Wire Instance Perception from RGBD Imagery with Mask R-CNN

Jingyi Xiang[1] | Graduate Mentor: Holly Dinkel[2]

[1]Department of Electrical & Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL

[2]Department of Aerospace Engineering, University of Illinois at Urbana-Champaign, Urbana, IL

Introduction

This work performs wire instance segmentation in RGBD imagery with the goal of developing deformable linear object perception and manipulation capabilities for intravehicular robots in space habitats.

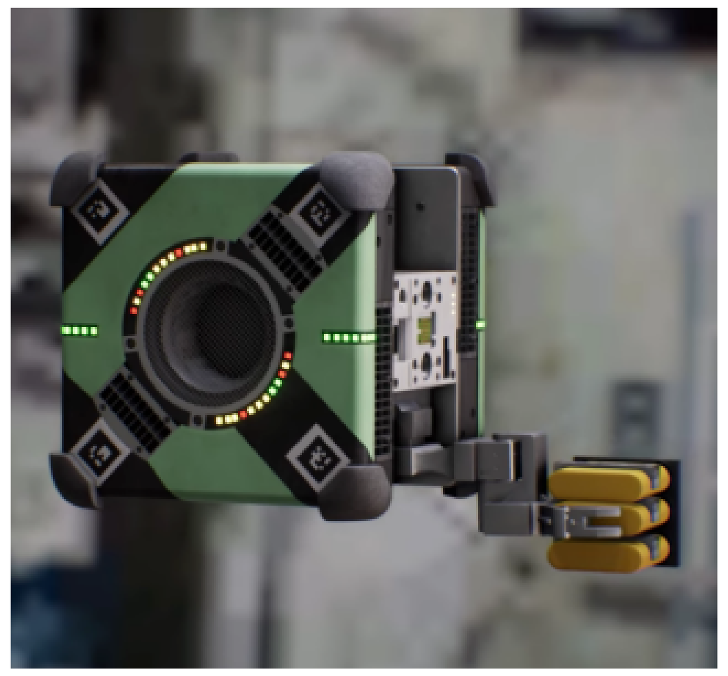

Figure 1: Rendering of the Astrobee free-flying robot in the International Space Station.

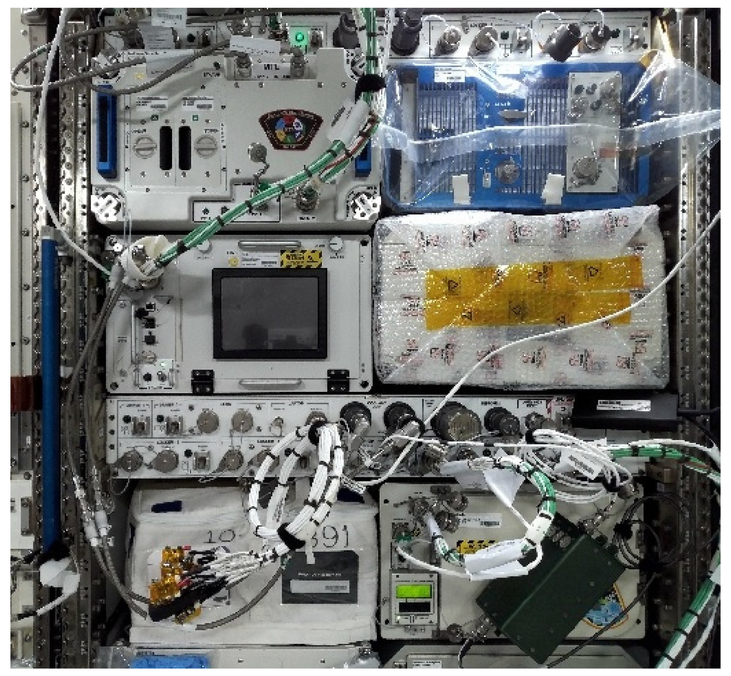

Figure 2: Wire-abundant environment in the International Space Station.

The problem can be approached with the following steps:

1. Perform instance segmentation on RGB images using Mask R-CNN
2. Use RGB instance segmentation masks to segment a point cloud
3. Construct a surface model of the objects to use in manipulation tasks

Step 1: Instance Segmentation in RGB Image

The Detectron2 implementation of Mask R-CNN was trained on the UIUCWires data set to perform wire instance segmentation.

Two types of annotation format were tested:

1. The Object Semantics (OS) format (shown in Figure 3), where all contours of an object are represented within one mask
2. The Object Segment Semantics (OSS) format (shown in Figure 4), where one contour of an object is represented within one mask

Figure 5 and 6 show predictions by Mask R-CNN trained with OS and OSS annotation format. The one trained with OSS annotation format shows cleaner and more complete instance segmentation masks.

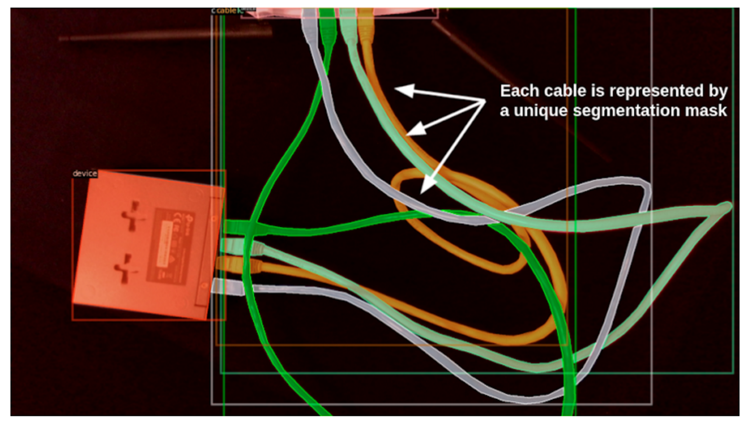

Figure 3: Object Semantics (OS) format.

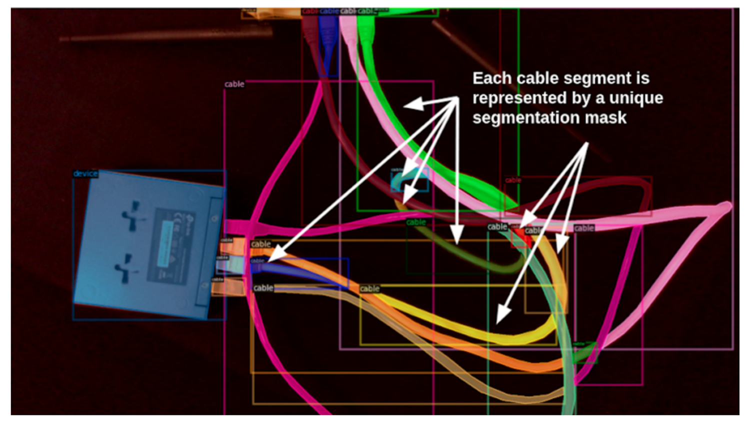

Figure 4: Object Segment Semantics (OSS) format.

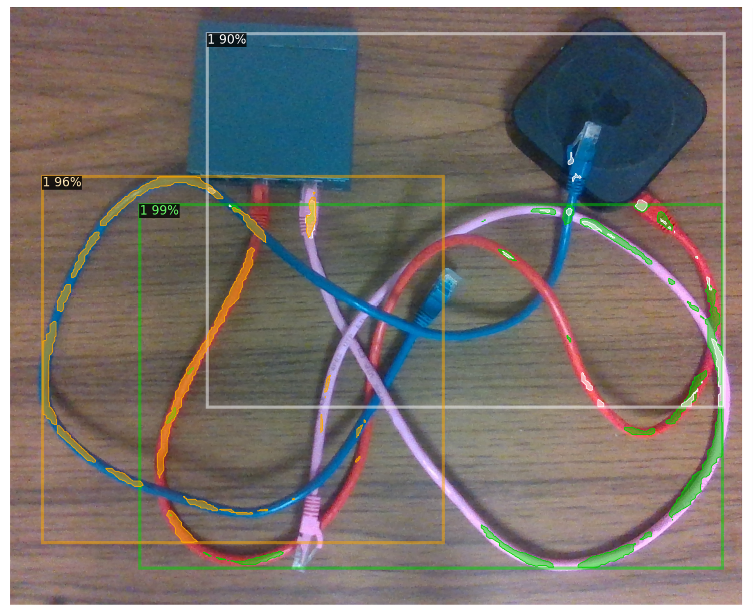

Figure 5: Predictions by Mask R-CNN trained with OS annotation format.

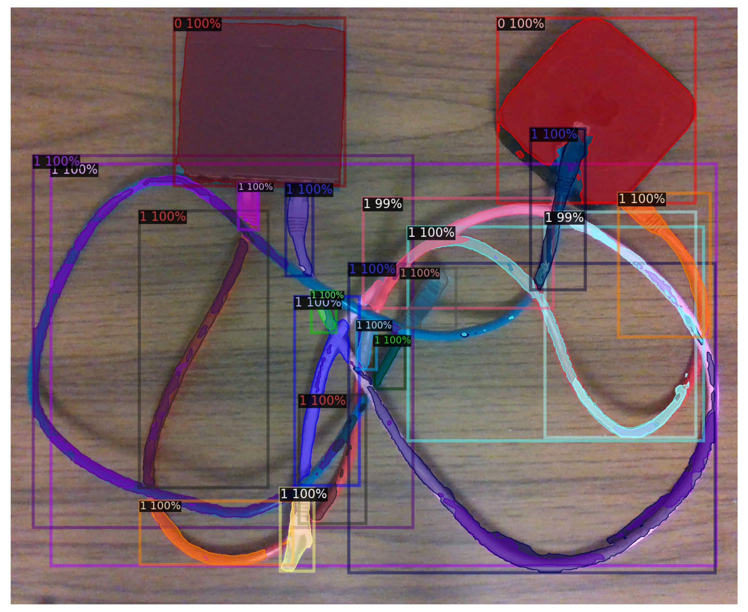

Figure 6: Predictions by Mask R-CNN trained with OSS annotation format.

Step 2: Instance Segmentation in Point Cloud

Figure 7: Intel® RealSense™ Depth Camera D435.

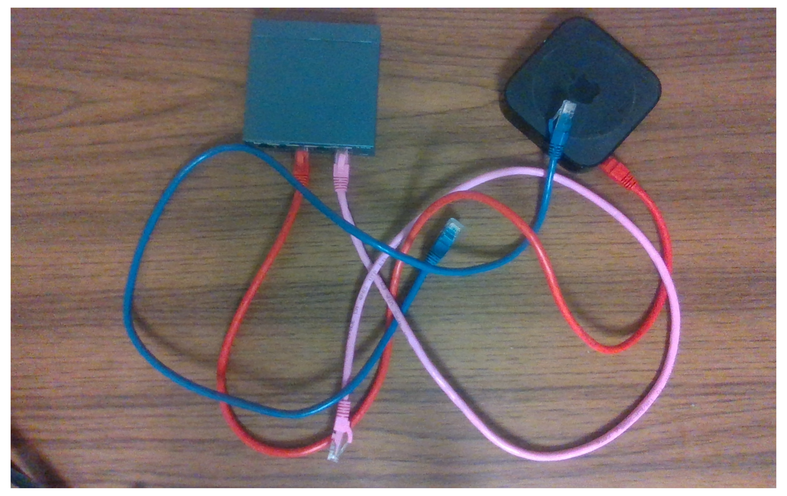

Figure 8: Color image of the scene.

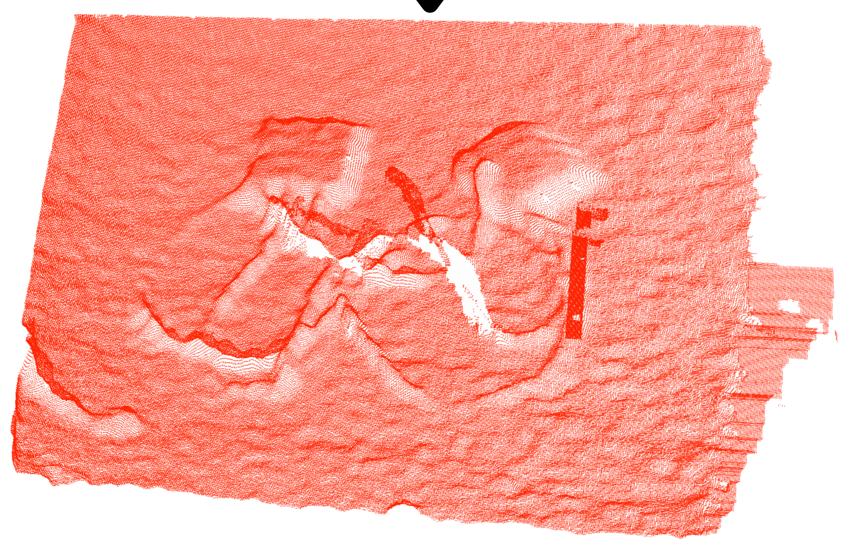

Figure 9: Raw point cloud of the scene.

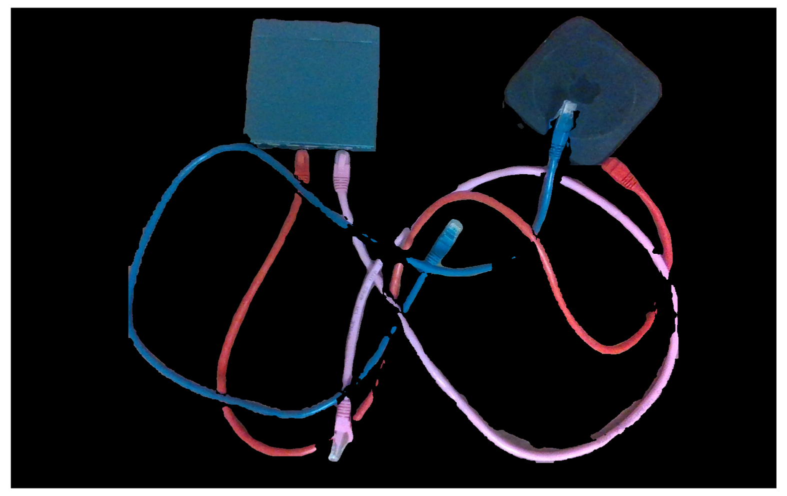

Figure 10: Instance segmentation masks predicted by Mask R-CNN.

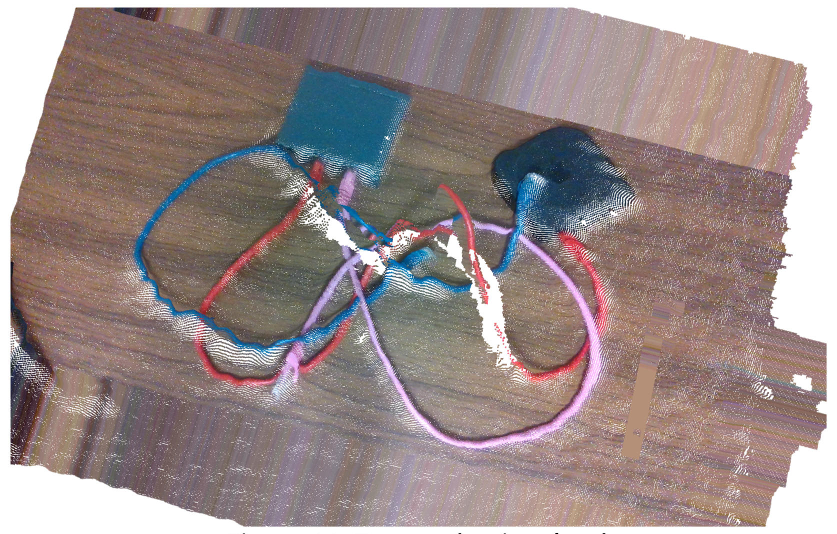

Figure 11: Textured point cloud.

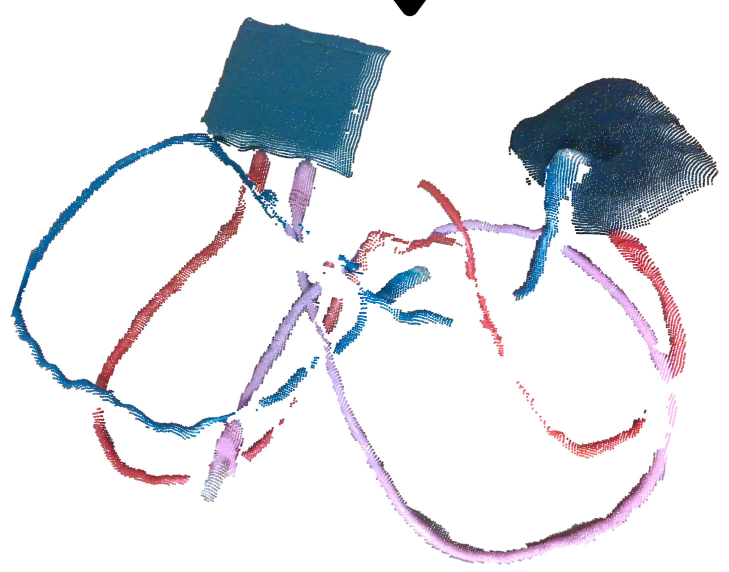

Figure 12: Segmented point cloud.

The Procedure:

1. Use the Intel RealSense D435 depth camera to collect the RGB image (Figure 8) and the point cloud (Figure 9) of the scene
2. Predict wire instances in the RGB image with Mask R-CNN (Figure 10)
3. Texturize the point cloud by mapping the color frame to the depth frame (Figure 11)
4. Apply the instance segmentation masks to the textured point cloud to obtain the 3D points that belong to each wire instances (Figure 12)

Test Results:

The above procedure was tested on two scenes, one contains two wires, the other three wires. The results (shown in Table 1) showed that training Mask R-CNN with the Object Segment Semantics annotation format improved the number of points recovered in depth segmentation by 153%.

Table 1: Number of points in instance-segmented point clouds

Annotation Format	Two-Wire Scene	Three-Wire Scene
Object Segment Semantics	21,688	26,366
Object Semantics	8,580	4,082

Step 3: Surface Reconstruction

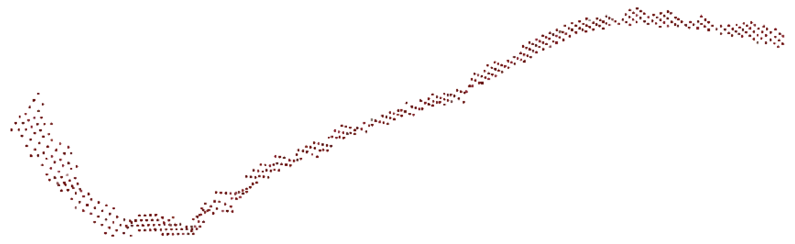

Figure 13: Segmented raw point cloud.

Segmented, noisy, raw point cloud

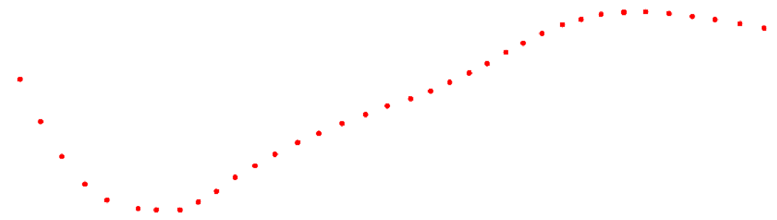

Figure 14: Smooth curve obtained from applying 1D MLS.

Apply 1D Moving Least Square algorithm (1D MLS) to obtain a smooth curve

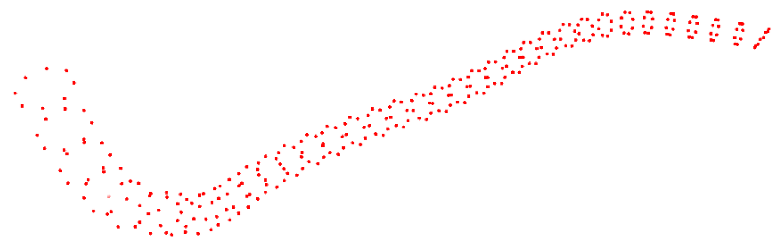

Figure 15: The new point cloud produced by estimating the wire as a cylindrical object.

Assume a wire radius and approximate the wire as a cylindrical object by constructing circles perpendicular to each line segment in the 1D MLS curve

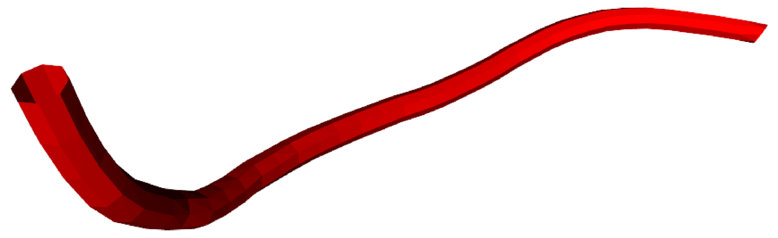

Figure 16: Mesh created from the new point cloud.

Traverse the resulting points to establish connectivity and create the mesh object

Conclusions and Future Research

This work demostrates three steps toward autonomous perception of wires:

1. Wire object instances can be identified in an RGB image using Mask R-CNN
2. Segmentation masks can be applied to an image-aligned point cloud to extract the 3D shape of the object
3. Wires can be represented as cylindrical mesh objects for robotics manipulation

This work exposes several new areas for contribution:

1. Point cloud obtained from multiple perspectives could be fused to reduce noise in scene reconstruction
2. The wire radius could be estimated directly from perception and replace the assumed radius in the mesh generation step

Acknowledgements

I would like to express my sincere gratitude to my mentor Holly Dinkel for her guidance and encouragement throughout the semester, project collaborators Harry Zhao, Brian Cotlin, Trey Smith, and Tim Bretl for their valuable technical support, and the Undergraduate Research Apprenticeship Program and NASA for their support of this work.

References

H. Dinkel, J. Xiang, H. Zhao, B. Coltin, T. Smith, and T. Bretl, "Wire Point Cloud Instance Segmentation from RGBD Imagery with Mask R-CNN," in *Under Submission*, 2022.

H. Dinkel, H. Zhao, J. Xiang, B. Coltin, T. Smith, and T. Bretl, "Benchmarking Wire Instance Segmentation with Mask R-CNN," in *Under Submission*, 2022.

L. Flückiger, K. Browne, B. Coltin, J. Fusco, T. Morse, and A. Symington. "Astrobee Robot Software: A Modern Software System for Space." In *Proceedings of the 14th International Symposium on Artificial Intelligence, Robotics and Automation in Space (iSAIRAS)* (2018).

K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," in *IEEE Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Oct. 2017, pp. 2980–2988.

Intel. "PyRealsense2: Python Wrapper for Intel Realsense SDK 2.0," https://pypi.org/project/pyrealsense2, 2022.

L. Keselman, J. Woodfill, A. Grunnet-Jepsen, A. Bhowmik, "Intel® RealSense™ Stereoscopic Depth Cameras", in *IEEE Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, July 2017, pp. 1-10

A. Kirillov, Y. Wu, K. He, and R. Girshick, "PointRend: Image Segmentation as Rendering," in *IEEE Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, June 2020, pp. 9796–9805.

M. Musy, G. Jacquenot, G. Dalmasso, neoglez, R. de Bruin, A. Pol- lack, F. Claudi, C. Badger, icemtel, Z.-Q. Zhou, B. Sullivan, B. Lerner, D. Hrisca, and D. Volpatto, "Vedo: A Python Module for Scientific Analysis and Visualization of 3D Objects and Point Clouds," https://vedo.embl.es, 2022.

NASA. "Astrobee-Released-Data," Feb. 2022. https://nasagov.app.box.com/s/4ign43svk39guhy9ev8t5xkzui6tqjm1