

分级包络对抗域适应和松-紧耦合行人轨迹 预测模型

李勇明¹, 胡 杰¹, 张小恒², 王 品¹, 李文正¹

(1. 重庆大学微电子与通信工程学院, 重庆 400044; 2. 重庆开放大学, 重庆 400044)

摘 要: 行人轨迹预测在自动驾驶、智能安防及智慧城市等领域有着广泛用途, 但是行人交互的复杂性和不确定性使得行人轨迹预测任务至今仍是一个充满挑战的课题. 现有行人轨迹预测模型存在以下共性局限性: (1) 仅考虑单一的社交耦合关系, 不但引入了冗余交互关系, 而且未能充分考虑不同场景下轨迹社交耦合关系不同以及行人间耦合关系不一的特性, 从而限制了模型对不同场景特征的深度挖掘与有效利用; (2) 未充分考虑域偏移问题, 极少数方法虽然考虑了域偏移问题, 但是采用基于统计准则的域分布对齐方式, 对预定义统计度量具有高度依赖性, 对复杂多变场景表现出明显的局限性. 为了解决上述问题, 本文提出了一种基于分级包络对抗域适应和松-紧耦合的行人轨迹预测模型. 首先, 设计包络样本变换机制构造包络样本并将其扩展到图结构; 其次, 结合局部域适应和全局域适应, 设计对抗域适应模块; 此外, 针对不同场景下不同耦合关系, 构建松-紧耦合包络样本构造模块. 实验部分采用了两个代表性的公共数据集进行有效性验证, 并与 6 个相关代表性算法模型进行综合对比. 实验结果表明, 本文模型比相关算法显著具有更高准确性, ADE (Average Displacement Error) 指标与 FDE (Final Displacement Error) 指标分别下降了 17.6% 和 19.1%, 时间开销满足实际需要, 这说明本文主要创新点有效.

关键词: 行人轨迹预测; 社交交互; 对抗域适应; 包络样本变换; 松-紧耦合

基金项目: 国家自然科学基金 (No.U21A20448, No.61771080, No.72001032)

中图分类号: TP391.4

文献标识码: A

文章编号: 0372-2112(2025)10-3640-19

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20250729

Hierarchical Envelope Adversarial Domain Adaptation with Loose-Tight Coupled Trajectory Prediction Model

LI Yong-ming¹, HU Jie¹, ZHANG Xiao-heng², WANG Pin¹, LI Wen-zheng¹

(1. School of Microelectronics and Communication Engineering, Chongqing University, Chongqing 400044, China;

2. Chongqing Open University, Chongqing 400044, China)

Abstract: Pedestrian trajectory prediction holds significant applications across autonomous driving, intelligent surveillance, and smart cities. However, the complexity and unpredictability of pedestrian interactions make this task a persistent challenge. Current models face two common limitations: (1) Considering only a single type of social coupling. This introduces redundant interactions. More critically, it fails to account for the varying nature of trajectory coupling across different scenarios and between different pedestrians. Consequently, models cannot deeply explore or effectively utilize diverse scene features; (2) Inadequate handling of domain shift. Although very few methods address domain shift, they rely on statistical criterion-based domain distribution alignment. Such approaches exhibit strong dependency on predefined statistical metrics. This leads to significant limitations in complex, dynamic environments. To address these issues, this paper proposes a hierarchical envelope adversarial domain adaptation with loose-tight coupled model. Firstly, an envelope sample transformation mechanism was designed. It constructs envelope samples and extends them into graph structures; Secondly, an adversarial domain adaptation module was developed. This integrates both local and global domain adaptation strategies; meanwhile, a loose-tight coupling envelope sample construction module was created. It dynamically adapts to diverse coupling relationships across scenarios. These innovations collectively enhance prediction accuracy and robustness. The experimental section employed two representative public datasets for validation and conducted comprehensive comparisons with six rele-

vant baseline algorithms. Results demonstrate that our model achieves significantly higher accuracy compared to existing methods, with the average displacement error (ADE) and final displacement error (FDE) metrics reduced by 17.6% and 19.1%, respectively. The time overhead meets practical requirements, which verifies the effectiveness of our key innovations.

Key words: pedestrian trajectory prediction; social interaction; adversarial domain adaptation; envelope sample transformation; loose-tight coupled

Foundation Item(s): National Natural Science Foundation of China (No.U21A20448, No.61771080, No.72001032)

1 引言

行人轨迹预测是指通过技术手段分析并预判行人未来的移动路径。行人轨迹预测模型是至今研究热点之一。复杂环境下的行人轨迹预测任务受多种因素影响^[1],包括行人的自身属性^[2]和行人间的交互^[3]。行人轨迹具有群体交互特性,个体轨迹不仅受自身运动意图支配,更受到周边行人轨迹不同程度的持续影响,因此行人的社交耦合关系存在异质性。早期研究中占据主导地位的为社会力模型与等速模型等知识模型体系^[4,5],虽具备数学可解释性,但在刻画多智能体非线性交互等方面存在显著瓶颈。随着深度学习技术的快速发展,行人轨迹预测领域出现了显著转向,基于深度学习的预测模型通过捕捉时序特征展现出优越的性能表现^[6-8]。但是,这些模型大多忽略了不同轨迹域间存在分布差异的问题。Zhang 等人^[9]用空间、时间和光谱域信息表示更通用行人运动模式来减小从源数据集中学习知识并将其转移到目标数据集的差距。Zhi 等人^[10]和 Liu 等人^[11]使用最大均值差异(Maximum Mean Discrepancy, MMD)损失函数对齐两个域之间的差异。Poi-breński 等人^[12]使用基于伪标签的域适应方法,以监督方式同时使用标记和未标记的数据进行训练。虽然这些研究考虑了域分布差异问题,但仍存在以下主要局限性:(1)仅考虑单一的社交耦合关系,忽视了场景内行人的社交耦合关系的异质性,这种单一化的建模方式不仅导致冗余交互关系的引入,还未能充分考虑不同场景下轨迹社交耦合关系的特性,从而限制了模型对场景特征的深度挖掘与有效利用;(2)域偏移问题考虑不充分;(3)仅考虑行人个体这一单一粒度。

为了解决上述问题,本文前期作了初步探索,提出了基于分级包络域适应的行人轨迹预测模型(Multi-level Envelope Sample Construction-Hierarchical Envelope Domain Adaptation, MESC-HEDA)^[13],取得了显著改进效果。但是,该方法仍然没有解决单一社交耦合关系问题,域偏移问题也依赖于预定义的统计度量,这限制了模型的泛化能力。针对这些不足,本文进一步做如下创新工作:

(1)构建松-紧耦合包络样本构造模块。针对行人的社交耦合关系存在异质性的问题,基于包络样本变换机制,设计松-紧耦合相结合的混合耦合方式,构建反映不

同耦合关系的包络样本用于建模,再通过决策融合进行轨迹预测。使得模型能够更好地适配各类不同的复杂场景,提高模型在稀疏场景与密集场景的预测准确性。

(2)构建对抗域适应模块。该类域适应方法能灵活地学习复杂场景多变下的域不变特征,避免人为设计距离度量缺乏自适应性的缺陷;通过生成器和判别器的博弈,从而学习到更具区分性和鲁棒性的动态特征表示。文末进行多组实验验证本文模型的性能,包括消融实验、参数分析、算法对比、轨迹可视化、特征分布可视化等。本文工作将现有“基于单一社交耦合关系的域适应行人轨迹预测模型”转化为“基于包络对抗域适应的松-紧耦合行人轨迹预测模型”,对相关研究具有一定的参考价值。

2 相关工作综述

2.1 非域适应的行人轨迹预测模型

现有大部分采用基于非域适应的模型预测行人轨迹。传统方法有社会力模型^[14]、卡尔曼滤波器和扩展卡尔曼滤波器模型^[15]等。但面对智能体交互复杂度提升,物理与统计模型的表征能力受限。深度学习技术的迅猛发展,催生了众多处理时序数据的模型,这些新兴模型能够有效地提取丰富的时空特征,同时建模行人间的复杂交互关系,在处理非线性数据上具有显著优势。目前,基于深度学习的行人轨迹预测模型主要基于以下3种架构:循环神经网络(Recurrent Neural Network, RNN)、生成对抗网络(Generative Adversarial Network, GAN)和图神经网络(Graph Neural Network, GNN)。

RNN方面,Alahi 等人^[16]提出了社交长短期记忆网络(Social Long Short-Term Memory, Social-LSTM)首次表征行人交互。Xue 等人^[17]采用3个不同尺度LSTM编码器对个体、邻近和场景信息进行编码。Hug 等人^[18]提出一种混合密度层结合粒子滤波的LSTM多模型行人轨迹预测模型。Bisagno 等人^[19]在社会交互建模框架中更强调行人群体分组的关键性特征。Yao 等人^[20]提出了一种基于门控循环单元的编码器-解码器模型。但RNN方法递归的计算方式使网络难以实现并行化运算,制约了模型训练的效率与预测的时效性。

GAN方面,Gupta 等人^[21]用改进多样性损失与最大池化分析交互。Sadeghian 等人^[22]结合注意力机制聚合智能体信息。Amirian 等人^[23]引入额外信息约束样本并

最大化生成器与编码器间的互信息来缓解 Social-GAN 模型训练不稳定的问题. Jun 等人^[24]整合个体与群体信息,联合训练实现高精度多样化预测. 但 GAN 方法存在训练易不稳定(如模式坍塌)、耗时耗资源、可解释性差等问题.

GNN 方面, Mohamed 等人^[25]提出了社交时空图卷积神经网络(Social Spatio-Temporal Graph Convolutional Neural Network, Social-STGCNN),通过时空图卷积提取时空信息,并结合时间外推器卷积预测轨迹. Li 等人^[26]为图的边分配注意力权重,并引入新的就近行人间注意函数来计算加权邻接矩阵. Chen 等人^[27]提出了可变形时空卷积神经网络(Deformable Spatial-Temporal Convolutional Neural Network, DSTCNN),克服了时空纠缠、固定卷积核的缺点. Zhang 等人^[28]提出了(Social Interaction-Weighted Spatio-Temporal Convolutional Neural Network, Social-IWSTCNN),通过池化机制和注意力权重模拟行人之间相互作用. Lv 等人^[29]提出了社交软注意力图卷积网络(Social Soft Attention Graph Convolution Network, SSAGCN)处理行人之间、行人与周围环境的交互关系. Chen 等人^[30]提出了可解释性掩码图卷积网络(Interpretable Masked Graph Convolution Network, IMGCN),利用行人视野区域、距离和运动方向等可解释信息来智能掩盖交互特征. Pi 等人^[31]利用推测的未来轨迹来提取交互特征并建模. Chen 等人^[32]提出了可变形时空交互图卷积网络(Deformable Spatial-Temporal Interaction Graph Convolution Network, DSTIGCN),通过多个不对称可变形卷积核在时空维度上的偏移自主学习行人的时空交互关系,并引入拉丁超立方采样提高模型的多模态预测效果. Mi 等人^[33]允许具有不同时间戳的节点相互交互. Chen 等人^[34]对行人相对重要性进行建模.

2.2 基于域适应网络的行人轨迹预测模型

前述模型在策略上大多都存在局限,主要表现在它们更强调对一般性运动模式的学习,而未能充分考虑训练集与测试集之间潜在的分布差异^[35]. 早期的学习策略在未考虑这些分布差异的情况下,将不同分布的样本一起学习,导致模型不可避免地产生了域偏差.

针对这一问题,近年来有学者提出了一些基于域适应方法预测行人轨迹,取得了一定进展,但相关研究仍较为有限. Xu 等人^[35]提出了一种可迁移图神经网络(Transferable Graph Neural Network, T-GNN)框架,采用域不变 GNN 减少领域特定知识结构,并行地进行轨迹预测和域对齐. Zhi 等人^[9]首次结合域适应和持续学习预测轨迹. 其中,域适应模块分析数据在源域和目标域的潜在差异,持续学习模块通过限制模型参数变化应对灾难性遗忘,从而提高了模型泛化性. Huang 等人^[36]提

出跨域轨迹预测网络(Cross-Domain Prediction Network, CTP-Net),借助编码器-解码器结构以及域适应方法,构建将目标域数据映射到源域的网络. Liu 等人^[11]使用 MMD 损失函数对齐两个域之间的差异. Poibrenski 等人^[11]使用基于伪标签的域适应方法,以监督方式同时使用标记和未标记的数据进行训练. Huang 等人^[37]利用运动选择预先固有的不变性来减轻环境变化带来的域偏移的不利影响. Tang 等人^[38]设计进化算法,鼓励在不同场景之间共享共识的同时,对每个场景的最优参数进行网络搜索. Geng 等人^[39]提出了一种可分层孪生网络(Transferable Hierarchical Transformer Network, T-HSTN),填补了异构智能体轨迹预测中多源迁移学习任务的解决方案和基准的空白. Yang 等人^[40]基于行为模式对齐特征分布来减少域差异.

3 分级包络对抗域适应和松-紧耦合行人轨迹预测模型

3.1 问题定义

给定场景中 N 个行人在时间段 T_1 到 T_{obs} 的观察轨迹表示,如式(1)所示:

$$\text{TR}_{\text{obs}} = \{\mathbf{tr}_1^i, \mathbf{tr}_2^i, \dots, \mathbf{tr}_{\text{obs}}^i \mid \forall i \in \{1, 2, \dots, N\}\} \quad (1)$$

通过可学习参数为 W^* 的模型 $f(\cdot)$, 预测 N 个行人在未来时间段 $T_{\text{obs}+1}$ 到 T_{pre} 内的轨迹,如式(2)所示:

$$\begin{aligned} \text{TR}_{\text{pre}} &= \{f(\text{TR}_{\text{obs}}) \mid W^*\} \\ &= \{\hat{\mathbf{tr}}_{\text{obs}+1}^i, \hat{\mathbf{tr}}_{\text{obs}+2}^i, \dots, \hat{\mathbf{tr}}_{\text{pre}}^i \mid \forall i \in \{1, 2, \dots, N\}\} \end{aligned} \quad (2)$$

其中, $\mathbf{tr}_t^i = (x_t^i, y_t^i)$ 表示第 i 个行人在 t 时刻的二维空间真实坐标; $\hat{\mathbf{tr}}_t^i = (\hat{x}_t^i, \hat{y}_t^i)$ 表示第 i 个行人在 t 时刻的二维空间预测坐标.

3.2 模型描述

3.2.1 包络样本定义

相关研究表明,通过样本变换的方式生成的新样本有利于提高模型的性能^[41]. 一方面,这背后的原理可以用“知识金字塔”理论来解释. 依据该理论,数据处理遵循从数据到信息再到知识的递进过程^[42]. Rowley 等人^[43]进一步论述了这一结构:数据层作为基础,虽包含大量的原始信息,但冗余度也较高. 而信息层介于数据层和知识层之间,含有对知识发现具有重要价值的信息. 知识层则代表了对样本未来行为的预测能力;另一方面,人类的认知机制研究表明,人类大脑识别样本不仅仅依赖样本本身,还依赖相似样本之间的关联信息^[44]. Bosancic 等人^[45]强调,信息层在数据与知识之间起着关键的桥梁作用,是将原始数据转化为可预测知识的重要媒介. 这一过程如图 1 所示,原始样本构成数据层,而预测的相关结果则属于知识层. 在预测任务

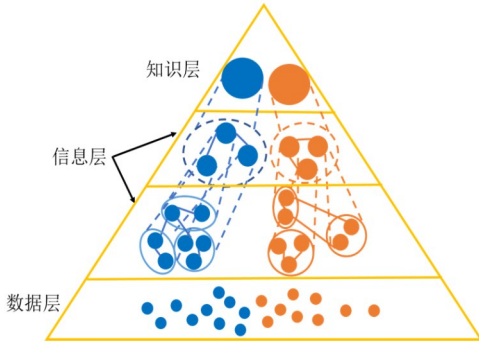


图1 知识金字塔

中,构建多层次包络样本能够充分利用信息层的关联特征,更好地捕捉特征信息,提高预测精度.

当前深度学习研究领域中,占据主导地位的是基于原始数据样本的直接建模方法,能够自动提取出反映数据本质的深层特征.本文所提出的创新性方法,融合了图结构与包络样本变换机制,致力于挖掘原始样本间潜在关联信息.通过该机制,实现将原始样本转化为多层次包络样本这一目标.

经过变换处理后获得的新样本集,不仅完整保留了原始数据的核心信息,尤为关键的是蕴含了样本间的隐含关联模式.这种处理方式本质上是对数据间复杂关系的抽象表达与高度概括.实例表明,该方法能够有效揭示系统演化的内在机理,从多维度实现对目标对象的全面表征.已有实验数据显示,相较于传统直接基于原始样本的建模方法而言,采用分层次包络样本建模策略有助于显著提升模型的预测精度^[46,47].

关于图结构包络样本变换机制的具体实施流程可作如下阐述:对于给定数据集 $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ 中的任意选定样本 $\mathbf{x}_i \in X$ 而言,首先需要通过特定算子 η 确定其邻近样本集合 $N(\mathbf{x}_i), N(\mathbf{x}_i) \subseteq X$; 随后将该初始样本 \mathbf{x}_i 与其邻近集合 $N(\mathbf{x}_i)$ 进行合并操作形成新的扩展样本集 $X_c = [\mathbf{x}_i, N(\mathbf{x}_i)]^T$. 在此过程中需要同步构建相应的邻接矩阵 $\mathbf{A}(X_c)$ 来描述各元素间的空间关系. 由此可见上述操作过程即为典型的样本包络化处理步骤; 进一步地,当对扩展后的样本集 $[X_c, \mathbf{A}(X_c)]$ 施加特定变换算子 Γ 时便会生成全新的包络样本集合 $X_e = \Gamma\{[X_c, \mathbf{A}(X_c)]\}$. 从初始集合 X 到最终包络集合 X_e 的整个转换过程被定义为完整的包络样本变换流程; 最终形成的多层次包络样本集 $(X_e^i)_{i=1}^L$ 与原有数据集 X 共同构成了分层的包络特征空间 $[X, (X_e^i)_{i=1}^L]$, 其中 $X_e^i = \Gamma^i\{[X_c, \mathbf{A}(X_c)]\}$. 实例表明通过这种方式原先基于原始数据的模型架构 $H(X)$ 得以转化为基于多层次包络特

征的新型模型 $H\{[X, (X_e^i)_{i=1}^L]\}$ 从而实现了图结构引导下的多层次包络样本构造目标.

3.2.2 模型整体介绍

本文提出的分级包络对抗域适应和松-紧耦合行人轨迹预测模型 (Hierarchical Envelope Adversarial Domain Adaptation with Loose-Tight Coupled Trajectory Prediction Model, HEADA-LTCTPM) 框图如图 2 所示. 给定源域和目标域的轨迹, 基于包络样本变换机制, 引入松-紧耦合包络样本构造模块, 将原始轨迹数据样本变换成松耦合与紧耦合的个体层次包络样本与局部层次包络样本, 而后通过特征提取与特征融合得到松耦合与紧耦合的双级包络样本. 在分级包络域适应的基础上引入对抗域适应模块, 通过生成器与判别器的对抗训练, 系统性地学习具有域不变性的深层特征, 避免了人为设计距离度量. 结合松-紧耦合决策融合机制, 进一步优化模型在复杂多变场景中的适应性和鲁棒性.

3.2.3 松-紧耦合包络样本构造模块

现有方法耦合缺乏弹性, 难以适应场景复杂多变特性, 包括冗余交互等问题. 本文设计松-紧耦合包络样本构造模块来解决这一问题. 松耦合关系指行人之间的交互关系松散. 在建模过程中, 松耦合关系通过设置较高的交互强度阈值来定义, 仅保留显著的行人交互关系. 其更适用于行人密度较低或交互较少的场景, 比如在稀疏街道场景中. 模型通过结合松耦合包络样本进行建模, 可以减少噪声干扰, 提高模型的预测精度. 紧耦合关系指行人之间的交互关系紧密. 在建模过程中, 紧耦合关系通过设置较低的交互强度阈值来定义, 会保留更多的行人交互关系. 正因如此, 其能够捕捉复杂密集场景中行人间的复杂社交行为, 适用于行人密度较高或交互频繁的场景, 比如在十字路口或地铁站等场景. 通过结合紧耦合包络样本进行建模, 可以更好地模拟行人之间的避让、跟随等行为, 提升模型在复杂场景的预测能力.

(1) 松-紧耦合局部层次包络样本构造模块

从空间维度上构造松-紧耦合局部层次包络样本, 如图 3 所示. 针对相同时刻下存在于同一场景内的各行人个体间的局部关系, 构造空间图结构, 用以表示行人轨迹的局部层次样本, 即将时间轴 t 时刻 N 个行人的坐标 $\text{TR}_t = \{\text{tr}_t^i | \forall i \in \{1, 2, \dots, N\}\}$ 处理为图表示 $G_{\text{loc}} = (V_{\text{loc}}, E_{\text{loc}})$, 其中 $E_{\text{loc}} = \{e_{T_1}^{ij}, e_{T_2}^{ij}, \dots, e_{T_{\text{obs}}}^{ij} | \forall i, j \in \{1, 2, \dots, N\}\}$ 是 G_{loc} 的边集, E_{loc} 中各条边由局部层次行人邻接关系矩阵 \mathbf{R}_{loc} 和邻接矩阵 \mathbf{A}_{loc} 表示, $V_{\text{loc}} = \{\text{tr}_{T_1}^i, \text{tr}_{T_2}^i, \dots, \text{tr}_{T_{\text{obs}}}^i | \forall i \in \{1, 2, \dots, N\}\}$ 为图表示 G_{loc} 的节点集. $\mathbf{F}_{\text{loc}} = [\mathbf{v}_{T_1}^i, \mathbf{v}_{T_2}^i, \dots, \mathbf{v}_{T_{\text{obs}}}^i]^T$ 表示各行人对应的特征矩阵, 其中 $\mathbf{v}_t^i = [x_t^i - x_{t-1}^i, y_t^i - y_{t-1}^i]$ 代表时间轴 t 时刻的第 i

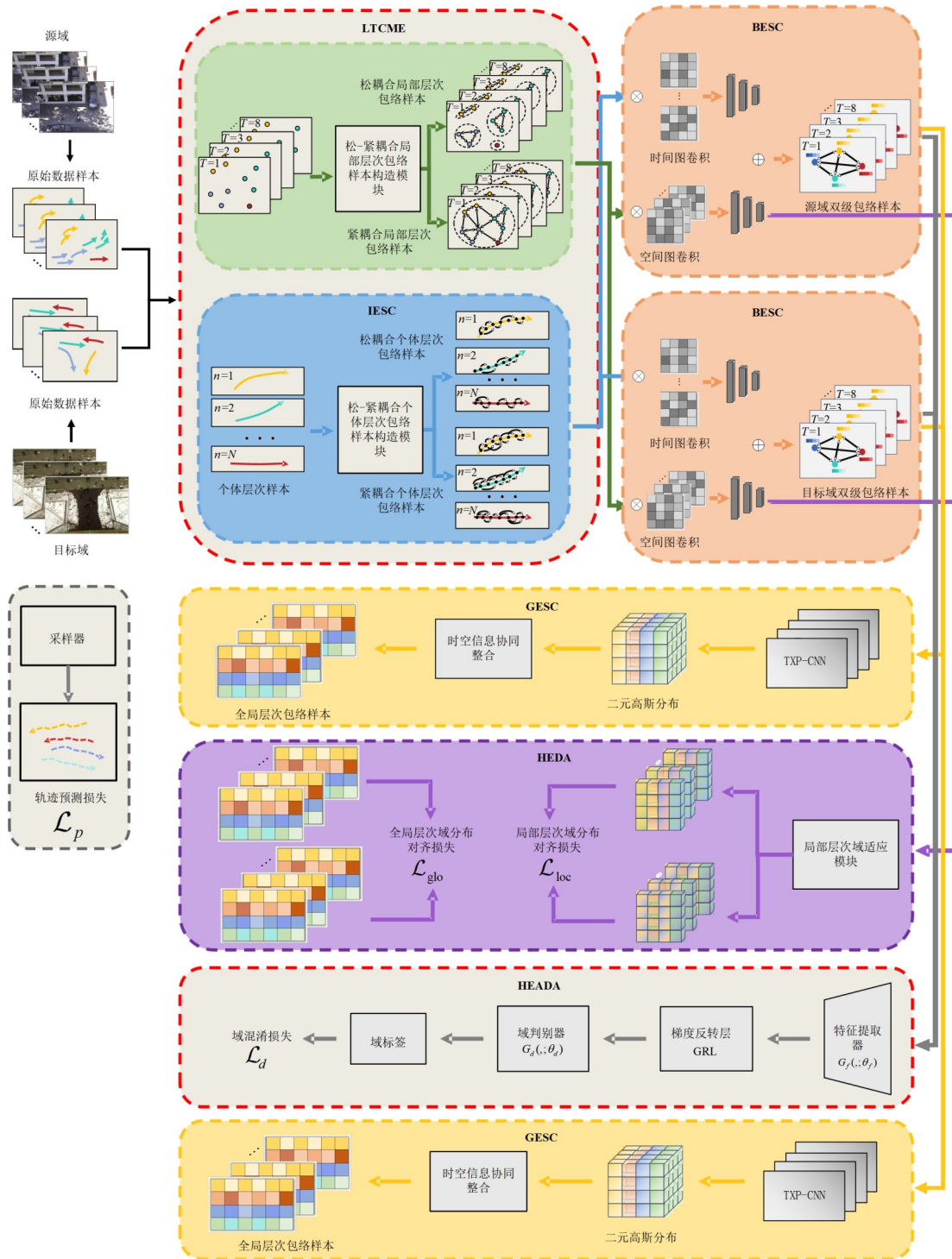


图2 HEADA-LTCTPM模型框图

个行人的速度矢量。

如图4所示,序号①~⑦代表7个行人的轨迹,局部层次行人邻接关系矩阵的构造需对行人群体的交互关系进行综合考量.两个半径不同的包围圈被定义为局部层次行人群体,分别为 r_1 和 r_2 .若目标行人与其他行人间的距离处于上述圆圈范围内,这些行人可被视作

一个局部层次包围样本.小于 r_1 的距离情况下,目标行人处于结伴状态整体性,该群体内部不存在局部层次交互关系(如行人①和行人②).当距离超过 r_2 时,行人与目标行人间的相互作用可忽略不计(如行人①和行人③). L_t^i 表示 t 时刻的行人群体包围圈内的行人 i 与行人 j 的欧式距离的倒数.

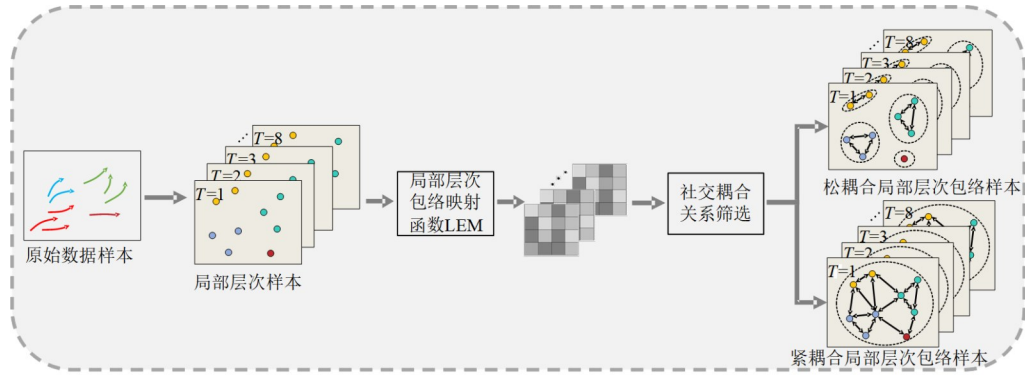


图3 松-紧耦合局部层次包络样本构造模块框图

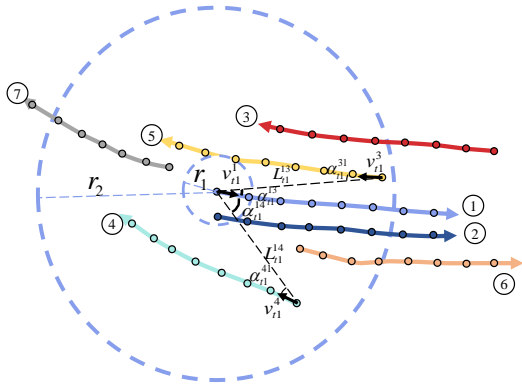


图4 局部层次包络样本交互示意图

$$L_i^{ij} = \begin{cases} \frac{1}{\sqrt{(x_i^t - x_j^t)^2 + (y_i^t - y_j^t)^2}}, & r_1 < \sqrt{(x_i^t - x_j^t)^2 + (y_i^t - y_j^t)^2} < r_2 \\ 0, & \text{Otherwise} \end{cases} \quad (3)$$

行人 i 速度方向与行人 j 当前位置连线间在 t 时刻的夹角为 α_i^{ij} , 其余弦函数值记为 $c_i^{ij} = \cos \alpha_i^{ij}$. 由此可见, t 时刻行人间速度夹角的综合度量通过 H_i^{ij} 实现.

$$H_i^{ij} = \max(0, |\mathbf{v}_i^t| c_i^{ij} + |\mathbf{v}_j^t| c_j^{ii}) \quad (4)$$

当行人 i 和行人 j 相向而行时, 夹角 α_i^{ij} 和 α_j^{ii} 均为锐角, 得到的 H_i^{ij} 大于 0, 表明行人 i 和行人 j 之间存在相互影响的局部层次交互关系 (如行人 ①、行人 ④ 和行人 ⑤). 当行人 i 和行人 j 同向而行时, 夹角 α_i^{ij} 为锐角, α_j^{ii} 为钝角, 行人 j 的速度大于行人 i 的速度, 得到的 H_i^{ij} 等于 0, 此时不存在交互作用, 即表明行人 i 和行人 j 不存在局部层次交互关系 (如行人 ① 和行人 ⑥). 当行人 i 和行人 j 反向而行时, 夹角 α_i^{ij} 和 α_j^{ii} 均为钝角, 得到的 H_i^{ij} 等于 0, 表明行人 i 和行人 j 不存在局部层次交互关系 (如行人 ① 和行人 ⑦).

为确保场景中停留驻足的行人不易受到局部层次行人群体的影响, 速度较快者则更易受到该群体作用, 本文定义了 $\tau_i^{ij} = \min(1, |\mathbf{v}_i^t|/|\mathbf{v}_j^t|)$. 综合来看, 多种因素被包含于局部层次包络样本映射函数 $F_{LEM}(\cdot)$ 之中, 如行人

速度值、间距度量以及相对运动方向上的夹角参数等. 由此可见, 局部层次的邻接关系矩阵 \mathbf{R}_{loc} 的各个元素 r_i^{ij} 的计算得以基于上述因素完成, 即

$$r_i^{ij} = F_{LEM}(\mathbf{v}_i^t, \mathbf{v}_j^t, \mathbf{tr}_i^t, \mathbf{tr}_j^t) = \begin{cases} L_i^{ij} H_i^{ij} \tau_i^{ij}, & i \neq j \\ 0, & \text{Otherwise} \end{cases} \quad (5)$$

为区分松-紧耦合关系, 引入不同的阈值超参数 η_1 和 η_2 实现社交耦合关系筛选, 生成松-紧耦合局部层次掩码矩阵 \mathbf{M}_{locloo} 和 \mathbf{M}_{locdig} , 即

$$\mathbf{M}_{locloo} = \mathbb{I}(\mathbf{R}_{loc} \geq \eta_1) \quad (6)$$

$$\mathbf{M}_{locdig} = \mathbb{I}(\mathbf{R}_{loc} \geq \eta_2) \quad (7)$$

通过逐元素乘法筛选不同耦合关系的局部层次行人邻接关系矩阵 \mathbf{R}_{loc} , 而后引入单位阵 \mathbf{I} , 得到松-紧耦合局部层次邻接矩阵 \mathbf{A}_{locloo} 和 \mathbf{A}_{locdig} , 其中 \odot 表示逐元素乘法, 即

$$\mathbf{A}_{locloo} = \mathbf{M}_{locloo} \odot \mathbf{R}_{loc} + \mathbf{I} \quad (8)$$

$$\mathbf{A}_{locdig} = \mathbf{M}_{locdig} \odot \mathbf{R}_{loc} + \mathbf{I} \quad (9)$$

而后, 通过 Zero-Softmax 操作对 \mathbf{A}_{locloo} 和 \mathbf{A}_{locdig} 中的每个元素 a_i^{ij} 归一化. 最终得到归一化邻接矩阵 $\mathbf{A}'_{locloo} = \text{Zero-Softmax}(\mathbf{A}_{locloo})$, $\mathbf{A}'_{locdig} = \text{Zero-Softmax}(\mathbf{A}_{locdig})$. 通过堆叠不同时刻的邻接矩阵 \mathbf{A}'_{locloo} 和 \mathbf{A}'_{locdig} 和对应行人特征矩阵 \mathbf{F}_{locloo} 和 \mathbf{F}_{locdig} , 最终构建松耦合局部层次包络样本 $[\mathbf{F}_{locloo}, \mathbf{A}'_{locloo}]$ 和紧耦合局部层次包络样本 $[\mathbf{F}_{locdig}, \mathbf{A}'_{locdig}]$.

(2) 松-紧耦合个体层次包络样本构造模块

从时间维度上构造松-紧耦合个体层次包络样本, 如图 5 所示. 针对同一场景不同时刻的同一行人的轨迹, 通过构建时间图结构表示该行人轨迹的个体层次样本; 将时间段 T_1 至 T_{obs} 内观察的行人 i 的坐标 $\text{TR}_i^t = \{\mathbf{tr}_i^t | \forall t \in \{T_1, T_2, \dots, T_{obs}\}\}$ 处理为图表示 $G_{ind} = (V_{ind}, E_{ind})$, 其中 $E_{ind} = \{e_{p,q}^1, e_{p,q}^2, \dots, e_{p,q}^N | \forall p, q \in \{T_1, T_2, \dots, T_{obs}\}\}$ 是 G_{ind} 的边集, E_{ind} 中各条边由个体层次行人邻接关系矩阵 \mathbf{R}_{ind} 和邻接矩阵 \mathbf{A}_{ind} 表示, $V_{ind} = \{\mathbf{tr}_i^1, \mathbf{tr}_i^2, \dots, \mathbf{tr}_i^N | \forall t \in \{T_1, T_2, \dots, T_{obs}\}\}$ 为 G_{ind} 的节点集.

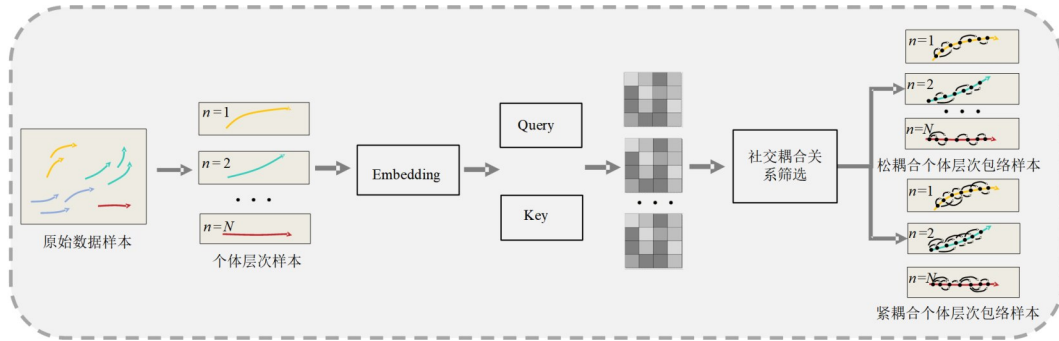


图5 松-紧耦合个体层次包络样本构造模块框图

$\mathbf{F}_{\text{ind}} = [\mathbf{v}_{T_1}^i, \mathbf{v}_{T_2}^i, \dots, \mathbf{v}_{T_{\text{obs}}}^i]^T$ 是 T_1 至 T_{obs} 内对应的行人特征矩阵, $\mathbf{v}_t^i = [x_t^i - x_{t-1}^i, y_t^i - y_{t-1}^i]$ 代表时间轴 t 时刻的第 i 个行人的速度矢量.

个体层次包络样本的行人邻接关系矩阵 \mathbf{R}_{ind} 由注意力机制求得, Fra_t 表示 t 时刻的时间帧数. $\mathbf{PE} = [\text{Fra}_{T_1}, \text{Fra}_{T_2}, \dots, \text{Fra}_{T_{\text{obs}}}]$ 对应单个行人轨迹 tr^i 在时间段 T_1 至 T_{obs} 内的位置编码张量. 对原始的个体层次样本特征向量拼接对应的位置编码张量 \mathbf{PE} , 并通过可学习参数 \mathbf{W}_{ind} 进行线性变换, 得到带有位置编码信息的高维张量 \mathbf{F}'_{ind} :

$$\mathbf{F}'_{\text{ind}} = (\mathbf{F}_{\text{ind}} \oplus \mathbf{PE}) \mathbf{W}_{\text{ind}} \quad (10)$$

然后, 通过可学习参数 \mathbf{W}_Q 和 \mathbf{W}_K 分别得到注意力机制的查询向量和键向量. 利用缩放因子 D 对查询向量和键向量的乘积进行放缩, 并通过 Softmax 归一化, 得到个体层次包络样本的 \mathbf{F}_{ind} 行人邻接关系矩阵 \mathbf{R}_{ind} , 有

$$\mathbf{Q}_{\text{ind}} = \mathbf{F}'_{\text{ind}} \mathbf{W}_Q \quad (11)$$

$$\mathbf{K}_{\text{ind}} = \mathbf{F}'_{\text{ind}} \mathbf{W}_K \quad (12)$$

$$\mathbf{R}_{\text{ind}} = \text{softmax} \left(\frac{\mathbf{Q}_{\text{ind}} \mathbf{K}_{\text{ind}}^T}{D} \right) \quad (13)$$

$$f(x, y | \boldsymbol{\mu}, \boldsymbol{\sigma}, \rho) = \frac{\exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_x}{\sigma_x} \right)^2 - 2\rho \frac{x-\mu_x}{\sigma_x} \frac{y-\mu_y}{\sigma_y} + \left(\frac{y-\mu_y}{\sigma_y} \right)^2 \right] \right\}}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \quad (18)$$

从二元正态分布中采样生成松耦合预测轨迹 $\hat{\text{TR}}_{\text{looptre}}$ 和紧耦合预测轨迹 $\hat{\text{TR}}_{\text{tigpre}}$. $\hat{\text{TR}}_{\text{looptre}} \sim \mathcal{N}(\hat{\boldsymbol{\mu}}_1, \hat{\boldsymbol{\sigma}}_1, \hat{\rho}_1)$, $\hat{\text{TR}}_{\text{tigpre}} \sim \mathcal{N}(\hat{\boldsymbol{\mu}}_2, \hat{\boldsymbol{\sigma}}_2, \hat{\rho}_2)$. 其中, $\boldsymbol{\mu} = [\mu_x, \mu_y]$ 为均值; $\boldsymbol{\sigma} = [\sigma_x, \sigma_y]$ 为标准差; ρ 为相关系数.

为了量化松耦合模型与紧耦合模型对预测轨迹的贡献程度, 本文引入松耦合模型的权重因子 w_1 和紧耦合模型的权重因子 w_2 , 分别用于衡量松耦合关系和紧耦合关系对预测轨迹的重要程度.

总的来说, 基于松耦合包络样本的预测模型主要捕捉行人之间较弱的交互关系, 所以适用于稀疏的行人场景; 基于紧耦合包络样本的预测模型更侧重于刻

为区分个体层次的松-紧耦合关系, 引入不同的阈值超参数 ψ_1 和 ψ_2 实现各个行人时间层次的耦合关系筛选, 生成松-紧耦合个体层次掩码矩阵 $\mathbf{M}_{\text{indloo}}$ 和 $\mathbf{M}_{\text{indtig}}$, 即

$$\mathbf{M}_{\text{indloo}} = \mathbb{I}(\mathbf{R}_{\text{loc}} \geq \psi_1) \quad (14)$$

$$\mathbf{M}_{\text{indtig}} = \mathbb{I}(\mathbf{R}_{\text{loc}} \geq \psi_2) \quad (15)$$

通过逐元素乘法运算将实现对不同耦合关系的个体层次行人邻接关系矩阵 \mathbf{R}_{ind} 的筛选, 得到松-紧耦合个体层次邻接矩阵 $\mathbf{A}_{\text{indloo}}$ 和 $\mathbf{A}_{\text{indtig}}$, 即

$$\mathbf{A}_{\text{indloo}} = \mathbf{M}_{\text{indloo}} \odot \mathbf{R}_{\text{ind}} \quad (16)$$

$$\mathbf{A}_{\text{indtig}} = \mathbf{M}_{\text{indtig}} \odot \mathbf{R}_{\text{ind}} \quad (17)$$

采用 Zero-Softmax 方法对邻接矩阵 $\mathbf{A}_{\text{indloo}}$ 和 $\mathbf{A}_{\text{indtig}}$ 进行归一化操作, 得到归一化的邻接矩阵 $\mathbf{A}'_{\text{indloo}}$ 和 $\mathbf{A}'_{\text{indtig}}$. 通过堆叠不同行人的邻接矩阵和对应行人特征矩阵 $\mathbf{F}_{\text{indloo}}$ 和 $\mathbf{F}_{\text{indtig}}$, 最终构建松耦合个体层次包络样本 $[\mathbf{F}_{\text{indloo}}, \mathbf{A}'_{\text{indloo}}]$ 和紧耦合个体层次包络样本 $[\mathbf{F}_{\text{indtig}}, \mathbf{A}'_{\text{indtig}}]$.

3.2.4 松-紧耦合轨迹融合模块

假设预测的行人轨迹服从二元正态分布^[16], 其概率密度函数如式(18)所示:

画行人间复杂的交互关系, 所以适用于密集交互的场景.

通过动态调整 w_1 和 w_2 的取值, 模型能够自适应地平衡两种耦合关系的权重, 更好地适应多样化的场景需求. 最终的行人轨迹预测结果通过加权融合得到, 如式(19)所示:

$$\hat{\text{TR}}_{\text{pre}} = w_1 \hat{\text{TR}}_{\text{looptre}} + w_2 \hat{\text{TR}}_{\text{tigpre}} \quad (19)$$

s.t. $w_1 + w_2 = 1$

其中, $\hat{\text{TR}}_{\text{looptre}}$ 和 $\hat{\text{TR}}_{\text{tigpre}}$ 分别表示松耦合模型和紧耦合模型的预测轨迹. 通过加权融合机制, 模型能够充分利用松-紧耦合包络样本的互补优势, 能更准确地预测未

来轨迹.

3.2.5 对抗域适应模块

通过引入对抗域适应模块,旨在学习源域和目标域的深层公共特征表示,如图6所示.

该模块由3个核心组件构成:特征提取器、轨迹预测模块和域判别器.特征提取器负责生成深度特征.该特征能够并行地作为轨迹预测模块和域判别器的输入.域判别器的主要功能是区分输入特征的领域来源,判断特征来自源域还是目标域.轨迹预测模块则用于预测未来的行人轨迹.

由双级包络样本特征融合模块得到源域双级包络

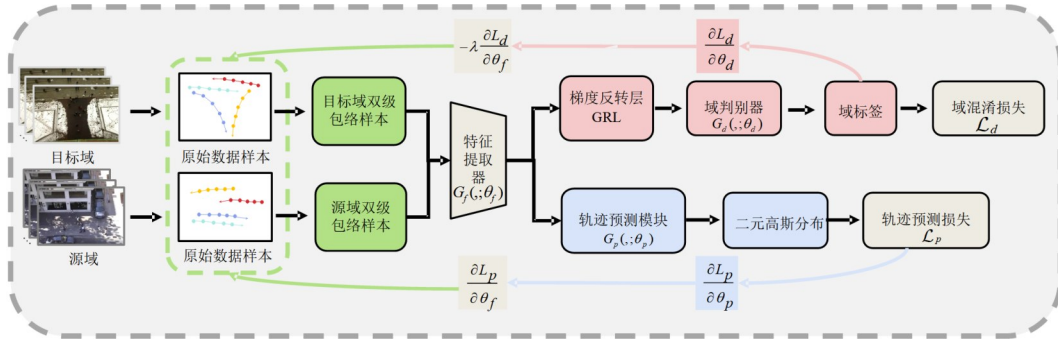


图6 对抗域适应模块

特征向量 f_{tef}^i 包含了某个双级包络样本中第 i 个人的时空信息.然而,源域和目标域的特征表示 F_{tefs} 和 F_{teft} 的尺度不同,传统的域适应方法难以直接进行特征对齐.因此采用注意力机制,学习目标行人与其周围行人之间的相关性,并确保不同域的特征表示具有相同的维度.注意力分数矩阵 G_S 和 G_T 的各个元素 $g_S^{i,j}$ 和 $g_T^{i,j}$ 计算方式如式(22)和式(23)所示:

$$g_S^{i,j} = \frac{\left(f_{\text{tefs}}^i (f_{\text{tefs}}^j)^T / \sqrt{D_a} \right)}{\sum_{j=1}^{N_s} \left(f_{\text{tefs}}^i (f_{\text{tefs}}^j)^T / \sqrt{D_a} \right)} \quad (22)$$

$$g_T^{i,j} = \frac{\left(f_{\text{teft}}^i (f_{\text{teft}}^j)^T / \sqrt{D_a} \right)}{\sum_{j=1}^{N_T} \left(f_{\text{teft}}^i (f_{\text{teft}}^j)^T / \sqrt{D_a} \right)} \quad (23)$$

通过计算双级包络样本中各个行人特征 f_{tef}^i 和 f_{tef}^j 的乘积,用以表示行人之间交互关系的权重.具体而言,计算特征向量的内积,求得行人 i 和行人 j 之间的关联强度.而后得到所有行人之间关联强度总和,并将特定两个行人之间的交互权重与总和的比例作为注意力分数.将对应的行人注意力分数与目标行人的局部层次特征相乘,最终得到源域和目标域的域不变的特征集中表达 J_{tefs} 和 J_{teft} :

样本的特征表示 $F_{\text{tefs}} \in \mathbb{R}^{N_s \times L_{\text{obs}} \times 2D}$ 和目标域双级包络样本的特征表示 $F_{\text{teft}} \in \mathbb{R}^{N_T \times L_{\text{obs}} \times 2D}$, N_s 是源域数据样本行人的数量, N_T 是目标域数据样本行人的数量, L_{obs} 为观测轨迹长度.为关注双级包络样本的特征分布集中表示,源域和目标域局部层次特征表示 F_{tefs} 和 F_{teft} 按不同行人进行分解,如式(20)和式(21)所示:

$$F_{\text{tefs}} = \left[f_{\text{tefs}}^1 \quad f_{\text{tefs}}^2 \quad \cdots \quad f_{\text{tefs}}^{N_s} \right] \quad (20)$$

$$F_{\text{teft}} = \left[f_{\text{teft}}^1 \quad f_{\text{teft}}^2 \quad \cdots \quad f_{\text{teft}}^{N_T} \right] \quad (21)$$

其中, f_{tefs}^i 和 f_{teft}^i 分别对应于源域和目标域中的双级包络样本的第 i 个行人的特征, $f_{\text{tefs}}^i \in \mathbb{R}^{L_{\text{obs}} \times 2D}$, $f_{\text{teft}}^i \in \mathbb{R}^{L_{\text{obs}} \times 2D}$. 将其展平为 D_a 大小的特征向量, $D_a = L_{\text{obs}} \times 2D$.

$$J_{\text{tefs}} = \frac{\sum_{i=1}^{N_s} \left(\sum_{j=1}^{N_s} \left(g_S^{i,j} f_{\text{tefs}}^i \right) \right)}{N_s} \quad (24)$$

$$J_{\text{teft}} = \frac{\sum_{i=1}^{N_T} \left(\sum_{j=1}^{N_T} \left(g_T^{i,j} f_{\text{teft}}^i \right) \right)}{N_T} \quad (25)$$

而后通过两个全连接层和 Sigmoid 激活函数进行降维处理,并表示为领域标签 d :

$$d = \text{Sigmoid} \left(\text{ReLU} \left(J_{\text{tef}} W_1 + b_1 \right) W_2 + b_2 \right) \quad (26)$$

梯度反转层是域适应任务中实现领域对抗思想的关键组件.为了形式化描述梯度反转层的作用,引入函数 $R_\lambda(x)$ 来表示其操作.式(27)和式(28)分别阐述了 GRL 在前向传播和反向传播中的行为,其中 I 为单位矩阵,即

$$R_\lambda(x) = x \quad (27)$$

$$\frac{dR_\lambda}{dx} = -I \quad (28)$$

对于输入 x ,域判别器 $G_d(\cdot; \theta_d)$ 用于判断特征 $G_f(x; \theta_f)$ 是来自源域分布 $P_s(G_f(x))$ 还是目标域分布 $P_t(G_f(x))$,即对 x 的领域来源进行分类.领域标签 d 用 1 (源域数据)和 0 (目标域数据)表示.轨迹预测损失和域判别损失分别定义如式(29)和式(30)所示:

$$\mathcal{L}_d^i(\theta_f, \theta_d) = \mathcal{L}_d(G_d(G_f(x_i; \theta_f); \theta_d), d_i) \quad (29)$$

$$\mathcal{L}_p^i(\theta_f, \theta_p) = \mathcal{L}_p(G_p(G_f(x_i; \theta_f); \theta_p), (\mathbf{tr}_{\text{obs}+1}^i, \dots, \mathbf{tr}_{\text{pre}}^i)) \quad (30)$$

$$\begin{aligned} E(\theta_f, \theta_p, \theta_d) &= \frac{1}{n} \sum_{i=1}^n \mathcal{L}_p^i(\theta_f, \theta_p) + \lambda \left(\frac{1}{n} \sum_{i=1}^n \mathcal{L}_d^i(\theta_f, \theta_d) + \frac{1}{m} \sum_{i=n+1}^{n+m} \mathcal{L}_d^i(\theta_f, \theta_d) \right) \\ &= -\frac{1}{n} \sum_{i=1}^n \sum_{t=T_{\text{obs}+1}}^{T_{\text{pre}}} \log \left(\mathbb{P} \left((x_t^i, y_t^i) \mid \hat{\boldsymbol{\mu}}_t^i, \hat{\boldsymbol{\sigma}}_t^i, \hat{\rho}_t^i \right) \right) + \lambda \left(\frac{1}{n} \sum_{i=1}^n \mathcal{L}_d^i(\theta_f, \theta_d) + \frac{1}{m} \sum_{i=n+1}^{n+m} \mathcal{L}_d^i(\theta_f, \theta_d) \right) \end{aligned} \quad (31)$$

其中, λ 为调节域判别任务与轨迹预测损失相对重要性的权重系数; n 是源域数据样本的数量; m 是目标域数据样本的数量。

本文采用高斯核函数 $K(x, y) = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right)$ 的

MMD 损失计算源域和目标域的局部层次域分布对齐损失 \mathcal{L}_{loc} 和全局层次域分布对齐损失 \mathcal{L}_{glo} 。

总体目标函数由 4 项损失组成, 包括轨迹预测损失 \mathcal{L}_p , 对抗域适应损失 \mathcal{L}_d , 局部层次域分布对齐损失 \mathcal{L}_{loc} 和全局层次域分布对齐损失 \mathcal{L}_{glo} :

$$\begin{aligned} \mathcal{L} &= \mathcal{L}_p + \lambda \mathcal{L}_d + \sigma \mathcal{L}_{\text{loc}} + \beta \mathcal{L}_{\text{glo}} \\ &= -\frac{1}{n} \sum_{i=1}^n \sum_{t=T_{\text{obs}+1}}^{T_{\text{pre}}} \log \left(\mathbb{P} \left((x_t^i, y_t^i) \mid \hat{\boldsymbol{\mu}}_t^i, \hat{\boldsymbol{\sigma}}_t^i, \hat{\rho}_t^i \right) \right) \\ &\quad + \lambda \left(\frac{1}{n} \sum_{i=1}^n \mathcal{L}_d^i(\theta_f, \theta_d) + \frac{1}{m} \sum_{i=n+1}^{n+m} \mathcal{L}_d^i(\theta_f, \theta_d) \right) \\ &\quad + \sigma \mathcal{L}_{\text{loc}} + \beta \mathcal{L}_{\text{glo}} \end{aligned} \quad (32)$$

其中, 轨迹预测损失 \mathcal{L}_p 只涉及源域样本; λ, σ 和 β 是平衡这 4 项损失的超参数。

4 实验结果与分析

本节组织了多种定量实验和定性实验进行方法的有效性验证, 包括消融实验、参数分析实验、方法对比实验、时间开销对比实验、轨迹可视化、特征分布可视化。

4.1 实验设置

数据集. 本文实验基于公开的且具有代表性的 ETH^[48] 与 UCY^[49] 数据集进行分析。这些数据集涵盖了城市的多种场景, 包括商业街、广场、大学区域等多样化的场景。具体而言, ETH 包含 ETH 和 HOTEL 两个场景; UCY 包含 UNIV、ZARA1 以及 ZARA2 三个场景。数据集经过标准化处理, 先逐帧标注行人位置信息, 而后统一转换至真实坐标系, 并采用标准化文本格式存储。

评估指标. 本文采用平均位移误差 (Average Displacement Error, ADE) 和最终位移误差 (Final Displacement Error, FDE) 作为评估指标, 用以评估轨迹预测模型的性能。ADE 指标表示模型总体预测轨迹与真实轨迹的差异, 计算公式如式 (33) 所示。FDE 指标表示模型

基于最小化负对数似然损失来衡量轨迹的预测损失 \mathcal{L}_p 。基于此, 对抗域适应模块的损失如式 (31) 所示:

在最后一帧的预测轨迹与真实轨迹的差异, 计算公式如式 (34) 所示:

$$\text{ADE} = \frac{1}{N \times (T_{\text{pre}} - T_{\text{obs}})} \sum_{i=1}^N \sum_{t=T_{\text{obs}+1}}^{T_{\text{pre}}} \sqrt{(\hat{x}_t^i - x_t^i)^2 + (\hat{y}_t^i - y_t^i)^2} \quad (33)$$

$$\text{FDE} = \frac{1}{N} \sum_{i=1}^N \sqrt{(\hat{x}_{T_{\text{pre}}}^i - x_{T_{\text{pre}}}^i)^2 + (\hat{y}_{T_{\text{pre}}}^i - y_{T_{\text{pre}}}^i)^2} \quad (34)$$

其中, N 表示行人数量; T_{obs} 表示观测的时间帧; T_{pre} 表示预测的时间帧; x, y 表示二维空间轨迹真实坐标; \hat{x}, \hat{y} 表示二维空间轨迹预测坐标。ADE 和 FDE 在量化轨迹预测的几何精度的同时, 还反映了模型在复杂交互场景中的适应能力。

实验环境. HEAD-A-LTCTPM 模型和所有其他比较模型都采用 Python 3.8, 在 CPU i9-13900H GPU 和 RTX A3000, 内存 64 GB 的服务器上进行模型的训练和测试。

实验设置. 本研究遵循了一套既定的、基于域适应的公认范式, 每个场景被视为一个轨迹域, 模型在一个轨迹域上训练, 在其他 4 个轨迹域上进行验证和测试。本文用 A、B、C、D、E 分别代表 ETH、HOTEL、UNIV、ZARA1、ZARA2 这 5 个场景的数据集。因此, 总共有 20 个轨迹预测任务: A→B/C/D/E, B→A/C/D/E, C→A/B/D/E, D→A/B/C/E, E→A/B/C/D。在训练阶段, 本文模型只能访问验证集中观察到的轨迹, 即验证集仅作为域适应用途^[50]。观测轨迹的时间长度为 8 帧 (3.2 s), 真实轨迹和预测轨迹的时间长度为 12 帧 (4.8 s)。

参数设置. 参考文献 [51], 局部层次包络圈半径 r_1 和 r_2 设置为 0.5 和 10; 局部层次包络样本的图卷积层数为 1; 松耦合局部层次包络样本的空间耦合关系力度筛选阈值设置为 0.1; 紧耦合局部层次包络样本的空间耦合关系力度筛选阈值设置为 0.05; 个体层次包络样本的图卷积层数为 2; 松耦合个体层次包络样本的时间耦合关系力度筛选阈值设置为 0.8; 紧耦合个体层次包络样本的时间耦合关系力度筛选阈值设置为 0.7; 双级包络样本特征融合后的时间外推卷积神经网络层数为 5; 特征维度为 64; 批量大小为 16; 迭代次数为 200; 初始学习率为 0.01 (100 代后下降为 0.005)。

4.2 消融实验

为验证本算法的对抗域适应模块和松-紧耦合包络样本构造模块的有效性,组织如下 HEADA-LTCTPM 消融实验.

(1) OSCDA (Original Sample Construction Domain Adaptation): 基于原始样本构建的域适应行人轨迹预测模型.

(2) MESC-HEDA (Multilevel Envelope Sample Construction-Hierarchical Envelope Domain Adaptation): 在 OSCDA 的基础上,引入分级包络样本构建模块.

(3) MESC-HEADA (Multilevel Envelope Sample Construction-Hierarchical Envelope Adversarial Domain Adaptation): 在 MESC-HEDA 的基础上,引入对抗域适应模块.

(4) HEADA-LTCTPM (Loosely-Tightly Coupled Multilevel Envelope-Hierarchical Envelope Adversarial Domain Adaptation): 在 MESC-HEADA 的基础上,引入松-紧耦合包络样本构造模块.

消融实验的结果如表 1 所示,其中 X 代表未使用该

模块,✓代表使用该模块,数据加粗为最优结果.

由表 1 可见,基于样本包络变换机制,MESC-HEDA 模型能够更好地挖掘行人在个体层次和局部层次的关联信息.较基于原始样本构造的模型 OSCDA,MESC-HEDA 模型的轨迹预测性能明显提升,平均位移误差 (ADE)降低了 8.9%,最终位移误差(FDE)降低了 9.3%.

进一步地,MESC-HEADA 模型引入“对抗域适应模块”后,模型能够有效提取具有跨域不变性的深层特征表示,在跨域任务中的表现明显提升.相较于依赖统计准则进行域适应的 MESC-HEDA 模型,MESC-HEADA 模型的轨迹预测性能明显提升,平均位移误差 (ADE)降低了 7.8%,最终位移误差(FDE)降低了 7.9%.

进一步地,HEADA-LTCTPM 模型引入“松-紧耦合包络样本构造模块”后,构建了不同耦合关系的包络样本,决策融合进行轨迹预测,能够减少冗余信息的影响,适应多样化的场景需求.与仅考虑单一耦合关系的 MESC-HEADA 模型相比,HEADA-LTCTPM 模型在各场景下的预测性能得到了明显的提升,ADE 降低了 10.6%,FDE 降低了 12.2%.

表 1 消融实验

单位:m

Variants	模块组件			AVE
	分级包络样本构建模块	对抗域适应模块	松-紧耦合包络样本构造模块	ADE/FDE
OSCD	X	X	X	0.56/0.96
MESC-HEDA	✓	X	X	0.51/0.89
MESC-HEADA	✓	✓	X	0.47/0.82
HEADA-LTCTPM	✓	✓	✓	0.42/0.72

4.3 参数对比实验

(1)不同的松-紧耦合局部层次包络样本的社交耦合关系力度筛选阈值

保持相同的松-紧耦合个体层次包络样本的社交耦合关系力度筛选阈值,本文设计了不同的松-紧耦合局部层次包络样本的社交耦合关系力度筛选阈值的组合来探究不同的筛选力度对模型的影响,实验结果如表 2 所示,其中加粗数据为最优结果.松耦合局部层次包络样本的空间耦合关系力度筛选阈值设置为 0.10,紧耦合局部层次包络样本的空间耦合关系力度筛选阈值设置为 0.050 时,松-紧耦合局部层次包络样本分别保留约 10% 和 20% 的社交关系(通过式(35)估算社交关系保

留百分比 ϖ),ADE 和 FDE 指标最小.当紧耦合交互阈值过小时,如 0.001 (约保留 50% 的社交关系),模型学习了大量的冗余信息,反而影响了模型的预测性能.当松耦合交互阈值过大时,如 0.12 (约保留 8% 的社交关系),模型无法充分学习到行人之间的交互影响,有

$$\varpi = \frac{1}{N_{\text{sample}}} \sum_{k=1}^{N_{\text{sample}}} \frac{\sum_{i=1}^N \sum_{j=1}^N \mathbb{I}(a^{ij} \neq 0)}{N * N} \quad (35)$$

其中, N 表示单个样本中行人的数量; a^{ij} 是该筛选阈值下的邻接矩阵中的元素; N_{sample} 是总的样本数.

(2)不同的松-紧耦合个体层次包络样本的社交耦合关系力度筛选阈值

保持相同的松-紧耦合局部层次包络样本的社交耦合关系力度筛选阈值,本文设计了不同的松-紧耦合个体层次包络样本的社交耦合关系力度筛选阈值的组合来探究不同的筛选力度对模型的影响,实验结果如表 3 所示,其中加粗数据为最优结果.当松耦合个体层次包络样本的时间耦合关系力度筛选阈值设置为 0.80,紧耦合个体层次包络样本的时间耦合关系力度筛选阈值设

表 2 局部层次包络样本松-紧耦合社交力度筛选实验结果(ADE/FDE)

紧耦合交互阈值	松耦合交互阈值			
	0.06	0.08	0.10	0.12
0.001	0.47/0.82	0.44/0.77	0.43/0.75	0.48/0.82
0.010	0.45/0.78	0.43/0.75	0.42/0.73	0.46/0.77
0.030	0.45/0.78	0.43/0.75	0.43/0.73	0.45/0.77
0.050	0.46/0.80	0.42/0.74	0.42/0.72	0.46/0.78

置为 0.70 时,松-紧耦合个体层次包络样本分别保留约 10% 和 50% 的交互信息,ADE 和 FDE 指标最小. 当紧耦合交互阈值过小时,如 0.50(约保留 99% 的交互信息),模型学习了大量的冗余信息,反而影响了模型的预测性能. 当松耦合交互阈值过大时,如 0.90(约保留 2% 的交互信息),模型无法充分学习到行人各个时间帧的交互信息.

表 3 个体层次包络样本松-紧耦合社交力度筛选实验结果(ADE/FDE)

紧耦合交互阈值	松耦合交互阈值			
	0.75	0.80	0.90	1.00
0.50	0.45/0.78	0.43/0.73	0.46/0.80	0.45/0.78
0.60	0.49/0.83	0.43/0.73	0.46/0.80	0.46/0.78
0.65	0.46/0.77	0.42/0.73	0.46/0.78	0.49/0.84
0.70	0.45/0.77	0.42/0.72	0.45/0.79	0.46/0.79

4.4 方法对比

在相同的实验设置下,对本文提出的 HEADA-LTCTPM 模型与 6 种相关模型进行比较. 20 个任务的 ADE 和 FDE 指标分别如表 4 和表 5 所示,其中加粗数据为每一行的最优结果. 在对比的所有模型中,本文提出的 HEADA-LTCTPM 模型在 ADE 和 FDE 指标上性能最好. HEADA-LTCTPM 模型的 ADE 指标为 0.42. 相较于 T-GNN 和 MESC-HEDA 模型分别降低了 36.3% 和 17.6%. HEADA-LTCTPM 模型的 FDE 指标为 0.72,相较于 T-GNN 和 MESC-HEDA 模型分别降低了 35.1% 和 19.1%. 在 20 项跨域任务中,HEADA-LTCTPM 模型在 18 项任务中取得了最优的 ADE 表现,在 17 项任务中实现了最优的 FDE 表现.

从目标域的角度来看:在稀疏场景下,以 HOTEL 场景为目标域的实验,HEADA-LTCTPM 模型在 4 项任务中均取得了最优的 ADE 和 FDE 指标,其平均 ADE 为 0.40,平均 FDE 为 0.67. 与次优的 MESC-HEDA 模型相比(平均 ADE 为 0.51,平均 FDE 为 0.86),HEADA-LTCTPM 模型的 ADE 和 FDE 指标分别降低了 22.0% 和 22.1%. 在密集场景下,以 ZARA1 场景为目标域的实验,HEADA-LTCTPM 模型在四项任务中均表现最优,其平均 ADE 为 0.34,平均 FDE 为 0.59. 相较于次优的 MESC-HEDA 模型(平均 ADE 为 0.39,平均 FDE 为 0.67),HEADA-LTCTPM 模型的 ADE 和 FDE 指标分别降低了 12.9% 和 11.9%. 该结果表明,从目标域的角度而言,无论是面向人流稀疏场景还是人流密集场景的跨域任务下,HEADA-LTCTPM 模型均取得较好的预测结果.

从源域的角度来看:在稀疏场景下,以 HOTEL 场景为源域的实验,HEADA-LTCTPM 模型在 4 项任务中均取得了最优的 ADE 和 FDE 指标,其平均 ADE 为 0.61,平均 FDE 为 1.09. 相较于次优的 MESC-HEDA 模型(平均 ADE 为 0.74,平均 FDE 为 1.34),HEADA-LTCTPM 模型

的 ADE 和 FDE 指标分别降低了 17.6% 和 18.7%. 在密集场景下,以 ZARA1 场景为源域的实验,HEADA-LTCTPM 模型在四项任务中有 3 项表现最优,其平均 ADE 为 0.37,平均 FDE 为 0.65. 对比次优的 MESC-HEDA 模型(平均 ADE 为 0.43,平均 FDE 为 0.81),HEADA-LTCTPM 模型的 ADE 和 FDE 指标分别降低了 14.0% 和 19.8%.

此外,在以 UNIV 场景为源域的实验,由于 UNIV 场景的数据量较大,可视为密集复杂的场景. HEADA-LTCTPM 模型在四项任务中的 3 项任务表现最优,其平均 ADE 为 0.32,平均 FDE 为 0.53. 与 Transformer 模型(平均 ADE 为 0.38,平均 FDE 为 0.60)相比,HEADA-LTCTPM 模型的 ADE 指标和 FDE 指标分别降低了 15.8% 和 11.7%. 这表明,在数据量大的密集场景中,即便 Transformer 模型具备较强的时空特征提取能力,HEADA-LTCTPM 模型能够通过对抗域适应模块和松-紧耦合决策融合机制,更好地适应多样化的人流稀疏场景或者是人流密集场景,显著优于 Transformer 模型.

综上所述,HEADA-LTCTPM 模型在人流稀疏场景和人流密集场景中表现优异,显著优于 Transformer 模型. HEADA-LTCTPM 模型的优异性主要归因于引入了松-紧耦合包络样本构造模块和对抗域适应模块. 通过构造松耦合与紧耦合包络样本,捕捉数据的内在时空规律,更好地适应复杂多变的场景. 对抗域适应的引入有效缓解了源域和目标域的深层次特征分布差异,增强了模型的迁移学习能力,解决了模型在复杂密集场景中的局限性,进一步验证了松-紧耦合包络样本和对抗域适应的有效性.

4.5 预测轨迹可视化对比

对 HEADA-LTCTPM、T-GCN 和 Social-STGCNN 三种模型进行定性分析. 图 7 展示了 3 种模型在稀疏场景和多个密集场景下行人轨迹的实际预测结果. 红色实线表示历史轨迹,蓝色实线表示真实轨迹,黄色虚线表示模型的预测轨迹.

图 7(a)展示了在稀疏场景中多人相遇和静止的情况,HEADA-LTCTPM 模型能够有效预测行人避让迎面相遇行人的轨迹,表现出较高的预测精度. 相比之下,T-GCN 模型在预测相遇行人时出现了接近相撞的情况,而尽管 Social-STGCNN 模型预测的轨迹避免了相撞情况的出现,但其预测的轨迹末端与真实轨迹存在较大的偏差,重合度较低.

图 7(b)和图 7(c)展示了在密集场景中多人并行和相遇的情况. 尽管源域和目标域之间的域差异相对较小,但该场景中行人密度较高,这成为行人间的交互关系更为复杂的诱因. HEADA-LTCTPM 模型能够准确预测行人结伴而行以及有效避让迎面相遇行人的情况,

表 4 HEADA-LTCTPM 模型与参考模型的 ADE 比较

单位:m

源域	目标域	SGCN ^[52]	STGAT ^[53]	Transformer ^[54]	Social-STGCNN ^[25]	T-GNN ^[35]	MESC-HEDA ^[13]	HEADA-TCTPM
A	B	0.84	1.19	1.04	1.13	0.92	0.90	0.63
	C	0.57	0.82	0.84	0.72	0.99	0.53	0.43
	D	0.45	0.54	0.65	0.67	0.77	0.44	0.37
	E	0.35	0.52	0.49	0.62	0.79	0.40	0.32
B	A	2.17	1.42	1.41	1.86	1.39	1.12	0.80
	C	0.81	0.84	1.04	0.97	0.83	0.56	0.55
	D	1.76	1.47	2.14	1.64	1.17	0.53	0.51
	E	1.06	0.95	1.29	1.10	0.88	0.79	0.56
C	A	0.79	0.88	0.53	0.68	0.66	0.66	0.57
	B	0.31	0.42	0.48	0.33	0.31	0.31	0.24
	D	0.30	0.39	0.27	0.33	0.33	0.29	0.25
	E	0.25	0.34	0.23	0.27	0.28	0.25	0.22
D	A	0.75	0.83	0.69	0.80	0.74	0.67	0.56
	B	0.59	0.73	0.81	0.67	0.48	0.42	0.35
	C	0.40	0.60	0.50	0.48	0.46	0.38	0.34
	E	0.28	0.34	0.22	0.36	0.32	0.26	0.23
E	A	0.75	0.65	0.70	0.70	0.72	0.60	0.51
	B	0.51	0.40	0.43	0.46	0.45	0.41	0.36
	C	0.41	0.60	0.41	0.45	0.42	0.39	0.34
	D	0.29	0.46	0.31	0.36	0.36	0.29	0.24
平均值		0.68	0.72	0.72	0.73	0.66	0.51	0.42

表 5 HEADA-LTCTPM 模型与参考模型的 FDE 比较

单位:m

源域	目标域	SGCN ^[52]	STGAT ^[53]	Transformer ^[54]	Social-STGCNN ^[25]	T-GNN ^[35]	MESC-HEDA ^[13]	HEADA-TCTPM
A	B	1.47	2.22	1.77	1.71	1.38	1.43	1.04
	C	0.93	1.61	1.16	0.85	1.56	0.78	0.64
	D	0.58	1.08	0.93	0.96	1.14	0.66	0.57
	E	0.56	1.05	0.64	0.78	1.13	0.56	0.53
B	A	3.01	2.18	2.25	2.61	2.40	1.89	1.42
	C	1.49	1.65	1.84	1.59	1.48	1.02	1.01
	D	3.31	2.76	3.91	2.93	2.02	0.98	0.92
	E	2.00	1.75	2.32	1.97	1.63	1.45	1.02
C	A	1.38	1.67	0.83	1.17	1.11	1.12	0.90
	B	0.59	0.88	0.73	0.55	0.53	0.51	0.37
	D	0.56	0.79	0.48	0.59	0.58	0.54	0.45
	E	0.50	0.71	0.35	0.49	0.52	0.45	0.39
D	A	1.27	1.55	1.03	1.28	1.23	1.28	1.01
	B	1.06	1.37	1.39	1.15	0.79	0.77	0.58
	C	0.73	1.24	0.84	0.82	0.84	0.72	0.60
	E	0.48	0.70	0.37	0.60	0.52	0.45	0.39
E	A	1.24	1.16	1.24	1.13	1.22	1.11	0.84
	B	0.85	0.79	0.79	0.77	0.77	0.75	0.68
	C	0.63	1.26	0.71	0.85	0.76	0.74	0.63
	D	0.52	0.88	0.55	0.63	0.64	0.51	0.42
平均值		1.16	1.36	1.21	1.17	1.11	0.89	0.72

与真实轨迹的重合度较高。相比之下,尽管T-GCN模型和Social-STGCNN模型在轨迹总体趋势上与真实轨迹大致吻合,但预测的轨迹和真实轨迹的重合度表现较差,特别是Social-STGCNN模型,预测轨迹与真实轨迹存在明显的偏差。

图7(d)和图7(e)展示了在复杂密集场景中多人并行、相遇、静止和急停等多种情况。受制于场景中行人密度较高,致使行人间的交互关系非常复杂。从图7(e)可以看出,尽管Social-STGCNN模型能够预测静止行人的状态,但在预测不同方向相遇行人的结伴而行时的精度较差,甚至个别行人的预测轨迹出现了相撞的情况。此外,受制于Social-STGCNN模型未考虑域适应问题,部分预测轨迹的长度远超真实轨迹,严重偏离了真实轨迹的情况。相比之下,HEADA-LTCTPM模型能够准确预测符合现实的行人轨迹,包括不同方向的行人

结伴而行、相遇、静止和急停的行为,充分体现了松-紧耦合包络样本构造模块和对抗域适应模块的有效性。

4.6 特征分布可视化对比

本文采用统一流形逼近与投影(Uniform Manifold Approximation and Projection, UMAP)方法对不同场景任务的目标域和源域的特征表示进行可视化。本文对复杂密集场景的轨迹预测任务(A→C、C→A、B→C、C→B、D→C、C→D)的目标轨迹域和源轨迹域的特征表示进行分析。源域和目标域样本分别用红色和蓝色标记。“w/o DA”表示未引入域适应模块的基准模型,“w/ DA”表示引入对抗域适应模块的改进模型。

图8显示,受制于ETH场景样本量较少,在未引入域适应模块(w/o DA)的情况下,源域和目标域的特征分布重叠十分有限,整体的分布差异明显。在引入对抗域适应模块(w/ DA)后,ETH场景的特征分布几乎包含

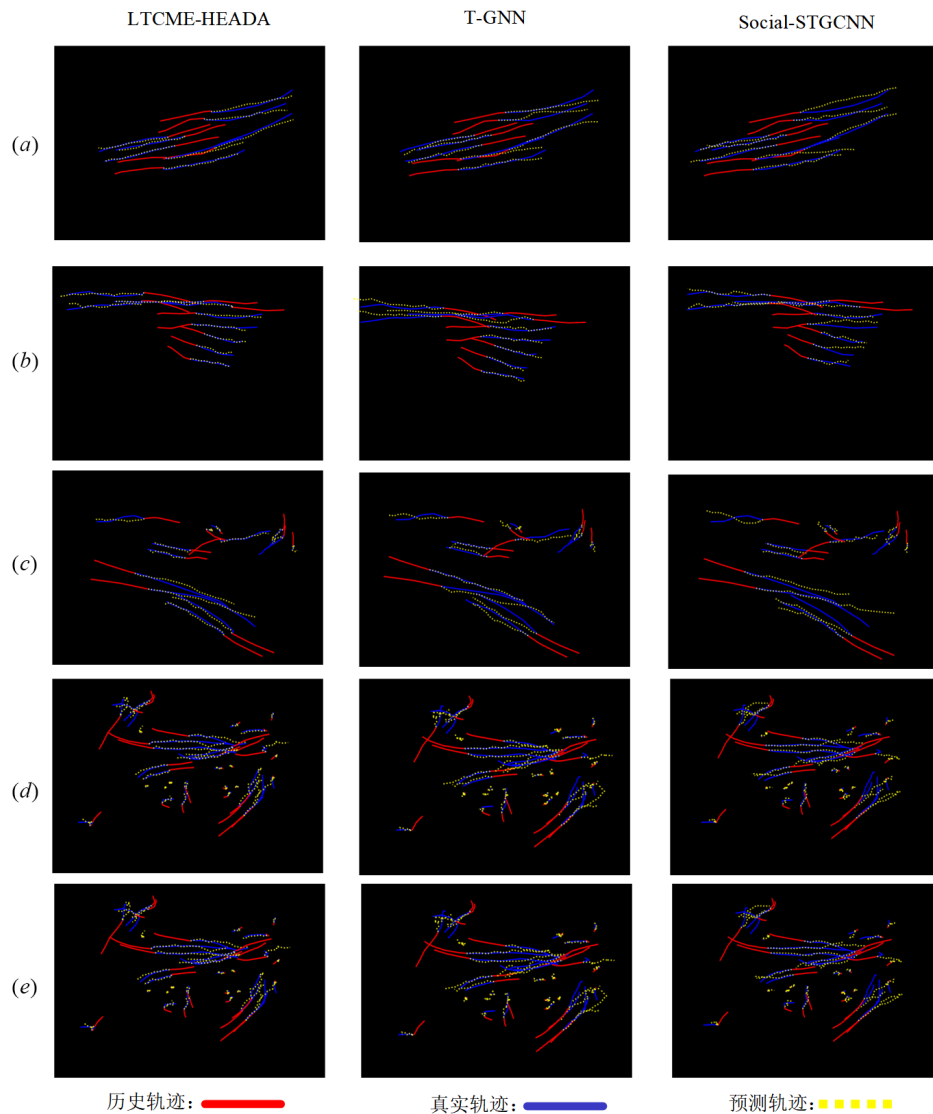


图7 预测轨迹的可视化

于 UNIV 场景的特征分布范围内,彰显了良好的改进效果.

图 9 显示,在未引入域适应模块(w/o DA)的情况下,源域和目标域的特征分布虽存在部分重叠区域,但由于 ZARA1 和 UNIV 场景都属于密集场景,故样本量

较大,整体分布差异依旧明显,尤其可以观察到目标域存在部分簇结构完全独立于源域的特征分布之外.在引入对抗域适应模块(w/ DA)后,源域和目标域的特征分布表现出了更高的重叠度,从而提升了目标域的预测精度.

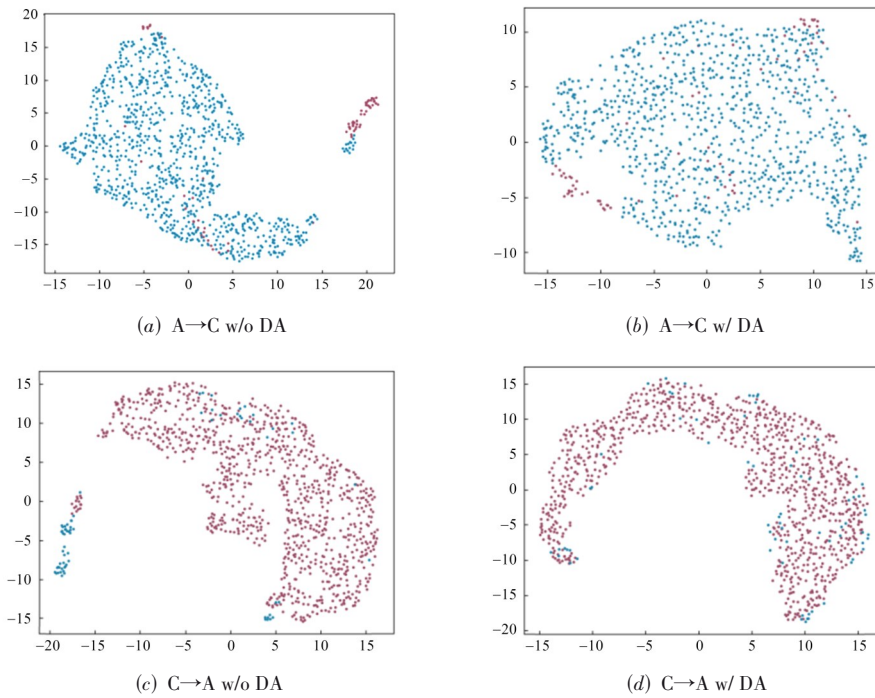


图 8 场景 A 和场景 C 特征的可视化结果

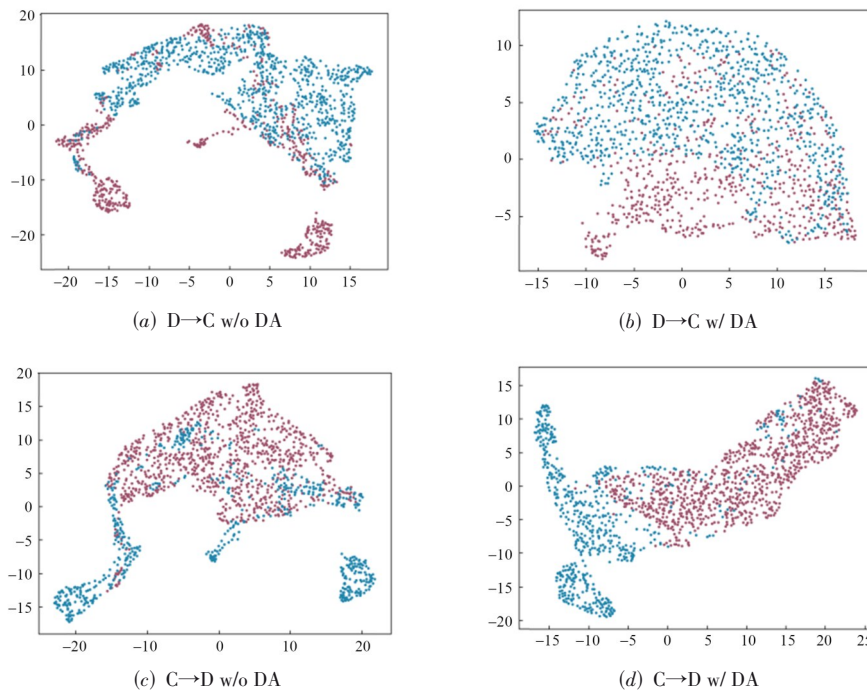


图 9 场景 C 和场景 D 特征的可视化结果

图 10 显示,在引入对抗域适应模块(w/ DA)后,源域和目标域的特征分布趋于一致, HOTEL 域的独立簇结构消失,预测精度提升, ADE 指标分别下降至 0.55 和 0.24.

通过多组跨域行人轨迹预测任务的特征分布可视

化验证了对抗域适应模块的有效性. 实验结果表明,无论是在稀疏场景还是密集场景中,引入对抗域适应模块后,能够提高源域和目标域特征分布的重叠度,有效缓解不同轨迹域之间的分布差异,提升模型的跨域轨迹预测的能力.

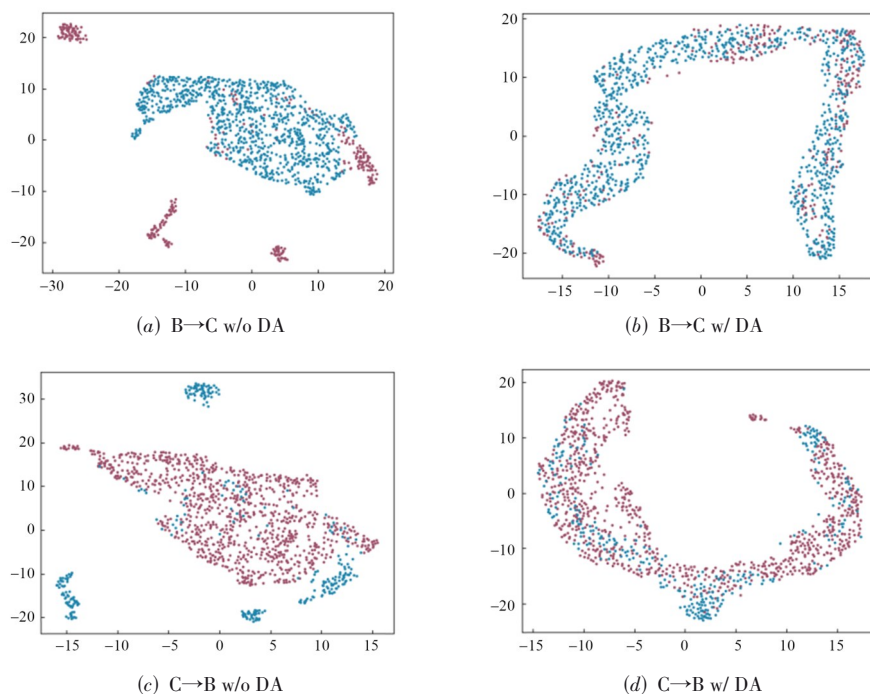


图 10 场景 C 和场景 B 特征的可视化结果

4.7 时间开销对比

在前述同样的实验设置下,对各模型的时间开销进行了对比分析,如表 6 所示.

从表 6 的实验结果可以看出, HEADA-LTCTPM 模型在保持较高预测精度的同时,测试时间远低于 Transformer 模型,与其他模型相当. 具体来说, HEADA-LTCTPM 模型在 20 项轨迹预测任务中的平均测试运行时间为 44.01 s,仅为 Transformer 模型平均训练时间的五分之一. 特别值得注意的是,在复杂场景的测试中,模型处理同一时段包含多个行人的轨迹样本平均用时 0.16 s,这一性能指标完全满足实际场景的应用需求.

5 结束语

针对现有行人轨迹预测方法存在的行人社交耦合关系异质性、依赖预定义统计度量等问题,本文提出了一种分级包络对抗域适应和松-紧耦合轨迹预测模型 (HEADA-LTCTPM). 该模型构建松耦合与紧耦合的个体层次与局部层次包络样本,决策融合实现轨迹预测,使得模型能够更好地适配各类不同的场景,缓解场景内行人社交耦合关系存在异质性的问题,显著提升了

人流密集场景中行人轨迹预测的准确性和鲁棒性. 进一步地,该方法通过引入对抗域适应模块,模型能够自适应地学习域不变特征,避免了传统方法中人为设计距离度量的局限性. 通过对抗训练的方式,生成器和判别器进行博弈,系统性地学习具有域不变性的深层特征,从而在特征层面实现跨域分布对齐的对抗学习目标.

实验结果表明, HEADA-LTCTPM 模型比 MESHEDA 模型具有更高准确性, ADE 指标与 FDE 指标分别下降了 17.6% 和 19.1%,在时间开销上也满足应用标准,综合性能优异. 定性实验的结果表明, HEADA-LTCTPM 模型在复杂密集场景中多人并行、相遇、静止和急停等多种情况下,预测的轨迹与真实轨迹的重合度明显高于 Social-STGCNN 模型和 T-GNN 模型,预测结果更贴近真实的轨迹. 此外,引入对抗域适应模块后的源域和目标域的特征分布重叠度明显增加,有效缓解不同轨迹域之间的分布差异. 定量实验的结果表明, HEADA-LTCTPM 模型相较于其他基线模型表现出了优异的预测性能. 在 20 项跨域任务中,共有 18 项任务中取得了最优的 ADE 表现. 具体而言, HEADA-

表 6 HEAD A-LTCTPM 模型与参考模型的测试运行时间比较

单位:s

源域	目标域	SGCN ^[52]	STGAT ^[53]	Transformer ^[54]	Social-STGCNN ^[25]	T-GNN ^[35]	MESC-HEDA ^[13]	HEAD A-LTCTPM
A	B	9.60	11.40	31.00	7.26	6.76	6.13	9.50
	C	49.95	26.59	944.75	47.56	47.67	49.43	147.94
	D	10.83	18.16	79.22	7.43	7.64	7.15	17.74
	E	19.21	24.30	217.13	16.19	16.05	14.66	40.24
B	A	3.18	6.29	11.86	0.84	0.84	0.87	1.64
	C	50.68	26.90	1 109.46	48.41	47.20	49.81	148.78
	D	11.05	18.16	86.09	7.22	7.88	8.18	18.26
	E	18.88	24.08	240.16	15.07	16.06	15.24	42.53
C	A	3.13	6.34	11.95	0.90	0.94	0.78	1.96
	B	6.35	11.36	43.83	3.59	3.79	4.31	9.94
	D	10.69	18.13	109.44	7.25	7.61	7.73	18.34
	E	21.84	24.08	243.73	15.44	15.84	15.61	43.16
D	A	4.07	6.65	12.01	0.89	0.89	0.88	1.85
	B	7.21	11.22	41.31	3.61	3.85	3.47	9.67
	C	54.28	26.78	1 192.60	50.86	47.47	49.88	148.94
	E	20.66	24.62	229.39	14.14	15.87	14.83	42.14
E	A	3.33	6.64	11.92	0.83	0.87	0.77	1.72
	B	6.78	11.58	42.50	3.36	3.93	3.47	9.62
	C	53.24	27.52	1 147.74	50.15	49.13	49.47	148.23
	D	11.26	18.48	81.63	7.32	7.55	7.24	17.93
平均值		18.81	17.46	294.39	15.42	15.39	15.50	44.01

LTCTPM 模型相较于 MESC-HEDA 模型, ADE 和 FDE 指标分别降低了 17.6% 和 19.1%; 相较于 T-GNN 模型, 性能提升更为显著, 两项关键指标分别降低了 36.3% 和 35.1%。在计算效率方面, HEAD A-LTCTPM 模型展现出一定的优势, 其训练和推理时间均大幅优于 Transformer 模型, 同时在最具挑战性的复杂密集场景测试中, 仍能保持出色的实时预测能力, 完全满足实际部署的时效性要求。综合评估预测精度与计算效能, HEAD A-LTCTPM 模型展现了当前最优的整体性能。

未来拟展开如下进一步研究: (1) 本文主要聚焦于俯瞰视角的预测, 后续工作将进一步拓展至第一人称视角, 以提升模型在多样化场景中的适用性; (2) 本文仅考虑行人之间的交互作用, 未来拟进一步探索异质对象的轨迹预测模型, 以应对复杂多变的现实环境。

参考文献

- [1] 孔玮, 刘云, 李辉, 等. 基于全局自适应有向图的行人轨迹预测[J]. 电子学报, 2022, 50(8): 1905-1916.
KONG W, LIU Y, LI H, et al. Pedestrian trajectory prediction based on global adaptive directed graph[J]. Acta Electronica Sinica, 2022, 50(8): 1905-1916. (in Chinese)
- [2] ZHANG H J, LIU Y J, WANG C, et al. Research on a pedestrian crossing intention recognition model based on nat-

ural observation data[J]. Sensors, 2020, 20(6): 1776.

- [3] MOUSSAÏD M, PEROZO N, GARNIER S, et al. The walking behaviour of pedestrian social groups and its impact on crowd dynamics[J]. PLoS One, 2010, 5(4): e10047.
- [4] HELBING D, MOLNÁR P. Social force model for pedestrian dynamics[J]. Physical Review E, 1995, 51(5): 4282-4286.
- [5] SCHNEIDER N, GAVRILA D M. Pedestrian path prediction with recursive Bayesian filters: A comparative study[M]// Pattern Recognition. Berlin: Springer, 2013: 174-183.
- [6] ZHU W J, LIU Y H, ZHANG M Y, et al. Reciprocal consistency prediction network for multi-step human trajectory prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(6): 6042-6052.
- [7] LI Y M, LIANG R Q, WEI W, et al. Temporal pyramid network with spatial-temporal attention for pedestrian trajectory prediction[J]. IEEE Transactions on Network Science and Engineering, 2022, 9(3): 1006-1019.
- [8] 桑海峰, 陈旺兴, 王海峰, 等. 基于多模式时空交互的行人轨迹预测模型[J]. 电子学报, 2022, 50(11): 2806-2812.
SANG H F, CHEN W X, WANG H F, et al. Pedestrian trajectory prediction model based on multi-mode spatio-temporal interaction[J]. Acta Electronica Sinica, 2022, 50(11): 2806-2812. (in Chinese)

- [9] ZHANG C, NI Z J, BERGER C. Spatial-temporal-spectral LSTM: A transferable model for pedestrian trajectory prediction[J]. *IEEE Transactions on Intelligent Vehicles*, 2024, 9(1): 2836-2849.
- [10] ZHI C Y, SUN H J, XU T. Adaptive trajectory prediction without catastrophic forgetting[J]. *The Journal of Supercomputing*, 2023, 79(14): 15579-15596.
- [11] LIU Y, ZHANG Y X, LI K M, et al. Knowledge-aware graph transformer for pedestrian trajectory prediction[C]//2023 IEEE 26th International Conference on Intelligent Transportation Systems. Piscataway: IEEE, 2024: 4360-4366.
- [12] POIBRENSKI A, NOZARIAN F, REZAEIANARAN F, et al. Uncertainty-aware pseudo labels for domain adaptation in pedestrian trajectory prediction[C]//2023 IEEE 26th International Conference on Intelligent Transportation Systems. Piscataway: IEEE, 2024: 5771-5777.
- [13] 李勇明, 李文正, 张小恒, 等. 基于分级包络域适应的行人轨迹预测模型[J]. *电子学报*, 2025, 53(4): 1308-1321.
LI Y M, LI W Z, ZHANG X H, et al. Pedestrian trajectory prediction model based on hierarchical envelope domain adaptation[J]. *Acta Electronica Sinica*, 2025, 53(4): 1308-1321. (in Chinese)
- [14] FERRER G, SANFELIU A. Behavior estimation for a complete framework for human motion prediction in crowded environments[C]//2014 IEEE International Conference on Robotics and Automation. Piscataway: IEEE, 2014: 5940-5945.
- [15] YANG B, ZHU J R, HU C, et al. Faster pedestrian crossing intention prediction based on efficient fusion of diverse intention influencing factors[J]. *IEEE Transactions on Transportation Electrification*, 2024, 10(4): 9071-9087.
- [16] 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. Piscataway: IEEE, 2016: 961-971.
- [17] XUE H, HUYNH D Q, REYNOLDS M. SS-LSTM: A hierarchical LSTM model for pedestrian trajectory prediction[C]//2018 IEEE Winter Conference on Applications of Computer Vision. Piscataway: IEEE, 2018: 1186-1194.
- [18] HUG R, BECKER S, HÜBNER W, et al. Particle-based pedestrian path prediction using LSTM-MDL models[C]//2018 21st International Conference on Intelligent Transportation Systems (ITSC). New York: ACM, 2018: 2684-2691.
- [19] BISAGNO N, ZHANG B, CONCI N. Group LSTM: Group trajectory prediction in crowded scenarios[M]//*Computer Vision - ECCV 2018 Workshops*. Cham: Springer, 2019: 213-225.
- [20] YAO Y, ATKINS E, JOHNSON-ROBERSON M, et al. BiTraP: Bi-directional pedestrian trajectory prediction with multi-modal goal estimation[J]. *IEEE Robotics and Automation Letters*, 2021, 6(2): 1463-1470.
- [21] GUPTA A, JOHNSON J, LI F F, et al. Social GAN: Socially acceptable trajectories with generative adversarial networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 2255-2264.
- [22] SADEGHIAN A, KOSARAJU V, SADEGHIAN A, et al. SoPhie: An attentive GAN for predicting paths compliant to social and physical constraints[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2020: 1349-1358.
- [23] AMIRIAN J, HAYET J B, PETTRÉ J. Social ways: Learning multi-modal distributions of pedestrian trajectories with GANs[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE, 2020: 2964-2972.
- [24] JUN O Y, SHI Q W, WANG X X, et al. Pedestrian trajectory prediction based on GAN and attention mechanism[J]. *Laser & Optoelectronics Progress*, 2020, 57(14): 141016.
- [25] MOHAMED A, QIAN K, 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. Piscataway: IEEE, 2020: 14412-14420.
- [26] LI K M, EIFFERT S, SHAN M, et al. Attentional-GCNN: Adaptive pedestrian trajectory prediction towards generic autonomous vehicle use cases[C]//2021 IEEE International Conference on Robotics and Automation. Piscataway: IEEE, 2021: 14241-14247.
- [27] CHEN W X, SANG H F, WANG J Y, et al. DSTCNN: Deformable spatial-temporal convolutional neural network for pedestrian trajectory prediction[J]. *Information Sciences*, 2024, 666: 120455.
- [28] ZHANG C, BERGER C, DOZZA M. Social-IWSTCNN: A social interaction-weighted spatio-temporal convolutional neural network for pedestrian trajectory prediction in urban traffic scenarios[C]//2021 IEEE Intelligent Vehicles Symposium. Piscataway: IEEE, 2021: 1515-1522.
- [29] LV P, WANG W T, WANG Y X, et al. SSAGCN: Social

soft attention graph convolution network for pedestrian trajectory prediction[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, 35(9): 11989-12003.

- [30] CHEN W X, SANG H F, WANG J Y, et al. IMGCN: Interpretable masked graph convolution network for pedestrian trajectory prediction[J]. *Transportmetrica B: Transport Dynamics*, 2024, 12(1): 2389896.
- [31] PI L, ZHANG Q, YANG L F, et al. Social interaction model enhanced with speculation stage for human trajectory prediction[J]. *Robotics and Autonomous Systems*, 2023, 161: 104352.
- [32] CHEN W X, SANG H F, WANG J Y, et al. DSTIGCN: Deformable spatial-temporal interaction graph convolution network for pedestrian trajectory prediction[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2025, 26(5): 6923-6935.
- [33] MI J, ZHANG X X, ZENG H H, et al. DERGCN: Dynamic-Evolving graph convolutional networks for human trajectory prediction[J]. *Neurocomputing*, 2024, 569: 127117.
- [34] CHEN W X, SANG H F, WANG J Y, et al. IGGCN: Individual-guided graph convolution network for pedestrian trajectory prediction[J]. *Digital Signal Processing*, 2025, 156: 104862.
- [35] XU Y, WANG L C, WANG Y Z, et al. Adaptive trajectory prediction via transferable GNN[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2022: 6510-6521.
- [36] HUANG P X, CUI Z H, LI J, et al. Cross-domain trajectory prediction with CTP-Net[M]//Artificial Intelligence. Cham: Springer, 2022: 80-92.
- [37] HUANG R, TOMPKINS A, PAGNUCCO M, et al. Towards single source domain generalisation in trajectory prediction: A motion prior based approach[C]//Proceedings of The 2nd Conference on Lifelong Learning Agents. Cambridge: PMLR, 2023: 227-243.
- [38] TANG X Q, SUN W G, HU S Y, et al. MS-net: A multi-path sparse model for motion prediction in multi-scenes[J]. *IEEE Robotics and Automation Letters*, 2024, 9(1): 891-898.
- [39] GENG M S, LI J Y, LI C J, et al. Adaptive and simultaneous trajectory prediction for heterogeneous agents via transferable hierarchical transformer network[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2023, 24(10): 11479-11492.
- [40] YANG H F, CHEN Y, CAI J H, et al. Cross-domain pedestrian trajectory prediction via behavioral pattern-aware multi-instance GCN[J]. *Knowledge-Based Systems*, 2025, 329: 114266.
- [41] MA J, ZHANG Y F, LI Y M, et al. Deep dual-side learning ensemble model for Parkinson speech recognition[J]. *Biomedical Signal Processing and Control*, 2021, 69: 102849.
- [42] MCCANDLESS D. *Knowledge is Beautiful*[M]. London: William Collins, 2014: 1-256.
- [43] ROWLEY J. The wisdom hierarchy: Representations of the DIKW hierarchy[J]. *Journal of Information Science*, 2007, 33(2): 163-180.
- [44] XIA S Y, LIU Y S, DING X, et al. Granular ball computing classifiers for efficient, scalable and robust learning[J]. *Information Sciences*, 2019, 483: 136-152.
- [45] BOSANCIC B. Information in the knowledge acquisition process[J]. *Journal of Documentation*, 2016, 72(5): 930-960.
- [46] LI F, ZHANG X H, WANG P, et al. Deep instance envelope network-based imbalance learning algorithm with multilayer fuzzy C-means clustering and minimum inter-layer discrepancy[J]. *Applied Soft Computing*, 2022, 123: 108846.
- [47] LI F, LI Y M, SHEN Y H, et al. Deep fuzzy envelope sample generation mechanism for imbalanced ensemble classification[J]. *IEEE Transactions on Fuzzy Systems*, 2024, 32(3): 1248-1262.
- [48] PELLEGRINI S, ESS A, SCHINDLER K, et al. You'll never walk alone: Modeling social behavior for multi-target tracking[C]//2009 IEEE 12th International Conference on Computer Vision. Piscataway: IEEE, 2010: 261-268.
- [49] LERNER A, CHRYSANTHOU Y, LISCHINSKI D. Crowds by example[J]. *Computer Graphics Forum*, 2007, 26(3): 655-664.
- [50] 李勇明, 李文正, 张小恒, 等. 一种基于分级包络域适应的行人轨迹预测方法: CN118918140A[P]. 2024-11-04. LI Y M, LI W Z, ZHANG X H, et al. A pedestrian trajectory prediction approach based on hierarchical envelope domain adaptation: CN118918140A[P]. 2024-11-04. (in Chinese)
- [51] LIU S H, LIU H B, WANG Y S, et al. MDST-DGCN: A multilevel dynamic spatiotemporal directed graph convolutional network for pedestrian trajectory prediction[J]. *Computational Intelligence and Neuroscience*, 2022, 2022(1): 4192367.
- [52] SHI L S, WANG L, LONG C J, et al. SGCN: Sparse

graph convolution network for pedestrian trajectory prediction[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2021: 8990-8999.

- [53] HUANG Y F, BI H K, LI Z X, et al. STGAT: Modeling spatial-temporal interactions for human trajectory predic-

tion[C]//2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2020: 6271-6280.

- [54] GIULIARI F, HASAN I, CRISTANI M, et al. Transformer networks for trajectory forecasting[C]//2020 25th International Conference on Pattern Recognition. Piscataway: IEEE, 2021: 10335-10342.

作者简介



李勇明 男, 1976年9月生, 四川绵阳人. 重庆大学微电子与通信工程学院教授, 博士生导师. 主要研究方向为医学信号处理、机器学习.
E-mail: yongmingli@cqu.edu.cn



王 品 女, 1979年11月生, 江苏盐城人. 博士, 副教授, 硕士生导师. 主要研究方向为图像处理与识别.
E-mail: wangpin@cqu.edu.cn



胡 杰 男, 2001年9月生, 四川遂宁人. 重庆大学微电子与通信工程学院硕士研究生. 主要研究方向为图神经网络、行人轨迹预测.
E-mail: 2357590375@qq.com



李文正 男, 1999年10月生, 福建漳州人. 重庆大学微电子与通信工程学院硕士研究生. 主要研究方向为图神经网络、行人轨迹预测.
E-mail: fjxmlwz9910@163.com



张小恒 男, 1980年10月生, 四川达州人. 博士研究生, 副教授. 主要研究方向为医学信号处理、机器学习、行人轨迹预测.
E-mail: 7818320@qq.com