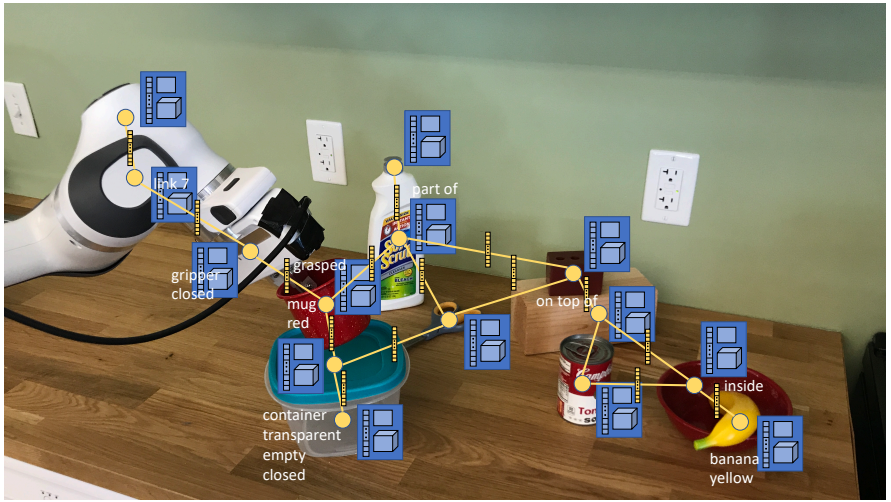
Graph Neural Networks for Multi-Object Perception and Manipulation

Tucker Hermans; University of Utah Dieter Fox; University of Washington

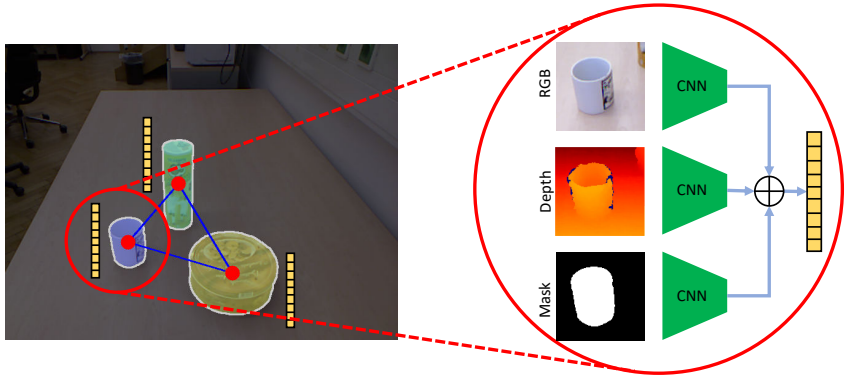
- Develop **Scene-GNNs**, a unified learning framework that provides robust solutions to key challenges in manipulation.
- Evaluate framework on segmentation, tracking, and multi-object attribute prediction from visio-tactile sensing.
- Plan using the framework for constrained and multi-step manipulation tasks.



Scene-GNN representing a manipulation scenario with objects, attributes, and relations.

Leverage Graph Neural Networks to Improve Unseen Object Instance Segmentation

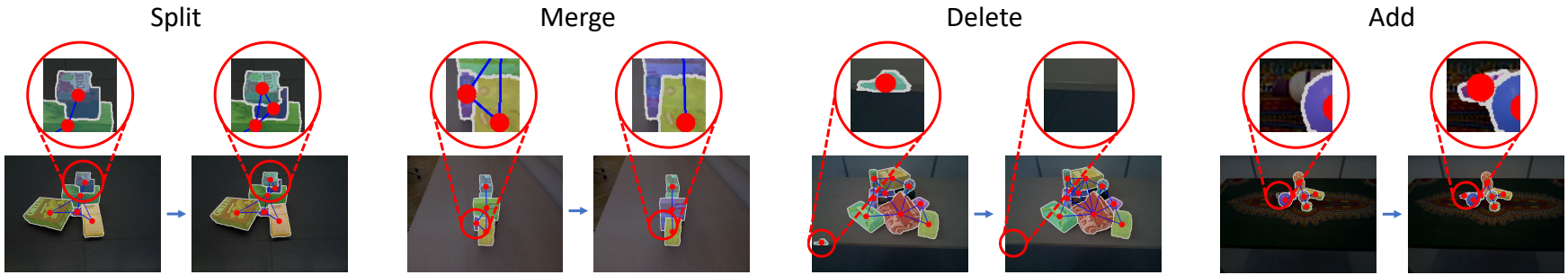
- Reasoning about object instances in cluttered environments is difficult.
- Can we represent such segmentations as graphs, and exploit the relational inductive bias that GNNs provide?



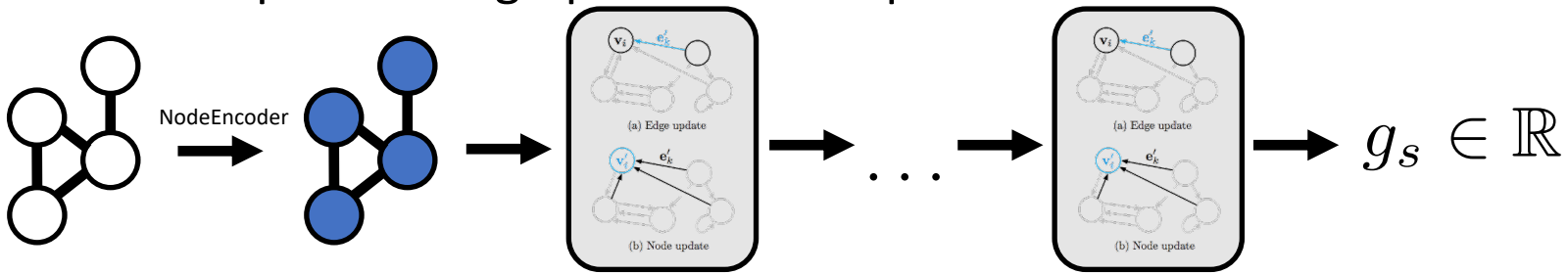
- Given initial segmentation, efficiently sample a tree of potentially better segmentations.

Algorithm Idea

- Efficiently sample a tree of potentially better graphs (segmentations)

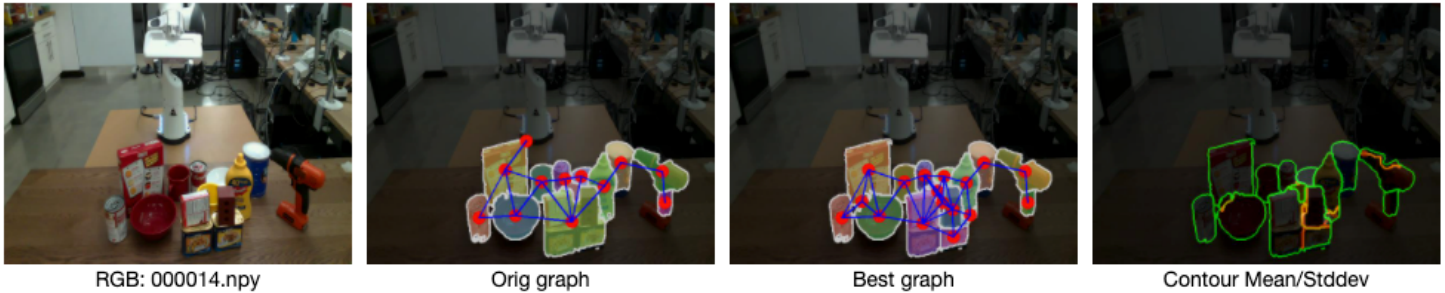


- Score the perturbed graphs with a Graph Neural Network



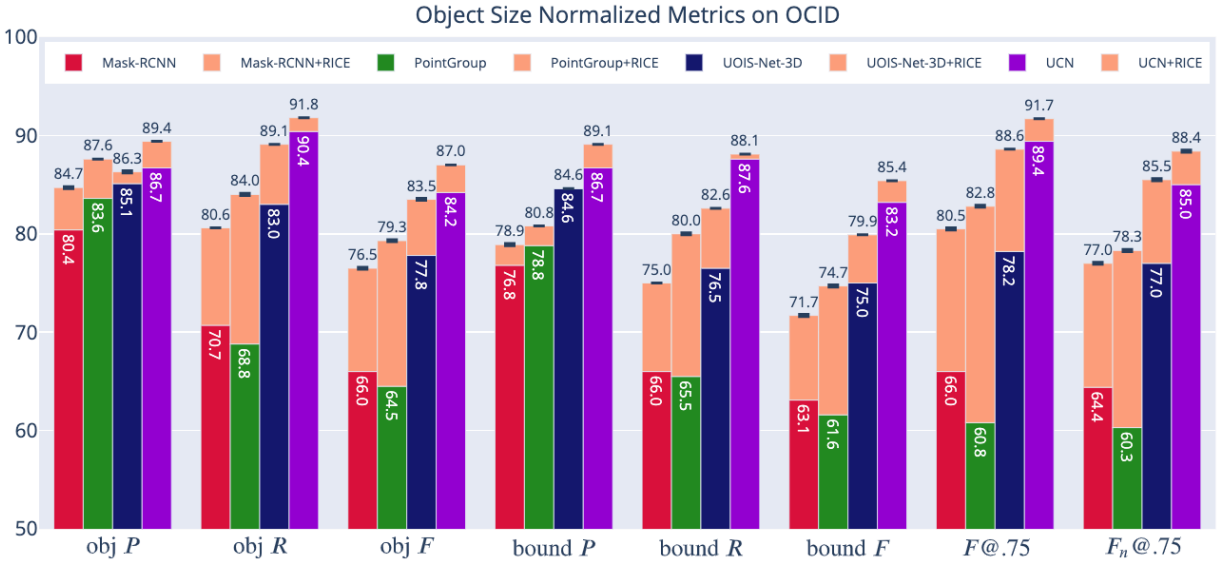
Example Results

- Given a sample tree, we can:
  1. Return the highest scoring segmentation graph
  2. Compute uncertainties of the segmentations
  3. Compute actions to resolve the uncertainty and better understand the scene.



Improvement over State-of-the-Art

- Reliably improves segmentations provided by other techniques.



- The F measures (the two metrics on far right) indicate that our method better identifies the correct number of objects.