

群组感知的行人轨迹预测方法研究

王汝言^{1,2,3}, 周玉蝶^{1,2,3}, 吴大鹏^{1,2,3}, 段昂^{1,2,3}, 崔亚平^{1,2,3}, 何鹏^{1,2,3}

(1. 重庆邮电大学通信与信息工程学院, 重庆 400065; 2. 先进网络与智能互联技术重庆市高校重点实验, 重庆 400065;
3. 泛在感知与互联重庆市重点实验室, 重庆 400065)

摘要: 自动驾驶场景下, 大多数方法没有对行人群体进行建模, 这样会对道路的安全造成影响。因此, 提出了一种针对群组感知的行人轨迹预测网络 (GPCNet)。具体来说, 在组内, 从个体层面学习行人之间的交互, 考虑不同行人的偏好问题。在组间, 从群体层面学习行人组间的交互, 使用社会力模型考虑行人轨迹的碰撞问题。仿真结果表明, 与常用的轨迹预测方法相比, GPCNet在ETH和UCY数据集上的性能提高了约75.4%。

关键词: 自动驾驶; 轨迹预测; 行人群体; 道路安全

中图分类号: TP18

文献标志码: A

DOI: 10.11959/j.issn.1000-436x.2024224

Pedestrian trajectory prediction method based on group perception

WANG Ruyan^{1,2,3}, ZHOU Yudie^{1,2,3}, WU Dapeng^{1,2,3}, DUAN Ang^{1,2,3}, CUI Yaping^{1,2,3}, HE Peng^{1,2,3}

1. School of Communications and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
2. Advanced Network and Intelligent Connection Technology Key Laboratory of Chongqing Education Commission of China, Chongqing 400065, China
3. Chongqing Key Laboratory of Ubiquitous Sensing and Networking, Chongqing 400065, China

Abstract: Most methods do not model the pedestrian groups in autonomous driving, which will have an impact on road traffic safety. Therefore, a group perception pedestrian trajectory prediction network called GPCNet was proposed. Specifically, in intra-group, the interaction between pedestrian was learned at the individual level and the preference issue of different pedestrian was considered. In inter-group, the interaction between pedestrian groups was learned at the group level and the collision issue of pedestrian trajectory was considered using the social force model. Simulation results demonstrate that GPCNet improves the performance on the ETH and UCY datasets by 75.4% compared to the commonly used trajectory prediction methods.

Keywords: autonomous driving, trajectory prediction, pedestrian group, road safety

收稿日期: 2024-07-10; 修回日期: 2024-11-05

通信作者: 崔亚平, cuiyp@cqupt.edu.cn

基金项目: 国家自然科学基金资助项目 (No.61801065, No.62271096, No.61871062, No.U20A20157, No.62061007, No.U24A20211); 重庆市教委科学技术研究基金资助项目 (No.KJQN202000603, No.KJQN202300621); 重庆市自然科学基金资助项目 (No. CSTB2024NSCQ-LZX0124, No. CSTB2022NSCQ-MSX0468, No.cstc2020jcyjzdxmX0024, No. CSTB2023NSCQ-LZX0134); 重庆市高校创新研究群体基金资助项目 (No.CXQT20017); 重邮信通青创团队支持计划基金资助项目 (No.SCIE-QN-2022-04); 四川省重点研发计划基金资助项目 (No.2024YFHZ0093)

Foundation Items: The National Natural Science Foundation of China (No.61801065, No.62271096, No.61871062, No.U20A20157, No.62061007, No.U24A20211), The Science and Technology Research Program of Chongqing Municipal Education Commission (No.KJQN202000603, No.KJQN202300621), The Natural Science Foundation of Chongqing (No.CSTB2024NSCQ-LZX0124, No. CSTB2022NSCQ-MSX0468, No.cstc2020jcyjzdxmX0024, No. CSTB2023NSCQ-LZX0134), The University Innovation Research Group of Chongqing (No.CXQT20017), The Youth Innovation Group Support Program of ICE Discipline of CQUPT (No.SCIE-QN-2022-04), Sichuan Science and Technology Program (No.2024YFHZ0093)

0 引言

行人轨迹预测在许多重要的应用领域,如自动驾驶车辆、服务机器人和高级监视系统受到了越来越多的关注^[1]。行人轨迹预测旨在根据观察到的历史轨迹来预测未来的位置。这些运动模式通常取决于行人的当前位置与历史运动轨迹,以及环境中其他行人和障碍物等的影响。随着无线通信技术的快速发展,自动驾驶在智能交通系统中越来越受欢迎^[2]。在自动驾驶中,行人轨迹预测是感知和决策模块的桥梁^[3]。行人轨迹预测的准确性直接影响车辆的安全性和行驶效率,能够帮助车辆更好地理解行人的行为意图,从而做出合理的驾驶决策。特别是在复杂的城市环境中,如图1所示,道路两旁的行人错综复杂,由于行人行为的多样性和随机性,这使得轨迹预测仍是一个具有挑战性的问题。行人轨迹预测在自动驾驶中发挥着不可或缺的作用。它不仅能够提高车辆的安全性和行驶效率,还能够使自动驾驶技术更好地适应复杂的交通环境,为人们带来更加便捷和安全的出行体验。然而,行人的运动轨迹往往受到多种因素的影响,包括个人的行为习惯、行人之间的人际关系、周围环境的变化等。因此,准确预测行人的轨迹仍具有挑战性。



图1 城市道路下的行人场景示意

现有的许多工作已经尝试使用各种不同的方法对行人的行为进行建模。基于物理的预测方法往往依赖于简化的数学模型,这类方法难以模拟复杂场景,在处理复杂环境和动态变化时存在局限性。例如,当环境中存在障碍物、车辆或其他动态物体时,行人的运动轨迹可能会突然发生变化,这种变化在基于物理的模型中难以准确预测,并且性能也

有待提高。近年来,随着深度学习的兴起,各种用于处理时序数据的模型被提出,使得基于神经网络的轨迹预测方法流行起来。通过使用循环神经网络(RNN, recurrent neural network)、生成对抗网络(GAN, generative adversarial network)和图卷积网络(GCN, graph convolutional network)等深度学习方法来对行人的轨迹进行研究。这些基于深度学习的模型能够从大量数据中自动提取行人轨迹的复杂特征,能够处理非线性数据和时序数据,有助于更准确地预测行人的轨迹,预测效果较传统轨迹预测方法也有了很大的提升。但是,深度学习模型通常具有较差的可解释性,这使得人们难以理解模型是如何做出预测的,从而导致对预测结果的信任度降低。并且,行人的行为通常会受到自身性格和习惯等的影响。同时,还会受到其他行人或同行者的影响。然而,大多数方法只考虑了外部因素,即行人与行人之间的社会交互。很少有方法将行人自身的习惯考虑在内。除此之外,大多数方法都没有考虑行人个体较为密集的场景,这样会导致所预测的轨迹不符合社会规范。

在密集人群环境中,大多数先前的工作是对行人个体进行建模,然而人群中多达70%的人 would 组成团体^[4]。尽管,这些方法可以从概念上捕捉到行人群体的运动行为。但是,在密集人群环境中此类方法难以同时学习个人和群体运动行为,难以更好地捕捉行人之间的社会交互,从而导致预测的性能较低。因此,本文考虑在密集人群场景下对行人个体进行建模,以同时学习个人和群体运动行为。这样既能将因个人习惯而形成的特定轨迹运动模式考虑在内,也能更好地学习行人之间社会交互和社交关系,从而提高行人轨迹预测的性能,生成更加符合社会规范的轨迹。因此,本文提出了一种针对群组感知的行人轨迹预测网络(GPCNet, group perception pedestrian trajectory prediction network),通过对组内和组间2个方面进行研究来学习行人的随机运动行为。首先,使用分组模块对行人个体进行分组,得到隶属于同一组的行人信息。传统轨迹预测方法通常关注单个行人或简单的行人之间的交互,而GPCNet引入分组模块能够更好地对密集人群环境下的行人个体进行研究。通过同时学习组内和组间的特征信息来提高轨迹预测的精度。其次,分别从组内和组间2个角度考虑行人个体,得到行

人的特征信息。传统的轨迹预测模型主要集中在行人与外部环境或其他行人之间的交互影响上,而忽略了行人个体的内在偏好,如行为习惯和个性化的运动模式等。而在 GPCNet 中,对于组内交互而言,考虑了行人的偏好问题,考虑因个人习惯而形成的特定运动模式这个因素。对于组间交互而言,通过使用物理方法与深度学习相结合的方式考虑行人轨迹的碰撞问题,减少行人之间碰撞发生的概率。除此之外,一般的深度学习模型往往缺乏可解释性,而 GPCNet 在提高轨迹预测性能的同时还增加了模型的可解释性。最后,通过基于编码器和解码器的轨迹预测框架来实现对行人轨迹的预测,生成行人未来的轨迹。

本文的主要贡献可概括如下。

1) 提出了一种针对群组感知的行人轨迹预测网络。将社会力模型引入基于深度学习的模型中来学习行人的运动模式,通过物理方法与深度学习相结合的方式提高模型的可解释性。

2) 提出了一种对行人群体进行分组的方法,从组内和组间 2 个角度捕获行人的特征信息。个体层面,将考虑不同行人的偏好问题。群体层面,使用碰撞预测网络来避免行人之间发生碰撞,减小碰撞发生的概率。

3) 在 ETH 和 UCY 数据集上的实验结果表明,GPCNet 在多个轨迹预测基准上实现了先进的性能。消融实验结果表明,GPCNet 中的每一个组件都有助于提高轨迹预测的性能,所有单独的组合都会导致性能降低。

1 相关工作

近年来,在轨迹预测方面取得了许多进展。先前许多关于轨迹预测的研究,有的使用基于物理的方法,例如通过运动学或者动力学模型来对行人的特征(速度、位置和角速度等)进行建模,从而预测出行人未来的轨迹。有的使用基于深度学习的方法利用观察到的运动轨迹,通过神经网络建立模型,从而在未来短时间内(如 5 s)预测行人的运动轨迹。但是很少有作品试图将基于物理和深度学习的方法结合起来进行预测。

1) 基于物理的方法

基于物理的方法通常使用基于牛顿定理的运动学或者动力学方程,生成行人未来的轨迹。例如,

文献[5]提出了社会力模型,通过使用目标的吸引力和其他智能体与障碍物之间的排斥力来对行人群体进行分析。对于人口密集环境中的许多任务,以前的方法只对人体运动做出了简单的假设。于是,文献[6]将社会力模型集成到多目标跟踪器中,描述群体行为分析等任务的个人和群体行为动态。文献[7]评估了行人的加速、减速和转弯等运动以及头部方向和不同的运动状态对性能的影响。文献[8]所使用的动态模型考虑了加速度、倾角、滚动和空气的驱动力与阻力的影响。文献[9]提出的动态贝叶斯网络运动模型,用于估计行人未来的位置、速度和穿越意图。文献[10]提出了一种将物理方法和序列方法相结合的轨迹预测方法。然而,基于物理的方法只能捕捉简单的交互特征,且在复杂密集的人群场景中所捕捉的交互特征会失效,难以完全学习复杂轨迹的结构特征和运动目标之间的交互特征。

2) 基于深度学习的方法

随着深度学习的发展,各种用于处理时序数据的模型被提出。对单个行人而言,文献[11]提出了基于社交的长短期记忆网络(LSTM, long short term memory),每个轨迹都使用一个 LSTM 模型,用于预测连续空间中的轨迹。对行人群体而言,文献[12]提出在密集人群环境中对群体关系进行建模,并研究此关系对轨迹预测的影响。具体来说,将行人之间的分组视为潜在变量,并与轨迹信息一起进行联合估计。最后,在给定的时间内生成可能的轨迹。文献[13]提出一种从群体层面来预测密集人群场景中行人未来轨迹的预测方法。该方法对具有相似运动趋势的轨迹进行聚类,通过这种方式,对同一组内的行人进行分类。然后,使用 LSTM 对未来的路径进行预测。虽然这些方法考虑了群体运动行为,但没有将行人的社会交互考虑在内。于是,文献[14]通过在密集人群环境中对行人进行分组来预测未来的轨迹。同时,还考虑了行人的社会交互和环境信息。文献[15]提出了一种基于群体的社会交互新模型,该方法使用 GCN 对群体进行建模,这样充分考虑了行人之间的社会交互。文献[16]提出一个新的大规模数据集用于研究密集人群环境下的行人轨迹预测。虽然,这些轨迹预测方法能够在密集人群环境下对行人群组进行研究。然而,不同行人会因其习惯而具有不同的运动模

式, 行人之间也应保持一定的距离从而避免碰撞的发生。

2 方法

本节介绍一种密集人群场景的行人轨迹预测方法, 从组内和组间 2 个方面来对目标行人的轨迹进行预测。

2.1 问题描述

行人轨迹预测问题可以被视为一个序列预测问题, 该问题从目标行人的历史轨迹和社会交互中提取线索, 以预测目标行人未来的路径。行人的轨迹通常可以分为 2 部分, 即观察到的轨迹和未来的轨迹。假设 N 个行人的历史轨迹为 $X = X_1, X_2, \dots, X_N$, 预测所得的未来轨迹为 $Y = Y_1, Y_2, \dots, Y_N$ 。因此, 行人 i 输入的轨迹定义为 $X_i^t = (x_i^t, y_i^t)$, 其中 $t = 1, \dots, t_{\text{obs}}$ 。未来的轨迹定义为 $Y_i^t = (x_i^t, y_i^t)$, 其中 $t = t_{\text{obs}} + 1, \dots, t_{\text{pred}}$ 。一般来说, 在图卷积网络中将节点图表示为 $\mathcal{G} = (V, E)$ 。其中, V 表示节点, E 表示边, 如图 2 所示。具体来说, 节点 $V = \{X_n | n \in [1, \dots, N]\}$ 表示行人的轨迹信息, 边 $E = \{e_{ij} | ij \in [1, \dots, N]\}$ 表示行人与行人之间的社会交互^[17]。

GPCNet 的整体框架如图 2 所示, 行人的历史轨迹经过编码器、分组模块、组内模块和组间模块, 最后通过解码器生成未来的轨迹。具体来说, 首先, 通过编码器对行人的坐标信息进行编码, 生成能够表示行人历史运动趋势的特征向量, 并从中提取出关键的时空特征, 这些特征包括行人的运动

模式 (如直线行走、曲线移动等)、速度变化、加速度等。编码器生成的特征表示作为分组模块的输入, 为之后的分组和进一步的关系建模提供数据支持。分组模块根据行人间的欧几里得距离来估计行人所属的分组信息。然后, 通过组内模块和组间模块来获取行人之间的交互信息。最后, 解码器将生成的特征信息, 并结合行人的历史轨迹信息预测出行人的未来轨迹。

2.2 分组模块

首先, 使用分组模块来估计行人所属的分组信息。其次, 根据每个行人的历史轨迹信息, 通过计算每对行人的欧几里得距离来测量所有行人对之间的特征相似性。然后, 通过卷积操作, 从输入信息中提取有用特征, 并将其映射到特征图上, 这些特征图可以用于进一步处理或估计阈值。接着, 通过计算得出的欧几里得距离, 选出所有隶属于同一组的行人。距离矩阵 A 和隶属于同一索引 γ 定义为

$$A_{ij} = \|F_{\phi}(X_i) - F_{\phi}(X_j)\|, \quad ij \in [1, \dots, N] \quad (1)$$

$$\gamma = \{\text{pair}(ij) | ij \in [1, \dots, N], i \neq j, A_{ij} \leq \pi\} \quad (2)$$

其中, $F_{\phi}(\cdot)$ 是可学习的卷积层, π 是可学习的阈值参数。

最后, 将属于同一个组的行人分到相关联的组中, 并分配他们的组索引。将组索引 G 表示为

$$G = \{G_k | G_k = \bigcup_{(ij) \in \gamma} \{ij\}\} \quad (3)$$

其中, $G_a \cap G_b = \emptyset$, 且 $a \neq b$ 。并且第 k 个组 G_k 是每组行人对集合的并集。

2.3 组内模块

为了更好地模拟行人之间的交互, 本文基于每

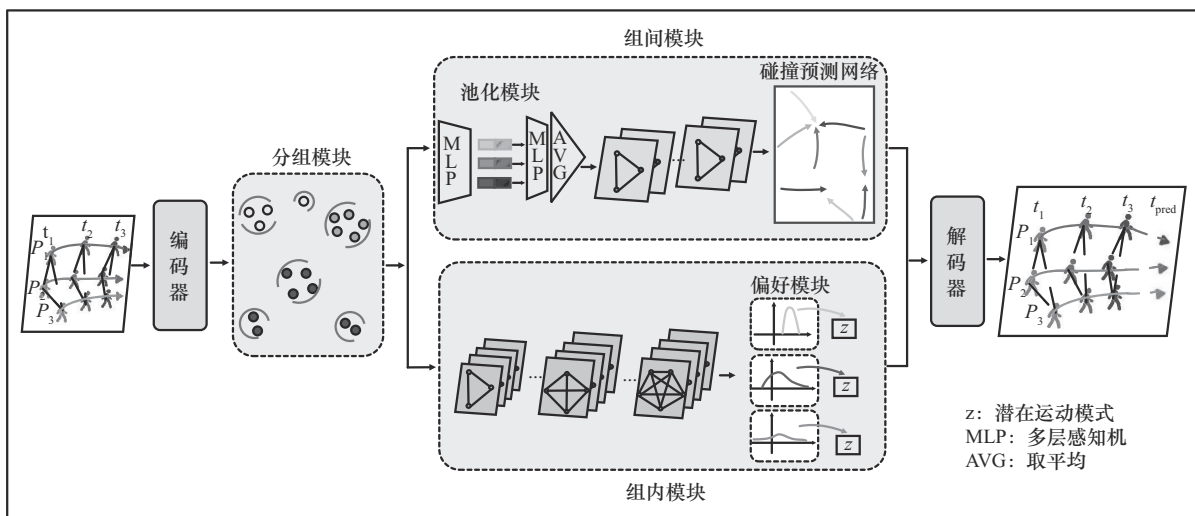


图 2 GPCNet 的整体框架

个行人及其邻域在每个时间步长的空间位置构建了一个交互图,如图3所示。使用GCN来捕捉组内行人之间的互动,并对组内交互进行建模。在每个时间步长 t ,空间图 $\mathcal{G}_s = (V_s, E_s)$ 表示每个行人及其相邻行人的位置坐标信息。同时, $V_s = \{X_n^t | t \in [1, \dots, T], n \in [1, \dots, N]\}$, $E_s = \{e_{ij}^t | i, j \in [1, \dots, N]\}$ 。其中, T 是观察到的轨迹的长度, X_n^t 表示行人的轨迹信息, e_{ij}^t 表示行人与行人之间的社会交互。

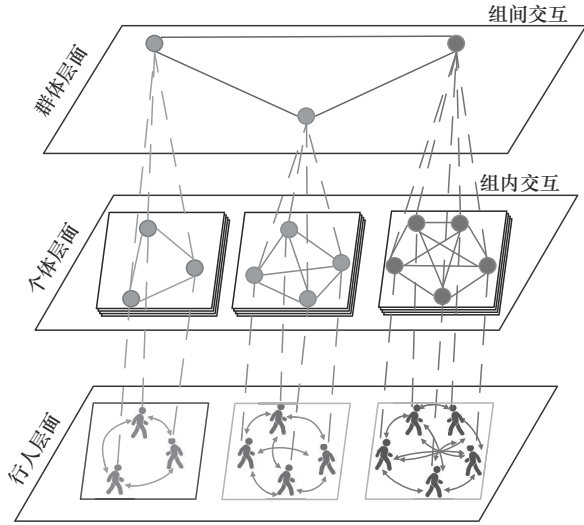


图3 组内和组间交互示意

首先,式(4)给出了一个GCN层用于更新图中每个节点的特征矩阵。

$$\mathbf{H}^{l+1} = \sigma \left(\sum_{i \in R} \hat{\mathbf{A}}^i \mathbf{H}^i \mathbf{W}^l \right) \quad (4)$$

其中, $\hat{\mathbf{A}}^i$ 表示图的归一化邻接矩阵, \mathbf{H}^i 是第 l 个GCN层的隐藏特征, \mathbf{W}^l 表示第 l 个GCN层的参数, $\sigma(\cdot)$ 表示非线性激活函数, R 表示实数。在这里,将每个节点的原始特征设置为行人的位置。图卷积网络的关键特征在于它可以在图中相邻节点之间传递信息,从而使有关行人交互信息的隐藏特征很容易被GCN捕获。在以行人位置为节点的GCN中,计算复杂度的核心在于信息传递的过程。GCN是基于邻接关系的节点进行特征聚合,模型会利用 $N \times N$ 的邻接矩阵将相邻节点的信息结合到当前节点的特征表示中,以捕捉行人之间的交互。当行人数量增加时,模型需要更多的计算资源来处理相邻节点之间的交互,导致计算效率显著下降。

由于不同成员会因其习惯而具有不同的运动模式,因此,本文还将在组内交互中考虑行人的偏好

问题,从个体层面学习特征信息。受轨迹预测方法的影响^[18-20],每个人的行为模式被建模为潜在分布,通过学习潜在分布来表示多模态分布并生成随机的未来状态。为了对行人个性化运动模式进行建模,认为同一运动模式的潜在变量分布相似,同时将不同运动模式的潜在变量分布分开。通过损失函数式(5)来学习潜在变量分布。

$$\begin{aligned} \mathcal{L}(\psi, \phi, \theta) = & \mathcal{L}_1(\phi, \theta) + \mathcal{L}_2(\phi, \psi) + \mathcal{L}_3(\phi) = \\ & \mathcal{L}_{\text{KL}}(q_\phi(z|x, y) || p_\theta(z|x)) - \\ & \lambda \mathbb{E}_{q_\phi(z|x, y)} [\lg g_\psi(y|x, z)] - \\ & \mu \mathbb{E}_{q_\phi(z)} \left[\lg \frac{h(z, x)}{\sum_{q_\phi(z)} h(z, x)} \right] \end{aligned} \quad (5)$$

其中, $\mathcal{L}_1(\phi, \theta)$ 是基于KL散度的损失项,用于确保历史轨迹和未来轨迹的潜在变量保持一致,保证模型对轨迹的潜在分布建模合理; $\mathcal{L}_2(\phi, \psi)$ 是重构误差项,通过生成分布 $g_\psi(y|x, z)$ 来对未来轨迹 y 进行建模,并通过对采样的潜在变量 z 进行优化,以生成最接近真实轨迹的预测; $\mathcal{L}_3(\phi)$ 表示分布区分项,通过密度比 $h(z, x)$ 来表示潜在变量 z 和历史轨迹 x 之间的互信息; $p(\cdot)$ 表示真实轨迹的分布, $q(\cdot)$ 表示模型预测的轨迹分布;通过最大化不同运动模式下潜在变量之间的差异,偏好模型能够更好地区分不同个体的运动习惯。

2.4 组间模块

首先,使用平均池化操作来获得每个组最具有代表性的组特征,其中相应节点特征被聚合为一个节点,以考虑群体行为,从群体层面学习特征信息。每个组的特征将输入场景中的组成员表示为单个节点,使行人图更简单。将合并后的组特征表示为

$$\mathbf{Z} = \{Z_k | k \in [1, \dots, K]\}, Z_k = \frac{1}{G_k} \sum_{i \in G_k} X_i \quad (6)$$

其中, k 为总的行人组数。

其次,通过使用去池化操作将分组图结构改变到其原始大小。这使得能够利用输出的智能体特征融合信息来预测每个行人的轨迹。行人组的去池化操作可以描述为

$$\bar{\mathbf{X}} = \{\bar{X}_n | n \in [1, \dots, N]\}, \bar{X}_n = Z_k \quad (7)$$

其中, $n \in G_k$, \bar{X}_n 表示从合并后的组特征 Z_k 中获得的智能体的轨迹特征信息。

本文使用GCN对行人之间的交互进行建模,

群体间的交互表示为 $\mathcal{G}_{\text{group}} = (V_{\text{group}}, E_{\text{group}})$ 。其中, 节点表示由池化操作所生成的组特征 $V_{\text{group}} = \{X_n | n \in [1, \dots, N]\}$, 边表示成对的组与组之间的交互 $E_{\text{group}} = \{e_{i,j} | i, j \in [1, \dots, N]\}$ 。

为了避免碰撞, 行人应与其他陌生人之间保持适当的距离。但行人经常转向以避免潜在的碰撞, 并保持其个人空间。社会力模型用于描述行人与周围其他行人之间的交互力^[5]。当行人靠近时, 模型会施加排斥力, 使得行人改变运动方向或速度以避免碰撞。因此, 本文将在组与组之间考虑排斥力, 将组 α 与组 β 之间的排斥力定义为

$$f_{\alpha\beta}(r_{\alpha\beta}) = -\nabla_{r_{\alpha\beta}} u_{\alpha\beta}(\|r_{\alpha\beta}\|) \quad (8)$$

$$u_{\alpha\beta}(\|r_{\alpha\beta}\|) = \Omega'_n k_{\alpha\beta} e^{-\|r_{\alpha\beta}\|/r_{\text{col}}} \quad (9)$$

其中, $r_{\alpha\beta} = X_\alpha^t - X_\beta^t$ 表示组 α 与组 β 之间的距离, 排斥力的方向是沿着两者的连线; Ω'_n 表示邻域; $\nabla_{r_{\alpha\beta}}$ 表示梯度, 指向行人间的相对位置, 用于计算排斥力的方向; $u_{\alpha\beta}$ 表示基于距离的排斥势场, 用于模拟行人之间的排斥效果, 是一个递减函数, 当两组行人间的距离越近时, 排斥力越大; $k_{\alpha\beta}$ 是一个可学习的随时间变化的参数, 用于控制行人之间的排斥力大小; r_{col} 表示超参数, 控制排斥力作用的有效范围。当距离小于某个阈值时, 排斥力开始生效, 促使行人产生规避行为。传统方法中, $u_{\alpha\beta}$ 采用对称椭圆场等简单的函数, 但这种简单函数不能准确描述复杂的行人交互。因此, 本文将 $u_{\alpha\beta}$ 建模为通过神经网络来进行学习的关于 $u_{\alpha\beta}$ 参数化的时变场。将 $k_{\alpha\beta}$ 设置为 $k_{\alpha\beta} = \text{asigmoid}(NN(q'_\alpha, q'_\beta)) + b$, a 和 b 是超参数, 用来确保所学习到的 $k_{\alpha\beta}$ 是有效的。组 α 与组 β 接近到一定程度时, 社会力模型会检测到潜在的碰撞风险。 q'_α 和 q'_β 分别表示 α 和 β 在 t 时刻的状态。如果在 t 时刻有 m 个组, 则组 α 所受到的排斥力为 $f_{\alpha\beta} = \sum_{\beta=0}^m f_{\alpha\beta}$ 。在图 4 的行人碰撞预测网络中, 对邻域 Ω'_n 中每个 q'_α 进行编码, 并与相邻 q'_β 连接。首先, q'_α 经过一个 [4, 64] 的线性层, 将输入状态映射到 64 维空间。接下来, 经过线性层的输出通过维度为 256 的 LSTM 层, 得到 α 当前时刻的隐藏层特征。然后, h_t 经过形状为 [256, 16] 的线性层, 将 256 维的隐藏状态转换为 16 维, 帮助模型学习更

低维度的特征表示。同样, q'_β 经过 [4, 64, 256, 16] 的 MLP 后得到当前时刻的隐藏层状态。最后, 将组 α 与组 β 的隐藏层状态通过结构为 [32, 512, 526, 512, 1] 的多层感知机 (MLP, multilayer perceptron) 后得到相应的结果。

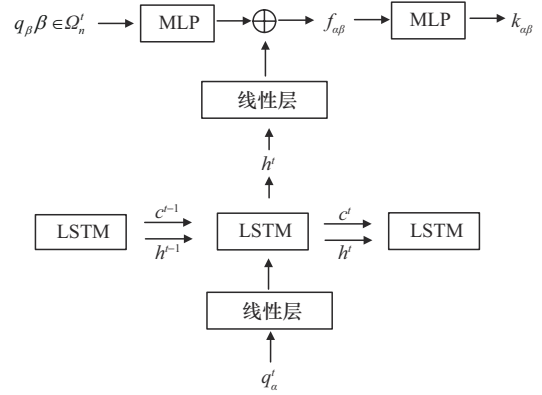


图 4 行人碰撞预测网络

2.5 预测模块

1) 编码过程

编码过程旨在生成近似后验分布的高斯参数, 并对其潜在分布进行建模。在数学上, 后验分布和先验分布的高斯参数由式(10)、式(11)给出, 将所有行人的轨迹经过全连接层。

$$V_x = \mathbb{R}(X_i), V_y = \mathbb{R}(Y_i) \quad (10)$$

$$u = F_e([V_x, V_y]), \sigma = F_e([V_x, V_y]) \quad (11)$$

其中, $\mathbb{R}(X_i)$ 和 $\mathbb{R}(Y_i)$ 是指通过上文所提到的分组模块后所得到的中间变量, F_e 是全连接层, u 和 σ 是后验分布的平均值和方差。

接下来, 由通过全连接层后所得到的 u 和 σ 所组成的高斯分布 $N(u, \sigma^2)$ 得到潜在变量 z 。通过对潜在变量 z 进行表示, 能够输出多种未来可能的路径。

2) 解码过程

解码过程使用一个残差解码器, 该解码器将编码过程生成的潜在变量 z 和过去的轨迹作为输入, 如图 5 所示。残差解码器的目的是预测未来的轨迹, 并对过去的轨迹进行重构, 避免了输入信息的丢失。残差解码器由具有相同结构的 2 个解码块组成, 每个解码块由一个用于编码序列的 LSTM 编码器和 MLP 用作输出结果的组成。

2.6 损失函数

在训练过程中, 采用了一种结合了行人个体的

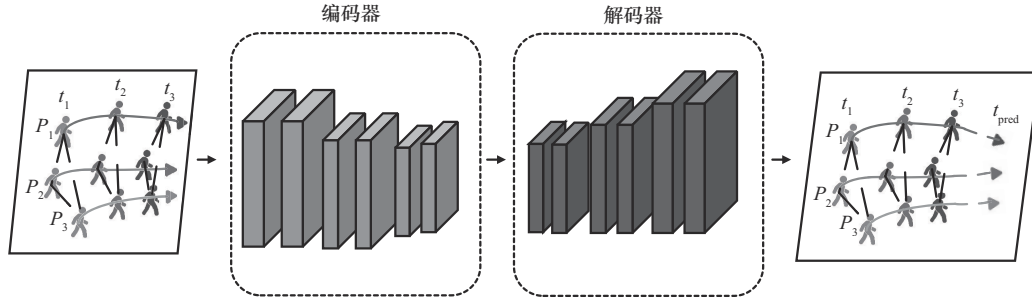


图5 轨迹预测编码器-解码器结构

预测损失和群体预测损失的加权损失函数。该损失函数不仅优化了行人轨迹的准确性，还鼓励模型学习行人之间的相互作用和群体运动行为。具体而言，行人个体的预测损失主要用于度量模型预测轨迹与真实轨迹之间的误差。对于每个行人 i ，计算预测轨迹 Y_i^t 和真实轨迹 X_i^t 之间的欧几里得距离 (L2 范数)，计算式为

$$L_{\text{traj}} = \sum_{i=1}^N \sum_{t=1}^T \|Y_i^t - X_i^t\|_2^2 \quad (12)$$

其中， N 是行人数量， T 是时间步长， $\|\cdot\|_2$ 表示欧几里得距离，即每个时间步的预测位置与真实位置之间的平方差。该损失函数的目的是最小化轨迹预测的误差，使得每个时间步上模型的预测轨迹尽可能接近真实轨迹。

群体预测损失用于建模行人之间的相互关系与群体运动行为。对于群体内的每对行人 c 和 d ，群体预测损失可以表示为

$$L_{\text{group}} = \sum_{c,d \in k} \sum_{t=1}^T \|Y_c^t - X_d^t\|_2^2 \quad (13)$$

其中， k 为总的行人组数，用于控制群体预测损失在总损失中的比重。群体预测损失的目标是通过最小化群体内成员之间的相对运动误差，来促进模型学习行人之间的协调行为。

因此，总的训练损失是行人个体的预测损失和群体预测损失的加权和。总损失 L_{total} 为

$$L_{\text{total}} = L_{\text{traj}} + \eta_{\text{group}} L_{\text{group}} \quad (14)$$

其中， η_{group} 是超参数。如果 η_{group} 较大，模型将更加关注群体行为的建模，可能会牺牲一些个体行为的预测精度。如果 η_{group} 较小，模型会更加关注个体行为的准确性，但可能忽略了群体互动模式，导致预测不符合实际的群体行为。 η_{group} 初始设定为 0.1~10，并观察模型在验证集上的表现进行调整。最终通过交叉验证得出，在行人群体较为密集的场景中，

η_{group} 为 3 时模型的性能最好，为最优的加权系数。通过优化总损失函数 L_{total} ，模型可以同时优化个体行人的轨迹预测和群体行为的建模。

3 实验

本节介绍所使用的数据集、评估指标和实施细节以及定量实验、消融实验、定性实验和复杂度分析的结果。

3.1 数据集

ETH^[21] 和 UCY^[22] 数据集常用于评估预测方法的性能，这 2 个数据集由真实世界的人类轨迹和丰富的人类互动场景组成，已被广泛用作行人轨迹预测的基准。ETH 数据集包含了 ETH 和 Hotel 这 2 个场景，UCY 数据集包含 Zara1、Zara2 和 Univ 这 3 个场景。本文遵循之前的工作^[23-24]，观察了 8 个时间步长 (3.2 s) 的轨迹，并同时预测了所有行人的 12 个时间步长 (4.8 s) 的轨迹。

3.2 评估指标和实施细节

遵循 Social GAN^[25] 中的标准评估方式，基于概率分布生成 20 个样本，然后从中选择最佳样本来测量评估指标。采用与文献[21]相同的评估指标，包括平均位移误差 (ADE, average displacement error) 和最终位移误差 (FDE, final displacement error)。ADE 计算预测轨迹和地面真实轨迹之间的欧几里得距离，而 FDE 则是计算预测的最终目的地和地面真实最终目的地之间的距离。

$$\text{ADE} = \frac{\sum_{i=1}^N \sum_{t=t_{\text{obs}}+1}^{t_{\text{pred}}} \|X_i^t - Y_i^t\|_2}{N t_{\text{pred}}} \quad (15)$$

$$\text{FDE} = \frac{\sum_{i=1}^N \|X_i^{t_{\text{pred}}} - Y_i^{t_{\text{pred}}}\|_2}{N} \quad (16)$$

其中, N 表示行人数量, t_{pred} 表示预测的时间步长。

在 Pycharm 平台上使用 Adam 优化器, 在 ETH 和 UCY 数据集上训练 300 个回合, 批次大小为 128, 初始学习率为 0.001, 所有 LSTM 隐藏层状态维度设置为 128。对于所有数据集, 随机噪声是从标准高斯分布中采样的 16 维向量。模型在 PyTorch 框架上实现, 在配备 RTX 3090 GPU(24 GB) 和 Intel(R) Xeon(R) Gold 5218 CPU(256 GB)的工作站上进行。

训练过程中的损失函数曲线如图 6 所示, 描述了模型的损失函数随迭代次数的变化情况。可以看出, 训练的损失值在迭代次数为 0~10 时迅速下降, 并在之后继续缓慢下降, 表明模型在早期阶段学习到了有效的特征。在训练过程中, 使用优化器根据损失曲线的变化动态调整学习率。随着训练的继续, 损失曲线逐渐趋于平稳, 损失值趋于 0.01。

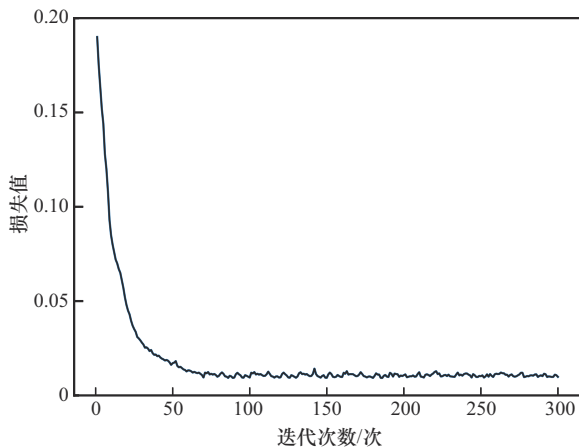


图6 训练过程中的损失函数曲线

3.3 实验结果与分析

3.3.1 定量分析

将 GPCNet 与线性模型 (Linear)、基于社交的长短期记忆网络 (Social LSTM)^[11]、基于社交的生成对抗网络 (Social GAN)^[25]、基于概率人群的生成对抗网络 (PCGAN)^[26]、基于时空图的轨迹预测框架 (STAR)^[27]、预测终点的条件网络 (PECNet)^[28]、分层时空注意力模型 (HSTA)^[29]、基于行为识别的行人轨迹预测模型 (BR-GAN)^[30]、基于多智能体的轨迹预测框架 (VNAGT)^[31] 和基于循环神经网络的条件变分自编码器框架 (Social LODE)^[32] 的基线方法进行了比较。表 1 显示了 ETH、Hotel、Univ、Zara1 和 Zara2 数据集的 ADE_{\min} 和 FDE_{\min} 。其中, 最佳性能用粗体标记。由表 1 可以看出, GPCNet 在大多数数据集上都取得了较好的结果, 性能也有所提高。与常用的基线方法 Social GAN 相比, 性能提高了约 75.4%。由此可以看出, GPCNet 在性能方面有了较大的提升。与同样考虑行人之间交互的 PCGAN 相比, GPCNet 在 5 个数据集上都有显著的改进, 性能提高了约 57.5%。这是因为不同行人会因其习惯而具有不同的运动模式, 而 GPCNet 将行人的偏好考虑在内。因此, 所生成的未来轨迹也更符合地面真实的轨迹。与目前较为先进的方法 Social LODE 相比, 性能提高了约 19.0%。由此可以看出, 使用物理方法与深度学习相结合的方式对行人的轨迹进行预测是可行的, 这样不仅增加了可解释性还提高了预测的精度。当比较单个数据集时, 在

表1 基线模型和 GPCNet 在 ETH 和 UCY 数据集上的 ADE_{\min} / FDE_{\min} 定量结果

方法	ETH	Hotel	Univ	Zara1	Zara2	平均
Linear	1.33/2.94	0.39/0.72	0.82/1.59	0.62/1.21	0.77/1.48	0.79/1.59
Social LSTM ^[11]	1.09/2.35	0.79/1.76	0.67/1.40	0.47/1.00	0.56/1.17	0.72/1.54
Social GAN ^[25]	0.81/1.52	0.72/1.61	0.60/1.26	0.34/0.69	0.42/0.84	0.58/1.18
PCGAN ^[26]	0.43/0.65	0.59/0.64	0.49/0.57	0.25/0.40	0.22/0.34	0.40 /0.52
STAR ^[27]	0.36/0.65	0.17/0.36	0.31/0.62	0.26/0.55	0.22/0.46	0.26/0.53
PECNet ^[28]	0.54/0.87	0.18/0.24	0.35/0.60	0.22/0.39	0.17/0.30	0.29/0.48
HSTA ^[29]	0.38/0.62	0.40/0.79	0.55/1.17	0.34/0.71	0.32/0.68	0.40/0.79
BR-GAN ^[30]	0.73/1.37	0.55/1.13	0.53/1.07	0.35/0.71	0.35/0.72	0.50/1.00
VNAGT ^[31]	0.52/0.88	0.16/0.25	0.27/0.51	0.23/0.44	0.18/0.33	0.27/0.48
Social LODE ^[32]	0.42/0.63	0.14/0.19	0.21/ 0.36	0.18/0.31	0.13/ 0.24	0.21/0.34
GPCNet	0.29/0.37	0.14/0.23	0.19/0.36	0.14/0.23	0.11/0.25	0.17/0.29

人群较为密集的场景，如 Univ 和 Zara1，GPCNet 相比最先进的方法 Social LODE 性能分别提高了约 9.5% 和 25.8%。这些结果表明，GPCNet 在预测高密度场景方面会具有更好的性能。

3.3.2 消融实验

对所提方法进行了广泛的消融研究，以此来研究关键技术组件在 GPCNet 中的贡献。

1) 模型 1 不包括行人偏好模块和行人碰撞预测网络模型，用于研究偏好模块和碰撞预测网络对行人轨迹预测精度的影响。

2) 模型 2 不包括行人偏好模块，用于研究在考虑行人碰撞预测网络模型时，行人偏好对组内交互的影响。

3) 模型 3 不包括碰撞预测网络模型，用于研究在考虑行人偏好模块时，碰撞预测网络模型对组间交互的影响。

4) 模型 4 是本文所提出的 GPCNet，同时考虑了行人偏好模块和行人碰撞预测网络模型。

表 2 和表 3 分别是通过消融实验所得到的 ADE 和 FDE 结果，结果表明，每一个组件都有助于提高轨迹预测的精度。模型 1 用于研究偏好模块和碰撞预测网络对行人轨迹预测精度的综合影响。实验结果显示，去除这 2 个模块后，模型的预测精度显著下降，性能降低了约 29% 和 22%。这表明偏好模块和碰撞预测网络对整体预测有至关重要的作用。模型 2 用于研究在包含碰撞预测网络的情况下，行人偏好对组内交互的影响。实验结果显示，去除该模块后模型的预测准确性有所下降，性能降低了约 23% 和 9%。偏好模块能够帮助模型捕捉个体间的行为差异，不同行人的偏好可能导致不同的轨迹选择。没有该模块的情况下，模型无法有效区分这些个体行为差异，从而降低了组内交互的精度。该结果表明，偏好模块对个体行为建模起到了关键作用，尤其在组内交互的细节捕捉中发挥了不可忽视的贡献。模型 3 用于研究在包含行人偏好模块的情况下，碰撞预测网络对组间交互的影响。通过移除碰撞预测网络，模型的预测精度有所下降，性能降低了约 19% 和 9%。碰撞预测网络的主要功能是在密集场景中捕捉行人之间的相互避让行为。去除该模块后，模型在组间交互中的表现下降，尤其在行人密集的情况下，无法有效预测行人之间的相互避让，导致多个行人轨迹出现潜在碰撞。实验表明，

碰撞预测网络在帮助模型理解行人组间的复杂交互、避免轨迹冲突方面起到了重要作用。模型 4 是本文提出的 GPCNet，包括行人偏好模块和碰撞预测网络。实验结果表明，这 2 个模块的结合可以有效捕捉行人个体差异和复杂交互行为，显著提高了轨迹预测的精度。

表 2 消融实验的 ADE 结果

模型	ETH	Hotel	Univ	Zara1	Zara2	平均
1	0.37	0.22	0.25	0.21	0.17	0.24
2	0.34	0.19	0.23	0.17	0.15	0.22
3	0.33	0.20	0.21	0.17	0.14	0.21
4	0.29	0.14	0.19	0.14	0.11	0.17

表 3 消融实验的 FDE 结果

模型	ETH	Hotel	Univ	Zara1	Zara2	平均
1	0.48	0.31	0.43	0.30	0.32	0.37
2	0.40	0.25	0.39	0.28	0.27	0.32
3	0.42	0.25	0.38	0.27	0.27	0.32
4	0.37	0.23	0.36	0.23	0.25	0.29

图 7 和图 8 分别描述了消融实验的可视化结果。由图 7 和图 8 可以更清楚地看出，模型 4 的 GPCNet 的 ADE 和 FDE 是最小的，轨迹预测的性能也是最好的。尤其是在人群较为密集的场景，如 Hotel 和 Zara 数据集下，GPCNet 的性能是最好的。结果表明，与同时考虑行人偏好和碰撞预测网络的方法相比，所有单独的行人偏好部分和碰撞预测网络模型的组合都会导致性能降低。

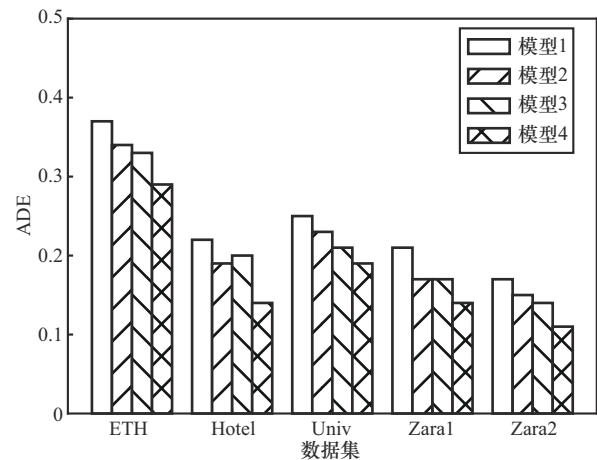


图 7 消融实验 ADE 结果对比

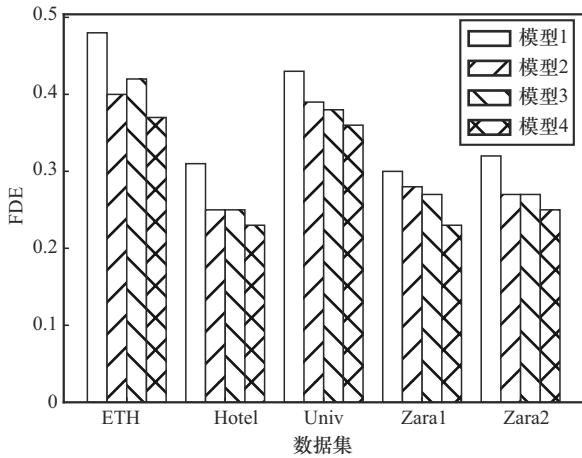


图8 消融实验FDE结果对比

3.3.3 定性分析

图9展示了行人较为密集的Zara、Univ和Hotel的一些场景，其中浅灰色表示观察到的行人的历史轨迹，黑色表示行人的真实轨迹，长虚线表示本文所提出的GPCNet所预测的行人的未来轨迹，

短虚线表示最常用的基准模型 Social GAN 所预测的行人的未来轨迹。

从图9可以看出，与基准模型相比，GPCNet所预测的轨迹更加接近地面的真实轨迹。对于结伴而行的行人来说，从图9(a3)中可以看到，当两组行人向相反的方向移动时，他们会采取行动避开对方。GPCNet不仅没有破坏他们群体之间的持续运动，还保留了组的结构和相干运动，同时避免了碰撞。基准模型所预测的轨迹虽然保持相同的行进方向，但轨迹误差较大。对于行人同向行走而言，如图9(b2)所示，有的人站在原地互相交谈或倾听。GPCNet所预测的轨迹绕过了站在原地的行人，从而避免发生碰撞，生成更加符合社会规范的轨迹。而基准模型的轨迹会与行人发生碰撞，存在安全隐患。对于单独行走的行人而言，如图9(b3)所示，GPCNet所预测出的轨迹与地面真实的轨迹较为接近。而基准模型所预测出的轨迹会朝着建筑物方向行走，最终会撞向障碍物。

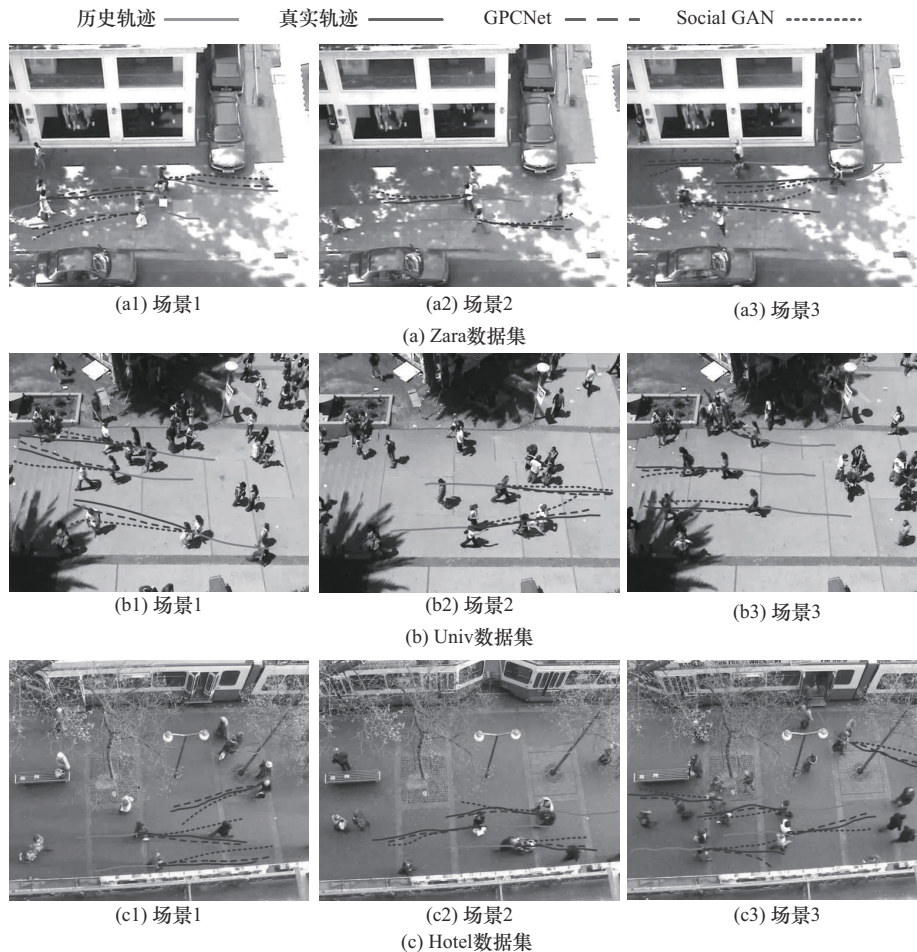


图9 Zara、Univ和Hotel数据集的可视化结果

这是因为行人会因为自身的习惯而具有一些特定的运动模式，而 GPCNet 将行人的偏好考虑在内，从而使预测的轨迹更符合真实情况。对于反向行走而言，如图 9(c2) 所示，几个行人成群结队地行走，他们的运动模式趋于一致变化，并且群体结构趋于稳定，GPCNet 所预测的轨迹接近地面实况的轨迹。然而，如图 9(c3) 所示，在某些场景下，模型预测出的轨迹会出现一定的偏差，个别行人的运动轨迹与真实轨迹存在差异，不符合实际社会规范。因此，当面对类似的场景时，模型倾向于依赖大多数行人遵循的常规轨迹预测，进而忽略了少数异常行为，从而产生偏差。为了应对数据中的类别不平衡问题，可以通过数据增强技术来增加模型对罕见行为的学习^[33]。例如，利用现有的正常行人轨迹数据，通过对行人行为进行合成和随机化（如随机生成反向行走、突然停顿等行为）来增强数据集的多样性，帮助模型更好地捕捉不同类型的行人行为。

3.3.4 复杂度分析

在自动驾驶领域，实时性对于算法在行业中的应用至关重要。自动驾驶汽车需要具备快速感知周围环境并进行决策的能力，以保证驾驶的安全和效率。在这一过程中，轨迹预测模型的推理速度直接影响系统的反应时间和决策准确性。为了更好地衡量 GPCNet 在自动驾驶场景中的可行性，设计了推理速度对比实验，比较了 GPCNet 与现有轨迹预测框架 Social GAN、PECNet、BR-GAN 和 GPCNet 在 GPU 和 CPU 环境下的表现，并取超过 10 次运行结果的平均值作为最终结果，单位为秒/步，如表 4 所示。通过实验，不仅衡量了推理速度，还进一步分析了模型的复杂度对计算性能的影响。

表 4 不同方法的推理速度对比

方法	CPU	CPU
Social GAN	2.027	3.581
PECNet	1.192	1.873
BR-GAN	0.556	1.048
GPCNet	0.073	0.125

实验结果表明，尽管 GPCNet 采用了较为复杂的交互建模机制，其在推理速度方面依然保持了较好的表现。这一优势使得 GPCNet 在自动驾驶、智

能交通等领域具备了广阔的应用前景，特别是在对实时性要求较高的应用场景中，能够有效提升车辆的感知和决策能力。虽然 GPCNet 已经在推理速度和精度之间取得了较好的平衡，但仍存在一定的优化空间。未来的研究可以进一步提高模型的计算效率，降低复杂度，提升其应用性和扩展性。随着多智能体系统和更高维度场景的不断发展，如何在保证高精度预测的同时，进一步降低模型的复杂度将是未来工作的一个重要方向。

4 结束语

本文提出了一种针对群组感知的行人轨迹预测网络，通过对个体层面和群体层面进行研究来学习行人的随机运动行为。具体地，首先使用分组模块得到行人所属的组信息。在组内，将行人的偏好问题考虑在内，考虑因个人习惯会形成特定的运动模式这个因素。在组间，通过使用物理方法与深度学习方法相结合的方式考虑行人轨迹的碰撞问题，使用碰撞预测网络减小行人之间碰撞发生的概率。实验结果表明，GPCNet 在广泛使用的 ETH 和 UCY 数据集上实现了比基准模型更好的性能，并且在高密度场景中的性能更为显著。在未来工作中，将考虑轨迹预测中所存在的长尾分布问题。用于解决数据尾部情况复杂，运动模式多样化，预测误差大的问题，进而提高轨迹预测的性能。

参考文献:

- [1] RUDENKO A, PALMIERI L, HERMAN M, et al. Human motion trajectory prediction: a survey[J]. The International Journal of Robotics Research, 2020, 39(8): 895-935.
- [2] CHEN X B, ZHANG H J, ZHAO F, et al. Vehicle trajectory prediction based on intention-aware non-autoregressive transformer with multi-attention learning for Internet of vehicles[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 2513912.
- [3] WANG A, MAVROGIANNIS C, STEINFELD A. Group-based motion prediction for navigation in crowded environments[C]//Proceedings of the 5th Conference on Robot Learning, New York: PMLR, 2022: 871-882.
- [4] RUDENKO A, PALMIERI L, LILIENTHAL A J, et al. Human motion prediction under social grouping constraints[C]//Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Piscataway: IEEE Press, 2018: 3358-3364.
- [5] HELBING D, MOLNÁR P. Social force model for pedestrian dynamics[J]. Physical Review E, 1995, 51(5): 4282-4286.
- [6] LUBER M, STORK J A, TIPALDI G D, et al. People tracking with human motion predictions from social forces[C]//Proceedings of the 2010

- IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2010: 464-469.
- [7] BROUWER N, KLOEDEN H, STILLER C. Comparison and evaluation of pedestrian motion models for vehicle safety systems[C]//Proceedings of the 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). Piscataway: IEEE Press, 2016: 2207-2212.
- [8] ZERNETSCH S, KOHNEN S, GOLDHAMMER M, et al. Trajectory prediction of cyclists using a physical model and an artificial neural network[C]//Proceedings of the 2016 IEEE Intelligent Vehicles Symposium (IV). Piscataway: IEEE Press, 2016: 833-838.
- [9] GU Y L, HASHIMOTO Y, HSU L T, et al. Motion planning based on learning models of pedestrian and driver behaviors[C]//Proceedings of the 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). Piscataway: IEEE Press, 2016: 808-813.
- [10] XIE G T, GAO H B, QIAN L J, et al. Vehicle trajectory prediction by integrating physics- and maneuver-based approaches using interactive multiple models[J]. IEEE Transactions on Industrial Electronics, 2018, 65(7): 5999-6008.
- [11] ALAHI A, GOEL K, RAMANATHAN V, et al. Social LSTM: human trajectory prediction in crowded spaces[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2016: 961-971.
- [12] PELLEGRINI S, ESS A, VAN GOOL L. Improving data association by joint modeling of pedestrian trajectories and groupings[C]//Proceedings of the 11th European Conference on Computer Vision: Part I. Berlin: Springer, 2010: 452-465.
- [13] BISAGNO N, ZHANG B, CONCI N. Group LSTM: group trajectory prediction in crowded scenarios[C]//Proceedings of the 17th European Conference On Computer Vision. Berlin: Springer, 2019: 213-225.
- [14] ZHANG W C, CHENG H, JOHORA F T, et al. ForceFormer: exploring social force and transformer for pedestrian trajectory prediction[C]//Proceedings of the 2023 IEEE Intelligent Vehicles Symposium (IV). Piscataway: IEEE Press, 2023: 1-7.
- [15] SUN J H, JIANG Q H, LU C W. Recursive social behavior graph for trajectory prediction[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2020: 657-666.
- [16] ROBICQUET A, SADEGHIAN A, ALAHI A, et al. Learning social etiquette: human trajectory understanding in crowded scenes[C]//Proceedings of the 14th European Conference on Computer Vision. Berlin: Springer, 2016: 549-565.
- [17] LI L H, PAGNUCCO M, SONG Y. Graph-based spatial transformer with memory replay for multi-future pedestrian trajectory prediction[C]//Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2022: 2221-2231.
- [18] LEE N, CHOI W, VERNAZA P, et al. DESIRE: distant future prediction in dynamic scenes with interacting agents[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2017: 2165-2174.
- [19] IVANOVIC B, PAVONE M. The trajectron: probabilistic multi-agent trajectory modeling with dynamic spatiotemporal graphs[C]//Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 2019: 2375-2384.
- [20] TANG Y C, SALAKHUTDINOV R R. Multiple futures prediction[C]//Proceedings of the 33rd International Conference on Neural Information Processing Systems. New York: ACM Press, 2019, 1-11.
- [21] PELLEGRINI S, ESS A, SCHINDLER K, et al. You'll never walk alone: modeling social behavior for multi-target tracking[C]//Proceedings of the 2009 IEEE 12th International Conference on Computer Vision. Piscataway: IEEE Press, 2009: 261-268.
- [22] LERNER A, CHRYSANTHOU Y, LISCHINSKI D. Crowds by example[J]. Computer Graphics Forum, 2007, 26(3): 655-664.
- [23] MANGALAM K, AN Y, GIRASE H, et al. From goals, waypoints & paths to long term human trajectory forecasting[C]//Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 2021: 15213-15222.
- [24] KOSARAJU V, SADEGHIAN A, MARTÍN-MARTÍN R, et al. Social-BiGAT multimodal trajectory forecasting using bicycle-gan and graph attention networks[J]. arXiv Preprint, arXiv: 1907.03395, 2019.
- [25] GUPTA A, JOHNSON J, LI F F, et al. Social GAN: socially acceptable trajectories with generative adversarial networks[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 2255-2264.
- [26] EIFFERT S, LI K M, SHAN M, et al. Probabilistic crowd GAN: multimodal pedestrian trajectory prediction using a graph vehicle-pedestrian attention network[J]. IEEE Robotics and Automation Letters, 2020, 5(4): 5026-5033.
- [27] YU C J, MA X, REN J W, et al. Spatio-temporal graph transformer networks for pedestrian trajectory prediction[C]//Proceedings of the 16th European Conference Computer Vision. Berlin: Springer, 2020: 507-523.
- [28] MANGALAM K, GIRASE H, AGARWAL S, et al. It is not the journey but the destination: endpoint conditioned trajectory prediction[C]//Proceedings of the 16th European Conference Computer Vision. Berlin: Springer, 2020: 759-776.
- [29] WU Y, CHEN G, LI Z J, et al. HSTA: a hierarchical spatio-temporal attention model for trajectory prediction[J]. IEEE Transactions on Vehicular Technology, 2021, 70(11): 11295-11307.
- [30] PANG S M, CAO J X, JIAN M Y, et al. BR-GAN: a pedestrian trajectory prediction model combined with behavior recognition[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(12): 24609-24620.
- [31] CHEN X B, ZHANG H J, HU Y, et al. VNAGT: variational non-autoregressive graph transformer network for multi-agent trajectory prediction[J]. IEEE Transactions on Vehicular Technology, 2023, 72(10): 12540-12552.
- [32] KE K X, YANG J, LIU Y J, et al. Social lode: human trajectory prediction with latent odes[C]//Proceedings of the 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway: IEEE Press, 2024: 5360-5364.
- [33] SUN L T, ROELOFS R, CAINE B, et al. CausalAgents: a robustness benchmark for motion forecasting[C]//Proceedings of the 2024 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE Press, 2024: 6820-6827.

[作者简介]



王汝言 (1969-), 男, 湖北浠水人, 博士, 重庆邮电大学教授、博士生导师, 主要研究方向为泛在网络、多媒体信息处理等。



段昂 (1991-), 男, 重庆人, 重庆邮电大学博士生, 主要研究方向为车联网、自动驾驶、边缘计算等。



周玉蝶 (2000-), 女, 四川宜宾人, 重庆邮电大学硕士生, 主要研究方向为自动驾驶、车联网等。



崔亚平 (1986-), 男, 河南新乡人, 博士, 重庆邮电大学副教授, 主要研究方向为车联网、目标检测和行为识别等。



吴大鹏 (1979-), 男, 黑龙江大庆人, 博士, 重庆邮电大学教授、博士生导师, 主要研究方向为泛在网络、人工智能等。



何鹏 (1990-), 男, 重庆人, 博士, 重庆邮电大学讲师, 主要研究方向为计算机视觉、数据挖掘和分析等。