Efficient Residual Learning with Mixture-of-Experts for Universal Dexterous Grasping

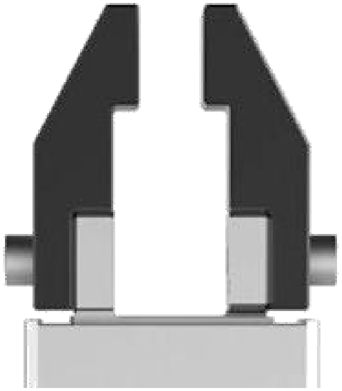
ICLR2025

ZiyeHuang, HaoqiYuan ,YuhuiFu ,ZongqingLu

Peking University

Introduction

- Universal Dexterous Grasping
  - Grasping a wide range of objects with a single policy
- Challenges
  - High degrees of freedom(DoFs)
    - parallel gripper: 1
    - dexterous hand: 12+
  - High variability in object geometry

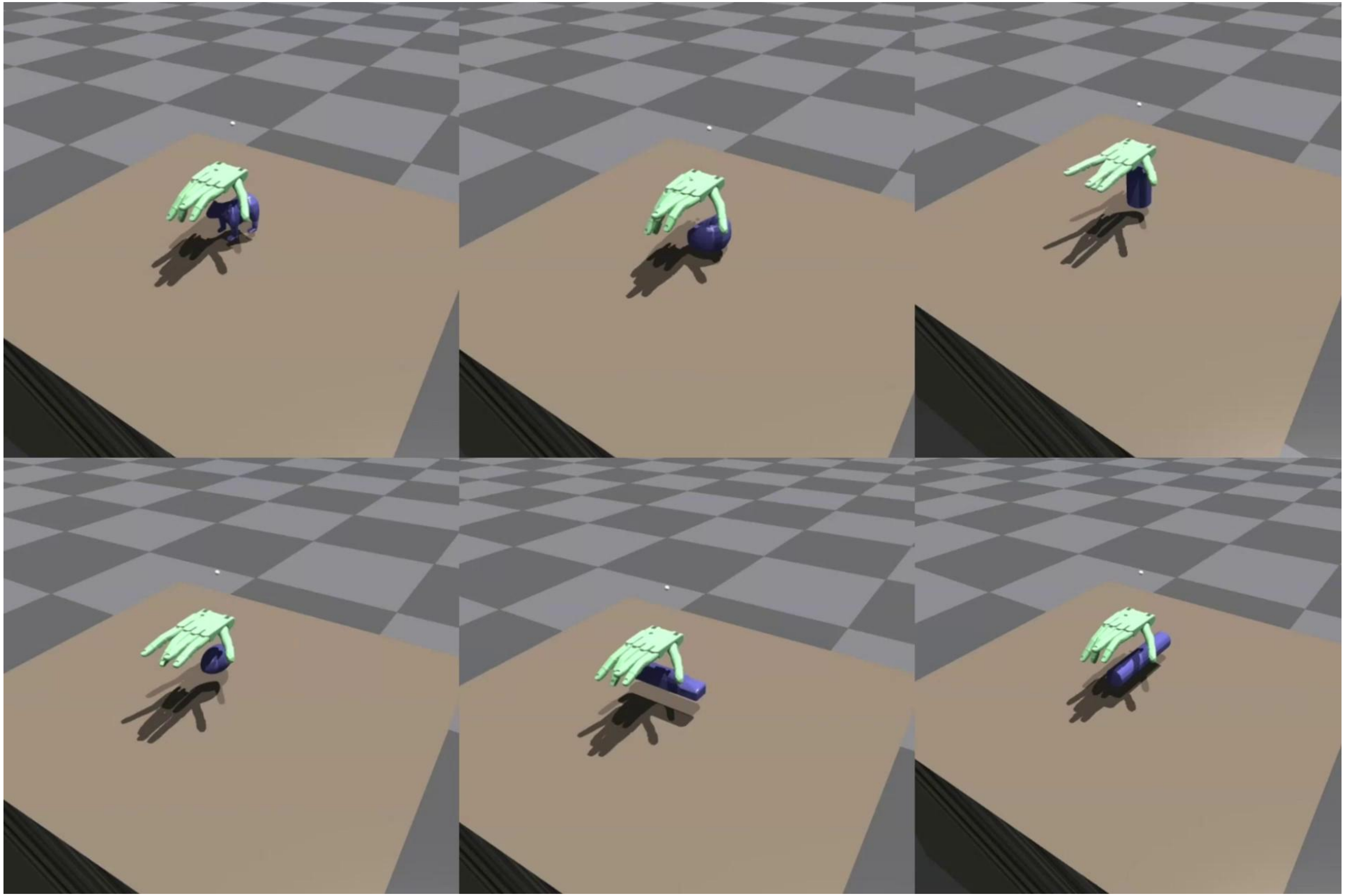


parallel gripper

shadow hand

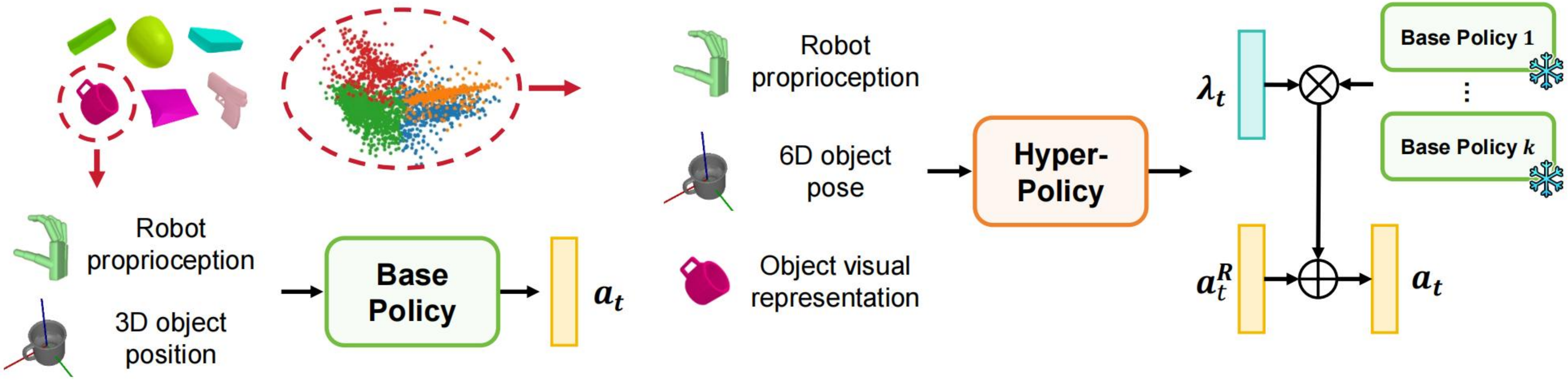
Introduction

- Online RL
- state-based and vision-based
  - state-based: object states are available
  - vision-based: only propriocpetion and vision input(point cloud, RGB, Depth...)
- From state-based to vision-based
  - Directly train a vision-based policy is difficult
  - Solution: teacher-student framework
    - use Dagger(an online imitation learning method)
- We only need a state-based universal grasping policy



Overview

Kmeans → **k objects** → **k base policies** → **1 hyper policy two-stage training**



1. Learning Geometry-Agnostic Experts

2. Residual Multi-Task RL with MoE

Learning Geometry-Agnostic Experts

- Motivation
  - a generalizable base policy facilitates exploration

- Empirical insights
  - Policies trained solely on robot proprioception generalize better
    - (Agarwal et al. 2023)
  - Limiting observations
    - Avoid <u>overfitting to specific object features</u>
    - Learn more generalizable grasping strategies

Learning Geometry-Agnostic Experts

- Principle: *Less observation, more generalizable*
- Use limited observation to train a single-task base policy
- Reward function

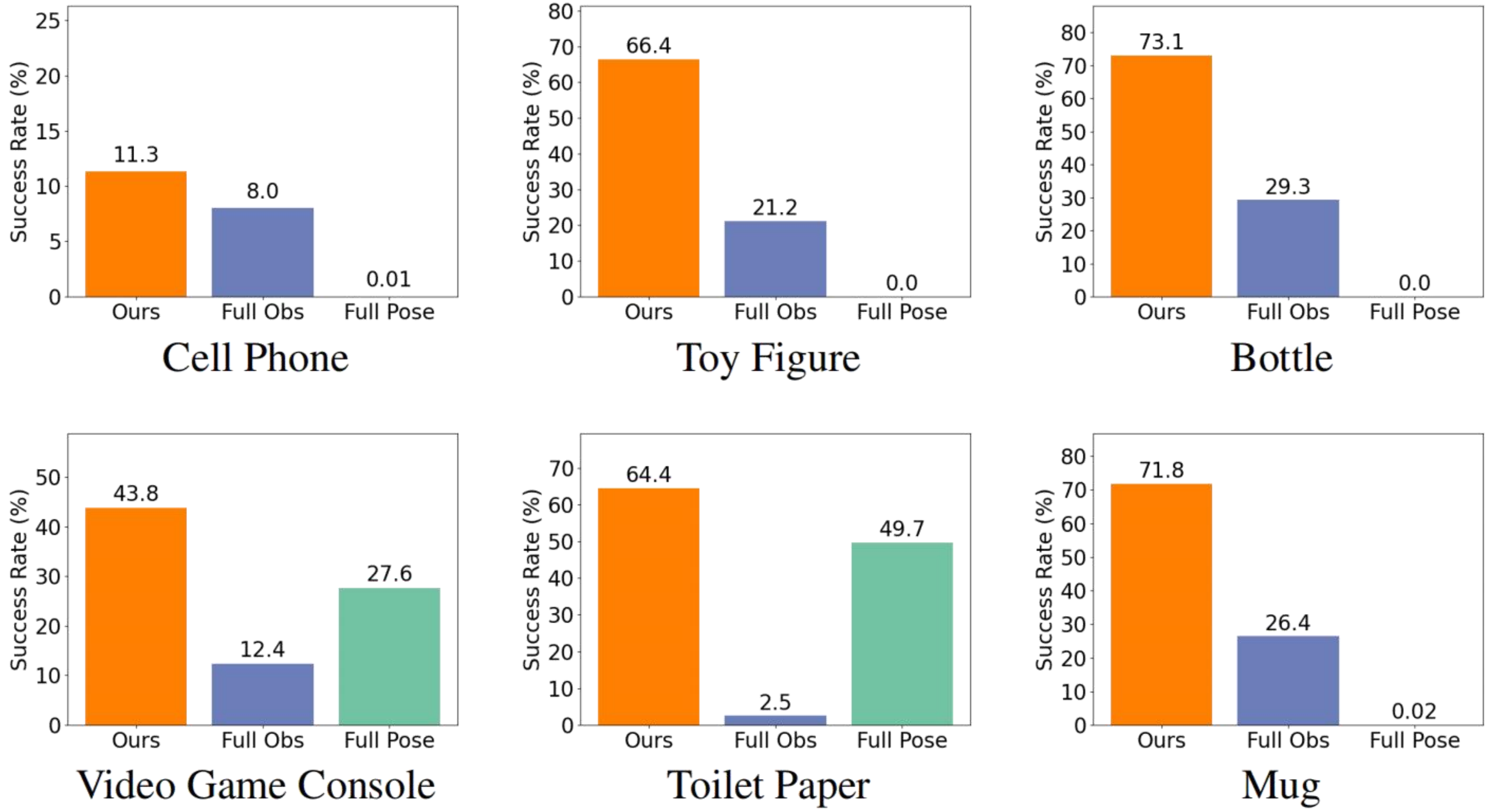
$$r = r^{pose} + r^{task}$$

$$r^{task} = r^{reach} + r^{lift} + r^{move} + r^{bonus}$$

$$r_t^{pose} = -||\boldsymbol{q} - \boldsymbol{q}_t||,$$

- Grasp pose **(R, t, q)**
  - full grasp pose leak object information("where to grasp"...)
  - only use **q** to strike a balance between performance and generalizability

Learning Geometry-Agnostic Experts



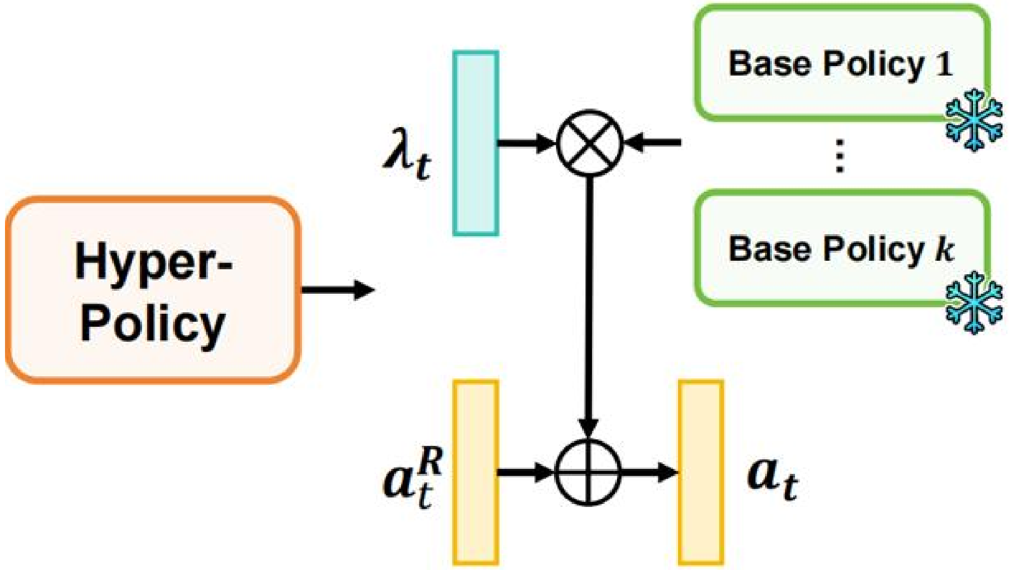
Success rates on the ***whole training set(3200 objects)***

Residual Multi-task RL with MoE

- Hyper policy
  - Combine k *fixed* base policies with normalized weights(MoE)
  - Add residual action

$$a_t = a_t^R + \frac{1}{\|\boldsymbol{\lambda}_t\|}\sum_{i=1}^{k} \lambda_{t,i} a_{t,i}^B,$$

- Residual Learning
  - refine objects that can be successfully grasped
  - explore with base actions
- Mixture of Experts
  - make grasping more natural



Main Results

Method	Train(%)	Test(%) Uns. Obj. Seen Cat.	Test(%) Uns. Cat.
UniDexGrasp	79.4	74.3	70.8
UniDexGrasp++	87.9	84.3	83.1
ResDex (stage-1)	90.6±0.6	89.7±0.8	90.9±0.1
ResDex (stage-2)	**94.6±1.6**	**94.4±1.7**	**95.4±1.0**

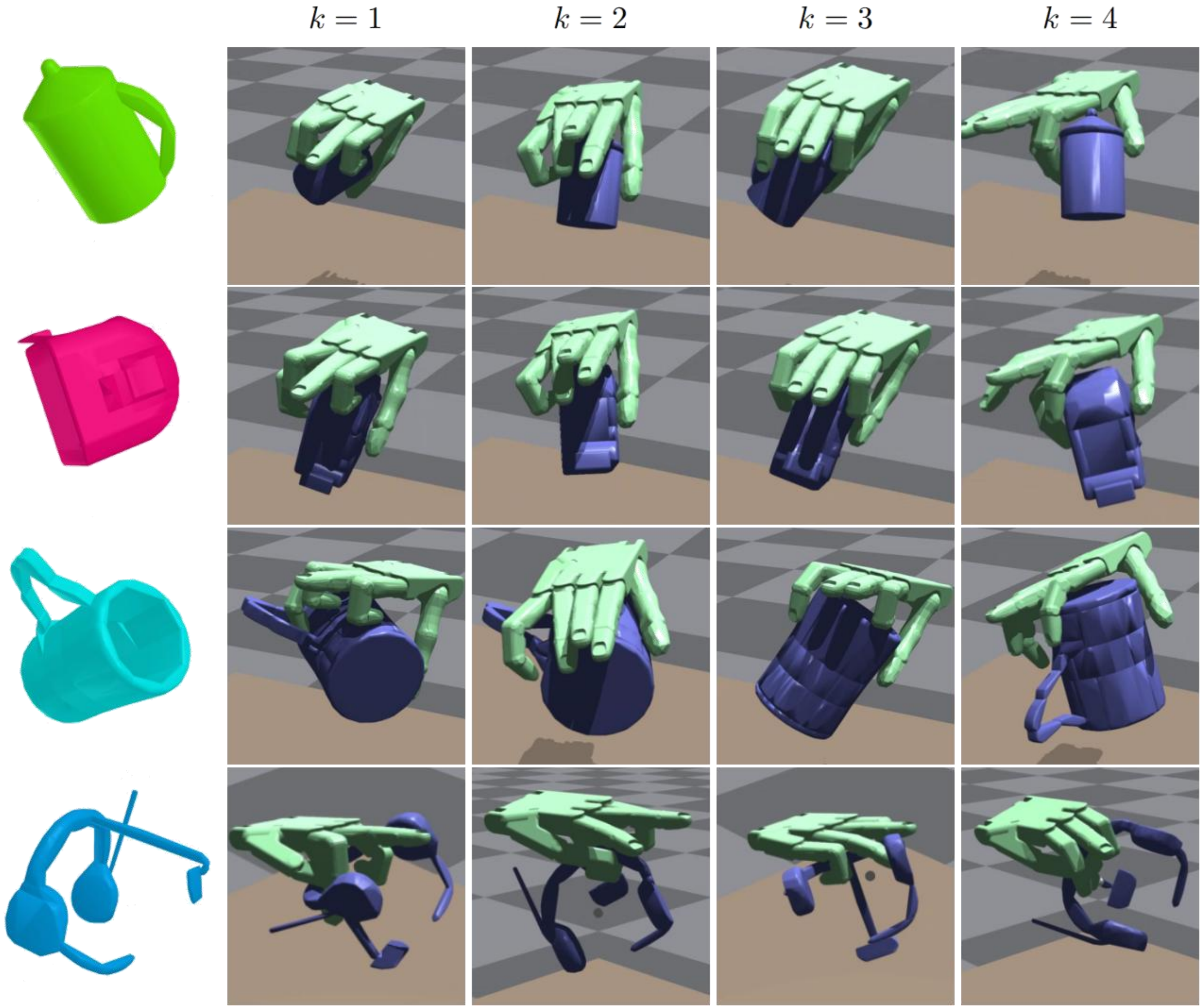
State-based policies(k=4)

Methods	Train(%)	Test(%) Uns. Obj. Seen Cat.	Test(%) Uns. Cat.
UniDexGrasp	73.7	68.6	65.1
UniDexGrasp++	85.4	79.6	76.7
ResDex	**88.8**	**88.5**	**87.2**

Vision-based policies

- Outperform baselines
- No generalization gap

Quality of Grasping poses

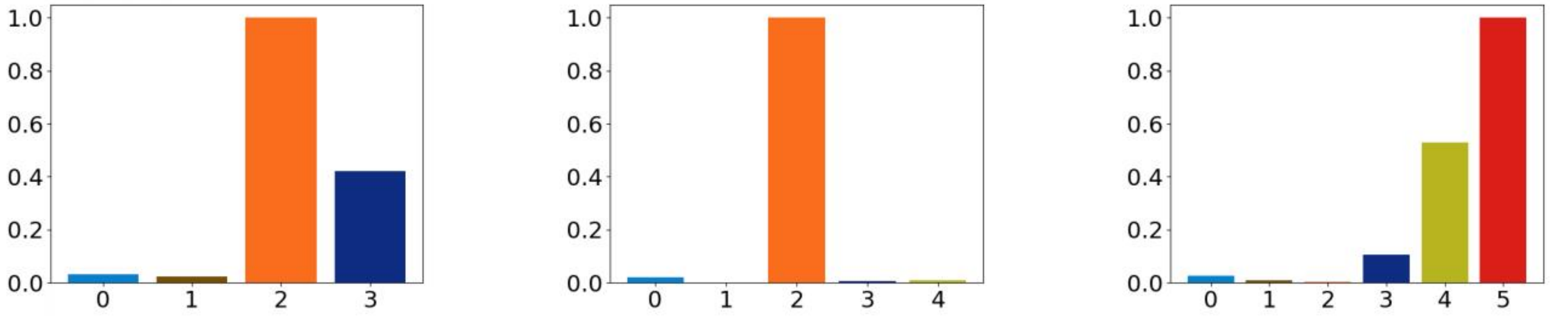


Quality of Grasping poses

$$D = -\sum_{t=1}^{T} r_t^{\text{pose}}$$

Lower D, higer quality of graping poses

Methods	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$
$D \downarrow$	223.6	174.5	194.3	176.3	204.6	176.1



(a) $k=4$ (b) $k=5$ (c) $k=6$

Diversity of the learned λ

Reference

1. Xu, Yinzhen, et al. "Unidexgrasp: Universal robotic dexterous grasping via learning diverse proposal generation and goal-conditioned policy." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023.

2. Wan, Weikang, et al. "Unidexgrasp++: Improving dexterous grasping policy learning via geometry-aware curriculum and iterative generalist-specialist learning." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023.

3. Ross, Stéphane, Geoffrey Gordon, and Drew Bagnell. "A reduction of imitation learning and structured prediction to no-regret online learning." Proceedings of the fourteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings, 2011.

4. Agarwal, Ananye, et al. "Dexterous functional grasping." arXiv preprint arXiv:2312.02975 (2023).

5. Johannink, Tobias, et al. "Residual reinforcement learning for robot control." 2019 international conference on robotics and automation (ICRA). IEEE, 2019.

Thank you!