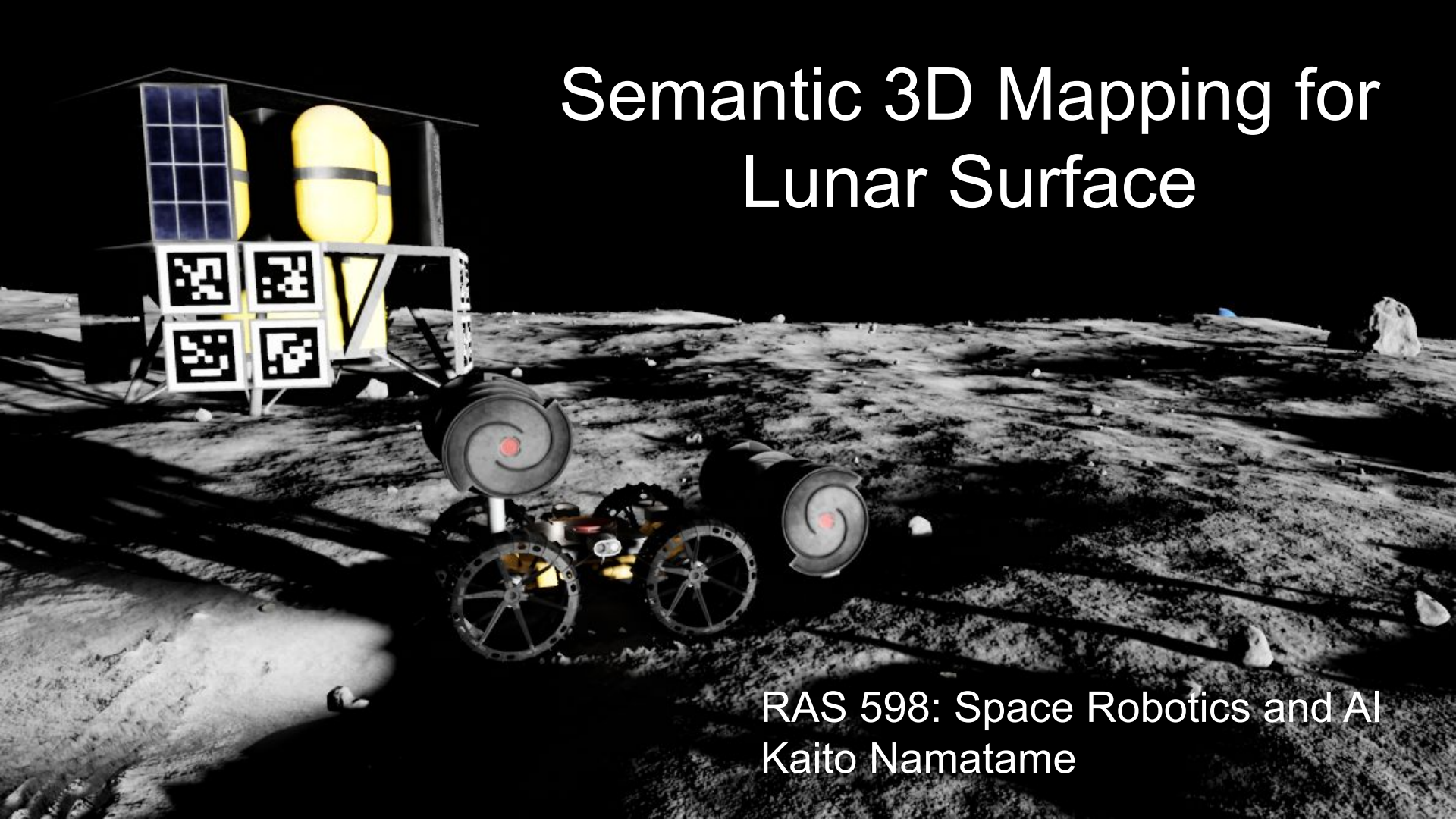
Semantic 3D Mapping for Lunar Surface



RAS 598: Space Robotics and AI

Kaito Namatame

Goal

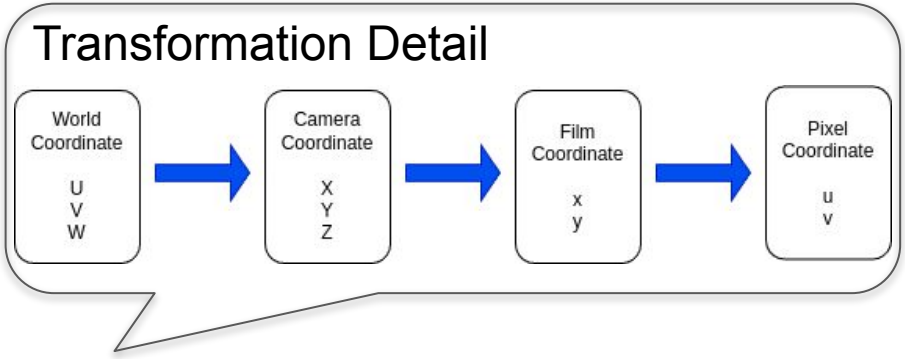
The goal of this project is to provide Semantic 3D point cloud by combining the semantic segmentation which was developed in the midterm project and ORB-SLAM.

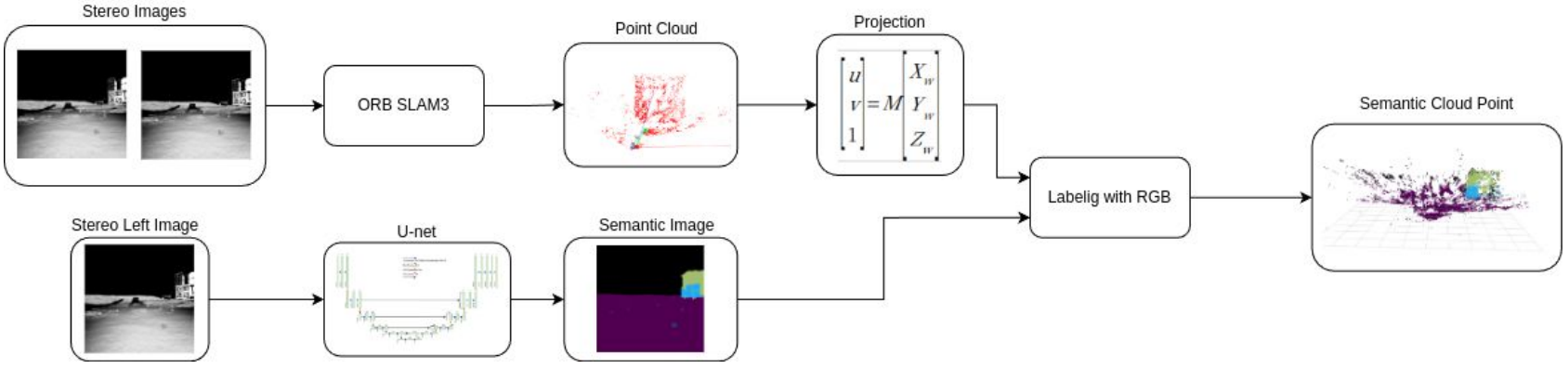
Semantic segmentation does not provide depth or 3D information. On the other hand, ORB-SLAM can generate 3D point cloud but lacks semantic understanding of the scene. To address this problem, combining those two helps this limitation.

By combining the strong point of both methods, I was able to generate a semantic 3D point cloud with more detailed information which overcomes limitation of both semantic segmentation and ORB-SLAM.

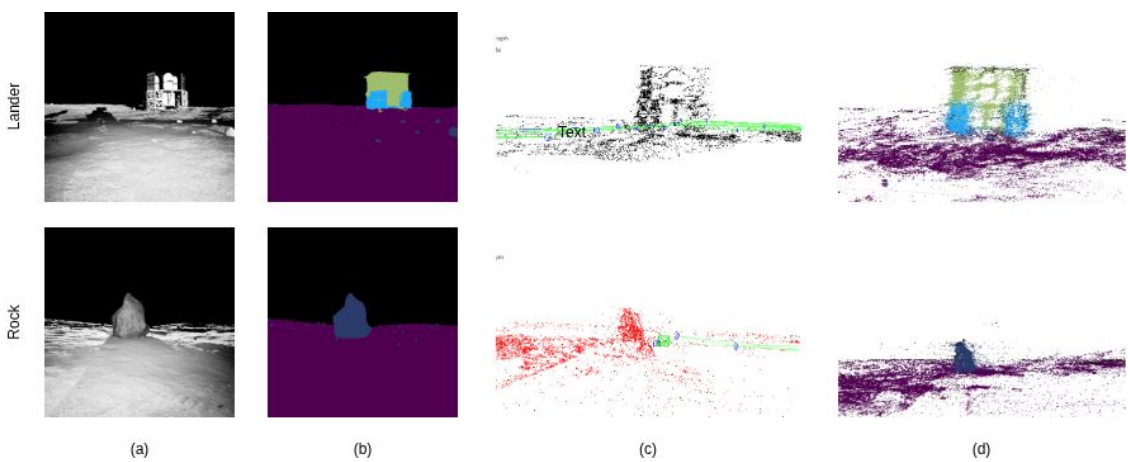
Methodology

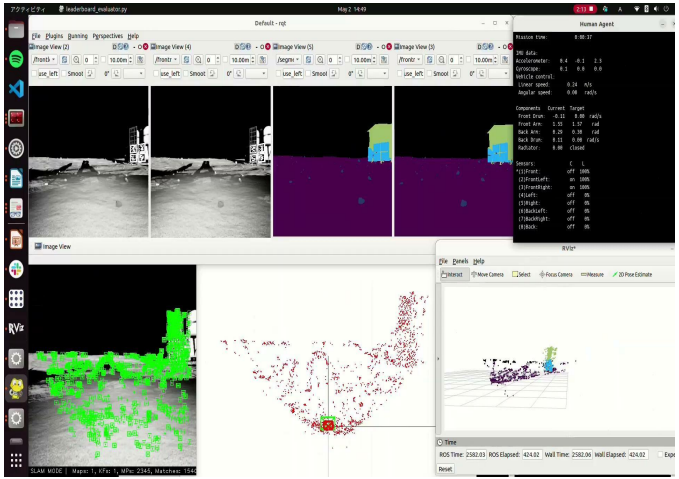
- Stereo images are processed by ORB-SLAM3 to estimate camera pose and generate a 3D point cloud.
- The left stereo image is segmented using U-Net model to produce a semantic mask.
- 3D points are projected onto the semantic mask using coordinate transformations.
- Semantic labels (RGB colors) are assigned to each 3D point based on the corresponding pixel's class.





Results

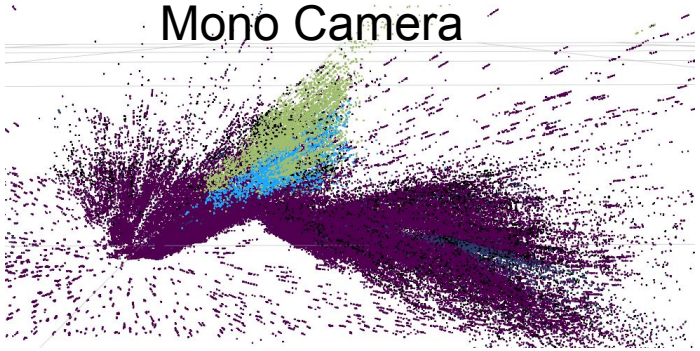




(a) is a raw image (b) is a corresponding semantic mask (c) is a 3d point cloud results of ORB-SLAM3 (d) is the results of semantic 3d point cloud.

Challenge and Future Plan

Mono Camera



Stereo Camera

1. The generated point cloud contains noise. Applying filtering to each ray could provide better results.

2. The U-Net architecture with VGG16 as the encoder is slow, providing 2D semantic images at only 4 FPS. This problem can be addressed by replacing the encoder with a more efficient backbone, such as EfficientNet or MobileNet.