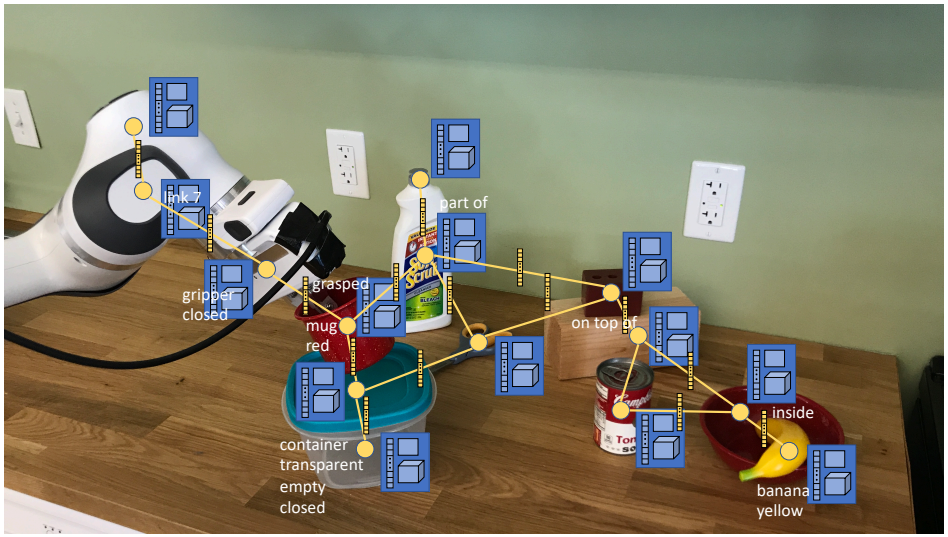
Graph Neural Networks for Multi-Object Perception and Manipulation

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
- Unknown object instance segmentation, tracking,



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

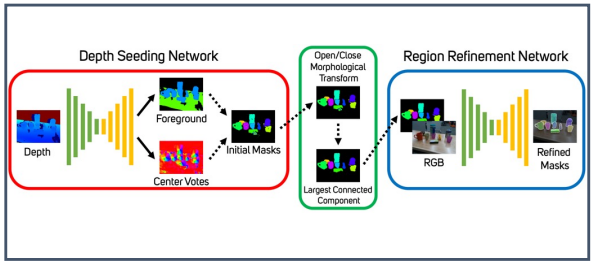
ROBOTICS CENTER
THE UNIVERSITY OF UTAH

PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

Unseen Object instance segmentation

Learn to Segment Objects in a Scene and Improve Segmentation via Graph Neural Networks

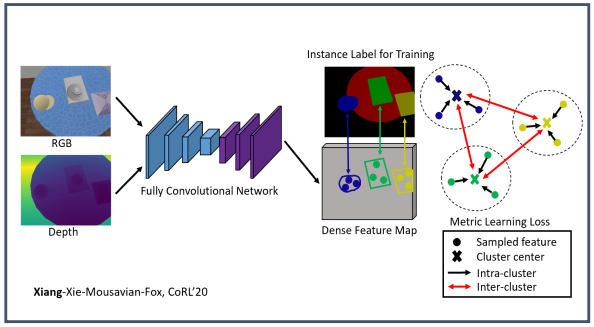
UOIS



- Foreground segmentation and 3D center voting on depth
- Refine segmentation masks using color information

[Xie-Xiang-Mousavian-Fox: CoRL 2019, T-RO-21]

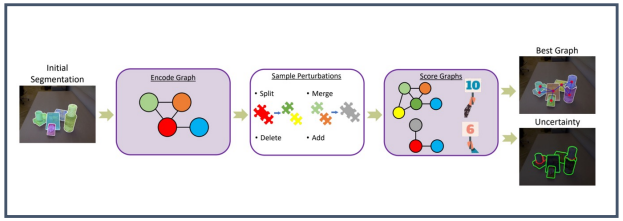
UCN



- Contrastive learning of object consistent features
- Mean shift clustering in feature space
- Repeat on individual masks for refinement

[Xiang-Xie-Mousavian-Fox: CoRL 2020]

RICE



- Convert initial segmentation into graph neural network
- Improve segmentation by sampling split/merge/add/delete operations
- Estimate segmentation uncertainties

[Xie-Xiang-Mousavian-Fox: CoRL-21]

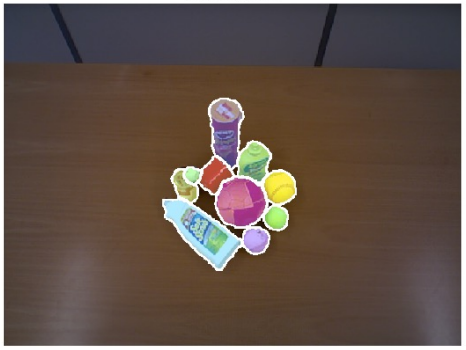
Input image



Feature map



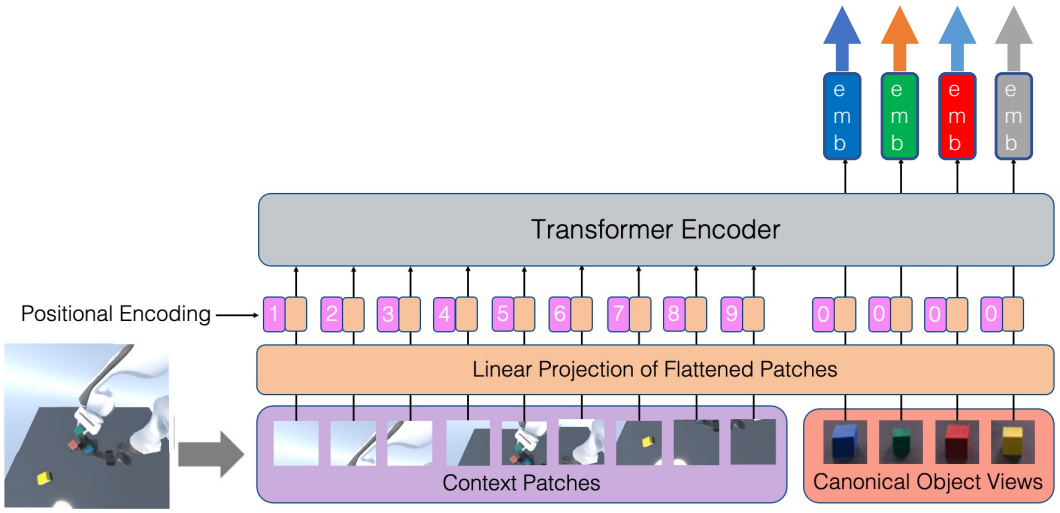
Segmentation



[Yuan-Paxton-Desingh-Fox: CoRL 2021]

SORNET: SPATIAL OBJECT-CENTRIC REPRESENTATION NETWORKS

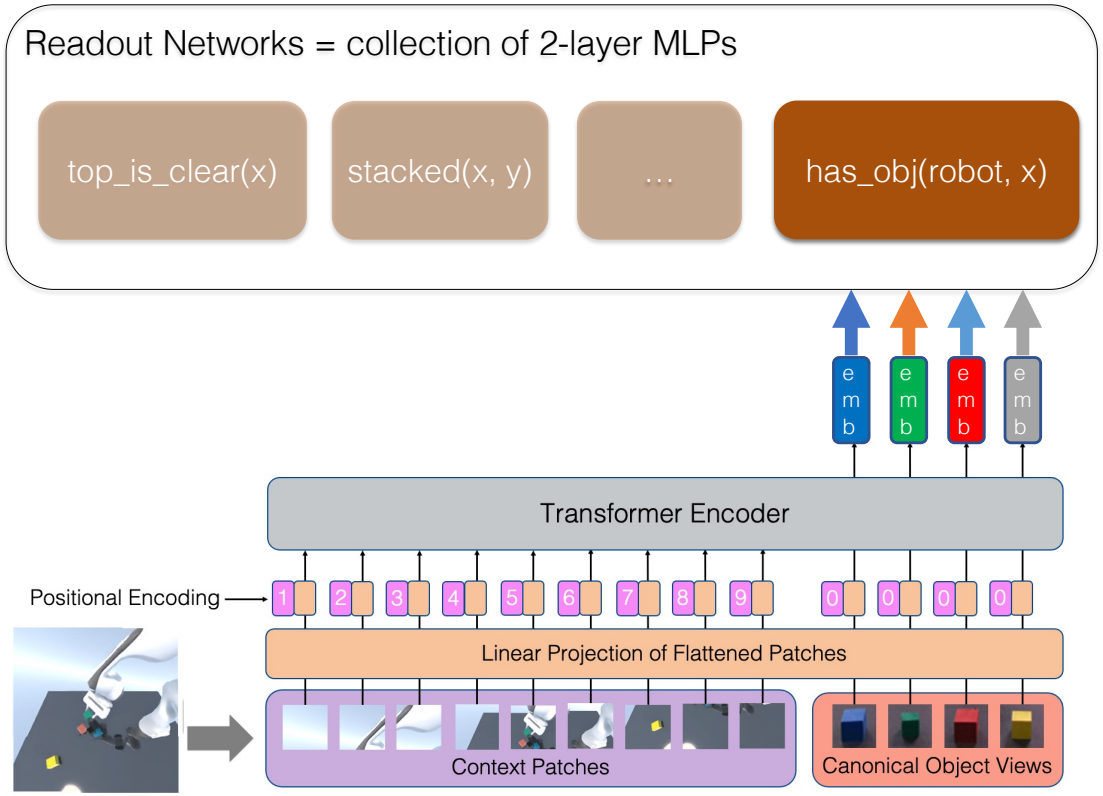
Learn Object-Centric Embeddings that Encode Spatial Relations



[Yuan-Paxton-Desingh-Fox: CoRL 2021]

SORNET: SPATIAL OBJECT-CENTRIC REPRESENTATION NETWORKS

Learn Object-Centric Embeddings that Encode Spatial Relations



Logical states

has_obj(robot,) = False

top_is_clear( ) = True

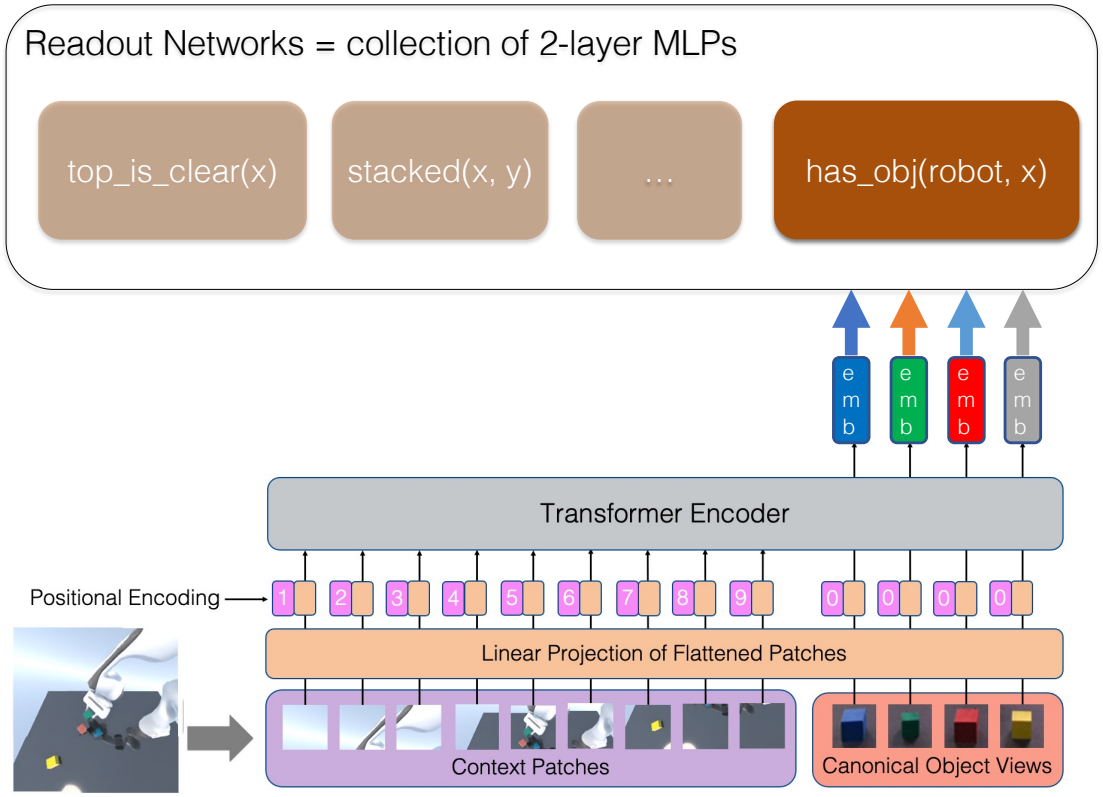
stacked(,) = False

…

[Yuan-Paxton-Desingh-Fox: CoRL 2021]

SORNET: SPATIAL OBJECT-CENTRIC REPRESENTATION NETWORKS

Learn Object-Centric Embeddings that Encode Spatial Relations



Logical states

has_obj(robot,) = False

top_is_clear( ) = True

stacked(,) = False

…