CableSort: Force-Guided Dual-Arm Sorting of Stacked Cables Towards Physics Action Models

Yuxin Fu*
Beijing Jiaotong University
Beijing, China
23221004@bjtu.edu.cn

Haoran Han*
Beijing Institute of Technology
Beijing, China
1120231688@bit.edu.cn

Boxian Yao*
Peking University
Beijing, China
296481922@qq.com

Peiyu Zhu
Peking University
Beijing, China
2501213267@stu.pku.edu.cn

Xuan Han
Peking University
Beijing, China
2501213219@stu.pku.edu.cn

Tianyou Zuo
Peking University
Beijing, China
zuo.you@stu.pku.edu.cn

LE ZHANG
Xi'an University of Architecture and Technology
Shaanxi, China
zhangle@xauat.edu.cn

Zhibo Pang†
Peking University
Beijing, China
zhibo.pang@pku.edu.cn

***Abstract*—Manipulating stacked deformable linear objects (DLOs), such as cables, is challenging due to occlusions, deformations, and contact uncertainty. This paper presents a modular perception-action pipeline for dual-arm stacked cable sorting using synthetic training data. The structure-aware perception module achieves an InstmIoU of 0.93 and converts instance masks and crossing relationships into graspable segments and dual-arm grasp points. An eagle-claw gripper pad improves post-lift lateral move success from 38.0% to 68.0%, while an adaptive bilateral wrist-force controller detects the transition from slack removal to cable tension without relying on fixed stretching distances. In complex simulated clutter, the full system achieves the best tested end-to-end success, suggesting a task-specific step toward Physics Action Models guided by physical semantics and closed-loop force feedback.The code is available at: https://github.com/HAORAN808/dual-arm-cable-manipulation.**

***Index Terms*—Deformable Linear Objects, Modular Perception-Action Pipeline, Force-Guided Control, Dual-Arm Manipulation**

I. Introduction

Deformable linear objects (DLOs), including cables, ropes, wires, and tubes, are common in industrial, domestic, and medical environments [1]. Although geometrically simple, they are difficult to manipulate because their shapes are highly deformable, underactuated, and sensitive to contact, friction, and external forces. Prior work has established important foundations in DLO perception, planning, and control across tasks such as routing, threading, untangling, suturing, and transport.

Existing DLO studies cover task-oriented grippers [2]–[5], instance or structure-aware perception [6]–[14], and force-aware manipulation for modeling, routing, or dual-arm control [15]–[19]. However, most methods focus on isolated cables, perception-centered tracing, structured routing, or specific topological objectives, rather than repeated sorting from a stacked cluttered scene.

Stacked cable sorting presents a different setting, where multiple cables are mutually occluded and must be repeatedly separated, straightened, and placed without disturbing the remaining stack [20]. The robot must therefore convert cluttered visual observations into actionable grasp representations, maintain stable contacts with thin cables, and detect when an initially slack cable becomes taut.

To address these challenges, we develop a modular dual-arm pipeline that integrates structure-aware cable perception, geometry-constrained grasping, and adaptive bilateral wrist-force feedback. Beyond solving a specific cable-sorting task, this pipeline also serves as a preliminary step toward Physics Action Model (PAM). In contrast to action models that primarily map visual-language instructions or learned world states to robot commands, a PAM emphasizes physically grounded action generation: robot actions should be conditioned on interpretable physical semantics, interactive sensing, real-time robot-object-environment state estimation, and feedback mechanisms that respect contact, deformation, topology, and force constraints. In this work, we instantiate a small but concrete component of this paradigm by converting cable topology and visible structure into grasp actions, using gripper geometry to enforce contact stability, and detecting force transitions to terminate straightening under deformation uncertainty.

The main contributions are:

1) A modular pipeline is proposed to address the task of stacked cable sorting, demonstrating improved robustness across randomized simulated stacked-cable scenes.
2) A structure-aware perception method is introduced for cluttered cable scenes, which achieves an InstmIoU of 0.93 for instance segmentation trained on 1000 synthetic images.
3) An adaptive force-jump straightening strategy is developed, achieving a 78.0% straightening success rate in post-grasp-and-lift trials and improving the full end-to-

*These authors contributed equally to this work.

This work is funded by the Beijing Natural Science Foundation International Scientists Project IS26042 and the State Key Laboratory of High-end Heavy-load Robots (Open Fund Project No. HHR2026010110).

†Corresponding author: Zhibo Pang (zhibo.pang@pku.edu.cn).

end task success rate to 48.0% among tested configurations.

II. Method

We propose an iterative perception-action pipeline for stacked cable sorting, as shown in Fig. 1. Given an RGB observation, the system selects a manipulable target cable, generates dual-arm grasp points, lifts it with eagle-claw grippers, stretches it under bilateral wrist-force feedback, and places it into the target region. This process is repeated until the cables are sequentially singulated, straightened, and arranged.

A. Gripper Design and Stable Cable Grasping

Following CaRoBio [5], we use an eagle-claw fingertip mounted on a parallel gripper. Its hook-like geometry provides partial lateral constraint for cables up to $15\,\mathrm{mm}$ in diameter, reducing sliding during pickup and lifting.

B. Cable Instance Segmentation and Grasp Point Generation

As illustrated by the perception-to-grasp module in Fig. 1, the RGB observation is converted into a grasp representation consisting of a target cable segment, two grasp points, and their local tangent directions. We use Mask2Former [7] for cable instance segmentation and train it on 1000 MuJoCo-rendered RGB images with 2–5 randomly placed cables per scene. Cable color, length, placement, and illumination are randomized during rendering, and the instance annotations are stored in RLE format. During inference, predictions are filtered by confidence, mask area, and duplicate mask IoU, and the remaining masks are treated as cable instances.

For each retained instance, we skeletonize the mask and keep the dominant connected component as the visible centerline. Crossing candidates are detected by nearest-neighbor matching between centerlines of different instances. Inspired by FASTDLO and RT-DLO [9], [10], we use RGB color continuity, including local mask continuity, reference-color consistency, and color variation, to infer the over-under relationship at crossings. Occluded crossing indices split bridged centerlines into graspable segments, while disconnected mask components are directly treated as visible segments.

The final grasp pair is generated on the selected valid segment by excluding endpoint-adjacent regions and crossing neighborhoods. A sliding-window search selects two sufficiently separated points on the same continuous segment, and the local tangent direction at each grasp point is estimated from neighboring centerline samples.

C. Force-Guided Cable Straightening

The grasp point generation stage provides two grasp points, g_1 and g_2, on the selected manipulable cable segment and their local tangent directions, $\mathbf{t}_1$ and $\mathbf{t}_2$. The robot aligns the gripper openings with these tangents, grasps the cable with both arms, and lifts it above the stack. From the lifted gripper positions $\mathbf{x}_0^L$ and $\mathbf{x}_0^R$, the horizontal stretch direction is

$$\mathbf{d} = \frac{\Pi_{xy}(\mathbf{x}_0^R - \mathbf{x}_0^L)}{\|\Pi_{xy}(\mathbf{x}_0^R - \mathbf{x}_0^L)\|_2}, \tag{1}$$

TABLE I

CABLE HANDLING PERFORMANCE OF DIFFERENT GRIPPER PADS OVER 50 TRIALS PER GRIPPER. GRASP, LIFT, AND MOVE DENOTE WHETHER THE TARGET CABLE REMAINS HELD AFTER CLOSURE, LIFTING, AND LATERAL TRANSLATION, RESPECTIVELY. SLIP/DISTURBANCE COUNTS TARGET-CABLE SLIP OR NOTICEABLE DRAGGING OF A NEIGHBORING CABLE.

Gripper	Grasp	Lift	Move	Slip/Disturb
Native flat pad	0.48	0.44	0.38	0.34
Eagle-claw pad	0.82	0.74	0.68	0.26

where $\Pi_{xy}(\cdot)$ projects onto the table plane. During stretching, the two grippers move in opposite directions along $\mathbf{d}$ while maintaining their lifted heights.

Before stretching, the robot records a wrist-force bias $\mathbf{b}^a$ for each arm $a \in \{L, R\}$. At step t, it computes the bias-compensated force magnitude $f_t^a = \|\mathbf{F}_t^a - \mathbf{b}^a\|_2$ and filters it with an exponential moving average:

$$\bar{f}_t^a = \alpha f_t^a + (1-\alpha)\bar{f}_{t-1}^a, \quad a \in \{L, R\}, \tag{2}$$

where α is the smoothing coefficient. The cable is considered tensioned when both arms observe a force jump:

$$\Delta\bar{f}_t^L > \tau_{\Delta f}, \quad \Delta\bar{f}_t^R > \tau_{\Delta f}, \tag{3}$$

where $\Delta\bar{f}_t^a = \bar{f}_t^a - \bar{f}_{t-1}^a$. Stretching stops at this event or at the maximum step limit. The bilateral condition rejects single-arm contact artifacts and captures the transition from slack removal to cable tension. The robot then moves the stretched cable to the target region, lowers it, and releases it with a tangent-aligned three-phase motion.

III. Experiments

A. Gripper Grasping Performance

We compare the eagle-claw fingertip with the native Robotiq-2F-85 flat pad in stacked-cable grasping. Both variants use the same gripper body, actuator, linkage, controller, scene generation, and grasp-point input; only the terminal pad geometry is changed. The two pad geometries are shown in Fig. 2. Each trial contains grasping, lifting, and a short lateral motion.

Table I shows improved handling with reduced slip/disturbance.

B. Cable Segmentation and Grasp Point Detection

Segmentation performance is evaluated by semantic IoU (SemIoU), instance-mask IoU (InstmIoU), F1 score, and inference speed measured in frames per second (FPS). FASTDLO and RT-DLO, evaluated with their original weights, often produce fragmented masks in complex stacked scenes. YOLOv8-seg recovers more foreground regions but may merge adjacent cables near intersections. Mask2Former better preserves instance separation and produces masks that are more consistent with the ground-truth annotations.

Table II shows that Mask2Former achieves the best instance-level accuracy, while YOLOv8-seg is less reliable near crossings.

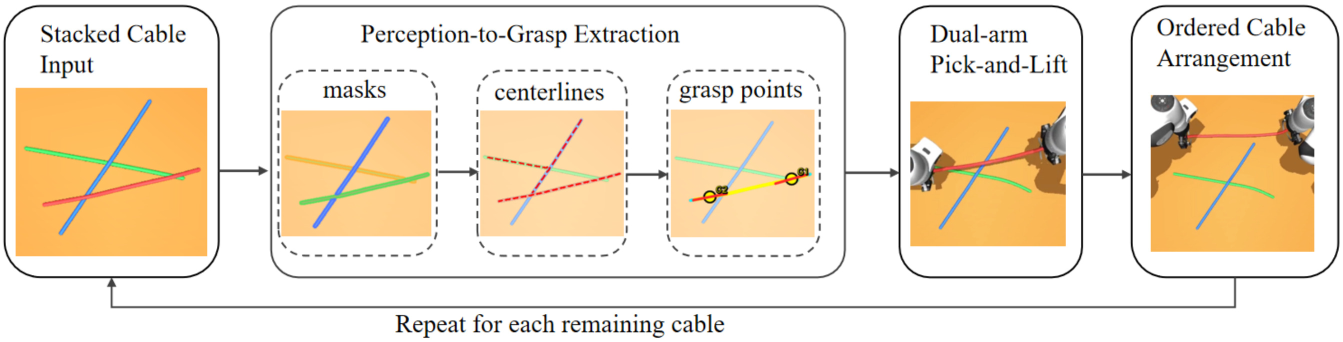
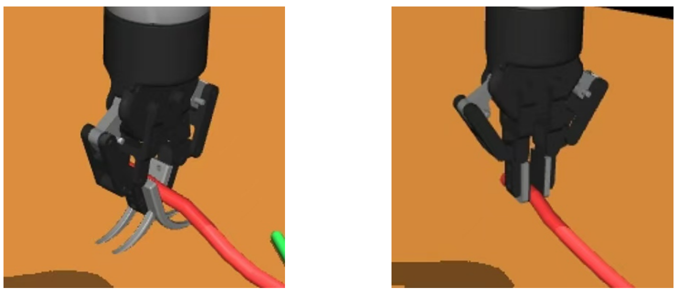


Fig. 1. Overall system pipeline for stacked cable sorting, including perception-to-grasp reasoning, dual-arm pick-and-lift, force-guided straightening, and iterative placement.



(a) Eagle-claw pad (b) Native flat pad

Fig. 2. Gripper pad comparison used in the grasping experiment: (a) eagle-claw fingertip and (b) native Robotiq-2F-85 flat pad.

TABLE II
QUANTITATIVE COMPARISON OF CABLE SEGMENTATION METHODS ON THE VALIDATION SET. FASTDLO AND RT-DLO USE THEIR ORIGINAL WEIGHTS, WHILE YOLOV8-SEG AND MASK2FORMER ARE TRAINED ON THE SAME SYNTHETIC SET. QUALITATIVE RESULTS ARE SHOWN IN FIG. 3.

Method	SemIoU	InstmIoU	F1	FPS
FASTDLO	0.55	0.16	0.32	1.87
RT-DLO	0.43	0.12	0.27	4.29
YOLOv8-seg	0.82	0.36	0.55	**52.19**
Mask2Former	**0.95**	**0.93**	**0.99**	1.91

C. Wrist-Force-Guided Straightening

We compare fixed-distance, fixed-force, and our force-jump termination rules. A straightening trial is successful if the tested termination condition is reached after the cable has been grasped and lifted. We also report the residual sag after stretching. Fig. 4 shows one representative force trace used to visualize the detected jump. Table III shows that our method achieves the highest straightening rate and the smallest residual sag magnitude.

D. Full Stacked Cable Management Task

Finally, we evaluate the full stacked-cable sorting task with five system configurations.

Table IV shows the highest task success with our full pipeline; failures mainly arise from reduced grasp reliability or less effective straightening.

TABLE III
STRAIGHTENING STRATEGY COMPARISON OVER 50 TRIALS PER METHOD. STRAIGHTEN REPORTS THE PERCENTAGE OF TRIALS REACHING THE STRAIGHTENED STATE BEFORE RELEASE. SAG DENOTES RESIDUAL VERTICAL SAG AFTER STRETCHING.

Method	Straighten	Residual sag magnitude (mm)
Fixed-distance	0.60	1.3
Fixed-force	0.64	0.5
Ours	0.78	0.3

TABLE IV
FULL STACKED-CABLE SORTING COMPARISON OVER 50 END-TO-END TRIALS PER CONFIGURATION. GRASP, STRAIGHTEN, AND TASK ARE MEASURED OVER ALL TRIALS; TASK REQUIRES GRASPING, STRAIGHTENING, AND FINAL RELEASE IN THE SAME TRIAL.

Configuration	Grasp	Straighten	Task
Mask2Former + Eagle-claw + Force jump	0.78	0.56	0.48
Mask2Former + Native pad + Force jump	0.40	0.32	0.30
Mask2Former + Eagle-claw + Fixed distance	0.68	0.40	0.26
Mask2Former + Eagle-claw + Fixed force	0.66	0.48	0.42
YOLOv8-seg + Eagle-claw + Force jump	0.46	0.34	0.28

IV. DISCUSSIONS

Our results support a modular, structure-aware approach for stacked cable sorting. Instead of reconstructing the full 3D deformable state, the system converts RGB observations into graspable segments and dual-arm grasp points, while bilateral wrist-force feedback detects the transition from slack removal to cable tension. This design is task-oriented: it extracts physical cues that are directly relevant to sorting, including visible topology, local grasp geometry, contact stability, and tension transition.

The current evaluation is conducted in simulation, and real-world validation remains an important next step. Although the proposed pipeline relies only on common sensing modalities, namely RGB images and wrist-force feedback, sim-to-real transfer may still be affected by camera calibration errors, lighting variation, cable material differences, contact friction, and gripper compliance. Future real-robot experiments will examine whether synthetic-data-trained segmentation and force-jump-based straightening can be transferred to physical cable piles, and whether additional domain randomization or limited real-world fine-tuning is required.

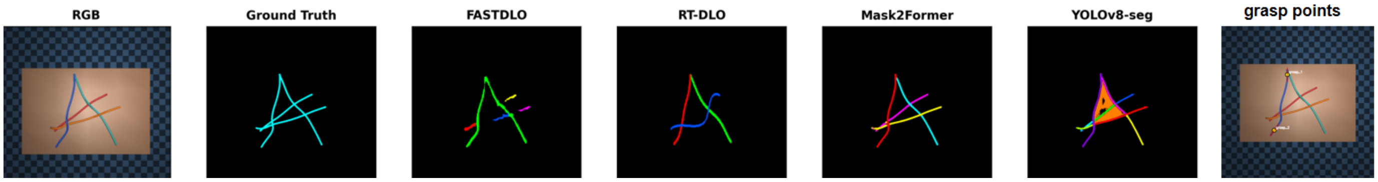


Fig. 3. Qualitative comparison of cable segmentation methods and downstream grasp point detection. Yellow circles indicate grasp points, and red line segments indicate local tangent directions.

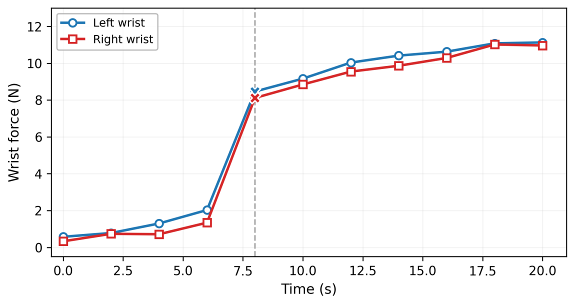


Fig. 4. Representative wrist-force response during cable straightening. The dashed vertical line marks the detected force-jump event.

The perception module relies on 2D RGB color continuity, which may fail in monochromatic cable stacks, severe occlusions, branching cables, or cases where cables are only partially visible in the camera image. Branching cables violate the current single-centerline assumption, while partial visibility may lead to incomplete graspable segments or incorrect target selection. Future work will extend the centerline representation to a graph-based DLO structure, where endpoints, junctions, crossings, and visible fragments are explicitly modeled. Multi-view perception or active viewpoint adjustment could further reduce ambiguity under partial visibility.

Another limitation is that the current experiments use cables with relatively consistent mechanical properties. Cable stiffness can affect both the observed shape and the force response during stretching. Softer cables may exhibit larger sag and delayed force increase, whereas stiffer cables may produce earlier force changes or resist local bending near the grasp points. Therefore, the force-jump threshold and bilateral termination condition should be further evaluated across cables with different diameters, materials, and bending stiffness. A promising extension is to make the termination criterion stiffness-adaptive by estimating cable compliance online from the force-displacement response during initial stretching.

The 48.0% full-task success rate also reflects compounding errors, where slips during pick-and-lift propagate to straightening and placement failures. Nevertheless, the results suggest that physically interpretable intermediate representations can provide a useful bridge between perception and action in deformable-object manipulation.

Toward Physics Action Model. This work represents an early step toward Physics Action Model (PAM) for deformable object manipulation (Fig. 5). A PAM generates actions from

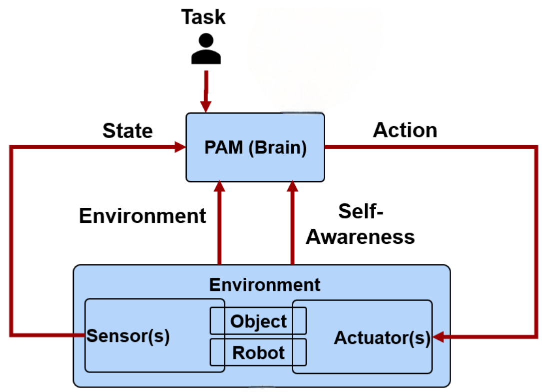


Fig. 5. Simplified Physics Action Model (PAM) concept. PAM maps task conditions and sensed robot-object-environment states to actions through physically meaningful representations, interactive perception, and closed-loop feedback. This paper instantiates a task-specific PAM-like pipeline for stacked cable sorting through topology-aware grasp generation, contact-stabilized grasping, and force-guided straightening.

physically meaningful representations of the robot, object, and environment, updated through interactive sensing and feedback. From this perspective, our pipeline provides a task-specific PAM-like instance: cable crossings encode topological cues for grasp selection, the eagle-claw fingertip promotes stable contact, and bilateral wrist-force feedback detects the transition from slack removal to cable tension. Future work will extend this direction with real-world validation, stiffness-aware physical state estimation, graph-structured DLO representations, disturbance-aware feedback, and closed-loop action refinement.

V. Conclusion

This paper presented a modular perception-action pipeline for sorting stacked deformable cables. By combining structure-aware perception, an eagle-claw gripper pad, and adaptive bilateral force-guided straightening, the system achieves an InstmIoU of 0.93 and a 48.0% full-task success rate in complex simulated clutter. The pipeline also offers a preliminary step toward PAM by using physical semantics such as topology, contact, grasp stability, and tension transition to guide robot actions. Future work will address real-robot deployment, sim-to-real transfer, and richer physical state representations.

References

[1] A. Caporali, I. Cuiral-Zueco, G. López-Nicolás, and G. Palli, "Robotic perception and manipulation of deformable linear objects: A survey,"

The International Journal of Robotics Research, 2026, onlineFirst. [Online]. Available: https://doi.org/10.1177/02783649261432253

[2] Y. She, S. Wang, S. Dong, N. Sunil, A. Rodriguez, and E. H. Adelson, "Cable manipulation with a tactile-reactive gripper," *The International Journal of Robotics Research*, vol. 40, no. 12–14, pp. 1385–1401, 2021.

[3] O. Azulay, K. Kondap, J. Drake, S. Xie, H. Li, S. Chitta, and K. Goldberg, "MOTORCYCLE 1.0: Automating bimanual cable routing around fixtures on the NIST task board," in *2025 IEEE 21st International Conference on Automation Science and Engineering (CASE)*, 2025, pp. 2636–2641. [Online]. Available: https://doi.org/10.1109/CASE58245.2025.11163872

[4] L. Manes, S. Fichera, H. Fakhruldeen, A. I. Cooper, and P. Paoletti, "A soft cable loop based gripper for robotic automation of chemistry," *Scientific Reports*, vol. 14, p. 8899, 2024.

[5] J. Zuo, B. Zhang, and F. Zhang, "CaRoBio: 3d cable routing with a bio-inspired gripper fingernail," in *2025 10th International Conference on Automation, Control and Robotics Engineering (CACRE)*, 2025, pp. 72–79. [Online]. Available: https://doi.org/10.1109/CACRE66141.2025.11119609

[6] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2980–2988. [Online]. Available: https://doi.org/10.1109/ICCV.2017.322

[7] B. Cheng, I. Misra, A. G. Schwing, A. Kirillov, and R. Girdhar, "Masked-attention mask transformer for universal image segmentation," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 1280–1289. [Online]. Available: https://doi.org/10.1109/CVPR52688.2022.00135

[8] A. Caporali, K. Galassi, and G. Palli, "Ariadne+: Deep learning-based augmented framework for the instance segmentation of wires," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 12, pp. 8607–8617, 2022. [Online]. Available: https://doi.org/10.1109/TII.2022.3154477

[9] A. Caporali, K. Galassi, R. Zanella, and G. Palli, "FASTDLO: Fast deformable linear objects instance segmentation," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 9075–9082, 2022. [Online]. Available: https://doi.org/10.1109/LRA.2022.3189791

[10] A. Caporali, K. Galassi, B. L. Žagar, R. Zanella, G. Palli, and A. C. Knoll, "RT-DLO: Real-time deformable linear objects instance segmentation," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 11, pp. 11 333–11 342, 2023. [Online]. Available: https://doi.org/10.1109/TII.2023.3245641

[11] A. Choi, D. Tong, B. Park, D. Terzopoulos, J. Joo, and M. K. Jawed, "mBEST: Realtime deformable linear object detection through minimal bending energy skeleton pixel traversals," *IEEE Robotics and Automation Letters*, vol. 8, no. 8, pp. 4863–4870, 2023. [Online]. Available: https://doi.org/10.1109/LRA.2023.3290419

[12] 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 *Proceedings of the 7th Conference on Robot Learning*, 2023.

[13] S. Kozlovsky, O. Joglekar, and D. Di Castro, "ISCUTE: Instance segmentation of cables using text embedding," 2024. [Online]. Available: https://arxiv.org/abs/2402.11996

[14] 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," in *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023.

[15] H. Wakamatsu and S. Hirai, "Static modeling of linear object deformation based on differential geometry," *The International Journal of Robotics Research*, vol. 23, no. 3, pp. 293–311, 2004.

[16] H. Nakagaki, K. Kitagaki, T. Ogasawara, and H. Tsukune, "Study of deformation and insertion tasks of a flexible wire," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, vol. 3, 1997, pp. 2397–2402.

[17] F. Süberkrüb, R. Laezza, and Y. Karayiannidis, "Feel the tension: Manipulation of deformable linear objects in environments with fixtures using force information," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022, pp. 11 216–11 222.

[18] J. Zhu, B. Navarro, P. Fraisse, A. Crosnier, and A. Cherubini, "Dual-arm robotic manipulation of flexible cables," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 479–484.

[19] M. Yu, K. Lv, C. Wang, M. Tomizuka, and X. Li, "A coarse-to-fine framework for dual-arm manipulation of deformable linear objects with whole-body obstacle avoidance," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 10 153–10 159.

[20] L. Zhang, K. Bai, Q. Li, Z. Chen, and J. Zhang, "A collision-aware cable grasping method in cluttered environment," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024, pp. 2126–2132.