arXiv:2304.00736v1 [cs.RO] 3 Apr 2023

TacGNN：使用分层图神经网络，让盲机器人学习基于触觉的手内操作

Linhan Yang[1]，Bidan Huang[2†], Qingbiao Li[3],Ya-Yen Tsai[4],Wang Wei Lee[2],Chaoyang Song[5],Jia Pan[6]

摘要—本文提出了一种全新的框架，用于让盲人形机器人手（即不依赖视觉感知）学习基于触觉的灵巧操作。首先，利用基于图的感知模型 TacGNN 从原始触觉信号中提取物体相关状态。随后，在第二阶段，将得到的触觉特征用于手内操作任务的策略学习。我们通过保定球任务对该方法进行验证：在手中同时操控两个球体，使其彼此绕行 180 度。我们使用强化学习算法（近端策略优化）开展了物体状态预测和手内操作实验。结果表明，与多层感知机、卷积神经网络和图卷积网络等其他方法相比，TacGNN 能有效预测操作过程中的物体相关状态，将预测均方根误差降低至 0.096 cm。最后，机器人手仅依靠机器人自身的感知能力——触觉感知和本体感觉——即可完成手内操作任务。此外，我们在三个难度等级不同的任务上测试了该方法，并将其迁移至真实机器人，且无需进一步训练。https://sites.google.com/view/tacgnn

索引词—灵巧操作、力与触觉感知、表征学习、强化学习

I. 引言

ENABLING 机器人实现人类水平的操作一直是机器人学家的梦想 [1]。人手在执行灵巧物体操作方面的能力无可匹敌。尽管人手的卓越能力很大程度上归功于高度进化的解剖结构所赋予的众多自由度，但高分辨率触觉反馈同样至关重要。抓取、手内操作和手指步态操作等任务无法……

稿件收到日期：2023年1月20日；接受日期：2023年3月17日。副编辑 Ashis Banerjee 根据副编辑和审稿人的意见评估后，推荐本文发表。

†表示通讯作者。[1]Linhan Yang 就职于南方科技大学机械与能源工程系，以及香港大学计算机科学系，中国香港。本研究是在他于腾讯机器人实验室X实习期间开展的。yanglh@connect.hku.hk[2]B. Huang 和 Wang Wei Lee 就职于腾讯机器人实验室X，中国bidanhuang@tencent.com, wwlee@tencent.com[3]Qingbiao Li 就职于英国剑桥剑桥大学计算机科学与技术系。ql295@cam.ac.uk[4]Y.-Y. Tsai 就职于伦敦帝国理工学院 Hamlyn机器人外科中心计算系，英国伦敦 SW7 2AZ。y.tsai17@imperial.ac.uk[5]Chaoyang Song 就职于南方科技大学机械与能源工程系，中国广东深圳 518055。songcy@ieee.org[6]Jia Pan 就职于香港大学计算机科学系，中国香港 999077。jpan@cs.hku.hk 数字对象标识符（DOI）：见本页顶部。

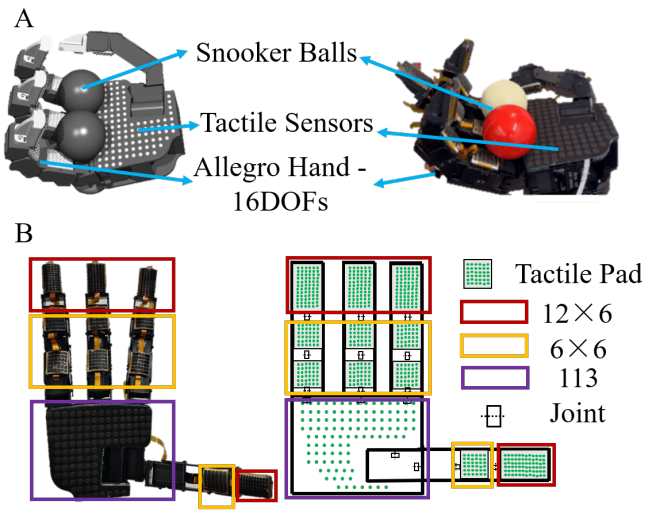

图1：硬件设置。（A）覆盖着分布式触觉传感器的 Allegro Hand 在仿真（左）和真实世界（右）中操作手内的两个斯诺克球。（B）Allegro Hand 上的传感器分布图。

由于可以根据接触面积分布推断手内物体状态，而接触的断开与重新建立则标示了操作过程中的关键过渡事件，因此无需触觉反馈也能高效完成操作 [2]。可以理解的是，在机器人手上实现类人的物体操作是一项重大挑战。

与人类不同，机器人操作长期以来依赖视觉来解读物体与手之间的交互。机器人视觉系统通常需要精确校准、多台摄像机、复杂算法或昂贵的动作捕捉系统 [3], [4], [5]。遗憾的是，视觉并非对力的直接观测，因此无法提供灵巧操作所需的接触信息。因此，尽管基于视觉的方法 [3] 近年来取得了值得肯定的进展，但由于触觉信息的应用尚未得到充分探索，物体操作仍然是一项开放挑战。

将触觉信号用于操作的一大障碍，是难以开发出能够捕捉时变触觉传感器读数与手部运动之间关系的解析模型。近期，人们开始探索基于学习的方法，使机器人能够直接从经验或示范中学习，以实现灵巧操作 [5], [6]。与力/力矩传感等其他感知模态不同，后者的信号可以简单地建模为时间序列向量 [7], ，大面积电子皮肤通常会产生高维触觉

信号，这些信号往往稀疏且结构不规则 [8]。传感器的形状和尺寸也可能因所用硬件而异（图 1）。因此，神经网络的结构对于利用触觉反馈进行学习式操作的效果至关重要。

早期研究曾尝试使用卷积神经网络（CNN）[6], [9]对高分辨率触觉输入进行建模。然而，卷积神经网络（CNN）仅接受预先定义尺寸的矩形输入，因此必须调整触觉输入的尺寸，这会丢失触觉像素之间固有的空间关系。为克服这一局限，研究人员引入了图卷积网络（GCN）[10] 来编码和解读触觉信息 [11], [12]。尽管图卷积网络（GCN）能够处理非结构化的触觉信号，但图的邻接矩阵（例如节点数量以及节点之间的连接）必须从一开始就指定，并在此后保持不变。因此，网络难以编码操作过程中机器人手可能采取的多种姿态所引发的触觉信号时变空间关系。此外，当系统中的触觉传感器配置不同时，图卷积网络（GCN）也无法适用，从而限制了该解决方案在不同平台上的适用性。

与上述方法不同，我们将触觉信号视为点集，其中每个点表示一个被激活的触觉传感器的三维位置。由于手上的接触区域分布不均，这些触觉点集在不同区域的密度也不均匀。我们使用分层图神经网络解析该触觉点集，以提取不同层级的特征 [13]。该触觉图称为 TacGNN，是动态的，即随着手指移动和接触区域变化，图的连接关系和节点数量也会改变。TacGNN 的独特之处在于能够适应任意传感器配置，并高效捕捉传感器之间的空间关系。基于 TacGNN 预测的物体位姿，我们仅依靠机器人感知，使用一种同策略强化学习算法（近端策略优化，PPO）训练了一个用于手内操作任务的自主控制策略（第三节）。

贡献。 总的来说，我们提出了一个用于学习和利用触觉特征以完成手内操作任务的框架。我们的主要贡献总结如下：

1) 我们提出了一种新方法 TacGNN，用于从分布式触觉感知中提取触觉特征。在物体位姿预测实验中，该方法优于 MLP、CNN、GCN 等已有方法。2) 我们结合 TacGNN 的预测输出和一种在线强化学习算法，完成了复杂的操作任务——保定球任务。3) 我们在简单、中等和困难任务上测试了该方法，成功率分别达到 94.71%、88.57% 和 79.78%，大幅超越了已有方法。4) 我们将仿真中学习到的策略迁移到真实机器人上，实验结果表明，真实环境中的表现与仿真结果一致。

II. 相关工作

A. 基于学习的手内操作方法

随着手的机构设计日益专用且复杂，要设计解析模型 [14] 来控制手完成各种任务变得越来越困难。深度学习的兴起为研究人员带来了巨大的潜力，使他们能够像人类一样通过经验实现灵巧操作 [5], [15], [16]。现有研究利用摄像头或动作捕捉系统提供的视觉数据来实现灵巧操作 [4], [11], [14]。例如，OpenAI 的研究人员提出了基于摄像头视觉数据的无模型强化学习方法，用于求解魔方 [5]。Bhatt等人 [14]提出了一种无需强化学习的简单模型，该模型基于动作捕捉系统，在操作各种形状、尺寸和重量的物体时表现出令人惊讶的鲁棒性。然而，人类无需任何视觉输入，仅凭手部触觉就能实现灵巧操作。灵巧操作中的这些额外约束促使研究人员积极探索如何处理触觉数据，以及如何基于这些数据设计操作控制算法。

B. 基于触觉的方法

研究人员一直在探索基于向量和基于图像的方法来处理触觉数据 [6], [17], [18]–[19]。Lee等人 [7] 将触觉信号简单地建模为一个向量。这种方法特别适用于机器人关节的本体感知力/扭矩。Funabashi等人 [17]提出了一种简单的前馈卷积神经网络，将触觉传感数据视为图像。然而，触觉传感器只有在接触物体时才能获取数据，因此随着物体移动，传感器读数在空间上十分稀疏。Sundaram等人 [6]提出了一种类似的基于卷积神经网络的方法，利用分布式触觉传感器进行物体分类。他们首先将触觉信号重塑为矩形尺寸，以适配卷积神经网络的输入，但这破坏了其固有的空间关系。这使得基于卷积神经网络的方法处理触觉传感器数据变得困难。

C. 用于手内操作的图神经网络

从分子生物学到多机器人路径规划的近期研究 [20] 表明，图神经网络能够有效捕捉图中节点之间的交互，并应用于节点与边分类以及消息传递。这一特性推动了利用图神经网络 [12], [21]捕捉触觉传感器数据、实现手内操作的研究。图神经网络在捕捉节点间反应方面的特性，启发研究人员将图神经网络用作控制算法。

近年来，图神经网络的优势已被应用于触觉感知。Garcia-Garcia等人 [21] 提出了一种基于图卷积网络的方法，用于提取 BioTac SP 触觉像素的读数，并将抓握二分类为稳定或易滑。Funabashi等人 [12]提出了一种基于图卷积网络的流程，用于从分布式触觉像素的触觉数据中提取测地特征。然而，他们的工作没有考虑图的拓扑结构如何随触觉像素读数的变化而改变。由于输入和输出的矩阵维度都是固定的，图卷积网络只能处理整个图，因此难以处理整个图的子集。

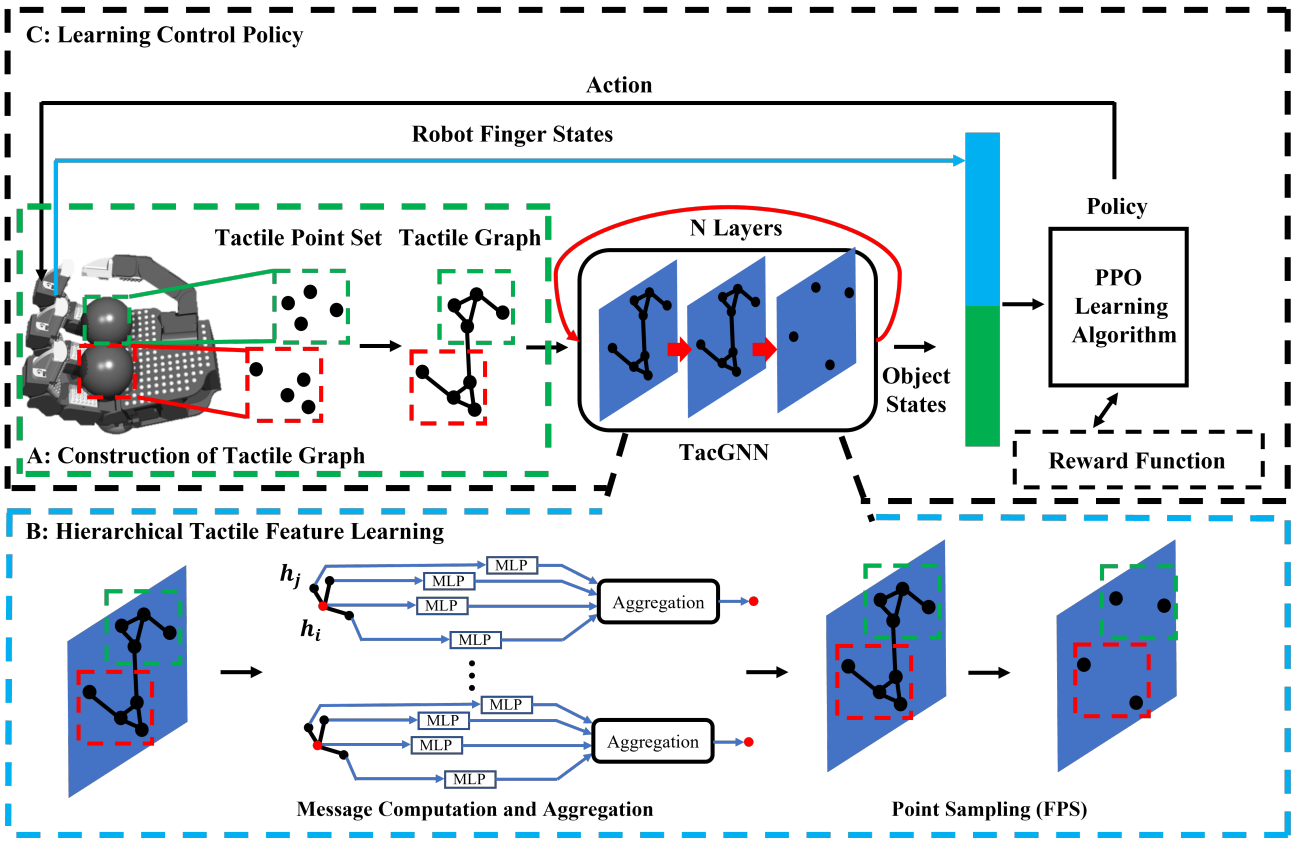

图 2：所提出的保定球任务控制方法示意图：在手中同时操纵两个球体，使它们彼此旋转 180 度。（A）首先，将触觉信息视为一个点集（三维位置）。然后，根据每个触觉传感器的位置生成触觉图。（B）学习触觉特征。首先，我们根据邻域计算每个图节点的嵌入，并对其进行聚合。然后，通过最远点采样（FPS）对点进行采样。这两个步骤在多个层中重复执行。（C）最后，将 TacGNN 预测的物体状态与其他机器人状态拼接，并输入控制策略（PPO）。控制策略根据状态向量生成动作，并根据奖励函数进行更新。

III. 方法

首先，我们将在第 III-A 节介绍图神经网络的基本概念，然后在第 III-B 节说明如何使用 TacGNN 从分布式触觉感知中解读物体状态。该模型称为感知模型。在第 IV 节的消融研究中，所有其他方法，如多层感知机、卷积神经网络和图卷积网络，均用作感知模型。随后，我们将在第 III-C 节阐述强化学习的形式化描述。借助学习算法，我们为一项手内操作任务学习连续控制策略。最后，我们将在第 III-D 节总结感知模型和控制策略，并展示训练过程的工作流程。

A. 图神经网络

本节简要回顾图神经网络的相关概念。图通常由节点和边组成。GNN 是一种可优化的变换，作用于图的所有属性（节点、边），同时保持图对称性（置换不变性）。[22] 受 PyTorch Geometric 文档的启发，[23], 我们可以将图像上的卷积算子推广为图卷积算子，用于聚合邻域信息或采用消息传递机制进行运算。将第 $l-1$ 层中节点 i 的节点特征定义为 $\mathbf{h}_i^{l-1} \in R$，并将来自节点 j 到节点 i 的边特征定义为 $\mathbf{e}_{j,\,i} \in R^D$，则消息传递图神经网络可表示为

$$\mathbf{h}_i^{(l)} = \gamma^{(l)}\left(\mathbf{h}_i^{(l-1)}, \mathcal{A}_{j\in\mathcal{N}_{\text{in}}(i)}\phi^{(l)}(\mathbf{h}_i^{(l-1)}, \mathbf{h}_j^{(l-1)}, \mathbf{e}_{j,i})\right) \quad (1)$$

其中，$\mathcal{A}$ 是任意聚合函数的通用符号，该函数具有可微且对排列不变的性质，例如求和、求均值或取最大值。此外，我们使用 γ 和 ϕ来定义可微函数，例如多层感知机（MLP）。

B. 触觉图神经网络（*TacGNN*）

本节介绍我们提出的感知模型——触觉图神经网络（TacGNN）：首先，根据每个触觉像素的位置构建触觉图；其次，采用三层分层图神经网络来解析触觉特征。

1）图构建：如图 2 所示，我们将触觉信息定义为一个点集 $P = \{p_1, p_2, p_3, ..., p_n\}$，其中 $p_i = (x_i)$ 是一个具有三维坐标 $x_i \in R^3$ 的触觉像素。

给定点云 P，我们通过为每个触觉节点查找 K 个最近邻来构建触觉图 $G = (P, E)$，这一过程称为 kNN 搜索算法。E 表示连接邻居的边。本文中 k 取 3。

我们方法的一个显著特点是只考虑已激活的触觉传感器。因此，图节点的数量会根据

机器人手与物体之间的交互状态而变化。因此，我们会在每一步重新构建图。得益于我们的方法只在局部节点之间传递信息，而无需考虑整个图的结构，这种方法得以实现。触觉学习模型的详细信息将在下一小节中介绍。

2）分层触觉特征学习：自然地，激活的触觉传感器呈稀疏分布，这意味着我们的触觉点集在不同区域的密度并不均匀。只有接触区域内的传感器会被激活。在我们的项目中，由于柔顺性，球与手每次接触通常会激活4至5个触觉传感器。仿真中的柔顺性基于真实机器人。

这种不均匀性给点集特征学习带来了巨大挑战。受卷积神经网络的启发，卷积神经网络能够通过池化操作，沿多分辨率层次结构逐步捕获尺度越来越大的特征。在较低层级，神经元的感受野较小；而在较高层级，感受野较大。因此，我们采用具有类似功能的分层图神经网络来处理这种触觉分布 [13],[24]。

我们的模型基本思路如下：首先，在较小的邻域内进行消息传播，以提取局部连接中的特征；然后，将图下采样为稀疏图，以便提取更大尺度上的特征。重复这一过程，直到获得整个触觉图的特征。最后，使用物体状态预测多层感知机得到最终输出。我们的层次结构由集合抽象操作——点采样构成，如图2(B)所示。

总而言之，我们的整个分层网络由三个图神经网络层组成，每层包含三个阶段：消息计算、消息聚合和点采样。最后添加一个多层感知机，以获取物体状态。

消息计算。在此阶段，我们使用一个以隐藏特征和相对位置为输入的多层感知机，为每条边计算传递的消息。输入维度为 $C \times 2+3$，输出维度为 C，其中 C 表示每个节点的特征通道数。我们将在第 IV-C 节进一步讨论 C 的选择。

$$\mathbf{h}_{i,j}^{l}=\phi^{(l)}\left(\mathbf{h}_{i}^{l},\mathbf{h}_{j}^{l},x_{j}-x_{i}\right) \tag{2}$$

其中 h_j^l 是传播层 l中节点 j 的隐藏特征，x_j，x_i 是点 i 的欧几里得坐标，j 而 $h_{i,\ j}^{l}$ 是由 j 传递给 i 的消息（j 是 i的邻居节点之一）。首先，将点的坐标转换到相对于质心点的局部坐标系中：$x_j - x_i$。使用相对坐标能够更高效地捕捉局部区域 [13]内点与点之间的关系。

消息聚合。对于每个节点，我们从邻居节点获取 $n \times C$特征。然而，每个节点的数量各不相同。为了保持维度一致，我们对每个点传递的消息进行聚合。输入维度为 $n \times C$，输出维度为 C，其中 C 表示每个节点的特征通道。

$$\mathbf{h}_{i}^{l+1}=\mathcal{A}\left(\mathbf{h}_{i,j}^{l} \mid j \in \mathcal{N}_{\text{in}}(i)\right) \tag{3}$$

其中，$h_{i,\ j}^{l}$ 是由 j 传播至 i 的消息（j 是 i的一个邻居节点），而 $\mathcal{N}_{\text{in}}$（i）包括所有邻居节点

节点 i的，并且 h_i^{l+1} 表示传播步骤 $l+1$时节点 i 的隐藏特征。本文使用的聚合函数 $\mathcal{A}$ 为最大值

点采样。为了提取层次特征，我们通过最远点采样（最远点采样） [25]对点集进行下采样。给定输入点 $\{p_1, p_2, ..., p_n\}$，我们使用最远点采样选择点的子集 $\{p_{i1}, p_{i2}, ..., p_{im}\}$，使得 p_{ij}是相对于其余点而言距离集合 $\{p_{i1}, p_{i2}, ..., p_{ij-1}\}$最远的点。使用迭代式最远点采样，网络能够从局部特征中提取全局特征。输入是维度为 $N \times C$ 的图，输出是维度为$N' \times C$ 的图，其中 $N' < N$。本文中，$N' = 0.5 \times N$**物体状态预测**。最后，我们聚合整个图的所有节点，得到一个图级特征。输入是维度为 $N \times C$ 的最终图，输出是一个节点，其特征通道数为C'。本文使用的聚合函数 $\mathcal{A}$ 为最大值。随后，使用多层感知机预测物体状态，即物体的位置和姿态。在操作任务中，预测两个球体的三维位置（一个 6D 向量）。训练细节将在第三-D 节中介绍。

$$\hat{o}=MLP\left(\mathcal{A}(\mathbf{h}_{i}^{L})\right). \tag{4}$$

C. 强化学习控制策略

本文的最终目标是学习手内操作任务的控制策略。在此，我们将控制问题表述为一个无限时域折扣部分可观测马尔可夫决策过程（POMDP）。我们将状态空间或观测空间定义为 S，将动作空间定义为A。为了与环境交互，智能体根据当前状态 s生成其随机策略 $\pi_\theta(a \mid s)$，其中a 是动作，θ 是策略函数的参数。另一方面，环境为智能体产生奖励 $r(s,\ a)$，而智能体的目标是找到一个能够最大化期望奖励的策略。需要注意的是，状态空间 S 是部分可观测的，即只能观测到机器人手部相关状态，包括机器人关节状态和触觉信号。

为了最大化期望奖励，我们采用近端策略优化（PPO）作为学习算法 [26]。我们将第三-B节中感知模型预测的物体状态与机器人手的本体感知状态（关节角度和速度）拼接后作为输入。输出为关节位置控制指令。

在强化学习算法中，RLU 被用作激活函数。动作网络和价值网络的 MLP 架构为 [512, 256,128]。我们使用 Adam优化器，学习率为 0.001。

D. 工作流程

我们的框架由两个阶段组成——感知和控制器。首先，感知模型通过 Kaiming 初始化 [27] 进行初始化并冻结。随后，使用强化学习算法训练控制策略。输入状态包括机器人手指角度、机器人手指速度、由感知模型推断出的物体状态，以及执行的上一个动作。输出为位置控制命令。奖励函数为：

$$R(s,a)=0.5\times R_{\text{angle}}+250\times R_{\text{success}}+(-100)\times R_{\text{fail}}, \tag{5}$$

其中， $R_{\text{angle}} = \|\text{Angle}_{\text{now}} - \text{Angle}_{\text{pre}}\|$ 表示两个球体的角度变化， $R_{\text{success}} = (\text{Angle}_{\text{now}} > 180)$表示任务是否成功， R_{fail} 表示球是否掉落（脱离手的控制）或时间是否耗尽（最大时间步数为 200）。

在此阶段记录所有触觉信息和物体状态。获得足够的触觉数据（25,000 步）后，使用该数据集通过监督学习训练感知模型。输入状态为触觉信息，输出为预测的物体状态。损失函数采用均方根误差（RMSE）计算。

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T}(\hat{o}_t - o_t)^2}{T}}, \tag{6}$$

其中， T 是物体状态的维度，在此任务中为 6。$\hat{o}$ 是预测输出， o 是对应的状态值。经过 10 个训练轮次后，该方法进入强化学习阶段。这两个阶段重复执行，直到达到最大迭代次数。

IV. 实验

A. 配备分布式传感器的 *Allegro Hand*

使用一只 Allegro Hand[1]，，它是一只四指仿人机器人手，每根手指有四个关节。因此，整个系统有 16 个自由度。

所有手指和手掌均覆盖了我们自行定制的电阻式触觉传感器 [28]。图 1(B) 展示了传感单元（触觉像素）的分布。每根手指的指尖覆盖一个 12×6触觉像素阵列，而每个内侧指骨连杆上都覆盖一个 6× 6触觉像素阵列。手掌上覆盖有 113 个触觉像素。每个触觉像素的读数是一维的，对应其所受的法向力。整个机器人系统配置如图 1 所示。使用两个尺寸相同的球作为被操作物体。

仿真实验在 Isaac Sim [29] 中进行。我们基于真实系统构建仿真模型，以应对仿真到现实迁移问题。仿真中的所有状态均添加了 10% 的噪声。GNN 模型使用 PyTorch-Geometric [23] 构建，并运行于配备 V100 GPU 的 Ubuntu 18.04 系统上。所有实验均使用不同的随机种子重复三次。

B. 物体状态预测

本节重点介绍如何通过监督学习获取触觉特征，并将其解读为物体状态。输入为触觉感知信息，输出为手中物体的位置和姿态。

为验证我们提取触觉特征方法的有效性和效率，我们针对一些形状规则的物体——立方体、圆柱体以及 YCB 物体集中的肉罐头 [30]——开展了物体状态预测任务。为了获得高接触数据集，我们对每个物体以随机初始姿态执行力量抓握 200,000 次。每次执行仅记录最后一步，记录数据包括触觉信号、物体位置和姿态。请注意，在某些情况下，物体的姿态具有旋转对称性，无法通过触觉感知感知到。例如，正方形绕自身旋转 90 度时，触觉传感器会获得完全相同的信号。因此，我们将立方体的旋转角度限制为 [0,90]。类似地，圆柱轴的旋转角度将被忽略。

表 I：物体状态预测实验结果。

物体	评估指标	MLP	CNN	GCN	TacGNN
Cube	位置损失 (cm)	0.187	0.891	0.195	**0.176**
	方向损失 (°)	12.31	15.87	10.33	**9.17**
圆柱体	位置损失 (cm)	0.142	1.424	0.114	**0.067**
	方向损失 (°)	14.25	27.76	12.32	**8.76**
肉罐头	位置损失（cm）	**0.134**	3.172	0.154	0.142
	姿态损失（°）	19.57	20.42	10.11	**9.76**

为证明我们方法的优越性，我们将所提出的方法与其他常用方法进行了比较：**多层感知机**模型直接将 653 个触觉值作为输入，并预测目标物体的状态；**卷积神经网络**模型首先将触觉信号重塑为矩形数组，并将其作为输入。**图卷积网络**将触觉信号作为静态图。它首先采用与 TacGNN 相同的方法，根据每个传感器的初始位姿构建图，并且图结构不会随时间变化。**TacGNN** 仅将已激活的触觉信号作为输入，并在每个步骤中根据交互状态重构图，如第 III 节所示。

每项实验使用 80% 的数据集进行训练，20% 用于测试。采用均方根误差（RMSE）评估预测精度。结果如表 I 所示。可以看出，所提出的方法在全部三个数据集上的表现均优于其他方法。多层感知机和卷积神经网络的结果较差，尤其是在圆柱体和肉罐头数据集上。

表 II：球位置预测实验结果。

网络	架构	训练损失（cm）	测试损失（cm）
MLP	[64,64,64]	0.230	0.250
卷积神经网络+多层感知机	[32,16],[784,128]	0.114	27.76
图卷积网络+多层感知机	[32,32,32],[32,6]	0.130	0.152
TacGNN+多层感知机	[32,32,32],[32,6]	**0.077**	**0.096**
TacGNN-32	[32,32,32],[32,6]	0.077	0.096
TacGNN-16	[16,16,16],[16,6]	0.340	0.251
TacGNN-64	[64,64,64],[64,6]	0.098	0.085
TacGNN-128	[128,128,128],[128,6]	**0.069**	**0.077**

C. 手内操作

在本节中，我们考虑一项手内操作任务——保定球任务。保定球任务是指同时操控两个球体，使它们在手中相互绕行。与以往工作 [4]，不同，我们仅将机器人手部状态作为输入，即机器人关节状态和触觉传感器信号。这项任务颇具挑战，因为两个球共用狭小的工作空间，真正考验了机械手的灵巧性。此外，物体间接触难以估计和推断。

[1]https://www.wonikrobotics.com/research-robot-hand

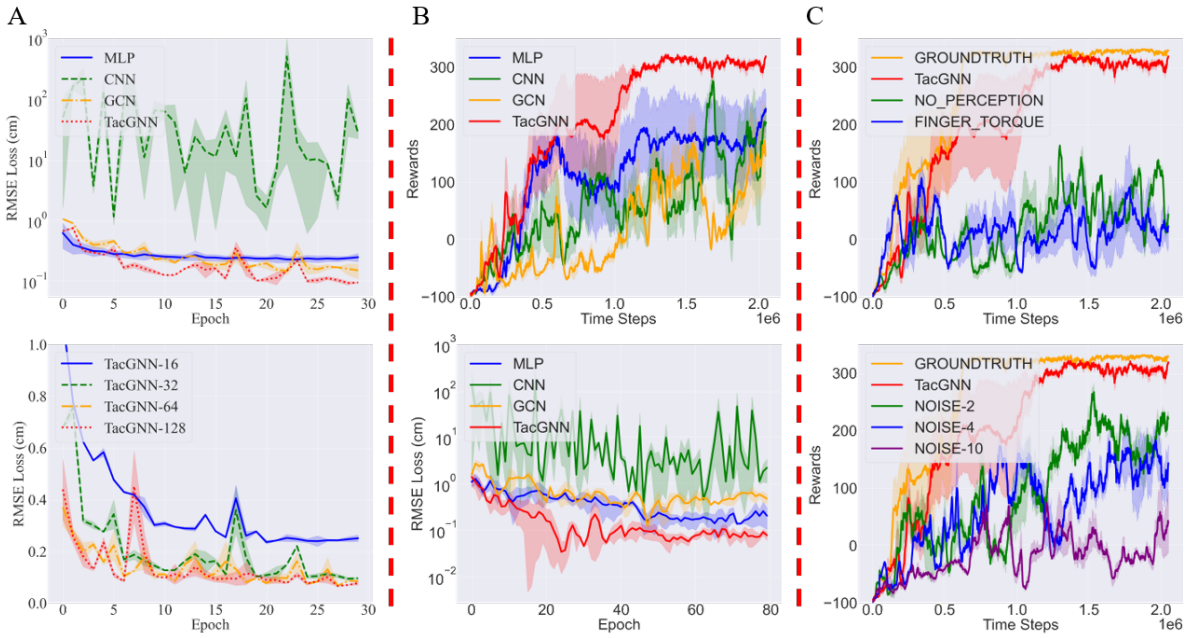

图3：手内操作任务的学习结果。(A) 使用不同感知模型——多层感知机、卷积神经网络、图卷积网络、TacGNN（上），以及不同层大小（下）时的球位置预测均方根误差损失（厘米）。(B) 使用多层感知机、卷积神经网络、图卷积网络、TacGNN作为感知模型时的保定球任务学习结果（上），以及训练过程中相应的预测均方根误差损失（下）。(C) 使用不同输入时的保定球任务学习结果。误差范围表示三次重复训练的标准差。

在这项任务中，我们首先基于触觉感知预测两个球体的位置。然后，将预测结果和其他机器人状态作为输入来训练控制策略。

1）采用不同架构的感知模型：为明确比较不同方法的感知效果，我们将感知模型从整个训练过程中分离出来。我们采用与之前相同的方法训练感知模型。模型的输入为触觉信息，输出为球体位置 $O = [P_1, P_2]$， $P \in R^3$。

为了获得高质量、即触觉接触丰富的数据集，我们首先训练一个策略，使手能够同时旋转两个球，并精确控制球的位置。只要物体位置足够准确，这也验证了该任务的可行性。利用训练好的模型，我们获得了 20000 条轨迹，每条轨迹由约 1000000 个按时间索引的步骤组成。在每个步骤中，记录原始触觉数据作为输入，并记录两个球的位置作为输出标签。

结果如图 3(A) 和表 II 所示。我们可以看到，我们的方法 TacGNN 的均方根误差（RMSE）最低。此外，卷积神经网络在测试数据集上的均方根误差（RMSE）远高于训练数据集，这意味着卷积神经网络在训练后更容易过拟合。

为了针对下游操作任务获得优化的网络架构，我们将通道数（即网络的层大小）调整为 16、64 和 128，如图 3(A) 和表 II 所示。结果表明，综合考虑预测准确度和计算资源消耗，32 个通道已经足够。

2）使用不同感知模型学习控制策略：在本节中，我们使用不同的感知模型训练控制策略：多层感知机、卷积神经网络、图卷积网络和 TacGNN。所有方法的感知模型架构与之前相同。

如图 3 (B) 所示，所提出的方法在整个训练过程中优于其他方法。其他方法

表 III：不同感知模型在手内操作任务上的成功率。

	Task	MLP	CNN	GCN	TacGNN
成功率	**简单**	52.44	42.71	70.62	**94.71**
	中等	45.57	63.75	64.51	**88.57**
	Hard	41.78	50.88	55.11	**79.78**
平均步数	**简单**	**76.41**	146.71	99.79	82.14
	中等	**77.472**	132.51	86.83	86.11
	Hard	**79.03**	119.11	102.79	93.93

在某些情况下可以取得良好表现，但方差较大。图 3 (B) 展示了训练期间的预测损失（均方根误差损失）。正如预期，我们提出的方法始终取得最低的损失。手内操作任务的奖励与感知模型的性能呈正相关。

我们在不进行进一步训练的情况下，使用三个难度等级的任务对这些训练好的模型进行了进一步测试。1）简单任务——同时使两个球体相互绕转超过180度，这与训练过程中使用的原始任务相同。2）中等任务——同时使两个球体相互绕转超过180度，并且每十步施加一次扰动力。力的大小为$F \sim \mathcal{N}$(0、0。1) N，方向完全随机。3）困难任务——除中等任务的要求外，该任务还要求机器人在完成旋转任务后保持5步，以防机器人意外达到旋转目标。该任务颇具挑战性，因为机器人必须在任务完成后防止球体掉落。

如表 III 所示，TacGNN 在

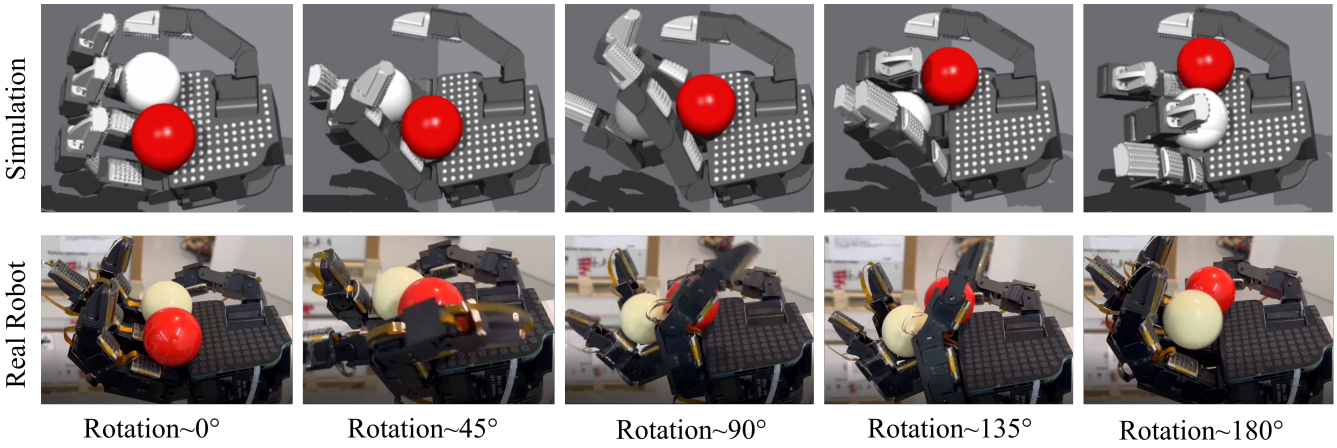

图 4：仿真环境和真实机器人上的手内操作实验关键帧。

全部三个任务中都取得了最高的成功率。尤其是，我们的方法在困难任务中的成功率约为 80%。需要注意的是，中等和困难任务对训练好的策略而言都是未曾见过的场景，且无需进行进一步训练。至于每次试验的平均步数，多层感知机方法始终所用步数最少，但较低的成功率表明，球体掉落会导致实验提前结束。

3）使用不同输入状态学习控制策略：为探究触觉信息对手内操作最终性能的影响，我们通过改变输入状态类型进行了实验。**真实值**：输入包括精确的物体位置、机器人关节角度和速度。**无感知**：输入仅包括机器人关节角度和速度。**手指扭矩**：输入包括机器人关节角度、速度和扭矩。**噪声 2、4、10**：输入包括添加均匀随机噪声 [-2,2], [-5,5], [-10,10] mm 的物体位置，以模拟由遮挡、标定误差等造成的感知误差，以及机器人关节角度和速度。

如图 3（C）所示，随着训练数据增多，TacGNN 最终能够达到真实值方法的性能。由于缺乏足够的交互相关信息，无感知和手指扭矩方法的奖励仍然较低。带噪声的物体信息会影响收敛速度和最终性能。

D. 真实机器人实验

最后，我们在真实机器人系统上验证所提出的方法，如图 1(B) 所示。我们采用域随机化 [31] ，使仅在仿真中训练的策略能够迁移到真实机器人。域随机化参数包括球的质量 [0.1、0.2]kg；摩擦力 [0.8、1.2]；重力 [7.0、12.0]m/s^2；关节刚度和阻尼 [0.8、1.2]× 初始值。关节刚度和阻尼的初始值通过网格搜索确定。首先，真实机器人执行一系列位置控制指令 (0°， 15°， 30°， 45°， 60°)，并记录关节角度。随后，在仿真中对关节刚度[1,3,5]N/m、关节阻尼 [0,0.1、0.5]Ns/m 进行网格搜索，并对每个关节执行相同的位置控制指令。最后，计算仿真与现实中记录的角度之间的损失。参数选择损失最低时对应的参数作为关节刚度和阻尼的初始值。

图4展示了真实机器人上的一次成功试验，其表现与仿真相似。这验证了学习到的触觉特征和策略可应用于真实机器人。

V. 讨论

1) 物体状态感知：机器人触觉感知为机器人与物体之间带来了三维物理关系。与向量和二维图像相比，图数据结构能够保留数据点之间的几何关系。这可能是图卷积网络和 TacGNN 在物体状态预测任务中表现更好的原因，尤其是对于圆柱体和肉罐头。

卷积神经网络将触觉阵列重塑为矩形形状，这会严重破坏触觉信号之间的关系。例如，进行力量抓握时，拇指和中指的指尖会彼此靠近，但在重塑后的图像中，它们却相距较远。这导致卷积神经网络在预测任务中表现最差且不稳定。与图卷积网络相比，TacGNN 仅以激活的触觉节点作为输入，并在每个步骤重建图，这有利于从多种手部姿态中编码空间关系。

2）手内操作：在感知方法的对比实验中，奖励与感知模型的性能呈正相关。采用 TacGNN 的控制策略获得了最高奖励，且收敛时间最短。

除了将两个球体简单旋转超过 180 度这一基础测试外，我们还加入了两个包含外部扰动和保持目标的进阶任务。我们的方法能够很好地处理这些任务，并且无需进一步训练即可泛化到这些任务。

在输入信息的对比实验中，我们可以看到，尽管真实值的收敛速度快于 TacGNN，但只要有足够的训练数据，我们的方法就能达到相近的性能。物体与手部交互信息对这一手内操作任务至关重要。在一定程度上，我们依靠机器人手自身的感知完成了这一灵巧操作任务。

3) 较小的仿真到现实差距：如图4所示，我们在仿真和现实中取得了相近的表现。第一帧展示了机器人和物体的初始位姿。随后，中指

伸展，无名指弯曲，以防止红球掉落。在接下来的几帧中，食指、中指和无名指形成凹面，将白球 удерж住。最后，无名指将红球推到另一侧，白球则被推到手掌中央。拇指是另一道“墙”，用于防止球体掉落。

VI. 结论与未来工作

总之，本文提出了一种用于基于触觉的手内操作任务的新型框架。首先，我们提出了一种新方法——TacGNN，作为感知模型，用于提取被操作物体的触觉特征和状态。然后，利用感知模型提取的触觉特征，我们采用近端策略优化强化学习算法训练自主控制策略，以完成手内操作任务。结果表明，在解读触觉信息方面，我们的方法优于其他常用方法（多层感知机、卷积神经网络、图卷积网络）。最后，我们的感知模型提取的特征能够为手内操作任务提供充足的交互信息，仅依靠机器人自身的感知能力即可实现此类灵巧技能。我们通过在真实机器人系统上实现该方法对其进行了验证，并取得了相当的性能表现。

然而，这项工作只是对基于图神经网络的触觉手内操作任务进行的初步探索。本研究仍存在一些局限。操作过程中会出现“触觉盲区”。例如，物体可能卡在触觉阵列之间的间隙中。在这种情况下，没有触觉传感器被激活，也没有触觉感知输入。一种可能的解决方案是利用时间序列信息，使机器人能够根据先前的数据推断出对物体的近似感知。

参考文献

[1] A. Bicchi，《用于灵巧操作和稳健抓取的手：迈向简洁的艰难道路》，*IEEE*机器人与自动化汇刊，第16卷，第6期，第652–662页，2000年。[2] R. S. Johansson 和 J. R. Flanagan，《物体操作任务中指尖触觉信号的编码与使用》。《自然》评论：神经科学，第10卷，第5期，第345–59页，2009年5月。[在线]。可访问：http://www.ncbi.nlm.nih.gov/pubmed/19532402

[3] O. M. Andrychowicz、B. Baker、M. Chociej、R. Jozefowicz、B. McGrew、J. Pachocki、A. Petron、M. Plappert、G. Powell、A. Ray、等，《学习灵巧的手内操作》，《国际机器人研究期刊》，第39卷，第1期，第3–20页，2020年。[4] A. Nagabandi、K. Konolige、S. Levine 和 V. Kumar，《用于学习灵巧操作的深度动力学模型》，载于机器人学习会议。PMLR，2020年，第1101–1112页。[5] I. Akkaya、M. Andrychowicz、M. Chociej、M. Litwin、B. McGrew、A. Petron、A. Paino、M. Plappert、G. Powell、R. Ribas、等，《用机器人手解魔方》，*arXiv*预印本 *arXiv:1910.07113*，2019年。[6] S. Sundaram、P. Kellnhofer、Y. Li、J.-Y. Zhu、A. Torralba 和 W. Ma-tusik，《利用可扩展触觉手套学习人类抓握特征》，《自然》，第569卷，第7758期，第698–702页，2019年。[7] M. A. Lee、Y. Zhu、K. Srinivasan、P. Shah、S. Savarese、L. Fei- Fei、A. Garg 和 J. Bohg，《理解视觉与触觉：面向高接触任务的多模态表征自监督学习》，载于*2019*年国际机器人与自动化会议（*ICRA*）。IEEE，2019年，第8943–8950页。[8] W. W. Lee、S. L. Kukreja 和 N. V. Thakor，《用于研究神经形态触觉时空特征的千赫兹千触点触觉传感器阵列》，*IEEE*生物医学电路与系统会议：健康心智与健全身体工程，*BioCAS 2015* 会议录，第c期，第1–4页，2015年。

[9] S. Funabashi、T. Isobe、S. Ogasa、T. Ogata、A. Schmitz、T. P. Tomo 和 S. Sugano，《使用三轴触觉传感器和卷积神经网络的低成本机器人手稳定手内操作》，载于 2020年 IEEE/RSJ智能机器人与系统国际会议（IROS）。IEEE，2020年，第 9166–9173 页。[10] T. N. Kipf 和 M. Welling，《基于图卷积网络的半监督分类》，arXiv 预印本 arXiv:1609.02907，2016年。[11] Z. Hu、Y. Zheng 和 J. Pan，《利用两阶段图强化学习抓取活体物体》，IEEE机器人与自动化快报，第 6 卷，第 2 期，第 1950–1957 页，2021年。[12] S. Funabashi、T. Isobe、F. Hongyi、A. Hiramoto、A. Schmitz、S. Sugano 和 T. Ogata，《利用图卷积网络和分布式触觉传感器，对具有不同属性的物体进行多指手内操作》，IEEE机器人与自动化快报，2022年。[13] C. R. Qi、L. Yi、H. Su 和 L. J. Guibas，《Pointnet++：在度量空间中的点集上进行深度层次特征学习》，神经信息处理系统进展，第 30 卷，2017。[14] A. Bhatt、A. Sieler、S. Puhlmann 和 O. Brock，《出乎意料地稳健的手内操作：一项实证研究》，arXiv 预印本 arXiv:2201.11503，2022年。[15] H. J. Charlesworth 和 G. Montana，《利用轨迹优化和强化学习解决具有挑战性的灵巧操作任务》，载于国际机器学习会议。PMLR，2021年，第 1496–1506 页。[16] Y. Niu、S. Jin、Z. Zhang、J. Zhu、D. Zhao 和 L. Zhang，《Goats：结合课程强化学习的舀取目标采样自适应》，arXiv 预印本 arXiv:2303.05193，2023年。[17] S. Funabashi、S. Morikuni、A. Geier、A. Schmitz、S. Ogasa、T. P. Torno、S. Somlor 和 S. Sugano，《使用三维触觉传感器的多指机器人手通过主动感知进行物体识别》，载于 2018年 IEEE/RSJ智能机器人与系统国际会议（IROS）。IEEE，2018年，第 2589–2595 页。[18] L. Yang、X. Han、W. Guo、F. Wan、J. Pan 和 C. Song，《用于刚柔交互抓取的学习型光电神经支配触觉手指》，IEEE机器人与自动化快报，第 6 卷，第 2 期，第 3817–3824 页，2021年。[19] W. Chen、Y. Yan、Z. Zhang、L. Yang 和 J. Pan，《基于聚合物的自校准光纤触觉传感器》，arXiv 预印本 arXiv:2303 00619，2023年。.[20] Z. Wu、S. Pan、F. Chen、G. Long、C. Zhang 和 S. Y. Philip，《图神经网络综合综述》，IEEE神经网络与学习系统汇刊，第 32 卷，第 1 期，第 4–24 页，2020年。[21] A. Garcia-Garcia、B. S. Zapata-Impata、S. Orts-Escolano、P. Gil 和 J. Garcia-Rodriguez，《TactileGCN：一种利用触觉传感器预测抓取稳定性的图卷积网络》，载于 2019 国际神经网络联合会议（IJCNN）。IEEE，2019年，第 1–8 页。[22] B. Sanchez-Lengeling、E. Reif、A. Pearce 和 A. B. Wiltschko，《图神经网络简明入门》，Distill，2021年，https://distill.pub/2021/gnn-intro。[23] M. Fey 和 J. E. Lenssen，《利用 PyTorch Geometric 进行快速图表示学习》，载于 ICLR图与流形表示学习研讨会，2019。[24] W. Shi 和 R. Rajkumar，《Point-GNN：用于点云三维物体检测的图神经网络》，载于 IEEE/CVF计算机视觉与模式识别会议论文集，2020年，第 1711–1719 页。[25] C. Moenning 和 N. A. Dodgson，《快速行进最远点采样》，剑桥大学计算机实验室，技术报告，2003年。[26] S. Fujimoto、H. Hoof 和 D. Meger，《解决演员-评论家方法中的函数逼近误差》，载于国际机器学习会议。PMLR，2018年，第 1587–1596 页。[27] K. He、X. Zhang、S. Ren 和 J. Sun，《深入研究整流器：在 ImageNet分类中超越人类水平性能》，载于 IEEE国际计算机视觉会议论文集，2015年，第 1026–1034 页。[28] Z. Ding、Y.-Y. Tsai、W. W. Lee 和 B. Huang，《利用触觉感知实现机器人操作的仿真到现实迁移》，载于 2021 IEEE/RSJ智能机器人与系统国际会议（IROS）。IEEE，2021年，第 6778–6785 页。[29] V. Makoviychuk、L. Wawrzyniak、Y. Guo、M. Lu、K. Storey、M. Macklin、D. Hoeller、N. Rudin、A. Allshire、A. Handa 和 G. State，《Isaac Gym：用于机器人学习的高性能基于GPU的物理仿真》，2021年。[30] B. Calli、A. Singh、A. Walsman、S. Srinivasa、P. Abbeel 和 A. M. Dollar，《YCB物体与模型集：迈向操作研究的通用基准测试》，载于 2015 国际先进机器人会议（ICAR）。IEEE，2015年，第 510–517 页。[31] J. Tobin、R. Fong、A. Ray、J. Schneider、W. Zaremba 和 P. Abbeel，《域随机化：将深度神经网络从

仿真迁移到现实世界”，载于 *2017 IEEE/RSJ*智能机器人与系统国际会议（*IROS*）。IEEE，2017年，第 23–30 页。