三维扩散策略：通过简单三维表示实现可泛化的视觉运动策略学习

3D Diffusion Policy: Generalizable Visuomotor Policy Learning via Simple 3D Representations

Yanjie Ze, Gu Zhang, Kangning Zhang, Chenyuan Hu, Muhan Wang, Huazhe Xu
RSS 2024 · arXiv:2403.03954v7
底本：https://arxiv.org/abs/2403.03954v7
许可：CC BY 4.0 · https://creativecommons.org/licenses/by/4.0/
整理与中文翻译：本地文献库 S16，2026-11-25

> 本册为完整中文改编译本，涵盖正文、附录、所有原图和表格；参考文献保留英文。翻译与重排为本次改动，原作者不为本译本背书。逐段英文对照、来源坐标与原始资产位于同目录“中英对照”。

阅读方式

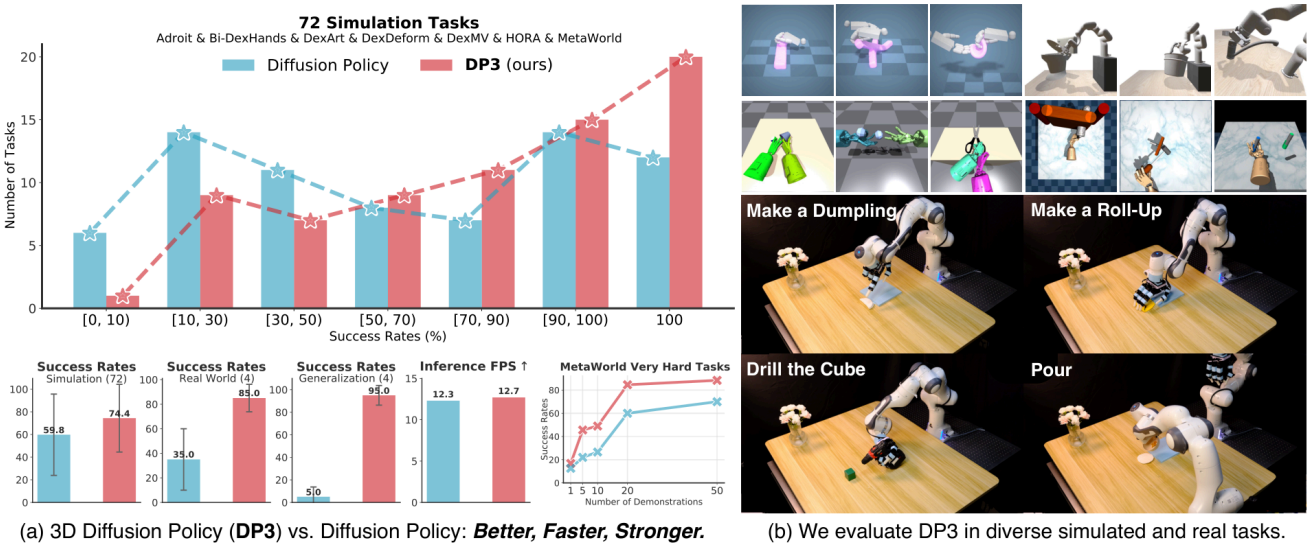
先读问题定义与方法，再用实验设计校准结论。每段小字编号对应原文 HTML 锚点；可在 paper.md 与 source_map.json 中检索同一编号。数学公式以原文 MathML 排印，图表数据不作重算或改写。

英文原文 17 页。以下页码是中文重排页码，与原文页码不同；来源标记中的 p. 指英文 PDF 页码。

p.1 · abstract

模仿学习能高效地教会机器人灵巧技能，但稳健、可泛化地学习复杂技能通常需要大量人类示范。为解决这一挑战，我们提出三维扩散策略（3D Diffusion Policy，DP3），把三维视觉表示的能力融入条件动作生成模型——扩散策略。核心设计是用高效点编码器从稀疏点云提取紧凑三维表示。在 72 项仿真任务中，DP3 仅用 10 条示范就能处理多数任务，相对基线提升 24.2%。在 4 项真实机器人任务中，每项仅给 40 条示范，平均成功率达到 85%，并展现空间、视角、外观、实例等多方面泛化。有趣的是，真实实验中 DP3 很少违反安全要求，而基线常需人工干预。广泛评估凸显了三维表示对真实机器人学习的重要性。代码与视频见 3d-diffusion-policy.github.io。

p.1 · S0.F1



(a) 3D Diffusion Policy (**DP3**) vs. Diffusion Policy: ***Better, Faster, Stronger.*** (b) We evaluate DP3 in diverse simulated and real tasks.

Fig. 1: 3D Diffusion Policy (DP3) is a visual imitation learning algorithm that marries 3D visual representations with diffusion policies, achieving surprising effectiveness in diverse simulation and real-world tasks, with a practical inference speed.

图 1：DP3 将三维视觉表示与扩散策略结合，在多种仿真与真实任务中取得显著效果，并保持实用的推理速度。

I 引言

p.1 · S1.p1

模仿学习能高效地教会机器人多种运动技能，如抓取 [68,60,82]、足式运动 [40]、灵巧操作 [1,16]、类人移动操作 [54]、移动操作 [57,12]。以图像或深度图等高维视觉观测为输入的视觉模仿学习，减少了任务专用状态估计的需求，因此广受关注 [10,60,82,11,20]。

p.1 · S1.p2

但视觉模仿学习的通用性往往以大量示范为代价 [16,10,11]。例如先进的 Diffusion Policy [10] 每项真实任务需要人工采集 100–200 条示范；长时程、易失败的采集过程可能持续数天。一种解法是在线学习 [16]，通过环境交互和从专家示范学得的奖励函数继续改进策略。但真实在线学习涉及安全、自动复位、人工干预和硬件成本。因此，如何使离线模仿学习用尽可能少的示范学到稳健、可泛化技能，是实际机器人学习的基础问题。

p.1 · S1.p3

我们提出简单有效的 DP3，将三维表示与扩散策略结合。它用直接、高效的 MLP 编码器把稀疏采样点云压成紧凑三维表示，再以该表示与机器人位姿为条件，将随机噪声去噪为连贯动作序列，同时利用三维模态固有的空间理解与扩散模型的表达能力。

p.2 · S1.p4

为全面评估，我们构建了来自 7 个领域的 72 项多样仿真任务及 4 项真实任务，其中包含可变形物体的困难灵巧操作。虽然概念简单，DP3 相对二维扩散策略与其他基线展现了若干优势：

p.2 · S1.p5

1. 效率与效果：示范和训练步数显著更少，精度仍更高。2. 泛化：三维特性促进空间、视角、实例、外观等方面的泛化。3. 安全部署：真实任务中，DP3 很少发出异常命令；基线却常出现不可预料行为，可能损伤硬件。

p.2 · S1.p6

我们进一步分析三维表示。有趣的是，BCRNN [35]、IBC [11] 等也受益于三维输入，但提升不及 DP3。DP3 持续优于深度图、体素等其他三维模态，以及 PointNeXt [46]、Point Transformer [84] 等点编码器。这说明成功不仅源于采用三维表示，也来自精心设计。

p.2 · S1.p7

贡献有四点：（1）提出少示范、多方面泛化的有效视觉运动策略 DP3；（2）为减小基准和任务选择造成的偏差，在广泛仿真与真实任务上验证通用性；（3）全面分析表示，说明简单点云优于复杂三维表示，且相对其他策略骨干更适合扩散策略；（4）仅凭 40 条示范即可用灵巧手操作真实可变形物体，说明少量人类数据也能处理复杂高维任务。

p.2 · S1.p8

DP3 强调在真实机器人学习中结合三维表示与扩散策略的力量。代码：https://github.com/YanjieZe/3D-Diffusion-Policy。

II 相关工作

II-A 机器人学中的扩散模型

p.2 · S2.SS1.p1

扩散模型逐步将随机噪声变为数据样本，已在高保真图像生成中取得成功 [23,63,51,62]。凭借强表达性，它们被用于强化学习 [70,2]、模仿学习 [10,39,50,72,64,41]、奖励学习 [25,37]、抓取 [71,66,61]、运动规划 [52,27]。本文遵循 [10,39]，把视觉运动策略表示为条件扩散模型，即扩散策略。以往主要以图像和状态为条件，本文率先引入三维条件。

II-B 视觉模仿学习

p.2 · S2.SS2.p1

模仿学习通常依赖专家示范的大量观测-动作对，使机器人高效获得类人技能。真实环境物体状态难以精确估计，图像等视觉观测因此成为实用替代。二维图像策略长期占主导 [38,11,10,35,16,56,68,15]，但三维的重要性正得到更多认可 [60,82,80,14,13,28,69]。

p.2 · S2.SS2.p2

PerAct [60]、GNFactor [82]、RVT [14]、ACT3D [13]、NeRFuser [74] 等三维策略已推动低维控制，但有两项主要挑战。（1）设置不够实用：通过关键帧位姿提取，将模仿学习变为预测加规划，虽有效却不适合高维控制。（2）推理慢：复杂结构使 PerAct 仅达 2.23 FPS，难以处理需要稠密命令的动态任务。相关工作 3D Diffuser Actor [28] 达 1.67 FPS，主要与语言词元注意力和任务设置不同有关。原文脚注特别说明：速度依赖相机数、像素大小、扩散步数和预测时域；两文原始设置不同，3D Diffuser Actor 作者在语言条件多任务

CALVIN 上用 DP3 代码复配，使 DP3 同样具备语言注意力，两相机反投影，同用 NVIDIA 2080 Ti。3D Diffuser Actor 推理 600 ms、预测 6 步；DP3 为 581 ms、预测 4 步。因此在该 CALVIN 设置，前者控制频率高于 DP3。相较这类工作，我们旨在构建通用且快速的三维策略，同时处理高维与低维控制。

II-C 灵巧技能学习

p.3 · S2.SS3.p1

使机器人具备类人操作技能是长期目标。强化学习已让灵巧手学会倒水 [47,81]、开门 [49,21,8]、旋转物体 [44,76,78,45]、重定向物体 [18,7,6]、转笔 [33]、抓工具 [1]、交接 [83,24]、搭乐高 [9]。模仿学习提供另一条路；DIME [3]、DexMV [47] 用重定向把人手运动转成机器人动作，从人类视频中学习。本文不同于这些专门设计：我们展示，只改进模仿学习算法本身，就能用少量示范获得复杂技能。

III 方法

p.3 · S3.p1

给定少量包含复杂技能轨迹的专家示范，目标是学习视觉运动策略 $\pi : \mathcal{O} \mapsto \mathcal{A}$，将视觉观测 $o \in \mathcal{O}$ 映射到动作 $a \in \mathcal{A}$，既复现技能，也超出训练数据泛化。DP3 主要有两部分：（a）感知：用高效点编码器把环境点云变为视觉特征；（b）决策：以表达能力强的 Diffusion Policy [10] 为骨干，在三维特征条件下生成动作序列。总览见图 2，后文分别介绍。

p.4 · S2.F2

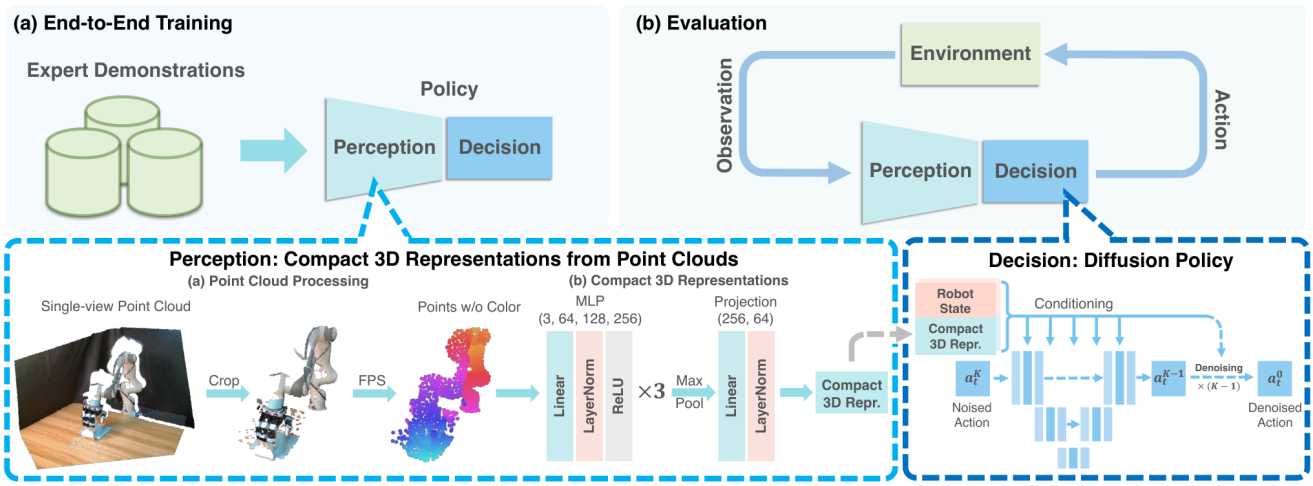


Fig. 2: Overview of 3D Diffusion Policy (DP3). Above: In the training phase, DP3 simultaneously trains its perception module and decision-making process in an end-to-end manner using expert demonstrations. During evaluation, DP3 determines actions based on visual observations from the environment. Below: DP3 perceives its environment through single-view point clouds. These are transformed into compact 3D representations by a lightweight MLP encoder. Subsequently, DP3 generates actions conditioning on these 3D representations and the robot's states, using a diffusion-based backbone.

图 2：DP3 总览。上：训练阶段根据专家示范端到端联合训练感知与决策；评估时从环境视觉观测决定动作。下：单视角点云经轻量 MLP 编码为紧凑三维表示，再与机器人状态一起作为扩散骨干的条件，生成动作。

III-A 一个启发性示例

p.3 · S3.SS1.p1

先用 MetaWorld Reach [77] 说明泛化能力：夹爪需要准确到达指定目标点。为检验不仅拟合训练点、也能泛化到新场景，图 3 可视化三维空间中的 · 训练点与 · 成功测试点。仅 5 个训练点，DP3 即能到达分布在三维空间的目标；二维 Diffusion Policy [10] 与 IBC [11] 主要局限于近似平面区域，BCRNN [35] 无法覆盖空间。该例说明 DP3 在数据有限时的泛化与效率优势。

p.4 · S3.F3

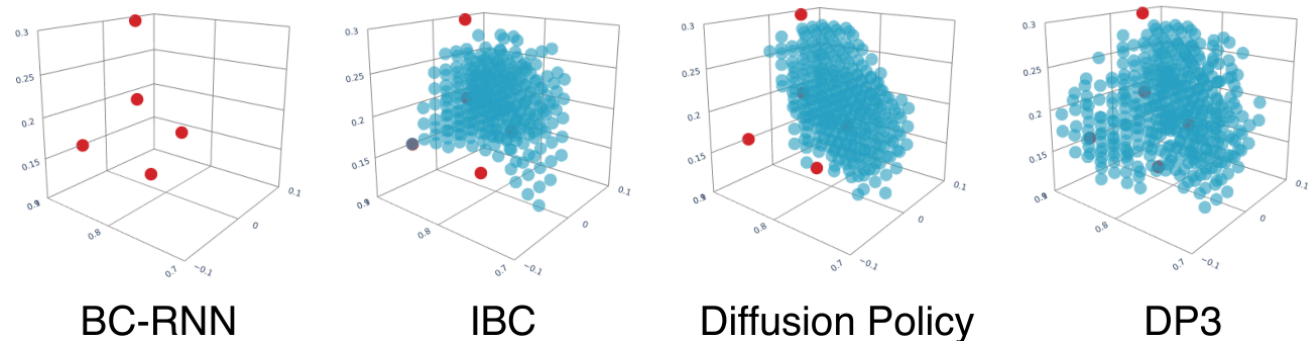


Fig. 3: Generalization in 3D space with few data. We use MetaWorld Reach as an example task, given only 5 demonstrations (visualized by ·). We evaluate 1000 times to cover the 3D space and visualize the · successful evaluation points. DP3 learns the generalizable skill in 3D space; Diffusion Policy and IBC [11] only succeed in partial space; BC-RNN [35] fails to learn such a simple skill with limited data. Number of successful trials from left to right: 0 / 285 / 327 / 415 .

图 3：少数据下的三维空间泛化。MetaWorld Reach 仅用 5 条示范（·），评估 1000 次覆盖空间，画出 · 成功点。DP3 学到三维可泛化技能；Diffusion Policy、IBC [11] 仅在部分空间成功，BC-RNN [35] 在有限数据下连这一简单技能也未学会。由左到右成功次数为 0 / 285 / 327 / 415。

III-B 感知

p.3 · S3.SS2.p1

DP3 的所有任务都仅使用单视角相机，不同于围绕机器人布置多相机的既有工作 [10,18]，主要出于真实任务的实用性。

p.3 · S3.SS2.p2

点云表示三维场景。三维场景可表示为 RGB-D、点云、体素 [7]、隐式函数 [36]、三维高斯 [29] 等。DP3 采用稀疏点云；表 IV 消融表明，相比 RGB-D、深度图、体素等显式表示，它更高效。

p.3 · S3.SS2.p3

仿真与真实环境均从单相机取得 84×84 深度图，再用相机内外参转换为点云。为更好地泛化外观，不使用颜色通道。

p.4 · S3.T1

	Adroit	Bi-DexHands	DexArt	DexDeform	DexMV	HORA	MetaWorld	MetaWorld	MetaWorld	MetaWorld	平均
算法 \ 任务	(3)	(6)	(4)	(6)	(2)	(1)	Easy (28)	Medium (11)	Hard (6)	Very Hard (5)	(72)
DP3	68.3	70.2	68.5	87.8	99.5	71.0	90.9	61.6	31.7	49.0	$74.4_{\pm 29.9}$ (↑ 24.2%)
Diffusion Policy	31.7	61.3	49.0	90.5	95.0	49.0	83.6	31.1	9.0	26.6	$59.8_{\pm 35.9}$

TABLE I: Main simulation results. Averaged over 72 tasks, DP3 achieves 24.2% relative improvement compared to Diffusion Policy, with a smaller variance. Success rates for individual tasks are in Appendix -C .

表 I：仿真主要结果。72 项任务平均，DP3 相比 Diffusion Policy 相对提高 24.2%，且方差更小。逐任务成功率见附录 C。

p.4 · S3.T2

	Adroit		MetaWorld				DexArt				
算法 \ 任务	锤击	开门	转笔	装配	拆卸	棍推	笔记本电脑	水龙头	马桶	桶	平均
DP3	$100_{\pm 0}$	$62_{\pm 4}$	$43_{\pm 6}$	$99_{\pm 1}$	$69_{\pm 4}$	$97_{\pm 4}$	$83_{\pm 1}$	$63_{\pm 2}$	$82_{\pm 4}$	$46_{\pm 2}$	74.4
Diffusion Policy	$48_{\pm 17}$	$50_{\pm 5}$	$25_{\pm 4}$	$15_{\pm 1}$	$43_{\pm 7}$	$63_{\pm 3}$	$69_{\pm 4}$	$23_{\pm 8}$	$58_{\pm 2}$	$46_{\pm 1}$	44.0
BCRNN	$0_{\pm 0}$	$0_{\pm 0}$	$9_{\pm 3}$	$3_{\pm 4}$	$32_{\pm 12}$	$45_{\pm 11}$	$3_{\pm 3}$	$1_{\pm 0}$	$5_{\pm 5}$	$0_{\pm 0}$	9.8

	Adroit		MetaWorld				DexArt				
算法 \ 任务	锤击	开门	转笔	装配	拆卸	棍推	笔记本电脑	水龙头	马桶	桶	平均
BCRNN+3D	$8_{\pm 14}$	$0_{\pm 0}$	$8_{\pm 1}$	$1_{\pm 5}$	$11_{\pm 6}$	$0_{\pm 0}$	$29_{\pm 12}$	$26_{\pm 2}$	$38_{\pm 10}$	$24_{\pm 11}$	14.5
IBC	$0_{\pm 0}$	$0_{\pm 0}$	$9_{\pm 2}$	$0_{\pm 0}$	$1_{\pm 1}$	$16_{\pm 2}$	$3_{\pm 2}$	$7_{\pm 1}$	$14_{\pm 1}$	$0_{\pm 0}$	5.0
IBC+3D	$0_{\pm 0}$	$0_{\pm 0}$	$10_{\pm 1}$	$18_{\pm 9}$	$3_{\pm 5}$	$50_{\pm 6}$	$1_{\pm 1}$	$7_{\pm 2}$	$15_{\pm 1}$	$0_{\pm 0}$	10.4

TABLE II: Comparing DP3 with more baselines in simulation. We include IBC, BCRNN, and their 3D variants, termed as IBC+3D and BCRNN+3D. The 3D variants use our DP3 Encoder for a fair comparison.

表 II：更多仿真基线，包括 IBC、BCRNN 及 IBC+3D、BCRNN+3D。为公平比较，三维版本均用 DP3 Encoder。

p.3 · S3.SS2.p4

点云处理。深度转换的点云可能包含桌面、地面等冗余点，我们裁剪，仅保留边界框内的点。

p.3 · S3.SS2.p5

进一步采用最远点采样 FPS [42] 下采样。相比均匀采样，它更充分覆盖三维空间、降低随机性。实践中，仿真与真实任务下采样到 512 或 1024 个点已足够。

p.3 · S3.SS2.p6

紧凑编码。点云经轻量 MLP 转为紧凑三维表示（图 2）。DP3 Encoder 由三层 MLP、用于汇聚点特征的最大池化和投影头组成，层间插入 LayerNorm 稳定训练 [22]。最终特征 v 仅 64 维。原文称池化为“order-equivariant”；就最终集合特征而言，最大池化对点排列不变。表 V 显示，这一简单编码器甚至优于预训练 PointNeXt [46]，与 [20] 的观察一致：视觉运动控制中，适当设计的小编码器可能优于预训练大编码器。

III-C 决策

p.3 · S3.SS3.p1

条件动作生成。DP3 决策模块是条件去噪扩散模型 [23,10,39]，以三维特征 v 与机器人位姿 q 为条件，将高斯噪声去噪成动作 a。从 a^K 开始，去噪网络 ϵ_θ 迭代 K 次，逐渐把随机噪声 a^K 变为无噪动作 a^0：

$$a^{k-1} = \alpha_k \left(a^k - \gamma_k \epsilon_\theta \left(a^k, k, v, q \right) \right) + \sigma_k \mathcal{N}(0, \mathbf{I}),$$

（1）。其中 $\mathcal{N}(0, \mathbf{I})$ 为高斯噪声，α_k、γ_k、σ_k 为依赖噪声调度器的 k 的函数。这也称反向过程 [23]。

p.3 · S3.SS3.p2

训练目标。训练去噪网络 ϵ_θ 时，从数据集中随机取数据点 a^0，施加扩散过程 [23] 得到第 k 次迭代的含噪数据 ϵ^k。目标是预测加入原始数据的噪声：

$$\mathcal{L} = \mathrm{MSE}\left(\epsilon^k, \epsilon_\theta(\bar{\alpha}_k a^0 + \bar{\beta}_k \epsilon^k, k, v, q) \right),$$

（2）。其中 $\bar{\alpha}_k$、$\bar{\beta}_k$ 为一次加噪的噪声调度 [23]。实现采用样本预测，见下一段；本段保留原文通用噪声预测表述。

p.4 · S3.SS3.p3

实现细节。采用卷积网络扩散策略 [10] 和 DDIM [62] 调度器。为更好生成高维动作，用样本预测替代 epsilon 预测；训练 100 步、推理 10 步。MetaWorld 较简单，训练 1000 个 epoch；其余仿真和真实任务训练 3000 个 epoch。DP3 和所有基线 batch 均为 128。

IV 仿真实验

IV-A 实验设置

p.4 · S4.SS1.p1

仿真基准。尽管仿真日趋逼真 [34,73,65,85]，与真实环境仍有明显差距 [80,30,7]，这强调了真实实验的重要性，也要求大规模多样仿真任务以进行更科学比较。我们收集 7 个领域共 72 项任务，覆盖双手操作 [8]、可变形物体 [31]、铰接物体 [5] 等困难任务，也包括平行夹爪操作 [77]。任务基于 MuJoCo [65]、Sapien [73]、IsaacGym [34]、PlasticineLab [26]，避免基准受制于单一仿真器。MetaWorld 按 [55] 分难度。概览见表 III，三维观测见图 4。

p.5 · S4.F4

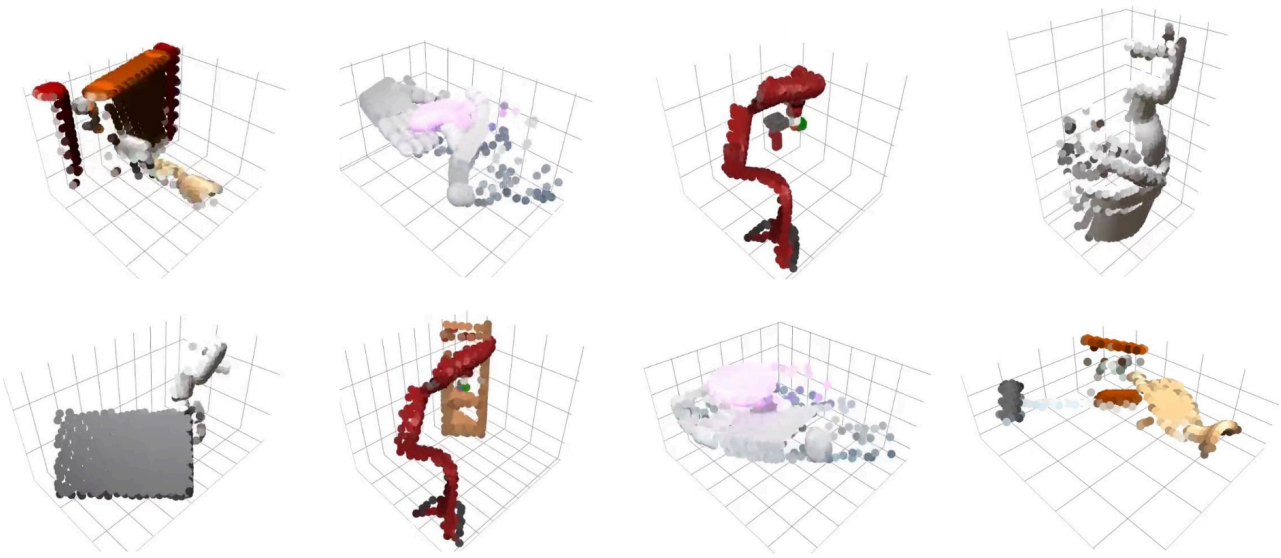


Fig. 4: 3D visual observations in simulation. We sample some simulated tasks and show the downsampled point clouds in these tasks.

图 4：仿真中的三维视觉观测。选取若干任务，展示下采样点云。

p.5 · S4.SS1.p2

专家示范。DexDeform 使用人工遥操；MetaWorld 使用脚本策略；其他领域使用强化学习智能体：Adroit 用 VRL3 [67]，其他用 PPO [53]。从智能体生成成功轨迹，确保各模仿算法使用相同示范。专家成功率见附录 C。

p.5 · S4.SS1.p3

基线。本文重点是三维模态在扩散策略中的意义，因此主要基线是图像扩散策略 [10]，简称 Diffusion Policy。还比较 IBC [11]、BCRNN [35] 及其三维版本；它们在困难任务中效果有限，因此只评估 10 项（表 II）。所有二维和三维方法使用相同图像、深度分辨率，以保证公平。

p.5 · S4.SS1.p4

指标。每项实验运行 3 个种子 $0, 1, 2$。每种子每训练 200 个 epoch 评估 20 回合，取最高的 5 个成功率求平均，再报告 3 种子的均值和标准差。

p.5 · S4.T3

仿真基准（72 项）

领域	机器人	物体	仿真器	ActD	任务数	示范数
Adroit	Shadow	刚体/铰接	MuJoCo	28	3	10
Bi-DexHands	Shadow	刚体/铰接	IsaacGym	52	6	10

DexArt	Allegro	铰接	Sapien	22	4	100
DexDeform	Shadow	可变形	PlasticineLab	52	6	10
DexMV	Shadow	刚体/流体	Sapien	30	2	10
HORA	Allegro	刚体	IsaacGym	16	1	100
MetaWorld	夹爪	刚体/铰接	MuJoCo	4	50	10

真实机器人基准（4 项）

任务	机器人	物体	ActD	示范数	描述
卷起	Allegro	可变形	22	40	包卷橡皮泥
包饺子	Allegro	可变形	22	40	包裹橡皮泥后用手指捏褶
电钻触碰	Allegro	刚体	22	40	抓电钻触碰方块
倾倒	夹爪	刚体	7	40	抓碗、倾倒并放回

TABLE III: Task suite of DP3, including Adroit [49] , Bi-DexHands [8] , DexArt [5] , DexDeform [31] , DexMV [47] , HORA [44] , MetaWorld [77] , and our real robot tasks. ActD: the highest action dimension for the domain. #Demo: Number of expert demonstrations used for each task in the domain. Art: articulated objects. Deform: deformable objects.

表 III：DP3 任务套件，包括 Adroit [49]、Bi-DexHands [8]、DexArt [5]、DexDeform [31]、DexMV [47]、HORA [44]、MetaWorld [77] 与自建真实任务。ActD 是领域最高动作维数；#Demo 是各任务专家示范数；Art 为铰接物体，Deform 为可变形物体。

IV-B 效率与效果

p.5 · S4.SS2.p1

DP3 的高效性体现在以下方面（原文先称三个方面，实际列出四项）。（1）精度高：概览见图 1(a)、逐领域见表 I。近 30 项成功率超过 90%，Diffusion Policy 不到 15 项；DP3 无任务低于 10%，后者有 10 余项低于 10%。多数任务仅用 10 条示范。（2）学习效率：为保证收敛各算法训练 3000 epoch，但 DP3 通常约 500 epoch 已收敛（图 5），二维策略更慢或收敛到次优。（3）示范扩展：Adroit 中两者均有效，DP3 用更少示范达到相近精度。Assembly、Disassemble 等高于简单等级的 MetaWorld 任务，在示范充分时 DP3 更精确，说明三维模态对某些任务可能是必要的。（4）推理有竞争力：图 1 中 DP3 略快于二维策略。与三维策略较慢的通常看法 [60,82,72] 相反，稀疏点云与紧凑表示让它实现高效推理。

p.6 · S4.F5

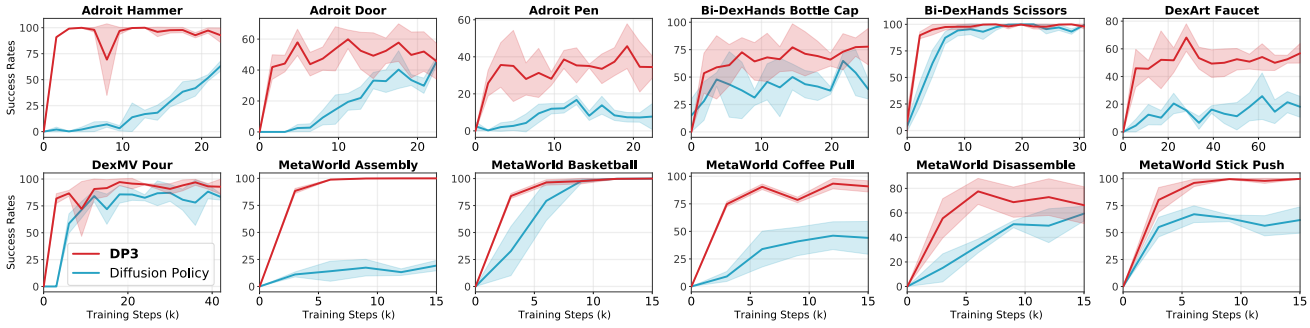


Fig. 5: Learning efficiency. We sample 12 simulation tasks and show the learning curves of DP3 and Diffusion Policy. DP3 demonstrates a rapid convergence towards high accuracy. In contrast, Diffusion Policy exhibits a slower learning progress and achieves notably lower convergence in most tasks.

图 5：学习效率。选取 12 项仿真任务展示学习曲线，DP3 快速收敛到高精度，Diffusion Policy 学习较慢，多数任务最终精度明显较低。

p.6 · S4.F6

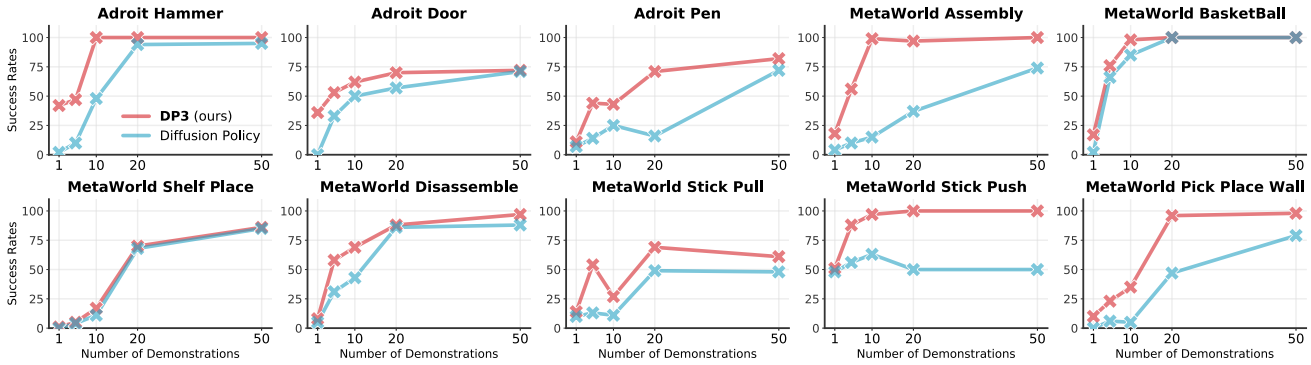


Fig. 6: Efficient scaling with demonstrations. We sample 10 simulation tasks and train DP3 and Diffusion Policy with an increasing number of demonstrations. DP3 addresses all these tasks well and generally improves the accuracy with more demonstrations. Diffusion Policy also scales well on some tasks while still falling short of accuracy.

图 6：随示范数高效扩展。在 10 项仿真任务中增加示范数训练。DP3 均能良好完成，通常示范越多精度越高；Diffusion Policy 在部分任务也能扩展，但精度仍不足。

IV-C 消融

p.5 · S4.SS3.p1

选取 6 项任务：Adroit 的 Hammer（H）、Door（D）、Pen（P），以及 MetaWorld 的 Assembly（A）、Disassemble（DA）、Stick-Push（SP），覆盖高低维控制，每项仅用 10 条示范。表格用这些缩写。

p.6 · S4.SS3.p2

三维表示选择。我们比较 RGB-D、深度、体素，以及除机器人位姿外包含物体状态、目标和机器人速度的真值状态。RGB-D、深度图用与 Diffusion Policy 相同的图像编码器；体素用 [7] 的 VoxelCNN。表 IV 显示这些表示均逊于 DP3。RGB-D 与深度图相近，但都不及点云，说明正确使用深度信息很重要。点云与真值状态竞争力相当，提示点云可能帮助从示范学到最优策略。

p.6 · S4.T4

表示	H	D	P	A	DA	SP	平均
真值状态	99 ± 2	61 ± 2	**44 ± 3**	94 ± 1	**72 ± 7**	91 ± 8	76.8
点云	**100 ± 0**	**62 ± 4**	43 ± 6	**99 ± 1**	69 ± 4	**97 ± 4**	78.3
图像	48 ± 17	50 ± 5	25 ± 4	15 ± 1	43 ± 7	63 ± 3	40.7
深度图	39 ± 15	49 ± 1	12 ± 3	15 ± 4	15 ± 2	62 ± 3	32.0
RGB-D	57 ± 14	47 ± 5	14 ± 2	15 ± 3	14 ± 1	61 ± 3	34.7
体素	54 ± 5	33 ± 3	18 ± 2	10 ± 2	17 ± 1	62 ± 6	32.3

TABLE IV: Ablation on 3D representations. We replace the visual observation and the corresponding encoder in DP3 to evaluate different 3D representations.

表 IV：三维表示消融。替换 DP3 的视觉输入和对应编码器，评估不同三维表示。

p.6 · S4.SS3.p3

点编码器选择。比较 DP3 Encoder 与 PointNet [42]、PointNet++ [43]、PointNeXt [46]、Point Transformer [84]，也纳入预训练 PointNet++、PointNeXt。表 V 中，这些复杂及预训练模型均不及 DP3 Encoder。

p.6 · S4.T5

编码器	H	D	P	A	DA	SP	平均
DP3 Encoder	**100** ± 0	**62** ± 4	**43** ± 6	**99** ± 1	**69** ± 4	**97** ± 4	78.3
PointNet	46 ± 8	34 ± 8	14 ± 4	0 ± 0	0 ± 0	0 ± 0	15.7
PointNet++	0 ± 0	0 ± 0	13 ± 3	0 ± 0	0 ± 0	0 ± 0	2.2
PointNeXt	0 ± 0	0 ± 0	14 ± 3	0 ± 0	0 ± 0	0 ± 0	2.3
Point Transformer	0 ± 0	0 ± 0	6 ± 5	0 ± 0	0 ± 0	0 ± 0	1.0
PointNet++ (pre-trained)	5 ± 9	19 ± 12	17 ± 6	0 ± 0	0 ± 0	0 ± 0	6.8
PointNeXt (pre-trained)	0 ± 0	36 ± 13	17 ± 6	0 ± 0	0 ± 0	0 ± 0	8.8

TABLE V: Ablation on point cloud encoders. We replace DP3 Encoder with other widely used encoders, including PointNet [42] , PointNet++ [43] , PointNeXt [46] , and Point Transformer [84] . We also include the pre-trained encoders.

表 V：点云编码器消融。替换 DP3 Encoder 为常见点编码器，并包括预训练版本。

p.6 · S4.SS3.p4

逐步修改 PointNet。为解释性能差异，逐步把 PointNet 改为 DP3 风格。表 VI 显示，T-Net 和 BatchNorm 是效率的主要阻碍。去掉二者后平均成功率 72.3，已可媲美 DP3 Encoder 的 78.3。一个可能解释是，控制任务使用固定相机，不需 T-Net 的特征变换。再降低高维特征维数，几乎不损性能（72.5 → 72.3），却提高速度。其他编码器失败原因留待未来。

p.7 · S4.T6

编码器	卷积	含 T-Net	含 BatchNorm	1024 维	平均
PointNet	✓	✓	✓	✓	15.7
	✗	✓	✓	✓	15.7
	✓	✗	✓	✓	16.0
	✗	✗	✓	✓	26.0
	✗	✓	✓	✗	18.2
性能转折	✓	✗	✗	✓	72.5
	✗	✗	✓	✗	19.8
	✗	✓	✗	✗	26.8
	✗	✗	✗	✗	72.3

TABLE VI: Gradually modifying a PointNet to a DP3-style encoder. Conv: use convolutional layers or linear layers. w/ T-Net: with or without T-Net. w/ BN: with or without BacthNorm layers. 1024 Dim: set feature dimensions before the projection layer to be 1024 or 256. Average success rates for 6 ablation tasks are reported.

表 VI：逐步将 PointNet 改为 DP3 风格。Conv：卷积层或线性层；w/ T-Net、w/ BN：是否使用 T-Net、BatchNorm；1024 Dim：投影前特征为 1024 还是 256 维。报告 6 项消融任务平均成功率。

p.7 · S4.SS3.p5

设计选择。除三维表示外，表 VII 的细节也贡献效果：（a）裁剪点云显著提高精度；（b）LayerNorm 有助于跨任务稳定训练 [22,4]；（c）采样器采用样本预测收敛更快（图 7）；（d）投影头降维加速，不损精度；（e）去掉颜色保证稳健外观泛化；（f）低维控制中 DPM-solver++ [32] 与 DDIM 相当，但前者难以处理高维控制。

p.7 · S4.F7

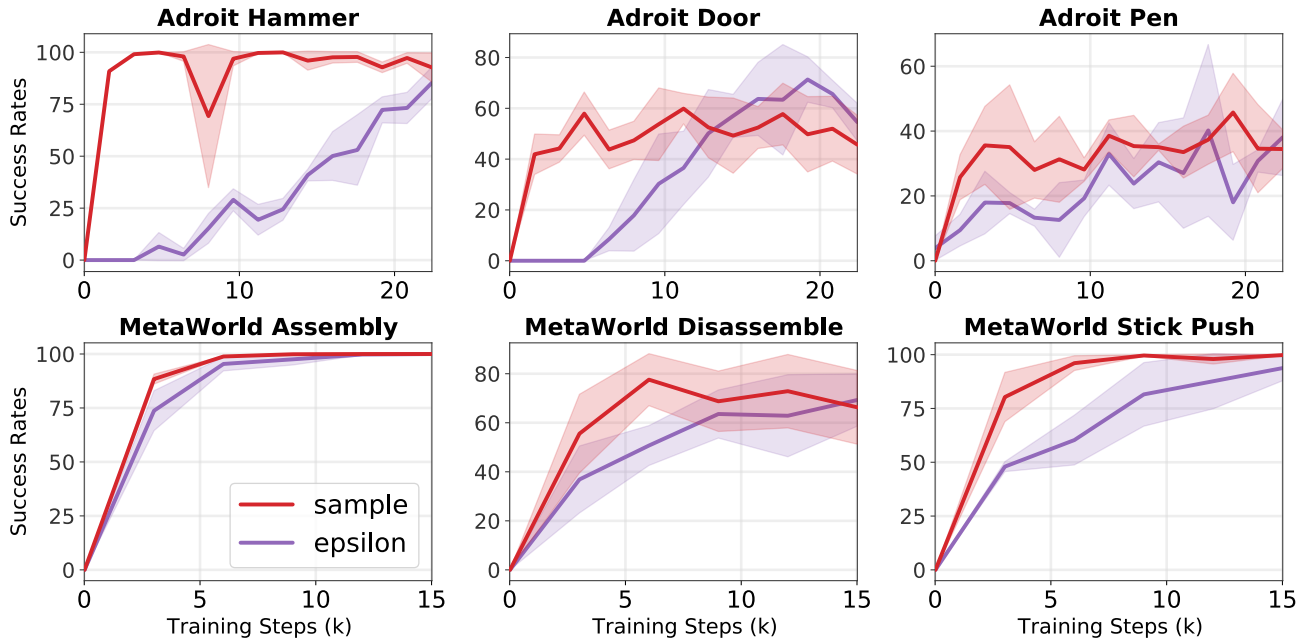


Fig. 7: Learning curves of DP3 with sample prediction and epsilon prediction. With sample prediction, DP3 generally converges faster, while epsilon prediction is also competitive.

图 7：样本预测与 epsilon 预测的学习曲线。样本预测通常收敛更快，但 epsilon 预测也有竞争力。

p.7 · S4.T7

设计	H	D	P	A	DA	SP	平均
DP3	**100 ± 0**	62 ± 4	43 ± 6	**99 ± 1**	69 ± 4	97 ± 4	78.3
去裁剪	98 ± 1	**69 ± 3**	14 ± 1	19 ± 9	32 ± 6	40 ± 2	45.3
去 LayerNorm	**100 ± 0**	56 ± 4	44 ± 3	96 ± 2	51 ± 3	91 ± 5	73.0
去样本预测	68 ± 3	67 ± 8	37 ± 12	96 ± 2	58 ± 9	76 ± 9	67.0
去投影头	**100 ± 0**	61 ± 2	**47 ± 3**	**99 ± 1**	60 ± 8	**99 ± 2**	77.7
含颜色	**100 ± 1**	67 ± 3	46 ± 4	76 ± 8	**75 ± 5**	68 ± 3	72.0
DDIM → DPM-solver++	12 ± 4	9 ± 5	26 ± 5	93 ± 3	58 ± 6	92 ± 14	48.3

TABLE VII: Ablation on design choices in DP3. Most of the design choices would not affect the accuracy but bring other benefits such as appearance generalization by removing color.

表 VII：DP3 设计消融。多数选择不直接影响精度，却带来其他好处，如去颜色改善外观泛化。

V 真实世界实验

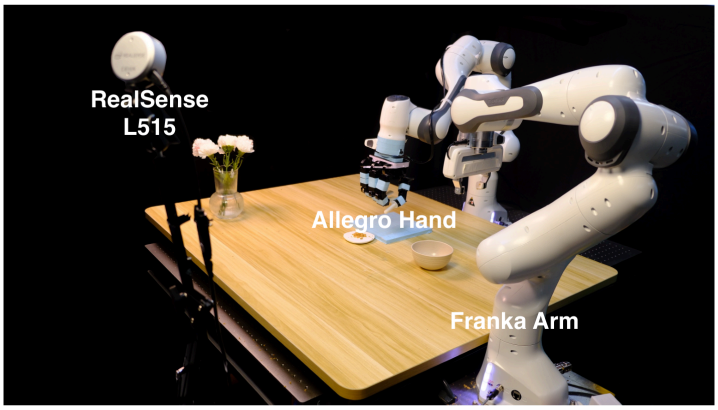
V-A 实验设置

p.7 · S5.SS1.p1

真实机器人基准。用 Allegro 手和夹爪两类机器人评估 4 项任务，通过一台 RealSense L515 获取视觉观测。任务见图 10、表 III，配置与日常物体见图 8。任务简述如下。



(a) Robots and objects used in DP3.



(b) Real-world experiment setup.

Fig. 8: (a) Our robots and objects. (b) Our real-world experiment setup. We use an Allegro hand and a gripper based on Franka arms and include diverse everyday objects in our manipulation tasks. A RealSense L515 camera is applied to capture visual observations.

图 8：（a）机器人与物体；（b）真实实验配置。Franka 机械臂安装 Allegro 手或夹爪，操作多种日常物体，RealSense L515 采集观测。

p.9 · S5.F10

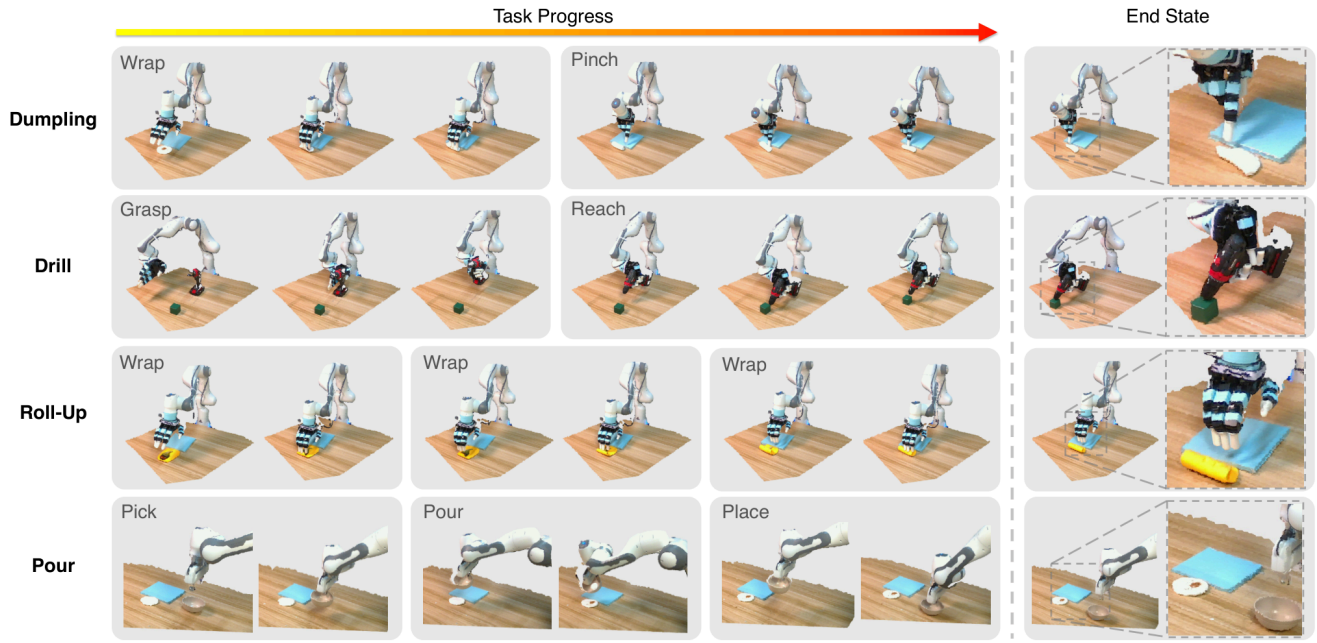


Fig. 10: Our real robot benchmark consists of 4 challenging tasks. The Allegro hand is required to make a Dumpling , Drill the cube, and make a Roll-Up . The gripper is required to Pour dried meat floss in the bowl. Each task contains multiple stages. We visualize the point clouds of the collected trajectories.

图 10：四项困难真实任务。Allegro 手完成包饺子、电钻触碰、卷起，夹爪完成倾倒肉松。每项均有多个阶段；图中展示采集轨迹的点云。

p.7 · S5.SS1.p2

1. Roll-Up（卷起）：Allegro 多次包卷橡皮泥形成卷。2. Dumpling（包饺子）：先包裹橡皮泥，再捏出褶。3. Drill（电钻触碰）：抓起电钻，移向绿方块，用电钻碰方块。4. Pour（倾倒）：夹爪抓碗，移向橡皮泥，将碗中干肉松倒出，再放回桌面。任务随机化见图 9：卷起与包饺子改变橡皮泥形状及上方物体外观；电钻与倾倒改变方块、电钻、碗的位置。

p.8 · S5.F9

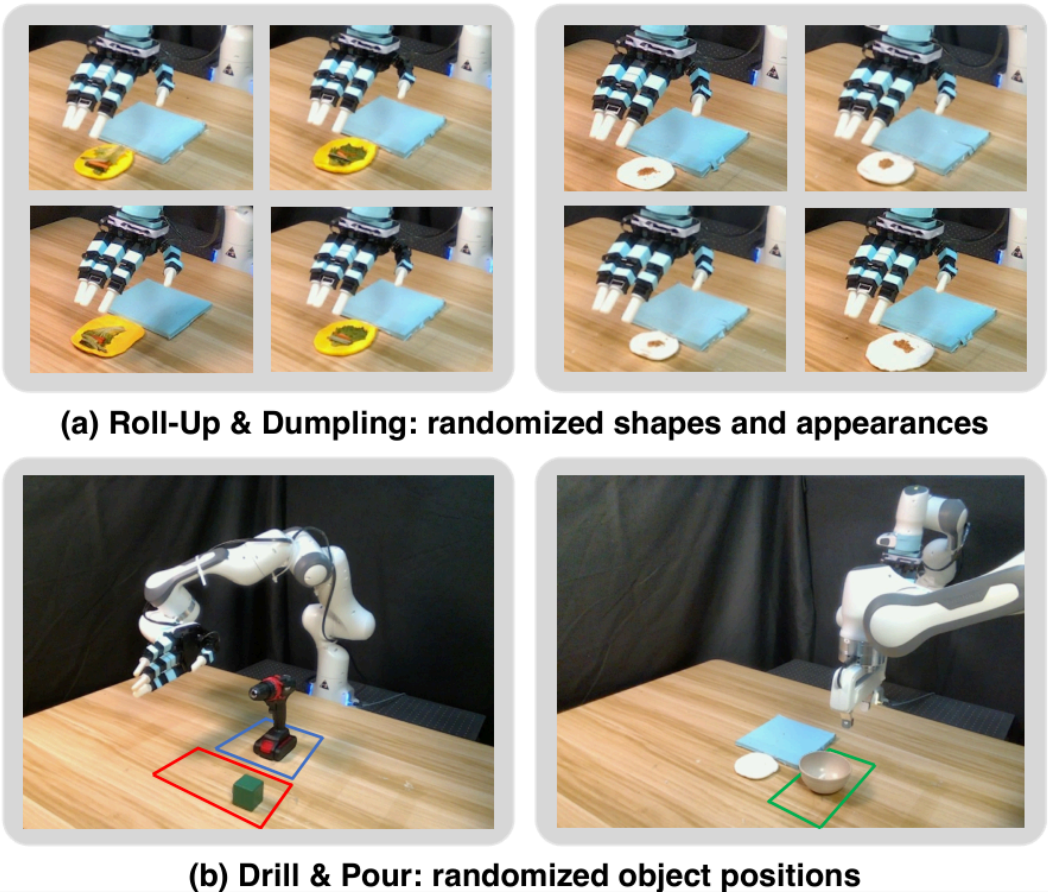


Fig. 9: Randomization in collected demonstrations for real-world tasks. Roll-Up : The shape of the plasticine and the vegetables on it varies in each trajectory. Dumpling : The shape of the plasticine and the distribution of the meat floss on it are different in each trajectory. Drill : The red and blue rectangles respectively mark the range of positions where the cube and drill can be placed. Pour : The green rectangle marks the range of positions of the bowl.

图 9：真实任务示范的随机化。卷起：每条轨迹橡皮泥形状及上方蔬菜不同；包饺子：橡皮泥形状与肉松分布不同；电钻：红、蓝框分别标出方块与电钻的位置范围；倾倒：绿框标出碗的位置范围。

p.8 · S5.SS1.p3

多指手任务特意展示其相对平行夹爪的优势：卷起与包饺子无需 RoboCook [59] 那样的额外工具即可包裹橡皮泥；真实电钻又大又重，夹爪难以抓取。

p.8 · S5.SS1.p4

专家示范由人遥操采集：键盘控制 Franka 臂与夹爪，基于视觉重定向 [48,17] 用人手控制 Allegro。任务多阶段，涉及复杂多指机器人与可变形物体，采集很费时，因此每项仅提供 40 条示范。

p.8 · S5.SS1.p5

基线。仿真显示图像、深度图扩散策略仍较强，因此选为真实实验基线。不同模态见图 11。

p.9 · S5.F11

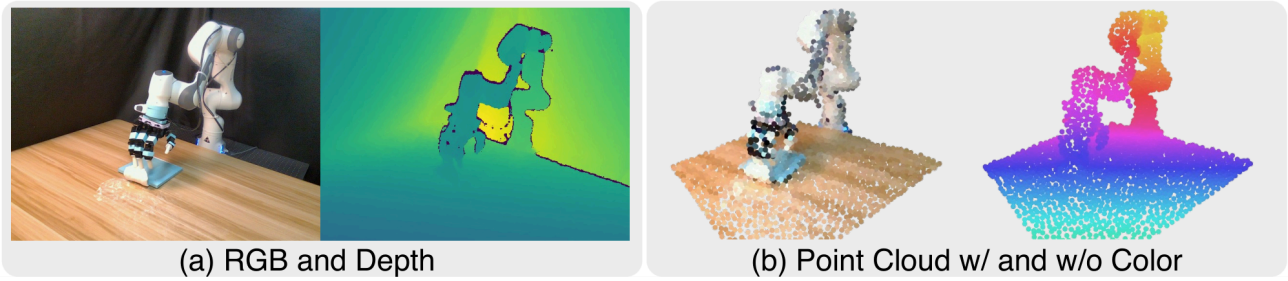


Fig. 11: Different vision modalities in the real world, include images, depths, and point clouds.

图 11：真实环境中的图像、深度图、点云视觉模态。

V-B 效果

p.8 · S5.SS2.p1

结果见表 VIII。与仿真一致，每项仅 40 条示范，DP3 仍能高成功率处理全部任务。图像与深度图策略虽平均精度较低，却各有强项：图像策略擅长电钻，却在卷起完全失败；深度策略在卷起达到 40% 成功率。

p.8 · S5.T8

真实机器人	卷起	包饺子	电钻触碰	倾倒	平均
Diffusion Policy	0	30	70	40	35.0 ± 25.0
Diffusion Policy (Depth)	40	20	10	10	20.0 ± 12.2
DP3	**90**	**70**	**80**	**100**	85.0 ± 11.2

TABLE VIII: Main results for real robot experiments. Each task is evaluated with 10 trials.

表 VIII：真实机器人主要结果。每项任务评估 10 次。

V-C 泛化

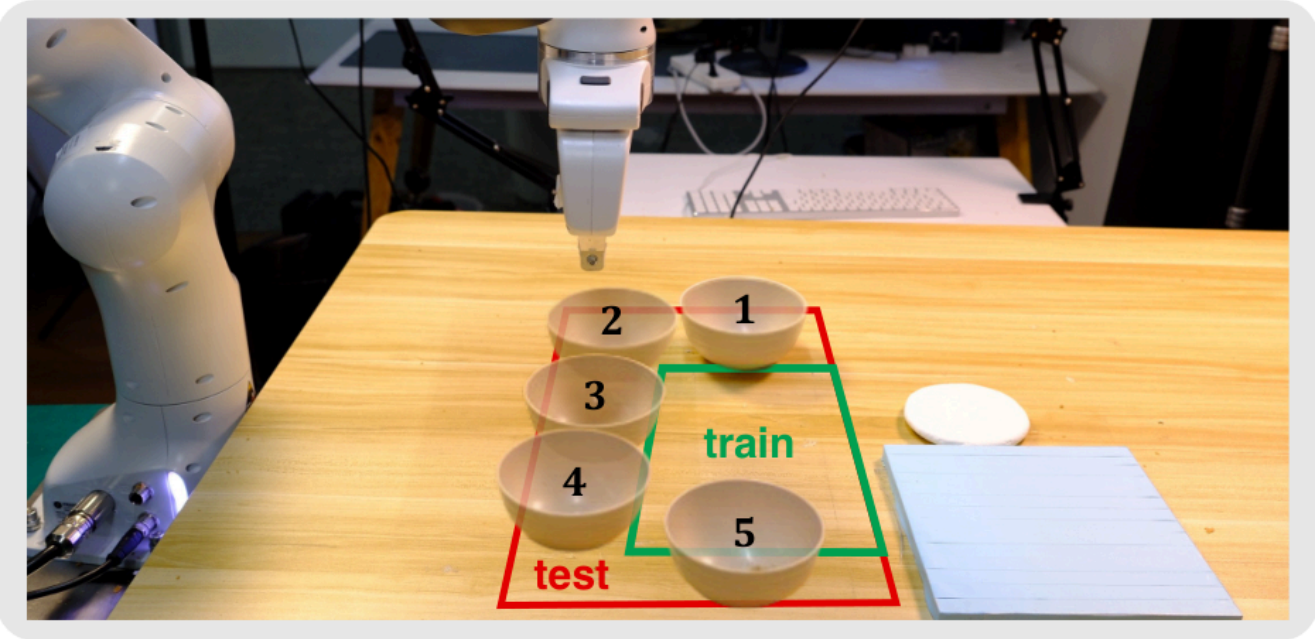
p.8 · S5.SS3.p1

除有效处理全部任务外，DP3 在真实环境展现强泛化，我们分为四方面逐项说明。

p.8 · S5.SS3.p2

空间泛化。启发性示例显示 DP3 更能在三维空间外推；表 IX 在真实任务验证。所有基线均无法泛化至测试位置，而 DP3 五次中成功四次。

p.8 · S5.T9



空间泛化	1	2	3	4	5
Diffusion Policy	✗	✗	✗	✗	✗
Diffusion Policy (Depth)	✗	✗	✗	✗	✗
DP3	✗	✓	✓	✓	✓

TABLE IX: Spatial generalization on Pour. We place the bowl at 5 different positions that are unseen in the training data. Each position is evaluated with one trial.

表 IX：倾倒任务空间泛化。将碗放在训练未见的 5 个位置，每位置评估 1 次。

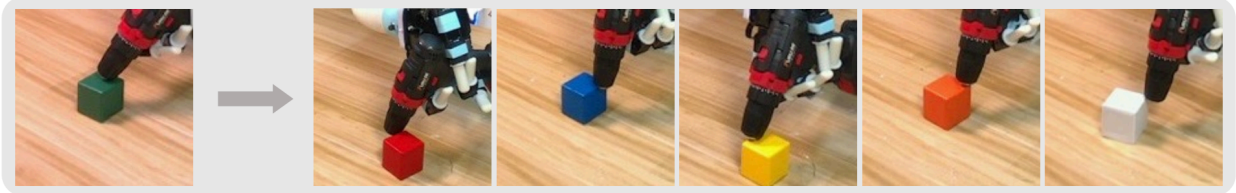
p.8 · S5.SS3.p3

外观泛化。DP3 不用颜色，因此自然有利于跨外观泛化。表 X 中，它对不同颜色方块持续泛化成功，基线未能做到。深度策略也无颜色，但在训练物体上精度较低，限制了泛化。

p.8 · S5.SS3.p4

图像方法可用强数据增强改善外观泛化 [20,19]，但可能妨碍学习 [79,19]。本文更重要的目标，是说明 DP3 即使不做任何数据增强，也能有效泛化，凸显三维表示的潜力。

p.9 · S5.T10



Apperance Generalization (■)	■	■	■	■	■
Diffusion Policy	✗	✗	✗	✗	✗
Diffusion Policy (Depth)	✗	✗	✗	✗	✗
DP3	✓	✓	✓	✓	✓

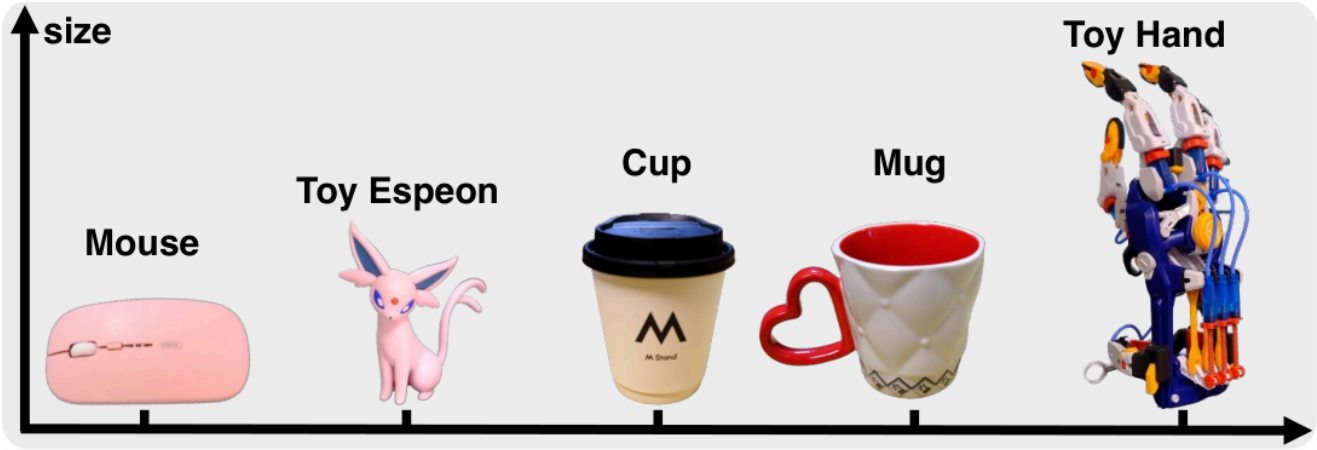
TABLE X: Appearance generalization on Drill. Algorithms are trained with the green cube only and evaluated on 5 different colored cubes. Each color is evaluated with one trial.

表 X：电钻任务外观泛化。仅用绿方块训练，在 5 种颜色方块上各评估 1 次。

p.8 · S5.SS3.p5

实例泛化。跨形状、尺寸、外观不同实例，比单纯外观泛化更难。表 XI 中 DP3 可处理广泛日常物体，主要归因于点云的固有性质：特别经下采样后，策略不易混淆，从而提高适应不同实例的能力。

p.9 · S5.T11



实例泛化	鼠标	太阳伊布玩偶	杯	马克杯	手形物
Diffusion Policy	✗	✗	✗	✗	✓
Diffusion Policy (Depth)	✗	✗	✓	✗	✗
DP3	✓	✓	✓	✓	✓

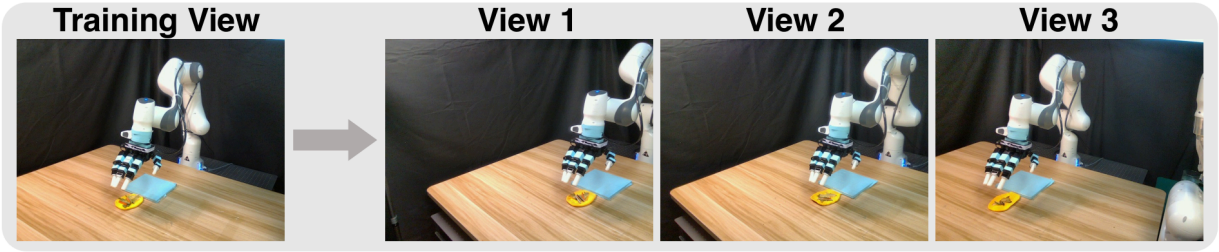
TABLE XI: Instance generalization on Drill. We replace the cube used in Drill with five objects in varied sizes from our daily life. Each instance is evaluated with one trial.

表 XI：电钻任务实例泛化。用日常生活中 5 种不同尺寸物体替换方块，每实例评估 1 次。

p.9 · S5.SS3.p6

视角泛化。图像策略跨视角泛化很困难 [75]，多视角采集又费时昂贵 [82,58]。表 XII 显示 DP3 能处理相机视角小幅改变。重要的是，视角变化后，我们手动变换点云并调整裁剪空间；网络稳健，因此不必精确变换。但它虽能泛化到小幅视角变化，显著变化仍可能困难。

p.10 · S5.T12



视角泛化	视角 1	视角 2	视角 3
Diffusion Policy	✗	✗	✗
Diffusion Policy (Depth)	✗	✗	✗
DP3	✓	✓	✓

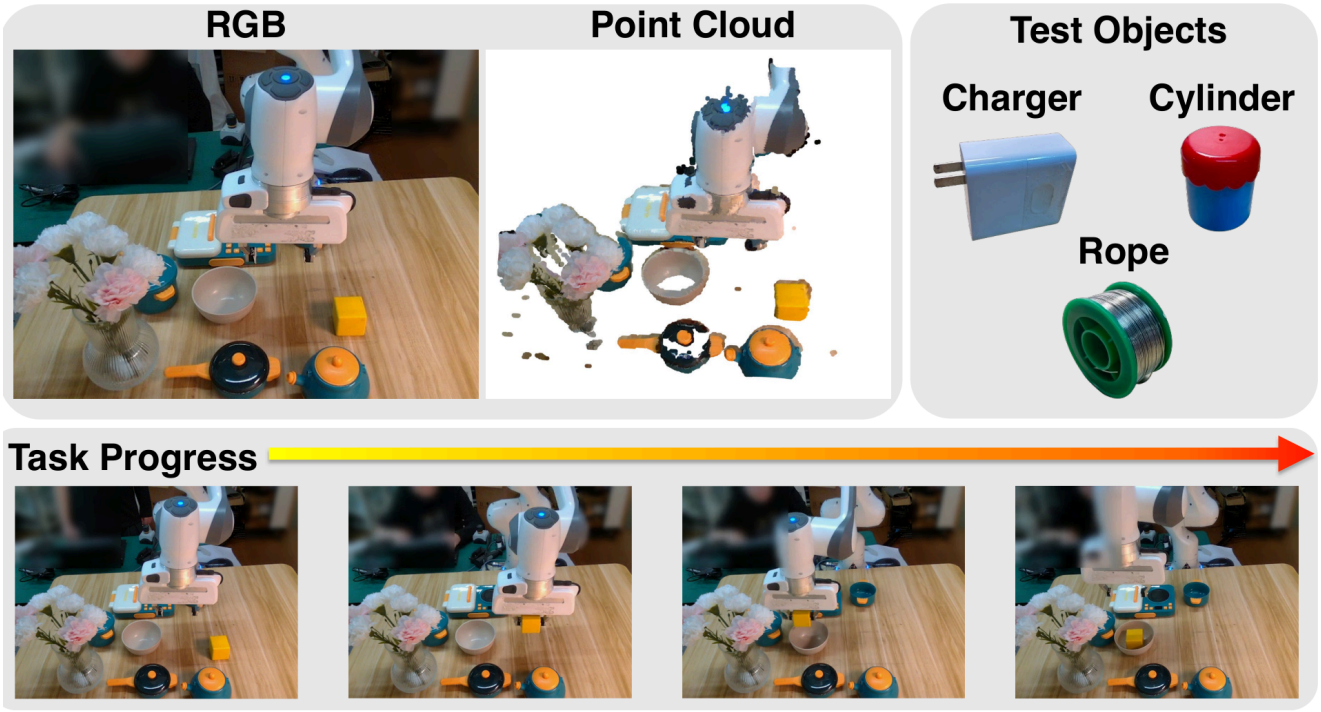
TABLE XII: View generalization on Roll-Up. Each view is evaluated with one trial.

表 XII：卷起任务视角泛化。每视角评估 1 次。

p.9 · S5.SS3.p7

杂乱场景。尽管编码器简单，DP3 也能处理复杂杂乱真实环境。我们设计夹爪抓方块并放碗中的任务，采集 50 条示范。表 XIII 中，DP3 高成功率完成；与仿真一致，换 PointNeXt 后失败。用彩色点云时，抓训练黄方块与原 DP3 相当，却难以处理其他颜色与物体，说明原 DP3 在复杂场景中的有效性与泛化。

p.10 · S5.T13



杂乱场景	Diffusion Policy	DP3 w/ PointNeXt	DP3 w/ color	DP3
成功率	60	0	**80**	**80**

Train with ■ Cube	■	■	■	充电器	圆柱	绳
DP3 w/ color	✗	✗	✗	✗	✗	✗
DP3	✓	✓	✓	✓	✓	✓

TABLE XIII: Results in cluttered scenes. Each algorithm is evaluated with 10 trials in the training color. Each out-of-domain color and object are evaluated with one trial.

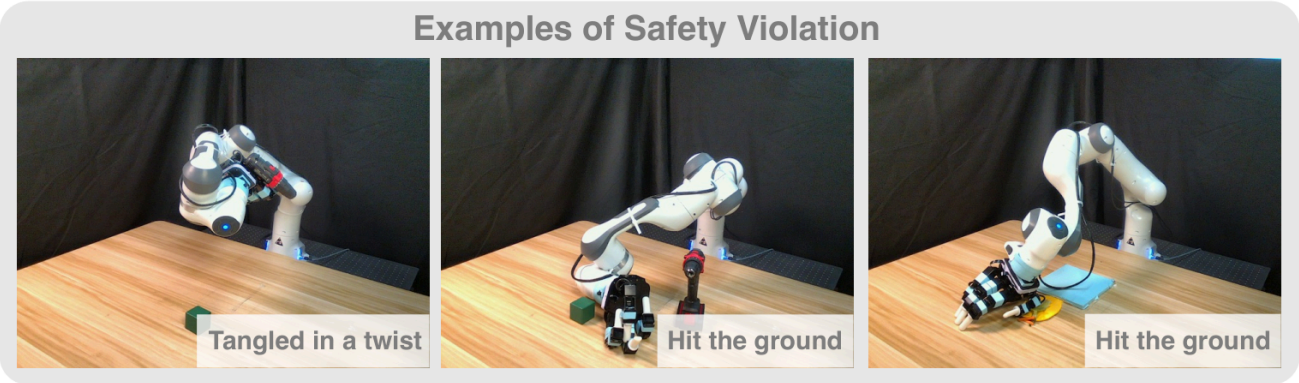
表 XIII：杂乱场景结果。训练颜色每算法评估 10 次；每种域外颜色、物体各评估 1 次。

V-D 部署安全观察

p.10 · S5.SS4.p1

真实实验中，图像、深度图扩散策略常出现不可预测行为，需要人工终止以保护机器人。我们将此定义为安全违规，统计主要真实实验中的比例（表 XIV）。令人意外的是，DP3 很少违规，显示其实用、对硬件友好。直观解释是机器人在三维中行动，直接观测三维信息有助避免碰撞。安全评估主要是定性的，更理论化的理解留待未来。

p.10 · S5.T14



安全违规率 ↓	卷起	包饺子	电钻触碰	倾倒	平均
Diffusion Policy	90	20	20	0	32.5
Diffusion Policy (Depth)	20	30	30	20	25.0
DP3	**0**	**0**	**0**	**0**	0.0

TABLE XIV: Safety violation rate. While conducting the main real-world experiments, we also count the times of safety violation and compute the rate.

表 XIV：安全违规率。在主要真实实验同时记录违规次数并计算比例。

VI 结论

p.10 · S6.p1

DP3 是高效视觉模仿算法，只需少量示范即可应对广泛仿真与真实任务。核心是精心设计的三维表示与扩散策略表达性相结合。72 项仿真中相对二维版本提升 24.2%；真实场景中，Allegro 手高精度执行复杂可变形物体操作。更重要的是，多方面泛化稳健，真实安全违规更少。

p.10 · S6.p2

局限。尽管架构高效，最优控制三维表示仍待发现；本文也未深入研究极长时域任务，留待未来探索。

致谢

p.10 · Sx1.p1

感谢 Zhecheng Yuan、Chen Wang、Cheng Lu、Jianfei Chen 的有益讨论。本工作受国家重点研发计划 2022ZD0161700 支持。

参考文献（保留英文原书目）

[1] Ananye Agarwal, Shagun Uppal, Kenneth Shaw, and Deepak Pathak. Dexterous functional grasping. In *CoRL*, 2023.

[2] Anurag Ajay, Yilun Du, Abhi Gupta, Joshua Tenenbaum, Tommi Jaakkola, and Pulkit Agrawal. Is conditional generative modeling all you need for decision-making? *arXiv preprint arXiv:2211.15657*, 2022.

[3] Sridhar Pandian Arunachalam, Sneha Silwal, Ben Evans, and Lerrel Pinto. Dexterous imitation made easy: A learning-based framework for efficient dexterous manipulation. In *ICRA*, 2023.

[4] Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E Hinton. Layer normalization. *arXiv*, 2016.

- [5] Chen Bao, Helin Xu, Yuzhe Qin, and Xiaolong Wang. Dexart: Benchmarking generalizable dexterous manipulation with articulated objects. In *CVPR*, 2023.
- [6] Tao Chen, Jie Xu, and Pulkit Agrawal. A system for general in-hand object re-orientation. In *CoRL*, 2022a.
- [7] Tao Chen, Megha Tippur, Siyang Wu, Vikash Kumar, Edward Adelson, and Pulkit Agrawal. Visual dexterity: In-hand reorientation of novel and complex object shapes. *Science Robotics*, 8(84):eadc9244, 2023a. doi: 10.1126/scirobotics.adc9244.

[8] Yuanpei Chen, Tianhao Wu, Shengjie Wang, Xidong Feng, Jiechuan Jiang, Zongqing Lu, Stephen McAleer, Hao Dong, Song-Chun Zhu, and Yaodong Yang. Towards human-level bimanual dexterous manipulation with reinforcement learning. *NeurIPS*, 2022b.

[9] Yuanpei Chen, Chen Wang, Li Fei-Fei, and C Karen Liu. Sequential dexterity: Chaining dexterous policies for long-horizon manipulation. *CoRL*, 2023b.

[10] Cheng Chi, Siyuan Feng, Yilun Du, Zhenjia Xu, Eric Cousineau, Benjamin Burchfiel, and Shuran Song. Diffusion policy: Visuomotor policy learning via action diffusion. *RSS*, 2023.

[11] Pete Florence, Corey Lynch, Andy Zeng, Oscar A Ramirez, Ayzaan Wahid, Laura Downs, Adrian Wong, Johnny Lee, Igor Mordatch, and Jonathan Tompson. Implicit behavioral cloning. In *CoRL*, 2022.

[12] Zipeng Fu, Tony Z. Zhao, and Chelsea Finn. Mobile aloha: Learning bimanual mobile manipulation with low-cost whole-body teleoperation. In *arXiv*, 2024.

[13] Theophile Gervet, Zhou Xian, Nikolaos Gkanatsios, and Katerina Fragkiadaki. Act3d: Infinite resolution action detection transformer for robotic manipulation. *arXiv preprint arXiv:2306.17817*, 2023.

[14] Ankit Goyal, Jie Xu, Yijie Guo, Valts Blukis, Yu-Wei Chao, and Dieter Fox. Rvt: Robotic view transformer for 3d object manipulation. *arXiv*, 2023.

[15] Huy Ha, Pete Florence, and Shuran Song. Scaling up and distilling down: Language-guided robot skill acquisition. In *Conference on Robot Learning*. PMLR, 2023.

[16] Siddhant Haldar, Jyothish Pari, Anant Rai, and Lerrel Pinto. Teach a robot to fish: Versatile imitation from one minute of demonstrations. *RSS*, 2023.

[17] Ankur Handa, Karl Van Wyk, Wei Yang, Jacky Liang, Yu-Wei Chao, Qian Wan, Stan Birchfield, Nathan Ratliff, and Dieter Fox. Dexpilot: Vision-based teleoperation of dexterous robotic hand-arm system. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020.

[18] Ankur Handa, Arthur Allshire, Viktor Makoviychuk, Aleksei Petrenko, Ritvik Singh, Jingzhou Liu, Denys Makoviichuk, Karl Van Wyk, Alexander Zhurkevich, Balakumar Sundaralingam, et al. Dextreme: Transfer of agile in-hand manipulation from simulation to reality. In *ICRA*, 2023.

[19] Nicklas Hansen, Hao Su, and Xiaolong Wang. Stabilizing deep q-learning with convnets and vision transformers under data augmentation. *Advances in neural information processing systems*, 2021.

[20] Nicklas Hansen, Zhecheng Yuan, Yanjie Ze, Tongzhou Mu, Aravind Rajeswaran, Hao Su, Huazhe Xu, and Xiaolong Wang. On pre-training for visuo-motor control: Revisiting a learning-from-scratch baseline. In *International Conference on Machine Learning (ICML)*, 2022.

[21] Nicklas Hansen, Yixin Lin, Hao Su, Xiaolong Wang, Vikash Kumar, and Aravind Rajeswaran. Modem: Accelerating visual model-based reinforcement learning with demonstrations. In *ICLR*, 2023a.

[22] Nicklas Hansen, Hao Su, and Xiaolong Wang. Td-mpc2: Scalable, robust world models for continuous control. *arXiv*, 2023b.

- [23] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *NeurIPS*, 2020.
- [24] Binghao Huang, Yuanpei Chen, Tianyu Wang, Yuzhe Qin, Yaodong Yang, Nikolay Atanasov, and Xiaolong Wang. Dynamic handover: Throw and catch with bimanual hands. *CoRL*, 2023a.
- [25] Tao Huang, Guangqi Jiang, Yanjie Ze, and Huazhe Xu. Diffusion reward: Learning rewards via conditional video diffusion. *arXiv*, 2023b.
- [26] Zhiao Huang, Yuanming Hu, Tao Du, Siyuan Zhou, Hao Su, Joshua B Tenenbaum, and Chuang Gan. Plasticinelab: A soft-body manipulation benchmark with differentiable physics. *arXiv*, 2021.
- [27] Michael Janner, Yilun Du, Joshua B Tenenbaum, and Sergey Levine. Planning with diffusion for flexible behavior synthesis. *arXiv*, 2022.
- [28] Tsung-Wei Ke, Nikolaos Gkanatsios, and Katerina Fragkiadaki. 3d diffuser actor: Policy diffusion with 3d scene representations. *Arxiv*, 2024.
- [29] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 2023.
- [30] Kun Lei, Zhengmao He, Chenhao Lu, Kaizhe Hu, Yang Gao, and Huazhe Xu. Uni-o4: Unifying online and offline deep reinforcement learning with multi-step on-policy optimization. *arXiv*, 2023.
- [31] Sizhe Li, Zhiao Huang, Tao Chen, Tao Du, Hao Su, Joshua B Tenenbaum, and Chuang Gan. Dexdeform: Dexterous deformable object manipulation with human demonstrations and differentiable physics. *arXiv*, 2023.
- [32] Cheng Lu, Yuhao Zhou, Fan Bao, Jianfei Chen, Chongxuan Li, and Jun Zhu. Dpm-solver++: Fast solver for guided sampling of diffusion probabilistic models. *arXiv*, 2022.
- [33] Yecheng Jason Ma, William Liang, Guanzhi Wang, De-An Huang, Osbert Bastani, Dinesh Jayaraman, Yuke Zhu, Linxi Fan, and Anima Anandkumar. Eureka: Human-level reward design via coding large language models. *arXiv*, 2023.
- [34] Viktor Makoviychuk, Lukasz Wawrzyniak, Yunrong Guo, Michelle Lu, Kier Storey, Miles Macklin, David Hoeller, Nikita Rudin, Arthur Allshire, Ankur Handa, et al. Isaac gym: High performance gpu-based physics simulation for robot learning. *arXiv*, 2021.
- [35] Ajay Mandlekar, Danfei Xu, Josiah Wong, Soroush Nasiriany, Chen Wang, Rohun Kulkarni, Li Fei-Fei, Silvio Savarese, Yuke Zhu, and Roberto Martín-Martín. What matters in learning from offline human demonstrations for robot manipulation. *arXiv*, 2021.
- [36] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 2021.
- [37] Felipe Nuti, Tim Franzmeyer, and João F Henriques. Extracting reward functions from diffusion models. *arXiv preprint arXiv:2306.01804*, 2023.
- [38] Jyothish Pari, Nur Muhammad Shafiullah, Sridhar Pandian Arunachalam, and Lerrel Pinto. The surprising effectiveness of representation learning for visual imitation. *arXiv preprint arXiv:2112.01511*, 2021.

[39] Tim Pearce, Tabish Rashid, Anssi Kanervisto, Dave Bignell, Mingfei Sun, Raluca Georgescu, Sergio Valcarcel Macua, Shan Zheng Tan, Ida Momennejad, Katja Hofmann, et al. Imitating human behaviour with diffusion models. *ICLR*, 2023.

[40] Xue Bin Peng, Erwin Coumans, Tingnan Zhang, Tsang-Wei Lee, Jie Tan, and Sergey Levine. Learning agile robotic locomotion skills by imitating animals. *arXiv*, 2020.

[41] Aaditya Prasad, Kevin Lin, Jimmy Wu, Linqi Zhou, and Jeannette Bohg. Consistency policy: Accelerated visuomotor policies via consistency distillation. In *Robotics: Science and Systems*, 2024.

[42] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *CVPR*, 2017a.

[43] Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *NeurIPS*, 2017b.

[44] Haozhi Qi, Ashish Kumar, Roberto Calandra, Yi Ma, and Jitendra Malik. In-hand object rotation via rapid motor adaptation. In *CoRL*, 2023a.

[45] Haozhi Qi, Brent Yi, Sudharshan Suresh, Mike Lambeta, Yi Ma, Roberto Calandra, and Jitendra Malik. General in-hand object rotation with vision and touch. In *CoRL*, 2023b.

[46] Guocheng Qian, Yuchen Li, Houwen Peng, Jinjie Mai, Hasan Hammoud, Mohamed Elhoseiny, and Bernard Ghanem. Pointnext: Revisiting pointnet++ with improved training and scaling strategies. *NeurIPS*, 2022.

[47] Yuzhe Qin, Yueh-Hua Wu, Shaowei Liu, Hanwen Jiang, Ruihan Yang, Yang Fu, and Xiaolong Wang. Dexmv: Imitation learning for dexterous manipulation from human videos. In *ECCV*, 2022.

[48] Yuzhe Qin, Wei Yang, Binghao Huang, Karl Van Wyk, Hao Su, Xiaolong Wang, Yu-Wei Chao, and Dietor Fox. Anyteleop: A general vision-based dexterous robot arm-hand teleoperation system. *arXiv preprint arXiv:2307.04577*, 2023.

[49] Aravind Rajeswaran, Vikash Kumar, Abhishek Gupta, Giulia Vezzani, John Schulman, Emanuel Todorov, and Sergey Levine. Learning complex dexterous manipulation with deep reinforcement learning and demonstrations. *arXiv*, 2017.

[50] Moritz Reuss, Maximilian Li, Xiaogang Jia, and Rudolf Lioutikov. Goal-conditioned imitation learning using score-based diffusion policies. *arXiv preprint arXiv:2304.02532*, 2023.

[51] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022.

[52] Kallol Saha, Vishal Mandadi, Jayaram Reddy, Ajit Srikanth, Aditya Agarwal, Bipasha Sen, Arun Singh, and Madhava Krishna. Edmp: Ensemble-of-costs-guided diffusion for motion planning. *arXiv*, 2023.

[53] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

[54] Mingyo Seo, Steve Han, Kyutae Sim, Seung Hyeon Bang, Carlos Gonzalez, Luis Sentis, and Yuke Zhu. Deep imitation learning for humanoid loco-manipulation through human teleoperation. *Humanoids*, 2023a.

- [55] Younggyo Seo, Danijar Hafner, Hao Liu, Fangchen Liu, Stephen James, Kimin Lee, and Pieter Abbeel. Masked world models for visual control. In *CoRL*, 2023b.
- [56] Nur Muhammad Shafiullah, Zichen Cui, Ariuntuya Arty Altanzaya, and Lerrel Pinto. Behavior transformers: Cloning k modes with one stone. *Advances in neural information processing systems*, 2022.
- [57] Nur Muhammad Mahi Shafiullah, Anant Rai, Haritheja Etukuru, Yiqian Liu, Ishan Misra, Soumith Chintala, and Lerrel Pinto. On bringing robots home. *arXiv*, 2023.
- [58] William Shen, Ge Yang, Alan Yu, Jansen Wong, Leslie Pack Kaelbling, and Phillip Isola. Distilled feature fields enable few-shot language-guided manipulation. *arXiv preprint arXiv:2308.07931*, 2023.
- [59] Haochen Shi, Huazhe Xu, Samuel Clarke, Yunzhu Li, and Jiajun Wu. Robocook: Long-horizon elasto-plastic object manipulation with diverse tools. *Proceedings of the 7th Conference on Robot Learning (CoRL)*, 2023.
- [60] Mohit Shridhar, Lucas Manuelli, and Dieter Fox. Perceiver-actor: A multi-task transformer for robotic manipulation. In *CoRL*, 2023.
- [61] Anthony Simeonov, Ankit Goyal, Lucas Manuelli, Lin Yen-Chen, Alina Sarmiento, Alberto Rodriguez, Pulkit Agrawal, and Dieter Fox. Shelving, stacking, hanging: Relational pose diffusion for multi-modal rearrangement. *arXiv preprint arXiv:2307.04751*, 2023.
- [62] Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. *ICLR*, 2021a.
- [63] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *ICLR*, 2021b.
- [64] Kaustubh Sridhar, Souradeep Dutta, Dinesh Jayaraman, James Weimer, and Insup Lee. Memory-consistent neural networks for imitation learning. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=R3Tf7LDdX4.
- [65] Emanuel Todorov, Tom Erez, and Yuval Tassa. Mujoco: A physics engine for model-based control. In *IROS*, 2012.
- [66] Julen Urain, Niklas Funk, Jan Peters, and Georgia Chalvatzaki. Se (3)-diffusionfields: Learning smooth cost functions for joint grasp and motion optimization through diffusion. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023.
- [67] Che Wang, Xufang Luo, Keith Ross, and Dongsheng Li. Vrl3: A data-driven framework for visual deep reinforcement learning. *Advances in Neural Information Processing Systems*, 2022.
- [68] Chen Wang, Linxi Fan, Jiankai Sun, Ruohan Zhang, Li Fei-Fei, Danfei Xu, Yuke Zhu, and Anima Anandkumar. Mimicplay: Long-horizon imitation learning by watching human play. *CoRL*, 2023a.
- [69] Chen Wang, Haochen Shi, Weizhuo Wang, Ruohan Zhang, Li Fei-Fei, and C Karen Liu. Dexcap: Scalable and portable mocap data collection system for dexterous manipulation. *arXiv preprint arXiv:2403.07788*, 2024.
- [70] Zhendong Wang, Jonathan J Hunt, and Mingyuan Zhou. Diffusion policies as an expressive policy class for offline reinforcement learning. *ICLR*, 2023b.

[71] Tianhao Wu, Mingdong Wu, Jiyao Zhang, Yunchong Gan, and Hao Dong. Learning score-based grasping primitive for human-assisting dexterous grasping. In *NeurIPS*, 2023.

[72] Zhou Xian, Nikolaos Gkanatsios, Theophile Gervet, Tsung-Wei Ke, and Katerina Fragkiadaki. Chaineddiffuser: Unifying trajectory diffusion and keypose prediction for robotic manipulation. In *CoRL*, 2023.

[73] Fanbo Xiang, Yuzhe Qin, Kaichun Mo, Yikuan Xia, Hao Zhu, Fangchen Liu, Minghua Liu, Hanxiao Jiang, Yifu Yuan, He Wang, et al. Sapien: A simulated part-based interactive environment. In *CVPR*, 2020.

[74] Ge Yan, Yueh-Hua Wu, and Xiaolong Wang. NeRFuser: Diffusion guided multi-task 3d policy learning, 2024. URL https://openreview.net/forum?id=8GmPLkO0oR.

[75] Sizhe Yang, Yanjie Ze, and Huazhe Xu. Movie: Visual model-based policy adaptation for view generalization. *Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2023.

[76] Zhao-Heng Yin, Binghao Huang, Yuzhe Qin, Qifeng Chen, and Xiaolong Wang. Rotating without seeing: Towards in-hand dexterity through touch. *RSS*, 2023.

[77] Tianhe Yu, Deirdre Quillen, Zhanpeng He, Ryan Julian, Karol Hausman, Chelsea Finn, and Sergey Levine. Meta-world: A benchmark and evaluation for multi-task and meta reinforcement learning. In *CoRL*, 2020.

[78] Ying Yuan, Haichuan Che, Yuzhe Qin, Binghao Huang, Zhao-Heng Yin, Kang-Won Lee, Yi Wu, Soo-Chul Lim, and Xiaolong Wang. Robot synesthesia: In-hand manipulation with visuotactile sensing. *arXiv*, 2023.

- [79] Zhecheng Yuan, Zhengrong Xue, Bo Yuan, Xueqian Wang, Yi Wu, Yang Gao, and Huazhe Xu. Pre-trained image encoder for generalizable visual reinforcement learning. *Advances in Neural Information Processing Systems*, 2022.
- [80] Yanjie Ze, Nicklas Hansen, Yinbo Chen, Mohit Jain, and Xiaolong Wang. Visual reinforcement learning with self-supervised 3d representations. *IEEE Robotics and Automation Letters*, 2023a.
- [81] Yanjie Ze, Yuyao Liu, Ruizhe Shi, Jiaxin Qin, Zhecheng Yuan, Jiashun Wang, and Huazhe Xu. H-index: Visual reinforcement learning with hand-informed representations for dexterous manipulation. In *Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2023b.
- [82] Yanjie Ze, Ge Yan, Yueh-Hua Wu, Annabella Macaluso, Yuying Ge, Jianglong Ye, Nicklas Hansen, Li Erran Li, and Xiaolong Wang. Gnfactor: Multi-task real robot learning with generalizable neural feature fields. *Proceedings of the 7th Conference on Robot Learning (CoRL)*, 2023c.
- [83] Gu Zhang, Hao-Shu Fang, Hongjie Fang, and Cewu Lu. Flexible handover with real-time robust dynamic grasp trajectory generation. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2023.
- [84] Hengshuang Zhao, Li Jiang, Jiaya Jia, Philip HS Torr, and Vladlen Koltun. Point transformer. In *ICCV*, 2021.
- [85] Yuke Zhu, Josiah Wong, Ajay Mandlekar, Roberto Martín-Martín, Abhishek Joshi, Soroush Nasiriany, and Yifeng Zhu. robosuite: A modular simulation framework and benchmark for robot learning. *arXiv preprint arXiv:2009.12293*, 2020.

附录 A 实现细节

p.14 · A0.SS1.p1

DP3 分为感知与决策，细节如下。官方实现：https://github.com/YanjieZe/3D-Diffusion-Policy。

p.14 · A0.SS1.p2

感知。输入含视觉观测与机器人位姿。观测为不含颜色的点云，用最远点采样从原始点云下采样；所有仿真与真实任务取 512 或 1024 点。DP3 Encoder 将其编码为紧凑表示。下面保留原 PyTorch 代码并逐行对照：channels=3 表示只用 xyz；同一编码器也可接收 channels=6 的 xyzrgb。逐点线性层依次 3→64→128→256，每层接 LayerNorm、ReLU；输入 B×N×3 经 MLP 得 B×N×256，沿点维取 max 得 B×256，再经 256→64 投影与 LayerNorm 得 B×64。机器人位姿另用 DimRobo→64→64 MLP，中间 ReLU。点云 64 维和位姿 64 维拼为 128 维，作为动作决策骨干条件。

```
⬇
class DP3Encoder(nn.Module):
 def __init__(self, channels=3):
 # We only use xyz (channels=3) in this work
 # while our encoder also works for xyzrgb (channels=6) in our experiments
 self.mlp = nn.Sequential(
 nn.Linear(channels, 64), nn.LayerNorm(64), nn.ReLU(),
 nn.Linear(64, 128), nn.LayerNorm(128), nn.ReLU(),
 nn.Linear(128, 256), nn.LayerNorm(256), nn.ReLU())
 self.projection = nn.Sequential(nn.Linear(256, 64), nn.LayerNorm(64))

 def forward(self, x):
 # x: B, N, 3
 x = self.mlp(x) # B, N, 256
```

```
  x = torch.max(x, 1)[0] # B, 256
  x = self.projection(x) # B, 64
  return x


⬇
# DimRobo is the dimension of the robot poses.
Sequential(
 (0): Linear(in_features=DimRobo, out_features=64, bias=True)
 (1): ReLU()
 (2): Linear(in_features=64, out_features=64, bias=True))
```

p.14 · A0.SS1.p3

决策。卷积扩散策略将高斯噪声变为连贯动作序列，基于 [10] 官方 PyTorch 实现。以 N_{obs} 个观测时间步预测 H 个动作，推理仅执行最后 N_{act} 个动作。DP3 及扩散基线均设 $H = 4, N_{obs} = 2, N_{act} = 3$。

p.14 · A0.SS1.p4

原 Diffusion Policy 通常采用更长时域，主要因其任务时间步更稠密。表 XV 显示本文任务长短时域无显著差别。考虑真实操作可能突然受干扰，我们采用较短时域。

p.14 · A0.SS1.p5

归一化。各动作维与观测维的最小最大值分别缩放至 $[-1, 1]$。DDPM、DDIM 预测时必须把动作归一化为 $[-1, 1]$，因为它们为稳定会将预测裁剪到 $[-1, 1]$。

附录 B 任务套件

p.14 · A0.SS2.p1

仿真任务。收集多样操作任务系统评估模仿算法，包括 Adroit [49]、Bi-DexHands [8]、DexArt [5]、DexDeform [31]、DexMV [47]、HORA [44]、MetaWorld [77]。完整名称见表 XVIII；原本不支持三维模态的环境，我们增加三维观测支持。

p.14 · A0.SS2.p2

真实任务。回合长度不固定，示范平均长度为：卷起 79.9、包饺子 113.5、电钻 71.4、倾倒 83.6。评估时发现任务完成、策略已无法成功、或动作可能损害硬件，均停止机器人。配置物体见图 8，随机化见图 9。卷起、包饺子改变橡皮泥形状及上方物体外观；电钻、倾倒改变方块、电钻、碗位置。

附录 C 更多仿真实验

p.15 · A0.SS3.p1

逐任务仿真结果。表 XVIII 补充正文表 I，报告 3 种子平均成功率。HORA 是手内旋转任务，以归一化回报衡量而非成功率。

p.15 · A0.SS3.p2

专家成功率。强化学习智能体生成示范；每个专家严格评估 200 回合，成功率见表 XIX。MetaWorld 另报告脚本策略结果。

p.15 · A0.SS3.p3

预测时域选择。DP3 与 Diffusion Policy 基线均采用较短预测、执行时域 $H = 4, N_{act} = 3$，以适应复杂与真实任务：人工干预可能改变环境，策略需立即切换动作。表 XV 中短时域可与长时域竞争。

p.15 · A0.T15

算法	H	D	P	A	DA	SP	平均
DP3	$100_{\pm 0}$	$62_{\pm 4}$	$43_{\pm 6}$	$99_{\pm 1}$	$69_{\pm 4}$	$97_{\pm 4}$	78.3

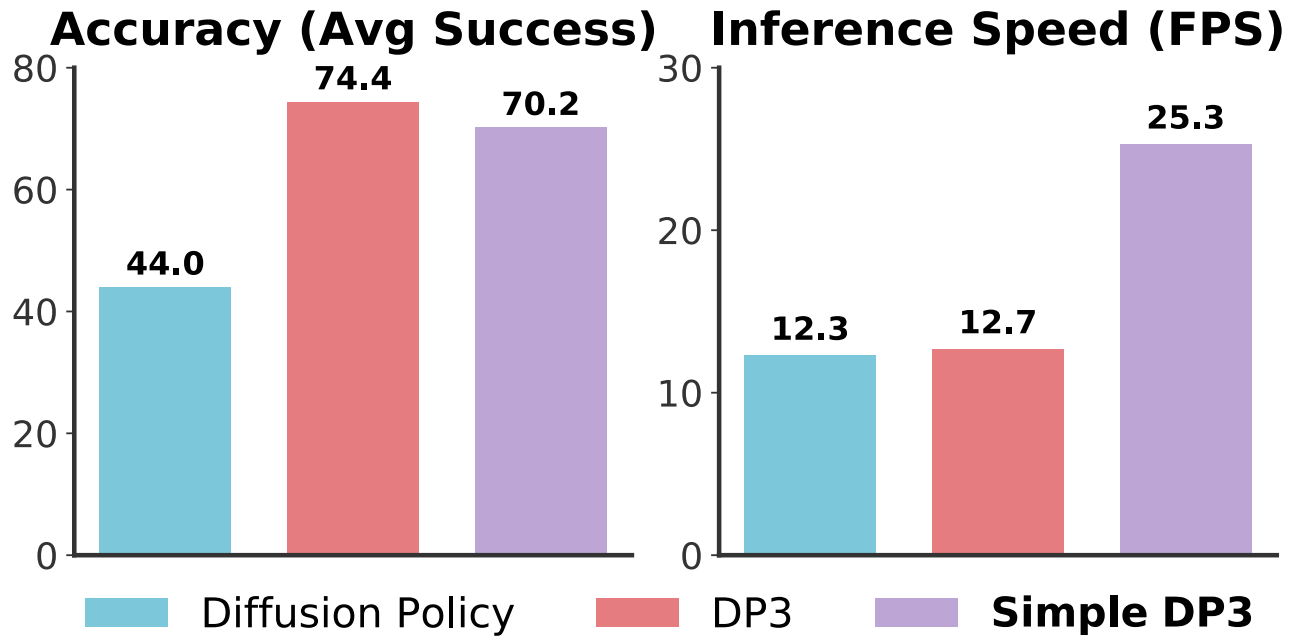
算法	H	D	P	A	DA	SP	平均
长时域	$100_{\pm 0}$	$64_{\pm 5}$	$46_{\pm 3}$	$99_{\pm 1}$	$75_{\pm 3}$	$85_{\pm 14}$	78.2
Diffusion Policy	$48_{\pm 17}$	$50_{\pm 5}$	$25_{\pm 4}$	$15_{\pm 1}$	$43_{\pm 7}$	$63_{\pm 3}$	40.7
长时域	$68_{\pm 11}$	$44_{\pm 4}$	$16_{\pm 2}$	$12_{\pm 3}$	$14_{\pm 1}$	$44_{\pm 5}$	33.0

TABLE XV: Ablation on prediction horizon. In this work, DP3 and Diffusion Policy uses a prediction horizon $H = 4, N_{act} = 3$. We test $H = 16, N_{act} = 8$ originally used in [10] for both methods.

表 XV：预测时域消融。本文 DP3 与 Diffusion Policy 使用 $H = 4, N_{act} = 3$，另测试 [10] 原用 $H = 16, N_{act} = 8$。

附录 D Simple DP3

p.15 · A0.SS4.fig1



无编号图：Simple DP3 简化 U-Net 冗余组件，以提高推理速度；原稿未单列图注。

p.15 · A0.SS4.p1

为提高真实机器人适用性，我们简化了影响速度的策略骨干，称为 Simple DP3。表 XVI 显示，它在保持高精度的同时约提升 2 倍推理速度，主要来自移除 U-Net 冗余组件。实现同官方代码库。

p.15 · A0.T16

算法	Diffusion Policy	DP3	Simple DP3
推理速度 (FPS)	12.3	12.7	25.3 (↑ **99%**)
精度（平均成功率）	44.0	74.4	70.2 (↓ **6%**)

TABLE XVI: Results of Simple DP3. Compared to DP3, Simple DP3 achieves nearly 2x inference speed without losing much accuracy. Full evaluation results are given in Table XVII .

表 XVI：Simple DP3 结果。相对 DP3，推理近快 2 倍、精度损失很小；完整结果见表 XVII。

p.15 · A0.T17

	Adroit			MetaWorld			DexArt				
算法 \ 任务	锤击	开门	转笔	装配	拆卸	棍推	笔记本电脑	水龙头	马桶	桶	平均
DP3	$100_{\pm 0}$	$62_{\pm 4}$	$43_{\pm 6}$	$99_{\pm 1}$	$69_{\pm 4}$	$97_{\pm 4}$	$83_{\pm 1}$	$63_{\pm 2}$	$82_{\pm 4}$	$46_{\pm 2}$	74.4

	Adroit			MetaWorld			DexArt			
算法 \ 任务	锤击	开门	转笔	装配	拆卸	棍推	笔记本电脑	水龙头	马桶	桶	平均
Diffusion Policy	48 ± 17	50 ± 5	25 ± 4	15 ± 1	43 ± 7	63 ± 3	69 ± 4	23 ± 8	58 ± 2	**46 ± 1**	44.0
Simple DP3	**100 ± 0**	58 ± 4	**46 ± 5**	79 ± 1	50 ± 3	**97 ± 5**	**84 ± 2**	**63 ± 3**	81 ± 6	44 ± 6	70.2

TABLE XVII: Full evaluation results of Simple DP3. We evaluate Simple DP3 on 10 tasks and compare it with DP3 and find that Simple DP3 could achieve results very competitive to DP3.

表 XVII：Simple DP3 完整评估。在 10 项任务与 DP3 比较，性能相当。

p.16 · A0.T18

	Adroit *[49]*				Bi-DexHands *[8]*				
Alg \ 任务	锤击	开门	转笔	堆积木	瓶盖	向外开门	抓取放置	交接	剪刀
DP3	**100 ± 0**	**62 ± 4**	**43 ± 6**	**24 ± 15**	**83 ± 10**	**100 ± 0**	**69 ± 22**	**45 ± 8**	**100 ± 0**
Diffusion Policy	45 ± 5	37 ± 2	13 ± 2	4 ± 4	61 ± 5	**100 ± 0**	65 ± 9	38 ± 0	**100 ± 0**

	DexArt *[5]*				DexDeform *[31]*					DexMV *[47]*		HORA *[44]*
Alg \ 任务	笔记本电脑	水龙头	马桶	桶	绳	包子	包饺子	包裹	翻转	折叠	倾倒	放入	旋转
DP3	**83 ± 1**	**63 ± 2**	**82 ± 4**	**46 ± 2**	93 ± 2	70 ± 9	**92 ± 0**	**94 ± 0**	97 ± 1	81 ± 2	**99 ± 2**	**100 ± 0**	**71 ± 31**
Diffusion Policy	69 ± 4	23 ± 8	58 ± 2	**46 ± 1**	**97 ± 0**	**76 ± 4**	**92 ± 0**	91 ± 0	**99 ± 0**	**88 ± 1**	90 ± 2	**100 ± 0**	49 ± 11

	Meta-World *[77]* (Easy)						
Alg \ 任务	按按钮	向下按按钮	隔墙向下按按钮	隔墙按按钮	咖啡机按钮	转旋钮	关门
DP3	**100 ± 0**	**100 ± 0**	**99 ± 2**	**99 ± 1**	**100 ± 0**	**66 ± 1**	**100 ± 0**
Diffusion Policy	99 ± 1	98 ± 1	96 ± 3	97 ± 3	99 ± 1	63 ± 10	**100 ± 0**

	Meta-World (Easy)								
Alg \ 任务	锁门	开门	解锁门	关抽屉	开抽屉	关水龙头	开水龙头	压把手	拉把手
DP3	**98 ± 2**	**99 ± 1**	**100 ± 0**	**100 ± 0**	**100 ± 0**	**100 ± 0**	**100 ± 0**	**100 ± 0**	**53 ± 11**
Diffusion Policy	86 ± 8	98 ± 3	98 ± 3	**100 ± 0**	93 ± 3	**100 ± 0**	**100 ± 0**	81 ± 4	27 ± 22

	Meta-World (Easy)							
Alg \ 任务	侧压把手	侧拉把手	拉杠杆	滑盘	向后滑盘	侧向后滑盘	侧滑盘	到达
DP3	**100 ± 0**	**85 ± 3**	**79 ± 8**	**100 ± 1**	**99 ± 0**	**100 ± 0**	**100 ± 0**	**24 ± 1**
Diffusion Policy	**100 ± 0**	23 ± 17	49 ± 5	83 ± 4	**99 ± 0**	**100 ± 0**	**100 ± 0**	18 ± 2

	Meta-World (Easy)					Meta-World (Medium)			
Alg \ 任务	隔墙到达	关窗	开窗	侧拔销	投篮	箱中抓取	关盒	拉咖啡杯	推咖啡杯
DP3	**68 ± 3**	**100 ± 0**	**100 ± 0**	**75 ± 5**	**98 ± 2**	**34 ± 30**	**42 ± 3**	**87 ± 3**	**94 ± 3**
Diffusion Policy	59 ± 7	100 ± 0	100 ± 0	74 ± 3	85 ± 6	15 ± 4	30 ± 5	34 ± 7	67 ± 4

	Meta-World (Medium)							Meta-World (Hard)		
Alg \ 任务	锤击	侧插销	隔墙推	足球	扫	扫入	装配	手插入	孔中取物	抓放
DP3	**76 ± 4**	**69 ± 7**	**49 ± 8**	**18 ± 3**	**96 ± 3**	**15 ± 5**	**99 ± 1**	**14 ± 4**	**14 ± 9**	**12 ± 4**
Diffusion Policy	15 ± 6	34 ± 7	20 ± 3	14 ± 4	18 ± 8	10 ± 4	15 ± 1	9 ± 2	0 ± 0	0 ± 0

	Meta-World (Hard)		Meta-World (Very Hard)					平均
Alg \ 任务	推	后推	放上架	拆卸	棍拉	棍推	隔墙抓放	
DP3	**51 ± 3**	**0 ± 0**	**17 ± 10**	**69 ± 4**	**27 ± 8**	**97 ± 4**	**35 ± 8**	74.4

	Meta-World (Hard)		Meta-World (Very Hard)					平均
Alg \ 任务	**推**	**后推**	**放上架**	**拆卸**	**棍拉**	**棍推**	**隔墙抓放**	
Diffusion Policy	30 ± 3	0 ± 0	11 ± 3	43 ± 7	11 ± 2	63 ± 3	5 ± 1	59.8

TABLE XVIII: Main results on 72 simulation tasks. Results for each task are provided in this table. A summary across domains is shown in Table I .

表 XVIII：72 项仿真主要逐任务结果。领域汇总见表 I。

p.17 · A0.T19

	Adroit *[49]*				Bi-DexHands *[8]*				
Alg \ 任务	锤击	开门	转笔	堆积木	瓶盖	向外开门	抓取放置	交接	剪刀
专家	99.0	100.0	97.0	83.5	100.0	100.0	100.0	77.0	99.5

	DexArt *[5]*				DexDeform *[31]*						DexMV *[47]*		HORA *[44]*
Alg \ 任务	笔记本电脑	水龙头	马桶	桶	绳	包子	包饺子	包裹	翻转	折叠	倾倒	放入	旋转
专家	86.5	58.0	66.5	80.0	100.0	100.0	100.0	100.0	100.0	100.0	88.5	64.5	80.5

	Meta-World *[77]* (Easy)						
Alg \ 任务	按按钮	向下按按钮	隔墙向下按按钮	隔墙按按钮	咖啡机按钮	转旋钮	关门
专家	100.0	100.0	100.0	98.5	100.0	100.0	100.0

	Meta-World (Easy)								
Alg \ 任务	锁门	开门	解锁门	关抽屉	开抽屉	关水龙头	开水龙头	压把手	拉把手
专家	100.0	98.5	100.0	100.0	100.0	100.0	100.0	100.0	100.0

	Meta-World (Easy)								
Alg \ 任务	侧压把手	侧拉把手	拉杠杆	滑盘	向后滑盘	侧向后滑盘	侧滑盘	到达	
专家	100.0	100.0	98.5	100.0	100.0	100.0	100.0	100.0	

	Meta-World (Easy)					Meta-World (Medium)			
Alg \ 任务	隔墙到达	关窗	开窗	侧拔销	投篮	箱中抓取	关盒	拉咖啡杯	推咖啡杯
专家	100.0	100.0	100.0	99.0	100.0	97.0	90.0	100.0	100.0

	Meta-World (Medium)						Meta-World (Hard)			
Alg \ 任务	锤击	侧插销	隔墙推	足球	扫	扫入	装配	手插入	孔中取物	抓放
专家	100.0	92.0	100.0	90.5	100.0	90.0	100.0	100.0	100.0	100.0

	Meta-World (Hard)		Meta-World (Very Hard)				
Alg \ 任务	推	后推	放上架	拆卸	棍拉	棍推	隔墙抓放
专家	100.0	0.0	99.5	92.5	95.0	100.0	99.5

TABLE XIX: Success rates of experts on 72 simulation tasks. We evaluate 200 episodes for each task. For MetaWorld tasks, we evaluate both BAC agents and the script policies provided officially in MetaWorld. For DexDeform tasks, the demonstrations are collected by human teleportation [31] and we record the success rates as 100%.

表 XIX：72 项仿真专家成功率，每任务评估 200 回合。MetaWorld 同时评估 BAC 智能体与官方脚本；DexDeform 示范由人工遥操 [31] 采集，记为 100%。原文写 teleportation，此处按该数据集采集语境译为遥操。

读者校注与工程映射（本册原创）

三维表示的收益先于生成器选型

DP3 的关键不是增加复杂三维网络，而是将已标定深度转为度量点云、裁掉背景，再用小编码器保留控制相关几何。表 IV 中点云平均成功率 78.3，图像 40.7，深度图 32.0，RGB-D 34.7，体素 32.3；真值状态为 76.8。这组 6 任务消融测的是表示与相应编码器的组合，不应读成所有任务上“点云严格优于真值”。

从像素到 128 维条件的维度链

对像素齐次向量 $x=(u,v,1)^{T}$，轴向深度 z，相机内参 K、相机到世界的刚体变换 (R,t)，反投影为 $p_w=R(zK^{-1}x)+t$。深度单位与平移单位必须一致。裁剪与 FPS 后得到 B×N×3；逐点 MLP 为 3→64→128→256；沿 N 取最大值得 B×256；线性投影与 LayerNorm 得 B×64。机器人本体位姿经 DimRobo→64→64 MLP，形成另一 B×64，拼接为 B×128 条件。两步观测再按决策骨干接口组合，不应误把点数 N 当时间长度。

max-pooling 对点排列不变。原文正文称它为“order-equivariant”，而对输出的单一集合特征，更精确的术语是 permutation invariant。本文编码器没有由此获得 SE(3) 等变性；相机位姿、坐标系和数据覆盖仍影响控制。

损失和实际参数化需要同时核对

正文式（2）展示噪声预测 MSE，但实现段明确使用 sample prediction。前者回归加入的高斯噪声，后者回归干净动作；使用同一噪声调度并不能让两种监督的权重自动相同。复现应检查 scheduler.prediction_type、归一化范围和裁剪设置，而不是仅把论文公式抄成 epsilon loss。原文训练 100 步、DDIM 推理 10 步；这不是“单步策略”。

动作、观测逐维 min-max 归一化到 [−1,1]。这一约定与 DDPM/DDIM 的预测裁剪耦合；若动作单位或工作区改变却复用旧归一化器，裁剪可能把合法动作当越界值。推理应保存并复用训练集统计，而不在每个测试块重新估计 min/max。

有效结果与条件各在哪里

72 任务领域均值：DP3 74.4，二维 Diffusion Policy 59.8，相对改善按 (74.4−59.8)/59.8 计算约为 24.4%；原表标注 24.2%，差异可能来自未四舍五入的底层值。本册保留原表 24.2%，不以重算的舍入值替换作者结果。

真实任务平均 85% 来自 4 项任务各 10 次，逐任务为卷起 90%、包饺子 70%、电钻 80%、倾倒 100%。视角泛化仅每视角 1 次，且作者手动变换点云、调整裁剪区域；显著视角变化仍被明确列为困难。因此，3 次视角成功是小样本、带几何适配的证据，不是任意相机变化的保证。

安全违规采用人工终止判断，原文也将其称为主要定性观察。表 XIV 的 0% 是这批真实实验的观测频率；没有给出形式化碰撞安全约束。不要用它代替独立碰撞检测或执行器限幅。

本机 S9 对“相机泛化”作了什么补充

V:/embodied-3d-affordance-policy/DELIVERY_REPORT.md 中，45° 相机偏移下，世界系点云 Flow 与点云 DDIM 都为 59/64（92.2%），而图像 Flow 为 0/64，平均终端位置误差 224.39 mm。高成功率主要来自当前外参与可见表面配准；单独几何观察器的定位误差已约 1.57 mm / 0.41°。不带模板配准的点云 Flow 在 45° 虽比图像好，但仍为 0/64 成功。

这使映射关系更清楚：DP3 的点云表示支持几何泛化；S9 的高成功率是“已标定点云+学习形状配准+动作模型”的混合系统效果，不能反过来声称裸 DP3 编码器已达到该成绩。

同一 S9 检查点新增球障碍时，关闭 SDF 引导 64/64 碰撞，完整引导后 64/64 成功、0/64 碰撞；5 步中位推理从 40.37 ms 到 184.36 ms。组件消融显示，单独流梯度回拉仍有 64/64 碰撞，步内近端干净路径修正才清除碰撞。这提供一个清晰扩展接口：DP3 负责从观测提出动作块，独立几何模块负责测试期路径可行性，并把计算成本计入端到端延迟。

另外三条本机线如何影响迁移

S3（V:/lerobot-flow-matching-v2/DELIVERY_REPORT.md）中的 4,200 条配对回合说明，开环专家数据训练的快速生成器仍会在闭环失接触后漂移。DP3 的滚动动作块可缓解时序问题，但若失接触恢复状态根本不在数据中，换三维编码器或扩散采样器并不能凭空补齐行为。

S7（V:/rectified-flow-distill-bench/DELIVERY_REPORT.md）的低 NFE 数值分析提示：减少 DDIM 步数、改用连续流或单步蒸馏，应在同一观测器和同一闭环门禁下比较。DP3 原文的高低维 DPM-solver++ 消融已表明，低维采样优势不能直接迁移到高维动作。

S8（V:/event-neural-flow/DELIVERY_REPORT.md）的 1.213 ms 设备内 CUDA Graph 光流通路不含采集、磁盘、主机上传；它不能直接作为 DP3 全策略时延。若事件估计为三维策略补充运动条件，应先校准时钟、相机内外参、深度尺度，再评价遮挡与速度变化；二维光流不是三维物体速度。

原文内部差异与复现实务

正文 IV-B 称没有任务成功率低于 10%，但附录表 XVIII 的 MetaWorld Push Back 为 0±0；本册均保留。图 1 与附录 Simple DP3 的速度依赖具体硬件和观察设置；相关工作脚注给出的 CALVIN 公平对照中，DP3 为 581 ms 预测 4 步，3D Diffuser Actor 为 600 ms 预测 6 步，不能仅比较 FPS 就断言控制频率更高。

Terminology ledger

English	中文与约定
point cloud	点云，原文默认只用 xyz、不含颜色
farthest point sampling, FPS	最远点采样；区别于 frames per second 的帧率缩写
sample prediction	样本预测，即干净动作参数化
epsilon prediction	噪声预测
proprioceptive state	本体感知状态
spatial / viewpoint generalization	空间 / 视角泛化，分开统计
deployment safety violation	部署安全违规，按作者人工终止标准
receding-horizon execution	滚动时域执行

English engineering takeaway

DP3 benefits from calibrated metric geometry and a deliberately small point encoder. Its max-pooled feature is permutation invariant, not an SE(3)-equivariant policy. Preserve the sample-prediction implementation detail, normalization and camera transforms when reproducing the method. Local S9 controls isolate registration from action generation and show why SDF path correction is a separate inference-time component. Local S3 failures motivate explicit recovery data rather than assuming better perception solves contact drift.