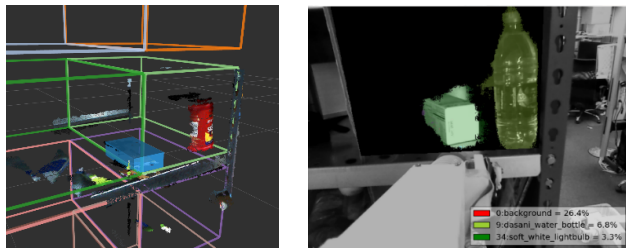


Find object in shelf-bin

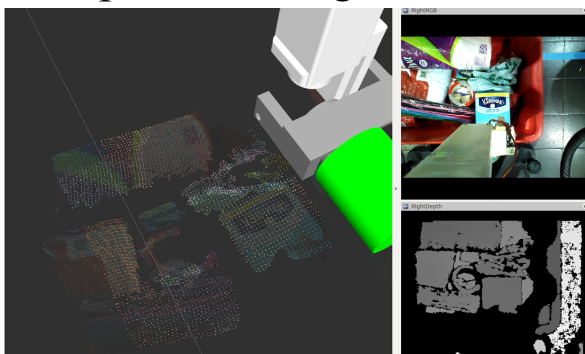
Bin box attention +
Fully convolutional networks



Manually annotated ~200 training data

Find object in tote

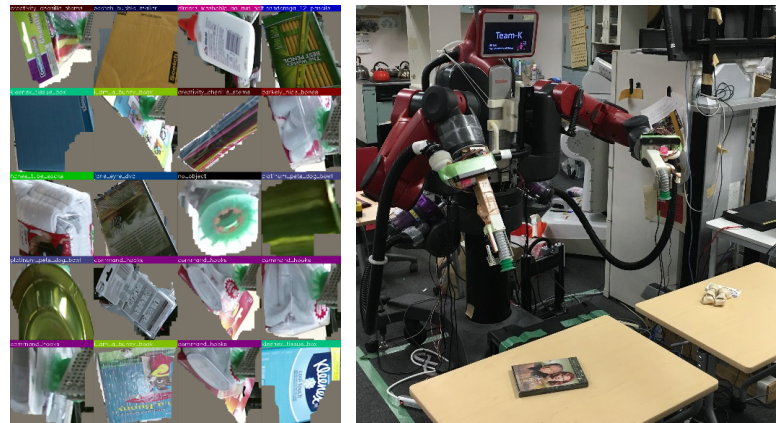
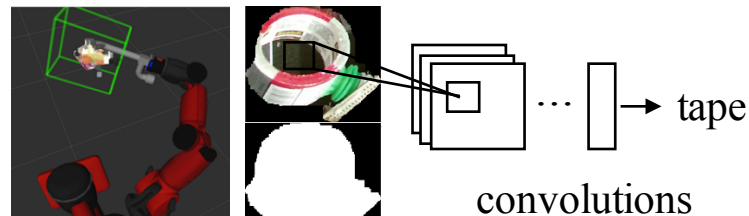
Tote box attention +
Super voxel segmentation



Grasp object

Verify/identify object in hand

In-hand box attention + VGG16 network



Autonomously obtained ~1K Training Data

Place/return object to tote/shelf-bin