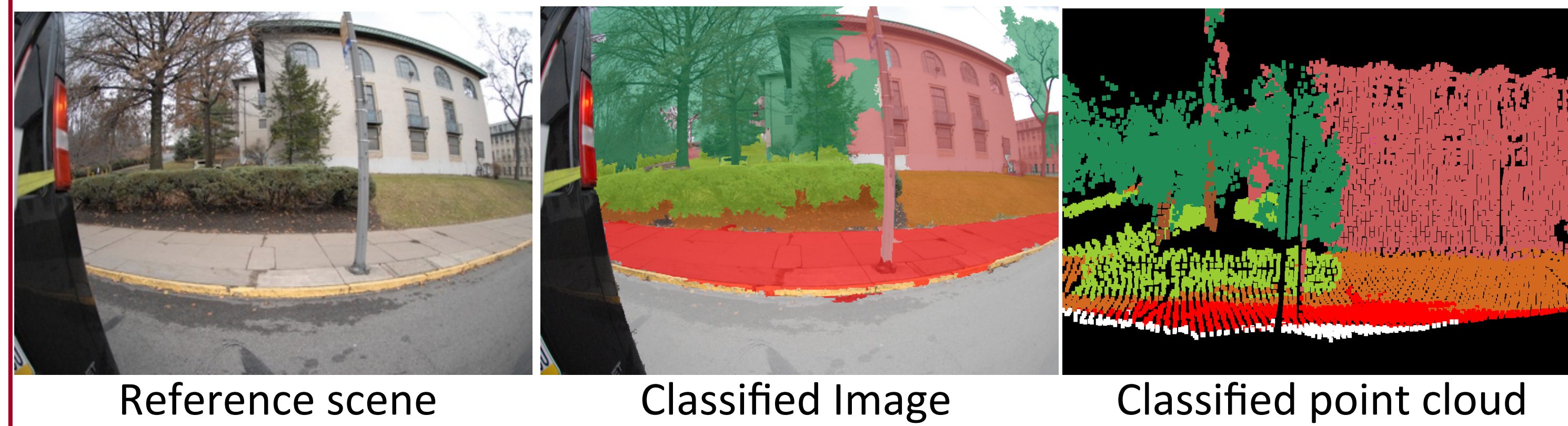
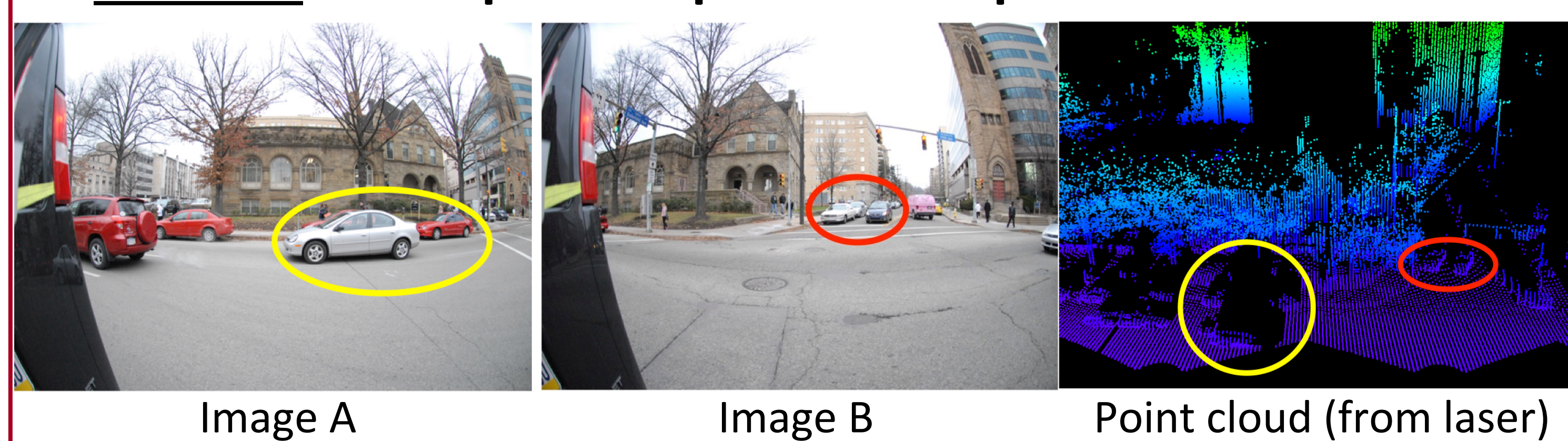


Multi-modal Scene Understanding

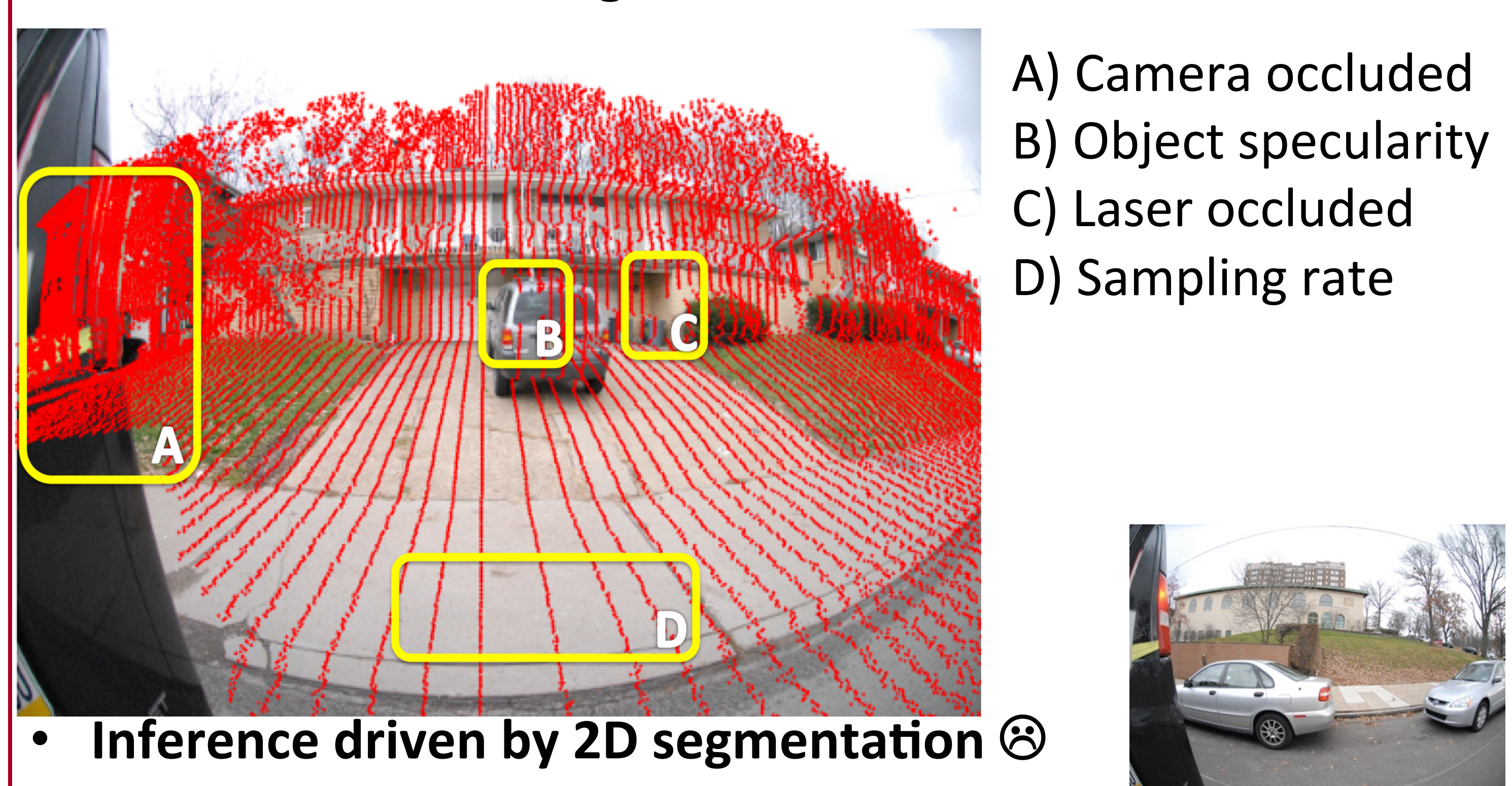
- Joint image and 3D point cloud scene understanding



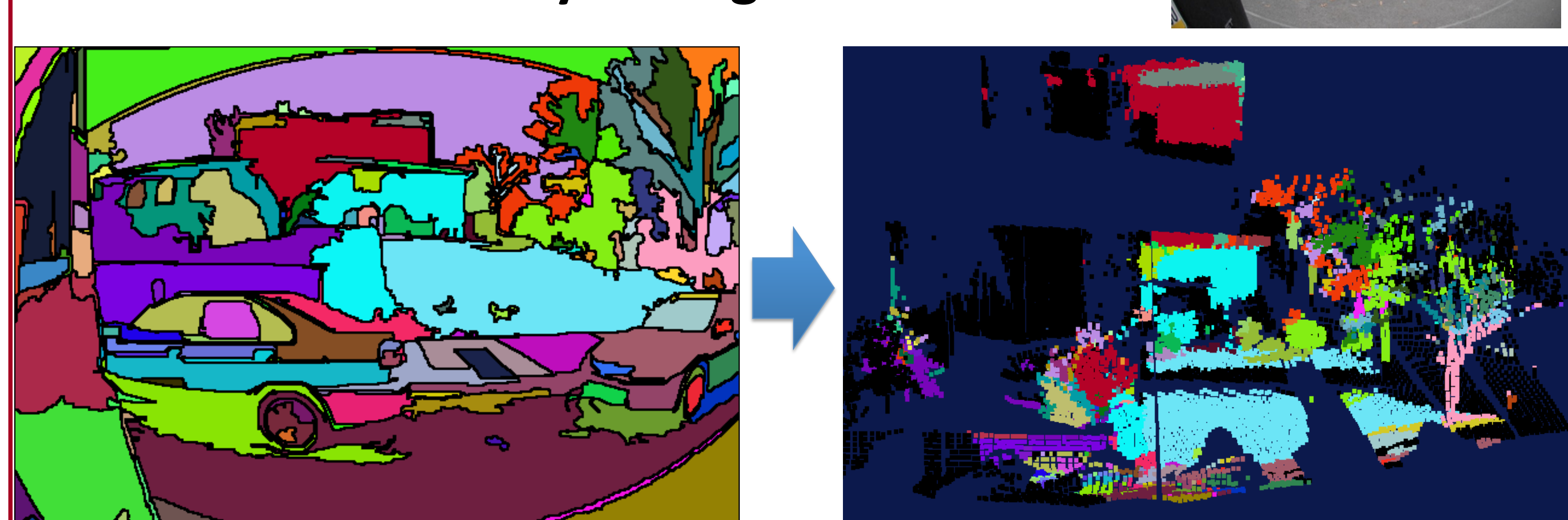
- Without 1-to-1 pixel-to-point correspondences



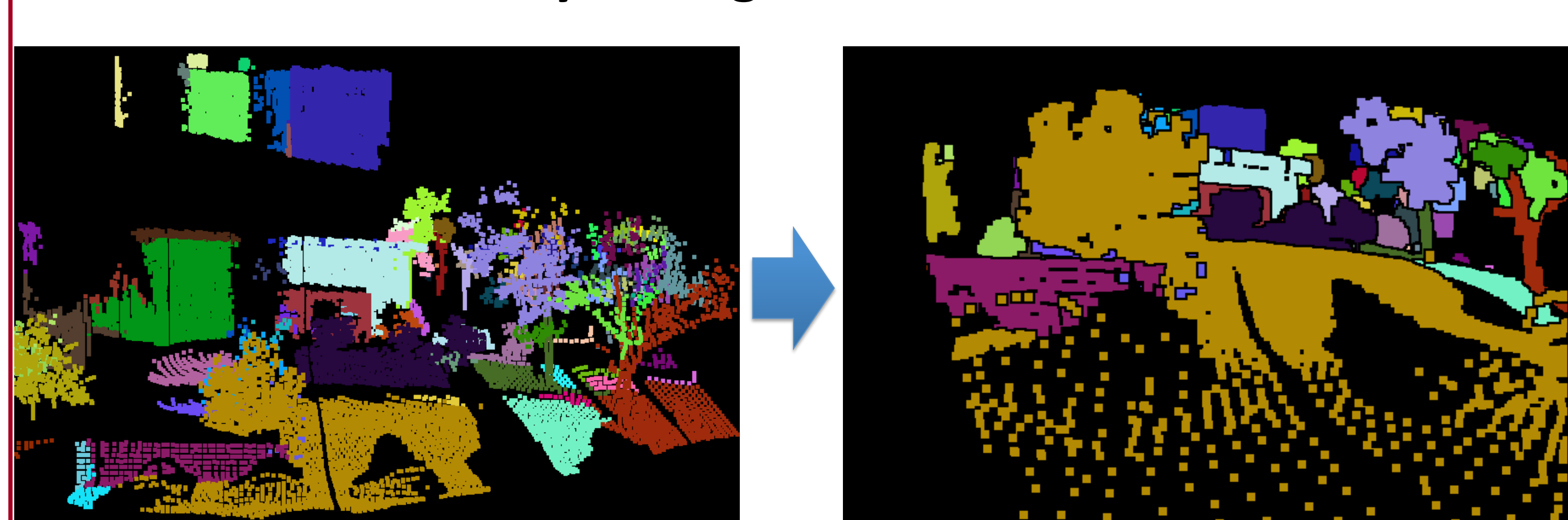
- Fundamental challenges



- Inference driven by 2D segmentation ☹️

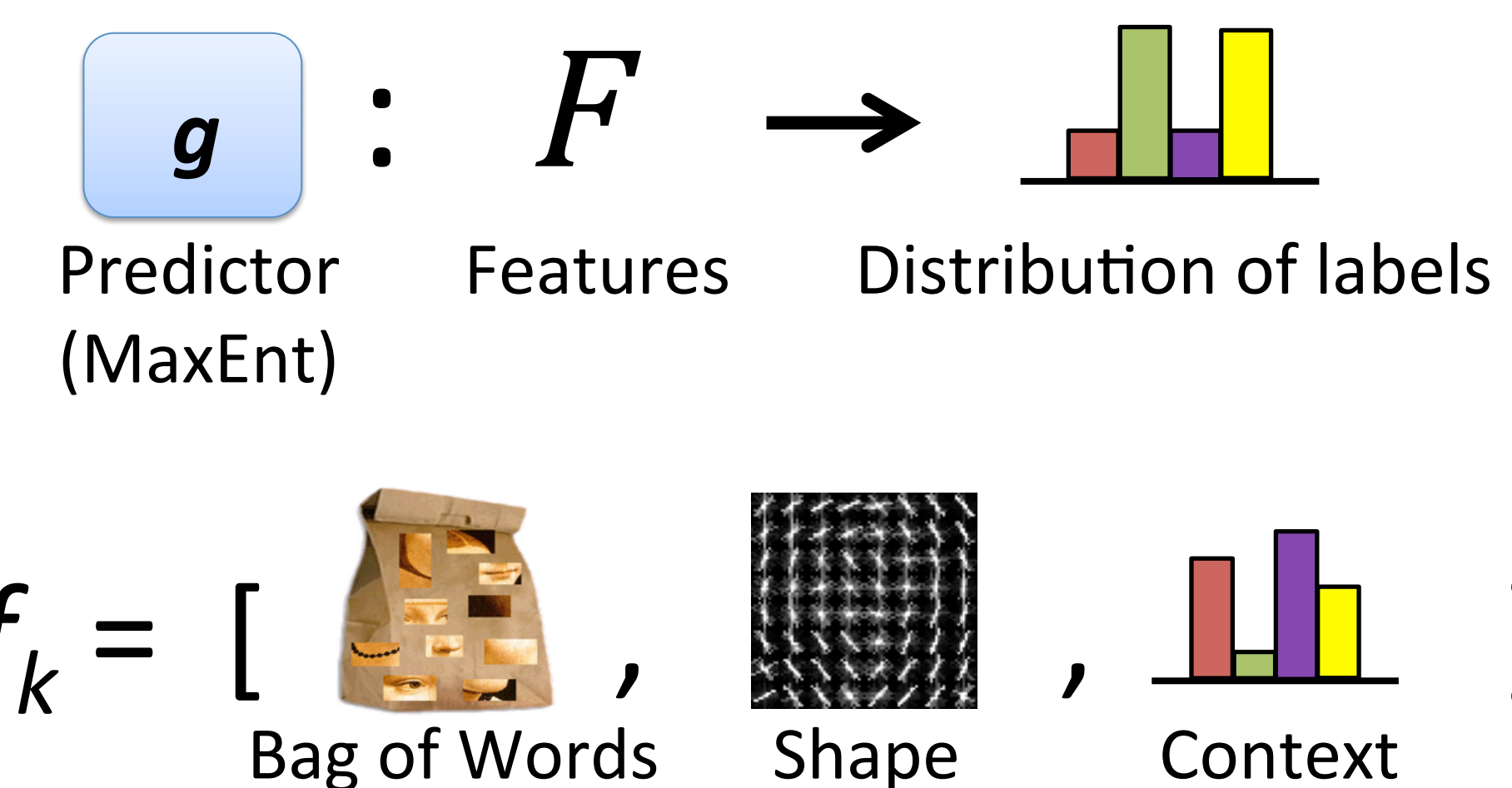
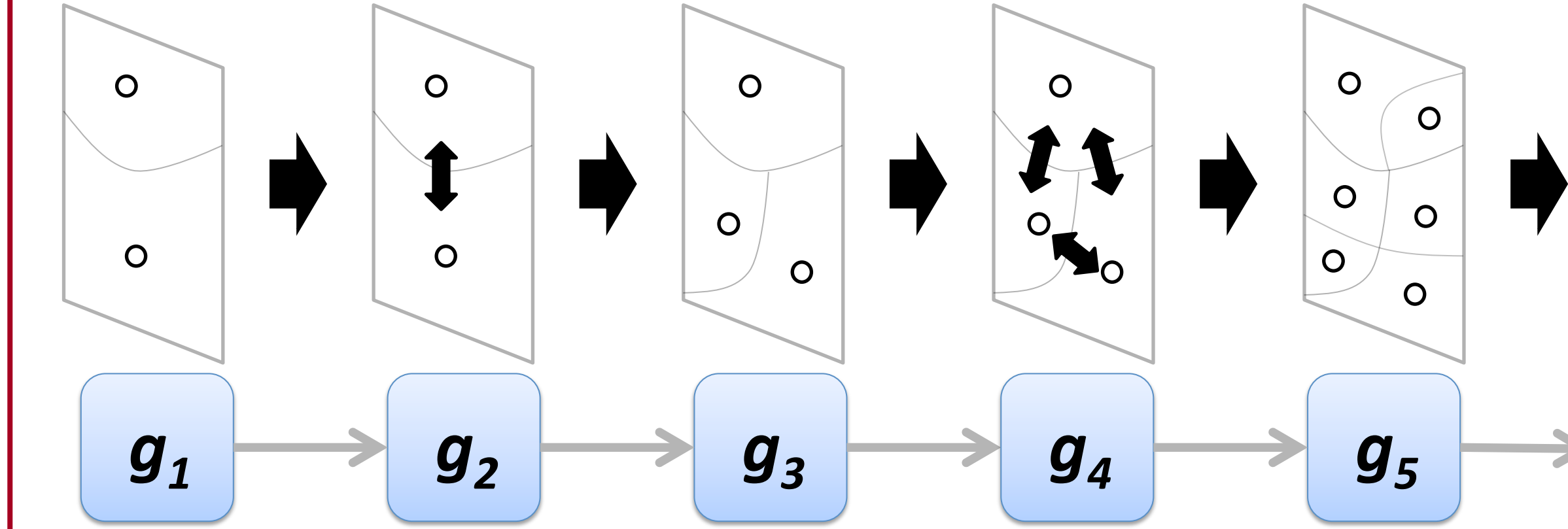


- Inference driven by 3D segmentation ☹️

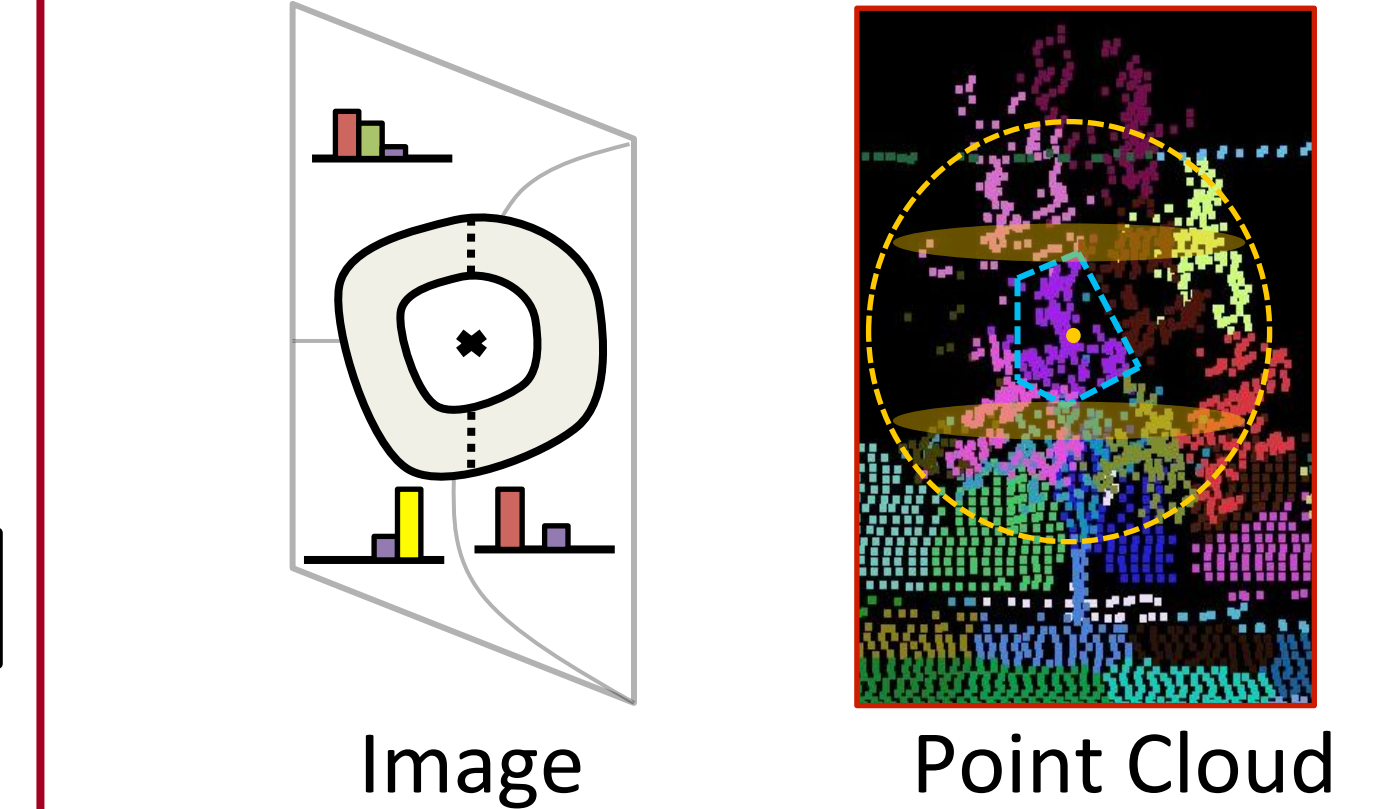


Inference Machines [1] [2]

- Structured prediction as a sequence of predictions

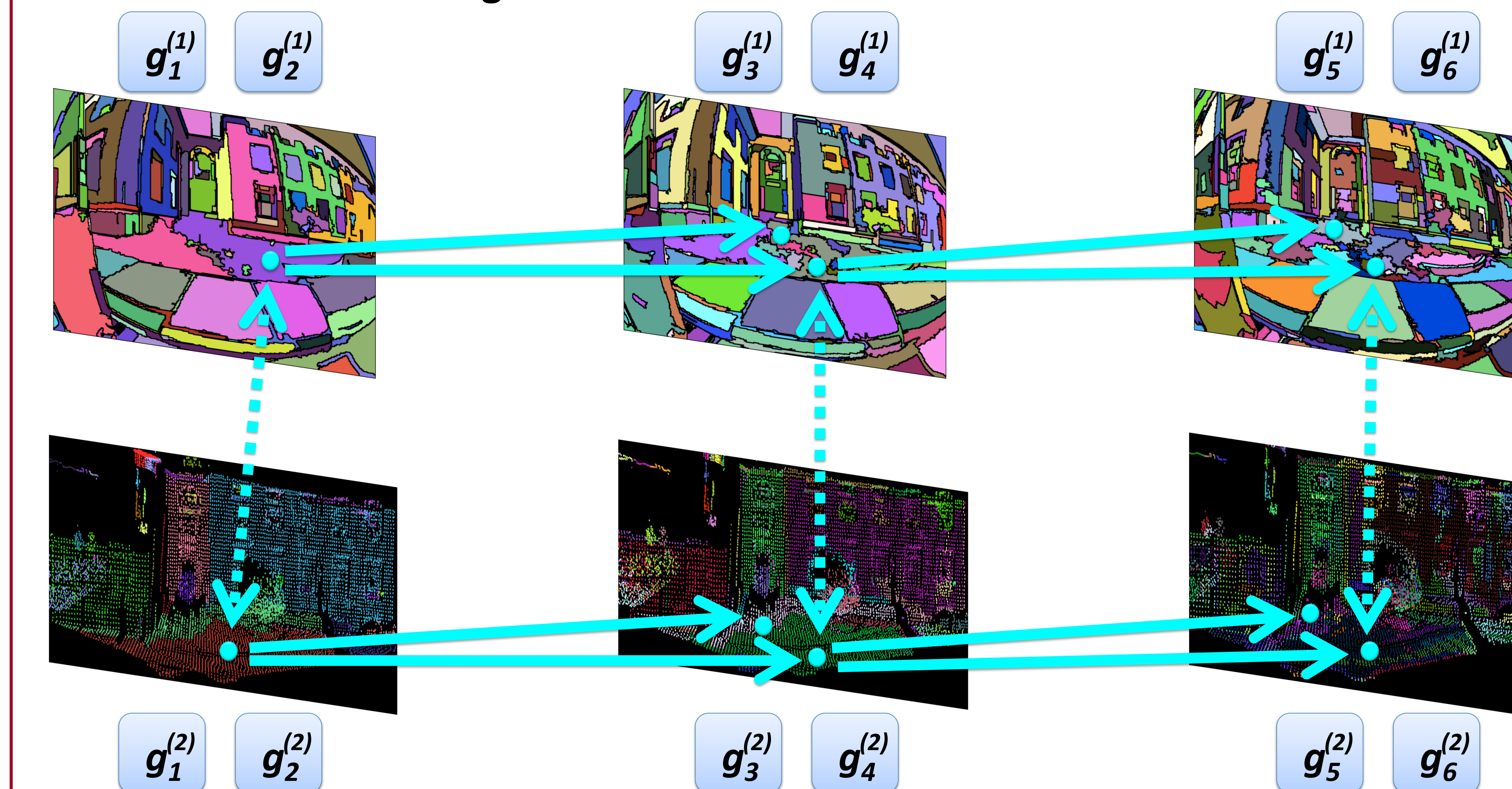


- Contextual information



Co-Inference Machines

- Joint inference over region hierarchies in each domain



- Feature for region x_k in domain i

$$\tilde{f}_k^{(i)} = [f_k^{(i)}, \text{Context from domain } j]$$

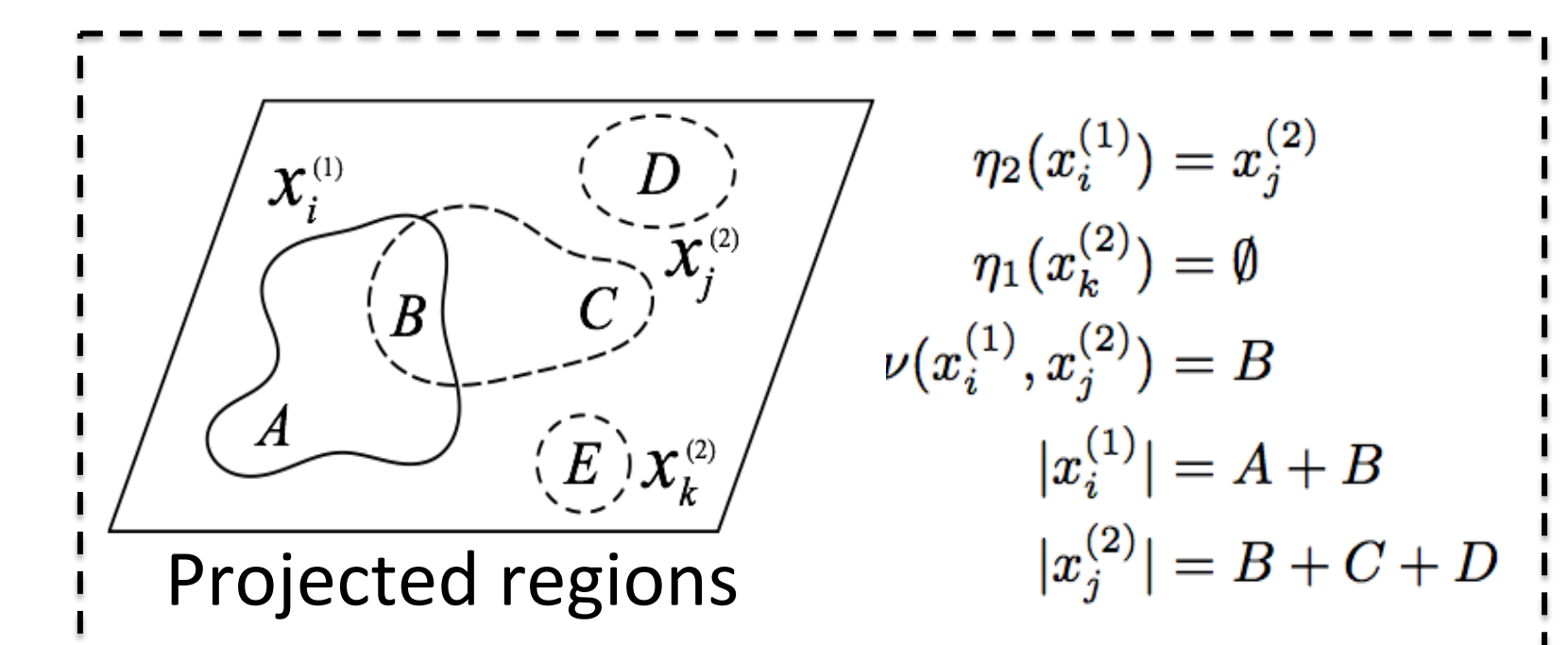
- Context from soft correspondences

$$\sum_{x_l^{(j)} \in \eta_j(x_k^{(i)})} \frac{\nu(x_k^{(i)}, x_l^{(j)})}{|x_k^{(i)}|} [b_{l,t-1}^{(j)}, 1]$$

Normalized overlap

Neighbors in domain j

$x_l^{(j)}$ label distribution from $g_{t-1}^{(j)}$



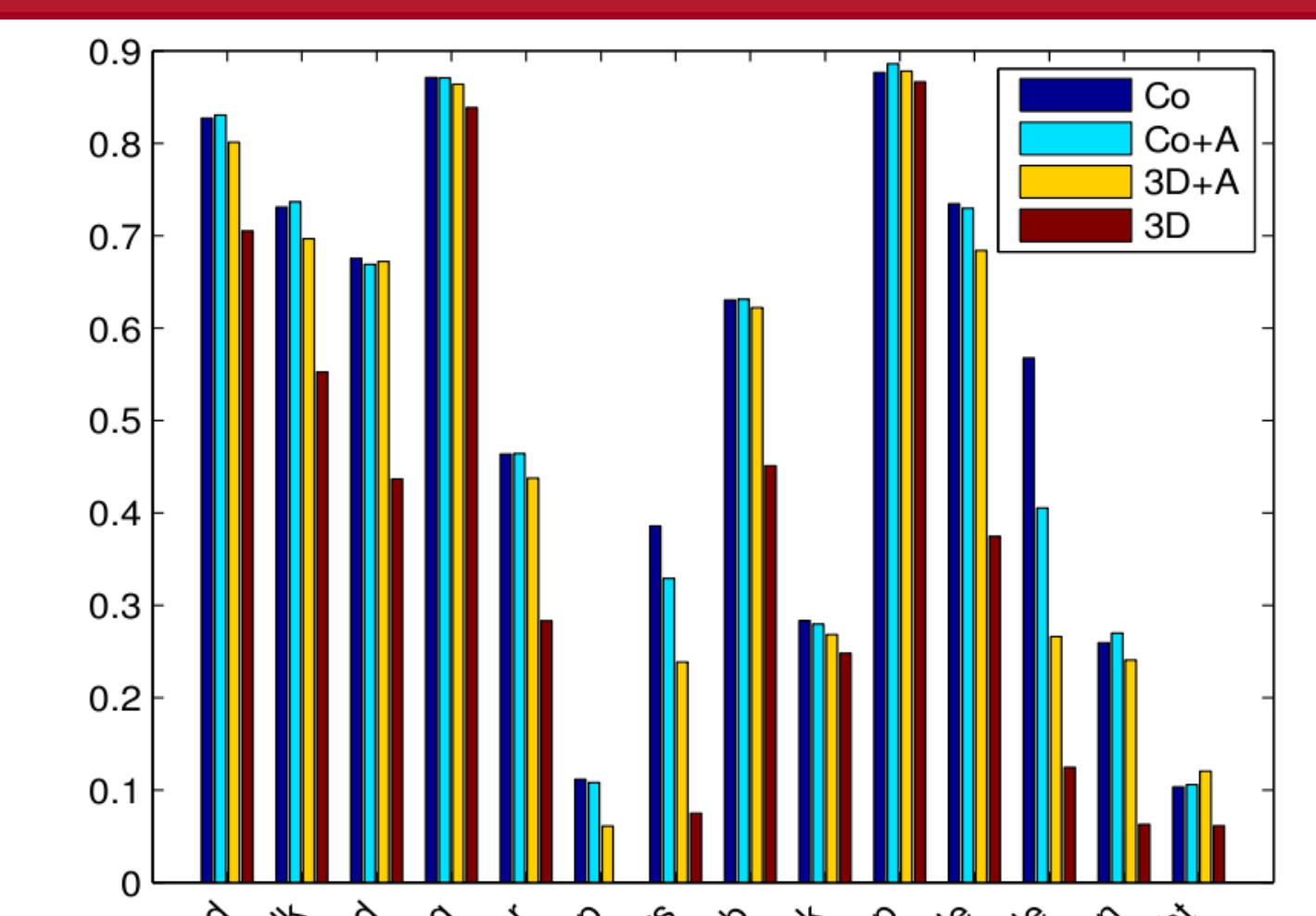
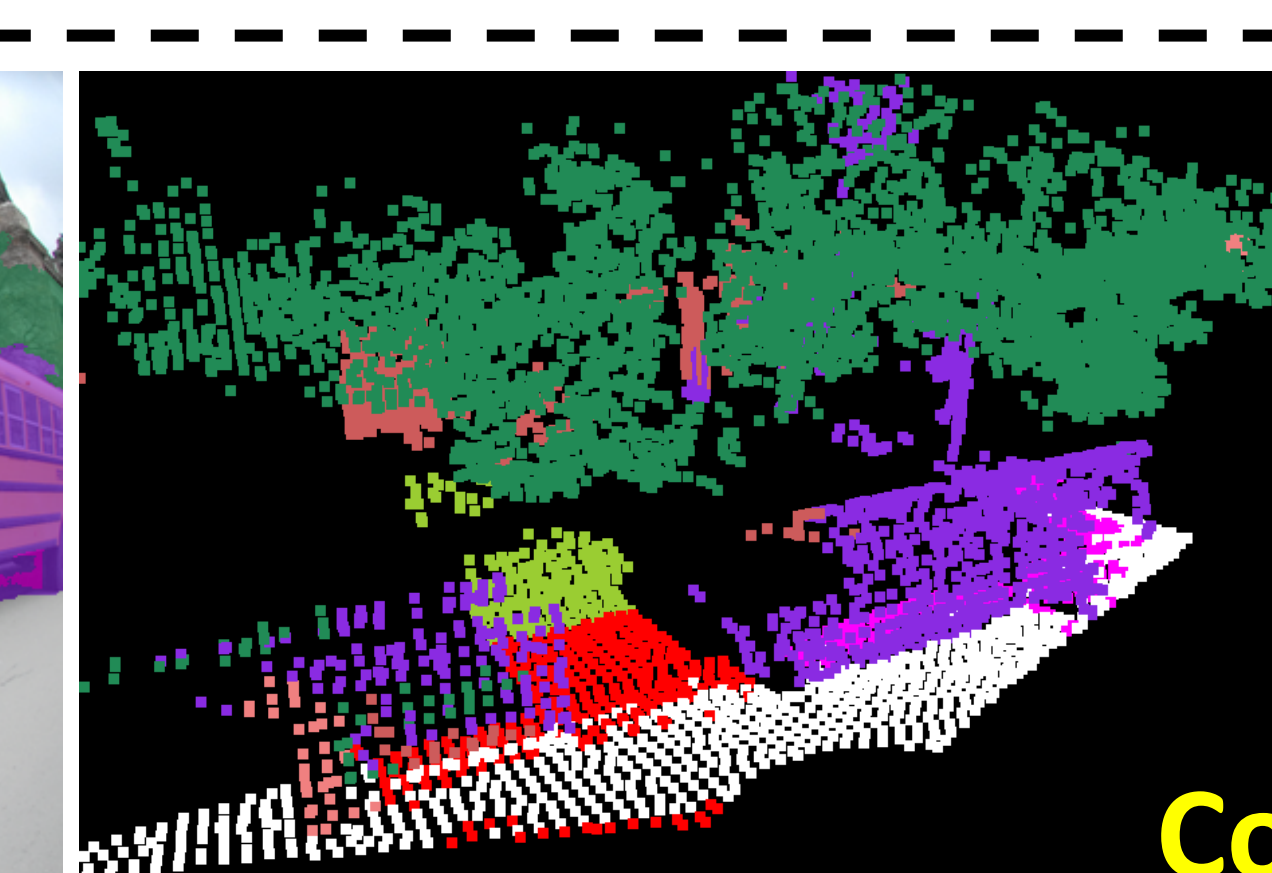
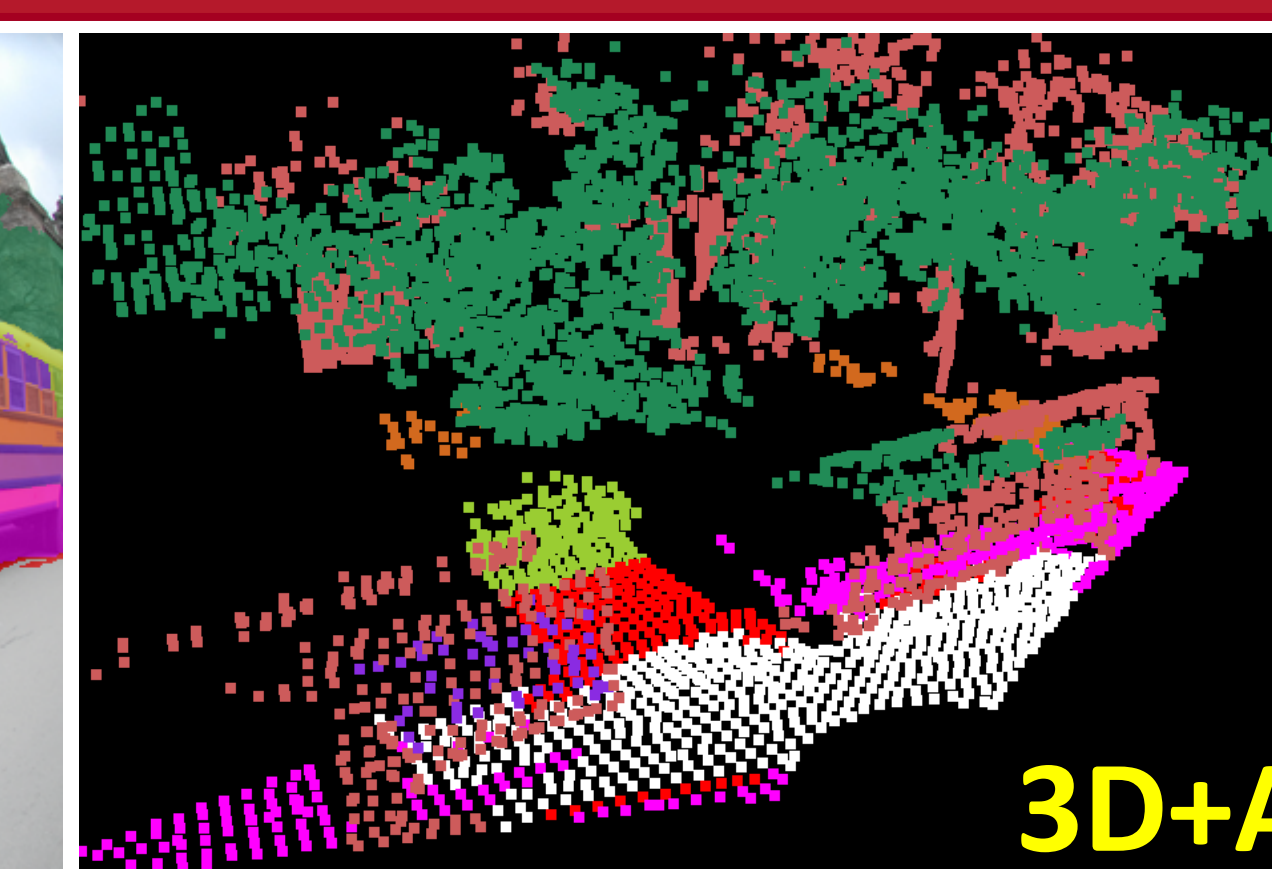
Experiments

- Dataset

- Camera + laser line scanner
- 29 semantic categories

- Comparisons

- 2D: Only image regions [1]
- 3D: Only point cloud regions [2]
- 2D+A: Augment w. 3D features
- 3D+A: Augment w. 2D features
- Co: Co-inference (this work)
- Co+A: Augment w. 2D/3D features



Label	Co	2D+A	Diff.	Label	Co	3D+A	Diff.
Barrier	.509	.521	-.012	Road	.827	.802	.026
Bus-stop	.163	.138	.025	Sidewalk	.731	.697	.034
Stairs	.339	.297	.042	Barrier	.464	.438	.026
Small-vehicle	.844	.825	.019	Bus-stop	.112	.061	.051
Person	.502	.391	.111	Stairs	.386	.239	.147
Post	.474	.465	.010	Tree-trunk	.284	.268	.015
Tall-light	.020	.005	.015	Small-vehicle	.735	.684	.051
Wire	.097	.072	.025	Big-vehicle	.568	.266	.302
Traffic-signal	.178	.282	-.104	Person	.260	.241	.019
				Tall-light	.103	.120	-.017

[1] D. Munoz, J. A. Bagnell, M. Hebert, *Stacked Hierarchical Labeling*, ECCV 2010

[2] X. Xiong, D. Munoz, J. A. Bagnell, M. Hebert, *3-D Scene Analysis via Sequenced Predictions over Points and Regions*, ICRA 2011