



面向未知域场景的车辆轨迹预测模型

卢一凡, 李焱鹏, 薛启凡

引用本文:

卢一凡, 李焱鹏, 薛启凡. 面向未知域场景的车辆轨迹预测模型[J]. 智能系统学报, 2024, 19(5): 1238–1247.

LU Yifan, LI Xuanpeng, XUE Qifan. Vehicle trajectory prediction model for unseen domain scenarios[J]. *CAAI Transactions on Intelligent Systems*, 2024, 19(5): 1238–1247.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202306046>

您可能感兴趣的其他文章

地理位置和时间感知的表示学习框架

A geography and time aware representation learning framework

智能系统学报. 2021, 16(5): 909–917 <https://dx.doi.org/10.11992/tis.202104011>

基于孪生变分自编码器的小样本图像分类方法

A small-sample image classification method based on a Siamese variational auto-encoder

智能系统学报. 2021, 16(2): 254–262 <https://dx.doi.org/10.11992/tis.201906022>

多层卷积特征的真实场景下行人检测研究

Research on pedestrian detection based on multi-layer convolution feature in real scene

智能系统学报. 2019, 14(2): 306–315 <https://dx.doi.org/10.11992/tis.201710019>

基于动态系统的机器人模仿学习方法研究

Research on robot imitation learning method based on dynamical system

智能系统学报. 2019, 14(5): 1026–1034 <https://dx.doi.org/10.11992/tis.201807018>

隐式特征和循环神经网络的多声部音乐生成系统

A polyphony music generation system based on latent features and a recurrent neural network

智能系统学报. 2019, 14(1): 158–164 <https://dx.doi.org/10.11992/tis.201804009>

基于深度学习的视频预测研究综述

Review of deep learning-based video prediction

智能系统学报. 2018, 13(1): 85–96 <https://dx.doi.org/10.11992/tis.201707032>

DOI: 10.11992/tis.202306046

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20240828.1010.016>

面向未知域场景的车辆轨迹预测模型

卢一凡, 李煊鹏, 薛启凡

(东南大学 仪器科学与工程学院, 江苏 南京 210096)

摘要: 自动驾驶技术随着科技革新迎来蓬勃发展, 轨迹预测已成为智能汽车软件系统不可或缺的关键组成部分。为了解决传统车辆轨迹预测模型中存在的泛化能力不足的问题, 提出一种基于泛化终点预测和地图场景的车辆轨迹预测方法。该方法采用基于不变风险最小化的条件变分自编码器生成轨迹终点, 并结合时序网络编码的地图场景特征, 提升了模型预测未知域数据的准确率。在交互式道路场景数据集 INTERACTION 上的实验结果证明该模型具有良好的泛化性能。本方法与效果最好的方法 REx 相比 1、2、3 s 处的 mADE 值 (越小越好) 分别下降 0%、36.59%、50.68%, 在未知测试域的预测轨迹准确度得到显著提升。

关键词: 轨迹预测; 域泛化; 不变风险最小化; 条件变分自编码器; 端点生成; 矢量地图; 场景上下文; 时序网络
中图分类号: TP391 **文献标志码:** A **文章编号:** 1673-4785(2024)05-1238-10

中文引用格式: 卢一凡, 李煊鹏, 薛启凡. 面向未知域场景的车辆轨迹预测模型 [J]. 智能系统学报, 2024, 19(5): 1238-1247.

英文引用格式: LU Yifan, LI Xuanpeng, XUE Qifan. Vehicle trajectory prediction model for unseen domain scenarios[J]. CAAI transactions on intelligent systems, 2024, 19(5): 1238-1247.

Vehicle trajectory prediction model for unseen domain scenarios

LU Yifan, LI Xuanpeng, XUE Qifan

(School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China)

Abstract: With the rapid development of autonomous driving technology, trajectory prediction has become an essential component of smart car software systems. To address the limitations in the generalization of traditional vehicle trajectory prediction models, this study proposes a vehicle trajectory prediction method based on generalized endpoint prediction and vector maps. This method employs a conditional variational autoencoder based on invariant risk minimization to generate trajectory endpoints. The accuracy of the model in predicting unseen domains is improved by integrating map scene features encoded by a time-series network. Experiments were conducted using the interaction dataset, which includes interactive driving scenarios. Experimental results showed that, compared with the best-performing state-of-the-art (SOTA) method REx, the mADE values (where lower is better) decreased by 0%, 36.59%, and 50.68% at 1, 2, and 3 s, respectively. The accuracy of the predicted trajectories in the unseen test domain was significantly improved.

Keywords: track prediction; domain generalization; invariant risk minimization; conditional variational auto encoder; endpoint generating; vector map; scene context; time-series networks

轨迹预测是无人驾驶系统中重要的功能组成, 是车辆对附近物体运动作出反馈的前提。近年来, 深度学习模型成为解决轨迹预测问题的一

种有效方法^[1]。但传统深度学习模型是一种基于数据独立同分布假设的机器学习方法, 无法适应源域和目标域分布差异大的情况, 导致模型性能下降^[2]。车辆轨迹数据由于采集地区、采集方式、处理方法不同, 数据域的分布偏差不可避免, 故存在模型泛化能力不足问题。在实际使用场景中, 模型泛化能力弱会带来极大的安全隐患。

不变风险最小化 (invariant risk minimization,

收稿日期: 2023-06-26. 网络出版日期: 2024-08-28.

基金项目: 国家重点研发计划项目 (2021YFB1600501); 国家自然科学基金项目 (61906038); 东南大学至善青年学者项目 & 中央高校基本科研业务费专项资金项目 (2242021R41184).

通信作者: 李煊鹏. E-mail: li_xuanpeng@seu.edu.cn.

IRM) 方法是近年来提出的针对泛化性问题的有效尝试^[3], 在图像分类等任务中表现良好。但是, 将 IRM 方法直接用于解决回归性问题效果不佳, 其原因包括: 1) 车辆轨迹是时序性数据, 若 IRM 方法直接作用于编码器拟合的是所有轨迹点的分布, 训练难度大, 需要利用 IRM 方法特性设计新模块; 2) 车辆轨迹数据相比图像数据更稀疏, 需要引入场景信息提升信息维度。为解决上述问题, 本文提出一种面向未知域场景的车辆轨迹预测方法, 引入基于 IRM 方法的终点预测模块 (end-point prediction module based on invariance risk minimization, EPI) 并结合地图场景特征, 通过在 INTERACTION 公开数据集上进行实验, 验证算法有效性, 获得泛化性能的提升。本文的主要贡献包含 3 个方面:

1) 提出一种用于提升车辆轨迹预测模型泛化能力的方法, 以编码器-解码器框架为基础, 结合基于不变性原理的终点预测模块和地图信息处理模块, 提升模型预测未知数据域车辆轨迹的准确度。

2) 基于不变性原理的终点预测模块引入不变风险最小化损失函数, 作用于学习目标终点分布的条件自编码器 (conditional variational auto encoder, CVAE), 捕获以域不变特征为条件的潜在分布, 提升预测未知域车辆终点结果的准确率。

3) 提出一种提取地图上下文场景信息的方法, 通过添加长短期记忆 (long short-term memory, LSTM) 网络编码器组成的矢量地图数据处理单元, 捕捉车道的方向性, 为预测的未来轨迹提供约束, 并且采用车道中心线搜索和排序的方式, 使轨迹预测模型能够更加有效利用地图场景信息。

1 相关工作

车辆轨迹预测模型的泛化问题是一种多要素信息耦合的综合性问题。不仅需要建模车辆轨迹以及周边的环境信息, 还应采用合适的域泛化方法提升模型泛化性。相关工作包括 2 个方面, 即车辆轨迹及地图编码方法和域泛化方法。

1.1 车辆轨迹及地图编码方法

1.1.1 车辆轨迹预测方法

时序网络由于结构的特殊性, 适合轨迹性数据的预测工作。时序网络可以作为提取历史轨迹特征的编码器, 也可以作为输出层的解码器。近年来时序网络演化出多种结构: 循环神经网络^[4] (recurrent neural network, RNN)、LSTM 网络^[5]、门控循环单元^[6] (gated recurrent unit, GRU)。其中

LSTM 和 GRU 解决了 RNN 梯度容易消失或爆炸的问题^[7]。基于深度学习的轨迹预测模型最初应用于行人轨迹预测^[8]。Park 等^[9]首次利用 LSTM 编码器-解码器架构处理时序数据。但单一时序网络不能捕获多个相关序列之间的关系, 因此社交池化机制被应用于时序网络^[10]。Social-LSTM^[11]引入社交池化层建模轨迹之间的交互, 并共享相关序列特征。Deo 等^[12]提出的 CS-LSTM 在社交池化层的基础上使用卷积操作, 捕获轨迹间的相对位置特征。社交池化层丰富了车辆轨迹数据, 成为提取轨迹之间关系的途径。

图神经网络也是编码轨迹数据的一种结构^[13], 其主要思想是将环境中的每个对象都视为图的节点。每个场景可以看作一个不规则的图, 每个图包含大小可变的无序节点^[14]。Diehl 等^[15]首次将图神经网络使用在轨迹预测上, 他们使用图卷积网络和图注意力网络分别进行基于车辆交互的轨迹预测。Mohamed 等^[16]提出 Social-STGCNN 将行人与周围物体的交互行为建模为图模型, 并通过一个核函数将行人之间的社交互动嵌入邻接矩阵中。GRIP (graph-based interaction-aware trajectory prediction)^[17]、GRIP++^[18]使用无向图来表示车辆之间的交互特性, 其中图的节点代表车辆, 车辆的位置信息建模为图的节点特征, 图的边代表车辆之间的交互性。Zhao 等^[19]利用图神经网络提出了一种称为 TNT (target-driven trajectory) 的目标驱动方法, 该方法定义稀疏的候选目标点并选择到达目标的最佳轨迹。在 TNT 的基础上, Dense-TNT^[20]通过获取更稠密的候选目标, 获得比 TNT 更好的结果。由于深度学习模型过拟合影响 IRM 的效果^[21], 参数量大的网络反而不利于 IRM 方法发挥效用^[22]。

1.1.2 地图编码方法

场景上下文可以为网络预测提供先验知识。Ziegler 等^[23]将高清地图渲染应用于轨迹预测, 并通过与车辆关联的车道信息搜索地图拓扑, 从而获得其沿车道的未来轨迹。Hong 等^[24]使用卷积网络提取场景上下文特征编码地图信息。然而, 卷积网络的感受野有限, 且地图编码需要人工干预。Gao 等^[25]直接从高清地图中学习有效的上下文表示, 提出了一种基于分层图神经网络来建模车辆与车辆以及车辆与环境之间的交互特性的方法, 即 VectorNet。该方法用矢量表示各个道路组件的空间局部性, 根据上下文恢复被随机掩盖的地图实体和车辆轨迹, 增强对地图特征的捕捉。Zeng 等^[26]使用 LaneRCNN 模型编码历史轨迹和

局部地图拓扑,并通过交互模块捕捉局部车道地图的交互。

图网络编码地图的方式注重车道间的拓扑关系,但是没有强调车道的方向性以及约束轨迹时存在的时序性。

1.2 域泛化方法

提升模型域泛化能力目前主要可以通过以下 3 种方式:数据操作、策略学习、表征学习^[27]。

数据操作主要是指在模型训练前,为了缩小数据集之间的偏差对输入数据进行处理,一般为数据增强^[28]和数据生成^[29]。这种方法本质是通过完善输入数据的分布来提升模型的泛化性。

策略学习通过训练策略提升泛化能力,包括集成学习和元学习。Segu 等^[30]采用集成学习的思想,假设测试域中的任何样本都可以被视为多源域的集成样本,因此整体预测应通过在不同域上训练出的模型的组合来推断。元学习方法^[31]将来自多个域的数据分为元训练数据和元测试数据,通过模拟域转移情况学习泛化性更好的数据表示方式。

表征学习在域泛化中是涉及到模型结构设计的一类方法,一部分的研究方法注重特征解耦,即将特征解耦为域共享和域独有 2 类。另一部分注重促进模型学习域不变的表示,核心思想是通过学习域不变表征缩小域之间偏差,从而提升未知域预测轨迹的准确度。域不变表征可以通过训练域不变的特征转换核获得^[32],也可以利用对抗学习的思想,从特征处理器和域判别器对抗训练中获得^[33]。

不变风险最小化也属于表征学习,其主要思想为:只有不变因果特征是预测目标应考虑的因素,并且在不同域中保持不变。可以利用域之间特征的差异性区别不变特征与变化的特征,通过学习不变特征的条件分布提高泛化性。Ar-jovsky 等^[3]提出 IRM 方法不寻求匹配所有域的表示分布,而是在特征空间强制最佳分类器在所有域中都相同。Krueger 等^[34]提出 REx(risk extrapolation)方法通过最小化源域之间的外推风险来促使不变性特征的学习。Bayesian IRM^[21]指出导致 IRM 方法在深度学习上效果不佳的原因是过拟合,提出了用贝叶斯后验分布拟合一个分类器来增强 IRM 作用效果。Sparse IRM^[22]中提出给网络加上稀疏性约束,训练一个稀疏的网络来抑制过拟合。

对于不变风险最小化的探索一直应用在模拟数据集和图像分类问题上,所以在时序真实数据

的泛化性问题上,需要设计合适的框架结构合理应用不变性原理。

2 模型结构

本文提出一种应用于未知域的车辆轨迹预测模型 (invariance principle and vector map based vehicle trajectory prediction model, IPVM), 设计基于不变性原理的终点预测模块与地图信息处理模块解决车辆轨迹预测模型泛化能力不足的问题。

2.1 问题表述

假设 n 个不同场景下的车辆轨迹数据域 $D = \{D_1, D_2, \dots, D_n\}$, 单个数据域内的信息包括 $D_n = \{T^n, H^n, C^n\}$, 其中 D_n 中的智能体的未来轨迹 $T^n = \{Y_1^n, Y_2^n, \dots, Y_m^n\}$, 智能体的历史轨迹 $H^n = \{X_1^n, X_2^n, \dots, X_m^n\}$, 其中 m 表示智能体的个数。单条未来轨迹定义为 $Y_m^n = \{y_{m,t+1}^n, y_{m,t+2}^n, \dots, y_{m,t+\alpha}^n\}$, α 表示未来轨迹的时间步长。单条历史轨迹定义为 $X_m^n = \{x_{m,t-\beta+1}^n, x_{m,t-\beta+2}^n, \dots, x_{m,t}^n\}$, β 表示历史轨迹的时间步长。数据域 D_n 的场景上下文信息为 $C^n = \{c_1^n, c_2^n, \dots, c_\tau^n\}$, 其中 c_1^n 为 D_n 场景中的车道中心线, τ 为场景中车道中心线的个数。其中定义 c_τ^n 内坐标为 $c_\tau^n = \{(a_1^n, b_1^n), (a_2^n, b_2^n), \dots, (a_r^n, b_r^n)\}$, r 为车道中心线的节点数。方法目标是根据被预测车辆和周边车辆历史轨迹 X_m^n 以及场景上下文信息 C^n 预测未来轨迹的坐标 Y_m^n 。本文将轨迹预测拆分为 2 个子任务, 首先预测终点的可能位置 $y_{m,t+\alpha}^n$, 接着根据终点位置预测其他轨迹点 $\{y_{m,t+1}^n, y_{m,t+2}^n, \dots, y_{m,t+\alpha-1}^n\}$ 。

2.2 基于不变性原理的终点预测模块

基于不变性原理的终点预测模块的主要结构是基于不变风险最小化原理的条件变分自编码器 (CVAE), 根据历史轨迹、场景上下文信息实现对终点 $y_{t+\alpha}$ 的分布 $P_q(y_{t+\alpha}|H^1, H^2, \dots, H^n, C^1, C^2, \dots, C^n)$ 的估计。具体实现方法是构建一个隐变量 Z , 假设后验分布 $P_q(Z|y_{t+\alpha}, H^1, H^2, \dots, H^n, C^1, C^2, \dots, C^n)$ 符合正态分布, 利用本模块拟合该分布的均值 μ 和方差 σ^2 , 使 $Z \sim N(\mu, \sigma^2)$ 接近 $N(0, 1)$ 分布, 在预测场景下利用对隐变量 Z 的采样生成终点位置分布。

IRM 风险项与 CVAE 联合构成基于不变性原理的终点预测模块。若将不变特征设为 $F_s = \{H_s^1, H_s^2, \dots, H_s^n, C_s^1, C_s^2, \dots, C_s^n\}$, 变化的特征设为 $F_r = \{H_r^1, H_r^2, \dots, H_r^n, C_r^1, C_r^2, \dots, C_r^n\}$, EPI 模块的编码器以 KL(kullback-leibler) 散度为损失函数, 使拟合的 $Q_q(Z|y_{t+\alpha}, F_s, F_r)$ 分布接近 $P_q(Z|y_{t+\alpha}, F_s, F_r)$ 的分布, 解码器学习 $P_\theta(y_{t+\alpha}|Z, F_s, F_r)$ 的分布。本文加入的 IRM 惩罚项 L_{irm}^e 公式为

$$L_{\text{irm}}^e = \|\nabla_{w|w=1.0} R^e(w \cdot \phi)\|^2 \quad (1)$$

式中: w 为一个线性回归器, ϕ 为基于不变性原理的终点预测模块, R^e 是环境 e 下终点预测值与真值的偏差。IRM 惩罚项促使终点预测模块最终学习到的 $Q_q(\mathbf{Z}|\mathbf{y}_{t+\alpha}, \mathbf{F}_s, \mathbf{F}_r)$ 分布会削弱与 \mathbf{F}_r 的相关性, 拟合仅以 \mathbf{F}_s 和 $\mathbf{y}_{t+\alpha}$ 为条件的隐变量 \mathbf{Z} 的分布, 使预测结果不受变化的特征影响。

IPVM 整体结构如图 1 所示, 其中绿色框代表

传输过程中的数据张量, 蓝色框代表模型编码解码模块。模型在编码器-解码器结构的基础上添加了基于不变性原理的终点预测模块, 在信息输入维度上增加地图信息和邻近车辆信息, 使用本文提出的地图信息处理模块对地图数据进行处理, 使用卷积社会池化模块提取车辆之间的交互信息, 为最终生成轨迹提供强约束, 增强模型的泛化能力。

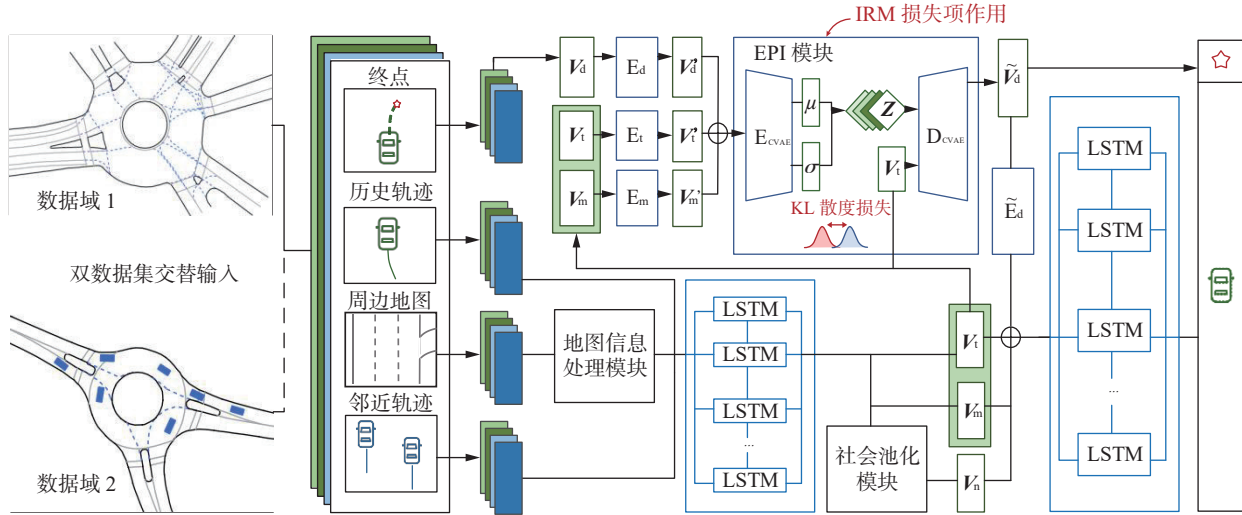


图 1 IPVM 整体结构

Fig. 1 Overall structure of IPVM

首先数据集被整理成车辆终点、历史轨迹、周边地图信息、邻近车辆轨迹 4 个部分。其中地图数据输入到地图信息处理模块中整合筛选, 再与历史轨迹、邻近车辆轨迹 3 个部分共同经过 LSTM 模块编码, 利用 LSTM 长短期记忆的特性提取时序性特征, 生成历史轨迹张量 V_t 、地图张量 V_m 和邻近车辆轨迹张量 V_n 。泛化终点预测模块需要输入地图信息、车辆终点信息以及历史轨迹信息。所以历史轨迹张量 V_t 、地图张量 V_m 、车辆终点张量 V_d 先经过各自的编码器模块 E_t 、 E_m 、 E_d 生成张量 V'_t 、 V'_m 、 V'_d , 整合成相应维度后, 再将 V'_t 、 V'_m 、 V'_d 3 个张量级联作为泛化终点预测模块的输入。在模块内部, 输入经过编码器 E_{CVAE} 拟合隐变量 $\mathbf{Z} \sim N(\mu, \sigma^2)$ 分布的均值 μ 和方差 σ^2 , 隐变量 \mathbf{Z} 与历史轨迹张量 V_t 级联输入到解码器 D_{CVAE} 中生成轨迹终点 \tilde{V}_d , \tilde{V}_d 张量组成最终预测结果的轨迹终点 $\hat{\mathbf{y}}_{m,t+\alpha}^n$ 。最后终点张量 \tilde{V}_d 经过编码器 \tilde{E}_d 编码, \tilde{E}_d 的输出与其他 3 部分历史轨迹张量 V_t 、地图张量 V_m 和邻近车辆轨迹张量 V_n 级联后输入到 LSTM 解码模块, 生成除终点外的其余轨迹。

训练过程中为使 IRM 惩罚项发挥作用, 需要至少 2 个不同域的数据, 交替输入模型如图 1 所

示, 并分别计算出对应数据域的 IRM 惩罚项, 最后对惩罚项求平均后反向传播。测试过程中由于车辆终点不再是已知信息, 所以跳过编码器生成和的步骤, 直接对隐变量 \mathbf{Z} 进行采样, 与历史轨迹张量 V_t 级联输入到解码器 D_{CVAE} 中生成未来终点 \tilde{V}_d 。

2.3 地图信息处理模块

车辆的未来轨迹与道路结构密切相关^[35]。本文处理场景上下文中的地图信息采用矢量地图抽取邻近道路中心线维度的信息。这种处理有利于编码器捕捉车道方向。在处理信息过程中由于矢量地图会出现左右车道线方向不一致的情况, 所以加入车道方向修正步骤; 最后, 按距离排序筛选出离车辆历史轨迹近的中心线。本文认为行驶方向的车道中心线在约束车辆轨迹时也具备时序特征, 所以将地图信息最终输入到 LSTM 编码器中提取时序性特征。

地图信息处理模块由 4 个步骤组成, 如图 2 所示。步骤 1 是提取地图信息。大部分矢量地图是以树的结构在文件中存储, 存储的内容包括各组件属性信息以及与其组件的关系。节点组件保存各节点的经纬度坐标。道路组件保存了组成道路车道线的节点以及道路属性, 如车速、

类型、颜色等等。关系 (relation) 组件利用道路组件描述地图中成员之间的联系, 例如 2 个道路组

件即可组成机动车行驶车道。图 2 左下方所示为矢量图可视化的效果。

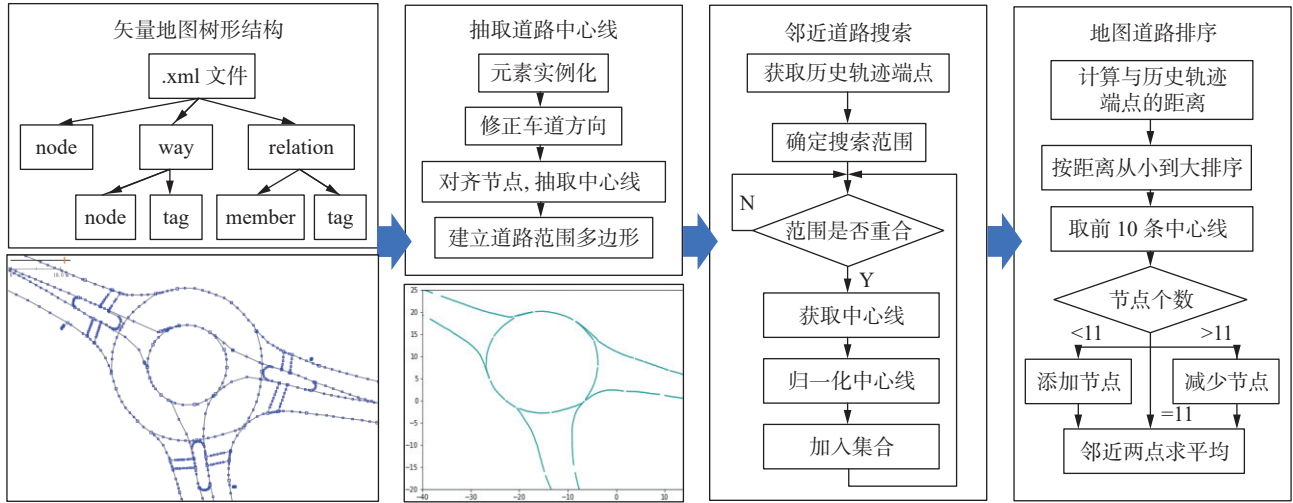


图 2 地图信息处理模块流程

Fig. 2 Map information processing flow chart

步骤 2 是抽取道路中心线。首先需要将地图的组件均实例化, 统一节点坐标与车辆轨迹坐标的坐标系。由于矢量地图会出现车道线方向不一致的情况, 扰乱后续计算道路中心线的步骤, 需要判断两车道线方向是否一致, 依情况将左车道线反向。另外车道的左右节点个数差异也会导致中心线节点计算错误。针对这个问题, 本模块将节点个数少的车道根据个数多的车道补充新节点。一次样条插值计算新节点的坐标的公式为

$$f[k_u, k_{u+1}] = \frac{j_{u+1} - j_u}{k_{u+1} - k_u}$$

$$L_{1,u}(k) = j_u + f[k_u, k_{u+1}](k - k_u)$$

$$\tilde{L}_1(k) = \begin{cases} L_{1,1}(k), k \in [k_1, k_2] \\ L_{1,2}(k), k \in [k_2, k_3] \\ \vdots \\ L_{1,u-1}(k), k \in [k_{u-1}, k_u] \end{cases}$$

式中: (k_u, j_u) 为矢量地图中的道路组件, 取中线后可得车道 c_r^n 中心线坐标 (a_u, b_u) 。最后根据机动车道的坐标范围, 构建道路范围多边形以供道路搜索模块使用, 道路多边形构建方式的公式为

$$S_r^n = \{\min(j_u), \min(k_u), \max(j_u), \max(k_u)\}$$

步骤 3 是邻近道路搜索。先根据当前车辆历史轨迹的端点, 以及设置的搜索范围超参数, 确定搜索道路的坐标范围。再依据搜索范围遍历道路范围多边形 S_r^n , 将有范围重合的道路中心线以历史轨迹端点为基准实现坐标归一化。最后将归一化的中心线加入到集合中。

步骤 4 是地图道路排序。遍历步骤 3 中归一化后的中心线集合。中心线按照地图相关距离 d 从小到大排序, 地图相关距离 d 定义为中心线起始点和末尾点分别于历史轨迹端点的欧氏距离,

由于坐标事先归一化, 历史轨迹端点即为原点

$$d = \sqrt{a_1^2 + b_1^2} + \sqrt{a_u^2 + b_u^2}$$

为了防止中心线长度不一致形成的末尾补零影响 LSTM 编码的效果, 采取对齐中心线节点个数的措施。最后将相邻两点求取平均值, 减少部分节点异常变化的影响。

2.4 社会池化模块

场景中其他动态交通参与者是影响被预测车辆未来轨迹的重要因素^[36]。本模型的社会池化模块旨在提取被预测车辆周边邻近车辆轨迹的交互特征。使用与 CS-LSTM^[12] 类似的卷积池化结构, 如图 3 所示, 池化模块通过设置空间网格结构组织社会池化张量 V_n 。依照邻近车辆所在的网格位置, 使用历史轨迹的编码填充网格。本模型将社会池化张量经过 2 层卷积层和 1 层池化层。卷积模块有利于捕捉邻近车辆之间的位置关系, 最大池化层可以增加局部平移不变性^[12]。为了更契合复杂交通环境交互特征的情况, 本模型将原本 3×13 的空间网格扩张为 13×13 , 扩大针对临近车辆的感受野。

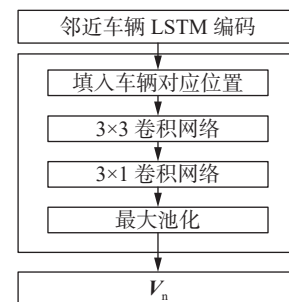


图 3 社会池化模块框架

Fig. 3 Structure of social pooling module

2.5 损失函数

本模型的损失函数由4个部分组成。基于不变性原理的终点预测模块包含3个损失项:预测重点目标的终点偏差损失 L_{dest}^e 、KL散度损失 L_{KL}^e 以及不变因果最小化损失 L_{irm}^e 。 L_{irm}^e 如式(1)所示。重构损失为终点预测模块矫正预测终点的准确度,计算终点预测值 $\hat{y}_{m,t+\alpha}^n$ 与真值 $y_{m,t+\alpha}^n$ 的 l_2 范数。KL散度损失用于约束隐变量 \mathbf{Z} 的分布 $N(\mu, \sigma^2)$ 更接近 $N(0,1)$ 分布。IRM损失促进终点预测模块削弱与变化的非因果特征的相关性。

$$L_{\text{dest}}^e = \frac{1}{m} \sum_{i \in m} \|\mathbf{y}_{i,t+\alpha}^e - \hat{\mathbf{y}}_{i,t+\alpha}^e\|_2, e \in n$$

$$L_{\text{KL}}^e = \text{KL}(N(\mu, \sigma^2) \| N(0, 1)) =$$

$$\frac{1}{2}(-\ln(\sigma^2) + \mu^2 + \sigma^2 - 1), e \in n$$

剩余轨迹点生成引出第4个损失项轨迹偏差损失 L_{rmse}^e 。轨迹偏差损失 L_{rmse}^e 用于衡量预测轨迹的准确性,校正轨迹预测结果,计算轨迹预测值与真实值的 l_2 范数,其计算公式为

$$L_{\text{rmse}}^e = \frac{1}{m} \sum_{i \in m} \|\mathbf{y}_{i,t+1}^e, \mathbf{y}_{i,t+2}^e, \dots, \mathbf{y}_{i,t+\alpha-1}^e -$$

$$(\hat{\mathbf{y}}_{i,t+1}^e, \hat{\mathbf{y}}_{i,t+2}^e, \dots, \hat{\mathbf{y}}_{i,t+\alpha-1}^e)\|_2, e \in n$$

由于训练过程中需要分别计算各个数据域的损失,求平均后共同反向传播,所以总损失函数 L 的计算公式为

$$L = \frac{1}{n} \sum_{e \in n} (\gamma L_{\text{dest}}^e + \delta L_{\text{KL}}^e + \eta L_{\text{rmse}}^e) + \lambda L_{\text{irm}}$$

式中相关超参数本文设置为 $\gamma = \delta = \eta = 1$ 。 λ 为训练过程中调整IRM损失作用强弱的超参数。

3 实验分析

3.1 数据集与评价指标

本模型的实验部分使用公开数据集 INTER-ACTION^[37],其包含4种不同道路场景,即环岛、无信号交叉路口、有信号交叉路口以及会车与变道,总共11个地图场景。数据采样频率为10 Hz,轨迹长度基本大于5 s。此数据集是具有高清矢量地图的交互式驾驶场景中的轨迹数据。本文实验在环岛、交叉路口2种复杂场景中进行2组实验,定义每个地图场景为1个数据域,选取1个作为测试集,其余作为训练集,测试域的选择根据轨迹数据量决定。测试集数据在训练过程中对模型不可见。因此环岛场景中使用FT地图数据和EPR地图数据训练的模型预测OF地图数据,交叉路口场景中使用MA地图数据和GL地图数据训练的模型预测EP地图数据。针对每个数据域,本文参考文献[19]将轨迹分割为5 s的片段,

包括2 s的历史轨迹和3 s的未来轨迹。车辆轨迹原始采样频率为10 Hz,本文为平滑真实轨迹并降低模型参数量,将轨迹下采样,下采样系数为2,即历史轨迹步长 $\beta=10$,未来轨迹步长 $\alpha=15$ 。

实验评价指标使用最小平均位移误差(minimum average displacement error, mADE)衡量车辆轨迹模型的泛化能力。mADE定义为生成的 k 个终点与真值插值最小的那条预测轨迹与真值轨迹点的平均 l_2 范数公式为

$$L_{\text{mADE}} = \min \left\{ \frac{\sum_{i \in m} \sum_{j=1}^{\alpha} \|\mathbf{y}_{m,t+j}^n - \hat{\mathbf{y}}_{m,t+j}^n\|_2}{m \times \alpha} \right\}_k \quad (2)$$

在实验网络参数的设置方面,学习率设置为0.001,优化器使用Adam优化器,所有实验统一训练25个回合,IRM损失项的系数 λ 取3 100。

3.2 定量结果与分析

本文首先针对IPVM对不变风险最小化方法、终点预测模块以及地图信息处理模块3个部分进行消融研究。接着将IPVM与域泛化的方法进行横向比较。实验指标为根据式(2)计算的mADE值。

3.2.1 消融实验

为验证IPVM各模块的性能,采用消融实验评估IRM损失、终点预测模块以及地图处理模块的作用。

2组实验的结果如表1~2所示,表中给出模型在预测时间1~3 s的mADE值,评价指标以m为单位。“√”代表使用该模块,“×”代表不使用该模块。实验结果表明各模块对模型预测车辆的准确度的提升均有积极贡献。

表1 FT+EPR预测OF的消融实验
Table 1 Prediction of ablation experiment of OF by FT+EPR

IRM	CVAE	MAP	1 s	2 s	3 s
×	√	√	0.23	0.56	1.16
√	×	√	0.23	0.62	1.39
√	√	×	0.28	0.53	0.86
√	√	√	0.18	0.41	0.72

表2 MA+GL预测EP的消融实验
Table 2 Prediction of ablation experiment of EP by MA+GL

IRM	CVAE	MAP	1 s	2 s	3 s
×	√	√	0.18	0.32	0.54
√	×	√	0.16	0.38	0.8
√	√	×	0.11	0.23	0.38
√	√	√	0.12	0.21	0.34

IRM 损失有效性 引入 IRM 框架有助于测试集上 1、2、3 s 的预测准确度提升。IRM 框架的使用促使模型内的编码器学习不变因果特征的分布,有助于模型泛化能力的提升。

终点预测模块有效性 引入终点预测模块对测试结果提升幅度最大。终点预测的机制模拟被预测车辆决策行驶方向的过程,并且防止预测误差随时间积累的问题。

地图信息模块有效性 引入地图信息处理模块的提升最小。地图信息的加入目的在于约束未来车辆轨迹,使车辆不仅考虑邻近车辆交互,还考虑场景地图的走势。但对地图处理过程中无法避免尖锐拐点、采样点疏密不一等问题,一定程度影响地图模块的最终效果。

3.2.2 定量结果分析

本节选择以下域泛化方法作为对比方法: 1) ERM^[38]: 在编码器-解码器架构上,应用传统平均风险最小化 ERM 的方法; 2) DANN^[39]: 使用的对抗学习的方法促进特征提取器的域不变性,特征提取器选择 CS-LSTM 网络; 3) MMD^[33]: 使用最大均值差异原理促使不同域特征提取器的分布接近,特征提取器同样选择 CS-LSTM 网络; 4) IRM^[3]: IRM 方法作用于整个骨干网络而不仅使用在终点预测模块中; 5) REx^[34]: 以最小化源域之间的外推风险的方式促进不变因果特征分布的学习; 6) IPVM: 本文提出的模型结构,包括泛化终点预测模块与地图信息处理模块。上述 6 种方法在数据集上 mADE 的实验结果如表 3~4 所示,评价指标以 m 为单位。

表 3 FT+EPR 预测 OF 的 mADE 实验结果
Table 3 Prediction of mADE performance of OF by FT+EPR m

时间/s	ERM	MMD	DANN	IRM	REx	IPVM
1	0.32	0.15	0.40	0.22	0.18	0.18
2	0.73	0.55	0.84	0.58	0.67	0.41
3	1.51	1.31	1.71	1.30	1.46	0.72

表 4 MA+GL 预测 EP 的 mADE 实验结果
Table 4 Prediction of mADE performance of EP by MA+GL m

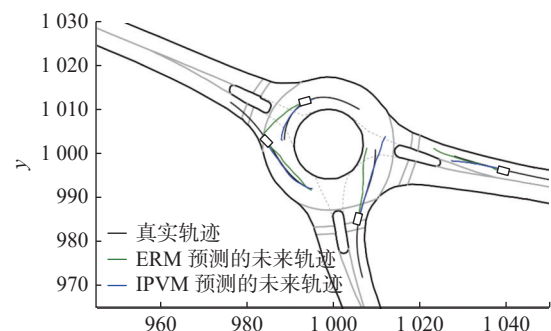
时间/s	ERM	MMD	DANN	IRM	REx	IPVM
1	0.19	0.13	0.24	0.24	0.12	0.12
2	0.39	0.32	0.56	0.45	0.32	0.21
3	0.78	0.70	1.16	0.85	0.73	0.36

根据表 3~4 的结果,首先 MMD 和 REx 2 种泛化方法在 2 组实验中的效果均比 ERM 方法优秀。而 IRM 方法的表现并不稳定,说明 IRM

方法直接作用于深度学习网络存在局限性。DANN 方法的结果最差,说明对抗学习在车辆轨迹预测问题上相与其余对比模型相比没有更好表现。分析原因可知车辆轨迹的域标签特征并不明确,导致域判别器无法促使特征提取器生成域不变分布。IPVM 模型的实验结果明显优于其余对比基线,与效果最好的方法 REx 相比 1、2、3 s 处的 mADE 值分别下降 0%、36.59%、50.68%,虽然相比 REx 方法 IPVM 在 1 s 处没有提升,但其在 2、3 s 处提升效果明显。实验结果证明本文提出的基于不变性原理的终点预测模块与地图信息处理模块可以带来优化效果。本方法与 IRM 作用于整体网络相比,需要学习的分布从未来轨迹所有点的分布简化为轨迹终点分布,训练的难度减小有助于 IRM 方法产生效用。另外模型通过联系场景信息丰富轨迹数据,为预测轨迹提供有效约束也有助于泛化性的提升。

3.2.3 定性结果分析

本文提出的模型旨在提升车辆轨迹预测模型的域泛化性,如图 4 所示为 OF 和 EP 2 个测试集的车辆轨迹预测结果,其中黑色方框为车辆当前位置,黑色轨迹为车辆真实轨迹,绿色轨迹为不考虑泛化性的 ERM 模型的预测结果,蓝色轨迹为本文模型的预测结果。从图 4 可以看出 IPVM 的预测结果明显优于 ERM,说明本文采取的措施有助于模型泛化性的提升。虽然 ERM 模型预测直线行驶的效果和 IPVM 相差不多,但在交叉路口和环形路口等复杂场景中,ERM 的效果会偏离航线。从图 4(b)、(d) 中可以看出 ERM 模型的预测结果已经延伸至不可行驶区域,而 IPVM 模型由于加入了场景地图信息约束轨迹避开场景边界;从图 4(a)、(c) 可以发现由于使用 LSTM 编码地图信息,考虑车道的方向性,避免了轨迹如 ERM 结果驶向逆行的车道。但是通过图 4(c) 仍可以发现,2 种模型均无法捕捉车辆减速的情况,预测轨迹均超过车辆停止线,所以模型捕捉车辆减速停车信息方面仍然存在不足。



(a) OF 地图场景下可视化结果 (时刻 1)

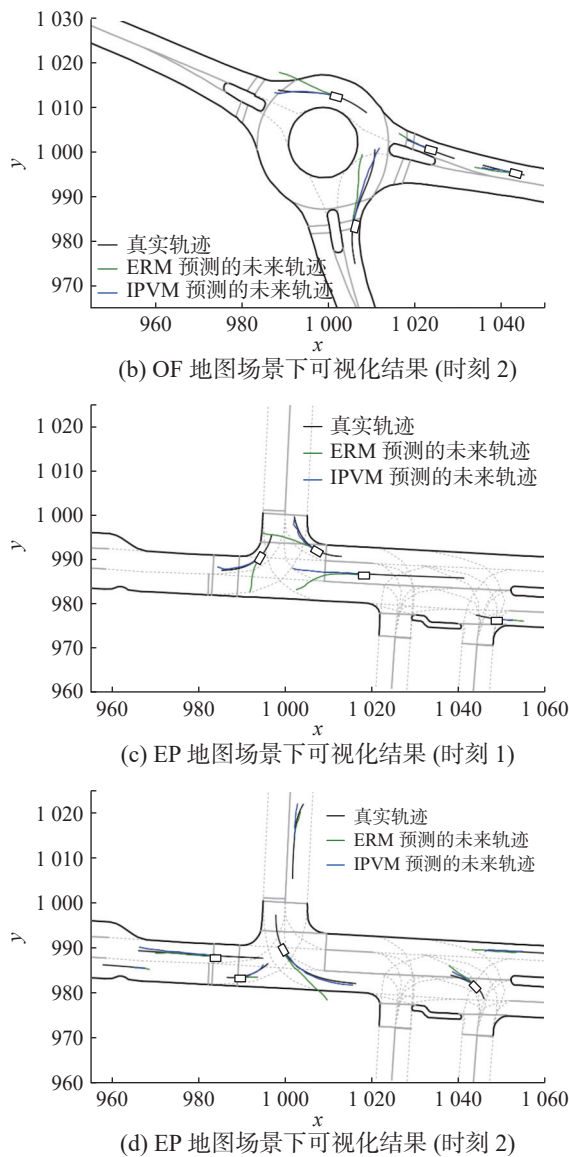


图 4 测试集车辆轨迹可视化

Fig. 4 Visualization of vehicle trajectory in test set

4 结束语

本文提出一种加入基于不变性原理的终点预测模块和地图信息的车辆轨迹预测框架,旨在解决车辆轨迹预测模型域泛化性差的问题。基于不变性原理的终点预测模块在传统 CVAE 模块上添加 IRM 损失项,促使 CVAE 学习的条件分布增强与不变特征的相关性,地图信息处理模块采用与传统方式不同的 LSTM 编码器处理车道中心线的方法,提升模型捕捉场景信息的能力。根据在 INTERACTION 数据集上的实验结果,本文提出的 IPVM 模型相比其他基线模型提高了未来轨迹预测的准确度。未来工作将尝试结合现有模型解决不变因果特征本身存在的分布偏差问题,进一步提升模型域泛化能力。

参考文献:

- [1] HUANG Yanjun, DU Jiatong, YANG Ziru, et al. A survey on trajectory-prediction methods for autonomous driving[J]. *IEEE transactions on intelligent vehicles*, 2022, 7(3): 652–674.
- [2] 王翎, 孙涵. 基于域适应的多场景车辆检测[J]. *计算机技术与发展*, 2019, 29(12): 158–161, 166.
WANG Ling, SUN Han. Vehicle detection in different scenes based on domain adaptation[J]. *Computer technology and development*, 2019, 29(12): 158–161, 166.
- [3] ARJOVSKY M, BOTTOU L, GULRAJANI I, et al. Invariant risk minimization[EB/OL]. (2019–07–05) [2023–06–26]. <https://arxiv.org/abs/1907.02893>.
- [4] ELMAN J L. Finding structure in time[J]. *Cognitive science*, 1990, 14(2): 179–211.
- [5] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. *Neural computation*, 1997, 9(8): 1735–1780.
- [6] CHO K, VAN MERRIENBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. Doha: ACL, 2014: 1724–1734.
- [7] 柴进. 基于聚类和 LSTM 算法的车辆轨迹预测模型研究[D]. 北京: 北京交通大学, 2020.
CHAI Jin. On vehicle trajectory prediction model based on combined clustering and LSTM algorithm[D]. Beijing: Beijing Jiaotong University, 2020.
- [8] 孔玮, 刘云, 李辉, 等. 基于深度学习的行人轨迹预测方法综述[J]. *控制与决策*, 2021, 36(12): 2841–2850.
KONG Wei, LIU Yun, LI Hui, et al. Survey of pedestrian trajectory prediction methods based on deep learning[J]. *Control and decision*, 2021, 36(12): 2841–2850.
- [9] PARK S H, KIM B, KANG C M, et al. Sequence-to-sequence prediction of vehicle trajectory via LSTM encoder-decoder architecture[C]//2018 IEEE Intelligent Vehicles Symposium. Changshu: IEEE, 2018: 1672–1678.
- [10] 赵靖文, 李煊鹏, 张为公. 车辆多目标交互行为建模的轨迹预测方法[J]. *智能系统学报*, 2023, 18(3): 480–488.
ZHAO Jingwen, LI Xuanpeng, ZHANG Weigong. Vehicle trajectory prediction method based on modeling of multi-agent interaction behavior[J]. *CAAI transactions on intelligent systems*, 2023, 18(3): 480–488.
- [11] 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. Las Vegas: IEEE, 2016: 961–971.
- [12] DEO N, TRIVEDI M M. Convolutional social pooling for

- vehicle trajectory prediction[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City: IEEE, 2018: 1468–1476.
- [13] 曾伟良, 陈漪皓, 姚若愚, 等. 时空图注意力网络在交叉口车辆轨迹预测的应用[J]. *计算机科学*, 2021, 48(S1): 334–341.
ZENG Weiliang, CHEN Yihao, YAO Ruoyu, et al. Application of spatial-temporal graph attention networks in trajectory prediction for vehicles at intersections[J]. *Computer science*, 2021, 48(S1): 334–341.
- [14] 赵靖文. 基于改进 Transformer 模型的车辆轨迹预测研究[D]. 南京: 东南大学, 2022.
ZHAO Jingwen. Research on vehicle trajectory prediction based on the improved transformer model. Nanjing: Southeast University, 2022.
- [15] DIEHL F, BRUNNER T, LE M T, et al. Graph neural networks for modelling traffic participant interaction[C]//2019 IEEE Intelligent Vehicles Symposium. Paris: IEEE, 2019: 695–701.
- [16] MOHAMED A, QIAN Kun, ELHOSEINY M, et al. Social-STGCNN: a social spatio-temporal graph convolutional neural network for human trajectory prediction[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 14412–14420.
- [17] LI Xin, YING Xiaowen, CHUAH M C. GRIP: graph-based interaction-aware trajectory prediction[C]//2019 IEEE Intelligent Transportation Systems Conference. Auckland: IEEE, 2019: 3960–3966.
- [18] LI Xin, YING Xiaowen, CHUAH M C. GRIP++: enhanced graph-based interaction-aware trajectory prediction for autonomous driving[EB/OL]. (2019–07–17) [2023–06–26]. <https://arxiv.org/abs/1907.07792>.
- [19] ZHAO Hang, GAO Jiyang, LAN Tian, et al. TNT: target-driven trajectory prediction[EB/OL]. (2020–08–19) [2023–06–26]. <https://arxiv.org/abs/2008.08294>.
- [20] GU Junru, SUN Chen, ZHAO Hang. DenseTNT: end-to-end trajectory prediction from dense goal sets[C]//2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021: 15283–15292.
- [21] LIN Yong, DONG Hanze, WANG Hao, et al. Bayesian invariant risk minimization[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022: 16000–16009.
- [22] ZHOU Xiao, LIN Yong, ZHANG Weizhong, et al. Sparse invariant risk minimization[C]//Proceedings of the 39th International Conference on Machine Learning. Baltimore: PMLR, 2022: 27222–27244.
- [23] ZIEGLER J, BENDER P, SCHREIBER M, et al. Making bertha drive—an autonomous journey on a historic route[J]. *IEEE intelligent transportation systems magazine*, 2014, 6(2): 8–20.
- [24] HONG J, SAPP B, PHILBIN J. Rules of the road: predicting driving behavior with a convolutional model of semantic interactions[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 8454–8462.
- [25] GAO Jiyang, SUN Chen, ZHAO Hang, et al. VectorNet: encoding HD maps and agent dynamics from vectorized representation[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 11522–11530.
- [26] ZENG Wenyuan, LIANG Ming, LIAO Renjie, et al. LaneRCNN: distributed representations for graph-centric motion forecasting[C]//2021 IEEE/RSJ International Conference on Intelligent Robots and Systems. Prague: IEEE, 2021: 532–539.
- [27] WANG Jindong, LAN Cuiling, LIU Chang, et al. Generalizing to unseen domains: a survey on domain generalization[J]. *IEEE transactions on knowledge and data engineering*, 2023, 35(8): 8052–8072.
- [28] PENG Xi, QIAO Fengchun, ZHAO Long. Out-of-domain generalization from a single source: an uncertainty quantification approach[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2024, 46(3): 1775–1787.
- [29] SOMAVARAPU N, MA C Y, KIRA Z. Frustratingly simple domain generalization via image stylization[EB/OL]. (2020–06–19) [2023–06–26]. <https://arxiv.org/abs/2006.11207>.
- [30] SEGU M, TONIONI A, TOMBARI F. Batch normalization embeddings for deep domain generalization[J]. *Pattern recognition*, 2023, 135: 109115.
- [31] LI Da, YANG Yongxin, SONG Yizhe, et al. Learning to generalize: meta-learning for domain generalization[J]. *Proceedings of the AAAI conference on artificial intelligence*, 2018, 32(1): 3490–3497.
- [32] GAN Chuang, YANG Tianbao, GONG Boqing. Learning attributes equals multi-source domain generalization[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 87–97.
- [33] LI Haoliang, PAN S J, WANG Shiqi, et al. Domain generalization with adversarial feature learning[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 5400–5409.
- [34] KRUEGER D, CABALLERO E, JACOBSEN J H, et al.

Out-of-distribution generalization *via* risk extrapolation (REx)[EB/OL]. (2020-03-02) [2023-06-26]. <https://arxiv.org/abs/2003.00688>.

- [35] 侯廉. 面向自动驾驶应用的周车轨迹群体交互式预测 [D]. 北京: 清华大学, 2019.

HOU Lian. Distributed and interactive prediction of future trajectories of surrounding vehicles for autonomous driving. Beijing: Tsinghua University, 2019.

- [36] 王少博. 动态场景下基于交互性预测的自动驾驶汽车轨迹规划方法研究 [D]. 合肥: 中国科学技术大学, 2020.

WANG Shaobo. Research on interaction-aware predictive trajectory planning for autonomous vehicle in dynamic environments. Hefei: University of Science and Technology of China, 2020.

- [37] ZHAN Wei, SUN Liting, WANG Di, et al. Interaction dataset: an international, adversarial and cooperative motion dataset in interactive driving scenarios with semantic maps[EB/OL]. (2019-10-30) [2023-06-26]. <https://arxiv.org/abs/1910.03088>.

- [38] VAPNIK V N. Statistical learning theory[M]. New York: Wiley, 1998.

- [39] GANIN Y, USTINOVA E, AJAKAN H, et al. Domain-

adversarial training of neural networks[J]. Journal of machine learning research, 2016, 17(1): 2096-2030.

作者简介:



卢一凡, 硕士研究生, 主要研究方向为道路场景下的模型领域泛化。
E-mail: 18502517268@163.com。



李煊鹏, 副教授, 博士, 主要研究方向为道路场景下的环境感知、行为预测及因果推理。主持和参与国家级及省部级项目 10 余项, 授权发明专利 7 项, 发表学术论文 30 余篇。E-mail: li_xuanpeng@seu.edu.cn。



薛启凡, 博士研究生, 主要研究方向为自动驾驶领域的轨迹预测和模型领域泛化。授权发明专利 4 项, 发表学术论文 5 篇。E-mail: xue_qifan@seu.edu.cn。