Robotic Assembly of Deformable Linear Objects via Curriculum Reinforcement Learning

Kai Wu[1*], *Member, IEEE*, Rongkang Chen[1], Qi Chen[1] and Weihua Li[2], *Senior Member, IEEE*
1. Shien-Ming Wu School of Intelligent Engineering, South China University of Technology, Guangzhou 510640, China
2. School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou 510640, China
Corresponding author: Kai Wu, whphwk@scut.edu.cn

***Abstract*—The automated assembly of flexible objects presents significant challenges. Although significant progress has been made in the assembly of rigid objects, the methods used for rigid objects cannot be directly applied to flexible objects due to their infinite degrees of freedom. This study proposes a reinforcement learning (RL) based method for deformable cable insertion tasks executed with a universal 2-finger gripper. Firstly, a vision-based detection method is employed to monitor the cable's state in real time, while a state classifier is introduced to provide real-time reward feedback for RL training. Secondly, an adaptive curriculum learning (CL) method is proposed to adjust the initial degree of cable bending through the success rate in the training process, allowing the RL agent to learn progressively from easier to more difficult tasks. The validation experiments were conducted on a type-C cable insertion task, where the robot grips the cable portion of the electrical connector. The results indicate that our method is capable of adapting to various degrees of cable bending, successfully handling cable configurations bent up to a maximum of 40° from its straight, unbent state, with an assembly success rate of over 90%.**

Keyword: Deformable linear objects, robotic assembly, curriculum learning.

I. INTRODUCTION

The peg-in-hole (PiH) problem has been widely studied, but most works focus on rigid objects while neglecting deformable linear objects (DLOs) such as cables and wires. DLOs are essential in consumer electronics, automotive, and aerospace industries, yet their infinite deformation degrees of freedom make automated assembly extremely challenging. Existing studies often avoid cable deformation by grasping rigid connectors, but in practical scenarios like electric vehicle battery production, cables differ in interfaces, dimensions, and flexibility, while installation spaces are highly constrained. These conditions limit robot grippers and end-effectors, highlighting the importance of developing universal and compact solutions for flexible cable assembly in confined environments.

Reinforcement learning (RL) has shown strong performance in robotic assembly but faces challenges with DLOs. Accurate simulation modeling is difficult, causing a gap between simulation and reality, while real-world RL training suffers from low efficiency. Moreover, dynamic connector poses complicate the design of dense reward functions, which are critical for RL performance.

Inspired by human strategies of manipulating cables using both hands and environmental contact, we propose an RL-based framework that exploits contact between the cable and environment to regulate pose and achieve assembly. Expert demonstrations and curriculum learning (CL) are introduced to alleviate reward design difficulties and improve data efficiency. This approach enables the agent to gradually acquire insertion strategies under sparse rewards, eventually achieving cable assembly with bend angles up to 40°.The main contributions are as follows:

1. An RL-based method for flexible cable connector assembly, allowing compliant insertion by leveraging environmental contact forces.
2. A vision-based monitoring approach and an assembly state classifier that provide real-time reward feedback during training.
3. An adaptive CL method that automatically adjusts initial cable bend angles based on success rates, improving training efficiency.
4. Real-world experiments on type-C cable insertion, validating the effectiveness of the proposed method.

II. TASK DESCRIPTION

This study addresses the task of inserting an electrical connector into a socket using a universal two-finger gripper. As shown in Fig. 1, the cable is initially bent along the x-axis in our task setup, which makes controlling its bending crucial for completing the task. As shown in Fig. 1(a), although the connector and socket are aligned, the fit between them is typically a transition fit, requiring sufficient force to complete the insertion. Even if the robot moves toward the socket, the cable tends to bend further, making it difficult to apply force effectively to the connector. As shown in Fig. 1(b), although the robot's z-axis aligns with the socket, the bending of the cable can easily cause jamming. In such cases, moving the robot toward the socket fails to complete the task and may damage the cable. Therefore, this study involves aligning the connector with the socket and controlling the shape of the cable (Fig. 1. (c)), thereby enabling successful insertion.

This study is based on the following assumption: the direction of the socket is known, the robot and the gripper are positioned perpendicular to the socket, the cable is held at the center of the gripper, and the cable is bent along the x-axis. However, the extent of the bending and the position of the socket are unknown.

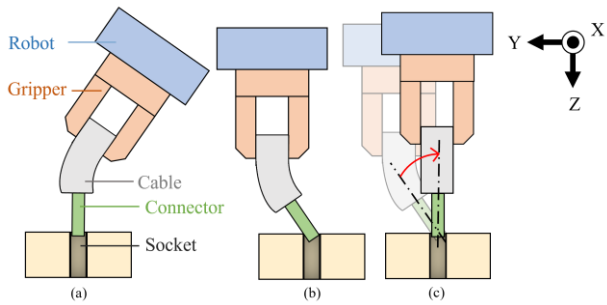

Fig. 1. Strategy for assembly tasks. (a) Connector and socket alignment often require force for insertion due to the transition fit. Additionally, cable bending can obstruct the process. (b) Cable bending causes jamming and potential damage. (c) Using interaction to straighten the cable facilitates alignment and insertion.

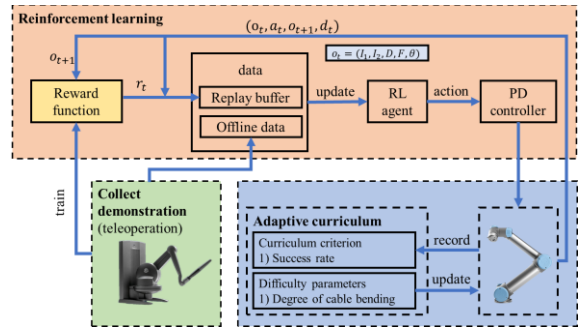

Fig. 2. System pipeline for the insertion task. The green section involves collecting teleoperated demonstrations for offline RL training and classifier development. The blue section adjusts task difficulty online based on success rate. The orange section integrates online and offline RL training, with a PD controller executing policy actions to control the robot.

III. METHOD

We proposed the curriculum RL-based assembly method for the assembly task. It mainly consists of the following three parts: i) A CL method that adaptively adjusts difficulty level based on task success rate; ii) A RL observation space specifically designed for cable assembly tasks; iii) A classifier-augmented reward function. Fig. 2. shows the overall pipline.

A. Adaptive Curriculum Learning

In our experimental setup, the curriculum is defined by the initial degree of cable bending, which increases as the task difficulty level rises. The curriculum difficulty is represented as a coefficient ranging from 0 to 1, which adjusts dynamically based on the agent's performance. Initially set to 0, the difficulty increases or decreases by 0.1 depending on the success rate over the most recent 20 episodes.

B. RL for Cable Insertion Tasks

We consider the insertion tasks as a Markov decision process (MDP). We chose RLPD [1], an off-policy RL algorithm, to train our assembly policy and utilized the implementation provided by [2] to conduct our experiments. RLPD uses symmetric sampling to handle offline data, which means the algorithm samples 50% of the data from the replay buffer and the remaining 50% from the offline data buffer. The offline data buffer was collected by expert demonstration with a teleoperation device.

- *Design of action space:* the action is defined as $a \in \mathbb{R}^3$, representing the displacement along the x, y, and z axes in the Cartesian coordinate system.
- *Design of Observation Space:* Our approach integrates multiple sensor modalities to capture the robot's state and task-related information. The observation o_t consists of two images I_1, I_2, the displacement D, the three-dimensional force F at the robot end-effector, the angle θ between the connector and the horizontal direction, representing the degree of cable bending, represented as $o_t = (I_1, I_2, D, F, \theta)$.
- *Visual-based Reward Function:* The reward consist of 3 parts. Firstly, a vision-based reward classifier (VBRC) is used to distinguish between two states: the connector not being in the socket and the connector being either partially or fully inserted into the socket. The VBRC takes images from two viewpoints as input and uses a convolutional neural network (CNN) to produce a value between 0 and 1, indicating whether the connector is partially inserted. Secondly, the reward function is defined as a function of the connector's angle, with higher rewards granted for increasing the angle, thereby encouraging the robot to straighten the cable. Finally, task completion is indicated when the socket receives power from the connector.

C. PD Controller

Sufficient force must be exerted on the connector to overcome the interference fit between the connector and the socket while protecting the cable from excessive external forces that could cause damage. Then, the actions a_t output by the RL agent is processed through a low-level proportional-derivative (PD) controller before being executed by the robot. Unlike traditional methods, the objective is not to keep a constant force contact between the connector and the socket, but to control the robot to move along the socket surface in a bouncing manner.

IV. EXPERIMENT

A. Experiment Setup

As shown in Fig. 3, the robot environment setup includes a UR5 robot, a Robotiq FT-300 sensor mounted on the robot end-effector, and a Robotiq 2f-85 gripper to hold the cable. For visual perception, Intel RealSense D405 and Imaging Source DFK 33GX264 are used as eye-in-hand and eye-to-hand cameras, respectively. For data collection, 3D Systems TouchX is used as the input interface for human teleoperation of the robot (Fig. 3(c)). Both data collection and model training are conducted on a laptop equipped with an Intel Core i9-11950H CPU and an NVIDIA A5000 GPU.

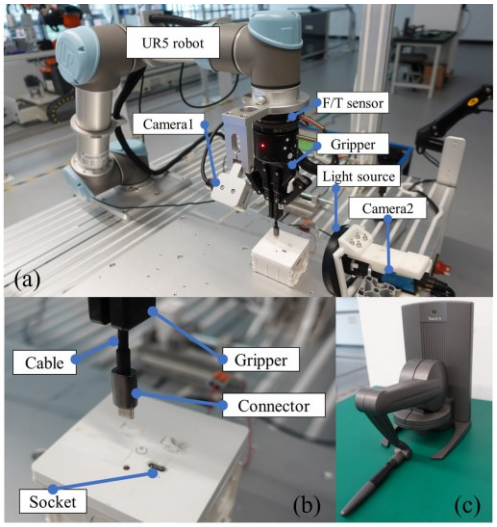

Fig. 3. Overview of the experiment platform.

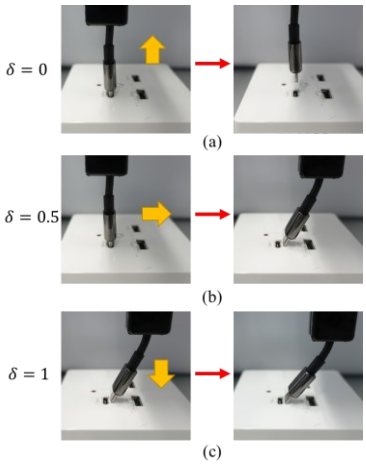

Fig. 4. Reset methods under different curriculum difficulty levels.

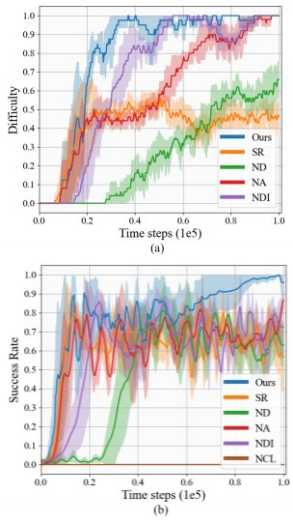

Fig. 5. Training process of the ablation experiments. (a) Difficulty variation curve, (b) Success rate variation curve.

B. Cable Bending Adjustment Method

A three-stage reset method (Fig. 4) automatically adjusts the initial cable bending. In Stage 1, the connector is reset by being vertically pulled from the socket. In Stage 2, before full removal, the robot moves along the y-axis to cause a collision between the connector and the socket, thereby inducing bending; as difficulty increases, greater displacement results in more pronounced bending. In Stage 3, following Stage 2, the robot moves along the z-axis to apply downward pressure on the cable, further deforming it, with the downward displacement increasing with difficulty. Depending on the task, some stages may be omitted—for example, Stage 3 is skipped at lower difficulty levcels. In our setup, the cable is bent nearly 40° from its straight state at the highest difficulty.

C. Result and analysis

We evaluate the aforementioned methods based on test results from three difficulty levels, including success rate (SR) and average time (AT). Each difficulty level was tested 50 times. Tasks that reach the maximum timestep without completion are considered failures. RL-based assembly method[3] and the imitation learning algorithm ACT[4] were used to compared. The results are presented in Table I.

Table I EXPERIMENTAL RESULTS OF DIFFERENT METHODS

Method	Difficulty	SR [%]	AT [s]	Method	Difficulty	SR [%]	AT [s]	Method	Difficulty	SR [%]	AT [s]
Ours	0	**100**	1.7	LBTSM [3]	0	96	1.65	ACT [4]	0	72	3.45
	0.5	**100**	**1.8**		0.5	0	-		0.5	54	3.63
	1	**98**	**3.23**		1	0	-		1	52	5.13

After 3 hours of training, our method outperformed others across all difficulty levels, achieving a 98% success rate even at the highest difficulty. It also demonstrated shorter completion times compared to the other methods.

Fig. 5. shows the training process. At the some time, 5 ablation experiment were designed to show the performance. Including No Curriculum Learning (NCL), No Demonstration (ND), Sparse Reward (SR), No Angle (NA) observation, No Displacement (NDI) observation. Our method is the only one that has achieved the highest difficulty level and a high success rate. Although the highest difficulty is reached after 30,000 steps, a high success rate is observed only around 80,000 steps. Despite implementing a mechanism to slow down difficulty updates, the results indicate that the agent remains undertrained at lower difficulty levels before the difficulty increases. Overall, the training efficiency of the other ablation methods is lower than ours.

REFERENCE

[1] P. J. Ball, L. Smith, I. Kostrikov, and S. Levine, "Efficient Online Reinforcement Learning with Offline Data," May 31, 2023, *arXiv*: arXiv:2302.02948. Accessed: Feb. 26, 2024. [Online]. Available: http://arxiv.org/abs/2302.02948

[2] J. Luo *et al.*, "SERL: A Software Suite for Sample-Efficient Robotic Reinforcement Learning," Feb. 12, 2024, *arXiv*: arXiv:2401.16013. Accessed: Feb. 25, 2024. [Online]. Available: http://arxiv.org/abs/2401.16013

[3] J. Zhao, Z. Wang, L. Zhao, and H. Liu, "A Learning-Based Two-Stage Method for Submillimeter Insertion Tasks With Only Visual Inputs," *IEEE Trans. Ind. Electron.*, vol. 71, no. 7, pp. 7381–7390, Jul. 2024, doi: 10.1109/TIE.2023.3299051.

[4] T. Z. Zhao, V. Kumar, S. Levine, and C. Finn, "Learning Fine-Grained Bimanual Manipulation with Low-Cost Hardware," Apr. 23, 2023, *arXiv*: arXiv:2304.13705. doi: 10.48550/arXiv.2304.13705.