Construction Robot Skill Learning for Fragile Object Installation with Low-Effort Demonstration and Sample-Efficient Hierarchical Reinforcement Learning Models

Vaidhyanathan Chandramouli, Hongrui Yu, and Ci-Jyun Liang

Abstract— **Construction robots are increasingly recognized as a key solution to address labor shortages and stagnant productivity in the construction industry. However, their widespread adoption faces several critical challenges: (1) limited skill sets—while many skill transfer studies exist, there are limited studies focusing on tasks involving collision-free fragile material manipulation; (2) the absence of a generalizable simulation environment compatible with major robot learning algorithms, hindering future benchmarking efforts; and (3) the high data demands typically required for training, which place a heavy burden on human workers. To tackle these issues, this paper presents three main contributions: (1) a robot skill learning algorithm that incorporates collision force awareness for improved handling of fragile materials, (2) a simulation environment built using MuJoCo, designed to support a wide range of robot learning algorithms, and (3) a hierarchical learning framework that significantly reduces the amount of demonstration data needed. Experimental results demonstrate the effectiveness of the proposed hierarchical approach in enhancing robot skill learning performance.**

I. Introduction

With severe labor shortages, the construction industry is in desperate need of alternative sources of labor [1]. Robots, with their high physical capabilities, are regarded as a promising solution [2][3]. However, the limited development of material manipulation skills in robots, combined with the extensive data requirements placed on human workers to instruct and demonstrate tasks, has significantly slowed the adoption of construction robotics [4][5]. This underscores the need for robot learning systems that are low in human workload, sample-efficient, and effective in skill abstraction and replication.

On the other hand, many current robot skill learning models emphasize the potential for generalization across platforms and tasks [6]. While valuable from a research perspective, this emphasis often increases the amount of training data required compared to more targeted approaches, such as joint-state-based learning. From a practical standpoint, a typical construction company is more likely to adopt a single robot platform that offers optimal cost-performance and focuses on training it for the most urgent or repetitive task. With such considerations, methods that prioritize focused learning on a specific platform with minimal data overhead are more aligned with the realistic constraints of industry deployment.

Additionally, fragile building materials such as glass panes and photovoltaic (PV) panels are ubiquitous in modern construction, yet current robotic learning approaches have barely addressed their manipulation [7]. Human workers still carry out most glass installation and PV panel placement tasks, which are both physically taxing and can lead to significant strain and risk of injury [8]. The study of replicating human's skills and delegating such tasks to robots is crucial for occupational safety and health. When performing such tasks, human workers rely on the tactile input and senses of force to determine the actions during installations and protect the material from being damaged. For a robot to replicate such skill and take over the installation tasks, a learning model incorporating the tactile input is necessary.

II. Literature Review

As outlined in the Introduction section, several challenges remain before robots can effectively take over fragile material manipulation tasks. While some of these challenges are grounded in practical realities, many can be addressed through advances in engineering design and system-level coding innovations. There are several algorithms and models advancing temporal and behavioral abstraction efficiency, including but not limited to the options framework [9][10], latent skill models [11], and diffusion-based policies [12]. The most critical bottleneck, however, lies in developing sample-efficient learning models capable of representing human manipulation skills. This section reviews recent progress in sample-efficient learning methods, which serve as the foundation for the proposed methodological framework.

Hierarchical models have become a popular option due to the potential for learned skill reuse and strong suitability for easily readable temporal abstraction of skills [13]. For example, ROMAN was proposed to perform hierarchical behavior cloning from human demonstrations with several gating networks [13]. It was able to achieve skill replication with only three demonstrations. Moreover, CRISP was proposed as a Hierarchical Reinforcement Learning (HRL) algorithm to replicate the long horizon trajectory-based skills in simulated environments [14]. It largely reduced the demonstration data needs to 5-10 trajectories. However, even though both models demonstrated success in skill replication,

V.C. Author is with the Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA 24061 USA (e-mail: vaidhyanathan@vt.edu).

H.Y. Author is with the Department of Civil and Environmental Engineering, Virginia Tech, Blacksburg, VA 24061 USA (corresponding author to provide phone: 540-231-0591; fax: 540-231-7532; e-mail: hryu42@vt.edu).

C.L. Author is with the Department of Civil Engineering, Stony Brook University, Stony Brook, NY 11794 USA (e-mail: ci-jyun.liang@stonybrook.edu).

they did not consider multimodal input data, such as combining both trajectory and tactile input data. Therefore, in this paper, one of the research goals is to validate if adding force data and improving demonstration modality will help with HRL model training.

III. METHODOLOGY

This paper addresses key challenges in the adoption of construction robots and robot learning, namely the lack of suitable simulation environments, high demonstration workload, and low sample efficiency. To tackle these issues, we propose the following contributions:

- Development of a collision force-aware simulation environment: We implemented a simulation environment in MuJoCo [15] that models the physical dynamics of the collision process, capturing forces between the panel and the track during installation tasks.
- Design of a low-effort demonstration system: To reduce the demonstration workload and improve sample efficiency, we introduced a mouse-based teleoperation system. This method allows users to intuitively demonstrate joint motions, enhancing ergonomic usability and data efficiency for individual robots.
- Implementation of HRL: To further enhance learning efficiency, we developed a hierarchical reinforcement learning framework that autonomously identifies subgoals and segments trajectories. Two Deep Q-Networks (DQNs) are employed to learn policies for each trajectory segment.

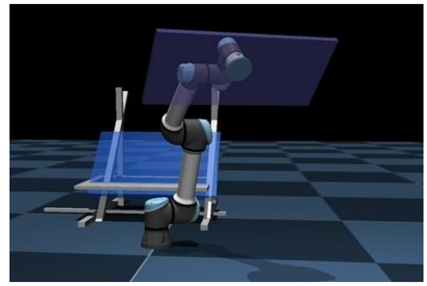
The technical details of these contributions are elaborated in the subsections below.

A. Simulation Environment

One common challenge faced by construction robotics researchers is the lack of publicly available robot simulation models that are directly compatible with the majority of robot learning models. Observing this challenge, the authors select MuJoCo as the base of the simulation environment. Construction task-related 3D models and robot models are added to the simulation environment. Gravity compensation was also enabled. The physical properties of objects are also modified based on real-world settings. A screenshot of the system is shown in Fig. 1.

The environment includes the installation object, target, a 3D model of a robot, and physical properties that enable the simulation of collision effects (e.g. visual arrows scaled to the magnitude of the collision force). Currently, the simulation has been utilized for solar PV panel installation and DQN-based reinforcement learning. However, it is designed to be extensible, allowing for easy integration of additional tasks and learning models [15].

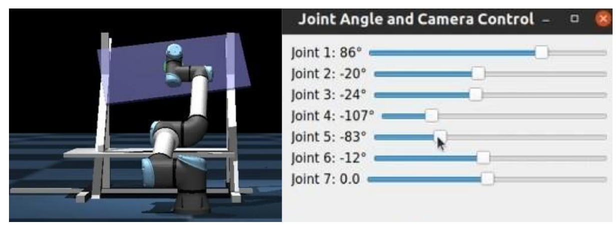
Figure 1. The robot picked up a panel in MuJoCo.



B. Demonstration System

This paper provides a simulation environment that allows human workers to teleoperate the robot. The workload of demonstration is minimized to use the mouse to drag the joint state controller pane, as shown in Fig. 2. This is a more ergonomic approach compared to asking workers to perform the tasks repeatedly [5].

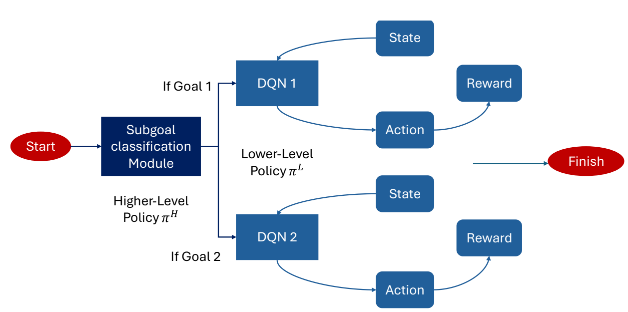
Figure 2. Teleoperation with joint manipulation in the MuJoCo simulation environment.



C. Hierarchical Reinforcement Learning Algorithm

This study adopts the hierarchical structure of the learning algorithm to improve sample efficiency. As mentioned in the Literature Review section, HRL is a commonly used method to separately train a model that works better for a small segment of the data and thus improves the model learning efficiency. The HRL model proposed in this paper is shown in Fig. 3. The HRL model has two components: a subgoal classification network and the reinforcement learning parts. The subgoal classification network is a multi-layer perceptron (MLP) that performs supervised learning based on the input of state spaces (force data and joint space data). A total of 160 trajectories were used to train this network. The DQN has three layers: linear, ReLU, and linear layer. The reward function was set to penalize the distance from the target and the collision force between the panel and the target track.

Figure 3. The proposed HRL architecture.



The reward function is formulated as follows:

$$\text{Stage 1 Pick-up: } R_1 = -1 \times \alpha \times D_1 \quad (1)$$

$$\text{Stage 2 Installation: } R_2 = -1 \times \beta \times D_2 - \gamma \times F \quad (2)$$

Where R_1 and R_2 are the reward functions calculated at each step, D_1 and D_2 are the distances toward the subgoal/target at one specific stage, F is the collision force between the target track and the held object, and α, β, γ are scaling coefficients. After each installation phase, the reward will be reset.

IV. Experimental Validation and Task

This paper presents a case study on solar panel installation, selected due to the panels' substantial weight and the increasing need to delegate such labor-intensive tasks to robotic systems. The manipulated object, the solar panel, was modeled with carefully calibrated physical and material properties to ensure realistic simulation. In particular, parameters such as maximum allowable force and elasticity were derived based on the known characteristics of glass materials.

To validate the proposed system, computational experimental evaluation was conducted. In the computational experiments, the evolution of the reward function was analyzed in relation to task progression. The authors further assessed the effectiveness of the hierarchical learning scheme by comparing model training performance with and without the hierarchical learning model.

V. Results

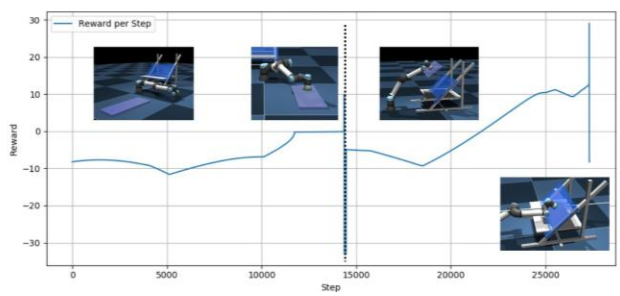
The results of the model training are summarized as follows. Initially, the proposed system underwent standard training and testing procedures. As illustrated in Fig. 4, the reward function demonstrated a clear upward trend as the robot progressed toward the goal during the first stage. This observed increase in reward validates the effectiveness of the designed reward function in guiding the robot toward successful task completion. (Note: the authors manually reset the step at the end of each phase hence the sudden drop of the reward.)

With the proposed reward functions validation, the authors adopted a metric to evaluate the model learning performance, meaning that a higher reward within the same training duration and with the same training effort will be regarded as a better algorithm. The authors compared the reward between DQN with and without the subgoal classification process. The results are shown in Table I. The hierarchy DQN achieved a higher final step reward and average reward, demonstrating the promising performance.

TABLE I. Comparison of Model Training Rewards

Evaluation Metric	**Algorithm**	
	DQN	***Hierarchical DQN***
Final Step Reward	-4.38	29.42
Average Reward	-13.55	-1.62

Figure 4. Reward change during installation and key robot states.



VI. Discussions and Conclusions

This paper is motivated by observing the slow adoption of construction robots in the construction industry. Several reasons were identified for the slow adoption, including the lack of robot skill learning models for construction, the data-intensive requirements of learning models that increase the demonstration workload of workers, and the lack of a benchmark simulation environment that connects construction tasks with advanced robot learning models. To address these challenges, this paper worked on the following technical advances:

- Developed a robot simulation environment using MuJoCo that allows convenient deployment and application of advanced robot learning models.
- Added force simulation to the MuJoCo environment to improve the sensory input modality and compatibility with construction tasks.
- Designed a hierarchical robot learning model with HRL and adopted a teleoperation-based demonstration method to minimize the data collection workload and improved the ergonomics of the demonstration process.

To validate the proposed system, computational and robotics experiments are performed. The authors compared the proposed hierarchical DQN with the traditional DQN method. The results demonstrated that the proposed hierarchical DQN can achieve a good training performance.

Admittedly, this study has several limitations that present opportunities for future work. First, the current evaluation is limited to a single model architecture with the MLP-DQN structure. While this structure demonstrated promising results, the dataset (high-dimensional joint state data), algorithm, and its components (e.g., reward function design) require more comprehensive validation through benchmarking against alternative learning frameworks on larger, standardized datasets. Second, the simulation environment currently supports only a single task: solar panel installation. Although this scenario reflects a relevant and challenging construction application, it restricts the generalizability of the approach. Ongoing efforts are focused on expanding the simulation to include a broader set of construction tasks. This will enable more diverse training scenarios and facilitate the development of more versatile and robust robotic policies for real-world deployment.

ACKNOWLEDGMENT

This research was conducted without any external financial support.

REFERENCES

[1] Hardy, N., Barbosa-Resende, C., Kereri, J., & Taghinezhad, A. (2024). Root and Causes of Skilled Labor Shortage in Construction Industry: Comprehensive Literature Review. *The Professional Constructor, 18*-36.

[2] Liang, C. J., Wang, X., Kamat, V. R., & Menassa, C. C. (2021). Human–robot collaboration in construction: Classification and research trends. *Journal of Construction Engineering and Management, 147*(10), 03121006.

[3] Park, S., Yu, H., Menassa, C. C., & Kamat, V. R. (2023). A comprehensive evaluation of factors influencing acceptance of robotic assistants in field construction work. *Journal of Management in Engineering, 39*(3), 04023010.

[4] Liang, C. J., Kamat, V. R., Menassa, C. C., & McGee, W. (2022). Trajectory-based skill learning for overhead construction robots using generalized cylinders with orientation. *Journal of Computing in Civil Engineering, 36*(2), 04021036.

[5] Yu, H., Kamat, V. R., & Menassa, C. C. (2024). Cloud-based hierarchical imitation learning for scalable transfer of construction skills from human workers to assisting robots. *Journal of Computing in Civil Engineering*, *38*(4), 04024019.

[6] Rana, M., Mukadam, M., Ahmadzadeh, S. R., Chernova, S., & Boots, B. (2018, January). Towards robust skill generalization: Unifying learning from demonstration and motion planning. *In Intelligent robots and systems*.

[7] Lee, S., & Moon, J. I. (2015). Case studies on glazing robot technology on construction sites. In ISARC. Proceedings of *the International Symposium on Automation and Robotics in Construction (Vol. 32*, p. 1). IAARC Publications.

[8] Guinsler, V., & Schwerha, D. (2018). An Examination of the Musculoskeletal Impact of Residential Solar Panel Installation.

[9] Sutton, R. S., Precup, D., & Singh, S. (1999). Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artificial intelligence*, *112*(1-2), 181-211.

[10] Yang, X., Ji, Z., Wu, J., Lai, Y. K., Wei, C., Liu, G., & Setchi, R. (2021). Hierarchical reinforcement learning with universal policies for multistep robotic manipulation. *IEEE Transactions on Neural Networks and Learning Systems*, *33*(9), 4727-4741.

[11] Eysenbach, B., Gupta, A., Ibarz, J., & Levine, S. (2018). Diversity is all you need: Learning skills without a reward function. *arXiv preprint arXiv:1802.06070*.

[12] Li, W. (2023). Efficient planning with latent diffusion. *arXiv preprint arXiv:2310.00311*.

[13] Triantafyllidis, E., Acero, F., Liu, Z., & Li, Z. (2023). Hybrid hierarchical learning for solving complex sequential tasks using the robotic manipulation network roman. *Nature Machine Intelligence, 5*(9), 991-1005.

[14] Singh, U., & Namboodiri, V. P. (2023). Crisp: Curriculum inducing primitive informed subgoal prediction. *arXiv e-prints*, arXiv-2304.

[15] Todorov, E., Erez, T., & Tassa, Y. (2012, October). Mujoco: A physics engine for model-based control. In *2012 IEEE/RSJ international conference on intelligent robots and systems* (pp. 5026-5033). IEEE.