RESEARCH ARTICLE

CNN-LSTM Driving Style Classification Model Based on Driver Operation Time Series Data

YINGFENG CAI[1], (Senior Member, IEEE), **RUIDONG ZHAO**[1],
HAI WANG[2,3], (Senior Member, IEEE), **LONG CHEN**[1], **YUBO LIAN**[4], **AND YILIN ZHONG**[4]

[1]Automotive Engineering Research Institute, Jiangsu University, Zhenjiang 212013, China
[2]School of Automotive and Traffic Engineering, Jiangsu University, Zhenjiang 212013, China
[3]Zhenjiang City Jiangsu University Engineering Technology Research Institute, Zhenjiang 212013, China
[4]BYD Auto Industry Company Ltd., Shenzhen 518118, China

Corresponding author: Hai Wang (wanghai1019@163.com)

ABSTRACT This paper aims to establish a driving style recognition method that is highly accurate, fast and generalizable, considering the lack of data types in driving style classification task and the low recognition accuracy of widely used unsupervised clustering algorithms and single convolutional neural network methods. First, we propose a method to collect the information on driver's operation time sequence in view of the imperfect driving data, and then extract the driver's style features through convolutional neural network. Then, for the collected temporal data, the Long Short Term Memory networks (LSTM) module is added to encode and transform the driving features, to achieve the driving style classification. The results show that the accuracy of driving style recognition reaches over 93%, while the speed is improved significantly.

INDEX TERMS Convolutional neural network (CNN), driving style classification, LSTM, neural network, time series data.

I. INTRODUCTION

Advanced driver assistance systems (ADAS) can improve driving comfort and safety, but there are still imperfections in their powerful features that lead to distrust, prejudice and limited vision reliance on ADAS systems by drivers [1] The roots of ADAS at this stage are in 'assistance' and the driver is still the main operator of the vehicle. However, the behavior and driving style of different drivers varies enormously, which places greater demands on the ability to personalize the vehicle in terms of driver tuning and the threshold for triggering ADAS capabilities.

Taking the above into consideration, the individual driving behavior of the driver should be considered while designing the vehicle system, which can have a significant impact on the safety performance of the vehicle. The development of

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

vehicle intelligence also requires the vehicles to be able to adapt to the driver's driving style and provide the appropriate assistance. However, the current development of vehicle parameters using different drivers is not only time-consuming and labor-intensive, but also subjectively influenced by the different driving styles. Therefore, the development of a system that can accurately identify the driving style of the vehicle driver is of great importance for the development of intelligent vehicles.

As of now, there exist many studies focusing on the driving style, and these studies generally rely on three methods, traditional questionnaire-based research methods, visual recognition-based methods, and non-visual driving signal-based research methods. Basically, the mainstream research methods are based on the vehicle driving signals, since the visual-based methods suffer from inherent problems that cannot be eradicated, including invasion of driver privacy etc., and are more influenced by the environmental light.

Deep learning methods have been applied to many aspects of the vehicles, such as YOLOv4-5D [2] is used for the vehicle detection and CenterPoint [3] is used for the LiDAR point cloud processing, both of which have achieved a good accuracy compared to traditional methods. However, deep learning is less commonly used in the processing of temporal data, and accordingly, this paper uses the features of CNNs in feature extraction to obtain the driver's operating characteristics, and then uses LSTM networks to classify the driving styles. Furthermore, the experimental results show that the model used in this paper achieves a high level of accuracy.

The key contributions of this paper are as follows:

To address the problem of insufficient driver operation data in the current driving style recognition dataset, this paper establishes a driving style dataset that contains both driver operation information and vehicle dynamics information. We also propose the idea of fusing accelerator pedal signal and brake pedal signal, to improve the classification accuracy, which is of great value for the research of driving style recognition.

For the temporal features of the data, this paper proposes a method to add an LSTM module after a single CNN network, to learn the before and after temporal information in the driving data stream. This essentially improves the accuracy of recognition, and ensures full utilization of all types of data as well as faster convergence compared with the existing CNN networks. Evident from the detection results of real vehicle data, the CNN-LSTM model exhibits a high generalization ability.

II. RELATED WORK
A. METHODS BASED ON QUESTIONNAIRE
The research on current questionnaire survey methods for driving styles primarily focus on the design of questionnaires, and the widely used such methods are Driver Style Questionnaire DSQ [4] and Multidimensional Driving Style Inventory MDSI [5]. The DSQ has been developed from the questionnaires and principal component analysis, and its validity has been verified by correlations with behavioral indicators as well as its usefulness in the representation of drivers. Since its release in 2004, the Multidimensional Driving Style Inventory MDSI has been applied globally to a diverse sample of drivers, thereby illustrating its usefulness in road safety. In one work, Sergio Useche et al. [6] used MDSI to study different driving styles among the professional drivers, validating the influence of work environment on the key factors of driver driving and expanding the direction of driving style research.

Questionnaire-based survey methods have yielded good results in analyzing the correlation of various parameters during the early stages of driving style research. However, recent advancements in detection technology have enabled the collection of real-time data for driver status and vehicle movement, owing to which the data analysis-based research methods have now become more popular.

B. METHODS BASED ON VISUAL DATA
Vision-based research methods focus on acquiring the driver image data to analyze the driver behavior and extract the corresponding style factors from it. Darnet, proposed by Streiffer et al. [7], collects the driver information and vehicle trajectories through in-vehicle cameras and IMUs on mobile phones, and classifies the driver behaviors with an accuracy of 87.02%.

Additionally, Galarza et al. [8] utilizes a mobile phone with Android operating system to detect the driver's drowsiness, by focusing on the driver's head posture, eye behavior and frequency of hiccups. Moreover, an accuracy of 93.37% was obtained under the natural lighting conditions. In conjunction with the emerging online taxi business, Ma et al. [9] used the information obtained from the camera to analyze the start and end of the taxi-hailing task. Next, the driver's driving style was analyzed using the obtained vehicle driving data. The authors used principal component analysis and K-means clustering to construct a detection model, and the results showed that changes in driving style during different driving tasks differed significantly for the turning, acceleration, and deceleration maneuvers.

Furthermore, driver data obtained using visual methods is heavily influenced by the lighting conditions, and despite the current rapid development of in-vehicle camera technology, variation in lighting conditions can reduce the accuracy, and the corresponding image processing requires greater computational resources, which is contrary to the original intent of vehicle design to reduce costs and maintain accuracy. At the same time, the use of driver detection technologies such as cameras in the vehicle can suffer from the issues such as privacy invasion, making it difficult to be accepted by the vast majority of drivers. In contrast, non-visual based driving signals consume less computational resources and are not affected by the driving environment, but their accuracy of recognition is low at this stage. To bridge this gap, this paper proposes a method to acquire vehicle CAN signals for driver style classification, which avoids the visual invasion of driver's privacy while maintaining a high detection accuracy.

C. METHODS BASED ON NON-VISUAL DRIVING DATA
Manzoni et al. [10] proposes a method for obtaining the driver's driving style using vehicle dynamics data. The method uses inertial and GPS sensors installed in the vehicle to obtain the data on vehicle's movement, in combination with the vehicle dynamics data from ECU, to describe the driver's driving style. They also tested it on a bus, and obtained the desired results. Ly et al. [11] also used the inertial sensors as a source of data, and the relevant analysis revealed that braking and cornering situations are more indicative of an individual's driving style compared to acceleration.

In terms of model selection for the analysis of data, Wang et al. [12] and Xu and Zhu [13] both used Hidden Dirichlet Process (HDP) and Hidden Markov Model (HMM) for feature extraction and classification, respectively.

Besides, Wenshuo Wang used a modified Semi-Hidden Markov Model (HSMM) to enhance the recognition, while Songlin Xu analyzed the driving styles of the identified self-driving cars and performed a risk assessment of the driving environment containing such driving styles. In the modelling of driving styles, Suzdaleva and Nagy [14], [15] has used both single-layer and two-layer pointer models to estimate the mixture parameters and the actual driving style by means of a recursive algorithm under a Bayesian approach. The two-layer pointer adds an internal pointer to the single-layer pointer model, to describe the current driving environment, which the authors classify as 'urban', 'rural' and 'highway'. This approach to differentiate the effects of different driving environments on the drivers is also reflected in the work of Karlsson et al. [16], which shows that the importance of driving attributes is influenced by the changes in driving environment.

Traditional machine learning algorithms are also widely used for the driving style classification tasks. Likewise, Tong et al. [17] used K-means and Gaussian mixture model clustering to obtain three driving styles. Similarly, Li et al. [18] and Mohammadnazar et al. [19] also used an unsupervised clustering algorithm for classifying the driving styles. The unsupervised algorithm obtained the driving style data with clearer boundaries, but lacked a practical theoretical interpretation. To address these drawbacks and speed up the convergence of the model, semi-supervised and supervised models were investigated. Correspondingly, Mingjun et al. [20] used k-means clustering to complete the labelling of data samples, and then used a supervised support vector machine (SVM) with a multiclassification semi-supervised learning algorithm (iMLCU) to construct a recognition model. Wang et al. [21] utilized a semi-supervised support vector machine S3V to obtain a classification model with 10% improvement in the accuracy compared to the baseline, while significantly reducing the labelling of samples. In another work, Chen et al. [22] achieved better results using Latent Diligree Allocation with Labelling (LLDA) to understand the underlying driving style of individuals' driving behavior.

The development of deep learning and artificial neural networks has also brought new research ideas for driver style recognition. Liu et al. [23] used Deep Sparse Autoencoder (DSAE) to extract the hidden features and visualize the driving behavior, where different driving behaviors and driving styles of drivers were represented by converting the features into RGB scale and mapping them on trajectories. Abdennour et al. [24] extracted the data from CAN bus and analyses the driving styles through residual convolutional networks (RCN), thereby eliminating the problem of user privacy invasion. In the work reported by Guo et al. [25], the original labels are obtained by voting on multiple clustering methods, and the classification results obtained by three supervised models are then voted on to derive the corresponding driving styles, which combines the advantages of different models and provides more convincing results.

In terms of recognition method, Shahverdya et al. [26] used neural network model (CNN) for the first time to extract the driver's driving features, then transformed the features into recursive graphs [27], [28], and finally obtained the driver's style type using a fully connected layer. Zhang et al. [29] extends the field of view to the surrounding vehicles, where CNN is applied to obtain the corresponding driving style from the driving data of surrounding vehicles; the final recognition results contribute to the trajectory planning of autonomous vehicles.

In contrast to the simple application of CNNs, some scholars have taken into account the temporal relationship of the data and improved the detection accuracy by adding LSTM networks after CNN networks. Mou et al. [30] used a CNN-LSTM model based on an attention mechanism to identify the driver's stress level. The multimodal data came from the driver's eye data, vehicle dynamic data and driving environment data. The authors concluded from their analysis that eye movement data accounts for a relatively large part of the recognition accuracy, but there are still privacy concerns regarding eye movement detection for drivers. Similarly, Mou et al. [31] used an attention-based multimodal fusion model to detect driver drowsiness by obtaining detection results on a newly created dataset of driver facial and head information. This is essentially a visual detection problem due to the use of image input, while issues remain in terms of privacy invasion. The approach used by Curaet al. [32] avoids the privacy invasion issue. The authors collected driving data from five drivers in a fixed test site environment and compared the performance differences in driver classification between LSTM and CNN networks, respectively. However, the bus they used was a more unique vehicle type compared to the cars most drivers drive, and the type and amount of data collected was less, making it more limited in terms of the level of data coverage.

Existing methods have achieved good results based on non-visual data; however, there still remain many problems that need to be addressed:

1. At this stage, most of the data used for driving style classification comes from perceptual data such as driver facial images and eye movement data from real cars and trajectory data including GPS and on-board three-axis accelerometers. Although these research methods have achieved a high level of accuracy, they involve the addition of many sensors to the vehicle and are susceptible to weather, vibration and other factors, making it difficult to guarantee accuracy in the real-world environment, as well as the risk of privacy violations.

As the direct operator of the vehicle, the driver's operation data can most intuitively reflect his driving style, while the kinematic data of the vehicle can reflect the driver's expectation of the vehicle's motion state. Therefore, this paper therefore proposes a method of collecting driver operation data and vehicle kinematic data to build our dataset, which can be parsed in real vehicles via the CAN protocol, with low
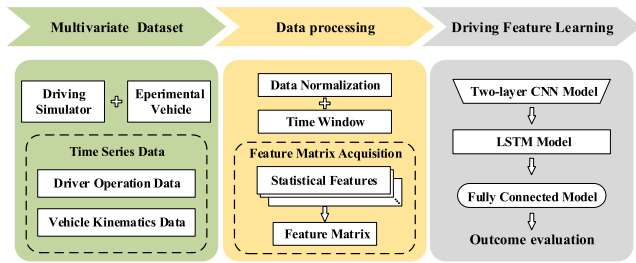
FIGURE 1. The framework of the proposed driving style classification method, including building multivariate datasets, data processing, and driving feature learning.

FIGURE 2. Schematic of the data collection pathway. This includes collecting driving data from driving simulators and real vehicle environments. In the simulator environment, we use CarSim to collect the data of the driver using the Logitech G29 driving simulator; in the real car environment, use CANoe to obtain the driving data.

FIGURE 3. Differences in accelerator pedal opening for different driver styles in the same driving scenario (vehicle starting from a standstill).

acquisition difficulty and high stability and accuracy in terms of style classification.

2. Previous studies have used a small amount of data and there is a large influence of different road environments on driver style factors in real life scenarios. This paper therefore uses a combination of simulator data and real car data to collect data. Compared to the real driving environment, the different driving scenarios in the simulated driving environment have less influence on the driver and more variation between the different styles. Therefore this paper uses most of the simulator data with a small proportion of the real scenario data when training the network and uses the real scenario data for evaluation at the end.

3. At this stage, common driving style classification algorithms use convolutional neural networks to perform full concatenation after feature extraction is complete, ignoring the temporal and pre-post correlation in the data. Given the temporal nature of the data, LSTM networks alone have also been used for driving style classification, but ignoring the extraction of features from the data. Because features in real-world driving situations are not the same, such as sudden acceleration and deceleration and large angle steering, these features occur with probabilities independent of each other, and a simple fully connected or temporal network would have the opposite effect on driving style recognition.

To address these issues, this paper combines CNN networks with LSTM networks, using CNN networks to extract data features and then LSTM networks to obtain the contextual relationships in the temporal data stream. The combination of the two networks improves the accuracy of driving style detection and demonstrates greater robustness in the dataset in real scenarios.

III. PROPOSED METHOD

Fig. 1 shows the framework of the driving style recognition method proposed in this paper. In this work, we first use the vehicle driving data obtained from the driving simulator with real car CANoe parsing, then process the data in segments and input the finished data into CNN-LSTM model for classification, and finally obtain the driving style corresponding to the current data.
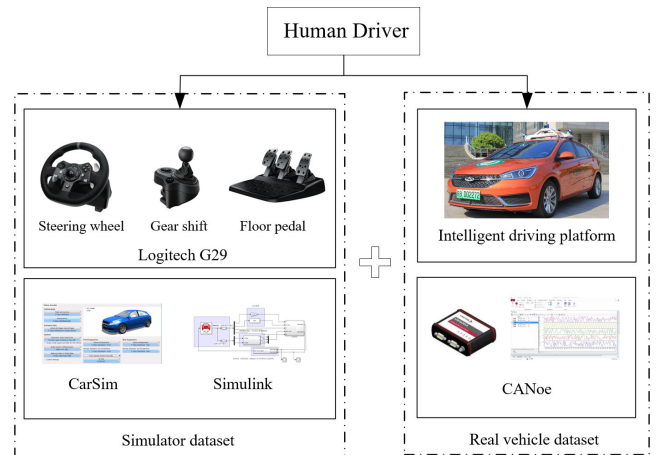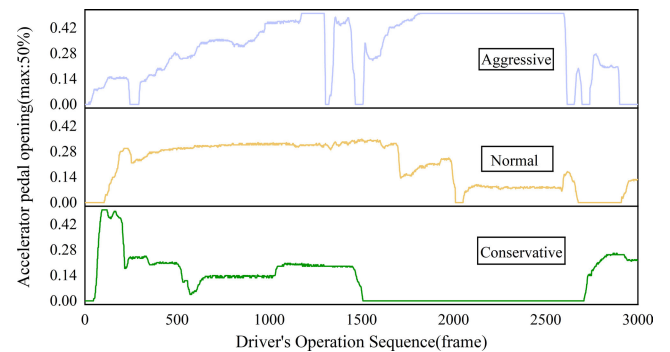
A. DATA TYPES AND ACQUISITION

Unlike previous studies, this work collects both simulator data and real car data to form a dataset. Compared to real car, simulator can build a comprehensive driving environment that fully simulates the driver's complex working conditions on city roads, at high speeds, etc. At the same time, the driver's surroundings are relatively fixed, which reduces the influence of environment on the driving style. However, the use of simulators alone limits the comprehensiveness of the data, thus this paper employs a fusion of driving simulator and real car data, where the simulator data is used as test and validation set for network training, and the real car data is used to evaluate the generalization performance after the network training is completed.

As shown in Fig. 2, the simulation data acquisition is based on the Logitech G29 driving simulator with the driver operating the steering wheel, pedals, and gear lever. For the test scenario, a circular map was first drawn based on the Carsim platform, and then the input driving signals were collected using MATLAB and transferred to Carsim for the control of experimental vehicle. The maps were drawn for a

TABLE 1. Data attributes and definitions.

Attribute name	Explanations	Sampling frequency(Hz)
Driver_ID	Driver ID number	100
Time	The time elapsed to collect data	100
Str_Angle	Steering wheel angle	50
Str_AngleSpd	Steering wheel speed	50
Accelerator_O	Accelerator pedal opening	100
Brake_O	Brake pedal opening	100
ω	Vehicle yaw rate	100
u_x	Longitudinal speed of the vehicle	100
u_y	Lateral speed of the vehicle	100
a_x	Longitudinal acceleration of the vehicle	100
a_y	Lateral acceleration of the vehicle	100
I	Combination of Accelerator_O and Brake_O	100

TABLE 2. Driving style clustering results.

Style categories / Algorithms	Aggressive	Normal	Conservative
FCM	14	31	15
SC	13	32	15

2km city road, a 3km country road, and a 6km highway road. The test recorded 60 drivers driving on the road, and to avoid unfamiliarity with the equipment affecting the results, the drivers first drove for 5mins on a simulated road to familiarize themselves with the environment. Moreover, the type of data collected and the sampling frequency were consistent with the real car data.

To verify the generalization ability of the model, this paper collects real vehicle data based on the Chery Arizawa 5e intelligent driving test platform. The operational data of six drivers in a real driving environment, the real vehicle dynamics data and the vehicle trajectory were recorded using the on-board CANoe and inertial guidance system.

Considering the type of driving data included in the dataset, the current research approaches focus on the vehicle trajectory data and vehicle dynamics data, and no research has yet been conducted by directly using the driver operation data. The steering wheel angle, brake pedal opening and accelerator pedal opening, which are the building blocks of driving directly controlled by human drivers while driving a vehicle, are the most direct indicators of the differences between the drivers' operations, that is, the driving styles of different drivers. As shown in Fig. 3, there is a large difference in throttle opening between the three styles of drivers in the same driving scenario, with the more aggressive drivers tending to keep the accelerator pedal open more when operating. Therefore, to improve the accuracy of driver style classification, we focused on four signals of driver operation, namely steering wheel angle, steering wheel speed, accelerator pedal opening and brake pedal opening. These signals are supplemented by vehicle dynamics data such as the angular velocity of vehicle's traverse and the longitudinal velocity acceleration, to describe the driver's driving style. The specific data attributes and explanations are provided in Table 1 below. In real life conditions, this data can be resolved via the CAN protocol, and no additional sensors are required.

When analyzing the acquired signals, we found that using the accelerator pedal signal or the brake pedal signal alone

can have a cut-off that affects the representation of the driver's behavior. For example, when driver operates the brake pedal to decelerate, the accelerator pedal has an opening of 0, but the brake pedal has a signal value. Hence, the accelerator pedal does not respond well to the driver's acceleration expectations. In this paper, the accelerator and brake pedal signals are combined to form a single signal, by adding the brake pedal open value of current moment to the accelerator pedal signal at the same moment as a negative value. In this way, two signals can be combined to characterize the driver's acceleration request in the longitudinal direction of the vehicle, expressed as signal "I". The practicality of this idea will also be demonstrated while analyzing the experimental results later.

B. DATA PROCESSING
Considering the difference in dimension of different data, we normalized the data. We project different data into the same fixed interval according to the threshold range of the data.

For the collected data, this paper uses fuzzy c-means algorithm (FCM) and spectral clustering (SC), to obtain the corresponding driving style labels of drivers. Firstly, using FCM and SC, the number of categories for clustering is set as k, the fuzzy mean clustering result as $C_f(i)$, the clustering center as m_j, and the clustering result for spectral clustering as $C_s(i)$. Meanwhile, i is the code name of the sample. For sample i, if $C_f(i) = C_s(i)$ in the case of FCM and SC clustering, the sample i is labelled with its type label and divided into the labelled data set X_i. If the two clustering results differ, then the driver's corresponding driving style label is obtained based on the subjective evaluation of the driver.

The results of driver styles obtained after FCM and spectral clustering are shown in the Table 2. As can be seen from the table, using FCM, the number of samples classified as normal style is 31, the number of samples classified as aggressive style is 14, and the number of samples classified as conservative is 15. The two clustering algorithms disagree on the definition of the style of driver 25, and combining our subjective assessment with the objective ride experience, we define the style of this driver as aggressive, so we use the results of FCM as the driving style label for the classification.

The traditional clustering methods have poor real-time performance in driver detection, and the center of clustering is prone to shifting when subjected to uneven samples. Accordingly, this paper trains a network model for driver driving style recognition, the results of which are not affected by
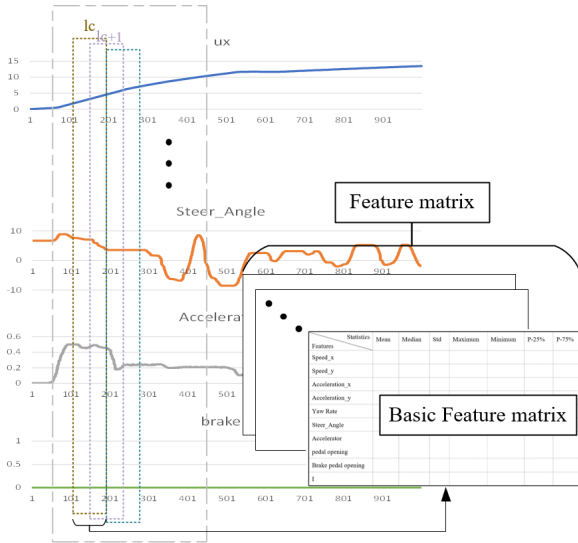
FIGURE 4. The schematic diagram of the sliding window. The larger dashed box represents the entire feature matrix, which is composed of smaller feature matrices.

FIGURE 5. CNN-LSTM structure diagram. It consists of the input feature matrix, the CNN network for feature extraction, the LSTM network for learning the timing features of driver operations, and the final fully connected network.

the overall bias of the input information, once the network is trained. Driving style, as a broad feature contained in a multidimensional driving signal, is difficult to represent clearly in terms of a specific signal type. Therefore, the collected raw data stream needs to be processed, and it is a common approach to split the data, count the statistical parameters of a particular piece of data in the data segment, and then obtain the driving style corresponding to that data from the statistical parameters of the overall data. This work also follows the same approach.

Based on the raw data, as shown in Fig. 4, this paper uses the concept of contextual window to focus the attention on a fixed time window. This time window lc is fixed at 8. In order to avoid sudden changes in the values due to sensor errors and to be able to describe the differences between the drivers, the statistical values of the data in this time window need to be calculated, which mainly include the mean, median, standard deviation, maximum, minimum, and 25% and 75% quartiles for a total of seven parameters. After the calculation is completed, the time window lc is moved backwards on the time axis by four units to obtain the next lc, for which the statistical values are then calculated, and so on. To avoid missing the style features due to small size of lc, a matrix of 128 lc's is used as a source of minimum features for the driver's driving style, correspondingly, la is 516 frames and 5.16 seconds.

In this paper, small windows are used to refine the specific numerical sizes, and a large range of feature matrices consisting of small windows are used to capture the style features contained in the transient changing numerical features. The double time window approach allows for the microscopic counting and transformation of features for transient driving behavior, and conveys the information about the driver's driving style at a macroscopic level. However, this information exists only as the statistical information in the matrix, and
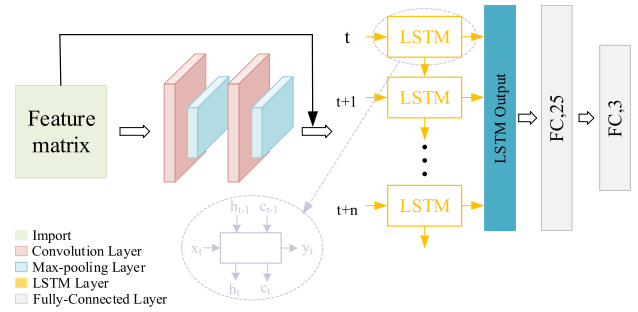
is not expressed in the form of data for a particular feature. Hence, the deep learning network is expected to learn and express the driver's driving style characteristics in a broad sense from this matrix input.

C. NETWORK STRUCTURE

Fig. 5 shows the schematic diagram of CNN-LSTM structure. Firstly, for the processed data, a convolutional neural network is used to extract the driving features in the driving data, such as abrupt acceleration and faster steering wheel speed, and pool the output to a smaller size. Subsequently, an LSTM network is used to obtain the dependencies between the driving features, which are then converted into the output labeled with a specific driving style. The input part of the network is the feature matrix of size $7Q \times la$ generated in the previous section, where Q is the type of data input (ten data types in total, as indicated in Table 1), and la is used as the time axis for this data matrix.

Convolutional neural networks are widely used in areas such as image processing and natural language processing, and can automatically learn the deep features from lower-level data structures, while different layers of the network can learn features at different levels. The shallow network layer has a small perceptual area and can learn features in the local domain of the input data (e.g., sudden changes in acceleration, and large angular velocities of transverse pendulums). On the other hand, the deeper network layer has a larger perceptual area and can better learn the abstract features in the input data (e.g., radical lane changing behavior, etc.). Convolutional neural networks are therefore a suitable choice for extracting such features of the driving behavior.

The convolutional neural network used in this paper consists of two convolutional layers and two maximum pooling layers, each of which is followed by a convolutional layer for down sampling. The feature matrix generated by the second pooling layer is superimposed to be used as the input for the next stage of the network. Since convolution on driving features is not practical, this paper proposes to apply one-dimensional convolution on the time axis. The kernel size of the first layer of convolution kernels is therefore set to $7Q \times 5$,

for a total of 32 convolution kernels. The second layer uses 64 convolutional kernels of size 1×3. The step size of both layers of the network is set to 1. Moreover, the activation functions of both convolutional layers are rectified linear units. The pooling operations after both layers are performed on the feature axes, with each pool of size 8×1 and a step size of 1. To maintain the same size of time axis in both convolutional and pooling layers, we use zero padding, where the dropout probability is set to 50%.

Previous studies [26], [29] confirmed the feasibility of convolutional neural networks for driving style classification tasks, however they suggest to enter the fully connected layer after CNN has finished extracting the features, whereas the driver's operating features should not be simply superimposed to obtain the driver's driving style. The occurrence of each different parameter feature in the time dimension of the actual driver driving the vehicle is stochastic, and full connectivity may lead to a wrong understanding of the driving style. Essentially, the driver operation data and the vehicle dynamics data have strong temporal characteristics, and thus, this paper uses an LSTM module to encode the dependencies between the driving features, and outputs the predicted driving style from the final fully-connected layer.

The lower half of Fig. 5 shows the sketch of LSTM structure, where x_t is the input at current moment and y_t is the output at current moment. Unlike the fully connected layer, the information passed between the LSTM modules is h_t and c_t. Here, h_t is the output of previous time step that is transferred to the next cell, enabling the transfer of information from the previous time step. Moreover, c_t is the state of cell in the previous time step, which affects the output and state in the next time step. Each LSTM cell controls the transmission state through the gating state present in it, remembering what needs to be remembered for a long time, e.g., the driver's driving style, while forgetting the unimportant information, e.g., changes in the driving information over a short period of time. The LSTM model is therefore better suited for the task of driving style classification, as opposed to the fully connected model.

The output of the convolutional network is stacked and concatenated with the initial data feature matrix, which together serve as the input to the LSTM network. The LSTM network has a hidden layer of 100 neurons, but for simple classification tasks, increasing the number of layers would make the training results worse.

In this paper, a two-layer fully connected network is used after the LSTM network, to get the final output. The first fully connected layer uses 25 neurons while the second layer uses 3 neurons and outputs the predicted probability values for different driving styles, using the SoftMax function.

To investigate the LSTM module's ability to understand the temporal data, this paper uses a combination of CNN and four-layer fully connected network as a benchmark to illustrate the superiority of the CNN-LSTM model. The kernel size of the first layer of convolutional kernels is set to
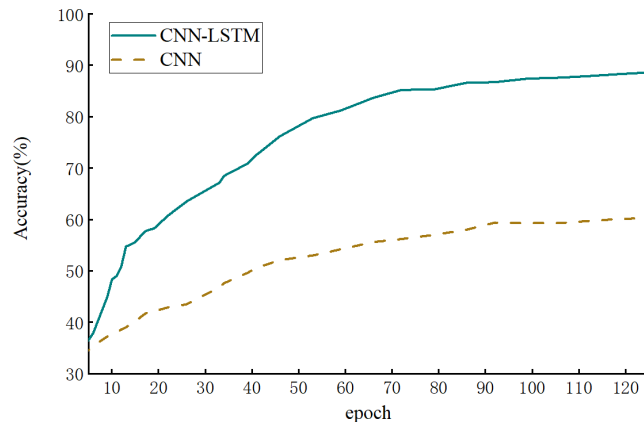

FIGURE 6. Model accuracy versus training epoch. The cyan solid line represents the change of the detection accuracy of the CNN-LSTM network with the training process, and the brown dotted line represents the change of the CNN network detection with the training process.

6Q×5, with a total of 64 convolutional kernels. Meanwhile, the second layer uses 128 convolutional kernels with a size of 6×15, and the third layer uses 64 convolutional kernels with a size of 1×3. There are 4 fully connected layers, with 50 neurons in the first layer, 128 neurons in the second layer, 32 neurons in the third layer, and 3 neurons in the final layer. Finally, the corresponding predicted probability of driving style is output using the SoftMax activation function.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we evaluate the stability of the proposed driving style recognition method, which detects a total of three different driving styles, namely aggressive, normal line, and conservative. For our own dataset, we first use the data collected by the driving simulator to train the network, then adjust the network parameters and obtain the recognition accuracy of the model, and lastly use the data collected from real cars to verify the accuracy and generalization ability of the model. In the network training, 80% of the data is used as the training set, while the rest 20% as the test set.

A. EXPERIMENTAL SETUP

The experiments are performed on a workstation with Intel i9 9900k CPU and Nvidia 1080ti GPU. The experimental environment is Ubuntu 18.04, and the network is built based on the TensorFlow open-source code library.

For the CNN part of the training process, Adam optimizer with a learning rate of 0.05, a decay rate of 1e-6, a batch size of 128 and 500 iterations was used. In addition, for the LSTM part, we used RMSProp optimizer with a learning rate of 1e-6 and $\rho = 0.9$. For the comparison model, Adam optimizer with a learning rate of 0.001 and a batch size of 16 was chosen for the network training, and the cross-entropy function was used as the loss function.

B. TRAINING RESULTS BASED ON SIMULATOR DATA

The Figure 6 shows the accuracy of the networks as a function of the number of training epochs. As seen from the figure, the

TABLE 3. Analysis of the effect of data type on the results.

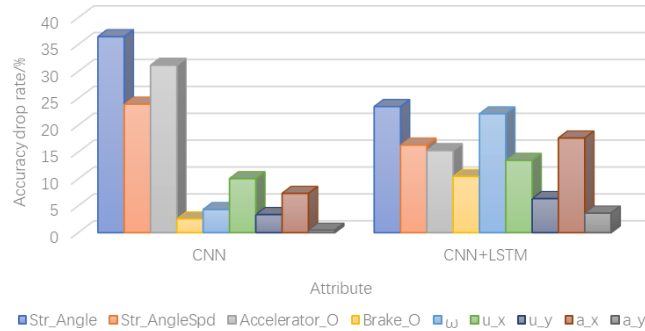| Type of data	Driver Operation Data	Vehicle Kinematics Data	Combination of the two data
CNN-LSTM	78.64%	60.32%	92.43%
CNN	67.39%	53.76%	77.2%

FIGURE 7. Model versus data relationship diagram. The figure shows the decline rate of network recognition accuracy caused by removing a certain driving parameter. The left side is the CNN network, and the right side is the CNN-LSTM network. The driving parameters used include steering wheel angle and rotational speed, accelerator pedal opening, brake pedal opening, vehicle yaw rate, total vehicle speed, lateral vehicle speed, longitudinal acceleration, and lateral acceleration.

CNN+LSTM model starts to converge after about the 73rd epoch, with higher accuracy and faster convergence than the CNN model.

1) COMPARATIVE TESTS FOR DIFFERENT TYPES OF DATA

To investigate the effect of adding the LSTM module on the temporal data, the reduction in recognition accuracy after removing each data item is compared. As shown in Figure 7, the addition of LSTM module allows the model to use the data more comprehensively, compared to the single CNN model. Additionally, the convolutional neural network relies more on the driver's operating parameters, particularly the steering wheel angle, which consequently causes a 36.47% reduction in the accuracy after its removal. On the other hand, the kinematic parameters of the vehicle do not affect the accuracy by more than 10%. The addition of LSTM module can thus increase the stability of the results, as it is important to combine more comprehensive information and reduce the reliance on a single piece of information for the complex task of assessing driver driving style.

In contrast to previous studies, we proposed the idea of collecting the driver operating data directly, and therefore needed to conduct the ablation experiments to assess the effect of different data types on the accuracy of driving style recognition. The vehicle kinematic data includes the angular and longitudinal velocities of the vehicle, and the acceleration data. The cross-sectional comparison shows that the use of driving operation data yields more accurate results than using the vehicle kinematic data, which thereby reflects

TABLE 4. Comparison of accuracy of different models.

Models	CNN-LSTM+I	CNN-LSTM	CNN
Train accuracy	94.37%	95.61%	80.63%
Test accuracy	93.81%	92.43%	77.2%
Inference time	13.599cs	19.993cs	9.596cs

that the driver's driving style is more hidden in the parameters of his/her direct operation of the vehicle. Besides, the longitudinal comparison also confirms that the addition of LSTM module ensures a better processing of the temporal data.

2) MODEL TRAINING RESULTS

Table 3 shows the final model training results, where different network parameters were adjusted to achieve the highest accuracy for both models. The results of the test set show that the CNN+LSTM model is more accurate than the CNN-only model, but it takes more time for training, owing to its more complex network structure. For the accelerator pedal and brake pedal fusion method proposed in 3.2, we verify its effectiveness through comparative experiments. The term I in the second column of the table below indicates that the two signals are fused together. Compared to the method without fusion, the classification accuracy of CNN-LSTM + I increased by 1.38%. It is also found that the training accuracy of the CNN-LSTM model is higher when the fusion method is not used; however the corresponding detection accuracy decreases, and the use of the fusion method improves the performance of the model.

C. TEST RESULTS BASED ON REAL VEHICLE DATA

The driving style of a driver in a real driving environment is generally different from that in the simulator environment, due to current weather conditions, occupants of the vehicle and so on, and the style classification model in fact needs to serve the actual vehicle. Therefore, to verify the generalization ability of the network, the accuracy was also calculated by inputting the collected real vehicle data into the network. It can be seen that the CNN+LSTM model experiences a drop in accuracy of about 5%, indicating that the driving simulator is slightly lacking in simulating the real driving scenarios. In addition, the CNN model suffers from a more significant drop in the accuracy of about 10%. That is to say,

TABLE 5. Comparison of accuracy of live vehicle simulators.

Data source	Driving simulator	Real vehicles
Accuracy- [CNN-LSTM+I]	93.81%	89.04%
Accuracy- [CNN]	77.2%	66.37%

the composite structure of CNN+LSTM has better robustness against the realistic data sets, and is applicable to a wider range of data.

V. CONCLUSION

In this paper, a driving style classification method based on driver operating signals and vehicle dynamics is presented. Driving data from different road conditions and different drivers in a simulator environment is collected in this paper, and then the driving style labels are obtained using a combination of unsupervised clustering and voting methods. A CNN+LSTM network was then trained using the labels and driving data, to realize the detection of driving styles. In the examination of real car data, the network proposed in this paper demonstrates high generalization ability, along with the advantages of low cost and high efficiency.

Finally, it is proposed that the driver of the car and the surrounding vehicles can be signaled to plan the driving route in time to improve efficiency. Future work includes optimizing the network structure to improve recognition accuracy and generalization of detection capabilities in different driving scenarios, analyzing the driving style of the corresponding driver by collecting form data from surrounding vehicles through the sensing system and incorporating driving style into the vehicle's ADAS functions to improve driver acceptance.

REFERENCES

[1] P. M. Greenwood, J. K. Lenneman, and C. L. Baldwin, "Advanced driver assistance systems (ADAS): Demographics, preferred sources of information, and accuracy of ADAS knowledge," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 86, pp. 131–150, Apr. 2022.

[2] Y. Cai, T. Luan, H. Gao, H. Wang, L. Chen, Y. Li, M. A. Sotelo, and Z. Li, "YOLOv4–5D: An effective and efficient object detector for autonomous driving," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–13, 2021.

[3] T. Yin, X. Zhou, and P. Krähenbühl, "Center-based 3D object detection and tracking," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 11779–11788.

[4] M. Ishibashi, M. Okuwa, S. Doi, and M. Akamatsu, "Indices for characterizing driving style and their relevance to car following behavior," in *Proc. SICE Annu. Conf.*, Sep. 2007, pp. 1132–1137.

[5] T. B. A. Orit, M. Mario, and G. Omri, "The multidimensional driving style inventory-scale construct and validation," *Accident Anal. Prevention*, vol. 36, no. 3, pp. 323–332, May 2004.

[6] S. A. Useche, B. Cendales, F. Alonso, J. C. Pastor, and L. Montoro, "Validation of the multidimensional driving style inventory (MDSI) in professional drivers: How does it work in transportation workers?" *Transp. Res. F, Traffic Psychol. Behav.*, vol. 67, pp. 155–163, Nov. 2019.

[7] C. Streiffer, R. Raghavendra, T. Benson, and M. Srivatsa, "DarNet: A deep learning solution for distracted driving detection," Presented at the 18th ACM/IFIP/USENIX Middleware Conf., Ind. Track, Las Vegas, NV, USA, 2017, doi: 10.1145/3154448.3154452.

[8] E. E. Galarza, F. D. Egas, F. M. Silva, P. M. Velasco, and E. D. Galarza, "Real time driver drowsiness detection based on driver's face image behavior using a system of human computer interaction implemented in a smartphone," in *Proc. Int. Conf. Inf. Technol. Syst. (ICITS)*. Cham, Switzerland: Springer, 2018, pp. 563–572.

[9] Y. Ma, W. Li, K. Tang, Z. Zhang, and S. Chen, "Driving style recognition and comparisons among driving tasks based on driver behavior in the online car-hailing industry," *Accident Anal. Prevention*, vol. 154, May 2021, Art. no. 106096.

[10] V. Manzoni, A. Corti, P. De Luca, and S. M. Savaresi, "Driving style estimation via inertial measurements," in *Proc. 13th Int. IEEE Conf. Intell. Transp. Syst.*, Sep. 2010, pp. 777–782.

[11] M. Van Ly, S. Martin, and M. M. Trivedi, "Driver classification and driving style recognition using inertial sensors," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2013, pp. 1040–1045.

[12] W. Wang, J. Xi, and D. Zhao, "Driving style analysis using primitive driving patterns with Bayesian nonparametric approaches," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 8, pp. 2986–2998, Aug. 2019.

[13] S. Xu and J. Zhu, "Estimating risk levels of driving scenarios through analysis of driving styles for autonomous vehicles," 2019, *arXiv:1904.10176*. Accessed: Apr. 1, 2019.

[14] E. Suzdaleva and I. Nagy, "An online estimation of driving style using data-dependent pointer model," *Transp. Res. C, Emerg. Technol.*, vol. 86, pp. 23–36, Jan. 2018.

[15] E. Suzdaleva and I. Nagy, "Two-layer pointer model of driving style depending on the driving environment," *Transp. Res. B, Methodol.*, vol. 128, pp. 254–270, Oct. 2019.

[16] F. Ekman, M. Johansson, M. Karlsson, H. Strömberg, and L.-O. Bligård, "Trust in what? Exploring the interdependency between an automated vehicle's driving style and traffic situations," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 76, pp. 59–71, Jan. 2021.

[17] L. Tong, F. Rui, Z. Mingfang, and T. Shun, "Study on driving style clustering based on *K*-means and Gaussian mixture model," *China Saf. Sci. J.*, vol. 29, no. 12, pp. 40–45, 2019.

[18] G. Li, Y. Chen, D. Cao, X. Qu, B. Cheng, and K. Li, "Extraction of descriptive driving patterns from driving data using unsupervised algorithms," *Mech. Syst. Signal Process.*, vol. 156, Jul. 2021, Art. no. 107589.

[19] A. Mohammadnazar, R. Arvin, and A. J. Khattak, "Classifying travelers' driving style using basic safety messages generated by connected vehicles: Application of unsupervised machine learning," *Transp. Res. C, Emerg. Technol.*, vol. 122, Jan. 2021, Art. no. 102917.

[20] L. Mingjun, Z. Zhenghao, S. Xiaolin, C. Haotian, and Y. Binlin, "Driving style classification model based on a multi-label semi-supervised learning algorithm," *J. Hunan Univ., Natural Sci.*, vol. 47, no. 4, pp. 10–15, 2020.

[21] W. Wang, J. Xi, A. Chong, and L. Li, "Driving style classification using a semisupervised support vector machine," *IEEE Trans. Human-Mach. Syst.*, vol. 47, no. 5, pp. 650–660, Oct. 2017.

[22] D. Chen, Z. Chen, Y. Zhang, X. Qu, M. Zhang, and C. Wu, "Driving style recognition under connected circumstance using a supervised hierarchical Bayesian model," *J. Adv. Transp.*, vol. 2021, Jun. 2021, Art. no. 6687378.

[23] H. Liu, T. Taniguchi, Y. Tanaka, K. Takenaka, and T. Bando, "Visualization of driving behavior based on hidden feature extraction by using deep learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 9, pp. 2477–2489, Sep. 2017.

[24] N. Abdennour, T. Ouni, and N. B. Amor, "Driver identification using only the CAN-bus vehicle data through an RCN deep learning approach," *Robot. Auto. Syst.*, vol. 136, Feb. 2021, Art. no. 103707.

[25] Y. Guo, X. Wang, Y. Huang, and L. Xu, "Collaborative driving style classification method enabled by majority voting ensemble learning for enhancing classification performance," *PLoS ONE*, vol. 16, no. 7, Jul. 2021, Art. no. e0254047.

[26] M. Shahverdy, M. Fathy, R. Berangi, and M. Sabokrou, "Driver behavior detection and classification using deep convolutional neural networks," *Expert Syst. Appl.*, vol. 149, Jul. 2020, Art. no. 113240.

[27] S. Spiegel, "Discovery of driving behavior patterns," in *Smart Information Systems: Computational Intelligence for Real-Life Applications*, F. Hopfgartner, Ed. Cham, Switzerland: Springer, 2015, pp. 315–343.

[28] S. Spiegel, J.-B. Jain, and S. Albayrak, "A recurrence plot-based distance measure," in *Translational Recurrences*. Cham, Switzerland: Springer, 2014, pp. 1–15.

[29] Y. Zhang, Q. Xu, J. Wang, K. Wu, Z. Zheng, and K. Lu, "A learning-based discretionary lane-change decision-making model with driving style awareness," 2020, *arXiv:2010.09533*.

[30] L. Mou, C. Zhou, P. Zhao, B. Nakisa, M. N. Rastgoo, R. Jain, and W. Gao, "Driver stress detection via multimodal fusion using attention-based CNN-LSTM," *Expert Syst. Appl.*, vol. 173, Jul. 2021, Art. no. 114693.

[31] L. Mou, C. Zhou, P. Xie, P. Zhao, R. Jain, W. Gao, and B. Yin, "Isotropic self-supervised learning for driver drowsiness detection with attention-based multimodal fusion," *IEEE Trans. Multimedia*, vol. 25, pp. 529–542, 2023.

[32] A. Cura, H. Küçük, E. Ergen, and I. B. Öksüzoglu, "Driver profiling using long short term memory (LSTM) and convolutional neural network (CNN) methods," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 10, pp. 6572–6582, Oct. 2021.

HAI WANG (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees from the School of Instrument Science and Engineering, Southeast University, Nanjing, China.

In 2012, he joined the School of Automotive and Traffic Engineering, Jiangsu University, where he is currently a Professor. He has published more than 50 articles in the field of machine vision-based environment sensing for intelligent vehicles. His research interests include computer vision, intelligent transportation systems, and intelligent vehicles.

YINGFENG CAI (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees from the School of Instrument Science and Engineering, Southeast University, Nanjing, China. In 2013, she joined the Automotive Engineering Research Institute, Jiangsu University, where she is currently a Professor. Her research interests include computer vision, intelligent transportation systems, and intelligent automobiles.

LONG CHEN received the Ph.D. degree in vehicle engineering from Jiangsu University, Zhenjiang, China, in 2002.

His research interests include intelligent automobiles and vehicle control systems.

YUBO LIAN received the M.S. degree from Nanjing University, Zhenjiang, China, in 2000.

His research interest includes intelligent connected vehicles.

RUIDONG ZHAO received the B.S. degree from Jiangsu University, Zhenjiang, China, where he is currently pursuing the M.S. degree.

His research interests include reinforcement learning and control of intelligent vehicles.

YILIN ZHONG received the Ph.D. degree in vehicle engineering from Jiangsu University, Zhenjiang, China, in 2010.

His research interests include intelligent automobiles and vehicle control systems.

· · ·