Evaluating the Efficacy of Real-Time Connected Vehicle Basic Safety Messages in Mitigating Aberrant Driving Behaviour and Risk of Vehicle Crashes: Preliminary Insights from Highway Scenarios

Nan Zhong[1], Munish Kumar Gupta[2], Orest Kochan[3], Xiangping Cheng[4,*]
[1]School of Electronics and Control Engineering, Chang'an University, Xi'an 710054, China
[2]Faculty of Mechanical Engineering, Opole University of Technology, Opole 45-758, Poland
[3]Measuring Engineering Department, Lviv Polytechnic National University, Lviv region 79000, Ukraine
[4]Applied Physics and Research Institute, Jiangxi Academy of Science, Nanchang 330029, China
zhongnan@cttic.cn; m.gupta@po.edu.pl; orest.v.kochan@lpnu.ua; [*]chengxiangping@jxas.ac.cn

***Abstract*—Connected vehicle (CV) technology has revolutionised the intelligent transportation management system by providing new perspectives and opportunities. To further improve risk perception and early warning capabilities in intricate traffic scenarios, a comprehensive field test was conducted within a CV framework. Initially, data for basic safety messages (BSM) were systematically gathered within a real-world vehicle test platform. Subsequently, an innovative approach was introduced that combined multimodal interactive filtering with an advanced vehicle dynamics model to integrate BSM vehicle motion data with observations from roadside units. In addition, a driving condition perception methodology was developed, leveraging rough sets and an enhanced support vector machine (SVM), to identify aberrant driver behaviours and potential driving risks effectively. Furthermore, this study integrated BSM data from various scenarios, including car-following, lane changes, and free driving within the CV environment, to formulate multidimensional driving state sequence patterns for short-term predictions (0.5 s) utilising the long short-term memory (LSTM) model framework. The results demonstrated the effectiveness of the proposed approach in accurately identifying potentially hazardous driving conditions and promptly predicting collision risks. The findings from this research hold substantial promise in advancing road traffic safety management.**

***Index Terms*—Connected vehicles; Basic safety messages; Advanced driver assistant systems; Intelligent vehicles; Artificial intelligence.**

I. INTRODUCTION

The rapid advancement of connected vehicle (CV) technology has opened up new perspectives in road safety, particularly in terms of perception of driving risk and collision avoidance. As vehicle-to-vehicle (V2V) communication, vehicle-to-infrastructure (V2I), and related technologies continue to evolve, they hold substantial promise for improving road safety, alleviating traffic congestion, and elevating driving comfort [1]. Consequently, research on vehicle safety within the CV realm has gained increasing attention. This study, rooted in the basic safety message (BSM) data set derived from CV under the Standard of Society of Automotive Engineers (SAE) J2735 protocol [2], [3], places a focal point on the perception of driving risk and early warnings. It meticulously selects and efficiently integrates CV standardised driving safety information, examining challenges encompassing vehicle motion data acquisition, real-time assessment of potential risk conditions, and risk alerts concerning potential forward collisions, all through the fusion of vehicle-to-infrastructure (V2I) data.

In typical human driving scenarios, drivers manoeuvre their vehicles at safe distances from adjacent vehicles primarily relying on their visual acumen and subjective judgments. In emergency situations, drivers engage in a sequence of braking manoeuvres to rapidly increase the safe distance from oncoming vehicles, thus immediately mitigating collision risks. In light of this, our study posits that emergency braking serves as a straightforward criterion reflecting the likelihood of a vehicle collision or a pre-collision scenario. Predicting and promptly warning the driver of such situations can significantly reduce the probabilities of accidents.

To address this challenge, we thoroughly gathered BSM data and integrated them into a real-world connected vehicle test system, capturing vehicle dynamics, motion status, and driver behaviour. In addition, artificial intelligence

Manuscript received 13 November, 2023; accepted 6 January, 2024.

algorithms were employed to forecast driver behaviour patterns and assess vehicle collision risks. In particular, a multidimensional long short-term memory (LSTM) model proved particularly adept at extracting meaningful insights from a continuous series of BSMs for predicting acceleration/deceleration due to emergency braking. The findings of this study lay the foundation for the advancement of the convergence of transportation technology, information technology, and automotive technology. Specifically, they serve as a reference point to enhance the precision and real-time capabilities of driver risk warning systems.

The remainder of this paper unfolds as follows. Section II conducts a comprehensive review of the relevant literature, elucidating the objectives of the study. In Section III, we introduce the proposed methodology and modelling framework. Section IV provides an in-depth description of the data sets utilised and delineates the detailed data processing procedures. Section V engages in a discussion of the results obtained, and, lastly, Section VI draws our conclusions.

II. LITERATURE REVIEW

Traditional vehicle trajectory tracking methods are mainly based on the combination of the vehicle dynamics model and Kalman filter algorithm. The three-dimensional kinematics model, based on the traditional vehicle dynamics model, mitigates the significant trajectory error often encountered when processing a single scene. These models have been enhanced through integration with a layered trajectory tracking system that incorporates an interactive particle filter [4], [5]. Kluga, Kluga, and Vecvagars [6] proposed a low-cost complex navigation system for land vehicles, which effectively improves the robustness of vehicle state estimation. Karamat, Atia, and Noureldin [7] introduced an improved error model for reduced inertial sensor systems (RISS) that takes into account vehicle tilt and accelerometer observation errors. The proposed method was tested on real trajectory data collected in the environment of GPS signal attenuation, and the navigation performance improved significantly compared to the traditional GPS system. To reduce the computational complexity of accurate positioning, Li, Gao, Zhang, and Qiu [8] used roadside devices to improve auxiliary positioning. Real-time update of the vehicle positioning at the lane level was implemented based on the Bayesian model using received signal strength (RSS) data used in all connected vehicle networks. The results showed that better measurement accuracy was achieved under the premise of reducing computational complexity in real time. Zhang, Hinz, Gulati, Clarke, and Knoll [9] developed a method for cooperative positioning of vehicle infrastructures based on the filter symmetric metric equation (SME) to solve the problem that large errors may occur even if there is uncertain or even missing associated data of observed vehicle targets in the vicinity of CV. Unknown associated measurement data can be converted into symmetric measurement equations to estimate the corresponding states, effectively solving data association problems in vehicle infrastructure scenarios.

In summary, existing movement state tracking methods mainly focus on autonomous vehicle positioning and collaborative positioning by roadside devices. In autonomous vehicle positioning, better filtering models are usually sought to achieve smaller tracking errors. Regarding collaborative positioning, the V2I collaborative positioning occupies an increasingly important position through the gradual application of CV, and the V2I collaborative positioning method compensates for the possible shortcomings in the perception of information from vehicle sensing.

In addition, a substantial number of sensors onboard have been used to detect acceleration, braking, and steering events in moving vehicles. Daza, Bergasa, Bronte, Yebes, Almazán, and Arroyo [10] proposed a method to detect driver fatigue in real time, and input indices were based on the driver's physiological state and driving behaviour. The sample data in the simulation environment were collected from advanced driver assistance systems (ADAS). The results confirmed that the driver fatigue detection index was within a narrow range of the established threshold. Bergasa, Almería, Almazán, Yebes, and Arroyo [11] designed an APP to monitor inappropriate driving behaviour. When sensor observations exceeded a preset threshold based on experience, they were input into a fuzzy set to assess whether the incidents were triggered. Ly, Martin, and Trivedi [12] used a conventional support vector machine (SVM) method to detect events related to vehicle movement behaviour. The optimal detection rate was 60 %, and the detection effect for acceleration events was lower than expected. Traditional methods to identify driving behaviour typically rely on information about vehicle movements, driver control changes, and psychological characteristics. However, given the rapid changes and complexity of road traffic, it is still difficult to identify dangerous driving behaviour efficiently and dynamically in real time.

It should be noted that detecting and providing early warning of the risk of vehicle collisions is critical to avoiding traffic accidents. Previous methods for assessing collision risk are mainly based on the safety distance model. The longitudinal minimum safety distance (LMSD) has been widely used as a key index to determine the risk of longitudinal collision. Wu, Peng, Huang, Zhong, and Chu [13] found that the simple LMSD model performed poorly in terms of accuracy and adaptability, and developed a fuzzy inference-based LMSD model. Meyer [14] concluded that the degree of collision risk can be better measured by calculating the safety time based on the safety distance. Therefore, safe speed is also important. An algorithm to monitor the speed information of autonomous vehicles has advantages in speed monitoring accuracy and energy consumption [15]. The indices used to measure safety time included mainly time to collision (TTC), time to brake (TTB), time to react (TTR), and time to right of way (THW) [14], [16]–[18]. In addition, big data are involved in various fields, and it also plays an important role in cloud computing in vehicles [19]. To investigate the emergence of the intention of drivers intention to change lanes, Thiemann, Treiber, and Kesting [20] used the vehicle lane data provided by the Next Generation Simulation (NGSIM) Open Data to perform a comparative analysis. The safe distance model and the safe time model are

widely used in adaptive cruise control (ACC), advanced emergency braking system (AEBS), and other technologies. However, uncertainty in data perception is usually ignored in the risk assessment process [21]. To solve this problem, Kim, Kim, Lee, Ko, and Yi [22] used the energy function to derive the expected motion state of the vehicle and proposed the artificial potential field to evaluate the potential collision risk of the surrounding V2V.

It is also popular to evaluate driving safety by calculating the probability of vehicle collisions. Toledo-Moreo and Zamora-Izquierdo [23] used multiple interactive models to construct the longitudinal and lateral motion of vehicles, and used different motion models to describe the motion states of vehicles. Valdés-Vela, Toledo-Moreo, Terroso-Sáenz, and Zamora-Izquierdo [24] proposed a vehicle collision avoidance system based on real-time vehicle behaviour detection using low-cost sensors and extracted behaviour rules from vehicle track data using a fuzzy logic model. Liu, Ozguner, and Ekici [25] also introduced the concept of vehicle-road cooperation in the development of a collision warning system.

In general, research to determine the risk of vehicle collisions is relatively mature. To further reduce the probability of vehicle collisions, it is necessary to maintain high precision in predicting models risks in dynamic and uncertain traffic scenarios, and it is still a challenging problem to continuously achieve higher prediction time, especially in view of the goal of "zero accidents" in autonomous driving in the future. In addition, more research is needed on how to ensure good interaction with CV.

III. PROPOSED METHOD

A. Vehicle Motion State Acquisition Model

Two basic vehicle motions, straight-ahead driving and turning, can be represented by the constant acceleration model (CA), the constant velocity model (CV), and the constant turning speed model (CT) [4], [13], [26], [27]. The interactive multi-model state-space equation describing the vehicle driving process is as follows

$$x(k+1)=\Phi_i(k)x(k)+G_i(k)(u(k)+W(k)), \tag{1}$$

where $x(k)$ is vehicle motion state, $u(k)$ is the driver control, $W(k)$ is the white Gaussian noise with zero mean and variance Q. In (1), i is the state space matrix corresponding to different dynamics models, e.g., i = 1 means that the advanced CA model is adopted, while i = 2 means that the CT model is adopted. In this study, the state matrices of CA model and CT model are denoted as $\phi_1(k)$ and $\phi_2(k)$ respectively, while the input matrices for CA and CT models are represented as $G_1(k)$ and $G_2(k)$. All of these matrices are parameterized at below:

$$\Phi_1(k)=\begin{bmatrix} 1 & 0 & T & 0 & T^2/2 & 0 \\ 0 & 1 & 0 & T & 0 & T^2/2 \\ 0 & 0 & 1 & T & 0 & 0 \\ 0 & 0 & 0 & 1 & T & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{2}$$

$$\Phi_2(k)=\begin{bmatrix} 1 & \frac{\sin\omega T}{\omega} & 0 & -\frac{1-\cos\omega T}{\omega} \\ 0 & \frac{1-\cos\omega T}{\omega} & 1 & \frac{\sin\omega T}{\omega} \\ 0 & \cos\omega T & 0 & -\sin\omega T \\ 0 & \sin\omega T & 0 & \cos\omega T \end{bmatrix}, \tag{3}$$

$$G_1(k)=\begin{bmatrix} \frac{T^2}{2} & 0 \\ 0 & \frac{T^2}{2} \\ T & 0 \\ 0 & T \end{bmatrix}, \tag{4}$$

$$G_2(k)=\begin{bmatrix} \frac{T^2}{2} & 0 \\ 0 & \frac{T^2}{2} \\ T & 0 \\ 0 & T \end{bmatrix}, \tag{5}$$

here T is the sampling interval. Suppose that the observation vector at time k is $Z(k)$, the RSU coordinate of the i^{th} roadside intelligent station is (x_i^r, y_i^r), and the position of the tested vehicle is $(x(k), y(k))$. The measurement noise during measurement is $V(k)$, and its covariance is R, $\rho_i(k)$ is the actual relative distance from the i^{th} RSU to the vehicle, $\theta_i(k)$ is the angle between the RSU radar sensor (north is the positive direction) and the vehicle under test. Then the measurement formula is shown as follows

$$Z(k)=\begin{pmatrix}\rho_i(k)\\ \theta_i(k)\end{pmatrix}=\begin{pmatrix}\sqrt{(x(k)-x_i^R)^2+(y(k)-y_i^R)^2}\\ \arctan(y(k)-y_i^R)/(x(k)-x_i^R)\end{pmatrix}+V(k). \tag{6}$$

To calculate the error in the estimation of the motion state at each step, this study compares the estimated value of the algorithm with the true value. Furthermore, the calculation results presented in this section are the average of 50 Monte Carlo trials, and the root mean square error is used to estimate the deviation from vehicle motion state tracking to evaluate whether the combined positioning method of vehicle motion state based on V2I communication can accurately estimate and predict the motion state of the vehicle. The definition of the root mean square error is shown as follows

$$RMSE=\frac{1}{n}\sqrt{\sum_{i=1}^{n}(x^*(i)-x(i))^2+(y^*(i)-y(i))^2}. \tag{7}$$

To assess the tracking efficacy of the combined positioning method for vehicle motion state based on V2I outlined in this section, a comprehensive manoeuvring scenario was chosen from the real vehicle experiments conducted on the Nanchang-Jiujiang Intelligent Highways. Moreover, an enhanced Constant Acceleration (CA) model structure is utilized, integrating the driver's behavioral control input into the vehicle kinematic model. This stands in contrast to the conventional CA kinematic model employed for predicting

the vehicle's motion states, which often neglects to fully consider the impact of the driver's intent. Typically assumed to be zero in order to ensure smooth driving, the driver's control input is incorporated into the process noise during processing in traditional approaches. As a result, a certain level of discrepancy exists between the predicted results and the actual outcomes. The refined CA model structure, detailed explicitly by Wu [13], effectively boosts the accuracy of the vehicle motion state prediction model by incorporating the driver's behavioral control input into the car kinematic model. In this study, the model time step is set to 1 s. The initial weight of each submodel was set as 0.5, and the transition probability matrix of the model was set as

$$\mu_{ij} = \begin{bmatrix} 0.99 & 0.01 \\ 0.01 & 0.99 \end{bmatrix}. \quad (8)$$

In this work, we set $X(k) = [x(k), y(k), \dot{x}(k), \dot{y}(k), \ddot{x}(k), \ddot{y}(k)]^T$, which obtained from the vehicle-mounted unit as the model input, and the $Z(k) = [\rho(k), \theta(k)]^T$, which obtained from the roadside unit as the observation quantity. As shown in Fig. 1, in the V2I environment of the vehicle network, the improved CA model, CT model, and the unscented Kalman filter (UKF)-based interacting multiple model (IMM) can be combined simultaneously with the two state variables. Furthermore, using the BSM driving state data collected in the real vehicle test, the trajectory accuracy of the proposed method on the expressway in the network environment is verified.

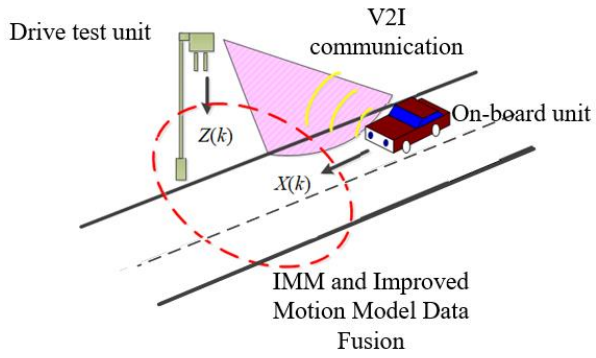


Fig. 1. Scenario of vehicle motion state collection based on V2I data fusion.

B. Driver Behaviour Identification

This study adopts the hybrid rough set and genetic algorithm (GA) to optimise the SVM to identify potentially dangerous driving states [28]. The following 11 variables will be selected as the initial input of this model: velocity v, velocity variation Δv, longitudinal velocity v_y, lateral velocity v_x, plane acceleration a, acceleration variation Δa, longitudinal acceleration a_y, lateral acceleration a_x, yaw angle ΔB, vertical acceleration a_z, and $\dot{a}_z$. Then, the rough set was applied to eliminate redundant features and select key features of control decisions on the premise of keeping the classification accuracy of original samples unchanged, thus simplifying samples and improving computational efficiency.

Furthermore, the GA method is combined to find the best value of penalty factors C_{SVM} and γ of SVM to identify the significant driver behaviour in the modelling process. The realisation process of the classification of potentially risky driving states is as follows.

1. Data preprocessing: Training samples and test samples are normalised in the interval [0, 1], respectively, and the normalised mapping formula is as follows

$$y = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \quad (9)$$

where $x, y \in R^n$ and $y \in [0,1], i = 1, 2, ..., n.$

2. Sample division: The BSM driving data corresponding to each type of driving state are divided into two groups, namely training samples and test samples.
3. Input feature selection: The rough set is used to process the selected training set, the variables that may have redundant properties are deleted, and the reduced attribute set is taken as the real input of the improved SVM.
4. Use the GA algorithm to optimise the parameters of the training set, and find the parameter CSVM and γ the best value.
5. The obtained CSVM and γ parameter SVM training model are used to predict and verify the test set.
6. Use confusion matrix to visualise the classification effect and evaluate its classification performance.

C. Vehicle Forward Collision Risk Early Warning Model

Compared with other neural networks, long short-term memory (LSTM) can solve the problem of gradient disappearance or explosion and preserve the presequence memory well. In this paper, the LSTM model is used to predict vehicle acceleration after 0.5 s using the driving state series at continuous moments as a data set. It is assumed that there are three basic hidden features, namely, vehicle acceleration, speed, and relative distance with the vehicle in front. Therefore, the sequence of three-dimensional dynamic driving behaviour $x_l(t)$ at time t can be defined as follows

$$x_l(t) = (d(t), v(t), a(t))^T, \quad (10)$$

where $a(t)$ represents the acceleration of the vehicle at time t, $v(t)$ represents the speed of the vehicle at time t, and $d(t)$ represents the relative distance between the vehicle under test and the vehicle ahead in the lane at time t.

Define the input sequence $X_l(t)$ at time t

$$X_l(t) = (x_l(t-4), x_l(t-3), x_l(t-2), x_l(t-1), x_l(t)). \quad (11)$$

Define the real output sequence $Y_l(t+5)$ of the time to be predicted at time t

$$Y_l(t+5) = y_l(t+5) X_l(t). \quad (12)$$

Finally, the prediction output at time t can be expressed as follows

$$\hat{Y}_l(t+5) = F_{\text{LSTM}}(X_l(1), X_l(2), \cdots, X_l(t)). \quad (13)$$

Thus, the inputs and outputs required to construct the

model are completed. In addition, the Adam optimisation method is used to optimise the training. Vehicle real-time deceleration can be obtained directly from the BSM data set, while TTC and THW can be obtained indirectly through the calculation of position information and speed information in the BSM data set. The calculation methods are shown in the following formula, respectively:

$$TTC = \frac{d}{v_h - v_p}, \tag{14}$$

$$THW = \frac{d}{v_h}, \tag{15}$$

where d is the relative distance between the front and rear cars, v_h is the current speed of the car, and v_p is the current speed of the car ahead. Since there is no universal standard of constant deceleration threshold for near-collision events [29], this section determines that the judgment criteria for potential collision risk at the moment are as follows:

– The "abrupt braking" state perceived by the improved SVM driving state perception method;
– THW is less than 0.5 s;
– TTC is less than 5 s.

A potential collision risk is considered if one of them is met. Furthermore, to facilitate subsequent data processing, the labels of events with potential collision risk were labelled as "1", and those without potential collision risk were labelled as "0", to build the real potential forward collision risk event database.

Through the above model, the predicted value of acceleration 0.5 s after the current time can be obtained. The next step is to select the appropriate acceleration value as the classification standard for the potential risk of collision. The optimal threshold will be selected from the LSTM acceleration prediction results using the receiver operating characteristic (ROC) curve and Youden index. During the evaluation of LSTM prediction results by the ROC curve, the area under the curve (AUC) was used as the evaluation index of the model. Firstly, the samples can be divided into True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN) according to their real category and the prediction category of the proposed model. Different sample accelerations were used as thresholds for calculation, and the corresponding true positive rate (TPR) and true negative rate (TNR) are successively taken as points on the coordinates to connect the ROC curve. The Youden index can be used to find the optimal critical value of the ROC curve [30], and can be used to find the acceleration threshold that can classify risks. Finally, the predicted value of the LSTM acceleration can be divided according to the potential risk threshold for collision. If it is below the threshold, it outputs the judgment "there is a risk of collision", and if it is above the threshold, it outputs "safe".

IV. DATA PROCESSING

A. Vehicle Motion Data Preprocessing

From the 2329 sets of driving motion data collected, a scene with a large deviation of the autonomous vehicle positioning method is selected as an example. The data must be predicted for subsequent model verification. Processing is mainly divided into two steps: the preliminary correction of the trajectory and the coordinate change. Given the error in the trajectory data caused by some confounding factors in the acquisition of the OBU data, the RSU map data can be combined to make a preliminary correction. The GPS Lane deviation appears in the data sequence and the data does not change, but the tachograph shows that the vehicle moves smoothly and forward on the expressway between 14:39:52 and 14:40:00. According to the actual type of road, (29.120715, 115.776571) are straight line segments. Therefore, you can use exponential smoothing to correct incorrect data, as shown in Table I.

Since the vehicle BSM collected by GPS is latitude and longitude information, it is based on the WGS-84 geodetic coordinate system and must be converted to a plane rectangular coordinate system before it can be input into the proposed model. In this paper, the seven-parameter method for the South China Sea is used to convert the latitude and longitude, and you get the plane rectangular coordinate data as shown in Fig. 2.

TABLE I. MODIFIED TRACK COLLECTION INFORMATION.

ID	Time	Longitude (°)	Latitude (°)	Corrected longitude (°)	Corrected latitude (°)
2080	2019/07/16 14:39:50	…29.118478	115.777417	29.118478	115.777417
2081	2019/07/16 14:39:51	29.118665	115.777348	29.118665	115.777348
	…	…	…	…	…
2093	2019/07/16 14:40:03	29.120998	115.776473	29.120998	115.776473
2094	2019/07/16 14:40:04	29.121198	115.776369	29.121198	115.776369

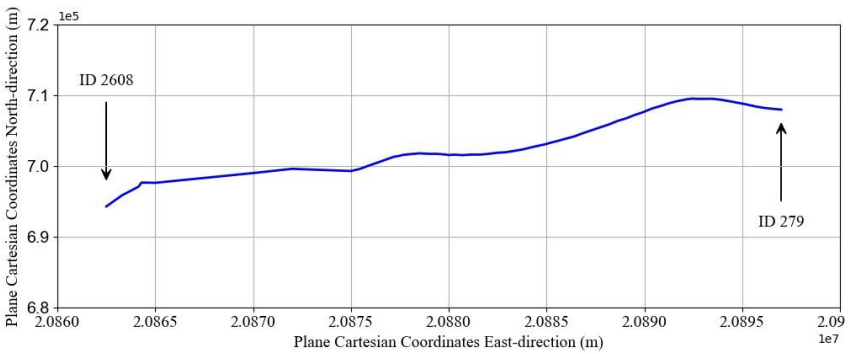


Fig. 2. Vehicle trajectory based on the plane Cartesian coordinate system.

B. *Vehicle Potential Risk Data Preprocessing*

During the driving process, drivers usually encounter driving scenes with potential risks, such as sudden acceleration or braking, lane change, and bump. Research on vehicle potential risk data is a multiclassification problem, and it is necessary to identify four driving states that can lead to risks during vehicle rapid acceleration, sudden braking, lane change, and bumpy driving

1. For sudden acceleration behaviour, the time period of increasing speed within at least 2 s was found to be a reasonable sudden acceleration event, and the specific time before and after the increase of speed was greater than 20 %, or the acceleration was greater than 2 m/s^2;
2. For sudden braking behaviour, the state of sharp deceleration is considered to be at least two consecutive moments before and after, and the vehicle speed reduction at the specific moments before and after is greater than 20 %, or the deceleration is less than -3 m/s^2;
3. For lane change behaviour, the lane change event was recorded as a lane change event when the midpoint of the nose crossed the lane line and entered another lane. The Lane change moment was identified from all BSM data, and the data between the beginning and the end of lane change were used as input for the lane change event;
4. For the phenomenon of vehicle bumps, combined with the real road conditions of the real vehicle experiment, the time is marked when the vehicle passes through the position with a large degree of fluctuation.

By selecting 2329 groups of real-vehicle BSM data, 181 groups of observed data were found to meet the discriminative conditions for the potential risk status of driving, and labelling categories were added to them, as shown in Table II.

To facilitate subsequent processing, 11 condition variables such as Δv, ΔB, v, v_x, v_y, a_z, a, Δa, a_x, a_y, a_z, etc. are, respectively, unified corresponding to the condition variable set $C = \{c1, c2, \ldots, c11\}$; the driving state output corresponds to the decision attribute d. The sample data of the driving status are shown in Table III.

TABLE II. SAMPLE PROFILES OF POTENTIALLY DANGEROUS DRIVING BEHAVIOURS.

Behaviour category	Amount	Label category (d)
Sudden acceleration	87	1
Sudden braking	15	2
Lane change	35	3
Bumps	44	4

Of all samples, 80 % of the samples are used for training and the remaining 20 % of the samples are used for testing. Since there are differences in the number of different types of operational behaviour, each category is extracted and split in a ratio of 8:2. The specific split is shown in Fig. 3.

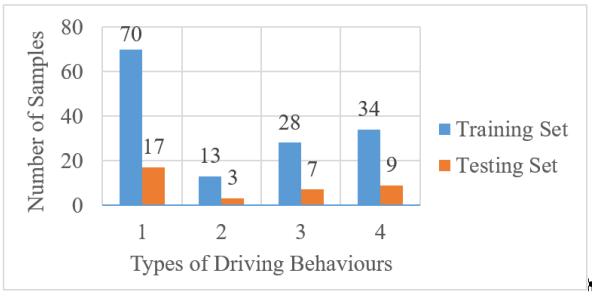


Fig. 3. Sample division.

Since the 11 initially selected variables are continuous data of the motion state generated during the driving process, the units and value ranges are quite different. Using the method of clustering of K-means, the value of the state space is divided into five intervals from small to large, and the range of values is replaced by {0, 1, 2, 3, 4}. The output d is confined to the value range {1, 2, 3, 4} as indicated in Table II. Prior to reduction, the decision table for driving state information is derived from this range. Subsequently, the table undergoes rough set attribute reduction, leading to the identification and elimination of three redundant motion states, namely c3, c7, and c11. Therefore, the eight attributes obtained are determined as the final input of the following model. Furthermore, after finalising the input of the model, according to (9), the samples of the test set can be normalised and calculated, and the results are shown in Table IV.

TABLE III. DRIVING STATE SAMPLE DATA.

	Attributes											
ID	**c1**	**c2**	**c3**	**c4**	**c5**	**c6**	**c7**	**c8**	**c9**	**c10**	**c11**	**d**
1	0.00	0.04	96.36	0.07	95.36	1.25	0.27	-0.63	0.00	0.27	11.04	4
…	…	…	…	…	…	…	…	…	…	…	…	…
57	2.12	0.17	72.79	0.22	72.79	0.26	1.83	-0.08	0.003	1.82	10.05	1
58	0.76	1.08	81.61	1.54	81.60	1.06	0.52	-0.18	0.01	0.52	10.85	4
59	0.76	-2.66	82.62	-3.83	82.53	1.07	0.71	-0.09	-0.03	0.70	10.86	4
…	…	…	…	…	…	…	…	…	…	…	…	…
181	-3.89	2.38	14.4	0.60	14.39	0.26	-2.71	1.32	-0.07	-2.71	10.05	2

TABLE IV. NORMALISATION RESULTS OF TEST SET SAMPLES AFTER FEATURE REDUCTION.

Attributes								
ID	**c1**	**c2**	**c4**	**c5**	**c6**	**c8**	**c9**	**c10**
107	0.601	0.500	0.371	0.679	0.617	0.656	0.500	0.662
108	0.441	0.498	0.344	0.867	0.382	0.431	0.508	0.434
109	0.646	0.503	0.397	0.846	0.626	0.682	0.518	0.688
…	…	…	…	…	…	…	…	…
181	0.327	0.503	0.373	0.000	0.540	0.195	0.441	0.197

.

C. Vehicle Motion Data Extraction in Typical Scenes

The test samples were also collected from real vehicles on the Nanchang-Jiujiang intelligent highways. The collection time was from 9:00 a.m. to 11:00 a.m. on July 16, 2019, and three sets of BSM data from real vehicles were selected in different time periods. To validate the model, the selected time period must include a data change between the three types of scenarios and cover as many collision risks as possible. The time, speed, acceleration, and position coordinates are all BSM base attribute data, and the relative distance from the vehicle ahead is calculated based on the BSM base data. The status of the driving scene is determined as defined in Section IV-B, where the scene of lane change is labelled "3", the vehicle following scene is labelled "5", and the free driving scene is labelled "6".

If there is no car in front of the tested vehicle, the relative distance to the vehicle in front is usually zero. To simplify data processing, it is assumed that there is a car 125 m in front of the free-running vehicle, and all relative distances that are zero values by default are replaced by 125 m in subsequent model processing. Table V shows the driving scenarios in some time periods containing the data of real vehicles changing from the pursuit state to the free-running state.

TABLE V. ACTUAL VEHICLE DATA IN SOME TIME PERIODS.

ID	Time (s)	Speed (m/s)	Acceleration (m/s^2)	D (m)	Driving scenarios
1	9:40:09.2	11.08862	0	58.36615	5
…	…	…	…	…	…
479	9:40:57.1	18.29714	0.042672	51.24602	5
480	9:40:57.2	18.2819	-0.27127	51.24298	5
481	9:40:57.3	18.20266	-1.16129	-	6
482	9:40:57.4	18.06245	-1.81661	-	6
483	9:40:57.5	17.9131	-1.46304	-	6
484	9:40:57.6	17.81556	-0.46634	-	6
485	9:40:57.7	17.79422	0.240792	-	6

V. DISCUSSION

A. Vehicle Motion State Collection Results

As can be seen in Figs. 4 and 5, speed and lane change events occurred in the data set during the period from 941 to 954. The black dotted line marked "+" in Fig. 5 is the actual lane of the vehicle during the period 941 to 954 in the data set for a total of 14 s; the blue solid line with a five-pointed star is the position estimate of the autonomous vehicle positioning method; the red solid line with a diamond is the position estimate of the method proposed in this section. In addition, the bold dotted green line is the test road lane where the vehicle is located. The lower right line is lane 1 and the upper left line is lane 2. The vehicle travels from the lower left direction of the coordinate system in the figure to the upper right direction, as shown by the direction indicated by the purple arrow. In Fig. 4, the driver starts to slow down in the sixth second in the figure and steers the vehicle in the eighth second. In the following three seconds, the vehicle gradually changes from lane 1 to lane 2, accelerating to the neighbouring lane. The above working conditions are also entered into the model, and after 100 Monte Carlo calculations, the results are shown in Figs. 6 and 7.

From Figs. 6 and 7, it can be seen that the lane change of the vehicle starts from the eighth second, the speed and position error of the comparative approach gradually become clear, the autonomous navigation and positioning of the vehicle behaviour have some hysteresis, and the V2I-based positioning method of the vehicle motion combination provides more accurate positioning due to its mature degree of real-time change of acceleration.

To test the accuracy of the proposed method in reproducing the real driving condition, the difference between the transverse and longitudinal velocities according to the proposed method and the real value is tested. First, the Kolmogorov-Smirnov (K-S) test is performed to check whether the proposed method corresponds to the Gaussian distribution, and then the possibility test of the paired T-test is performed.

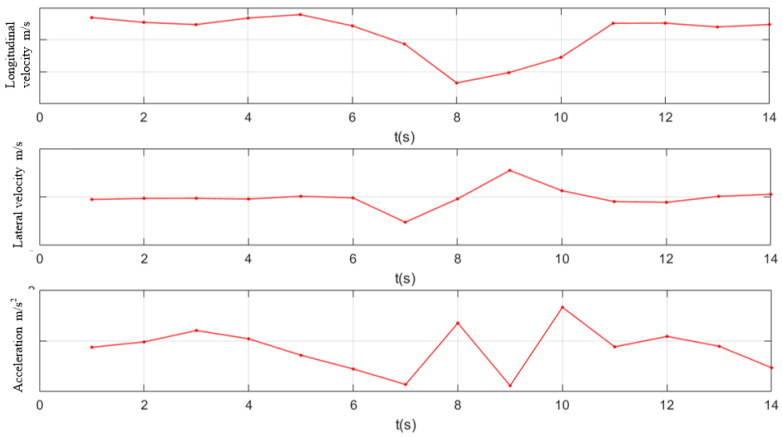


Fig. 4. Speed and acceleration conditions.

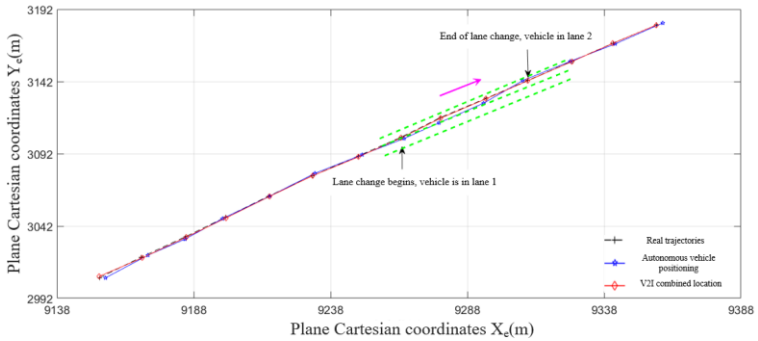


Fig. 5. Comparison of trajectory distribution.

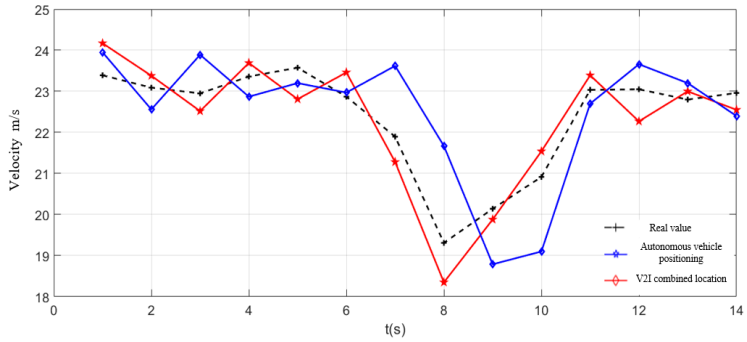


Fig. 6. Comparison of velocity tracking.

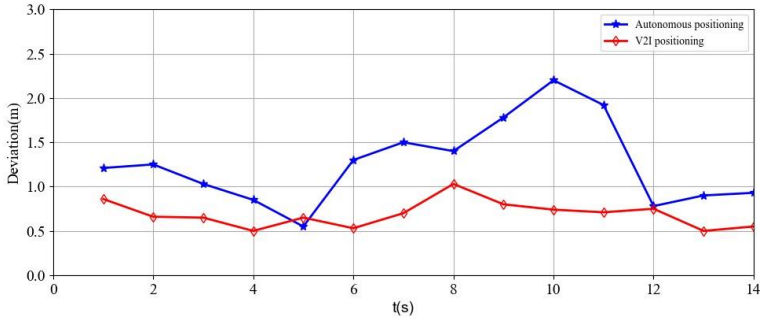


Fig. 7. RMSE deviation of position.

As shown in Table VI, the critical values of the K-S test for the four items are greater than 0.05, indicating that the random errors correspond to a Gaussian distribution. Second, for the T-test, all critical values are higher than the T-statistics, indicating that the proposed method can better reflect the motion characteristics of the real vehicle. In summary, the method proposed in this paper achieves better results in vehicle track location compared to autonomous vehicle navigation and positioning.

TABLE VI. ERROR TEST ANALYSIS.

Inspection	**Lane change**	
	Lateral velocity difference (m/s)	**Longitudinal velocity difference (m/s)**
K-S statistics	0.121	0.14
K-S test critical value	0.391	0.169
T-statistics	0.166	0.739
T-test critical value	1.828	1.953

B. Vehicle Potential Risk Determination Results

The eight attributes are entered as the training set for the improved SVM and the CSVM parameters are optimised by GA. The CSVM penalty factor value range is [0, 100], the γ value range is [0, 100], the maximum GA iteration number is set to 200 generations, the maximum population number is set to 20, the crossover probability is set to 0.9, the mutation probability is set to 0.01, and the cross-validation parameter is set to 5. Figure 8 shows the parameter optimisation process. The combination of optimal parameters is CSVMbest = 7.2313, γ = 5.3225, and the cross-validation rate is 93.7931 %, indicating that the training model has the best classification ability.

At the same time, the two parameters are entered into the SVM for calculation and 69 support vectors are obtained. The number of support vectors corresponding to each type is shown in Table VII, and the coefficients of the SVM decision function are shown in Table VIII.

TABLE VII. NUMBER OF SUPPORT VECTORS.

Traffic status type	**Number of support vectors**
1	19
2	16
3	9
4	25

TABLE VIII. SUPPORT VECTOR AND DECISION FUNCTION COEFFICIENT RESULTS.

Support vector	Coefficient	Support vector	Coefficient
x_1	0, 0, 0.0265	x_2	0, 0, 1.9208
x_3	0, 0.9829, 0	x_4	2.2326, 0, 1.0069
…	…	…	…
x_{69}	-2.6228, 0, -1.9766	-	-

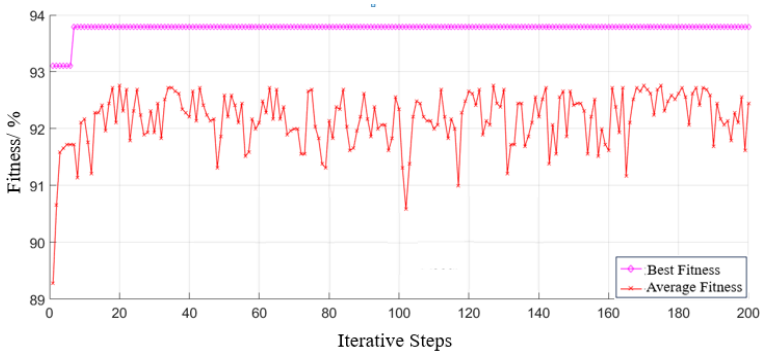
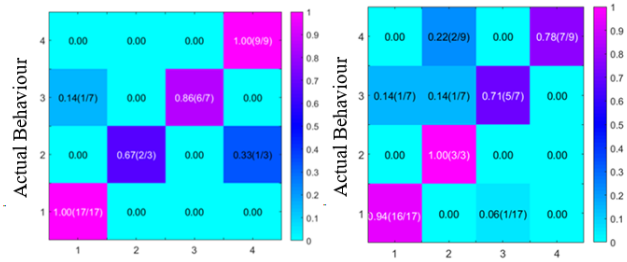


Fig. 8. Iteration of optimisation of the parameters of the genetic algorithm.

Moreover, the improved rough set model, namely the variable precision rough set (VPRS), is selected as a comparison model. Compared to SVM, VPRS can be extracted from many unordered data without providing any prior information other than that required for the problem, making it an effective tool for uncertainty classification. Therefore, VPRS is selected as the comparative model for the GA-SVM model. The VPRS model is trained with the same training set and, finally, the test set is classified. To evaluate the multiple classification results more intuitively, the confusion matrix visualisation method is used. Each column represents the driving condition category predicted by SVM and each row represents the actual driving condition category. Compared with the superiority of the receiver operating characteristic curve (ROC) in binary classification problems, the confusion matrix can perform the visualisation and evaluation of multiclassification problems well. The prediction results of the two models are, respectively, input to the confusion matrix, and the results are shown in Fig. 9.



(a) GA-SVM Prediction Behaviour (b) VPRS Prediction Behaviour

Fig. 9. Results of confusion matrix visualisation: (a) GA-SVM; (b) VPRS.

As shown in Fig. 9(a), the overall representation of the GA-SVM model is clearer and more intuitive. Of the samples originally labelled as the rapid acceleration state, 16 samples are correctly classified and one sample is incorrectly classified as the lane change state. Among the samples originally labelled as sudden braking condition, two samples are correctly classified and one sample is incorrectly classified as jerking; Among the samples originally labelled as lane change condition, one sample is correctly classified and one sample is incorrectly classified as lane change condition; Among the samples originally labelled as turbulence, all samples are correctly classified. In general, 34 of the 36 samples are correctly classified and the overall classification accuracy is 94.44 %. In comparison, 31 of the 36 samples in Fig. 9(b) are correctly classified, and the overall classification accuracy is 86.11 %. In summary, compared to the VPRS model, GA-SVM can efficiently and accurately identify potentially risky driving conditions from BSM data.

C. Vehicle Collision Warning Judgment Results

To assess the predictive effectiveness of the multidimensional LSTM neural network, the one-dimensional LSTM model is used for comparison. 60 % of each set of samples is used for training, and the rest are used for testing. At the same time, the relevant model parameters are fixed. After the appropriate modification, the learning rate is 0.01, the number of iterations is 500, and the number of hidden neurons is 20. Unlike multidimensional LSTM, one-dimensional LSTM takes only the vehicle acceleration sequence as a single variable as input, and the internal parameterisation of the model is the same as that of multidimensional LSTM. Finally, the acceleration is predicted after 0.5 s, as shown in Fig. 10.

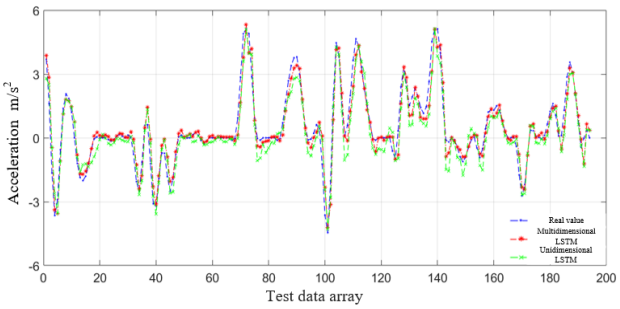


Fig. 10. Comparison of the results of the model prediction.

Figure 10 shows the results of the comparison. During this period, the general driving situation is relatively safe because the vehicle is in the following state most of the time. Between Group 50 and Group 160, e.g., the vehicle is generally in a state of intermittent acceleration, except for a brief braking near the Group 100 data. In general, the prediction results for the multidimensional LSTM and the one-dimensional LSTM are good in the steady state. However, a close examination of the data from groups 75 to 80, 105 to 110, and 140 to 150 reveals a common phenomenon: when the driver quickly

releases the gas pedal after acceleration, there is no deceleration, and the one-dimensional LSTM would misjudge the more obvious deceleration behaviour. In contrast, the multidimensional LSTM is always closer to the real condition. The multivariable LSTM can predict the acceleration of driving well in advance.

Training sets and test sets from three different time periods (different driving scenarios) are selected and combined, with a total of 779 training samples and 554 test samples. The analysis yields 668 groups of actual safety condition events and 111 groups of actual potential collision risk events in the training set. Thus, there are 668 groups of negative samples labelled "1" and 111 groups of positive samples labelled "0". In the test set, there are 480 groups of actual safety state events and 74 groups of actual potential collision risk events, i.e., 480 groups of negative samples labelled "1" and 74 groups of positive samples labelled "0". Then, the thresholds are evaluated from small to large according to the predicted acceleration values of the multivariable LSTM and the single-variable LSTM, and the ROC curve coordinates of TPR and TNR under each threshold are calculated. The results are shown in Table IX.

TABLE IX. ROC CURVE COORDINATE RESULTS.

Test result variable	Greater than or equal to this value is positive	*TPR*	*TNR*
Multidimensional LSTM	-5.79154427	1.000	1.000
	-5.48007631	1.000	0.981
	…	…	…
	5.882681435	0.000	0.000
Unidimensional LSTM	-5.6685484	1.000	1.000
	-5.3578658	1.000	0.981
	…	…	…
	5.635484	0.000	0.000

The ROC curve is shown in Fig. 11.

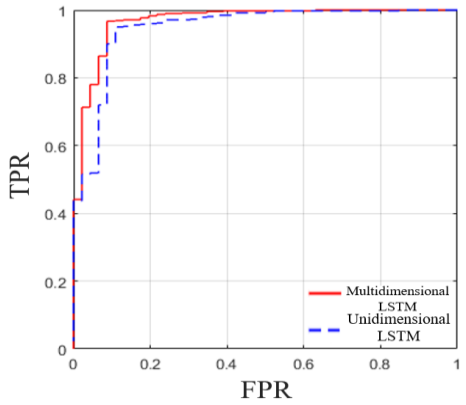


Fig. 11. Comparison of the results of the ROC curve model.

The solid red line represents the predicted value of the multivariable LSTM and the dotted blue line represents the predicted value of the univariate LSTM. The AUC and standard error of the two were further calculated as shown in Table X. It is obvious that the AUC of the multivariable LSTM (0.968) is significantly larger than that of the univariate LSTM (0.947), indicating that the selection of the multivariable LSTM prediction model has better performance.

Finally, the critical value of the ROC curve is determined by utilizing equation (10) to calculate the Youden index for each prediction generated by the multidimensional LSTM. The optimal value is then identified by selecting the threshold that corresponds to the maximum Youden index. When the acceleration threshold is configured at -1.8298, the Youden index attains its peak value of 0.786, marking the optimal critical state for the ROC. At this juncture, the false positive rate (FPR) stands at 0.949, while the false negative rate (FNR) is 0.163. Table XI illustrates the classification performance of the training set at this threshold. The overall accuracy for the training set achieves 93.332%, rising to 94.91% specifically in scenarios where the current driving condition is deemed safe. Notably, the accuracy in predicting potential collision risks within the training set reaches 83.784%. With the threshold set at -1.8298, the test set's actual data and predicted results are distinctly categorized. Therefore, this method can use BSM data to identify most potential collision risk events and give timely warnings of 0.5 s in advance.

TABLE X. AUC CALCULATION RESULTS.

Test result variable	AUC	Standard error	Progressive significance	Asymptotically 95 % confidence interval	
				Low limit	Upper limit
Multidimensional LSTM	0.968	0.014	0.000	0.940	0.997
Unidimensional LSTM	0.947	0.019	0.000	0.909	0.985

TABLE XI. TRAINING SET CLASSIFICATION VERIFICATION.

Real category	Prediction category		Accuracy
	0	1	
0	634	34	94.910 %
1	18	93	83.784 %
Overall accuracy	-	-	93.332 %

VI. CONCLUSIONS

This study conducted field tests using real vehicles on an expressway. The aim was to meticulously select the basic safety message (BSM) data from various angles and execute effective fusion techniques with the objective of improving risk perception and early warnings. The study systematically addressed several challenges: estimating and optimising the

real-time motion state of the vehicle in conjunction with vehicle and road data, real-time identification of hazardous driving behaviours within the CV environment, and providing warnings regarding the risk of rear end collisions. In this investigation, an enhanced vehicle motion model, capable of reflecting the real-time control dynamics of the driver, was integrated with an interactive multimodal approach. This hybrid model introduced a novel positioning method to determine the state of the vehicle's motion based on the vehicle-to-infrastructure (V2I) communication of CV. The BSM data set was streamlined using a coarse set, and the data were used as inputs for a support vector machine (SVM) model optimised by genetic algorithms (GA). The verification was carried out using 181 sets of real vehicle test data, revealing a model accuracy rate of 94.4 %, surpassing the comparison model VPRS, which achieved 86.11 %. The proposed method demonstrated its adaptability to dynamic driving scenarios in real vehicles and standard vehicle networking environments, thus enhancing real-time detection of potentially risky driving behaviours.

For prototypical scenarios that present potential forward collision risks, three sets of multidimensional traffic condition data were selected at different time intervals as inputs for the long short-term memory (LSTM) model, predicting the acceleration value 0.5 s ahead. Leveraging real vehicle BSM data, the prediction accuracy of the proposed model was verified within a dynamic driving environment. The receiver operating characteristic (ROC) curve was used to compare the prediction results with the actual values, producing an area under the curve (AUC) value of 0.968 for the proposed model, outperforming the one-dimensional LSTM method, which scored 0.948. Furthermore, the ROC curve and the Youden index were utilised to identify -1.8298 m/s^2 as the optimal risk threshold for predicting the acceleration of time. Consequently, the proposed method could anticipate 87.838 % of possible collision risks in advance based on CV's BSM data, thus facilitating timely warnings of collision risks. Future endeavours should emphasise intensifying real vehicle tests and expanding the pool of example data. This expansion would enable a more comprehensive exploration of complex driving scenarios such as overtaking, prolonged downhill drives, side slips, or rollovers. A wider spectrum of feature sets can be scrutinised to effectively identify, detect, or predict various driving events.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] K. Gilly, S. Filiposka, S. Alcaraz Carrasco, and A. Mishev, "Dynamic resource management of real-time edge services for intelligent vehicular networks: A case study", *Elektronika ir Elektrotechnka*, vol. 25, no. 4, pp. 58–61, 2019. DOI: 10.5755/j01.eie.25.4.23971.

[2] C. Zhai, R. Zhang, T. Peng, C. Zhong, and H. Xu, "Heterogeneous lattice hydrodynamic model and jamming transition mixed with connected vehicles and human-driven vehicles", *Physica A: Statistical Mechanics and its Applications*, vol. 623, art. 128903, 2023. DOI: 10.1016/j.physa.2023.128903.

[3] L. Sun, A. Jafaripournimchahi, and W. Hu, "A forward-looking anticipative viscous high-order continuum model considering two leading vehicles for traffic flow through wireless V2X communication in autonomous and connected vehicle environment", *Physica A: Statistical Mechanics and its Applications*, vol. 556, art. 124589, 2020. DOI: 10.1016/j.physa.2020.124589.

[4] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking. Part I. Dynamic models", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1333–1364, 2003. DOI: 10.1109/TAES.2003.1261132.

[5] C. Zhang and J. Eggert, "A probabilistic method for hierarchical 2D-3D tracking", in *Proc. of 2010 International Joint Conference on Neural Networks*, 2010, pp. 1–8. DOI: 10.1109/IJCNN.2010.5596336.

[6] J. Kluga, A. Kluga, and V. Vecvagars, "Magnetometer error models of low-cost land vehicle navigation system", *Elektronika ir Elektrotechnika*, vol. 22, no. 6, pp. 57–60, 2016. DOI: 10.5755/j01.eie.22.6.17225.

[7] T. B. Karamat, M. M. Atia, and A. Noureldin, "An enhanced error model for EKF-based tightly-coupled integration of GPS and land vehicle's motion sensors", *Sensors*, vol. 15, no. 9, pp. 24269–24296, 2015. DOI: 10.3390/s150924269.

[8] J. Li, J. Gao, H. Zhang, and T. Z. Qiu, "RSE-assisted lane-level positioning method for a connected vehicle environment", *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 7, pp. 2644–2656, 2019. DOI: 10.1109/TITS.2018.2870713.

[9] F. Zhang, G. Hinz, D. Gulati, D. Clarke, and A. Knoll, "Cooperative vehicle-infrastructure localization based on the symmetric measurement equation filter", *Geoinformatica*, vol. 20, pp. 159–178, 2016. DOI: 10.1007/s10707-016-0244-3.

[10] I. G. Daza, L. M. Bergasa, S. Bronte, J. J. Yebes, J. Almazán, and R. Arroyo, "Fusion of optimized indicators from Advanced Driver Assistance Systems (ADAS) for driver drowsiness detection", *Sensors*, vol. 14, no. 1, pp. 1106–1131, 2014. DOI: 10.3390/s140101106.

[11] L. M. Bergasa, D. Almería, J. Almazán, J. J. Yebes, and R. Arroyo, "Drive safe: An app for alerting inattentive drivers and scoring driving behaviors", in *Proc. of IEEE Intelligent Vehicles Symposium*, 2014, pp. 240–245. DOI: 10.1109/IVS.2014.6856461.

[12] M. Van Ly, S. Martin, and M. M. Trivedi, "Driver classification and driving style recognition using inertial sensors", in *Proc. of 2013 IEEE Intelligent Vehicles Symposium (IV)*, 2013, pp. 1040–1045. DOI: 10.1109/IVS.2013.6629603.

[13] C. Wu, L. Peng, Z. Huang, M. Zhong, and D. Chu, "A method of vehicle motion prediction and collision risk assessment with a simulated vehicular cyber physical system", *Transportation Research Part C: Emerging Technologies*, vol. 47, part 2, pp. 179–191, 2014. DOI: 10.1016/j.trc.2014.07.002.

[14] F. Meyer, "Time-to-collision from first-order models of the motion field", *IEEE Transactions on Robotics and Automation*, vol. 10, no. 6, pp. 792–798, 1994. DOI: 10.1109/70.338534.

[15] Y. Tang, "Monitoring algorithm for speed information of autonomous vehicles based on magnetoresistive sensor", *Jordan Journal of Mechanical and Industrial Engineering*, vol. 14, no. 1, pp. 43–52, 2020.

[16] Z. Huang, Y. He, Y. Wen, and X. Song, "Injured probability assessment in frontal pedestrian-vehicle collision counting uncertainties in pedestrian movement", *Safety Science*, vol. 106, pp. 162–169, 2018. DOI: 10.1016/j.ssci.2018.03.009.

[17] C.-Y. Kim, D. H. Wiznia, L. Averbukh, A. Torres, E. Kong, S. Kim, and M. P. Leslie, "PROMIS Computer Adaptive Tests compared with time to brake in patients with complex lower extremity trauma", *Journal of Orthopaedic Trauma*, vol. 30, no. 11, pp. 592–596, 2016. DOI: 10.1097/BOT.0000000000000645.

[18] S. M. Iranmanesh, H. N. Mahjoub, H. Kazemi, and Y. P. Fallah, "An adaptive Forward Collision Warning framework design based on driver distraction", *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 12, pp. 3925–3934, 2018. DOI: 10.1109/TITS.2018.2791437.

[19] M. S. Alsayfi, M. Y. Dahab, F. E. Eassa, R. Salama, S. Haridi, and A. S. Al-Ghamdi, "Big data in vehicular cloud computing: Review, taxonomy, and security challenges", *Elektronika ir Elektrotechnika*, vol. 28, no. 2, pp. 59–71, 2022. DOI: 10.5755/j02.eie.30178.

[20] C. Thiemann, M. Treiber, and A. Kesting, "Estimating acceleration and lane-changing dynamics from Next Generation Simulation trajectory data", *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2088, no. 1, pp. 90–101, 2008. DOI: 10.3141/2088-10.

[21] R. Schubert, K. Schulze, and G. Wanielik, "Situation assessment for automatic lane-change maneuvers", *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 3, pp. 607–616, 2010. DOI: 10.1109/TITS.2010.2049353.

[22] K. Kim, B. Kim, K. Lee, B. Ko, and K. Yi, "Design of integrated risk management-based dynamic driving control of automated vehicles", *IEEE Intelligent Transportation Systems Magazine*, vol. 9, no. 1, pp. 57–73, 2017. DOI: 10.1109/MITS.2016.2580714.

[23] R. Toledo-Moreo and M. A. Zamora-Izquierdo, "Collision avoidance support in roads with lateral and longitudinal maneuver prediction by fusing GPS/IMU and digital maps", *Transportation Research part C: Emerging Technologies*, vol. 18, no. 4, pp. 611–625, 2010. DOI: 10.1016/j.trc.2010.01.001.

[24] M. Valdés-Vela, R. Toledo-Moreo, F. Terroso-Sáenz, and M. A. Zamora-Izquierdo, "An application of a fuzzy classifier extracted from data for collision avoidance support in road vehicles", *Engineering Applications of Artificial Intelligence*, vol. 26, no. 1, pp. 173–183, 2013. DOI: 10.1016/j.engappai.2012.02.018.

[25] Y. Liu, O. Ozguner, and E. Ekici, "Performance evaluation of intersection warning system using a vehicle traffic and wireless simulator", in *Proc. of IEEE Intelligent Vehicles Symposium*, 2005, pp. 171–176. DOI: 10.1109/IVS.2005.1505097.

[26] G. Zhai, H. Meng, and X. Wang, "A Constant Speed Changing Rate and Constant Turn Rate model for maneuvering target tracking", *Sensors*, vol. 14, no. 3, pp. 5239–5253, 2014. DOI: 10.3390/s140305239.

[27] J. Sorstedt, L. Svensson, F. Sandblom, and L. Hammarstrand, "A new vehicle motion model for improved predictions and situation assessment", *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1209–1219, 2011. DOI: 10.1109/TITS.2011.2160342.

[28] G. Castignani, T. Derrmann, R. Frank, and T. Engel, "Driver behavior profiling using smartphones: A low-cost platform for driver monitoring", *IEEE Intelligent Transportation Systems Magazine*, vol. 7, no. 1, pp. 91–102, 2015. DOI: 10.1109/MITS.2014.2328673.

[29] L. Peng, M. A. Sotelo, Y. He, Y. Ai, and Z. Li, "Rough set based method for vehicle collision risk assessment through inferring driver's braking actions in near-crash situations", *IEEE Intelligent Transportation Systems Magazine*, vol. 11, no. 2, pp. 54–69, 2019. DOI: 10.1109/MITS.2019.2903539.

[30] Y. Zhu, T. Zhang, and C. Chen, "Study on probability estimation of haze in Beijing based Logistic Regression Model", *Journal of Geoscience and Environment Protection*, vol. 5, no. 6, pp. 37–41, 2017. DOI: 10.4236/gep.2017.56005.