IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 18 July 2024, accepted 11 August 2024, date of publication 28 August 2024, date of current version 10 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3450935

RESEARCH ARTICLE

A Hierarchical Network-Based Method for Predicting Driver Traffic Violations

MINGZE WANG[1,2] **AND NAIWEN LI**[1]
[1]School of Business Management, Liaoning Technical University, Huludao 125000, China
[2]School of Business, Nanning University, Nanning 530000, China

Corresponding author: Mingze Wang (wangmingze@unn.edu.cn)

This work was supported by the Basic Research Project of Liaoning Provincial Department of Education under Grant LN2020JCL029.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institute of Science and Technology, Liaoning Technical University, and performed in line with the Declaration of Helsinki.

ABSTRACT Through the cleaning and filtering of real data from driving studies, a traffic violation dataset was constructed in order to determine the relationship between driver factors and the occurrence of traffic violations. Driver factors were analyzed using the indicator significance method and a multidimensional indicator set for predicting driver traffic violations was created. On this basis, we propose a hierarchical network-based method for predicting driver traffic violations. First, time sequences data preprocessing is performed to the multidimensional input data and a framework is presented for analyzing traffic violation data using the convolutional neural network(CNN) and long short term memory network (LSTM). CNN is used to obtain the time sequences of factors related to traffic violations and LSTM is adopted to acquire the temporal characteristics of these sequences, thus completing the initial calibration of the relationship between driver factors and the occurrence of traffic violations. Finally, the types of driver traffic violations are predicted using an improved attention network that adds two plug-and-play modules, including spatio-temporal interaction and deep convolutional feature extraction, which accomplishes the self-learning recalibration function of the indicator weights for calculating the probability of the occurrence of a certain traffic violation in the future period. On the traffic violation dataset, the proposed method was evaluated, and it increases prediction accuracy when compared to non-hierarchical methods and existing joint methods, thus contributing to the synergy of smart connected vehicle systems.

INDEX TERMS Traffic violations, hierarchical network, spatio-temporal interaction, deep convolutional feature extraction, prediction accuracy.

I. INTRODUCTION

Road traffic accidents has become a major global problem affecting traffic safety, and the number of injuries and deaths due to this cause remains chronically high. China faces even more serious traffic safety problems. According to the National Bureau of Statistics of China [1], the number of road traffic accidents has been increasing year by year, and the three absolute indicators of deaths, injuries and economic losses are still at a high level, as shown in Table 1. Therefore, technical methods are needed to reduce the number of accidents and avoid road traffic accidents with serious consequences.

TABLE 1. Indicators of road traffic accidents in china.

Year	Number of accidents	Number of injuries	Economic loss/RM B10K	Number of deaths
2022	250409	263621	123926	60676
2021	273098	281447	145036	62218
2020	244674	250723	131361	61703
2019	247646	256101	134618	62763
2018	244937	258532	138456	63194

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

© 2024 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

Numerous researchers and scholars at home and abroad have conducted a series of studies on the causes of road traffic accidents, and have found that the factors affecting traffic accidents include human, vehicles, roads and the environment, as well as the synergistic effect of the above four aspects, with the human factor being particularly important compared with other factors. More than 90 percent of road traffic accidents are caused by drivers violating traffic laws. For instance, 50 percent of fatal road traffic accidents in Europe and 75 percent of fatal road traffic accidents in China are caused mainly by driver traffic violations [2].

The characterization and evolution of driver traffic violation is a hotspot in the field of traffic safety [3]. Accurate prediction of traffic violations can enhance the synergy of the current smart connected vehicle systems [4] and at the same time, it can promptly alert drivers to potential violations, thus reducing the probability of traffic accidents, protecting the personal safety of traffic participants, and reducing the loss of property. With the advancement of fundamental research on driver traffic violations as well as the development and iteration of artificial intelligence and deep learning theories, the study of traffic violation prediction models and methods has opened up new potential [5], [6], [7], [8]. The traditional mathematical model based traffic violations are general motor model [9], stimulus response model [10], collision avoidance model [11] and Psycho physical model [12]. With the rise of machine learning techniques, SVM [13], decision tree [14], k-means [15], hidden markov model [16], generative adversarial networks [17], ant colony optimization [18] are gradually applied in the field of traffic accident data analysis and driver behavior recognition. In addition, the rapid development of traffic-related video devices and the application of computer vision methods, such as Mixture of Gaussian [19], Convolutional Neural Networks [20] and Yolo [21] are capable of recognizing driver relevant driving behaviors and thus predicting the driver's intention. However, the diversity and variability of drivers in terms of age, gender, driving experience and personality leads to the following challenges in predicting traffic violations. (1) The difficulty of collecting real data on traffic violations and the small number of trainable samples lead to a lack of consistency in the results of the constructed prediction models. Meanwhile, there are some other methods to collect datasets, such as questionnaires, simulations, accident information data, natural driving data, in- and out-of-vehicle videos, etc., which suffer from subjectivity, lack of authenticity and inaccessibility due to data privacy reasons. (2) The selection of indicators for traffic violation prediction is subjective, and the subjectivity of the participants in the data collection process makes the determination of the type of traffic violations uncertain. Existing studies generally focus on the road traffic accident itself, lacking synergistic studies of external factors, and there is the problem of difficulty in recording the main traffic violations that led to the accident. (3) Existing methods ignore the regularity of time sequence related to traffic violations and lack a targeted time window design.

The driver's own factors are a major cause of traffic violations in four main areas: age, gender, driving experience and personality. Studies at home and abroad have shown that only 7% of young drivers between the ages of 18 and 24 believe that their driving skills are poor and that they are prone to traffic violations when driving [22], while the remaining 93% believe that their driving skills are average or better. The truth, however, is quite the opposite: younger drivers are more prone than older drivers to make mistakes behind the wheel and to commit more violations, particularly in circumstances that call for good driving judgment. Although female drivers commit more traffic violations than male drivers do as a result of having less skill, the results of male and female drivers' self-evaluations of their driving abilities show that there is not much of a difference in the probability of committing traffic violations due to gender differences [23]. Existing studies on driving experience suggests that novices are unable to analyze potential hazards in the most adaptive mode and take a rational approach to avoiding traffic violations compared to experienced drivers [24]. There is a correlation between drivers' personality factors and traffic violations. For example, dutifulness and openness are positively correlated with avoiding traffic violations, while personality such as neuroticism, impulsivity, irritability, and depression are negatively correlated with avoiding traffic violations. Moreover, drivers' personality factors have a long-lasting impact on the occurrence of traffic violations.

Based on the above analysis, existing studies suffer from difficulties in obtaining datasets, overemphasis on vehicle and environmental factors, and analyzing only a single factor while ignoring the synergistic effects of multiple factors. Therefore, this paper constructs a traffic violation dataset based on the data from the Traffic Management Bureau of the Public Security Department of Liaoning Province, China, in which the source data contains four types of indicators describing driver factors, i.e., age, gender, driving experience, and personality, as well as 55 types of traffic violations. Through the relationship between driver factors and occurrence of traffic violations described in this dataset, a driver traffic violation prediction method is proposed. This method is able to proactively intervene in the driver behavior of smart connected vehicles to avoid traffic violations, thus providing a credible basis for improving the synergy of smart connected vehicle systems. Specially, the main contributions of the paper can be summarized as follows:

(1) A hierarchical network that can predict driver traffic violations is designed based on a traffic violation dataset. The network describes the characteristics of driver factors and calculates the probability of a traffic violation occurring in the future. It is also more advantageous than the non-hierarchical network in three evaluation metrics: prediction accuracy, mean absolute error and prediction precision.

(2) The proposed network utilizes the powerful feature extraction function of neural networks to obtain the time sequences of factors related to traffic violations from the dataset, and at the same time characterizes temporal

characteristics of these sequences using a long and short-term memory network to calibrate the relationship between traffic violations and driver factors.

(3) The attention network is improved by designing two plug-and-play modules including spatio-temporal interacttion and deep convolutional feature extraction. The former solves the problem of matching the dimensionality of the relevant indicator matrices, and the latter reduces the complexity of calculating the interrelationships among different indicators. The addition of these two modules overcomes the problem of indicator weight calibration error caused by the global average pooling operation and realizes the self-learning recalibration function of indicator weights.

The remaining part of the paper is structured as follows: Section II describes the related work. Section III gives the prediction framework for driver traffic violations. Section IV shows experimental setting and performance evaluation. Finally, Section V concludes this paper.

II. RELATED WORK

In terms of the relevant factors of driver traffic violations, the construction of the prediction model is generally realized through the analysis of the violation data set, i.e., the questionnaire is used to inquire drivers about their driving operations and psychological activities at the time of the violation, so as to extract the correlation indicators. With the introduction of sensors, images, positioning and other devices into traffic activities, it has become possible to collect vehicle trajectories and obtain data on the entire process of vehicles, but legal and cost restrictions prevent access to data on the characteristics of drivers at the time of the violation, resulting in a lack of relevance and comprehensiveness of prediction results. Li et al. [25] constructed vehicle-road collaboration scenarios using simulated driving technology and analyzed the multidimensional change characteristics of driving behaveiors in different scenarios, but they failed to form universal scenarios and applications due to the deviation of simulated driving technology from real scenarios. Chen and Chen [26] collected natural driving data to form an experimental database and discussed risky driving behavior and the relationships between the factors in it. Deng et al. [27] and Payyanadan et al. [28] showed that reaction time to traffic hazards and the desire to drive illegally decreases with age and experience. In addition, the inability to accurately assess and identify roadway risks leads to a decrease in cautious driving behavior during vehicle operation, increasing the risk of driving traffic violations. Gelmini et al. [29] showed that thrill-seeking or other driving intentional risk-taking behaviors also contribute to the occurrence of traffic violations.

In terms of related prediction technologies, road traffic violations are dominated by driver violations while driver behavior is monitored by exploring driving patterns. In the 1950s, a classical mathematical model reflecting current driving behavior was obtained by analyzing the linear driving behavior between neighboring vehicles on one-way streets, represented by the general motor model, which was applied to describe the phenomenon of frequent deceleration of vehicles during driving. And then, with collision avoidance models based on the relative distance between adjacent vehicles as the main factor, a safe distance and reaction delay time were introduced and a model based on this with speed of travel, safe distance, front and rear speed constraints emerges in order to simulate driver behavior leading to dangerous collisions. Later, safety distance and reaction delay time were introduced to establish a collision avoidance model based on the relative distance between neighboring vehicles as the primary factor. Based on this, constraint models with driving speed, safe distance, and front and rear vehicle speeds emerged to simulate driver behavior. In the 1960s, the driving psychological model was constructed using the effect of poor driver vision on the judgement of the driving environment as a factor, and six models of Weidmann driving behavior were formed from this. Then Fancher added driver psychological reaction thresholds to quantify the driver's perception of the moving lead vehicle. In the 1990s, Bando proposed the OV model that uses the speed difference to react to the driver's behavioral delay phenomenon and set up a speed optimization function to control the behavioral output of the rear-guided vehicle. After entering the 21st century, machine learning methods have made great progress, and techniques such as SVM, ant colony algorithm, genetic algorithm, recurrent neural network, GoogleLeNet, VGG and ResNet, have been successively applied to the field of driver behavior recognition. For example, Menno obtains two-bit features of facial thermal images by extracting the driver's facial features and combining them with Gaussian mixture model to identify the driving behaviour under the drinking condition. Yeo and Chuang used SVM method as the core, which can obtain the EEG signal features characterizing fatigue driving behavior and the EEG change features in different regions respectively. Liu extracted the driver's eye and head running information features by semi-supervised learning method and improved the binary discrimination accuracy by semi-supervised extreme learning machine method. Masood used VGG16 network to analyze the behavioral images during driving, and achieved better results in detecting distracted driving on a small-scale dataset.

Deep learning methods such as Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM) have the technical advantage of automatically extracting high-level features from low-level features to describe the constituents of driver traffic violations. This has led to deep learning methods for predicting driver traffic violations gradually becoming mainstream [30], [31], [32], [33]. Okafuji et al. [34] designed a behavior recognition scheme based on CNN fusion model, which achieves good performance on public datasets, but only considered the spatial characteristics of the data and ignored the temporal characteristics of the behavior. Li et al. [35] presented a behavioral pattern recognition model combining LSTM and recursive

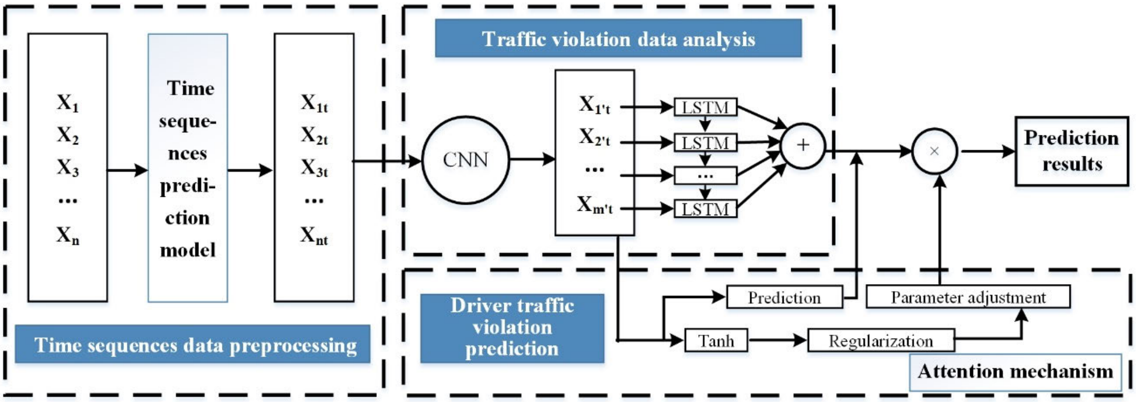

FIGURE 1. The overview of the hierarchical network-based method.

neural network algorithm, and the experimental results on public dataset showed that the model is adapted to data with temporal characteristics, and suffers from the defect of insufficient spatial information mining. Cura et al. [36] fused CNN and LSTM to extract spatio-temporal fusion characteristics and achieved high recognition rates for time-continuous behaviors, but the accuracy of the recognition results leaves something to be desired due to the feature extraction setup with the same weights. The attention mechanism (AT) [37], [38], [39] uses an automatic weighting mechanism that different modules in the model can be correlated with each other in a weighted way so that the model focuses its attention on specific parts of the input sequence. The BiLSTM-MCNN model proposed by Zhu et al. [40] introduced an attention mechanism for analyzing personal emotions and achieving accurate recognition. Gao et al. [41] used a lightweight network incorporating an attention mechanism to localize key features of the data. Chen and Gong [42] improved the attention module with matrix manipulation of spatio-temporal interactions and depth-separable convolutional methods, and integrated the module into an existing network to increase the accuracy of recognition. Developments in computer vision technology drive the prediction of traffic violations. A seatbelt detection method based on the Gaussian Mixture Model (GMM) and cascade classifier (Adabooast) was proposed by Chen in 2015. However, because of the high pixel requirements, this method has limited utility. Meng et al. [43] proposed the AdaVit framework, which combines strategies for the use of image blocks, self-attentive heads and transformer blocks to improve the efficiency of vision transformers. The framework can be used for driver behavioural reasoning. Li et al. [44] and Talukdera et al. [45] studied multi-scale feature representation of transformer models for image classification and their results could provide an effective way to classify driver behavior. Manzari et al. [46] examined the potential direct application of vision transformers for traffic sign recognition and evaluated the benefits and drawbacks of using vision transformers in conjunction with convolutional neural networks for this purpose.

III. THE HIERARCHICAL NETWORK-BASED METHOD FOR PREDICTING DRIVER TRAFFIC VIOLATIONS

The hierarchical network-based method for predicting driver traffic violations is divided into three parts: time sequences data preprocessing, traffic violation data analysis, and driver traffic violation prediction, as shown in Figure 1. The first part forms the multidimensional time sequences input data through data cleaning and fliting, time window and variable structure adjustment, as detailed in Section III-A. The second part constructs a framework for predicting traffic violations using CNN and LSTM, as detailed in Section III-B. The last part outputs prediction results by optimizing the model parameters through the attention layer and configuring the fully connected layer according to the regular and hierarchical input structure, as detailed in Section III-C.

Since the factors leading to the current driver's traffic violations are a continuation of the previous period and the LSTM prediction results are only influenced by historical data, it is necessary to extract time-series characteristic parameters of the driver's personality. Thus, the output of the CNN layer is used as the input to the LSTM layer, which is computed to output the time-series feature vector of the driver's personality at each moment.

Specifically, first traffic violation data is input in the form of a matrix, and this matrix is convolved with the convolution kernel K_i and the RELU is used as the activation function, so as to obtain the results of the convolution operation and the computed results after the nonlinear mapping. A maximum pooling operation is performed on these results and the pooled results are spliced. Secondly, x_t and h_{t-1} are fed into the LSTM layer to obtain the hidden layer information h_t. Next, the hidden layer data is convolved with K_i to compute the multidimensional features at different time moments in order to improve the network layer representation. Finally, the final hidden layer information of the LSTM is used as the query vector of the attention mechanism, and the multidimensional features are used as the key and value vectors of the attention mechanism, while the similarity between the query vectors and the outputs of the attention mechanism is then measured by the dot-product operation, which maps the importance of each time-step feature on the prediction results of the driver's illegal behaviors.

A. TIME SEQUENCES DATA PREPROCESSING

Driver traffic violation data are usually time sequences data, including basic driver information (such as age, fatigue status, etc.), driving experience, and personality changes that occur while driving (such as mood swings). Due to issues with missing values, anomalies, and noise throughout the data collection process, it is necessary to deal with the data that is currently accessible by interpolating the missing values and identifying the anomalies.

In the case of missing data values due to cognitive biases of the respondents, the mean value method is used to deal with the missing values. Let $X = \{x_1, x_2, \ldots, x_n\}$ denotes the traffic violation data, and X has missing value x_j, which can be calculated by (1):

$$x_j = \frac{x_i - x_k}{k - i} j + x_i \tag{1}$$

where x_i and x_k denote the pre-data and post-data of x_j, respectively, $k \in \{1, 2, \cdots, k - q - 1\}$.

There are obvious outliers in the dataset due to respondents' thinking biases, and the mainstream methods include Pauta criterion, Chauvenet criterion, Dixon criterion and Grubbs criterion. Among them, the Grubbs criterion can detect outliers when the overall standard deviation is unknown and the probabilistic significance is clear, which is calculated as follows:

$$|x_i - \overline{x}| > g\sigma(x) \tag{2}$$

In (2), x_i is the verification point, and $\overline{x}$ and $\sigma(x)$ are the mean and standard deviation of the verification sample, respectively, where the verification sample consists of a number of points extracted forward and backward starting at x_i, g is the value at certain defined confidence probability.

Redundant data can lead to slow and poor model fitting [47]. To overcome this shortcoming, significance analysis is used to explore the relationship between each indicator and the occurrence of traffic violations, removing redundant indicators and retaining significantly related indicators as the final inputs to the model. As shown in Figure 1, in the part of time sequences data preprocessing, the data before the current moment is taken as input, and after the prediction model with the neural network algorithm as the core, the current moment value is predicted, and then the intermediate input of each data $\{x_{1t}, x_{2t}, \ldots, x_{nt}\}$ as inputs for the next process.

The time window is the span of time between a change in a driver basic information or personality and the occurrence of a traffic violation. Setting of the time window includes two kinds of methods: fixed time window width and variable time window width, the former refers to the forward collection of characterization indicators over a fixed period of time, counting from the time of the occurrence of the traffic violation; the latter refers to the selection of different time window widths by taking into account the different characteristics of drivers. It has been shown that the time window within 1-3 seconds has a high prediction accuracy [48], [49], so the time window can be set to increment in 0.5 second intervals within 1-3 seconds and the appropriate time window width can be selected according to the experimental results.

B. TRAFFIC VIOLATION DATA ANALYSIS

1) THE CNN LAYER

Given the advantages of CNN in feature extraction, we use it to classify and recognize the basic information, driving experience and personality of drivers, as shown in Figure 2. The CNN contains four layers, the first of which is the input layer. Its data structure is a 4-tuple (D 1 H W), where D is the driver indicator data related to a single traffic violation with channel number 1, and H and W are the corresponding height and width of the indicator data. The feature information from the indicator data is extracted using layers 2 through 4, which are CNN layers with convolutional kernels K1, K2, and K3, respectively. Each layer has a batch normalization, activation layer, pooling layer, and fully connected layer, and the output format is a 4-tuple (D K_2 H_2 W_2), (D K_3 H_3 W_3) and (D K_4 H_4 W_4), respectively.

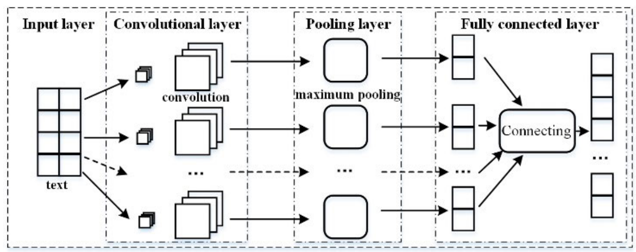

FIGURE 2. The CNN layer structure.

CNN related rules are as follows: The convolution layer extracts local features through convolution operation to reduce the dimensionality of input data; the pooling layer reduces the dimensionality by removing some of the features for the purpose of feature data filtering; the fully connected layer is used for feature weighting to complete the classification function. In (3), the process of convolution operation is displayed.

$$\mathrm{D} \times \omega = \sum_{k=1}^{C} \sum_{j=1}^{W} \sum_{i=1}^{H} \left[\mathrm{D}^k(i,j)\, \omega^k(i,j) \right] \tag{3}$$

where D is the convolutional layer input data, ω is the weight of the convolutional kernel, and C, W and H are the number of channels, width and height of the convolutional kernel. Typically, an activation function is added to the convolutional layer, often using the Leaky Relu function using (4).

$$LeakyRelu(x) = \begin{cases} x & x > 0 \\ ax & x \le 0 \end{cases} \tag{4}$$

The above function solves the negative-zero gradient problem by assigning a very small linear classification of x

to the negative input ax, where a is a very small positive number.

C. THE LSTM LAYER

In order to overcome the gradient explosion or vanishing problem of CNN, LSTM is adopted for extracting the temporal characteristics of the traffic violation data and predicting the type of traffic violations through model training, as shown in FIGURE3. LSTM contains forget gate f_t, input gate i_t, and output gate o_t, all of which are functions of the previous state h_{t-1} and the new input x_t as independent variables. Memory cell c_t and candidate memory cell $\overline{c_t}$ are related to the memory cell of previous moment c_{t-1} and are adjusted by i_t and f_t. The method of updating the gate unit and memory cell is described in (5)-(10).

$$f_t = \sigma\left(W_f\left(h_{t-1}, x_t\right) + b_f\right) \tag{5}$$
$$i_t = \sigma\left(W_i\left(h_{t-1}, x_t\right) + b_i\right) \tag{6}$$
$$o_t = \sigma\left(W_o\left(h_{t-1}, x_t\right) + b_o\right) \tag{7}$$
$$\overline{c_t} = tanh\left(W_c\left(x_t, h_{t-1}\right) + b_c\right) \tag{8}$$
$$c_t = f_t c_{t-1} + i_t\overline{c_t} \tag{9}$$
$$h_t = o_t * tanh\left(c_t\right) \tag{10}$$

where, W_f, W_i, W_o an W_c are the weight parameters, b_f, b_i, b_o and b_c are the bias parameters, and f_t, i_t and o_t use the sigmoid function as the activation function and normalize the state of the memory cell. The sigmoid function is defined as in (11).

$$sigmoid\left(x\right) = \frac{1}{1+e^{-x}} \tag{11}$$

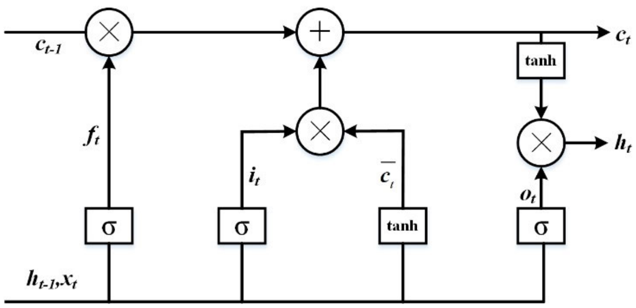

FIGURE 3. The STM layer structure.

D. DRIVER TRAFFIC VIOLATION PREDICTION

1) THE ATTENTION LAYER

The attention mechanism is the brain's signal processing mechanism for dealing with visual signals, which focuses on assigning different weights to different aspects of a things in order to reduce the significance of unimportant information. Traditional traffic violation prediction methods ignore the proportion of weights between variables, but it has been found that different indicators have different relevance to the occurrence of traffic violations. By exploring the internal relationships of the violation data, an attention layer is designed to automatically optimize the parameters, thus improving the prediction accuracy. Therefore, the indicators that are significantly related to traffic violations are filtered out using significance analysis, and then the values of the significance indicators of the hidden vectors output from LSTM are calculated and ranked by the attention mechanism, so as to improve the accuracy of the prediction method, see FIGURE 4.

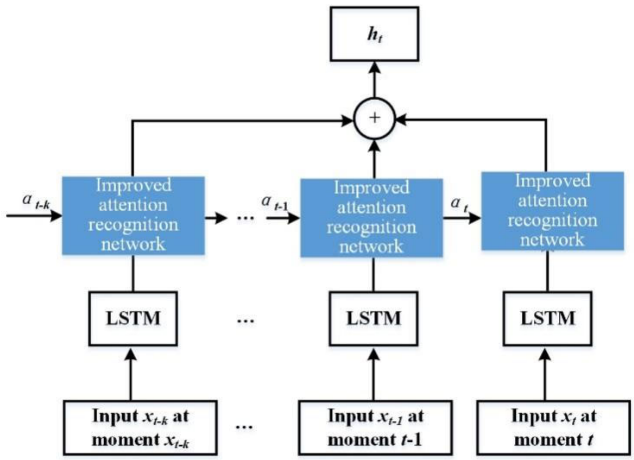

FIGURE 4. The structure of the attentional layer.

In the proposed method, the output of the LSTM layer using significance analysis is taken as the input of the attention layer, and matrix multiplication is used to process the correlation indicators. We find that the current attention mechanism is somewhat maladaptive, as the global average pooling operation used does not assign weights based on indicator characteristics, reinforcing unimportant information to some extent. Therefore, considering the assignment of different learnable weights based on the indicator characteristics, the existing attention mechanism is improved by designing a spatio-temporal interaction module and a deep convolutional feature extraction module. These two modules fulfill the plug-and-play requirement and can be added directly to an existing attention layer network to form a new recognition network.

The spatio-temporal interaction module aims to solve the problem of matching the dimensions of the matrix of relevant indicators, and the main processes are as follows. First, the input dimension $[D \times T \times H \times W]$ is converted to $[DT \times HW]$ using matrix multiplication to obtain the first median value, where D is the value of the indicator sequence, T is the number of channels, H denotes the height of the sequence, and W denotes the width of the sequence. Secondly, the feature information of the dimensions is extracted using the Reshape-Conv method, after which the Softmax method is used to assign different learnable weights to different indicators to obtain the second median value. Then, a matrix multiplication operation is performed on these two median values, and later the different indicator weights are weighted to the corresponding original indicator characteristics using incentive weighting. Finally, the original feature completes the recalibration operation. To address the shortcomings of the existing incentive weighting functions that rely on batches in calculating the mean and variance, i.e., too small a batch value will result in a mean and variance that are insufficient to fully express the entire distribution of the data, the Group-Norm method is used instead of the BatchNorm method so as to eliminate the influence of the batch value on the results.

The deep convolutional feature extraction module aims to reduce the computational complexity when calculating the interrelationships between different indicators, and the main process is as follows. Deep convolution is utilized to achieve the compression of the input dimension, assuming that the input is changed to $(DT\times H\times W)$ using the Permute function, the convolution kernel is $(K_1\times K_2)$, the number of convolution kernels is L. In addition, the inputs are grouped based on the indicators related to the traffic violations of the drivers (basic information, driving experience, and personality), and the number of groups is G. Using the usual convolutional computation method, the number of parameters is $DT\times K_1\times K_2\times L$. After using grouped convolution, the number of parameters is one-third of the original. When $L=G$, the grouped convolution uses the Depthwise-Conv method; when $L=G$ and $K_1=H$, $K_2=W$, global pooling can be achieved while assigning learnable weights to different variability indicators. The entire process of the improved attention recognition network can be seen in FIGURE 5.

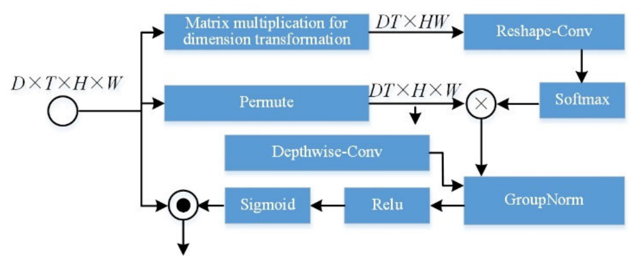

FIGURE 5. The improved attention recognition network.

Since the factors that lead to different traffic violations are not the same, however the feature data is computed using CNN and LSTM layers, the weights of each factor are balanced. Considering the fact that the importance of irrelevant information leading to traffic violations is somehow enhanced while the importance of relevant information is suppressed, the above two modules assign learnable weights to the feature data, which improves the prediction accuracy.

2) PREDICTION RESULTS OF DRIVER TRAFFIC VIOLATIONS

The hierarchical network of proposed method is mainly divided into 8 layers. Specifically, layers 1 through 4 are CNN layers (see Section III-B1). Layer 5 is an LSTM layer with k hidden units, which is used to learn the nonlinear relationship between the type of traffic violations and related indicator, and the temporal dependence of indicator sequences. Layer 6 is the attention layer, which acquires the traffic violation data for each time period and weights the hidden states output from the LSTM layer. Layer 7 is a fully connected layer that filters the traffic violation features and maps them to the sample space through a weight matrix. Layer 8 is the output layer, which is used to obtain the percentage of traffic violations types, i.e., to predict the probability of occurrence of a particular type of traffic violations.

IV. EXPERIMENTS

A. EXPERIMENTAL SETTINGS

The experiment was simulated using an HP server with an Intel(R) Core(TM) i9-10900X @3.7GHz CPU, 128G of RAM, and an NVIDIA GeForce RTX 2090 graphics card. The dataset was obtained from the Traffic Management Bureau of the Public Security Department of Liaoning Province, China. The data were collected in the form of a questionnaire, which was divided into three main aspects: basic information, driving experience and personality, as shown in Table 2. Meanwhile, traffic violations were classified into five types with a total of 55 items, including 12 items in the first type, 15 items in the second type, 12 items in the third type, 12 items in the fourth type, and 4 items in the fifth type. The questionnaire was completed by drivers when they processed their traffic violation record at the Traffic Management Bureau.

TABLE 2. A partial presentation of the questionnaire structure.

	Basic information		Driving experience			
Number	Age	Gender	Driving years	Occupation	Education	Driving skill
Personality						
Obsessive-compulsive disorder	Sensitivity	Depression	Psychosis	Anxiety	Hostility	Stubbornness
Driver traffic violations						
Type 1 (12 items)	Type 2 (15 items)		Type 3 (12 items)	Type 4 (12 items)		Type 5 (4 items)

The experimental parameters involved in the first 5 layers of our hierarchical network are K_1, K_2, K_3 and k. In our experiments, we found that the prediction accuracy of traffic violation results is higher when $K_1=32$, $K_2=128$, $K_3=64$ and $k=64$. Meanwhile, the size of the convolution kernel is set to (1, 11), (1, 9) and (1, 7), the step size is set to (1, 2), the padding values in the convolution are set to (0, 5), (0, 4), and (0, 3), the number of encoder and decoder layers of LSTM layer is 4, the number of encoder units and decoder units are 64 and 128 respectively, the number of hidden units is 128, the maximal pooling layer window length is set to (1, 3), and the pooling step size is set to (1, 2). Cross entropy is used as a loss function during CNN forward propagation; Adam's method is used to update the weights and biases of the CNN during error back propagation. Additionally, the training batch size is 128, the initial value of the learning rate is 0.001, the number of iterations is 100, the weight decay value is 0.0001, the dropout value is 0.2 and the input time window width is (1.0s, 1.5s, 2.0s, 2.5s, 3.0s). The evaluation metrics such as precision, F1 score, and precision are the same as defined in the [50].

B. ANALYSIS OF EXPERIMENTAL RESULTS

In order to verify the prediction accuracy of the proposed method, a significance analysis was conducted on the indicators related to driver traffic violations, as shown in TABLE 3. Individual indicators are used as independent inputs into the time sequence prediction model, and the output time prediction sequence are used as inputs to the next layer. Meanwhile, the other indicators remain unchanged and go directly to the CNN layer.

TABLE 3. Significance values for some indicators.

Indicator	Mean value	Variance	Significance value
Age	0.79	0.25	<0.05
Gender	0.27	0.43	<0.01
Distraction driving	2.05	1.21	>0.05
Fatigue driving	2.00	0.40	>0.01
Driving skill	2.30	0.27	>0.05
Sensitivity	0.21	0.19	<0.01
Hostility	1.85	0.63	>0.05
Depression	0.15	0.10	<0.01
Stubbornness	1.13	0.42	<0.05

The temporal sequences into the CNN layer and LSTM layer are fitted to output the classification results of traffic violations. An Adam method is used for optimization and an exit threshold of 0.5 is set to prevent overfitting. Also, the model error was verified with the mean absolute error MAE, which is defined as:

$$\mathrm{MAE}=\frac{1}{n}\sum_{r=1}^{n}\left|b_r-b_r'\right| \tag{12}$$

where b_r is the prediction result of driver traffic violations and b_r' is the actual value of traffic violations. Two performance evaluation metrics, prediction accuracy and MAE, are used to validate the effectiveness of the proposed method by comparing the non-hierarchical network-based method (LSTM), the hierarchical network-based method (CNN+ LSTM), and the method with the addition of an attention mechanism (CNN+LSTM+ ATTENTION).

As can be seen from TABLE 4, the addition of the attention layer can effectively improve the prediction accuracy of traffic violations and reduce the MAE value. The hierarchical network structure can better improve the prediction accuracy, with the best effect when the time window is 2.0s. The reason for this may be that the time of 2.0s coincides with the time difference between the change in the temporal indicator of the traffic violations and the occurrence of the traffic violations.

The F1-score comparison results are shown in TABLE 5 and it can be seen that the proposed CNN+LSTM+ ATTENTION method has the highest value, which indicates that it has some advantages.

TABLE 4. Prediction accuracy and MAE values for different methods under different time window widths.

Time window width (s)	NON-HIERARCHICAL NETWORK-BASED METHOD		Hierarchical network-based method			
	LSTM		CNN+LSTM		CNN+LSTM+ATTENTION	
	Accuracy	MAE	Accuracy	MAE	Accuracy	MAE
1.0	0.732	0.300	0.744	0.281	0.792	0.221
1.5	0.749	0.318	0.763	0.230	0.791	0.221
2.0	0.824	0.244	0.861	0.187	0.898	0.152
2.5	0.786	0.263	0.830	0.200	0.882	0.177
3.0	0.788	0.220	0.826	0.224	0.875	0.185

TABLE 5. F1-score for different methods.

Method	CNN	LSTM	CNN-LSTM	SVM	CNN+LSTM+ATTENTION
F1-score	0.72	0.75	0.79	0.70	0.85

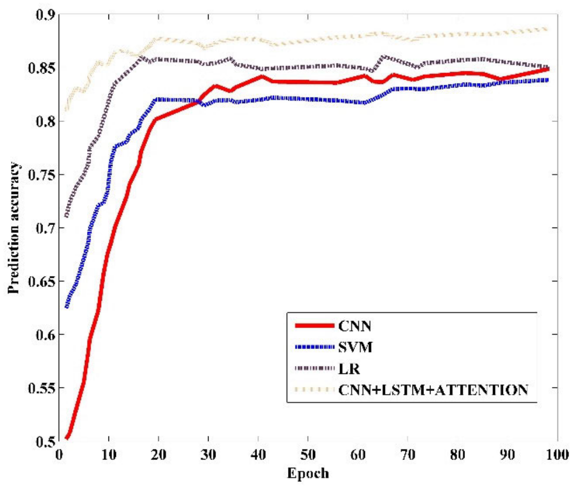

FIGURE 6. The prediction accuracy curve during the training period of the dataset.

FIGURE 6 shows the prediction accuracy curve of the proposed method during the training period of the dataset. From the figure, it can be seen that the prediction accuracy tends to increase with the increase in the number of iterations and finally reaches a steady state, and the prediction accuracy of the proposed method is higher than that of the comparison methods. That is, the model converges faster and predicts driver traffic violations with higher accuracy during the training process.

Our method is compared with existing models, including Logistic Regression(LR), Support Vector Machine (SVM), CNN, [50] and [51] to verify the effectiveness of the proposed method. The comparison results are shown in Table 6. As can be seen from the table, the prediction precision of the proposed method (CNN+LSTM+ ATTENTION) is about 10% higher than the other methods, and then comparing the prediction precision of the different models at each time window width, it is found that the performance when the time window width is 2.0s is generally higher than the other time windows. Therefore, the proposed hierarchical method can be more effective in predicting driver traffic violations, and at the same time, the time window width of 2.0s gets the best prediction effect.

TABLE 6. Prediction precision of the different methods at different time intervals.

METHOD\ WINDOW WIDTH TIME	1.0s	1.5s	2.0s	2.5s	3.0s
LR	0.752	0.781	0.835	0.823	0.828
SVM	0.715	0.732	0.754	0.701	0.711
CNN	0.610	0.614	0.658	0.636	0.688
[50]	0.735	0.727	0.770	0.770	0.742
[51]	0.717	0.732	0.783	0.720	0.723
CNN+LSTM+ ATTENTION	0.769	0.789	0.885	0.861	0.860

The computation time for the prediction of a single data item is shown in Table 7. When working with larger datasets, LSTM performs better than CNN, LR, and SVM; nevertheless, it takes longer because it needs more parameter. This lays the foundation for our method in this paper, but we also add an improved attention module, increasing calculation time.

TABLE 7. Calculation time for different methods.

Method	CNN	LSTM	LR	SVM	CNN+LSTM+ ATTENTION
Time(s)	0.002	0.066	0.061	0.064	0.109

V. CONCLUSION

Most road traffic accidents are related to human factors, such as driver traffic violations. In order to explore the relationship between driver factors and traffic violations, and thus to predict the occurrence of driver traffic violations, a hierarchical network-based method is proposed to enhance the synergy of the current smart connected vehicle systems. It is challenging to access the traffic violation dataset because it is only stored in the traffic management department. We constructed the traffic violation dataset by collecting data from the Traffic Management Bureau of Liaoning Provincial Public Security Department. On this basis, we use neural networks to obtain the time series of factors related to traffic violations, while extracting the time sequences features through long and short-term memory networks. In order to predict drivers' traffic violations, an attention network is introduced, but in the process, it was discovered that the network has the problem of indicator weight calibration error. As a result, two plug-and-play modules are designed: spatio-temporal interaction and deep convolutional feature extraction. The former solves the problem of mismatch in the dimensionality of the relevant indicator matrices, and the latter reduces the computational complexity in measuring the interrelationships between different indicators. The experimental results show that the proposed method can improve the prediction accuracy of driver traffic violations, which in turn provides a credible basis for smart connected vehicles to actively intervene in driver behavior, avoid traffic violations, and improve the safety of traffic participants.

Due to the self-preference of the drivers interviewed, the data items in the dataset are more favorable to them when it comes to road traffic accidents. This led to some bias in the attention module of the proposed method when analyzing and processing the importance of individual factors. Further experiments are required in the future to rectify the self-preferences of the drivers. Additionally, the trade-off between recognition accuracy and model complexity must be carefully considered.

Future work will focus on two main aspects. On the one hand, the sample size is increased much more. The current dataset includes 55 traffic violations and four types of indicators that describe characteristics of drivers. The study discovered that some of the drivers interviewed unable to adequately describe their self-personality and that not all personalities were included in the dataset. So the current questionnaire needs to be improved in order to refine and expand the dataset. On the one hand, in order to suppress the issue of drivers' subjective preferences, new techniques and methods can be investigated. For example, using computer vision techniques for real-time video monitoring of driving behavior with privacy concerns, and exploring fusion solutions for current hierarchical network and computer vision networks.

ACKNOWLEDGMENT

Mingze Wang would like to thank the data collection personnel of the Transportation Department of Liaoning Province for the data support.

REFERENCES

[1] (2023). *National Data About Road Traffic Accidents*. [Online]. Available: https://data.stats.gov.cn/easyquery.htm?cn=C01&zb=A0S0D01&sj=2023
[2] Picks. *Road Accidents in Europe—Statistics & Facts*. Accessed: 2023. [Online]. Available: https://www.statista.com/topics/4188/road-accidents-in-europe/
[3] S. A. Elsagheer Mohamed, "Automatic traffic violation recording and reporting system to limit traffic accidents : Based on vehicular ad-hoc networks (VANET)," in *Proc. Int. Conf. Innov. Trends Comput. Eng. (ITCE)*, Aswan, Egypt, Feb. 2019, pp. 254–259.
[4] Y. Xing, C. Lv, X. Mo, Z. Hu, C. Huang, and P. Hang, "Toward safe and smart mobility: Energy-aware deep learning for driving behavior analysis and prediction of connected vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 4267–4280, Jul. 2021.

[5] X. Zhu, J. Chang, W. Li, B. Yu, L. Zhao, and X. Wang, "Driver violation prediction based on neural network," in *Proc. 9th Int. Conf. Dependable Syst. Their Appl. (DSA)*, Wulumuqi, China, Aug. 2022, pp. 143–152.

[6] G. Baicang, J. Lisheng, S. Jian, and Z. Shunran, "A risky prediction model of driving behaviors: Especially for cognitive distracted driving behaviors," in *Proc. 4th CAA Int. Conf. Veh. Control Intell. (CVCI)*, Hangzhou, China, Dec. 2020, pp. 103–108.

[7] H. Seo, J. Shin, K.-H. Kim, C. Lim, and J. Bae, "Driving risk assessment using non-negative matrix factorization with driving behavior records," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 11, pp. 20398–20412, Nov. 2022.

[8] P. P, S. C. Y, A. Srivastava, C. P. K J, and A. Sharma, "A conceptual model simulation to detect and report city traffic violations using distributed intelligent agents," in *Proc. 6th Int. Conf. Converg. Technol. (I2CT)*, Maharashtra, India, Apr. 2021, pp. 1–6.

[9] D. Wang, J. Peng, J. Zhao, Y. Teng, W. Xue, and X. Tao, "Dual-transformer: A general model for traffic accident prediction," in *Proc. IEEE 98th Veh. Technol. Conf. (VTC-Fall)*, Hong Kong, Oct. 2023, pp. 1–5.

[10] H. J. Cho and M. C. Hwang, "A stimulus-response model of day-to-day network dynamics," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 1, pp. 17–25, Mar. 2005.

[11] Z. Wang, G. Xu, M. Zhang, and Y. Guo, "Collision avoidance models and algorithms in the era of Internet of Vehicles," in *Proc. IEEE 3rd Int. Conf. Safe Prod. Informatization (IICSPI)*, Chongqing, China, Nov. 2020, pp. 123–126.

[12] X. Yang, Y. Zhou, and Q. Fu, "A lane-changing and off-ramp model of vehicles on enclosed road considering driving psychology," *J. Highway Transport. Res. Dev.*, vol. 7, no. 32, pp. 112–119, Jul. 2015.

[13] T. Zhang, J. Xu, S. Cong, C. Qu, and W. Zhao, "A hybrid method of traffic congestion prediction and control," *IEEE Access*, vol. 11, pp. 36471–36491, 2023.

[14] Y.-J. Kim and E.-K. Hong, "Identification of mobile traffic characteristics in university area with decision tree learning model," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2021, pp. 854–857.

[15] M. A. Mondal and Z. Rehena, "Identifying traffic congestion pattern using K-means clustering technique," in *Proc. 4th Int. Conf. Internet Things Smart Innov. Usages (IoT-SIU)*, Ghaziabad, India, Apr. 2019, pp. 1–5.

[16] Z. Li, C. Chen, Y. Min, J. He, and B. Yang, "Dynamic hidden Markov model for metropolitan traffic flow prediction," in *Proc. IEEE 92nd Veh. Technol. Conf. (VTC-Fall)*, Victoria, BC, Canada, Nov. 2020, pp. 1–5.

[17] G. Byun, V.-V. Vo, S. M. Raza, D.-T. Le, H. Yang, and H. Choo, "Proactive network traffic prediction using generative adversarial network," in *Proc. Int. Conf. Inf. Netw. (ICOIN)*, Ho Chi Minh, Vietnam, Jan. 2024, pp. 156–159.

[18] P.-S. Shih, S. Liu, and X.-H. Yu, "Ant colony optimization for multi-phase traffic signal control," in *Proc. IEEE 7th Int. Conf. Intell. Transp. Eng. (ICITE)*, Beijing, China, Nov. 2022, pp. 517–521.

[19] L. Xie, T. Guo, J. Chang, C. Wan, X. Hu, Y. Yang, and C. Ou, "A novel model for ship trajectory anomaly detection based on Gaussian mixture variational autoencoder," *IEEE Trans. Veh. Technol.*, vol. 72, no. 11, pp. 13826–13835, Nov. 2023.

[20] M. Yang, "Traffic flow prediction optimization model based on convolutional neural network of time-domain graph," in *Proc. IEEE 3rd Int. Conf. Electr. Eng., Big Data Algorithms (EEBDA)*, Changchun, China, Feb. 2024, pp. 20–24.

[21] N. B. Bahadure, P. D. Patil, R. Birewar, P. Nayyar, A. Shrivastav, and M. Oberoi, "Traffic signal detection and recognition from real-scenes using Yolo," in *Proc. IEEE Eng. Informat.*, Nov. 2023, pp. 1–5.

[22] G. Dou, X. Song, and R. Chang, "Effect of personality on the driver's hazard perception skill," *J. Liaoning Normal Univ. Social Sci. Ed.*, vol. 39, no. 4, pp. 46–50, Jul. 2016.

[23] H. Momma, J. Suzurikawa, R. Takeshima, Y. Sawada, Y. Harada, and T. Kondo, "Assessment of driving skills of a mobility scooter using driving operation logs," in *Proc. IEEE 42nd Annu. Int. Conf. Eng. Med. Biol. Soc. (EMBC)*, Montreal, QC, Canada, Jul. 2020, pp. 3212–3215.

[24] S. Gan, Q. Li, Q. Wang, W. Chen, D. Qin, and B. Nie, "Constructing personalized situation awareness dataset for hazard perception, comprehension, projection, and action of drivers," in *Proc. IEEE Int. Intell. Transp. Syst. Conf. (ITSC)*, Sep. 2021, pp. 1697–1704.

[25] L. Shuguang, L. Zhonglin, W. Wenbo, Z. Yang, H. Jierui, and C. Hong, "Vehicle-in-the-loop intelligent connected vehicle simulation system based on vehicle-road-cloud collaboration," in *Proc. IEEE 25th Int. Conf. Intell. Transp. Syst. (ITSC)*, Beijing, China, Oct. 2022, pp. 1711–1716.

[26] L.-W. Chen and H.-M. Chen, "Driver behavior monitoring and warning with dangerous driving detection based on the Internet of Vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 11, pp. 7232–7241, Nov. 2021.

[27] Y. Deng, J. Cauffman, M. Y. Lau, E. Johnson, A. Avr, and C. Cunningham, "Driver hazard response when processing on-road and in-vehicle messaging of non-safety-related information," in *Proc. IEEE Int. Conf. Human-Mach. Syst. (ICHMS)*, Rome, Italy, Sep. 2020, pp. 1–8.

[28] R. P. Payyanadan, F. A. Sanchez, and J. D. Lee, "Influence of familiarity on the driving behavior, route risk, and route choice of older drivers," *IEEE Trans. Human-Mach. Syst.*, vol. 49, no. 1, pp. 10–19, Feb. 2019.

[29] S. Gelmini, S. C. Strada, M. Tanelli, S. M. Savaresi, and V. Biase, "Online assessment of driving riskiness via smartphone-based inertial measurements," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 9, pp. 5555–5565, Sep. 2021.

[30] B. M. Albaba and Y. Yildiz, "Driver modeling through deep reinforcement learning and behavioral game theory," *IEEE Trans. Control Syst. Technol.*, vol. 30, no. 2, pp. 885–892, Mar. 2022.

[31] J. Xu, S. Pan, P. Z. H. Sun, S. Hyeong Park, and K. Guo, "Human-factors-in-driving-loop: Driver identification and verification via a deep learning approach using psychological behavioral data," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 3, pp. 3383–3394, Mar. 2023.

[32] T. Huang, R. Fu, Y. Chen, and Q. Sun, "Real-time driver behavior detection based on deep deformable inverted residual network with an attention mechanism for human-vehicle co-driving system," *IEEE Trans. Veh. Technol.*, vol. 71, no. 12, pp. 12475–12488, Dec. 2022.

[33] I. Jegham, A. B. Khalifa, I. Alouani, and M. A. Mahjoub, "Soft spatial attention-based multimodal driver action recognition using deep learning," *IEEE Sensors J.*, vol. 21, no. 2, pp. 1918–1925, Jan. 2021.

[34] Y. Okafuji, T. Sugiura, and T. Wada, "Preliminary investigation of visual information influencing driver's steering control based on CNN," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Toronto, ON, Canada, Oct. 2020, pp. 2263–2268.

[35] D. C. Li, M. Y. Lin, and L.-D. Chou, "Macroscopic big data analysis and prediction of driving behavior with an adaptive fuzzy recurrent neural network on the Internet of Vehicles," *IEEE Access*, vol. 10, pp. 47881–47895, 2022.

[36] 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.

[37] I. Kotseruba and J. K. Tsotsos, "Attention for vision-based assistive and automated driving: A review of algorithms and datasets," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 11, pp. 19907–19928, Nov. 2022.

[38] K. Zhang, L. Zhao, C. Dong, L. Wu, and L. Zheng, "AI-TP: Attention-based interaction-aware trajectory prediction for autonomous driving," *IEEE Trans. Intell. Vehicles*, vol. 8, no. 1, pp. 73–83, Jan. 2023.

[39] H. Wang, L. Xu, A. Bezerianos, C. Chen, and Z. Zhang, "Linking attention-based multiscale CNN with dynamical GCN for driving fatigue detection," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2021.

[40] X. Zhu, S. Yin, and Z. Chen, "Attention based BiLSTM-MCNN for sentiment analysis," in *Proc. IEEE 5th Int. Conf. Cloud Comput. Big Data Anal. (ICCCBDA)*, Chengdu, China, Apr. 2020, pp. 170–174.

[41] Z. Gao, J. Xu, J. Y. Zheng, H. Fan, R. Yu, J. Zong, and X. Li, "A lightweight VK-Net based on motion profiles for hazardous driving scenario identification," in *Proc. IEEE 23rd Int. Conf. High Perform. Comput. Commun., 7th Int. Conf. Data Sci. Syst., 19th Int. Conf. Smart City, 7th Int. Conf. Dependability Sensor, Cloud Big Data Syst. Appl. (HPCC/DSS/SmartCity/DependSys)*, Hainan, China, Dec. 2021, pp. 908–913.

[42] Y. Chen and S. Gong, "Human action recognition network based on improved channel attention mechanism," *J. Electr. Informat. Technol.*, vol. 43, no. 12, pp. 3538–3545, Dec. 2021.

[43] L. Meng, H. Li, B.-C. Chen, S. Lan, Z. Wu, Y.-G. Jiang, and S.-N. Lim, "AdaViT: Adaptive vision transformers for efficient image recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, New Orleans, LA, USA, Jun. 2022, pp. 12299–12308.

[44] L. Li, Z. Wu, J. Liu, L. Wang, Y. Jin, P. Jiang, J. Feng, and M. Wu, "Cross-attention based multi-scale feature fusion vision transformer for breast ultrasound image classification," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Las Vegas, NV, USA, Dec. 2022, pp. 1616–1619.

[45] S. Tummala, S. Kadry, S. A. C. Bukhari, and H. T. Rauf, "Classification of brain tumor from magnetic resonance imaging using vision transformers ensembling," *Current Oncol.*, vol. 29, no. 10, pp. 7498–7511, Oct. 2022.

[46] O. N. Manzari, A. Boudesh, and S. B. Shokouhi, ''Pyramid transformer for traffic sign detection,'' in *Proc. 12th Int. Conf. Comput. Knowl. Eng. (ICCKE)*, Mashhad, Iran, Nov. 2022, pp. 112–116.
[47] J. Park, H. Son, J. Lee, and J. Choi, ''Driving assistant companion with voice interface using long short-term memory networks,'' *IEEE Trans. Ind. Informat.*, vol. 15, no. 1, pp. 582–590, Jan. 2019.
[48] V. Urošević and S. Dimitrijević, ''Optimum input sequence size for a sliding window-based LSTM neural network used in short-term electrical load forecasting,'' in *Proc. 29th Telecommun. Forum (TELFOR)*, Belgrade, Serbia, Nov. 2021, pp. 1–4.
[49] J. Zhao, X. Zhang, H. Wang, and D. Mu, ''Load trend prediction analysis based on LSTM and sliding time window,'' in *Proc. IEEE 5th Adv. Inf. Manage., Communicates, Electron. Autom. Control Conf. (IMCEC)*, vol. 5, Chongqing, China, Dec. 2022, pp. 1239–1242.
[50] S. Jia, F. Hui, S. Li, X. Zhao, and A. J. Khattak, ''Long short-term memory and convolutional neural network for abnormal driving behaviour recognition,'' *IET Intell. Transp. Syst.*, vol. 14, no. 5, pp. 306–312, May 2020.
[51] I. E. Mallahi, A. Dlia, J. Riffi, M. A. Mahraz, and H. Tairi, ''Prediction of traffic accidents using random forest model,'' in *Proc. Int. Conf. Intell. Syst. Comput. Vis. (ISCV)*, Fez, Fez, Morocco, May 2022, pp. 1–7.

MINGZE WANG received the B.S. degree in accounting from Liaoning University, Liaoning, China, in 2013, and the M.S. degree in accounting from the Central University of Finance and Economics, Beijing, China, in 2015. She is currently pursuing the Ph.D. degree in management science and engineering with Liaoning Technical University, Liaoning.

Since 2019, she has been a Lecturer with the Business School, Nanning University. Her master's focused on research in enterprise and organization management. Her current research interests include safety psychology and behavior management. She is a participant in a research project on coal mine safety management and geological disaster prediction recently (Basic Research Project of Liaoning Provincial Department of Education).

NAIWEN LI is currently a Professor, a Doctoral Supervisor, and the Deputy Dean of the School of Business Administration, Liaoning Technical University. Since 1991, he has been teaching with the School of Business Administration, Liaoning Technical University. He presided over the research and training of related horizontal and vertical topics, such as the National Natural Science Foundation of China ''Research on Coal Mine Safety Accidents and Miners Burnout'' and the Enterprise Cooperation Project ''Research on Safety Competency of Special Coal Mine Operators.'' He is one of the most influential scholars in the field of behavioral safety in China. His research interests include human resource management and behavioral safety management.

He received honors including: Experts in the Humanities and Social Sciences Expert Database of the Ministry of Education and the second Liaoning Provincial Philosophy and Social Science Achievement Award Subject Experts of the Review Team. He was also selected for the ''Hundred Talents'' level of Liaoning ''Hundred and Thousand Talents Project.''