



分析力学和图神经网络的轨迹预测方法

李明晗, 肖阳, 邢向磊

引用本文:

李明晗, 肖阳, 邢向磊. 分析力学和图神经网络的轨迹预测方法[J]. *智能系统学报*, 2025, 20(6): 1355-1365.

LI Minghan, XIAO Yang, XING Xianglei. Trajectory prediction methods based on analytical mechanics and graph neural networks[J]. *CAAI Transactions on Intelligent Systems*, 2025, 20(6): 1355-1365.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202501020>

您可能感兴趣的其他文章

视觉协同的违规驾驶行为分析方法

A visual collaborative analysis method for detecting illegal driving behavior

智能系统学报. 2021, 16(6): 1158-1165 <https://dx.doi.org/10.11992/tis.202101024>

一种基于2D时空信息提取的行为识别算法

A behavioral recognition algorithm based on 2D spatiotemporal information extraction

智能系统学报. 2020, 15(5): 900-909 <https://dx.doi.org/10.11992/tis.201906054>

旅游知识图谱特征学习的景点推荐

Tourism knowledge-graph feature learning for attraction recommendations

智能系统学报. 2019, 14(3): 430-437 <https://dx.doi.org/10.11992/tis.201810032>

基于自组织递归模糊神经网络的PM2.5浓度预测

Prediction of PM2.5 concentration based on self-organizing recurrent fuzzy neural network

智能系统学报. 2018, 13(4): 509-516 <https://dx.doi.org/10.11992/tis.201710007>

高斯核函数卷积神经网络跟踪算法

Convolutional neural network tracking algorithm accelerated by Gaussian kernel function

智能系统学报. 2018, 13(3): 388-394 <https://dx.doi.org/10.11992/tis.201612040>

基于深度学习的视频预测研究综述

Review of deep learning-based video prediction

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

DOI: 10.11992/tis.202501020

网络出版地址: <https://link.cnki.net/urlid/23.1538.tp.20250911.0946.002>

分析力学和图神经网络的轨迹预测方法

李明哈, 肖阳, 邢向磊

(哈尔滨工程大学智能科学与工程学院, 黑龙江哈尔滨 150001)

摘要: 轨迹预测旨在通过分析智能体的历史运动数据来预测轨迹。然而, 现有深度学习因忽略物理约束和运动规律导致可解释性不足。针对此问题, 提出一种融合分析力学与图神经网络的轨迹预测模型, 利用图神经网络、卷积神经网络和图注意力机制提取目标时空动态特征, 结合欧氏距离和相对运动推断交互力, 并引入拉格朗日力学对动力学过程和约束进行建模, 以生成符合物理规律的轨迹。在 Spring-balls 数据集上的实验表明, 模型在 5 球 10 帧短期预测中精度提升 14.29%。在更具挑战性的 50 帧长期预测中, 5 球和 10 球场景下的精度分别提升 6.25% 和 4.81%。此外, 在人体运动预测中, 该模型平均多关节位置误差 (mean per joint position error, MPJPE) 均优于主流方法, 验证了其在长期预测中的更高精度和有效性。

关键词: 轨迹预测; 图神经网络; 拉格朗日力学; 交互作用力; 深度学习; 多层感知机; 运动学模型; 分析力学
中图分类号: TP391.41 **文献标志码:** A **文章编号:** 1673-4785(2025)06-1355-11

中文引用格式: 李明哈, 肖阳, 邢向磊. 分析力学和图神经网络的轨迹预测方法 [J]. 智能系统学报, 2025, 20(6): 1355-1365.

英文引用格式: LI Minghan, XIAO Yang, XING Xianglei. Trajectory prediction methods based on analytical mechanics and graph neural networks[J]. CAAI transactions on intelligent systems, 2025, 20(6): 1355-1365.

Trajectory prediction methods based on analytical mechanics and graph neural networks

LI Minghan, XIAO Yang, XING Xianglei

(College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin 150001, China)

Abstract: Trajectory prediction seeks to forecast the future motion of intelligent agents by analyzing their past trajectories. While deep learning methods have been demonstrated to capture complex features, they frequently neglect physical constraints, thereby constraining interpretability. To address this, a trajectory prediction model is proposed that integrates analytical mechanics with graph neural networks (GNNs). The model combines GNNs, convolutional neural networks, and graph attention to extract spatiotemporal dynamics, infers interaction forces from Euclidean distance and relative motion, and incorporates Lagrange mechanics to enforce physical laws. Experiments on the Spring-balls dataset demonstrate the superior performance of the proposed model in comparison to traditional models, exhibiting a 14.29% accuracy gain in 10-frame short-term prediction for the 5-ball case and improvements of 6.25% and 4.81% in 50-frame long-term scenarios. In the domain of human motion prediction, our model demonstrates a reduction in mean position error (MPJPE) when compared to prevailing approaches for a wide range of actions. This finding signifies enhanced long-term accuracy and validates the efficacy of the model.

Keywords: trajectory prediction; graph neural networks; Lagrangian mechanics; interaction force; deep learning; multi-layer perceptron; kinematic model; analytical mechanics

随着计算机和人工智能的快速发展, 人工智能技术在各个领域得到了广泛的应用, 显著提高了人们的工作效率。自动驾驶、智能设备、智能

交通等涉及动态环境和多智能体交互的领域对理解和模拟人类行为的需求日益增加, 准确预测智能体轨迹对推动相关领域的技术进步和应用发展具有重要的实际意义。因此, 设计一个准确、鲁棒且可解释的预测模型显得尤为重要。

目前, 轨迹预测的方法可以分为基于知识、

收稿日期: 2025-01-26. 网络出版日期: 2025-09-11.

基金项目: 国家自然科学基金项目 (62076078, 61703119).

通信作者: 邢向磊. E-mail: xingxl@hrbeu.edu.cn.

基于学习和基于混合的方法。基于知识的轨迹预测算法利用先验知识、物理规律或数学模型对人类行为进行建模,如建立运动学模型、动力学模型、社会力模型等,模型根据目标的当前状态、速度、加速度等信息,结合环境约束条件推断未来的轨迹。Barth等^[1]采用包含位置、方向、速度、加速度和偏航率的运动学模型描述车辆的运动状态,通过扩展卡尔曼滤波器(extended Kalman filter, EKF)融合车辆运动信息,预测车辆的未来行驶轨迹。Kooij等^[2]从车辆视角使用线性运动学模型描述行人的运动状态,并计算行人未来位置的概率分布。Karamouzas等^[3]基于统计力学(statistical mechanical)计算行人之间相互作用的能量,描述人类在不同场景、速度和密度下的相互作用行为,进而模拟人群流动和预测行人行为。Batz等^[4]根据车辆的位置、速度、偏航率等状态信息,通过改进的卡尔曼滤波器(Kalman filter, KF)预测一定范围内其他车辆的轨迹,用来实时地检测危险情况。Pang等^[5]使用漏桶基于能量概率模型^[6](leaky bucket-energy based model, LB-EBM),从学习到的LB-EBM中采样或优化会产生一个信念向量,该信念向量可用于制定路径规划,进而帮助预测长距离轨迹。然而当前采用单模态高斯分布作为输出分布,限制了对更复杂分布(如混合高斯分布)的建模能力。

基于学习的方法适用于物理模型难以捕捉的复杂系统,但需要大量的高质量数据,模型性能受限于数据的代表性和多样性,可能出现过拟合或泛化能力差的问题,导致预测结果不符合实际情况。在轨迹预测领域,早期的研究多依赖于传统的机器学习模型,如支持向量回归(support vector regression, SVR)结合优化算法(如遗传算法)已被用于提高预测精度^[7]。Liu等^[8]使用卷积神经网络(convolutional neural networks, CNN)学习人体运动序列中隐藏的时间和空间结构信息。Aksan等^[9]提出一种时空自注意力机制,结合Transformer架构,显式地从当前和过去的信息中捕捉时间依赖性和空间结构,使模型更好地理解运动的动态变化。Mao等^[10]利用离散余弦变换(discrete cosine transform, DCT)表示时间信息,并设计可学习的图卷积网络以捕获人体关节的长距离依赖关系。然而,关节角度表示可能导致多解性问题,即不同角度组合可能对应相同姿态,这使得基于角度的评价结果可能不完全准确。Dang等^[11]首次提出从细粒度到粗粒度的降维路径,再通过从粗粒度到细粒度的升维路径,使在不同尺

度下提取的特征相互补充。虽然该方法在高分辨率3D运动捕捉数据上表现出色,但其对低质量或噪声数据的适应性仍需深入研究,以确保在实际应用中的鲁棒性。Li等^[12]将人体分解为多个部分分别进行特征提取,并通过可训练的频带滤波器实现图谱的多频段分解,捕捉低频和高频特征,从而保留丰富的运动信息。但是自适应图散射和多部分建模增加了模型复杂性,同时尽管采用了多频段分解和部分分离,但未深入分析不同频段或部分的特定贡献,使得模型的可解释性降低。上述基于学习的方法能够从轨迹数据中学习复杂的模式,能够提供更好的预测性能,可以适应不同数据集和环境,但解释性较差,难以理解模型的决策过程^[13]。

由此可见,轨迹预测存在一些关键挑战:1)基于知识的轨迹预测算法通常建立在物理方程或知识基础之上,预测结果具有较强的可解释性,但对复杂场景的适应性差;2)基于数据驱动的深度学习方法能够提供良好的预测性能,但可解释性较差^[14-15];3)目标之间存在复杂的动态交互,会对未来轨迹产生影响。

本文通过拉格朗日力学描述系统的动力学行为,通过参数化分析力学中的物理量实现物理知识与深度学习方法的结合,提高模型的可解释性。针对多目标之间的交互作用会对未来轨迹产生影响,本文利用图神经网络推断场景中多目标之间的交互强度,设计了空时特征提取模块和交互作用力推理模块,对目标之间交互关系和约束条件进行建模,模型实现多个场景下目标未来轨迹的准确预测并表现出了优秀的长期预测性能,实现了从数据中推断出目标的动力学信息,并能应用于受物理约束的系统。同时针对传统保守场物理模型建模的局限,学习非保守力约束条件,并融入拉格朗日函数,更准确地描述智能体行为^[16-17]。本文通过在两种主流数据集上进行轨迹预测实验,并与主流模型进行对比,验证了所提出模型的有效性。

1 相关知识

1.1 拉格朗日力学建模

拉格朗日力学是经典力学的一种表述形式,提供了一种更为简洁的方法来描述物体的运动。拉格朗日力学是一种以能量和广义坐标为基础的力学理论,其核心思想是通过建立系统的动力学方程来描述物体的运动,而不是直接考虑力的作用^[18]。

在拉格朗日力学中,使用拉格朗日函数来描

述系统的动力学行为, 每一个力学系统都可以由一个确定的拉格朗日函数 $L(\mathbf{q}, \dot{\mathbf{q}}, t)$ ^[19] 所表征。拉格朗日力学基于最小作用量原理, 该原理认为一个物理系统在一段时间的运动轨迹会使作用量取极小值, 作用量表示为

$$S = \int_{t_1}^{t_2} L(\mathbf{q}, \dot{\mathbf{q}}, t) dt \quad (1)$$

式中: \mathbf{q} 表示广义坐标, $\dot{\mathbf{q}}$ 表示广义速度。由式 (1) 取极小值的条件, 通过变分法最小作用原理表示为

$$\delta S = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial \mathbf{q}} \delta \mathbf{q} + \frac{\partial L}{\partial \dot{\mathbf{q}}} \delta \dot{\mathbf{q}} \right) dt = 0$$

经过分部积分将进一步表示为

$$\delta S = \frac{\partial L}{\partial \dot{\mathbf{q}}} \delta \mathbf{q} \Big|_{t_1}^{t_2} + \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial \mathbf{q}} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{q}}} \right) \delta \mathbf{q} dt = 0$$

由此, 推导出拉格朗日方程:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{q}}} - \frac{\partial L}{\partial \mathbf{q}} = 0$$

该方程描述了系统的广义坐标随时间的演化过程。因此, 拉格朗日方程描述了系统的运动规律, 包括速度、加速度以及广义坐标与时间的关系。通过求解这些方程, 可以了解系统在给定条件下的轨迹、速度和加速度等信息。当系统中出现摩擦力、空气阻力等非保守力时, 拉格朗日方程为

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{q}}} - \frac{\partial L}{\partial \mathbf{q}} = \mathbf{Q} \quad (2)$$

式中 \mathbf{Q} 为非保守力。拉格朗日函数通常可以表示为动能 T 和势能 V 的差:

$$L(\mathbf{q}, \dot{\mathbf{q}}, t) = T(\dot{\mathbf{q}}) - V(\mathbf{q}) \quad (3)$$

将式 (2) 代入到式 (3) 中可以得到:

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{\mathbf{q}}} \right) - \frac{\partial T}{\partial \dot{\mathbf{q}}} = \frac{\partial V}{\partial \mathbf{q}} + \mathbf{Q} \quad (4)$$

式 (4) 将势能和动能统一起来, 适用于描述具有多个自由度的复杂系统的运动。在系统受到复杂约束时, 拉格朗日力学通过广义坐标可以更方便地处理约束条件。

1.2 GRU 循环神经网络

GRU(gated recurrent unit) 即门控循环单元, 是一种特殊的循环神经网络 (recurrent neural network, RNN) 变体。传统的 RNN 在处理长序列数据时, 容易出现梯度消失或梯度爆炸的问题, 导致模型难以学习到序列中的长期依赖关系。而 GRU 通过引入门控机制, 有效地缓解了这一问题。

在本研究中选择 GRU 网络, 有以下原因: 在处理大量轨迹数据时, 由于 GRU 结构相对简单,

相较于一些复杂的循环神经网络, 其门控单元数量更少, 计算过程中所需参数也更少, 这使得它在计算时速度更快, 有助于提高研究效率。此外, 轨迹数据本质上是时间序列数据, 其中包含长期依赖关系, 例如人体运动轨迹中, 当前时刻的动作可能受到之前较长时间动作的影响。GRU 能够很好地捕捉这种长期依赖关系, 使得模型能根据当前输入和前一时刻信息动态更新隐藏状态, 从而准确预测目标的未来轨迹。同时, GRU 的门控机制还有效缓解了梯度消失问题, 在训练过程中, 它能让梯度更稳定地反向传播, 避免梯度在传播过程中过度衰减, 进而增强模型的学习能力, 提升模型的性能和稳定性。

2 基于分析力学和图神经网络的轨迹预测

在轨迹预测领域, 常规的物理网络基于对常微分方程 (ordinary differential equation, ODE) 的数值计算, 通过明确的动力学模型对系统运动规律进行精确描述和预测。然而, 一些常规处理常微分方程的方法如: 龙格-库塔法 (Runge-Kutta)^[20] 在面对复杂场景和多智能体交互时通常计算复杂度高, 且对系统参数敏感, 鲁棒性较低。同时, 基于神经网络的数据驱动方法能够以较低计算开销直接从数据中学习复杂运动模式, 但在特征提取过程中容易忽略系统的物理约束和运动规律, 导致模型的可解释性较低。因此, 将分析力学与图神经网络相结合, 通过神经网络对物理方程进行参数化学习, 不仅能够降低传统数值积分方法的计算难度, 还能够通过网络优化提升对复杂场景的适应性与鲁棒性。这样的结合使得预测模型既具备物理约束的解释性, 又在复杂系统建模中表现出更高的灵活性、计算效率和可扩展性。

本文所提出的模型主要包含 3 部分: 空时特征提取模块、交互作用力推理模块以及分析力学模块, 模型整体结构如图 1 所示。首先将轨迹序列数据输入到空时特征提取模块中, 通过图卷积层、时间特征提取网络和图注意力层提取轨迹序列的空时特征, 然后将模型的输入和提取到的特征通过交互作用力推理模块推断场景中每个目标受到的交互作用力, 再通过分析力学模块根据每个目标的初始运动信息结合交互作用力推断出状态变化量, 经过 r 层交替训练后由分析力学模块输出最终预测结果。接下来的小节将对组成轨迹预测模型的基本模块以及损失函数进行详细阐述。

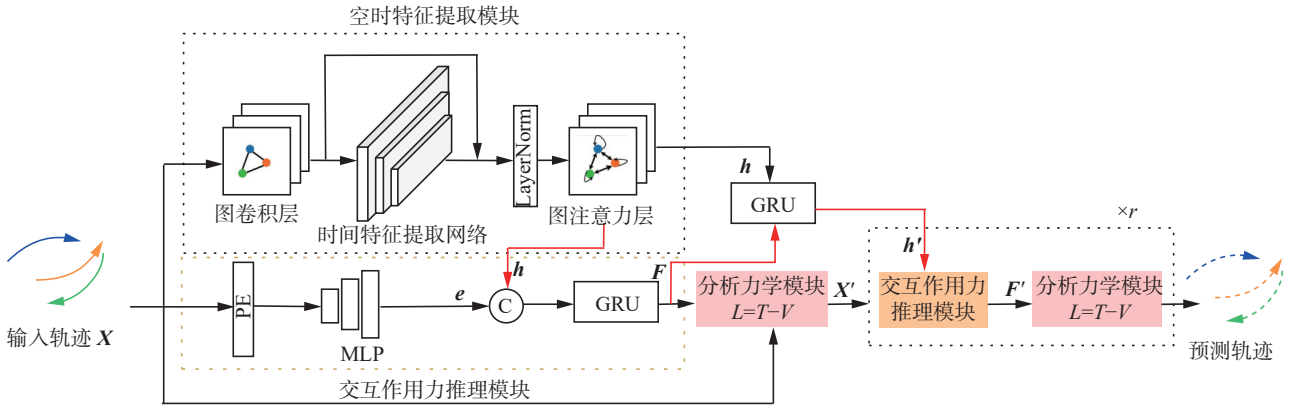


图 1 基于拉格朗日力学和图神经网络的轨迹预测模型

Fig. 1 Trajectory prediction model based on Lagrangian mechanics and graph neural networks

2.1 拉格朗日力学神经网络建模

在轨迹预测任务中,深度学习的方法通常缺乏先验知识或物理方程的约束,可能导致模型预测的不稳定性和不可解释性。例如,在没有适当约束的情况下,深度学习模型可能会生成不符合实际物理规则的预测,如穿过障碍物或者生成的轨迹不平滑。在轨迹预测任务中,场景中多目标交互会对未来轨迹产生影响。目标之间的交互作用会影响每个目标的运动决策和轨迹规划,导致未来轨迹不同于单个目标情况下的预期轨迹。

将物理方程和神经网络相结合可以改善轨迹预测模型的性能,这种方法利用神经网络来参数化物理方程中的未知参数或边界条件,既利用了神经网络强大的学习能力来拟合复杂的运动模式,又保留了物理方程中的约束和先验知识,从而实现对系统运动的准确建模和预测^[21]。为了提升模型的可解释性,本文将拉格朗日力学引入到轨迹预测中,作为物理先验约束模型的预测过程。本文以多个质点目标在前 N 帧中的轨迹位置作为输入,利用位置序列计算出广义速度和加速度

等信息,进而通过拉格朗日力学构造动能和势能模型,并通过神经网络进行优化学习。最终,输出多个目标在未来 T 帧的轨迹位置,实现受物理约束的高精度预测。Lutter 等^[22]通过引入拉格朗日力学作为物理先验,证明了这一方法在捕捉系统动力学特征方面的有效性。

传统拉格朗日力学中,作用量的定义如式(1)所示,其中 L 为动能 T 和势能 S 的差。但在行人运动系统中存在交互作用力(非保守力),运动受到各种约束条件影响。约束条件分为两种:一种是行人间的交互作用力和摩擦力、空气阻力等非保守力作为广义力物理约束;另一种与行人的意图和动机有关,包含心理、社会和环境等非物理量约束,这些约束随着行人的位置和速度而变化,因此所做的功可能会因为路径的不同而不同。传统的拉格朗日函数不足以描述复杂的行人交互,需要引入额外项来表示约束条件,而这种条件难以显示表达,适合用神经网络来进行学习。如图 2 所示,将 $\psi(q; \gamma)$ 作为运动轨迹输入 MLP(multi-layer perception)网络建模的约束条件,来学习笛卡尔坐标系下的拉格朗日函数。

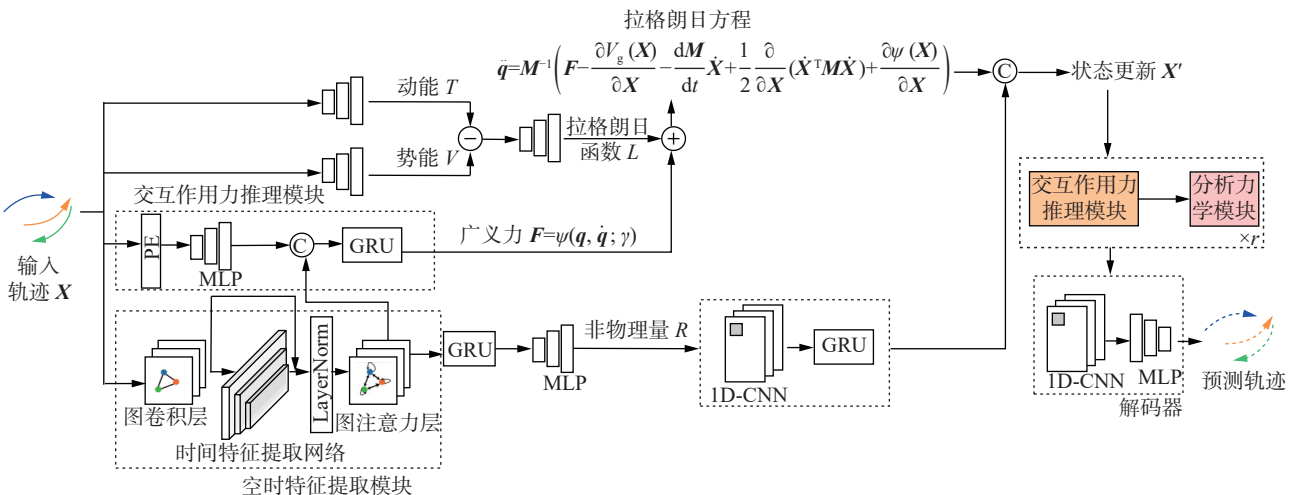


图 2 拉格朗日力学神经网络模型结构

Fig. 2 Structure of the Lagrangian mechanics neural network model

系统的作用量表示为

$$S = \int_{t_1}^{t_2} L(\mathbf{q}, \dot{\mathbf{q}}; \alpha, \beta) dt$$

$L(\mathbf{q}, \dot{\mathbf{q}}; \alpha, \beta)$ 为不含约束条件的拉格朗日函数。由最小作用原理, 推导出非保守系统拉格朗日方程为

$$\frac{d}{dt} \left(\frac{\partial (L(\mathbf{q}, \dot{\mathbf{q}}; \alpha, \beta))}{\partial \dot{\mathbf{q}}} \right) - \frac{\partial (L(\mathbf{q}, \dot{\mathbf{q}}; \alpha, \beta))}{\partial \mathbf{q}} = \mathbf{F} \quad (5)$$

式中: \mathbf{F} 为广义力, 如场景中摩擦力等非保守力。基于视频内容的轨迹预测中, 非保守力 \mathbf{F} 无法通过传感器等工具进行测量, 因此为了计算出广义加速度 $\ddot{\mathbf{q}}$, 并更新出轨迹预测位置 \mathbf{X} , 本文利用交互作用力推理模块对广义力 \mathbf{F} 进行建模, $\mathbf{F} = \psi(\mathbf{q}, \dot{\mathbf{q}}; \gamma)$ 式 (5) 中拉格朗日函数 $L = L(\mathbf{q}, \dot{\mathbf{q}}; \alpha, \beta)$ 用动能和势能来表示:

$$L = T - V = \frac{1}{2} \dot{\mathbf{q}}^T \mathbf{M}(\mathbf{q}; \alpha) \dot{\mathbf{q}} - V_g(\mathbf{q}; \beta) \quad (6)$$

式中: V_g 为 MLP 将势能建模为位移的函数。 T 为动能, 由速度和质量矩阵计算得到。 \mathbf{M} 为质量矩阵, 为确保质量矩阵是对称正定的, 将 \mathbf{M} 定义为

$$\mathbf{M} = \mathbf{J}_\alpha^T(\mathbf{q}) \mathbf{J}_\alpha(\mathbf{q}) + \lambda \mathbf{I}$$

式中: \mathbf{J}_α 为 MLP 进行端到端的学习, λ 为超参数, \mathbf{I} 为单位矩阵。在人体运动预测和行人轨迹预测中, 假设质量和约束关系 (如关节连接关系和关节长度) 不随时间变化, 将式 (6) 代入式 (5) 中, 得到:

$$\ddot{\mathbf{q}} = \mathbf{M}^{-1} \left(\mathbf{F} - \frac{\partial V_g(\mathbf{q})}{\partial \mathbf{q}} - \frac{d\mathbf{M}}{dt} \dot{\mathbf{q}} + \frac{1}{2} \frac{\partial}{\partial \mathbf{q}} (\dot{\mathbf{q}}^T \mathbf{M} \dot{\mathbf{q}}) \right) \quad (7)$$

其中 $\ddot{\mathbf{q}}$ 为广义加速度。

受到文献 [22] 的启发, 式 (7) 中求导和求偏导运算通过链式法则来计算, 计算过程为

$$\begin{aligned} \frac{d\mathbf{M}}{dt} &= \frac{d}{dt} (\mathbf{J}_\alpha^T(\mathbf{q}) \mathbf{J}_\alpha(\mathbf{q})) = \\ &= \mathbf{J}_\alpha^T(\mathbf{q}) \frac{d\mathbf{J}_\alpha(\mathbf{q})}{dt} + \frac{d\mathbf{J}_\alpha^T(\mathbf{q})}{dt} \mathbf{J}_\alpha(\mathbf{q}) \\ \frac{d\mathbf{J}_\alpha(\mathbf{q})}{dt} &= \frac{\partial \mathbf{J}_\alpha(\mathbf{q})}{\partial \mathbf{q}} \dot{\mathbf{q}} = \frac{\partial \mathbf{J}_\alpha(\mathbf{q})}{\partial \mathbf{h}_{l-1}} \frac{\partial \mathbf{h}_{l-1}}{\partial \mathbf{h}_{l-2}} \dots \frac{\partial \mathbf{h}_1}{\partial \mathbf{q}} \dot{\mathbf{q}} \\ \frac{\partial \mathbf{h}_i}{\partial \mathbf{h}_{i-1}} &= \text{diag}(\sigma'_i(\mathbf{W}_i^T \mathbf{h}_{i-1} + \mathbf{b}_i)) \mathbf{W}_i \end{aligned}$$

式中: \mathbf{h}_i 为 i 个网络中隐状态, 网络由激活函数 σ_i 和线性变换 \mathbf{W}_i 、 \mathbf{b}_i 组成。

2.2 空时特征提取模块

2.1 节详细地介绍了如何利用拉格朗日力学进行神经网络建模, 强调了在系统动力学建模中引入物理约束的重要性。然而, 仅靠拉格朗日建模无法充分捕获轨迹的多目标交互和动态特性。因此引入空时特征提取模块, 如图 2 所示, 该模块提取目标间的空间拓扑关系与时间动态信息, 生成的特征表示作为后续交互作用力推理模块的输入。在图 2 中, 该部分以 GCN (graph convolutional network)、1D-CNN 与多头注意力网络形式展现, 整体负责建模目标在时空域内的潜在结构依

赖。通过结合图神经网络和时间特征提取网络, 高效地从轨迹数据中学习隐含的非物理量, 为后续的交互作用力推理和动力学建模提供了准确且鲁棒的输入特征。这一设计显著提升了模型对复杂场景的适应性和预测精度。

为了有效地捕捉多目标之间的空间关系, 本文采用了图神经网络 (graph neural network, GNN) 来提取轨迹数据中的交互特征。与刘建敏等 [23] 和连静等 [24] 提出的基于图模型和注意力机制的车辆轨迹和场景预测方法在技术思路中具有相似性。有研究将基于网格的图神经网络应用于瞬态多相流场的高效时空建模和预测, 实现了对复杂物理过程的高效仿真 [25]。在相关研究中, 也常采用联合局部多尺度和全局上下文特征的方法来提升模型的表现力 [26]。这种方法受到 Kipf 等提出的 NRI (neural relational inference) 模型 [27] 的启发, 通过结合图神经网络和一维卷积网络来提取轨迹序列的空间和时间特征。由于数据中的空间关系对于预测结果有重要影响, 本文首先使用 GCN [11] 处理轨迹数据的空间关系:

$$\mathbf{h}^l = \sigma(\mathbf{A}^{l-1} \mathbf{h}^{l-1} \mathbf{W}^{l-1} + \mathbf{b}^{l-1})$$

式中: l 是 GCN 层数; \mathbf{A}^{l-1} 、 \mathbf{W}^{l-1} 和 \mathbf{b}^{l-1} 是可学习参数; σ 选用 SELU (scaled exponential linear unit) 激活函数; $\mathbf{h}^l = (h_1^l, h_2^l, \dots, h_N^l)$ 表示第 l 层卷积层的节点特征向量, N 是节点数目。在 $l=1$ 时, 有 $\mathbf{h}^0 = \mathbf{X}$, 即初始节点特征向量是输入轨迹序列。通过 l 层 GCN 能从轨迹数据中有效地聚合节点信息, 能充分提取到多目标之间的空间关系。为了增强空间和时间特征之间的关联性, 在最后一层图卷积网络后使用时间特征提取网络提取序列时间特征, 计算公式为

$$\hat{\mathbf{h}} = \text{LN}(\text{TFN}(\mathbf{h}^l) + \mathbf{h}^l) \quad (8)$$

式中: $\hat{\mathbf{h}} = (\hat{h}_1, \hat{h}_2, \dots, \hat{h}_N)$ 为更新后的特征向量, LN 为层归一化 (layer normalization)。时间特征提取网络 TFN (temporal feature network), 包括 1D-CNN、SELU 激活函数、最大池化层以及 Dropout 层, 结构如图 3 所示。

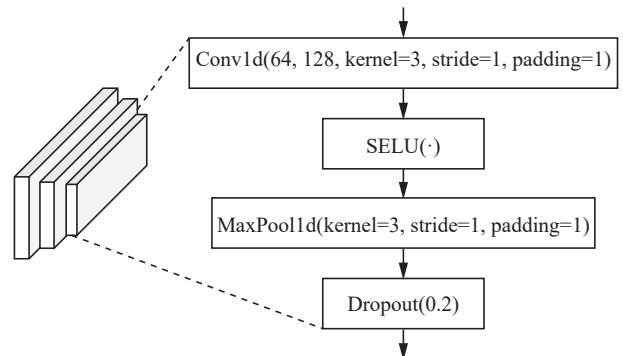


图 3 时间特征提取网络结构

Fig. 3 Temporal feature extraction network structure

一维卷积网络通过在时间序列数据上滑动卷积核来提取特征,能够捕捉到数据中的时间依赖性和局部特征,可以帮助模型学习到如速度和加速度等动态变化的模式和规律。一维卷积网络中输入通道数为 64,输出通道数为 128,卷积核大小设置为 3,卷积步长设置为 1,使用大小为 1 的填充。相比使用 RELU(rectified linear unit)激活函数,SELU 激活函数具有自归一化特性,可以确保网络输出的均值和方差在训练过程中保持稳定,减轻梯度消失问题。最大池化层有助于模型在减少计算量的同时,保持对关键信息的敏感性。本文使用一维最大池化层,池化核大小设置为 3,滑动步长设置为 1,使用大小为 1 的填充。在式(8)中使用了残差连接来更有效地传递和利用空间特征,减少空间信息的丢失,同时能缓解梯度消失和梯度爆炸问题。在轨迹预测任务中,不同特征的尺度可能差异很大,层归一化可以消除这种影响,使得模型更加关注特征之间的相对关系,而不是绝对值。

为了进一步学习节点间的复杂交互关系,本文使用多头图注意力动态地为每个节点分配不同的权重,加强重要的空间和时间特征表示,从而更好地捕获轨迹之间的空时依赖关系,计算公式为

$$\tilde{\mathbf{h}}_i = \text{GAT}(\mathbf{h}_j) = \sigma \left(\frac{1}{K} \sum_{k=1}^K \sum_{j \neq i} \alpha_{ij}^k \theta^k \hat{\mathbf{h}}_j \right) \quad (9)$$

式中: θ^k 为第 k 个注意力层的可学习参数, $\hat{\mathbf{h}}$ 为经过层归一化后得到的特征向量,交互权重 α_{ij}^k 由计算得到:

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(a[\mathbf{W}\hat{\mathbf{h}}_i \parallel \mathbf{W}\hat{\mathbf{h}}_j]))}{\sum_{k \in \mathcal{N}_i} \exp(\text{LeakyReLU}(a[\mathbf{W}\hat{\mathbf{h}}_i \parallel \mathbf{W}\hat{\mathbf{h}}_k]))}$$

经过多头注意力机制处理,轨迹序列中各个时间步和空间位置之间的关联程度能够在整个轨迹序列中捕捉全局的空间和时间依赖关系,提高了模型对序列数据的理解能力。

2.3 交互作用力推理模块

在 2.2 节中,我们通过空时特征提取模块有效捕获了轨迹数据的空间和时间特征,为多目标之间交互关系的建模奠定了基础。然而,仅提取特征还不足以全面描述目标间的动态交互对轨迹预测的影响。因此,进一步设计了交互作用力推理模块,如图 2 所示。交互作用力推理模块接收空时特征提取模块输出的节点特征,通过图神经网络和多层感知机 MLP^[28]推断目标之间的交互强度和作用力,输出结果作为拉格朗日模型中的

广义力项参与后续计算。这一模块能够动态地刻画目标之间的相互影响,为分析力学建模提供关键输入,同时提升了模型在复杂交互场景中的表现能力。

在交互作用力推理模块中,本文通过图神经网络推断多目标间的交互强度。类似于 Ha 等^[29]在异质交互强度学习中的方法,本模块综合考虑了目标间的相对距离和方向,以推测相对交互力的大小和方向。为了衡量场景中多目标之间的交互强度,在得到空时特征提取模块编码的轨迹特征后,将模型的输入和提取到的特征通过交互作用力推理模块计算多目标之间的交互作用力。交互作用力的计算公式为

$$\mathbf{e}_{ij} = \phi(\text{PE}(\|\mathbf{X}_i^{r-1} - \mathbf{X}_j^{r-1}\|_2; |\theta_{ij}|)) \quad (10)$$

$$\mathbf{F}_i^r = \sum_{j \neq i} \varphi_1(\tilde{\mathbf{h}}_i^{r-1}, \tilde{\mathbf{h}}_j^{r-1}, \mathbf{e}_{ij}) \quad (11)$$

$$\tilde{\mathbf{h}}_i^r = \varphi_2(\tilde{\mathbf{h}}_i^{r-1}, \mathbf{F}_i^r) \quad (12)$$

式中: ϕ 为 MLP,通过将 MLP 的输入进行 PE(positional encoding)位置编码,得到 \mathbf{e}_{ij} 。 φ_1 和 φ_2 为 GRU 网络进行端到端的训练, $\tilde{\mathbf{h}}_i^{r-1}$ 和 $\tilde{\mathbf{h}}_j^{r-1}$ 分别是目标 i 和 j 在前一时刻经过处理的隐特征,将隐特征与式(10)得到的交互度量值 \mathbf{e}_{ij} 共同作为 GRU 网络的输入,得到 \mathbf{F}_i^r (\mathbf{F}_i^r 为第 r 层网络计算的节点 i 受到的交互作用力),并在式(12)和 $\tilde{\mathbf{h}}_i^{r-1}$ 共同作为 GRU 网络 φ_2 的输入,得到目标 i 在 r 时刻新的隐特征,用于下一时刻的计算迭代, $\|\cdot\|_2$ 为 L2 范数(表示节点间相对距离), $[\cdot; \cdot]$ 为张量拼接。当 $r=1$ 时, $\tilde{\mathbf{h}}_i^0$ 为式(9)计算出的特征向量。 $|\theta_{ij}| = |\theta_i - \theta_j|$ 为节点 i 相对节点 j 运动方向的角度差,本文不使用数据集中角度坐标作为模型输入,每个节点的角度由不同时刻位置坐标计算得来。

与同样基于图网络计算作用力的 GMN(graph mechanics network)模型^[30]不同,式(11)和(12)中节点嵌入特征向量充分考虑了数据中空间和时间特征,此外本文模型不仅仅考虑了节点间相对位置关系,式(10)还考虑相对运动方向会影响节点间交互作用强度,并加入了 PE 位置编码,通过位置编码帮助模型更好理解序列中各个点的顺序。 \mathbf{e}_{ij} 通过 MLP 学习节点之间空间上的约束和现实的交互情况,例如,对于行人轨迹预测,即使两个行人距离很近,通过计算两个行人运动方向的角度差,如果他们的行进方向不同,他们之间的交互可能相对较小。对于人体运动预测,当一个关节绕着另一个关节摆动时,关节之间的运动角度有一定的差值,但关节之间的距离始终为固定值,则会降低相连接关节之间交互关系权重。

e_{ij} 通过位置编码确保模型能够理解数据的时间结构, 不同时间节点之间距离和方向的变化会改变交互权重。

2.4 基于分析力学的轨迹预测模块

图 2 中右侧的分析力学模块是整个模型的核心计算单元, 它结合目标初始状态 (位置与速度) 以及由前述模块推理出的交互力, 计算系统的广义加速度, 并更新预测轨迹。模块内部基于拉格朗日方程形式进行端到端优化学习, 输出符合物理约束的预测结果, 完成从时空特征提取到未来轨迹生成的闭环流程。通过式 (11) 得到交互作用力后, 由式 (7) 分析力学计算加速度 $\ddot{\mathbf{q}}_i^r$, 然后对节点状态更新, 计算公式为

$$\begin{aligned} \ddot{\mathbf{q}}_i^r &= \mathbf{M}^{-1} \left(\mathbf{F}_i - \frac{\partial V_{\mathbf{g}}(\mathbf{X}_i^{r-1})}{\partial \mathbf{X}_i^{r-1}} - \frac{d\mathbf{M}}{dt} \mathbf{v}_i^{r-1} + \right. \\ &\quad \left. \frac{1}{2} \frac{\partial}{\partial \mathbf{X}_i^{r-1}} \left((\mathbf{v}_i^{r-1})^T \mathbf{M} \mathbf{v}_i^{r-1} \right) \right) \\ \mathbf{v}_i^r &= \mathbf{v}_i^{r-1} + \ddot{\mathbf{q}}_i^r \Delta t \\ \mathbf{X}_i^r &= \mathbf{X}_i^{r-1} + \varphi_3(\mathbf{v}_i^r) \end{aligned}$$

式中: $\mathbf{F}_i = \psi(\mathbf{X}_i^{r-1}, \mathbf{v}_i^{r-1}; \gamma)$, φ_3 为 MLP。当 $r=1$ 时, \mathbf{X}_i^0 为输入轨迹数据, $\mathbf{v}_i^0 = \dot{\mathbf{X}}_i^0$ 。通过交互作用力模块和分析力学模块交替更新 r 次后, 输出最终预测结果 $\hat{\mathbf{Y}}_i = \mathbf{X}_i^r$ 。本文通过神经网络对拉格朗日方程参数化, 在笛卡尔坐标系下隐式编码了约束条件, 相比使用极坐标系、球坐标系等表示轨迹数据, 实现了更自然地描述现实物理系统。相比使用牛顿第二定律更新系统状态, 使用拉格朗日方程能更好地扩展到受约束系统的运动预测, 提高了模型的可解释性。

2.5 损失函数

对于人体运动预测, 模型轨迹预测的结果由 $\hat{\mathbf{Y}}$ 表示, $\mathbf{X}_i^t = (x_i^t, y_i^t, z_i^t)$, $\hat{\mathbf{Y}}_i^t = (\hat{x}_i^t, \hat{y}_i^t, \hat{z}_i^t)$, 其中 (x_i^t, y_i^t, z_i^t) 是第 i 个关节在 t 时刻的三维坐标。

在得到网络预测结果后, 拉格朗日力学神经网络优化为

$$\alpha^*, \beta^*, \gamma^*, \tau^* = \operatorname{argmin}_{\alpha, \beta, \gamma, \tau} \| Y - L(\mathbf{q}, \dot{\mathbf{q}}; \alpha, \beta, \gamma, \tau) \|_2$$

式中: α 、 β 和 γ 为拉格朗日力学神经网络参数, τ 为交互作用力推理模块的网络参数, Y 为真实未来轨迹, $L(\mathbf{q}, \dot{\mathbf{q}}; \alpha, \beta, \gamma, \tau)$ 为基于分析力学深度神经网络的预测结果, $\| \cdot \|_2$ 为 L2 距离, 公式通过优化参数 α 、 β 、 γ 、 τ , 最小化 $L(\mathbf{q}, \dot{\mathbf{q}}; \alpha, \beta, \gamma, \tau)$ 与真实轨迹 Y 之间的差距。本文模型损失函数由 3 部分组成, 计算公式为

$$L_{\text{loss}} = \lambda_1 \cdot L_d + \lambda_2 \cdot L_c + \lambda_3 \cdot L_r$$

$$L_d = \frac{1}{N \cdot T_f} \sum_{i \in N} \sum_{t=T_p+\Delta t}^{T_f} \| Y_i^t - \hat{Y}_i^t \|_2 \quad (13)$$

$$L_c = \sum_{t=T_p+\Delta t}^{T_f} \max_{\forall j, k \in N, j \neq k} \left(\frac{1}{\| \hat{d}_{jk}^t \|_2 + 1} \right) \quad (14)$$

$$L_r = \frac{1}{N_r \cdot T_f} \sum_{j, k \in N_r} \sum_{t=T_p+\Delta t}^{T_f} (\| L_{jk}^t \|_2 - \| \hat{L}_{jk}^t \|_2) \quad (15)$$

式中: λ_1 、 λ_2 和 λ_3 为超参数, N 为场景中目标数量, T_p 和 T_f 分别是输入轨迹序列帧数和预测帧数, Δt 为轨迹序列的采样时间间隔。式 (13) 是预测未来轨迹和真实未来轨迹之间的误差值, 用预测轨迹结果 \hat{Y} 和真实轨迹 Y 之间的 L2 损失定义。式 (14) 是考虑现实情况, 避免生成的未来轨迹中目标发生碰撞, 用节点间相对距离 \hat{d}_{ij} 定义。 \hat{d}_{ij} 为某一时刻目标 i 和 j 的欧氏距离。存在一个最小距离阈值 d_{safe} , 如果 $\hat{d}_{ij}(t)$ 小于 d_{safe} , 则认为在 t 时刻发生了碰撞。式 (15) 是考虑约束条件下, 预测结果中关节 j 和关节 k 之间的长度 \hat{L}_{jk} 与真实关节长度 L_{jk} 之间的误差值, N_r 为有连接的关节节点, 在行人轨迹预测中不包含此项。

3 实验分析

实验使用 ADAMW (adaptive moment estimation with weight decay) 优化器训练模型, 设置初始学习率为 0.002, 每迭代 10 次学习率衰减 10%。实验训练 120 轮后进行测试, 设置的批大小为 256。

3.1 数据集与评价指标

为了验证本文所提方法的有效性, 本文在两个公开轨迹数据集上进行实验, 包括: Spring-balls^[27]、Human 3.6 Million^[31]。

本文使用均方误差 (mean squared error, MSE) 作为性能指标来评价 Spring-balls 数据集的多目标交互运动预测性能。其计算公式为

$$E_{\text{MS}} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

本文使用平均关节位置误差 (mean per joint position error, MPJPE) 作为性能指标来评价模型人体运动预测性能^[31]。对于所有 (K 个) 关节, MPJPE 是每个关节位置误差的平均值, 计算公式为

$$E_{\text{MPJP}} = \frac{1}{K} \sum_{i=1}^K \sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + (z_i - \hat{z}_i)^2} \quad (16)$$

3.2 实验结果与分析

3.2.1 物理场景预测结果与分析

在 Spring-balls 数据集上, 根据现有的工作, 在该数据集中使用 5 个球-弹簧和 10 个球-弹簧

模型进行实验,采用前 50 帧历史轨迹数据作为输入,后 50 帧的未来轨迹数据作为真值,使用 MSE 作为评价指标,并与现有 3 种主流模型的性能进行对比,主流模型有: GATv2 模型^[32]、NRI 模型^[27]和 RAIN 模型^[29]。表 1 给出了本文模型与上述模型在数据集 Spring-balls 上定量比较,给出了 5 个球-弹簧系统和 10 个球-弹簧系统的 MSE 值。从表 1 的预测结果可以观察到,本文模型在两种场景上都取得了更好的表现,预测精度好于同样采用了图注意力机制的 GATv2 模型。在 5 球场景 10 帧短期预测上与 RAIN 模型相比具有 14.29% 的精度提升,在 10 球场景 10 帧短期预测上与 NRI 模型相比具有 1.64% 的精度提升,在预测难度更高的 50 帧长期预测上与 RAIN 模型相比两种场景分别具有 6.25% 和 4.81% 的精度提升。为了进一步验证模型性能,图 4 给出了本文模型在数据集 Spring-balls 上的可视化结果,不同颜色表示不同的球,图中半透明圆形表示输入轨迹的前 50 帧,不透明圆形表示真实未来轨迹的 50 帧,叉形表示预测未来轨迹的 50 帧。从可视化结果中可以观察到,本文模型预测精度好于 RAIN 模型以及其他主流模型,证明了本文模型推断模拟物理系统中交互强度的有效性。

表 1 不同模型在数据集 Spring-balls 上预测的 MSE(10⁻²)
Table 1 MSE predictions of different models on the Spring-balls dataset (10⁻²)

模型	5球			10球		
	10帧	30帧	50帧	10帧	30帧	50帧
GATv2 ^[32]	2.30	20.81	42.52	13.09	26.34	41.92
NRI ^[23]	1.22	13.95	35.73	0.61	8.66	24.40
RAIN ^[29]	0.84	7.14	21.93	0.69	8.92	23.51
本文模型	0.72	6.38	20.56	0.60	8.13	22.38

注: 加粗表示本列最优结果。

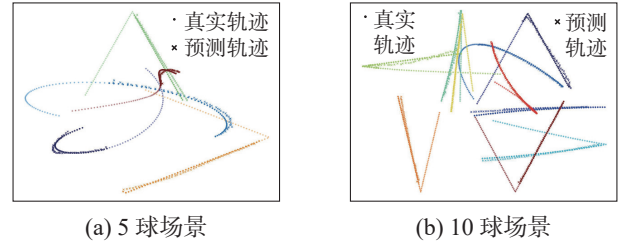


图 4 Spring-balls 数据集上 5 球和 10 球可视化结果
Fig. 4 Visualization results on Spring-balls dataset with 5 balls and 10 balls

3.2.2 人体运动预测结果与分析

本文在数据集 Human 3.6 Million 上评估所提出模型在人体运动预测中的性能,并与现有模型的性能进行比较。根据现有的工作,本文在短期预测 (80~400 ms) 和长期预测 (400~1 000 ms) 场景上进行对比实验,使用式 (16) 计算的 MPJPE 作为评价指标 (单位 mm),该指标值越低代表模型性能越好。在训练阶段,将所有动作的数据一起输入到网络中,仅在测试阶段按动作类别进行评估。

在数据集 Human 3.6 Million 上,本文与现有 4 种基于 GCN 的主流模型性能进行对比,主流模型有: LTD 模型^[10]、MSRGCN 模型^[11]、SPGSN 模型^[12]和 PGBIG 模型^[33]。表 2 和表 3 分别给出了本文模型与上述模型在数据集 Human 3.6 Million 上短期和长期预测的定量比较,展示了打电话、进食、遛狗 3 种动作在不同预测时间下的 MPJPE。由于不同动作的复杂性不同,模型在对各种动作的预测精度上存在差异。从表 2 和表 3 的预测结果可以观察到,在几乎所有动作上,本文模型都取得了更好或者接近的表现,而且全部动作 MPJPE 的平均值优于同样基于 GCN 的 4 种主流模型。可以认为,本文模型在处理运动模式较规律的动作和随机性较强的动作上均可表现出较好的长期预测性能和较强的鲁棒性,说明本文结合 GCN 和分析力学的预测模型的有效性。

表 2 不同模型在数据集 Human 3.6 Million 上短期预测 MPJPE
Table 2 MPJPE for short-term predictions of different models on Human 3.6 Million dataset mm

模型	打电话时间/ms				进食时间/ms				遛狗时间/ms			
	80	160	320	400	80	160	320	400	80	160	320	400
LTD ^[10]	10.2	21.0	42.5	52.3	8.4	16.9	33.2	40.7	23.4	46.2	83.5	96.0
MSRGCN ^[11]	10.1	20.7	41.5	51.3	8.4	17.1	33.0	40.7	20.7	42.9	80.4	93.3
SPGSN ^[12]	8.7	18.3	38.7	48.5	7.1	14.9	30.5	37.9	17.8	37.2	71.7	84.9
PGBIG ^[33]	8.3	18.3	38.7	48.4	7.0	15.1	30.6	38.1	18.8	39.3	73.7	86.4
本文模型	8.2	17.9	38.9	47.6	6.8	14.5	29.8	36.9	17.6	37.1	71.5	83.7

注: 加粗表示本列最优结果。

表 3 不同模型在数据集 Human 3.6 Million 上长期预测 MPJPE

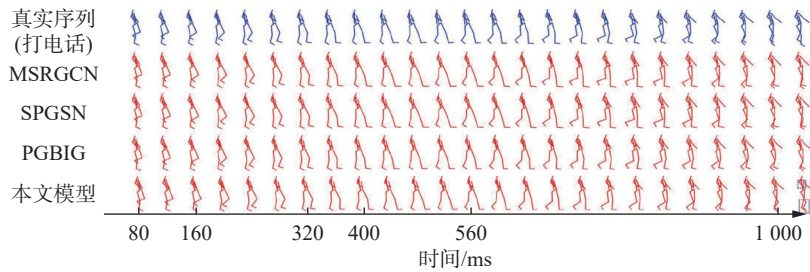
Table 3 MPJPE for long-term predictions of different models on Human 3.6 Million dataset mm

模型	打电话时间/ms		进食时间/ms		遛狗时间/ms	
	560	1000	560	1000	560	1000
LTD ^[10]	69.2	103.1	53.4	77.8	111.9	148.9
MSRGCN ^[11]	68.3	104.4	52.5	77.1	111.9	148.2
SPGSN ^[12]	66.7	102.5	49.8	73.4	102.4	138.0
PGBIG ^[33]	69.5	102.7	51.1	76.0	104.7	139.8
本文模型	66.5	100.6	49.2	73.2	99.2	137.8

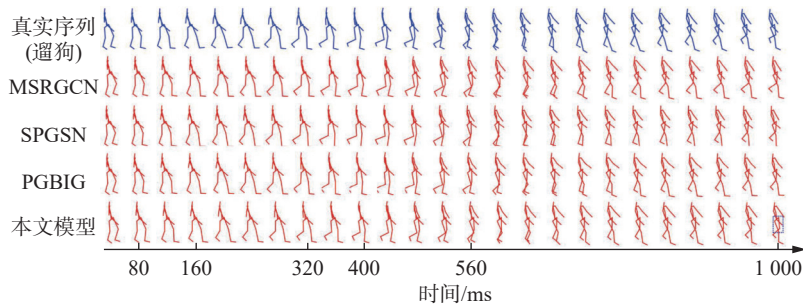
注: 加粗表示本列最优结果。

为了进一步验证模型性能, 图 5 给出了打电

话和遛狗动作的可视化预测结果, 其中打电话代表运动模式较规律的动作类别, 遛狗代表随机性较强的动作类别。每张子图第一行是真实动作序列, 其余每行是不同模型的预测结果。从可视化结果中可以观察到, 本文模型预测的未来姿态更接近真实序列, 特别是在胳膊和腿上取得了更好的效果。例如在打电话动作中, 本文模型在第 1000 ms 的预测结果胳膊和腿部动作 (图中虚线框表示) 与真实序列更相似。从长期预测 (560~1000 ms) 结果可以观察到, 本文模型具有更好的表现, 更能捕捉到胳膊和腿部动作的大幅度变化, 可以说本文分析力学模块有效地从关节序列中捕捉到动力学特征。



(a) 打电话动作可视化结果



(b) 遛狗动作可视化结果

图 5 不同模型在数据集 Human 3.6 Million 上可视化结果

Fig. 5 Visualization results of different models on Human 3.6 Million dataset

3.3 消融实验

通过消融实验探究本文模型重要模块的作用, 表 4 给出了本文模型在数据集 Human 3.6 Million 上预测时间 80~1000 ms 的 MPJPE。

表 4 在数据集 Human 3.6 Million 上消融实验结果

Table 4 Ablation study results on Human 3.6 Million dataset mm

ST	AM	预测时间/ms					
		80	160	320	400	560	1000
		20.1	36.4	66.5	78.1	95.8	135.8
√		16.4	29.1	57.2	69.1	88.9	124.6
	√	12.9	26.0	51.7	62.2	80.0	115.2
√	√	9.9	21.8	46.4	57.8	75.6	108.0

3.3.1 基于图神经网络和卷积神经网络的空时特征提取模块

为验证所提出的基于 GNN 和 CNN 的空时特征提取模块 (spatial and temporal feature extraction module, ST) 在提取序列特征方面的作用, 在消融实验中将该模块替换为 MLP。如表 4 中第 1 行和第 2 行结果所示, 使用该模块可大幅提高预测精度。

3.3.2 基于分析力学的预测模块

为验证所提出的基于分析力学的预测模块 (prediction module based on analytical mechanics, AM) 的作用, 在消融实验中将该模块替换为牛顿第二定律, 即得到交互作用力后计算加速度, 然后直接更新速度和位置。如表 4 中第 1 行和第 3 行结果所示, 使用该模块可大幅提高预测精度。

通过消融实验,得到各模块的作用:1)基于GNN和CNN的空时特征提取模块有助于编码数据中空间和时间特征;2)基于分析力学的预测模块可以捕获动力学特征和空间约束关系。如表4中第4行结果所示,结合物理知识和数据驱动的方式有利于生成准确的预测。

4 结束语

本文提出了一种基于分析力学和图神经网络的轨迹预测模型,该模型通过图神经网络、一维卷积网络和图注意力机制学习多目标之间的交互关系,通过参数化分析力学从数据中推断系统的动力学特征,实现了将物理方程与数据驱动方法的更好结合,提高了模型的可解释性。本文在一个模拟物理系统数据集、一个人体运动数据集上进行实验,取得了比主流模型更好或者接近的效果。通过实验发现,本文模型能够从数据中推断出物理系统的动力学信息,在人体运动数据集上也表现出更好的综合效果和长期预测性能,验证了其受几何约束系统的有效性。

参考文献:

- [1] BARTH A, FRANKE U. Where will the oncoming vehicle be the next second? [C]//2008 IEEE Intelligent Vehicles Symposium. Eindhoven: IEEE, 2008: 1068–1073.
- [2] KOOIJ J F P, SCHNEIDER N, GAVRILA D M. Analysis of pedestrian dynamics from a vehicle perspective[C]//2014 IEEE Intelligent Vehicles Symposium Proceedings. Dearborn: IEEE, 2014: 1445–1450.
- [3] KARAMOUZAS I, SKINNER B, GUY S J. Universal power law governing pedestrian interactions[J]. *Physical review letters*, 2014, 113(23): 238701.
- [4] BATZ T, WATSON K, BEYERER J. Recognition of dangerous situations within a cooperative group of vehicles[C]//2009 IEEE Intelligent Vehicles Symposium. Xi'an: IEEE, 2009: 907–912.
- [5] PANG Bo, ZHAO Tianyang, XIE Xu, et al. Trajectory prediction with latent belief energy-based model[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 11814–11824.
- [6] LECUN Y, CHOPRA S, HADSELL R, et al. Energy-based models[M]//Predicting Structured Data. [S. l.]: The MIT Press, 2007: 191–246.
- [7] 姜立超, 尚晓兵, 金豹, 等. 基于遗传算法-v支持向量回归的船舶轨迹预测[J]. 哈尔滨工程大学学报, 2024, 45(10): 2001–2006.
JIAN Lichao, SHANG Xiaobing, JIN Bao, et al. Ship trajectory prediction based on genetic algorithm-v support vector regression[J]. *Journal of Harbin Engineering University*, 2024, 45(10): 2001–2006.
- [8] LIU Xiaoli, YIN Jianqin, LIU Jin, et al. TrajectoryCNN: a new spatio-temporal feature learning network for human motion prediction[J]. *IEEE transactions on circuits and systems for video technology*, 2021, 31(6): 2133–2146.
- [9] AKSAN E, KAUFMANN M, CAO Peng, et al. A spatio-temporal transformer for 3D human motion prediction[C]//2021 International Conference on 3D Vision. London: IEEE, 2021: 565–574.
- [10] MAO Wei, LIU Miaomiao, SALZMANN M, et al. Learning trajectory dependencies for human motion prediction[C]//2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 9489–9497.
- [11] DANG Lingwei, NIE Yongwei, LONG Chengjiang, et al. MSR-GCN: multi-scale residual graph convolution networks for human motion prediction[C]//2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021: 11447–11456.
- [12] LI Maosen, CHEN Siheng, ZHANG Zijing, et al. Skeleton-parted graph scattering networks for 3D human motion prediction[C]//Computer Vision-ECCV 2022. Cham: Springer Nature Switzerland, 2022: 18–36.
- [13] ZHOU Jincao, BAI Xin, FU Weiping, et al. Pedestrian intention estimation and trajectory prediction based on data and knowledge-driven method[J]. *IET intelligent transport systems*, 2024, 18(2): 315–331.
- [14] 窦慧, 张凌茗, 韩峰, 等. 卷积神经网络的可解释性研究综述[J]. 软件学报, 2024, 35(1): 159–184.
DOU Hui, ZHANG Lingming, HAN Feng, et al. Survey on convolutional neural network interpretability[J]. *Journal of software*, 2024, 35(1): 159–184.
- [15] 杨朋波, 桑基韬, 张彪, 等. 面向图像分类的深度模型可解释性研究综述[J]. 软件学报, 2023, 34(1): 230–254.
YANG Pengbo, SANG Jitao, ZHANG Biao, et al. Survey on interpretability of deep models for image classification[J]. *Journal of software*, 2023, 34(1): 230–254.
- [16] 曹健, 陈怡梅, 李海生, 等. 基于图神经网络的行人轨迹预测研究综述[J]. 计算机工程与科学, 2023, 45(6): 1040–1053.
CAO Jian, CHEN Yimei, LI Haisheng, et al. A survey of pedestrian trajectory prediction based on graph neural network[J]. *Computer engineering & science*, 2023, 45(6): 1040–1053.
- [17] 冯唐思捷, 梁伟. 基于物理信息神经网络的薄壁结构屈曲分析[J]. 力学学报, 2023, 55(11): 2539–2553.
FENG Tangsijie, LIANG Wei. The buckling analysis of thin-walled structures based on physicsinformed neural networks[J]. *Chinese journal of theoretical and applied mechanics*, 2023, 55(11): 2539–2553.
- [18] 刘才山, 郭永新. 多体系统动力学与分析动力学专题序[J]. 力学学报, 2020, 52(4): 916.

- LIU Caishan, GUO Yongxin. Special topics on dynamics of multi-body systems and analytical dynamics[J]. *Chinese journal of theoretical and applied mechanics*, 2020, 52(4): 916.
- [19] LANDAU L D, LIFSHITZ E M. Course of theoretical physics[M]. British: Pergamon Press, 1980.
- [20] BUTCHER J C. A history of Runge-Kutta methods[J]. *Applied numerical mathematics*, 1996, 20(3): 247–260.
- [21] 张心祎, 谭耀, 邢向磊. 基于物理先验的深度特征融合水下图像复原[J]. *智能系统学报*, 2023, 18(6): 1185–1196. ZHANG Xinyi, TAN Yao, XING Xianglei. Deep feature fusion for underwater-image restoration based on physical priors[J]. *CAAI transactions on intelligent systems*, 2023, 18(6): 1185–1196.
- [22] LUTTER M. Deep lagrangian networks: Using physics as model prior for deep learning[EB/OL]. (2019–07–10) [2024–09–19]. <https://arxiv.org/abs/1907.04490>.
- [23] 刘建敏, 林晖, 汪晓丁. 基于图注意力机制的无地图场景轨迹预测方法[J]. *计算机工程*, 2024, 50(7): 144–153. LIU Jianmin, LIN Hui, WANG Xiaoding. Graph attention mechanism-based method for trajectory prediction in map-free scenes[J]. *Computer engineering*, 2024, 50(7): 144–153.
- [24] 连静, 丁荣琪, 李琳辉, 等. 基于图模型和注意力机制的车辆轨迹预测方法[J]. *兵工学报*, 2023, 44(7): 2162–2170. LIAN Jing, DING Rongqi, LI Linhui, et al. Vehicle trajectory prediction method based on graph models and attention mechanism[J]. *Acta armamentarii*, 2023, 44(7): 2162–2170.
- [25] 郝伟琛, 谢心喻, 丁家琦, 等. 瞬态多相流场图神经网络时空预测方法研究[J]. *哈尔滨工程大学学报*, 2024, 45(9): 1761–1769. HAO Yichen, XIE Xinyu, DING Jiaqi, et al. Spatiotemporal prediction method for the transient multiphase flow field via graph neural network[J]. *Journal of Harbin Engineering University*, 2024, 45(9): 1761–1769.
- [26] 李浩森, 张含笑, 邢向磊. 联合局部多尺度和全局上下文特征的步态识别[J]. *智能系统学报*, 2024, 19(4): 853–862. LI Haomiao, ZHANG Hanxiao, XING Xianglei. Gait recognition with united local multiscale and global context features[J]. *CAAI transactions on intelligent systems*, 2024, 19(4): 853–862.
- [27] KIPF T, FETAYA E, WANG K C, et al. Neural relational inference for interacting systems[C]//International conference on machine learning. Stockholm: PMLR, 2018: 2688–2697.
- [28] RIEDMILLER M. Multi layer perceptron[EB/OL]. [2024–10–11]. http://machine-learning-lab.com/_media/documents/teaching/ss12/ml/05_mlps.printer.pdf.
- [29] HA S, JEONG H. Learning heterogeneous interaction strengths by trajectory prediction with graph neural network[EB/OL]. (2022–01–31)[2024–10–15]. <https://arxiv.org/abs/2105.14491>.
- [30] HUANG W, HAN J, RONG Y, et al. Equivariant graph mechanics networks with constraints[EB/OL]. (2022–03–12) [2024–12–15]. <https://arxiv.org/abs/2208.13179>.
- [31] IONESCU C, PAPAVALAS D, OLARU V, et al. Human3.6M: large scale datasets and predictive methods for 3D human sensing in natural environments[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2014, 36(7): 1325–1339.
- [32] BRODY S, ALON U, YAHAV E. How attentive are graph attention networks? [EB/OL]. (2021–05–30)[2024–12–15]. <https://arxiv.org/abs/2105.14491>.
- [33] MA Tiezheng, NIE Yongwei, LONG Chengjiang, et al. Progressively generating better initial guesses towards next stages for high-quality human motion prediction[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022: 6427–6436.

作者简介:



李明哈, 主要研究方向为轨迹预测与图像识别。E-mail: liminghan721@gmail.com。



肖阳, 硕士研究生, 主要研究方向为轨迹预测、基于物理模型的深度学习。E-mail: xiaoy@hrbeu.edu.cn。



邢向磊, 教授, 博士生导师, 主要研究方向为模式识别与计算机视觉。获得黑龙江省高校科学技术奖(自然科学类)一等奖, 获得《智能系统学报》优秀论文奖。发表学术论文 60 余篇。E-mail: xingxl@hrbeu.edu.cn。