applied sciences

MDPI

*Article*

# Spatio-Temporal-Attention-Based Vehicle Trajectory Prediction Considering Multi-Vehicle Interaction in Mixed Traffic Flow

**Jie Zeng [1,2], Yue Ren [3,*], Kan Wang [1,2], Xiong Hu [1,2] and Jiufa Li [4]**

[1] China Merchants Testing Vehicle Technology Research Institute Co., Ltd., Chongqing 401329, China; cjzengjie@cmhk.com (J.Z.); cjhuxiong@cmhk.com (X.H.)
[2] Chongqing Key Laboratory of Industry and Informatization of Automotive Active Safety Testing Technology, Chongqing 401329, China
[3] College of Engineering and Technology, Southwest University, Chongqing 400715, China
[4] College of Artificial Intelligence, Southwest University, Chongqing 400715, China
* Correspondence: renyueok@hotmail.com; Tel.: +86-15223081321

**Abstract:** As a link connecting the environmental perception system and the decision-making system, accurate obstacle trajectory prediction provides a reliable guarantee of correct decision-making by autonomous vehicles. Oriented toward a mixed human-driven and machine-driven traffic environment, a vehicle trajectory prediction algorithm based on an encoding–decoding framework composed of a multiple-attention mechanism is proposed. Firstly, a directed graph is used to describe vehicle–vehicle motion dependencies. Then, by calculating the repulsive force between vehicles using a priori edge information based on the artificial potential field theory, vehicle–vehicle interaction coefficients are extracted via a graph attention mechanism (GAT). Subsequently, after concatenating the vehicle–vehicle interaction feature with the encoded vehicle trajectory vectors, a spatio-temporal attention mechanism is applied to determine the coupling relationship of hidden vectors. Finally, the predicted trajectory is generated by a gated recurrent unit (GRU) decoder. The training and evaluation of the proposed model were conducted on the NGSIM public dataset. The test results demonstrated that compared with existing baseline models, our approach has fewer prediction errors and better robustness. In addition, introducing artificial potential fields into the attention mechanism causes the model to have better interpretability.

**Keywords:** trajectory prediction; encoding–decoding framework; vehicle–vehicle interaction; spatio-temporal attention



**Citation:** Zeng, J.; Ren, Y.; Wang, K.; Hu, X.; Li, J. Spatio-Temporal-Attention-Based Vehicle Trajectory Prediction Considering Multi-Vehicle Interaction in Mixed Traffic Flow. *Appl. Sci.* **2024**, *14*, 161. https://doi.org/10.3390/app14010161

Academic Editor: Mohammed Chadli

Received: 25 October 2023
Revised: 12 December 2023
Accepted: 14 December 2023
Published: 24 December 2023



## 1. Introduction

Correct decision-making and precise control are fundamental to the safety of autonomous vehicles, which plays a vital role in their widespread adoption. The accurate perception and prediction of the motion of surrounding obstacles greatly contribute to autonomous vehicles making safe and comfortable decisions. However, it is still too early to implement fully autonomous driving. Autonomous vehicles will remain under testing in dynamic interactive scenarios of human-driven and machine-driven mixed traffic flow for a long time. Due to differences in driving skills, driving styles, and degrees of autonomous driving, predicting the future trajectories of surrounding vehicles and their drivers' intents is one of the major challenges for autonomous vehicles, which has become a research hotspot in recent years.

Early vehicle motion prediction methods can be mainly divided into two categories, namely physical models and behavioral models. Traditional physical models assume that the motion trend of an object remains unchanged over a short period of time; these models include the constant acceleration model, extended Kalman filter (EKF), etc. [1,2]. Behavioral models are generally based on statistical theory, using the Gaussian mixture model [3] and hidden Markov processes [4] to estimate a driver's intended behavior and a

vehicle's motion trends. Such methods are suitable for some simple scenarios, but have poor long-term prediction performance when facing a complex environment.

Thanks to the rapid development of deep learning in recent years, it now achieves excellent performance in the fields of object detection, pattern recognition, and behavioral prediction owing to its strong feature extraction capabilities. For vehicle trajectory prediction, since changes in the motion of vehicles or humans are continuous and reflect significant temporal coupling characteristics, related networks such as recurrent neural networks (RNNs) [5] and long short-term memory (LSTM) [6–9] are commonly adopted to extract the temporal features of target trajectories. At the same time, with the continuous enrichment of the autonomous driving dataset [10,11], its scene coverage has been gradually improved, which also provides a strong support for model training. Furthermore, benefiting nowadays from the decreasing cost of sensors and the gradual popularization of V2X technology, autonomous vehicles can obtain richer environmental information. In order to characterize potential spatio-temporal interactions in complex environments, encoding–decoding frameworks based on a variety of aggregation modes have been widely adopted, including the generative adversarial network (GAN) [12], graphic neural network (GNN) [13], and convolutional neural network (CNN) [14]. However, the vast number of spatio-temporal interactions contain a lot of redundant information. To more effectively extract important features from a large amount of spatio-temporal information, attention mechanisms based on encoding–decoding frameworks have become mainstream approaches in recent years.

To improve predictive accuracy and interpretability, an encoding–decoding framework that incorporates spatio-temporal information, allowing for a more comprehensive extraction of potential interaction relationships between the target vehicle and surrounding vehicles, is proposed in this paper. As shown in Figure 1, our approach contains three modules. The vehicle–vehicle interaction module embeds vehicle motion states via a graph message and extracts spatial interaction features using a graph attention mechanism (GAT). The historical trajectory encoder adopts BiLSTM to extract temporal trajectory features and combines them with vehicle–vehicle interaction coefficients to carry out spatio-temporal information aggregation. The decoder module finally generates the predicted trajectory by decoding the hidden variables.

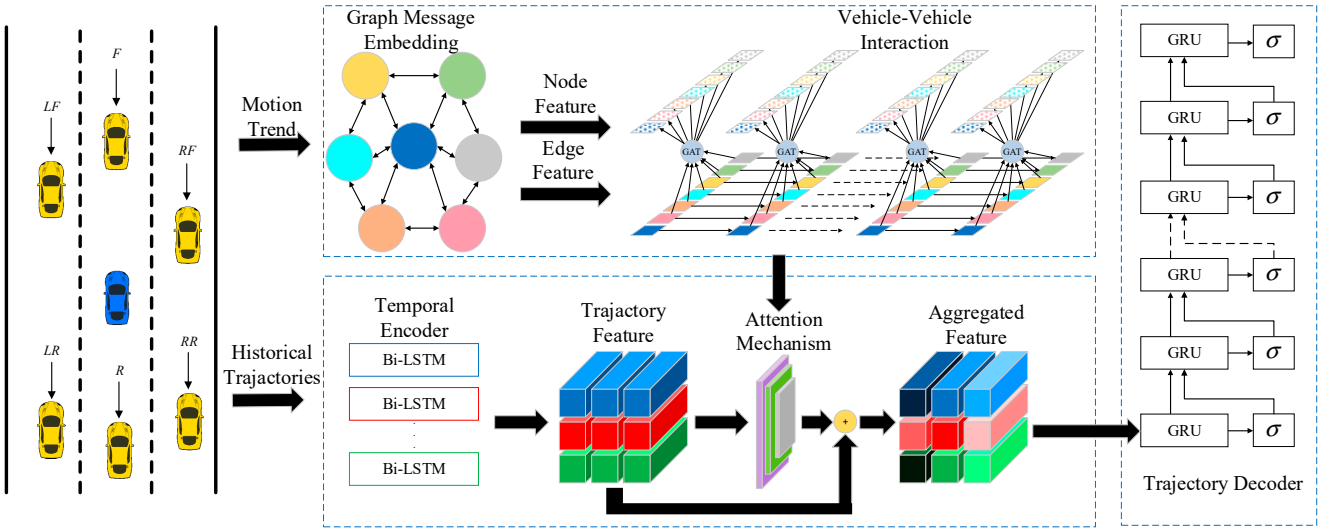



**Figure 1.** The encoding–decoding framework for vehicle trajectory prediction.

The remainder of the paper proceeds as follows. Section 2 is the literature review of the related work. Section 3 gives the description of the vehicle trajectory problem and the definition of the basic scenario parameters. Then, the encoding–decoding trajectory prediction model considering vehicle–vehicle interaction is outlined in Section 4. The training

process is presented and the prediction performance analyzed in Section 5. Section 6 draws the conclusion for the whole paper.

## 2. Related Work

Conventional trajectory prediction methods mainly rely on kinematics and dynamics models. The constant velocity (CV), constant acceleration (CA), and their combinations were adopted in the early years of research [15]. Barrios et al. employed the Kalman filter for vehicle trajectory prediction [16]. Considering model nonlinearity, Schubert adopted the unscented Kalman filter in a constant-turn-rate model for better prediction performance [17]. Other scholars attempted to analyze vehicle behavioral characteristics from historical motion states. Inspired by the human driver's visual system, Xia et al. implemented a mixed Gaussian model combined with a hidden Markov model to derive earlier lane-changing intentions of drivers by judging based on sudden changes in the speed of the target vehicle [18]. Li et al. exploited multiple predictive features including historical states of vehicles and the road structures, which were entered into a dynamic Bayesian network to infer the probability of each maneuver of the target vehicle [19]. Such methods are simple to compute. However, they mainly rely on the current motion information of the target vehicle. Model uncertainties and changes in the driver behavior are not covered. Furthermore, environmental interactions are also not considered. So, a physical model is only suitable for short-term prediction. Although some other approaches [20–22] integrated physical and behavioral models through multi-model interaction to enhance the prediction accuracy, this still has insufficient reliability for long-term prediction, especially in complex traffic environments.

The essence of vehicle trajectory prediction is the regression of time-series data. LSTM has been subject to significant advances in sequence generation, meaning it has been widely adopted for natural language processing, target tracking, and trajectory prediction in recent years. Alahi proposed an LSTM-based trajectory aggregating method in which they synthesized all the information through a pooling operation and applied a decoder to generate predicted trajectories [7]. Other forms of LSTM incorporated with different networks have also been utilized to better represent multi-object interactions. Gao incorporated the graph representation learning module into an LSTM encoder–decoder model in order for it to precisely learn the spatial interactions between vehicles [8]. Sheng processed spatio-temporal trajectory information through a combination of a GNN and CNN [14]. Li applied graph convolutional blocks to represent the interactions of close objects [23]. Xu et al. established a multi-scale heterogeneous network for varying numbers of moving objects, and the prediction of targets was realized by decoding the hidden features [24]. Ce et al. proposed an interaction-aware Kalman neural network (IaKNN)-based multi-layer architecture to resolve a high-dimensional traffic environment [25]. To capture the importance of spatio-temporal coupling information more accurately and further enhance the prediction performance, Messaoud applied a multi-head attention mechanism that encodes the vehicle dynamics and category information, and they formulated a multimodal trajectory prediction framework [26]. Mo presented a multi-vehicle trajectory prediction based on establishing a heterogeneous edge-enhanced graph attention network describing the multi-vehicle interaction mechanism and introducing a gate-based multi-objective selective map sharing mechanism [13]. To extract feature information more efficiently, Li introduced reinforcement learning into the graph attention mechanism in order to determine the relative significance of each node in continuous training interactions [27]. Some other forms of attention mechanisms from different perspectives have also been adopted in similar studies, which effectively improved the prediction performance [28–30]. Although the attention mechanism has good performance in selecting the important spatio-temporal information, most of the existing studies have adopted the generalized attention mechanism for vehicle trajectory prediction, which directly concatenates or produces a weighted sum of the node features to obtain the attention coefficients, while failing to introduce evaluation indexes

that characterize the interaction risk in the actual driving scenario, which results in poor interpretability of the prediction model.

To tackle the above limitations, an attention-based vehicle trajectory prediction model considering multi-vehicle interaction is proposed in this paper. The main contributions are as follows.

The multi-vehicle motion features are extracted via graph messaging embedding and the GAT is employed to describe the vehicle–vehicle spatial interaction relationships. In addition, these features are concatenated with temporal trajectory features via a self-attention mechanism to achieve spatio-temporal information aggregation.

The potential field model is introduced as the prior information for the edge feature when obtaining the vehicle–vehicle interaction feature, which differs from the traditional model, calculating the edge weights by directly concatenating the node information.

## 3. Problem Description and Basic Definition

It is believed that in a complex traffic environment, the motion states of surrounding vehicles have a significant impact on the future motion trends of the target vehicle. The vehicles in the current and adjacent lanes that are closest to the target vehicle in front and behind are defined as the surrounding vehicles. As shown in the left of Figure 1, the blue car is the target vehicle $V_o$ whose motion needs to be predicted. Yellow cars are the surrounding vehicles, expressed as $V_i, i \in \mathcal{N}(V_o) = [F, R, LF, LR, RF, RR]$, where $\mathcal{N}(V_o)$ represents the set of the surrounding vehicles. The encoder proposed in this paper consists of two components, the vehicle–vehicle interaction module and trajectory encoding module. For the vehicle–vehicle interaction module, the motion states of vehicles are primarily focused. We define the motion states of the object vehicle and surrounding vehicles at time $t$ as $s_o^t = \left[u_o^t, v_o^t, a_o^t\right]$ and $s_i^t = \left[u_i^t, v_i^t, a_i^t\right]$, where $u_o^t, v_o^t, a_o^t$ are the longitudinal velocity, lateral velocity, and longitudinal acceleration of the target vehicle and $u_i^t, v_i^t, a_i^t$ are the longitudinal velocity, lateral velocity, and longitudinal acceleration of the $i$th surrounding vehicle, respectively. Similarly, for the trajectory encoding module, we define the trajectories of the target vehicle and surrounding vehicles at time $t$ as $\tau_o^t = \left[x_o^t, y_o^t\right]$ and $\tau_i^t = \left[x_i^t, y_i^t\right]$, where $x_o^t, y_o^t, x_i^t, y_i^t$ represent the longitudinal and lateral positions of the target vehicle and $i$th surrounding vehicle. Based on these, the historical observation can be formulated as

$$\begin{aligned} S_o^t &= [s_o^{t-t_{obs}}, s_o^{t-t_{obs}+1}, \cdots, s_o^{t-1}, s_o^t] \\ S_i^t &= [s_i^{t-t_{obs}}, s_i^{t-t_{obs}+1}, \cdots, s_i^{t-1}, s_i^t] \\ \Gamma_o^t &= [\tau_o^{t-t_{obs}}, \tau_o^{t-t_{obs}+1}, \cdots, \tau_o^{t-1}, \tau_o^t] \\ \Gamma_i^t &= [\tau_i^{t-t_{obs}}, \tau_i^{t-t_{obs}+1}, \cdots, \tau_i^{t-1}, \tau_i^t] \end{aligned} \tag{1}$$

where $t_{obs}$ is the observation horizon.

Note that the trajectories and motion states of all vehicles need to be unified in the same coordinate. Here, we define the projection position on the centerline of the lane where the object vehicle is located at time $t$ as the origin point. The $x$-axis corresponds to the tangent of that point on the lane's centerline, and the $y$-axis is perpendicular to the $x$-axis. The coordinate system is shown in Figure 2.

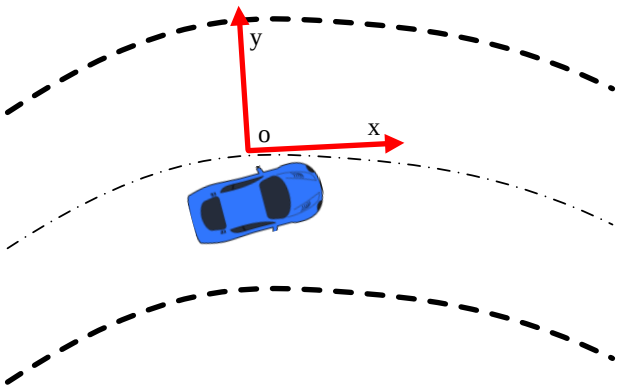



**Figure 2.** The coordinate system for trajectory prediction.

## 4. The Vehicle Trajectory Prediction Model

To effectively aggregate the interaction features between the target vehicle and the surrounding environment and thereby achieve more reliable trajectory prediction, an attention-based encoding–decoding framework for trajectory prediction is proposed. As illustrated in Figure 1, the whole framework can be divided into the vehicle–vehicle interaction module, trajectory encoding module, and decoding module.

### *4.1. Vehicle–Vehicle Interaction Module*

When vehicles move on highways or urban roads, there are significant interactions between a target vehicle and surrounding obstacles, especially moving targets. Such interactions are crucial for determining the future trajectory of the target vehicle. By observing the historical motion states of traffic participants and extracting their temporal features, we may obtain relationship coefficients between the target vehicle and the surrounding vehicles via a graph attention mechanism.

A mixed traffic scenario consisting of multiple vehicles can be regarded as a multi-agent system expressed as the directed graph $\mathcal{G} = (\mathcal{V}, E)$, where $\mathcal{V} = (V_o, V_i), i \in \mathcal{N}(V_o)$ represents the nodes consisting of vehicles and $E \subset V \times V$ is the set of edges, denoting the interactions between nodes. Note that the edge information only exists for neighboring nodes, indicating that the motion of the vehicle is affected by the other vehicles nearby. Firstly, based on the motion states of the object vehicle and surrounding vehicles, the node information is encoded at each sampling time $t$ as

$$\begin{aligned} m_o^t &= \sigma\left(W_v^o s_o^t\right) \\ m_i^t &= \sigma\left(W_v^i s_i^t\right) \end{aligned} \tag{2}$$

where $\sigma(\bullet)$ is the nonlinear activation function, and $W_v^o$ and $W_v^i$ are the weight matrixes for the object vehicle and surrounding vehicles, respectively. Considering that the trajectory and motion states of the vehicle are continuously time-varying parameters, the interactions between the motion states of all surrounding vehicles and the target vehicle also show temporal variations. Therefore, the LSTM is adopted here to process the historical motion states of the target vehicle and the surrounding vehicles, to obtain their potential temporal features, which are represented as temporal motion feature vectors in the following equations:

$$\begin{aligned} h_o^t &= \mathrm{LSTM}\left(h_o^{t-1}, m_o^t; W_h\right) \\ h_i^t &= \mathrm{LSTM}\left(h_i^{t-1}, m_i^t; W_h\right) \end{aligned} \tag{3}$$

where $h_o^t$ and $h_i^t$ are the hidden feature vectors of the target and surrounding vehicles at time $t$, while $W_h$ is the LSTM encoder weight. Although different vehicles use separate LSTM networks, they share the same network parameters.

Based on the directed graph containing nodes and edges, the interaction features can be extracted by aggregating the node and edge information, which effectively indicates the vehicle–vehicle motion interaction mechanism. Here, the GAT is adopted to execute aggregation operations on neighboring nodes for adaptive matching of different node weights. The basic principle of GAT is shown in Figure 3.

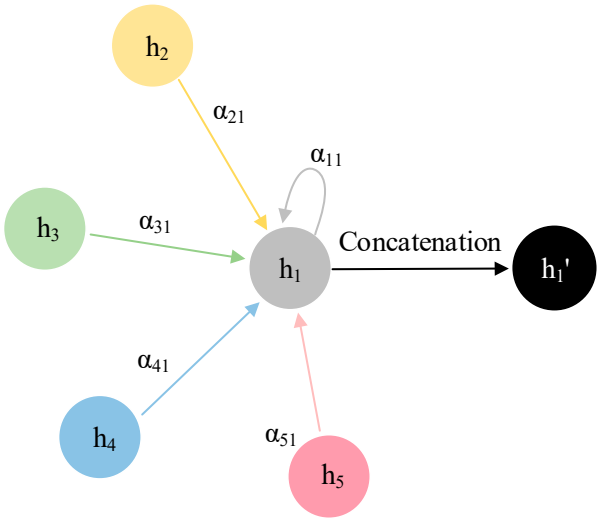



**Figure 3.** The principle of the graph attention mechanism.

The general GAT architecture directly concatenates the node information to obtain the spatial weight coefficients. However, the node feature only contains the vehicles' own motion information, and it is difficult to represent the relative vehicle motion relationships via direct concatenation of each node. Moreover, it is impossible to explicitly express the physical collision constraints between vehicles. In fact, there are a number of well-established evaluation metrics that can effectively quantify the vehicle collision risk and provide a reliable basis for vehicle decision-making systems. For example, the time to collision (TTC) and time head way (THW) are widely adopted in adaptive cruise control (ACC) and autonomous emergency braking (AEB) for longitudinal collision risk assessment, which are also incorporated into the vehicle motion features to establish the cognitive model of driver behavior [31].

In complex environments, vehicle motion trends are influenced by a combination of multidimensional factors. The artificial potential field method is widely used to establish multi-agents' interactions, which represents the combined effect of obstacles in the environment on the target, as determined by adding the attractive and repulsive fields together. As for vehicle–vehicle interaction, the repulsive field is the most suitable to characterize the effect of surrounding vehicles on the target vehicle, which generates the repulsive force need for the target vehicle to adjust its future motion trends in order to avoid potential collision risks and ensure a safe driving space [32–34]. Thus, we adopt the repulsive potential field as the priori information about the edge features. The repulsive force between vehicles at time $t$ can be presented as

$$\begin{aligned} F_{ij}^t &= c_{ij} e^{-(c_{xij}^t x_{ij}^{t\,2} + c_{yij}^t y_{ij}^{t\,2})} \\ c_{xij}^t &= \frac{1}{u_{ij}^t + p} \\ c_{yij}^t &= \frac{1}{W_j} \end{aligned} \tag{4}$$

where $F_{ij}^t$ is the repulsive force on the $i$th vehicle generated by the $j$th vehicle. $c_{ij}$ is the scaling factor determined by the vehicle type. $x_{ij}^t$,$y_{ij}^t$ denote the relative longitudinal and

lateral distances between the $i$th and the $j$th vehicles. $u_{ij}^t$ is the relative longitudinal velocity. Note that compared with longitudinal collision, lateral collision is more dangerous. Even if the collision speed is low, it is still prone to causing dangerous conditions such as vehicle side slip and spinning. Hence, we introduce $c_{xij}^t$ and $c_{yij}^t$ as the shape factors defining the vehicle collision risk at different scales for the longitudinal and lateral directions. $p$ is a very small positive number and $W_j$ is the width factor for the $j$th vehicle. For the partial derivatives of $F_{ij}^t$ with respect to $x_{ij}^t$ and $y_{ij}^t$, the longitudinal and lateral repulsive force gradients of the vehicle can be obtained as

$$\begin{aligned} g_{xij}^t &= \frac{\partial F_{ij}^t}{\partial x_{ij}^t} \\ g_{yij}^t &= \frac{\partial F_{ij}^t}{\partial y_{ij}^t} \end{aligned} \tag{5}$$

After summing up the repulsive force gradients generated by the surrounding vehicles, the total gradient of the $i$th vehicle is

$$\begin{aligned} g_{xi}^t &= \sum g_{xij}^t, j \in \mathcal{N}(V_i) \\ g_{yi}^t &= \sum g_{yij}^t, j \in \mathcal{N}(V_i) \end{aligned} \tag{6}$$

When applying an artificial potential field, we regard the longitudinal and lateral repulsive force gradients as the priori edge information and concatenate that with vehicles' own motion feature parameters to form new node features, as

$$\begin{aligned} \widehat{h}_o^t &= \left( h_o^t \middle\| g_{xo}^t \middle\| g_{yo}^t \right), j \in \mathcal{N}(V_o) \\ \widehat{h}_i^t &= \left( h_i^t \middle\| g_{xi}^t \middle\| g_{yi}^t \right), j \in \mathcal{N}(V_i) \end{aligned} \tag{7}$$

where $\|$ defines the concatenation operation. Then, the attention coefficients between individual nodes can be calculated as

$$\alpha_{ij}^t = \frac{\exp\left( \text{LeakyReLU}\left( a^T \left[ W\widehat{h}_i^t \middle\| W\widehat{h}_j^t \right] \right) \right)}{\sum_{k \in \mathcal{N}(i)} \exp\left( \text{LeakyReLU}\left( a^T \left[ W\widehat{h}_i^t \middle\| W\widehat{h}_k^t \right] \right) \right)} \tag{8}$$

where $\alpha_{ij}^t$ is the attention coefficient of node $j$ to node $i$, $W$ is the weight matrix of the trainable linear transformation for each node, $a$ is the weight vector of the single-layer feed-forward neural network, and LeakyReLU($\bullet$) is the LeakyReLU activation function. Subsequently, the weighted sum of node features is derived based on the graph attention coefficients. The output of GAT can be reformulated as the new node feature

$$\hat{h}_i^t = \sigma\left( \sum_{j \in \mathcal{N}(i)} \alpha_{ij}^t W\widehat{h}_j^t \right) \tag{9}$$

Equation (9) illustrates the vehicle motion features containing spatial vehicle–vehicle interactions. Due to the continuity of vehicle motion, such features at different sampling times are still temporally coupled. Thus, another LSTM network is utilized here to implement spatio-temporal feature fusion, denoted as

$$\begin{aligned} p_o^t &= \text{LSTM}\left( p_o^{t-1}, \hat{h}_o^t; W_g \right) \\ p_i^t &= \text{LSTM}\left( p_i^{t-1}, \hat{h}_i^t; W_g \right) \end{aligned} \tag{10}$$

Similar to Equation (3), each LSTM network shares the same learnable weight matrix $W_g$. Finally, the multilayer perceptron (MLP) is used to aggregate the spatio-temporal fusion features over the observation horizon, and the vehicle interaction fusion coefficients can be obtained:

$$\begin{aligned} \overline{p}_o^t &= \mathrm{MLP}\left(p_o^{t-t_{obt}}, p_o^{t-t_{obt}+1}, \cdots, p_o^{t-1}, p_o^t\right) \\ \overline{p}_i^t &= \mathrm{MLP}\left(p_i^{t-t_{obt}}, p_i^{t-t_{obt}+1}, \cdots, p_i^{t-1}, p_i^t\right) \end{aligned} \tag{11}$$

### *4.2. Vehicle Historical Trajectory Encoder*

In addition to the vehicle–vehicle motion interaction, the vehicle historical trajectory is also very important and should not be ignored. Particularly when there are only a few traffic participants around, its historical trajectory can reflect the drivers' behavioral characteristics to a certain extent, which can provide a valuable reference for future trajectory prediction. Therefore, the historical trajectories of the target vehicle and the surrounding vehicles are encoded here first.

Considering its spatio-temporal coupling properties, we used the BiLSTM network to process the vehicle historical trajectory information. Different from the general LSTM network stated as Equation (3), the BiLSTM network contains two layers of LSTM networks. By adding an inverse LSTM layer to the LSTM network, BiLSTM can further deal with the information from the future. The bi-directional characteristic of BiLSTM cannot only better tackle the long-term dependency problem but also improves the prediction accuracy due to the increased number of networks.

As shown in Figure 4, BiLSTM can be formulated as

$$\begin{aligned} \left(h_{fi}^t, c_f^t\right) &= \mathrm{LSTM}\left(h_{fi}^{t-1}, \tau_i^t; c_f^{t-1}\right) \\ \left(h_{bi}^{t-t_{obs}}, c_b^{t-t_{obs}}\right) &= \mathrm{LSTM}\left(h_{bi}^{t-t_{obs}+1}, \tau_i^{t-t_{obs}}; c_b^{t-t_{obs}+1}\right) \end{aligned} \tag{12}$$

where $c_f$ and $c_b$ denote the feedforward and backward network parameters. Considering that the deep BiLSTM architecture is prone to lead optimization bottleneck and gradient disappearance, a deep residual network is used to connect the inputs and outputs of the BiLSTM layers. We sum the hidden variable with the input sequence and concatenate them for forward and backward networks. We can then obtain the output of the BiLSTM network:

$$h_{pi}^t = \left(h_{fi}^t + \tau_i^t \middle\| h_{bi}^{t-t_{obs}} + \tau_i^{t-t_{obs}}\right) \tag{13}$$

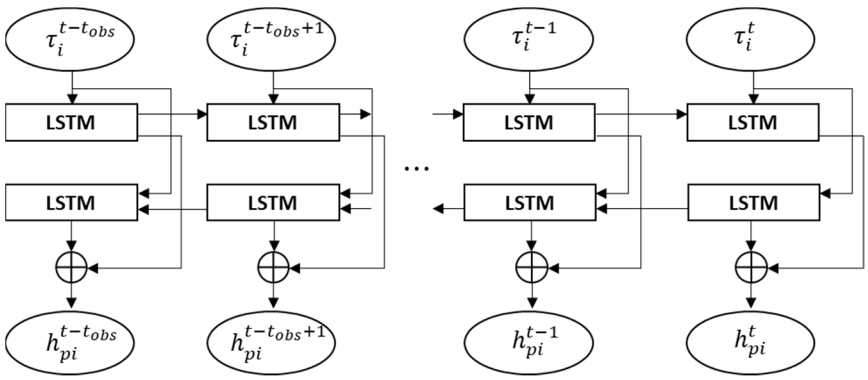



**Figure 4.** The diagram of BiLSTM.

After encoding the trajectory sequence, a large number of historical trajectory feature vectors can be obtained. Subsequently, based on Equations (11) and (13), we carry out

weighted aggregation of the trajectory feature with the vehicle–vehicle interaction feature to generate a new feature vector $r_i^t$ as

$$r_i^t = \tan h\left(W_p h_{pi}^t + W_v \overline{p}_i^t\right) \tag{14}$$

where $W_p$ and $W_v$ are weight matrixes.

Then, the dual-attention mechanism is adopted here for key information extraction. The spatial attention mechanism is mainly used to handle the influences of different surrounding vehicles on the target vehicle. For the vehicles' trajectories, the interaction is closely related to the relative distance. The closer obstacles always have more significant influences on the target vehicle. Therefore, the cosine distance is used here to calculate the correlation coefficients of encoded vectors between vehicles.

$$f(r_o, r_i) = 1 - \frac{r_0 \cdot r_i}{\|r_o\|_2 \cdot \|r_i\|_2} \tag{15}$$

To ensure that the weight scales are the same, we normalize $f(r_o, r_i)$ to obtain the spatial attention coefficients.

$$s_i = \frac{f(r_o, r_i)}{\sum\limits_{i \in \mathcal{N}(V_o)} f(r_o, r_i)} \tag{16}$$

By multiplying the hidden feature vectors and attention coefficients, the context vectors including spatial correlation can be obtained as

$$\overline{r}_{so} = \sum_{i \in \mathcal{N}(V_o)} s_i r_i \tag{17}$$

To determine the importance of the historical information at each sampling time, the temporal attention mechanism is applied here. Due to the nonlinearity of the vehicle model and the unpredictability of the driver behavioral characteristics, the relevance of the temporal features cannot be characterized using quantitative measurements such as the spatial attention mechanism. Thus, the inter-relationships of historical trajectory features are described by the linear transformation as

$$g\left(r_o^t, r_o^k\right) = \tan h\left(r_o^i W_t r_o^k\right), k \in [t - t_{obs}, t - t_{obs} + 1, \ldots, t - 1] \tag{18}$$

where $W_t$ is the trainable weight. Then, we adopt the softmax function to normalize the temporal attention distribution

$$\alpha_k = \frac{\exp\left(g\left(r_o^t, r_o^k\right)\right)}{\sum\limits_k \exp\left(g\left(r_o^t, r_o^k\right)\right)} \tag{19}$$

Similar to Equation (17), the context vectors including temporal correlation can be obtained as

$$\overline{r}_{to} = \sum_k \alpha_k r_o^k \tag{20}$$

Finally, we utilize a convolution pooling layer to integrate the encoded trajectory feature, spatial attention, and temporal attention of the object vehicle, to derive the overall feature

$$\tilde{h}_o^t = W_1 r_o^t + W_2 \overline{r}_{so} + W_3 \overline{r}_{to} \tag{21}$$

### *4.3. Future Trajectory Prediction Decoder*

Based on the aggregated feature vectors, the decoder predicts the future trajectory of the vehicle. Considering that the training process involves learning from real-world

trajectory datasets, and in order to generate diverse samples, we add noise to the aggregated features to establish the initial hidden states of the decoder

$$n_o^t = \tilde{h}_o^t \| z \tag{22}$$

where $z$ is the noise satisfying $z \sim N(0,1)$. Then, the GRU network is applied here for trajectory decoding. At time $t$, we input the initial hidden state to the GRU obtain the hidden state at the next moment and calculate the predicted trajectory via another nonlinear activation function:

$$n_o^{t+1} = \mathrm{GRU}\left(n_o^t, \tau_o^t; W_g\right) \tag{23}$$

$$\hat{\tau}_o^{t+1} = \sigma\left(n_o^{t+1}\right) \tag{24}$$

Subsequently, we re-input the predicted motion states obtained at each moment into the GRU for iteration, to obtain the target vehicle trajectory in the predicted horizon:

$$n_o^i = \mathrm{GRU}\left(n_o^{i-1}, \hat{\tau}_o^t; W_g\right), i \in \left[2 : t_{pre}\right] \tag{25}$$

$$\hat{\tau}_o^i = \sigma\left(n_o^i\right), i \in \left[2 : t_{pre}\right] \tag{26}$$

where $t_{pre}$ is the prediction horizon.

## 5. Experiment and Analyses

### *5.1. Training Details*

The next-generation simulation (NGSIM) dataset, comprising vehicle trajectory data from the US101 and I-180 highways collected through a network of synchronized digital cameras [10], was adopted for model training in this study. The vehicle trajectory data records are provided for every one-tenth of a second. This dataset contains a large number of vehicle-following and lane-changing scenarios in dense traffic, which effectively reflects interaction between vehicles and other traffic participants. For the prediction model proposed in this paper, the vehicle historical data of the past 3 s is applied to predict the vehicle trajectory for the next 5 s.

To quantitatively illustrate the accuracy and reliability of the prediction model proposed in this paper, the root mean squared error (RMSE) metric between the predicted position and ground truth trajectory position within the prediction range for the prediction horizon is chosen as the evaluation index used for evaluation

$$L_{\mathrm{RMSE}} = \sqrt{\frac{1}{t_{pre}} \sum_{1}^{t_{pre}} \left\| \tau_g^t - \hat{\tau}_o^t \right\|^2} \tag{27}$$

where $\tau_g^t$ means the position of the ground truth trajectory at time $t$, while $\hat{\tau}_o^t$ denotes the position generated by the prediction model at time $t$.

### *5.2. Quantitative Comparison with Baselines*

To reveal the advantage of the proposed model, some benchmark models are listed here for quantitative comparison.

CV: single-vehicle trajectory prediction using a Kalman filter at a constant velocity [35].

IA-KNN: a multi-layer architecture of interaction-aware Kalman neural networks (IaKNNs), which involves an interaction layer for resolving high-dimensional traffic environmental observations and a filter layer for future trajectories' estimation [26].

S-LSTM: sharing the information between multiple LSTMs to capture the interactions within the neighborhood corresponding to the neighboring trajectories [11].

GR-LSTM: adopting an LSTM to handle the temporal sequence and a graph representation learning module to precisely represent the spatial interaction between vehicles [27].

S-GAN: using a recurrent sequence-to-sequence model to observe the motion histories and aggregate the information via a pooling mechanism. The plausible future trajectories are predicted through adversarial training against a recurrent discriminator [28].

GRIP++: applying a graph to represent the interactions of neighboring objects with several graph convolutional blocks to extract features, while an encoder–decoder LSTM model is designed to make the predictions [29].

DAM: a dual attention mechanism is introduced for trajectory prediction by analyzing the influence between the neighboring vehicle and target vehicle, which is combined with the temporal hidden trajectory feature of the target vehicle to reduce the uncertainty of the potential trajectory [17].

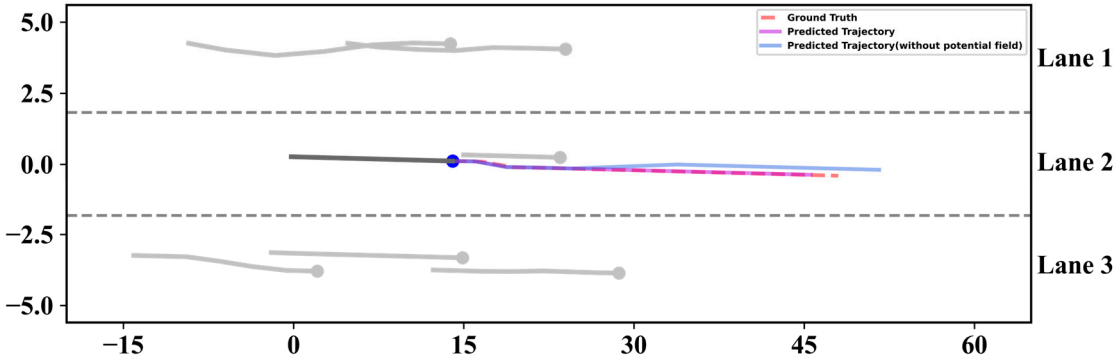
The experiment results and the comparisons between benchmarks are listed in Table 1. The CV model has the worst prediction performance when compared with other algorithms both for prediction error and its growth rate, which illustrates that vehicle nonlinear characteristics, different driver behaviors, and environmental changes have a significant effect on the future trajectory changes of vehicles. The rest of the learning-based prediction algorithms enhanced the prediction performance by extracting hidden features from the dataset. Note that the original S-LSTM was used for trajectory prediction of pedestrians. Although it analyzed the relative interactions of different objects, it did not have satisfactory performance for vehicle trajectory prediction, illustrating the significant differences between pedestrian–pedestrian and vehicle–vehicle interaction. GR-LSTM, S-GAN, and DAM used different methods to further explore the potential influence mechanism of spatio-temporal features on the future trajectories of vehicles, all of which improved the prediction accuracy. GRIP++ has the best prediction accuracy among the existing algorithms, revealing the advantage of a graph for the representation of vehicle–vehicle interactions. Compared with the aforementioned methods, our model not only represents the correlation of historical positions of the object vehicle and the surrounding vehicles through the spatio-temporal attention mechanism, but the relative dynamic motion interaction features between vehicles are also extracted through GAT combined with the artificial potential field. Thus, it achieves the lowest average RMSE of 1.45 m for the prediction trajectory, which improves the best baseline (GRIP++) by 5.9%. For the entire prediction horizon, our method also has the better performance at all prediction moments except for at 1 s. It is not all that worse than the best method, and is acceptable for a vehicle decision system.

**Table 1.** Comparison of the prediction performances of different models.

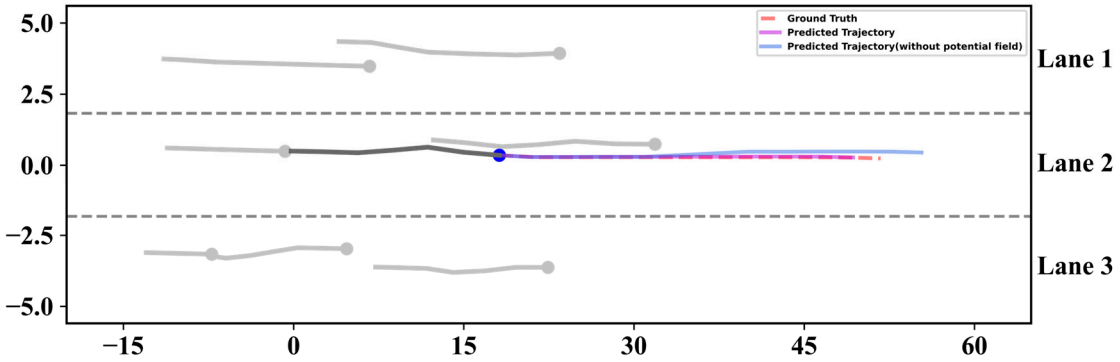
| Prediction Horizons (s) | 1 | 2 | 3 | 4 | 5 | Average |
|---|---|---|---|---|---|---|
| CV | 0.7 | 1.78 | 3.13 | 4.78 | 6.68 | 3.42 |
| IA-KNN | 0.62 | 1.03 | 1.97 | 2.93 | 4.12 | 2.13 |
| S-LSTM | 0.65 | 1.31 | 2.16 | 3.25 | 4.55 | 2.81 |
| GR-LSTM | 0.68 | 1.17 | 1.74 | 2.64 | 3.32 | 1.91 |
| S-GAN | 0.57 | 1.32 | 2.22 | 3.26 | 4.4 | 2.35 |
| GRIP++ | 0.38 | 0.89 | 1.45 | 2.14 | 2.94 | 1.56 |
| DAM | 0.5 | 1.11 | 1.78 | 2.69 | 3.93 | 2.0 |
| Ours | 0.42 | 0.79 | 1.32 | 2.03 | 2.64 | 1.45 |
| Comparison | +10.5% | −11% | −9% | −5.1% | −10.2% | −5.9% |

### *5.3. Prediction Performances under Different Scenarios*

Vehicle lane-keeping and lane-changing maneuvers are the two most common scenarios in highway driving. To illustrate the prediction performance and demonstrate the advantage of applying the potential field method to describe a priori information of vehicle–vehicle interactions, we introduced another prediction model for comparison, which has the same network architecture as the proposed model in this paper but without the potential field module. Comparisons of the prediction performances for these two models under these two key scenarios are shown in Figures 5 and 6.

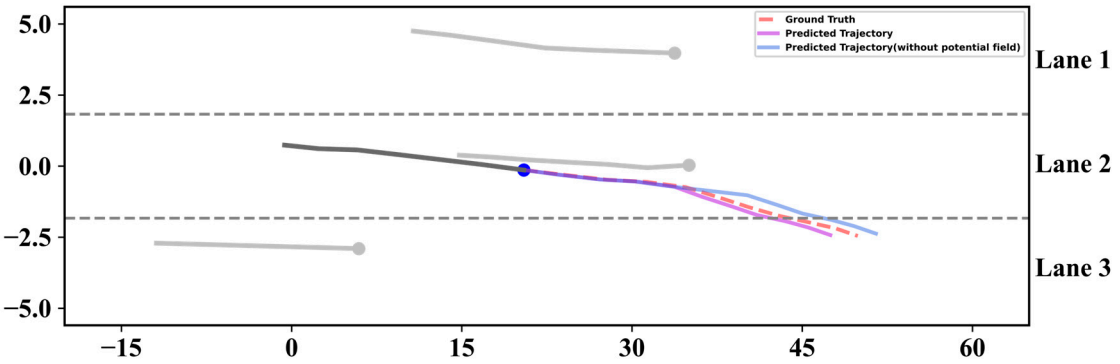



(**a**) Lane-keeping scenario with six surrounding vehicles

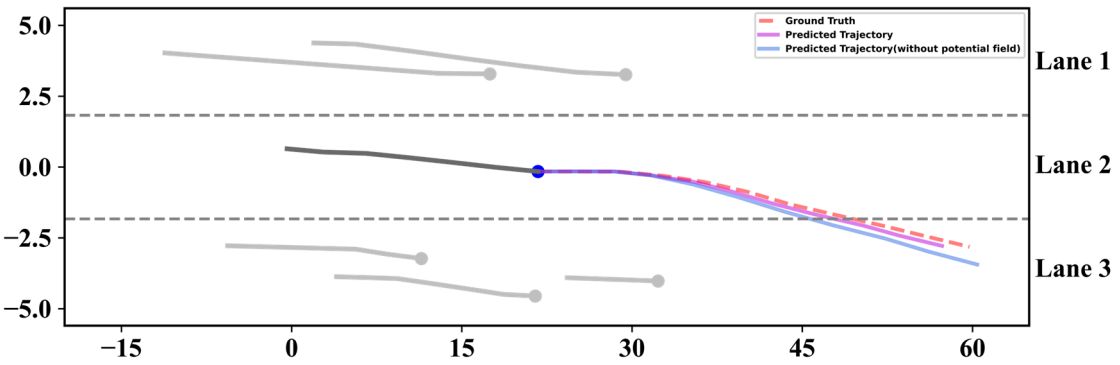



(**b**) Lane-keeping scenario with seven surrounding vehicles

**Figure 5.** Trajectory prediction for lane-keeping maneuver.




(**a**) Lane-changing scenario with three surrounding vehicles




(**b**) Lane-changing maneuver with five surrounding vehicles

**Figure 6.** Trajectory prediction for lane-changing maneuver.

Figure 5 illustrates the prediction results for the lane-keeping maneuver. The gray and blue dots denote the surrounding vehicles and the target vehicle, respectively. The gray solid line is the historical trajectory of the target and surrounding vehicles. The pink dashed line is the future ground truth trajectory. The purple solid line is the predicted trajectory for the proposed model in this paper and the blue solid line is the predicted trajectory generated by the prediction model without the potential field module. Figure 5a is a congested traffic scenario. The vehicles are accelerating at a low initial speed. Note that there are six moving obstacles around the target vehicle. Although the speed of the front vehicle is slower and blocks the target vehicle, there are surrounding vehicles in both the left and right adjacent lanes, which move side-by-side with the target vehicle within the observation horizon. The free space for lane changing is insufficient. Therefore, both prediction models judge that the vehicle will sustain lane keeping. The model without a potential field module is not sensitive enough to predict the change in the vehicle speed, which leads to a relatively large final distance error (FDE). In contrast, our approach analyzes the collision threat of the front vehicle to the target vehicle more accurately, enabling a better FDE performance. In Figure 5b, the traffic flow is still congested but vehicles are moving steadily. During the observation horizon, no obvious acceleration, deceleration, or lane-changing maneuver occurs, meaning the safety of the object vehicle is not threatened. So, the vehicle trajectory prediction results are continual lane keeping for both prediction models. Similar to Figure 5a, the proposed algorithm predicts the future speed of the vehicle more accurately and the predicted trajectory is closer to the ground truth.

In Figure 6a, there are three obstacle vehicles around the object vehicle. Due to its slower speed, the target vehicle is gradually approaching the leading vehicle. But at this time, the right lane is relatively empty in front, with no obstacles. And the historical trajectory of the target vehicle has a tendency of moving to the right. So, it is predicted that the vehicle will execute a lane change action. It is worth noting that due to the insufficient extraction of vehicle–vehicle interaction features for the model without a potential field, it judges that the vehicle will first continue to move according to the historical motion trend, and implement the lane change operation later when it is too close to the preceding vehicle. In comparison, the proposed model predicts the lane change trend of the target vehicle earlier under the effect of the repulsive force, resulting in smoother prediction trajectories and lower prediction errors. In Figure 6b, there is no obstacle in the current lane of the target vehicle, and the other vehicles in the adjacent lanes have no lane-changing tendencies, meaning they have no obvious impact on the target vehicle. The prediction of the future trajectory for the target vehicle at this time relies more on its historical motion state. When analyzing its historical trajectory, it is predicted that the target vehicle will change lanes to the right. Note that the predicted trajectory is not directly offset to the right lane, but keeps going straight for a period of time and then moves to the right. Particularly for the proposed model in this paper, by precisely considering the potential threat of the obstacle vehicle in front and to the right of the object vehicle at the predicted moment, it generates a more conservative lane change trajectory that leaves a larger safety distance than the model without a potential field module. It can be seen that both for the vehicle lane-keeping and lane-changing scenarios, the predicted trajectories and the real trajectories maintain a high degree of consistency. Therefore, it seems that by using the prediction model proposed in this paper, autonomous vehicles can accurately predict the future trajectories of surrounding vehicles in front or in adjacent lanes by capturing their historical motion information, so as to make the best corresponding decisions and thus ensure driving safety.

## 6. Conclusions

In this paper, an attention-based vehicle trajectory prediction architecture considering multi-vehicle interactions has been proposed. Firstly, the motion interaction between the target vehicle and surrounding vehicles is described as the directed graph and the GAT is

adopted to extract the interaction features. Unlike the tradition method, which calculates the edge attention weight by directly concatenating the node features, we introduce the repulsive force between vehicles as a priori edge information to better represent the potential collision risk. Then, the vehicle–vehicle interaction features are integrated into the embedded trajectory vectors and the spatio-temporal attention mechanism is applied to excavate the significance of hidden features, guiding the decoder to generate a plausible future trajectory. The NGSIM public dataset was chosen for training and evaluation of the proposed prediction model. The test results reveal that compared with the existing prediction models, our method can reduce the prediction error and maintain good robustness both for short and long prediction horizons. It is also demonstrated that even in a dense traffic environment, the proposed model can accurately predict the future trajectory of the objecti vehicle even when there are lane-keeping and lane-changing maneuvers, meaning it provides reliable information for decision-making systems of autonomous vehicles.

In future work, we will expand the network by exploring richer interaction features, that is, not only vehicle–vehicle interaction but also vehicle–pedestrian interaction in dynamic driving environments, enabling better scenario adaptation and robustness of the prediction model. In addition, map information has not yet been utilized for the prediction model. With the gradual expansion of high-precision maps, we will incorporate map data into the model to further improve its prediction accuracy.

**Author Contributions:** Conceptualization, J.Z. and Y.R.; methodology, Y.R.; software, Y.R. and J.L.; validation, J.Z. and Y.R.; formal analysis, K.W.; investigation, X.H.; resources, K.W.; data curation, X.H.; writing—original draft preparation, J.Z.; writing—review and editing, Y.R.; visualization, J.L.; supervision, Y.R.; project administration, K.W.; funding acquisition, J.Z. and K.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Science and Technology Innovation Key R&D Program of Chongqing, no. cstc2021jscx-cylhX0006; a National Key Research and Development Program of China Grant, no. 2022YFF0604900; and the Open Fund of the Chongqing Key Laboratory of Industry and Information of Automotive Active Safety Testing Technology, no. 22AKC03.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the commercial privacy.

**Conflicts of Interest:** Authors Jie Zeng, Kan Wang and Xiong Hu were employed by the company China Merchants Testing Vehicle Technology Research Institute Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relation-ships that could be construed as a potential conflict of interest.

## References

1. Qin, Z.; Chen, L.; Fan, J.; Xu, B.; Hu, M.; Chen, X. An improved real-time slip model identification method for autonomous tracked vehicles using forward trajectory prediction compensation. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–12. [CrossRef]
2. Ess, A.; Schindler, K.; Leibe, B.; Van Gool, L. Object detection and tracking for autonomous navigation in dynamic environments. *Int. J. Robot. Res.* **2010**, *29*, 1707–1725. [CrossRef]
3. Wiest, J.; Höffken, M.; Kreßel, U.; Dietmayer, K. Probabilistic trajectory prediction with Gaussian mixture models. In Proceedings of the 2012 IEEE Intelligent Vehicles Symposium, Madrid, Spain, 3–7 June 2012; pp. 141–146.
4. Qiao, S.; Shen, D.; Wang, X.; Han, N.; Zhu, W. A self-adaptive parameter selection trajectory prediction approach via hidden Markov models. *IEEE Trans Intell. Transp. Syst.* **2014**, *16*, 284–296. [CrossRef]
5. Qin, Y.; Song, D.; Chen, H.; Cheng, W.; Jiang, G.; Cottrell, G.W. A dual-stage attention-based recurrent neural network for time series prediction. *arXiv* **2017**, arXiv:02971.
6. Zhang, P.; Ouyang, W.; Zhang, P.; Xue, J.; Zheng, N. Sr-lstm: State refinement for lstm towards pedestrian trajectory prediction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 12085–12094.

7. Alahi, A.; Goel, K.; Ramanathan, V.; Robicquet, A.; Fei-Fei, L.; Savarese, S. Social lstm: Human trajectory prediction in crowded spaces. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 27–30 June 2016; pp. 961–971.
8. Gao, Z.; Sun, Z. Modeling spatio-temporal interactions for vehicle trajectory prediction based on graph representation learning. In Proceedings of the 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), Indianapolis, IN, USA, 19–22 September 2021; pp. 1334–1339.
9. Deo, N.; Trivedi, M.M. Multi-modal trajectory prediction of surrounding vehicles with maneuver based lstms. In Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV), Suzhou, China, 26–30 June 2018; pp. 1179–1184.
10. Punzo, V.; Borzacchiello, M.T.; Ciuffo, B. On the assessment of vehicle trajectory data accuracy and application to the Next Generation SIMulation (NGSIM) program data. *Transp. Res. Part C Emerg. Technol.* **2011**, *19*, 1243–1262. [CrossRef]
11. Caesar, H.; Bankiti, V.; Lang, A.H.; Vora, S.; Liong, V.E.; Xu, Q.; Krishnan, A.; Pan, Y.; Baldan, G.; Beijbom, O. Nuscenes: A multimodal dataset for autonomous driving. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 11621–11631.
12. Gupta, A.; Johnson, J.; Fei-Fei, L.; Savarese, S.; Alahi, A. Social gan: Socially acceptable trajectories with generative adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 2255–2264.
13. Mo, X.; Huang, Z.; Xing, Y.; Lv, C. Multi-agent trajectory prediction with heterogeneous edge-enhanced graph attention network. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 9554–9567. [CrossRef]
14. Sheng, Z.; Xu, Y.; Xue, S.; Li, D. Graph-based spatial-temporal convolutional network for vehicle trajectory prediction in autonomous driving. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 17654–17665. [CrossRef]
15. Xu, W.; Pan, J.; Wei, J.; Dolan, J.M. Motion planning under uncertainty for on-road autonomous driving. In Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA), Hong Kong, China, 31 May–5 June 2014; pp. 2507–2512.
16. Barrios, C.; Motai, Y.; Huston, D.J.I.T.o.I.E. Trajectory estimations using smartphones. *IEEE Trans. Ind. Electron.* **2015**, *62*, 7901–7910. [CrossRef]
17. Schubert, R.; Richter, E.; Wanielik, G. Comparison and evaluation of advanced motion models for vehicle tracking. In Proceedings of the 2008 11th International Conference on Information Fusion, Cologne, Germany, 30 June–3 July 2008; pp. 1–6.
18. Xia, Y.; Qu, Z.; Sun, Z.; Li, Z. A human-like model to understand surrounding vehicles' lane changing intentions for autonomous driving. *IEEE Trans. Veh. Technol.* **2021**, *70*, 4178–4189. [CrossRef]
19. Li, J.; Dai, B.; Li, X.; Xu, X.; Liu, D. A dynamic Bayesian network for vehicle maneuver prediction in highway driving scenarios: Framework and verification. *Electronics* **2019**, *8*, 40. [CrossRef]
20. Xie, G.; Gao, H.; Qian, L.; Huang, B.; Li, K.; Wang, J. Vehicle trajectory prediction by integrating physics-and maneuver-based approaches using interactive multiple models. *IEEE Trans. Ind. Electron.* **2017**, *65*, 5999–6008. [CrossRef]
21. Xiong, L.; Fu, Z.; Zeng, D.; Leng, B. Surrounding vehicle trajectory prediction and dynamic speed planning for autonomous vehicle in cut-in scenarios. In Proceedings of the 2021 IEEE Intelligent Vehicles Symposium (IV), Nagoya, Japan, 11–17 July 2021; pp. 987–993.
22. Gao, H.; Qin, Y.; Hu, C.; Liu, Y.; Li, K. An interacting multiple model for trajectory prediction of intelligent vehicles in typical road traffic scenario. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, *34*, 6468–6479. [CrossRef] [PubMed]
23. Li, X.; Ying, X.; Chuah, M.C. Grip++: Enhanced graph-based interaction-aware trajectory prediction for autonomous driving. *arXiv* **2019**, arXiv:07792.
24. Xu, C.; Li, M.; Ni, Z.; Zhang, Y.; Chen, S. Groupnet: Multiscale hypergraph neural networks for trajectory prediction with relational reasoning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 6498–6507.
25. Ju, C.; Wang, Z.; Long, C.; Zhang, X.; Chang, D.E. Interaction-aware kalman neural networks for trajectory prediction. In Proceedings of the 2020 IEEE Intelligent Vehicles Symposium (IV), Long Beach, CA, USA, 19 October–13 November 2020; pp. 1793–1800.
26. Messaoud, K.; Yahiaoui, I.; Verroust-Blondet, A.; Nashashibi, F. Attention based vehicle trajectory prediction. *IEEE Trans. Intell. Veh.* **2020**, *6*, 175–185. [CrossRef]
27. Li, J.; Yang, F.; Ma, H.; Malla, S.; Tomizuka, M.; Choi, C. Rain: Reinforced hybrid attention inference network for motion forecasting. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 16096–16106.
28. Guo, H.; Meng, Q.; Cao, D.; Chen, H.; Liu, J.; Shang, B. Vehicle trajectory prediction method coupled with ego vehicle motion trend under dual attention mechanism. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–16. [CrossRef]
29. Liang, M.; Yang, B.; Hu, R.; Chen, Y.; Liao, R.; Feng, S.; Urtasun, R. Learning lane graph representations for motion forecasting. In Proceedings of the Computer Vision–ECCV 2020, Glasgow, UK, 23–28 August 2020; pp. 541–556.
30. Wang, R.; Song, X.; Hu, Z.; Cui, Y. Spatio-Temporal Interaction Aware and Trajectory Distribution Aware Graph Convolution Network for Pedestrian Multimodal Trajectory Prediction. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 1–11. [CrossRef]
31. Li, L.; Zhao, W.; Wang, C.; Chen, Q.; Chen, F. BRAM-ED: Vehicle Trajectory Prediction Considering the Change of Driving Behavior. *IEEE/ASME Tran. Mechatr.* **2022**, *27*, 5690–5700. [CrossRef]

32. Ji, J.; Khajepour, A.; Melek, W.W.; Huang, Y. Path planning and tracking for vehicle collision avoidance based on model predictive control with multiconstraints. *IEEE Trans. Veh. Technol.* **2017**, *66*, 952–964. [CrossRef]
33. Rasekhipour, Y.; Khajepour, A.; Chen, S.-K.; Litkouhi, B. A Potential Field-Based Model Predictive Path-Planning Controller for Autonomous Road Vehicles. *IEEE Trans. Intell. Transp. Syst.* **2017**, *18*, 1255–1267. [CrossRef]
34. Ren, Y.; Zheng, L.; Yang, W.; Li, Y. Potential field–based hierarchical adaptive cruise control for semi-autonomous electric vehicle. *Proc. Inst. Mech. Eng. Part D J. Automob. Eng.* **2019**, *233*, 2479–2491. [CrossRef]
35. Deo, N.; Trivedi, M.M. Convolutional social pooling for vehicle trajectory prediction. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 1468–1476.