This CVPR paper is the Open Access version, provided by the Computer Vision Foundation.
Except for this watermark, it is identical to the accepted version;
the final published version of the proceedings is available on IEEE Xplore.

# GeCo-SRT: Geometry-aware Continual Adaptation for Cross-Task Sim-to-Real Transfer

Wenbo Yu[1,*] Wenke Xia[2,3,4,*] Weitao Zhang[2] Di Hu[2,3,4,†]
[1]Beijing Forestry University, Beijing, China
[2]Gaoling School of Artificial Intelligence, Renmin University of China, Beijing, China
[3]Beijing Key Laboratory of Research on Large Models and Intelligent Governance
[4]Engineering Research Center of Next-Generation Intelligent Search and Recommendation, MOE

## Abstract

*Bridging the sim-to-real gap is important for applying low-cost simulation data to real-world robotic systems. However, previous methods are severely limited by treating each transfer as an isolated endeavor, demanding repeated, costly tuning and wasting prior transfer experience. To move beyond isolated sim-to-real, we build a continual cross-task sim-to-real transfer paradigm centered on* ***knowledge accumulation across iterative transfers, thereby enabling effective and efficient adaptation to novel tasks.*** *Thus, we propose GeCo-SRT, a geometry-aware continual adaptation method. It utilizes domain-invariant and task-invariant knowledge from local geometric features as a transferable foundation to accelerate adaptation during subsequent sim-to-real transfers. This method starts with a geometry-aware mixture-of-experts module, which dynamically activates experts to specialize in distinct geometric knowledge to bridge observation sim-to-real gap. Further, the geometry-expert-guided prioritized experience replay module preferentially samples from underutilized experts, refreshing specialized knowledge to combat forgetting and maintain robust cross-task performance. Leveraging knowledge accumulated during iterative transfer, the GeCo-SRT method not only achieves 52% average performance improvement over the baseline, but also demonstrates significant data efficiency for new task adaptation with only 1/6 data. We hope this work inspires approaches for efficient, low-cost cross-task sim-to-real transfer. Our code is released at this https URL*

## 1. Introduction

Leveraging scalable and low-cost simulation data to train real-world robotic systems offers an ideal pathway to significantly accelerate robot learning and deployment [6, 9, 13, 28, 43]. However, the inevitable discrepancies between the simulated model and physical reality, such as unrealistic visual rendering and simplified physical dynamics [7, 14, 22, 35], cause policies learned in simulation to suffer significant performance degradation upon real-world deployment [2, 44]. This phenomenon, known as the "sim-to-real gap", presents a critical and long-standing challenge that impedes the acceleration of real-world robot learning through simulation [43, 44].

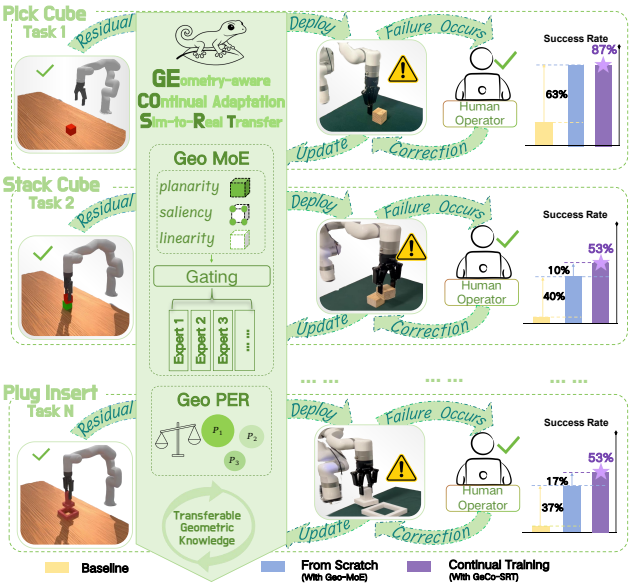



Figure 1. Our GeCo-SRT method accumulates transferable geometric knowledge from sequential sim-to-real tasks (Task 1...$N$), which facilitates effective and efficient transfer to novel tasks.

Several strategies have been proposed to address the sim-to-real gap [10, 23, 44], but they often face limitations in efficiency and generalization. For instance, System Identification seeks to model real-world physics but is labor-intensive and struggles with complex dynamics [25, 34].

*Equal contribution.
†Corresponding author.

While Domain Randomization offers an alternative by training policies across diverse simulations [24, 36], its effectiveness hinges on manually-configured randomization ranges that may not fully capture real-world properties. Although recent data-driven transfer methods [12, 27] mitigate the reliance on manual design, they remain limited to single-task adaptation. This highlights a common drawback across existing approaches: they treat each sim-to-real transfer as an isolated process, which inevitably demands repeated and costly efforts for new task adaptation. This limitation points to a critical but under-explored challenge: how to *leverage past experiences to facilitate effective and efficient sim-to-real transfer on new tasks*, rather than requiring isolated, task-specific tuning from scratch for each.

To move beyond the current paradigm of isolated sim-to-real transfer, we propose a new continual cross-task sim-to-real transfer paradigm. As shown in Figure 1, this paradigm is designed to achieve continual knowledge accumulation across iterative transfers from previous experiences (e.g., "pick cube"), thereby enabling effective and efficient adaptation to novel tasks (e.g., "stack cube"). This process hinges on a robust medium for transferable knowledge that bridges both tasks and domains. To this end, we leverage local geometric features as the primary source of transferable knowledge due to their crucial dual invariance. On the one hand, local geometric features are domain-invariant, offering a robust solution to the sim-to-real gap. Different from visual perception, which is highly sensitive to domain shifts in texture and material properties, local geometric features (e.g., surface normals) derived from 3D data are structurally consistent across simulation and reality. On the other hand, there are also task-invariant, as these primitives are shared across diverse manipulation tasks. For example, the same flat surface of the cube that informs a grasping pose for the "pick cube" task is also the critical target feature for the "stack cube" task. This dual invariance makes local geometric features a robust knowledge medium, enabling effective generalization and continual knowledge accumulation.

Building on this premise, we introduce the GeCo-SRT method, which is a geometry-aware continual adaptation method to achieve continual knowledge accumulation across iterative sim-to-real transfers. Concretely, our GeCo-SRT method employs the human-in-the-loop method [16] to reframe sim-to-real transfer as learning from human correction trajectories. This process quantifies the sim-to-real gap into a continuously growing replay buffer, enabling knowledge accumulation across tasks. Further, we propose the Geometry-aware Mixture-of-Experts (Geo-MoE) module, which serves as a reusable perception residual network for point cloud inputs to bridge the observation gap. This module first extracts local geometric features from the input point cloud and utilizes these local geometric features as cues to dynamically activate experts, thereby achieving distinct knowledge reuse across different tasks. Furthermore, to protect this specialized geometric knowledge within the experts during continual learning, we introduce a Geometry-expert-guided Prioritized Experience Replay (Geo-PER) module. This strategy combats catastrophic forgetting by dynamically prioritizing historical samples based on expert utility, ensuring the model retains its ability to process key geometric structures from previous tasks while adapting to new ones. Through our GeCo-SRT method, we achieve effective and efficient sim-to-real transfer across various tasks by continually leveraging geometric knowledge extracted from prior sim-to-real deployments.

In this work, we establish a sequence of 4 sim-to-real robotic manipulation tasks to evaluate continual sim-to-real transfer. The results of single-task and continual cross-task sim-to-real transfer experiments prove the effectiveness and efficiency of our geometry-aware continual adaptation method. Specifically, our method ultimately outperforms the baseline by an average of 52% while demonstrating remarkable data efficiency, matching the baseline's success rate with only 16.7% of the data. This validates that our geometry-aware method can effectively extract reusable knowledge for continual sim-to-real transfer. We hope this work will inspire future research on effective and efficient cross-task sim-to-real transfer method exploration.

## 2. Related Work

### 2.1. Robot Learning via Sim-to-Real Transfer

Bridging the gap between simulation and reality remains a fundamental and extensively studied challenge in robot learning [3, 4, 26, 31]. While System Identification (SysID) seeks to model the physical parameters of the real system, it relies on pre-defined model structures and extensive real-world data collection for parameter tuning [25]. However, this process is highly labor-intensive and proves inadequate for modeling complex, non-linear dynamics. Alternatively, Domain Randomization trains policies across a wide distribution of simulated variations, aiming to treat the real world as another variant [5, 24]. This approach requires expert-driven tuning of the randomization ranges, which may fail to effectively cover the real-world system dynamics. Alternatively, Transic [16] employs online correction trajectories to achieve domain-agnostic sim-to-real transfer without explicit modeling. However, these approaches are limited to single-task sim-to-real transfer, which makes the adaptation of each new task from simulation to the real world a costly endeavor. In this work, we propose a geometry-aware continual adaptation method to accumulate geometric knowledge from previous sim-to-real transfer for efficient novel task adaptation.

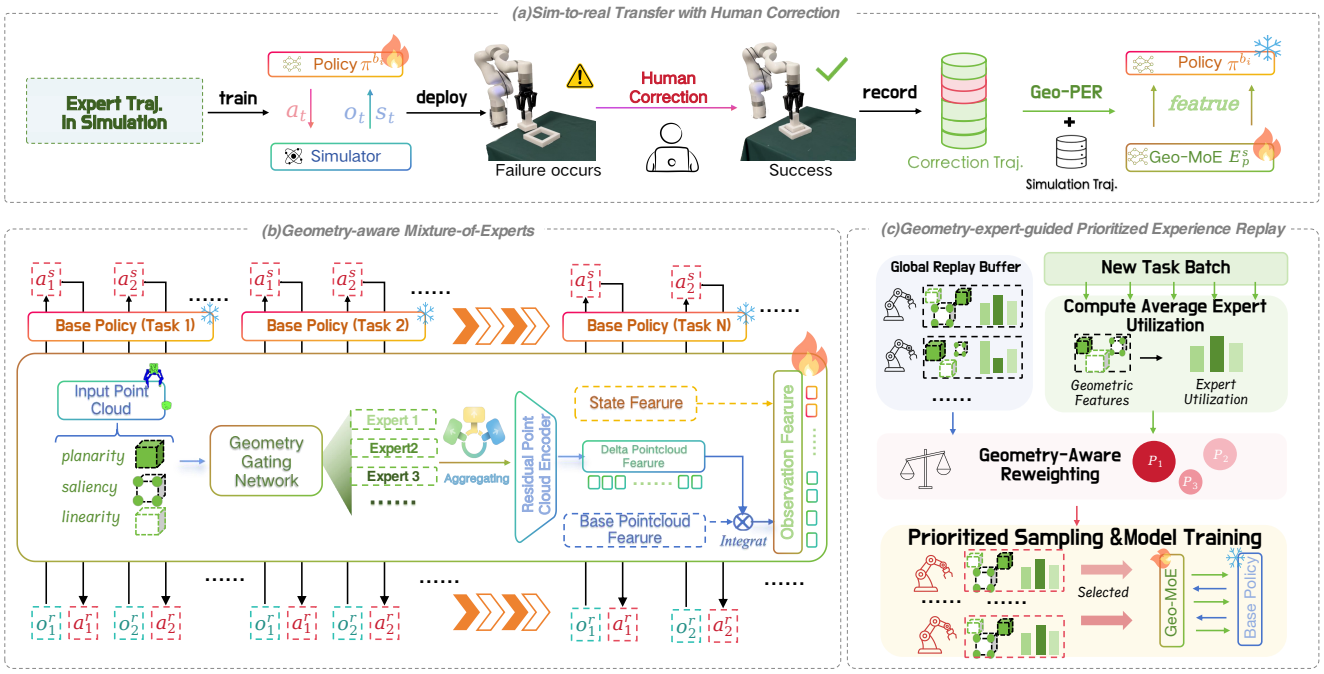



Figure 2. Overview of our geometry-aware continual adaptation method. (a) demonstrates the sim-to-real transfer with human correction pipeline, which translates sim-to-real heuristics into learnable human correction trajectories. (b) represents the Geometry-aware Mixture-of-Experts module, which extracts local geometric features to dynamically activate experts, enabling them to specialize in distinct geometric knowledge for effective sim-to-real transfer. (c) illustrates the Geometry-expert-guided Prioritized Experience Replay, which leverages expert utilization to prioritize historical data, thereby mitigating catastrophic forgetting and promoting knowledge reuse.

### 2.2. Continual Learning in Robotic Manipulation

Continual learning enables robots to acquire sequential skills without catastrophic forgetting, which is crucial for adaptation in dynamic environments. To mitigate forgetting, replay-based methods [32, 33, 44] interleave stored past data during new task training, while generative approaches [11, 18, 39–41] utilize pseudo-samples to reduce memory overhead. Alternatively, regularization-based methods [17, 19] protect prior knowledge by constraining updates to parameters critical for previous tasks. However, the high-dimensional spaces and long-horizon dependencies of robotic manipulation complicate the direct application of these methods. While recent efforts have introduced dedicated benchmarks like LIBERO [21] and specialized algorithmic paradigms [15, 29, 37, 42], they focus primarily on imitation learning. In contrast, this work introduces CL into cross-task sim-to-real transfer. We propose a geometry-expert-guided prioritized experience replay module that leverages geometric knowledge to reweight samples, facilitating efficient novel task adaptation while simultaneously alleviating catastrophic forgetting.

## 3. Geometry-aware Continual Adaptation

The persistent sim-to-real gap presents a significant challenge for robotic policy sim-to-real transfer. While prior work often addresses sim-to-real transfer as an individual, isolated process, we propose GeCo-SRT, a geometry-aware continual adaptation method, which is designed to continually accumulate knowledge across iterative sim-to-real transfers for effective and efficient adaptation to novel tasks. To achieve this, our method is designed to address two fundamental questions:

- How to convert human-designed sim-to-real heuristics into a learnable format for neural networks?
- What specific knowledge is transferable across the iterative sim-to-real transfer process?

To address the first problem, our GeCo-SRT method proposes a human-in-the-loop method to convert sim-to-real transfer as learning from human correction trajectories, making the process suitable for iterative continual learning as shown in Section 3.1.

In response to the second question, we posit that local geometric features serve as the core transferable knowledge. We therefore propose the Geometry-aware Mixture-of-Experts (Geo-MoE) module to extract and reuse these features for efficient cross-task adaptation as shown in Section 3.2. Finally, we introduce the Geometry-expert-guided Prioritized Experience Replay (Geo-PER) in Section 3.3 to protect this specialized knowledge by dynamically prioritizing historical samples, ensuring its retention during continual adaptation. The method is shown in Figure 2.

### 3.1. Sim-to-real Transfer with Human Correction

For each task $i$, we train a base policy $\pi^{b_i}$ in simulation using expert trajectories with point cloud observations. This policy consists of a point cloud encoder $E_p^{b_i}$ and a diffusion policy head $\pi_h^{b_i}$. We then deploy the base policy to the real-world environment and collect the human corrections dataset $\mathcal{D}_{real}^{h_i}$ via a shared-autonomy framework. Finally, to bridge the sim-to-real gap, we introduce a shared perception residual module $E_p^s$. During the sim-to-real transfer, this shared module is continually updated using a hybrid replay buffer that draws from both the simulation dataset and the newly collected real-world correction dataset $\mathcal{D}_{real}^{h_i}$.

#### 3.1.1. Base Policy Training in Simulation

To bypass the drawbacks of RGB (e.g., vulnerability to camera poses as well as discrepancies between synthetic and real images), we use 3D point clouds as the primary visual modality to better facilitate sim-to-real transfer.

For task $i$, we generate an expert dataset $D_{sim}^{e_i}$ using a motion planner in simulation. For observations, we capture depth images from two cameras, which are then transformed and aggregated into a complete scene point cloud $P$ in the robot's base frame. This process mirrors our real-robot setup to minimize the sim-to-real gap. The resulting $P$ is then downsampled via Farthest Point Sampling to serve as the policy network input.

We train our base diffusion policy $\pi^{b_i}$, using behavior cloning on the expert dataset $D_{sim}^{e_i}$. This policy is parameterized as the noise prediction network $\epsilon_\theta$. It is optimized using the simplified L2 diffusion loss:

$$\mathcal{L}_{\text{diff}} = \mathbb{E}_{(a_t,o_t)\sim\mathcal{D}_{sim}^{e_i},\epsilon\sim\mathcal{N}(0,I)}\left[\left\|\epsilon - \epsilon_\theta(a_t^k, k, o_t)\right\|^2\right] \quad (1)$$

where $o_t$ includes the proprioceptive state $s_t$ and point cloud $P$, $a_t$ is the corresponding expert action.

#### 3.1.2. Continual Learning with Human Correction

For each task $i$, we deploy the base policy $\pi^{b_i}$ in the real world under human supervision via a shared-autonomy framework. At each timestep $t$, if the operator anticipates a failure, they intervene with a teleoperated action $a_t \leftarrow a_t^h$, setting an indicator $I_t^h$ to true. Otherwise, the base policy acts, $a_t \leftarrow \pi^{b_i}(o_t)$, and $I_t^h$ is set to false. We collect these tuples $(o_t, a_t, I_t^h)$ to form the real-world human correction dataset $\mathcal{D}_{real}^{h_i}$. Further, we merge the correction dataset $\mathcal{D}_{real}^{h_i}$ with the simulation trajectories $D_{sim}^{e_i}$ to form the replay buffer $D_{buf}^{m_i}$. This buffer is used to train the perception residual module $E_p^s$, which aligns the point cloud observation feature to bridge the sim-to-real gap.

As shown in Figure 2(a), while the parameters of the base simulation policy $\pi^{b_i}$ are frozen, the perception residual module $E_p^s$ is maintained as a shared module across all tasks. This allows it to progressively learn and consolidate reusable knowledge from the buffer $D_{buf}^{m_i}$ of each task, enabling effective iterative sim-to-real transfer.

### 3.2. Geometry-Aware Mixture-of-Experts

We posit that local geometric information is the ideal primary source of transferable knowledge for continual sim-to-real transfer due to the crucial dual invariance:

- Domain Invariance: Unlike textures or dynamics, local geometry (e.g., planarity and linearity) remains consistent between simulation and reality, providing a stable bridge for the sim-to-real gap.
- Task Invariance: Geometric primitives (e.g., edges, corners) are shared across distinct manipulation tasks (e.g., "pick cube" vs. "stack cube"), enabling the continual accumulation of knowledge.

Based on this premise, we propose the Geometry-aware Mixture-of-Experts (Geo-MoE) as our perception residual module $E_p^s$. It processes local geometric discrepancies to dynamically activate specialized experts, thereby adapting the feature representation and enabling the reuse of geometric knowledge across tasks.

Concretely, our Geo-MoE module first samples local point groups $g_i$ from the input point cloud $P$ using the k-Nearest Neighbors (k-NN) method [8]. For each group $g_i$, we estimate its local geometric features—planarity, linearity, and saliency—using the local PCA calculation method [1]. The group $g_i$ is then processed by a lightweight gating network $G$ to produce routing weights $w_i = \text{Softmax}(G(g_i))$ for $M$ parallel expert networks. The final processed feature for each group is the weighted sum:

$$g_i' = \sum_{j=1}^{M} w_{i,j} \cdot \text{Expert}_j(g_i). \quad (2)$$

These per-group features $\{g_i'\}$ are ultimately aggregated to form the single corrective residual vector $g_{res}'$. This observation residual, $g_{res}'$, is then concatenated with the output of the frozen base encoder $E_p^{b_i}$ to form a corrected feature vector: $\hat{f} = \text{Concat}(E_p^{b_i}(o), g_{res}')$. Finally, this corrected feature $\hat{f}$ is passed to the diffusion policy head $\pi_h^{b_i}$ to produce the final action $\hat{a} = \pi_h^{b_i}(\hat{f})$.

During training, we keep the base policy $\pi^{b_i}$ frozen and exclusively update the Geo-MoE module $E_p^s$ using the mixed replay buffer $D_{buf}^{m_i}$:

$$L_{total} = MSE(\hat{a}, a) + \alpha L_{balance}, \quad (3)$$

where $L_{balance}$ is a standard auxiliary loss [30, 38] to ensure balanced expert utilization and prevent gating collapse.

### 3.3. Geometry-Expert-Guided Prioritized Experience Replay

In our continual sim-to-real setting, the model adapts sequentially across $N$ tasks. After $N-1$ tasks, the data is

stored in a unified replay buffer $\mathcal{R}$, the union of all preceding per-task mixed buffers: $\mathcal{R} = \bigcup_{j=1}^{N-1} D_{buf}^{m_j}$.

When adapting to the new N-th task, the key challenge is preventing expert-level catastrophic forgetting. Standard Prioritized Experience Replay (PER) methods, which prioritize samples based on task loss, are ill-suited for this. By ignoring expert-specific utility, they neglect idle experts, whose specialized knowledge is subsequently forgotten. To address this deficiency, we introduce Geometry-expert-guided Prioritized Experience Replay (Geo-PER). Our strategy fundamentally shifts the prioritization metric: from task loss to expert utilization. This ensures all experts are periodically refreshed with relevant data from R, even when inactive in the current task N.

Our Geo-PER mechanism operates as follows: First, we store the expert activation vector $W_i = \{w_{i,1}, \ldots, w_{i,M}\}$ associated with each historical sample $i \in \mathcal{R}$. When processing new data from the current task $N$, we compute the average expert utilization $U^{\text{new}} = \{u_1^{\text{new}}, \ldots, u_M^{\text{new}}\}$. We then dynamically update the sampling priority $P_i$ for all historical data $i \in \mathcal{R}$:

$$P_i \propto \sum_{j=1}^{M} w_{i,j} \cdot \frac{1}{u_j^{\text{new}} + \epsilon}, \tag{4}$$

where $w_{i,j}$ is the stored activation of expert $j$ for sample $i$, $u_j^{\text{new}}$ is the current task's average utilization of expert $j$, and $\epsilon$ is a stability constant. This prioritization scheme functions as an active counter-balance to expert imbalance. If an expert $j$ is under-utilized by the new task ($u_j^{\text{new}}$ is low), its reciprocal term $(u_j^{\text{new}} + \epsilon)^{-1}$ becomes large. Consequently, Geo-PER assigns a high sampling probability to historical samples (from tasks $1..N-1$) that strongly activated that expert ($w_{i,j}$ is high). This ensures that the parameters of idle experts are continuously included in gradient updates, thereby mitigating catastrophic forgetting.

# 4. Experiments

To comprehensively evaluate our method, we propose experiments to answer the following questions:

- **Q1:** Could our geometry-aware method promote efficient continual sim-to-real transfer across different tasks? (Section 4.2)
- **Q2:** What are the specific contributions of our Geo-MoE and Geo-PER method on continual cross-task sim-to-real transfer? (Section 4.3)
- **Q3:** Would the task similarity influence the efficiency of knowledge transfer when adapting between different sim-to-real tasks? (Section 4.4)

## 4.1. Implementation Details

For empirical validation, we utilize four robotic manipulation tasks, spanning both the ManiSkill simulation environment and the real world (Fig. 3 (a)): *Pick Cube*, *Stack Cube*, *Pick Banana*, and *Plug Insert*.

Our sim-to-real pipeline starts by training an initial policy in simulation using imitation learning on 2,000 expert trajectories. We then deploy this policy in the real world and gather 60 human correction trajectories. Finally, to ensure effective transfer and mitigate catastrophic forgetting, we train the final policy on a mixed dataset combining these new corrections with a replay buffer from prior tasks.

### 4.1.1. Real-world Setup

Our hardware (Fig. 3 (b)) comprises an Xarm5 arm and a Rotiq2F140 gripper. We utilize a RealSense D435 and a RealSense D435i to obtain the point cloud observation; their point clouds are transformed to the robot's base frame, fused, and cropped within a bounding box to remove background. Human correction trajectories are collected by operators providing real-time interventions via a 3Dconnexion SpaceMouse during execution.

### 4.1.2. Training Details

Our training involves two stages: simulation policy training and continual sim-to-real transfer.

In the simulation phase, we first train the diffusion policy with 3D point cloud encoder as the base simulation policy. The main hyperparameters for this stage were set as follows: learning rate of $3 \times 10^{-4}$, denoise steps are 10, and action chunking size is 8.

During the continual sim-to-real transfer, we employ a continual learning approach to acquire generalizable geometric knowledge to continually bridge the sim-to-real gap. To retain knowledge from previous tasks, we set the mixing in 10% of data from prior tasks (including both simulation and real-world trajectories). Human-corrected data constitutes 95% of the training data in this stage. In this process, the residual learning rate is set to $1 \times 10^{-3}$.

For the model architecture and experience replay strategy, each expert within our Geo-MoE module is trained with learning rate of $1 \times 10^{-3}$. For our Geo-PER strategy, the parameter controlling the degree of prioritization in sampling is set to 0.6. Additionally, a smoothing term to ensure non-zero sampling probability is set to $1 \times 10^{-6}$, and the EMA coefficient for priority updates is set to 0.4.

### 4.1.3. Evaluation Metrics

To comprehensively evaluate the performance of our continual cross-task sim-to-real approach, we employ two key metrics: Success Rate (SR) and Normalized Negative Backward Transfer (N-NBT). SR measures the cross-task forward transfer ability, while N-NBT quantifies the degree of catastrophic forgetting. For each task, we evaluate methods with 30 trials and report the average SR and N-NBT.

Normalized Negative Backward Transfer: Negative Backward Transfer (NBT) [20] fails to accurately reflect

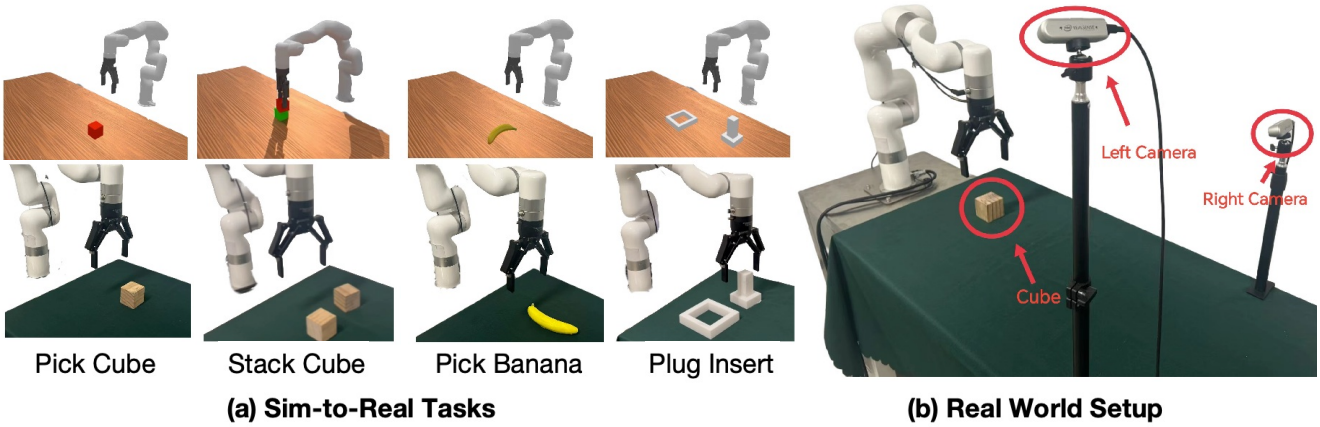



Figure 3. We demonstrate the task setup and real-world setup. (a) illustrates 4 robotic manipulation task in simulation and real-world scenarios. (b) shows that the real-world setup, which consists of a XARM equipped with a robotiq-140 gripper.

forgetting when the initial success rate is low, as is common in sim-to-real. We therefore propose Normalized Negative Backward Transfer (N-NBT) to evaluate the *relative* performance drop.

$$\text{N-NBT}_k = \frac{1}{K-k} \sum_{\tau=k+1}^{K} \left( \frac{C_{\tau,k} - C_{k,k}}{C_{k,k}} \right). \quad (5)$$

A lower N-NBT value indicates better knowledge retention from prior tasks and thus less catastrophic forgetting.

## 4.2. Comparison Results

To evaluate our proposed Geometry-guided Continual Sim-to-real Transfer method, we conduct comprehensive comparisons against several baseline approaches. The evaluation is performed on two primary fronts: (1) single-task sim-to-real transfer and (2) cross-task continual learning.

### 4.2.1. Single-Task Sim-to-Real Transfer

We first evaluate single-task adaptation against the following three baselines:

- **Direct Deploy:** The simulation base policy directly deploys in the real-world environment without fine-tuning.
- **Action Residual:** The action residual module is a residual diffusion policy head network, which takes the observation feature from frozen point cloud encoder and return the residual action. It serves as a baseline to quantify the sim-to-real gap through only action level.
- **Transic [16]:** Transic achieves sim-to-real transfer by leveraging human correction trajectories via a residual network architecture. It serves as a comparison method without specific geometry-aware knowledge.

As shown in Table 1, the Direct Deploy baseline's 3.1% average success rate confirms a substantial sim-to-real gap. The action residual method offers only marginal improvement (7.5%), which strongly suggests that applying residuals only in the action space, without concurrently addressing the significant observation sim-to-real gap, is insufficient to overcome the reality gap.

Table 1. Single-task sim-to-real transfer performance. Our Geo-MoE significantly outperforms all baselines, demonstrating the advantage of specific activated geometric knowledge in bridging the sim-to-real gap.

| Method | Pick Cube | Stack Cube | Pick Banana | Plug Insert | Avg. SR |
|---|---|---|---|---|---|
| Direct Deploy | 5.7% | 0.0% | 6.7% | 0.0% | 3.1% |
| Action Residual | 16.7% | 3.3% | 10.0% | 0.0% | 7.5% |
| Transic[1] | 66.7% | 30.0% | 23.3% | 33.3% | 38.3% |
| Geo-MoE | **80.0%** | **43.3%** | **40.0%** | **36.7%** | **50.0%** |

Nevertheless, our Geo-MoE module consistently outperforms all baselines with a 50.0% average success rate. We attribute this to its dynamic use of geometric knowledge. Unlike methods that address the gap using the overall point cloud, our approach prioritizes distinct geometric properties. This allows the model to learn more fine-grained, geometrically-aware knowledge for precise adaptation.

### 4.2.2. Cross-Task Continual Sim-to-Real Transfer

We further evaluate the method's ability to sequentially learn the sim-to-real gaps of distinct tasks while mitigating catastrophic forgetting. In this continual cross-task sim-to-real transfer setup, the model trains on tasks sequentially, utilizing a dataset that combines current task trajectories with replayed experiences from previous tasks

We compare our method against several baselines:

- **Naive Fine-tuning**: This baseline fine-tunes on the new task's data without any continual learning mechanism.
- **Transic + PER**: Transic combined with standard PER.
- **Geo-MoE + EWC**: Our Geo-MoE module combined with Elastic Weight Consolidation (EWC).
- **Geo-MoE + PER**: Our Geo-Moe module with standard Prioritized Experience Replay.

[1]We reproduced this method, adapting it to align with the distinct point cloud configuration utilized in our approach.

Table 2. Continual learning performance across four sequential sim-to-real tasks (learned in the order: Pick Cube → Stack Cube → Pick Banana → Plug Insert). Our method (GeCo-SRT) achieves the best overall performance, with the highest average success rate and lowest average forgetting. This performance demonstrates Geo-MoE's advantage in leveraging specific geometric knowledge to bridge the sim-to-real gap, combined with Geo-PER's ability to mitigate catastrophic forgetting by continuously updating idle experts.

| Method | Pick Cube | | Stack Cube | | Pick Banana | | Plug Insert | | Avg. | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $SR\uparrow$ | $N-NBT\downarrow$ | $SR\uparrow$ | $N-NBT\downarrow$ | $SR\uparrow$ | $N-NBT\downarrow$ | $SR\uparrow$ | $N-NBT\downarrow$ | $SR\uparrow$ | $N-NBT\downarrow$ |
| Direct Deploy | 5.7% | / | 0.0% | / | 6.7% | / | 0.0% | / | 3.1% | / |
| Naive Fine-tuning | 16.7% | 100% | 3.3% | 100.0% | 13.3% | 100% | 3.3% | / | 9.2% | 75.0% |
| Transic + PER | 76.7% | 81.2% | 30.0% | 72.3% | 20.0% | 66.5% | 33.3% | / | 40.0% | 55.0% |
| Geo-MoE + PER | 83.3% | 34.6% | 50.0% | 76.8% | 46.7% | 7.1% | 43.3% | / | 55.8% | 29.6% |
| Geo-MoE + EWC | 80.0% | 70.8% | 36.7% | 77.3% | 20.0% | 50.0% | 16.7% | / | 38.3% | 49.5% |
| Ours (GeCo-SRT) | **86.7%** | **28.3%** | **53.3%** | **72.0%** | **60.0%** | **5.5%** | **53.3%** | / | **63.3%** | **26.5%** |

The results in Table 2 show that the Naive Fine-tuning baseline suffers from catastrophic forgetting, as new task data overwrites prior knowledge. While conventional continual learning methods like Geo-MoE + EWC and Transic + PER remain suboptimal, exhibiting significant forgetting and limited average success rates (e.g., $\leq 40.0\%$).

A critical insight emerges when comparing Transic + PER with Geo-MoE + PER. Simply replacing the backbone with our Geo-MoE module significantly boosts the average success rate (from 40.0% to 55.7%) and reduces forgetting. This strongly suggests the Geo-MoE architecture is inherently superior for continual learning. Its design, routing information via local geometric features, enables specialized experts to be activated and adapted, facilitating more effective knowledge reuse across tasks.

However, this specialized knowledge requires active protection. Our GeCo-SRT method integrates our proposed Geo-PER strategy, achieving the best-in-class performance (63.3% average success). This refinement works because Geo-PER is tailored to the MoE structure. By shifting the prioritization metric from task-level loss to expert-level utilization, it ensures all experts—even those idle on the current task—are periodically refreshed from the buffer. This actively preserves their specialized knowledge and leads to robust retention.

In summary, our GeCo-SRT method combines a superior architecture for knowledge reuse (Geo-MoE) with a targeted strategy for knowledge protection (Geo-PER), achieving the most robust continual sim-to-real performance.

#### 4.2.3. Cross-Task Transfer Efficiency

We evaluate sample efficiency in Figure 4. We compare two training scenarios: training from scratch (using only the new task's data) and our continual learning approach (transferring from a pre-trained task). Both scenarios are evaluated using 20, 40, and 60 correction trajectories.

The results clearly illustrate the significant benefit of our method. For the Pick Cube task, the model trained via continual learning achieves 76.6% success with only 20 trajectories. This performance dramatically surpasses the from-scratch baseline at the same data point and nearly matches the performance the from-scratch model achieves with 3× the data (60 trajectories). This trend is consistent for the Plug Insert task, where our method with 20 trajectories achieves a performance level that the from-scratch model requires more than 60 trajectories to reach.

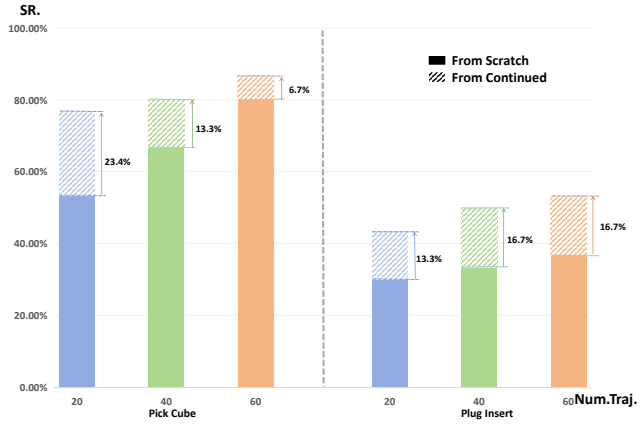



Figure 4. Data efficiency comparison of training "From Scratch" vs. "From Continued". The performance gain (hatched area) highlights our method's ability to utilize transferable geometric knowledge for cross-task sim-to-real transfer, which enhances data efficiency when adapting to new tasks, especially in low-data regimes.

This experiment confirms that our continual learning method significantly improves sample efficiency. The agent's ability to adapt so rapidly is not merely due to parameter initialization; it demonstrates that our method successfully learns and reuses transferable geometric knowledge about the sim-to-real gap. By leveraging this shared cross-task knowledge, the agent can adapt to new tasks with limited correction demonstrations, which is critical for real-world applications.

### 4.3. Ablation Results

Table 3 highlights each component's contribution. The baseline model (Observation Residual ×, MoE ×) performs poorly, and sees negligible improvement when only the MoE module is added (×, ✓). This suggests that the expert architecture is ineffective without informative features to bridge the observation sim-to-real gap.

The most significant gain comes from introducing our Observation Residual component (✓, ×). This change alone causes the average success rate to leap to 45.8% and

Table 3. Ablation study on core components of Geo-MoE. For the Observation Residual component, ✓ indicates using our point cloud encoder to align the observation sim-to-real gap, whereas × denotes using the action residual module. For the MoE component, ✓ indicates using our mixture-of-experts architecture to obtain distinct knowledge, while × represents the baseline without it.

| Observation Residual | MoE | Pick Cube | | Stack Cube | | Pick Banana | | Plug Insert | | Avg. | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $SR\uparrow$ | $N-NBT\downarrow$ | $SR\uparrow$ | $N-NBT\downarrow$ | $SR\uparrow$ | $N-NBT\downarrow$ | $SR\uparrow$ | $N-NBT\downarrow$ | $SR\uparrow$ | $N-NBT\downarrow$ |
| × | × | 16.7% | 100% | 3.3% | 100.0% | 10.0% | 100.0% | 0.0% | / | 7.5% | 75.0% |
| × | ✓ | 16.7% | 86.7% | 3.3% | 100.0% | 13.3% | 75.2% | 3.3% | / | 9.2% | 65.5% |
| ✓ | × | **83.3%** | 37.3% | 46.7% | 96.5% | 23.3% | 14.2% | 30.0% | / | 45.8% | 37.0% |
| ✓ | ✓ | **83.3%** | **34.6%** | **50.0%** | **76.8%** | **46.7%** | **7.1%** | **43.3%** | / | **55.8%** | **29.6%** |

average N-NBT to plummet to 37.0%. This provides a stable feature bridge for visual alignment, effectively mitigating the sim-to-real gap.

Building on this strong Observation Residual component, adding the MoE module (✓, ✓) achieves the best performance. This highlights the MoE's synergistic role: it takes the stable geometric features and enables specialized handling. By routing geometric patterns to distinct experts, the model processes local discrepancies in a more targeted manner. This specialization improves performance on later tasks and further mitigates forgetting.

In summary, the ablation confirms our components are complementary: the Observation Residual component extracts a stable, transferable feature representation based on sim-real invariant geometry, while the MoE module builds on this to specialize adaptation and preserve knowledge.

### 4.4. Cross-task Transfer Results

Finally, we investigate how task similarity impacts the effectiveness of knowledge transfer, particularly in a data-scarce scenario (using only 10 correction trajectories). We compare training "From scratch" versus transferring from various pre-trained source tasks, with results in Table 4.

The results indicate that task geometric similarity is a critical factor for our geometry-aware continual adaptation method. This is evidenced by significant positive transfer when learning the Stack Cube task after the similar Pick Cube task, far outperforming the baseline. Conversely, negative transfer occurs when transferring from the dissimilar PlugInsert task. Notably, the transfer from PickBanana to PlugInsert is also beneficial, which we attribute to the shared geometric challenge of grasping non-cuboid objects.

While the primary insight concerns task similarity, this positive transfer is also highly data-efficient. For instance, the success rate on Stack Cube with only 10 trajectories nearly matches the performance of the from-scratch model trained with 60 trajectories.

This experiment confirms our geometry-aware method learns transferable geometric knowledge. The model's ability to reuse knowledge from geometrically similar tasks, while failing with dissimilar ones, demonstrates it captures and leverages foundational geometric understanding for rapid, low-data adaptation.

Table 4. Impact of task similarity on knowledge transfer. All models use 10 trajectories unless otherwise specified.

| Evaluate Task | Pre-trained Task | SR |
|---|---|---|
| StackCube | PickCube | 40.0% |
| | PickBanana | 33.3% |
| | PlugInsert | 16.7% |
| | From scratch | 26.7% |
| | From scratch (60 traj.) | 43.3% |
| PlugInsert | PickCube | 33.3% |
| | StackCube | 40.0% |
| | PickBanana | 30.0% |
| | From scratch | 23.3% |
| | From scratch (60 traj.) | 36.7% |

## 5. Conclusion

In this work, we build a continual cross-task sim-to-real transfer paradigm for knowledge accumulation across iterative transfers, enabling effective and efficient adaptation to novel tasks. To achieve this, we propose the GeCo-SRT, a geometry-aware continual adaptation method. It extracts domain-invariant and task-invariant knowledge from local geometric features to accelerate subsequent adaptation. This method consists of a geometry-aware mixture-of-experts module to obtain specialized transferable knowledge, and a geometry-expert-guided prioritized experience replay strategy to mitigates catastrophic forgetting with the guidance of expert utilization. Results show a 52% average performance improvement over the baseline and demonstrate significant data efficiency, adapting to new tasks with only 16.7% data. We hope our work provides insight into continual sim-to-real transfer by effectively leveraging local geometric knowledge.

**Discussion and Future Work.** While GeCo-SRT has demonstrated effectiveness and efficiency in cross-task sim-to-real transfer with continual knowledge accumulation, our method primarily focuses on bridging the observation gap by leveraging local geometric features. This specific focus may limit its applicability in non-geometric sim-to-real gaps (e.g., complex dynamics). Furthermore, we believe exploring multiple modalities, such as the semantic relationships between tasks, could provide a promising avenue for effective cross-task sim-to-real transfer.

## 6. Acknowledgement

This work was supported by Beijing Natural Science Foundation (4262050) and the fund for building world-class universities (disciplines) of Renmin University of China.

## References

[1] Evangelos Alexiou, Irene Viola, and Pablo César. Pointpca: Point cloud objective quality assessment using pca-based descriptors. *CoRR*, abs/2111.12663, 2021. 4

[2] Elie Aljalbout, Jiaxu Xing, Angel Romero, Iretiayo Akinola, Caelan Reed Garrett, Eric Heiden, Abhishek Gupta, Tucker Hermans, Yashraj Narang, Dieter Fox, Davide Scaramuzza, and Fabio Ramos. The reality gap in robotics: Challenges, solutions, and best practices, 2025. 1

[3] Karol Arndt, Murtaza Hazara, Ali Ghadirzadeh, and Ville Kyrki. Meta reinforcement learning for sim-to-real domain adaptation, 2019. 2

[4] Konstantinos Bousmalis, Alex Irpan, Paul Wohlhart, Yunfei Bai, Matthew Kelcey, Mrinal Kalakrishnan, Laura Downs, Julian Ibarz, Peter Pastor, Kurt Konolige, Sergey Levine, and Vincent Vanhoucke. Using simulation and domain adaptation to improve efficiency of deep robotic grasping, 2017. 2

[5] Xiaoyu Chen, Jiachen Hu, Chi Jin, Lihong Li, and Liwei Wang. Understanding domain randomization for sim-to-real transfer, 2022. 2

[6] Longchao Da, Justin Turnau, Thirulogasankar Pranav Kutralingam, Alvaro Velasquez, Paulo Shakarian, and Hua Wei. A survey of sim-to-real methods in rl: Progress, prospects and challenges with foundation models, 2025. 1

[7] Ruoxuan Feng, Jiangyu Hu, Wenke Xia, Tianci Gao, Ao Shen, Yuhao Sun, Bin Fang, and Di Hu. Learning unified static-dynamic representation across multiple visuo-tactile sensors. In *The Thirteenth International Conference on Learning Representations*, 2025. 1

[8] Gongde Guo, Hui Wang, David Bell, Yaxin Bi, and Kieran Greer. Knn model-based approach in classification. In *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE*, pages 986–996, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg. 4

[9] Tyler Han, Preet Shah, Sidharth Rajagopal, Yanda Bao, Sanghun Jung, Sidharth Talia, Gabriel Guo, Bryan Xu, Bhaumik Mehta, Emma Romig, Rosario Scalise, and Byron Boots. Wheeled lab: Modern sim2real for low-cost, open-source wheeled robotics, 2025. 1

[10] Tairan He, Jiawei Gao, Wenli Xiao, Yuanhang Zhang, Zi Wang, Jiashun Wang, Zhengyi Luo, Guanqi He, Nikhil Sobanbab, Chaoyi Pan, Zeji Yi, Guannan Qu, Kris Kitani, Jessica Hodgins, Linxi "Jim" Fan, Yuke Zhu, Changliu Liu, and Guanya Shi. Asap: Aligning simulation and real-world physics for learning agile humanoid whole-body skills, 2025. 1

[11] Nick Heppert, Minh Quang Nguyen, and Abhinav Valada. Real2gen: Imitation learning from a single human demonstration with generative foundational models. In *ICRA 2025 Workshop on Foundation Models and Neuro-Symbolic AI for Robotics*, 2025. 3

[12] Xinyue Huo, Lingxi Xie, Hengtong Hu, Wengang Zhou, Houqiang Li, and Qi Tian. Domain-agnostic prior for transfer semantic segmentation, 2022. 2

[13] Stephen James, Paul Wohlhart, Mrinal Kalakrishnan, Dmitry Kalashnikov, Alex Irpan, Julian Ibarz, Sergey Levine, Raia Hadsell, and Konstantinos Bousmalis. Sim-to-real via sim-to-sim: Data-efficient robotic grasping via randomized-to-canonical adaptation networks, 2019. 1

[14] Rae Jeong, Yusuf Aytar, David Khosid, Yuxiang Zhou, Jackie Kay, Thomas Lampe, Konstantinos Bousmalis, and Francesco Nori. Self-supervised sim-to-real adaptation for visual robotic manipulation, 2019. 1

[15] Zeng Jia, Bu Qingwen, Wang Bangjun, Xia Wenke, Chen Li, Dong Hao, Song Haoming, Wang Dong, Hu Di, Luo Ping, Cui Heming, Zhao Bin, Li Xuelong, Qiao Yu, and Li Hongyang. Learning manipulation by predicting interaction. In *Proceedings of Robotics: Science and Systems (RSS)*, 2024. 3

[16] Yunfan Jiang, Chen Wang, Ruohan Zhang, Jiajun Wu, and Li Fei-Fei. Transic: Sim-to-real policy transfer by learning from online correction, 2024. 2, 6

[17] Hyunyoung Jung, Zhaoyuan Gu, Ye Zhao, Hae-Won Park, and Sehoon Ha. Ppf: Pre-training and preservative fine-tuning of humanoid locomotion via model-assumption-based regularization. *IEEE Robotics and Automation Letters*, 10 (11):11466–11473, 2025. 3

[18] Pushkal Katara, Zhou Xian, and Katerina Fragkiadaki. Gen2sim: Scaling up robot learning in simulation with generative models. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6672–6679, 2024. 3

[19] Luzia Knoedler, Chadi Salmi, Hai Zhu, Bruno Brito, and Javier Alonso-Mora. Improving pedestrian prediction models with self-supervised continual learning, 2022. 3

[20] Bo Liu, Yifeng Zhu, Chongkai Gao, Yihao Feng, Qiang Liu, Yuke Zhu, and Peter Stone. Libero: Benchmarking knowledge transfer for lifelong robot learning. *Advances in Neural Information Processing Systems*, 36:44776–44791, 2023. 5

[21] Bo Liu, Yifeng Zhu, Chongkai Gao, Yihao Feng, Qiang Liu, Yuke Zhu, and Peter Stone. Libero: Benchmarking knowledge transfer for lifelong robot learning, 2023. 3

[22] Jingxian Lu, Wenke Xia, Yuxuan Wu, Zhiwu Lu, and Di Hu. When would vision-proprioception policies fail in robotic manipulation?, 2026. 1

[23] Liqian Ma, Jiaojiao Meng, Shuntao Liu, Weihang Chen, Jing Xu, and Rui Chen. Sim2real$^2$: Actively building explicit physics model for precise articulated object manipulation, 2023. 1

[24] Bhairav Mehta, Manfred Diaz, Florian Golemo, Christopher J. Pal, and Liam Paull. Active domain randomization, 2019. 2

[25] Marius Memmel, Andrew Wagenmaker, Chuning Zhu, Patrick Yin, Dieter Fox, and Abhishek Gupta. Asid: Active exploration for system identification in robotic manipulation, 2024. 1, 2

[26] Tongzhou Mu, Zhan Ling, Fanbo Xiang, Derek Yang, Xuanlin Li, Stone Tao, Zhiao Huang, Zhiwei Jia, and Hao Su. Maniskill: Generalizable manipulation skill benchmark with large-scale demonstrations, 2021. 2

[27] Martin Pecka, Karel Zimmermann, Matěj Petrlík, and Tomáš Svoboda. Data-driven policy transfer with imprecise perception simulation. *IEEE Robotics and Automation Letters*, 3 (4):3916–3921, 2018. 2

[28] Xue Bin Peng, Marcin Andrychowicz, Wojciech Zaremba, and Pieter Abbeel. Sim-to-real transfer of robotic control with dynamics randomization. In *2018 IEEE international conference on robotics and automation (ICRA)*, pages 3803–3810. IEEE, 2018. 1

[29] Sam Powers, Abhinav Gupta, and Chris Paxton. Evaluating continual learning on a home robot, 2023. 3

[30] Samyam Rajbhandari, Conglong Li, Zhewei Yao, Minjia Zhang, Reza Yazdani Aminabadi, Ammar Ahmad Awan, Jeff Rasley, and Yuxiong He. DeepSpeed-MoE: Advancing mixture-of-experts inference and training to power next-generation AI scale. In *Proceedings of the 39th International Conference on Machine Learning*, pages 18332–18346. PMLR, 2022. 4

[31] Erica Salvato, Gianfranco Fenu, Eric Medvet, and Felice Andrea Pellegrino. Crossing the reality gap: A survey on sim-to-real transferability of robot controllers in reinforcement learning. *IEEE Access*, 9:153171–153187, 2021. 2

[32] James Smith, Yen-Chang Hsu, Jonathan Balloch, Yilin Shen, Hongxia Jin, and Zsolt Kira. Always be dreaming: A new approach for data-free class-incremental learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 9374–9384, 2021. 3

[33] James Smith, Yen-Chang Hsu, Jonathan Balloch, Yilin Shen, Hongxia Jin, and Zsolt Kira. Always be dreaming: A new approach for data-free class-incremental learning, 2021. 3

[34] Nikhil Sobanbabu, Guanqi He, Tairan He, Yuxiang Yang, and Guanya Shi. Sampling-based system identification with active exploration for legged robot sim2real learning, 2025. 1

[35] Marcel Torne, Anthony Simeonov, Zechu Li, April Chan, Tao Chen, Abhishek Gupta, and Pulkit Agrawal. Reconciling reality through simulation: A real-to-sim-to-real approach for robust manipulation. *arXiv preprint arXiv:2403.03949*, 2024. 1

[36] Riccardo Volpi, Diane Larlus, and Gregory Rogez. Continual adaptation of visual representations via domain randomization and meta-learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4443–4453, 2021. 2

[37] Weikang Wan, Yifeng Zhu, Rutav Shah, and Yuke Zhu. Lotus: Continual imitation learning for robot manipulation through unsupervised skill discovery, 2024. 3

[38] Lean Wang, Huazuo Gao, Chenggang Zhao, Xu Sun, and Damai Dai. Auxiliary-loss-free load balancing strategy for mixture-of-experts, 2024. 4

[39] Xinyi Wang, Zitao Wang, and Wei Hu. Serial contrastive knowledge distillation for continual few-shot relation extraction, 2023. 3

[40] Haojun Xia, Zhen Zheng, Yuchao Li, Donglin Zhuang, Zhongzhu Zhou, Xiafei Qiu, Yong Li, Wei Lin, and Shuaiwen Leon Song. Flash-llm: Enabling cost-effective and highly-efficient large generative model inference with unstructured sparsity, 2023.

[41] Wenke Xia, Ruoxuan Feng, Dong Wang, and Di Hu. Phoenix: A motion-based self-reflection framework for fine-grained robotic action correction. *arXiv preprint arXiv:2504.14588*, 2025. 3

[42] Wenke Xia, Yichu Yang, Hongtao Wu, Xiao Ma, Tao Kong, and Di Hu. Human-assisted robotic policy refinement via action preference optimization. *arXiv preprint arXiv:2506.07127*, 2025. 3

[43] Wenshuai Zhao, Jorge Peña Queralta, and Tomi Westerlund. Sim-to-real transfer in deep reinforcement learning for robotics: a survey. In *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 737–744, 2020. 1

[44] Wenshuai Zhao, Jorge Peña Queralta, and Tomi Westerlund. Sim-to-real transfer in deep reinforcement learning for robotics: a survey. *CoRR*, abs/2009.13303, 2020. 1, 3