# Sim-to-Real Peg-in-Hole Assembly Learning from Peg-out-Hole Disassembly

Yongqiang Zhao, Zhuo Chen, Xuyang Zhang, and Shan Luo

***Abstract*— Peg-in-hole (PiH) assembly is a fundamental yet challenging robotic manipulation task. Although reinforcement learning (RL) offers potential solutions, it require extensive exploration and carefully-designed reward functions. In this paper, we propose a novel PiH skill-learning framework that leverages peg-out-hole (PoH) disassembly to facilitate PiH learning. To this end, we first collect a large dataset of PoH trajectories in simulation, which are then inverted to generate training data for PiH. To bridge the Sim-to-Real gap, the policy is fine-tuned with tactile measurements to compensate for real-world peg-hole misalignment. Compared to direct RL approaches that train PiH policies from scratch, our method achieves a twofold improvement in both learning speed and Sim-to-Sim success rates. Extensive experiments across various peg-hole geometries and robotic setups, validate the effectiveness of our framework, achieving an average success rate of 90.4% across all tested objects.**

## I. INTRODUCTION

Peg-in-hole (PiH) task is pivotal in both industrial applications and everyday activities such as gear assembly and USB insertion [1]. It can be formulated as a target localization problem [2], which requires precise insertion of a peg into a hole. Various methods, including reinforcement learning (RL)[1], supervised learning[2], and control-based approaches [3], have been explored to address this challenge. However, RL and supervised learning approaches demand large datasets and precise supervision, while control-based methods are hindered by target variability and intricate contact dynamics in PiH tasks. Despite humans excel at adapting new skills from related experiences [4], most research on PiH assembly overlooks its inherent connection to the inverse peg-out-hole (PoH) task, which could simplify exploration and enhance insertion strategies.

In this work, we propose a novel skill learning framework for peg-in-hole assembly that leverages a trained peg-out-hole policy to generate PiH trajectories in simulation, which are then transferred to the real world. To account for the distinct force feedback between PiH and PoH [5], we integrate tactile sensing to compensate peg-hole misalignment. Our proposed framework significantly accelerates learning compared to direct RL approaches that train the PiH policies from scratch. Furthermore, extensive Sim-to-Real experiments on both single-arm and dual-arm robotic setups, as well as various peg and hole geometries, demonstrate that the generalizability of our proposed framework.

Department of Engineering, King's College London, Strand, London, WC2R 2LS, United Kingdom, {yongqiang.zhao, zhuo.7.chen, xuyang.zhang, shan.luo}@kcl.ac.uk.

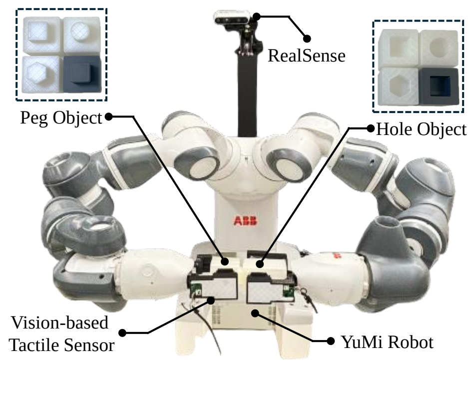


Fig. 1. The real-world dual-arm peg-in-hole setup.

## II. METHODOLOGY

### A. Peg-out-hole Policy Learning

Figure 2 (a) illustrates that the peg-out-hole (PoH) policy $\overrightarrow{\pi}$ is developed through reinforcement learning (RL), leveraging privileged access to object state information in simulation. This facilitates the accumulation of an extensive dataset of PoH trajectories $\overrightarrow{\tau} = \{(\overrightarrow{S}_t, \overrightarrow{A}_t)_{t=1}^T\}$, where $\overrightarrow{S}_t$ and $\overrightarrow{A}_t$ denote the input state and output action, respectively. The input state comprises the robot's kinematic state $\overrightarrow{R}_t$, privileged object state $\overrightarrow{O}_t$, and the desired goals $\overrightarrow{G}$.

### B. Peg-in-hole Policy Learning

By reversing the PoH process, where the robot execues inverse actions to ones in peg-in-hole (PiH), the task transitions into a PiH process. As shown in Figure 2 (b), a subset of the collected PoH trajectories is inverted to create the training dataset $D_{\overleftarrow{\tau}} = \{\overleftarrow{\tau}_i\}_{i=1}^N = \{\{(\overleftarrow{S}_t^i, \overleftarrow{A}_t^i)\}_{t=1}^T\}_{i=1}^N$ for the PiH task. Domain randomization is employed during the collection of these trajectories to increase the policy's robustness. In this setup, $\overleftarrow{S}_t = (\overrightarrow{R}_{T-t}, \overrightarrow{I}_{T-t})$ represents the input state, and $\overleftarrow{A}t = -\overrightarrow{A}_{T-t-1}$ denotes the labeled action for training $\overleftarrow{\pi}$, where $\overrightarrow{I}$ indicates visual observation.

### C. Tactile-based Peg-in-hole Policy Fine-tuning

After training the simulated PiH policy, we transfer it to the real world, where tactile-based fine-tuning compensates for peg-hole misaglignment during deployment, as is shown in Fig. 2 (c). Specifically, we discretize the misalignment between the peg and hole objects into 9 contact quadrants based on the patterns of the marker motion in the tactile images. This allows us to determine the appropriate movement directions for the robot's grippers.

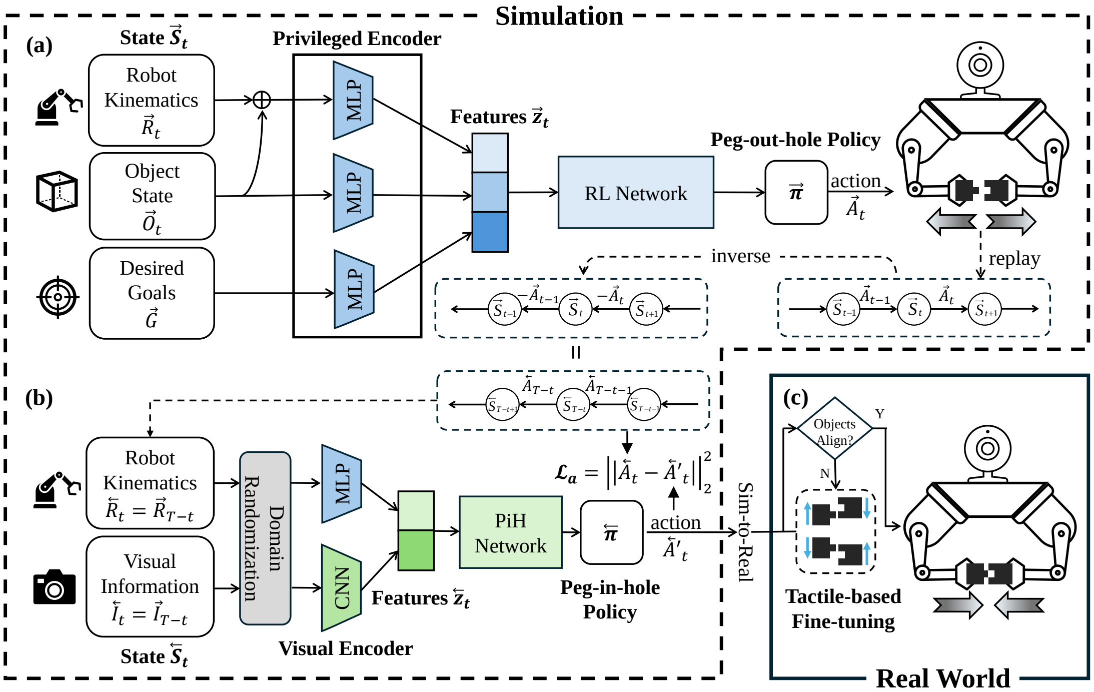


Fig. 2. Overview of the proposed peg-in-hole (PiH) skill learning framework.

TABLE I

SIM-TO-REAL SUCCESS RATES OF PEG-IN-HOLE TASK

| Object Types | Single-arm | | Dual-arm | |
|---|---|---|---|---|
| | *w/o Tactile* | *w/ Tactile* | *w/o Tactile* | *w/ Tactile* |
| White Cube | 53.3% | **93.3%** | 46.7% | **93.3%** |
| White Cylinder | 60.0% | **96.7%** | 53.3% | **93.3%** |
| White Hexagon | 46.7% | **86.7%** | 36.7% | **83.3%** |
| Black Cube | 50.0% | **90.0%** | 40.0% | **86.7%** |

## III. EXPERIMENT RESULTS & ANALYSIS

The real-world setup features an ABB YuMi dual-arm robot equipped with four GelSight-like vision-based tactile sensors on its fingers and an Intel RealSense camera mounted on its body (Fig. 1). The robot's right gripper holds the peg objects, while the left holds the hole objects. The camera is employed to capture visual information of the environment, while the tactile sensors are utilized to fine-tune the peg-in-hole policy in real-world scenarios. During the experiments, both robotic arms initiate movement from random positions and coordinate through joint control to achieve alignment between the peg and the hole.

After training the PiH policies for both single-arm (left gripper fixed) and dual-arm setups in simulation, we transfer them to the real world. The quantitative results (Table I), compare the performance of the policies with and without tactile-based fine-tuning. We identified two failure scenarios: (1) a small misalignment between the peg and hole, which accounts for the majority of failures, and (2) a large misalignment where the peg and hole fail to make contact. Tactile feedback can determine the contact configurations and provides an adjustment strategy for real robots in the first case, while the system terminates directly in the second case. These findings demonstrate that incorporating tactile sensing enhances PiH policy performance in both single-arm and dual-arm setups. Furthermore, the results indicate that the proposed skill-learning framework enables different types of real robots to acquire the peg-in-hole skill and generalize it across objects with various geometries.

## IV. CONCLUSIONS

In this work, we propose a Sim-to-Real skill learning framework for robot peg-in-hole assembly that allows robots to efficiently learn insertion skills. The key module of this framework is a PoH-to-PiH policy learning method that leverages trajectories from PoH tasks to accelerate PiH skill learning. Meanwhile, tactile sensing is utilized to fine-tune the peg-hole misalignment. Sim-to-Real experiments across different robot configurations and various peg and hole shapes demonstrate the effectiveness of the proposed framework. In the future, we plan to extend this method to a broader range of robot manipulation tasks.

## REFERENCES


[1] Y. Wang, L. Zhao, Q. Zhang, R. Zhou, L. Wu, J. Ma, B. Zhang, and Y. Zhang, "Alignment method of combined perception for peg-in-hole assembly with deep reinforcement learning," *Journal of Sensors*, vol. 2021, no. 1, p. 5073689, 2021.

[2] S. Dong and A. Rodriguez, "Tactile-based insertion for dense box-packing," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 7953–7960.

[3] J. Xu, K. Liu, Y. Pei, C. Yang, Y. Cheng, and Z. Liu, "A noncontact control strategy for circular peg-in-hole assembly guided by the 6-dof robot based on hybrid vision," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–15, 2022.

[4] A. Walsman, M. Zhang, K. Kotar, K. Desingh, A. Farhadi, and D. Fox, "Break and make: Interactive structural understanding using lego bricks," in *European Conference on Computer Vision*. Springer, 2022, pp. 90–107.

[5] F. Goli, Y. Zhang, M. Qu, Y. Zang, M. Saadat, D. T. Pham, and Y. Wang, "Jamming problems and the effects of compliance in dual peg-hole disassembly," *Proceedings of the Royal Society A*, vol. 480, no. 2286, p. 20230364, 2024.