SNAP: Towards Segmenting aNything in Any Point Cloud

Aniket Gupta[1], Hanhui Wang[1], Charles Saunders[2], Aruni RoyChowdhury[2], Hanumant Singh[1], Huaizu Jiang[1]

[1]Northeastern University [2]The Mathworks Inc.

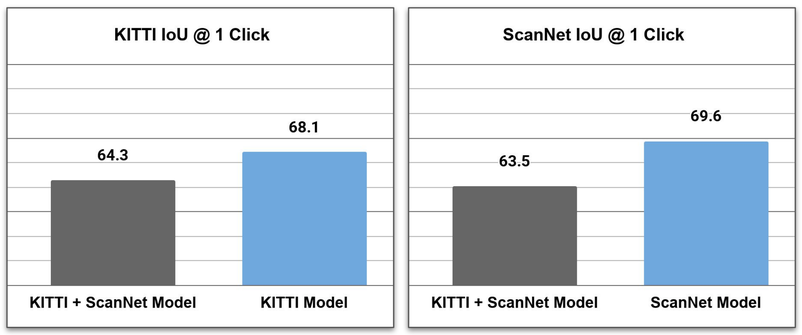
THE PROBLEM

High Variability in 3D data

Point Clouds come from diverse sources including LiDAR, Stereo, SfM etc and exhibit massive variability in point density, modality and ranges.

Prior works on interactive point cloud segmentation operate on point clouds from a single domain, i.e. either Indoor point clouds or Outdoor point clouds. For new domains, we need to train a completely new model from scratch.

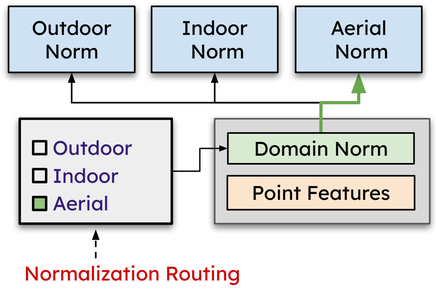
Negative Transfer is the bottleneck!



Naive pooling degrades performance due to significant statistical divergence between datasets. - we call this "negative transfer"

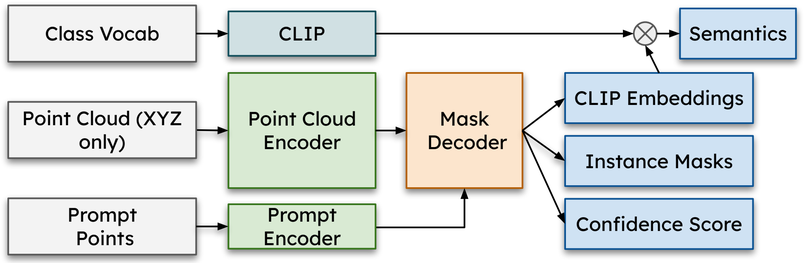
SNAP Framework

Domain Norm

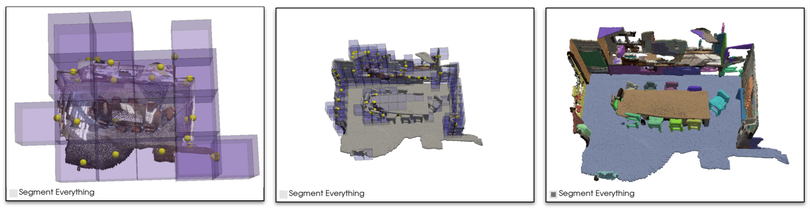


SNAP is trained with Domain-specific normalization, this mitigates negative transfer and leads to better zero-shot results.

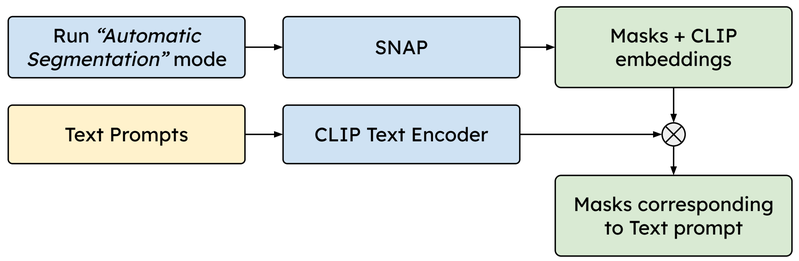
SNAP can automatically label masks



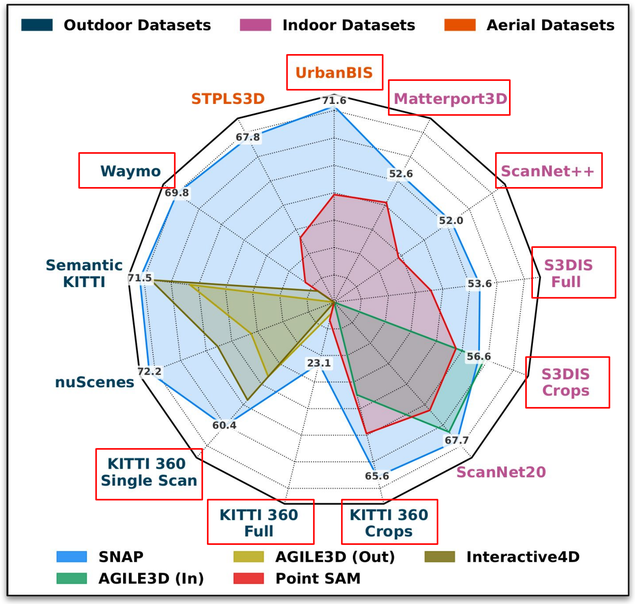
Automatic Segmentation



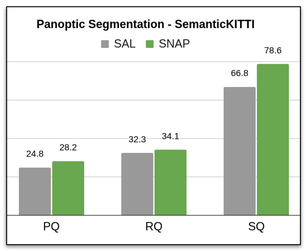
Supports both spatial and textual prompts

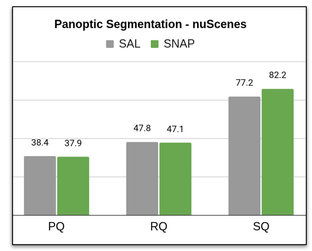


Results

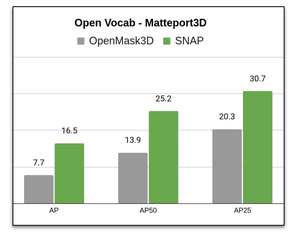


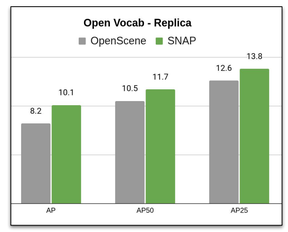
SNAP achieves an average of **20.6%** improvement in IoU@1 click experiments, and achieves a new SOTA on 8/9 zero-shot benchmarks using only a *single set of weights*





SNAP outperforms SAL by 3.4% on SemanticKITTI and remains comparable on nuScenes dataset





SNAP outperforms image-augmented models on Matterport3D and non-visual baselines like OpenScene on Replica.