Front. Eng. Manag. 2024, 11(1): 107–127
https://doi.org/10.1007/s42524-023-0293-x

RESEARCH ARTICLE

Cheng CHANG, Jiawei ZHANG, Kunpeng ZHANG, Yichen ZHENG, Mengkai SHI, Jianming HU, Shen LI, Li LI

CAV driving safety monitoring and warning via V2X-based edge computing system

Abstract Driving safety and accident prevention are attracting increasing global interest. Current safety monitoring systems often face challenges such as limited spatiotemporal coverage and accuracy, leading to delays in alerting drivers about potential hazards. This study explores the use of edge computing for monitoring vehicle motion and issuing accident warnings, such as lane departures and vehicle collisions. Unlike traditional systems that depend on data from single vehicles, the cooperative vehicle-infrastructure system collects data directly from connected and automated vehicles (CAVs) via vehicle-to-everything communication. This approach facilitates a comprehensive assessment of each vehicle's risk. We propose algorithms and specific data structures for evaluating accident risks associated with different CAVs. Furthermore, we examine the prerequisites for data accuracy and transmission delay to enhance the safety of CAV driving. The efficacy of this framework is validated through both simulated and real-world road tests, proving its utility in diverse driving conditions.

Received Jul. 23, 2023; revised Sep. 25, 2023; accepted Dec. 12, 2023

Cheng CHANG, Jiawei ZHANG, Jianming HU, Li LI (✉)
Department of Automation, Tsinghua University, Beijing 100084, China
E-mail: li-li@tsinghua.edu.cn

Kunpeng ZHANG
Department of Automation, Tsinghua University, Beijing 100084, China; College of Electrical Engineering, Henan University of Technology, Zhengzhou 450001, China

Yichen ZHENG, Mengkai SHI
Nebula Link Technology Co., Ltd., Beijing 100080, China

Shen LI (✉)
Department of Civil Engineering, Tsinghua University, Beijing 100084, China
E-mail: sli299@tsinghua.edu.cn

This work was supported in part by the National Key Research and Development Program of China (Grant No. 2021YFB2501200).

Keywords driving safety, accident prevention, connected and automated vehicles, edge computing

1 Introduction

Driving safety and accident prevention represent significant concerns within the domain of transportation systems and vehicle-related research (Kang et al., 2021; He et al., 2021). However, mitigating accidents remains a formidable challenge, given that even a brief lapse of attentiveness can result in severe collisions. In recent years, researchers have increasingly directed their efforts toward monitoring vehicle movements to anticipate potential hazards and proactively avert accidents.

Driving safety monitoring systems can be categorized into two distinct categories: Vehicle-side systems and road-side systems. A pivotal component among vehicle-side systems is the Advanced Driver Assistance Systems (ADASs). These systems have undergone rigorous testing in recent years. ADAS can comprehend the surrounding driving environment through sensory data, furnish the driver with feedback regarding potential risks, and, if necessary, assume control of the vehicle. Noteworthy functions include Lane Departure Warning (Cualain et al., 2012), Adaptive Cruise Control (Shang and Stern, 2021), and Lane Keeping Aid (Sentouh et al., 2019). Nevertheless, prevailing ADASs primarily rely on data collected by the vehicle's sensors. Given the intricacies and dynamism of the driving environment, achieving a comprehensive and precise understanding of driving scenarios proves to be challenging (Tapia-Espinoza and Torres-Torriti, 2013). ADAS occasionally misjudges routine situations, yielding unnecessary warnings, or overlooks potential risks, resulting in no response. Consequently, ADAS assessments tend to be myopic and may not invariably guarantee safety.

Another established monitoring system deployed on road side is the video-based system (Fernández-Caballero et al., 2008; Shehata et al., 2008; Saligrama et al., 2010).

This system assesses driving safety by leveraging vehicle information acquired via cameras installed alongside roadways. Road-side systems possess a more comprehensive grasp of the driving environment, thanks to the cameras' capability to concurrently monitor multiple vehicles. However, these methods still contend with limitations concerning monitoring range, leading to less precise predictions of driving behavior.

With the development of vehicle-to-everything (V2X) communication techniques, a new approach to this issue emerges. V2X communication modes primarily include vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication. Concerning these communication modes, many studies have concentrated on the development of V2V-based monitoring systems (Tan and Huang, 2006; Wang et al., 2011; Lyu et al., 2022) but have yet to fully exploit the potential of road-side systems. Road-side infrastructure systems offer broader fields of view, enabling the collection of a wider range of information (Jiang et al., 2021) for comprehensive monitoring. Regarding data transmission, some studies utilize raw sensory data for information exchange (Miucic et al., 2018; Figueiredo et al., 2022), such as images and point clouds acquired by cameras and Lidar. However, raw data consumption entails greater storage space and demands increased transmission bandwidth and processing time. In contrast, vehicle movement data necessitates fewer computational resources and simplifies transmission, facilitating the direct acquisition of vehicle states.

By deploying edge computing devices, road infrastructure systems (Yu et al., 2023) can effectively gather position data from Connected and Automated Vehicles (CAVs) across a wide spatial area via V2I communication. The data can subsequently be utilized to comprehensively assess the risk associated with each vehicle. The advantages of these approaches primarily include:

(1) We can exclusively collect vehicle movement and trajectory data. Dependable position and trajectory data can be furnished by the existing on-board position and navigation systems.

(2) The approach holds the potential to markedly diminish the computational time expenses associated with sensory processing since environmental data are systematically cataloged within road infrastructure databases. Consequently, concerns regarding vehicle occlusion, perceptual inaccuracies, and identification failures become superfluous. This method permits the direct consideration of surrounding vehicles.

(3) A substantial volume of position and state data can be proficiently amassed, rendering it invaluable for trajectory prediction and collision advisories.

While researchers have made endeavors to embrace V2X-based road-side systems for driving safety monitoring (Wang et al., 2020a; Jo et al., 2021; Chang et al., 2022; Miao et al., 2022), specific questions necessitate attention. Firstly, prevailing monitoring systems primarily prioritize the safety of a particular vehicle type and may struggle to accommodate different levels of vehicle automation effectively. These systems frequently overlook potential risks tied to the road environment, such as lane departures and non-compliance with traffic regulations. Consequently, existing V2X-based frameworks might lack universality and generalizability. Secondly, a comprehensive analysis of the ramifications of data measurement errors and transmission delays, crucial for the practical implementation of algorithms, is conspicuously absent. Thirdly, testing of these systems typically confines itself to a single type of simulated or real scenario, warranting further validation.

In this paper, we propose a novel monitoring framework via a V2X-based edge computing system. The contributions are outlined as follows:

(1) We have devised proficient algorithms and data structures adept at addressing diverse accident scenarios, including lane departures and vehicle collisions. This framework accommodates the examination of varying levels and categories of CAVs while accounting for inter-vehicle interactions.

(2) We conduct an in-depth evaluation of the fundamental prerequisites pertaining to trajectory data precision and transmission delay within the safety monitoring system. Subsequently, we investigate the performance characteristics and differences of algorithms tailored for different CAV levels.

(3) We have substantiated the framework's efficacy through comprehensive testing across a spectrum of driving environments, including scenarios such as ramps and intersections. The extensive evaluation comprises simulation assessments and real-world road trials, consistently affirming the system's efficiency. Notably, our findings underscore the superior warning performance of the V2X-based edge computing approach in comparison to the conventional single-vehicle-based approach.

The remaining sections of the article will be organized as follows. Section 2 will provide a comprehensive summary of the prerequisites for an optimal safety monitoring system. In Section 3, we will expound upon the proposed monitoring system and outline the algorithms designed for CAVs in rigorous detail. Section 4 will enumerate the results of our experimental trials, substantiating the efficacy of the proposed approach across a spectrum of typical scenarios. In Section 5, a comparative analysis will be conducted to assess the performance disparities between V2X-based edge computing methods and single-vehicle methods. Finally, Section 6 will draw the article to a conclusion.

2 Requirements of the desired safety monitoring framework

In essence, the necessary requirement is an outstanding

safety monitoring framework for driving, distinguished by remarkable comprehensiveness, swift computational efficiency, high precision, and strong compatibility. The particular prerequisites for this desired monitoring framework are as follows.

(1) Comprehensiveness

The monitoring system must possess the capability to comprehensively acquire vehicle data and environmental information across extensive spatial and temporal domains. Traditional methods, including the aforementioned ADAS and roadside video-based systems, exhibit limited spatial coverage, thus impeding precise motion prediction and scenario comprehension. Consequently, researchers have embarked on the development of monitoring systems grounded in the V2X communication technique. Regarding comprehensiveness, road-side units (RSUs) employing V2I communication offer broader fields of view in comparison to vehicle-centric V2V communication methods (Wang et al., 2011; Jiang et al., 2021). Furthermore, these RSUs can pre-store global map data within their control area, markedly enhancing the depth of comprehension regarding the road environment.

(2) Efficiency

The monitoring system must possess the capability to swiftly execute data collection, storage, and collision/risk prediction within an acceptable duration. To optimize time efficiency, ensuring real-time monitoring and timely warnings, the following two critical considerations should be taken into account during system design. Firstly, the selection of an appropriate data transmission format is paramount. Compared to raw sensory data, the transmission of trajectory and movement data entails reduced bandwidth, computational resource, and storage space requirements. The direct acquisition of state data additionally mitigates the risk of perceptual failures. Secondly, the system's data structures for storage and predictive algorithms must demonstrate noteworthy efficiency. The system should adeptly manage the real-time storage and extraction of vehicle data with minimal time complexity. This efficiency facilitates expeditious accident detection and the provision of prompt warnings.

(3) Accuracy

The algorithms within the safety monitoring framework must demonstrate a level of accuracy that enables precise and timely warnings. Accuracy entails not only the early detection of genuine collisions but also the mitigation of false alarms and unnecessary disruptions. Some studies have relied on Time-To-Collision threshold parametric methods (Lee and Peng, 2005; Li et al., 2016) to determine impending collisions. Nevertheless, fixed thresholds often lack precision in specific driving environments (Wang et al., 2020b). Setting the threshold too low may result in vehicles failing to perceive potential risks promptly, while setting it too high can lead to a surplus of false warnings, potentially distracting drivers. In contrast, safety monitoring based on trajectory and collision prediction yields more robust results and attains more accurate metrics.

(4) Compatibility

The safety monitoring framework should possess the capacity to accommodate CAVs operating at varying levels of automation in diverse conditions, thereby enhancing overall compatibility. Firstly, scenarios may involve the coexistence of different levels of autonomous vehicles sharing the road (Di and Shi, 2021; Li et al., 2022b). In some instances, driving patterns are well-defined, whereas in others, they remain uncertain. The system must demonstrate adaptability across a spectrum of scenarios, employing different algorithms tailored to the specific safety monitoring requirements of various CAV types. Secondly, diverse potential accident types, such as vehicle collisions and lane departures, must be included within the monitoring system's compatibility framework to ensure comprehensive vehicle safety.

In line with these outlined requirements, we propose a novel V2X-based driving safety monitoring framework. In pursuit of comprehensiveness, we have embraced a V2X communication-based edge computing system to comprehensively monitor the safety of a wide spectrum of vehicles. To enhance efficiency, we have rigorously crafted data structures and algorithms for data storage, prediction, and warning, with a particular focus on reducing time complexity. For the crucial facet of accuracy, our designed prediction and detection algorithms consistently maintain high precision, ensuring the provision of timely warnings, a fact corroborated through experimental testing. Lastly, our framework embodies compatibility through the incorporation of diverse algorithms tailored to different CAV types, including collisions with both dynamic and stationary objects.

3 Framework

3.1 Data flow and data structures

The monitoring system includes four critical functions: Vehicle and map data storage, motion prediction, accident detection, and accident warning. These four functions and their respective sub-modules are executed sequentially. Figure 1 illustrates the relevant data structures and data flow for each sub-module.

Unlike previous ADASs that collect data in an event-triggered manner, our monitoring system operates either on an RSU or in the cloud. This configuration allows for the continuous monitoring of all vehicles within a specified spatial area, facilitating the direct recording of any changes in their states. Furthermore, the system can actively receive and integrate other types of messages, including signal states and vehicle driving intentions.

To accommodate the ongoing movement of vehicles entering and departing from the control area, our system

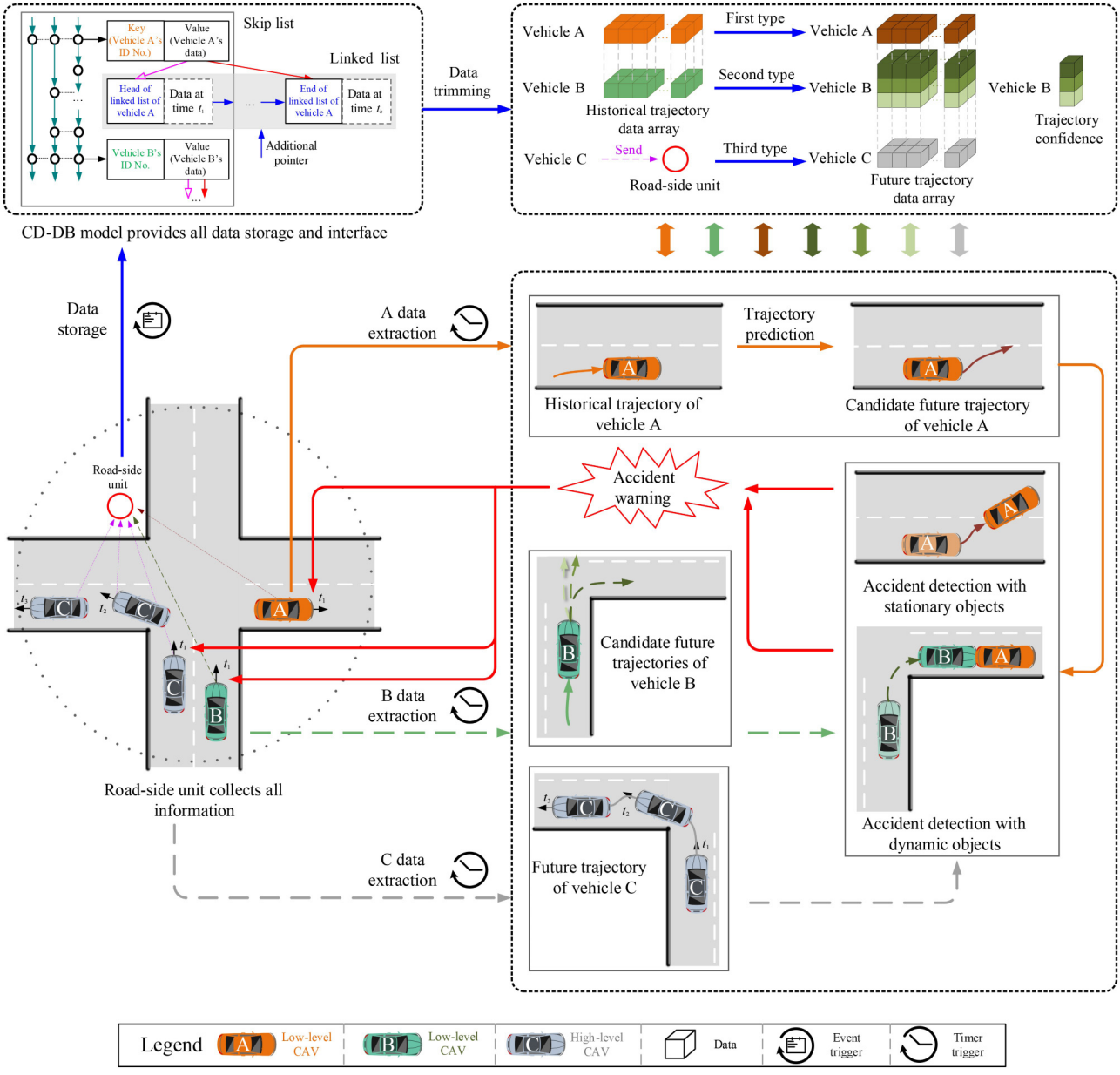


Fig. 1 The illustration for data flow in the framework.

leverages a database designed for cooperative driving to store and index raw data from CAVs. Building upon prior research (Yu et al., 2023), we have adopted the data storage model for cooperative driving (CD-DB) data structure. CD-DB consists of a skip list and multiple linked lists, effectively organizing the data of active vehicles. Within the CD-DB structure, the skip list streamlines the storage and indexing of vehicles, while concrete vehicle positions and states are housed within individual linked lists. Building upon our previous research (Yu et al., 2023), we underscore the advantages and merits of the CD-DB storage model as follows:

(1) With the composite data structure, specifically a skip list coupled with numerous linked lists, we can accomplish query and addition operations with a time complexity of $O(\log n)$, where n represents the number of CAVs within the control area. In comparison to alternative storage models, this structure exhibits notably superior performance.

(2) By segregating the storage of active vehicle data from inactive vehicle data, we ensure the real-time read and write operations for active data while efficiently storing inactive data through pointer copying rather than memory copying.

(3) As vehicle data accumulates to a certain threshold, it is transmitted to a remote cloud center for persistent storage and refined learning and analysis. Additionally, the incorporation of a data recovery mechanism enhances data security.

(4) Within the safety monitoring framework, a

noteworthy feature of CD-DB involves the inclusion of an additional pointer. This pointer references the last several nodes containing information from the past 5 to 10 s, given that the latest data is primarily pertinent for prediction. The incorporation of this specialized pointer significantly reduces unnecessary traversal time.

CAVs collect their sensor data, including global positioning system (GPS) and inertial measurement unit (IMU) information, and transmit the data to the RSU using on-board units (OBUs). The communication protocols employed for CAVs and RSUs include IEEE 802.11p, LTE, and 5G (Tahir and Katz, 2022). The raw data packages generated by CAVs are transmitted in specific protocol formats, such as XML, JSON, and protobuf (Feng and Li, 2013). Notably, the protobuf format excels at data compression through binary transfer, reducing both data volume and transmission delay. Each CAV's information is serialized into the aforementioned formats, enabling the parsing of data packages into critical data fields, as presented in Table 1. In this paper, our primary focus centers on data aggregation and algorithm utilization, along with the analysis of data measurement and transmission influences in our experiments. Detailed discussions regarding specific sensors and protocols are beyond the scope of our investigation.

Vehicles cannot guarantee the transmission of data to RSUs at regular intervals. Despite being programmed to operate periodically, the packages received by RSUs may not be precisely timing-aligned due to unforeseen circumstances. Factors such as sensor device malfunctions or data transmission congestion can disrupt the regularity of data transmission. To streamline subsequent tasks, our framework incorporates interpolation methods to ensure that the data are aligned with respect to timing. Through the theoretical derivation, the normal linear interpolation can satisfy the requirement. Suppose the received vehicle position at timestamp t_0 is (x_{t_0}, y_{t_0}), and position at timestamp t_1 is (x_{t_1}, y_{t_1}). The Lagrange interpolation remainder of the linear interpolation on x-axis is $x''(t)(t_0-t_1)^2/8$, where $t \in (t_0, t_1)$, $x''(t)$ is the acceleration on x-axis and it will not exceed 8 m/s^2 normally, and $(t_0-t_1)^2$ will be nearly 10^{-2}. So the reminder is of the 10^{-3} order, which is much less than the localization error itself. For the application of trajectory prediction and collision detection, the trimmed data can be used effectively. Moreover, the fitting errors of the trimmed data will be reflected in position errors of each CAV, and the impact will be further discussed in Section 4.

Table 1 Data fields contained in packages from CAV to RSU

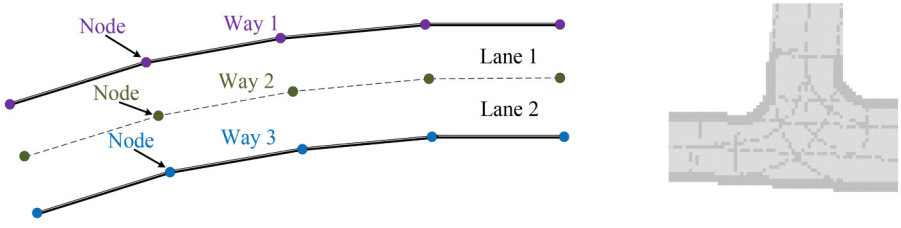
Time stamp	Vehicle ID	Vehicle type
Vehicle speed	Vehicle position x	Vehicle position y
Vehicle acceleration	Vehicle orientation	Vehicle size

The remaining sub-modules, including trajectory prediction, accident detection, and accident warning, will operate in a timer-triggered manner. When retrieving historical data for vehicles, our proposed framework efficiently utilizes the additional pointer to acquire recent historical position data.

For the road environment within the control area, the RSU will pre-store map data for subsequent use in trajectory prediction and accident detection. To ensure compatibility with mainstream map standards such as Lanelets (Poggenhans et al., 2018), OpenStreetMap (Haklay and Weber, 2008), and OpenDrive (Dupuis et al., 2010), the RSU will store the position array of boundary ways and their corresponding nodes, which collectively define the road network structure. The boundary composed of a dense sequence of discrete nodes facilitates collision detection with stationary objects. Furthermore, Bird's-Eye-View (BEV) map formats offer rich semantic information, enhancing the representation and fusion of driving environment details. Consequently, the roadside system will also pre-store binary grid occupancy matrices representing various road environment elements from a BEV. The BEV map data will play a pivotal role in extracting and learning the interactions between vehicles and the road environment during the training of prediction algorithms. Figure 2 illustrates the content of these two types of road environment data.

3.2 The trajectory prediction algorithms

The trajectory prediction function takes historical position data from vehicles as input. Subsequently, the prediction sub-module generates groups of candidate future trajectories. These output trajectories are systematically stored in



(a) Road network structure representation (b) Road environment BEV

Fig. 2 The types of road environment data.

several aligned two-dimensional arrays to facilitate subsequent detection. Furthermore, the confidences associated with the candidate trajectories are also stored in a one-dimensional array.

While some researchers have employed various methods, such as Kalman filtering (Barrios and Motai, 2011), to predict the most probable future trajectory for lane keeping, such approaches are primarily suited for simpler environments, like straight highways. However, when addressing more complex scenarios, it becomes crucial to consider the influence of multiple potential future trajectories (Zhang et al., 2022) on effective accident detection.

Based on the automation levels of CAVs, we have designed three distinct algorithms as follows.

(1) The first algorithm

The first algorithm is tailored for low-level CAVs where only the latest historical data is available. Previous research (Wang et al., 2014) has demonstrated the efficacy of polynomial curves in fitting trajectory data. In this context, we employ the polynomial extrapolation algorithm to estimate one candidate future trajectory for CAVs.

We denote the trimmed historical position data as (τ_0, x_0, y_0), (τ_1, x_1, y_1), …, (τ_n, x_n, y_n). A Bezier spline curve (Choi et al., 2008) of two dimensions has been applied for efficiently estimating the future trajectory. For example, the nth Bezier curve can be formulated as:

$$\begin{cases} x(\tau) = \sum_{i=0}^{n} \binom{n}{i} \tau^i (1-\tau)^{n-i} x_i \\ y(\tau) = \sum_{i=0}^{n} \binom{n}{i} \tau^i (1-\tau)^{n-i} y_i \end{cases} \tau \in [0,\ 1]. \quad (1)$$

The parameters of the Bezier curve can be derived through the fitting of the latest position samples. Subsequently, the Bezier curve is extrapolated to predict the future trajectory.

The order of the Bezier curve is determined by the number of selected sampling points from the historical trajectory. During the Bezier curve prediction process, two key considerations should be taken into account.

First, when a vehicle initially enters the control area of the RSU, there may not be sufficient accumulated trajectory data to fit a high-order Bezier curve. In such instances, it is advisable to employ lower-order (2nd or 3rd) polynomial curves for predicting the future trajectory. This approach enables the system to promptly initiate monitoring services for the vehicle, issuing timely warnings to mitigate potential risks.

Second, as historical trajectory data accumulate, the selection of a higher order (4th) curve for trajectory fitting becomes pertinent. Attention should also be given to the settings of the historical sampling time interval. The choice of historical sampling time intervals primarily adheres to the principles of minimizing fitting errors based on historical trajectory samples, as outlined in Algorithm 1.

Figure 3 provides an illustrative example of Bezier curve prediction. In practice, prediction results may be influenced by both data measurement errors and fitting errors. Generally, measured data do not exhibit significant deviations. Thus, we primarily consider fitting errors when selecting historical parameters for the Bezier curve. A detailed analysis of the effect of measurement errors will be provided in Section 4.

In summary, the first prediction algorithm employs a dynamic approach based on the Bezier curve. This dynamic method offers the advantage of predicting trajectories earlier and with greater accuracy compared to the original Bezier curve. The candidate set of historical sampling time intervals invariably includes several alternative values, enhancing the algorithm's adaptability and precision. Without increasing the order of time cost, the complexity of the dynamic algorithm will also be $O(n)$, where n is the amount of trajectory samples.

(2) The second algorithm

The second algorithm is also designed for low-level CAVs scenarios. In these instances, we have access to both historical vehicle states and existing motion patterns, signifying that we have amassed a significant volume of sample trajectories. We leverage deep learning-based models tailored for the precise estimation of the vehicles' future trajectory in such cases.

Deep learning models for trajectory prediction include Convolutional Neural Networks (CNNs) (Cui et al., 2019), Recurrent Neural Networks (RNNs) (Xin et al., 2018; Hou et al., 2023), and attention mechanisms (Messaoud et al., 2021; Kim et al., 2021). Recent research has emphasized that future trajectories of traffic participants exhibit a multimodal nature due to diverse driving behaviors. The attention mechanism-based model has attracted considerable attention from researchers and has demonstrated its efficacy in multimodal trajectory prediction (Zhang and Li, 2022).

Here we provide an attention-based multi-modal prediction network. This model draws inspiration from our previous work (Zhang et al., 2022; Chang et al., 2023b), which has been validated to deliver exceptional performance in comparison to other conventional models, such as CNN and RNN based models. Therefore, in this paper, we employ the BEV and attention-based model as an example for trajectory prediction and accident warning.

In this model, as illustrated in Fig. 4, historical features F^{h} for each vehicle are extracted as external input, while pre-stored BEV map data is embedded as internal input. Additionally, utilizing vehicle positions and global map information, the RSU generates historical scene BEVs as external input. These scene BEVs facilitate the capture of spatiotemporal interactions among scenario elements and enhance the understanding of the environmental context.

Historical trajectory samples serve as queries. The historical scene BEV and global map BEV are embedded

Algorithm 1 Dynamic prediction method based on Bezier curve

Input: Extracted historical trajectory sequences $Traj_{\mathrm{H}}$

Candidate set of historical sampling time interval T_{S}

Output: Predicted future trajectory $Traj_{\mathrm{P}}$

1. According to the minimal time historical sampling interval $t_{\min}$ from T_{S}, calculate the maximum number of sampling points $n_{\max}$ that can be extracted from $Traj_{\mathrm{H}}$.
2. **If** $2t_{\min} < n_{\max} \le 3t_{\min}$ **then**
   2nd Bezier curve with sampling interval $t_{\min}$ is used.
3. **If** $3t_{\min} < n_{\max} \le 4t_{\min}$ **then**
   3rd Bezier curve with sampling interval $t_{\min}$ is used.
4. **If** $n_{\max} > 4t_{\min}$ **then**
   4th Bezier curve is used. Init variable leastErr as inf.
5. **For** t_s in T_{S} **do**
6. **If** $n_{\max} > 4t_s$ **then**
   Calculate the Bezier fitting error of the historical trajectory data samples under the time sampling interval t_s.

$$Err = \sum_{\lambda=1}^{n-1}\left(\sum_{i=0}^{n}\binom{n}{i}\left(\frac{\lambda}{n}\right)^{i}\left(1-\frac{\lambda}{n}\right)^{n-i}x_i - x_\lambda\right)^2 + \left(\sum_{i=0}^{n}\binom{n}{i}\left(\frac{\lambda}{n}\right)^{i}\left(1-\frac{\lambda}{n}\right)^{n-i}y_i - y_\lambda\right)^2 \quad (n=4)$$

7. **If** leastErr > Err **then**
   Temporally set sampling interval as t_s. Set leastErr as Err value.
8. Use the appropriate order Bezier curve and the selected historical sampling time interval t_s to extract corresponding trajectory samples from $Traj_{\mathrm{H}}$.
9. Fit the parameters of Bezier curve following Eq. (1), then extrapolate and output the predicted future trajectory samples $Traj_{\mathrm{P}}$.

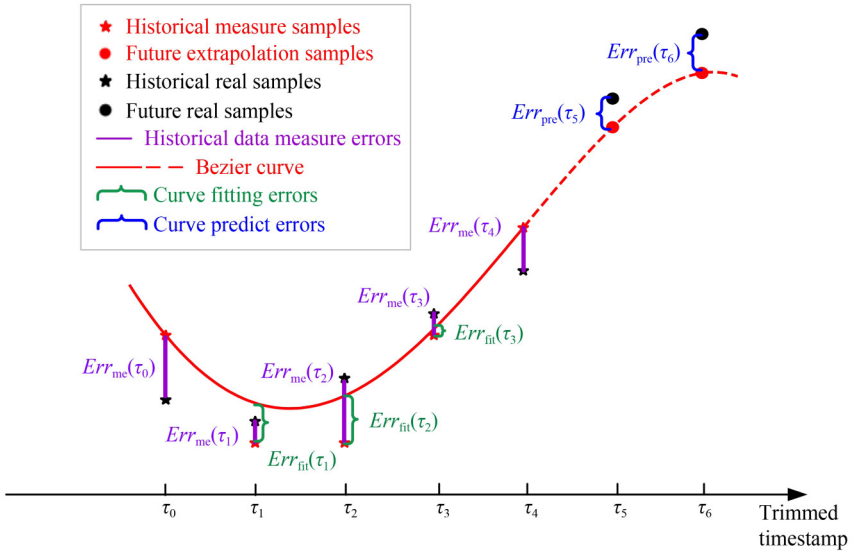


Fig. 3 Illustration of Bezier curve prediction method.

using CNN models. These tensors are subsequently employed as keys and values in the calculation of spatial and temporal attention. Leveraging an attention-based encoder and decoder architecture, spatial and temporal features are efficiently extracted and predicted. Notably, the decoder can generate long-time-sequence trajectories in a single forward procedure, eliminating the time-consuming iterations in conventional architectures. Then, the output of the decoder is fed into three prediction heads to infer different possible future trajectories V^{P}.

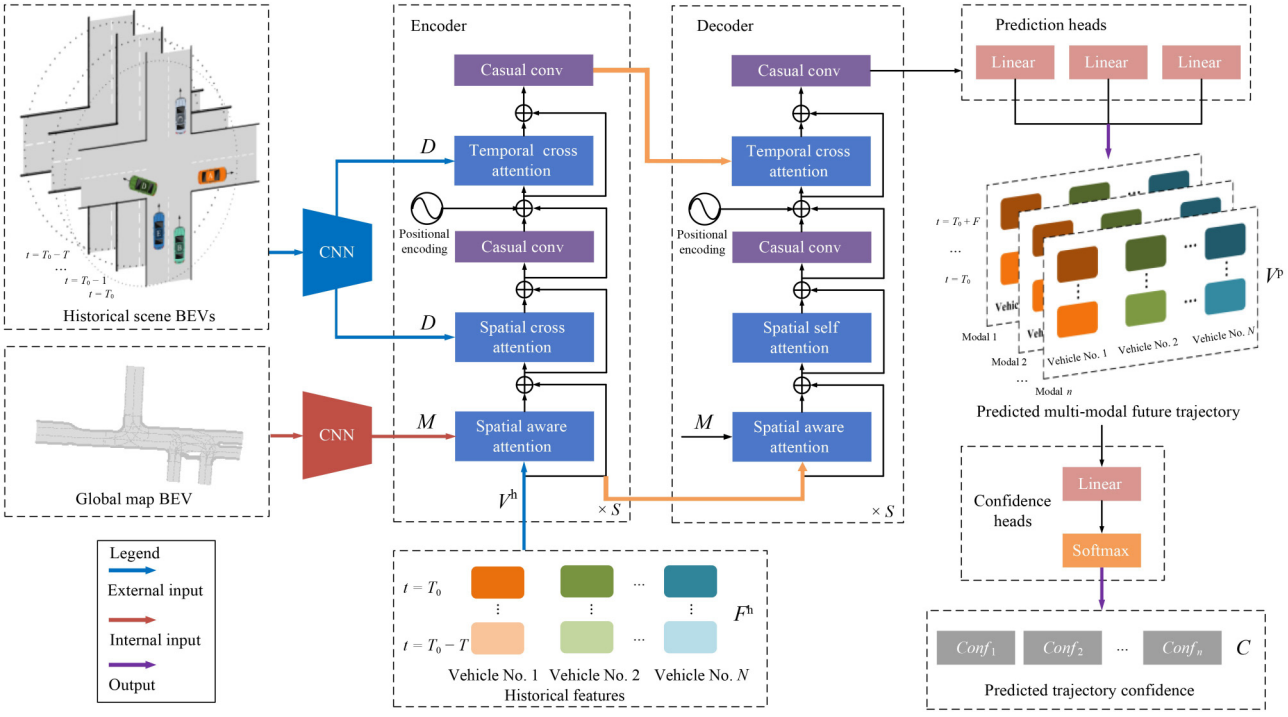


Fig. 4 An attention-based model for trajectory prediction.

The predicted trajectories are also fed into a linear layer with softmax function to further compute the confidence scores. The model can be efficiently trained with the Mean Absolute Error (MAE) and Negative Log-Likelihood (NLL) loss function. The training and optimization details are also similar.

We further consider the computation complexity of the model. Assume that the size of BEV is $H * W$, the size of the embedded tensor is E, and the length of the historical samples is T. The complexity of the convolution and embedding model is $O(HW + E^2)$, the complexity of the spatiotemporal attention model is $O(E^2 + T^2E)$, and the complexity of prediction heads and confidence heads is $O(E^2)$. In contrast to the high-dimensional BEV features, the embedded tensor results in a substantial reduction in time complexity. Furthermore, the attentions associated with historical vehicle trajectories are computed using the same set of scene BEV data. This allows for the parallel prediction of multiple future vehicle trajectories, thereby reducing inference time. A comprehensive analysis of time costs and their effect will be provided in Section 4.

(3) The third algorithm

The third algorithm is employed in scenarios involving high-level CAVs. In these instances, vehicles directly transmit predicted future positions and intentions to the RSU. Consequently, the prediction process is conducted by the vehicles themselves rather than the monitoring system. High-level CAVs, typically Level-4 or Level-5 vehicles, generally possess stronger capabilities for scenario prediction. We do not impose restrictions on specific algorithms, as long as they can accurately estimate future trajectories. This approach ensures that various autonomous vehicles from different manufacturers can be uniformly accommodated within our monitoring framework by periodically receiving their prediction results.

It is worth noting that the three prediction algorithms may introduce certain types of errors, primarily measurement errors and prediction errors. Section 4 will provide an analysis of how these errors impact the monitoring performance of vehicles.

To evaluate trajectory prediction, we utilize two metrics: Average Displacement Error (ADE) and Final Displacement Error (FDE). ADE represents the mean Euclidean distance between ground truth positions and predicted positions:

$$ADE = \frac{1}{F}\sum_{t=1}^{F}\frac{1}{N}\sum_{i=1}^{N}\sqrt{\left(p_t^{(x,y)}(i) - \hat{p}_t^{(x,y)}(i)\right)^2}, \tag{2}$$

and FDE is mean Euclidean distance between the final ground truth positions and predicted positions.

$$FDE = \frac{1}{N}\sum_{i=1}^{N}\sqrt{\left(p_F^{(x,y)}(i) - \hat{p}_F^{(x,y)}(i)\right)^2}, \tag{3}$$

where $p_t^{(x,y)}(i)$ is the ground truth positions of ith vehicle at future timestamp t, $\hat{p}_t^{(x,y)}(i)$ is the predicted positions of ith vehicle at future timestamp t, F is the prediction time horizon, and N is the amount of vehicles. For the second prediction model, the final metrics are calculated between the predicted candidate trajectory with the highest confidence and the ground truth trajectory.

3.3 The accident detection algorithm

The accident detection function utilizes the predicted trajectory data of each vehicle as inputs and generates outputs indicating the presence of potential accidents associated with the vehicles' future positions. This sub-module calculates and assesses potential risks, subsequently consolidating the detection results for the warning sub-module.

Accidents can be broadly categorized into two types: Those involving dynamic objects, such as surrounding vehicles, and those involving stationary objects, such as road boundaries and temporarily restricted zones indicated by signs or signals. By estimating the trajectory of each vehicle, it becomes feasible to establish a series of rectangles representing the vehicle's future positions within specific time intervals. Similarly, the boundaries of stationary objects are represented as line segments. To evaluate whether a vehicle will collide with other objects within the future time intervals, we employ the line intersection algorithm (Antonio, 1992). An illustration of the algorithm is provided in Fig. 5.

For line segment AB and CD, initially, a fast exclusion method is employed to determine whether the lines projections onto the x and y coordinates will overlap. If no overlap is detected, it can be concluded that there is no intersection between the line segments. Conversely, further cross product checks are performed as follows:

$$\left(\vec{AB}\times\vec{AC}\right)\cdot\left(\vec{AB}\times\vec{AD}\right)<0,\ \left(\vec{CD}\times\vec{CA}\right)\cdot\left(\vec{CD}\times\vec{CB}\right)<0. \tag{4}$$

If true, we will judge the existence of line overlapping.

For stationary objects, the accident detection process includes two steps. Firstly, the candidate future trajectories of vehicles are scrutinized to determine whether they will intersect with or traverse the boundaries of lanes or roads. Secondly, the system assesses whether vehicles are likely to disregard traffic regulations, such as violating traffic lights. To streamline these evaluations, the proposed system maintains and indexes the boundaries of both permanently and temporarily restricted zones within a one-dimensional array. This approach ensures efficient accident detection through enumeration. The operation complexity is $O(mn)$, where m represents the amount of vehicles and n represents the amount of stationary objects.

Regarding dynamic objects, accident detection entails a one-by-one comparison of the candidate trajectories for each vehicle. The line segments representing these future trajectories are aligned in time, enabling consistent evaluations at specific time intervals. The algorithm assesses whether the trajectories and vehicle rectangles intersect at the corresponding future time intervals. The operation complexity is $O(pf)$, where p represents the amount of vehicle pair and f represents the amount of predicted future position points.

3.4 The accident warning algorithm

The determination of warnings varies depending on the CAV levels. In the case of low-level CAVs that utilize the first prediction algorithm, as well as high-level CAVs that produce a single candidate trajectory as the prediction model output, warnings are promptly activated if the system detects an impending accident based on the calculated risks. However, for low-level CAVs employing the second algorithm, the system assesses every pair of candidate trajectories for each vehicle pair to confirm whether they meet a specific condition (Eq. (5)). If the condition is met, dynamic warnings are triggered. Similarly, for stationary warnings, we assign a confidence value of 1 to the positions of lane or road boundaries, and warnings are determined in a similar manner. The collision detection results are weighted by the confidences of the candidate trajectories. With the assistance of the threshold method, the proposed system effectively filters out false warnings associated with low-confidence trajectories.

$$\begin{aligned}&\min\left(conf_{A_i},\ conf_{B_j}\right)\cdot I(A_i,\ B_j)>TH\\&I(A_i,\ B_j)=\begin{cases}1 & \text{accident detected for } A_i \text{ and}\\ & B_j \text{ candidate trajectory}\\0 & \text{accident not detected}\end{cases},\\&1\leqslant i,\ j\leqslant n,\ \ conf_{A_i},\ conf_{B_j}\in(0,\ 1),\ \ TH\in(0,\ 1)\end{aligned} \tag{5}$$

where n is the candidate trajectory number, $conf_{A_i}$ is the confidence value of the ith trajectory for vehicle A, $conf_{B_j}$ is the confidence value of the jth trajectory for vehicle B, $I(A_i,\ B_j)$ is characteristic function, and TH is threshold adapted to scenarios.

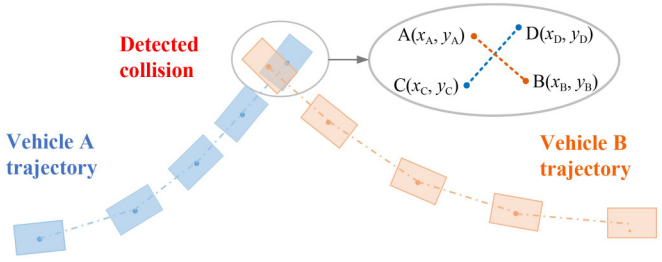


Fig. 5 Accident detection with predicted positions.

Upon confirming warning information, the module generates data packages containing the current states of the objects predicted to be involved in collisions with the assessed vehicles. These packages also incorporate the predicted accident timestamp and the corresponding positions of potential accidents. Subsequently, the monitoring system transmits messages to CAVs via V2I communication. Upon receiving the warning information, the CAVs execute necessary control and maneuvers to avert accidents.

To assess the warning performance of the safety monitoring framework, we employ the confusion matrix (Batista et al., 2000). The terms "Positive" and "Negative" pertain to the involvement of each vehicle in an accident or its absence during the scenario's progression. If an accident occurs, the corresponding sample is labeled as positive; conversely, if no accident occurs, it is labeled as a negative sample. Similarly, the terms "Positive" and "Negative" in the warnings refer to whether the system has the ability to deliver precise and rapid-response warnings when a vehicle is on the verge of collision.

After evaluating the system warnings and comparing them to actual accidents, the warnings can be categorized into four types as follows:

(1) True Positive (TP), where the accident is positive and the warning is also positive.

(2) False Positive (FP), where the accident is negative while the warning is positive.

(3) True Negative (TN), where the accident is negative and the warning is also negative.

(4) False Negative (FN), where the accident is positive while the warning is negative.

Category percentages can be calculated to provide a comprehensive assessment. For example, the True Positive Rate (TPR) and False Positive Rate (FPR) are as follows:

$$TPR = \frac{TP}{TP + FN}, \quad FPR = \frac{FP}{FP + TN}. \tag{6}$$

4 Simulation results

To assess the effectiveness of the proposed system, a series of experiments have been conducted. These experiments are outlined as follows:

Determination of System Parameters: The system's parameters are appropriately configured to ensure optimal performance.

Processing Speed Evaluation: The processing speed of the system is tested under various algorithms and scenarios to assess efficiency.

Error Analysis: The relationship between measurement and prediction errors and the system's warning performance is analyzed to understand their impact.

Data Transmission Delay Study: The impact of data transmission delay on the driving safety monitoring system is examined to assess its implications.

Comparison of Prediction Algorithms: Different prediction algorithms are compared to evaluate their performance. Due to space constraints, the experiments focus on low-level CAVs and specifically investigate the first and second prediction algorithms.

To conduct these experiments, simulations are carried out using various datasets that meet specific criteria. The selected datasets should have the following characteristics:

(1) Sufficient Data Duration: The datasets should contain vehicle motion data with relatively long continuous durations to accumulate enough data for simulation and validation.

(2) Large Spatial Area: The spatial area covered by the datasets should be relatively large to demonstrate the advantages and significance of V2X-based methods effectively.

(3) Diverse Scenarios and Road Maps: The datasets should include various types of scenarios and corresponding road maps to support a wide range of experiments.

As a representative dataset for these experiments, the Interaction dataset (Zhan et al., 2019) has been chosen.

The dataset used for experimentation includes interactive scenarios, providing both tracking information for vehicles and high definition (HD) map data. The proposed framework is validated in two specific scenarios: Ramp merging and urban intersections, corresponding to the DR_CHN_Merging_ZS and DR_USA_Intersection_EP0 data folders, respectively. As the initial datasets do not originally contain dynamic collisions among vehicles, a process is employed to accelerate or decelerate a subset of vehicles, and authentic accidents are selected and simulated. The data processing is carried out using the MetaScenario framework (Chang et al., 2023a). Figure 6 illustrates scenes from these scenarios, showcasing examples of both stationary and dynamic collisions.

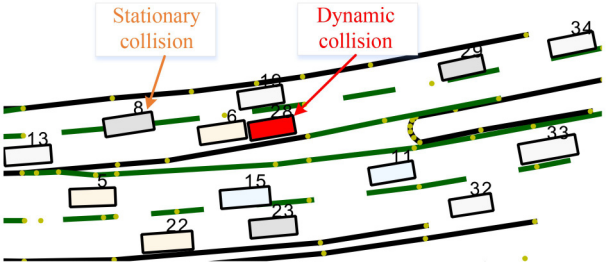
During the experiments, the machine utilized is equipped with an Intel 10900X CPU and 64 GB of RAM. The server runs on Ubuntu 18.04 LTS. The primary programming language used for implementing the main framework is C++, which interfaces with the deep learning model through PyTorch.

4.1 Selection of system parameters

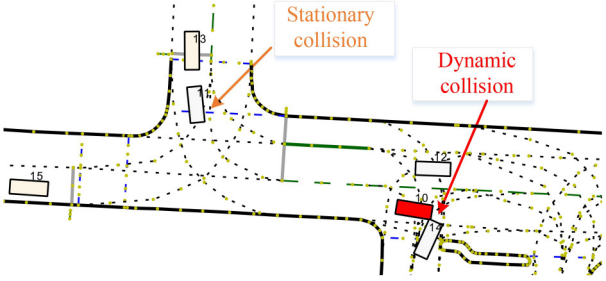
The system operates in a timer-triggered manner, and several timer parameters are crucial for its effective functioning. These parameters include:

(1) System Check Time Interval: This represents the time interval between system-executed cycles of accident prediction and detection.

(2) Historical Trajectory Time Horizon: This specifies the duration for which historical trajectories of vehicles are considered.



(a) Accidents in the ramp merging scenario



(b) Accidents in the intersection scenario

Fig. 6 Examples of accident scenes in our experiment.

(3) Historical Sampling Time Interval: This sets the frequency at which the system samples and extracts historical data.

(4) Prediction Time Horizon: This determines the time span into the future for which the system predicts the movements of vehicles.

(5) Prediction Sampling Time Interval: This sets the frequency at which the system samples and predicts the future positions of vehicles.

Figure 7 provides visual explanations of these timer parameters.

Balancing warning performance and computational complexity is essential when selecting these parameters. The first prediction algorithm, with lower precision compared to the others, requires careful consideration when determining the appropriate parameter settings. To illustrate the parameter selection process, we will focus on the ramp merging scenario with the first algorithm, which typically involves a higher number of vehicles and exhibits more noticeable trends.

There are several considerations when selecting the timer parameters for the system:

Historical Time Horizon and Historical Sampling Time Interval: Setting an excessively long historical time horizon may pose challenges due to the limited calculation capability of prediction models. The dynamic prediction method based on the Bezier curve selects historical time horizon and historical sampling interval parameters dynamically, making the most of the past 5 s of trajectory data with the least error. The historical sampling time interval is adaptively selected according to Algorithm 1.

Prediction Time Horizon: Considering CAV kinetic parameters (Wang et al., 2021) and findings of Brännström et al. (2013) and Lee and Yeo (2016), the typical prediction time horizon for risky driving behaviors falls within the range of 1.2 to 3 s. The majority of events in the Interaction Merging dataset occur around 2 s (Wang et al., 2022). Therefore, setting the prediction horizon to 2 s is reasonable to effectively monitor risky driving behaviors during the simulation.

System Check Frequency and Prediction Sampling Time Interval: Careful selection of these parameters is crucial to ensure CAV safety and minimize false alarms. In the experiments, various combinations of parameters are tested. The check time interval is varied within the range of 0.1 to 1.5 s, while the prediction time interval is varied within the range of 0.2 to 1.0 s.

The system uses TPR and FPR metrics as criteria for parameter selection. These metrics were introduced in Section 3.4. Figure 8 displays the changes in these metrics.

The observations regarding the selection of timer parameters are as follows:

(1) Check Frequency: Lowering the check frequency (decreasing the check time interval) leads to an increase in TPR for both dynamic and stationary collisions. However, it also results in an increase in FPR. This indicates that lower check frequency hinders timely accident detection in vehicles, reducing safety assurance but minimizing disturbance to vehicles.

(2) Prediction Sampling Time Interval: For dynamic collisions, increasing the prediction sampling time interval significantly reduces TPR metrics. If the prediction

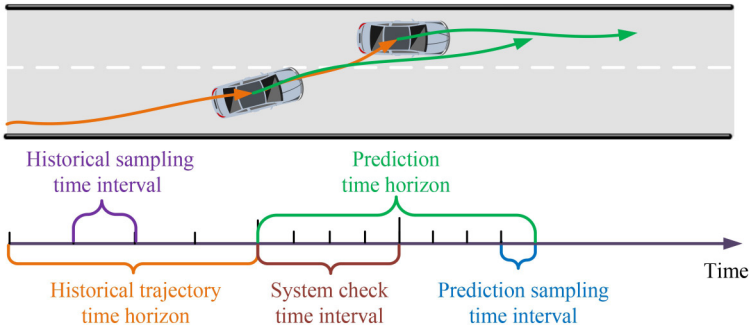


Fig. 7 A visual illustration of system timer parameters.

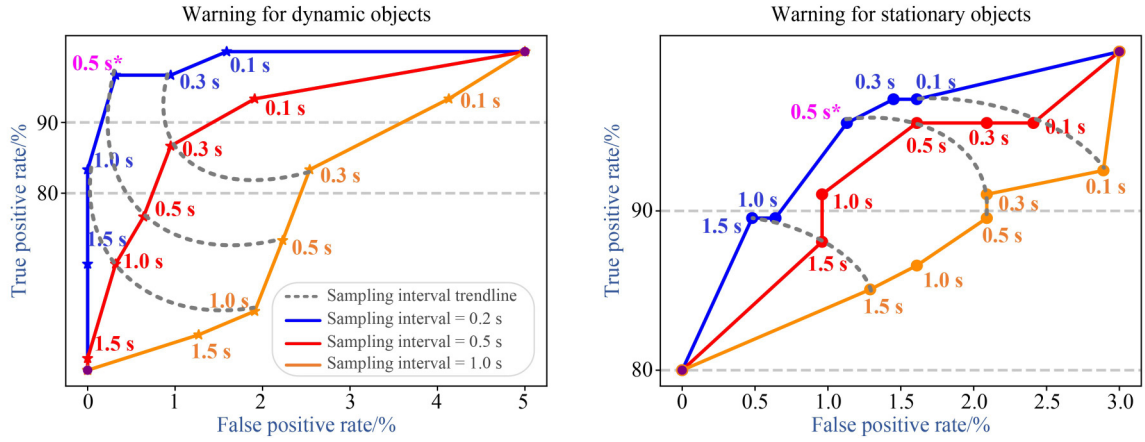


Fig. 8 Variation of warning TPR and FPR metrics with check time interval and prediction time interval parameters (the check time intervals are annotated near the data points, and the left and right plots share the same legend).

sampling interval is relatively high (e.g., 1.0 s), it also leads to an obvious increase in FPR metrics. For stationary collisions, an increase in the sampling time interval primarily results in an increase in FPR metrics.

(3) Dynamic Bezier Curve Prediction: Compared to the original Bezier curve method (Chang et al., 2022), the dynamic Bezier curve prediction method improves TPR significantly for dynamic objects without increasing FPR. For stationary objects, the TPR and FPR metrics change only slightly.

The parameter combinations located around the upper-left curve region are generally considered more effective. Taking computational time cost into account, the check frequency is set to 2 Hz (0.5 s check time interval), and the prediction time interval is set to 0.2 s. These parameters are also suitable for the intersection scenario dataset and subsequent simulation comparisons.

4.2 Process efficiency testing of the framework

The time cost of the proposed framework is distributed among several components: Data storage, movement prediction, and accident detection. In the ramp merging scenario, the RSU controls an area within the range of 150 m × 50 m, and approximately 100 vehicles engage in continuous information interaction with the system. The spatial distribution of vehicles in the ramp merging scenario is relatively dense. In the intersection scenario, the RSU controls an area within the range of 140 m × 80 m, and approximately 20 vehicles perform data transmission and interaction simultaneously in the intersection area.

For both scenarios, the processing speed is tested using the first and second prediction methods. The average time for processing per 1 s data stream is provided in Table 2.

As indicated in Table 2, efficient data storage can be realized through the utilization of the CD-DB data structure. Detailed discussions regarding alternative data storage models have already been conducted in our prior work (Yu et al., 2023), obviating the need for further elaboration here. The historical position data retrieval operates with expediency, with the first algorithm's processing time in the ramp merging scenario quantified at 1.462 ms. In contrast, the accident detection sub-module scrutinizes vehicle positions and road boundaries exhaustively, resulting in a relatively greater time expenditure, specifically 3.798 ms.

Table 2 Average processing time (ms) of each sub-module for per 1 s data stream

Type	Data storage	Trajectory prediction	Accident detection	Sum
Ramp first	0.126	1.462	3.798	5.386
Ramp second	0.126	66.802	3.798	70.726
Intersection first	0.079	0.432	0.482	0.993
Intersection second	0.079	31.350	0.482	31.911

In the context of the intersection scenario, where the vehicle count is comparatively lower, the temporal overhead for both the initial prediction algorithm and collision detection algorithm is notably reduced, measuring at 0.432 and 0.482 ms, respectively — figures that are lower than their ramp merging scenario counterparts.

Concerning the second trajectory prediction algorithm, deep learning models necessitate increased computational resources and demand more time from RSUs. Nevertheless, in the proposed attention-based multi-modal prediction algorithm, we mitigate the inference time by enabling concurrent attention calculations on the same group of BEV scenes and the prediction of multiple vehicles in parallel. The inference time is confined to a maximum of 100 ms when processing a per 1 s data stream, a duration deemed acceptable in comparison to prior research (Zhang et al., 2022).

Furthermore, it is crucial to clarify that the training of deep learning models incurs higher time costs than the inference stage. However, training can be conducted offline. Upon accumulating a substantial volume of vehicle

movement data, it can be uploaded to a remote cloud center to initiate the training process for the development of a valuable prediction model. Although this process is time-intensive, the temporal burden is borne by cloud training and analysis rather than practical real-time application. Consequently, the provided statistics in this context pertain exclusively to inference time. Additionally, as more vehicle data accumulates within the control area, the prediction model stands to benefit from further refinement through training with novel data in the cloud environment.

It is worth noting that the temporal expenditure is contingent on parameter selection. Fine-tuning parameters may lead to increased time consumption for prediction and detection functions. However, judicious parameter selection, as expounded in Section 4.1, empowers the framework to effectively manage cooperative driving scenarios within an acceptable temporal envelope.

4.3 Impact of measurement error

Measurement errors may arise from diverse sources, including inaccuracies in sensors, the GPS system, and data transmission. Following the insights of Meng et al. (2018), it is assumed that these errors can be reflected as deviations in vehicle positions.

In our investigation of their impact, we utilize the unaltered data as the baseline for accuracy comparison. Gaussian noise, characterized by a mean of zero and varying standard deviations, is introduced to the position data to replicate measurement errors. In our evaluation of vehicle safety, our attention is concentrated on TPR metrics.

As depicted in Fig. 9, the statistical results afford insight into the following three key observations:

(1) At lower levels of noise intensity, the framework exhibits a notable capacity to maintain a relatively high TPR metric. However, as noise intensity escalates, corresponding to increased deviations, there is a substantial decline in TPR metrics for dynamic collisions. The TPR for stationary warnings also experiences a reduction, albeit less pronounced than that observed for dynamic warnings.

(2) Introducing measurement noise, we observe that warning metrics reliant on the second prediction algorithm markedly outperform those predicated on the first prediction algorithm. In the context of dynamic warnings, disparities in noise scales within each subgraph reveal that the deep-learning-based model displays superior

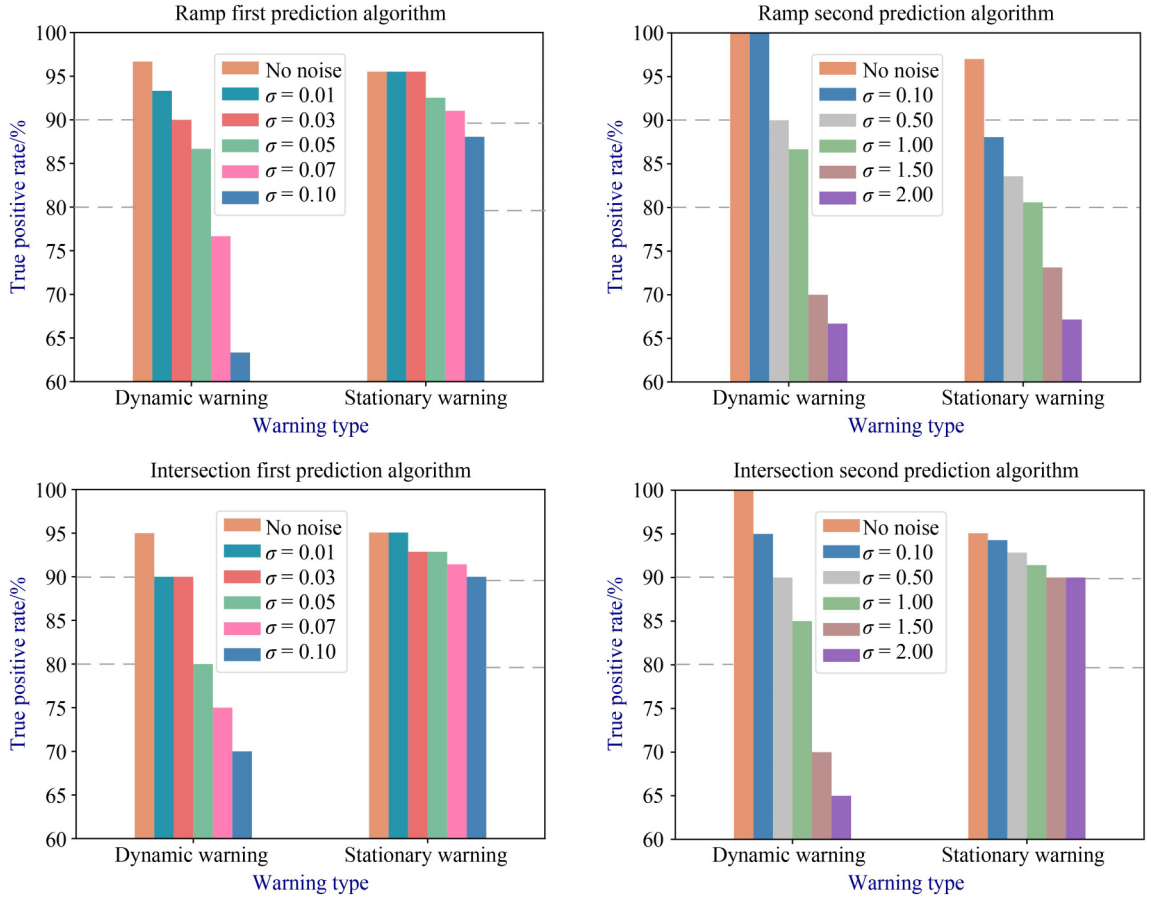


Fig. 9 TPR changes with noises of different deviations for warnings in different scenario datasets (the σ standard deviation scales are different in the first and second prediction algorithms).

noise tolerance compared to the Bezier polynomial prediction model, with an increase of an order of magnitude. For stationary warnings, as indicated in Table 3, it becomes evident that both lane width and road width in intersection scenarios surpass those in ramp scenarios. In scenarios with broader roads (such as intersection scenarios), warning performance for stationary objects proves less susceptible to the influence of vehicle position noise. This phenomenon arises from the fact that on narrow roads, disturbances or noise can significantly alter the judgment results concerning contact with stationary boundaries, whereas the situation reverses on wider roads. Consequently, in ramp scenarios, there exists no substantial difference in noise tolerance between the two prediction methods, while in intersection scenarios, the second algorithm's performance surpasses that of the first, which can be obtained by comparing the results of two algorithms at 0.1 m standard deviation noise.

(3) Notably, it is evident that within the error intensity of 0.05 m standard deviation for the first prediction algorithm and 1.0 m standard deviation for the second prediction algorithm, our proposed system consistently achieves a TPR exceeding 80% across various scenarios. These results align favorably with performance discussions in existing research studies (Lee and Peng, 2005; Guyonvarch et al., 2020; Wang et al., 2020b).

It is worth highlighting that global navigation satellite system (GNSS)-based localization typically provides a position accuracy of approximately 5 m (Tavani et al., 2020). Based on the measurement error analysis, the framework advocates equipping CAVs with real-time-kinematic (RTK) equipment, capable of achieving centimeter-level positional accuracy, to enhance prediction precision.

Table 3 Comparison for lane/road widths in different road environment

Environment	Lane width/m	Road width/m
Ramp	3.30	7.70
Intersection	4.27	9.97

4.4 Impact of transmission delay

In the course of data exchange and transmission between CAVs and RSUs, the occurrence of data transmission delays is an inherent inevitability.

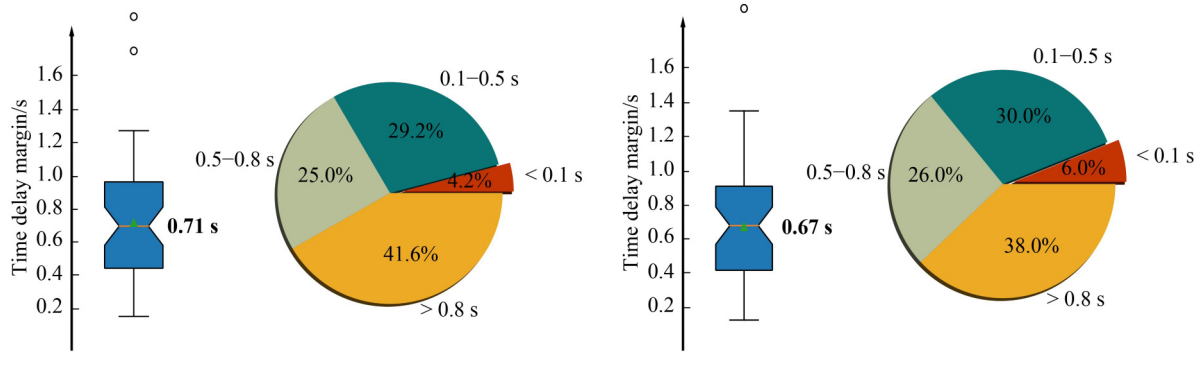
In the context of data transmission from CAVs to RSUs, the system integrates a timestamp examination process during data package collection. Subsequently, employing an interpolation method, the data are both trimmed and temporally synchronized along the time axis. Consequently, the repercussions of this time delay align with the same factors previously addressed concerning data measurement inaccuracies.

Regarding data transmission from RSUs to CAVs, an assessment of time delay margins is conducted. The margin refers to the available time span between the timestamp when the CAV receives the warning and the timestamp at which accident avoidance maneuvers are effectively executed, as elucidated by Li et al. (2022a). Notably, the time incurred for data processing and model inference is duly accounted for and is excluded from the results pertaining to time margins.

By scrutinizing the statistical data illustrated in Fig. 10, two key observations can be outlined:

(1) The mean tolerable time delay margin for CAVs across the aforementioned scenarios surpasses 0.67 s. Remarkably, a minimum of 94% of warning instances can accommodate a time transmission delay as brief as 0.1 s. It is noteworthy that, in accordance with contemporary advancements in the V2X communication protocol, as discussed by Chen et al. (2016), the maximum latency is expected to remain below 0.1 s. This substantiates the framework's capability to withstand transmission delays, as corroborated by experimental results.

(2) The second prediction method necessitates a more substantial computational time. This aligns with the requisites of deploying algorithms on the roadside, where enhanced computational resources can be availed. Consequently, the residual time delay margin experiences a minor reduction when contrasted with the first algorithm in instances of correctly-triggered warnings. Opting for



(a) First prediction method

(b) Second prediction method

Fig. 10 Distributions for time delay margin of correctly-triggered dynamic warnings.

the first prediction method enhances real-time performance and mitigates warning failures arising from data transmission delays. Nonetheless, it is crucial to acknowledge that the accuracy and scope of risky scenarios addressed by the first algorithm are inferior to those of the second, a topic that will be subject to further discussion in subsequent discourse.

4.5 Comparison between different prediction algorithms

For the first algorithm, the parameters are derived through Algorithm 1, which minimizes fitting errors on trajectory samples. In contrast, the second algorithm necessitates comprehensive training of the dataset to discern driving patterns and forecast future trajectories. During the training stage, we configure the batch size as $B = 64$, the embedding dimension as $E = 8$, and employ the Adam optimizer with an initial learning rate of 0.001. It is crucial to note that when applying the model to other scenario datasets, parameter adjustments should be made judiciously. The training regimen includes approximately 300 epochs, with the dataset partitioned into training, validation, and test sets in an 8:1:1 ratio. The optimal model is determined based on its performance on the validation set, subsequently culminating in a final evaluation on the test set.

We conduct a comparative analysis of prediction performance using different methods, evaluating the ADE and FDE metrics. Furthermore, we assess system warning performance through the TPR and FPR, with the results presented in Table 4. To gain deeper insights, we also perform error distribution statistics and examine how the two algorithms' performance evolves over prediction time. The errors are modeled by a gamma distribution to facilitate quantitative analysis, and the results for the ramp scenario dataset are provided as an illustration in Fig. 11.

Combining the findings from Table 4 and Fig. 9 with the aforementioned experimental statistics, we compare the two prediction algorithms along the following dimensions:

(1) Within a 2 s prediction time horizon, it is evident that the mean displacement error remains below 0.6 m in the ramp scenario and 0.8 m in the intersection scenario. In terms of warning metrics, particularly TPR and FPR, the proposed system effectively fulfills accident warning tasks within these prediction error thresholds.

(2) In comparison to the first prediction algorithm, the second algorithm exhibits diminished prediction errors. Furthermore, it enhances warning performance by increasing TPR by 2% while reducing FPR by nearly half. As time progresses, the parameters of the fitted gamma distribution $gamma(\alpha, \beta)$ will change, which mainly reflects in the significant decrease of β parameter. When α parameter is almost constant, the decrease of β will lead to an increase in both the mean and variance of prediction errors. Additionally, the second prediction algorithm demonstrates superior prediction accuracy, especially as the prediction time horizon extends.

(3) Drawing upon the aforementioned experimental results, it becomes evident that the second prediction algorithm offers enhanced accuracy and greater tolerance to noise compared to the first algorithm. However, it does exhibit relatively weaker real-time warning capabilities. The selection of the appropriate prediction algorithm should consider the trade-offs between these performance aspects across various conditions.

Table 4 Comparison for different prediction methods

Algorithm	ADE/m	FDE/m	TPR/%	FPR/%
Ramp first	0.25	0.58	96.06	0.73
Ramp second	0.18	0.39	98.43	0.40
Intersection first	0.31	0.76	95.45	3.69
Intersection second	0.28	0.65	97.27	2.35

5 Performance comparison and analysis of single vehicles and V2X-based edge computing system

5.1 Comparison on naturalistic scenario datasets

The warning performances of single vehicles generally

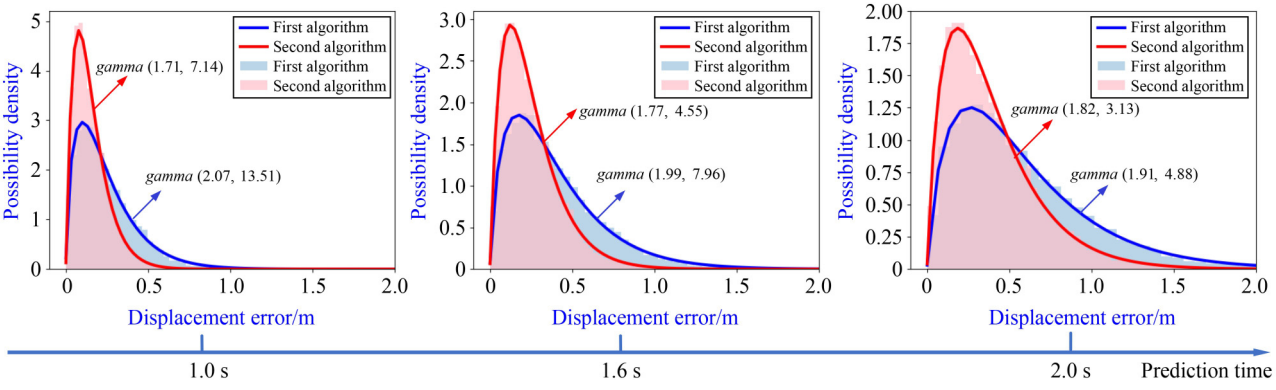


Fig. 11 Error distributions of different trajectory prediction algorithms over time.

fall below those of CAVs, primarily due to a range of factors including constrained sensing capabilities, perceptual inaccuracies, and instances of recognition failures. Within this section, we investigate the conceivable disparities in warning performance arising from the restricted sensing range of single vehicles. Figure 12 offers a visual juxtaposition of the sensing ranges between CAVs and single vehicles.

Within the simulated environment, where vehicles are closely spaced, and scenarios include diverse instances of vehicle congestions and road occlusions, we establish the sensing range for individual vehicles as a square area with a radius of 30 m around each vehicle. Other parameters remain consistent with the previously configured settings. The simulation includes both single vehicles and CAVs operating within the same system environment. Trajectory prediction is executed using the second algorithm, noted for its superior detection rates. The results are presented in Table 5.

Upon examining Table 5, it becomes apparent that TPR metrics for stationary collisions demonstrate similarity between single vehicles and CAVs. However, in the context of dynamic accident detection, CAVs exhibit superior performance compared to single vehicles, showcasing significantly improved warning capabilities in both types of scenarios. Furthermore, the system's performance for CAVs surpasses that of certain previously published collision warning algorithms (Lee and Peng, 2005; Lee and Yeo, 2016; Lyu et al., 2022).

In the context of dynamic warnings, CAVs demonstrate the capability to deliver timely warnings for all dynamic objects across both scenario types. In contrast, single vehicles, constrained by their limited sensing range, may exhibit insensitivity to certain dynamic collisions, resulting in delayed detection.

Additionally, we conducted an analysis to identify the specific categories of dynamic warnings that elude detection by single vehicles. As depicted in Fig. 13, single vehicles encounter challenges in providing timely warnings in approximately 12% of dynamic collision scenarios. Among the instances of missed warnings, it is observed that the inter-vehicle gaps and speed differentials among interacting vehicles, measured 2 s prior to the collision, are markedly higher than in cases where warnings are triggered accurately. This observation suggests that the failures in warnings by single vehicles can be ascribed to two primary factors. Firstly, single vehicles may possess limited capacity to preemptively perceive impending accidents, particularly when substantial gaps exist between vehicles. Secondly, single vehicles may struggle to activate timely warnings when faced with significant speed differentials among the involved vehicles.

In Fig. 14, we present a representative illustration of

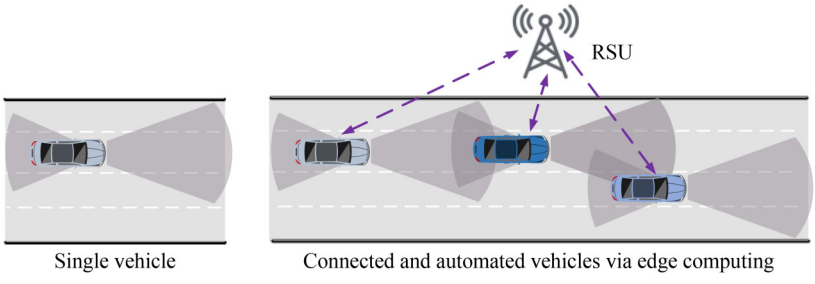


Fig. 12 Sensing range comparison for single vehicle and CAV.

Table 5 Warning performances comparison for CAVs and single vehicles

Ramp System dynamic		Data			
		Positive		Negative	
Positive	CAVs	60	100.00%	2	0.32%
	Single vehicles	54	90.00%	2	0.32%
Negative	CAVs	0	0.00%	627	99.68%
	Single vehicles	6	10.00%	627	99.68%

Ramp System stationary		Data			
		Positive		Negative	
Positive	CAVs	65	97.01%	3	0.48%
	Single vehicles	64	95.52%	2	0.32%
Negative	CAVs	2	2.99%	619	99.52%
	Single vehicles	3	4.48%	620	99.68%

Intersection System dynamic		Data			
		Positive		Negative	
Positive	CAVs	40	100.00%	2	1.22%
	Single vehicles	34	85.00%	2	1.22%
Negative	CAVs	0	0.00%	162	98.78%
	Single vehicles	6	15.00%	162	98.78%

Intersection System stationary		Data			
		Positive		Negative	
Positive	CAVs	67	95.71%	5	3.73%
	Single vehicles	65	92.86%	4	2.99%
Negative	CAVs	3	4.29%	129	96.27%
	Single vehicles	5	7.14%	130	97.01%

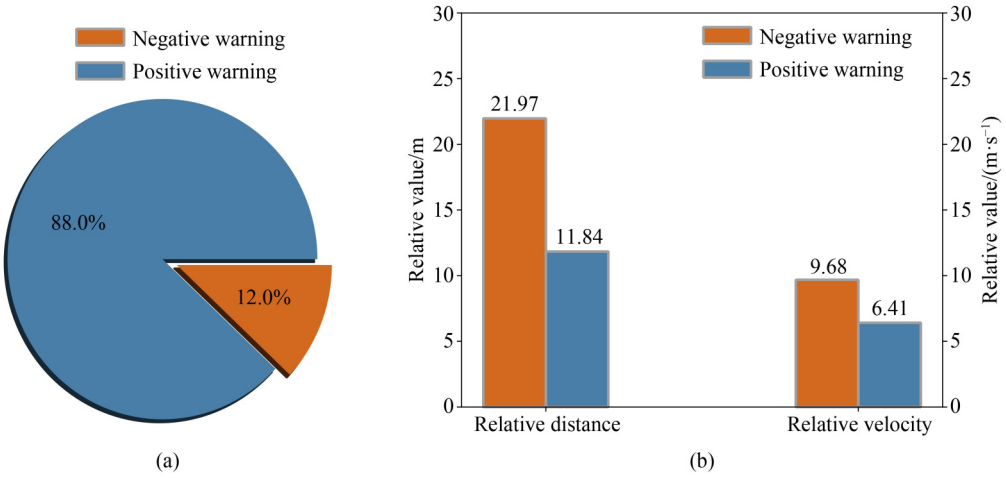


Fig. 13 (a) The distribution about whether single vehicles trigger timely warning for dynamic objects; (b) indicators comparisons for the positive and negative warning cases.

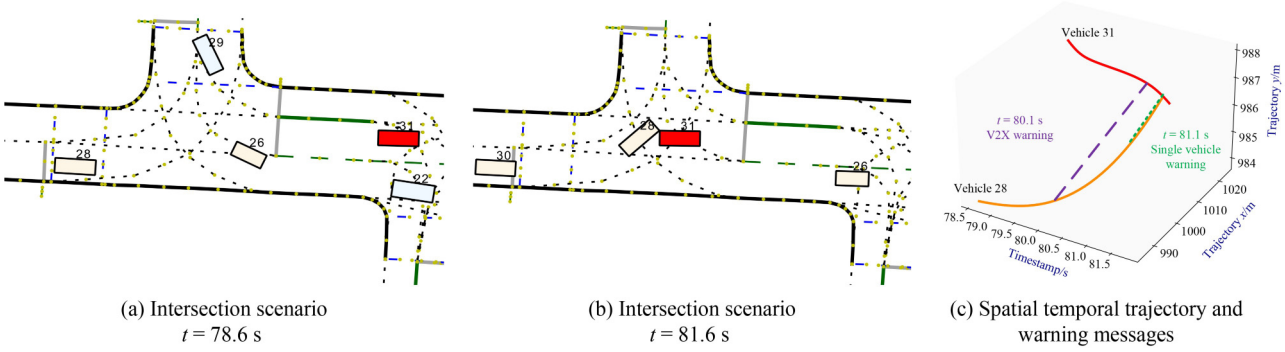


Fig. 14 A typical example of risky warnings by CAVs and single vehicles.

risky warnings within the scenario dataset. Owing to the constraints imposed by the intersection's geometry and the visibility obstructions caused by vehicle 26, 31, and 28, these vehicles are unable to establish mutual perception promptly by themselves. However, with the aid of a V2X-based edge computing system, the data and contextual scenario information can be swiftly extracted to facilitate accident detection. Consequently, the V2X warning is issued 1 s in advance of the warning generated by a single vehicle, affording more time and distance margin for the adjustment of vehicle trajectories.

5.2 Comparison in real world environment

We also conducted real-world experiments to evaluate the proposed V2X-based monitoring system at the Shunyi testing field in Beijing, China. The experiments were conducted as follows.

Devices: For the experiments, we arranged 10 intelligent vehicles equipped with OBUs and one RSU. As illustrated in Fig. 15, the OBU devices incorporate multiple modules, including GPS, WiFi, Bluetooth, and a V2X communication module. These vehicles employ the OBU to transmit and receive data packets with the RSU. In general, the communication signals between OBU and RSU can cover a range of up to 500 m. The position data of OBUs is subscribed, and the data transmission interval between OBU and RSU is set at 0.1 s. Each data package contains core vehicle state information, occupying approximately 54 bytes.

Scenario: As depicted in Fig. 16, we orchestrated the movement of 10 vehicles through the 150 m × 150 m intersection. Six vehicles were positioned to traverse from the west side to the east side, while four vehicles were designated to travel from the south side to the north side. Throughout the progression of this driving scenario, the vehicles were required to remain vigilant for side and forward collision warnings.

Experiments: We collected and processed the data from both OBU and RSU sources, subsequently constructing consistent scenarios in CAVSim (Zhang et al., 2023), a simulator tailored for CAVs environments. As depicted in Fig. 17(a), two of the ten vehicles approached the intersection from the west and south sides individually. Due to the constraints posed by road geometry and limited visibility, these vehicles encountered challenges in perceiving each other directly. However, they were able to access comprehensive global scenario

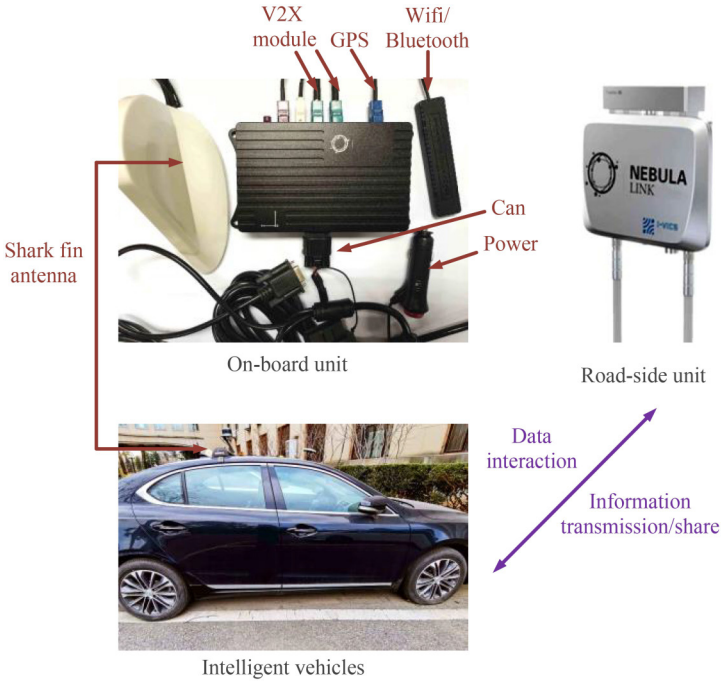
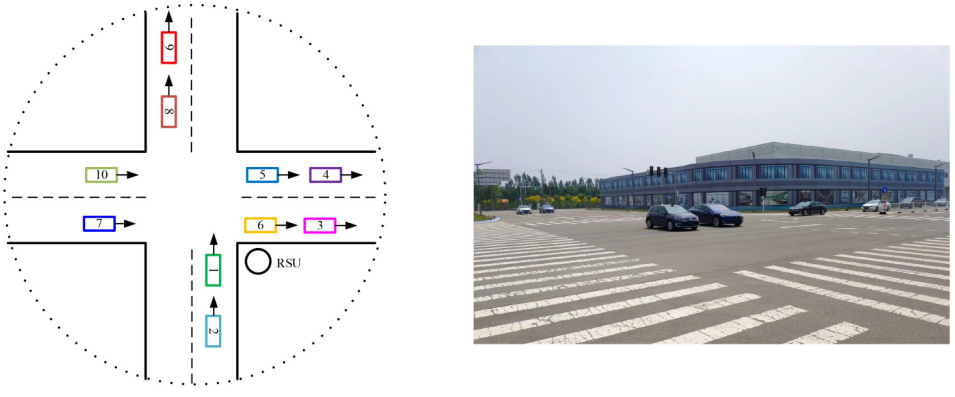


Fig. 15 The devices that support real-world vehicle testing.



(a) The testing scenario from BEV (b) The testing scenario from cameras

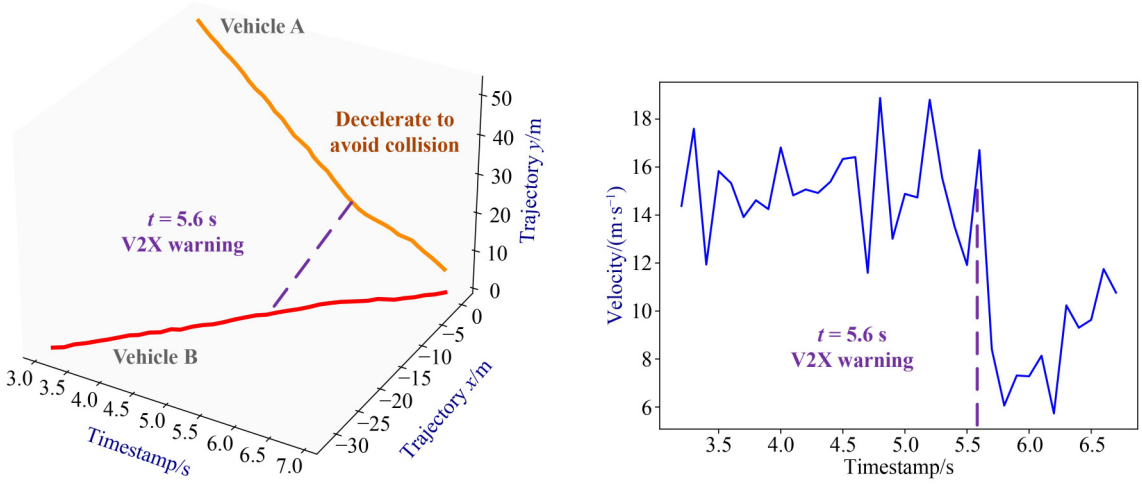
Fig. 16 The scenarios that under real-world vehicle testing.

information through data interactions between their OBUs and the RSU. In this scenario, our proposed monitoring framework successfully issued an early warning at 5.6 s, as illustrated in Fig. 18, with signals transmitted to the vehicles to convey risk warnings. Analyzing the trajectory changes in Fig. 17(a) and velocity alterations in Fig. 17(b), it becomes apparent that the vehicle executed reactive deceleration maneuvers, successfully averting a collision. These experiments serve to substantiate the practical efficacy of our framework.

6 Conclusions

This paper presents a novel safety monitoring framework tailored for CAVs. The framework seamlessly integrates efficient algorithms and data structures to assess risks using real-time data. It includes various trajectory prediction algorithms adaptable to different CAV levels. To ensure streamlined execution, the system is implemented on the roadside, capitalizing on the computational capabilities afforded by edge computing.

The monitoring framework effectively meets the criteria of comprehensiveness, efficiency, accuracy, and compatibility. Through extensive simulations, the framework's effectiveness is empirically verified. By maintaining the measurement error within the range of zero-mean Gaussian noise with a standard deviation of 0.05 m for the first prediction algorithm and 1.0 m for the second prediction algorithm, and by judiciously selecting parameters, the framework attains a remarkable TPR of over 80% for accident detection. Furthermore, approximately 96% of



(a) Evolution of spatial temporal trajectory (b) Velocity changes of vehicle A

Fig. 17 The motion process and spatial temporal trajectories of the two vehicles.

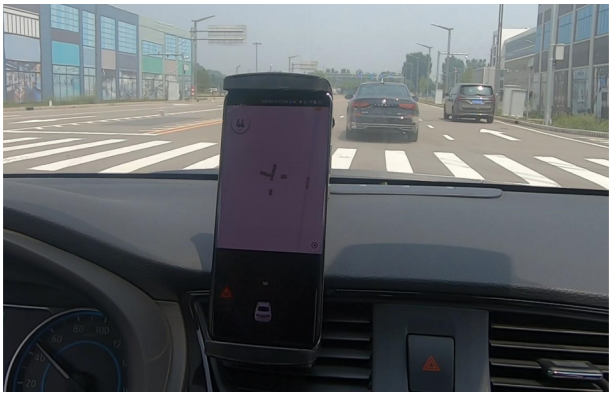


Fig. 18 The warning signals that transmit to the vehicle.

system warnings are robust against transmission delays exceeding 0.1 s. The system demonstrates its efficiency in CAV environments with minimal time overhead. Additionally, we elucidate the appropriate selection of prediction algorithms under diverse conditions. Both simulation and real-world testing substantiate that the V2X-based edge computing approach surpasses conventional single-vehicle approaches in terms of warning performance.

It is worth noting that certain aspects are not extensively explored due to space limitations. In real-world driving scenarios, it is common for not all vehicles to be CAVs (Ma et al., 2023). Accessing data from human-driven vehicles can be challenging, and the presence of pedestrians and non-motorized vehicles in the transportation system (Gao et al., 2023) introduces additional complexities that necessitate consideration. Our future research endeavors will concentrate on developing methods to facilitate safe cooperative driving under these conditions.

Competing Interests The authors declare that they have no competing interests.

References

Antonio F (1992). Faster line segment intersection. In: Kirk D, ed. Graphics Gems III (IBM Version). San Francisco, CA: Morgan Kaufmann, 199–202

Barrios C, Motai Y (2011). Improving estimation of vehicle's trajectory using the latest global positioning system with Kalman filtering. IEEE Transactions on Instrumentation and Measurement, 60(12): 3747–3755

Batista G E A P A, Carvalho A C P L F, Monard M C (2000). Applying one-sided selection to unbalanced datasets. In: Mexican International Conference on Artificial Intelligence. Mexico: Springer, 315–325

Brännström M, Sandblom F, Hammarstrand L (2013). A probabilistic framework for decision-making in collision avoidance systems. IEEE Transactions on Intelligent Transportation Systems, 14(2): 637–648

Chang C, Cao D, Chen L, Su K, Su K, Su Y, Wang F Y, Wang J, Wang P, Wei J, Wu G, Wu X, Xu H, Zheng N, Li L (2023a). Metascenario: A framework for driving scenario data description, storage and indexing. IEEE Transactions on Intelligent Vehicles, 8(2): 1156–1175

Chang C, Zhang J, Zhang K, Zhong W, Peng X, Li S, Li L (2023b). BEV-V2X: Cooperative birds-eye-view fusion and grid occupancy prediction via V2X-based data sharing. IEEE Transactions on Intelligent Vehicles, 8(11): 4498–4514

Chang C, Zhang K, Zhang J, Li S, Li L (2022). Driving safety monitoring and warning for connected and automated vehicles via edge computing. In: IEEE 25th International Conference on Intelligent Transportation Systems (ITSC). Macao: IEEE, 3940–3947

Chen S, Hu J, Shi Y, Zhao L (2016). LTE-V: A TD-LTE-based V2X solution for future vehicular network. IEEE Internet of Things Journal, 3(6): 997–1005

Choi J W, Curry R, Elkaim G (2008). Path planning based on Bézier curve for autonomous ground vehicles. In: Advances in Electrical and Electronics Engineering – IAENG Special Edition of the World Congress on Engineering and Computer Science. San Francisco,

CA: IEEE, 158–166

Cualain D O, Hughes C, Glavin M, Jones E (2012). Automotive standards-grade lane departure warning system. IET Intelligent Transport Systems, 6(1): 44–57

Cui H, Radosavljevic V, Chou F C, Lin T H, Nguyen T, Huang T K, Schneider J, Djuric N (2019). Multimodal trajectory predictions for autonomous driving using deep convolutional networks. In: International Conference on Robotics and Automation (ICRA). Montreal, QC: IEEE, 2090–2096

Di X, Shi R (2021). A survey on autonomous vehicle control in the era of mixed-autonomy: From physics-based to AI-guided driving policy learning. Transportation Research Part C: Emerging Technologies, 125: 103008

Dupuis M, Strobl M, Grezlikowski H (2010). OpenDRIVE 2010 and beyond: Status and future of the de facto standard for the description of road networks. In: Proceedings of the Driving Simulation Conference Europe. Paris: INRETS, 231–242

Feng J, Li J (2013). Google protocol buffers research and application in online game. In: IEEE Conference Anthology. Chongqing: IEEE, 1–4

Fernández-Caballero A, Gomez F J, Lopez-Lopez J (2008). Road-traffic monitoring by knowledge-driven static and dynamic image analysis. Expert Systems with Applications, 35(3): 701–719

Figueiredo A, Rito P, Luís M, Sargento S (2022). Mobility sensing and V2X communication for emergency services. Mobile Networks and Applications, in press, doi:10.1007/s11036-022-02056-9

Gao Z, Huang H J, Guo J, Yang L, Wu J (2023). Future urban transport management. Frontiers of Engineering Management, 10(3): 534–539

Guyonvarch L, Lecuyer E, Buffat S (2020). Evaluation of safety critical event triggers in the UDrive data. Safety Science, 132: 104937

Haklay M, Weber P (2008). Openstreetmap: User-generated street maps. IEEE Pervasive Computing, 7(4): 12–18

He Q, Xu J, Wang T, Chan A P (2021). Identifying the driving factors of successful megaproject construction management: Findings from three Chinese cases. Frontiers of Engineering Management, 8(1): 5–16

Hou L, Li S E, Yang B, Wang Z, Nakano K (2023). Integrated graphical representation of highway scenarios to improve trajectory prediction of surrounding vehicles. IEEE Transactions on Intelligent Vehicles, 8(2): 1638–1651

Jiang L, Molnár T G, Orosz G (2021). On the deployment of V2X roadside units for traffic prediction. Transportation Research Part C: Emerging Technologies, 129: 103238

Jo Y, Jang J, Park S, Oh C (2021). Connected vehicle-based road safety information system (CROSS): Framework and evaluation. Accident: Analysis and Prevention, 151: 105972

Kang L, Li H, Li C, Xiao N, Sun H, Buhigiro N (2021). Risk warning technologies and emergency response mechanisms in Sichuan–Tibet Railway construction. Frontiers of Engineering Management, 8(4): 582–594

Kim B, Park S H, Lee S, Khoshimjonov E, Kum D, Kim J, Kim J S, Choi J W (2021). Lapred: Lane-aware prediction of multi-modal future trajectories of dynamic agents. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, TN: IEEE, 14631–14640

Lee D, Yeo H (2016). Real-time rear-end collision-warning system using a multilayer perceptron neural network. IEEE Transactions on Intelligent Transportation Systems, 17(11): 3087–3097

Lee K, Peng H (2005). Evaluation of automotive forward collision warning and collision avoidance algorithms. Vehicle System Dynamics, 43(10): 735–751

Li L, Zhao C, Wang X, Li Z, Chen L, Lv Y, Zheng N, Wang F Y (2022a). Three principles to determine the right-of-way for AVs: Safe interaction with humans. IEEE Transactions on Intelligent Transportation Systems, 23(7): 7759–7774

Li Y, Pan B, Xing L, Yang M, Dai J (2022b). Developing dynamic speed limit strategies for mixed traffic flow to reduce collision risks at freeway bottlenecks. Accident: Analysis and Prevention, 175: 106781

Li Y, Zhang L, Song Y (2016). A vehicular collision warning algorithm based on the time-to-collision estimation under connected environment. In: 14th International Conference on Control, Automation, Robotics and Vision (ICARCV). Phuket: IEEE, 1–4

Lyu N, Wen J, Duan Z, Wu C (2022). Vehicle trajectory prediction and cut-in collision warning model in a connected vehicle environment. IEEE Transactions on Intelligent Transportation Systems, 23(2): 966–981

Ma Y, Liu Q, Fu J, Liufu K, Li Q (2023). Collision-avoidance lane change control method for enhancing safety for connected vehicle platoon in mixed traffic environment. Accident: Analysis and Prevention, 184: 106999

Meng Y, Li L, Wang F Y, Li K, Li Z (2018). Analysis of cooperative driving strategies for nonsignalized intersections. IEEE Transactions on Vehicular Technology, 67(4): 2900–2911

Messaoud K, Yahiaoui I, Verroust-Blondet A, Nashashibi F (2021). Attention based vehicle trajectory prediction. IEEE Transactions on Intelligent Vehicles, 6(1): 175–185

Miao L, Chen S F, Hsu Y L, Hua K L (2022). How does C-V2X help autonomous driving to avoid accidents? Sensors, 22(2): 686

Miucic R, Sheikh A, Medenica Z, Kunde R (2018). V2X applications using collaborative perception. In: IEEE 88th Vehicular Technology Conference (VTC-Fall). Chicago, IL: IEEE, 1–6

Poggenhans F, Pauls J H, Janosovits J, Orf S, Naumann M, Kuhnt F, Mayr M (2018). Lanelet2: A high-definition map framework for the future of automated driving. In: 21st International Conference on Intelligent Transportation Systems (ITSC). Maui, HI: IEEE, 1672–1679

Saligrama V, Konrad J, Jodoin P M (2010). Video anomaly identification. IEEE Signal Processing Magazine, 27(5): 18–33

Sentouh C, Nguyen A T, Benloucif M A, Popieul J C (2019). Driver-automation cooperation oriented approach for shared control of lane keeping assist systems. IEEE Transactions on Control Systems Technology, 27(5): 1962–1978

Shang M, Stern R E (2021). Impacts of commercially available adaptive cruise control vehicles on highway stability and throughput. Transportation Research Part C: Emerging Technologies, 122: 102897

Shehata M S, Cai J, Badawy W M, Burr T W, Pervez M S, Johannesson R J, Radmanesh A (2008). Video-based automatic incident detection for smart roads: The outdoor environmental challenges regarding false alarms. IEEE Transactions on Intelligent Transportation Systems, 9(2): 349–360

Tahir M N, Katz M (2022). Performance evaluation of IEEE 802.11p, LTE and 5G in connected vehicles for cooperative awareness. Engineering Reports, 4(4): e12467

Tan H S, Huang J (2006). DGPS-based vehicle-to-vehicle cooperative collision warning: Engineering feasibility viewpoints. IEEE Transactions on Intelligent Transportation Systems, 7(4): 415–428

Tapia-Espinoza R, Torres-Torriti M (2013). Robust lane sensing and departure warning under shadows and occlusions. Sensors, 13(3): 3270–3298

Tavani S, Pignalosa A, Corradetti A, Mercuri M, Smeraglia L, Riccardi U, Seers T, Pavlis T, Billi A (2020). Photogrammetric 3D model via smartphone GNSS sensor: Workflow, error estimate, and best practices. Remote Sensing, 12(21): 3616

Wang C, Xie Y, Huang H, Liu P (2021). A review of surrogate safety measures and their applications in connected and automated vehicles safety modeling. Accident: Analysis and Prevention, 157: 106157

Wang H, Wang W, Yuan S, Li X, Sun L (2022). On social interactions of merging behaviors at highway on-ramps in congested traffic. IEEE Transactions on Intelligent Transportation Systems, 23(8): 11237–11248

Wang Q, Li Z, Li L (2014). Investigation of discretionary lane-change characteristics using next-generation simulation data sets. Journal of Intelligent Transport Systems, 18(3): 246–253

Wang S, Wang Y, Zheng Q, Li Z (2020a). Guidance-oriented advanced curve speed warning system in a connected vehicle environment. Accident: Analysis and Prevention, 148: 105801

Wang X, Liu J, Qiu T, Mu C, Chen C, Zhou P (2020b). A real-time collision prediction mechanism with deep learning for intelligent transportation system. IEEE Transactions on Vehicular Technology, 69(9): 9497–9508

Wang Y, Wenjuan E, Tian D, Lu G, Yu G, Wang Y (2011). Vehicle collision warning system and collision detection algorithm based on vehicle infrastructure integration. In: 7th Advanced Forum on Transportation of China. Beijing: IEEE, 216–220

Xin L, Wang P, Chan C Y, Chen J, Li S E, Cheng B (2018). Intention-aware long horizon trajectory prediction of surrounding vehicles using dual LSTM networks. In: 21st International Conference on Intelligent Transportation Systems (ITSC). Maui, HI: IEEE, 1441–1446

Yu H, Chang C, Li S, Li L (2023). CD-DB: A data storage model for cooperative driving. IEEE Transactions on Intelligent Vehicles, 8(1): 492–501

Zhan W, Sun L, Wang D, Shi H, Clausse A, Naumann M, Kummerle J, Konigshof H, Stiller C, de La Fortelle A, Tomizuka M (2019). Interaction dataset: An international, adversarial and cooperative motion dataset in interactive driving scenarios with semantic maps. arXiv preprint. arXiv:1910.03088

Zhang J, Chang C, He Z, Zhong W, Yao D, Li S, Li L (2023). CAVSim: A microscopic traffic simulator for evaluation of connected and automated vehicles. IEEE Transactions on Intelligent Transportation Systems, 24(9): 10038–10054

Zhang K, Chang C, Zhong W, Li S, Li Z, Li L (2022). A systematic solution of human driving behavior modeling and simulation for automated vehicle studies. IEEE Transactions on Intelligent Transportation Systems, 23(11): 21944–21958

Zhang K, Li L (2022). Explainable multimodal trajectory prediction using attention models. Transportation Research Part C: Emerging Technologies, 143: 103829