ISSN 1648-4142 / eISSN 1648-3480

VILNIUS TECH Vilnius Gediminas Technical University

TRANSPORT

2024
Volume 39
Issue 3
Pages 229–239
https://doi.org/10.3846/transport.2024.22749

Original Article

STARTING DRIVING STYLE RECOGNITION OF ELECTRIC CITY BUS BASED ON DEEP LEARNING AND CAN DATA

Dengfeng ZHAO[1], Zhijun FU[2]✉, Chaohui LIU[3], Junjian HOU[4], Shesen DONG[5], Yudong ZHONG[6]

[1,2,4,6] *Zhengzhou University of Light Industry, Zhengzhou, China*
[3] *Research Institute of Technology, Yutong Bus Co. Ltd., Zhengzhou, China*
[5] *Zhengzhou Senpeng Electronic Technology Co. Ltd., Zhengzhou, China*

Highlights:

- a new method proposed to identify the starting driving style of electric city buses based on deep learning and CAN data;
- an end-to-end deep learning framework is used to investigate methods for recognising driving styles during the start-up phase of electric buses, which have a high rate of frontal collision accidents;
- the deep learning method proposed in this article can automatically extract deep spatiotemporal features of multi-channel time series data and achieve end-to-end data processing with higher accuracy and generalisation capability;
- the proposed deep learning framework directly processes raw in-vehicle can bus data for end-to-end training through simple pre-processing, which makes it universally applicable.

Article History:

- submitted 13 August 2022;
- resubmitted 29 December 2023, 6 February 2024;
- accepted 15 April 2024.

Abstract. Drivers with aggressive driving style driving electric city buses with rapid response and high acceleration performance characteristics are more prone to have traffic accidents in the starting stage. It is of great importance to accurately identify the drivers with aggressive driving style for preventing traffic accidents of city buses. In this article, a starting driving style recognition method of electric city bus is firstly proposed based on deep learning with in-vehicle Controller Area Network (CAN) bus data. The proposed model can automatically extract the deep spatiotemporal features of multi-channel time series data and achieve end-to-end data processing with higher accuracy and generalization ability. The sample data set of driving style is established by pre-processing the collected in-vehicle CAN bus data including the status of driving and vehicle motion, the data pre-processing method includes data cleaning, normalization and sample segmentation. Data set is labelled with subjective evaluation method. The starting driving style recognition method based on Convolutional Neural Network (CNN) model is constructed. Multiple sets of convolutional layers and pooling layers are used to automatically extract the spatiotemporal characteristics of starting driving style hidden in the data such as velocity and pedal position etc. The fully connected neural network and incentive function *Softmax* are applied to establish the relationship mapping between driving data characteristics and the starting driving styles, which are categorized as cautious, normal and aggressive. The results show that the proposed model can accurately recognize the starting driving style of electric city bus drivers with an accuracy of 98.3%. In addition, the impact of different model structures on model performance such as accuracy and F_1 scores was discussed, and the performance of the proposed model was also compared with Support Vector Machine (SVM) and random forest model. The method can be used to accurately identify drivers with aggressive starting driving style and provide references for driver's safety education, so as to prevent accidents at the starting stage of electric city bus and reduce crash accidents.

Keywords: CAN bus data, deep learning, driving style, electric city bus, recognition.

✉Corresponding author. E-mail: *fuzhijun2008@163.com*

Notations

CAN – controller area network;
CNN – convolutional neural network;
DT – decision tree;
GPS – global positioning system;
HMM – hidden Markov model;
KNN – *k*-nearest neighbours;
NB – naive Bayes;
PCA – principal component analysis;

ReLU – rectified linear unit;
SAX – symbolic aggregate approximation;
SOC – state of charge;
SVM – support vector machine.

1. Introduction

Electric city buses are an important part of the urban public transport system. Its driving safety is very important for it always drives on fixed routes with large passenger flow and complex road environment. In order to facilitate passengers to ride and wait for the signal light to temporarily stop, accelerating from the stop and turning the wheel back to driving lane are the main driving conditions of electric city buses. The proportion of city bus collision accidents caused by drivers is the highest, and the proportion is as high as 80% (Han, Zhao 2020). Driving style reflects the driver's attitude and decision-making preference for road traffic safety, and directly reflects the characteristics of the driver's driving behaviour. Researchers always classify driving styles into 3 categories: cautious, normal and aggressive (Chen *et al.* 2019). Aggressive driving styles are prone to cause collisions (Ma *et al.* 2021). When aggressive drivers drive electric city buses with rapid start, high acceleration or sharp turn, the risk of collisions will be further increased, which will have a great impact on urban public transportation safety. Studying on the starting driving style is of great significance for preventing traffic accidents of electric city buses and promoting the healthy development of the new energy vehicles.

There are 3 types of driving style recognition methods, the 1st is the subjective evaluation recognition method (Hu *et al.* 2021), where the driving style is subjectively evaluated by observing the driver's driving behaviour or actually riding the vehicle to experience the driving process. This method is relatively intuitive, but the problem is that the on-site evaluation is likely to cause greater psychological pressure on the driver, and the implementation process is time-consuming, labour-intensive and costly (Magaña *et al.* 2021). The 2nd is the rule-based recognition method (Ding *et al.* 2019), which classifies the driving style according to the predefined thresholds. The drawback is that the recognition accuracy is greatly affected by the thresholds; it is very complicated and difficult to set a reasonable threshold due to the complexity and uncertainty of driving behaviours (Liu *et al.* 2020). The 3rd is based on the data driven recognition method (Bosurgi *et al.* 2013). Different styles of drivers will show different driving behaviours, such as the opening degree of stepping on the accelerator pedal, the angular velocity of turning the steering wheel, etc. This correspondingly will bring about the difference in the driving state of the vehicle, such as vehicle velocity and acceleration (Deng *et al.* 2022). Therefore, the in-vehicle CAN bus data contains a wealth of driving style information, which can be well used for driving behaviour identification (Li *et al.* 2022; Abdennour *et al.* 2021).

The basic steps of driving style recognition based on driving data can be divided into 4 steps, i.e., data collection, data processing, driving style identification and results application, as shown in Figure 1. The 1st is to collect various types of data related to different drivers for further processing. 2nd, take appropriate steps to pre-process the raw data, such as data cleaning, slicing, and regularization, to ensure that the data is of high quality and in an appropriate format. Besides, certain methods (such as correlation coefficient method, PCA, etc.) are adopted to screen and reduce the dimension of high-dimensional feature data to reduce the redundancy of data information. Then, based on driving data, key characteristics of each driver's driving process are captured to train an efficient and accurate driving style recognition model. Finally, the model is applied to identify and improve the driver's driving style.

The task of driving style recognition is to analyse a series of time series data, capture its key features, and assign the matching driving style, which can be regarded as a multivariable time series classification problem with driving style as a label. It is important to select key features from these actual driving data and find a combination of features (Zhao *et al.* 2022). Commonly used methods mainly include SVM (Amsalu *et al.* 2015), random forest (Zhao *et al.* 2012), NB (Bian *et al.* 2018), KNN (Li *et al.* 2020), DT (Montella *et al.* 2012), SAX (Chaovalit *et al.* 2013) and HMM (You *et al.* 2022), *k*-mean (Van Ly *et al.* 2013). However, these methods are used to identify individual driving behaviours, and it is difficult to capture the complex temporal characteristics of driving behaviours. The good news is that CNNs demonstrate great advantages in feature learning and have proved to be an effective way to identify driving behaviour, in which convolution and

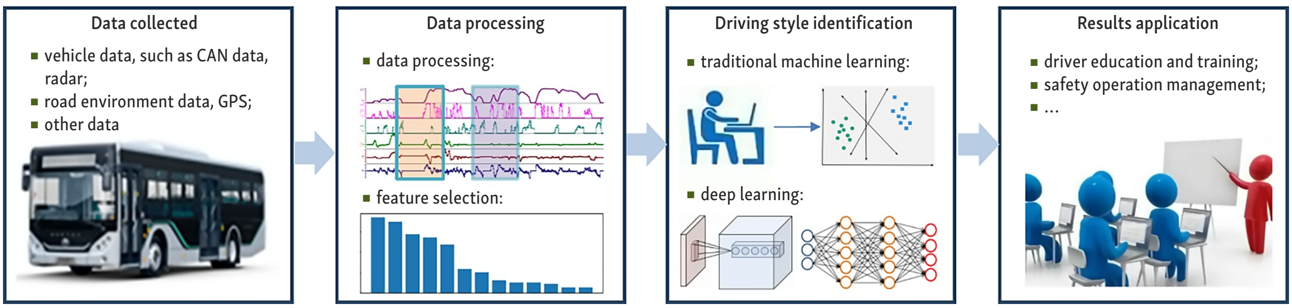

Figure 1. Process of driving style recognition

pooling operations are applied along the time dimension of sensor signals (Alkinani *et al.* 2020; Liu *et al.* 2019). This architecture can capture the time dependence of the features extracted by convolution operation. In addition, Furthermore, in most recent studies on CNNs related to action recognition, the use of 1D/2D convolutions in a single time series is often employed to capture the local dependencies of sensor signals along the temporal dimension (Sajid *et al.* 2021; Zhang *et al.* 2019b) to achieve temporal best performance for sequence data recognition. The main motivation of this study is to take advantage of CNN to design an end-to-end deep learning framework that is more suitable for processing time series CAN bus data, and to identify the starting driving style recognition of electric city buses. The proposed method can automatically identify the characteristics of driving style and extract the time characteristics without professional knowledge of feature modelling. In addition, the method can capture significant structural characteristics of high-dimensional data and explore the correlation between multi-sensor data to obtain rich feature information of driving style.

The notable contributions of the study are listed as follows:

- to the best of our knowledge, this is the 1st time to use an end-to-end deep learning framework to investigate a method for identifying the driving style during its starting stage of an electric city bus, which is a high incidence period for frontal collision accidents;
- compared with traditional machine learning methods, the deep learning method proposed in this article can automatically extract deep spatiotemporal features of multi-channel time series data, and achieve end-to-end data processing with higher accuracy and generalization ability;
- the proposed deep learning framework can perform end-to-end training without any function selection and directly process raw on-board CAN bus data through simple pre-processing, making it universally applicable.

The remainder of this article is organized as follows. Section 1 is introduction. Section 2 details the data collection process. Section 3 analyses the data processing. In Section 4, the proposed starting driving style recognition based on the deep learning is given. Besides, implementation and results are also presented. The discussion is carried out in Section 5, and the conclusions for the whole article are drawn in Section 6.

2. Data collection

This article collects the CAN bus data of an electric city bus using data sampling instrument (model type: *Kvaser Memorator Professional HS/HS*) in the natural driving state as the driving data source. The bus travels a route of about 12 km one-way, passing through 16 platforms, 21 intersections, and 24 traffic lights, as shown in Figure 2. A total of 10 electric city buses operates on this route every day, all models are ZK6126BEVG, and the vehicle length is 12 m. The bus is equipped with GPS, gyroscope, angle, displacement, radar, camera and other sensors and the vehicle CAN bus system, as shown in Figure 3. The system can collect data from nearly 100 channels, such as vehicle ON fire signal, gear position signal, GPS position, vehicle velocity, accelerator pedal opening positon, brake pedal opening position, steering wheel angle, longitudinal acceleration, lateral acceleration, passenger door switch signal, CAN bus data such as motor torque, motor speed, battery SOC, etc. A total of 24 drivers are responsible for driving the vehicle for operation. The age distribution ranges from 21 to 50 years old, and the driving experience ranges from 0.5 years to 25 years. The drivers are numbered 1…24.

3 professional evaluators (marked as A, B and C) adopted the subjective evaluation method to observe the driver's driving behaviour during the starting process by taking a passenger car on the spot, and experience the change range of the velocity, acceleration and other motion states of the passenger car. The starting driving styles of the 24 drivers were independently evaluated, and the driver states were marked according to 3 types of states (cautious, normal and aggressive, corresponding states: 0, 1 and 2). In order to ensure the reliability of the evaluation results, the principle of "the same status of the 3 evaluation marks as the final result" is adopted to determine the final evaluation results of the starting driving style of 24 drivers, as shown in Table 1.

As can be seen from Table 1, according to the above evaluation principle, the evaluation results of 3 evaluators from a total of 16 drivers (serial numbers 2, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 17, 18, 19, 21, 23) are consistent, which means that the final evaluation results of these drivers meet the requirements for driving data selection. The subjective evaluation results of the 3 evaluators of the remaining 8 drivers controversial on driving style. Considering the deep learning model is a supervised learning model that requires a large number of high-quality labelled driving data samples for training, in order to ensure the data sample quality for supervised learning of recognition models, so the driving data of the 16 drivers is adopted for training the driving style recognition model and the driving data of the other 8 drivers is not used.

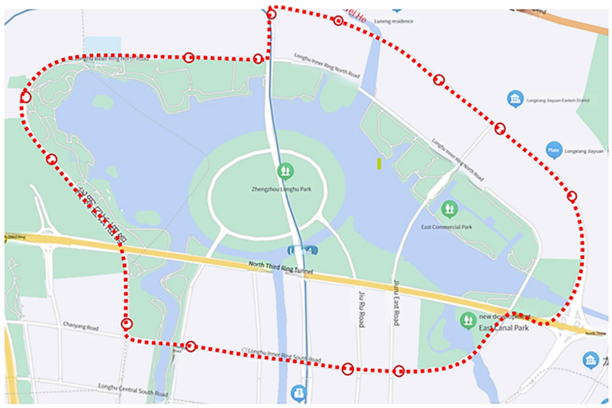

Figure 2. Driving routes of electric city bus

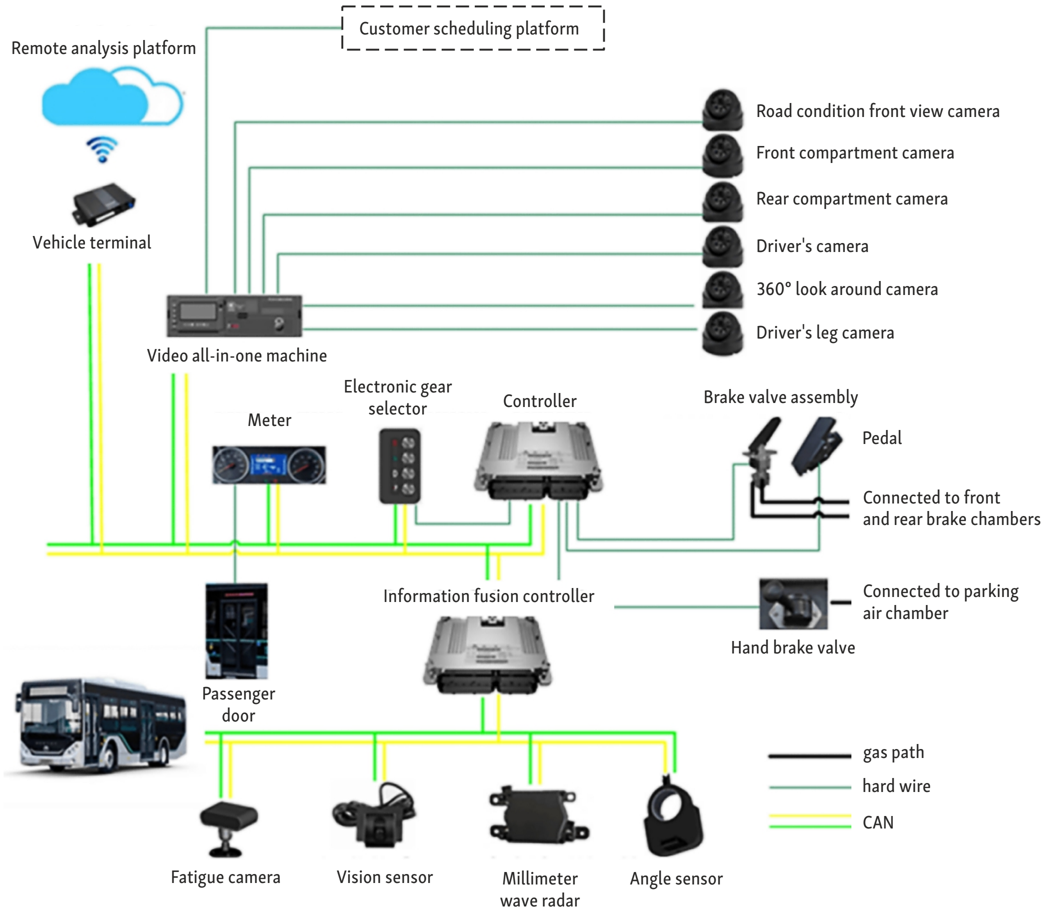

Figure 3. Sensors of electric city bus

Table 1. Evaluation results of driver driving style

Result	Number of drivers																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	0	1	1	2	1	0	1	0	1	2	0	1	2	0	1	2	1	1	0	0	0	1	2	1
B	0	1	2	2	1	0	1	0	2	2	0	1	2	0	1	1	1	1	0	1	0	2	2	0
C	1	1	1	1	1	0	1	0	2	2	0	1	2	0	1	2	1	1	0	1	0	2	2	0
Final	–	1	–	–	1	0	1	0	–	2	0	1	2	0	1	–	1	1	0	–	0	–	2	–

3. Data processing

This article selects the CAN bus data collected during the natural driving of electric city buses by 16 drivers from October 2020 to November 2020 as the basic data source for building the driving style recognition model. The original feature data is cleaned, filtered, segmented, regularized (normalized), and labelled. The article applies data cleaning to reduce data noise with the methods such as data resampling with equal time intervals, missing data interpolating using the front and back neighbouring means, deleting outlier data, and focuses on feature data screening, normalization, segmentation and labelling:

- ***data screening***: this article mainly studies the driving style in the process of starting driving. In order to filter the characteristic data highly related to the starting driving style, combined with professional knowledge in the field of engineering, 7 characteristic data reflecting the characteristics of the starting process were preliminarily selected, including vehicle velocity, accelerator pedal position, brake pedal position, motor output torque, longitudinal acceleration, lateral acceleration, steering wheel angle. Pearson correlation coefficient was used to test the correlation between feature data. The correlation coefficients between the 7-feature data are shown in Figure 4;
- ***data normalization***: data normalization is done to unify the data scales of all features each of which has a different range such that the recognition model treats them equally. Usually, it is the process of transforming the features into values between 0 and 1. Normalization of a unique feature data $X(t)$ is defined (García *et al.* 2015), such that:

$$X(t) = \frac{X(t) - X_{\min}}{X_{\max} - X_{\min}}, \quad (1)$$

where: X_{min}, X_{max} represents the minimum and the maximum values of the feature data $X(t)$. This normalization equation is applied to all the CAN data before loading them to the model;

- ***data segmentation***: the data resampling and dynamic sliding window method are common methods for segmenting data and has been successfully applied in driving behaviour recognition. In order to ensure that the multi-channel data has the same spatiotemporal correspondence, the feature data is resampled before data segmentation, with a sampling period of 50 ms.

  Another step of data preparation is performed applying the sliding window technique. CAN bus data is not a single data point. Instead, it can be considered as a sequence of values ordered in time, i.e., time series. Sliding window is a well-established method to be used for data segmentation, especially in time series, in order to guarantee the proper data segmentation. The time series data are divided into continuous data segments employing a fixed window size and a time step (i.e., the number of samples the window is shifted over the data). As shown in Figure 5, sliding window employs overlapping windows to better capture the progressive nature of the time series data and to extract the underlying features of the data. In this article, sliding window method is used for feature data segmentation, and the method recommended by Zhang *et al.* (2019a) is used to determine the key parameters, i.e., window length and sliding step size. At the same time, with the statistical results of the starting driving period as reference, considering effectively capture the characteristic information of driving data, the window size is set to10 s, and the sliding step of the window is set 1 s. Therefore, instead of having 3 segments of size 10 s for 30 s data, the data are divided into 29 segments exploiting a sliding window with 30 s. It is worth mentioning that since the data resampling period is 50 ms in the dataset, a window size of 10 s contains exactly 201 data points. Figure 5 shows the process of segmenting the resampled vehicle velocity, accelerator pedal position, brake pedal position and longitudinal acceleration data using a dynamic sliding window;
- ***labelling***: in order to facilitate the training of the recognition model, the labels are set by one-hot encoding for 3 types of starting driving styles (cautious, normal and aggressive), as shown in Table 2. According to the driving style evaluation final results in Table 1, the driving data of 16 drivers are marked with corresponding driving style labels.

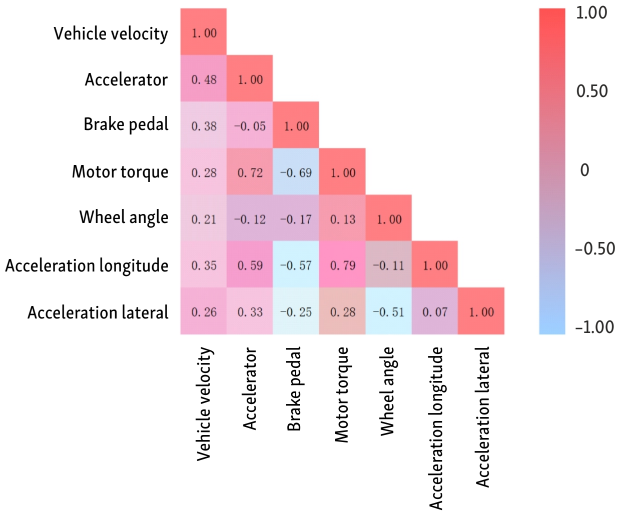

Figure 4. Correlations between the feature data

Table 2. Labels of starting driving styles

Style category	Labels
Cautious	1 0 0
Normal	0 1 0
Aggressive	0 0 1

4. Starting driving style recognition

The driving operation data and vehicle driving state data during the starting process of the electric city bus are multi-channel time series data, which have a certain period of reciprocation, and the data of each channel has a nonlin-

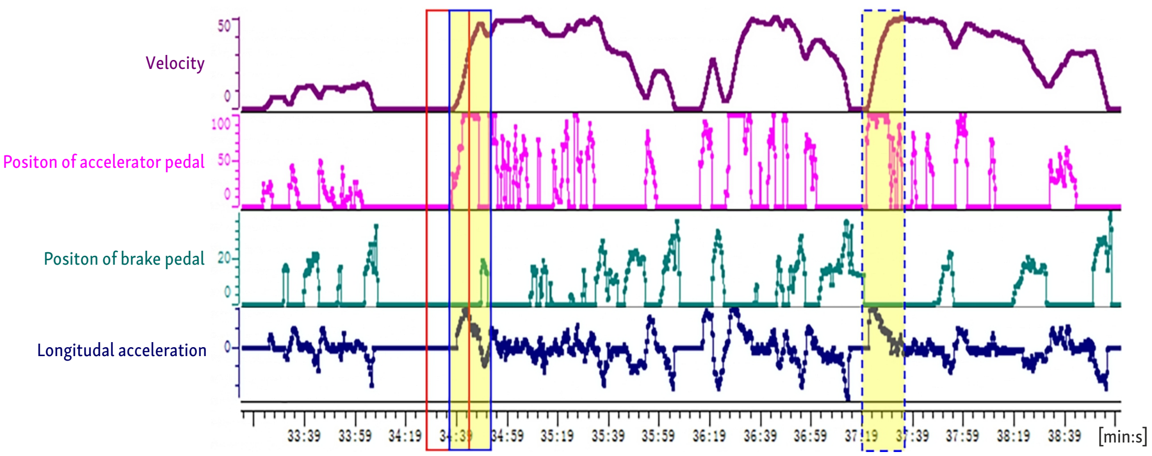

Figure 5. Dynamic sliding window slicing data

ear and complex coupling relationship. It is difficult to design some reasonable features only by expert experience. In recent years, popular deep learning models can extract deep spatiotemporal features of multi-channel time series data, and have been successfully applied in many fields (Zhang *et al.* 2021; Chao *et al.* 2020; Minhas *et al.* 2022). This article intends to build a model based on the classic CNN in deep learning to identify the starting driving style.

4.1. The proposed deep learning framework

The proposed deep learning framework in the article mainly includes an input layer, 5 convolutional layers, 6 pooling layers, a flatten layer, a fully connected layer and an output layer, as illustrated in Figure 6.

The input layer is the interface for data input. After data pre-processing, each data record contains one-dimensional multi-channel time-domain data records of a certain length of time. In the convolution layer, the convolution checks the output of the previous layer for convolution, extracts the spatial features of the local region, and obtains the feature map with a certain width, height and depth. In this process, a nonlinear activation function is generally used to construct the output characteristics, and its mathematical model is described as (Zhang *et al.* 2021):

$$y_i^{l+1}\left(j\right) = w_i^l \cdot x^l\left(j\right) + b_i^l; \tag{2}$$

$$z_i^{l+1}\left(j\right) = f\left(y_i^{l+1}\left(j\right)\right), \tag{3}$$

where: w_i^l represents the weight of the ith convolution kernel in layer l; b_i^l denotes the offset of the ith convolution kernel in layer l; $x^l\left(j\right)$ represents the input of the ith neuron in layer l; $y_i^{l+1}\left(j\right)$ represents the input of the jth neuron in layer l + 1; $z_i^{l+1}\left(j\right)$ represents the output of the ith kernel of the jth neuron in l + 1. The symbol represents the dot product of the convolution kernel with the local region.

After the convolution operation, the logical value output of each convolution is nonlinearly transformed through an activation function, such as *ReLU* function, to transform the originally linear inseparable multi-dimensional features to another space, so as to enhance the linear separability of these features. The *ReLU* function is expressed as:

$$f\left(x\right) = \begin{cases} x, & \text{if } x > 0; \\ k \cdot x, & \text{if } x \le 0, \end{cases} \tag{4}$$

where: k represents the functional gradient and is a random variable taken from the probability model of continuous uniform distribution.

The purpose of the pooling layer is to reduce the complexity of the output, prevent over fitting of the data, and reduce the amount of computation by reducing the length of the data by down sampling. Usually, the maximum or average value of the receptive field is used as the output of the feature map. The flattening layer flattens the output features of the last pooling layer into a one-dimensional vector, which is used as the input of the fully connected layer, and then establishes a fully connected network between the input and output of the fully connected layer. Fitting mapping is performed between the extracted feature information and the output layer. The output layer is the output of the fully connected layer activated by the *Softmax* function. Each output node corresponds to a type of label, and the output result is the probability value of each type. The sum of the probabilities corresponding to each type of label is equal one, and the label corresponding to the maximum probability value is taken as the result of identification and classification.

In this article, the CAN bus data of electric city bus are used to identify the starting driving style. The collected CAN bus data of electric city bus driving operation and vehicle driving state are fed input into the model after pre-processing. The potential features can be extracted through convolution pooling operation, identified and classified by neural network, and then output the recognition results of starting driving style, so as to avoid feature extraction of signals before neural network recognition, and realize end-to-end starting driving style recognition. The input data mainly includes 4 channel time-synchronized sequence data of vehicle velocity, longitudinal acceleration, accelerator pedal opening, and brake pedal opening. The data of each channel is fused and feature extracted through convolution layer and pooling layer. The size of the convolution kernel used in the 1st convolutional layer is 2, and the sliding window step size of the convolution kernel is 1. Convolution and maximum pooling operations are performed on the time series data of different channels, and then multiple convolution layers and maximum pooling are performed. The mapping feature data is obtained through average pooling, so as to complete the fusion and feature extraction of driving

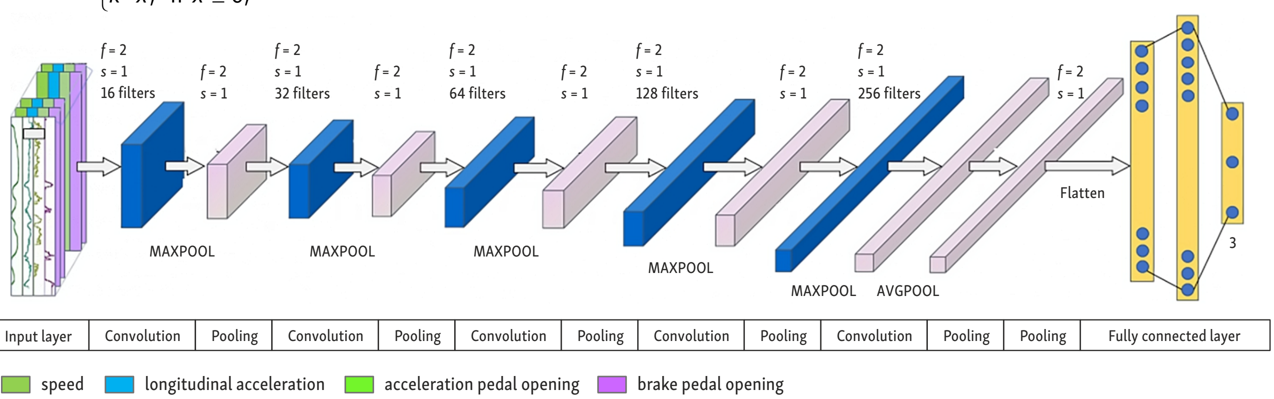

Figure 6. The proposed deep learning framework

operation data and driving state data. Finally, the feature data is flattened and transformed into one-dimensional vector data, which is input into the full connection layer of CNN. The main parameters of the proposed CNN model are shown in the Table 3.

4.2. Model training and evaluation

The vehicle driving data of 16 drivers for 5 weeks was selected as the data source, and the time series data of accelerator pedal position, brake pedal position, vehicle velocity and longitudinal acceleration corresponding to each driver's starting condition was extracted to form a sample data set. Finally, 50824 samples were obtained, each containing 201 time series data values in 4 channels, corresponding to a duration of 10 s. The samples are divided into 3 groups, which are used for model training, model validation and model testing respectively. The data distribution ratio between the training samples, validation samples, and test samples is about 50%, 30%, 20%, and the number of samples in each group is shown in Table 4.

For the multi-classification, the best loss function of the CNN model is the cross-entropy, which measures the distance between the predicted probability distribution by network and the true distribution of the labels. The cross-entropy J for multi-class classification is calculated by:

$$J = -\frac{1}{S} \cdot \sum_{s=1}^{S} \sum_{c=1}^{C} 1\left[pred^{(s)} = c \right] \cdot \log\left(p_c^{(s)} \right), \tag{5}$$

where: S denotes the total number of samples; $1[\cdot]$ represents an indicator function returning 1 for true condition or 0 for false condition; $pred^{(s)}$ denotes the prediction result for the sth sample; C denotes the number of driving style types; $p_c^{(s)}$ denotes the probability that the sth sample bellongs to type c.

Adam optimization algorithm is adopted to update the network model parameters during model training. Adam is an optimization algorithm that excels at solving non-convex optimization problems containing large-scale data and parameters. It can efficiently iterate and update neural network model parameters based on training data. The learning rate is set to 0.001. The random deactivation regularization method is introduced in the full connection layer to avoid over fitting the training data, and the dropout ratio is 0.5. Dropout assigns zero weights by randomly selecting neurons in the network to reduce the sensitivity of the neural network to small changes in the data and further improve the accuracy of invisible data processing. In this article, the *ReLU* function is used as the activation function to improve the convergence speed and reduce the over fitting phenomenon. The batch size is set to 100. We select all these settings and parameters after various trial and error experimentation. The impact of dropout rate on model performance as a sample is discussed in Section 5.2 in detail.

The performance of deep learning model needs to be evaluated by a series of indicators (Hastie *et al.* 2001), including accuracy, F_1 score and loss function value et al., which can be calculated by Equations (6), (7) and (5), respectively:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}; \tag{6}$$

$$F_1 score = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}, \tag{7}$$

where: TP (true positive) is the cases where the prediction and the target are both true; TN (true negatives) is the case where the prediction and the target are both false; FP (false positive) is when the prediction is positive while the target is false; FN (false negative) is when the prediction is negative while the target is true.

Table 3. Parameters of CNN model

Layer		Data size	Kernel	Number	Step	Padding
No	Name					
1	Input layer	201×4×*n*	–	–	–	–
2	Convolution 1	201×4×16	1×2	16	1	same
3	Pool layer 1	100×4×16	2×1	–	1	–
4	Convolution 2	100×4×32	2×1	32	1	same
5	Pool layer 2	50×4×32	2×1	–	1	–
6	Convolution 3	50×4×64	2×1	64	1	same
7	Pool layer 3	25×4×64	2×1	–	1	–
8	Convolution 4	25×4×128	2×1	128	1	same
9	Pool layer 4	12×4×128	2×1	–	1	–
10	Convolution 5	12×4×256	2×1	256	1	same
11	Pool layer 5	6×4×256	2×1	–	1	–
12	Pool layer 6	3×4×256	2×1	–	1	–
13	Flattening layer	3072×1	–	–	–	–
14	Full connection	1000	–	–	–	–
15	Output layer	3	–	–	–	–

Table 4. Classification of driving style sample data

Style category	Samples	Training samples	Validation samples	Test samples	Sample size
Cautious type	12038	6019	3612	2407	201×4
Normal type	28862	14431	8658	5773	201×4
Aggressive type	9924	4962	2977	1985	201×4
Subtotal	50824	25412	15247	10165	–

4.3. Results and model validation

The 1st group is taken as the training sample, and the 2nd group is taken as the validating sample, which is fed into the CNN recognition model for training. The accuracy of model training and the change of loss function value are shown in Figure 7. As can be seen from Figure 7, the CNN model proposed in this article can accurately identify the starting driving style of electric bus. In the process of 30 iterations of the model, the accuracy increases rapidly and converges to one stable value 98.3% nearly, and the value of loss function decreases rapidly near zero. The result of F_1 score is 98.2%. It indicates that the model has remarkable effect on feature mining and learning of driving data, and it can accurately recognize the starting driving style of electric city bus.

In order to further verify the performance of the deep learning recognition model proposed in this article, the 3rd group of sample data is used as the test data set to verify its generalization and validity. The 3rd group including 10165 samples were randomly fed into the CNN model as the sample data of unknown tags for starting driving style prediction, and the corresponding starting driving style identification results of each sample data were obtained. The results were compared with the real tag state and the confusion matrix was drawn as shown in Figure 8.

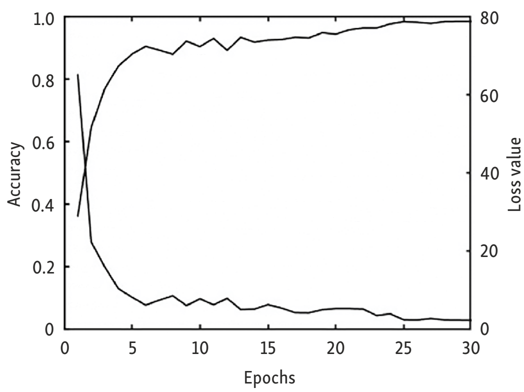

Figure 7. Model accuracy and loss value

True classification	Forecast classification: cautious type	normal type	aggressive type
aggressive type	2364 98.25%	42 1.75%	0 0%
normal type	74 1.30%	5572 98.19%	29 0.51%
cautious type	0 0%	16 0.81%	1969 99.19%

Figure 8. Confusion matrix of the test data set

As can be seen from Figure 8, although there is a certain confusion among cautious, normal style and aggressive driving style, but the average accuracy is still over 98%, and the accuracy of the aggressive driving style reaches at 99.19%, indicating that the method proposed in this article demonstrate good effect in identifying aggressive starting driving styles for electric city bus.

To sum up, in practical application scenarios, the operation data and vehicle status data during the natural driving process of different drivers can be selected, and after data pre-processing, they are fed into the CNN starting driving style recognition model proposed in this article, then the starting driving style categories of electric city bus drivers can be identified, which can provide a reliable reference for the maintenance managers of urban bus operating companies to carry out driver safety management in a targeted manner.

5. Discussion

5.1. The impact of network structure on model performance

To investigate the impact of the number of convolutional and pooling layers of network structure on model recognition performance, using the network structure in Table 3 as the benchmark Model A, the structural model without the layers of the 10th (Convolution 5) and 11th (Pool layer 5) is labelled as Model B, and the structural model without the layers of the 8th (Convolution 4), 9th (Pool layer 4), 10th, and 11th is labelled as Model C. Keep the other structural and training parameters of Model B and Model C the same as those of Model A. Model A, Model B, and Model C were trained respectively. The comparison of the recognition results of 3 model is shown in Table 6. From Table 6, it

Table 6. Comparison on performance of models

Model	Accuracy [%]	F_1 score [%]
A	98.3	98.2
B	97.8	97.6
C	92.4	93.2

can be seen that Model A has the best performance, with a maximum accuracy of 98.3%. Compared with Model B and Model C, the accuracy of Model A is better than 0.5% and 5.9%, respectively. The F_1 score of Model A are basically equivalent to the performance of Model B, and is above 5.0% better than Model C. This is mainly due to the significant reduction in model training parameters after the reduction of the network structure layers of Model C, which leads to a decrease in the learning ability.

5.2. The impact of dropout rate on model performance

To investigate the influence of training parameters on the performance of the starting driving style recognition model, the article compares the effects of different dropout rates by setting it to 0.3, 0.4, and 0.5, respectively. The accuracy and loss values with different dropout rates are shown in Figure 9. It can be seen from Figure 9, during the training iteration convergence process, the dropout rate has a certain impact on the convergence process of model accuracy and loss values. However, at the last epochs of training, the values of accuracy under different dropout rates are close each other, as are the loss values, indicating that the performance of the model is basically equivalent at this point. Comparing the loss values and accuracy corresponding to different dropout rates, it can also be found that when the dropout rate is set to 0.5, the fluctuation of the accuracy and loss values is relatively small during the training iteration convergence process.

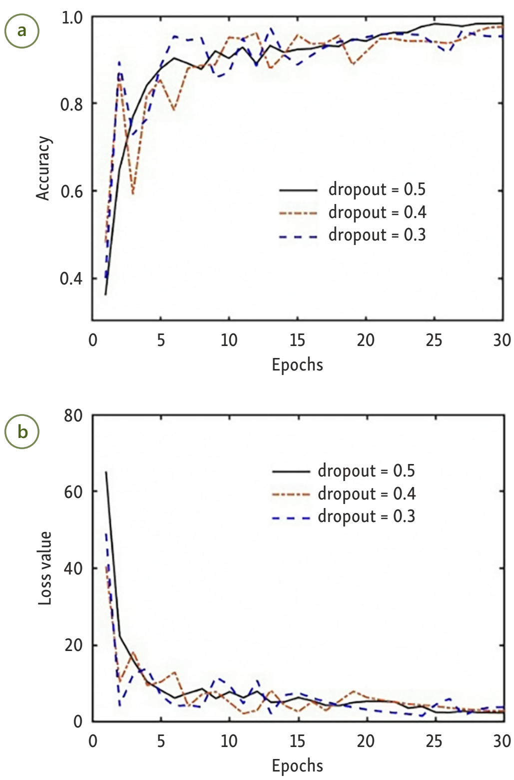

Figure 9. Impact of dropout rates parameter on model performance:
(a) – accuracy; (b) – loss value

5.3. Comparison of model performance

In order to further verify the performance of the proposed model, SVM model and random forest model was chosen for comparative analysis. The comparison results with common SVM and random forest algorithms are shown in Table 7, which shows the accuracy and F_1 score of the proposed model is higher than SVM and random forest. This is because the proposed model is better at analysing multi-channel time series driving data with automatically extract deep spatiotemporal features.

To sum up, the accuracy of model classification prediction is not only related to the quality of the neural network model structure and driving style dataset, but also closely related to the selection of model training hyperparameters and training algorithms. In practical application scenarios, the operation data and vehicle status data during natural driving of different drivers can be selected, and after data pre-processing, they are fed into the CNN starting driving style recognition model proposed in this article, then the starting driving style categories of electric city bus drivers can be identified. Based on the recognition results, the driver's starting driving style type can be determined, which can provide a reliable reference for the maintenance mangers of urban bus companies to carry out driver safety management in a targeted manner.

Table 7. Performance Comparison of models

Model	Accuracy [%]	F_1 score [%]
SVM	88.4	85.1
Random forest	96.2	92.6
Proposed model	98.3	98.2

6. Conclusions

Drivers with aggressive driving styles are more prone to traffic accidents during the starting stages of electric city buses. CNNs based on deep learning have powerful capabilities for automatically extracting features from multichannel time data and end-to-end learning of datasets. In order to prevent electric city bus accidents caused by aggressive driving styles of drivers, a deep learning model was proposed that accurately identifies the driver's starting driving style in the article. The proposed model uses some convolutional layers and pooling layers to automatically extract deep spatiotemporal features of starting driving style of electric city bus by feeding it with multi-channel time series data, then attaches a fully connected neural network to complete the prediction of starting driving style types. The main conclusions are summarized as follows:

- 1st, the CNN model based on fusion of driving operation data and vehicle state data can accurately identify the starting driving style category of electric city bus and the accuracy is near 98.3%, which proves that the proposed method is effective and feasible;
- 2nd, compared with the traditional machine learning model such as SVM and random forest model, the pro-

posed model in the article can automatically extract the deep spatiotemporal features of multi-channel time series data and realize end-to-end data processing with higher accuracy and generalization ability;
- 3rd, the performance of the model is greatly affected by its structure and key parameters of the model.

Dramatically decrease on the number of convolutional and pooling layers of the proposed model would cause a decline on accuracy. The parameter dropout ratio has some impact on the convergence process of model training.

In addition, driving style is also affected by road conditions, traffic flow levels, and complex and varied driving environments, which makes the driving styles identification highly challenging. Study on starting driving style recognition of electric city bus based on CNN provides the possibility for promoting the application of artificial intelligence technology in the field of driving safety technology for electric city buses. Next step, in order to continuously improve the application effect of this method, more experiments on the driving style recognition methods of electric city bus would be carried out, the impact of more factors such including selection of model input data, model structural parameters, training algorithms, etc. on the performance of the model should be studied further, and the corresponding safety control measures such as aggressive driver safety education should also be investigated.

Funding

This work is supported by:
- the Key Research and Development Projects of Henan Province (Grants No 221111240200, 231111241200 and 241100240300);
- the National Natural Science Foundation of China (Grant No 62073298);
- the Doctoral Fund of Zhengzhou University of Light Industry (Grants No 2021BSJJ021 and 2019BSJJ002);
- the Application for Key Scientific and Technological Project of Henan Province (Grants No 232102221036 and 242102241409);
- the Henan Foreign Scientist Studio (Grant No GZS2023011).

Author contributions

Dengfeng Zhao and *Zhijun Fu* conceived the study and were responsible for the design and development of the data analysis.

Chaohui Liu, *Junjian Hou* and *Shesen Dong* were responsible for data collection and analysis.

Yudong Zhong and *Zhijun Fu* were responsible for data interpretation.

Dengfeng Zhao wrote the 1st draft of the article.

Disclosure statement

The authors declare that there are no competing financial, professional, or personal interests from other parties.

References

Abdennour, N.; Ouni, T.; Ben Amor, N. 2021. Driver identification using only the CAN-BUS vehicle data through an RCN deep learning approach, *Robotics and Autonomous Systems* 136: 103707. https://doi.org/10.1016/j.robot.2020.103707

Alkinani, M. H.; Khan, W. Z.; Arshad, Q. 2020. Detecting human driver inattentive and aggressive driving behavior using deep learning: recent advances, requirements and open challenges, *IEEE Access* 8: 105008–105030. https://doi.org/10.1109/ACCESS.2020.2999829

Amsalu, S. B.; Homaifar, A.; Afghah, F.; Ramyar, S.; Kurt, A. 2015. Driver behavior modeling near intersections using support vector machines based on statistical feature extraction, in *2015 IEEE Intelligent Vehicles Symposium* (*IV*), 28 June – 1 July 2015, Seoul, South Korea, 1270–1275. https://doi.org/10.1109/IVS.2015.7225857

Bian, Y.; Yang, C.; Zhao J. L.; Liang, L. 2018. Good drivers pay less: a study of usage-based vehicle insurance models, *Transportation Research Part A*: *Policy and Practice* 107: 20–34. https://doi.org/10.1016/j.tra.2017.10.018

Bosurgi, G.; D'Andrea, A.; Pellegrino, O. 2013. What variables affect to a greater extent the driver's vision while driving?, *Transport* 28(4): 331–340. https://doi.org/10.3846/16484142.2013.864329

Chao, Q.; Tao, J.; Wei, X.; Wang, Y.; Meng, L.; Lu, C. 2020. Cavitation intensity recognition for high-speed axial piston pumps using 1-D convolutional neural networks with multi-channel inputs of vibration signals, *Alexandria Engineering Journal* 59(6): 4463–4473. https://doi.org/10.1016/j.aej.2020.07.052

Chaovalit, P.; Saiprasert, C.; Pholprasit, T. 2013. A method for driving event detection using SAX on smartphone sensors, in *2013 13th International Conference on ITS Telecommunications* (*ITST*), 5–7 November 2013, Tampere, Finland, 450–455. https://doi.org/10.1109/ITST.2013.6685587

Chen, K.-T.; Chen, H.-Y. W. 2019. Driving style clustering using naturalistic driving data, *Transportation Research Record*: *Journal of the Transportation Research Board* 2673(6): 176–188. https://doi.org/10.1177/0361198119845360

Deng, Z.; Chu, D.; Wu, C.; Liu, S.; Sun, C.; Liu, T. 2022. A Probabilistic model for driving-style-recognition-enabled driver steering behaviors, *IEEE Transactions on Systems*, *Man*, *and Cybernetics*: *Systems* 52(3): 1838–1851. https://doi.org/10.1109/TSMC.2020.3037229

Ding, N.; Ma, H.; Zhao, C.; Ma, Y.; Ge, H. 2019. Data anomaly detection for internet of vehicles based on traffic cellular automata and driving style, *Sensors* 19(22): 4926. https://doi.org/10.3390/s19224926

García, S.; Luengo, J.; Herrera, F. 2015. *Data Preprocessing in Data Mining*. Springer. 320 p. https://doi.org/10.1007/978-3-319-10247-4

Han, W.; Zhao, J. 2020. Driver behaviour and traffic accident involvement among professional urban bus drivers in China, *Transportation Research Part F*: *Traffic Psychology and Behaviour* 74: 184–197. https://doi.org/10.1016/j.trf.2020.08.007

Hastie, T.; Friedman, J.; Tibshirani, R. 2001. Model assessment and selection, in T. Hastie, J. Friedman, R. Tibshirani (Eds.). *The Elements of Statistical Learning*: *Data Mining*, *Inference*, *and Prediction*, 193–224. https://doi.org/10.1007/978-0-387-21606-5_7

Hu, L.; Zhou, X.; Zhang, X.; Wang, F.; Li, Q.; Wu, W. 2021. A review on key challenges in intelligent vehicles: safety and driver-oriented features, *IET Intelligent Transport Systems* 15(9): 1093–1105. https://doi.org/10.1049/itr2.12088

Li, X.-S.; Cui, X.-T.; Ren, Y.-Y.; Zheng, X.-L. 2022. Unsupervised driving style analysis based on driving maneuver intensity, *IEEE Access* 10: 48160–48178. https://doi.org/10.1109/ACCESS.2022.3171347

Li, Z.; Wu, C.; Tao, P.; Tian, J.; Ma, L. 2020. DP and DS-LCD: a new lane change decision model coupling driver's psychology and driving style, *IEEE Access* 8: 132614–132624. https://doi.org/10.1109/ACCESS.2020.3010409

Liu, W.; Deng, K.; Zhang, X.; Cheng, Y.; Zheng, Z.; Jiang, F.; Peng, J. 2020. A semi-supervised Tri-CatBoost method for driving style recognition, *Symmetry* 12(3): 336. https://doi.org/10.3390/sym12030336

Liu, Y.; Wang, J.; Zhao, P.; Qin, D.; Chen, Z. 2019. Research on classification and recognition of driving styles based on feature engineering, *IEEE Access* 7: 89245–89255. https://doi.org/10.1109/ACCESS.2019.2926593

Ma, Y.; Gu, X.; Yu, Y.; Khattakc, A. J.; Chen, S.; Tang, K. 2021. Identification of contributing factors for driver's perceptual bias of aggressive driving in China, *Sustainability* 13(2): 766. https://doi.org/10.3390/su13020766

Magaña, V. C.; Pañeda, X. G.; Garcia, R.; Paiva, S.; Pozueco, L. 2021. Beside and behind the wheel: factors that influence driving stress and driving behavior, *Sustainability* 13(9): 4775. https://doi.org/10.3390/su13094775

Minhas, A. A.; Jabbar, S.; Farhan, M.; Najam ul Islam, M. 2022. A smart analysis of driver fatigue and drowsiness detection using convolutional neural networks, *Multimedia Tools and Applications* 81(19): 26969–26986. https://doi.org/10.1007/s11042-022-13193-4

Montella, A.; Aria, M.; D'Ambrosio, A.; Mauriello, F. 2012. Analysis of powered two-wheeler crashes in Italy by classification trees and rules discovery, *Accident Analysis* & *Prevention* 49: 58–72. https://doi.org/10.1016/j.aap.2011.04.025

Sajid, F.; Javed, A. R.; Basharat, A.; Kryvinska, N.; Afzal, A.; Rizwan, M. 2021. An efficient deep learning framework for distracted driver detection, *IEEE Access* 9: 169270–169280. https://doi.org/10.1109/ACCESS.2021.3138137

Van Ly, M.; Martin, S.; Trivedi, M. M. 2013. Driver classification and driving style recognition using inertial sensors, in *2013 IEEE Intelligent Vehicles Symposium* (*IV*), 23–26 June 2013, Gold Coast, Australia, 1040–1045. https://doi.org/10.1109/IVS.2013.6629603

You, H.; Byun, S.-H.; Choo, Y. 2022. Underwater acoustic signal detection using calibrated hidden Markov model with multiple measurements, *Sensors* 22(14): 5088. https://doi.org/10.3390/s22145088

Zhang, J.; Wu, Z.; Li, F.; Xie, C.; Ren, T.; Chen, J.; Liu, L. 2019a. A deep learning framework for driving behavior identification on in-vehicle CAN-BUS sensor data, *Sensors* 19(6): 1356. https://doi.org/10.3390/s19061356

Zhang, Y.; Li, J.; Guo, Y.; Xu, C.; Bao, J.; Song, Y. 2019b. Vehicle driving behavior recognition based on multi-view convolutional neural network with joint data augmentation, *IEEE Transactions on Vehicular Technology* 68(5): 4223–4234. https://doi.org/10.1109/TVT.2019.2903110

Zhang, K.; Robinson, N.; Lee, S.-W.; Guan, C. 2021. Adaptive transfer learning for EEG motor imagery classification with deep convolutional neural network, *Neural Networks* 136: 1–10. https://doi.org/10.1016/j.neunet.2020.12.013

Zhao, C. H.; Zhang, B. L.; He, J.; Lian, J. 2012. Recognition of driving postures by contourlet transform and random forests, *IET Intelligent Transport Systems* 6(2): 161–168. https://doi.org/10.1049/iet-its.2011.0116

Zhao, D.; Zhong, Y.; Fu, Z.; Hou, J.; Zhao, M. 2022. A review for the driving behavior recognition methods based on vehicle multisensor information, *Journal of Advanced Transportation* 2022: 7287511. https://doi.org/10.1155/2022/7287511