

Interaction-Aware and Driving Style-Aware Trajectory Prediction for Heterogeneous Vehicles in Mixed Traffic Environment

Qixiang Zhang, *Graduate Student Member, IEEE*, Yang Xing, *Senior Member, IEEE*, Jinxiang Wang, *Member, IEEE*, Zhenwu Fang, *Member, IEEE*, Yahui Liu, *Senior Member, IEEE*, and Guodong Yin, *Senior Member, IEEE*

Abstract—Trajectory prediction (TP) of surrounding vehicles (SVs) is crucial for autonomous vehicles (AVs) to understand traffic situations and achieve safe-efficient decision-making and motion planning. However, different drivers' personalized driving preferences will bring uncertainties for long-term TP in the mixed traffic environment. To this end, this paper proposes a TP model with interaction awareness and driving style awareness for long-term TP of heterogeneous SVs. Firstly, the driving conditions in the highD dataset are distinguished, and three different driving styles of the vehicle in the car-following condition are obtained based on an unsupervised clustering algorithm. Then, an encoder-decoder architecture based on novel lane attention and multi-head attention mechanisms is proposed, where the encoder analyzes historical trajectory patterns and the decoder generates future trajectory sequences. The lane attention mechanism enhances the spatial perception capability of vehicles towards the target lane, and the multi-head attention mechanism extracts high-dimensional global interaction information about the heterogeneous vehicle group (HVG) surrounding the target vehicle (TV). Experimental results show that the proposed model outperforms state-of-the-art models in root-mean-square-error (RMSE) for long-term TP and exhibits excellent adaptability to diverse driving tasks. Moreover, this paper verifies that the driving style topology within the HVG has multiple impacts on the TP accuracy of the TV.

Index Terms—Trajectory prediction, mixed traffic environment, heterogeneous vehicles, personalized driving, vehicle interactions.

I. INTRODUCTION

AUTONOMOUS driving technology has witnessed substantial advancements in the last decade, with a steady rise in the market share of conditionally autonomous vehicles (AVs) [1]. However, due to incomplete road infrastructure, inadequate regulatory standards, and drivers' low trust in new technologies, large-scale market application of AVs will be relatively slow [2]. The mixed traffic environment (MTE)

where human-driven vehicles (HVs) and AVs coexist will dominate road traffic for the foreseeable future [3], [4].

A. Motivation

Trajectory prediction (TP) plays a crucial role in improving driving safety, road efficiency, comfort, and energy efficiency [5]. For example, research in [6] demonstrates that autonomous driving (AD) systems integrated with TP can reduce collision rates and traffic rule violations to 1.80% and 1.24%, respectively, while maintaining the Jerk at about 0.15m/s^3 . AVs need to predict the motion trajectories of SVs, like human drivers, to accurately understand long-term future traffic conditions. However, the complex nature of MTE increases the challenges of the TP [7], [8]: (1) uncertainty in HVs' behavior, such as random lane changes (LC); (2) diversity in HVs' behavior arose from different driving preferences; (3) behavioral patterns differ markedly across vehicle types (e.g., cars and trucks); (4) asymmetry interaction information due to differing perceptual limits of AVs and HVs; (5) non-human-like AV behavior heightens human driver insecurity. In general, TP in MTE is not always easy, even for human drivers, owing to many complex factors, including dynamic traffic states, driving preferences, and vehicle types [9].

Despite significant advances in vehicle trajectory prediction, existing research still faces two major limitations [10], [11]. First, most existing methods focus exclusively on the direct interaction between the AV and the target vehicle (TV), neglecting the broader interactions between the TV and its SVs. This limited perspective reduces the model's ability to capture the full complexity of the MTE, thereby affecting TP accuracy. Second, the majority of studies overlook the impact of personalized driving preferences, which are critical for improving both TP performance and human-like driving capabilities. To address these challenges, this paper introduces the concept of a heterogeneous vehicle group (HVG), which defines a basic traffic flow unit consisting of the TV and its eight directly adjacent heterogeneous vehicles that directly affect the TV. By modeling multi-vehicle interactions within the HVG and introducing a lane attention mechanism, the proposed framework bridges gaps in existing TP methods. Additionally, this paper explores the influence of driving style diversity and its topology on TP accuracy, offering deeper insights into personalized driving behavior. Integrating interaction and driving style awareness, the model enhances TP accuracy while improving AVs' safety, traffic efficiency, and human-like driving capabilities.

This work was supported by National Natural Science Foundation of China under Grants 52372410 and 52025121; State Key Laboratory of Intelligent Green Vehicle and Mobility under Project No. KFY2415; and SEU Innovation Capability Enhancement Plan for Doctoral Students under Grant CXJH_SEU 25062. (Corresponding author: Jinxiang Wang.)

Q. Zhang, J. Wang, and G. Yin are with the School of Mechanical Engineering, Southeast University, Nanjing 211189, China (e-mail: zhangqx@seu.edu.cn; wangjx@seu.edu.cn; ygd@seu.edu.cn).

Y. Xing is with the School of Aerospace, Transport, and Manufacturing, Cranfield University, U. K. (e-mail: yang.x@cranfield.ac.uk).

Z. Fang is with the Department of Civil and Environmental Engineering, National University of Singapore, Singapore 119077 (e-mail: mechinefang@gmail.com).

Y. Liu is with the School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China (email: liuyahui@tsinghua.edu.cn).

B. Related Works

In recent years, various approaches have been proposed to predict vehicle motion trajectories, classified into physics-based, maneuver-based, and interaction-aware models [12], [13]. Physics-based models employ explicit vehicle modeling methods, which can only predict the motion state in short time horizons and cannot characterize vehicle interactions. Maneuver-based models can achieve vehicle TP under specific behaviors but are only suitable for weak interaction driving scenarios. In contrast, interaction-aware models can model the interactions between traffic participants and predict trajectories over extended time horizons. In [14], [15], the dynamic interactions between vehicles are translated into optimal control problems through optimization methods or game theory. In [16], [17], the feature sequence of vehicles is input into the Long Short-Term Memory (LSTM) networks, and the spatial interactions between traffic participants are extracted through cost function or pooling module with fully connected layers. In these studies, different cost functions need to be designed for different driving scenarios, limiting the model's generalization ability. Furthermore, the pooling layer ignores spatial location features, which are key for vehicles to predict safe trajectories in dynamic and uncertain environments [18].

Recent efforts with interaction-aware models have increasingly focused on leveraging deep learning methods to enhance TP capabilities in dynamic and complex driving environments. Deep learning-based models stand out for their adeptness in capturing many vehicle interaction characteristics, making them pivotal in addressing the challenges of TP [19]. Nachiket et al. [20] utilized the convolution and social pooling layer to replace the fully connected layer to extract the spatial dependence between vehicles, which has translation invariance when processing input data. The authors of [21] employed the Convolutional Neural Network (CNN) to learn the interaction between the TV and other environmental vehicles and obtain accurately predicted trajectories. Based on the graph-structured data transformed from the spatial structural information among vehicles, the Graph Neural Network (GNN) can capture complex interactions among traffic participants and predict trajectories [22]. The work in [23] proposed a Graph Convolutional Neural Network (GCN) with vehicle edge feature enhancement, which improves the algorithm's scalability in dynamic driving environments with different numbers of vehicles. However, some studies have pointed out that the memory mechanism of LSTM cannot process data in parallel, limiting the model's generalization capability [24], [25]. To this end, the Transformer network based on the attention mechanism is proposed to capture the long-range dependencies in the sequence and perform global parallel computation [26]. Mo et al. [27] introduced the Transformer method into GCN to explicitly extract the interaction features between traffic participants, achieving TP for multiple heterogeneous vehicles.

The attention-based transformer network performs well in handling time-series tasks and has received widespread attention in vehicle TP in recent years. Messaoud et al. [28] combined convolutional layers with the non-local multi-head attention mechanism to capture the relative importance of SVs

concerning the TV. In [29], a multi-head attention module was employed to handle embedded trajectory and motion information and vehicle class features and achieved significant results in TP. The work in [30] utilized convolutional social pooling and multi-head attention mechanism to quantify the social interactions of TVs in traffic environments and generate the distribution of its future trajectories. The authors of [31] proposed a dual-attention TP method based on a temporal attention mechanism and a spatial attention mechanism to maximize the utilization of temporal information and reduce trajectory uncertainty. However, these studies often use complex grid data processing methods, which increase the computational complexity and make it difficult to adapt to actual driving conditions [32]. In addition, the model's performance in long-term prediction needs to be improved.

Kim et al. [33] combined the embedded vehicle historical trajectory information and lane information into the input of the multi-head attention layer and achieved good results in a short-term prediction scenario of 3 s. In [34], a TP network based on temporal pattern attention was proposed to formulate the interaction model among vehicles. The author of [35] proposed an LSTM model with spatial-temporal attention mechanisms to improve the interpretability of vehicle TP. Zhang et al. [36] utilized a graph attention network to capture vehicle interactions, achieving the prediction of heterogeneous vehicle trajectories in different traffic scenarios. In [37], a non-autoregressive Transformer model with multi-attention learning was designed to attain multimodal vehicle TP. Hu et al. [38] introduced a TP network that combines temporal pattern attention and graph convolutional sequence encoding to mitigate data-dependent uncertainty such as machine-perceptual errors and observation noise. The input features in the above studies focus on the vehicle class and the vehicle's historical time series information, such as coordinates, speed, and acceleration, while the potential impacts of different vehicle driving styles is ignored. In fact, due to the influence of driving experience, gender, personality, and other factors, there are significant differences in the driving behaviors of different drivers [39]. Research in [40] shows that the vehicle's driving style will significantly affect TP.

C. Contribution and Paper Organization

Although interaction-aware motion prediction models have been proposed, their accuracy in long-term TP needs to be improved, and different driving modes between vehicles need to be considered. The rational thinking paradigm behind the human task-solving process involves focusing on lane information based on the driving intention, judging the driving mode of SVs based on their motion features and assigning differentiated weights to SVs in different lanes. Existing research has shown that human driving task processing (especially skilled/safer drivers) is the optimal mode [9], [41]. Inspired by the paradigm observed in human driving task processing, this paper proposes an interaction-aware and driving style-aware TP model to enhance the AV's scene-understanding capability. The proposed model improves prediction accuracy with multi-head attention mechanisms, and despite the increased computational cost compared to simpler models, the performance

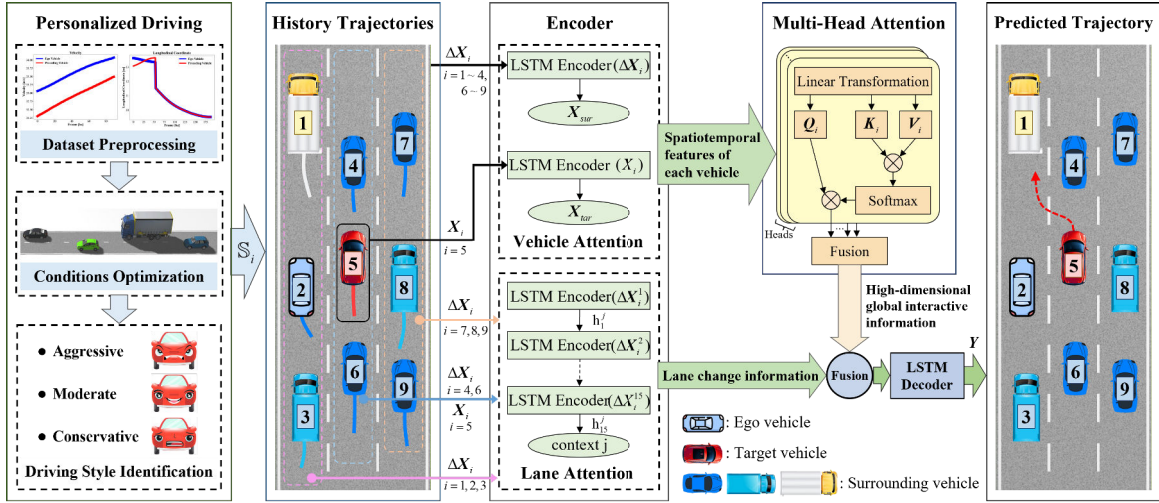


Fig. 1. The proposed interaction-aware and driving style-aware vehicle trajectory prediction framework: The model inputs, X_i and ΔX_i , are 15-dimensional vectors that encapsulate 3 s of historical motion data. The model output, Y , is a 25-dimensional vector representing the TV's trajectory over the next 5 s.

enhancement in long-term predictions and complex scenarios ensures its efficiency in real-time prediction. The contributions of this paper are summarized as follows:

(1) An interaction-aware and driving style-aware TP system suitable for TV in the HVG is proposed, including personalized driving style identification and interaction mechanism modeling of the TV and SVs.

(2) Based on the optimized driving conditions, vehicle driving styles with rich information are identified and utilized as input for the TP model to achieve data dimensionality reduction. Furthermore, the evaluation method for the overall driving style of the HVG is proposed, revealing for the first time the multiple impacts of the topological driving styles within the HVG on the TP accuracy of the TV.

(3) A novel vehicle interaction mechanism that fully utilizes the spatiotemporal coupling information within the HVG is proposed to enhance the performance of TP. The lane attention mechanism prioritizes LC information, reducing computational complexity; the multi-head attention mechanism extracts high-dimensional global interaction information to enhance the vehicle's perception capability of traffic situations.

The rest of this paper is organized as follows. Section II introduces the proposed interaction-aware and driving-style-aware vehicle TP method. Section III presents the experimental results and analyses. Finally, Section IV concludes this paper.

II. METHODOLOGY

A. System Framework

The proposed framework of the interaction-aware and driving style-aware vehicle TP is shown in Fig. 1. The general system architecture incorporates key modules, including personalized driving style identification, fine-grained historical trajectory encoding, and global dynamic interaction modeling, which enables more accurate capturing of complex interactions between the TV and SVs within the HVG. Specifically, the personalized driving style identification module focuses on generating the driving style labels for the eight directly adjacent vehicles by dataset preprocessing and driving conditions

optimization. The obtained vehicle driving style information is one of the key inputs to the TP model. In the encoder part, to ensure the persistence of historical trajectory encoding, a time-series encoding mechanism using LSTM networks is employed, enabling the capture of temporal dependencies and maintaining consistency across time steps. The encoding at each time step integrates the previous state, ensuring the continuity and coherence of trajectory information. The model continuously updates the latent encoding throughout the prediction. The LSTM encoder processes vehicle data sequentially, capturing temporal dependencies and ensuring continuity. The multi-head attention mechanism refines the encoding by integrating global context, adjusting it dynamically at each step. The decoder LSTM, initialized with the encoder's hidden states, updates the encoding over 25 steps, reflecting temporal and interactive dependencies for accurate predictions. The input of lane attention is on a lane basis, including the historical trajectory information of the three vehicles on each lane, and the input to the vehicle attention mechanism involves encoding the historical trajectory information of two sets of vehicles: the TV and its eight directly adjacent vehicles. It should be noted that the sampling frequency of this paper is 5 Hz, so the 3 s historical trajectory is 15 frames. Subsequently, the lane-changing information obtained by the lane attention and the high-dimensional global interaction information captured by the multi-head attention are fused. Finally, the single predicted trajectory for the TV (No. 5) can be obtained by the LSTM decoder, as illustrated by the red dashed line in Fig. 1. It should be noted that vehicle 2 is the ego vehicle, it needs to predict the trajectories of all vehicles directly adjacent to it to accurately assess driving risks and identify safe and feasible driving areas. To avoid confusion and better highlight the focus of this paper, Fig. 1 only presents the relationship between the ego vehicle (No. 2) and the TV. The definitions of mathematical symbols in Fig. 1 are detailed in Section II-C.

1) *Driving Scenario*: The driving scenario studied involves an HVG consisting of one TV and eight SVs, occupying three lanes with three vehicles per lane, including cars and

trucks. As shown in Fig. 1, the solid line behind each vehicle represents its past 3 s trajectory, while the dashed line ahead indicates the 5 s predicted trajectory of the TV.

2) *Driving Style Awareness*: By considering personalized driving styles, the TP model captures driver behaviors and preferences, resulting in more accurate and reliable TP in various driving scenarios. The suitable road section data is selected from the highD dataset, and data preprocessing is performed on three aspects to extract effective data: vehicle speed, abnormal data, and trajectory duration. Then, condition optimization is conducted, and data under cruising conditions (CC) are excluded to enhance data quality and more effectively extract vehicle features. The Gaussian Mixture Model (GMM) unsupervised clustering algorithm is employed to assign the optimal clustering label to the eight directly adjacent vehicles. A detailed description of the personalized driving style identification will be introduced in Section II-B.

3) *Interaction Awareness*: The novel vehicle interaction mechanism employs an encoder-decoder architecture comprising lane attention and multi-head interaction mechanisms. The lane attention mechanism prioritizes and enhances the focus on the target lane. In contrast, the multi-head interaction mechanism captures the complex interactions and dependencies between SVs on different lanes and TVs. By fully utilizing the spatiotemporal coupling information of the HVG, the vehicle interaction mechanism can comprehensively and accurately understand the driving environment. A detailed description of the vehicle interaction mechanism will be elaborated in Section II-C. The evaluation and comparison of the TP model will be performed in Section III.

B. Identification of Personalized Driving Styles

1) *Preprocessing of the highD Dataset*: The proposed model is trained and tested on the trajectories extracted from the highD dataset, a high-precision natural vehicle trajectory dataset extracted from German highways [42], as shown in Fig. 2. The sampling rate is 25 Hz, and the mean visible duration for each vehicle is 13.6 s. Considering various factors such as the coherence of research, the scale of the data, and the richness of road features, recordings with location ID 1 are selected, covering 11.07 hours and involving 85,962 vehicles, with cars and trucks making up 81.14% and 18.86%, respectively. The studied sample accounts for 61.67% (as shown in Table I) of the highD dataset, making it highly representative. The highD dataset, with its bird's-eye view, large volume, high accuracy, and ease of processing, ensures data integrity, making it ideal for the studied task.

Besides, the raw highD dataset needs to be preprocessed to address the specific challenges of TP more effectively. Vehicles with a mean velocity below 15 m/s are excluded due to homogenized behaviors in severe traffic congestion. Anomalous data, such as zero values in distance headway (DHW) and time headway (THW), are replaced with the last observed non-zero values, assuming THW remains unchanged when a vehicle exits the observed section. During the operation of an autonomous vehicle, it is essential to evaluate the driving styles of the SVs and predict their trajectory based on the historical trajectory data from both themselves and the SVs.

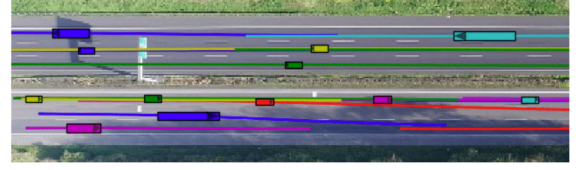


Fig. 2. Highway drone dataset highD [42]: recording covers about 420 m of two-way roads.

TABLE I
DATA DISTRIBUTION OF THE HIGHD DATASET

Location ID	Recording ID	Percentage	Road type	Velocity limit
1	11-14, 25-57	61.67%	Six-lane	120 km/h
2	1, 2, 3	5.00%	Four-lane	/
3	4, 5, 6	5.00%	Six-lane	130 km/h
4	7, 8, 9, 10	6.66%	Six-lane	/
5	15-24	16.67%	Four-lane	/
6	58, 59, 60	5.00%	Six-lane	/

For consistency with similar studies, the historical trajectory duration is set to 3 s, and the predicted duration to 5 s, excluding vehicles with a total duration of less than 8 s.

2) *Optimization of Driving Conditions*: The driving conditions in the highD are optimized to improve the data quality and ensure the accuracy of TP. highD contains many realistic highway scenarios but does not distinguish between driving conditions. In fact, under the CC, vehicles typically have no intention to change lanes and are largely unaffected by SVs. Therefore, it is necessary to eliminate the unsuitable CC for studies of personalized driving styles. Firstly, vehicles that did not change lanes while driving are extracted, accounting for more than 89% of the highD dataset. Then, car-following conditions (CF) and CC are further distinguished. There is no uniform standard for the strategy in this area, but it is mainly considered from the two aspects of THW and time-to-collision (TTC). This paper employs a hybrid strategy, combining THW and TTC, to fully account for the influence of the preceding vehicle's (PV) motion state on driving conditions. The approach aligns more closely with real-world scenarios and ensures an accurate division between CF and CC.

$$d_t = \max \{t_{hc}v_e + t_{cc}(v_e - v_p), d_0\} \quad (1)$$

where t_{hc} represents the constant of THW, v_e signifies the velocity of the TV, t_{cc} is the constant of TTC, v_p signifies the velocity of PV, d_0 represents the minimum switching distance, and d_t denotes the switching distance. According to the Highway Capacity Manual of the U. S., vehicles are considered to be in CC when THW exceeds 5 s [43]. Besides, the traffic flow theory states that CF occurs when THW is less than 100 meters [44]. Therefore, in this work, $t_{hc} = 5$ s, and $d_0 = 100$ m. When $t_{cc} = 1$ s, CC accounted for 7.4% with 5311 vehicles, while for $t_{cc} = 2$ s, CC accounted for 6.2% with 4453 vehicles. Considering the limited impacts of PV velocity on the TV motion state during CC, this paper chooses $t_{cc} = 1$ s. Based on the optimized driving condition, 7.4% of the vehicles in the study sample are classified as CC, which is relatively low. This phenomenon can be attributed to two main reasons: 1) the road segments used by the highD dataset are relatively short, limiting data collection on cruising vehicles with large DHW; 2) the road surveys are conducted exclusively on weekdays, with data collection concentrated

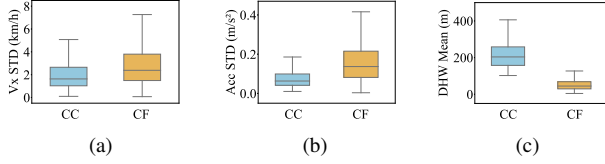


Fig. 3. Boxplots of driving behavior under cruising and car-following conditions: (a) Standard deviation of longitudinal velocity; (b) Standard deviation of longitudinal acceleration; (c) Mean distance headway.

during the peak hours of 8:30-9:30, 11:00-12:30, and 16:30-18:00, resulting in higher traffic volumes with few cruising vehicles. The proportion of cruising vehicles on the highway tends to be higher during non-peak hours on weekends.

Statistical analysis is conducted to obtain driving behavior data under CC and CF, as shown in Fig. 3. It can be found that the medians for the standard deviation (STD) of longitudinal velocity in CC and CF are 1.63 km/h and 2.38 km/h, respectively. The interquartile range (IQR) for CF is notably higher, 1.42 times greater than that of CC. Additionally, the data distribution for the STD of longitudinal acceleration in CC and CF is consistent with Fig. 3(a), which indicates that vehicle speed tends to be relatively stable under CC. Conversely, in CF, frequent acceleration and deceleration events resulting from the influence of PV can lead to substantial speed variations. As illustrated in Fig. 3(c), the lower and upper quartiles for the mean of DHW in CC and CF are 157.6 m, 258.5 m, 29.4 m, and 68.5 m, respectively, demonstrating a significant differentiation in the following distance between these two driving conditions. The proposed strategy has effectively distinguished between cruising and car-following conditions. It should be noted that the differences in behavior between cruising and car-following are general across the dataset.

3) *Driving Style Recognition Based on GMM*: After excluding CC data from the samples, the personalized driving styles are identified based on GMM, including aggressive, moderate, and conservative. In CF, the TV's speed is greatly influenced by SVs, making it less effective in distinguishing driving styles. However, features related to speed fluctuations and following distance are more informative in reflecting the personalized preferences of different drivers [40]. In practice, drivers can intuitively perceive the acceleration through visual and auditory cues, such as relative speed changes, engine sounds, and brake lights. Besides, the vehicle's sophisticated sensors can accurately measure the acceleration of surrounding traffic vehicles [45]. The multidimensional statistical measures illustrated in Table II are applied to enhance the accuracy and robustness of the personalized driving style recognition algorithm. Specifically, the STD of velocity directly reflects the variability in speed, while the maximum, minimum, STD, and mean values for acceleration and Jerk indirectly indicate the speed fluctuations. The associated features of distance headway and time headway provide a comprehensive reflection of the spatial and temporal spacing between vehicles. Note that the highD dataset lacks Jerk data, which measures vehicle motion smoothness. Jerk can be calculated as follows:

$$Jerk(t) = \frac{acc(t + \Delta t) - acc(t)}{\Delta t} \quad (2)$$

TABLE II
STATISTICS FOR DRIVING STYLE RECOGNITION

Features	Statistic Values	
Velocity	Velocity STD	
Acceleration	Max Acceleration	Min Acceleration
	Acceleration STD	Mean Acceleration
Jerk	Max Jerk	Min Jerk
	Jerk STD	Mean Jerk
Distance Headway	Max Distance Headway	Min Distance Headway
	Distance Headway STD	Mean Distance Headway
Time Headway	Max Time Headway	Min Time Headway
	Time Headway STD	Mean Time Headway

Based on the vehicle statistical features, GMM is employed to generate driving styles with significant differences. Unlike K-means or hierarchical clustering methods, GMM can model arbitrary-shaped data by combining multiple Gaussian distributions, making it well-suited for classifying diverse driving styles. GMM has been widely utilized in intelligent vehicles and AD systems [46], [47], which can be expressed as:

$$p(\chi_m | \Theta) = \sum_{n=1}^K \pi_n \mathcal{N}(\chi_m | \mu_n, \Sigma_n) \quad (3)$$

where K represents the number of Gaussian components, π_n is the weight of the n -th Gaussian component, Θ represents the parameter set of the GMM, which can be expressed as $\Theta = \{\pi_n, \mu_n, \Sigma_n\}$, $p(\chi_m | \Theta)$ represents the probability density function of the sample vector χ_m , and $\mathcal{N}(\chi_m | \mu_n, \Sigma_n)$ denotes the Gaussian distribution with mean vector μ_n and covariance matrix Σ_n , which can be described as:

$$\mathcal{N}(\chi_m | \mu_n, \Sigma_n) = \frac{\exp\left(-\frac{1}{2}(\chi_m - \mu_n)^T \Sigma_n^{-1}(\chi_m - \mu_n)\right)}{\sqrt{(2\pi)^d |\Sigma_n|}} \quad (4)$$

where d is the dimension of the sample vector, GMM estimates the probability distribution parameters by the Expectation-Maximization (EM) algorithm, which iteratively optimizes the likelihood function to reach a local maximum. Moreover, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are applied to assess the performance of GMM, which aids in identifying the most suitable model for fitting the data while avoiding overfitting issues.

TV evaluates the driving styles of SVs based on their motion states within a 3 s time window. Therefore, the motion information from the first 3 s of each vehicle in the sample data is inputted into the GMM for clustering analysis. GMM divides driving styles into three categories and names them based on their dynamic performance: conservative, moderate, and aggressive. In the study sample of 76,098 vehicles, the distribution of driving styles is as follows: 23,861 vehicles (31.36%) are classified as aggressive type, 45,809 vehicles (60.19%) fall into the category of moderate type, and 6,428 vehicles (8.45%) are identified as conservative type. Notably, the tripartite classification of driving styles has been shown to align with psychological and behavioral theories of drivers and has been widely accepted. Utilizing three driving styles as input for TP model accomplishes input data dimensionality reduction, reducing the model's computational complexity and enhancing its generalization capability [39], [40], [48].

TABLE III
COMPARISON OF PERSONALIZED DRIVING STYLES

Parameter	Conservative	Moderate	Aggressive
Velocity STD (m/s)	0.287	0.164	0.258
ACC Max (m/s ²)	0.191	0.086	0.176
ACC Min (m/s ²)	-0.264	-0.081	-0.232
ACC STD (m/s ²)	0.146	0.057	0.136
ACC Mean (m/s ²)	-0.050	0.002	-0.028
Jerk Max (m/s ³)	0.012	0.010	0.010
Jerk Min (m/s ³)	-0.015	-0.010	-0.011
Jerk STD (m/s ³)	0.007	0.005	0.006
Jerk Mean (m/s ³)	-0.002	-0.001	-0.001
DHW Max (m)	108.571	60.611	53.018
DHW Min (m)	51.041	55.952	48.233
DHW STD (m)	21.442	1.369	1.420
DHW Mean (m)	78.234	58.369	50.679
THW Max (s)	3.496	2.050	1.771
THW Min (s)	1.662	1.899	1.626
THW STD (s)	0.681	0.045	0.043
THW Mean (s)	2.523	1.976	1.699

The comparison of statistical features for the three driving styles is shown in Table III, where all the features are measured in international units. Aggressive and conservative vehicles exhibit significant speed fluctuations, whereas moderate vehicles demonstrate the least variation in speed. This phenomenon can be attributed to the driving behavior preferences of aggressive and conservative drivers. Aggressive drivers prioritize efficiency, exhibiting frequent speed adjustments due to offensive behaviors. Conservative drivers prioritize safety, reacting to disturbances from surrounding vehicles with frequent speed changes. Moderate drivers, valuing comfort, show more stability and less sensitivity to external disturbances, resulting in smaller speed variations. Both aggressive and conservative vehicles exhibit negative mean acceleration. The aggressive drivers' preference for higher speeds often leads to frequent deceleration when following other vehicles. Conversely, conservative drivers tend to enhance their sense of safety by decelerating. Notably, conservative vehicles demonstrate relatively higher maximum acceleration values, which may be attributed to novice drivers within the conservative group due to less precise control over the throttle pedal. From a spatiotemporal perspective, aggressive vehicles exhibit the smallest mean DHWs and fluctuations, as they are inclined to follow PV closely. Conservative vehicles exhibit the largest mean and fluctuations in THWs, as their cautious driving behaviors maintain larger distances to account for potential hazards from PVs. Their frequent speed and distance adjustments based on PV actions result in greater THW fluctuations. The proposed driving style evaluation system demonstrates a high level of recognition, providing reliable input support for vehicle TP.

C. Trajectory Prediction with Vehicle Interaction Mechanism

1) *Problem Formulation*: The essence of vehicle TP is a time series problem, i.e., predicting the future trajectory of the TV based on its own and SVs' historical trajectories. The historical trajectories can be expressed as:

$$\mathbf{H} = [\mathbf{h}^{(t-t_h+1)}, \dots, \mathbf{h}^{(t-1)}, \mathbf{h}^{(t)}] \quad (5)$$

where $\mathbf{H}$ is the history trajectories, which contains a sequence of vehicle information within a time window, t_h signifies the number of time points in the historical trajectories, and $\mathbf{h}^{(t)} =$

$[\mathbf{X}_1^{(t)}, \dots, \mathbf{X}_i^{(t)}, \mathbf{X}_9^{(t)}]$ denotes the trajectory information of the HVG at time t , $i = 1, 2, \dots, 9$. $\mathbf{X}_i^{(t)}$ is the information of the i -th vehicle at time t , which can be expressed as follows:

$$\mathbf{X}_i^{(t)} = [x_i^{(t)}, y_i^{(t)}, v_{xi}^{(t)}, v_{yi}^{(t)}, a_{xi}^{(t)}, a_{yi}^{(t)}, \mathcal{T}_i, \mathbb{S}_i] \quad (6)$$

where $x_i^{(t)}$, $y_i^{(t)}$, $v_{xi}^{(t)}$, $v_{yi}^{(t)}$, $a_{xi}^{(t)}$, $a_{yi}^{(t)}$ represent the longitudinal position, lateral position, longitudinal velocity, lateral velocity, longitudinal acceleration, and lateral acceleration of the i -th vehicle at time t , respectively. $\mathcal{T}_i$ signifies the vehicle type (car or truck) of the i -th vehicle. $\mathbb{S}_i$ denotes the driving style (aggressive, moderate, and conservative) of the i -th vehicle.

Analyzing the relative motion of SVs helps TV recognize and adapt to different traffic scenarios, achieving safer, more efficient, and more comfortable driving behaviors. Thus, the historical information of relative motion is selected as the model's input to capture the interaction and dynamic changes between vehicles and improve the accuracy of TP results.

$$\mathbf{X} = [\Delta \mathbf{X}_1^t, \dots, \Delta \mathbf{X}_i^t, \Delta \mathbf{X}_8^t, \mathbf{X}_5^t] \quad (7)$$

where $\mathbf{X}$ is the input vector of the model, $\Delta \mathbf{X}_i^t = \mathbf{X}_i^t - \mathbf{X}_5^t$ is the relative motion information of SVs to TV. i ranges from 1 to 9, excluding 5, and $t = [t, t-1, \dots, t-t_h+1]$ is the time steps. $\mathbf{X}_5^t$ is the historical motion information of the TV. The model's output is the TV's future trajectory coordinates.

$$\mathbf{Y} = [\mathbf{Y}_5^{(t+1)}, \dots, \mathbf{Y}_5^{(t+t_f)}, \mathbf{Y}_5^{(t+t_f)}] \quad (8)$$

where $\mathbf{Y}$ is the output vector, t_f is the number of time points in the predicted trajectories, $t_j = 1, 2, \dots, t_f$, and $\mathbf{Y}_5^{(t+t_j)} = [x_5^{(t+t_j)}, y_5^{(t+t_j)}]$.

2) *Lane Attention*: Human drivers prioritize vehicles' trajectories and intentions in the target lane, such as giving more attention to the right lane when turning right. This lane-dependent focus, especially prominent in highway scenarios, motivates the lane attention mechanism, which dynamically adjusts lane weights to enhance emphasis on relevant lane information. Fig. 4-(a) illustrates the workflow of the lane attention mechanism. It can be seen that in HVG, three vehicles in the same lane are grouped. There are three groups, and each group is handled independently. Firstly, the encoder processes the information from each group through LSTM, generating three distinct lane context vectors, namely $\mathbf{h}_{lh}^1$, $\mathbf{h}_{lh}^2$, and $\mathbf{h}_{lh}^3$, are the left, middle, and right lanes, respectively. These vectors are three-dimensional and contain information from all time points in the historical trajectories. Then, the weights of $\mathbf{h}_{lh}^1$, $\mathbf{h}_{lh}^2$, and $\mathbf{h}_{lh}^3$ are determined based on these three lane context vectors and $\mathbf{Y}^{(t-1)}$. It should be noted that $\mathbf{Y}^{(t-1)}$ is the predicted position of TV at the last moment, which is a two-dimensional vector. On this basis, the weighted lane context vectors $\mathbf{C}_t^1$, $\mathbf{C}_t^2$, and $\mathbf{C}_t^3$ can be obtained, corresponding to left lane, middle lane, and right lane, respectively, which are three-dimensional vectors, where t ranges from 1 to t_f . Finally, the total context vector can be described as:

$$\mathbf{C}_t = \Phi(\text{cat}(\mathbf{C}_t^1, \mathbf{C}_t^2, \mathbf{C}_t^3)) \quad (9)$$

where Φ represents a linear layer used for dimension reduction and information fusion, *cat* is utilized to concatenate the data, and $\mathbf{C}_t$ denotes the total context vector.

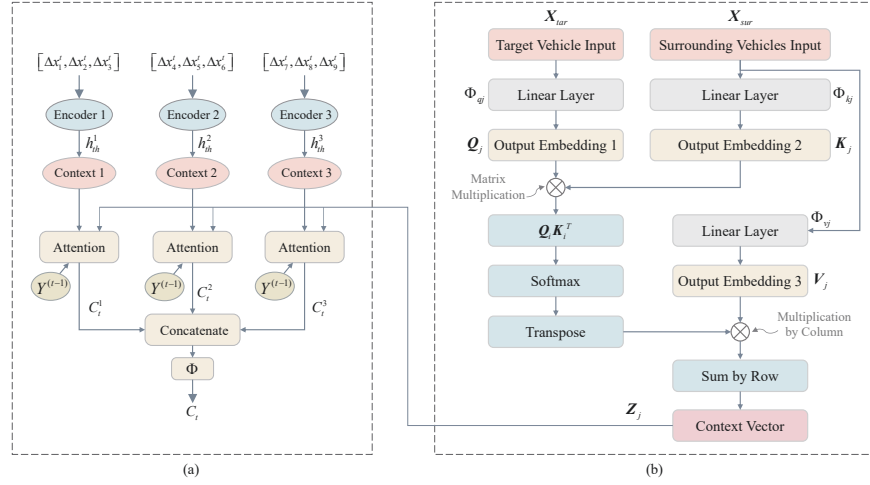


Fig. 4. Vehicle interaction mechanism: (a) Lane attention mechanism; (b) Cross-attention mechanism.

3) *Multi-Head Attention*: The lane attention mechanism provides spatial information about lane occupancy but does not distinguish the impacts of different vehicles within the same lane on the TV. Therefore, the multi-head attention mechanism is introduced to extract high-dimensional global interaction information and investigate the complex relationship between the TV and SVs. It complements the lane attention mechanism by enhancing the AV's ability to perceive traffic situations, particularly in scenarios where surrounding vehicles exhibit erratic or non-standard behaviors, such as sudden LC or unpredictable braking. To showcase the mechanism of multi-head attention, single-layer multi-head attention, known as cross-attention, is taken as an illustrative example. The operational process of the cross-attention mechanism is depicted in Fig. 4-(b). In a manner reminiscent of the classical Transformer structure [27], the cross-attention mechanism utilizes dot product operations to capture the interactive information among vehicles. In pursuit of this goal, the input information of the cross-attention mechanism is classified into two types: TV data, represented as X_{tar} , and SVs data, expressed as X_{sur} . Then, X_{tar} and X_{sur} are processed through three distinct linear layers, Φ_{qj} , Φ_{kj} , and Φ_{vj} , which play a crucial role in extracting key features and enhance representation capability. The “query” (Q), “keys” (K), and “values” (V) can be represented as:

$$Q_j = \Phi_{qj}(X_{tar}) \quad (10)$$

$$K_j = \Phi_{kj}(X_{sur}) \quad (11)$$

$$V_j = \Phi_{vj}(X_{sur}) \quad (12)$$

where Φ_{qj} , Φ_{kj} , and Φ_{vj} are three different linear layers used for dimension reduction and information fusion. Q_j carries the relevant information of TV, while K_j and V_j carry the data regarding SVs. The output of the cross-attention can be obtained through the *softmax* operation.

$$Z_j = \text{softmax} \left(\frac{Q_j K_j^T}{\sqrt{d_k}} \right) V_j \quad (13)$$

where d_k is the column dimension of the Q_j and K_j , and Z_j is the weighted information of the j -th head, indicating the attention weights of SVs in HVG to TV. Finally, the data extracted

from all heads are concatenated and further processed through dimension reduction and information fusion, culminating in the output of the multi-head attention model.

$$Z = \Phi(\text{cat}(Z_1, \dots, Z_j)) \quad (14)$$

where $j = 1, 2, \dots, m$ is the j -th head, m is the number of the heads, the attention head counts (4–12) are tested, and 8 heads are found to be optimal, balancing prediction accuracy and computational cost. Z refers to the output of the multi-head attention model, which encapsulates the high-dimensional interactive information between the SVs and the TV. The output of the multi-head attention model is then prepared for information fusion with the lane attention model. Multi-head attention is capable of learning detailed relationships between every two vehicles. It enhances the lane attention model because it does not consider the weight difference in the same lane while they should still differ. The multi-head attention is introduced to refine the weight distribution for the basic lane attention mechanism since it captures global correlation as a reinforcement after information fusion. During training, the lane attention and multi-head attention mechanisms operate in parallel, fusing their outputs through the context attention module. Furthermore, optimization using a unified loss function ensures balanced contribution from both mechanisms, improving model's ability to capture local and global interactions, thereby enhancing prediction accuracy.

4) *Implementation Details*: To enhance the predictive performance and interpretability of the model, conducting a correlation analysis on input data is necessary to avoid redundancy in the information. The correlation coefficient is a statistical measure that quantifies the strength of the linear relationship between two variables. In this context, the Pearson correlation coefficient is widely employed and is defined as:

$$r_{\zeta\vartheta} = \frac{\sum (\zeta_i - \bar{\zeta})(\vartheta_i - \bar{\vartheta})}{\sqrt{\sum (\zeta_i - \bar{\zeta})^2 \sum (\vartheta_i - \bar{\vartheta})^2}} \quad (15)$$

where ζ_i and ϑ_i represent the observed values of two variables, while $\bar{\zeta}$ and $\bar{\vartheta}$ denote the mean values of the two variables, and $r_{\zeta\vartheta}$ is the Pearson correlation coefficient. The correlation

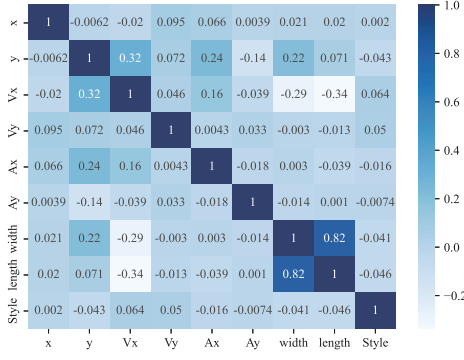


Fig. 5. Correlation coefficient matrix of the input data.

coefficient matrix of the data input to the model is shown in Fig. 5. This paper adopts the widely recognized criteria for interpreting Pearson correlation coefficients: $r_{\zeta\vartheta} > 0.7$ indicates a strong correlation, while $r_{\zeta\vartheta} < 0.4$ signifies a weak correlation [49], [50]. The width and length of vehicles show a strong correlation ($r_{\zeta\vartheta} = 0.82$), vehicle length is chosen to represent vehicle type. This approach minimizes feature redundancy while retaining key information about vehicle characteristics. Other data exhibit weak correlations ($r_{\zeta\vartheta} < 0.32$), ensuring independent information across variables. The HVG represents an ideal scenario due to its rarity in real-world driving situations. Therefore, virtual vehicles are introduced to ensure the consistency of the input data format and enhance the model's adaptability to various driving scenarios. Specifically, the introduced virtual vehicles have the same features as the TV, except for the coordinates. In addition, the relative distance between the two vehicles is guaranteed to be far enough (300 m in the x-axis direction, 0 or 3.75 m in the y-axis direction) to effectively eliminate the influence of the virtual vehicles on the TV. Given that the number of lane-keeping (LK) samples in the highD dataset far exceeds that of lane-changing samples, directly training a model on this dataset may not sufficiently learn and discriminate lane-changing scenarios. Therefore, the data resampling strategy is applied to address the challenge of imbalanced data, which ensures that the model can better understand and recognize various driving scenarios. To this end, the LK samples are downsampled as follows:

$$\mathbf{x}_{frame}^{k2} = \{\mathbf{x}_{frame}^{k1} \mid frame \% 5 = 0\} \quad (16)$$

where $\mathbf{x}_{frame}^{k1}$ and $\mathbf{x}_{frame}^{k2}$ represent the original and downsampled LK data, respectively; $frame$ denotes the sampled frames of the highD dataset. “%” refers to the modulo operator. The lane-changing samples are upsampled five times to achieve data enhancement, which can be expressed as follows:

$$\mathbf{x}_{frame}^{c2} = \{\mathbf{x}_{frame}^{c1} \mid frame \% 5 = 0, 1, 2, 3, 4\} \quad (17)$$

where $\mathbf{x}_{frame}^{c1}$ and $\mathbf{x}_{frame}^{c2}$ signify the original and upsampled lane-changing data, respectively. It should be noted that each of the five trajectories has distinct starting points, with an average speed of 113 km/h for lane-change samples. The longitudinal distance between the starting sampling points of the five trajectories is approximately 1.3 meters, resulting in varied trajectory distributions. Here, each trajectory is unique with variations. Due to the diversity of the high-dimensional

dataset, oversampling does not compromise performance [38], [51]. Finally, the sampling frequency of the study data becomes 5 Hz, and the ratio between LK trajectory and lane-changing trajectory becomes 8:5. Therefore, $t_h = 15$, $t_f = 25$. The model weights are initialized with a truncated Gaussian distribution in the range of $[-0.15, 0.15]$, whose mean and standard deviation are 0 and 0.05, respectively. The Adam optimizer is employed for optimization with an initial learning rate of 0.025 and a weight decay of 0.00003. A learning rate that is too high may cause instability, while one that is too low can slow convergence, increase training time, and risk local minima. The training consists of 160 epochs, with model parameters selected based on the lowest test loss. The model is implemented in Python using the PyTorch framework.

III. EXPERIMENTAL RESULTS AND ANALYSIS

A. Evaluation Metrics and Comparison Models

To evaluate the performance of the TP models, the root mean square error (RMSE), which emphasizes overall prediction accuracy, is adopted to measure the difference between predicted and truth values. RMSE is chosen for its ability to effectively measures the overall error of the trajectory, aligning with the precision requirements of this paper and established practices in the field (e.g., [29], [32], [38], [40], [51]–[53]). The expression of RMSE is as follows:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N ((x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2)} \quad (18)$$

where N represents the number of samples, $\hat{x}_i$ and $\hat{y}_i$ denote the longitudinal and lateral coordinates of the predicted trajectory, respectively. Subsequently, several representative models are selected for comparative analysis, as listed below:

- *V-LSTM* [53]: An LSTM-based encoder-decoder model, where the decoder generates predicted trajectories using spatiotemporal features extracted by the encoder.
- *L-RRNN* [53]: An RRNN-based encoder-decoder model that partitions the driving scene into lane-based input matrices.
- *Sc-RRNN* [53]: An RRNN-based encoder-decoder model that takes the entire driving scene as an input vector.
- *MHA-LSTM* [29]: An LSTM-based encoder-decoder model with multi-head attention to capturing spatial dependencies via grid-based road partitioning, using coordinates, velocity, acceleration, and vehicle type as inputs.
- *TPA-GCSE* [38]: A deep neural network model accounting for spatiotemporal topology and heteroskedastic uncertainty.
- *IATP*: This model applies the proposed interaction mechanism without considering vehicle driving styles.
- *DATP*: The proposed double-aware TP model.

B. Prediction Result and Analysis

1) *Quantitative Results*: The TP results are mainly compared quantitatively from three aspects: total, longitudinal, and lateral RMSE, as shown in Table IV and Fig. 6. It can be seen that V-LSTM can generate reasonable TP results within 2 s but exhibits the weakest predictive performance among the six methods, which can be attributed to LSTM's inherent limitations in capturing long-term spatial dependencies between SVs

TABLE IV
COMPARISON OF RMSE FOR DIFFERENT METHODS @HIGHD

Class	Prediction Horizon(s)	V-LSTM	L-RRNN	Sc-RRNN	MHA-LSTM	TPA-GCSE	IATP	DATP
Total	1	0.31	0.22	0.29	0.06	0.05	0.12	0.10
	2	0.81	0.65	0.69	0.09	0.07	0.18	0.16
	3	1.51	1.31	1.33	0.24	0.18	0.29	0.24
	4	2.48	2.22	2.22	0.59	0.47	0.51	0.43
	5	3.71	3.38	3.33	1.18	0.92	0.97	0.81
	AVG	1.76	1.56	1.57	0.43	0.34	0.41	0.34
Longitudinal	1	0.27	0.22	0.27	0.11	0.06	0.11	0.09
	2	0.74	0.63	0.66	0.17	0.09	0.17	0.15
	3	1.44	1.29	1.30	0.33	0.21	0.27	0.23
	4	2.42	2.19	2.16	0.56	0.51	0.49	0.41
	5	3.65	3.34	3.27	1.42	1.00	0.94	0.78
	AVG	1.70	1.53	1.53	0.52	0.37	0.40	0.33
Lateral	1	0.10	0.05	0.08	0.02	0.04	0.04	0.04
	2	0.32	0.14	0.18	0.09	0.07	0.06	0.06
	3	0.46	0.26	0.30	0.14	0.12	0.10	0.08
	4	0.57	0.37	0.43	0.34	0.22	0.15	0.12
	5	0.65	0.48	0.53	0.59	0.37	0.22	0.17
	AVG	0.42	0.26	0.30	0.24	0.16	0.11	0.10

Note1: AVG is the average RMSE over the prediction horizon of 5 seconds.
Note2: In [29], [38], the longitudinal and lateral RMSEs are calculated from the mean values for each driving maneuver.

and TV. An interesting finding is that L-RRNN and Sc-RRNN TP results are relatively similar. They perform well in the first 2 s, but the TP performance drops rapidly from the 3 s, with RMSE errors exceeding 3.33 m by the 5-th s. The shortcoming of these models is that more attention is paid to time series data than spatial interaction information between vehicles. Another interesting observation is that the lateral errors of L-RRNN are consistently lower than those of Sc-RRNN, suggesting that the approach of generating input matrices based on lane-divided scenarios significantly enhances the model's understanding of interdependencies between lanes. MHA-LSTM performs well in the first 3 s, and its predictive capability does not exhibit a noticeable degradation after that. This phenomenon indicates that the multi-head attention mechanism can effectively establish long-term higher-dimension interactions between SVs and TV. TPA-GCSE performs spatiotemporal topological modeling of the driving environment and addresses the data-dependent uncertainties, which achieves the best average TP performance within 5 s TP horizon among the five baseline methods.

The collaborative synergy of the lane attention mechanism and the multi-head interaction mechanism enables IATP to outperform MHA-LSTM from the 4-th s. Specifically, under similar execution times, IATP improves the TP accuracy by 13.2% and 17.9% for the fourth and 5-th s, respectively, compared to MHA-LSTM. Furthermore, IATP outperforms MHA-LSTM and TPA-GCSE in terms of lateral error starting from the 2-th s, indicating that the lane attention mechanism significantly enhances the model's understanding and TP capabilities for lateral traffic situations such as vehicle LC. However, there is still room for improvement in the TP accuracy of IATP in the first 3 s. The possible reason for this issue is that within short TP horizons, the most recent motion can be accurately predicted through vehicle dynamics [32]. Therefore, simple models are sufficient to ensure excellent TP performance. In comparison, the IATP model is more complex, and its short-term predictive performance experiences a decline. Nonetheless, as the TP range increases, the vehicle's behavior becomes more complex, and the influence of interference factors such as noise also intensifies. At this

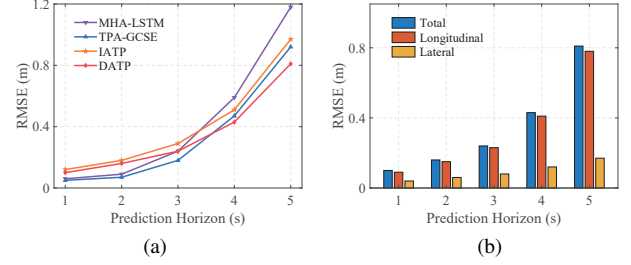


Fig. 6. Quantitative results of the prediction error: (a) Comparison of total prediction error with four different methods; (b) Prediction error under DATP.

TABLE V
COMPUTING COST FOR DIFFERENT MODELS

Metric	CS-LSTM	IATP	DATP
Training time(s)	3656	7137	8079
Inference time(ms)	0.021	0.041	0.047

time, the advantages of the IATP in long-term TP become prominent. Additionally, optimizing the driving conditions of the highD dataset enhances the reliability of data samples, which has also contributed to improving TP results.

Compared to IATP, DATP demonstrates a notable improvement in predictive accuracy across various TP horizons, with a more pronounced enhancement as the TP horizon extends. DATP exhibits a 16.5% reduction in TP error at the 5-th s and an average error reduction of 15.1% across the entire TP horizon compared to IATP. On the other hand, although DATP's TP results for the first 3 s are not as good as those of TPA-GCSE, its precision remains within 0.24 meters, which is already satisfactory for practical application [24], [54], [55]. It is worth noting that the TP performance of DATP in the fourth and 5-th s is 8.5% and 11.9% higher than that of TPA-GCSE, indicating that it has more advantages in long-term TP. In addition, DATP achieves the best average performance in lateral TP, enhancing its overall TP effectiveness. These results all demonstrate that the proposed DATP can not only capture and utilize global high-dimensional dynamic interactions between vehicles within the HVG but also understand and simulate the behavior of vehicles under different driving styles. These advantages of DATP significantly enhance its TP accuracy across diverse driving scenarios.

To assess the real-time performance of the proposed model, the CS-LSTM is used as a reference to compare the computing costs of CS-LSTM, IATP, and DATP. Execution times correspond to TP over a 5 s horizon, as detailed in Table V. The training cost of DATP is 13.2% and 121.0% higher than that of IATP and CS-LSTM, respectively, due to its higher data input dimensions and more complex model structure. However, higher training costs lead to better TP results. Inference times for all models, measured in milliseconds, demonstrate real-time TP capability. For example, STDAN [32], a real-time TP model, achieves an inference time of 0.061 ms, while DATP slightly outperforms it with an even faster inference time. All training and testing were conducted on an NVIDIA GeForce RTX 4070 Ti GPU. It should be noted that hardware performance significantly influences model testing speed, with higher-performance devices enabling faster testing.

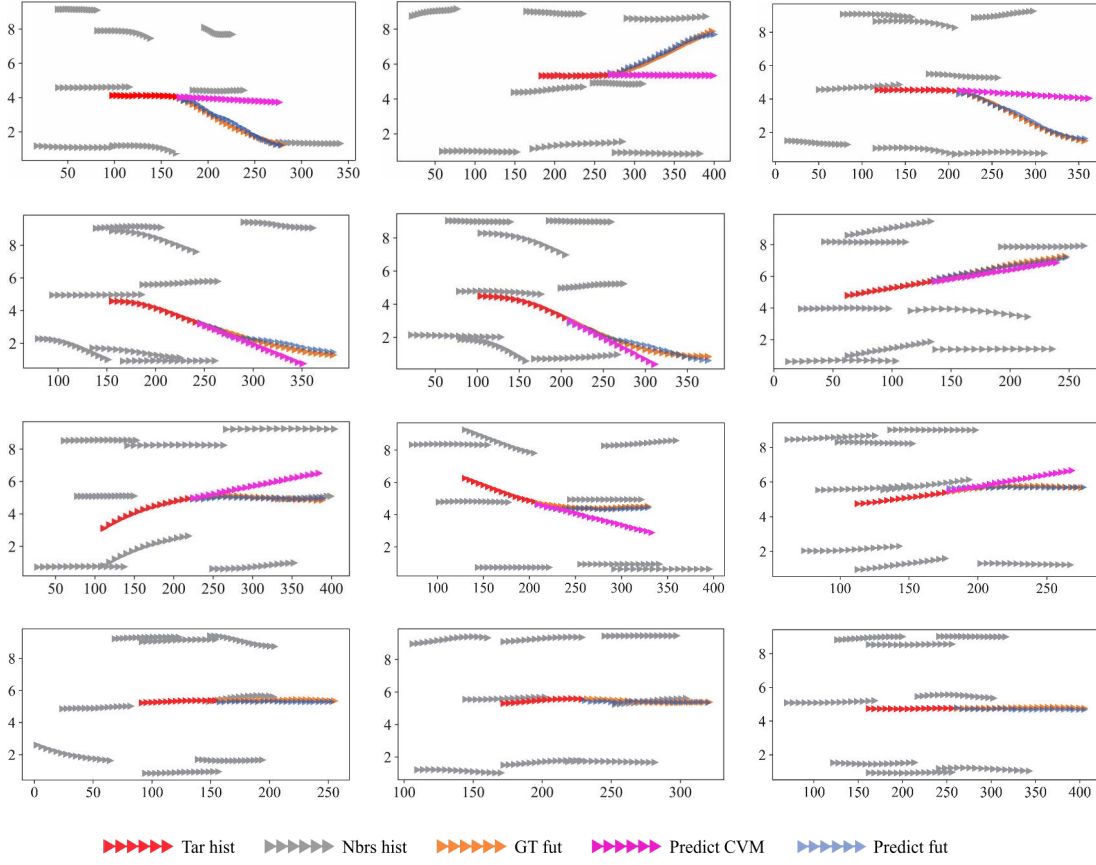


Fig. 7. Visualization examples of prediction results within a 5 s prediction horizon for different driving scenarios. The x-axis and y-axis are longitudinal and lateral positions, respectively, measured in meters. The top row is the TP results before LC, the second row is the TP during LC, the third row exhibits case examples after LC, and the bottom row illustrates TP results for LF. The first three rows all consider the two scenarios of left LC and right LC. Different trajectories are distinguished by color. Tar hist: the history track of the TV; Nbrs hist: the history tracks of SVs; GT fut: the ground truth future trajectory of the TV; Predicted CVM: the predicted trajectory of implemented the constant velocity model; Predicted fut: the predicted trajectory of implemented DATP

2) *Results Visualization*: To gain deeper insights into the model's predictive performance across dynamic, highly interactive scenarios, visualized TP results of the DATP and the constant velocity model (CVM) are presented, as depicted in Fig. 7. lane following (LF) and lane-changing (LC) are the most fundamental driving behaviors. Given that the LC durations on highways range from 3.5 to 8.5 s, with an average of 5.8 s [56], LC is subdivided into three scenarios to verify the model's effectiveness: before lane change, during lane change, and after lane change. Fig. 7 shows in detail the historical trajectories of each vehicle in the HVG, the ground truth future trajectory of the TV, the predicted trajectory output by CVM, and the predicted trajectory output by DATP. Notably, all vehicles share the same sampling frequency for their trajectory data, so their speed can be inferred from the trajectory length.

When TV suddenly changes lanes while driving in a straight line, the DATP model is able to effectively predict the trajectory reflecting the vehicle's lane change intention. In contrast, CVM fails to capture this behavior. The results demonstrate that CVM is unable to recognize the TV's intent to change lanes, whereas interaction-aware models, including the proposed approach, successfully predict trajectories that account for lane change. These comparisons highlight that the proposed method outperforms CVM, particularly in dynamic and highly interactive driving scenarios. Specifically,

the traffic scenarios before lane changes indicate that DATP can capture the lane-changing intentions of the TV based on the surrounding traffic conditions and provide predicted trajectories consistent with the TV's driving intentions. For example, in the driving scenario depicted in the first row and second column of Fig. 7, the unpredictable braking of PV in the TV's lane triggers the TV's lane-changing motivation to avoid potential congestion and continue a smooth journey. Simultaneously, the adjacent vehicle in the right lane travels faster. For safety reasons, this reduces the opportunity for the TV to change lanes to the right. On the other hand, PV in the left lane is moving at a faster speed, while the neighboring vehicle on the left is moving at a slower speed. Therefore, the left lane change presents a reasonable choice. Besides, when vehicles undergo lane changes, as shown in the second row of Fig. 7, the traffic scenario rapidly evolves, such as a swift reduction in the lateral displacement between the TV and the vehicles within the target lane. DATP predicts the behavior of the TV through real-time interaction with SVs to ensure a smooth lane change process, demonstrating its capability to adapt to dynamic and time-varying driving tasks.

As shown in the third row in Fig. 7, in the later stage of vehicle lane change, the lateral relative positions of the TV and SVs within the traffic scenario are stable and no longer show a trend of rapid change. DATP can determine that the

TABLE VI
EVALUATING INDICATORS FOR DRIVING STYLE TOPOLOGY

Description	Math Expression
Overall aggressiveness of HVG	$D_g = \sum_{i=1}^9 \alpha_i S_i$
Proportion of TV's aggressiveness to D_g	$R_t = \frac{\alpha_t S_t}{D_g}$
Aggressiveness ratio (same style as TV to D_g)	$R_s = \sum_{i=1}^{N_s} \frac{\alpha_i S_i}{D_g}$
Lane aggressiveness ratio (TV lane to left lane)	$R_l = \frac{\sum_{i=1}^3 \alpha_e S_e}{\sum_{i=1}^3 \alpha_l S_l}$
Lane aggressiveness ratio (TV lane to right lane)	$R_r = \frac{\sum_{i=1}^3 \alpha_e S_e}{\sum_{i=1}^3 \alpha_r S_r}$

vehicle has completed the lane change maneuver based on the aforementioned dynamic interaction mechanisms and provide a reasonably predicted trajectory for the TV after the lane change. The last row in Fig. 7 displays the TP results in the LF scenarios. An important observation is that the TP performance of DATP is superior in LF scenarios compared to LC situations. This phenomenon is due to LF involving a simpler and smoother motion pattern, while the diversity and uncertainty associated with LC behaviors increase prediction complexity. It is evident that the proposed DATP can comprehensively understand the global high-dimensional dynamic interactions between vehicles to predict various driving behaviors and provide reasonable motion trajectories. Besides, in scenarios with minimal vehicle interactions, such as when TV follows a stable lane, CVM can yield competitive results, particularly for short-term predictions where the trajectory is mainly influenced by the vehicle's current velocity and direction.

3) *Comparison Between Different Driving Styles*: Evaluating the model's performance across different driving styles is essential for assessing its robustness and effectiveness. The data for each driving style are divided into training and testing sets, and the average total RMSE for 5-s TP is computed. The results are as follows: aggressive (0.38m), moderate (0.32m), and conservative (0.47m). Specifically, the highest RMSE is observed for conservative driving, likely due to its smaller sample size, which may lead to overfitting, along with the inherent high randomness in conservative behavior, such as hesitation and prolonged maneuvers. In contrast, the larger dataset for moderate driving provides greater diversity, and the more stable, deliberate behavior of moderate drivers leads to the best TP accuracy. Although aggressive drivers exhibit high randomness, their larger sample size improves TP accuracy, yielding an RMSE between that of conservative and moderate drivers. These findings demonstrate the model's robustness in handling diverse driving behaviors while highlighting its sensitivity to sample size and driving styles. It performs optimally with a balanced dataset and stable driving patterns, yet still achieves competitive results with more random behaviors.

4) *Sensitivity Analysis*: Three scenarios are tested to evaluate the impact of driving style misclassification on model performance: (1) misclassifying moderate drivers as aggressive, (2) misclassifying conservative drivers as moderate, and (3) misclassifying aggressive drivers as conservative. The results show that the average total RMSE increased by 21.4%, 29.3%,



Fig. 8. Analysis of the importance of driving style-related indicators.

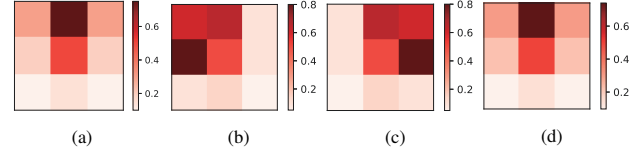


Fig. 9. Weight distribution of each vehicle in the HVG: (a) Straight-line driving. (b) Change lane to the left. (c) Change lane to the right. (d) Integrate various driving behaviors.

and 43.6%, respectively, compared to the correct classification. However, in all cases, the average total RMSE remained below 0.8m, indicating that even with misclassification, the model's performance still offers practical value for AD applications. These findings underscore the significance of driving style differences, which have a substantial impact on accuracy. Besides, the results show that the vehicle interaction mechanism can effectively capture spatiotemporal interactions between vehicles, ensuring model robustness despite misclassification. Future research will explore semi-supervised learning and anomaly detection techniques to mitigate misclassification effects and enhance model reliability and accuracy.

5) *Impacts of driving style topology*: To further explicitly reveal how the driving style of vehicles with HVG impacts the TP of TV, the following quantified performance indicators are designed to evaluate the characteristics of HVG's driving style, as shown in Table VI. Here, α_i represents the weight of the driving style for the i -th vehicle, S_t denotes the TV's driving style, and α_t is the TV's style weight. N_s is the number of vehicles with the same driving style as the TV in the HVG, while S_e and S_l represent the driving styles of vehicles in the ego and left lanes, respectively. α_e and α_l are the weights for vehicles in the ego and left lanes, while S_r and α_r correspond to the right lane. The correlation among the five indicators is shown in Fig. 8. Except for R_l and R_r , the correlation among the indicators is weak. Thus, D_g , R_t , R_s , and R_l are selected as the final driving style indicators. The aggressive, moderate, and conservative driving styles are assigned quantified labels with 3, 2, and 1, respectively. The weight of the vehicle's driving style is influenced by the driving behavior of the TV and its position in the HVG, as illustrated in Fig. 9. When the TV is in a straight-line driving state, it places more emphasis on the front vehicles, as shown in Fig. 9-(a). However, during a lane-changing maneuver, besides being influenced by the leading vehicle, the TV pays closer attention to the adjacent vehicle within the target lane, as depicted in Figs. 9-(b) and (c). It is worth noting that drivers do not frequently change lanes during

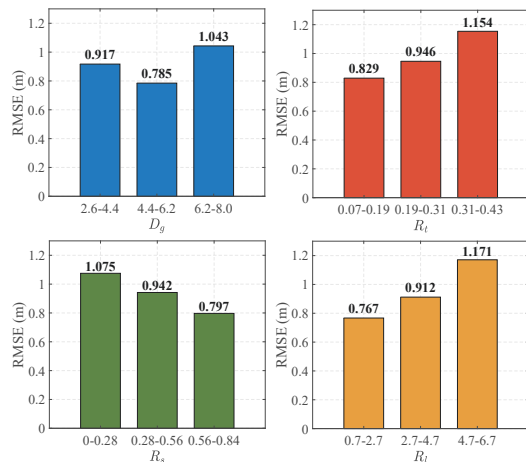


Fig. 10. Trajectory prediction error for different driving style indicators.

real driving scenes, and most of the time is spent in straight-line driving. The vehicle behavior distribution in the highD dataset collected based on real driving scenes also reflects this phenomenon. Specifically, the ratio of straight-line driving, left lane-changing, and right lane-changing is 0.9:0.05:0.05. Therefore, based on these proportions, this paper obtains the weight distribution of each vehicle in the HVG considering different driving behaviors, as shown in Fig. 9-(d).

As shown in Fig. 10, the RMSE of vehicle TP varies with driving styles. TV is prone to experiencing higher TP errors in a more aggressive HVG, characterized by a higher proportion of aggressive vehicles. Aggressive driving behaviors, such as emergency braking and frequent lane changes, introduce instability and unpredictability to HVG dynamics, increasing TP errors. The TP error for TV in a less aggressive HVG exceeds that of a moderately aggressive HVG. Conservative drivers tend to hesitate in choosing paths, exhibiting higher randomness in their driving behavior. Therefore, it is challenging to maintain smooth traffic flow and accurately estimate vehicle trajectories in HVG. The proportion of TV's aggressiveness increases in HVG, its TP error also grows accordingly. This situation highlights the aggressiveness of TV more prominently, simultaneously amplifying the uncertainty of its behavior. Fig. 10 demonstrates that the more SVs with a driving style consistent with TV, the more accurate TP results. The same driving style means that TV and SVs have similar environmental awareness and consistent driving behavior patterns, which improves the level of trust and interaction efficiency between vehicles and enhances the smoothness and predictability of HVG. Another interesting finding is that the accuracy of TP decreases as the disparity in aggressiveness between TV's lane and the target lane increases. The difference in aggressiveness between lanes reflects the inconsistency between the current driving environment of TV and the surrounding driving environment. The greater the inconsistency, the greater the behavioral differences between vehicles and the uncertainty within the driving environment.

6) *Model Generalizability*: To further evaluate the generalizability of the proposed method, the NGSIM dataset is utilized, which has a sampling rate of 10Hz and includes vehicle trajectories spanning at least 8 s. The dataset is preprocessed

TABLE VII
COMPARISON OF RMSE FOR DIFFERENT METHODS @NGSIM

Prediction Horizon(s)	MHA-LSTM	TPA-GCSE	DATP	Error reduction
1	0.41	0.24	0.23	4.17%
2	1.01	0.67	0.69	-2.90%
3	1.74	1.35	1.24	8.15%
4	2.67	2.24	1.92	14.29%
5	3.83	3.32	2.57	22.59%
AVG	1.93	1.56	1.33	14.74%

following the methodology in [40]. Table VII presents a comparative analysis of DATP with MHA-LSTM and TPA-GCSE. Compared to existing baseline methods, DATP achieves the lowest prediction error across nearly all prediction horizons, except the 2-th s prediction horizon. However, the 2-th s prediction error of 0.69m is very close to the best method's error of 0.67m. Notably, compared to the strongest baseline, TPA-GCSE, DATP reduces the prediction error by 22.59% at the 5-s prediction horizon and achieves a 14.74% reduction in average error across the entire prediction horizon. These results demonstrate that the DATP method exhibits exceptional performance in long-term TP. It is worth noting that, as shown in Tables IV and VII, the prediction accuracy on the highD is higher than that on the NGSIM, indicating that the quality of the dataset also plays a role in the prediction results. Overall, the experimental results show that the proposed DATP model considering the interaction and driving style outperforms the compared methods and demonstrates high accuracy and reliability across different datasets.

IV. CONCLUSIONS AND FUTURE WORK

In this paper, an interaction-aware and driving style-aware TP model is proposed for TVs within HVG. By analyzing temporal information such as position, speed, acceleration, vehicle type, and driving style, the model enhances TP accuracy. Three distinct driving styles are generated based on spatiotemporal coupling data of each vehicle in HVG over the past 3 seconds. The DATP model, incorporating lane attention and multi-head interaction mechanisms, emphasizes lateral motion and efficiently extracts high-dimensional global interactions with SVs. Compared to the IATP model, which focuses only on interaction awareness, DATP, which considers driving style, yields superior results, particularly in long-term TP. The model outperforms state-of-the-art models and demonstrates strong adaptability across various dynamic driving tasks. Additionally, this paper investigates the impact of driving style topology in HVG on the TP accuracy of TV.

However, nine vehicles within HVG have limitations, necessitating further investigation to optimize vehicle numbers for different densities. Additional metrics, such as trustworthiness and transparency, should be explored to evaluate the model's performance. Then, external factors like weather, road conditions, and geographic variations in driving behavior will be integrated to enhance trajectory prediction accuracy and model adaptability. Next, we will improve latent encoding regularization, explore advanced encoding techniques, process heterogeneous sensor data, and use synthetic data generation and scenario simulation to enhance model robustness and

adaptability to dynamic and diverse traffic environments. Future work will integrate physics-based and data-driven models to improve accuracy in data-scarce scenarios, extend the prediction horizon beyond 5 seconds to meet the demands of specific long-term TP tasks and visualize spatiotemporal features to enhance model interpretability. Subsequently, the proposed model will be enhanced with GNNs or reinforcement learning, and validated in urban and rural environments using diverse datasets, transfer learning, and real-time traffic data for improved accuracy. Future work will also address deployment challenges through multi-sensor fusion, noise filtering, and edge computing, while also extending the model to include non-motorized road users by integrating pedestrian kinematics with visual information.

REFERENCES

- [1] L. Chen, Y. Li, C. Huang, B. Li, Y. Xing, D. Tian, L. Li, Z. Hu, X. Na, Z. Li, S. Teng, C. Lv, J. Wang, D. Cao, N. Zheng, and F.-Y. Wang, "Milestones in autonomous driving and intelligent vehicles: Survey of surveys," *IEEE Trans. Intell. Veh.*, vol. 8, no. 2, pp. 1046–1056, Feb. 2023.
- [2] Y. Huang, S. Yang, L. Wang, K. Yuan, H. Zheng, and H. Chen, "An efficient self-evolution method of autonomous driving for any given algorithm," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 1, pp. 602–612, Jan. 2024.
- [3] Y. Yan, L. Peng, T. Shen, J. Wang, D. Pi, D. Cao, and G. Yin, "A multi-vehicle game-theoretic framework for decision making and planning of autonomous vehicles in mixed traffic," *IEEE Trans. Intell. Veh.*, vol. 8, no. 11, pp. 4572–4587, Nov. 2023.
- [4] T. Kessler, K. Esterle and A. Knoll, "Mixed-Integer Motion Planning on German Roads Within the Apollo Driving Stack," *IEEE Trans. Intell. Veh.*, vol. 8, no. 1, pp. 851–867, Jan. 2023.
- [5] Y. Huang, J. Du, Z. Yang, Z. Zhou, L. Zhang, and H. Chen, "A survey on trajectory-prediction methods for autonomous driving," *IEEE Trans. Intell. Veh.*, vol. 7, no. 3, pp. 652–674, Sep. 2022.
- [6] Z. Huang, H. Liu, J. Wu, and C. Lv, "Differentiable integrated motion prediction and planning with learnable cost function for autonomous driving," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 11, pp. 15222–15236, Nov. 2024.
- [7] I. Mahdini, A. Mohammadnazar, R. Arvin, and A. J. Khattak, "Integration of automated vehicles in mixed traffic: Evaluating changes in performance of following human-driven vehicles," *Accid. Anal. Prev.*, vol. 152, Art. no. 106006, Mar. 2021.
- [8] S. Wang, Z. Wang, R. Jiang, R. Yan, and L. Du, "Trajectory jerking suppression for mixed traffic flow at a signalized intersection: A trajectory prediction based deep reinforcement learning method," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 10, pp. 18989–19000, Oct. 2022.
- [9] P. Hang, Y. Zhang, and C. Lv, "Brain-inspired modeling and decision-making for human-like autonomous driving in mixed traffic environment," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 10, pp. 10420–10432, Oct. 2023.
- [10] X. Mo, Y. Xing, and C. Lv, "Graph and recurrent neural networkbased vehicle trajectory prediction for highway driving," in *Proc. IEEE Int. Intell. Transp. Syst. Conf. (ITSC)*, Oct. 2021, pp. 1934–1939.
- [11] J. Zhao, T. Qu, X. Gong, and H. Chen, "Interaction-aware personalized trajectory prediction for traffic participant based on interactive multiple model," *IEEE Trans. Intell. Veh.*, vol. 8, no. 3, pp. 2184–2196, Mar. 2023.
- [12] D. E. Benrachou, S. Glaser, M. Elhenawy, and A. Rakotonirainy, "Use of social interaction and intention to improve motion prediction within automated vehicle framework: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 12, pp. 22807–22837, Dec. 2022.
- [13] H. Xu, Z. Niu, B. Jiang, Y. Zhang, S. Chen, Z. Li, M. Gao, and M. Zhu, "ERRT-GA: Expert genetic algorithm with rapidly exploring random tree initialization for multi-UAV path planning," *Drones*, vol. 8, no. 8, pp. 367, Aug. 2024.
- [14] V. Lefkopoulou, M. Menner, A. Domahidi, and M. N. Zeilinger, "Interaction-aware motion prediction for autonomous driving: A multiple model kalman filtering scheme," *IEEE Robot. Automat. Lett.*, vol. 6, no. 1, pp. 80–87, Jan. 2021.
- [15] A. Martin, "Interactive motion prediction using game theory," M. S. thesis, Dept. Automat. Eng., Univ. Degli Studi di Padova, Italy, 2013.
- [16] A. Zyner, S. Worrall, and E. Nebot, "A recurrent neural network solution for predicting driver intention at unsignalized intersections," *IEEE Robot. Automat. Lett.*, vol. 3, no. 3, pp. 1759–1764, Jun. 2018.
- [17] A. Gupta, J. Johnson, L. Fei-Fei, S. Savarese, and A. Alahi, "Social GAN: Socially acceptable trajectories with generative adversarial networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018, pp. 2255–2264.
- [18] Y. Wang, J. Wang, J. Jiang, S. Xu, and J. Wang, "Sa-lstm: A trajectory prediction model for complex off-road multi-agent systems considering situation awareness based on risk field," *IEEE Trans. Veh. Technol.*, vol. 72, no. 11, pp. 14016–14027, Nov. 2023.
- [19] Z. Ding and H. Zhao, "Incorporating driving knowledge in deep learning based vehicle trajectory prediction: A survey," *IEEE Trans. Intell. Veh.*, vol. 8, no. 8, pp. 3996–4015, Apr. 2023.
- [20] N. Deo and M. M. Trivedi, "Convolutional social pooling for vehicle trajectory prediction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2018, pp. 1468–1476.
- [21] T. Zhao, Y. Xu, M. Monfort, W. Choi, C. Baker, Y. Zhao, Y. Wang, and Y. N. Wu, "Multi-agent tensor fusion for contextual trajectory prediction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, 2019, pp. 12126–12134.
- [22] T. Westny, J. Oskarsson, B. Olofsson, and E. Frisk, "Mtp-go: Graphbased probabilistic multi-agent trajectory prediction with neural odes," *IEEE Trans. Intell. Veh.*, vol. 8, no. 9, pp. 4223–4236, Sep. 2023.
- [23] H. Jeon, J. Choi, and D. Kum, "SCALE-Net: Scalable vehicle trajectory prediction network under random number of interacting vehicles via edge-enhanced graph convolutional neural network," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Feb. 2021, pp. 2095–2102.
- [24] W. Luo, B. Yang, and R. Urtasun, "Fast and furious: Real time end-to-end 3d detection, tracking and motion forecasting with a single convolutional net," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018, pp. 3569–3577.
- [25] D. Debnath, F. Vanegas, S. Boiteau, and F. Gonzalez, "An integrated geometric obstacle avoidance and genetic algorithm TSP model for UAV path planning," *Drones*, vol. 8, no. 7, pp. 302, Jul. 2024.
- [26] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 5998–6008.
- [27] X. Mo, Z. Huang, Y. Xing, and C. Lv, "Multi-agent trajectory prediction with heterogeneous edge-enhanced graph attention network," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 9554–9567, Jul. 2022.
- [28] K. Messaoud, I. Yahiaoui, A. Verroust-Blondet, and F. Nashashibi, "Non-local social pooling for vehicle trajectory prediction," in *Proc. IEEE Intell. Veh. Symp. (IV)*, Paris, France, Jun. 2019, pp. 975–980.
- [29] K. Messaoud, I. Yahiaoui, A. Verroust-Blondet, and F. Nashashibi, "Attention based vehicle trajectory prediction," *IEEE Trans. Intell. Veh.*, vol. 6, no. 1, pp. 175–185, Mar. 2021.
- [30] Z. Zuo, X. Wang, S. Guo, Z. Liu, Z. Li and Y. Wang, "Trajectory prediction network of autonomous vehicles with fusion of historical interactive features," *IEEE Trans. Intell. Veh.*, vol. 9, no. 1, pp. 2171–2183, Jan. 2024.
- [31] H. Guo, Q. Meng, D. Cao, H. Chen, J. Liu, and B. Shang, "Vehicle trajectory prediction method coupled with ego vehicle motion trend under dual attention mechanism," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–16, Apr. 2022.
- [32] X. Chen, H. Zhang, F. Zhao, Y. Hu, C. Tan, and J. Yang, "Intention-aware vehicle trajectory prediction based on spatial-temporal dynamic attention network for internet of vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 10, pp. 19471–19483, Oct. 2022.
- [33] H. Kim, D. Kim, G. Kim, J. Cho, and K. Huh, "Multi-head attention based probabilistic vehicle trajectory prediction," in *Proc. IEEE Intell. Veh. Symp. (IV)*, Oct. 2020, pp. 1720–1725.
- [34] H. Hu, Q. Wang, M. Cheng, and Z. Gao, "Trajectory prediction neural network and model interpretation based on temporal pattern attention," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 3, pp. 2746–2759, Nov. 2022.
- [35] L. Lin, W. Li, H. Bi, and L. Qin, "Vehicle trajectory prediction using lstms with spatial-temporal attention mechanisms," *IEEE Intell. Transp. Syst. Mag.*, vol. 14, no. 2, pp. 197–208, Apr. 2022.
- [36] K. Zhang, L. Zhao, C. Dong, L. Wu, and L. Zheng, "AI-TP: Attention-based interaction-aware trajectory prediction for autonomous driving," *IEEE Trans. Intell. Veh.*, vol. 8, no. 1, pp. 73–83, Jan. 2023.
- [37] X. Chen, H. Zhang, F. Zhao, Y. Cai, H. Wang, and Q. Ye, "Vehicle trajectory prediction based on intention-aware non-autoregressive trans-

former with multi-attention learning for internet of vehicles,” *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–12, Jul. 2022.

- [38] H. Hu, Q. Wang, L. Du, Z. Lu, and Z. Gao, “Vehicle trajectory prediction considering aleatoric uncertainty,” *Knowl.-Based Syst.*, vol. 255, no. 109617, Nov. 2022.
- [39] C. Lv, X. Hu, A. Sangiovanni-Vincentelli, Y. Li, C. M. Martinez and D. Cao, “Driving-style-based codesign optimization of an automated electric vehicle: A cyber-physical system approach,” *IEEE Trans. Ind. Electron.*, vol. 66, no. 4, pp. 2965–2975, Apr. 2019.
- [40] Y. Xing, C. Lv, and D. Cao, “Personalized vehicle trajectory prediction based on joint time-series modeling for connected vehicles,” *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 1341–1352, 2020.
- [41] S. Chen, S. Zhang, J. Shang, B. Chen, and N. Zheng, “Brain-inspired cognitive model with attention for self-driving cars,” *IEEE Trans. Cognit. Develop. Syst.*, vol. 11, no. 1, pp. 13–25, Mar. 2019.
- [42] R. Krajewski, J. Bock, L. Kloecker, and L. Eckstein, “The highD dataset: A drone dataset of naturalistic vehicle trajectories on german highways for validation of highly automated driving systems,” in *Proc. IEEE Int. Intell. Transp. Syst. Conf. (ITSC)*, Nov. 2018, pp. 2118–2125.
- [43] H. C. Manual, “Highway capacity manual (6th edition),” in *Transp. Res. Board, Nat. Res. Council.*, Washington, DC, USA, Tech. Rep., 2016.
- [44] H. M. Zhang, R. Kühne, and P. Michalopoulos, “Traffic flow theory-a state-of-the-art report: revised monograph on traffic flow theory,” *Turner-Fairbank Highway Research Center*, Washington, DC, USA, 2002.
- [45] Y. Zhu, M. Jiang, T. Yamamoto, N. Ding, H. Shinkai, H. Aoki, and K. Shimazaki, “A framework for combining lateral and longitudinal acceleration to assess driving styles using unsupervised approach,” *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 1, pp. 638–656, Jan. 2024.
- [46] R. W. Wolcott and R. M. Eustice, “Robust LIDAR localization using multiresolution Gaussian mixture maps for autonomous driving,” *Int. J. Robot. Res.*, vol. 36, no. 3, pp. 292–319, 2017.
- [47] Y. Yan, D. Han, Q. Zhang, J. Wang, D. Pi, D. Chu, and G. Yin, “Event-triggered personalized driving based on passenger’s subjective risk evaluation,” *IEEE Trans. Intell. Transp. Syst.*, vol. 26, no. 12, pp. 1982–1998, Feb. 2025.
- [48] Y. Zhang, Q. Xu, J. Wang, K. Wu, Z. Zheng, and K. Lu, “A learning-based discretionary lane-change decision-making model with driving style awareness,” *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 1, pp. 68–78, Jan. 2023.
- [49] R. J. Hunt, “Percent agreement, Pearson’s correlation, and kappa as measures of inter-examiner reliability,” *J. Dental Res.*, vol. 65, no. 2, pp. 128–130, 1986.
- [50] H. Xiong, S. Shekhar, P. N. Tan, and V. Kumar, “Exploiting a support-based upper bound of Pearson’s correlation coefficient for efficiently identifying strongly correlated pairs,” in *Proc. 10th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2004, pp. 334–343.
- [51] Z. Peng, J. Yan, H. Yin, Y. Wen, W. Ge, T. Watzel, and G. Rigoll, “Efficient interaction-aware trajectory prediction model based on multi-head attention,” *Automotive Innovation*, vol. 7, pp. 258–270, Apr. 2024.
- [52] L. Li, W. Zhao, C. Wang, Q. Chen, and F. Chen, “BRAM-ED: Vehicle trajectory prediction considering the change of driving behavior,” *IEEE/ASME Trans. Mechatron.*, vol. 27, no. 6, pp. 5690–5700, Dec. 2022.
- [53] K. Messaoud, I. Yahiaoui, A. Verroust-Blondet, and F. Nashashibi, “Relational recurrent neural networks for vehicle trajectory prediction,” in *Proc. IEEE Intell. Transp. Syst. Conf. (ITSC)*, Oct. 2019, pp. 1813–1818.
- [54] S. Jia, Y. Zhang, X. Li, X. Na, Y. Wang, B. Gao, B. Zhu, and R. Yu, “Interactive decision-making with switchable game modes for automated vehicles at intersections,” *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 11, pp. 11785–11799, Nov. 2023.
- [55] M. Tian, Q. Zhang, D. Tian, L. Jin, J. Li, et al., “Pre-stability control for in-wheel-motor-driven electric vehicles with dynamic state prediction,” *IEEE Trans. Intell. Veh.*, vol. 9, no. 3, pp. 4541–4554, Mar. 2024.
- [56] T. Toledo and D. Zohar, “Modeling duration of lane changes,” *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1999, no. 1, pp. 71–79, Jan. 2007.



Qixiang Zhang (Graduate Student Member, IEEE) received the M.S. degree (cum laude) in vehicle engineering from Jilin University, Changchun, China, in 2022. He is currently working toward the Ph.D. degree in vehicle engineering, Southeast University, Nanjing, China. His research interests include vehicle dynamics and control, chassis-by-wire, and decision-making and motion planning for autonomous vehicles.



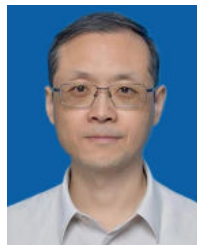
Yang Xing (Senior Member, IEEE) received the Ph.D. degree from Cranfield University, Cranfield, U.K., in 2018. In 2021, he joined Cranfield University as a Lecturer in applied AI for engineering. Before that, he finished the Postdoctoral Fellowship with the University of Oxford, Oxford, U.K., and Nanyang Technological University, Singapore, respectively. His research interest focusses on applied artificial intelligence on human-autonomy collaboration and human-centred autonomous vehicles. He is an Associate Editor for IEEE TRANS. ON INTELLIGENT VEHICLE, IEEE TRANS. ON NEURAL NETWORKS AND LEARNING SYSTEMS. He was also a Guest Editor of IEEE INTERNET OF THINGS JOURNAL, and IEEE ITS Magazine.



Jinxiang Wang (Member, IEEE) received the B.S. degree in mechanical engineering and automation and the Ph.D. degree in vehicle engineering from Southeast University, Nanjing, China, in 2002 and 2010, respectively. From 2014 to 2015, he was a Visiting Research Scholar with the Department of Mechanical and Aerospace Engineering, The Ohio State University, Columbus, OH, USA. He is currently a Professor with the School of Mechanical Engineering, Southeast University. His research interests include vehicle dynamics and control, assisted-driving system, and control on autonomous vehicles.



Zhenwu Fang (Member, IEEE) received the B.S. degree in mechanical engineering and automation and the M.S. and Ph. D degrees in vehicle engineering in 2017, 2020 and 2024, respectively, from the School of Mechanical Engineering, Southeast University, Nanjing, China. He is currently a Research Fellow with the Department of Civil and Environmental Engineering, National University of Singapore. His research interests include driver-vehicle system dynamics, driver state recognition, and human-machine shared control.



Yahui Liu (Senior Member, IEEE) received the B.S. degree in automotive engineering from Jilin University, Changchun, China, in 2003, and the Ph.D. degree in automotive engineering from Beihang University, Beijing, China, in 2009. He is currently an Associate Professor with the School of Vehicle and Mobility, Tsinghua University, Beijing, China. His research interests include driver-vehicle system dynamics, driver-vehicle automation collaboration, and shared control. He serves as an Associate Editor for JOURNAL OF DYNAMIC SYSTEMS, and MEASUREMENT AND CONTROL, an Editorial Board Member for MACHINES and CYBER-PHYSICAL SYSTEMS.



Guodong Yin (Senior Member, IEEE) received the M.S. and Ph.D. degrees in vehicle engineering from Southeast University, Nanjing, China, in 2003 and 2007, respectively. From August 2011 to August 2012, he was a Visiting Scholar with the Department of Mechanical and Aerospace Engineering, Ohio State University, Columbus, OH, USA. He is currently a Professor with the School of Mechanical Engineering, Southeast University. His current research interests include vehicle dynamics and control, chassis-by-wire, connected vehicles, and multi-agent control. He serves as an Associate Editor or Editorial Board Member for journals such as IEEE TRANS. ON INTELLIGENT VEHICLES, JOURNAL OF INTELLIGENT AND CONNECTED VEHICLES. He was the recipient of the National Science Fund for Distinguished Young Scholars and the Fellow of the China Society of Automotive Engineers.



DSPACE

<https://dspace.org/>

Interaction-aware and driving style-aware trajectory prediction for heterogeneous vehicles in mixed traffic environment

Zhang, Qixiang; Xing, Yang; Wang, Jinxiang; Fang, Zhenwu; Liu, Yahui ...

2025-07

Institute of Electrical and Electronics Engineers (IEEE)

<https://doi.org/10.1109/tits.2025.3553697>

Zhang Q, Xing Y, Wang J, et al., (2025) Interaction-aware and driving style-aware trajectory prediction for heterogeneous vehicles in mixed traffic environment. IEEE Transactions on Intelligent Transportation Systems, Volume 26, Issue 7, July 2025, pp. 10710-10724