2026 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
September 27 - October 1, 2026. Pittsburgh, PA, USA

Relational Dexterity: Transferring Human Dexterity via Relational Geometric Prior

Jiaying Chen[1], Bingheng Zhang[1], Qingtao Liu[1], Yiyang Ge[1], Wenzhe Ouyang[1], Anjun Chen[2] and Qi Ye[1,†]

***Abstract*— While large-scale human demonstrations provide valuable guidance for learning dexterous manipulation, transferring these skills to robotic hands remains challenging due to the inherent embodiment gap between human and robot morphologies. Existing methods typically rely on absolute keypoint representations that discard fine-grained finger synergies, and employ explicit kinematic retargeting to accelerate the learning process, which can induce kinematic inconsistency and physical conflicts. We propose ReDex, a framework that shifts the focus from absolute spatial tracking to intrinsic interaction geometry. ReDex first extracts relational manipulation vectors from human data and compresses them into a smooth latent space, establishing a unified geometric prior that is invariant across embodiments and capture hand synergies. Guided by this prior, we introduce a curriculum reinforcement learning strategy with a state pool constrained by the latent space for episode initialization, which confines exploration to physically valid, human-like motion manifolds and stabilizes policy learning. Extensive experiments demonstrate that ReDex consistently outperforms current state-of-the-art methods across diverse manipulation modes, with robust cross-embodiment transfer and real-world feasibility. For more information, please visit `https://psycho811.github.io/ReDex`.**

I. Introduction

Equipping robotic hands with human-like dexterity remains one of the core challenges in Embodied AI. Given the high-dimensional action spaces and complex contact dynamics, leveraging large-scale human demonstrations has emerged as a mainstream paradigm [1]–[3]. However, such skill transfer is fundamentally hindered by the profound embodiment gap between human and robotic hands, encompassing significant discrepancies in kinematics, degrees of freedom, and physical properties.

To bridge the gap without collecting robot-specific demonstrations, recent works [4]–[6] increasingly formulate skill transfer as tracking hand-object interaction (HOI) trajectories derived from human demonstrations. Within this paradigm, most methods rely on keypoints to describe states and goals of the hands [7]–[10]. While keypoints can capture global spatial variations by treating joints independently in absolute space, they struggle to represent the fine-grained, structural finger synergies that are critical for dexterous manipulation.

Furthermore, these approaches typically resort to explicit kinematic retargeting to accelerate policy convergence, due to the large exploration space induced by high-dimensional hand–object interactions. A common practice [10]–[13] is to map human hand poses into robot-specific joint trajectories, utilizing them as reference signals for policy optimization. While recent advancements, such as ManipTrans [9], no longer treat these retargeted poses as training signals, they still depend on them for state initialization to facilitate early exploration. Retargeting is prone to introduce mapping errors, such as severe self-collisions, object penetrations, and missed contacts. Forcing the robot to begin episodes in these erroneous poses often places it in physically infeasible configurations, which misguides policy exploration and degrades sample efficiency.

To address these limitations, we propose ReDex, a novel framework that shifts the focus from absolute spatial tracking to intrinsic interaction geometry. Specifically, we extract Relational Manipulation Vectors (RMVs) from human demonstrations. These vectors characterize the hand poses through coordinated spatial relations that are naturally invariant across different embodiments. Since movements of individual fingers induce coupled variations across multiple relations, directly using the raw vectors leads to a highly entangled observation space. We therefore learn a compact latent representation of RMVs, termed RMVR, using a tailored Variational Autoencoder (RMV-VAE). This low-dimensional latent space captures the underlying variation patterns of the coupled RMVs, effectively mitigating the state complexity for downstream policy learning.

We further employ a guided curriculum learning strategy with an RMVR-guided state pool to enable sample-efficient exploration from physically valid states, without explicit retargeting. A preliminary kinematic anchoring phase first decouples hand coordination from object dynamics to acquire stable synergistic motion patterns. For the subsequent full-task exploration, the state pool actively filters physically valid states to serve as safe training episode initializations, confining exploration within the human-like motion manifold.

In summary, our contributions are as follows:

- We propose ReDex, which formulates RMVR as a unified geometric prior to effectively capture and transfer

[1]College of Control Science and Engineering, Zhejiang University, Hangzhou, 310027, China.
[2]School of Informatics, Xiamen University, Xiamen, Fujian, 361102, China.
[†]Qi Ye (Corresponding author, qi.ye@zju.edu.cn) is with the College of Control Science and Engineering, Zhejiang University and Zhejiang Key Laboratory of Airspace Perception and Autonomous Unmanned.
This work was supported in part by Brain Science and Brain-like Intelligence Technology-National Science and Technology Major Project (2025ZD0218800), NSFC under Grants (No.62233013, 62293511), the State Key Laboratory of Industrial Control Technology, China (Grant No. ICT2026C04), and Science and Technology Bureau of Ningbo City under Grant No. 2025Z001.

Copyright ©2026 IEEE

fine-grained, structured synergistic relationships from human demonstrations to robotic hands.
- We propose a curriculum RL strategy with an RMVR-guided state filter, leveraging the learned human prior to evaluate hand pose quality for safe, sample-efficient exploration.
- Extensive experiments demonstrate that our method outperforms existing state-of-the-art approaches, and validate its feasibility across four dexterous hands and in real-world deployments.

II. Related Works

A. Dexterous Manipulation with Demonstration Guidance

Learning high-dimensional dexterous manipulation from scratch is highly challenging. To endow robots with human-like manipulation capabilities, demonstration guidance has emerged as a prevailing paradigm [14]–[16]. Early imitation learning methods primarily relied on teleoperation systems to collect robot-specific kinematic data [17]–[19]. However, acquiring such data requires customized hardware and is difficult to scale. To improve scalability, recent work has started to use readily available human demonstrations, including motion capture (MoCap) [20], [21] and monocular videos [11], [22]. Since human motions cannot be directly executed on robotic hardware, existing methods typically employ reinforcement learning (RL), using human demonstrations as reference trajectories [4]–[6], [8]. Nevertheless, the pronounced embodiment gap between human and robotic hands still remains the challenge. Although current approaches [9], [13] utilize hierarchical tracking and residual learning to incorporate demonstrations into RL reward mechanisms, they do not explicitly investigate which geometric features from human demonstrations are intrinsically stable and efficiently transferable to the targeted robotic kinematic structure.

B. Scale-Invariant Representations for Human-Robot Transfer

Many human-to-robot transfer pipelines rely on motion retargeting or correspondence estimation [23], [24], where a fundamental bottleneck is the choice of state representation used for alignment and learning. Early methods often formulate the transfer as an optimization problem, computing target joint angles based on predefined absolute 3D keypoints [25], [26]. However, tracking these isolated, scale-sensitive points introduces mapping errors and ignores the holistic hand synergies crucial for interaction. As an alternative, methods like DexPilot [19] and DexRepNet [27] utilize relative directional vectors or spatial distances, sharing a similar geometric motivation with our approach. While these vectors provide scale invariance, directly utilizing such highly coupled representations in downstream policies exacerbates exploration difficulties and inevitably degrades sample efficiency. To address this feature coupling, our RMV-VAE encodes relative spatial vectors into a compact, shared latent space.

C. Curriculum Learning

To prevent direct RL optimization from falling into local optima due to continuous action spaces and complex contact dynamics in contact-rich manipulation, curriculum learning is widely used to simplify early-stage exploration [28]–[30]. Prior works generally scale task difficulty along various dimensions, such as progressively expanding object geometric diversity [31], [32] or transitioning from simplified physical conditions to actual dynamics. For example, ManipTrans [9] initially relaxes physical constraints like gravity and friction. UniBYD [10] and Dexmachina [8] introduce an object dynamics curriculum using virtual proportional-derivative (PD) controllers. However, virtual PD control is easily overpowered by erratic collisions during early exploration, knocking the object off its trajectory and breaking hand-object coordination, which ultimately misguides policy learning. To address this, we propose a two-stage curriculum framework that leverages a state initialization pool filtered by RMVR. This provides physically valid starting configurations, preventing early failures and accelerating learning without altering the underlying dynamics.

III. Method

As illustrated in Fig. 1, we propose ReDex, a three-stage framework that enables robots to reproduce object trajectories from human demonstrations while preserving human-like hand synergies. First, we extract a consistent, universal geometric representation from human demonstrations based on RMVs (Sec. III-A). Second, to capture the inherent coordination of human fingers, we project these RMVs into a unified latent space via an RMV-VAE module, yielding the RMVR (Sec. III-B). Finally, building upon this shared representation, we employ guided curriculum reinforcement learning to achieve efficient and physically plausible skill transfer to the robotic hand (Sec. III-C).

A. Relational Manipulation Vectors (RMVs)

Prevailing retargeting paradigms, which predominantly rely on direct keypoint-wise correspondence across heterogeneous morphologies [12], [26], frequently suffer from kinematic and physical inconsistencies. These issues stem from the inherent discrepancies in degrees of freedom (DoFs) and linkage proportions between human and robotic hands. Drawing inspiration from vector-based formulations of hand geometry [19], [27], we adopt a relation-centric representation, Relational Manipulation Vectors (RMVs), to characterize intra-hand structural relationships. This formulation projects human demonstrations into a morphology-agnostic manipulation space defined by relative functional vectors, rather than anchoring them to absolute joint or keypoint coordinates.

Specifically, as illustrated in the upper right of Fig. 1, we select a set of 17 essential manipulation-space vectors, denoted as $\mathcal{V} = \{\mathbf{v}_i\}_{i=1}^{17}$, derived from large-scale human interaction datasets [33]–[35]. These vectors are organized into four functional groups G_g ($g \in \{1, \ldots, 4\}$) (Tab. I), to capture both the global wrist-relative spatial configuration

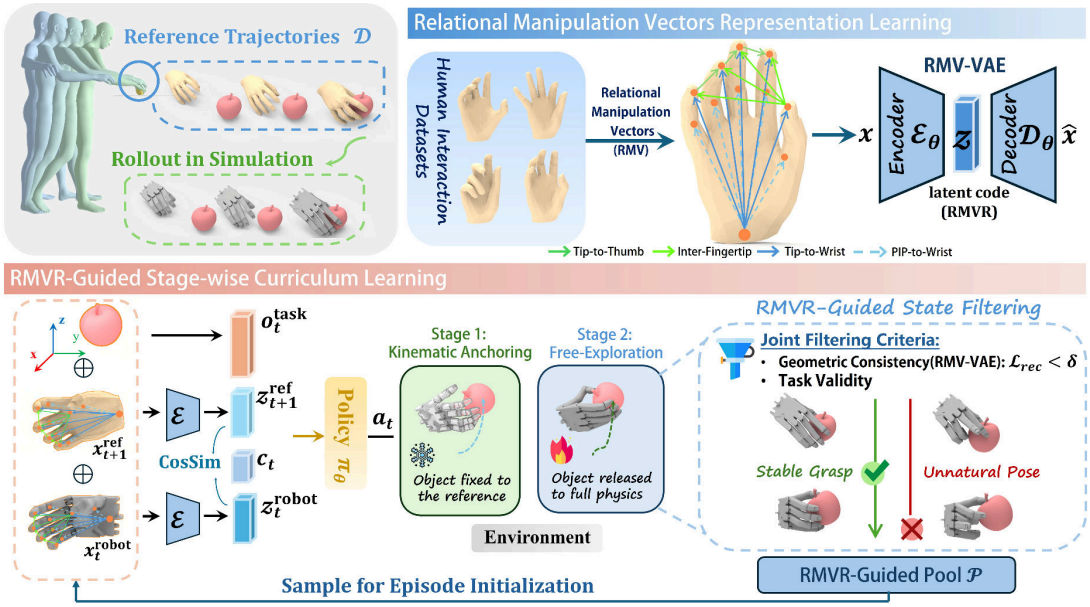


Fig. 1: Overview of our ReDex framework. ReDex extracts RMVs from human demonstrations to model human-like synergistic priors (RMVR) via the RMV-VAE. These priors guide a stage-wise curriculum to train a dexterous policy, utilizing an RMVR-guided state pool to initialize training episodes strictly with geometrically valid and physically stable poses.

TABLE I: Categories of RMVs and their synergistic roles. Tip refers to the fingertip, and PIP refers to the proximal interphalangeal joint.

Category	Num.	Synergistic Purpose
G_1: Tip-to-Wrist	5	Global hand scale and structural closure
G_2: PIP-to-Wrist	5	Proximal bending and palm arch state
G_3: Tip-to-Thumb	4	Precision grasping and opposition relations
G_4: Inter-Fingertip	3	Fine-grained spacing and collision avoidance

and fine-grained inter-digital synergies. This hierarchical formulation preserves essential interaction intents while remaining robust to specific finger lengths and joint limits.

To ensure representation consistency, all vectors are first transformed into a unified wrist-centric local coordinate frame. Each vector $\mathbf{v}_i \in \mathcal{V}$ is decoupled into a directional unit vector $\mathbf{d}_i \in \mathbb{S}^2$ and a normalized scalar length $l_i \in \mathbb{R}$. These components are concatenated to formulate the geometric feature tensor $\mathbf{x} = \{[\mathbf{d}_i^\top, l_i]\}_{i=1}^{17} \in \mathbb{R}^{17\times 4}$.

By encoding directional geometric relations within this manipulation space, RMVs establish a consistent representation that remains invariant across diverse robotic end-effectors, regardless of their specific kinematic structures.

B. Variational Latent Synergy Modeling

In the defined manipulation space $\mathbb{R}^{17\times 4}$, small fingertip displacements simultaneously affect multiple relational vectors. Although this coupling faithfully captures the coordinated structure of the hand, it also produces a highly correlated observation space. Directly feeding such coupled features into reinforcement learning makes it difficult for the policy to extract stable interaction patterns. To resolve this, we design the RMV-VAE module, a latent prior extractor that employs a VAE [36] to project the high-dimensional geometric features $\mathbf{x} \in \mathbb{R}^{17\times 4}$ onto a smooth latent manifold $\mathbf{z}$, where $\dim(\mathbf{z}) = 16$. By distilling intrinsic coordination patterns from extensive human interaction data, this module effectively decouples the observation space and provides a structured manipulation prior. This latent representation captures the underlying synergy structure of human manipulation, providing a compact and structured feature space that facilitates more stable and efficient policy learning in the subsequent reinforcement learning stage.

To ensure the latent space precisely preserves interaction-relevant geometric cues, we adopt a decoupled reconstruction objective for the RMV-VAE. Rather than optimizing for raw coordinate error, the total reconstruction loss $\mathcal{L}_{rec}$ is formulated to separately supervise the orientation and magnitude of each vector. Specifically, for each vector $\mathbf{v}_i \in \mathcal{V}$, we compute a cosine-based directional loss $\mathcal{L}_{dir,i}$ and a length loss $\mathcal{L}_{len,i}$:

$$\mathcal{L}_{\text{dir},i} = 1 - \frac{\mathbf{d}_i \cdot \hat{\mathbf{d}}_i}{\|\mathbf{d}_i\|\|\hat{\mathbf{d}}_i\|}, \quad \mathcal{L}_{\text{len},i} = \text{SmoothL1}(l_i, \hat{l}_i), \quad (1)$$

where $\hat{(\cdot)}$ denotes the reconstructed output.

Since long-range wrist-relative vectors typically exhibit larger magnitudes than fine-grained inter-finger relations, we employ homoscedastic uncertainty modeling [37] to automatically balance the reconstruction losses without manual weight tuning:

$$\mathcal{L}_{\text{rec}} = \sum_{g=1}^{4} \sum_{i\in\mathcal{I}_g} \left(\frac{w_{\text{d}}\mathcal{L}_{\text{dir},i}}{\exp(s_{g,\text{d}})} + s_{g,\text{d}} + \frac{w_{\text{l}}\mathcal{L}_{\text{len},i}}{\exp(s_{g,\text{l}})} + s_{g,\text{l}} \right), \quad (2)$$

where $\mathcal{I}_g$ denotes the index set for functional group G_g, $w_{\mathrm{d}}, w_{\mathrm{l}}$ are fixed base weights, and $s_{g,\mathrm{d}}, s_{g,\mathrm{l}}$ are learnable log-variance parameters.

In addition, a KL divergence regularization term is introduced, yielding the overall objective $\mathcal{L}_{\mathrm{VAE}} = \mathcal{L}_{\mathrm{rec}} + \beta \mathcal{L}_{\mathrm{KL}}$. Once pre-trained, both the RMV-VAE encoder $\mathcal{E}_\theta(\cdot)$ and decoder $\mathcal{D}_\theta(\cdot)$ are frozen for downstream tasks.

C. Guided Curriculum Learning for Dexterous Manipulation

1) ***Problem Formulation****:* The problem is defined as follows: given a reference human-object interaction trajectory $\mathcal{D} = \{(\mathbf{H}_t, \mathbf{O}_t)\}_{t=1}^{T}$, where $\mathbf{H}_t$ denotes the human hand pose and $\mathbf{O}_t$ denotes the object state (pose and velocity) at time step t, the goal is to learn a control policy π_θ that tracks the target object trajectories while preserving hand coordination patterns consistent with the human demonstration.

We model the task as a Markov Decision Process (MDP) $\mathcal{M} = \langle \mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \gamma \rangle$, where $\mathcal{S}$ represents the state space, $\mathcal{A}$ the action space, $\mathcal{T}$ the transition dynamics, $\mathcal{R}$ the reward function, and $\gamma \in (0, 1]$ the discount factor. We employ Proximal Policy Optimization (PPO) [38] to facilitate policy learning. The objective is to maximize the expected discounted return $J(\pi_\theta) = \mathbb{E}_{\pi_\theta}\left[\sum_{t=0}^{T} \gamma^t r_t\right]$.

2) ***State Space with Pre-trained Representations****:* We formulate the state space $\mathcal{S}$ by augmenting the physical task state with latent geometric priors. Utilizing the frozen encoder $\mathcal{E}_\theta(\cdot)$ from Sec. III-B, the state at time t is defined as a tuple:

$$\mathbf{s}_t = \langle \mathbf{o}_t^{\mathrm{task}}, \underbrace{\mathcal{E}_\theta(\mathbf{x}_t^{\mathrm{robot}})}_{\mathbf{z}_t^{\mathrm{robot}}}, \underbrace{\mathcal{E}_\theta(\mathbf{x}_{t+1}^{\mathrm{ref}})}_{\mathbf{z}_{t+1}^{\mathrm{ref}}}, c_t \rangle, \tag{3}$$

where $\mathbf{o}_t^{\mathrm{task}} \in \mathbb{R}^d$ comprises object states, BPS features [39] and fingertip-to-object distances. Here, $\mathbf{x}_t^{\mathrm{robot}}$ and $\mathbf{x}_{t+1}^{\mathrm{ref}}$ denote the RMVs extracted from the robot state and the reference human demonstration. Finally, $c_t = \mathrm{CosSim}(\mathbf{z}_t^{\mathrm{robot}}, \mathbf{z}_{t+1}^{\mathrm{ref}})$ acts as an explicit alignment heuristic in the latent space, where CosSim denotes the cosine similarity, guiding the policy toward human-consistent synergy patterns.

3) ***Action Space****:* The action space $\mathcal{A}$ consists of target position commands for the K actuated finger joints, denoted as $\mathbf{a}_t \in \mathbb{R}^K$. The wrist pose is kinematically driven by a reference trajectory from the human demonstration. This setup enables the policy to prioritize the learning of dexterous finger manipulation.

4) ***Reward Function****:* Following [9], we design a multi-objective reward to guide policy learning, formulated as a weighted sum of three components:

$$r_t = w_{\mathrm{task}} \cdot r_t^{\mathrm{task}} + w_{\mathrm{imi}} \cdot r_t^{\mathrm{imi}} + w_{\mathrm{smooth}} \cdot r_t^{\mathrm{smooth}}. \tag{4}$$

The task tracking reward r_t^{task} minimizes the 6-DoF pose and velocity tracking errors of the manipulated object. The imitation reward r_t^{imi} applies sparse kinematic supervision by tracking only the fingertips and proximal interphalangeal (PIP) joints via exponential decay kernels. To ensure physically safe and smooth motions, r_t^{smooth} applies a direct penalty on high joint velocities and excessive joint torques.

5) ***Stage-wise Training Strategy****:* Simultaneously mastering faithful motion imitation and dynamic object manipulation presents a formidable challenge. The high-dimensional action space and complex contact dynamics often cause the policy to be trapped in suboptimal local minima, in which the hand visually mimics the motion but fails to establish stable physical interactions. To mitigate this, we propose a two-stage curriculum learning mechanism that progressively scales task complexity to guide policy convergence.

Stage I: Kinematic Anchoring Phase. During this initial stage, the object's pose is directly set to the reference trajectory, remaining unaffected by gravity or contact impulses. Specifically, we set the $w_{\mathrm{task}} = 0$, forcing the robot to focus exclusively on maximizing the r_t^{imi}. This phase allows the robot to internalize human-like synergistic priors and achieve stable interactions without failing prematurely.

Stage II: Free-Exploration Phase with RMVR-Guided State Pool. Stage II begins when the imitation success rate stably exceeds a threshold or a step limit is reached. Here, the object is released into full physical dynamics, and r_t^{task} is activated to encourage active manipulation. To ensure efficient exploration across the entire trajectory, we introduce the RMVR-guided state pool.

Resetting the environment to the initial frame makes learning later trajectory phases inefficient, while Reference State Initialization (RSI) with imperfect retargeted states often leads to premature episode termination due to contact artifacts [40]. Thus, we adopt Physical State Initialization (PSI), sampling starting states from a buffer of past successful rollouts. However, standard PSI uses reward-based filtering. In dexterous manipulation, policies exploit this by adopting unnatural, non-anthropomorphic poses to maximize task rewards.

To enforce human-like coordination, we introduce the pool equipped with a geometric consistency filter. A state $\mathbf{s}_t$ is admitted into the pool $\mathcal{P}$ only if it satisfies two criteria:

- Task Validity: The manipulated object remains stable without dropping.
- Geometric Consistency: The reconstruction loss satisfies $\mathcal{L}_{\mathrm{rec}}(\mathbf{x}_t^{\mathrm{robot}}, \hat{\mathbf{x}}_t^{\mathrm{robot}}) < \delta$. Since the RMV-VAE is trained solely on human demonstrations, a low reconstruction loss indicates that the current configuration lies close to the learned human synergy manifold. This prevents the policy from drifting toward unnatural hand poses.

During Stage II, each training episode is initialized by sampling a state from the pool $\mathcal{P}$ with probability 0.8, or by resetting to the initial frame $t = 0$ otherwise.

IV. Experiments

The experiments begin with the dataset setup, evaluation metrics, and implementation details (Sec. IV-A). We then benchmark ReDex against state-of-the-art (SOTA) methods (Sec. IV-B), conduct ablation studies to analyze its core components (Sec. IV-C, IV-D), exhibit cross-embodiment generalization (Sec. IV-E), and validate real-world deployment (Sec. IV-F).

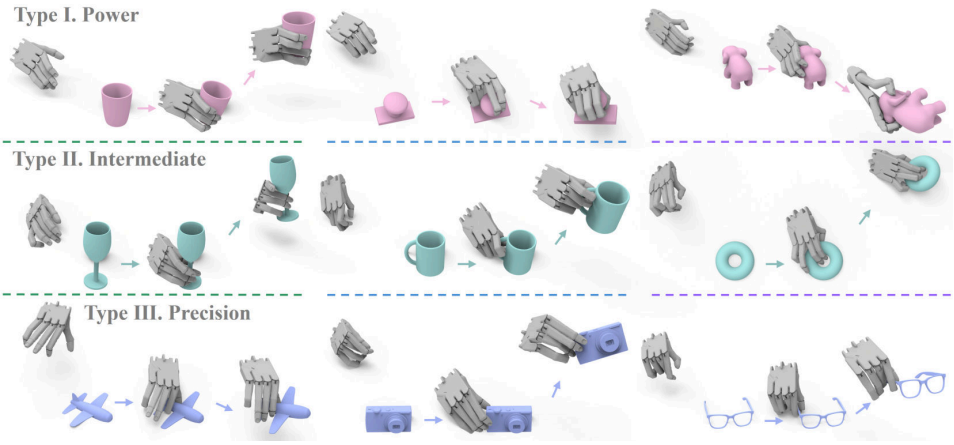


Fig. 2: Qualitative results of ReDex. From top to bottom, the rows illustrate grasping sequences for three distinct manipulation modes: Type I. Power, Type II. Intermediate, and Type III. Precision.

TABLE II: Quantitative comparison against baselines.

Method	Type I: Power			Type II: Intermediate			Type III: Precision		
	SR (%) ↑	E_{pos} **(cm)** ↓	E_{rot} (°) ↓	**SR (%)** ↑	E_{pos} **(cm)** ↓	E_{rot} (°) ↓	**SR (%)** ↑	E_{pos} **(cm)** ↓	E_{rot} (°) ↓
AnyTeleop	0.2	2.99	29.96	8.6	2.74	27.93	12.0	2.64	27.31
ManipTrans	46.1 ± 3.7	1.46	14.55	63.8 ± 8.9	0.29	8.94	44.0 ± 0.3	1.99	19.93
Ours	**81.8 ± 4.2**	**0.37**	**9.79**	**88.6 ± 3.5**	**0.13**	**6.54**	**78.8 ± 5.5**	**0.37**	**9.20**

A. Experiment Setup

*1) **Datasets Setup**:* We select reference human-object interaction trajectories from the GRAB dataset [33], encompassing 27 unique objects with diverse geometries. These tasks are categorized into three manipulation modes based on contact characteristics [41]: *Power* (high contact area involving the palm), *Intermediate* (hybrid contact with the thumb and 3–4 fingers), and *Precision* (fingertip-only sparse contact).

*2) **Metrics**:* We train each policy across three independent random seeds and evaluate manipulation performance over 100 trials per trajectory. The primary metric is the **Success Rate (SR)**, defined as the percentage of trials in which the object remains within the predefined tolerance thresholds ($E_{\text{pos}} < 3\,\text{cm}$ and $E_{\text{rot}} < 30°$) at all timesteps. To characterize the attainable precision of the learned policies, we report the **Position Error** E_{pos} and **Orientation Error** E_{rot} from the best-performing seed. These metrics compute the deviation of the object's achieved pose from the reference pose, where E_{pos} measures the mean Euclidean distance (cm) and E_{rot} calculates the average minimum rotation angle (degrees). Failed trials are assigned the maximum threshold values (3 cm and 30°) for error computation.

*3) **Implementation Details**:* All simulations are conducted in Isaac Gym [42] with 4,096 parallel environments, running on a workstation equipped with an Intel Xeon Gold 6326 CPU and an NVIDIA RTX 3090 GPU. Our RMV-VAE (a 3-layer MLP) is pre-trained using AdamW with a learning rate of 1e-3 for 800 epochs, taking 3 hours. Quantitative benchmarks use the 22-DoF Shadow Hand [43], while cross-embodiment validations employ the 12-DoF XHand [44], 20-DoF Wuji Hand [45], and 6-DoF Inspire Hand [46].

B. Main Results: Comparison with SOTA

We benchmark our approach against two representative methods: AnyTeleop [25], which relies on open-loop explicit kinematic retargeting, and ManipTrans [9], the SOTA framework employing residual reinforcement learning. As demonstrated in Tab. II, our method achieves superior performance across all three manipulation modes. The explicit retargeting baseline, AnyTeleop, performs poorly in dynamic tasks (SR $\leq 12.0\%$), as solely minimizing kinematic distances fails to account for the embodiment gap and the intricate contact dynamics essential for stability. While ManipTrans improves performance via residual learning, it struggles in sparse-contact and high-precision scenarios. This limitation can be attributed to two factors: first, tracking absolute keypoints overlooks the relative inter-digital coordination vital for precision; second, direct kinematic mapping for Reference State Initialization (RSI) introduces persistent physical conflicts, leading to premature policy divergence.

In contrast, by anchoring demonstrations in a shared RMV latent space, ReDex effectively mitigates physical inconsistencies and achieves robust success rates across all categories: 81.8% in contact-rich *Power*, 88.6% in *Intermediate*, and 78.8% in highly demanding *Precision* tasks. Notably, in the most challenging *Precision* mode, ReDex achieves a +34.8% absolute margin over ManipTrans. Furthermore, ReDex consistently yields the lowest tracking errors ($E_{\text{pos}} \leq 0.37\,\text{cm}$, $E_{\text{rot}} \leq 9.79°$), highlighting its effectiveness in achieving high-fidelity and physically plausible manipulation

transfer.

C. The Effect of RMVs and RMVR

To isolate the structural advantages of our framework, we investigate how the consistent geometric features (RMVs) and their latent encoding (RMVR) contribute to the stability and precision of learned skills. We ablate the RL state inputs into four representative variants (Fig. 3) while maintaining an identical backbone architecture:

- **Raw Keypoint** directly utilizes absolute 3D keypoint coordinates (relative to the wrist) and their temporal deltas, mirroring standard tracking baselines [9];
- **Latent Keypoint** compresses the absolute keypoints via a similarly structured VAE $\mathcal{E}_{\text{kp}}(\cdot)$;
- **RMVs** feed the uncompressed 17 relational vectors $\mathbf{x} = \{(\mathbf{d}_i, l_i)\}_{i=1}^{17}$ directly into the policy network;
- **RMVR (Ours)** projects the RMVs into the smooth latent manifold $\mathbf{z}$, representing our full proposed pipeline.

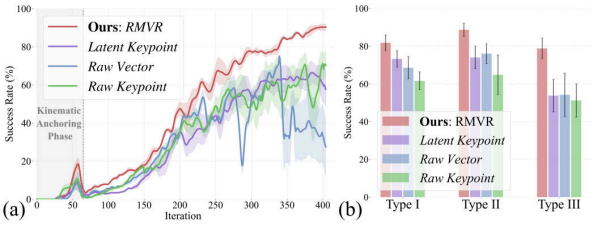


Fig. 3: Ablation study on geometric representations. (a): Training process on the wineglass toast task; (b): Evaluation results across three manipulation modes: Type I (Power), Type II (Intermediate), and Type III (Precision).

*1) **Relational Vector vs. Absolute Keypoint**:* A limitation of absolute keypoint representations is that tracking errors accumulate along the kinematic chain, so that small deviations at proximal joints can induce amplified spatial variance at distal fingertips.

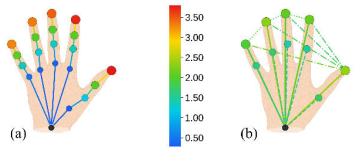


Fig. 4: Spatial Error Distribution (mm). (a) Absolute keypoint errors encoded by sphere scale and color, with edges showing adjacent averages. (b) Relational vector errors along structural connections, categorized by line style. The average error for Tip and PIP is 2.16mm in (a) and 1.88mm in (b).

As illustrated in Fig. 4a, although homoscedastic uncertainty modeling is introduced to balance loss magnitudes across joints, projecting unconstrained absolute keypoints into a VAE creates an optimization mismatch. Under this formulation, the model tends to assign larger variance parameters to distal joints, effectively down-weighting the fingertip reconstruction errors in the overall objective. The keypoint-VAE essentially trades off fine-grained precision to accommodate the massive spatial fluctuations inherent in absolute coordinates.

By comparison, our relational vectors (RMVs) represent local synergies rather than absolute kinematics, effectively mitigating error propagation along the kinematic chain (Fig. 4b). By keeping spatial variance bounded within comparable scales, the uncertainty-weighted objective can better coordinate supervision across functional groups. This structural prior leads to a 25.0% absolute increase in success rate over keypoint encoding in contact-sparse *Precision* tasks (Fig. 3b).

*2) **Latent Encoding vs. Raw Input**:* Compared to high-dimensional raw inputs, the smooth RMV-VAE latent manifold $\mathbf{z}$ significantly accelerates convergence and improves peak RL performance.

As illustrated in Fig. 3a, directly feeding the policy with structurally coupled raw vectors leads to higher training variance and more fluctuating convergence. By contrast, the RMVR provides a regularized and decoupled search space that resolves inherent kinematic entanglement, allowing the agent to focus on task-relevant geometric structure and achieve more stable policy optimization.

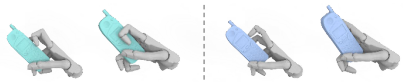


Fig. 6: Comparison of Hand Postures. Left: +KA+SP w/o vae loss. Right: Ours.

D. The Effect of Curriculum Learning

As shown in Tab. III, we evaluate the necessity of each curriculum component through a series of ablation studies, analyzing how stage-wise task progression and manifold-based filtering improve exploration in dexterous manipulation.

*1) **The Role of Kinematic Anchoring (+KA)**:* Learning the target policy directly from scratch (**Base**) frequently results in exploration collapse, with a success rate of only 16.4% (*Power*). The simultaneous learning of contact discovery and full task dynamics hinders effective exploration. By introducing the Kinematic Anchoring phase (**+KA**), we decouple hand coordination from object dynamics. This allows the policy to first acquire the coordinated motion patterns in a simplified environment, providing a robust initialization for subsequent dynamic tasks and increasing the SR to 73.0%.

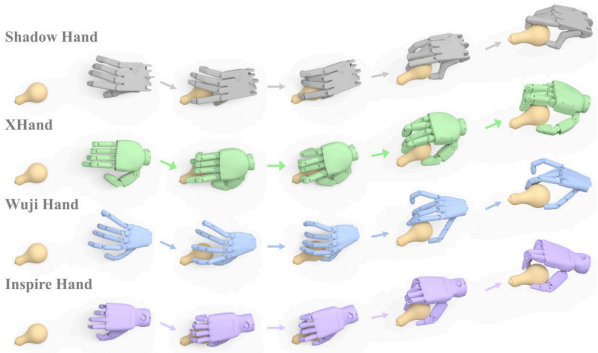


Fig. 7: Cross Embodiment Results: Passing lightbulb.

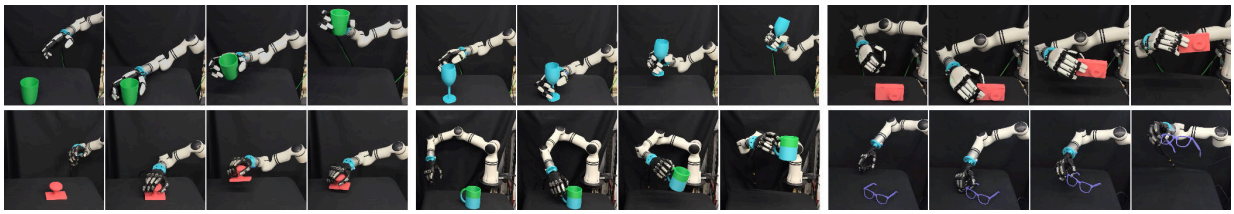


Fig. 5: Real-world experiments. Hardware deployments using the XHand (top) and the Wuji Hand (bottom). The columns, from left to right, correspond to power, intermediate, and precision tasks.

TABLE III: Ablation study on the stage-wise curriculum learning strategy.

Mode	Training Strategy	SR (%) ↑	E_{pos} (cm) ↓	E_{rot} (°) ↓
Type I. Power	Base	16.4 ± 6.2	2.02	22.27
	+KA Phase	73.0 ± 5.2	0.38	9.72
	+KA+SP (w/o VAE loss)	70.1 ± 5.9	0.61	9.85
	Ours (Full)	**81.8 ± 4.2**	**0.37**	**9.79**
Type II. Intermediate	Base	40.9 ± 9.4	1.22	15.90
	+KA Phase	69.1 ± 5.1	0.61	10.29
	+KA+SP (w/o VAE loss)	77.6 ± 5.1	0.35	7.32
	Ours (Full)	**88.6 ± 3.5**	**0.13**	**6.54**
Type III. Precision	Base	17.5 ± 3.8	1.97	21.34
	+KA Phase	77.3 ± 6.4	0.41	10.10
	+KA+SP (w/o VAE loss)	77.9 ± 5.2	**0.35**	9.44
	Ours (Full)	**78.8 ± 5.5**	0.37	**9.20**

*2) **The Role of RMVR-Guided State Pool (+KA+SP w/o VAE loss)**:* Although a physical state initialization (PSI) buffer facilitates exploration, an unconstrained state pool (**+KA+SP w/o VAE loss**) may encourage unintended reward-driven behaviors. For example, in phone manipulation (Fig. 6), the policy converges to physically valid but anatomically implausible grasps. While satisfying basic task rewards, these unnatural states yield unstable and non-generalizable solutions.

Using the RMV-VAE reconstruction loss as a geometric consistency filter restricts the state pool to human-like coordination manifolds. Initializing exploration from these kinematically plausible states increases the SR to 81.8% (*Power*), highlighting the importance of this structural prior for both policy robustness and generalization.

E. Cross-Embodiment Validation

To demonstrate the extensibility of ReDex, we train policies for four dexterous hand embodiments (Fig. 7) using a unified learning pipeline. Encoding relational geometry (RMVs) rather than absolute keypoints or joint angles, our representation facilitates cross-morphology transfer. Consequently, ReDex achieves consistent, human-like manipulation performance across all hands, without explicit kinematic retargeting or hand-specific reward tuning.

F. Real World Experiment

We deploy ReDex on a physical 7-DoF RealMan arm equipped with either the XHand or the Wuji Hand to validate real-world transferability. The arm is controlled through inverse kinematics, while the hand joint angles are directly sent to the hardware controllers. As shown in Fig. 5, the system achieves fluid, precise, and coordinated manipulation behaviors in real-world settings. Please refer to our supplementary video for detailed demonstrations of the system's real-world performance.

V. Conclusion

In this paper, we present ReDex for transferring dexterous manipulation skills from human demonstrations to robotic hands. The key contributions of our work are twofold: 1) formulating a compact latent representation of RMVR as a unified geometric prior to effectively capture and transfer fine-grained, structured synergistic relationships; 2) proposing a curriculum reinforcement learning strategy featuring an RMVR-guided state pool to enable efficient and physically feasible skill transfer by geometrically filtering out unnatural poses. Comprehensive experiments demonstrate that ReDex significantly outperforms existing SOTA baselines, notably in tracking precision, while also exhibiting favorable cross-embodiment transferability and feasibility for real-world deployment.

Discussion and Limitations While effective, our framework currently relies on high-quality hand-object demonstrations to provide reference trajectories for its downstream skill transfer. Additionally, the scope of this work is restricted to single-hand manipulation. In future work, we will extend the RMVR representation and our curriculum strategy to support bimanual coordination, aiming to tackle more complex collaborative manipulation tasks.

References

[1] O. M. Andrychowicz, B. Baker, M. Chociej, R. Jozefowicz, B. McGrew, J. Pachocki, A. Petron, M. Plappert, G. Powell, A. Ray, *et al.*, "Learning dexterous in-hand manipulation," *The International Journal of Robotics Research*, vol. 39, no. 1, pp. 3–20, 2020.

[2] Y.-H. Wu, J. Wang, and X. Wang, "Learning generalizable dexterous manipulation from human grasp affordance," in *Conference on robot learning*. PMLR, 2023, pp. 618–629.

[3] H. Luo, Y. Feng, W. Zhang, S. Zheng, Y. Wang, H. Yuan, J. Liu, C. Xu, Q. Jin, and Z. Lu, "Being-h0: vision-language-action pretraining from large-scale human videos," *arXiv preprint arXiv:2507.15597*, 2025.

[4] S. Dasari, A. Gupta, and V. Kumar, "Learning dexterous manipulation from exemplar object trajectories and pre-grasps," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 3889–3896.

[5] B. Zhou, H. Yuan, Y. Fu, and Z. Lu, "Learning diverse bimanual dexterous manipulation skills from human demonstrations," *arXiv preprint arXiv:2410.02477*, 2024.

[6] S. Xu, Y.-W. Chao, L. Bian, A. Mousavian, Y.-X. Wang, L. Gui, and W. Yang, "Dexplore: Scalable neural control for dexterous manipulation from reference scoped exploration," in *Conference on Robot Learning*. PMLR, 2025, pp. 2184–2199.

[7] Y. Chen, C. Wang, Y. Yang, and C. K. Liu, "Object-centric dexterous manipulation from human motion data," *arXiv preprint arXiv:2411.04005*, 2024.

[8] Z. Mandi, Y. Hou, D. Fox, Y. Narang, A. Mandlekar, and S. Song, "Dexmachina: Functional retargeting for bimanual dexterous manipulation," *arXiv preprint arXiv:2505.24853*, 2025.

[9] K. Li, P. Li, T. Liu, Y. Li, and S. Huang, "Maniptrans: Efficient dexterous bimanual manipulation transfer via residual learning," in *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2025.

[10] T. Yuan, B. Guan, W. Ye, Z. Tian, Y. Yang, W. Zhou, Y. Huang, P. Wang, C. Zhao, and J. Wang, "Unibyd: A unified framework for learning robotic manipulation across embodiments beyond imitation of human demonstrations," *arXiv preprint arXiv:2512.11609*, 2025.

[11] Y. Qin, Y.-H. Wu, S. Liu, H. Jiang, R. Yang, Y. Fu, and X. Wang, "Dexmv: Imitation learning for dexterous manipulation from human videos," in *European Conference on Computer Vision*. Springer, 2022, pp. 570–587.

[12] C. Wang, H. Shi, W. Wang, R. Zhang, L. Fei-Fei, and K. Liu, "Dexcap: Scalable and portable mocap data collection system for dexterous manipulation," in *RSS 2024 Workshop: Data Generation for Robotics*.

[13] X. Liu, J. Adalibieke, Q. Han, Y. Qin, and L. Yi, "Dextrack: Towards generalizable neural tracking control for dexterous manipulation from human references," in *The Thirteenth International Conference on Learning Representations*.

[14] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Robotics and autonomous systems*, vol. 57, no. 5, pp. 469–483, 2009.

[15] P. Englert and M. Toussaint, "Learning manipulation skills from a single demonstration," *The International Journal of Robotics Research*, vol. 37, no. 1, pp. 137–154, 2018.

[16] J. Ye, J. Wang, B. Huang, Y. Qin, and X. Wang, "Learning continuous grasping function with a dexterous hand from human demonstrations," *IEEE Robotics and Automation Letters*, vol. 8, no. 5, pp. 2882–2889, 2023.

[17] A. Rajeswaran, V. Kumar, A. Gupta, G. Vezzani, J. Schulman, E. Todorov, and S. Levine, "Learning complex dexterous manipulation with deep reinforcement learning and demonstrations," *arXiv preprint arXiv:1709.10087*, 2017.

[18] S. Christen, S. Stevšić, and O. Hilliges, "Guided deep reinforcement learning of control policies for dexterous human-robot interaction," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 2161–2167.

[19] A. Handa, K. Van Wyk, W. Yang, J. Liang, Y.-W. Chao, Q. Wan, S. Birchfield, N. Ratliff, and D. Fox, "Dexpilot: Vision-based teleoperation of dexterous robotic hand-arm system," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 9164–9170.

[20] Z. Q. Chen, K. Van Wyk, Y.-W. Chao, W. Yang, A. Mousavian, A. Gupta, and D. Fox, "Dextransfer: Real world multi-fingered dexterous grasping with minimal human demonstrations," *arXiv preprint arXiv:2209.14284*, 2022.

[21] L. Shi, Y. Liu, L. Zeng, B. Ai, Z. Hong, and H. Su, "Learning adaptive dexterous grasping from single demonstrations," in *2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2025, pp. 9456–9463.

[22] K. Shaw, S. Bahl, and D. Pathak, "Videodex: Learning dexterity from internet videos," in *Conference on Robot Learning*. PMLR, 2023, pp. 654–665.

[23] X. Lin, K. Yao, L. Xu, X. Wang, X. Li, Y. Wang, and M. Li, "Dexflow: A unified approach for dexterous hand pose retargeting and interaction," *arXiv preprint arXiv:2505.01083*, 2025.

[24] C. Xin, M. Yu, Y. Jiang, Z. Zhang, and X. Li, "Analyzing key objectives in human-to-robot retargeting for dexterous manipulation," *IEEE Robotics and Automation Practice*, 2026.

[25] Y. Qin, W. Yang, B. Huang, K. Van Wyk, H. Su, X. Wang, Y.-W. Chao, and D. Fox, "Anyteleop: A general vision-based dexterous robot arm-hand teleoperation system," in *Robotics: Science and Systems*, 2023.

[26] H. Zhang, S. Hu, Z. Yuan, and H. Xu, "Doglove: Dexterous manipulation with a low-cost open-source haptic force feedback glove," *arXiv preprint arXiv:2502.07730*, 2025.

[27] Q. Liu, Y. Cui, Q. Ye, Z. Sun, H. Li, G. Li, L. Shao, and J. Chen, "Dexrepnet: Learning dexterous robotic grasping network with geometric and spatial hand-object representations," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2023, pp. 3153–3160.

[28] P. Ojaghi, R. Mir, A. Marjaninejad, A. Erwin, M. Wehner, and F. J. Valero-Cuevas, "Curriculum is more influential than haptic feedback when learning object manipulation," *Science Advances*, vol. 11, no. 14, p. eadp8407, 2025. [Online]. Available: https://www.science.org/doi/abs/10.1126/sciadv.adp8407

[29] A. S. Chiappa, P. Tano, N. Patel, A. Ingster, A. Pouget, and A. Mathis, "Acquiring musculoskeletal skills with curriculum-based reinforcement learning," *Neuron*, vol. 112, no. 23, pp. 3969–3983, 2024.

[30] Z. Sun, Z. Shi, J. Chen, Q. Liu, Y. Cui, J. Chen, and Q. Ye, "Vtao-bimanip: Masked visual-tactile-action pre-training with object understanding for bimanual dexterous manipulation," in *2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2025, pp. 3201–3208.

[31] W. Wan, H. Geng, Y. Liu, Z. Shan, Y. Yang, L. Yi, and H. Wang, "Unidexgrasp++: Improving dexterous grasping policy learning via geometry-aware curriculum and iterative generalist-specialist learning," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 3891–3902.

[32] K. Zhu, F. Bai, Y. Xiang, Y. Cai, X. Chen, R. Li, X. Wang, H. Dong, Y. Yang, X. Fan, *et al.*, "Dexflywheel: A scalable and self-improving data generation framework for dexterous manipulation," in *The Thirty-ninth Annual Conference on Neural Information Processing Systems*.

[33] O. Taheri, N. Ghorbani, M. J. Black, and D. Tzionas, "GRAB: A dataset of whole-body human grasping of objects," in *European Conference on Computer Vision (ECCV)*, 2020. [Online]. Available: https://grab.is.tue.mpg.de

[34] Y. Liu, H. Yang, X. Si, L. Liu, Z. Li, Y. Zhang, Y. Liu, and L. Yi, "Taco: Benchmarking generalizable bimanual tool-action-object understanding," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 21 740–21 751.

[35] X. Zhan, L. Yang, Y. Zhao, K. Mao, H. Xu, Z. Lin, K. Li, and C. Lu, "OakInk2: A dataset of bimanual hands-object manipulation in complex task completion," in *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024, pp. 445–456.

[36] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*, 2013.

[37] A. Kendall, Y. Gal, and R. Cipolla, "Multi-task learning using uncertainty to weigh losses for scene geometry and semantics," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7482–7491.

[38] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[39] S. Prokudin, C. Lassner, and J. Romero, "Efficient learning on point clouds with basis point sets," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 4331–4340.

[40] S. Xu, H. Y. Ling, Y.-X. Wang, and L. Gui, "Intermimic: Towards universal whole-body control for physics-based human-object interactions," in *CVPR*, 2025.

[41] X. Xu, J. Sun, S. Chen, L. Ma, K. Sun, B. Zhao, J. Yuan, S. Yi, H. Zhu, Y. Lu, *et al.*, "Dexcanvas: Bridging human demonstrations and robot learning for dexterous manipulation," *arXiv preprint arXiv:2510.15786*, 2025.

[42] V. Makoviychuk, L. Wawrzyniak, Y. Guo, M. Lu, K. Storey, M. Macklin, D. Hoeller, N. Rudin, A. Allshire, A. Handa, *et al.*, "Isaac gym: High performance gpu-based physics simulation for robot learning," *arXiv preprint arXiv:2108.10470*, 2021.

[43] D. Sharma, K. Tokas, A. Puri, and K. Sharda, "Shadow hand," *Journal of Advance Research in Applied Science ISSN*, vol. 2208, p. 2352, 2014.

[44] Robotera, "Xhand1 official website." [Online]. Available: https://www.robotera.com/en/goods1/4.html

[45] W. TECH, "Wuji hand official website." [Online]. Available: https://wuji.tech/

[46] Inspire-Robots, "Inspire hand official website." [Online]. Available: https://inspire-robots.store/