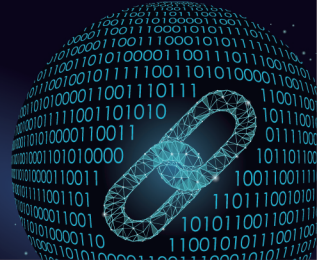




中国科技期刊卓越行动计划项目入选期刊

# 控制与决策

CONTROL AND DECISION



## 基于动态进化图的行人轨迹预测

牟菁, 张旭秀, 闫涵

引用本文:

牟菁, 张旭秀, 闫涵. 基于动态进化图的行人轨迹预测[J]. 控制与决策, 2024, 39(7): 2345–2353.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2022.1907>

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 基于深度学习的行人轨迹预测方法综述

Survey of pedestrian trajectory prediction methods based on deep learning

控制与决策. 2021, 36(12): 2841–2850 <https://doi.org/10.13195/j.kzyjc.2020.1841>

#### 基于图卷积网络的行为识别方法综述

A survey of action recognition methods based on graph convolutional network

控制与决策. 2021, 36(7): 1537–1546 <https://doi.org/10.13195/j.kzyjc.2020.0514>

#### Anchor-free的尺度自适应行人检测算法

Anchor-free scale adaptive pedestrian detection algorithm

控制与决策. 2021, 36(2): 295–302 <https://doi.org/10.13195/j.kzyjc.2020.0124>

#### 基于多尺度特征表示的行人再识别

Multi-scale feature representation for person re-identification

控制与决策. 2021, 36(12): 3015–3022 <https://doi.org/10.13195/j.kzyjc.2020.0952>

#### 考虑退化轨迹差异性与相似性的轴承RUL预测

Prediction of bearing remaining useful life involving difference and similarity of degradation trajectories

控制与决策. 2021, 36(11): 2832–2840 <https://doi.org/10.13195/j.kzyjc.2020.1028>

# 基于动态进化图的行人轨迹预测

聿菁<sup>1</sup>, 张旭秀<sup>1†</sup>, 闫涵<sup>2</sup>

(1. 大连交通大学 自动化与电气工程学院, 辽宁 大连 116021;

2. 大连理工大学 控制科学与工程学院, 辽宁 大连 116024)

**摘要:** 行人轨迹预测在自动驾驶和社交机器人等领域有着广泛的应用. 对行人间复杂的交互关系进行有效建模是提高轨迹预测准确性的关键问题. 然而, 基于图神经网络的方法建模行人间的复杂交互时, 存在行人间交互关系不会随着时间推移而改变, 并且图模型无法自适应地调整网络参数, 导致预测轨迹与真实轨迹偏差较大. 为此, 提出基于动态进化图的行人轨迹预测方法, 设计动态特征更新 (DFU) 以定义行人间的动态特性, 对行人间动态交互进行建模以构建时间域的网络动态性, 提升对行人间复杂交互关系建模的能力. 采用进化图卷积单元优化编码器, 灵活进化图模型网络参数, 增强图模型的自适应能力. 研究表明, 在预测 8 个时间步长下, 与 STGAT 模型相比, 所提出模型在两个公开数据集 (ETH 和 UCY) 上取得了更好的性能, 平均位移误差降低 12.26%, 最终位移误差降低 14.10%.

**关键词:** 动态图; 进化图; 图卷积; 轨迹预测; 行人轨迹

中图分类号: TP391

文献标志码: A

DOI: 10.13195/j.kzyjc.2022.1907

引用格式: 聿菁, 张旭秀, 闫涵. 基于动态进化图的行人轨迹预测[J]. 控制与决策, 2024, 39(7): 2345-2353.

## Pedestrian trajectory prediction based on dynamic evolving graph

MI Jing<sup>1</sup>, ZHANG Xu-xiu<sup>1†</sup>, YAN Han<sup>2</sup>

(1. College of Automation & Electrical Engineering, Dalian Jiaotong University, Dalian 116021, China; 2. College of Control Science and Engineering, Dalian University of Technology, Dalian 116024, China)

**Abstract:** Pedestrian trajectory prediction has a wide range of applications in areas such as autonomous driving and social robotics. Effective modeling of complex interactions between pedestrians is a key issue to improve trajectory prediction accuracy. However, when modeling the complex interactions between pedestrians based on graph neural networks, there are problems that the interactions between pedestrians do not change over time and the graph model cannot adjust the network parameters adaptively, resulting in large deviations between predicted and real trajectories. Therefore, this paper proposes a pedestrian trajectory prediction method based on dynamic evolving graphs, and designs dynamic feature update (DFU) to define the dynamic characteristics between pedestrians and model the dynamic interactions between pedestrians to build the network dynamics in the time domain, which improves the ability to model the complex interactions between pedestrians. An evolving graph convolution unit optimization encoder is used to flexibly evolve the graph model network parameters and enhance the adaptive capability of the graph model. The results show that the proposed model achieves better performance on two publicly available datasets (ETH and UCY) with 12.26% reduction in average displacement error and 14.10% reduction in final displacement error compared with the STGAT model at predicted 8 time steps.

**Keywords:** dynamic graph; evolving graph; graph convolution; trajectory prediction; pedestrian trajectory

## 0 引言

行人轨迹预测被广泛地应用在各种领域, 例如, 自动驾驶<sup>[1]</sup>、移动机器人和导航等. 早期的轨迹预测方法有 高斯过程回归法、运动学、动态法<sup>[2]</sup> 和贝叶斯网络法<sup>[3]</sup> 等. 这些方法基于先验概率、回归残差和核

函数的数学推导, 未考虑行人间交互. 在实际工程应用中, 行人间会相互影响, 从而产生复杂的行为模式, 这导致早期轨迹预测方法的预测精度不理想, 影响了早期轨迹预测方法在实际工程中的应用. 因此, 对行人间潜在的交互关系进行建模成为了一项具有挑战

收稿日期: 2022-11-04; 录用日期: 2023-04-15.

基金项目: 国家自然科学基金项目 (62103074); 辽宁省重点研发计划项目 (2022020594-JH1/108).

†通讯作者. E-mail: zhangxuxiu@163.com.

\*本文附带电子附录文件, 可登录本刊官网该文“资源附件”区自行下载阅览.

性的任务。

近年来,许多研究人员为捕获行人间的交互关系,提出了许多交互建模的方法,如池化机制<sup>[4-5]</sup>,图神经网络<sup>[6-8]</sup>(graph neural network, GNN)等,这些方法可概括为通过信息聚合来实现隐式交互建模.文献[9]提出采用长短期记忆网络(long short-term memory, LSTM)对行人轨迹进行建模,所有的行人轨迹相互独立. Social LSTM<sup>[9]</sup>在LSTM的基础上增加了一个社交池机制,对行人间的交互关系进行建模,并在每个时间步长上,将邻居的隐藏状态合并. Social GAN<sup>[10]</sup>则对 Social LSTM进行了改进,提出了一个新的池化模块来对行人间的交互关系进行建模,并且引入了多样性损失函数鼓励产生多种社会可接受的轨迹. STGAT<sup>[11]</sup>利用图注意力网络<sup>[12]</sup>(graph attention networks, GAT)来聚合LSTM的隐藏状态,从而对行人间的空间交互作用进行建模,并采用了一个额外的LSTM来模拟交互作用的时间相关性.文献[13]提出了一种基于时空图Transformer模型框架,该框架通过注意力机制来处理轨迹预测;文献[14-15]基于GAT和GCN通过构造全局特征更新和全局特征选择以提升空间域和时间域的网络感受范围.另一类是显式的对潜在交互图进行关系推理,例如NRI<sup>[16]</sup>和 EvolveGraph<sup>[17]</sup>,其中NRI中的交互图在训练过程中是静态的; EvolveGraph则通过一个潜在交互图在多智能体间进行动态的显式交互建模.

上述研究取得了令人满意的结果,然而在实际应

用中,图模型中的节点特征是动态变化的.例如,社交网络的用户会随着时间的推移而发展友谊,因此用户的向量表示应该相应地更新,以反映用户的社会关系的时间演变.行人间交互关系会随时间而发生改变,因此行人交互关系可以视为是动态的.以往的研究在时间域上构造图模型时,不同行人的交互关系视为不变的,未考虑行人间交互关系的动态性,使得网络模型不能有效模拟行人间的现实互动.行人的频繁出现和消失造成图模型无法自适应地调节网络参数,导致图模型的自适应能力不强,也无法自适应地进化改变网络模型参数.为此,本文提出了基于动态进化图的行人轨迹预测方法(dynamic-evolving graph convolutional network, DEGCN).

针对以上问题,本文的研究主要如下:1)设计动态特征更新DFU,提取行人的动态特征,赋予图模型的动态学习能力,强化网络对行人间交互关系的建模能力,提高网络学习并模拟现实场景的能力.2)采用进化图卷积单元(evolving graph convolution unit, EGCU)优化网络模型,进化并减少网络参数,使特征提取不依赖于节点嵌入,有效降低行人频繁出现和消失对图模型带来的影响,提高模型的学习效率和自适应能力.3)建模行人交互关系为动态图,定义行人间动态交互关系,减少不必要的交互链接,提高网络的学习能力,增强网络的动态性.

### 1 算法描述

基于动态进化图的DEGCN模型如图1所示.其

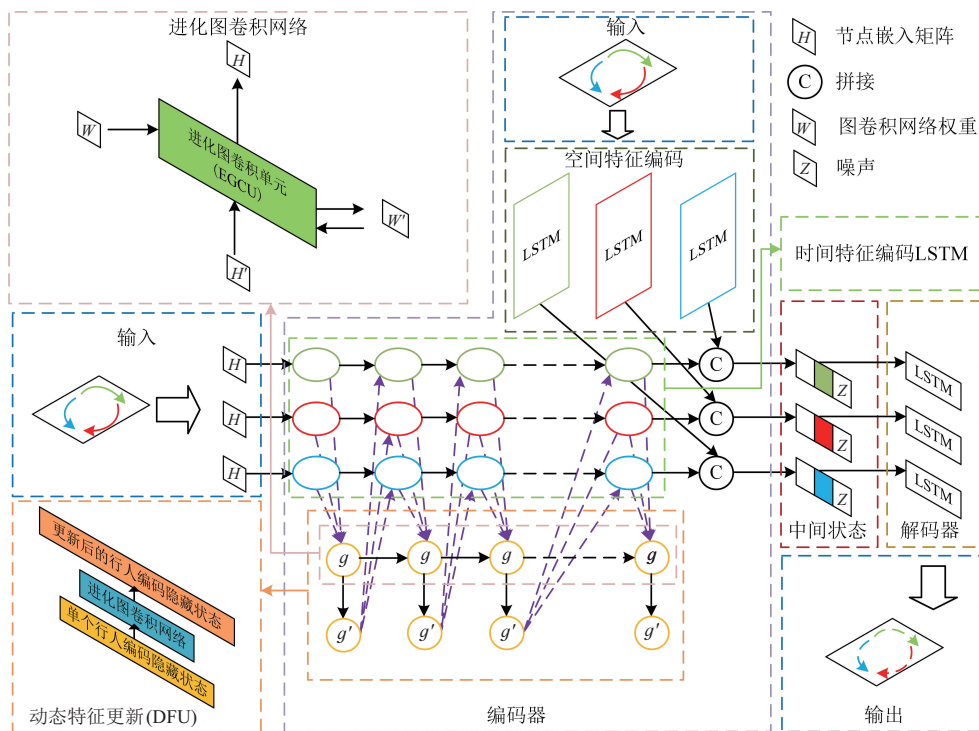


图1 基于动态进化图的DEGCN模型

中,DEGCN模型由编码器、中间状态、解码器3个部分组成.编码器分为空间特征编码和时间特征编码.空间特征编码的输入为行人的位置信息,利用LSTM<sub>1</sub>聚合行人的空间特征,输出行人空间特征隐藏状态.时间特征编码由DFU和EGCU组成,捕捉行人间动态的交互信息,利用LSTM<sub>2</sub>聚合行人的时间特征,输出行人的时间特征隐藏状态.中间状态包含所有轨迹的时空特征信息.解码器根据中间状态和时空特征矩阵来预测未来轨迹.其中,编码器中的DFU、EGCU优化LSTM和采用动态图建模行人交互是本文主要创新点.

### 1.1 问题定义

假设一个场景中有 $N$ 位行人,设为 $p_1, p_2, \dots, p_n$ ,在时间步长 $t$ 时,假设行人 $p_i$ 的二维坐标为 $S_i^t = (x_i^t, y_i^t)$ .然后给出行人 $i = 1, 2, \dots, N$ 在时间步长 $t = 1, 2, \dots, T_{\text{obs}}$ 时的位置 $S_i^t$ ,目标是预测时间步长 $t = T_{\text{obs}+1}, \dots, T_{\text{pred}}$ 的未来位置 $S_i^t$ .

### 1.2 一个行人的轨迹编码

首先,计算每个行人与前一个时间步长的相对位置,即

$$\Delta x_i^t = x_i^t - x_i^{t-1}, \quad (1)$$

$$\Delta y_i^t = y_i^t - y_i^{t-1}. \quad (2)$$

将每个时间步长的相对位置嵌入到一个固定长度的向量 $e_i^t$ 中,并将这些向量作为LSTM<sub>1</sub>的输入,即

$$e_i^t = \phi(\Delta x_i^t, \Delta y_i^t, W_{ee}), \quad (3)$$

$$m_i^t = \text{LSTM}_1(m_i^{t-1}, e_i^t, W_m). \quad (4)$$

其中: $\phi(\cdot)$ 是非线性激活函数ReLU嵌入层(embedding layer)网络; $W_{ee}$ 是嵌入的权重; $W_m$ 是LSTM<sub>1</sub>的权重,场景中的行人共享权重; $m_i^t$ 是LSTM<sub>1</sub>在时间步长 $t$ 时的隐藏状态.

### 1.3 基于动态进化图的行人交互建模

根据文献[9],仅仅使用一个LSTM并不能捕获行人间的交互.为了在复杂的场景中,共享不同行人间的交互信息,并且捕获行人间的动态交互关系,本节首先采用动态图对行人间交互关系进行建模,然后设计DFU捕获图序列背后的动态交互信息,最后利用EGCU优化LSTM,进化网络参数和节点嵌入矩阵,从而灵活调整网络参数,提高图模型的自适应能力.行人间动态交互关系如图2所示.其中,一个场景中行人在每个时间步长上都视为图上的节点,每位行人对应一个特征矩阵 $H_i^t$ .

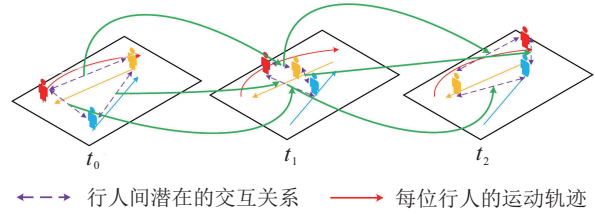


图2 行人间动态交互关系

#### 1.3.1 图卷积网络(GCN)

一个GCN由多层图卷积组成,类似于一个感知器,GCN还有一个由光谱卷积驱动的邻域聚合步骤.本文选用的GCN层数为2,在 $t$ 时刻,第1层以邻接矩阵 $A_i^t$ 和节点嵌入矩阵 $H_i^t$ 作为输入,并使用权值矩阵 $W^t$ 将节点嵌入矩阵 $H_i^t$ 更新至第2层节点嵌入矩阵 $\hat{H}_i^t$ ,即

$$\hat{H}_i^t = \text{GCONV}(A_i^t, H_i^t, W^t) = \sigma(\hat{A}_i^t, H_i^t, W^t). \quad (5)$$

其中:GCONV是图卷积单元. $\hat{A}_i^t$ 是 $A_i^t$ 的标准化,定义为(为清晰起见,省略时间指数)

$$\hat{A} = \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}}, \tilde{A} = A + 1, \tilde{D} = \text{diag}\left(\sum_j \tilde{A}_{ij}\right); \quad (6)$$

$\sigma$ 是除输出层外的所有层的激活函数.初始嵌入矩阵来自于节点特征矩阵 $F_i^t \in \mathbb{R}^{N \times d}$ , $N$ 为问题定义中一个场景的 $N$ 位行人, $F_i^t$ 的每一行都对应节点的一个 $d$ 维特征向量.对于输出层,函数 $\sigma$ 可以被认为恒等式,在这种情况下, $\hat{H}_i^t$ 包含了从初始特征转换而来的图节点的更高一级的表示;函数 $\sigma$ 也可以是用于节点分类的softmax函数,在这种情况下, $\hat{H}_i^t$ 包含预测概率.

#### 1.3.2 进化图卷积单元(EGCU)

动态图在时间域和更新策略上对静态图进行了扩展.动态图神经网络(dynamic graph neural network, DGNN)一般采用GNN和循环神经网络的组合.GNN处理图<sup>[18-21]</sup>的信息,循环神经网络则处理图的动态信息.动态图是将 $W^t$ 作为动态系统的隐藏状态.本文使用门控循环单元(gate recurrent unit, GRU)更新在时间 $t$ 时输入系统的隐藏状态.节点嵌入矩阵 $H_i^t$ 作为系统的输入信息.动态图表示如下:

$$\underbrace{W^t}_{\text{hidden state}} = \text{GRU}\left(\underbrace{H_i^t}_{\text{input}}, \underbrace{W^{t-1}}_{\text{hidden state}}\right). \quad (7)$$

结合式(5)中图卷积单元GCONV和式(7)中的GRU,便得到了EGCU. EGCU如下:

$$W^t = \text{GRU}(H_i^t, W^{t-1}), \quad (8)$$

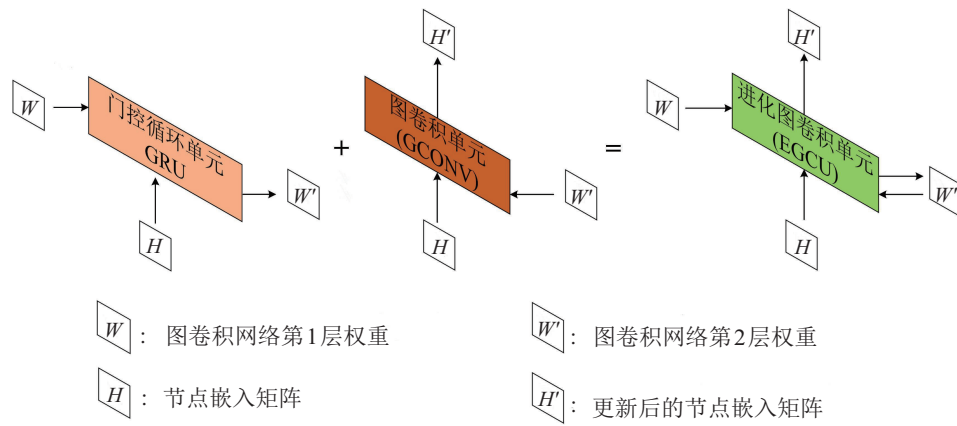


图3 进化图卷积单元(EGCU)

$$\hat{H}_i^t = \text{GCONV}(A_i^t, H_i^t, W^t), \quad (9)$$

$$[\hat{H}_i^t, W^t] = \text{EGCU}(A_i^t, H_i^t, W^{t-1}). \quad (10)$$

循环架构的隐藏状态是GCN权值,EGCU随着时间的推移利用GRU和GCONV进化权重矩阵.将单元自上向下连接,从而得到一个时间步长的多层GCN.图3所示为每一层的计算与 $H_i^t$ 、 $W^t$ 和 $\hat{H}_i^t$ 之间的关系.

1.3.3 动态特征更新(DFU)

在每个时间步长上,DFU结合式(8)中的GRU更新隐藏状态,更新后的隐藏状态作为时间域编码器LSTM<sub>2</sub>的输入.图4是DFU的示意图,其中每个时间步长上包含一个按时间索引的GCN,GCN的参数是时间步长 $t$ 和权重矩阵 $W^{t-1}$ .更新后的权重矩阵 $W^t$

作为LSTM<sub>2</sub>的输入,以更新LSTM<sub>2</sub>的隐藏状态.节点频繁出现和消失,导致LSTM网络学习这种不规则行为的效果不佳,因此本文采用EGCU进化网络参数.网络参数 $W^t$ 和 $\hat{H}_i^t$ 随着时间不断进化,使得图模型不依赖于当前时间步长的节点嵌入矩阵,当节点嵌入矩阵频繁改变时,网络参数可以自行地进行调整,以提高网络的自适应能力,即

$$r_i^t = \text{LSTM}_2(r_i^{t-1}, \hat{H}_i^t, W^t). \quad (11)$$

其中: $\hat{H}_i^t$ 如式(5)所示, $W^t$ 是更新后的权重矩阵.在时间步长 $T_{\text{obs}}$ 中,每个行人都有两个隐藏变量( $m_i^{T_{\text{obs}}}$ ,  $r_i^{T_{\text{obs}}}$ ),把这两个隐藏变量输入到两个不同的多层感知器(MLP)中,将行人的空间特征隐藏状态和时间特征隐藏状态进行结合,即

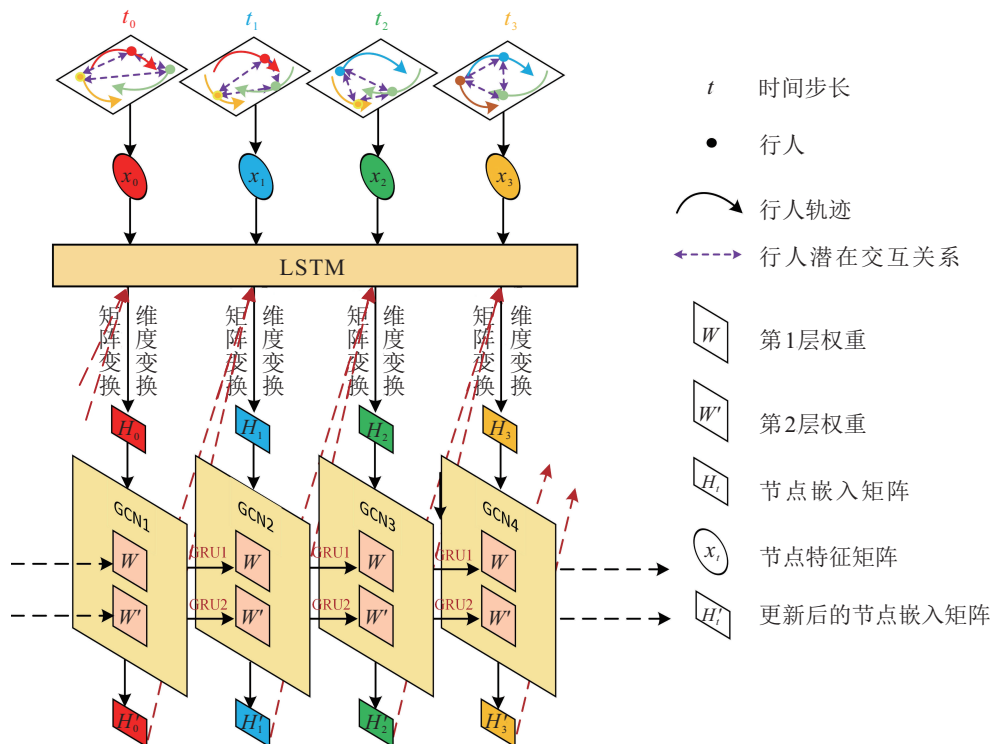


图4 动态特征更新(DFU)的示意图

$$\bar{m}_i = \delta_1(m_i^{T_{\text{obs}}}), \quad (12)$$

$$\bar{r}_i = \delta_2(r_i^{T_{\text{obs}}}), \quad (13)$$

$$h_i = \bar{m}_i \parallel \bar{r}_i. \quad (14)$$

其中:  $\delta_1$  和  $\delta_2$  为两个 MLP,  $\parallel$  是连接操作.

#### 1.4 生成未来的轨迹

为了能够生成多种真实合理的轨迹, 文献[10]提出了一种多样性损失函数来鼓励网络产生不同的样本, 并验证其方法的有效性. 本文引入文献[10]的策略来模拟行人运动的多模态特性.

中间状态的计算公式如下所示:

$$d_i^{T_{\text{obs}}} = h_i \parallel z. \quad (15)$$

其中:  $z$  表示噪声, 中间状态  $d_i^{T_{\text{obs}}}$  作为解码器 LSTM<sub>3</sub> 的初始隐藏状态. 预测相对位置为

$$d_i^{T_{\text{obs}}+1} = \text{LSTM}_3(d_i^{T_{\text{obs}}}, e_i^{T_{\text{obs}}}, W_d), \quad (16)$$

$$(\Delta x_i^{T_{\text{obs}}+1}, \Delta y_i^{T_{\text{obs}}+1}) = \delta_3(d_i^{T_{\text{obs}}+1}). \quad (17)$$

其中:  $W_d$  为 LSTM<sub>3</sub> 的权重;  $\delta_3$  为一个线性层;  $e_i^{T_{\text{obs}}}$  如式(3)所示, 在时间步长  $T_{\text{obs}+1}$  得到预测的相对位置后. 根据最后预测的相对位置计算 LSTM<sub>3</sub> 的后续输入.

对于每个行人, 该模型通过从  $N(0, 1)$  (标准正态分布) 中随机抽样  $z$ , 产生多个预测轨迹; 然后选择  $L_2$  意义下的最佳预测 (距离真实轨迹距离最小的轨迹) 作为模型输出, 计算多样性损失函数, 即

$$L_{\text{variety}} = \min_k \|Y_i - \hat{Y}_i^{(k)}\|_2. \quad (18)$$

其中:  $Y_i$  为行人  $i$  的地面真实轨迹,  $\hat{Y}_i^{(k)}$  为 DEGCN 模型产生的轨迹,  $k$  为超参数. 通过考虑最佳轨迹, 损失函数鼓励网络对冲赌注, 并覆盖符合过去轨迹的输出空间.

## 2 仿真实验

### 2.1 测试问题与性能指标

为验证所提出的算法和模型的有效性, 本节选用两个著名的行人轨迹数据预测数据集 ETH<sup>[22]</sup> 和 UCY<sup>[23]</sup> 进行仿真实验. 其中, ETH 包含两个场景, 分别表示为 ETH 和 HOTEL. UCY 包含 3 个场景, 分别表示为 ZARA1、ZARA2 和 UNIV. 两个数据集中的样本都是在 8 s 内以 0.4 s 为一个时间步长进行采样. 这些数据涵盖了拥挤场景下的 1 536 名行人和多种具有挑战性的场景, 如人群交叉、避免碰撞、群体形成和分散等. 为了与其他方法进行公平的比较, 该方法的实验设置遵循 Social-GAN. 在训练和评估过程中, 使用前 3.2 s (8 帧) 作为观测历史, 其余 4.8 s (12 帧) 作为真

实轨迹. 评价采用了与 Social-GAN 一样的指标, 即平均位移误差 (ADE) 和最终位移误差 (FDE). 这两个指标分别定义如下.

平均位移误差 (ADE): 在所有预测的时间步长上, 真实轨迹与预测轨迹之间的平均  $L_2$  距离.

最终位移误差 (FDE): 在预测周期  $T_{\text{pred}}$  结束时, 预测的最终目的地与真实的最终目的地之间的距离.

### 2.2 实施环境与评估方法

实验在 Pytorch 1.10.2 的环境中运行, 训练过程使用 Nvidia GTX-3060Ti GPU. 该模型被训练了 400 个 epoch, batch-size 为 64. 初始学习速率设置为 0.001. 其中 GCN 有两层, 两个 GCN 层的嵌入大小设置相同, 以减少超参数调优的工作量.

本文根据文献[9-11], 对预测性能的评估方法采用留一法. 该模型在 4 个数据集上进行训练, 并在其余的数据集上进行测试, 并且观察一个 3.2 s (8 个时间步长) 的轨迹, 然后预测接下来 3.2 s (8 个时间步长) 和 4.8 s (12 个时间步长) 的轨迹.

### 2.3 实验结果与分析

本节首先对行人轨迹预测模型 DEGCN 进行消融研究; 然后将该模型在 ETH 和 UCY 测试集上与 LSTM、S-LSTM、SGAN 和 STGAT 模型进行 ADE 和 FDE 指标的对比; 最后对 DEGCN 模型的多样性损失函数、运行速度和定性结果进行对比分析.

#### 2.3.1 消融研究

为了验证 DEGCN 模型的性能, 本节在 ETH 和 UCY 数据集上进行消融研究. 本文将提出的轨迹预测模型完整方法称为 DEGCN- $kV$ - $N$ , 其中  $kV$  表示模型是否使用多样性损失函数 ( $k = 1$  表示没有多样性损失函数,  $N$  表示在测试期间内从模型中取样的时间数), 这里的参数含义与 STGAT、SGAN 模型的参数含义一样. 在测试时, DEGCN 模型进行多次采样, 并选择  $L_2$  意义下的最佳预测进行定量评价 ( $L_2$  指生成的样本离实际真实轨迹的距离).

#### 2.3.2 定量评价

实验选用 4 种模型与 DEGCN 模型进行对比. 表 1 和表 2 显示了 5 种模型在  $T_{\text{pred}} = 8$  和  $T_{\text{pred}} = 12$  时的两个误差指标 ADE 和 FDE. 实验结果表明: DEGCN 模型的预测性能明显优于 LSTM 和 S-LSTM. 当  $k = 1, N = 1$  时, DEGCN-1V-1 模型的预测性能优于 STGAT-1V-1 模型和 SGAN-1V-1 模型的预测性能. 这是因为 DFU 能捕获行人的动态关系, 使模型能够更好地对多智能体 (行人) 的交互关系进行建模, 从而提高了模型的预测性能. 当  $k = 1, N = 20$  时,

DEGCN-1V-20模型的预测性能优于STGAT-1V-20模型的预测性能,并且由于在测试期间内对模型进行了取样,DEGCN-1V-20模型的预测性能优于DEGCN-1V-1模型的预测性能.当 $k = 20$ ,  $N = 20$ 时,由于使用了多样性损失函数,DEGCN-20V-20模型的预测性能最优.与SGAN-20V-20模型相比,在预测未来8

和12个时间步长下,DEGCN-20V-20模型ADE性能指标分别下降了30.0%和38.1%,FDE的性能分别提高了34.5%和37.2%.与STGAT-20V-20模型相比,在预测未来12个时间步长下,DEGCN-20V-20模型的最终位移误差稍逊于STGAT-20V-20模型,但总体上显示出了模型的竞争力.

表1 实验结果比较(预测未来12s轨迹)

方法	ETH				UCY						平均误差		
	ETH		HOTEL		UNIV		ZARA1		ZARA2		AVG		
	ADE	FDE	ADE	FDE	ADE	FDE	ADE	FDE	ADE	FDE	ADE	FDE	
LSTM	1.09	2.41	0.86	1.91	0.61	1.31	0.41	0.88	0.53	1.11	0.70	1.52	
S-LSTM	1.09	2.35	0.79	1.76	0.67	1.40	0.47	1.00	0.56	1.17	0.72	1.54	
SGAN	1V-1	1.13	2.21	1.01	2.18	0.60	1.28	0.42	0.91	0.52	1.11	0.74	1.54
	20V-20	0.81	1.52	0.72	1.61	0.60	1.26	0.34	0.69	0.42	0.84	0.58	1.18
	20VP-20	0.87	1.62	0.67	1.37	0.76	1.52	0.35	0.68	0.42	0.84	0.61	1.21
STGAT	1V-1	<b>0.88</b>	<b>1.66</b>	0.56	1.15	<b>0.52</b>	<b>1.13</b>	<b>0.41</b>	<b>0.91</b>	<b>0.31</b>	<b>0.68</b>	0.54	1.11
	1V-20	<b>0.80</b>	<b>1.53</b>	0.52	1.08	<b>0.51</b>	<b>1.12</b>	<b>0.39</b>	<b>0.87</b>	<b>0.30</b>	<b>0.64</b>	0.50	1.05
	20V-20	<b>0.65</b>	<b>1.12</b>	0.35	0.66	0.52	1.10	0.34	0.69	<b>0.29</b>	<b>0.60</b>	0.43	<b>0.83</b>
DEGCN	1V-1	<b>0.88</b>	1.93	<b>0.32</b>	<b>0.62</b>	<b>0.52</b>	1.16	0.43	0.95	0.33	0.74	<b>0.50</b>	<b>1.08</b>
	1V-20	0.81	1.77	<b>0.30</b>	<b>0.58</b>	0.52	1.14	0.41	0.92	0.32	0.71	<b>0.47</b>	<b>1.02</b>
	20V-20	0.66	1.28	<b>0.28</b>	<b>0.53</b>	<b>0.50</b>	<b>1.08</b>	0.36	0.75	0.31	0.68	<b>0.42</b>	0.86

表2 实验结果比较(预测未来8s轨迹)

方法	ETH				UCY						平均误差		
	ETH		HOTEL		UNIV		ZARA1		ZARA2		AVG		
	ADE	FDE	ADE	FDE	ADE	FDE	ADE	FDE	ADE	FDE	ADE	FDE	
LSTM	0.70	1.45	0.55	1.17	0.36	0.77	0.25	0.53	0.31	0.65	0.43	0.91	
S-LSTM	0.73	1.48	0.49	1.01	0.41	0.84	0.27	0.56	0.33	0.70	0.45	0.91	
SGAN	1V-1	0.79	1.61	0.71	1.44	0.37	0.75	0.25	0.53	0.32	0.66	0.49	1.00
	20V-20	0.61	1.22	0.48	0.95	0.36	0.75	0.21	0.42	0.27	0.54	0.39	0.78
	20VP-20	0.60	1.19	0.52	1.02	0.44	0.84	0.22	0.43	0.29	0.58	0.41	0.81
STGAT	1V-1	0.75	1.55	0.43	0.88	<b>0.31</b>	<b>0.66</b>	<b>0.25</b>	<b>0.53</b>	<b>0.21</b>	<b>0.44</b>	0.39	0.81
	1V-20	0.74	1.52	0.42	0.85	<b>0.31</b>	<b>0.65</b>	<b>0.24</b>	<b>0.50</b>	<b>0.20</b>	<b>0.42</b>	0.38	0.79
	20V-20	0.56	1.10	0.27	0.50	0.32	0.66	0.21	<b>0.42</b>	<b>0.20</b>	<b>0.40</b>	0.31	0.62
DEGCN	1V-1	<b>0.67</b>	<b>1.35</b>	<b>0.26</b>	<b>0.48</b>	<b>0.31</b>	0.67	<b>0.25</b>	0.54	<b>0.21</b>	0.46	<b>0.34</b>	<b>0.70</b>
	1V-20	<b>0.61</b>	<b>1.23</b>	<b>0.24</b>	<b>0.45</b>	<b>0.31</b>	0.66	<b>0.24</b>	0.52	<b>0.20</b>	0.44	<b>0.32</b>	<b>0.66</b>
	20V-20	<b>0.54</b>	<b>1.01</b>	<b>0.23</b>	<b>0.42</b>	<b>0.30</b>	<b>0.63</b>	0.22	0.44	<b>0.20</b>	0.42	<b>0.30</b>	<b>0.58</b>

### 2.3.3 多样性损失函数评估

由于行人运动的多样性,DEGCN模型使用多样性损失函数来产生多个社会可接受轨迹.在表1和表2中表示了DEGCN模型的3种不同的控制设置. DEGCN-1V-20模型和DEGCN-20V-20模型可以产生多种未来的发展轨迹.利用多样性损失函数,在预

测8和12个时间步长下,相比于DEGCN-1V-20模型,DEGCN-20V-20模型的ADE分别降低了6.7%和11.9%,FDE分别降低了13.8%和18.6%.这是因为多样性损失函数鼓励网络产生不同的样本.同时,以DEGCN-1V-N模型为例,在训练中为每个序列抽取一个样本,在测试中抽取样本 $N$ ,如图5所示.实验结

果表明,随着抽取样本  $N$  数量的增加,ADE和FDE显著降低,但模型运行速度显著下降. 因此,在实际应用中选用合适的样本数量是尤为重要的.

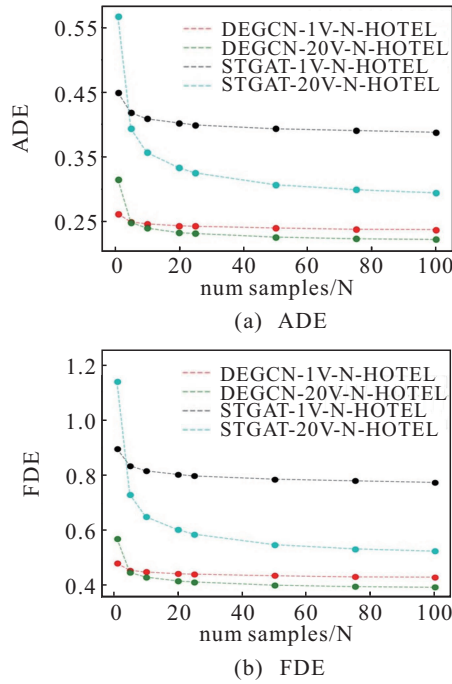


图5 多样性损失函数对DEGCN的影响

另外,为测试参数  $k$  对模型性能的影响,选取  $k = 50$  对模型进行训练,训练和测试均采用 HOTEL 数据集,结果如表3所示. 实验结果表明,当  $k = 50$  时,DEGCN模型的预测效果明显优于STGAT,并且随着  $k$  的增加,DEGCN的预测精度也随之提升,但训练时间明显加长. 其原因是通过只考虑最佳的轨迹,多样性损失函数鼓励网络对冲赌注,并覆盖符合过去轨迹的输出空间,以提高模型的预测性能.

表3  $k$  对模型的影响

模型	ADE	FDE
DEGCN-50V-1	0.31	0.56
DEGCN-50V-20	0.24	0.43
DEGCN-50V-50	0.23	0.41
STGAT-50V-1	0.64	1.28
STGAT-50V-20	0.33	0.61
STGAT-50V-50	0.31	0.55

### 2.3.4 推断时间比较

在实际应用中,比如汽车的自动驾驶需要对行人的行为进行准确的预测,因此模型推断时间是至关重要的. 本节将 DEGCN 与 SGAN 和 STGAT 进行以秒为单位的推断时间比较,所有的模型都在同一数据集和一个 Nvidia GTX 3060TiGPU 上进行基准测试,如表4所示. 由表4可知,DEGCN模型推断时间比STGAT减少了29.52%,与SGAN的推断时间相

当. 由此可见,DEGCN模型中的DFU模块比STGAT模型中的GAT模块推断时间更快,更加适合实际的应用. 这主要得益于DEGCN采用了EGCU,减少了网络参数,提高了模型的学习效率.

表4 推断时间比较

时间步长	S-LSTM	SGAN-P	STGAT	DEGCN
8	798.206	17.837	25.028	19.948
12	970.659	18.595	24.482	18.191
AVG	884.433	18.216	24.755	19.070
推断时间	48.55x	1x	1.36x	1.05x

### 2.3.5 定性评价

在行人运动场景中,行人间会产生多种潜在的交互关系,从而在整个系统层面上产生复杂的行为模式,包括形成群体、跟随他人、改变方向、避免碰撞等,因此模拟行人的行为如何影响其他行人的行为至关重要. 在现实世界中,行人通常是平稳移动的,只有在紧急情况下或经过拐角时,才会发生突然的运动变化.

本文提出的模型使用数据驱动方法,通过动态图卷积网络来模拟人与人之间的交互关系. 本节分析了DEGCN模型在不同的运动模式和避免碰撞的场景下所做出的预测,并且选择了有多个行人和复杂交互关系的场景,让DEGCN-20V-20与STGAT-20V-20在6种不同的情况下进行了比较,结果如图6所示. 实验结果表明,STGAT (STGAT-20V-20)模型可以捕获行人间的交互关系,并在大多数情况下生成社会可以接受的轨迹,但是STGAT预测的轨迹有时会在不合理的转弯,这未能考虑时间的连续性. 相比于STGAT,DEGCN由于考虑到了时间的依赖性和连续性,使得模型可以预测更合理且平滑的轨迹.

为了观察DEGCN模型生成的各种预测轨迹,本节选用一个具有挑战性的场景,在这个场景中,有多个行人并且行人与行人之间有着复杂的交互关系,如图7所示. 图7选择了DEGCN模型生成的4种不同的预测,其中图7(d)代表了最接近真实轨迹的预测(称之为最佳预测). 可以看出,DEGCN也可以生成非线性轨迹,通过模型的不断训练,DEGCN模型的预测精度逐渐上升,生成轨迹的社会接受度也逐渐上升. 这些定性结果直观地表明,考虑到行人相互作用的动态关系时,预测的轨迹变的更加准确(特别是在预测线性轨迹时),并且能够生成更加可靠的行人轨迹.

针对现实世界中行人运动轨迹的特殊情况(如行人突然停止、突然折返和避开障碍物等), DEGCN

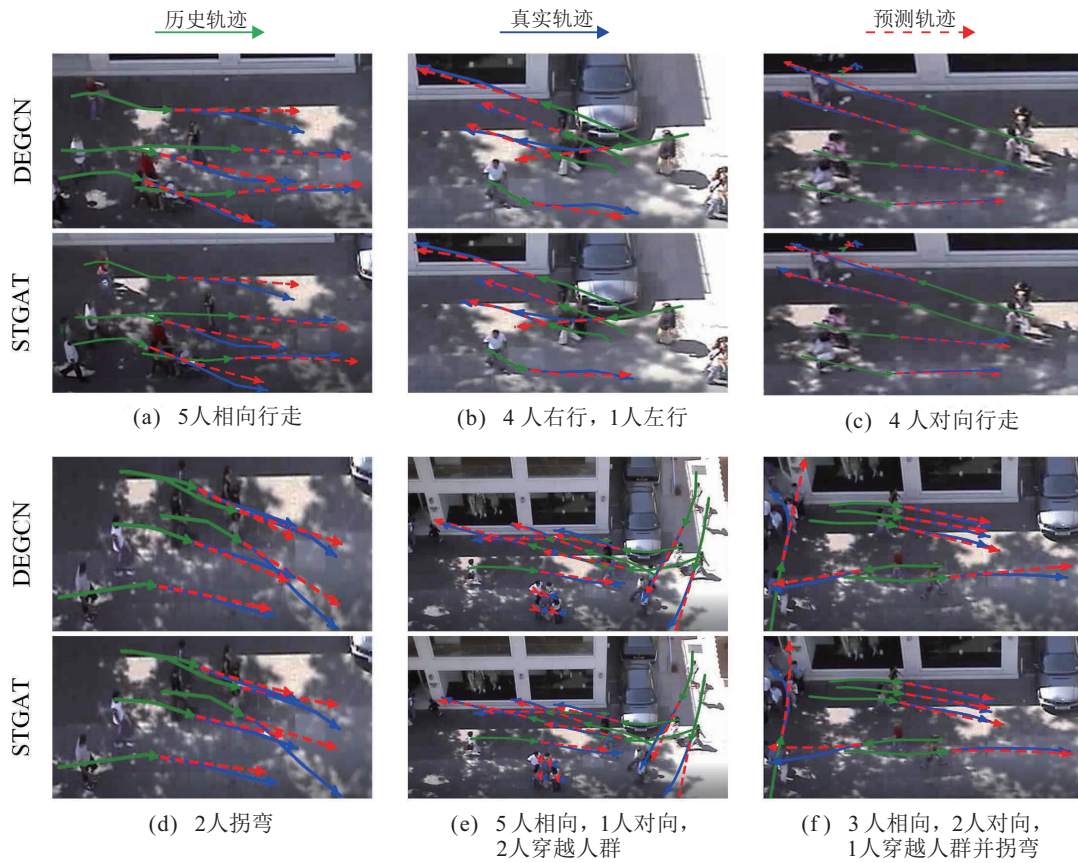


图6 DEGCN-20V-20模型与STGAT-20V-20模型的在6种场景下的比较

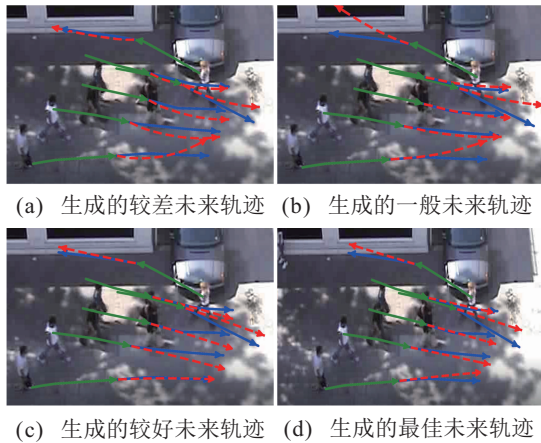


图7 DEGCN模型生成的不同未来轨迹

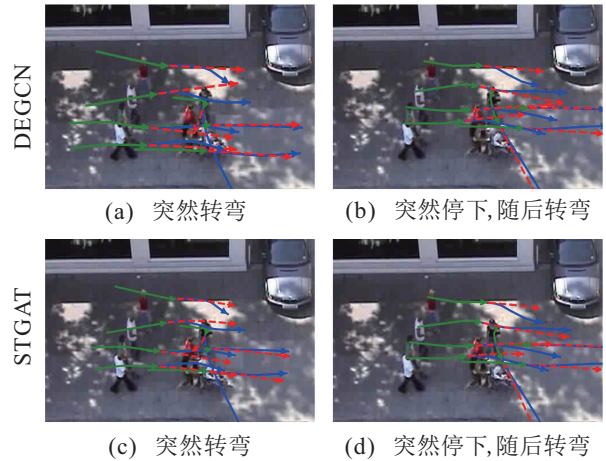


图8 特殊场景的比较

模型和STGAT模型都无法预测正在行走的行人突然折返、转弯或停下,如图8所示.这是由于模型进行预测依赖于过去观测的轨迹.但在行人折返或者转弯后,DEGCN模型对行人未来的前进方向的判断优于STGAT模型.图8显示了一位行人在向前行走的过程中,突然停下,随后转弯,可以看出STGAT模型因为受限于注意力机制的影响,会产生非线性的轨迹,这不太符合现实生活中行人实际的路线轨迹.而DEGCN模型生成的轨迹更加符合真实世界中行人的前进方向和行动路线,这是因为DEGCN模型另外关注了行人的特征在时间维度上的更新.

### 3 结论

本文提出了一种基于动态进化图的行人轨迹预测方法(DEGCN),目的是解决在人群运动场景中,交互关系无法随着时间推移而改变,图模型无法自适应地调整图模型参数的问题.在广泛使用的ETH和UCY测试数据上,将DEGCN模型与4种方法进行了比较,在不采用多样性损失函数时,多数场景中的ADE和FDE的表现均优于其他方法;采用多样性损失函数时,模型在5个开放的场景数据集上取得了优异的实验性能;在预测8个时间步长时,预测精度最优,平均位移误差降低12.26%,最终位移误差降低

14.10%。实验结果验证了DFU能够提高模型的交互建模能力,通过EGCU进化图模型网络参数,可提高网络的稳定性和自适应能力,预测更加可靠的行人轨迹。基于动态进化图在解决人群互动场景中,进一步表明了复杂的轨迹预测问题上的竞争力。但在预测12个时间步长时,本文方法的预测效果还有提升空间,在接下来的工作中,对编码-解码框架和图神经网络进行改进,设法提高模型的长期预测效果,以此提高本文方法在各场景中的预测精度。

#### 参考文献(References)

- [1] 任柯燕, 谷美颖, 袁正谦, 等. 自动驾驶3D目标检测研究综述[J]. 控制与决策, 2023, 38(4): 865-889. (Ren K Y, Gu M Y, Yuan Z Q, et al. 3D object detection algorithms in autonomous driving: A review[J]. Control and Decision, 2023, 38(4): 865-889.)
- [2] Toledo-Moreo R, Zamora-Izquierdo M A. IMM-based lane-change prediction in highways with low-cost GPS/INS[J]. IEEE Transactions on Intelligent Transportation Systems, 2009, 10(1): 180-185.
- [3] Lefèvre S, Laugier C, Ibañez-Guzmán J. Exploiting map information for driver intention estimation at road intersections[C]. 2011 IEEE Intelligent Vehicles Symposium. Baden-Baden, 2011: 583-588.
- [4] Peng Y S, Zhang G F, Shi J, et al. SRAI-LSTM: A social relation attention-based interaction-aware LSTM for human trajectory prediction[J]. Neurocomputing, 2022, 490: 258-268.
- [5] Rossi L, Paolanti M, Pierdicca R, et al. Human trajectory prediction and generation using LSTM models and GANs[J]. Pattern Recognition, 2021, 120: 108136.
- [6] Zhou H, Ren D C, Xia H X, et al. AST-GNN: An attention-based spatio-temporal graph neural network for Interaction-aware pedestrian trajectory prediction[J]. Neurocomputing, 2021, 445: 298-308.
- [7] Liang J W, Jiang L, Murphy K, et al. The garden of forking paths: Towards multi-future trajectory prediction[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, 2020: 10505-10515.
- [8] Kong W, Liu Y, Li H, et al. GSTA: Pedestrian trajectory prediction based on global spatio-temporal association of graph attention network[J]. Pattern Recognition Letters, 2022, 160: 90-97.
- [9] 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, 2016: 961-971.
- [10] 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. Salt Lake City, 2018: 2255-2264.
- [11] Huang Y F, Bi H K, Li Z X, et al. STGAT: Modeling spatial-temporal interactions for human trajectory prediction[C]. 2019 IEEE/CVF International Conference on Computer Vision. Seoul, 2020: 6271-6280.
- [12] Veličković P, Cucurull G, Casanova A, et al. Graph attention networks[J/OL]. 2017, arXiv: 1710.10903.
- [13] Yu C J, Ma X, Ren J W, et al. Spatio-temporal graph transformer networks for pedestrian trajectory prediction[C]. European Conference on Computer Vision. Cham: Springer, 2020: 507-523.
- [14] 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 W, Liu Y, Li H, et al. Survey of pedestrian trajectory prediction methods based on deep learning[J]. Control and Decision, 2021, 36(12): 2841-2850.)
- [15] 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.)
- [16] Kipf T, Fetaya E, Wang K C, et al. Neural relational inference for interacting systems[J/OL]. 2018, arXiv: 1802.04687.
- [17] Li J C, Yang F, Tomizuka M, et al. EvolveGraph: Multi-agent trajectory prediction with dynamic relational reasoning[J/OL]. 2020, arXiv: 2003.13924.
- [18] Huang X H, Ye Y M, Ding W H, et al. Multi-mode dynamic residual graph convolution network for traffic flow prediction[J]. Information Sciences, 2022, 609: 548-564.
- [19] Sankar A, Wu Y H, Gou L, et al. DySAT: Deep neural representation learning on dynamic graphs via self-attention networks[C]. Proceedings of the 13th International Conference on Web Search and Data Mining. New York, 2020: 519-527.
- [20] Pareja A, Domeniconi G, Chen J, et al. EvolveGCN: Evolving graph convolutional networks for dynamic graphs[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(4): 5363-5370.
- [21] Tang H W, Wei P, Li J P, et al. EvoSTGAT: Evolving spatiotemporal graph attention networks for pedestrian trajectory prediction[J]. Neurocomputing, 2022, 491: 333-342.
- [22] 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. Kyoto, 2010: 261-268.
- [23] Lerner A, Chrysanthou Y, Lischinski D. Crowds by example[J]. Computer Graphics Forum, 2007, 26(3): 655-664.

#### 作者简介

聿菁(1998—), 女, 硕士生, 从事多智能体轨迹预测、深度学习等研究, E-mail: 18303847790@163.com;  
张旭秀(1968—), 女, 教授, 博士, 从事多智能体轨迹预测、信号处理等研究, E-mail: zhangxuxiu@163.com;  
闫涵(1994—), 男, 博士生, 从事深度学习、图像识别等研究, E-mail: yan\_han@mail.dlut.edu.cn.