501机器人 潘组例行周会
3D Diffusion Policy论文分享

李佩泽 2025.07.26

论文分享 Section

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

Yanjie Ze[1*] Gu Zhang[12*] Kangning Zhang[12] Chenyuan Hu[13] Muhan Wang[13] Huazhe Xu[314]

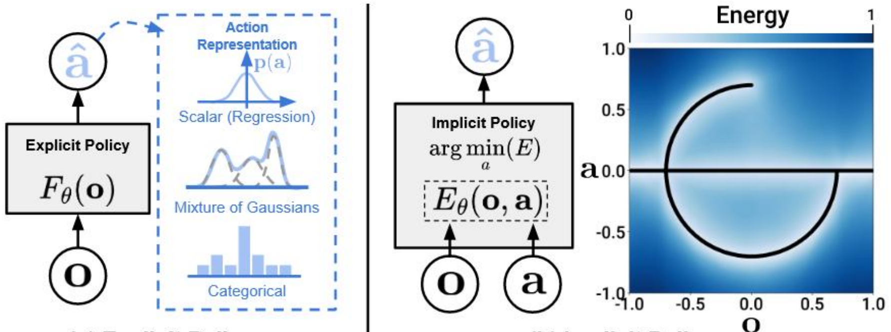
Abstract—**Imitation learning provides an efficient way to teach robots dexterous skills; however, learning complex skills robustly and generalizablely usually consumes large amounts of human demonstrations. To tackle this challenging problem, we present 3D Diffusion Policy (DP3), a novel visual imitation learning approach that incorporates the power of 3D visual representations into diffusion policies, a class of conditional action generative models. The core design of DP3 is the utilization of a compact 3D visual representation, extracted from sparse point clouds with an efficient point encoder. In our experiments involving 72 simulation tasks, DP3 successfully handles most tasks with just 10 demonstrations and surpasses baselines with a 24.2% relative improvement. In 4 real robot tasks, DP3 demonstrates precise control with a high success rate of 85%, given only 40 demonstrations of each task, and shows excellent generalization abilities in diverse aspects, including space, viewpoint, appearance, and instance. Interestingly, in real robot experiments, DP3 rarely violates safety requirements, in contrast to baseline methods which frequently do, necessitating human intervention. Our extensive evaluation highlights the critical importance of 3D representations in real-world robot learning. Code and videos are available on 3d-diffusion-policy.github.io .**

In summary, our contributions are four-fold:

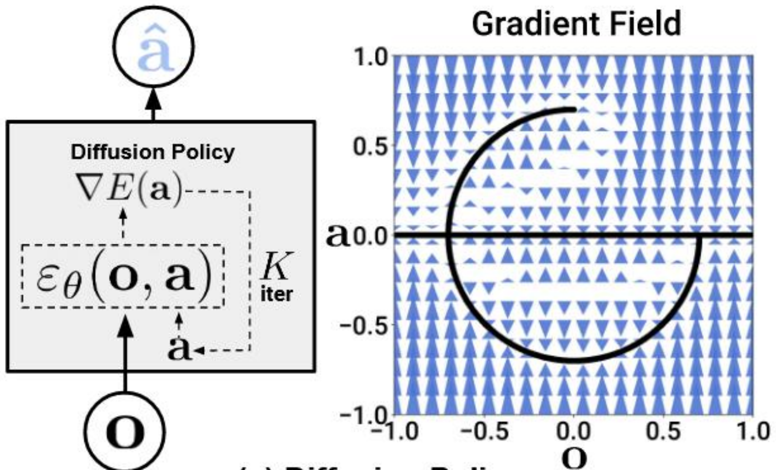
1) We propose 3D Diffusion Policy (DP3), an effective visuomotor policy that generalizes across diverse aspects with few demonstrations.
2) To reduce the variance brought by benchmarks and tasks, we evaluate DP3 in a broad range of simulated and real-world tasks, showing the universality of DP3.
3) We conduct comprehensive analyses on visual representations in DP3 and show that a simple point cloud representation is preferred over other intricate 3D representations and is better suited for diffusion policies over other policy backbones.
4) DP3 is able to perform real-world deformable object manipulation using a dexterous hand with only 40 demonstrations, demonstrating that complex high-dimensional tasks could be handled with little human data.

DP3 基座：Diffusion Policy

Cheng Chi,et al. Diffusion policy: Visuomotor policy learning via action diffusion. RSS, 2023.



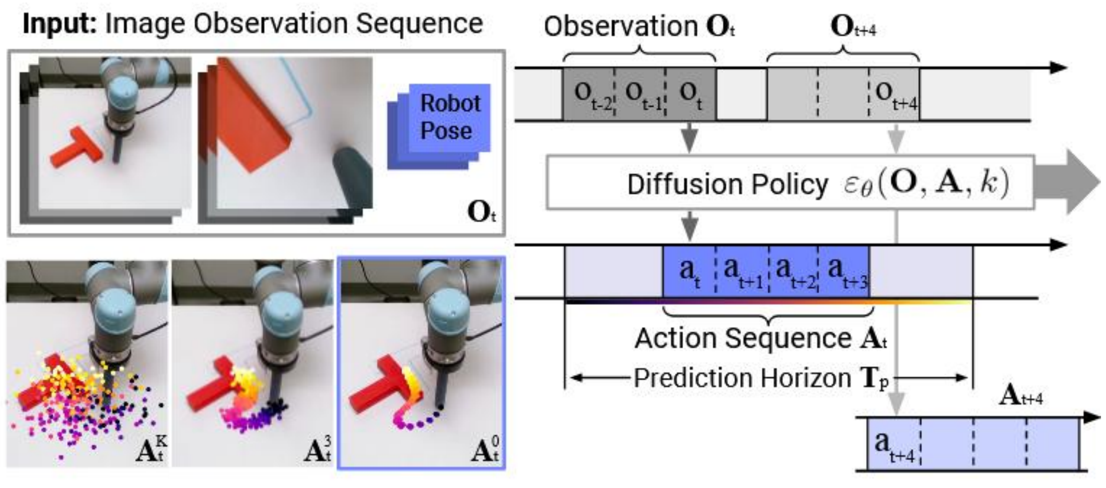
(a) Explicit Policy (b) Implicit Policy



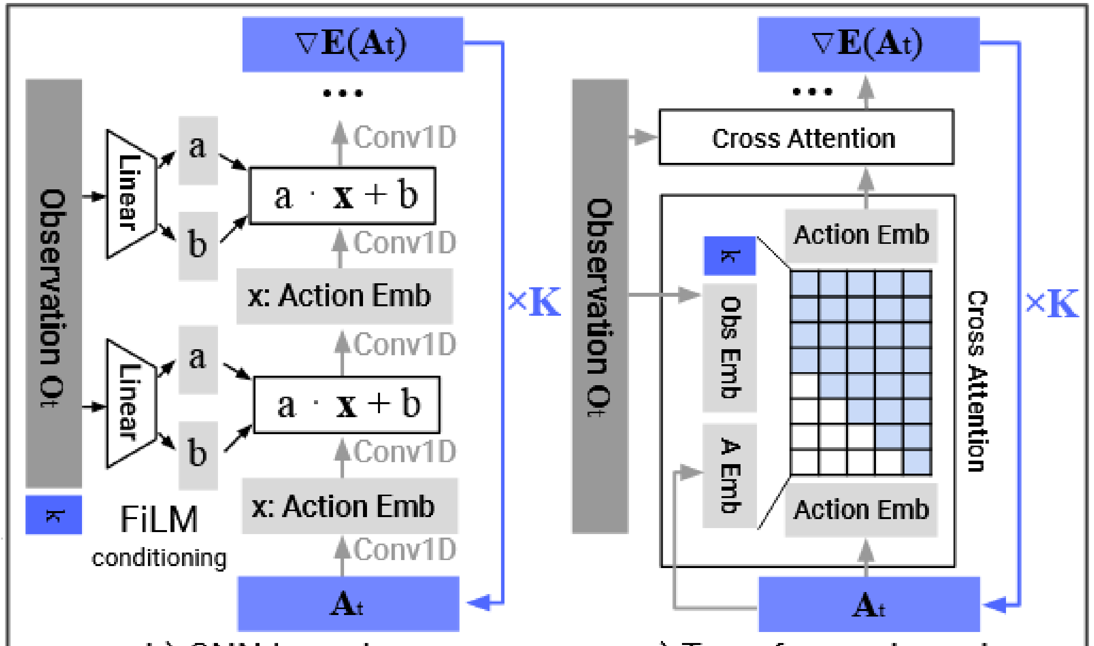
(c) Diffusion Policy

$$\mathbf{x}^{k-1} = \alpha(\mathbf{x}^k - \gamma\varepsilon_\theta(\mathbf{x}^k,k) + \mathcal{N}(0,\sigma^2 I)), \quad (1)$$

$$\mathbf{x}' = \mathbf{x} - \gamma\nabla E(\mathbf{x}), \quad (2)$$

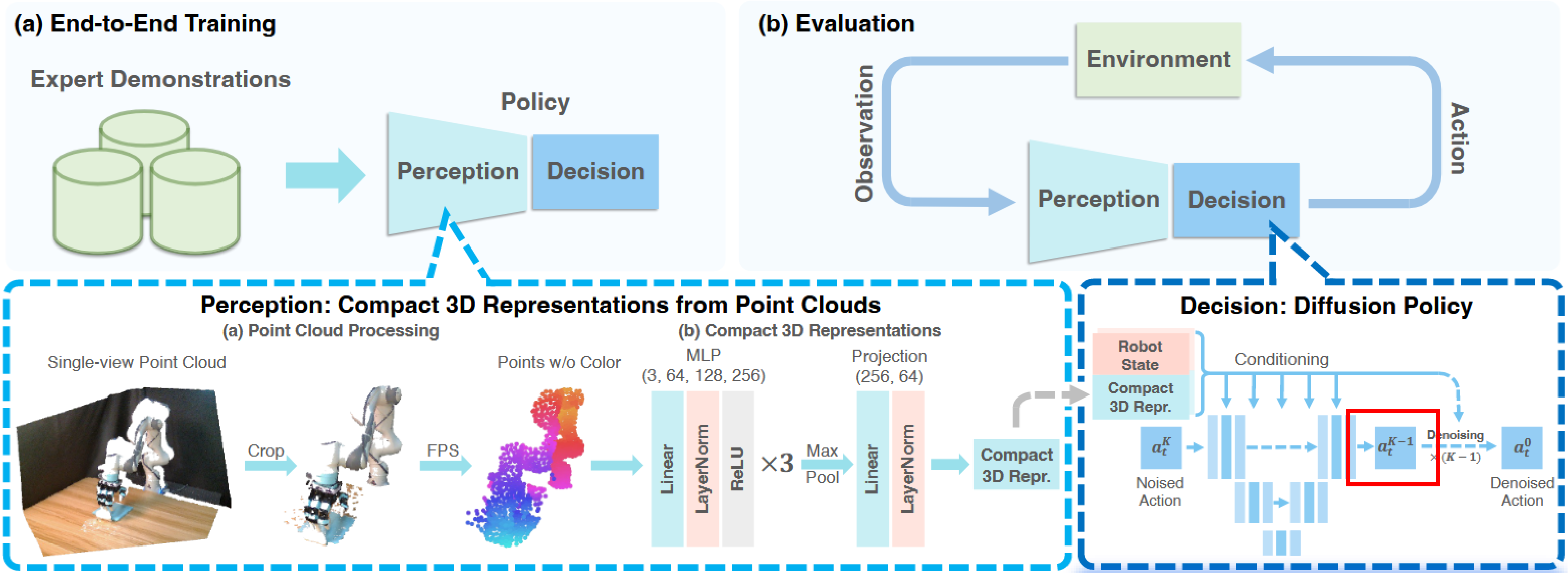


a) Diffusion Policy General Formulation



b) CNN-based c) Transformer-based

DP3 模型实现架构



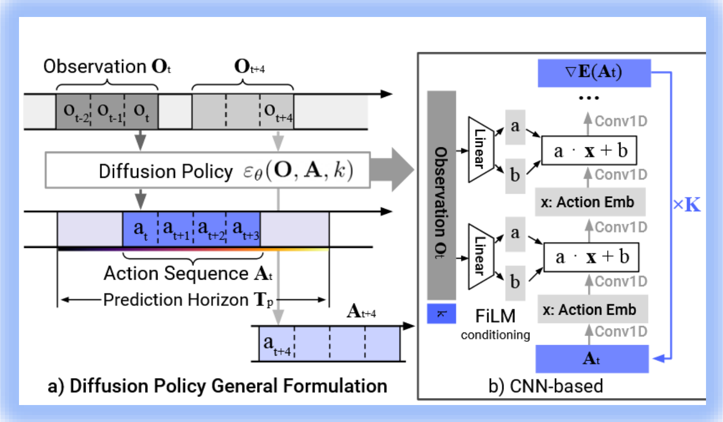
Crop:

只保留一个硬编码的箱范围内的点云

FPS:

FPS通过迭代最大化采样点间的距离实现均匀覆盖，步骤如下：

1. **初始化选择**：随机选取一个点作为首个采样点（如点云中心或索引0的点）[2][3][6]。
2. **距离计算**：计算所有点到**当前采样点集**中最近点的距离（如欧氏距离），生成距离数组。
3. **选取最远点**：选择距离数组中值最大的点加入采样集 [2][6]。
4. **迭代更新**：重复步骤2-3，直至达到目标采样点数 *K*。



DP3 模型横评实验

A. A Motivating Example

To better illustrate the generalization ability of DP3, we first give a straightforward example. We use the MetaWorld Reach task [77] as our testbed. In this task, the goal is for the gripper to accurately reach a designated target point. To evaluate the effectiveness of imitation learning algorithms in not only fitting training data but also generalizing to new scenarios, we visualize the ● training points and the ● successful evaluation points in 3D space, as shown in Figure 3. We observe that with merely five training points, DP3 reaches points distributed over the 3D space, while for 2D-based methods, Diffusion Policy [10] and IBC [11] learn to reach within a plane-like area, and BCRNN [35] fails to cover the space. This example demonstrates the superior generalization and efficiency of DP3, particularly in scenarios where available data is limited.

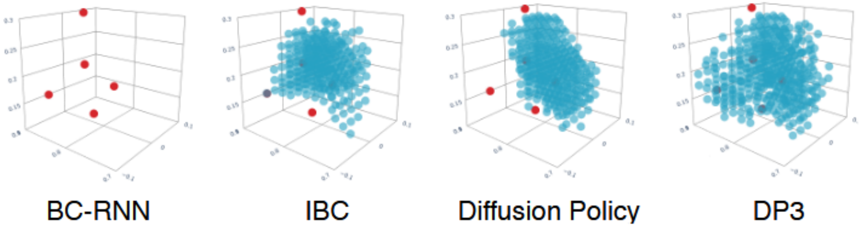
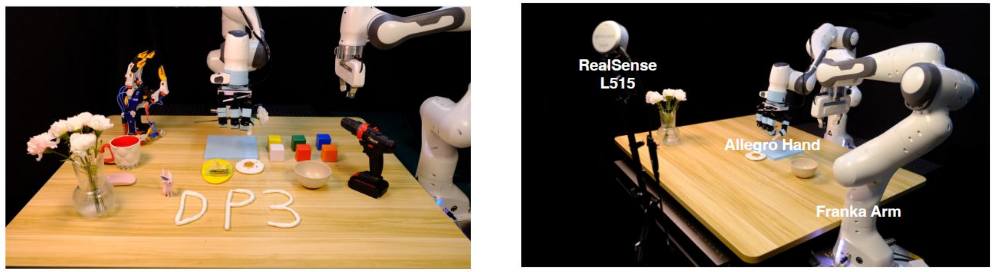


Fig. 3: **Generalization in 3D space with few data.** We use MetaWorld Reach as an example task, given only 5 demonstrations (visualized by ●). We evaluate 1000 times to cover the 3D space and visualize the ● successful evaluation points. **DP3** learns the generalizable skill in 3D space; **Diffusion 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.

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.

Simulation Benchmark (72 Tasks)						
Domain	Robo	Object	Simulator	ActD	#Task	#Demo
Adroit	Shadow	Rigid/Art	MuJoCo	28	3	10
Bi-DexHands	Shadow	Rigid/Art	IsaacGym	52	6	10
DexArt	Allegro	Art	Sapien	22	4	100
DexDeform	Shadow	Deform	PlasticineLab	52	6	10
DexMV	Shadow	Rigid/Fluid	Sapien	30	2	10
HORA	Allegro	Rigid	IsaacGym	16	1	100
MetaWorld	Gripper	Rigid/Art	MuJoCo	4	50	10

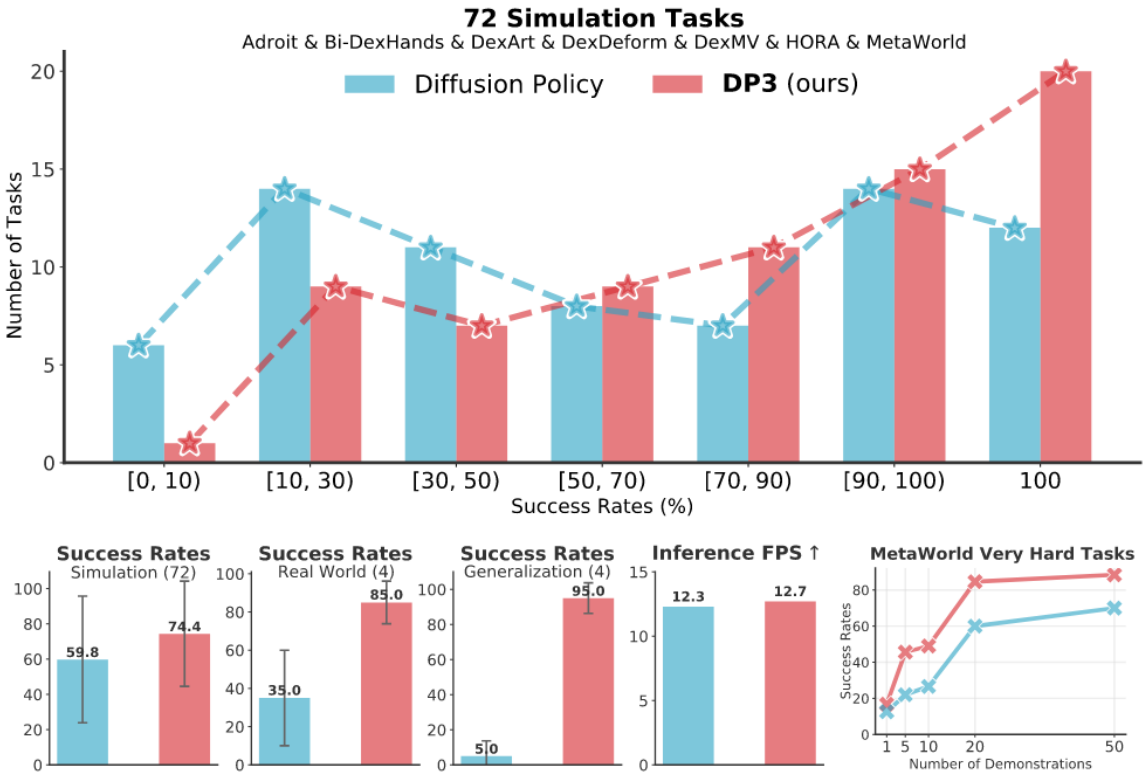
Real Robot Benchmark (4 Tasks)					
Task	Robo	Object	ActD	#Demo	Description
Roll-Up	Allegro	Deform	22	40	Wrap plasticine to make a roll-up
Dumpling	Allegro	Deform	22	40	Wrap plasticine and pinch with fingers
Drill	Allegro	Rigid	22	40	Grasp the drill and touch the cube
Pour	Gripper	Rigid	7	40	Pick the bowl, pour, and place



(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.

DP3 模型横评实验



(a) 3D Diffusion Policy (**DP3**) vs. Diffusion Policy: ***Better, Faster, Stronger.***

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.

Algorithm \ Task	Adroit (3)	Bi-DexHands (6)	DexArt (4)	DexDeform (6)	DexMV (2)	HORA (1)	MetaWorld Easy (28)	MetaWorld Medium (11)	MetaWorld Hard (6)	MetaWorld Very Hard (5)	Average (72)
DP3	**68.3**	**70.2**	**68.5**	87.8	**99.5**	**71.0**	**90.9**	**61.6**	**31.7**	**49.0**	**74.4±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±35.9

BCRNN: Behavior Cloning RNN 输入状态s 预测动作a

IBC: Implicit Behavior Cloning 隐式行为克隆

评估条件动作误差E(o,a) 取最小处生成动作

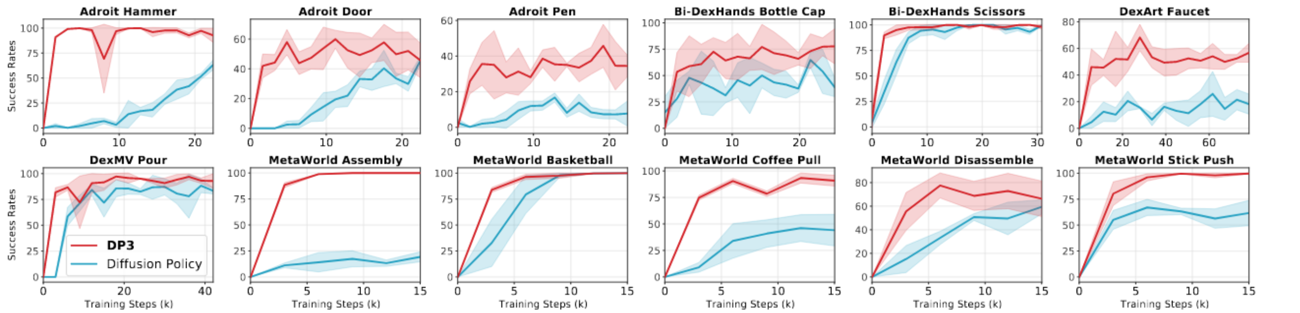


Fig. 5: **Learning efficiency.** We sample 12 simulation tasks and show the learning curves of DP3 and Diffusion Policy. DP3 demonstrates a rapid convergence towards high accuracy. In contrast, Diffusion Policy exhibits a slower learning progress and achieves notably lower convergence in most tasks.

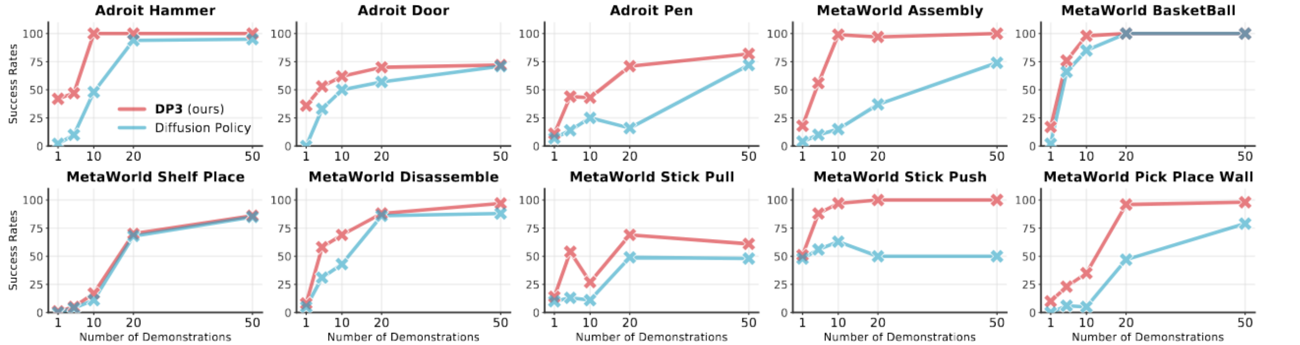


Fig. 6: **Efficient scaling with demonstrations.** We sample 10 simulation tasks and train DP3 and Diffusion Policy with an increasing number of demonstrations. DP3 addresses all these tasks well and generally improves the accuracy with more demonstrations. Diffusion Policy also scales well on some tasks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

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.

	Adroit			MetaWorld			DexArt				
Algorithm \ Task	Hammer	Door	Pen	Assembly	Disassemble	Stick-Push	Laptop	Faucet	Toilet	Bucket	**Average**
DP3	**100±0**	**62±4**	**43±6**	**99±1**	**69±4**	**97±4**	**83±1**	**63±2**	**82±4**	**46±2**	**74.4**
Diffusion Policy	48±17	50±5	25±4	15±1	43±7	63±3	69±4	23±8	58±2	**46±1**	44.0
BCRNN	0±0	0±0	9±3	3±4	32±12	45±11	3±3	1±0	5±5	0±0	9.8
BCRNN+3D	8±14	0±0	8±1	1±5	11±6	0±0	29±12	26±2	38±10	24±11	14.5
IBC	0±0	0±0	9±2	0±0	1±1	16±2	3±2	7±1	14±1	0±0	5.0
IBC+3D	0±0	0±0	10±1	18±9	3±5	50±6	1±1	7±2	15±1	0±0	10.4

讨论问题：IBC和BCRNN在二维图像输入和3D点云输入之间性能表现出现明显差异的可能性是什么？

DP3 模型消融实验

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

Repr.	H	D	P	A	DA	SP	Average
Oracle State	99±2	61±2	**44±3**	94±1	**72±7**	91±8	76.8
Point cloud	**100±0**	**62±4**	43±6	**99±1**	69±4	**97±4**	**78.3**
Image	48±17	50±5	25±4	15±1	43±7	63±3	40.7
Depth	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
Voxel	54±5	33±3	18±2	10±2	17±1	62±6	32.3

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.

Designs	H	D	P	A	DA	SP	Average
DP3	**100±0**	62±4	43±6	**99±1**	69±4	97±4	**78.3**
w/o cropping	98±1	**69±3**	14±1	19±9	32±6	40±2	45.3
w/o LayerNorm	**100±0**	56±4	44±3	96±2	51±3	91±5	73.0
w/o sample pred	68±3	67±8	37±12	96±2	58±9	76±9	67.0
w/o projection	**100±0**	61±2	**47±3**	**99±1**	60±8	**99±2**	77.7
w/ color	**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

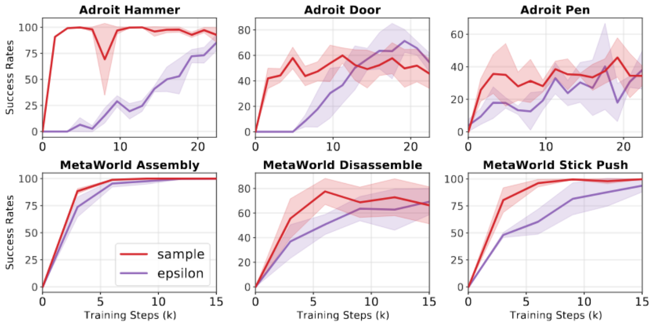


Fig. 7: **Learning curves of DP3 with sample prediction and epsilon prediction.** With sample prediction, DP3 generally converges faster, while epsilon prediction is also competitive.

Sample:
去噪过程对样本进行预测

Epsilon:
去噪过程对噪声进行预测

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.

Encoders	H	D	P	A	DA	SP	Average
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 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.

Encoders	Conv	w/ T-Net	w/ BN	1024 Dim	Average
PointNet	✓	✓	✓	✓	15.7
	✗	✓	✓	✓	15.7
	✓	✗	✓	✓	16.0
	✗	✗	✓	✓	26.0
	✗	✓	✓	✗	18.2
Turnaroud!	✓	✗	✗	✓	**72.5**
	✗	✗	✓	✗	19.8
	✗	✓	✗	✗	26.8
	✗	✗	✗	✗	**72.3**

T-Net:
Transformation Network
进行点云空间变换的网络

BN:
Batch Normalization
一定会损失信息

DP3 模型泛化实验

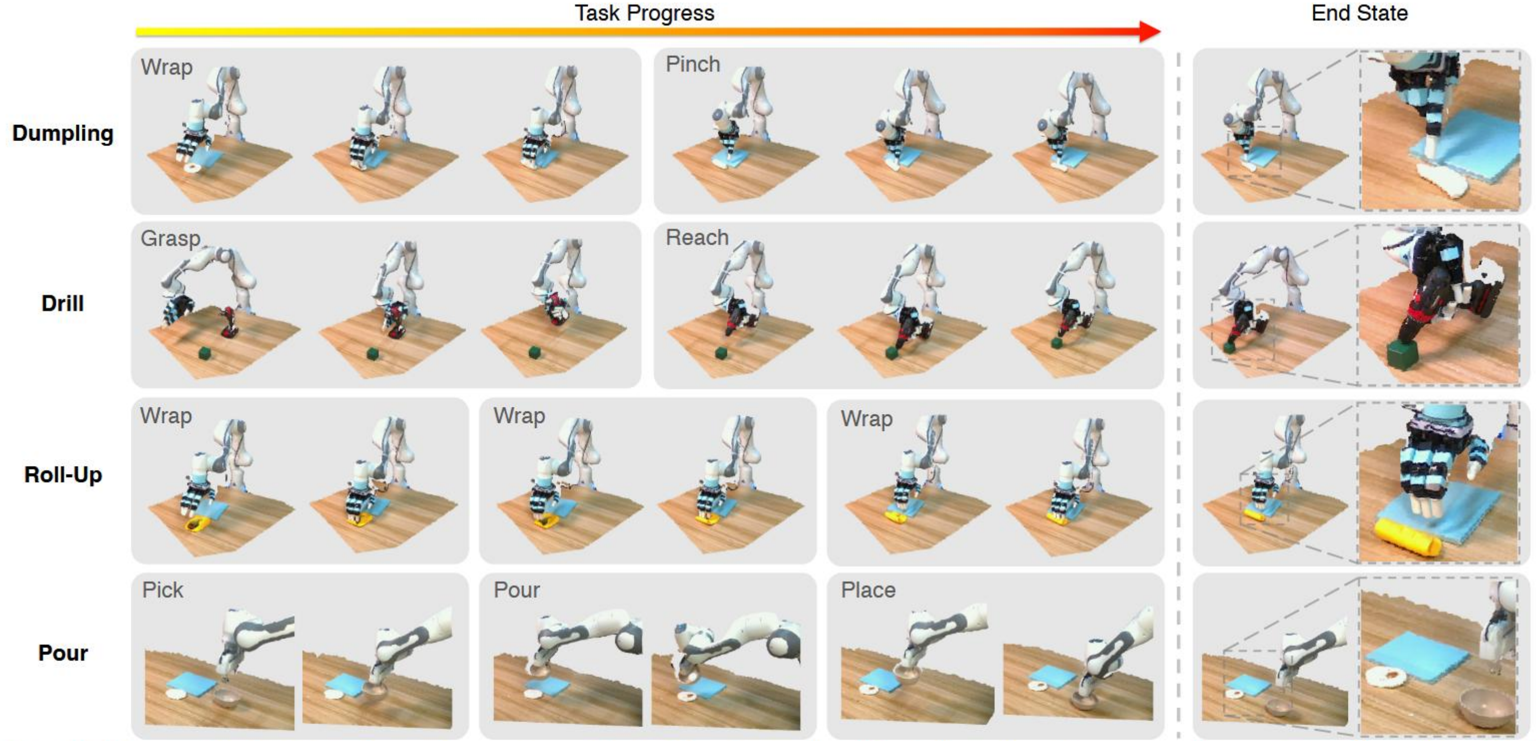
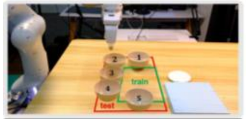


Fig. 10: **Our real robot benchmark** consists of 4 challenging tasks. The Allegro hand is required to make a ***Dumpling***, ***Drill*** the cube, and make a ***Roll-Up***. The gripper is required to ***Pour*** dried meat floss in the bowl. Each task contains multiple stages. We visualize the point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

TABLE VIII: **Main results for real robot experiments.** Each task is evaluated with 10 trials. given only *40* demonstrations.

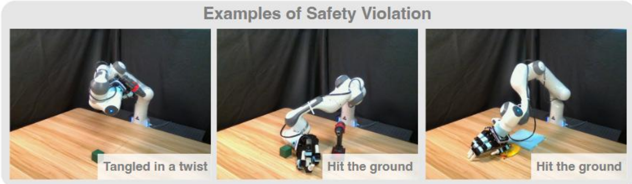
Real Robot	Roll-Up	Dumpling	Drill	Pour	Average
Diffusion Policy	0	30	70	40	35.0±25.0
Diffusion Policy (Depth)	40	20	10	10	20.0±12.2
DP3	**90**	**70**	**80**	**100**	**85.0±11.2**

TABLE 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.



Spatial Generalization	1	2	3	4	5
Diffusion Policy	✗	✗	✗	✗	✗
Diffusion Policy (Depth)	✗	✗	✗	✗	✗
DP3	✗	✓	✓	✓	✓

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



Safety Violation Rate ↓	Roll-Up	Dumpling	Drill	Pour	Average
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 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.

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

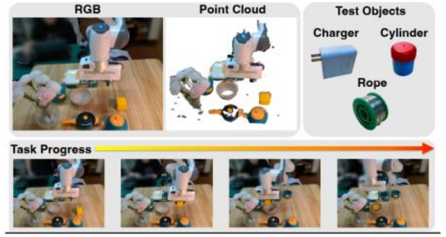
TABLE XI: **Instance generalization on Drill.** We replace the cube used in Drill with five objects in varied sizes from our daily life. Each instance is evaluated with one trial.

Instance Generalization	Mouse	Espeon	Cup	Mug	Hand
Diffusion Policy	✗	✗	✗	✗	✓
Diffusion Policy (Depth)	✗	✗	✓	✗	✗
DP3	✓	✓	✓	✓	✓

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

View Generalization	View 1	View 2	View 3
Diffusion Policy	✗	✗	✗
Diffusion Policy (Depth)	✗	✗	✗
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.



Cluttered Scenes	Diffusion Policy	DP3 w/ PointNeXt	DP3 w/ color	DP3
Success Rate	60	0	**80**	**80**

Train with ■ Cube	■	■	■	Charger	Cylinder	Rope
DP3 w/ color	✗	✗	✗	✗	✗	✗
DP3	✓	✓	✓	✓	✓	✓

Two Questions for Discussion

- **1. 批判性问题:**
  - 关于PointNet中各个组成部分的消融实验中，选择不引入T-Net的底层原因是因为选择了相对固定的点云描述视角。如果在实际工况下，单一视角不能支持所需的点云获取，而是选择将多视角点云转换至同一坐标系下（如机器人躯干系下），DP3是否还能保持其高性能？若不能，如何改善？

- **2. 建设性问题：**
  - 作业机器人头部的NEO相机可以直接输出结构光点云图；手眼相机为双目、内外参已经标定。是否可以将DP3部署于作业机器人上，完成简单单任务？如果可以，需要增加哪些工作？

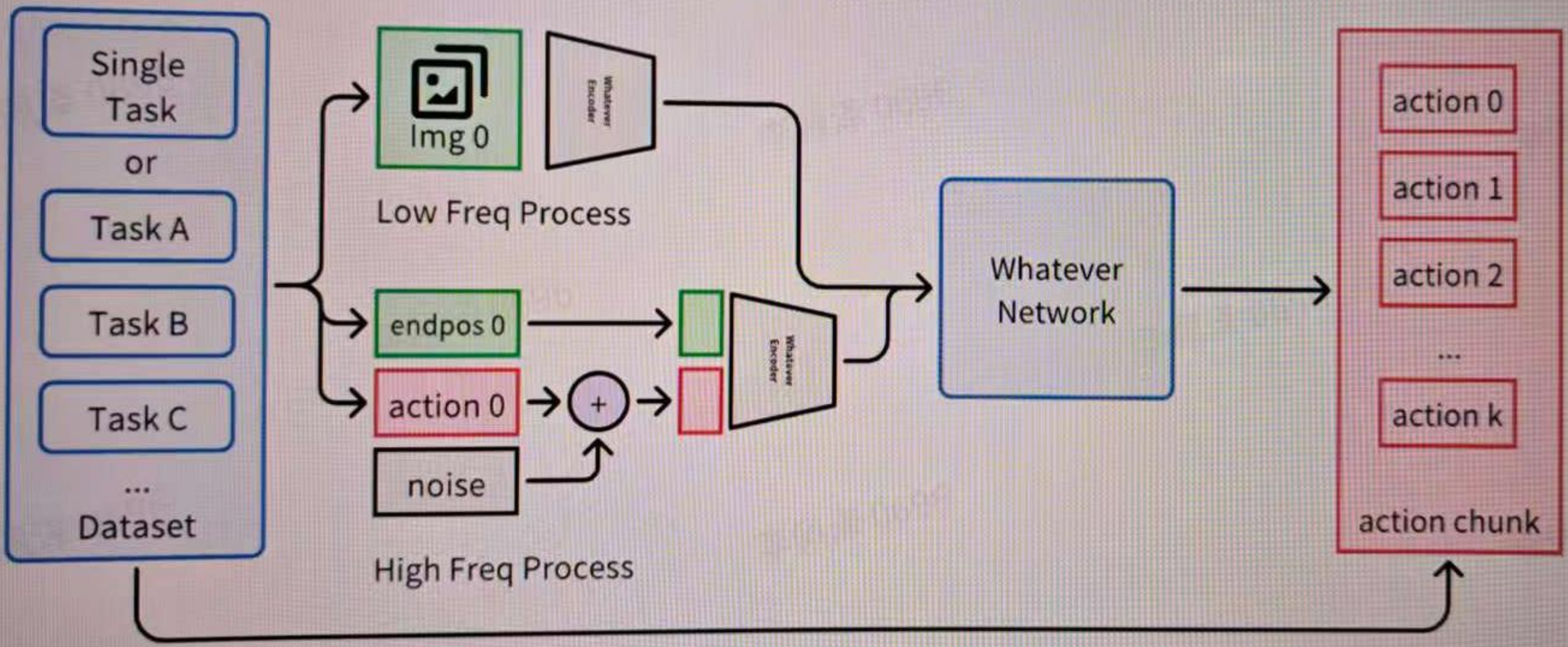
Q & A Section

个人科研分享

Ideal Real-time TeleOp System

Action Chunk Prediction Model

Supervised Training



Action Chunk Prediction Model Deployment (Single Tasks)