arXiv:2303.04320v2 [cs.RO] 6 Aug 2023

SG-LSTM：用于密集人群中机器人导航的社交群体长短期记忆网络

Rashmi Bhaskara、Maurice Chiu 和 Aniket Bera，普渡大学计算机科学系，美国。包含代码和数据集的补充版本见https://ideas.cs.purdue.edu/research/sg-lstm

摘要—随着个人机器人日益普及且价格亲民，其应用范围已从大型企业仓库和工厂扩展到多种管控程度较低的环境，在这些环境中，机器人会与更多人群互动。在此类情境下，确保安全与效率、减轻对人类可能造成的不良心理影响，以及遵守不成文的社会规范，至关重要。本研究旨在应对这些挑战，开发一种先进模型，以预测拥挤环境中的行人运动及交互。为此，我们提出一种名为社交群体长短期记忆网络（SG-LSTM）的新方法，可有效捕捉密集环境中人类群体行为及交互的复杂性。通过将社会感知融入长短期记忆网络架构，我们的模型显著提升了轨迹预测效果。SG-LSTM模型的实现使导航算法能够更快、更准确地计算无碰撞路径，尤其适用于复杂且拥挤的场景。为推动社交导航研究进一步发展，我们贡献了一个包含已标注行人群体的大型视频数据集，并向更广泛的研究群体开放。为全面评估该方法的性能，我们在包括ETH、Hotel和MOT15在内的多个数据集上进行了大量实验。我们比较了多种预测方法，如LIN、长短期记忆网络、O-LSTM和S-LSTM，并严格评估其运行时性能。

I. 引言

社交导航是社交机器人学中的关键能力，它使机器人能够以符合社会规范且有效的方式在人的环境中移动并与人互动。社交导航指机器人在空间中移动时，能够考虑社会规范和期望。这包括绕行人移动、避开障碍物，以及遵循社会习俗，例如排队或为他人让路。将社交导航融入设计后，社交机器人便能更流畅、自然地在人类环境中移动，从而提升其社会接受度和有效性。这对于设计用于公共空间、与人互动的机器人尤为重要，例如在机场、购物中心或医院提供引导或协助的机器人。社交导航还能让机器人理解并响应社交线索和惯例，从而更有效地与人互动。例如，在繁忙的医院走廊中导航的机器人，必须识别并避免与迎面而来的行人相撞，还要懂得何时应让急赴急诊的医院工作人员优先通行。现有的社会规范导航算法旨在帮助服务机器人安全地穿行于复杂环境，同时尊重社会规范并避免与行人碰撞。然而，部分算法将行人视为彼此独立的障碍物，因此不适用于拥挤区域 [33], [35],[36] ，因为在这些区域中，行人的社会动态发挥着更重要的作用，并会影响行人动力学。图2展示了任何导航算法都会面临的问题。具备社会意识的导航算法会选择绕开群组，而不是将其拆散；而将行人视为彼此独立障碍物的常规导航算法，则可能导致机器人停滞不前，或决定穿过群组。

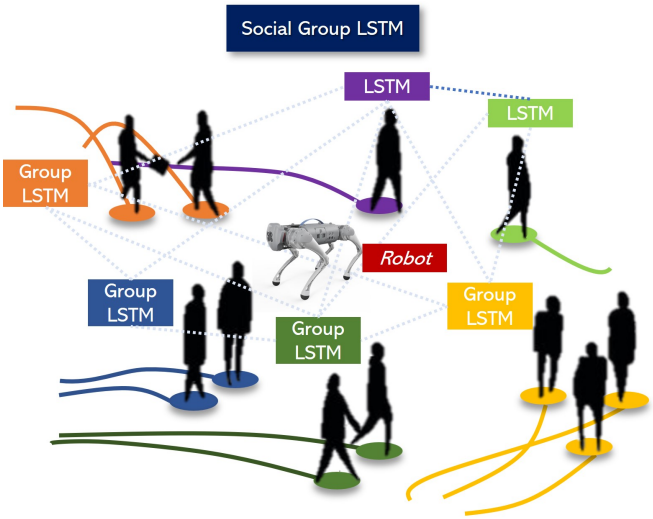


图 1：***SG-LSTM***：我们提出一种符合社会规范的群组长短期记忆网络（*SG-LSTM*）模型，将群体动力学融入行人预测。我们的分层模型在群组和个体层面解耦行人，并能够在这两个层面上建模个体动态。

近年来，融入群体凝聚力的算法在提升智能体于高密度人群环境中的导航能力方面展现出良好效果。这些算法依赖于计算群体凝聚力，即分析群组中个体的速度、方向和邻近程度等因素，以预测他们彼此靠近同行的程度。然而，这些方法存在一些局限性，包括需要进行耗费资源的在线计算，以及在特定条件下准确率降低 [1]。例如，人群密度较高时，计算多种特征的速度会显著变慢，从而对导航性能产生负面影响。此外，当某些特征无法检测时，这些算法的效果也会受到限制，导致群组定义不明确。为应对这些挑战，本文提出一种创新方法，利用长短期记忆网络（LSTM）架构识别 crowd videos 中的感知群体。这种符合社会规范的轨迹预测算法能够改善导航，相关示例见图1。我们的**主要贡献**如下：

- 社会群体长短期记忆网络（SG-LSTM）模型将群体动力学融入行人预测。

- 我们的分层模型分别在群组和个体层面对行人进行建模，从而有效捕捉个体动态。
- 我们的群体优化方法将计算时间显著缩短了 50% 以上。
- 此外，我们提供了一个大型数据集，其中包含超过 30,000+ 个人工标注帧，并附有详细的群体信息。

在本文其余部分中，我们将在第 II 节概述相关工作。第 III 节将详细介绍系统的整体架构和方法论。接着，第 IV 节将介绍我们的数据集。在第 V 节中，我们将讨论评估指标、结果和分析。最后，第 VI 节将探讨未来方向和局限性，并总结我们的方法。

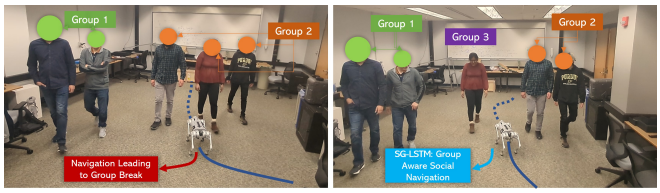


图 2：将彼此邻近程度较高的行人识别为群组。机器人将规划路径，避免将行人群体拆散。

II. 相关工作

本节回顾了现有的群组识别和行人轨迹预测方法。

A. 人类群体学习

研究广泛关注于利用视频输入识别群体特征和发现群体。值得注意的是， [2]采用 YOLOv2 进行目标检测，将检测结果转换为特征向量，并对空间位置邻近的个体进行聚类。对象最初被视为彼此独立的簇，并通过反复计算欧几里得距离和更新质心来形成群体。类似地， [4] 根据行人的邻近程度、速度和运动方向定义社会群体。此外，Sathyamoorthy 等人 [1] 提出了行人群体表征的数学模型，将步行速度、交互、群体规模和邻近程度视为定义因素。

除了人群中的位置和距离等物理属性外，群体识别还可以考虑集体行为 [3]。Shao 等人利用行人轨迹片段的集体转移（CT）来形成簇，并根据与锚定轨迹片段之间的高速度相关性定义群体。该方法为理解人群动态和社会群体划分提供了有益见解。这些研究凸显了多种策略，通过利用目标检测、特征向量化、聚类和集体行为分析，有效地从视频数据中识别并分类人群中的群体。

B. 社交导航

本综述探讨了多种研究成果，这些成果致力于提升复杂场景中高效、安全导航的不同方面。Cheung 等人 [46] 提出了一种基于驾驶员行为分类的创新车辆导航方法，而 Chandra 等人 [47] 则推出了 CMetric，一种利用中心性函数的驾驶员行为指标。Randhavane 等人 [48]提出了行人主导性建模，用于具备社会意识的机器人导航，并考虑了行人在拥挤环境中的主导性。Bera 等人[49]开发了 Adapt，这是一种实时自适应行人跟踪系统；Bera 和 Manocha [50] 则提出了一种利用互惠速度障碍的实时多层级人群跟踪方法。此外，Bera 等人 [51]推出了 GLMP，这是一种同时结合全局和局部运动模式的实时行人路径预测模型。Bera 等人[52],研究了拥挤场景中的实时异常检测；Bera 等人 [53]则提出了用于数据驱动的人群仿真和内容生成的在线参数学习方法。Murino 等人 [54] 的研究探讨了计算机视觉领域中的群体与人群行为分析技术。此外，Bera 等人 [55]深入探讨了“社会隐形机器人”这一概念，旨在使机器人能够在社会环境中有效导航和互动。另外，Bera 等人 [56] 提出了一个交互式人群行为学习系统，可支持监控和训练任务。Chandra 等人 [57] 提出了 GraphRQI，这是一种利用图谱对驾驶员行为进行分类的方法；Cheung 等人 [58] 则通过轨迹特征识别驾驶员行为，以提升车辆导航和安全性。这些研究共同推动了各种场景下导航与行为分析的发展，为具备社会意识的机器人和人群管理系统的开发提供了宝贵见解。

计算机视觉领域为行人行为建模作出了重大贡献 [5]。Helbing 和 Molnar 提出的基础社会力模型 [5] 已应用于人群仿真 [6] 和异常行为检测 [9]。后续扩展包括对行人轨迹与群体划分进行联合建模 [11]，以及为多人跟踪建模社会行为和群体行为 [7]。值得注意的是，已有研究利用社会力进行运动预测来跟踪行人 [8]，并检测视频数据中的社交群体 [10]。此外，研究人员还探索了将基于图像的运动上下文用于多人跟踪[12]。近期研究探索了多种人类活动预测技术，包括使用 PORCA 进行轨迹预测 [41],、利用 SocioSense 实现具备社会意识的导航[40], ，以及使用 LSTM-CNN混合网络对智能体交互建模 [39]。Bera 等人 [38] 预测复杂环境中的行人路径。相关研究还涉及从视频流中进行早期活动识别 [31], 、从部分观测视频中识别人类活动 [27], ，以及从静态场景中预测动作 [28]。研究人员提出了数据驱动的活动预测算法 [29],, Vondrick等人 [30] 则提出了一个基于深度学习的框架，用于预测视觉表征。此外，研究还进一步探讨了静止人群中的行人行为，分析其关系 [17], 、预测显著性 [18], 并进行具备社会意识的预测 [19]。相关技术包括数据驱动分析 [20], 、高斯过程回归流 [21],和轨迹学习 [22]。预测模型涵盖行为识别 [23], 、意图学习 [24], 和动作预测 [25]。Kitani等人 [26] 根据视觉输入预测未来动作。

当前的社交导航方法往往忽视了场景中检测并维持社交群体凝聚力这一关键方面。由于忽略了这类群体的存在及其动态变化，这些方法可能无法达到最佳预测准确率，也无法考虑个体的社会行为。为了解决这一局限，我们提出了全新的 SG-LSTM 方法，为轨迹预测引入了一种关注群组的范式。我们的方法

通过在多个数据集上进行大量实验，展现了其有效性，证明该方法能够预测符合社会规范的轨迹，同时避免与检测到的群体发生碰撞。SG-LSTM 在轨迹预测中考虑群体动力学，代表了该领域的一项重大进展，有望在复杂的真实场景中实现更可靠、更符合人类需求的机器人导航。

III. 方法论

A. 概述

SG-LSTM 将视频中的原始 RGB 帧和深度帧作为输入。随后，我们的基于 CNN 的群体学习算法学习预测密集人群中的群组，为轨迹预测流程提供群组的空间坐标。其架构如图 3 所示，本文中使用的所有符号均汇总于表I。

表 I：SG-LSTM 中使用的符号及其定义

d	群体平均深度
$(x_b,\ y_b)$	边界框质心
FOV_{cam}	相机视场
w	图像帧宽度
ϕ	机器人的朝向
ϕ_g	群体相对于机器人的角位移
h_t^i	i^{th} 群组的长短期记忆网络隐藏层
e_t^i	对输入坐标进行向量嵌入，并 该 i群组的隐藏张量 th 在时间 t
$(x_g,\ y_g)_t^i$	群组相对于 i^{th} 的空间坐标 在时间 t 的机器人
$(x_r,\ y_r)_t$	机器人在时间 t 的位置
u_s	机器人的速度
u_ϕ	机器人的转向角

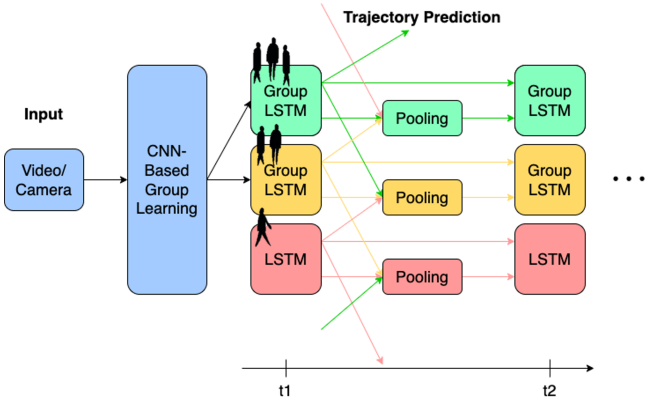


图 3：我们的社会规范群组长短期记忆网络模型概览。摄像机捕获的 RGB帧将经过群体检测，从而推断出空间坐标，并将其作为输入馈送至架构。

B. 轨迹预测流程

我们的轨迹预测算法以 Social-LSTM 模型为基础 [32]，该模型专为拥挤社交场景中的人类轨迹预测而设计。Social-LSTM 将社会交互和个体的过往轨迹纳入考量，以预测未来运动模式。尽管 Social-LSTM 是用于拥挤社交场景中人类轨迹预测的强大模型，但仍存在局限性。行人密度较高或更多智能体组成群组时，其准确率往往会下降。通过引入更优越的群体动力学模型，我们在两个关键方面提升了最先进水平：实现更准确的轨迹预测，并增强模型应对检测或传感器输出不准确的能力。在对群组进行建模时，我们将所有行人视为一个群组，因此即使难以跟踪群组中的部分智能体，预测仍可依靠群组轨迹。这种鲁棒性在行人检测算法难以准确提取空间坐标时尤为重要，尤其是当人们成群行走或场景中存在遮挡时。

扩展 Social-LSTM 或类似模型以处理包含大量个体的复杂大型场景颇具挑战，因为捕捉场景的社会背景需要大量计算资源和内存。由于模型使用长短期记忆网络来表示每个人员，包含多个个体的场景就需要相应数量的长短期记忆网络，从而增加计算开销。此外，在处理静止人员或作为群组共同移动的个体时，这些资源可能无法得到最优利用。因此，越来越需要更高效、可扩展的轨迹预测模型，能够处理此类场景，同时保持预测准确率和计算效率。

C. 问题表述

在我们的方法中，我们将每个检测到的群组在任意给定时刻 t的空间坐标定义为 $(x_g,\ y_g)_t^i$ 。对于不属于任何群组的行人，我们将其视为仅包含一个成员的独立群组。我们以 30 fps 的帧率收集连续五个时间步（从 t_1 到 t_5）的空间坐标，并将其输入 Social-Group-LSTM 模型，以预测随后五个时间步（从 t_6 到 t_{10}）的轨迹。这种时间步配置使我们能够有效计算并预测群组行人和个体行人的移动模式，从而提高轨迹预测准确率，并实现拥挤环境中的社会规范导航。

D. 群组定位

我们的方法是在标注数据集上训练一个基于卷积神经网络的群组学习模型。该模型基于单阶段目标学习算法，并利用空间金字塔池化（SPP）模块检测不同尺度和尺寸的物体。这一能力有助于识别形状和尺寸各异的群组。此外，我们的方法通过结合空间特征和注意力机制，有效处理遮挡情况，即群组中的物体彼此部分或完全遮挡，从而确保群组内目标检测的准确性。

训练后的模型会在拥挤场景中为群组输出边界框。这些群组边界框与数据集中已有的、属于任何群组的个体行人边界框相结合。随后，将这些边界框投影到各场景的深度图上，并利用边界框质心计算每个行人或群组的深度。通过使用像素深度值（d）、帧宽度以及以弧度表示的摄像机视场角（FOV_{cam}），我们计算所有行人群体和未分组行人的空间坐标。利用边界框质心（$x_b,\ y_b$）和摄像机视场角，可以确定各行人/群组相对于机器人的角位移，其中 $\phi_g = (x_b/w) \cdot FOV_{cam}$。行人/群组的空间坐标计算如下：$(x_g,\ y_g) = d \cdot [cos\phi_g, sin\phi_g]$。

E. 社会交互建模

在我们的方法中，我们采用了一个类似于 Social-LSTM 的社交池化层来分析场景中的群体交互。该池化层由一个用于 i^{th}轨迹的张量 H_t^i 定义，其形式如下：

$$H_t^i = \sum_{j \in \mathcal{N}i} 1m, n[(x_g, y_g)_t^j - (x_g, y_g)_t^i] h_{t-1}^j \qquad (1)$$

其中，$^j\ h_{t-1}$ 表示时间步 $t-1$ 时对应于 j^{th} 群组的隐藏状态，而$1_{m,\ n}$ 是一个指示函数，用于检查表示为集合 $\mathcal{N}_i$ 的 i^{th} 群组的邻近群体对应的坐标（x_g，y_g）是否位于网格的（m，n）单元格中。该层有效地捕捉了场景中所有群组之间的社会交互。

Social-Group LSTM 模型中的社会池化层在对场景中的行人交互进行建模时，将每个群组视为一个独立实体。这种有利的方法使模型更容易扩展到更密集的人群，因为与 Social-LSTM 不同，背景区域中的每个人员并不会都由一个长短期记忆网络表示。因此，这显著减少了对密集人群场景进行建模所需的参数数量。此外，通过将人员分组，该模型克服了计算背景区域中被遮挡行人的空间坐标时可能出现的不准确性，从而进一步提高了轨迹预测的整体准确率。

模型层：我们的模型层和权重与 Social-LSTM 模型类似。输入坐标和隐藏张量被嵌入到向量 e_t^i中，并作为具有权重 W 的 LSTM 单元隐藏状态（h_t^i）的输入，该 LSTM 单元对应于 i^{th} 群组在当前时间步的状态，即 $h_t^i = LSTM（h_{t-1}^i，\ e_t^i；\ W）$。

轨迹估计：预测轨迹被假定服从正态分布，因此表示为：$(\hat{x}_g，\ \hat{y}_g)_t^i \sim \mathcal{N}（\mu_t^i，\ \sigma_t^i，\ \rho_t^i）$。此外，通过最小化每条 i^{th} 轨迹的负对数似然损失来训练模型参数。

$$L^i(W) = -\sum_{t_6}^{t_{10}} logP((x_g, y_g)_t^i | \mu_t^i, \sigma_t^i, \rho_t^i) \qquad (2)$$

模型返回的轨迹（x_g，y_g）被传递给导航系统，后者可利用这些轨迹为机器人计算最优路径。由于我们采用群组检测模型，利用行人之间的社会交互来预测轨迹，因此可以说，机器人能够借此部署一个符合社会规范的导航系统，从而避免穿越群组时造成阻碍。

F. 机器人导航

我们将算法部署在一只 Unitree 机器狗上。前文介绍的路径预测算法也可用于在密集人群或行人中穿行，避免发生碰撞。该方法基于广义速度障碍物（GVO）[37]。路径预测与 GVO 相结合后，可用于类汽车机器人，同时考虑其动态约束。

在此情境中，我们采用类似汽车的运动学约束，并假设机器人能够检测附近动态障碍物（例如行人和其他机器人）的位置，尽管存在传感器噪声。我们已在机器人上实现此方法，使其能够穿过拥挤人行道，前往预定目的地。

GVO导航方法是一种利用速度障碍物为具有运动学约束的机器人进行导航的技术。在我们的情境中，我们采用类似汽车的运动学约束，并假设机器人能够检测附近移动障碍物（例如行人）的位置，尽管检测结果会受到一些噪声影响。机器人在穿过拥挤人行道前往预定目的地时，会运用我们的方法预测每位行人的移动。

机器人采用传统运动学模型 [37] 。机器人的状态由其位置（x_r，y_r）和朝向 ϕ 表示。机器人可控制自身的速度和转向角，分别用 u_s 和 u_ϕ表示。

机器人的构型由其位置（x_r，y_r）和朝向 ϕ 表示；机器人可控制速度和转向角，分别用 u_s 和 u_ϕ表示。

假设在给定的时间间隔内控制量保持不变，则可通过以下表达式确定机器人在特定时刻 t的位置：

T机器人的约束规定如下　　流：

$$R(t,u) = \begin{pmatrix} \frac{1}{\tan(u_\phi)} \sin(u_s \tan(u_\phi)t) \\ -\frac{1}{\tan(u_\phi)} \cos(u_s \tan(u_\phi)t) + \frac{1}{\tan(u_\phi)} \end{pmatrix}. \qquad (3)$$

更多详情请参阅 [37]。

机器人利用我们开发的预测方案，预测行人和行人群体的路径，并避免采取可能导致碰撞的转向动作。我们假设行人不会主动采取措施避免与机器人碰撞，这意味着机器人有责任确保避碰。

IV. 数据集

A. 我们的 *SG-LSTM* 数据集与数据标注

我们的数据集（图4）基于在西拉法叶的普渡大学校园采集的数据，以及公开可用的三维点云和 RGB-D 行人数据集 [42]。在本研究中，我们从 RGB 帧中手动标注了拥挤场景中视觉上可见的行人群体。我们采用参考算法辅助标注，该算法



图 4：此处简要展示了我们数据集中行人群体的标签（由边界框表示）。

涉及处理带有检测到的行人二维边界框及其对应深度信息的 RGB 帧。我们使用 [1]中提出的方法，推算行人相对于机器人坐标系的位置。此外，我们还计算了多种特征，例如群体规模、步行速度以及检测到的行人的邻近程度，这些特征对于确定群体凝聚力至关重要。为了将行人归类为群体，我们参考了 [1]中提出的条件。随后，我们使用 CVAT.ai 为所有识别出的群体准确生成二维边界框。接着，我们按照第III节所述的方法，使用标注数据训练模型。标注数据集包含经过人脸模糊处理、时间同步的彩色图像和深度帧，以及行人群体和单个行人的边界框，并将随本文一同发布。

表 II：各方法在不同数据集上的定量结果。

指标和方法	数据集		
平均位移误差	**ETH**	**MOT15**	**我们的数据集**
线性	0.80	0.93	0.65
LSTM	0.60	0.67	0.43
O-LSTM	0.49	0.59	0.32
S-LSTM	0.50	0.57	0.38
SG-LSTM（我们的方法）	0.35	0.40	**0.23**
最终位移误差	**ETH**	**MOT15**	**我们数据集**
线性	1.31	1.01	0.91
LSTM	1.31	0.70	0.58
O-LSTM	1.06	0.66	0.41
S-LSTM	1.07	0.69	0.42
SG-LSTM（我们的方法）	0.68	0.48	**0.27**

五、评估

本节总结了我们的方法与其他方法在所选评估指标上的表现，这些指标用于评估我们的 SG-LSTM 数据集、ETH 和 Hotel 数据集以及 MOT15 数据集。以下是我们测试的方法：

- **线性模型**：我们假设行人沿直线路径行走
- **长短期记忆网络**：不使用社会池化且不进行分组的原始长短期记忆网络
- **O-LSTM**：使用占据图的长短期记忆网络
- **S-LSTM**：使用社会池化的长短期记忆网络
- **SG-LSTM**（我们的方法）：使用社会池化和分组的长短期记忆网络

此外，我们使用以下指标来衡量准确率：

- **平均位移误差**：我们按照 [7]，计算场景中每个人员的轨迹中所有预测点的均方误差。
- **最终位移误差**：我们计算预期轨迹终点与每个人员实际终点之间的距离。

A. 预测轨迹的位移误差

假设结伴同行的行人遵循相同的轨迹，我们就无需计算每个行人的轨迹。其他将行人视为个体的轨迹预测方法，在高密度拥挤环境中效率较低，因为它们需要为场景中的每个人都添加轨迹。而且，这些方法在计算某个人的轨迹时，可能会将周围的人视为独立的障碍物。相比之下，我们的模型通过识别这些人属于同一个群组，避免了这些耗费大量计算的操作。我们通过计算预测轨迹与真实值之间的平均位移误差和最终位移误差，证明了关于群组

表 III：在拥挤场景中测得的各模型计算一条轨迹的平均运行时间

方法	**平均运行时间（毫秒）**
线性	8.4
S-LSTM [32]	101
O-LSTM [32]	63
原始长短期记忆网络	57
SG-LSTM（我们的方法）	45

中的行人共享相同轨迹的假设成立；其中，真实值是场景中每个人的轨迹。表II 中计算出的误差证实了这一假设成立，且不会影响模型的准确率。

B. 运行时性能

我们发现，我们的方法 ***SG-LSTM*** 平均耗时比 S-LSTM 少 55.4%（表III） [32]。平均而言，我们跟踪的行人中有 30%～50% 属于某个群组，这显著缩短了计算时间，因此更适合用于机器人等边缘设备。较低的计算需求意味着硬件和软件开发成本也更低。这使机器人导航系统更经济实惠，也能让更多用户使用。与 Social-LSTM 相比，该方法在准确率和速度方面表现更佳。

VI. 结论 ，局限性 ，与未来工作

我们提出了一种社会群体长短期记忆网络（SG-LSTM）模型，将群体动力学融入行人预测。我们的分层模型将群组层面和个体层面的行人解耦，并能在这两个层面上对个体动态进行建模。我们还表明，预测轨迹可用于构建社会效率更高的导航系统，并利用这些轨迹为机器人规划最优路径。我们还提供了运行时数据，并证明群体优化方法可将计算时间缩短 50% 以上。

与此同时，我们的方法也存在一些局限性。像我们这样的人行人预测模型通常只依赖行人的位置和运动，而不考虑天气、一天中的时间或障碍物是否存在等其他上下文因素。此外，我们的模型仅适用于短距离，例如机器人周围几米范围内。这可能会限制机器人规划与行人长期交互的能力。许多行人预测模型都很复杂，计算开销也很大，因此难以部署在资源受限的机器人平台上。

未来，我们希望着手解决这些问题。我们的模型依赖来自单一传感器模态的数据，例如摄像头。未来工作可以探索融合视觉、激光雷达和雷达等多模态数据，以提高行人预测模型的准确率和鲁棒性。未来，我们可以考虑更广泛的上下文信息，例如天气条件、一天中的时间以及行人行为模式。这有望在各种场景下带来更准确、更可靠的预测。我们还可以探索如何将这些模型扩展到更长的时间范围，例如预测行人在接下来一分钟或更长时间内的轨迹。

参考文献

[1] A. J. Sathyamoorthy、U. Patel、M. Paul、N. K. S. Kumar、Y. Savle 和 D. Manocha，“CoMet：在人群拥挤场景中为社会规范机器人导航建模群体凝聚力”，IEEE RA-L，第7卷，第2期，第1008–1015页，2022年4月

[2] M. Yang 等人，“基于聚类的人群移动行为检测”，载于 2018 2018, 1 8 DICTA，12月，第–页。J. Shao、C. C. Loy 和 X. Wang，“人群中的场景无关群体画像”，载于 2014 2014, 2227 2234 IEEE计算机视觉与模式识别会议，6月，第–页。A. K. Chandran、L. A. Poh 和 P. Vadakkepat，“识别行人群体视频中的社交群体”，载于图像分析与模式识别国际会议，1月，2015, 1 6 第–页。D. Helbing 和 P. Molnar，“行人动力学的社会力模型”，《物理评论E》，第51卷，第5期，第4282–4286页，1995年5月，doi：10.1103/PhysRevE.51.4282。[6] A. Lerner、Y. Chrysanthou 和 D. Lischinski，“《示例驱动的人群》”，《计算机图形论坛》，第26卷，第3期，第655–664页，2007年，doi：10.1111/j.1467-8659.2007.01089.x。[7] S. Pellegrini、A. Ess、K. Schindler 和 L. van Gool，“你永远不会独自行走：面向多目标跟踪的社会行为建模”，载于 2009 2009, 261 268 IEEE国际计算机视觉会议，9月，第–页。M. Luber、J. A. Stork、G. D. Tipaldi 和 K. O. Arras，“利用基于社会力的人体运动预测进行人员跟踪”，载于IEEE国际机器人与自动化会议，5月，2010, 464 469 第–页。R. Mehran、A. Oyama 和 M. Shah，“使用社会力模型进行异常人群行为检测”，载于计算机视觉与模式识别会议，2009年6月，第935–942页。[10] K. Yamaguchi、A. C. Berg、L. E. Ortiz 和 T. L. Berg，“你和谁在一起，又要去哪里？”，载于2011年计算机视觉与模式识别会议，2011年6月，第1345–1352页。doi：10.1109/CVPR.2011.5995468。[11] S. Pellegrini、A. Ess 和 L. Van Gool，“通过联合建模行人轨迹与群体划分改进数据关联”，载于欧洲计算机视觉会议，2010, 2010, 452 465 柏林，海德堡，第–页。L. Leal-Taixé、M. Fenzi、A. Kuznetsova、B. Rosenhahn 和 S. Savarese，“为多人跟踪学习基于图像的运动上下文”，载于 2014 2014, 3542 3549 IEEE计算机视觉与模式识别会议，6月，第–页。A. Treuille、S. Cooper 和 Z. Popovic，“连续体人群”，《ACM图形学汇刊》，第25卷，第3期，第1160–1168页，2006年7月。[14] J. M. Wang、D. J. Fleet 和 A. Hertzmann，“人体运动的高斯过程动力学模型”，IEEE模式分析与机器智能汇刊，第30卷，第2期，第283–298页，2008年2月。[15] G. Antonini、M. Bierlaire 和 M. Weber，“行人步行行为的离散选择模型”，《交通研究B辑：方法论》，第40卷，第8期，第667–687页，2006年9月。[16] M. K. C. Tay 和 C. Laugier，“使用高斯过程对平滑路径建模”，载于《野外与服务机器人：第六届国际会议成果》，C. Laugier 和 R. Siegwart 编。柏林，海德堡：Springer，2008年，第381–390页。[17] S. Yi、H. Li 和 X. Wang，“从静态人群群体中理解行人行为”，载于计算机视觉与模式识别会议，2015年6月，第3488–3496页。[18] H. S. Park 和 J. Shi，“社会显著性预测”，载于计算机视觉与模式识别会议，2015年6月，第4777–4785页。[19] A. Alahi、V. Ramanathan 和 L. Fei-Fei，“具备社会感知能力的大规模人群预测”，载于计算机视觉与模式识别会议，2014年6月，第2211–2218页。[20] M. Rodriguez、J. Sivic、I. Laptev 和 J.-Y. Audibert，“视频中的数据驱动人群分析”，载于IEEE国际计算机视觉会议，2011年11月，第1235–1242页。[21] K. Kim、D. Lee 和 I. Essa，“用于运动轨迹分析的高斯过程回归流”，载于IEEE国际计算机视觉会议，2011年11月，第1164–1171页。[22] B. T. Morris 和 M. M. Trivedi，“面向活动理解的轨迹学习：无监督、多层级和长期自适应方法”，IEEE模式分析与机器智能汇刊，第33, 11, 2287 2301, 2011 卷，第期，第–页，11月。J. Azorin-Lopez、M. Saval-Calvo、A. Fuster-Guillo 和 A. Oliver-Albert，“一种基于轨迹表示的人类行为识别预测模型”，载于2014年IJCNN，2014年7月，第1494–1501页。[24] J. Elfring、R. van de Molengraft 和 M. Steinbuch，“学习意图以改进人体运动预测”，载于2013年ICAR，2013年11月，第1–7页。[25] Y. Kong、D. Kit 和 Y. Fu，“一种具有多个时间尺度的动作预测判别模型”，载于《计算机视觉——欧洲计算机视觉会议2014》，Cham，2014年，第596–611页。[26] K. M. Kitani、B. D. Ziebart、J. A. Bagnell 和 M. Hebert，“活动预测”，载于《计算机视觉——欧洲计算机视觉会议2012》，柏林，海德堡，2012年，第201–214页。[27] Y. Cao 等人，“从部分观测视频中识别人类活动”，载于2013年IEEE计算机视觉与模式识别会议，2013年6月，第2658–2665页。[28] T.-H. Vu、C. Olsson、I. Laptev、A. Oliva 和 J. Sivic，“从静态场景预测动作”，载于《计算机视觉——欧洲计算机视觉会议2014》，Cham，2014年，第421–436页。[29] B. Minor、J. R. Doppa 和 D. J. Cook，“数据驱动的活动预测：算法、评估方法论与应用”，载于ACM SIGKDD知识发现与数据挖掘国际会议，2015年，第805–814页。[30] C. Vondrick、H. Pirsiavash 和 A. Torralba，“从无标签视频中预测视觉表征”，载于2016年IEEE计算机视觉与模式识别会议，2016年6月，第98–106页。doi：10.1109/CVPR.2016.18。[31] M. S. Ryoo、T. J. Fuchs、L. Xia、J. K. Aggarwal 和 L. Matthies，“使用起始表征从第一人称视频中早期识别人类活动”。arXiv，2015年7月6日。[32] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese，《Social LSTM：拥挤空间中的人类轨迹预测》，发表于 2016 2016, 961 971 IEEE计算机视觉与模式识别会议，6月，第–页。A. J. Sathyamoorthy, U. Patel, T. Guan, and D. Manocha，《Frozone：在人群中实现无冻结、行人友好的导航》，IEEE机器人与自动化快报，第5卷，第3期，第4352–4359页，2020年7月。[34] L. Tai, J. Zhang, M. Liu, and W. Burgard，《通过原始深度输入和生成对抗模仿学习实现社会规范导航》。arXiv，2月。[35] 26, 2018 Y. F. Chen, M. Everett, M. Liu, and J. P. How，《使用深度强化学习进行具备社会意识的运动规划》，发表于 IEEE/RSJ智能机器人与系统国际会议，2017年9月，第1343–1350页。[36] Y. F. Chen, M. Liu, M. Everett, and J. P. How，《使用深度强化学习实现非通信多智能体的分散式避碰》。arXiv，2016年9月28日。[37] Wilkie, D., Van Den Berg, J., and Manocha, D.（2009年10月）。广义速度障碍物。发表于 IEEE/RSJ智能机器人与系统国际会议（第5573–5578页）。IEEE。[38] A. Bera, S. Kim, T. Randhavane, S. Pratapa, and D. Manocha，《GLMP：利用全局和局部运动模式实时预测行人路径》，发表于 IEEE国际机器人与自动化会议，2016年5月，第5528–5535页。[39] R. Chandra, U. Bhattacharya, A. Bera, and D. Manocha，《TraPHic：利用加权交互预测密集异质交通中的轨迹》，发表于 IEEE计算机视觉与模式识别会议，2019年6月，第8475–8484页。[40] A. Bera, T. Randhavane, R. Prinja, and D. Manocha，《SocioSense：在社会和心理约束下实现行人间的机器人导航》，发表于 IEEE/RSJ智能机器人与系统国际会议，2017年9月，第7018–7025页。[41] Y. Luo, P. Cai, A. Bera, D. Hsu, W. S. Lee, and D. Manocha，《PORCA：在人流众多的环境中建模和规划自动驾驶》，IEEE机器人与自动化快报，第3卷，第 4, 3418 3425, 2018 期，第–页，10月。[42] Diego Paez-Granados, Yujie He, David Gonon, Lukas Huber, Aude Billard, 2021年1月27日，《机器人在人群中导航时行人的三维点云和RGBD数据：检测与跟踪》，IEEE Dataport。[43] D. Paez-Granados, Y. He, D. Gonon, L. Huber, and A. Billard，《自动人群导航中的行人与机器人交互：数据集与指标》，第3页。[44] A. Ess, B. Leibe, and L. Van Gool，《移动场景分析中的深度与外观》，发表于 IEEE国际计算机视觉会议 2007，第1–8页。[45] L. Leal-Taix´e, A. Milan, I. Reid, S. Roth, and K. Schindler，《MOTChallenge 2015：迈向多目标跟踪基准》。arXiv，2015年4月8日。[46] E. Cheung, A. Bera, and D. Manocha，《基于驾驶员行为分类的高效安全车辆导航》，发表于 IEEE/RSJ智能机器人与系统国际会议，2012年，第4472–4479页。[47] R. Chandra, U. Bhattacharya, T. Mittal, A. Bera, and D. Manocha，《CMetric：利用中心性函数度量驾驶行为》，发表于 IEEE/RSJ智能机器人与系统国际会议，2016年，第3943–3950页。[48] T. Randhavane, A. Bera, E. Kubin, A. Wang, K. Gray, and D. Manocha，《面向具备社会意识的机器人导航的行人主导性建模》，发表于 IEEE/RSJ智能机器人与系统国际会议，2016年，第3162–3169页。[49] A. Bera, N. Galoppo, D. Sharlet, A. Lake, and D. Manocha，《Adapt：拥挤场景中的实时自适应行人跟踪》，发表于 IEEE 计算机视觉与模式识别会议，2015 年，第 2125–2134 页。[50] A. Bera and D. Manocha，《利用互惠速度障碍实现实时多层级人群跟踪》，IEEE图形学汇刊，第23卷，第1期，第521–530页，2017年。[51] A. Bera, S. Kim, T. Randhavane, S. Pratapa, and D. Manocha，《GLMP——利用全局和局部运动模式实时预测行人路径》，发表于 IEEE/RSJ智能机器人与系统国际会议，2016年，第3935–3942页。[52] A. Bera, S. Kim, and D. Manocha，《利用轨迹级人群行为学习进行实时异常检测》，发表于 IEEE国际计算机视觉会议，2017年，第2214–2222页。[53] A. Bera, S. Kim, and D. Manocha，《面向数据驱动人群仿真和内容生成的在线参数学习》，ACM图形学汇刊，第36卷，第6期，第193页，2017年。[54] V. Murino, M. Cristani, S. Shah, and S. Savarese，《面向计算机视觉的群体与人群行为》，IEEE模式分析与机器智能汇刊，第35卷，第12期，第2875–2891页，2012年。[55] A. Bera, T. Randhavane, E. Kubin, A. Wang, K. Gray, and D. Manocha，《社会隐形机器人：利用机器人实体性在社会世界中导航》，发表于 IEEE/RSJ智能机器人与系统国际会议，2019年，第4193–4200页。[56] A. Bera, S. Kim, and D. Manocha，《用于监控和训练的交互式人群行为学习》，发表于 IEEE计算机视觉与模式识别会议，2018年，第8044–8053页。[57] R. Chandra, U. Bhattacharya, X. Mittal, Trisha, Li, A. Bera, and D. Manocha，《GraphRQI：利用图谱分类驾驶员行为》，IEEE智能交通系统汇刊，第19卷，第5期，第1588–1598页，2017年。[58] E. Cheung, A. Bera, E. Kubin, K. Gray, and D. Manocha，《利用轨迹特征识别车辆导航中的驾驶员行为》，IEEE智能交通系统汇刊，第13卷，第3期，第1181–1191页，2012年。