

# 基于图时空注意力的多车交互轨迹预测模型\*

张新锋<sup>1,2</sup> 赵娟<sup>1</sup> 刘国华<sup>1</sup> 刘鹏菲<sup>1</sup>

(1. 长安大学汽车学院, 西安 710064; 2. 新疆农业大学交通与物流工程学院, 乌鲁木齐 830052)

**【摘要】**为有效提取高速交通场景下车辆间的交互特征, 从而准确预测动态障碍轨迹, 基于编-解码框架, 提出基于图时空注意力的多车交互轨迹预测模型。结合斥力场和图模型建立车-车图交互场, 利用节点和邻接特征矩阵表征车辆之间的动态交互, 通过图空间注意力和时间多头注意力提取深层时空交互, 获取图时空融合编码; 将车辆横纵向行为意图独热编码与其拼接, 实现目标车辆多模态轨迹预测。利用NGSIM数据集进行验证, 相较于其他6种模型, 该模型RMSE和NLL值最低; 通过消融实验进一步验证图交互场的有效性, 结果表明, 该模型能够有效提高车辆轨迹预测精度。

**关键词:** 多车交互 斥力场 注意力机制 图模型 轨迹预测

中图分类号: TP399 文献标志码: A DOI: 10.19620/j.cnki.1000-3703.20240734

## Multi-Vehicle Interaction Trajectory Prediction Model Based on Graph Spatial-Temporal Attention

Zhang Xinfeng<sup>1,2</sup>, Zhao Juan<sup>1</sup>, Liu Guohua<sup>1</sup>, Liu Pengfei<sup>1</sup>

(1. School of Automobile, Chang'an University, Xi'an 710064; 2. School of Transportation and Logistics Engineering, Xinjiang Agricultural University, Urumqi 830052)

**【Abstract】**In order to effectively extract interaction features among vehicles in high-speed traffic scenarios, thus accurately predict the trajectories of dynamic obstacles, this paper proposes a multi-vehicle interaction trajectory prediction model using the coding-decoding framework based on the graph spatial-temporal attention mechanism. The vehicle-to-vehicle graph interaction field is established by combining the repulsive force field and the graph model, the node feature matrix and the adjacency feature matrix are used to characterize the dynamic interaction between the vehicle and the surrounding vehicles, and the deep spatial-temporal interaction features are extracted by the graph spatial attention and temporal polytope attention to obtain the graph spatial-temporal fusion coding features. The one-hot encoding of the longitudinal and lateral behavior intentions of the vehicles is concatenated with the encoding to achieve multimodal trajectory prediction for the target vehicles. Validation using the NGSIM dataset shows that, compared with 6 other models, the proposed model achieves the lowest RMSE and NLL values. Ablation experiments further validate the effectiveness of the graph interaction field, demonstrating that the model can significantly improve the accuracy of vehicle trajectory prediction.

**Key words:** Multi-vehicle interaction, Repulsive fields, Attentional mechanisms, Graph modeling, Trajectory prediction

**【引用格式】**张新锋, 赵娟, 刘国华, 等. 基于图时空注意力的多车交互轨迹预测模型[J]. 汽车技术, 2025(3): 30-38.

ZHANG X F, ZHAO J, LIU G H, et al. Multi-Vehicle Interaction Trajectory Prediction Model Based on Graph Spatial-Temporal Attention[J]. Automobile Technology, 2025(3): 30-38.

## 1 前言

准确预测交通参与者未来运动轨迹有助于自动驾驶系统更好地适应复杂动态交通场景, 提高决策准确性和驾驶安全性<sup>[1]</sup>。

针对自动驾驶轨迹预测, 已有大量研究。Carvalho等<sup>[2]</sup>提出综合环境建模和随机模型预测控制(Model Predictive Control, MPC)框架, 利用交互式多模型和卡尔曼滤波预测目标车辆轨迹, 但依赖当前状态迭代计算。Danielsson等<sup>[3]</sup>使用蒙特卡洛方法预测交通场景中

\*基金项目: 陕西省重点研发计划项目资助(2022GY-303); 西安市科技计划项目资助(2022GXFW0152)。

通信作者: 张新锋(1976—), 男, 副教授, 研究生导师, 主要研究方向为自动驾驶、智能网联汽车与交通等, zhxf@chd.edu.cn。

的潜在危险,但需消耗大量计算资源生成随机样本,且结果准确性过于依赖样本数量和质量。

Mandalia等<sup>[4]</sup>首次利用支持向量机(Support Vector Machine, SVM)通过轨迹坐标、速度等特征实现车辆换道意图预测识别。Izquierdo等<sup>[5]</sup>采用SVM实现对车辆未来动作的预测,且预测精度得到显著提升。Deng等<sup>[6]</sup>通过模糊逻辑对驾驶场景分类,并针对不同驾驶场景实现车辆驾驶行为的预测。Gao等<sup>[7]</sup>采用隐马尔可夫模型(Hidden Markov Model, HMM)对车辆轨迹进行预测。然而,这类方法忽略了环境和交互因素,并不适用于所有交通环境。

随着深度学习的应用,车辆轨迹预测方法取得显著进展。Dai等<sup>[8]</sup>考虑车辆时空特征利用长短期记忆网络(Long Short-Term Memory, LSTM)实现车辆轨迹预测,并通过NGSIM数据集进行验证。季学武等<sup>[9]</sup>利用LSTM编码器将车辆轨迹信息编码为上下文向量,并将其未来位置表示为概率分布,实现了车辆轨迹的多模态预测。Messaoud等<sup>[10]</sup>利用注意力机制为周车分配权重,提出基于多头注意力机制的LSTM轨迹预测模型。Phan-Minh等<sup>[11]</sup>将图像作为输入,通过卷积神经网络(Convolutional Neural Network, CNN)生成车辆未来轨迹。近年来,图神经网络发展迅速,Sheng等<sup>[12]</sup>利用图卷积网络提取车辆间时空交互特征,实现车辆未来轨迹预测。尽管当前深度学习轨迹预测方法已取得较为理想的预测效果,但多数模型直接将车间欧式距离和相对运动状态拼接输入注意力机制提取潜在特征变量,忽视了动态车辆间相对运动的影响机理。

因此,本文引入斥力场和图模型表征动态车辆之间的空间相互作用机理,并结合时空注意力机制,基于图神经网络和长短时记忆网络提出基于图时空注意力机制的多车交互轨迹预测模型(Graph Spatial-Temporal Attention Mechanism, GSTAM)。

## 2 车辆轨迹预测问题描述

自动驾驶汽车在行驶过程中受到周围车辆的影响,目标车辆未来轨迹的准确预测是自动驾驶汽车规划未来路径和确保行驶安全的基础。本文通过获取车辆行驶时的交通信息以及目标车辆与周围车辆之间的交互关系,来生成准确、安全的未来轨迹。

车辆轨迹预测问题描述为:在 $t$ 时刻,将目标车辆与周围车辆的历史轨迹特征作为模型的输入,包括车辆横纵坐标、速度及加速度等,得到目标车辆未来预测时刻横纵坐标。将车辆 $i$ 的历史轨迹 $X_i$ 表示为:

2025年 第3期

$$X_i = (p_i^1, p_i^2, \dots, p_i^T) \quad (1)$$

式中: $T$ 为观测历史轨迹序列的长度, $p_i^t = (x_i^t, y_i^t, v_i^t, a_i^t)$ 为车辆 $i$ 在时刻 $t$ 的横坐标、纵坐标、速度及加速度。

将车辆 $i$ 在 $H$ 预测时长内的轨迹真值 $Y_i$ 与预测值 $\hat{Y}_i$ 表示为:

$$Y_i = (p_i^{T+1}, p_i^{T+2}, \dots, p_i^{T+H}) \quad (2)$$

$$\hat{Y}_i = (\hat{p}_i^{T+1}, \hat{p}_i^{T+2}, \dots, \hat{p}_i^{T+H}) \quad (3)$$

式中: $\hat{p}_i^t = (\hat{x}_i^t, \hat{y}_i^t)$ 为车辆 $i$ 在时刻 $t$ 的预测值。

把目标车辆与周围车辆 $1 \sim T$ 时刻的历史运动状态输入网络模型,可得到目标车辆在 $(T+1) \sim (T+H)$ 时刻的未来预测轨迹坐标 $\hat{Y}_i$ 。

## 3 基于图时空注意力机制的多车交互轨迹预测模型架构

本文所提出的车辆轨迹预测模型主要包括编码器和解码器两部分,其模型架构如图1所示,编码器由斥力场模块、车-车图交互场模块、图时空注意力编码模块组成,主要对车辆节点和车辆历史运动信息进行编码,形成特征向量。解码器通过对特征向量进行解码处理,最终生成多模态预测轨迹。

### 3.1 车-车图交互场构建

#### 3.1.1 斥力场模块

人工势场模型是一种用于路径规划和避障的算法,其基本原理是通过构建一个虚拟的力场,使车辆能够自动避开障碍物并朝向目标点移动<sup>[13]</sup>。其中障碍物被视为斥力源,对车辆产生排斥力,迫使车辆避开周围障碍物。当车辆处于多车交互的复杂交通环境时,通过这种斥力机制可以揭示其潜在的碰撞风险,对轨迹进行调节,从而保证自车安全的行车空间。

周围车辆对目标车辆的影响随距离的增加而逐渐变小,这种影响与车辆运动状态信息高度相关。本文结合斥力场理论,采用改进二元正态分布将目标车辆与周围车辆的交互作用表示为:

$$F_{\text{rep}} = R \cdot e^{\left( a \cos \beta - \left( \frac{(x_i - x_j)^2}{2\sigma_x^2} + \frac{(y_i - y_j)^2}{2\sigma_y^2} \right) \right)} \quad (4)$$

$$R = \lambda(L_i + W_i)(1.566v^{6.687} \times 10^{-14} + 0.3345) \quad (5)$$

式中: $x_i$ 、 $y_i$ 、 $x_j$ 、 $y_j$ 分别为目标车辆与周车的横、纵坐标; $R$ 为目标车辆的危险度函数; $L_i$ 、 $W_i$ 分别为目标车辆的长度和宽度; $\beta$ 为周车与目标车辆中心点的连线与 $x$ 轴正方向的夹角; $\sigma_x$ 、 $\sigma_y$ 为交互作用力场纵横分布因子,与目标车辆 $x$ 、 $y$ 方向的速度和加速度 $v_x$ 、 $v_y$ 、 $a_x$ 、 $a_y$ 以及速度和加速度影响因子 $\gamma$ 、 $\delta$ 相关,计算公式如下:

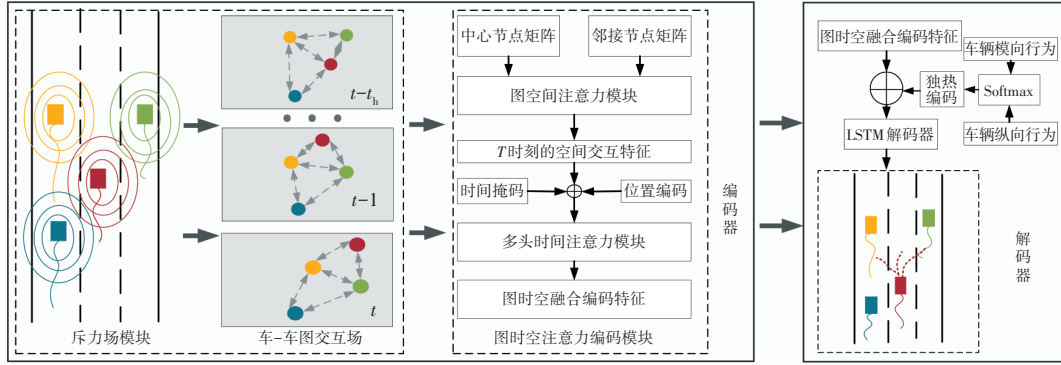


图1 基于图时空注意力机制的多车交互轨迹预测模型架构

$$\sigma_x = L_t \sqrt{1 + \gamma a_x^2 + \delta v_x^2} + \delta_x \quad (6)$$

$$\sigma_y = W_t \sqrt{1 + \gamma a_y^2 + \delta v_y^2} + \delta_y \quad (7)$$

### 3.1.2 车-车图交互场

图模型是一种采用图形来表示对象之间的关系或结构的数学模型,其中节点表示对象,边表示它们之间的关系。采用图结构来表征多车交互交通场景下目标车辆与周围车辆之间的交互作用,建立的车-车图交互场如图2所示。

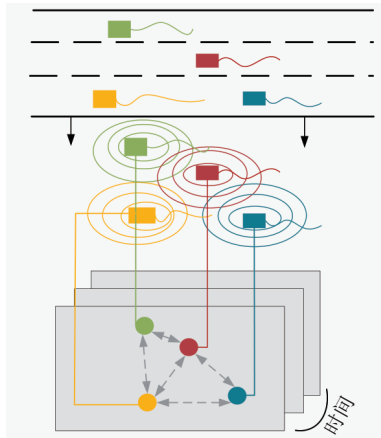


图2 车-车图交互场

利用节点特征矩阵和邻接特征矩阵对车辆节点以及车-车之间的交互关系进行表征,时刻 $t$ 车辆节点特征矩阵 $H_V^t$ 为:

$$H_V^t = [\text{node}_i^t]_{i \in \Theta_{\text{vehicle}}} \quad (8)$$

$$\text{node}_i^t = [\mathbf{x}_i^t - \mathbf{x}_i^{t-1}, \mathbf{y}_i^t - \mathbf{y}_i^{t-1}, v_i^t] \quad (9)$$

式中: $\Theta_{\text{vehicle}}$ 为当前场景下所有车辆, $\mathbf{x}_i^t$ 、 $\mathbf{y}_i^t$ 、 $v_i^t$ 为 $t$ 时刻目标车辆节点 $i$ 的横、纵坐标和速度, $\mathbf{x}_i^t - \mathbf{x}_i^{t-1}$ 、 $\mathbf{y}_i^t - \mathbf{y}_i^{t-1}$ 为时刻 $t$ 目标车辆节点 $i$ 的横、纵轴轨迹向量分量。

对于邻接特征矩阵,结合斥力场模块,可得时刻 $t$ 邻接特征矩阵 $X_V^t$ 为:

$$X_V^t = [A_{ji} \cdot e_{j \rightarrow i}^t]_{i \in \Theta_{\text{vehicle}}, j \in \Theta_{\text{nbr}}^i} \quad (10)$$

$$e_{j \rightarrow i}^t = [F_{\text{rep}}^{j \rightarrow i}, \mathbf{x}_j^t - \mathbf{x}_i^t, \mathbf{y}_j^t - \mathbf{y}_i^t] \quad (11)$$

式中: $\Theta_{\text{nbr}}^i$ 为目标车辆节点 $i$ 的所有邻接车辆节点; $F_{\text{rep}}^{j \rightarrow i}$ 为邻接车辆节点 $j$ 与目标车辆节点 $i$ 的斥力场值,由3.1.1节计算得到; $A_{ji}$ 为邻接矩阵。

在车-车图交互场构建中,车辆节点之间的交互作用通过算法实现自动标注。斥力场根据车辆间的距离及其相对速度和加速度量化交互强度,当斥力场值 $F_{\text{rep}}$ 超过模型损失最小的斥力场值 $F_{\text{rep}}^*$ (由模型训练得到)时,即判定两车间存在交互作用,模型将邻接矩阵的相应元素置为1,否则为0。邻接矩阵计算公式如下:

$$A_{ji} = \begin{cases} 1, & \text{节点 } j \text{ 到 } i \text{ 存在交互} \\ 0, & \text{节点 } j \text{ 到 } i \text{ 没有交互} \end{cases} \quad (12)$$

### 3.2 图时空注意力编码模块

图注意力(Graph Attention)能够对图结构数据进行有效处理,并利用注意力机制来改进图神经网络<sup>[4]</sup>(Graph Neural Network, GNN)的消息传递方式,每个节点根据其邻接节点的重要性来计算加权平均特征。本文利用图时空注意力编码模块,深度提取被预测车辆与周围车辆之间的时空耦合交互。

#### 3.2.1 空间图注意力编码模块

在多车交互的交通场景下,目标车辆的驾驶行为与周围车辆的驾驶行为高度相关。在空间维度上,设计空间图注意力编码模块,如图3所示,用于提取目标车辆与周围车辆在空间维度的高度关联性。

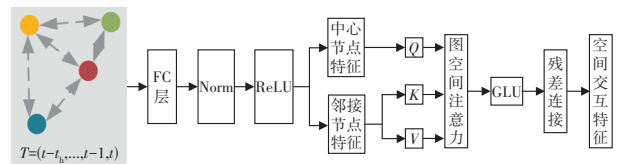


图3 空间图注意力编码模块

首先,将车-车图交互场得到的目标车辆节点特征 $\text{node}_i^t$ 与邻接车辆节点特征 $\text{node}_j^t$ 进行维度变换,以确保后续图空间注意力机制能够深度提取车辆之间的空间交互关系,具体公式如下:

$$h_i^t = \phi_{\text{center}}(\text{node}_i^t) \quad (13)$$

$$h_{ij}^t = \phi_{nbr}(node_j^t, e_{j \rightarrow i}^t) \quad (14)$$

式中:  $\phi_{center}$ 、 $\phi_{nbr}$  为特征向量经过连续的全连接层、层归一化及 ReLU 激活变换,  $h_i^t$  为  $t$  时刻目标车辆节点特征编码,  $h_{ij}^t$  为  $t$  时刻目标车辆节点  $i$  与其邻接车辆节点  $j$  的交互特征编码。

在研究目标车辆与周围车辆的空间交互作用时,仅考虑它们之间的欧式距离是不够合理的。虽然欧式距离是一个基本的衡量标准,但它忽略了影响车辆交互的动态信息。因此,本文基于建立的车-车图交互场模型,结合注意力机制<sup>[15]</sup>,形成空间图注意力机制,深度提取车辆在空间上的交互行为。将目标车辆节点特征编码  $h_i^t$  和交互特征编码  $h_{ij}^t$  输入至空间注意力模块,注意力机制通过 Query(查询)与 Key(键)的注意力汇聚实现对 Value(值)的注意力权重分配,生成最终的输出结果,其具体的实现过程为:

$$Q^t = W^q \cdot h_i^t, \quad K^t = W^k \cdot h_{ij}^t, \quad V^t = W^v \cdot h_{ij}^t \quad (15)$$

式中:  $Q^t$ 、 $K^t$ 、 $V^t$  为  $t$  时刻注意力机制的查询、键、值,  $W^q$ 、 $W^k$ 、 $W^v$  为空间注意力中对应的线性权重矩阵。

计算  $h_{ij}^t$  对  $h_i^t$  的注意力值为:

$$Z_i^t = \text{softmax} \left( \frac{Q^t}{\sqrt{d_k}} (K^t)^T \right) \cdot V^t \quad (16)$$

将维度  $d$  等分  $m$  份,  $d/m=d_k$ , 利用门控函数对目标车辆节点与邻接车辆节点的交互信息进行更新与聚合,具体的实现过程为:

$$\text{gate}_i^t = \text{sigmoid}(W_{center} h_i^t + W_{nbr} \cdot Z_i^t) \quad (17)$$

$$\hat{Z}_i^t = \text{gate}_i^t \odot W_{self} \cdot h_i^t + (1 - \text{gate}_i^t) \odot Z_i^t \quad (18)$$

式中:  $\text{gate}_i^t$  为  $t$  时刻门控机制的输出,  $\hat{Z}_i^t$  为更新聚合后的目标车辆节点特征,  $W_{center}$ 、 $W_{nbr}$ 、 $W_{self}$  为对应的权重矩阵,  $\odot$  为矩阵的乘积。

将历史轨迹时长内的车辆运动特征通过空间图注意力编码模块可得目标车辆与周围车辆的空间交互特征  $S_{enc} \in \mathbb{R}^{N \times T \times D}$ ,  $N$  为当前交通场景所含车辆节点数量总和,  $T$  为历史轨迹时长,  $D$  为特征维度。

### 3.2.2 时间多头注意力模块

在当前时刻,车辆自身状态与其过去的状态紧密相关,且车辆在行驶中与周围车辆的空间交互随时间动态变化。因此,设计时间多头注意力模块,如图4所示,用于捕捉车辆之间随时间推移而变化的空间交互的内在相关性。

首先,将空间交互特征编码  $S_{enc}$  作为输入,并引入时间掩码矩阵,以此确保仅计算当前时刻的时间依赖特征。定义时间掩码矩阵为上三角矩阵,设置需要掩盖的位置为负无穷或者极小值,使得其在后续操作中注意力权重趋近

于零。此外,构建位置编码以便模型理解输入历史轨迹特征的顺序性,采用三角函数对位置编码,可表示为:

$$PoE(t, 2i) = \sin \left( \frac{t}{10000^{\frac{2i}{D}}} \right) \quad (19)$$

$$PoE(t, 2i + 1) = \cos \left( \frac{t}{10000^{\frac{2i}{D}}} \right) \quad (20)$$

式中:  $PoE(t, 2i)$  和  $PoE(t, 2i + 1)$  分别为时刻  $t$  历史轨迹空间特征在  $2i$  和  $(2i + 1)$  维度的位置编码。

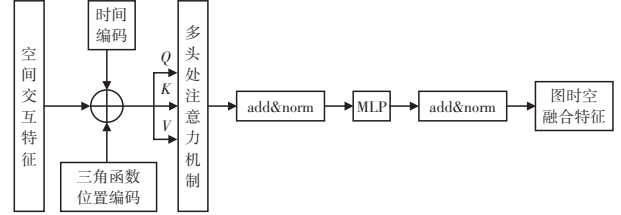


图4 时间多头注意力模块

将位置编码与空间交互特征编码结合,结合多头注意力机制,提取车辆在时间维度上的交互变化。具体步骤如下:

$$Q_i = W^{Q_i}(S_{enc} + PoE), \quad K_i = W^{K_i}(S_{enc} + PoE), \quad (21)$$

$$V_i = W^{V_i}(S_{enc} + PoE)$$

式中:  $W^{Q_i}$ 、 $W^{K_i}$ 、 $W^{V_i}$  为时间注意力机制中的线性权重矩阵。

计算时刻  $t$  的注意力值  $head_i$ :

$$\text{head}_i = \text{softmax} \left( \frac{Q_i}{\sqrt{d_k}} (K_i)^T \right) \cdot V_i \quad (22)$$

多头注意力机制能够避免注意力过度集中在自身位置,并对轨迹时间特征从不同层面进行提取,因此将式(22)平行计算  $h$  次, Con 为拼接操作,将  $h$  个注意力头进行合并:

$$\text{MhA}(q_i, k_i, v_i) = \text{Con}(\text{head}_1, \text{head}_2, \dots, \text{head}_h) \quad (23)$$

在特征维度上对  $Q_i$ 、 $K_i$ 、 $V_i$  均分并求其注意力值  $head_i$ , 随后进行拼接并经过层归一化、残差连接及 MLP 得到其完整的图时空融合特征编码  $ST_{enc} \in \mathbb{R}^{N \times D}$ ,  $N$  为当前交通场景所含车辆节点数量总和。

### 3.3 解码器

车辆行驶轨迹通常呈现多模态特征,通过将驾驶员的意图建模为横向与纵向行为来进行表征。横向行为包括当前车道行驶、左换道、右换道,而纵向行为则涵盖恒定速度、加速、减速。将横向和纵向行为类型组合构成驾驶行为集  $N$ , 即:  $N = \{n_i | i = 1, 2, \dots, 9\}$ 。

将图时空注意力编码模块输出的图时空融合特征  $ST_{enc}$  经过含 Softmax 激活函数的全连接(Fully

Connected, FC)层,可得横向行为和纵向行为意图概率 $P_H, P_V$ ,进一步计算每种驾驶行为意图的概率分布 $P(n_i|X)$ 。具体可表示为:

$$\begin{cases} P_H = \text{Softmax}(S_{\text{Tenc}}\mathbf{W}_H + b_H) \\ P_V = \text{Softmax}(S_{\text{Tenc}}\mathbf{W}_V + b_V) \end{cases} \quad (24)$$

式中, $P_H, P_V$ 为车辆横、纵向行为意图概率; $S_{\text{Tenc}}$ 为图时空注意力编码模块输出的图时空融合特征; $\text{Softmax}$ 为一种激活函数, $\mathbf{W}_H, \mathbf{W}_V$ 为对应的权重矩阵, $b_H, b_V$ 为对应的偏置。

利用二元高斯分布对未来轨迹的不确定性进行建模,可表示为:

$$P(Y|X) = \prod_i P_R(Y|n_i)P(n_i|X), n_i \in N \quad (25)$$

式中: $R=[R_{t+1}, \dots, R_{t+T}]$ 为二元高斯分布参数, $R^t = (\boldsymbol{\mu}^t, \boldsymbol{\Sigma}^t)$ 为车辆未来位置在时刻 $t$ 的均值和协方差。

然后,将横纵向行为意图的独热编码向量与图时空注意力编码模块的输出 $ST_{\text{enc}}$ 拼接,通过LSTM解码器得到目标车辆的未来预测轨迹,具体表示为:

$$e^t = FC_e(\text{Concat}(S_{\text{Tenc}}, P_H, P_V), \mathbf{W}_e) \quad (26)$$

$$h_d^t = \text{LSTM}(e^t, h_d^{t-1}, \mathbf{W}_{\text{decoder}}) \quad (27)$$

$$\theta^t = (\boldsymbol{\mu}^t, \boldsymbol{\sigma}^t, \boldsymbol{\rho}^t) = FC_\theta(h_d^t, \mathbf{W}_\theta) \quad (28)$$

式中: $\mathbf{W}_e$ 和 $\mathbf{W}_\theta$ 为对应FC层的可学习权重矩阵, $\mathbf{W}_{\text{decoder}}$ 为解码器的可学习权重矩阵, $\boldsymbol{\mu}^t, \boldsymbol{\sigma}^t, \boldsymbol{\rho}^t$ 分别为未来轨迹预测值的均值、方差和相关系数。

$\boldsymbol{\mu}^t$ 和 $\boldsymbol{\Sigma}^t$ 为车辆在未来时刻的坐标位置和不确定性程度,可表示为:

$$\boldsymbol{\mu}^t = \begin{pmatrix} \mu_x^t \\ \mu_y^t \end{pmatrix}, \boldsymbol{\Sigma}^t = \begin{pmatrix} (\sigma_x^t)^2 & \sigma_x^t \sigma_y^t \rho^t \\ \sigma_x^t \sigma_y^t \rho^t & (\sigma_y^t)^2 \end{pmatrix} \quad (29)$$

### 3.4 损失函数

均方误差(Mean Squared Error, MSE)反映了模型预测值与实际观测值之间的差异程度,是衡量模型预测性能的重要指标。在模型训练前 $n$ 轮次的初期阶段,由于模型参数尚未收敛,使用MSE能够帮助模型快速优化,减少初始的预测误差,因此将均方误差作为模型初始训练损失函数。MSE损失函数可表示为:

$$M = \sum_{t=1}^{T_1} \left[ (x_{\text{te}}^t - x_{\text{pd}}^t)^2 + (y_{\text{te}}^t - y_{\text{pd}}^t)^2 \right] \quad (30)$$

式中: $(x_{\text{te}}^t, y_{\text{te}}^t)$ 和 $(x_{\text{pd}}^t, y_{\text{pd}}^t)$ 分别为车辆在时刻 $t$ 的实际坐标和预测坐标。

车辆轨迹预测问题本质上是一个多模态问题,第 $n$ 轮次之后,单纯的MSE难以刻画轨迹的多样性和不确

定性。而极大似然估计的负对数似然(Negative Log-Likelihood, NLL)可以通过对预测轨迹的概率分布建模,更好地处理多模态输出,提高预测的准确性和稳定性。通过对训练损失值和验证集的监测,当模型的MSE损失在约10轮训练后趋于稳定,因此将NLL作为模型训练后期阶段的损失函数。NLL损失函数可表示为:

$$L = -\log(P_R(Y|n_{\text{te}}, X)P(n_{\text{te}}|X)) \quad (31)$$

这种根据不同阶段选择合适的损失函数的训练策略,提高了模型的预测精度和收敛速度。

## 4 基于NGSIM数据集的实验验证与分析

### 4.1 数据集及其衡量方式

本文采用美国联邦公路管理局(Federal Highway Administration, FHWA)收集的下一代仿真<sup>[16]</sup>(Next Generation Simulation, NGSIM)数据集,选取其中的US101、I-80高速公路路段车辆轨迹数据集验证第3节提出的轨迹预测模型性能,数据集中以10 Hz的频率统计了车辆及前后车辆的详细运动信息,包括车辆尺寸、横纵坐标、车道信息、速度以及加速度等,研究区域路段如图5所示。

由于I-80和US-101数据集是从高速公路上方多个摄像头采集的视频中提取得到,因此存在一定的噪声和误差。首先利用多步滤波法和对数指数平均对车辆轨迹数据进行预处理,并利用处理后的位置数据重构速度、加速度等信息<sup>[17]</sup>。利用预处理后的数据集提取车辆轨迹,其中每条轨迹被分为8 s的片段,前3 s为输入模型的历史轨迹时长,后5 s则用于未来轨迹的预测。将数据集按7:1:2的比例划分为训练数据、评估数据和测试数据。

### 4.2 训练参数设置

利用Pytorch深度学习框架对本文提出的基于图时空注意力的多车交互轨迹预测模型进行训练与评估。模型的超参数根据文献[18]~文献[22]选取,设置批量大小为128,初始学习率为0.000 5,丢弃率为0.1,权重衰减为0.000 1,并通过Adam优化器对模型进行参数优化。

### 4.3 试验结果与分析

采用以下两种指标对模型轨迹预测质量与效果进行合理评估。

均方根误差(Root Mean Square Error, RMSE)是车辆轨迹预测的常用评价指标之一,时刻 $t$ 的RMSE计算公式为:

$$R = \sqrt{\frac{1}{N} \sum_{i=1}^N \left[ (\hat{x}_i^t - x_i^t)^2 + (\hat{y}_i^t - y_i^t)^2 \right]} \quad (32)$$

式中: $N$ 为车辆样本数, $(\hat{x}_i^t, \hat{y}_i^t)$ 和 $(x_i^t, y_i^t)$ 分别对应车辆 $i$ 在时刻 $t$ 的预测位置和地面真实位置。

真实轨迹的NLL侧重于评估多模态轨迹预测的准确性。计算公式为:

$$N = -\log \left( \sum_i P_{\theta}(Y|m_i, X) P(m_i|X), m_i \in M \right) \quad (33)$$

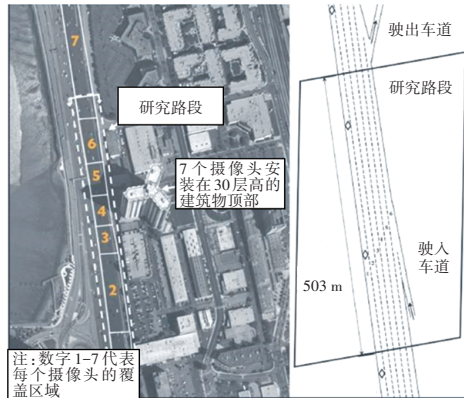
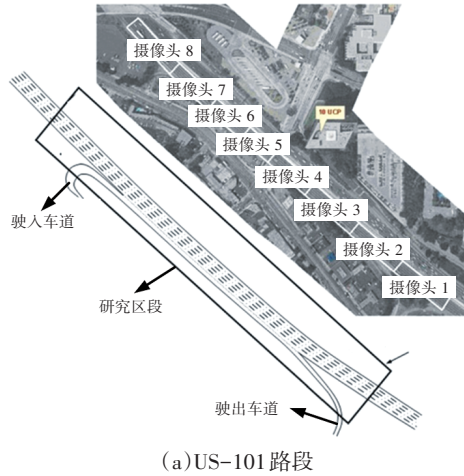


图5 US-101和I-80研究路段

### 4.3.1 测试集实验结果

在完成模型训练后,采用测试数据对模型的预测效果进行评估,模型在1~5 s预测时长内的RMSE和NLL指标值如表1所示。

表1 预测模型实验结果

预测时长/s	RMSE/m	NLL
1	0.37	1.16
2	0.96	2.86
3	1.65	3.80
4	2.53	4.46
5	3.68	5.05

对表1进行分析可知,在1~5 s的预测时长上所提出的预测模型均具有较低的RMSE和NLL值,这表明该模

型可以对车辆轨迹进行较为准确的预测,此外,可以看出,在5 s的长时预测中,该模型依然展现出良好的性能。

### 4.3.2 模型对比实验结果

为了验证模型的预测精度,选择CV、CS-LSTM<sup>[18]</sup>、PiP<sup>[19]</sup>、S-LSTM<sup>[20]</sup>、ADI-DCS-GRU(M)<sup>[21]</sup>、TSMHA<sup>[22]</sup>6种模型与GSTAM模型进行对比分析,所有模型的超参数与4.2节相同。对比模型如下:

- a. CV:通过恒速卡尔曼滤波器得到车辆的未来预测轨迹;
- b. CS-LSTM:利用编-解码结构,并叠加卷积和池化层捕捉车辆时空交互,生成目标车辆的多模态轨迹预测;
- c. PiP:考虑自车和周围目标车辆的空间关系,利用注意力机制实现轨迹预测;
- d. S-LSTM:利用LSTM编-解码器预测车辆未来轨迹;
- e. ADI-DCS-GRU(M):利用注意力机制和深度交互GRU算法实现车辆未来轨迹的多模态预测;
- f. TSMHA:考虑时空维度特征影响,利用多头注意力机制实现车辆未来轨迹的预测。

利用RMSE和NLL两个指标对本文模型与6种模型的结果<sup>[18-22]</sup>进行对比分析,结果如图6、图7所示。

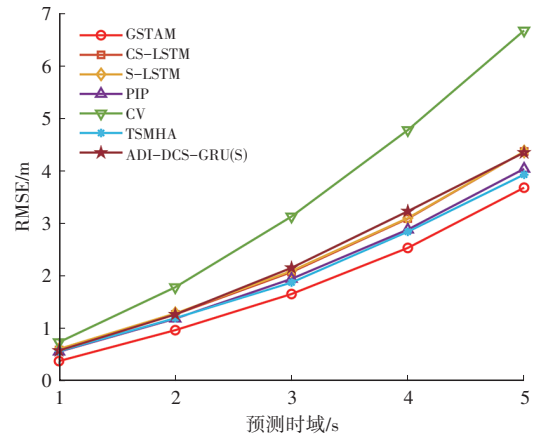


图6 RMSE指标对比

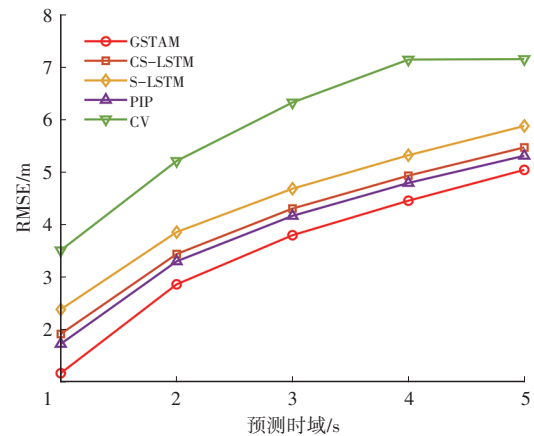


图7 NLL指标对比

对图6、图7分析可知,本文提出的模型拥有最低的

RMSE 和 NLL 值,相较于其他模型,能够更加准确地对目标车辆的未来轨迹进行预测。

#### 4.4 消融实验

消融实验通过逐步剔除或更改模型的特定组成部分,验证各个模块对模型整体性能的重要性程度。

将组成该预测模型的车-车图交互场、空间图注意力交互、时间多头注意力模块分别进行消融,选用 5 s 长时多模态预测效果进行评估,其结果如表 2 所示。

表 2 消融实验结果

策略	图交互场	空间	时间	RMSE/m	NLL
1	×	√	√	4.10	5.23
2	√	×	√	3.79	5.09
3	√	√	×	3.91	5.15
4	√	√	√	3.68	5.05

对表 2 进行分析可得,车-车图交互场、空间图注意力交互、时间多头注意力模块对该模型轨迹预测性能的提高均表现为积极作用。对车-车图交互场进行消融,模型具有最高的 RMSE 和 NLL 值,其模型预测性能降低最为明显,这是因为车-车图交互场能够有效表征目标车辆与周围车辆的交互作用,为模型的后续模块学习车辆历史轨迹之间的时空交互提供了丰富的先验知识;相较于其他两个模块,空间图注意力交互模块对于模型预测性能的贡献程度较低,这是因为车-车图交互场在此之前已经对车辆之间的空间交互作用进行了有效提取;除此之外,时间多头注意力模块对模型的预测性能也具有较大的影响,这主要是因为车辆运动状态不断随着时间进行变化,且车辆当前时刻的运动状态与其历史运动状态高度相关。

#### 4.5 预测结果可视化

为了更加直观地评估该预测模型的性能,对多车交互场景下的多模态轨迹预测结果进行可视化操作。选用 NGSIM 数据集中拥挤程度不同的交通场景和换道交通场景,对预测结果进行可视化分析。将目标车辆作为坐标原点,黑色虚线为车辆历史轨迹,红色虚线为目标车辆的多模态预测轨迹,黑色实线为目标车辆的真实轨迹。

##### 4.5.1 轻度拥挤交通场景

从数据集中提取轻度拥挤场景对多模态的预测结果进行可视化,结果如图 8 所示。

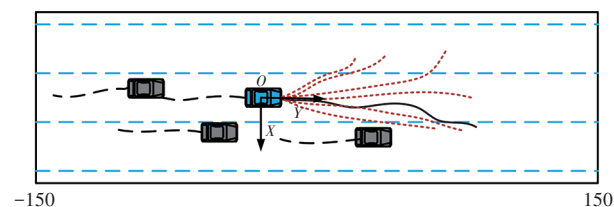


图 8 轻度拥挤交通场景多模态预测结果

从得到的多模态预测轨迹中提取最优预测轨迹,对测试集中直行、左换道、右换道的交通场景进行可视化,结果如图 9 所示。

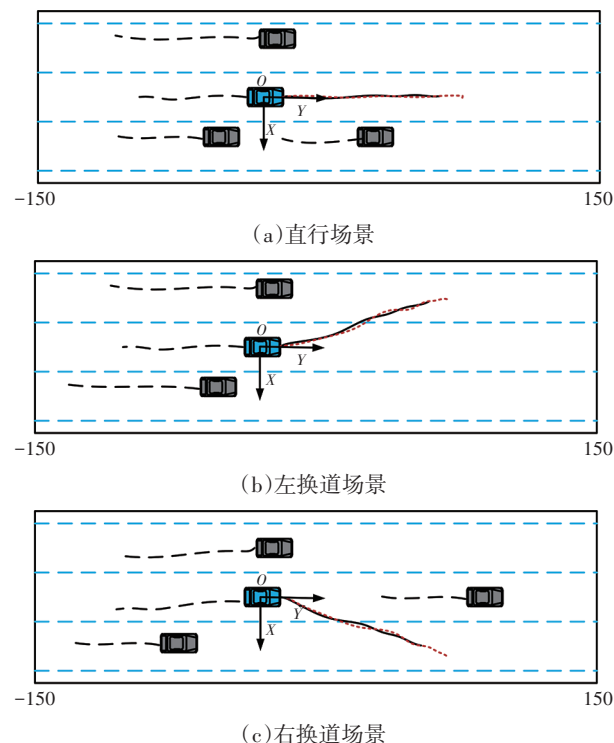


图 9 轻度拥挤交通场景可视化结果

由图 9 分析可得,在轻度拥挤交通场景下,分布在目标车辆的周围车辆较少,无论是直行场景,还是换道场景,GSTAM 模型均能够精确预测目标车辆的未来轨迹。

##### 4.5.2 中度拥挤交通场景

从数据集中提取中度拥挤场景对多模态的预测结果进行可视化,结果如图 10 所示。

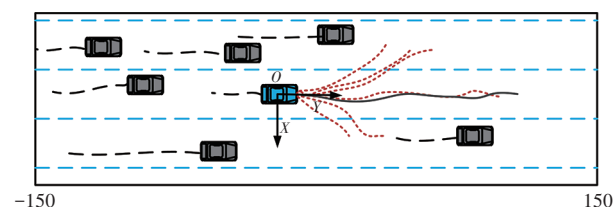


图 10 中度拥挤交通场景多模态预测结果

从得到的多模态预测轨迹中提取最优预测轨迹,对测试集中直行、左换道、右换道的交通场景进行可视化,结果如图 11 所示。

在中度拥挤交通场景下,车辆分布较为密集,由图 11a 可知,在当前直行场景下,目标车辆左侧与右侧均有车辆行驶,对于目标车辆,此时最好的行驶策略为保持当前车道行驶,预测出的未来轨迹与目标车辆的真实轨迹基本吻合。在换道场景下,分析图 11b 和图 11c 可知,在当前左换道场景下,周围车辆主要分布在目标车辆左后方和右前方,此时目标车辆进行左换道,预测出

的未来轨迹与其历史轨迹一致。同理,右换道场景下,目标车辆的未来预测轨迹与其真实轨迹依然保持一致。

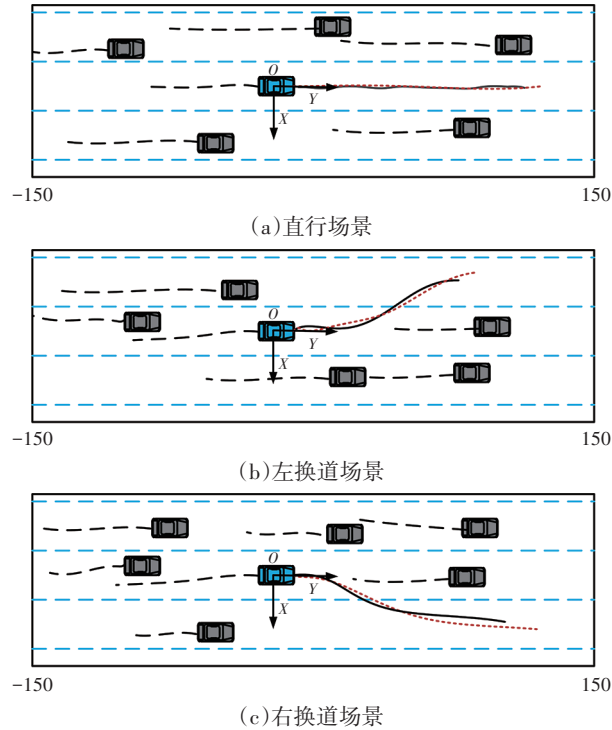


图11 中度拥挤交通场景可视化结果

#### 4.5.3 重度拥挤交通场景

从数据集中提取重度拥挤场景对多模态的预测结果进行可视化,结果如图12所示。

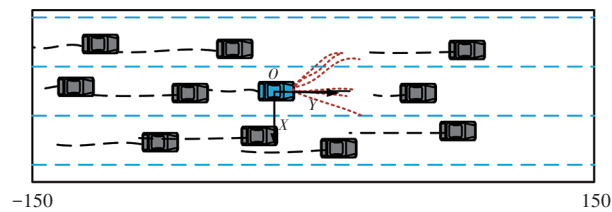
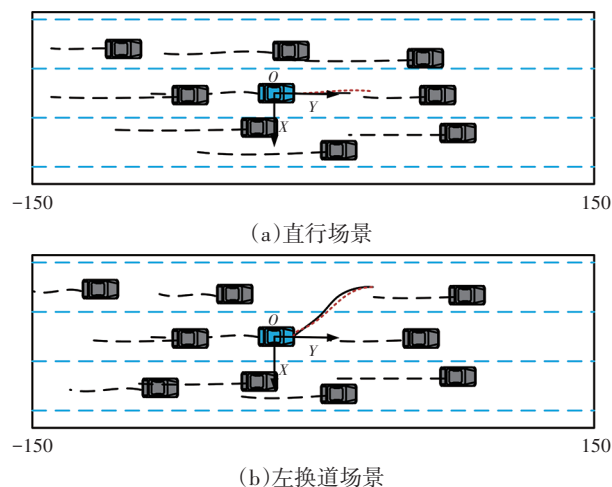


图12 重度拥挤交通场景多模态预测结果

从得到的多模态预测轨迹中提取最优预测轨迹,对测试集中直行、左换道、右换道的交通场景进行可视化,结果如图13所示。



(b)左换道场景

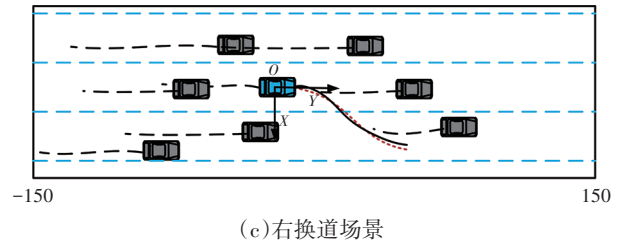


图13 重度拥挤交通场景可视化结果

在重度拥挤交通场景下,目标车辆周围分布的车辆十分密集,目标车辆趋向于跟驰行驶。由图13可得,在直行场景下,此时目标车辆左侧与右侧车辆十分密集,目标车辆无法进行换道,目标车辆保持当前车辆行驶,预测出的车辆轨迹与真实轨迹一致。在换道场景下,预测出的目标车辆未来轨迹与真实轨迹也基本一致。

分析图8~图13的可视化结果可知,在3种不同拥挤程度的交通场景下,本文提出的模型均能对目标车辆的未来轨迹进行较为精确地预测。

## 5 结束语

本文针对多车交互场景下自动驾驶车辆轨迹预测问题,利用斥力场和图模型表征车辆之间的交互关系,且结合时空注意力机制深度提取目标车辆与周围车辆的动态时空交互特征,并将其与车辆横纵向行为意图融合,提出了GSTAM车辆轨迹预测模型,基于NGSIM数据集对模型进行验证与分析,结果表明:

与CS-LSTM、S-LSTM、ADI-DCS-GRU(M)、TSMHA等模型相比,GSTAM模型具有最低的RMSE和NLL值,表明该模型能够对自动驾驶车辆轨迹进行更为精确地预测。

对模型进行消融实验,发现移除车-车图交互场模型性能降低明显,可见GSTAM模型引入的车-车图交互场能够为轨迹预测模型后续提取车辆间时空交互提供丰富的先验知识,有效提高了轨迹预测的精确度。

选用不同拥挤程度下的场景对预测结果进行可视化,包括直行、左换道以及右换道,模型预测出的未来轨迹能够切合目标车辆的真实轨迹。

### 参考文献

- [1] 郭景华,肖宝平,王靖瑶,等. 基于Residual BiLSTM网络的车辆切入意图预测研究[J]. 汽车工程, 2021, 43(7): 971-977. GUO J H, XIAO B P, WANG J Y, et al. Study on Vehicle Cut-in Intention Prediction Based on Residual BiLSTM Network[J]. Automotive Engineering, 2021, 43(7): 971-977.
- [2] CARVALHO A, GAO Y, LEFEVRE S, et al. Stochastic Predictive Control of Autonomous Vehicles in Uncertain Environments[C]// 12th International Symposium on

- Advanced Vehicle Control. Tokyo: Society of Automotive Engineers of Japan, 2014: 1–9.
- [3] DANIELSSON S, PETERSSON L, EIDEHALL A. Monte Carlo Based Threat Assessment: Analysis and Improvements[J]. Intelligent Vehicles Symposium. Istanbul: IEEE, 2007: 233–238
- [4] MANDALIA H M, SALVUCCI M D D. Using Support Vector Machines for Lane-Change Detection[C]// Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Los Angeles: SAGE Publications Sage CA, 2005: 1965–1969.
- [5] IZQUIERDO R, PARRA I, MUNOZ-BULNES J, et al. Vehicle Trajectory and Lane Change Prediction Using Ann and Svm Classifiers[C]// 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC). Yokohama: IEEE, 2017: 1–6.
- [6] DENG Q, SOFFKER D. Improved Driving Behaviors Prediction Based on Fuzzy Logic-Hidden Markov Model (F1-Hmm)[C]// 2018 IEEE Intelligent Vehicles Symposium (IV). Changshu, China: IEEE, 2018: 2003–2008.
- [7] GAO H, QIN Y, HU C, et al. An Interacting Multiple Model for Trajectory Prediction of Intelligent Vehicles in Typical Road Traffic Scenario[J]. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34(9): 6468–6479.
- [8] DAI S, LI L, LI Z. Modeling Vehicle Interactions via Modified LSTM Models for Trajectory Prediction[J]. IEEE Access, 2019: 38287–38296.
- [9] 季学武, 费聪, 何祥坤, 等. 基于 LSTM 网络的驾驶意图识别及车辆轨迹预测[J]. 中国公路学报, 2019, 32(6): 34–42.
- JI X W, FEI C, HE X K, et al. Intention Recognition and Trajectory Prediction for Vehicles Using LSTM Network[J]. China[J]. China Journal of Highway and Transport, 2019, 32(6): 34–42.
- [10] MESSAOUD K, YAHIAOUI I, VERROUST-BLONDET A, et al. Attention Based Vehicle Trajectory Prediction[J]. IEEE Transactions on Intelligent Vehicles, 2020, 6(1): 175–185.
- [11] PHAN-MINH T, GRIGORE E C, BOULTON F A, et al. CoverNet: Multimodal Behavior Prediction Using Trajectory Sets[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020: 14074–14083.
- [12] SHENG Z, XU Y, XUE S, et al. Graph-Based Spatial-Temporal Convolutional Network for Vehicle Trajectory Prediction in Autonomous Driving[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(10): 17654–17665.
- [13] KHATIB O. Real-Time Obstacle Avoidance for Manipulators And Mobile Robots[J]. The International Journal of Robotics Research, 1986, 5(1): 90–98.
- [14] GORI M, MONFARDINI G, SCARSELLI F. A New Model for Learning in Graph Domains[C]// 2005 IEEE International Joint Conference on Neural Networks. Montreal: IEEE, 2005, 2: 729–734.
- [15] VASWANI A, SHAZEER N, PARMAR N, et al. Attention Is All You Need[J]. Advances in Neural Information Processing Systems, 2017: 1–15
- [16] ALEXIADIS V, COLYAR J, HALKIAS J, et al. The Next Generation Simulation Program[J]. Institute of Transportation Engineers ITE Journal, 2004, 74(8): 22–26.
- [17] 张新锋, 王万宝, 柳欢, 等. 高速动态交通场景下自动驾驶车辆换道意图识别模型研究[J]. 汽车技术, 2023(4): 8–15.
- ZHANG X F, WANG W B, LIU H, et al. Research on Lane Change Intention Recognition Model of Automated Vehicle in High-Speed Dynamic Traffic Scenario[J]. Automobile Technology, 2023(4): 8–15.
- [18] TRIVEDI M M, DEO N. Convolutional Social Pooling for Vehicle Trajectory Prediction[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Salt Lake City, UT, USA: IEEE, 2018: 1468–1476.
- [19] SONG H, DING W, CHEN Y, et al. Pip: Planning-Informed Trajectory Prediction for Autonomous Driving[C]// Computer Vision – ECCV 2020 16th European Conference. Glasgow, UK: Springer, 2020: 598–614.
- [20] ALAHI A, GOEL K, RAMANATHAN V, et al. Social LSTM: Human Trajectory Prediction in Crowded Spaces [C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA: IEEE, 2016: 961–971.
- [21] 田彦涛, 黄兴, 卢辉迺, 等. 基于注意力与深度交互的周车多模态行为轨迹预测[J]. 吉林大学学报(工学版), 2023(5): 1474–1480.
- TIAN Y T, HUANG X, LU H Q, et al. Multi-Mode Behavior Trajectory Prediction of Surrounding Vehicle Based on Attention and Depth Interaction[J]. Journal of Jilin University (Engineering and Technology Edition), 2023(5): 1474–1480.
- [22] 宋秀兰, 董兆航, 单杭冠, 等. 基于时空融合的多头注意力车辆轨迹预测[J]. 浙江大学学报(工学版), 2023, 57(8): 1636–1643.
- SONG X L, DONG Z H, SHAN H G, et al. Vehicle Trajectory Prediction Based on Temporal-Spatial Multi-Head Attention Mechanism[J]. Journal of Zhejiang University (Engineering Science), 2023, 57(8): 1636–1643.

(责任编辑 王 一)

修改稿收到日期为2024年10月28日。

汽车技术