

Learning for Deformable Linear Object Insertion Leveraging Flexibility Estimation from Visual Cues

Mingen Li¹ and Changhyun Choi¹

Abstract—Manipulation of deformable Linear objects (DLOs), including iron wire, rubber, silk, and nylon rope, is ubiquitous in daily life. These objects exhibit diverse physical properties, such as Young’s modulus and bending stiffness. Such diversity poses challenges for developing generalized manipulation policies. However, previous research limited their scope to single-material DLOs and engaged in time-consuming data collection for the state estimation. In this paper, we propose a two-stage manipulation approach consisting of a material property (e.g., flexibility) estimation and policy learning for DLO insertion with reinforcement learning. Firstly, we design a flexibility estimation scheme that characterizes the properties of different types of DLOs. The ground truth flexibility data is collected in simulation to train our flexibility estimation module. During the manipulation, the robot interacts with the DLOs to estimate flexibility by analyzing their visual configurations. Secondly, we train a policy conditioned on the estimated flexibility to perform challenging DLO insertion tasks. Our pipeline trained with diverse insertion scenarios achieves an 85.6% success rate in simulation and 66.67% in real robot experiments. Please refer to our project page: <https://1meee.github.io/DLOInsert/>

I. INTRODUCTION

Interacting with highly deformable objects is challenging due to the infinite number of degrees of freedom. Moreover, their physical properties vary significantly among different kinds of ropes. Among manipulation tasks with deformable linear objects (DLOs), DLO insertion into a hole has extensive applications in healthcare, manufacturing, and households. Imagine suturing a surgical incision with a deformable thread, routing cables inside an automobile frame, or inserting cables into a desk grommet. These tasks involve DLOs with vastly different physical properties, requiring different insertion strategies. For instance, the grasping point of a stiff iron wire differs significantly from that of a flexible nylon rope, where the latter requires a closer grip near the rope head for successful insertion (Fig. 1). For successful insertion regardless of the DLO types, we have to address the following questions: (1) How to estimate the material properties (e.g., flexibility) from visual observations?, (2) How to select the pick position for DLOs with different flexibility?, and (3) How should the agent generate a feasible manipulation trajectory to insert the DLO to a target? Previous research [1, 2] often fails to account for the diversity of DLO types. Lv *et al.*[1] rely on simplified rigid-body models that may yield inaccurate properties like elasticity

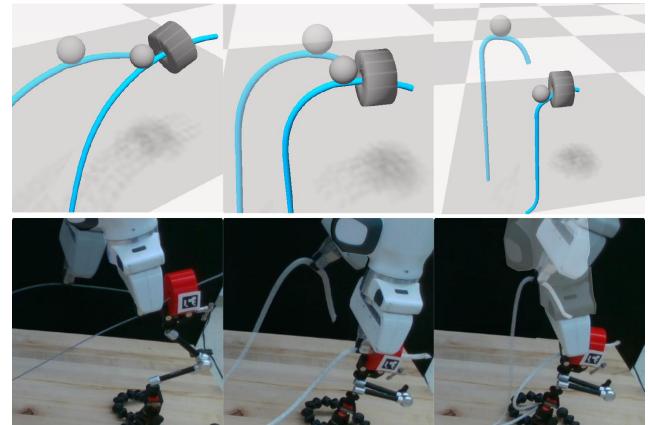


Fig. 1: Flexibility-aware DLO insertion. The real DLO insertion demonstration (lower) and corresponding simulation motion (upper) with estimated flexibility of the real DLOs.

and compliance. Additionally, these studies do not consider the various target hole configurations, such as different orientations and radii of holes, commonly encountered in real-world insertion tasks commonly encountered in real-world insertion tasks. Moreover, a heuristic design that the agent always grasps the tail-end, restricts the system from finding an optimal solution. To address these limitations, we propose a two-stage manipulation approach: the material property estimation of DLOs and an insertion policy learning conditioned on the estimated material property. In the initial phase, we collect ground truth data regarding the DLO’s state and flexibility via a single robotic interaction in a simulation and leverage a Graph Neural Network (GNN) to capture the underlying physics properties. We leverage SoftGym [3] as a simulator platform that uses position-based dynamics (PBD) as the backbone and offers fast and accurate modeling of dynamics of DLO and motion. Our flexibility estimation model is trained with ground truth data in the simulation. The estimated flexibility serves as a vital observation for reinforcement learning. We employ a motion primitive for grasping the object and inserting it into a target, and the parameters of the motion primitive are optimized through a RL. In real-world experiments, a robotic system initiates by grasping a DLO and takes a single interaction step (grasping) to estimate its flexibility. It then executes the insertion task guided by the learned policy as shown in Fig. 1.

The primary contributions of this paper are as follows:

- We introduce a one-step flexibility estimation model for DLOs.
- We propose a reinforcement learning framework for

¹The authors are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 {li002852, cchoi}@umn.edu

[†]This work was supported in part by NSF Award 2143730 and MnDRIVE Initiative on Robotics, Sensors, and Advanced Manufacturing.

DLO grasping and insertion with the estimated flexibility.

- Our pipeline demonstrates its effectiveness in handling diverse DLOs, including iron wire, rubber, and nylon ropes in the real world without any further training for the RL policy.

II. RELATED WORK

A. DLO Manipulation

DLO manipulation has been extensively studied with applications in robot-assisted surgery [4, 5], knot tying [6, 7], and cable handling [8]–[10]. Some works employ supervised learning for state estimation using visual observation to address challenges like occlusion [11], or dynamic DLO manipulation [12]. Several works employ goal-conditioned techniques for rope configuration [13, 14], or knot-tying [6]. [15] utilizes a generative model to generate a sequence of transient images from goal observation. However, relying on predefined goals can restrict the system’s adaptability and hinder its performance in dynamic scenarios. Another approach [16] is to employ demonstration-based techniques to address DLO manipulation challenges, but this approach requires time-consuming data collection and struggles to generalize to diverse DLOs. Unlike prior works, our approach does not rely on demonstration data or prior knowledge of DLO properties, making it adaptable to a wide range of DLOs.

The DLO-in-Hole task has been studied for both needle threading [1, 17, 18] and assembly task [19]. [4] leverages tactile information for thread-tail finding and eyelet insertion. While these works rely on a limited range of material properties and typically assume tail-end grasping, our work can effectively estimate flexibility for various DLOs and approach optimal insertion.

B. Reinforcement Learning for Deformable Manipulation

Reinforcement learning (RL) is frequently used in robotic manipulation for its adaptability and performance improvement in uncertain environments. As deformable simulators become more accessible [3, 20, 21], researchers have introduced specialized frameworks for RL, such as AMBF-RL[22], SurRoL[23], dVRL[24], and SoftGym[3]. Several studies have employed model-based RL to perform deformable manipulation tasks, including needle-threading [1], goal-conditioned deformable manipulation with video prediction [25], and 3D shape control of DLOs [26]. Lin *et al.* [27] acquire DLO manipulation skills from RGB images using goal-reaching RL. Singh *et al.* [28] employ hierarchical RL for DLO manipulation through attainable sub-goals. Yu *et al.* [29] use offline learning for the initial estimation of the DLO trajectory, followed by online updates to compensate for errors. While these approaches obtain favorable results in deformable object manipulation, their generalization across diverse objects and scenarios requires further study. Our methods can directly transfer to real experiments with diverse environmental settings and a wide range of DLO types without further training or fine-tuning.

III. METHOD

A. Problem Formulation

In our task setup, the robot needs to insert a DLO into a target ring. We randomize ring properties (e.g. radius, position, and orientation) to improve the generalization of our model. In addition, we employ a deep ring hole, prompting the robot agent to consider a proper grasping point before insertion. To account for DLO variation, we introduce flexibility f , a material property closely associated with DLO manipulation. Flexibility is used to characterize a wide range of daily-life DLOs, as discussed in Section III-B. Furthermore, we elaborate on the flexibility estimation in Section III-C.

We formulate the DLO insertion task problem as a Markov Decision Process (MDP) $(\mathbf{S}, \mathbf{A}, r, \gamma)$. Other than properties of target rings, the state \mathbf{S} consists of positions $\mathbf{p}_{1:n}$ of a DLO modeled with n particles. As for action space \mathbf{A} , we use a gripper with 3D translation motion and 1D rotation around one axis. We leverage a motion primitive for our grasp point selection and insertion, which generates long-horizon trajectories for the gripper as discussed in Section III-D.

B. Defining Flexibility

To characterize various types of DLOs, we seek a property that encapsulates task-related information. Such a property should satisfy the following requirements: 1) The property should be directly relevant to the insertion task, helping agents acquire valuable knowledge to improve training performance. 2) It can be obtained or represented by visual observation.

Before making contact with the ring, we noticed that the DLO could be treated as quasi-static under the influence of gravity. This pre-insertion phase is crucial as insertion success depends on the relationship between the orientation of the target ring and the DLO’s state $\mathbf{p}_{i_p:n}$ where i_p is the index of the picked particle. This insight led to developing a property linked to the DLO’s state under gravity, potentially enhancing policy training during insertion.

To simplify the property estimation, we assume the DLO always lies in the same plane perpendicular to the ground due to gravity. This plane has a normal vector $\hat{\mathbf{n}} = [x_n, y_n, 0]^T$, while the z-axis of the gripper coordinate system is $\hat{\mathbf{z}} = [0, 0, -1]^T$. We use Eq. (1) to convert DLO particles \mathbf{p}_i from world coordinates to the gripper coordinate system within the plane:

$$\begin{pmatrix} x_i^{proj} \\ y_i^{proj} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} (\hat{\mathbf{n}} \times \hat{\mathbf{z}})^T \\ \hat{\mathbf{n}}^T \\ \hat{\mathbf{z}}^T \end{pmatrix} \mathbf{p}_i. \quad (1)$$

Building on these considerations and the concept of curvature [30], we define flexibility f as shown below:

$$f = -\frac{\frac{\Delta_y}{\Delta_x}(i_p + 3, i_p + 1) - \frac{\Delta_y}{\Delta_x}(i_p + 1, i_p)}{(x_{i_p+1}^{proj} - x_{i_p}^{proj})(1 + \frac{\Delta_y}{\Delta_x}(i_p + 1, i_p)^2)^{1.5}} \quad (2)$$

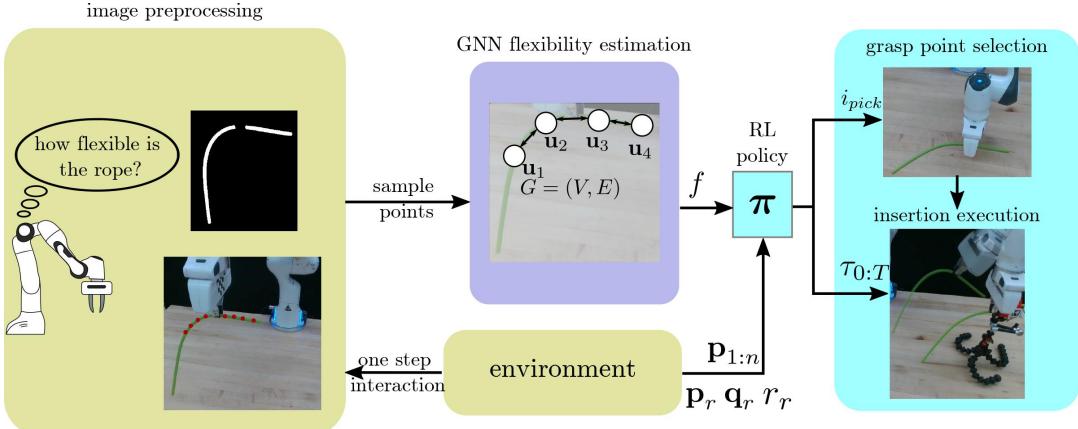


Fig. 2: **Pipeline for flexibility-aware DLO insertion in real experiment.** We start by asking the robot to grasp the testing DLO using a predefined pose and sample particles on the DLO with image prepossessing techniques. The DLO states are input to the GNN-based estimation module to predict flexibility. Taking flexibility, DLO states and ring configuration as input, the policy can output a predicted grasping point and a control trajectory for insertion.

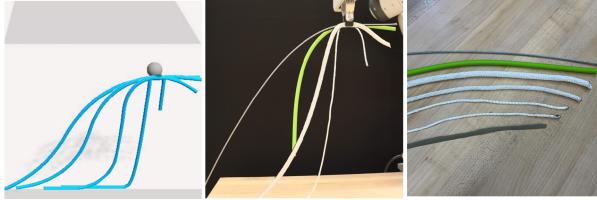


Fig. 3: **Predefined grasping poses for DLO flexibility estimation** in simulation (left) and real experiment (middle). The right image shows the DLOs used for the real experiment.

where $\frac{\Delta_y}{\Delta_x}(a, b) = \frac{y_a^{proj} - y_b^{proj}}{x_a^{proj} - x_b^{proj}}$. Unlike curvature, flexibility can take negative values, indicating a highly stiff DLO.

As shown in Fig. 3, we have observed that, when consistently applying the same grasping pose to the DLO, it consistently and uniformly maintains a static bending configuration. This phenomenon remains consistent whether the DLO is rigid or exceedingly flexible and deformable. To capture such subtle deformations at specific points along the DLO, we employ the concept of curvature. Although curvature represents sharpness at a singular focal point, it proves adequate in our ability to devise a consistent manipulation approach for characterizing the bending configuration of the entire length of the DLO. The discrete version of the curvature calculation is shown in Eq. (2).

C. GNN-based Flexibility Estimation

After defining flexibility, we use a Graph Neural Network (GNN)[31] to estimate the DLO flexibility. A graph $G = (V, E)$ is constructed naturally by using sampled particles as vertices V and directly connecting adjacent particles as edges E , forming a representation for DLO. Specifically, each particle becomes a vertex: $v_i = [\mathbf{p}_i]$, and adjacent vertices are connected with undirected edges: $E = \{(v_i, v_j) | |i - j| = 1\}$. We employ a four-layer message-passing network and a linear layer to predict flexibility values. We augment the collected simulation flexibility data to enhance data diversity and improve performance in noisy real-world scenarios. This involves introducing a scaling factor $\lambda \in [0.85, 1.0]$ and

adding translational Gaussian noise $\mathbf{n}_{1:n} \sim \mathcal{N}(0, 0.003)$ to $\mathbf{p}_{1:n}$, resulting in modified positions $\mathbf{p}'_{1:n} = \lambda(\mathbf{p}_{1:n} + \mathbf{n}_{1:n})$ [32].

D. Reinforcement Learning with Motion Primitive

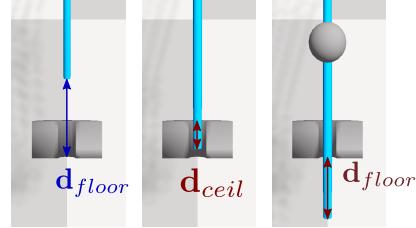


Fig. 4: **Distance reward visualization.** d_{floor} contributes negatively to reward when the DLO is not inserted. After insertion, a positive distance reward is given when the DLO is halfway or completely through.

This subsection describes details of the motion primitive and an RL method for insertion conditioned on the previously acquired flexibility estimation. **Task Definition:** Insertion involves the robot inserting different types of DLO through a ring until the DLO tip emerges on the other side, as depicted in Fig. 4. **Observation Space:** The observation space contains the position of DLO particles, $\mathbf{p}_{1:n}$, ring's position \mathbf{p}_r , angle \mathbf{q}_r , and radius r_r . The ground truth flexibility value of the DLO f is optionally provided depending on the baseline setting. **Motion Primitives and Action Space:** In the SoftGym environment, we use a gripper with position \mathbf{p}_p^t and orientation \mathbf{q}_p^t to enable the interaction with the DLO. To perform long-horizon planning, we leverage a motion primitive that takes grasped particle index i_p , starting pose $(\mathbf{p}_p^0, \mathbf{q}_p^0)$ and ending pose $(\mathbf{p}_p^T, \mathbf{q}_p^T)$ as input in Fig. 5. The output trajectory includes DLO grasping and insertion motion. A customized PD controller is utilized to generate the control trajectory $\tau_{1:T}$ for the insertion. The insertion is conducted in a planar plane formed by the DLO and the target ring. The agent observes the environment only at the beginning of the episode, and the insertion is executed in

an open-loop manner. **Policy Training:** We use Soft Actor-Critic (SAC) [33] for training with MLP for both actor and critic networks. We train with three random seeds using the default parameters from in `stable-baselines3`. **Reward Function:** As shown in Fig. 4, we define the reward function for three distinct stages: the initial stage (left), the halfway point (middle), and the successful passage (right) of the DLO through the ring. These stages are delineated by binary values of the `rope_in` and the `rope_out` variables. Eq. (3) incorporates both a stage-based reward component and a penalty term r_{pen} to account for abnormal stretching of the elastic DLOs during simulation:

$$r = 0.5(rope_in + rope_out) + r_{pen} + r_{dist}$$

$$r_{dist} = \begin{cases} -10 \cdot d_{floor} & \text{if not } rope_in \\ 5 \cdot d_{ceil} & \text{if } rope_in \text{ and not } rope_out \\ 10 \cdot d_{floor} & \text{if } rope_out \end{cases} \quad (3)$$

If the DLO segments are stretched to 120% of the original length during simulation, it is considered abnormal and the penalty $r_{pen} = -2$ is applied.

IV. EXPERIMENT AND RESULT

In this section, we detail an experiment about flexibility estimation and how flexibility can affect RL policy training and insertion task competence.

A. Experiment Detail

1) *Flexibility Estimation:* In this part, we illustrate our flexibility estimation functions and baselines.

Data Collection: To simulate DLOs with extreme flexibility value, we adopt the DLO configuration in SoftGym [3] built upon the PyFlex [34]. Specifically, we modify DLO stiffness by adding extra springs, reducing the constraint solver iterations, and decreasing bending stiffness. To obtain ground truth flexibility data, we grasped the middle of DLOs (particle number $n = 20$) and recorded their stationary state positions using Eq. (2). We collected a consistent dataset with linearly sampled DLO flexibilities.

Data Augmentation: With our labeled flexibility DLO data, we aim to capture more realistic DLO interactions to estimate flexibility. To address variations in DLO length, we introduce an unbalanced state by using a longer DLO ($n = 40$) and biasing the grasping point to $i_p = 10$, as depicted in Fig. 3 (left). We generate a dataset for training the flexibility estimation module by augmenting part of the state $\mathbf{p}_{21:40}$ with noise and pairing it with flexibility values.

Evaluation: To evaluate flexibility estimation with actual DLOs, we replicate the same DLO interaction demonstrated in Fig. 3 (middle). We create a DLO mask using SAM [35], unproject the center-line onto a predefined gripper coordinate plane to sample DLO particles $\hat{\mathbf{p}}_{21:40}$. We sample particles with a distance of 1.2cm, corresponding to the rest length of the DLO segment in SoftGym. Our evaluation dataset comprises 31 frames from 7 different DLOs. The simulated DLO with the closest estimated flexibility $\mathbf{p}_{21:40}$ is used

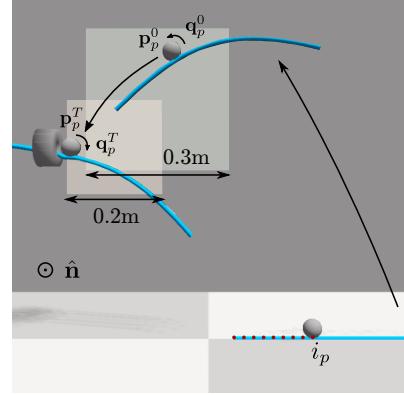


Fig. 5: **Action space visualization.** The upper right square shows the range of \mathbf{p}_p^0 and the lower left square shows the range of \mathbf{p}_p^T . The rotation axis $\hat{\mathbf{n}}$ is pointing outward.

for evaluation. Performance is assessed using the point-point distance described below:

$$d_{pp} = \frac{1}{20} \sum_{i=21}^{40} |\mathbf{p}_i - \hat{\mathbf{p}}_i|. \quad (4)$$

Baselines: We evaluate flexibility estimation using various baselines:

- **Analytic:** The analytical method computes the theoretical flexibility using Eq. (2) and the state extracted from real images.
- **Multi-Layer Perceptron (MLP):** MLP has two linear layers with a hidden dimension of 8.
- **GNN without data augmentation:** Since GNN yields the best performance among all baselines, we test GNN without data augmentation to demonstrate its effectiveness of it.

2) *Simulation Setup:* In both simulation and real experiments, we assume insertion takes place within the two-dimensional plane, resulting in a 7-dimensional action space $(i_p, \mathbf{p}_p^0, \mathbf{p}_p^T \in \mathbb{R}^2, \mathbf{q}_p^0, \mathbf{q}_p^T)$ as shown in Fig. 5. In practice, we implement $\mathbf{q}_p^0, \mathbf{q}_p^T$ as rotation with respect to axis $\hat{\mathbf{n}}$. Since the agent knows the ring's position and orientation, the action space is defined as a rectangle relative to the ring's local coordinates to ensure fast and informed exploration during RL.

Randomization: We randomize the position of the ring within a 10cm square and the initial DLO position within 10cm inside the plane. We also vary the ring radii from challenging small (1cm) to easier large (2.5cm) sizes, and the ring angle θ within the range of $[0, \frac{3}{4}\pi]$ rad, covering positions from facing upward to a 45 deg downward angle.

Baselines: We conduct various baseline experiments for our simulation:

- **Fixed angle (fix θ) and ring angle variation (rand θ):** Fixed angle variation provides the RL policy with one fixed challenging angle environment ($\theta = \pi/2$ rad), while ring angle variation (θ) provide the target ring angle change within a limit ($\theta \in [0, 3\pi/4]$ rad).
- **Flexibility observation (Ours w/wo f):** Whether providing estimated flexibility as an observation or not.

Method	Evaluation	Ours w/ f	Ours w/o f	VB w/ f	VB w/o f	Random f
rand θ	Success Rate (%) \uparrow	78 \pm 2	56 \pm 1	40	32	21 \pm 4
	Avg. dis. (cm) \uparrow	3.92 \pm 0.51	0.00 \pm 0.16	-0.98	-0.12	-5.35 \pm 0.48
fix θ	Success Rate (%) \uparrow	71 \pm 1	39 \pm 5	31	38	10 \pm 3.0
	Avg. dis. (cm) \uparrow	2.00 \pm 0.46	-2.72 \pm 0.79	-1.56	-1.42	-7.00 \pm 0.44

TABLE I: **Evaluation of all baselines in simulation.** For baselines trained with ring angle variation, random angle values are assigned for environments with different flexibility to perform evaluation. The ring angle is always $\pi/2$ rad for the fixed θ . The ring radius used for evaluation is 1cm.

- Random flexibility (Random f): Training the agent with flexibility but introducing a random flexibility value during evaluation.
- Visual baseline without flexibility (VB w/o f): A heuristic method where the DLO is grasped at point $i_p = 5$ and held horizontally. We calculate the angle formed by the ring, gripping point, and DLO tip and subsequently perform a rotation using this angle, assuming that it will align the DLO in parallel with the ring hole.
- Visual baseline with flexibility (VB w/ f): Similar to the previous baseline, but introducing flexibility to grasp particle i_p using a square root function.

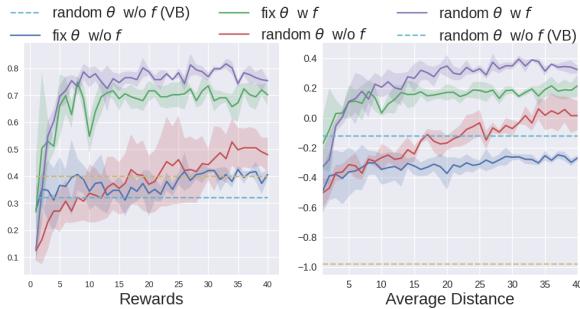


Fig. 6: **The training curves (left) and the average distance (right).** Average distance (avg. dis.) is the average signed endpoint distance. Similar to the reward definition, it equals to d_{floor} when $rope_out$ and $-d_{floor}$ otherwise.

Training Details: We have 4 sets of training settings (fix or random θ , Ours w/wo f). We train each of them with 3 random seeds, totaling 40,000 episodes. We only have a 40k state-action pair for dynamics primitive and the policy training is converged (Fig. 6). In our evaluation, successful insertion is defined as the DLO penetrating the ring’s floor (Fig. 4). We evaluate the performance by calculating the endpoint distance based on d_{floor} . Prior to success, the distance is assigned as negative ($-d_{floor}$), while a high endpoint distance indicates strong insertion performance.

B. Simulation Result Analysis

In this part, we discuss how our flexibility estimation and motion primitive affects training performance.

Accuracy of Flexibility Estimation: The baselines’ evaluation results are summarized in Table II. Directly calculating flexibility in Eq. (2) with local DLO state only is not accurate due to noise in real experiments. Also, it fails to consider the complete DLO states which offer more informative insight into flexibility. Table II highlights GNN’s superior

model	d_{pp} (mm)
Analytic	11.80 \pm 13.52
MLP	9.69 \pm 12.01
GNN w/o aug	9.76 \pm 12.7
GNN (Ours)	9.16 \pm 11.6

TABLE II: **Evaluation result for using different backbone or the analytical method for flexibility estimation model.** GNN achieves the smallest point-point distance and the smallest standard deviation.

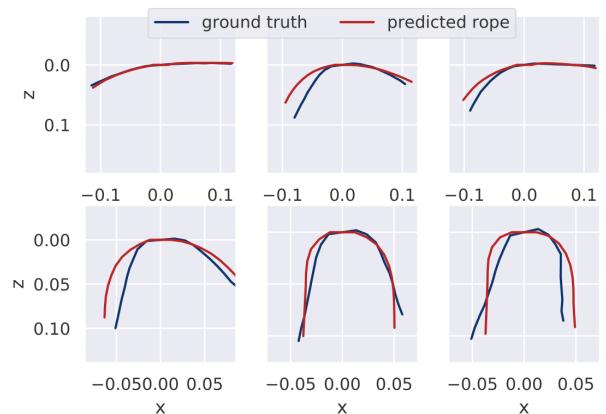


Fig. 7: **Visualization of real DLO flexibility estimation.** We estimate flexibility for the real DLOs and compare with the simulated DLOs which have the closest flexibility.

performance with the smallest average distance and standard deviation, showing the strength of the graph-based neural network. In Fig. 7, we compare the shapes of the estimated and real DLOs in the gripper’s plane.

Necessity of Flexibility Estimation: To explore the necessity of flexibility estimation, we compared RL training with and without flexibility in Table I. Compared with training without flexibility, Our w/ f obtained a 22% and 32% higher success rate on average for fix θ and rand θ , respectively. The substantial gap underscores that the policy without flexibility is underperforming to the diverse DLOs in our environment. Furthermore, our approach achieves a significantly larger average distance, suggesting that it is closer to obtaining the optimal solution compared to cases without flexibility. Random f shows the worst performance among all the methods because of significant divergence in optimal grasping points and insertion motion primitives for DLOs with various flexibility. These divergences prevent the achievement of uniformly superior outcomes in various flexibility situations. Consequently, without proper consideration of flexibility, achieving the optimal insertion actions for different DLO

Methods	Ring angle (deg)	≈66		≈90		≈110	
		Success (%)↑	Avg. Dis. (cm)↑	Success (%)↑	Avg. Dis. (cm)↑	Success (%)↑	Avg. Dis. (cm)↑
VB w/ f	rand θ	29	8.10	29	3.60	0	0.00
Ours w/ f	fix θ	—	—	57	2.98	—	—
Ours w/ f	rand θ	86	2.21	86	2.04	29	1.71

TABLE III: **Real experiment results for insertion.** The ring angle above is an approximated value from image perception. Perception noise and potential robot collision may cause ring angle deviation during the experiment.

types is unachievable.

Comparing Against the Visual Baselines: As shown in Table I, our method with flexibility (Ours w/ f) outperforms the visual baselines (VB w/ f) with a large margin. This is because our mapping between flexibility and grasp point selection is not as trivial as a simple function, but may involve higher-order terms. Moreover, the conventional approach of parallel insertion is sub-optimal due to the deformable nature of the objects and fails once DLOs collide with the ring wall. DLOs, unlike rigid objects with well-defined transformations, tend to deform during motion, contradicting the initial assumption of parallel alignment. Additionally, parallel insertion often leads to head-on collisions with the thick ring wall, rendering contact-based insertion impossible. In contrast, our method tilts the DLO slightly, as illustrated in Fig. 8, increasing the likelihood of a successful insertion.

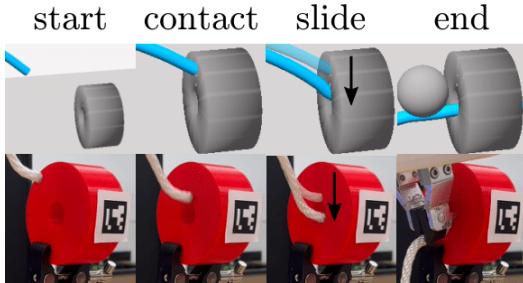


Fig. 8: Inserting by sliding in simulation (upper) and real experiment (lower). The DLO tip first contacts with the upper ring wall and slides into the ring.

Comparing Performance Under Different Target Ring Angles We evaluate how the target ring angle influences the insertion task using a different ring angle range as shown in Fig 9, the task success rate drops as the ring angle increases, indicating that the task difficulty is dependent on the ring angle. Nevertheless, our method achieves the highest level of success among all tested scenarios, demonstrating its proficiency in addressing the manipulation challenges posed by varying ring angles.

C. Real Robot Evaluation

We also conducted real experiments to verify the capability of our model in the real world.

Experiment Setup: We employed an Intel Real Sense camera to capture RGB images of the DLO and determined the orientation and position of the ring using Aruco markers[36]. Our manipulation was carried out using a Franka Emika robot with custom fingertips for DLO handling. For DLOs with limited depth information due to noise in the depth channel, we conducted flexibility estimation interactions at the plane positioned at $x = 0.35m$ and unprojected this data

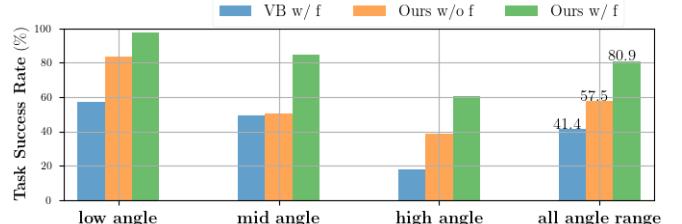


Fig. 9: Evaluation of baselines with different ranges of target ring angles in simulation. Low-angle $\theta \in [0, \pi/4]$ rad, mid-angle $\theta \in [\pi/4, \pi/2]$ rad, high-angle $\theta \in [\pi/2, 3\pi/4]$ rad.

using the camera transformation matrix. The insertion motion occurred at the plane corresponding to the ring’s position. We examined our method with 7 different testing DLOs in real experiment evaluation, including wire rope, rubber rope, thick nylon, median nylon, thin nylon, soaked thin nylon, and silk rope. We conducted 14 experiments for fix θ and 21 experiments for rand θ . The target ring had the same configuration as the one in the simulation, having an inner radius of 1cm, ring depth of 4cm, and outer radius of 4. We tested with three different ring angles.

Result Analysis: Table III summarizes the real experiment performance, and visualization of real DLO insertion is shown in Fig. 1. We observe a sharp success rate drop as the target ring angle increases from 90 deg to 110 deg, which is consistent with our findings in the simulation. Our model trained with random θ achieves the best result and outperforms that trained with fix θ where the ring angle is 90 deg. While our method trained in a ring angle-variant environment adapts to different angles (66 and 110 deg), our method trained with fix θ overfits to a single ring environment and hence is more sensitive to ring angle deviation. The VB has the worst success rate, characterized by frequent head-on collisions. The success heavily depends on the rigid characteristics of the object. The wire rope plays a pivotal role for VB and consistently achieves the longest average distance compared to other methods. Furthermore, VB is also significantly influenced by ring angle detection.

V. CONCLUSION

We presented a general robotic insertion task encompassing diverse DLOs and ring configurations. To accomplish this, we defined flexibility as a property that can be visually assessed and that guides the insertion task. We found that our flexibility-aware policy outperforms those without flexibility and could find a reliable trajectories than the considered baselines. Additionally, our policy exhibited good sim2real performance for various real DLO types.

REFERENCES

[1] J. Lv, Y. Feng, C. Zhang, S. Zhao, L. Shao, and C. Lu, "Sam-rl: Sensing-aware model-based reinforcement learning via differentiable physics-based simulation and rendering," 2023.

[2] P. M. Scheikl, B. Gynes, R. Younis, C. Haas, G. Neumann, M. Wagner, and F. Mathis-Ullrich, "Lapgym - an open source framework for reinforcement learning in robot-assisted laparoscopic surgery," *arXiv preprint arXiv:2302.09606*, 2023.

[3] X. Lin, Y. Wang, J. Olkin, and D. Held, "Softgym: Benchmarking deep reinforcement learning for deformable object manipulation," in *Conference on Robot Learning*, 2020.

[4] Z. Yu, W. Xu, J. Ren, T. Tang, Y. Li, S. Yao, G. Gu, and C. Lu, "Precise robotic needle-threading with tactile perception and reinforcement learning," in *7th Annual Conference on Robot Learning*, 2023. [Online]. Available: <https://openreview.net/forum?id=B7PnAw4ze0l>

[5] J. Grannen, P. Sundaresan, B. Thananjeyan, J. Ichnowski, A. Balakrishna, M. Hwang, V. Viswanath, M. Laskey, J. E. Gonzalez, and K. Goldberg, "Untangling dense knots by learning task-relevant keypoints," 2020.

[6] P. Sundaresan, J. Grannen, B. Thananjeyan, A. Balakrishna, M. Laskey, K. Stone, J. E. Gonzalez, and K. Goldberg, "Learning rope manipulation policies using dense object descriptors trained on synthetic depth data," 2020.

[7] V. Viswanath, K. Shivakumar, M. Parulekar, J. Ajmera, J. Kerr, J. Ichnowski, R. Cheng, T. Kollar, and K. Goldberg, "HANDLOOM: Learned tracing of one-dimensional objects for inspection and manipulation," in *7th Annual Conference on Robot Learning*, 2023. [Online]. Available: <https://openreview.net/forum?id=WWiKBdepNd>

[8] Y. She, S. Wang, S. Dong, N. Sunil, A. Rodriguez, and E. Adelson, "Cable manipulation with a tactile-reactive gripper," 2020.

[9] Y. Yang, H. Yu, X. Lou, Y. Liu, and C. Choi, "Attribute-based robotic grasping with data-efficient adaptation," *IEEE Transactions on Robotics*, vol. 40, pp. 1566–1579, 2024.

[10] A. Wilson, H. Jiang, W. Lian, and W. Yuan, "Cable routing and assembly using tactile-driven motion primitives," 2023.

[11] K. Lv, M. Yu, Y. Pu, X. Jiang, G. Huang, and X. Li, "Learning to estimate 3-d states of deformable linear objects from single-frame occluded point clouds," 2023.

[12] H. Zhang, J. Ichnowski, D. Seita, J. Wang, H. Huang, and K. Goldberg, "Robots of the lost arc: Self-supervised learning to dynamically manipulate fixed-endpoint cables," 2021.

[13] D. Seita, P. Florence, J. Tompson, E. Coumans, V. Sindhvani, K. Goldberg, and A. Zeng, "Learning to rearrange deformable cables, fabrics, and bags with goal-conditioned transporter networks," 2023.

[14] M. Yan, Y. Zhu, N. Jin, and J. Bohg, "Self-supervised learning of state estimation for manipulating deformable linear objects," 2020.

[15] A. Wang, T. Kurutach, K. Liu, P. Abbeel, and A. Tamar, "Learning robotic manipulation through visual planning and acting," 2019.

[16] S. Li, Z. Huang, T. Chen, T. Du, H. Su, J. B. Tenenbaum, and C. Gan, "Dexdeform: Dexterous deformable object manipulation with human demonstrations and differentiable physics," 2023.

[17] D. De Gregorio, R. Zanella, G. Palli, S. Pirozzi, and C. Melchiorri, "Integration of robotic vision and tactile sensing for wire-terminal insertion tasks," *IEEE Transactions on Automation Science and Engineering*, vol. 16, no. 2, pp. 585–598, 2019.

[18] W. Wang, D. Berenson, and D. Balkcom, "An online method for tight-tolerance insertion tasks for string and rope," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 2488–2495.

[19] K. Galassi, A. Caporali, and G. Palli, "Cable detection and manipulation for dlo-in-hole assembly tasks," in *2022 IEEE 5th International Conference on Industrial Cyber-Physical Systems (ICPS)*, 2022, pp. 01–06.

[20] C. Gan, J. Schwartz, S. Alter, D. Mrowca, M. Schrimpf, J. Traer, J. D. Freitas, J. Kubilius, A. Bhandwaldar, N. Haber, M. Sano, K. Kim, E. Wang, M. Lingelbach, A. Curtis, K. Feiglisis, D. M. Bear, D. Gutfreund, D. Cox, A. Torralba, J. J. DiCarlo, J. B. Tenenbaum, J. H. McDermott, and D. L. K. Yaminis, "Threedworld: A platform for interactive multi-modal physical simulation," 2021.

[21] M. Li, Z. Ferguson, T. Schneider, T. Langlois, D. Zorin, D. Panozzo, C. Jiang, and D. M. Kaufman, "Incremental potential contact: Intersection- and inversion-free large deformation dynamics," *ACM Trans. Graph. (SIGGRAPH)*, vol. 39, no. 4, 2020.

[22] V. M. Varier, D. K. Rajamani, F. Tavakkolmoghaddam, A. Munawar, and G. S. Fischer, "Ambf-rl: A real-time simulation based reinforcement learning toolkit for medical robotics," in *2022 International Symposium on Medical Robotics (ISMР)*, 2022, pp. 1–8.

[23] J. Xu, B. Li, B. Lu, Y.-H. Liu, Q. Dou, and P.-A. Heng, "Surrol: An open-source reinforcement learning centered and dvrk compatible platform for surgical robot learning," 2021. [Online]. Available: <https://arxiv.org/abs/2108.13035>

[24] F. Richter, R. K. Orosco, and M. C. Yip, "Open-sourced reinforcement learning environments for surgical robotics," 2019. [Online]. Available: <https://arxiv.org/abs/1903.02090>

[25] F. Ebert, C. Finn, S. Dasari, A. Xie, A. Lee, and S. Levine, "Visual foresight: Model-based deep reinforcement learning for vision-based robotic control," 2018.

[26] Y. Yang, J. A. Stork, and T. Stoyanov, "Online model learning for shape control of deformable linear objects," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022, pp. 4056–4062.

[27] X. Lin, H. S. Baweja, and D. Held, "Reinforcement learning without ground-truth state," 2019.

[28] U. Singh and V. P. Namboodiri, "Crisp: Curriculum inducing primitive informed subgoal prediction for hierarchical reinforcement learning," 2023.

[29] M. Yu, H. Zhong, and X. Li, "Shape control of deformable linear objects with offline and online learning of local linear deformation models," 2022.

[30] E. H. Gilbert Strang, *Calculus Volume 3*. OpenStax, 2016. [Online]. Available: <https://openstax.org/books/calculus-volume-3/pages/3-3-arc-length-and-curvature>

[31] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *International Conference on Learning Representations*, 2017. [Online]. Available: <https://openreview.net/forum?id=SJU4ayYg1>

[32] Y. Shao, C. C. Loy, and B. Dai, "Transformer with implicit edges for particle-based physics simulation," in *Computer Vision - ECCV 2022 - 17th European Conference*, 2022.

[33] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," *CoRR*, vol. abs/1801.01290, 2018. [Online]. Available: <http://arxiv.org/abs/1801.01290>

[34] M. Macklin, M. Müller, N. Chentanez, and T.-Y. Kim, "Unified particle physics for real-time applications," *ACM Transactions on Graphics (TOG)*, vol. 33, no. 4, pp. 1–12, 2014.

[35] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, P. Dollár, and R. Girshick, "Segment anything," *arXiv:2304.02643*, 2023.

[36] S. Garrido-Jurado, R. Muñoz-Salinas, F. Madrid-Cuevas, and M. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognition*, vol. 47, no. 6, pp. 2280–2292, 2014.