aggressive.

4. Method

4.1 Research strategy

In a naturalistic driving environment, due to the influence of road conditions, traffic conditions, driver characteristics and other impactful factors, drivers will make myriad operations, such as accelerating, decelerating, parking, approaching, following and more. However, because different drivers have different driving styles, they make different operations under the same conditions. Therefore, the driver's operating performance under these different driving conditions can be used to identify that driver's style.

Figure 4 Driving style labeled results



As shown in Figure 5, this research firstly identified different driving styles and then labeled DOCs based on naturalistic driving data. Then, operating parameters were extracted under different DOCs and four machine learning classification (MLC) methods were used to predict driving style; the prediction performance of the models was then evaluated.

4.2 Label method of driving operation conditions

Previous studies have shown that relative distance and relative speed are two important indicators of longitudinal driving; they can be used to simulate driver behavior by taking them as elements of a regression function in longitudinal driving scenarios and models (Itkonen et al., 2020). Therefore, in this study, speed and THW were selected as the label basis of the

DOCs. The following sections discuss the DOC labels and the labeling process is shown in Figure 6. The labeling of longitudinal driving behavior conditions was performed in two steps:

4.2.1 Label acceleration and deceleration segments

Taking an acceleration segment as an example, a sliding time window was adopted. From the initial moment when the vehicle entered the expressway, a fixed sampling threshold was set to 50 frames.

As shown in Figure 7, the abscissa represents the number of frames, and the ordinate represents the speed. Within the 50-frames range of the sliding time window ($t_2 - t_1 \geq 50$), if the speed increased, the driving segment of (t_1, t_2) was temporarily marked as an accelerating segment, otherwise, the driving

Figure 5 Research strategy of driving style identification

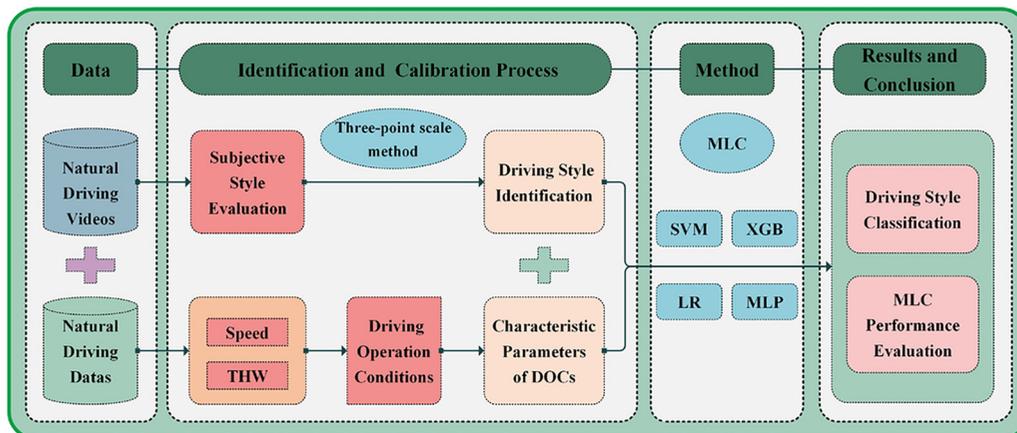


Figure 6 Label process of DOCs

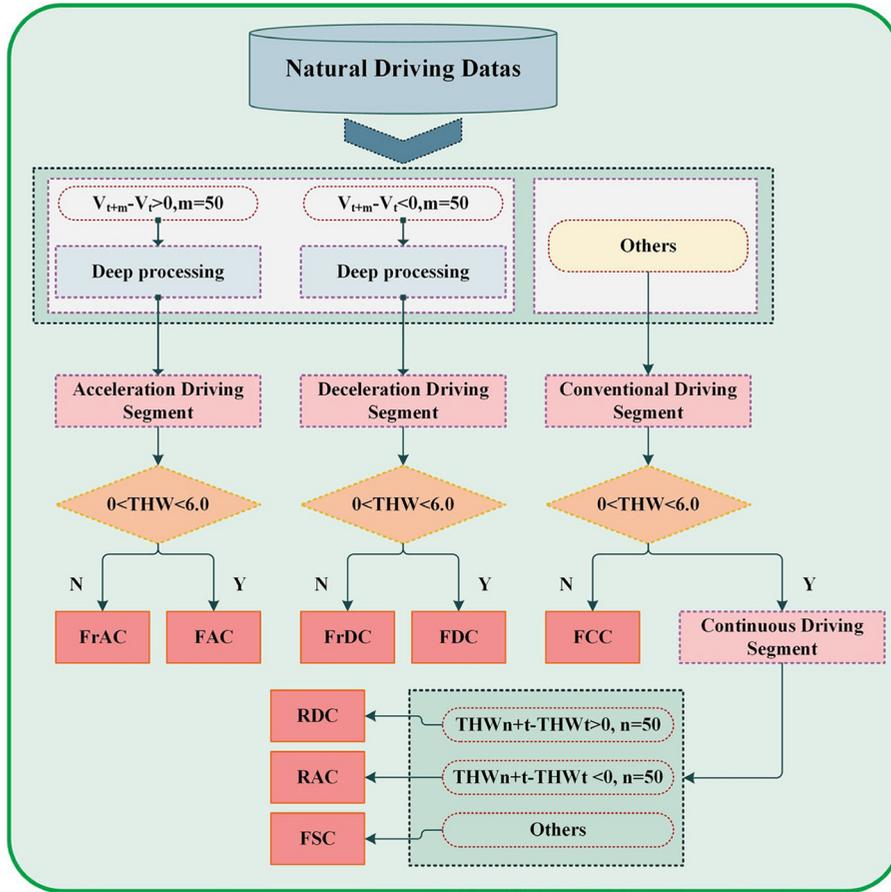
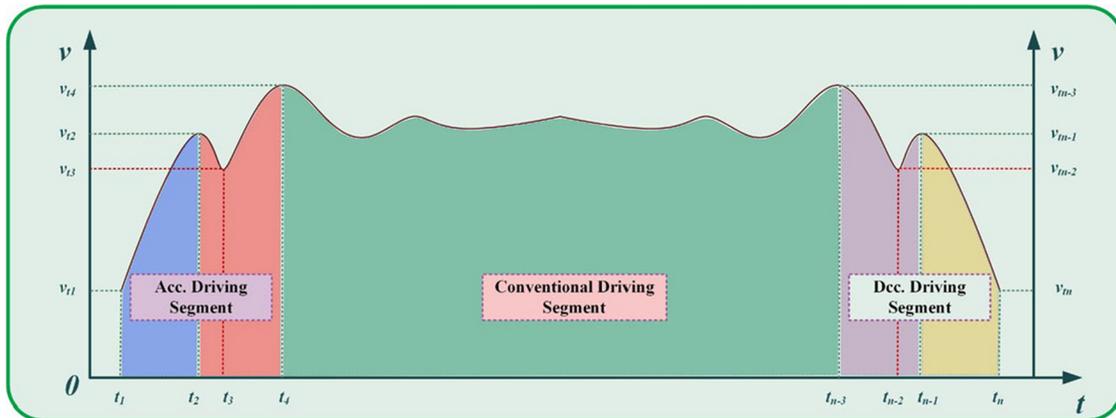


Figure 7 Schematic diagram of acceleration and deceleration segment label



segment of (t_1, t_2) was marked as a conventional driving segment. If the speed decreased in the range of (t_2, t_3) , subsequent processing was required. The subsequent processing followed key principles:

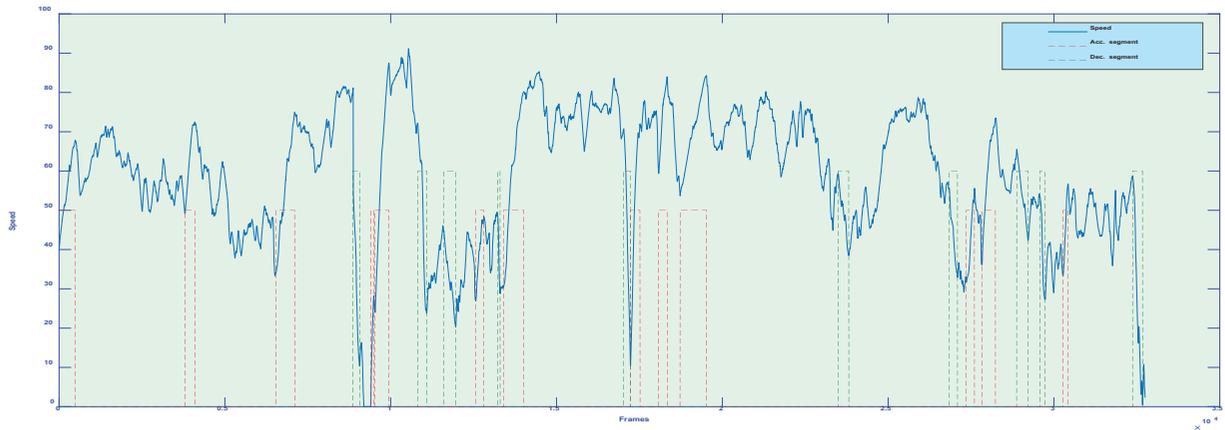
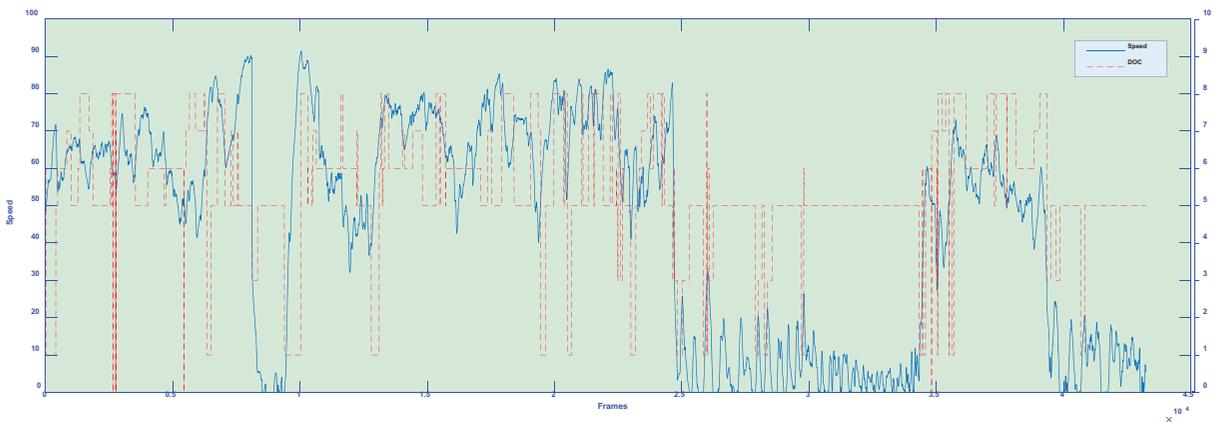
When the speed decreased at t_2 but started to rise at t_3 and the speed reaches its peak at t_4 :

If $t_3 - t_2 \geq 5$, $v_{t_4} > v_{t_2}$, then (t_1, t_4) was marked as an accelerating segment;

If $t_3 - t_2 > 5$, $t_4 - t_3 \geq 50$, then (t_1, t_2) and (t_3, t_4) were marked as accelerating segments and (t_1, t_2) was marked as a conventional driving segment for a further label;

If $t_3 - t_2 > 5$, $t_4 - t_3 < 50$, then (t_1, t_2) was marked as an accelerating segment and (t_2, t_4) was marked as a conventional driving segment for a further label.

Then, THW was used to determine whether the vehicle was following a car in the time window and the driving segment was

Figure 8 Acceleration and deceleration segment label**Figure 9** The label result of longitudinal DOCs

labeled as either a following acceleration condition (FAC) or a free acceleration condition (FrAC). Because of the detection equipment, a 0 in the THW data meant that there was no leading vehicle and a non-zero meant that there was a leading vehicle detected ahead. In addition, the accelerating segment with $\text{THW} \geq 6\text{s}$ was also marked as a FrAC because when $\text{THW} \geq 6$, the vehicle was in a relatively safe driving state. The label process for deceleration conditions was similar.

4.2.2 Label other conventional driving segments

The other conventional driving conditions included a free cruise condition (FCC), following steady condition (FSC), relatively distant condition (RDC) and a relative approximation condition (RAC). The sliding window was used to identify and label these continuous driving segments – all except for the FCC, which was labeled based on a $\text{THW} > 6\text{s}$ or $\text{THW} = 0$ – and the threshold and methods were similar to the acceleration label process described above. Within the sampling threshold, an increasing or decreasing THW was determined and the FSC, RDC and RAC were automatically labeled by MATLAB.

This study did not consider the impact of latitudinal vehicle operations (i.e. lane-changing). Only longitudinal driving conditions were considered. To sum up, the eight longitudinal driving conditions are defined as follows.

The FrAC and FrDC indicate that the speed of the host vehicle increased or decreased, respectively, within the sliding window detection time of 50 frames and either no leading vehicle was in front or the headway time between the front and rear vehicles was more than 6.0 s.

The FAC and FDC (following deceleration condition) indicate that the speed of the host vehicle increased or decreased, respectively, within the sliding window detection time of 50 frames and a leading vehicle was in front and the headway between the front and rear cars was within 6.0 s.

The FCC indicates that the speed of the host vehicle changed repeatedly within the sliding window detection time of 50 frames and either a leading vehicle was not detected in front or the headway between the front and rear vehicles was more than 6.0 s.

The RDC and RAC indicate that the speed of the host vehicle changed repeatedly and alternately within the sliding window detection time of 50 frames and a leading vehicle was in front and the headway of the front and rear vehicles was within 6.0 s. Within the 50-frame sliding window detection time, the headway time showed an increasing RAC or a decreasing RDC.

The FSC indicates that the speed of the host vehicle changed repeatedly within the sliding window detection time of 50 frames and there was a leading vehicle in front and the headway

between the front and rear vehicles was within 6.0 s. Within the sliding window detection time of 50 frames, the headway time showed repeated and alternate changes.

4.2.4 Measurement of index

Drawing on the 10 observable driving style indices described in existing literature (Itkonen et al., 2020), the longitudinal driving behavior analysis indices and particular index measurement, including speed (V), LA, THW, the count backward of TTC and DHW, were selected to characterize the driving style. For each driving condition, the index measurement was different. For example, for FCC, because only parameters of the vehicle were relevant, only speed and LA were calculated. The particular analysis indices and index measurement values of the DOCs are shown in Table 4.

The studied naturalistic driving data was captured from 44 participants driving on experimental road section 1 and experimental road section 3. The speed limit of road 1 was 70 Km/h and the speed limit of road 3 was 80 Km/h. In addition, the length of the two roads and the traffic flow on each road also differed, as observed through video. Therefore, to ensure high quality data analysis, the data of roads 1 and 3 were divided into independent analysis units and data from the whole process of driving on each section from the beginning to the exit was divided into small units with equal time intervals according to $t = 600$ frames. Then, the statistical index values in the small units that were split in different sections were analyzed, as shown in Table 3. Fragmented data less than 10 min was removed and subsequent analysis was not carried out. In this way, the naturalistic driving data from the 44 drivers on the two tested expressways was divided into 229 driving segments, and all the statistical analysis indicators were summarized to form a 229×211 driving condition index analysis matrix.

4.3 Machine learning classification methods

This study aimed to test the feasibility of using longitudinal DOCs to identify driving styles through MLC algorithms. To achieve these goals and based on previous literature, it was found that MLC models, namely, SVM, XGB, LR and MLP, have shown relatively good predictive performance in existing practical applications. Therefore, this study evaluated the prediction performance of these four machine learning models based on the label analysis of the DOCs:

(1) Support vector machines (SVMs):

Support vector machines (SVM) are one of the most widely used supervised classification methods in the field of machine

learning and artificial intelligence. The SVM proposed by Cortes and Vapnik made full use of the structural risk minimization theory, thus ensuring the strong generalization ability of the model (Cortes and Vapnik, 1995). SVMs are a supervised learning method to predict the labels of points in the test data set by learning the model of the training data set. This method is well-known in computer science and has been widely used in the field of transportation engineering, such as traffic accident prediction (Tang et al., 2020; Zhang et al., 2018), road risk prediction (Basso et al., 2018), vehicle trajectory state recognition (Siddique et al., 2019) and path selection (Sun et al., 2017), driving behavior prediction (Wang et al., 2017) and driving state recognition (Chai et al., 2019; Allahviranloo, 2013). SVMs have generally good predictive performance.

(2) Extreme gradient boosting (XGB):

XGB is an integrated machine learning model based on many decision trees that use an optimized gradient boosting system. It has the advantages of performing parallel processing, approximate greedy search and improving the learning process in the shortest time without overfitting. It has been proven that XGB has superior predictive performance and processing time compared with the random forest model (Chen and Boost, 2016). In recent years, XGB models have been proven to have good performance in traffic flow prediction (Mahmoud et al., 2021), rail defects prediction (Mohammadi et al., 2019), driving behavior prediction (Ayoub et al., 2021) and road risk identification and prediction (Das et al., 2020).

(3) Logistic regression (LR):

LR is generally used to model the relationship between a categorical dependent variable and categorical/dichotomous/continuous independent variables. These models predict the probability of occurrence of the dependent variable using a set of given independent variables (Venkata et al., 2020). LR is a generalized linear model and has been widely used in accident prediction (Venkata et al., 2020; Dong et al., 2018) and conflict risk prediction (Costela et al., 2020) in traffic safety research. At the same time, it is used in traffic system performance tests (Cafiso et al., 2020; Liu et al., 2018) and behavior prediction (Farooq et al., 2021; Ghasemzadeh et al., 2018).

(4) Multi-layer perceptron (MLP):

The use of NNs and deep learning optimization algorithms to enhance discrete selection models is an active research area, which has shown encouraging results (Zargamezhad et al., 2019). In recent years, experimental cases of deep learning methods in discrete choice models have been explored, such as personal travel mode prediction (Omran, 2015), path tracking prediction (Ge et al., 2021), driving behavior feature recognition (Jasper et al., 2018) and more. As a basic three-layered back-propagation MLP model was used to develop the first NN (Clark, 1993), MLP has been developed into a novel non-parametric approach based on an MLP NN and has been demonstrated to be successful in complex behavioral data modeling (Costa et al., 1997).

4.4 Model prediction performance evaluation

After parameter adjustment and model training, it was necessary to evaluate the generalization ability of the model on an independent test set. To evaluate the performance of the prediction model, a confusion matrix was introduced. Taking

Table 4 The measures of driving style used in the analysis

DOCs	Analysis index	Measurement of index
FrAC,	V, LA, THW,	Mean, standard deviation, quartile
FrDC, FAC,	TTC-1, DHW	(15%, 50%, 85%), mode (except the parameter of TTC-1), maximum, minimum
FDC, FSC,		
RDC, RAC		
FCC	V, LA	Mean, standard deviation, quartile (15%, 50%, 85%), mode, maximum, minimum

the dichotomy problem as an example, the confusion matrix is shown in Table 5.

True positive indicates that the number of the true value was positive and the predicted value was positive. False negative indicates that the number of the true value was positive, but the predicted value was negative. False positive indicates that the number of the true value was negative, but the predicted value was positive. True negative indicates that the number of the true value was negative and the predicted value was negative.

The indicators of accuracy (ACC), precision (PPV), sensitivity or recall rate (TPR), FPR, specificity TNR and the F1-score were used to evaluate the performance of the models. The calculation formula and meaning of the evaluation indices are shown in Table 6.

5. Results

5.1 Calibration results of longitudinal driving operation conditions

Naturalistic driving data from 47 Kms of the expressway (test route 1 and 3) was extracted and the label method described in the previous section was used to identify the DOCs from 44 drivers on the tested expressway. Figure 10 describes the DOCs frequency distribution from different drivers. It can be observed that unlike the label results on the entire experimental section, only six DOCs, namely, FAC, FDC, FCC, RDC, RAC and FSC, appeared on the expressway for all drivers, while FrAC and FrDC did not appear at all. By definition, FrAC and FrDC generally do not appear on expressways and by reviewing the natural driving video data, it was also confirmed that FrAC and FrDC are not present on the tested expressway.

Table 5 Confusion matrix

	Predictive value	
	Positive	Negative
Real value		
Positive	TP	FN
Negative	FP	TN

Table 6 Model prediction result evaluation index

Evaluation index	Formula	Meaning
Accuracy (ACC)	$ACC = \frac{TP}{TP + TN + FP + FN}$	The proportion of all the correct results of the classification model to the total observed values
Precision (PPV)	$PPV = \frac{TP}{TP + FP}$	Among all the results where the model prediction was positive, the proportion of correct model predictions
Sensitivity (TPR)	$TPR = \frac{TP}{TP + FN}$	Among all the results where the true value was positive, the proportion of correct model predictions
False positive (FPR)	$FPR = \frac{FP}{FP + TN}$	Among all the results that the true value was negative, the proportion that was incorrectly predicted
Specificity (TNR)	$TNR = \frac{TN}{TN + FP}$	Among all the results where the true value was negative, the proportion of correct model predictions
F1-score	$Score = \frac{TP}{TP + (FN + FP)/2}$	Integrate the results of precision and recall's output. The value ranged from 0 to 1. 1 represents the best output of the model, and 0 represents the worst

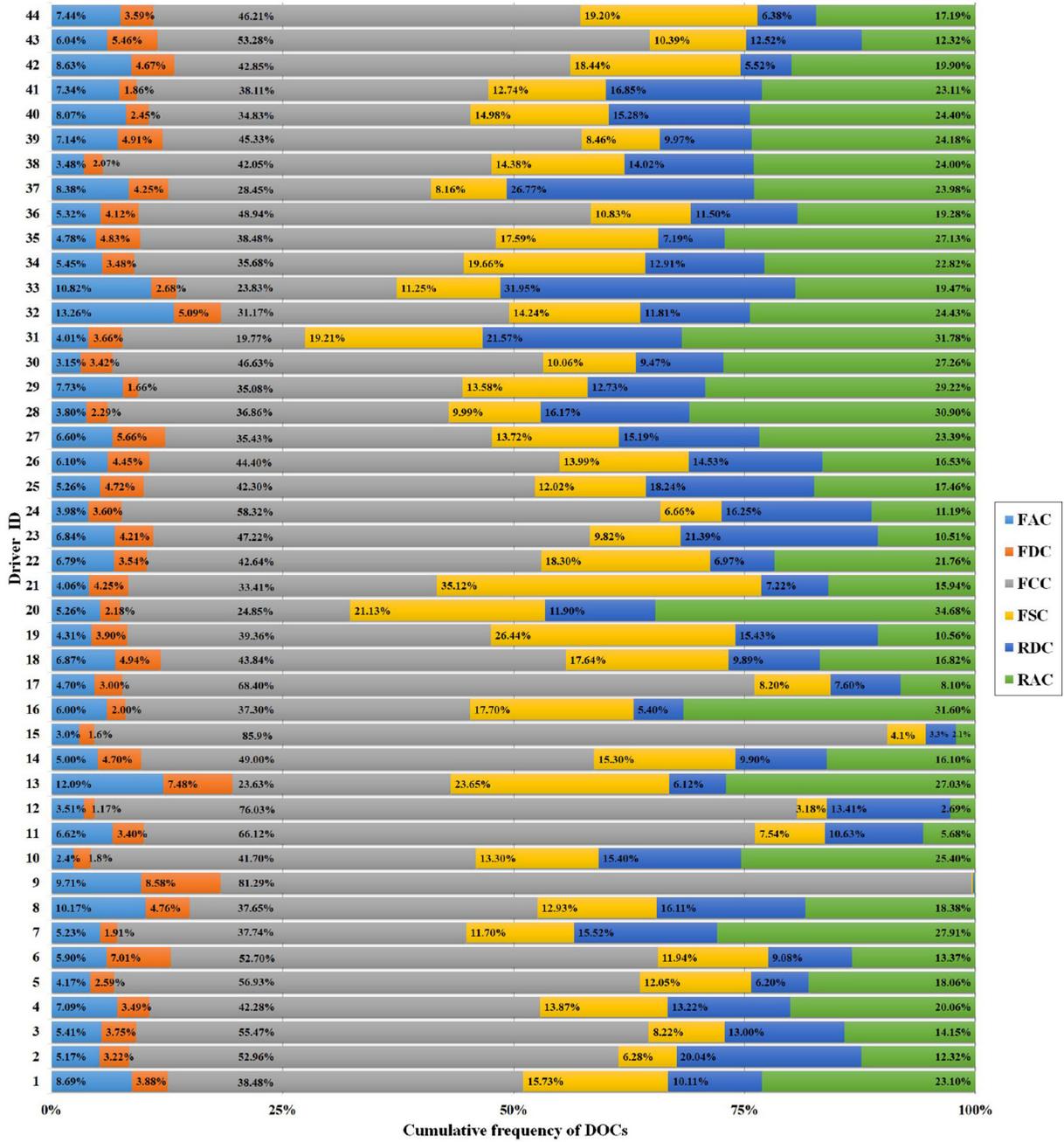
It can be observed in Table 7 that the FCC occurred most frequently, indicating that, when driving on the expressway, drivers were most likely to adopt FCC and less likely to adopt FAC and FDC. The reason may be that when formulating the criteria for labeling the DOCs, the model was established based on the naturalistic driving data of the entire experimental road section. The data input took into account driving data from multiple types of roads, whereas only two types of roads were actually analyzed. In addition, the overall law of DOCs distribution among all drivers was roughly the same, but the mean and variance of each DOC ratio were different, which reflects the heterogeneity of the frequency distribution of the different DOCs.

5.2 Driving style identification with different machine learning classification methods

With reference to the four machine learning models, a sample set was established to distinguish driving style. The difference in this study is that the samples were divided into driving style labels – namely, conservative driving style, moderate driving style and aggressive driving style – in the data aggregation stage. The sample set was divided into 70% training set and 30% test set. At the data level, the problem of the imbalanced number of samples for conservative drivers, moderate drivers and aggressive drivers was addressed. The ENN method was used to undersample the normal samples in the training set. Then, the five-fold cross-validation method was used to train and verify the data of the training set and finally the model was tested on an independent test set. Table 8 shows the confusion matrix predicted by the established model to distinguish driving style on the independent test set. The values in Table 8 represent the number of driving segments in the test set.

Table 9 shows that the MLP model had the highest overall accuracy. The most accurate prediction models of aggressive driving style, moderate driving style and conservative driving style were XGB, MLP, MLP ($PPV_{Aggressive} = 1.000$, $PPV_{Ordinary} = 0.659$, $PPV_{Conservative} = 0.867$), respectively. From the perspective of sensitivity (TPR), the detection rate of moderate driving style was higher than that of aggressive driving style and conservative driving style. This shows that these models had better predictive ability for moderate driving

Figure 10 Frequency cumulative distribution of different DOCs



styles. The FPR of moderate driving style was higher than that of aggressive driving style and conservative driving style. In the point of view of the F1-score, apart from the LR model, the

Table 7 Statistical analysis of different longitudinal DOCs

DOC	FAC	FDC	FCC	FSC	RDC	RAC
Mean (%)	6.27	3.78	44.61	13.50	12.61	19.24
Median (%)	6.00	3.66	42.30	13.30	12.61	19.47
Standard deviation	2.39	1.56	14.28	6.21	5.95	8.16
Minimum (%)	2.40	1.17	19.77	0.12	0.18	0.12
Maximum (%)	13.26	8.58	85.90	35.12	31.95	34.68

other prediction results exceeded 0.5, indicating that the overall output performance of the model was general under this sample size. Because it was difficult to clearly define a moderate driving style, its recognition rate was not high, which affected the overall recognition level of all the models.

Table 9 also shows that, under the current sample size, a small number of extracted longitudinal driving conditions can be used to effectively identify driving styles through MLC models, and with the increase of sample size, the accuracy of driving style identification will significantly improve. However, different MLC models differ in performance in the identification of driving style. It was found that the four models

Table 8 Different machine learning model prediction results

	SVM				XGB		
	Aggressive	Moderate	Conservative		Aggressive	Moderate	Conservative
<i>Real value</i>							
Aggressive	5	4	0	Aggressive	3	6	0
Moderate	1	30	2	Moderate	0	27	6
Conservative	0	17	10	Conservative	0	13	14
	LR				MLP		
	Aggressive	Moderate	Conservative		Aggressive	Moderate	Conservative
<i>Real value</i>							
Aggressive	2	6	1	Aggressive	6	2	1
Moderate	3	23	7	Moderate	3	29	1
Conservative	1	15	11	Conservative	1	13	13

Table 9 The prediction results of different machine learning models

Driving style	Aggressive					Moderate					Conservative					ACC
	PPV	TPR	FPR	TNR	F1-Score	PPV	TPR	FPR	TNR	F1-Score	PPV	TPR	FPR	TNR	F1-Score	
SVM	0.833	0.556	0.017	0.983	0.667	0.588	0.909	0.583	0.417	0.714	0.833	0.37	0.048	0.952	0.513	0.652
XGB	1.000	0.333	0	1.000	0.500	0.587	0.818	0.528	0.472	0.684	0.700	0.519	0.143	0.857	0.596	0.638
LR	0.333	0.222	0.067	0.933	0.267	0.523	0.697	0.583	0.417	0.597	0.579	0.407	0.190	0.810	0.478	0.522
MLP	0.600	0.667	0.067	0.933	0.631	0.659	0.879	0.417	0.583	0.753	0.867	0.481	0.048	0.952	0.619	0.696

all showed good performance in the prediction of driving style. However, in terms of accuracy, precision, recall and F1-score, the MLP model had the best prediction results.

6. Discussion

6.1 Statistical analysis of parameters based on different driving styles

According to the calibration results of the DOCs, a scatter diagram of average longitudinal driving behavior parameters was drawn, as shown in Figure 11. It can be observed that the scatter distribution of THW and DHW was significantly different.

Compared with cautious and moderate drivers, aggressive drivers adopted a smaller THW and DHW during the natural driving experiment, indicating that THW and DHW showed high significance for the identification of driving style. However, the significance of the other three parameters for the identification of driving style needed to be further analyzed.

According to normality and lognormality tests, it was found that the longitudinal driving behavior parameters of different driving styles do not conform to the Gaussian distribution. Therefore, a non-parametric test was adopted to analyze the correlation of longitudinal driving behavior parameters to different driving styles. As the number of drivers who exhibited different driving styles was imbalanced, as were the different DOCs parameters, the sample size of each group was asymmetrical. Therefore, the Kruskal Wallis test method was used for a non-parametric one-way ANOVA of the population sample. Meanwhile, Dunn's Multiple Index test method was also selected for the non-parametric one-way ANOVA comparative analysis of driving data from drivers who exhibited different driving styles. The results are shown in Table 10.

The four longitudinal driving behavior parameters of speed, THW, TTC and DHW showed significant differences in driving style identification, while the LA showed no significant difference in driving style identification. In particular, THW, TTC and DHW showed highly significant differences in driving style identification. This also indicated that the driver's subjective perception of LA during natural driving was far less strong than the objective factors of speed, THW, TTC and DHW. This distinction is useful for ADAS-equipped vehicles, which can display THW, TTC and DHW in real time through the onboard intelligent display terminal, so that drivers can easily respond to this data and adopt different driving strategies – also in real time.

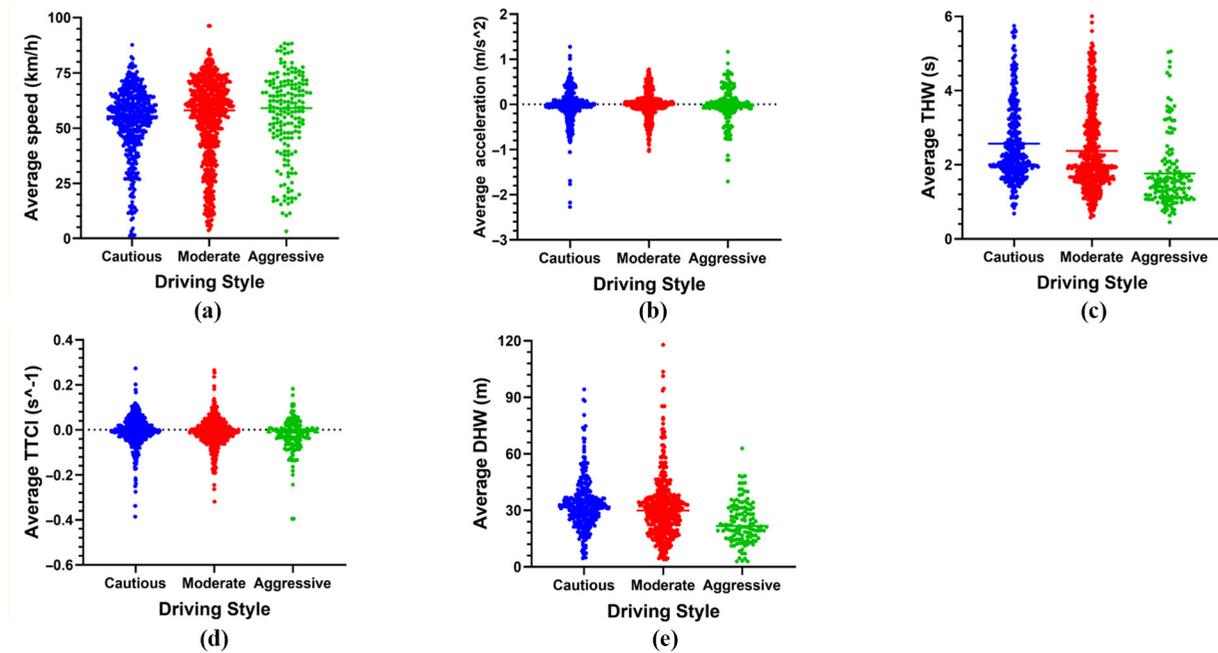
From the results of multiple comparison analyzes, LA showed no significant difference between the three driving styles. At the same time, speed and TTC showed no significant difference between moderate and aggressive drivers. This also indirectly shows that there was little difference between moderate and aggressive drivers.

6.2 Statistical analysis of parameters based on different longitudinal driving operation conditions

Based on the 229 segments of naturalistic driving data, the box plots of the mean values of speed, LA, THW, TTCi and DHW were drawn according to the six DOCs, as shown in Figure 12. It should be noted that the FCC lacked the statistics of THW, TTCi and DHW.

It can be observed in Figure 12 that, among all the DOCs, the mean speed of the FDC was the lowest (FAC = 54.3 Km/h, FDC = 42.4 Km/h, FCC = 47.3 Km/h, FSC = 57.5 Km/h, RDC = 58.3 Km/h and RAC = 60.1 Km/h). This shows that when drivers were in the FDC, most drove at a low following

Figure 11 Data scatter diagram of longitudinal DOCs of drivers with different driving styles



Notes: (a) Average speed scatter diagram; (b) Average acceleration scatter diagram; (c) Average THW scatter diagram; (d) Average TTCi scatter diagram; (e) Average DHW scatter diagram

Table 10 Non-parametric test of one-way ANOVA results

Non-parametric test method	Speed	Acceleration	P-value summary		
			THW	TTCi	DHW
Kruskal-Wallis test	0.011*	0.784	<0.001***	0.004**	<0.001***
<i>Dunn's multiple comparisons test</i>					
Cautious vs moderate	0.048*	0.980	0.001***	0.020*	0.006**
Cautious vs aggressive	0.014*	>0.999	<0.001***	0.006**	<0.001***
Moderate vs aggressive	0.450	>0.999	<0.001***	0.422	<0.001***

speed to maintain safety. However, the average speed was higher in the FAC, which indicates that the following vehicle accelerated when the lead vehicle accelerated. The speed distribution of the FSC, RDC and RAC was relatively uniform. The FCC had the largest range of speed fluctuations. This may be related to the fact that the vehicle entered an expressway from an urban road with a relatively low average speed. During this process, vehicles were required to accelerate.

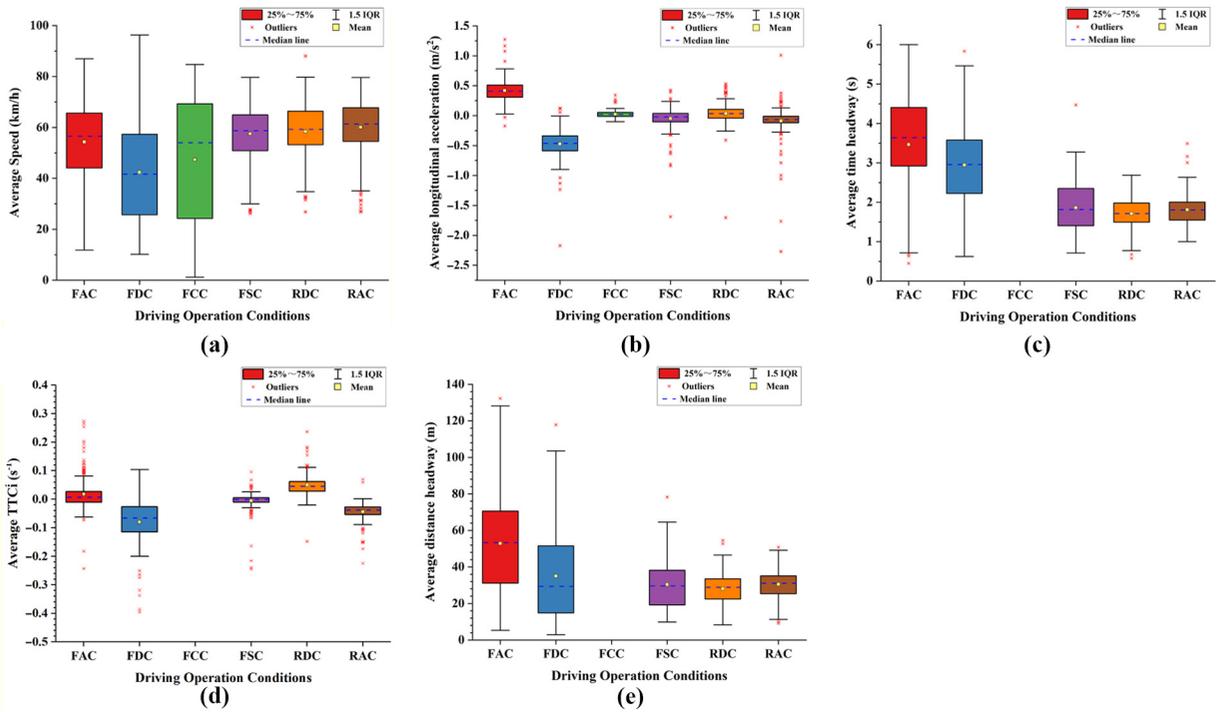
The average value of LA of FAC and FDC had similar distributions and the average absolute value of the LA between FAC and FDC (FAC = 0.42, FDC = 0.47) had little difference, but the absolute value of the maximum value of FDC was slightly larger than FAC (FAC = 0.78, FDC = 0.90) and significantly higher than the other DOCs. This shows that the driver had obvious acceleration or deceleration under these two DOCs, but the driving operation under the other DOCs was relatively smooth. The abnormal value of LA also illustrated the operating performance of aggressive drivers under different DOCs.

Meanwhile, the interquartile range of the average THW of FAC (Q3 = 4.40 s, Q1 = 2.92 s, IQR = 1.48 s, mean = 3.46 s)

was bigger than FDC (Q3 = 3.58 s, Q1 = 2.22 s, IQR = 1.36 s, mean = 2.94 s), which indicates that drivers generally maintained a larger THW when following accelerating vehicles than when following decelerating vehicles. This shows that when a rear vehicle followed a front accelerating vehicle, the rear vehicle showed a delay effect. When the rear vehicle followed a front decelerating vehicle, the rear vehicle showed aggressive behavior, resulting in a small THW. This can also be observed from the LA index of FDC. It can be observed from the interquartile range of the average THW of FSC (Q3 = 2.35 s, Q1 = 1.41 s, IQR = 0.94 s, mean = 1.86 s), RDC (Q3 = 1.98 s, Q1 = 1.49 s, IQR = 0.49 s, mean = 1.71s) and RAC (Q3 = 2.00 s, Q1 = 1.55 s, IQR = 0.45 s, mean = 1.80 s) that when the vehicle was in these three DOCs, although the vehicle was still following, it did not rapidly accelerate or decelerate, but the THW was already less than 3.0s, which is consistent with existing research conclusions (Xu et al., 2015; Suzdaleva and Nagy, 2018).

In general, although drivers exhibited different driving styles, they all maintained a large TTC when driving on the expressway. While the TTC index has been widely used for

Figure 12 Box plot for different DOCs



Notes: (a) Box plot of average speed for different DOCs; (b) Box plot of average LA for different DOCs; (c) Box plot of average THW for different DOCs; (d) Box plot of average TTCi for different DOCs; (e) Box plot of average DHW for different DOCs

potential risk assessment, the abnormal value of TTC under different DOCs reflects the behavior of different driving styles; in particular, the TTC of aggressive drivers fluctuated greatly.

Apart from the FCC, the interquartile distribution distance of the FAC (Q3 = 70.6 m, Q1 = 31.2 m, IQR = 39.4 m, mean = 52.8 m) and FDC (Q3 = 51.5 m, Q1 = 14.8 m, IQR = 36.7 m, mean = 35.1 m) were much larger than that of the FSC (Q3 = 38.2 m, Q1 = 19.2 m, IQR = 19 m, mean = 30.3 m), RDC (Q3 = 33.5 m, Q1 = 22.4 m, IQR = 11.1 m, mean = 28.0 m) and RAC (Q3 = 35.1 m, Q1 = 25.4 m, IQR = 9.7 m, mean = 30.5 m), indicating that the DHW of all drivers regardless of their dominant style was significantly different under the FAC and FDC, while the DHW difference was not significant under the FSC, RDC and RAC. In addition, from the perspective of mean distribution, the mean DHW of the FAC was higher than that of the FDC (FAC = 52.8, FDC = 35.1), which indicates that all drivers regardless of their dominant style were more inclined to follow a vehicle with a larger distance under the FAC.

According to normality and lognormality tests, it was found that the longitudinal driving control data of different DOCs did not conform to the Gaussian distribution, so a non-parametric test and analysis was adopted.

As the number of driving segments was consistent, the Friedman test method was used for a non-parametric one-way ANOVA of the sample population. At the same time, Dunn's multiple comparisons test method was selected to perform a non-parametric one-way ANOVA comparison analysis on the driving segment data from different DOCs. The analysis results

are presented in Table 11, which shows that longitudinal driving behavior parameters showed highly significant differences in the calibration of longitudinal DOCs ($p < 0.001$).

6.3 Frequency of longitudinal driving operation conditions based on different driving styles

As shown in Figure 13 and Table 11, the results of the DOC calibrations were classified and statistically analyzed according to driving style. In this naturalistic driving test, all drivers regardless of their dominant style preferred FCC. In addition to the influence of road factors (such as less crowded traffic flow and better road alignment), it showed that all drivers preferred free cruising conditions and attempted to avoid complex following conditions.

It can be observed in Table 11 that cautious drivers took the largest proportion of FCC and the one-way ANOVA showed no difference ($P = 0.073$), indicating that cautious drivers tended to maintain FCC for a long time. On the contrary, there were significant differences between moderate and aggressive drivers, indicating that they will change their driving strategies according to the changes of driving environment in the process of naturalistic driving. In particular, the proportion of FSC and RAC by aggressive drivers was higher, indicating that aggressive drivers tended to challenge complex driving conditions.

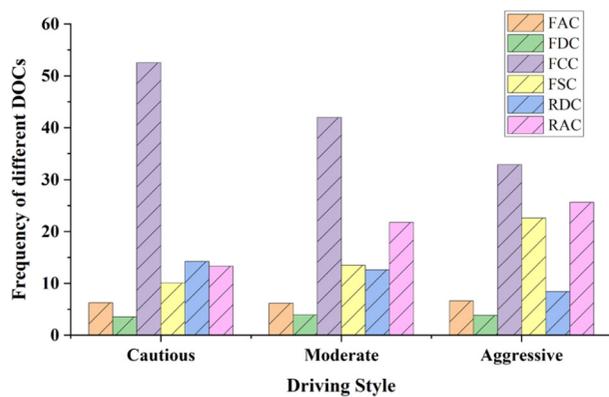
6.4 Discussion of model recognition results

With reference to the four machine learning models, a sample set was established to distinguish driving styles. The difference

Table 11 Non-parametric test of one-way ANOVA results

Non-parametric test method	P-value summary				
	Average speed	Average longitudinal acceleration	Average THW	Average TTCi	Average DHW
Friedman test	<0.001***	<0.001***	<0.001***	<0.001***	<0.001***
<i>Dunn's multiple comparisons test</i>					
FAC vs FDC	<0.001***	<0.001***	<0.142	<0.001***	<0.001***
FAC vs FCC	0.817	<0.001***	–	–	–
FAC vs FSC	>0.999	<0.001***	<0.001***	0.040*	<0.001***
FAC vs RDC	>0.999	<0.001***	<0.001***	<0.001***	<0.001***
FAC vs RAC	0.008**	<0.001***	<0.001***	<0.001***	<0.001***
FDC vs FCC	<0.001***	<0.001***	–	–	–
FDC vs FSC	<0.001***	<0.001***	<0.001***	<0.001***	>0.999
FDC vs RDC	<0.001***	<0.001***	<0.001***	<0.001***	0.288
FDC vs RAC	<0.001***	<0.001***	<0.001***	>0.999	>0.999
FCC vs FSC	0.054	0.003**	–	–	–
FCC vs RDC	0.032*	>0.999	–	–	–
FCC vs RAC	<0.001***	<0.001***	–	–	–
FSC vs RDC	>0.999	0.001**	0.359	<0.001***	0.529
FSC vs RAC	0.194	0.094	>0.999	<0.001***	>0.999
RDC vs RAC	0.302	<0.001***	0.714	<0.001***	0.022*

Figure 13 Frequency of different DOCs between different driving style



is that the samples were divided into driving style labels in the data aggregation stage, namely, conservative driving style, moderate driving style and aggressive driving style. The sample set was divided into 70% training set and 30% test set. At the data level, the problem of the unbalanced number of conservative driving style, moderate driving style and aggressive driving style samples was addressed. The ENN method was used to undersample the normal samples in the training set. Then, the five-fold cross-validation method was used to train and verify the data of the training set and finally the model was

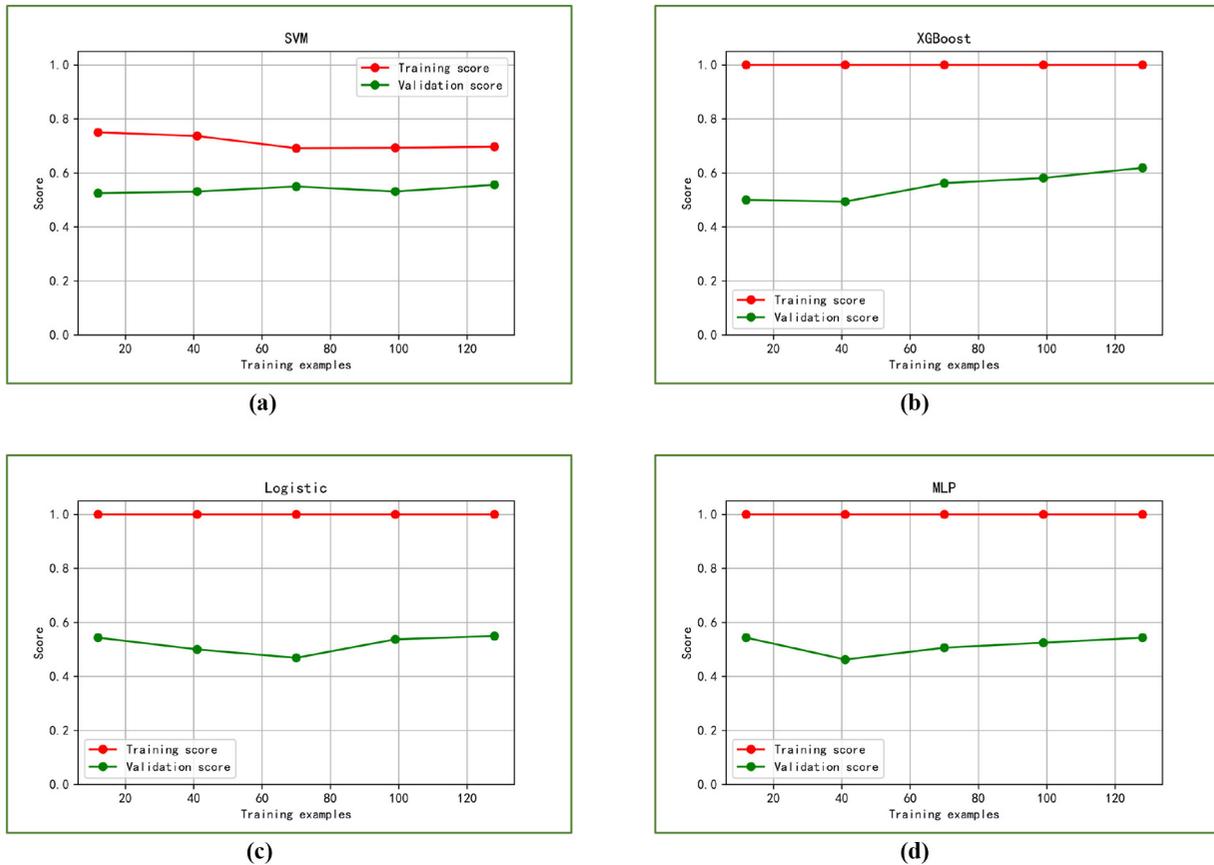
tested on an independent test set. Table 9 shows the confusion matrix predicted by the established model to distinguish driving style on the independent test set. The values in Table 9 represent the number of driving segments in the test set.

Figure 14 shows the variation trend and overfitting of the prediction accuracy of the training set and validation set with the increased sample training number in the cross-validation process of the four machine learning models, namely, SVM, XGB, LR and MLP. Table 9 shows the comparison of the prediction results of these models on the test set. For multiple classification problems, the evaluation index of the model was redefined. The accuracy of the model was the same as that of the binary classification problem, which was still the right proportion of the correctly classified samples to all the samples. As the confusion matrix of the three-way classification was different from that of the dichotomy, the PPV, TPR, FPR, TNR, F1-score were also different. In this study, to directly reflect the prediction of different driving styles, when calculating the evaluation index of any type of driving style prediction, the two types of driving styles were merged as one situation and then it was regarded as a binary classification problem.

Figure 14 shows that the fitting accuracy of the SVM model on the training set was less than 80%, while the fitting accuracy of the other three models on the training set reached 100%. Moreover, with the gradual increase of the number of samples,

Table 12 Statistical analysis of frequency of different longitudinal DOCs

Driving style	Frequency of different DOCs (%)						Mean (%)	SD	Sig
	FAC	FDC	FCC	FSC	RDC	RAC			
Cautious	6.25	3.55	52.57	10.08	14.23	13.32	16.67	18.05	0.073
Moderate	6.19	3.93	42.02	13.50	12.58	21.79	16.67	13.91	0.033*
Aggressive	6.61	3.82	32.92	22.59	8.42	25.64	16.67	11.95	0.019*

Figure 14 Cross-validation results of different machine learning models

Notes: (a) SVM model cross-validation results; (c) LR model cross-validation results; (d) MLP model cross-validation results

the performance of the SVM model on the training set worsened. In other words, the SVM model tended to be suitable for the training of data sets with a small sample size. From the point of view of the validation score, all the models were over-fitting. However, as the sample size gradually increased, the scores of all the models on the test set showed an upward trend and the change was most obvious for the XGB model. With the increase of the test sample size, the problem of overfitting of each classification model was gradually alleviated. Compared with other models, the overfitting problem of the SVM model had a smaller gap, but this was because the performance of the SVM model increased on the test set but decreased on the training set. That is to say, the SVM model relied on the decrease of accuracy on the training set and the increase of accuracy on the test set to solve the over-fitting problem, which is completely inconsistent with the performance of the other three models. Therefore, after analyzing the cross-validation results of the different machine learning models, the hierarchical performance ranking of the four models on the test set and training set was $XGB > MLP > LR > SVM$. Considering model cross-validation results and prediction results, the overall hierarchical prediction performance ranking of the four machine learning models under the current sample data set was $XGB > MLP > LR > SVM$.

7. Conclusions

The driving style of each driver is not fixed; it is affected by driving environment, traffic state, psychological state and myriad other influencing factors. This exemplifies the characteristics of temporal and spatial instability and segment heterogeneity. If a real-time evaluation method of driving style based on driving segment change can be constructed, it is of great significance for formulating personalized driving strategies, improving driving safety and reducing fuel consumption. The purpose of this research was to identify DOCs based on longitudinal driving behavior data and rapidly predict and label driving styles through MLC models. The main contributions of this research are as follows:

- Based on the longitudinal driving behavior parameters of naturalistic driving data, six DOCs of naturalistic driving on expressways were calibrated by formulating reasonable calibration rules, and the feasibility of the DOC calibrations was verified by naturalistic driving video data.
- Compared with cautious and moderate drivers, aggressive drivers adopted a smaller THW and DHW during naturalistic driving. THW, time-to-collision (TTC) and DHW, three well-established longitudinal driving behavior parameters, showed highly significant differences in driving style identification, while LA showed no

significant difference in driving style identification. At the same time, speed and TTC showed no significant difference between moderate and aggressive drivers.

- Cautious drivers undertook the largest proportion of FCC, while aggressive drivers primarily undertook FCC, FSC and RAC, which indicated that cautious drivers preferred free cruising, but aggressive drivers tended to challenge complex driving conditions.
- Four MLC methods, namely, SVM, XGB, LR and MLP, were used to classify and predict driving style based on the six DOCs. In consideration of the cross-validation results and model prediction results, the overall hierarchical prediction performance ranking of the four machine learning models under the current sample data set was $XGB \succ MLP \succ LR \succ SVM$.

The contribution of this research is to propose a criterion and solution for using longitudinal driving behavior data to label longitudinal DOCs and rapidly identify driving styles based on those DOCs and MLC models. This study provides a reference for real-time online driving style identification in vehicles equipped with onboard data acquisition equipment, such as ADAS.

However, there are still some directions to be further studied:

- Naturalistic driving data was heterogeneous due to different road types; as a result, the threshold criterion for the label of the DOCs based on driving data from different road types may not be portable nor extensive. Therefore, the DOCs calibration criteria developed in this study may not be fully applicable to driving style identification on all types of road scenes. In addition, the problems of endogeneity among various DOCs and the spatiotemporal correlation also needs to be further studied.
- The influence of lateral driving behavior was simplified in this research, which may affect the training and test performance of the model. This research was an attempt to *quickly* label driving style. The multi-dimensional data of the vehicle's longitudinal and lateral driving behavior will be worth considering for modeling in future research.
- The amount of sample input in this study was insufficient, which is reflected in the fact that the problem of overfitting was common in the process of model training and testing and the generalization error was large. Future research will carry out more naturalistic driving data collection to verify the model. At the same time, it is also necessary to carry out multi-scenario testing to study the applicability of the model under multiple scenarios.

References

- Allahviranloo, M., et al. (2013), "Daily activity pattern recognition by using support vector machines with multiple classes", *Transportation Research Part B: Methodological*, Vol. 58, pp. 16-43.
- Ayoub, J., et al (2021), "Modeling dispositional and initial learned trust in automated vehicles with predictability and explainability", *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 77, pp. 102-116.
- Baer, T., et al (2011), "Probabilistic driving style determination by means of a situation-based analysis of the vehicle data", *IEEE International Conference on Intelligent Transportation Systems-ITSC*, Washington, DC.
- Bao, S., et al (2020), "An examination of teen drivers' car-following behavior under naturalistic driving conditions: with and without an advanced driving assistance system", *Accident Analysis & Prevention*, Vol. 147, p. 105762.
- Basso, F., et al (2018), "Real-time crash prediction in an urban expressway using disaggregated data", *Transportation Research Part C: Emerging Technologies*, Vol. 86, pp. 202-219.
- Bellem, H., et al (2016), "Objective metrics of comfort: developing a driving style for highly automated vehicles", *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 41, pp. 45-54.
- Cafiso, S., et al (2020), "Safety effectiveness and performance of lane support systems for driving assistance and automation—experimental test and logistic regression for rare events", *Accident Analysis & Prevention*, Vol. 148 No. 105791.
- Chai, M., et al (2019), "Drowsiness monitoring based on steering wheel status", *Transportation Research Part D: Transport and Environment*, Vol. 66, pp. 95-103.
- Chen, T. and Boost, X.G. (2016), "A scalable tree boosting system", *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785-794.
- Chow, S. (2007), *Sample Size Calculations in Clinical Research*, Taylor & Francis, London.
- Clark, et al. (1993), "The use of neural networks and time series models for short term traffic forecasting: a comparative study", *PTRC 21st Summer Annu. Meet*, pp. 151-162.
- Cortes, C. and Vapnik, V. (1995), "Support-vector networks", *Machine Learning*, Vol. 20 No. 3, pp. 273-297.
- Costa, Á., et al. (1997), "Evaluating public transport efficiency with neural network models", *Transportation Research Part C: Emerging Technologies*, Vol. 5 No. 5, pp. 301-312.
- Costela, F.M., et al. (2020), "Risk prediction model using eye movements during simulated driving with logistic regressions and neural networks", *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 74, pp. 511-521.
- Das, S., et al. (2020), "Vehicle involvements in hydroplaning crashes: applying interpretable machine learning", *Transportation Research Interdisciplinary Perspectives*, Vol. 6 No. 100176.
- Dong, C., et al. (2018), "An innovative approach for traffic crash estimation and prediction on accommodating unobserved heterogeneities", *Transportation Research Part B: Methodological*, Vol. 118, pp. 407-428.
- Elander, J., West, et al. (1993), "Behavioral correlates of individual differences in road-traffic crash risk: an examination of methods and findings", *Psychological Bulletin*, Vol. 113 No. 2, pp. 279-294.
- Evans, L. (1996), "The dominant role of driver behavior in traffic safety", *American Journal of Public Health*, Vol. 86 No. 6, pp. 784-786.
- Farooq, M.U., et al. (2021), "A statistical analysis of the correlates of compliance and defiance of seatbelt use", *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 77, pp. 117-128.
- Ge, J., et al. (2021), "A robust path tracking algorithm for connected and automated vehicles under i-VICS",

- Transportation Research Interdisciplinary Perspectives*, Vol. 9 No. 100314.
- Ghasemzadeh, A., et al. (2018), "Utilizing naturalistic driving data for in-depth analysis of driver lane-keeping behavior in rain: non-parametric MARS and parametric logistic regression modeling approaches", *Transportation Research Part C: Emerging Technologies*, Vol. 90, pp. 379-392.
- Guo, F. and Fang, Y. (2013), "Individual driver risk assessment using naturalistic driving data", *Accident Analysis & Prevention*, Vol. 61, pp. 3-9.
- Ishibashi, M., et al. (2007), "Indices for characterizing driving style and their relevance to car following behavior", *46th SICE Annual Conference*, Kagawa University (Japan).
- Itkonen, T.H., et al. (2020), "Characterisation of motorway driving style using naturalistic driving data", *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 69, pp. 72-79.
- Jasper, S.W., et al. (2018), "Identifying behavioural change among drivers using long Short-Term memory recurrent neural networks", *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 53, pp. 34-49.
- Kondoh, T., et al. (2008), "Identification of visual cues and quantification of drivers' perception of proximity risk to the lead vehicle in car-following situations", *Journal of Mechanical Systems for Transportation and Logistics*, Vol. 1 No. 2, pp. 170-180.
- Kusano, K.D., et al. (2015), "Population distributions of time to collision at brake application during car following from naturalistic driving data", *Journal of Safety Research*, Vol. 54, pp. 95-104.
- Lajunen, T. and Özkan, T. (2011), "Self-report instruments and methods", In: Porter, B.E. (Ed.), *Handbook of Traffic Psychology*, Elsevier, London, pp. 43-59.
- Li, G., et al. (2017), "Estimation of driving style in naturalistic highway traffic using maneuver transition probabilities", *Transportation Research Part C: Emerging Technologies*, Vol. 74, pp. 113-125.
- Liu, K., et al. (2018), "Heterogeneity in the effectiveness of cooperative crossing collision prevention systems", *Transportation Research Part C: Emerging Technologies*, Vol. 87, pp. 1-10.
- Lu, Q.L., et al. (2021), "Exploring the influence of automated driving styles on network efficiency", *Transportation Research Procedia*, Vol. 52 No. 9, pp. 380-387.
- Mahmoud, N., et al. (2021), "Predicting cycle-level traffic movements at signalized intersections using machine learning models", *Transportation Research Part C: Emerging Technologies*, Vol. 124 No. 102930.
- Mensing, F., et al. (2014), "Eco-driving: an economic or ecologic driving style?", *Transportation Research Part C: Emerging Technologies*, Vol. 38, pp. 110-121.
- Mohammadi, R., et al. (2019), "Exploring the impact of foot-by-foot track geometry on the occurrence of rail defects", *Transportation Research Part C: Emerging Technologies*, Vol. 102, pp. 153-172.
- Müller, T., Hajek, H., Radic-Weissenfeld, L. and Bengler, K. (2013), "Can you feel the difference? The just noticeable difference of longitudinal acceleration", *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 57 No. 1, pp. 1219-1223.
- Omrani, H. (2015), "Predicting travel mode of individuals by machine learning", *Transportation Research Procedia*, Vol. 10, pp. 840-849.
- Orlovska, J., et al. (2020), "Effects of the driving context on the usage of automated driver assistance systems (ADAS) - Naturalistic driving study for ADAS evaluation", *Transportation Research Interdisciplinary Perspectives*, Vol. 4, pp. 100093.
- Plöchl, M., et al. (2007), "Driver models in automobile dynamics application", *Vehicle System Dynamics*, Vol. 45 Nos 7/8, pp. 699-741.
- Qi, G., et al. (2019), "Recognizing driving styles based on topic models", *Transportation Research Part D: Transport and Environment*, Vol. 66, pp. 13-22.
- Reiser, C., et al. (2008), "Kundenfahrverhalten im fokus der fahrzeugentwicklung", *Atz - Automobiltechnische Zeitschrift*, Vol. 110 No. 7-8, pp. 684-692.
- Rezaei, M., Saadati, M., et al. (2021), "Gender differences in the use of ADAS technologies: a systematic review", *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 78, pp. 1-15.
- Sagberg, F., et al. (2015), "A review of research on driving styles and road safety", *Hum. Factors*, Vol. 57 No. 7, pp. 1248-1275.
- Siddique, C., et al. (2019), "State-dependent self-adaptive sampling (SAS) method for vehicle trajectory data", *Transportation Research Part C: Emerging Technologies*, Vol. 100, pp. 224-237.
- Simons-Morton, B.G., et al. (2015), "Naturalistic teenage driving study: findings and lessons learned", *Journal of Safety Research*, Vol. 54, pp. 41-48.
- Sun, B., et al. (2017), "Route choice modeling with support vector machine", *Transportation Research Procedia*, Vol. 25, pp. 1806-1814.
- Suzdaleva, E. and Nagy, I. (2018), "An online estimation of driving style using data-dependent pointer model", *Transportation Research Part C: Emerging Technologies*, Vol. 86, pp. 23-36.
- Suzdaleva, E. and Nagy, I. (2018), "An online estimation of driving style using data-dependent pointer model", *Transportation Research Part C: Emerging Technologies*, Vol. 86, pp. 23-36.
- Tang, J., et al. (2020), "Statistical and machine-learning methods for clearance time prediction of road incidents: a methodology review", *Analytic Methods in Accident Research*, Vol. 27 No. 100123.
- Taubman-Ben-Ari, O., et al. (2004), "The multidimensional driving style inventory - scale construct and validation", *Accident Analysis & Prevention*, Vol. 36 No. 3, pp. 323-332.
- Taubman-Ben-Ari, O., et al. (2016), "The multidimensional driving style inventory a decade later: review of the literature and re-evaluation of the scale", *Accident; Analysis and Prevention*, Vol. 93, pp. 179-188.
- Toledo, T., et al. (2008), "In-vehicle data recorders for monitoring and feedback on drivers' behavior", *Transportation Research Part C: Emerging Technologies*, Vol. 16 No. 3, pp. 320-331.
- Toledo, T., et al. (2007), "Integrated driving behavior modeling", *Transportation Research Part C: Emerging Technologies*, Vol. 15 No. 2, pp. 96-112.

- Transportation Research Board, The Highway Capacity Manual (2010), Transportation Research Board, The Highway Capacity Manual, Transportation Research Board of the National Academies, Washington, DC.
- Van Huysduynen et al (2018), "The relation between self-reported driving style and driving behaviour: a simulator study", *Transportation Research Part F: Traffic Psychology & Behaviour*, Vol. 56, pp. 245-255.
- van Huysduynen, H.H., Terken, J. and Eggen, B. (2018), "The relation between self-reported driving style and driving behaviour. A simulator study", *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 56, pp. 245-255.
- Venkata, R.D., et al. (2020), "Variable categories influencing single-vehicle run-off-road crashes and their severity", *Transportation Engineering*, Vol. 2 No. 100038.
- Wang, E., et al. (2017), "Modeling the various merging behaviors at expressway on-Ramp bottlenecks using support vector machine models", *Transportation Research Procedia*, Vol. 25, pp. 1327-1341.
- Wang, J., et al. (2015), "Driving risk assessment using near-crash database through data mining of tree-based model", *Accident Analysis & Prevention*, Vol. 84, pp. 54-64.

- Xiong, H., et al. (2012), "Use patterns among early adopters of adaptive cruise control", *Human Factors: The Journal of the Human Factors and Ergonomics Society*, Vol. 54 No. 5, pp. 722-733.
- Xu, L., et al. (2015), "Establishing Style-Oriented driver models by imitating human driving behaviors", *IEEE Transactions on Intelligent Transportation Systems*, Vol. 16 No. 5, pp. 2522-2530.
- Zargamezhad, S., et al. (2019), "Predicting vehicle fuel consumption in energy distribution companies using ANNs", *Transportation Research Part D: Transport and Environment*, Vol. 74, pp. 174-188.
- Zhang, Z., et al. (2018), "A deep learning approach for detecting traffic accidents from social media data", *Transportation Research Part C: Emerging Technologies*, Vol. 86, pp. 580-596.
- Zhao, X., et al. (2020), "Evaluation of the effect of RPMs in extra-long tunnels based on driving behavior and visual characteristics", *China Journal of Highway and Transport*, Vol. 33, pp. 29-41.

Corresponding author

Chaozhong Wu can be contacted at: chaozhongwu@126.com