NRI: Self-Supervised Object Detection and Visual Navigation Award # IIS 1925231,

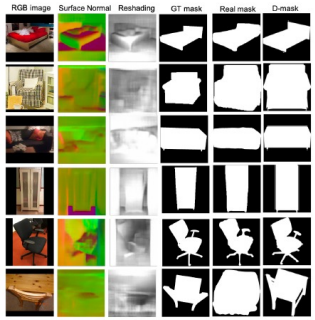
Jana Kosecka, George Mason University
http://cs.gmu.edu/~kosecka

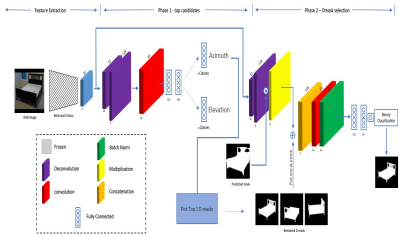
Challenge

1) Learning visual representations for navigation and object search in previously unseen environments
2) Zero–shot object pose estimation for Semantic SLAM
3) Self-supervised adaptation of pre-trained models for sematic segmentation

Task 2. Object Pose Estimation Using Mid-level Representations

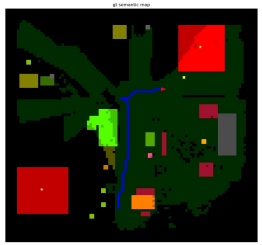
- Pose estimation for unseen instances of objects is challenging.
- Highly occluded objects in real-world indoor environments
- Train the pose estimation and object retrieval model on the top of mid—level visual representation
- Zero-shot transfer to to real-world unseen objects
- Competitive performance on low training data regime

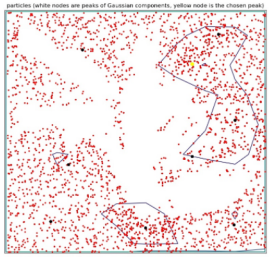




Architectures for retrieval and pose estimation

Task 1. Target driven navigation

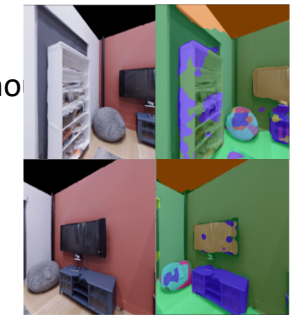




- Predict subgoal using object/room co-occurrence
- Short-range motion planning to reach the subgoal
- Learning semantic priors on the relative
- arrangement of objects and areas
- Select peaks in the belief space as sub-goals
- Improves SOTA on target driven navigation.
- Success 0.97% and 0.64 SPL on Matterport 3D object
- goal landmark

Task 3. Self-supervised adaptation of pre-trained models sematic segmentation

- Use temporal consistency as self-supervision signal for semantic segmentation



Broader Impacts

Improve robustness and enable better functionality of ho robots, semantic SLAM

New pose estimation benchmark

6337 objects' pose, and 3D bounding box are labeled

Students

Negar Nejatishahidin nnejatis@gmu.edu
Yimeng Li yli44@gmu.edu, Sulabh Shresta sshres2@gmu.edu