

2nd Place Solution to the HANDS 2026 Workshop Challenge-Dexterous Grasp Motion Track: Single-Shot Trajectory Warping for Grasp Motion Generation

Muneeb A. Khan¹ Woojin Kim¹ Shinwoo Kim¹
 Muhammad Munsif¹ Binod Bhattarai^{2,3} Seungryul Baek¹
¹UNIST, South Korea ²University of Aberdeen, UK ³Fogsphere (Redev AI Ltd, UK)

Abstract

This report describes our 2nd place solution to the HANDS 2026 workshop challenge (Dexterous Grasp Motion track) in conjunction with ECCV 2026. In this challenge, we address grasp motion generation for the 12-DoF LinkerHand O6, aiming to produce physically plausible reach-and-lift trajectories for unseen objects from randomized initial hand poses in simulation. This task is particularly challenging because each grasp requires a per-step policy to make approximately 70 twelve-dimensional decisions, with errors accumulating over time, while test objects and physical dynamics may differ from those encountered during training. To address these challenges, we propose editing a single successful GraspM3 demonstration instead of generating the motion step by step: a policy observes the object once and outputs a 12-D warp of the demonstration, which is then replayed open-loop. Moreover, we train the warp policy with one-step PPO over all 4,824 training objects in parallel. As a result, our method achieved success rates of 94.61% on the easy track, the highest of all submissions, and 57.18% on the hard track of the private test set.

1. Introduction

Dexterous grasping is a prerequisite for general-purpose manipulation. A multi-fingered hand covers a far wider range of objects and grasp types than a parallel-jaw gripper, but the added dexterity brings a high-dimensional action space and contact-rich dynamics that are hard to model [4]. Two developments have made data-driven approaches practical, including large synthetic corpora [2] that pair objects with hand poses or complete grasp motions, and massively parallel simulators [6] that amortize exploration over thousands of objects at once. Combining them is nontrivial, because a corpus prescribes what the hand should do for one object instance, whereas a deployed policy must decide what to do for an unseen object from an arbitrary starting pose. The Dexterous Grasp Motion track [3] is built around

this gap. It builds on GraspM3 [4, 9], retargeted from the Shadow Hand to the LinkerHand O6 consist of six actuated finger joints on a free-floating wrist, so that every control step is a 12-D command (wrist translation, wrist rotation, six finger targets). Training data covers 5,048 ShapeNet [1] objects with $\sim 110k$ trajectories; replaying them in Isaac Gym [6] and keeping the successful ones leaves 4,824 objects and 52,718 trajectories. Episodes start 15–20 cm from a randomly oriented object and succeed on a 30 cm lift. Submissions are scored on private Objaverse [2] objects on an *easy* track (mass 20 g, friction 20, as in training) and a *hard* track (mass $\sim \mathcal{U}[20, 200]$ g, friction $\sim \mathcal{U}[1, 20]$). The official baseline is behavior cloning on DexRep [5] features (23.93%).

Per-step policies face significant hurdles in this context, requiring approximately 70 twelve-dimensional decisions per grasp from a 1,757-D observation, with compounding error. DemoGrasp [10] showed that neither is strictly necessary for the geometric phase of grasping: one successful demonstration, edited by a policy that decides *where* the hand goes and *how far* it closes, generalizes across objects when the editor is trained as a single-step MDP. We port this formulation to the O6 task and GraspM3 data. Our contributions are (i) an object-centric warp of one demonstration with position-only alignment, verbatim lift preservation, and proportional finger rescaling (Sec. 2.2); (ii) one-step PPO over all 4,824 training objects in parallel with trajectory rotation (Sec. 2.4); and (iii) an analysis of the easy-to-hard gap that an open-loop warp cannot close (Sec. 3).

2. Method

Figure 1 shows the pipeline. An episode provides an object o (initial pose and point cloud) and a randomized initial hand pose. The policy π_θ observes the scene once and outputs $a \in [-1, 1]^{12}$; a warps a single demonstration τ into a full trajectory, which is replayed open-loop, and the episode returns $r = \mathbb{I}[\text{lift} \geq 0.3 \text{ m}]$. Learning is the contextual ban-

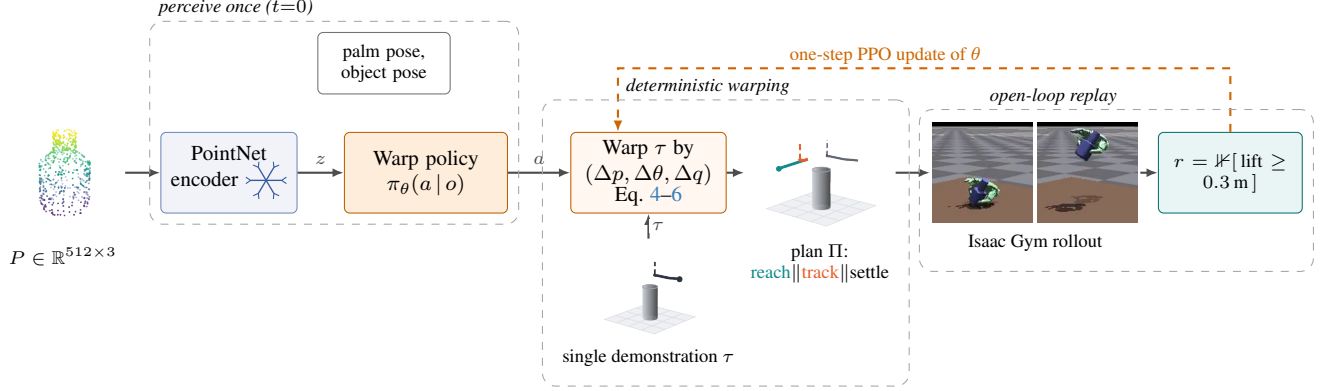


Figure 1. Overview of the trajectory warping pipeline. A frozen (*) PointNet embeds the object point cloud P ; with the palm and object poses, and the embedding z is passed to the warp policy π_θ once per episode.

dit

$$\max_{\theta} \mathbb{E}_{o \sim \mathcal{D}, a \sim \pi_\theta(\cdot|o)} [r(o, a)], \quad (1)$$

with no credit assignment across time.

2.1. Object-centric demonstration

We utilize a successful GraspM3 bottle grasp demonstration, verified via a single replay in Isaac Gym to ensure successful object lifting (Fig. 2). We record the *commanded* actions; replaying the resulting physical states would reintroduce a one-step execution lag: $\tau = \{u_t\}_{t=1}^T$, $u_t = (x_t, r_t, q_t) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^6$, $T = 69$, with lift onset $t_\ell = 34$ (first frame with object rise > 2 cm). The wrist command maps to the palm pose as $p_t = x_t + c$, $R_t = R_x(\phi_t)R_y(\theta_t)R_z(\psi_t)$ (intrinsic XYZ), with a constant offset c measured at reset. We verified this convention numerically; Isaac Gym’s extrinsic Euler helper is off by 5 cm and would corrupt every warped frame. The demonstration is expressed relative to its initial object *position* only,

$$\tilde{p}_t = p_t - o_0^{\text{demo}}, \quad \tilde{R}_t = R_t. \quad (2)$$

Aligning to the object rotation as well drops no-warp replay success from 23.8% to 4.8%: GraspM3 object rotations are near-random per trial ($\|\Delta R\|_F \in [0.77, 2.83]$ within one object) and carry no approach information.

2.2. Demonstration warping

Following the DemoGrasp scales, the action is split into a wrist shift, a wrist rotation, and a finger offset,

$$\Delta p = s_p a_{1:3}, \quad \Delta \theta = s_\theta a_{4:6}, \quad \Delta q = s_q a_{7:12}, \quad (3)$$

with $(s_p, s_\theta, s_q) = (0.05 \text{ m}, 1.57 \text{ rad}, 1.0 \text{ rad})$ and $R_\Delta = R_x(\Delta \theta_1)R_y(\Delta \theta_2)R_z(\Delta \theta_3)$.

Wrist. Pre-lift frames are rigidly warped in the object-centric frame, rotating the approach about the object and shifting it:

$$\hat{p}_t = R_\Delta \tilde{p}_t + \Delta p, \quad \hat{R}_t = R_\Delta \tilde{R}_t \quad (t < t_\ell). \quad (4)$$

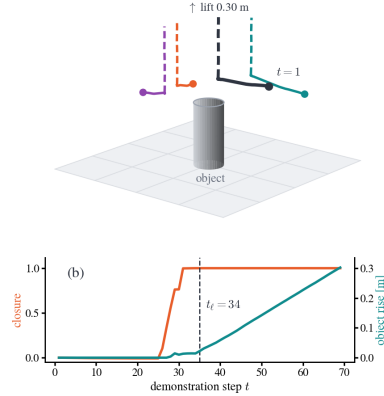


Figure 2. A trajectory demonstration and sampled warps. (a) Palm path of τ (black; solid: approach and grasp, dashed: lift) and three warps (colors): $\Delta \theta$ rotates the approach about the object, Δp shifts it, and the lift is copied verbatim (Eq. (5)). (b) Replay in simulation: fingers close over steps 25–31, lift onset $t_\ell = 34$, 0.30 m rise by $t = 69$.

Lift frames copy the demonstration’s relative displacement onto the warped grasp pose, so the vertical motion that made the demonstration succeed is never rotated or scaled:

$$\hat{p}_t = \hat{p}_{t_\ell-1} + (\tilde{p}_t - \tilde{p}_{t_\ell-1}) \quad (t \geq t_\ell). \quad (5)$$

The path returns to the world frame of the new episode (object at o) and to the action encoding: $p'_t = \hat{p}_t + o$, $x'_t = p'_t - c$, $r'_t = \text{Euler}_{XYZ}^{-1}(\hat{R}_t)$.

Fingers. Let q_1 be the open pre-grasp pose and q_{t_g} , $t_g = t_\ell - 1$, the closed grasp. The target grasp, $q^* = \text{clip}(q_{t_g} + \Delta q, q_{\min}, q_{\max})$ is achieved by scaling the closing motion

on a per-joint basis:

$$\lambda = (q^* - q_1) \odot (q_{t_g} - q_1), q'_t = \begin{cases} q_1 + \lambda \odot (q_t - q_1) & t < t_g \\ q^* & t \geq t_g, \end{cases} \quad (6)$$

so the closure morphs continuously into q^* with unchanged timing and the fingers keep squeezing through the lift.

Trajectory synthesis. The randomized reset pose (x_0, R_0) is generally far from the warped first frame (x'_1, R'_1) . We prepend a reach segment (LERP on position, SLERP on orientation) of $n = \max(\lceil \|x'_1 - x_0\|/0.01 \rceil, \lceil \angle(R_0, R'_1)/0.1 \rceil)$ steps, i.e. at most 1 cm and 0.1 rad per step, with fingers held at q'_1 . The plan $\Pi = [\text{reach}_{1:T_r} \parallel \text{track}_{1:T} \parallel \text{settle}_{1:10}]$ is executed by sending each frame as an absolute 12-D target to the task’s PD controller; the policy is never called again.

2.3. Warp policy and one-step PPO

The observation is the palm pose (7), the object pose (7), and a fixed-index subsample of 512 points of the task’s object-local cloud, encoded by the ShapeNet PointNet [7] auto-encoder shipped with the DexRep baseline (frozen); only a linear projection to $z \in \mathbb{R}^{128}$ is trained. Actor and critic are separate MLPs [1024, 1024, 512, 512] with ELU and orthogonal initialization. The policy is a tanh-squashed Gaussian, $a = \tanh(\tilde{a})$, $\tilde{a} \sim \mathcal{N}(\mu_\theta(o), \text{diag } \sigma^2)$, with σ learned (initialized at 0.8) and the tanh Jacobian correction in $\log \pi_\theta$. With horizon one there is no bootstrapping: the return is r , $V_\phi(o)$ is a variance-reducing baseline, and $\hat{A}$ is $r - V_\phi(o)$ standardized over the batch. We minimize the clipped PPO [8] objective

$$\mathcal{L} = \mathbb{E} \left[\max(-\hat{A}\rho_\theta, -\hat{A} \text{clip}(\rho_\theta, 1-\varepsilon, 1+\varepsilon)) \right] + c_v \mathbb{E}[(r - V_\phi(o))^2], \quad (7)$$

where $\rho_\theta = \pi_\theta(a|o)/\pi_{\theta_{\text{old}}}(a|o)$ denotes the policy ratio. We use a clipping threshold of $\varepsilon = 0.2$ and a value coefficient of $c_v = 2$. To ensure training stability, we clip gradients to a maximum norm of 1.0. Furthermore, we employ an adaptive learning rate strategy based on the Kullback-Leibler (KL) divergence. The learning rate is scaled by a factor of 1.5 to maintain a target KL divergence of 0.016, and is strictly bounded within the interval $[10^{-5}, 10^{-2}]$.

2.4. Large-scale training

Each iteration runs all 4,824 training objects in parallel, one environment per object, sharded over three GPUs (11.5 s per iteration); gradients, the KL statistic, and the advantage statistics are all-reduced, so the run equals a single 4,824-environment batch. Because object scale is fixed at simulation build time, we rebuild the simulation every K iterations with the next trajectory index, covering 51,659 of 52,718 trajectories (98%). The training process converged

Team (submission)	Easy [%]	Hard [%]	Rank
HandsDown (2)	91.29	77.45	1
UVLL HandDex (ours)	94.61	57.18	2
J-Team	87.20	43.88	3
Perception	68.80	23.98	4
OverGrasp	34.37	8.08	5
HCBHand	32.23	8.77	6
BC baseline (DexRep-TCN) [5]	23.93 [†]	–	–

Table 1. **Leaderboard** (private Objaverse test set, as of 1 Sept 2026): each team’s highest-ranked submission, ranked by team. [†]Reported by the organizers on their demo split.

Checkpoint	Evaluation split	Objects	Success [%]
Final (stage 2)	held-out traj. (valid)	4,502	90.98
Final (stage 2)	training traj. (train)	4,824	90.78
Stage 1	held-out traj. (valid)	4,502	86.81
64-object model	64 held-out objects	64	57.27 [‡]

Table 2. Local evaluation (deterministic, one rollout per object). The valid split shares no grasp sequence with train, but 4,487 of its 4,502 objects appear in training. [‡]Mean over 10 rounds on held-out objects; not comparable to the rows above.

over 4,000 iterations (approximately 16 hours). The initial 2,000 iterations utilized trajectory indices 0–3, rotated every 500 steps at an initial learning rate of 3×10^{-4} , then 2,000 with indices 0–19 rotated every 100. Resuming the converged stage-1 policy at the default learning rate collapsed it from 80.3% to 45.6% training success within 24 iterations; restarting at the rate the KL controller had settled on (10^{-5}) raised held-out success monotonically from 86.81% to 90.98%.

3. Experiments

Setup. We use three RTX 3090 GPUs. Local evaluation is deterministic (mean action, one rollout per object) in chunks of 1,500 objects, so peak memory is independent of object count; leaderboard numbers come from the organizers. Per the rules, our entrypoint changes four settings without editing the task: full 12-D absolute targets every step, the pose observation in place of per-step DexRep, vision enabled, and start-position canonicalization disabled. Initial poses and the success test are unchanged. Training uses only GraspM3 objects; the sole pretrained component is the baseline’s frozen PointNet.

Results. Table 1 compares teams. Our final submission scores 94.61% on the easy track, the highest easy-track number of any submission, and 57.18% on the hard track; HandsDown wins overall with 91.29/77.45. Scaling the training set from 64 objects (our second submission, 63.91/39.70) to all 4,824 added +30.7 pp (easy) and +17.5 pp (hard). Table 2 reports the local evaluation: held-

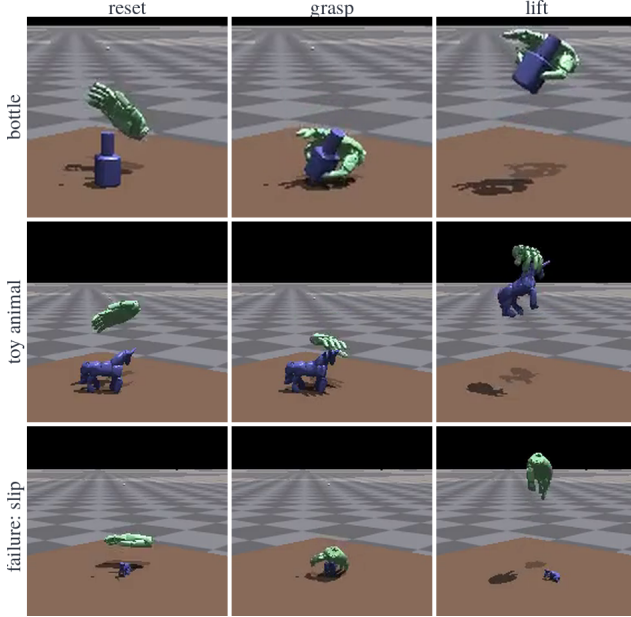


Figure 3. Qualitative evaluation of the final policy in simulation.

out trajectories score 0.2 pp *higher* than training trajectories, i.e. there is no measurable overfitting to the trained trajectories, and stage 1 alone reaches 86.81%, so object coverage rather than iteration count is the main lever.

Qualitative results and ablations. Figure 3 shows rollouts on a bottle and a toy animal: the warp places the palm on the surface, closes to a stable grasp, and the preserved lift completes the task. Failures follow one pattern: the object slips during the lift, and the plan proceeds without it. Three design choices had the largest effect: position-only alignment (23.8% vs. 4.8% no-warp replay on 64 objects), object coverage (64 $\rightarrow$ 4,824 objects: 63.9/39.7 $\rightarrow$ 94.6/57.2), and the learning rate on resumption (3×10^{-4} collapses a converged policy; 10^{-5} improves it monotonically).

Easy versus hard track. We retain 60.4% of our easy-track success on the hard track; HandsDown retains 84.8%. The gap is structural. (i) The plan is fixed before first contact, so an object that slides during the lift cannot be re-grasped. (ii) Mass and friction are invisible to the warp, which sees only geometry, and training used easy-track physics only. (iii) Grip force is inherited from the demonstration: finger targets are PD-tracked positions, and Δq changes the shape of the grasp, not how hard it holds under ten times the load. The single decision therefore solves the geometric problem, *where* and *how* to grasp, while the remaining gap is dynamic and requires feedback.

4. Conclusion

We ported DemoGrasp’s single-demonstration, one-step RL formulation to the LinkerHand O6 and scaled it to every

GraspM3 training object. One demonstration and one 12-D decision per episode reach 94.61% on unseen Objaverse objects, the highest easy-track score in the challenge, though only 57.18% under randomized mass and friction: the single decision solves the geometric problem, while closing the dynamic gap requires feedback.

References

- [1] Angel X Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, Jianxiong Xiao, Li Yi, and Fisher Yu. ShapeNet: An information-rich 3D model repository. *arXiv preprint arXiv:1512.03012*, 2015.
- [2] Matt Deitke, Dustin Schwenk, Jordi Salvador, Luca Weihs, Oscar Michel, Eli VanderBilt, Ludwig Schmidt, Kiana Ehsani, Aniruddha Kembhavi, and Ali Farhadi. Objaverse: A universe of annotated 3D objects. *arXiv preprint arXiv:2212.08051*, 2022.
- [3] HANDS Workshop Organizers. Dexterous grasp motion challenge 2026 (hands workshop at eccv 2026). <https://github.com/DexGraspMotionChallenge/DexGraspMotionChallenge2026>, 2026.
- [4] Haoming Li, Qi Ye, Yuchi Huo, Qingtao Liu, Shijian Jiang, Tao Zhou, Xiang Li, Yang Zhou, and Jiming Chen. TPGP: Temporal-parametric optimization with deep grasp prior for dexterous motion planning. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pages 18106–18112. IEEE, 2024.
- [5] Qingtao Liu, Yu Cui, Qi Ye, Zhengnan Sun, Haoming Li, Gaofeng Li, Lin Shao, and Jiming 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)*, pages 3153–3160. IEEE, 2023.
- [6] Viktor Makoviychuk, Lukasz Wawrzyniak, Yunrong Guo, Michelle Lu, Kier Storey, Miles Macklin, David Hoeller, Nikita Rudin, Arthur Allshire, Ankur Handa, and Gavriel State. Isaac Gym: High performance GPU-based physics simulation for robot learning. *arXiv preprint arXiv:2108.10470*, 2021.
- [7] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. PointNet: Deep learning on point sets for 3D classification and segmentation. *arXiv preprint arXiv:1612.00593*, 2016.
- [8] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.
- [9] Qi Ye, Haoming Li, Qingtao Liu, Shijian Jiang, Tao Zhou, Yuchi Huo, and Jiming Chen. Contact2Motion: Contact guided dexterous grasp motion generation with synergy embedded optimization. *The International Journal of Robotics Research*, page 02783649251364392, 2025.
- [10] Haoqi Yuan, Ziye Huang, Ye Wang, Chuan Mao, Chaoyi Xu, and Zongqing Lu. DemoGrasp: Universal dexterous grasping from a single demonstration. *arXiv preprint arXiv:2509.22149*, 2025.