arXiv:2603.01436v1 [cs.RO] 2 Mar 2026

PhysGraph：用于双手灵巧手–工具–物体操作的基于物理的图Transformer策略

Runfa Blark Li、David Kim、Xinshuang Liu、Keito Suzuki、Dwait Bhatt、Nikola Raicevic、Xin Lin、Ki Myung Brian Lee、Nikolay Atanasov、Truong Nguyen 加州大学圣地亚哥分校

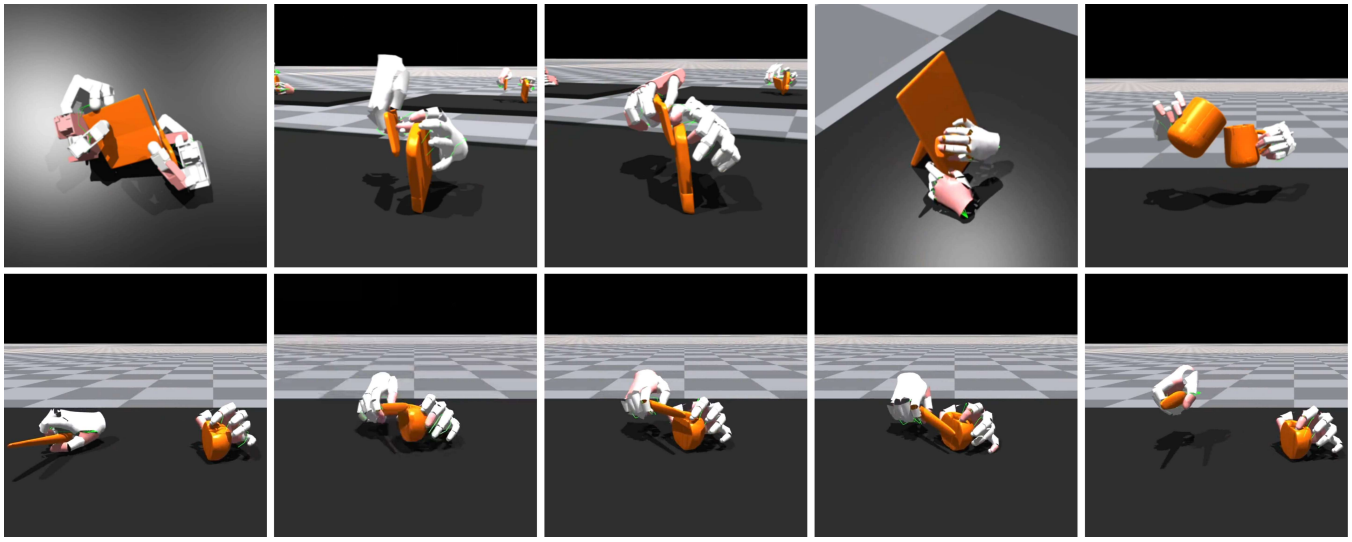

图 1：***PhysGraph*** 是一种基于物理的图 Transformer 策略，用于双手手部—工具—物体操作。 **上排**：在使用不同机器人手（Allegro、ArtiMANO、Shadow）的多种双手工具使用任务上训练和测试的 PhysGraph 策略。 **下排**：将策略零样本迁移到类似任务中未见过的工具和物体，例如，在“用砍刀切面包”任务上训练，并在“用水果刀切苹果”任务上测试。我们的代码和更多实验结果见：https://blarklee.github.io/PhysGraph_website_official/

摘要—由于状态空间维度高、接触动力学复杂，面向工具使用的双手灵巧操作仍是机器人领域的一项巨大挑战。现有方法往往将整个系统状态简单地表示为单个构型向量，忽略了关节式手固有的丰富结构和拓扑信息。我们提出 PhysGraph，这是一种专为具有挑战性的双手手部—工具—物体操作而设计的基于物理的图 Transformer 策略。与先前工作不同，我们将双手机器人系统表示为运动学图，并引入逐连杆标记化，以保留细粒度的局部状态信息。我们提出一种物理基础偏置生成器，将结构先验直接注入注意力机制，其中包括运动学空间距离、动态接触状态、几何邻近性和解剖属性。这样，策略便能显式推理物理交互，而非从稀疏奖励中隐式学习。大量实验表明，在仅使用 ManipTrans 51% 参数量的情况下，PhysGraph 在操作精度和任务成功率方面显著优于基线 ManipTrans。此外，我们架构固有的拓扑灵活性在定性上展现了向未见过的工具/物体几何形状进行零样本迁移的能力，并且具有足够的通用性，可在三种机器人手（Shadow、Allegro、Inspire）上进行训练。

I. 引言

双手灵巧操作是机器人学长期追求的目标，旨在使机器人能够执行人类水平的多接触协同操作。尽管近年来基于学习的方法在灵巧操作方面取得了显著进展，双手工具使用仍然尤其具有挑战性，因为机器人必须保持稳定抓握、调节接触力，并协调双手，同时由工具介导双手与物体之间的交互。与平行颚夹爪相比，多指手具有更高的灵巧性，并能更好地适应物体几何形状的变化，但其状态空间和动作空间维度高，运动学拓扑也很复杂。因此，在实践中学习可靠的双手工具使用策略仍然十分困难。

本文提出，关键瓶颈之一在于控制策略如何表征并推理手、工具和物体的运动学结构。许多近期的灵巧操作策略将手、工具和物体的状态视为一个展平的单一向量 [1]–[4] ，并输入多层感知机（MLP），以隐式地发现 (i) 每只手内部的运动学依赖关系，(ii) 跨

手协调模式，以及 (iii) 手、工具和物体之间随时间变化的交互。模型必须从头开始隐式学习“食指指尖”与“食指指关节”在运动学上相邻。然而，双手工具使用本质上具有结构性：系统可以自然地描述为一张图，其节点（连杆）和边（关节/接触）定义了信息流和物理影响的可行路径。忽视这种结构会迫使策略仅凭稀疏奖励信号学习物理推理，往往导致其在具有挑战性的接触密集型任务中表现脆弱且不符合物理规律。

为解决这一问题，我们提出了 PhysGraph，这是一种用于双手灵巧手—工具—物体操作、以物理为基础的图变换器策略，如图 1 所示。我们不将状态展平，而是构建图结构，其中节点表示各个刚体（连杆、工具、物体），边表示物理耦合（关节、接触）。我们的方法引入了两项关键创新：（i）逐连杆标记化策略；（ii）以物理为基础的头特定变换器偏置生成器。首先，我们不将状态汇总为全局嵌入，而是将每个连杆的状态作为独立标记进行处理，从而保留细粒度的局部属性。其次，不同于使用抽象图距离的通用图变换器（例如 Graphormer [5]），我们将基于学习的头特定复合偏置直接注入注意力机制。复合偏置包括空间偏置（运动学链距离）、边偏置 （关节与接触）、几何偏置 （位置邻近性）以及 解剖先验（串联／协同运动学）。这些偏置项使我们的策略能够显式推理物理连接性与接触逻辑，并将注意力集中于正在接触的手指或协调关节。我们的主要贡献总结如下。

- 我们提出 PhysGraph，这是首个面向高自由度、具有挑战性的双手灵巧工具使用任务的图 Transformer 策略。它显式地将手部—工具—物体交互建模为连杆—关节—接触图，并处理逐连杆的标记化观测。
- 我们通过一种新颖的物理基础偏置生成器，将结构先验注入 PhysGraph Transformer。该生成器通过头特定掩码提供拓扑、边、距离和解剖学偏置。
- 针对具有挑战性的双手工具使用任务开展的大量实验表明，PhysGraph 在成功率和运动保真度方面显著优于最先进基线 ManipTrans [1]，而参数量仅为 ManipTrans 的***51%***；此外，它支持向未见过的工具／物体进行零样本迁移，并兼容多种主流灵巧机器人手（Shadow [6],Allegro [7], Inspire [8]）。

II. 相关工作

双手灵巧操作。近期研究主要聚焦于从人类示范中学习多样化技能。DexMimicGen [9]和 DexMan [4]利用自动化数据生成和基于视频的学习，解决交接和拾取与放置等基础双手任务。同样，Bi-DexHands [2] 和 DexMachina [3] 探索功能性重定向，将人手关键点映射到机器人手。 [10] 进一步通过优先考虑物体状态表示来提升协调能力。ManipTrans [1] 是最先进基线中的佼佼者，它采用残差学习架构，高效地将人类技能迁移到机器人手上。然而，这些方法通常依赖全局状态，往往无法捕捉我们所研究的复杂工具使用任务所需的细粒度动力学。

灵巧工具使用与仿真到现实挑战。尽管基于夹爪的工具使用已得到广泛研究，灵巧工具使用仍处于起步阶段。同期最先进的研究优先关注仿真到现实差距，而非任务复杂度。例如， [11] 成功展示了在真实硬件上对关节式工具进行手内操作。然而，为实现仿真到现实迁移，他们将任务简化为手与工具的单手任务，不涉及工具-物体交互。类似地， [12] 实现了面向人形机器人双手任务的稳健仿真到现实迁移，但只关注抓取与伸手行为，而非精确的工具-物体交互。DexWild [13]和 ClutterDexGrasp [14]利用可扩展的神经控制，解决了非结构化杂乱环境中的操作问题。我们采取互补的立场：我们不为实现仿真到现实迁移而简化任务，而是首先专注于在仿真中解决高复杂度的双手工具-物体交互问题（例如精细剪纸或切面包），并提供高保真教师策略。我们的 PhysGraph 应对高自由度双手工具-物体交互的全部复杂性，建立高性能教师策略，为未来的仿真到现实研究奠定基础。

从参考中学习。一种日益兴起的趋势是利用人类视频数据，为人形机器人和手合成物理上合理的运动。SkillMimic [15] 最近展示了从示范中学习高度动态技能——打篮球的能力。类似地，TokenHSI [16]提出了一种标记化方法，用于统一人-场景交互，这启发了我们采用逐连杆标记化策略。尽管全身学习表明，人形智能体无需向策略提供明确的参考状态输入，通常也能学会多种交互，但对于灵巧手而言，这种“无参考”探索并不足够，因为手指与物体之间的接触动力学复杂得多。如 DexMan [4],所强调的，多指手的高维动作空间和灵巧性需要更强的引导。因此，最先进水平的灵巧策略，包括 DexTrack[17],、DexMachina [3],、OmniGrasp [18], 和 ManipTrans[1], ，均采用参考条件学习，使训练变得可行。PhysGraph 遵循这一成熟范式，利用参考输入引导探索；但它通过具备拓扑感知能力的图架构，明确建模双手工具使用中的动态拓扑，而非采用侧重单智能体交互的扁平向量，从而与之区别开来。

结构化与基于图的策略学习。为应对高自由度系统的复杂性，人们对形态感知架构的兴趣日益浓厚。图神经网络在机器人控制中的应用可追溯至 NerveNet [19],，它将机器人的运动学

图显式嵌入策略网络，以创建形态感知控制器。Graphormer [5] 随后展示了注意力机制在图表示方面的强大能力。GCNT [20] 和 GET-Zero [21] 近期提出了 Graph-Transformer 策略，以适应不同的机器人形态，并证明图结构能够跨具身形态泛化。然而，这些方法将图边限定为静态且不可变的运动或运动学关系，例如静态的骨骼／肢体 [20]，[21] 和化学键 [5]。PhysGraph 基于这一“图故事”发展而来，并将其拓展至操作任务。我们引入了专为工具使用设计的动态接触边和解剖学偏置。尽管这些方法使用静态图表示形态，PhysGraph 却会根据接触状态动态重构图拓扑，使策略能够明确地将工具视为手部运动学链的临时延伸并进行推理。

III. PHYSGRAPH 概述

A. 问题陈述

我们研究双手灵巧手–工具–物体操作，其中两只关节式手必须在丰富接触动力学下协调配合，与工具和物体进行交互。我们将该系统建模为一个马尔可夫决策过程（MDP），由元组 $(\mathcal{S},\ \mathcal{A},\ p,\ r,\ \gamma,\ T)$ 定义，其中 $\mathcal{S}$ 是状态空间，$\mathcal{A}$ 是动作空间，$p(s_{t+1}|s_t,\ a_t)$ 是由刚体物理和接触诱导的转移密度函数，$r(s,\ a)$ 是奖励函数，$\gamma \in (0,1]$ 是折扣因子，T 是时域。

在每个时间步 t，智能体接收状态 $s_t \in \mathcal{S}$，该状态由基于状态的特征（本体感觉和交互相关信号）组成，同时接收源自人类示范的参考运动状态 s_t^{ref}。将参考状态作为目标十分重要，也是最先进水平的灵巧策略训练工作中的标准做法 [1]–[4], [10]。策略输出动作 $a_t \in \mathcal{A}$：

$$a_t \sim \pi_\theta(\cdot \mid s_t). \tag{1}$$

目标是学习一种策略，使其最大化初始状态分布下的期望折扣回报，以此表示我们的价值函数：

$$V^{\pi_\theta}(\boldsymbol{s}_0) = \mathbb{E}_{\tau\sim\pi_\theta}\left[\sum_{t=0}^{T-1}\gamma^t r(\boldsymbol{s}_t,\boldsymbol{a}_t)\right], \tag{2}$$

其中，期望是针对从策略中采样的长度为 T 的轨迹 τ 计算的。奖励 $r(s_t,a_t)$ 包含任务完成项和跟踪项。由于参考轨迹 s_t^{ref} 是状态的一部分，跟踪项会根据 [1]，计算机器人当前姿态与参考姿态之间的偏差。由此将问题构建为参考跟踪强化学习任务。我们的贡献在于设计了一种新颖的架构，以有效地参数化 π_θ。

B. PhysGraph 架构

图2 概述了 PhysGraph：我们首先构建运动学图，随后使用基于物理的复合偏置设计图 Transformer。

运动学图构建。我们设计了一种策略架构，将双手操作的物理结构建模为图 $\mathcal{G} = (\mathcal{V},\ \mathcal{E})$，其节点 $\mathcal{V}$对应刚体，例如手部连杆、工具和物体；边 $\mathcal{E}$ 则对应运动学（关节）或动力学（接触）交互。

标记化与 Transformer 主干网络。我们提出一种新颖的逐连杆标记化方法，以保留交互的局部结构：接触发生在手部连杆处，稳定性取决于远端关节，而误差校正沿运动学链传播。将刚体连杆表示为独立标记，使 Transformer 能够学习结构化的消息传递，例如手指连杆 $\leftrightarrow$、工具 $\leftrightarrow$ 和物体。令 $\{x_i\}_{i=1}^N$ 表示从结构化的逐连杆状态 s_t 中提取的逐标记输入，其中标记对应右手状态、右手逐连杆状态、左手状态、左手逐连杆状态、工具状态和物体状态。每个逐标记输入经分词器 f_i 处理后，再进行层归一化，从而嵌入到 d 维空间：

$$h_i^{(0)} = \mathrm{LN}(f_i(x_i)) \in \mathbb{R}^d, \qquad i = 1,\dots,N. \tag{3}$$

我们在前面添加一个可学习的策略令牌：

$$h_0^{(0)} = \mathbf{e}_{\text{POL}} \in \mathbb{R}^d, \tag{4}$$

并构造令牌矩阵：

$$H^{(0)} = [h_0^{(0)}, h_1^{(0)}, \dots, h_N^{(0)}]^\top \in \mathbb{R}^{(N+1)\times d}. \tag{5}$$

我们采用一个 L 层的 Transformer 编码器，将 $H^{(\ell-1)} \mapsto H^{(\ell)}$ 映射为 $\ell = 1, \dots, L$。对于每一层 ℓ 和注意力头 h，注意力逻辑值计算如下：

$$A^{(\ell,h)} = \frac{Q^{(\ell,h)}K^{(\ell,h)\top}}{\sqrt{d_h}} + B^{(h)}, \tag{6}$$

其中，$Q^{(\ell,h)}$、$K^{(\ell,h)} \in \mathbb{R}^{(N+1)\times d_h}$ 和 $B^{(h)} \in \mathbb{R}^{(N+1)\times(N+1)}$ 是特定于注意力头的复合偏置。注意力权重和输出遵循标准 Transformer 的做法 [22]。

基于物理的复合偏置。我们将每个注意力头的复合偏置定义为四种偏置中具有物理意义的组成部分的加权和：

$$B^{(h)} = \lambda_{\text{sp}}B_{\text{sp}}^{(h)} + \lambda_{\text{edge}}B_{\text{edge}}^{(h)} + \lambda_{\text{geo}}B_{\text{geo}}^{(h)} + \lambda_{\text{anat}}^{(h)}B_{\text{anat}}^{(h)}. \tag{7}$$

所有系数 λ 均为可学习参数，因此网络能够调整其对各项偏置和先验的依赖程度。这种设计实现了有益的区分：在 (6) 中，基于内容的点积项捕获一般的任务相关性，而 $B^{(h)}$ 则编码与任务无关的物理合理性。我们的核心贡献是对复合偏置项进行创新设计，下一节将对此进行详细介绍。

IV. 基于物理的偏置

图 3 详述了四种偏置：空间偏置、边偏置、几何偏置和解剖偏置。以下内容中，偏置均在物理图上的节点空间 中生成 $\mathcal{G}$，然后通过固定映射提升至令牌空间，以匹配令牌序列。

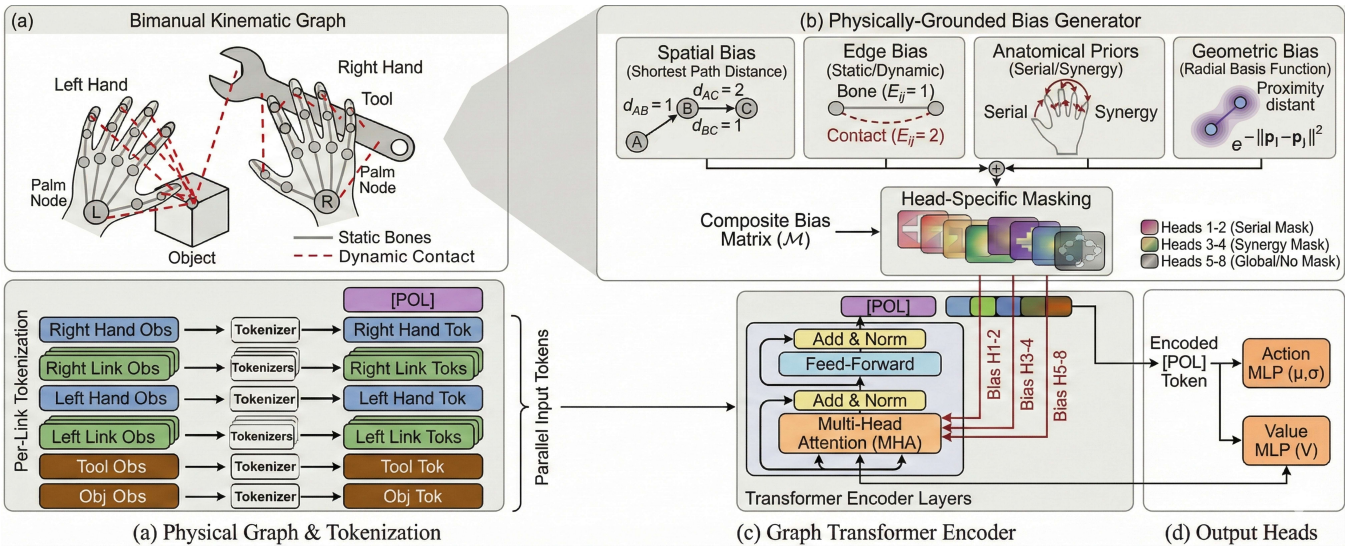

图 2： **PhysGraph 概览：（a）物理图与标记化：** 将双手工作空间建模为一张图，其中节点表示左手/右手、工具和物体的连杆，边表示运动学（关节）或动力学（接触）交互。将每个连杆的基于状态的多模态观测处理为并行输入词元。 **（b）物理基础偏置生成器：**该模块计算四种不同的偏置（详见图3），并将其汇总为复合偏置矩阵 $\mathcal{M}$。通过头特定掩码应用这些偏置，使不同注意力头能够关注特定的物理关系。 **（c）图 Transformer 编码器：** 标记化后的输入由 Transformer 编码器处理，其中多头注意力（MHA）受（b）中生成的偏置调制。 **（d）输出头：** 全局编码后的 $[POL]$标记被传递至多层感知机头，以预测策略动作分布（μ， σ）和价值函数（V）。

空间偏置（运动学图距离）。沿关节链的运动学影响主要是局部的：远端运动强烈依赖近端关节，而控制修正通常会沿相邻连杆传播。在稀疏奖励强化学习中，仅凭数据让注意力自行发现这种结构效率低下。空间偏置注入了一种先验，鼓励手部运动学图中距离较近的节点进行信息交互。

令 $d(u,\ v)$ 表示图上节点 u 与节点 v 在 $\mathcal{G}$ 上之间的最短图跳数距离。空间偏置定义为：

$$B_{\mathrm{sp}}^{(h)}(u,v) = b_{\mathrm{sp}}^{(h)}(\min(d(u,v), D_{\max})), \tag{8}$$

其中，$b_{\mathrm{sp}}^{(h)}(k) \in \mathbb{R},\ k \in \{0,1,\ldots,D_{\max}\}$ 是离散跳数距离上的头特定嵌入，$D_{\max}$ 是用于截断的上限。这实际上提供了一种可学习的“图位置偏置”，能够反映运动学拓扑。将距离截断至 D_{max} 可稳定训练，并防止偏置词表随图直径增长。

边偏置（关节 + 接触）。并非所有处于相同距离的节点对都同等重要。有些关系具有特殊意义：自环边保留节点身份，骨骼邻接体现刚性运动学耦合，而接触边则通过传递力和力矩来刻画瞬时物理约束。边偏置将这些语义边类型直接编码到注意力中。

我们定义离散边类型 $\tau(u,\ v)$，使其满足：$\tau(u,\ v) \in \{0,1,2,3\}$，分别表示（0）不连通、（1）骨骼邻接、（2）接触和（3）自身。具体而言，使用二元接触指示变量 $C_t(u,\ v)$：

$$\tau(u,v) = \begin{cases} 3 & u = v, \text{self} \\ 2 & C_t(u,v) = 1, \text{contact} \\ 1 & (u,v) \in \mathcal{E}_{\text{bone}}, \text{adjacency} \\ 0 & \text{disconnected.} \end{cases} \tag{9}$$

利用定义的边类型，我们将边偏置定义为：

$$B_{\text{edge}}^{(h)}(u,v) = b_{\text{edge}}^{(h)}(\tau(u,v)), \tag{10}$$

其中，$b^{(h)}{}_{\text{边}}(c) \in \mathbb{R},\ c \in \{0,1,2,3\}$ 是头特定的边嵌入。为提高效率，可以预先计算静态边索引并将其存储在缓冲区中，同时在运行时将动态接触作为稀疏子集添加。这样可以避免为这些边构建完整的 $N \times N$ 邻接矩阵，这对于实现高吞吐量非常重要。

在工具使用过程中，接触是主要的耦合机制：手指/手掌一旦接触工具，控制便取决于对局部身体状态、工具/物体运动及其响应的整合。我们提出的边偏置使注意力能够系统地访问这一通路。

几何偏置（位置邻近性）。仅凭运动学图拓扑无法建模交互局部性。在抓握过程中，来自不同手指（或不同手）的两个手指尖即使在运动学图中相距较远，也可能在空间上彼此相邻。同样，工具–物体和手–工具交互取决于欧几里得邻近性。我们提出一种几何偏置项，促使注意力聚焦于空间局部对，从而以可微分的方式近似物理交互可能性。

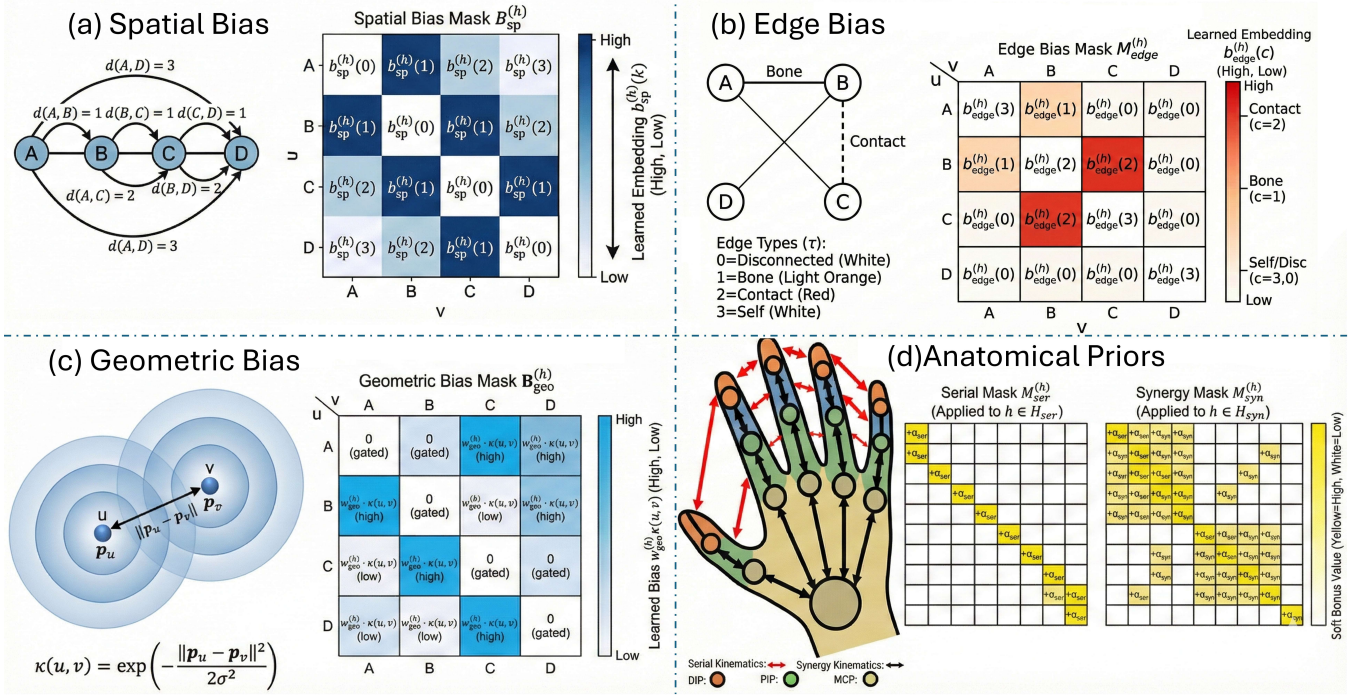

图 3：**基于物理的偏置的详细信息。（a）空间偏置：**通过将手部图中节点之间的最短路径距离 $d(u, v)$映射为学习得到的标量嵌入 $b_{sp}^{(h)}$，对手部图的拓扑结构进行编码，从而捕捉运动学链中的“跳数”距离。**（b）边偏置：**通过为连接两个节点的边类型 τ 分配不同的学习嵌入 $b_{edge}^{(h)}$，注入结构信息。**（c）几何偏置：**纳入笛卡尔空间中的空间邻近性。它利用径向基函数（RBF）核 $\kappa(u, v)$，根据节点位置之间的欧几里得距离 $\|\mathbf{p}_u - \mathbf{p}_v\|^2$ 对注意力进行加权。**（d）解剖先验：**对手部解剖运动学相关的知识进行编码。串联掩码（M_{ser}）突出显示单根手指沿运动学链的依赖关系，而协同掩码（M_{syn}）则促进不同手指之间具有相同解剖层级的对应连杆进行注意力交互。

设节点 u 的位置为 $p_u \in \mathbb{R}^3$。我们计算 RBF核：

$$\kappa(u, v) = \exp\left(-\frac{\|p_u - p_v\|_2^2}{2\sigma^2}\right), \tag{11}$$

带有可学习带宽 σ。我们使用学习得到的权重 $w_{\text{geo}}^{(h)}$ 将该邻近度标量投影为特定注意力头的偏置：

$$B_{\text{geo}}^{(h)}(u, v) = \mathbb{I}(d(u, v) > d_0)\, w_{\text{geo}}^{(h)}\, \kappa(u, v). \tag{12}$$

表达式的第一部分（即 $d(u, v) > d_0$）会对拓扑距离较远的节点对启用几何偏置，以避免与拓扑偏置冗余。

刚体交互从根本上由邻近关系驱动。RBF 项提供了一种平滑的归纳偏置：无论运动学距离有多远，笛卡尔位置相近的节点都应交换信息，从而提升对新出现的接触以及围绕工具和物体的双手协调的响应能力。

解剖先验（串联 + 协同）。手具有超出通用图的解剖结构。每根手指构成一条运动学串联链，而各手指在同一解剖层级表现出协同作用，例如掌指关节／近端指间关节的协调屈合／张开与相关运动。因此，强迫单个注意力头表征所有模式可能效率低下。我们转而对注意力头进行专门化：一部分关注串联运动学，一部分关注协同耦合，其余则保持全局推理。

我们在节点空间中定义了两个二元关系掩码：**(i)** 串联有效性 $S(u, v)$：对于运动学邻居附近的交互为真（例如 $d(u, v) \leq 1$）。**(ii)** 协同有效性 $Y(u, v)$：对于每只手内解剖对应组之间的关系为真（例如不同手指之间的指尖到指尖对应关系；掌指关节到掌指关节；近端指间关节到近端指间关节）。

我们分配注意力头子集：$\mathcal{H}_{\text{ser}}$、$\mathcal{H}_{\text{syn}}$、$\mathcal{H}_{\text{glob}}$。我们不采用硬掩码（$-\infty$），而是添加软奖励，以鼓励但不强制这些模式：

$$B_{\text{anat}}^{(h)}(u, v) = \begin{cases} \alpha_{\text{ser}}\, S(u, v) & h \in \mathcal{H}_{\text{ser}}, \\ \alpha_{\text{syn}}\, Y(u, v) & h \in \mathcal{H}_{\text{syn}}, \\ 0 & h \in \mathcal{H}_{\text{glob}}. \end{cases} \tag{13}$$

硬约束可能会损害灵巧性，而软奖励则能在保留模型灵活性的同时提供归纳偏置。

V. 实验

我们进行了大量实验，以评估 PhysGraph 从专家示范中学习复杂的双手手—工具—物体操作策略的能力。在第 V-A节中，我们在不同的动态、高接触密集型双手工具使用任务上，将我们的方法与最先进基线进行比较，并且

表 I：**Oakink2 数据集双手工具使用任务的定量结果 [23]**。SR：成功率；E t：工具与物体平移误差；E j：手部关节误差；E ft：指尖误差。ArtiMANO 手部。粉色表示最佳表现。

任务 ID	描述	工具（右）/物体（左）	方法	成功率（%）↑	E t（cm）↓	E j（cm）↓	E ft（cm）↓
083f7@0	用刀切面包	刀 / 面包	ManipTrans	55.69	1.31	2.42	2.01
			PhysGraph（无偏置）	82.84	1.02	2.26	1.92
			PhysGraph（完整版）	90.05	0.69	2.17	1.42
97055@0	舀入碗中	勺子 / 碗	ManipTrans	74.99	1.76	3.10	2.16
			PhysGraph（无偏置）	82.58	1.58	2.70	2.48
			PhysGraph（完整）	86.95	1.40	2.38	2.24
1292e@0	将液体从杯子倒入杯子	马克杯 / 马克杯	ManipTrans	45.60	6.78	3.50	3.35
			PhysGraph（无偏置）	46.58	6.21	3.65	4.07
			PhysGraph（完整版）	49.77	5.93	3.07	4.33
817fb@0	白板刷	刷子 / 白板	ManipTrans	50.05	1.55	2.17	2.24
			PhysGraph（无偏置）	58.84	1.60	1.70	1.94
			PhysGraph（完整）	62.24	1.33	1.54	1.86
e1fa6@0	剪纸	剪刀 / 纸	ManipTrans	15.16	1.18	3.31	2.25
			PhysGraph（无偏置）	27.16	1.43	2.92	1.97
			PhysGraph（完整）	35.84	1.35	2.52	1.83
9b728@0	涂抹在面包上	勺子 / 面包	Maniptrans	47.22	1.05	3.08	1.76
			PhysGraph（无偏置）	61.69	1.01	3.20	1.86
			PhysGraph（完整）	86.64	0.98	3.17	1.52

开展消融研究，以验证所提出的基于物理的偏置所作出的贡献。在第V-B节中，我们展示了对具有未见过的工具和物体几何形状的不同运动轨迹进行零样本迁移的能力。最后，我们在第 V-C节中展示 PhysGraph 在各种灵巧手具身形式中的可扩展性。

数据集。在训练和评估中，我们使用 OakInk2 数据集 [23]。双手操作本质上比单手任务更复杂，因为除了手、工具、物体之间复杂的接触序列外，还需要优势手（使用工具）与非优势手（稳定物体）之间进行精细协调。我们考虑了六项极具挑战性的任务，这些任务需要持续的动态接触和精细的施力。相关任务详见表I。

指标。为评估操作精度、任务遵循度和模仿保真度，我们考察以下指标：(i) 工具与物体的平均平移误差（E_t）和旋转误差（E_r）：轨迹展开中的工具/物体与专家参考轨迹之间三维位置的平均欧几里得距离（单位：cm）及角度差，反映任务遵循度；(ii) 平均手部关节误差（E_j）：所有手部活动关节上的均方误差（单位：cm），用于衡量总体运动学模仿质量；(iii) 平均指尖误差（E_{ft}）：专门评估五个指尖跟踪情况的欧几里得误差（单位：cm）这对于评估接触丰富的灵巧操作保真度至关重要；(iv) 成功率（SR）：判定为成功的试验所占百分比。只有当上述误差严格低于特定阈值时，轨迹才被视为成功。

基线。我们将 PhysGraph 与 ManipTrans[1]进行基准比较。据我们所知，ManipTrans 在双手灵巧手操作任务中表现最佳，因此我们将其作为唯一的基线。它采用基于多层感知机的残差学习架构，将所有观测拼接成一个状态向量。

实现细节。我们使用八个 Transformer 注意力头，其中两个采用串联掩码（$\mathcal{H}_{\text{ser}} \in \{1,2\}$），两个采用协同掩码（$\mathcal{H}_{\text{syn}} \in \{3,4\}$），其余注意力头采用全局掩码（$\mathcal{H}_{\text{glob}} \in \{5,6,7,8\}$）。我们使用 PPO [24] 算法训练策略，并使用比例微分控制器进行正向动力学计算。为确保对策略架构进行严格公平的比较，所有环境配置、奖励函数和物理动力学均与基线 ManipTrans完全相同。所有实验均在 Isaac Gym [25],中运行，每项任务同时使用 4096 个并行环境，时间步长为 1/60 秒，并配备 NVIDIA RTX 3090 GPU。为公平比较收敛情况，所有结果均取自最多 8k 个训练轮次内表现最佳的检查点；这一轮次数远超确保所有方法收敛所需的轮次。

A. 双手工具使用评估

定量结果。 表I 总结了评估结果。PhysGraph（完整版）在所有任务上的成功率均持续高于 ManipTrans，在切割和涂抹等高度接触敏感行为上的提升尤为显著。PhysGraph 在大多数设置下也展现出更高的保真度指标，表明交互轨迹更为稳定，且在接触转换附近的控制更为精细。重要的是，用于 ArtiMANO 的 ManipTrans 策略的参数数量为 ***3785187***，而 PhysGraph 的对应参数数量为 ***1950792***。这意味着 PhysGraph 以少得多的模型参数（$1950792/3785187 = \mathbf{51\%}$），充分体现了 PhysGraph 的有效性与效率。

定性结果。 图4 对滚动过程进行了可视化。与 ManipTrans 相比，PhysGraph 展现出更可靠的双手协调：ManipTrans 经常出现不自然的手部动作，导致工具掉落（刀、剪刀、马克杯），或因不合理的力

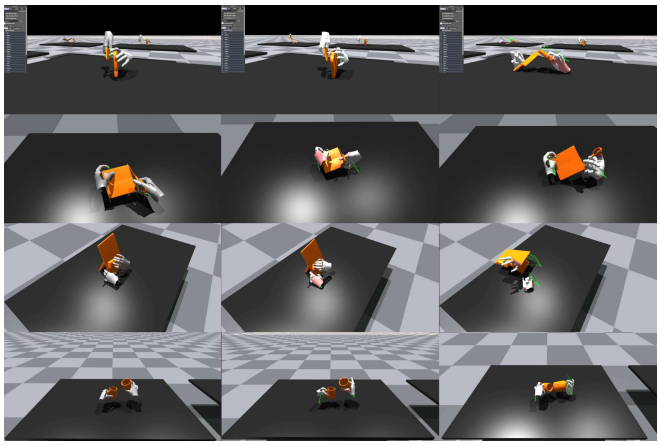

图 4：**Oakink2 数据集双手工具使用任务的定性结果**。从左至右：真实值、PhysGraph（我们的方法）、ManipTrans。从上至下：0837f@0、e1fa6@0、817fb@0、1292e@0。请参阅我们网站上的相应视频。

调节（刷子推动白板使其倾倒）。PhysGraph 产生高度自然的闭合抓握，并以物理上合理的力在整个任务过程中保持紧密接触。请查看我们网站上的相应视频以及补充材料以了解更多细节。

基于物理的偏置消融实验。为验证架构设计的作用，我们在表I中对基于物理的偏置组件进行了消融。值得注意的是，即使没有显式的物理偏置，PhysGraph（无偏置）在所有任务上的表现仍显著优于基于多层感知机的 ManipTrans。这从实证角度验证了逐连杆标记化策略的重要性：与会抹平结构信息的全局多层感知机不同，将双手机器人系统表示为离散的连杆标记能够保留局部几何保真度，即使采用标准 Transformer 主干网络，也能实现更优的特征学习。然而，移除偏置会导致所有评估指标大幅下降。无偏置变体难以维持动态工具使用运动学，迫使注意力机制仅凭数据发现潜在的运动学结构，导致样本效率极低。这表明，所提出的物理偏置项能够提供关键指导，使注意力头立即沿有效的物理路径进行信息交互。实验结果支持了我们的初衷：显式编码运动学拓扑、边语义、几何邻近性和解剖学信息，能够为富接触双手工具操作提供有效的归纳偏置。

B. 零样本策略迁移

我们通过将冻结策略部署到未见过的工具/物体及不同任务实例上，评估 PhysGraph 实现零样本迁移的能力，且不进行微调。如图5所示，面对实例变化，PhysGraph 表现出更强的鲁棒性，能够保留正确的交互结构，并更可靠地完成任务。如图5 所示，仅针对大型砍刀和长方形

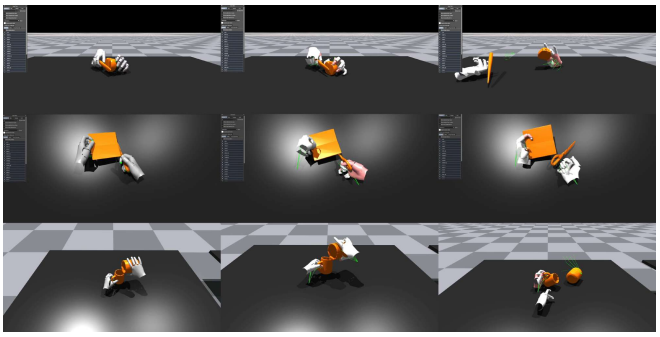

图5：**零样本策略迁移** 无需微调。从左到右：真实值、PhysGraph（本文方法）、ManipTrans。从上到下（训练序列 → 部署序列）：0837f@0→ 9fc3e@0（砍刀 → 水果刀；面包 → 苹果），e1fa6@0 → 66c7f@0，1292e@0 → b9695@0。

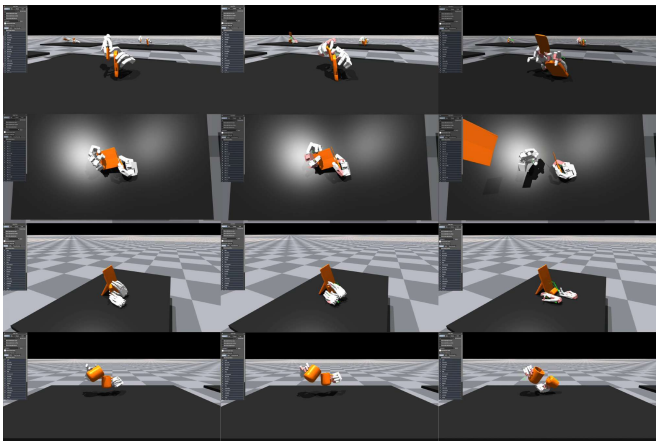

图6：**多种具身形态上的定性结果**。 从左到右：真实值、PhysGraph（本文方法）、ManipTrans。从上到下：0837f@0（Shadow）、e1fa6@0（Allegro）、817fb@0（Shadow）、1292e@0（Shadow）。

面包训练的策略可以无缝迁移到小得多的水果刀和球形苹果上。这表明，我们基于物理的偏置具有内在的空间不变性；由于我们显式编码的是相对空间邻近性，而非依赖绝对状态坐标，因此模型在训练中未见过的特定实例几何形状上，也能实现更出色的泛化。

C. 具身兼容性验证

我们证明，PhysGraph 架构兼容多种机器人手部拓扑结构。我们在多种灵巧手具身形式上训练并评估 PhysGraph（Shadow [6], Inspire [8], Allegro [7]）。表 II 显示，与基线相比，PhysGraph 在不同具身形式上始终能提升成功率和交互保真度。图 6 展示了定性结果：PhysGraph 能够维持物理上合理的接触点和交互力，表现优于基线。由于从 MANO 重定向会给非 MANO 具身形式带来额外的不匹配，这些结果也表明，该方法能够应对现实的

表 II：**跨具身定量结果**。粉色表示最佳表现。

任务 ID	Hand	方法	SR ↑	E.t ↓	E.j ↓	E.ft ↓
083f7@0	Inspire	ManipTrans	45.07	2.02	3.83	2.92
		PhysGraph	62.14	1.62	2.67	2.3
	Shadow	ManipTrans	22.34	1.58	3.54	2.67
		PhysGraph	54.77	1.25	3.03	2.26
	Allegro	ManipTrans	19.67	3.60	4.56	3.17
		PhysGraph	29.34	2.95	4.30	2.51
97055@0	Inspire	ManipTrans	52.95	2.16	4.20	2.91
		PhysGraph	78.45	1.17	3.24	1.97
	Shadow	ManipTrans	18.37	2.52	3.29	2.71
		PhysGraph	32.44	2.26	2.04	2.84
	Allegro	ManipTrans	23.37	3.51	4.86	3.25
		PhysGraph	36.28	2.85	4.08	2.41

轨迹噪声和具身特定差异。

指尖误差（E_{ft}）异常。在表I 和 II中，我们观察到，在某些任务（例如倾倒瓶子和剪纸）中，PhysGraph 的指尖误差 E_{ft} 高于 ManipTrans。这是由于真实演示数据存在局限，而非策略失效。我们使用[26]提取专家轨迹，方法是遵循 ManipTrans [1],将人体关键点对齐到 MANO 模型，再将其重定向到机器人具身形式。然而，这一重定向过程不可避免地引入运动学噪声并降低接触保真度，常常导致机器人指尖与工具/物体之间出现不符合物理规律的相互穿透（例如，图6和视频中，指尖穿入白板和马克杯）。由于缺乏物理理解，ManipTrans 会盲目过拟合这些有误的目标姿态，以最小化 E_{ft} 。另一方面，PhysGraph 基于物理的偏置会惩罚不符合物理规律的状态，使策略偏离带噪声的真实关键点，从而探索物理上合理的动作。剪纸任务中的一个例子如图（4，第二行）所示：重定向后的演示中，拇指从剪刀柄外侧握住手柄，这对人类而言并不合理。然而，PhysGraph 成功找到了稳定抓握方式，拇指像人手一样正确穿过手柄。PhysGraph 偏离有缺陷的演示以遵循物理约束，尽管 E_{ft}有所下降，仍实现了更高的 SR。

D. 局限性与未来工作

与 ManipTrans 和其他最先进的灵巧操作方法 [2]–[4], [10],类似，PhysGraph 的关键局限在于，策略在推理时需要参考轨迹。虽然以参考轨迹为条件能够稳定高维控制动作的学习，并应对复杂的接触动力学，但在缺少参考轨迹时，也会限制其部署。未来的工作应消除或减少对参考演示的依赖。

VI. 结论

我们提出了 PhysGraph，用于双手灵巧手–工具–物体操作。我们的结果表明，显式的物理结构对于学习具有挑战性的丰富接触技能至关重要。更广泛地说，这项工作指出，面向关节式物理系统的 Transformer策略可以受益于更能体现其底层拓扑、运动学和动力学结构的标记化与注意力机制。

参考文献

[1] K. Li、P. Li、T. Liu、Y. Li 和 S. Huang，“ManipTrans：通过残差学习实现高效的灵巧双手操作迁移”，发表于 CVPR，2025年。[2] B. Zhou、H. Yuan、Y. Fu 和 Z. Lu，“从人类示范中学习多样化的双手灵巧操作技能”，发表于 AAAI，2026年。[3] Z. Mandi、Y. Hou、D. Fox、Y. Narang、A. Mandlekar 和 S. Song，“Dexmachina：面向双手灵巧操作的功能重定向”，2025年。 [在线]。可访问：https://arxiv.org/abs/2505. 24853[4] J. Hsieh、K.-H. Tu、K.-H. Hung 和 T.-W. Ke，“Dexman：从人类视频和生成视频中学习双手灵巧操作”， 2025 [ ] 2510 08475 。[5] 在线。可访问：https://arxiv.org/abs/。C. Ying、T. Cai、S. Luo、S. Zheng、G. Ke、D. He、Y. Shen 和 T.-Y. Liu，“Transformer 在图表示方面真的表现很差吗？”，发表于 NeurIPS，2021年。[6] “ShadowRobot”，https://www.shadowrobot.com/ dexterous-hand-series，2005年。[7] “Allegro Hands”，https://www.allegrohand.com，2013年。[8] “Inspire Hands”，https://en.inspire-robots.com/product-category/the-dexterous-hands，2019年。[9] Z. Jiang、Y. Xie、K. Lin、Z. Xu、W. Wan、A. Mandlekar、L. Fan 和 Y. Zhu，“DexMimicGen：通过模仿学习实现双手灵巧操作的自动化数据生成”，发表于 ICRA，2025年。[10] Y. Chen、C. Wang、Y. Yang 和 K. Liu，“从人类运动数据中进行以物体为中心的灵巧操作”，发表于 CoRL，2024年。[11] S. Atar、D. Huang、F. Richter 和 M. Yip，“使用灵巧机器人手对关节式工具进行手内操作并实现仿真到现实迁移”，2025 [ ] 2509 23075 。在线。可访问：[12] https://arxiv.org/abs/。T. Lin、K. Sachdev、L. Fan、J. Malik 和 Y. Zhu，“面向人形机器人的基于视觉的灵巧操作仿真到现实强化学习”，发表于 CoRL，2025年。[13] T. Tao、M. K. Srirama、J. J. Liu、K. Shaw 和 D. Pathak，“DexWild：面向真实环境机器人策略的灵巧人类交互”，RSS，2025年。[14] Z. Chen、Q. Yan、Y. Chen、T. Wu、J. Zhang、Z. Ding、J. Li、Y. Yang 和 H. Dong，“ClutterDexGrasp：用于杂乱场景中通用灵巧抓握的仿真到现实系统”，CoRL，2025年。[15] Y. Wang、Q. Zhao、R. Yu、H. W. Tsui、A. Zeng、J. Lin、Z. Luo、J. Yu、X. Li、Q. Chen、J. Zhang、L. Zhang 和 P. Tan，“SkillMimic：从示范中学习篮球交互技能”，发表于 CVPR，2025年。[16] L. Pan、Z. Yang、Z. Dou、W. Wang、B. Huang、B. Dai、T. Komura 和 J. Wang，“TokenHSI：通过任务标记化统一生成物理人-场景交互”，发表于 CVPR，2025年，第 5379–5391 页。[17] X. Liu、J. Adalibieke、Q. Han、Y. Qin 和 L. Yi，“DexTrack：从人类参考动作中实现面向灵巧操作的可泛化神经跟踪控制”，发表于 ICLR，2025年。[18] Z. Luo、J. Cao、S. Christen、A. Winkler、K. Kitani 和 W. Xu，“OmniGrasp：使用仿真人形机器人抓握多样化物体”，发表于 NeurIPS，2024年。[19] T. Wang、R. Liao、J. Ba 和 S. Fidler，“NerveNet：使用图神经网络学习结构化策略”，发表于 ICLR，2018年。 [在线]。可访问：https://api.semanticscholar.org/CorpusID:65051725 [20] Y. Luo、M. Yao 和 X. Xiao，“GCNT：面向形态无关强化学习的基于图的 Transformer 策略”，发表于 IJCAI，2025年。[21] A. Patel 和 S. Song，“GET-Zero：面向零样本具身泛化的图具身 Transformer”，发表于 ICRA，2025年。[22] A. Vaswani、N. Shazeer、N. Parmar、J. Uszkoreit、L. Jones、A. N. Gomez、L. Kaiser 和 I. Polosukhin，“Attention Is All You Need”，发表于 NeurIPS，2017年。

[23] X. Zhan、L. Yang、Y. Zhao、K. Mao、H. Xu、Z. Lin、K. Li 和 C. Lu，“OAKINK2：复杂任务完成中的双手物体操作数据集”，发表于 *CVPR*，2024年6月，第 445–456 页。[24] J. Schulman、F. Wolski、P. Dhariwal、A. Radford 和 O. Klimov，“近端策略优化算法”，2017年。 [在线]。可访问：https://arxiv.org/abs/1707 06347 。V. Makoviychuk、L. Wawrzyniak、Y. [25] Guo、M. Lu、K. Storey、M. Mack- lin、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。[26] Y. Qin、W. Yang、B. Huang、K. Van Wyk、H. Su、X. Wang、Y.-W. Chao 和 D. Fox，“AnyTeleop：一种通用的基于视觉的灵巧机器人臂-手遥操作系统”，发表于 *RSS*，2023年。