Agentic Visuomotor Policy Learning via Real-to-Sim-to-Real

Kaifeng Zhang[1,2*] **Chao-Yuan Wu**[2] **Yunzhu Li**[1,2]
[1]Columbia University [2]World Labs

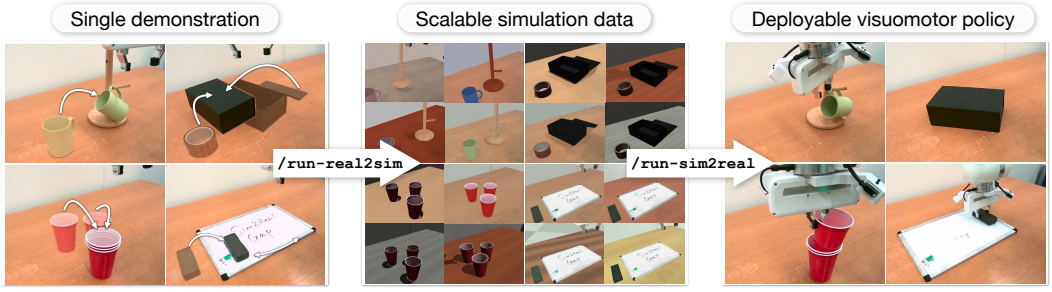


Figure 1: From a single real demonstration per task, a coding agent reconstructs the scene in simulation, generates randomized demonstrations at scale, and trains a visuomotor policy that runs on the real robot.

Abstract: Coding agents increasingly understand 3D scenes and solve manipulation tasks by calling motion planners and pretrained policies. Yet training the visuomotor policies underlying many such tasks still demands substantial human expertise. Agents must coordinate data generation, training, evaluation, and diagnosis, none of which can be resolved simply by tool calling; they require understanding and organizing the techniques behind each procedure. In this work, we present a harness designed to enable end-to-end agentic visuomotor policy learning. We leverage the idea of real-to-sim-to-real to translate a single real-world demonstration into simulation, and carry out the entire policy learning loop in simulation. Experiments demonstrate successful real-to-sim reconstruction and zero-shot sim-to-real transfer using our fully agentic pipeline, completing manipulation tasks on real hardware spanning rigid, articulated, and deformable objects. Project page: https://agentic-real2sim2real.github.io.

Keywords: Agentic robotics, visuomotor policy learning, real-to-sim-to-real

1 Introduction

Coding agents recently got significant improvements in their spatial understanding and robotic capabilities, unlocking surprising capabilities such as 3D modeling and simulation environment construction. Furthermore, by composing perception, motion planning, and control tools, they can turn task instructions into executable behavior and reusable skills [1, 2, 3, 4].

Pre-trained visuomotor policies remain a competitive alternative compared to purely agentic systems due to its real-time execution speed and reactive nature. For the same reason, they can serve as tool candidates for agents to call. However, training those visuomotor policies remains a bottleneck. Visuomotor learning requires expert choices about demonstrations, training settings, evaluation, and diagnosis throughout the process. Recent agentic systems automate parts of policy improvement [5], but reliably coordinating these choices end-to-end remains challenging, and doing so without exposing the robots to dangerous real-world failures is critical.

In this work, we apply the idea of real-to-sim-to-real [6, 7] and let agents train visuomotor policies entirely on data from simulations they build themselves. We introduce a harness (Fig. 2) that gives a

*Work partially done while interning at World Labs.

10th Conference on Robot Learning (CoRL 2026), Austin TX, USA.

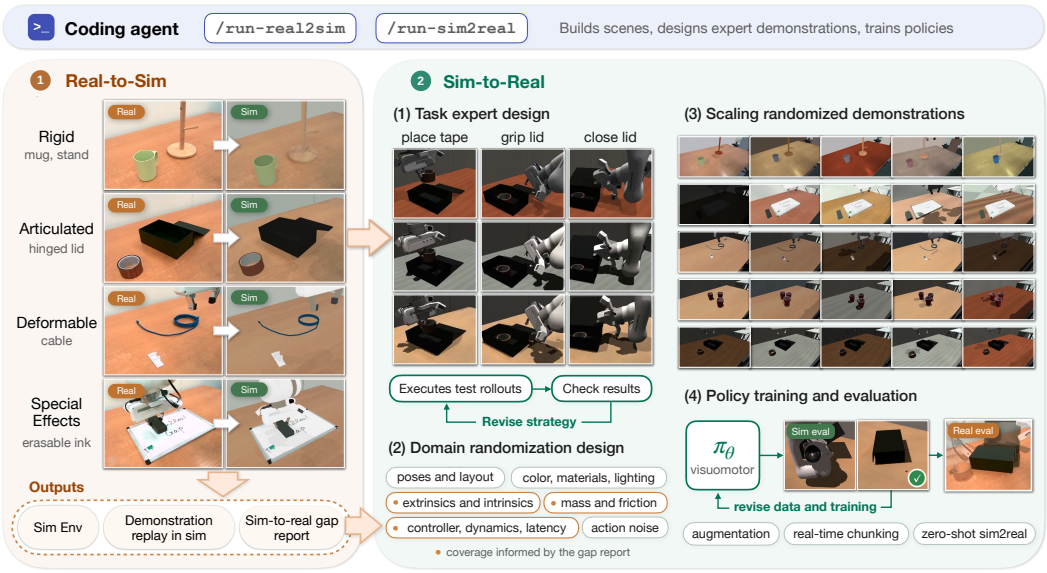


Figure 2: Harness overview. **Real-to-sim** (`/run-real2sim`): the agent reconstructs the scene from one demonstration, replays it in simulation, and reports sim-to-real gaps. **Sim-to-real** (`/run-sim2real`): it writes a task expert, designs domain randomization informed by the gap report, generates demonstrations at scale, and trains and evaluates the policy in closed loop. Finally, we deploy the policy zero-shot to the real robot.

coding agent two skills, `/run-real2sim` and `/run-sim2real`, and show that the resulting policies transfer zero-shot to the real robot on rigid, articulated, and deformable tasks.

2 Method

Real-to-sim. From calibrated recordings of a single real demonstration, the agent reconstructs a MuJoCo [8] scene and replays it in simulation. Beyond reconstructing a static scene, the agent must reproduce the entire task-completing trajectory, which requires physically plausible object parameters rather than visual appearance alone. It also reports remaining sim-to-real discrepancies, mostly from imperfect simulators and renderers, in a gap report for domain randomization.

Sim-to-real. The agent first writes a task expert and revises it from test rollouts, ensuring that the expert is robust under perturbations and varying initial conditions. It then designs domain randomization over layout, appearance, cameras, physics, controller dynamics, and action noise, and generates randomized demonstrations at scale. Next, it trains a visuomotor policy, evaluates it in closed loop, and revises the data strategy and training hyperparameters if needed. After evaluation in simulation, the agent runs the policy zero-shot on the real robot with safety measures; the policy executes without the agent in the loop, preserving real-time behavior. Throughout, the harness supplies procedural guidance and human-expert annotations, such as validation checks.

3 Experiments

Setup. We evaluate five tasks, each reconstructed from one real demonstration. For each task, the agent trains an ACT policy [9] only on simulated MuJoCo demonstrations. The policy observes four RGB cameras and proprioception and predicts 15-step chunks of end-effector poses and gripper commands at 15 Hz. Each checkpoint is evaluated both in simulation and on the real robot.

Results. For real-to-sim, we find that agents automatically determine object materials, both physical and visual, and implement effects the simulator does not support, such as erasable ink. For sim-to-real, agents automatically diagnose their demonstration-generation and domain-randomization strategies and adapt them to each task. All policies transfer zero-shot to the real robot: mug 88%, tape 72%, cups 80%, whiteboard 52%, and cable 76%. More results and discussion are available on our website: https://agentic-real2sim2real.github.io.

References

[1] J. Liang, W. Huang, F. Xia, P. Xu, K. Hausman, B. Ichter, P. Florence, and A. Zeng. Code as policies: Language model programs for embodied control. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 9493–9500. IEEE, May 2023. doi: 10.1109/icra48891.2023.10160591. URL http://dx.doi.org/10.1109/icra48891.2023.10160591.

[2] L. Fu, J. Yu, K. El-Refai, E. Kou, H. Xue, H. Huang, W. Xiao, L. Fei-Fei, G. Shi, J. Wu, S. S. Sastry, Y. Zhu, K. Goldberg, and L. Fan. CaP-X: A framework for benchmarking and improving coding agents for robot manipulation. In T. Zhang, M. Dudik, M. Jaggi, A. Agarwal, S. Li, D. Schuurmans, J. Zhu, F. Berkenkamp, H. Dong, and A. Bietti, editors, *Proceedings of the 43rd International Conference on Machine Learning*, volume 306 of *Proceedings of Machine Learning Research*, pages 32155–32212. PMLR, 06–11 Jul 2026. URL https://proceedings.mlr.press/v306/fu26u.html.

[3] R. Lu, Y. Wu, E. Kou, H. Lin, L. Fu, W. Xiao, A. Mandlekar, Y. Xu, G. Shi, K. Goldberg, A. Chen, M. Chowdhury, J. Kautz, Y. Zhu, L. J. Fan, and G. Wang. ASPIRE: Agentic /skills discovery for robotics, 2026. URL https://arxiv.org/abs/2607.00272.

[4] J. Hu, R. Martin-Martin, P. Stone, R. Duan, Z. Jiang, and G. Shi. SimEX: Simulation-integrated robotics autoresearch, 2026. URL https://arxiv.org/abs/2609.38982.

[5] W. Xiao, J. Xie, T. Zhang, H. Lin, L. M. Fu, H. Xue, J. Lu, Y. Yang, C. Dai, Z. Wang, J. Wu, G. Wang, S. S. Sastry, K. Goldberg, L. J. Fan, Y. Zhu, and G. Shi. ENPIRE: Agentic robot policy self-improvement in the real world, 2026. URL https://arxiv.org/abs/2606.19980.

[6] M. T. Villasevil, A. Simeonov, Z. Li, A. Chan, T. Chen, A. Gupta, and P. Agrawal. Reconciling Reality through Simulation: A Real-To-Sim-to-Real Approach for Robust Manipulation. In *Proceedings of Robotics: Science and Systems*, Delft, Netherlands, July 2024. doi:10.15607/RSS.2024.XX.015.

[7] T. Dai, J. Wong, Y. Jiang, C. Wang, C. Gokmen, R. Zhang, J. Wu, and L. Fei-Fei. Automated creation of digital cousins for robust policy learning. In P. Agrawal, O. Kroemer, and W. Burgard, editors, *Proceedings of The 8th Conference on Robot Learning*, volume 270 of *Proceedings of Machine Learning Research*, pages 4912–4943. PMLR, 06–09 Nov 2025. URL https://proceedings.mlr.press/v270/dai25a.html.

[8] E. Todorov, T. Erez, and Y. Tassa. MuJoCo: A physics engine for model-based control. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5026–5033. IEEE, Oct. 2012. doi:10.1109/iros.2012.6386109. URL http://dx.doi.org/10.1109/iros.2012.6386109.

[9] T. Z. Zhao, V. Kumar, S. Levine, and C. Finn. Learning Fine-Grained Bimanual Manipulation with Low-Cost Hardware. In *Proceedings of Robotics: Science and Systems*, Daegu, Republic of Korea, July 2023. doi:10.15607/RSS.2023.XIX.016.