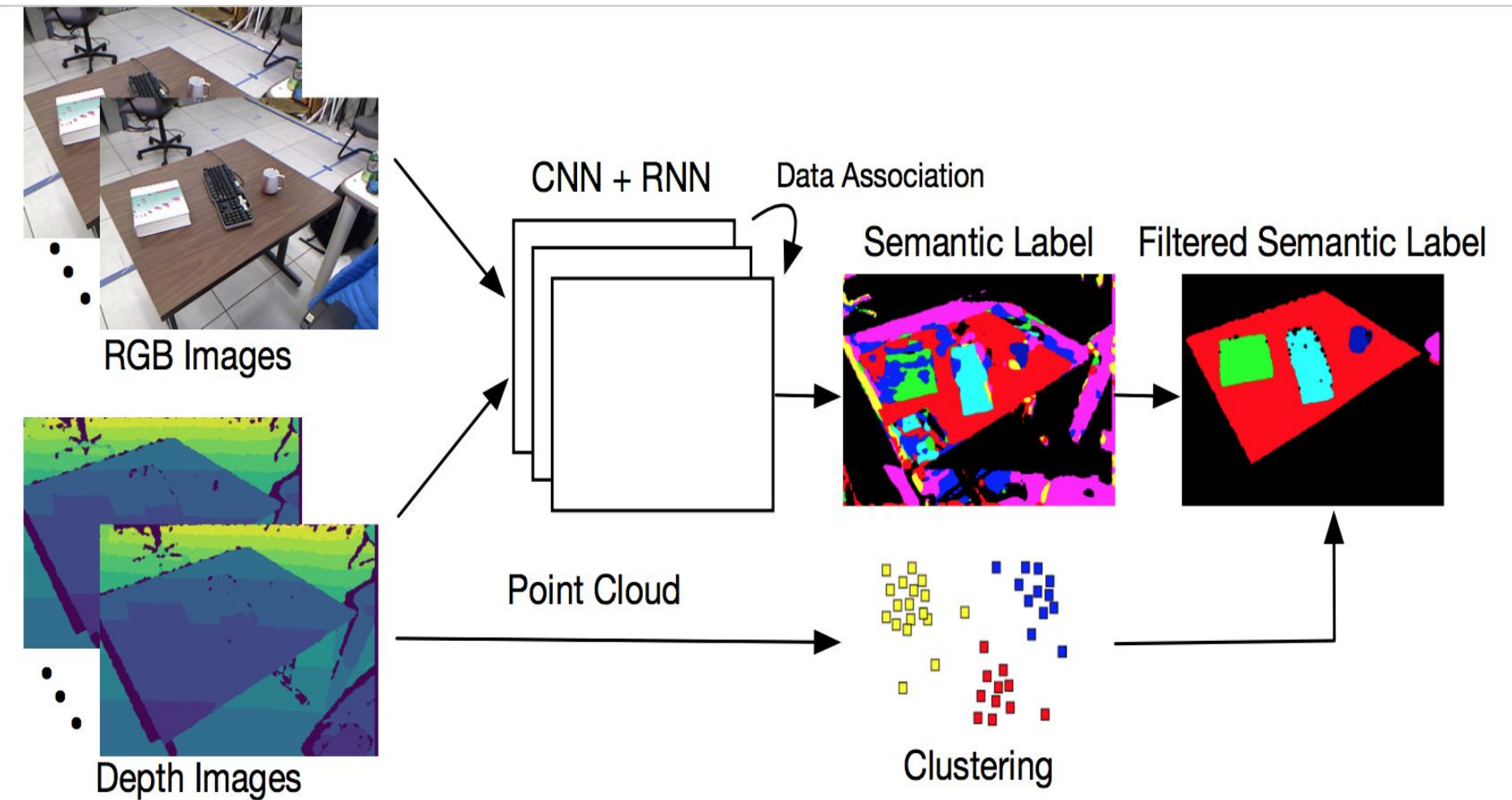


Enhanced Recurrent Neural Network Semantic Labeling with Point Cloud Processing

Wei Zhang, Iretiayo Akinola, David Watkins and Peter Allen,
Columbia University

Overview: Semantic grasping and manipulation requires a good understanding of the objects in a scene but many learned semantic labeling systems give noisy labels when deployed in real world. This work combines domain randomization and a point cloud post-processing step to enhance semantic labeling of a table top scene.

- Model Inputs: RGB and depth image.
- The model is trained on virtual 3D dataset and tested in a real scene.
- Semantic labels from model is usually noisy because it is generally hard to generalize from virtual scene to real scene.

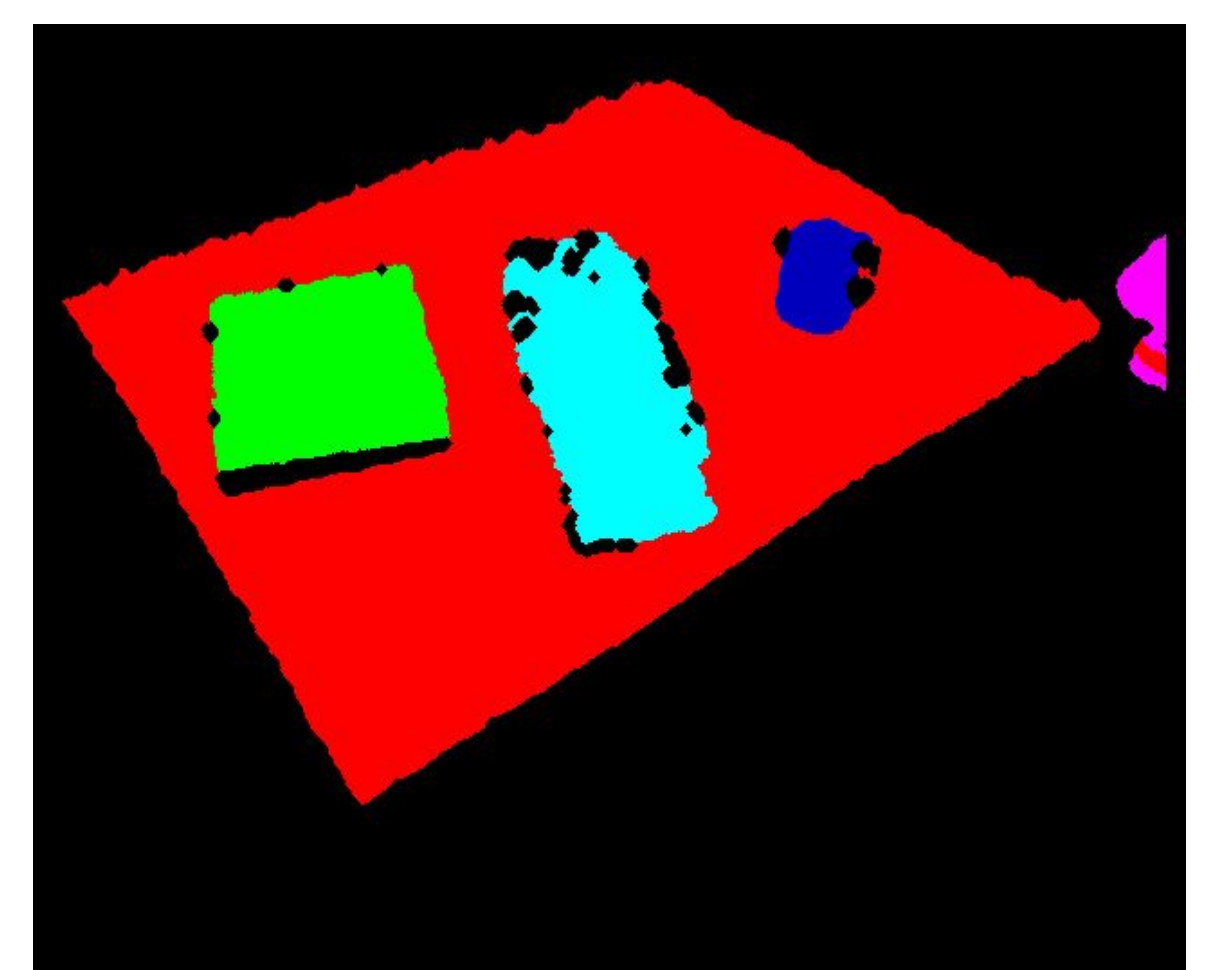
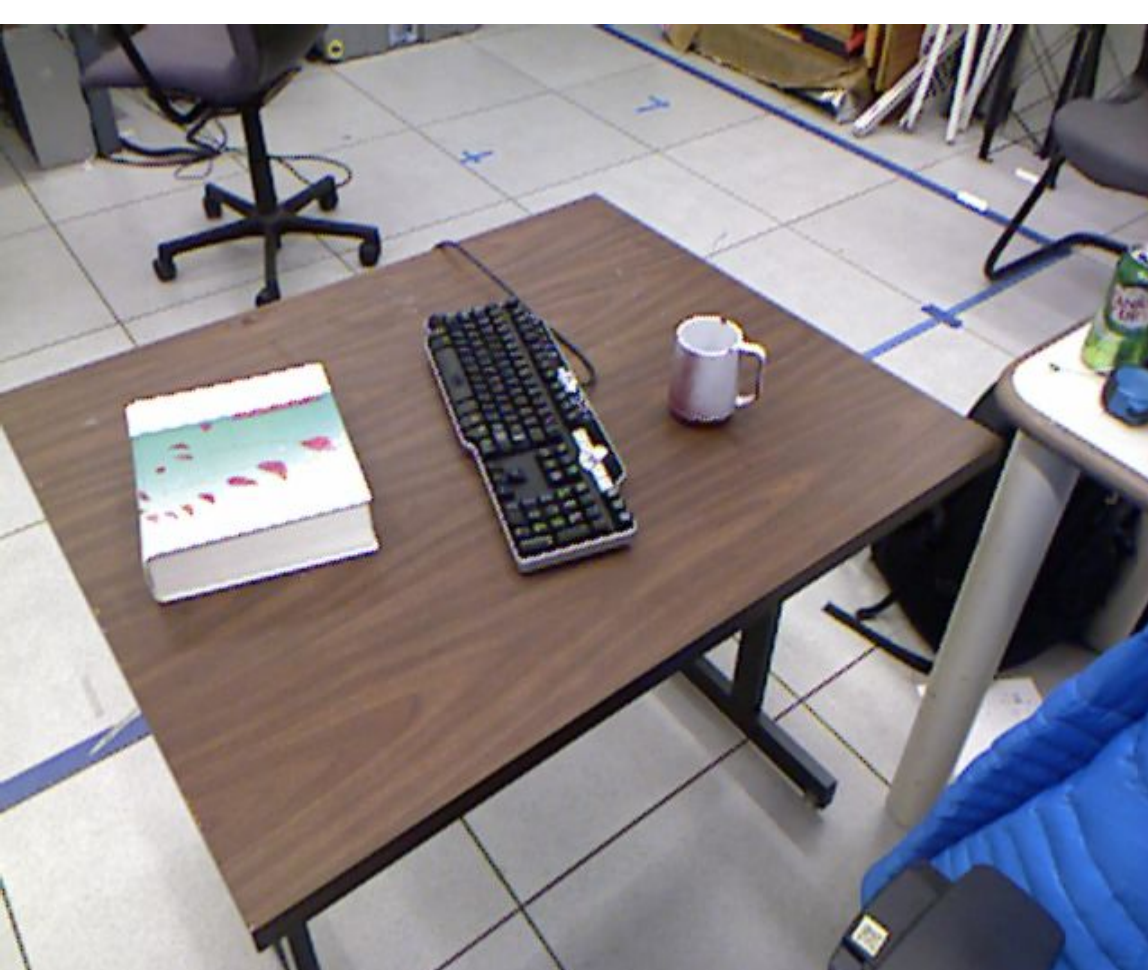
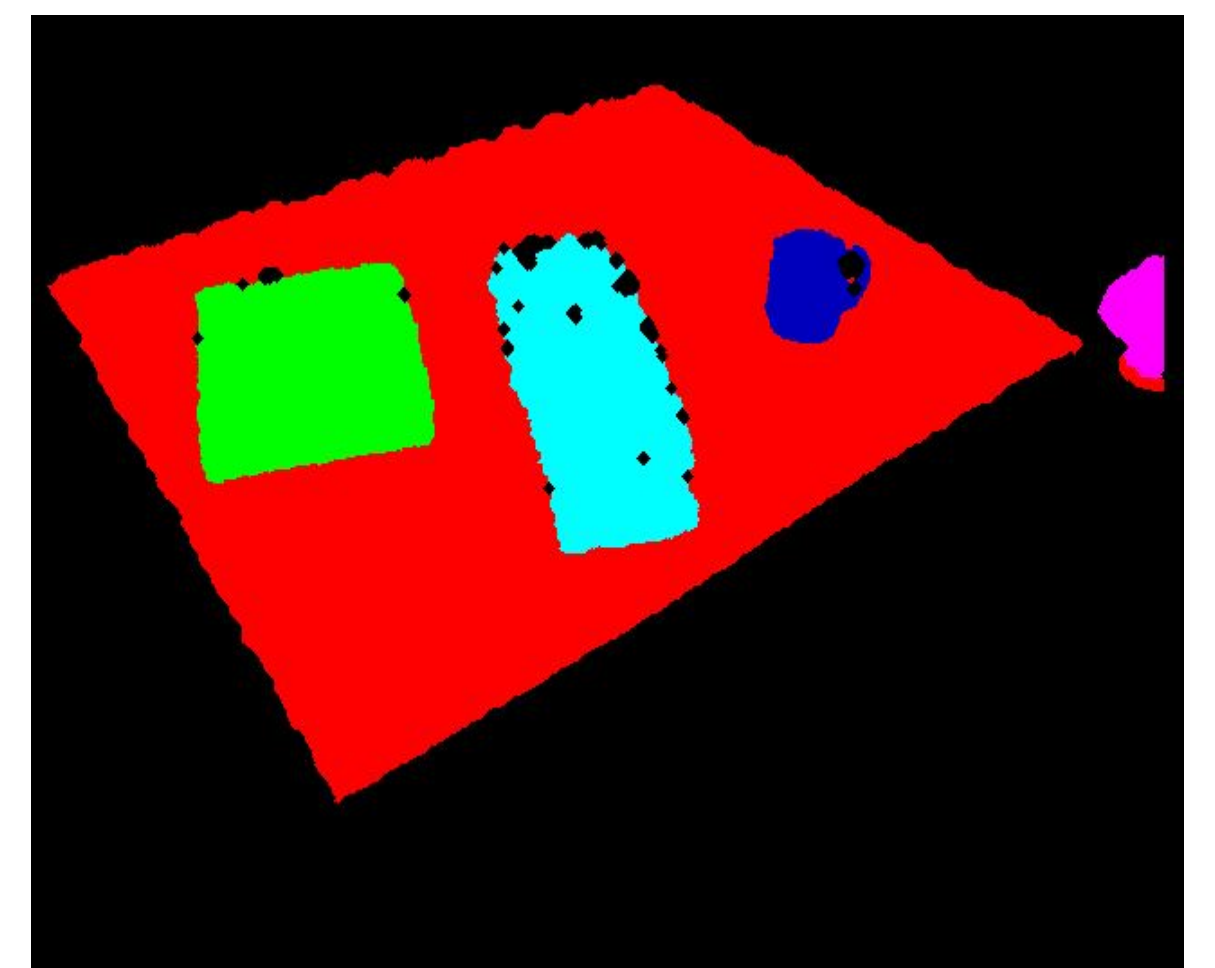
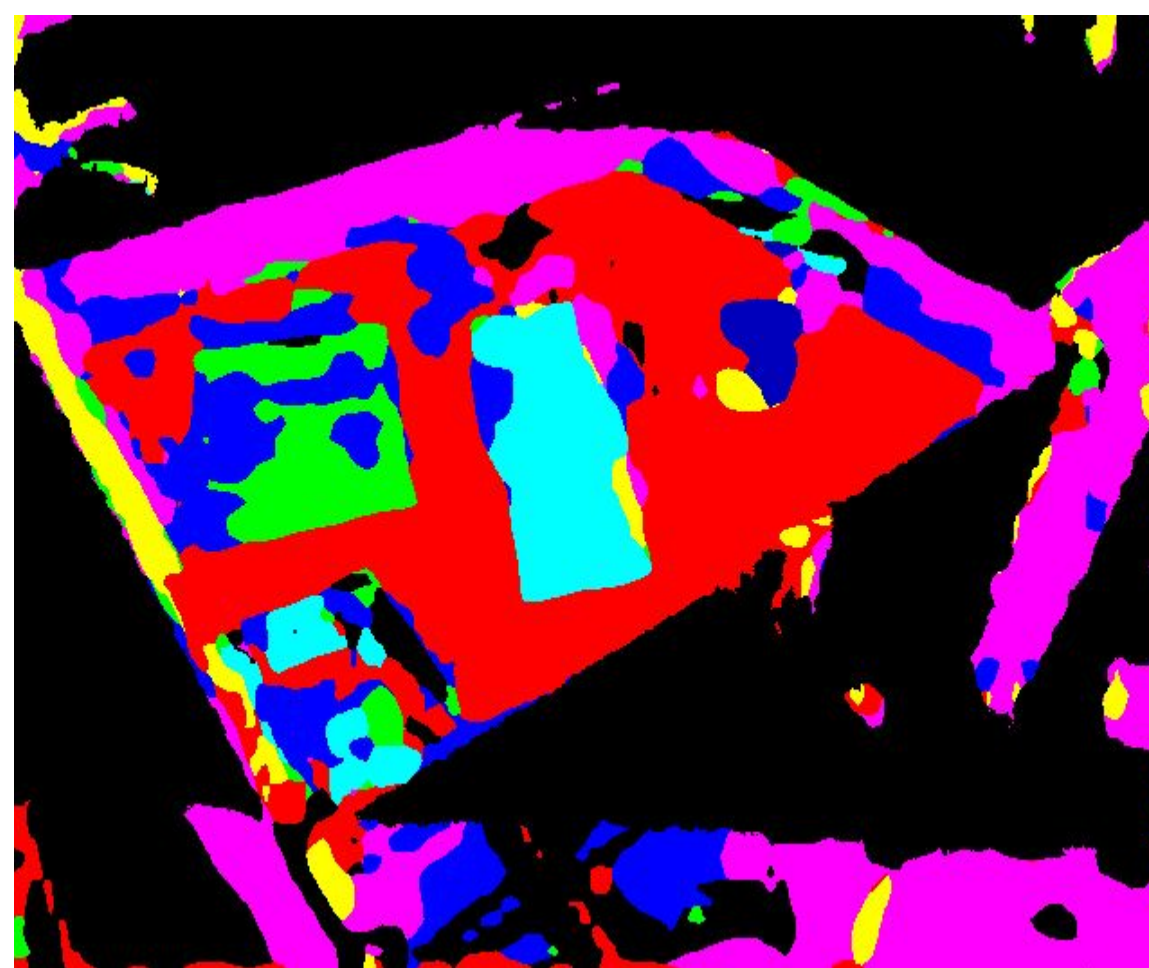
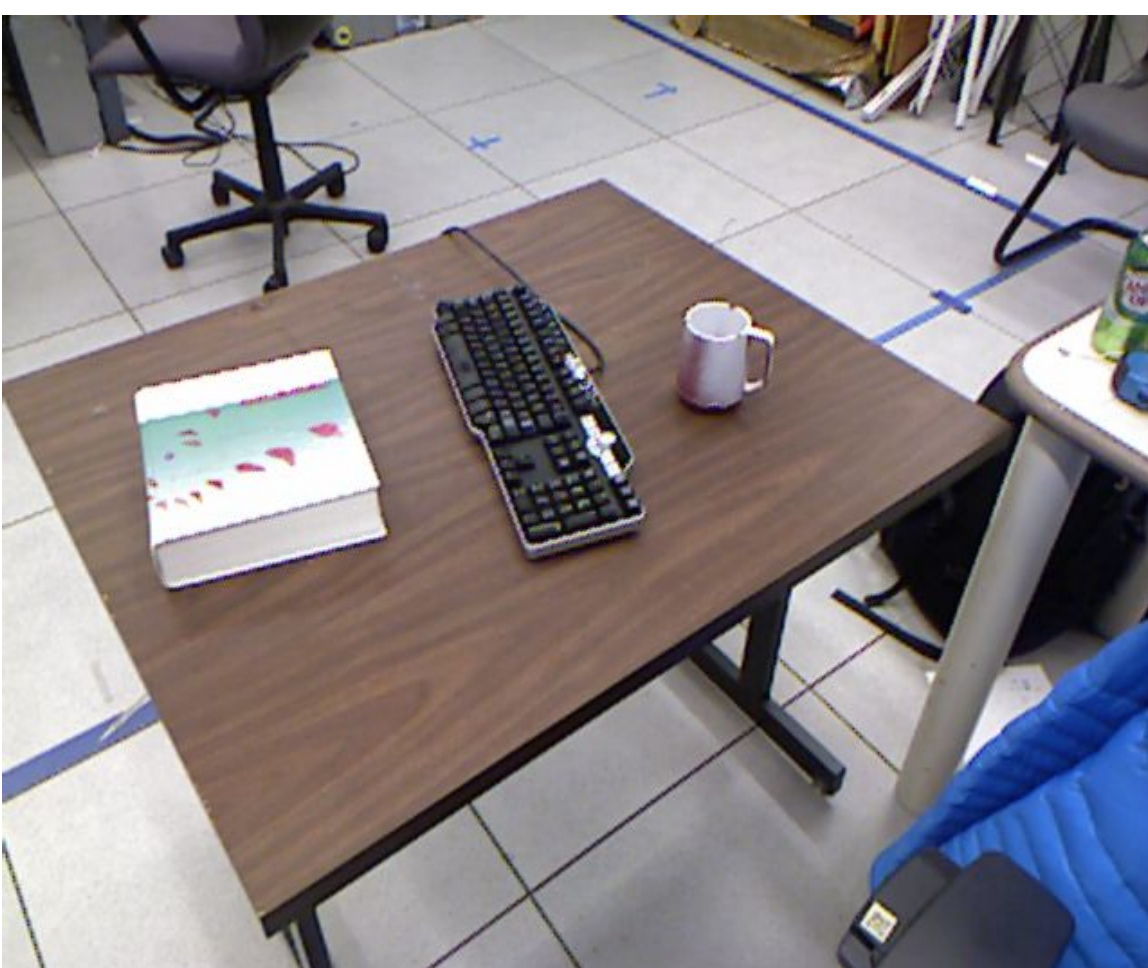


Pipeline of Enhanced Semantic Label

Proposed Methods: To improve, in addition to increasing the diversity of the training dataset using domain randomization, we propose to use point cloud to post filter the semantic labels from the model.

1. Generate semantic labeling image from the model
2. Acquire point cloud for the current frame
3. Perform plane finder algorithm and use the found plane coordinates as cluster centroid for clustering
4. Project 3D point back to 2D pixel and perform flood fill search algorithm
5. Use soft voting mechanism to vote for each block

Experiments: We train a model on virtual 3D scene dataset and use point cloud for post-processing offline. Some demonstrative examples can be shown below.



RGB Image

Semantic Labeling

Post-process

Reference: Xiang, Y., & Fox, D. (2017). DA-RNN: Semantic mapping with data associated recurrent neural networks. *arXiv preprint arXiv:1703.03098*.