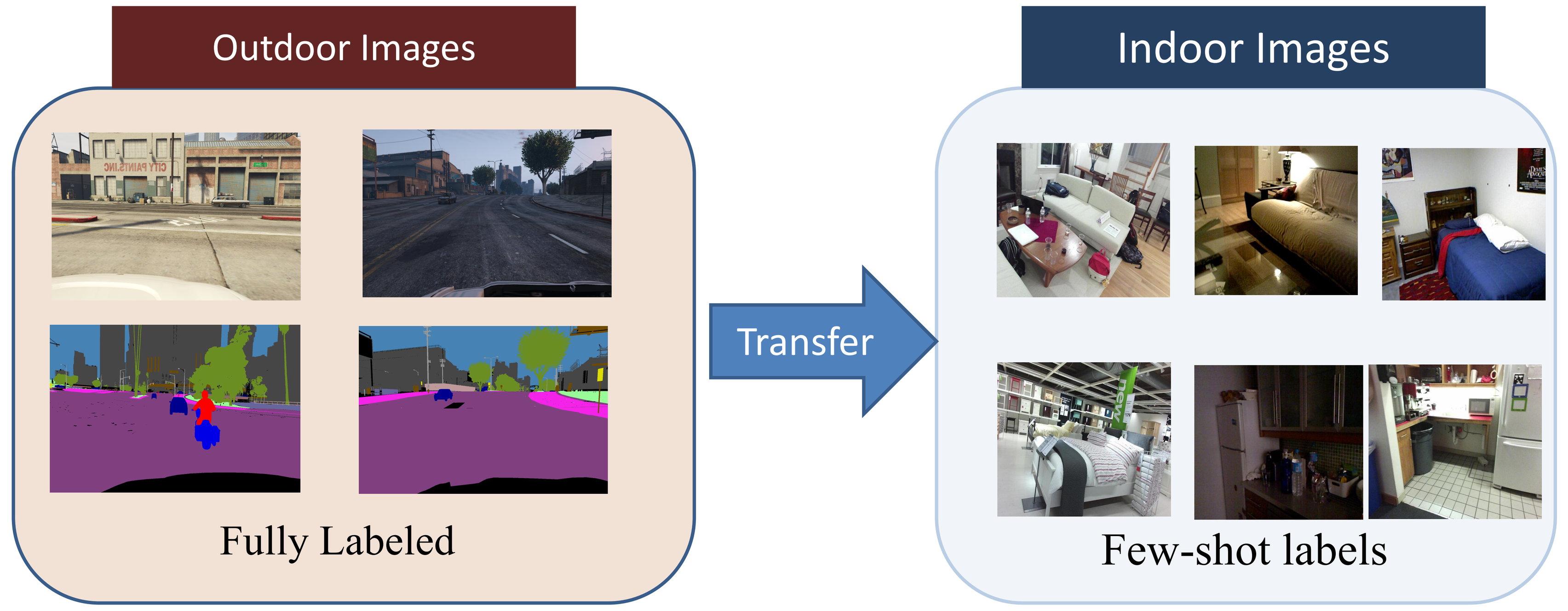
Cluster-to-adapt: Few Shot Domain Adaptation for Semantic Segmentation across Disjoint Labels

Tarun Kalluri *Manmohan Chandraker*

UC San Diego

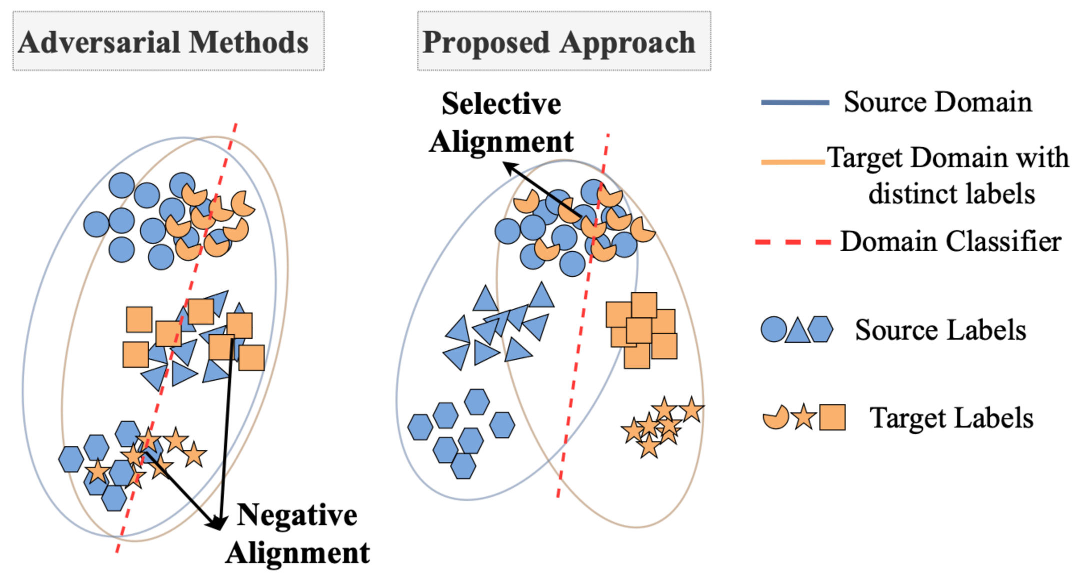
Domain Adaptive Semantic Segmentation

- Transfer a trained model for segmentation from a labeled source domain, such as synthetic images, to an unlabeled target domain, such as real images.
- Synthetic images are easier and cheaper to collect and label compared to real images.
- Existing adaptation works cannot transfer between two domains with *disparate label spaces.*
  - Example: Outdoor scenes to indoor scenes.



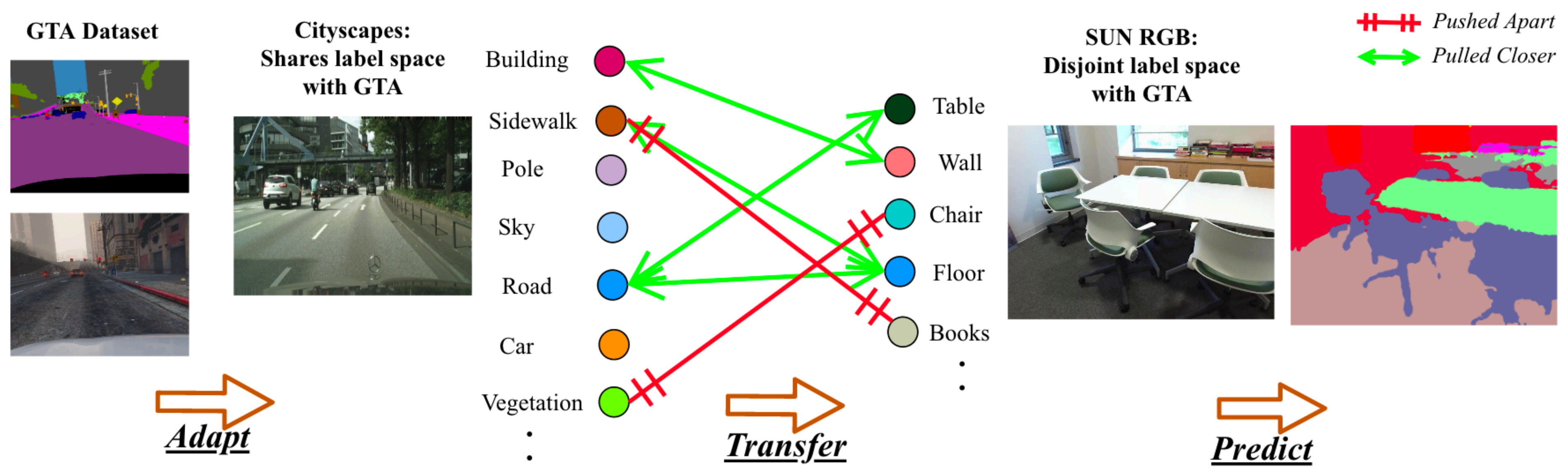
- We consider the problem of **domain adaptation across disjoint labels** with few labels in the target.

Global Domain Alignment vs. Selective Alignment

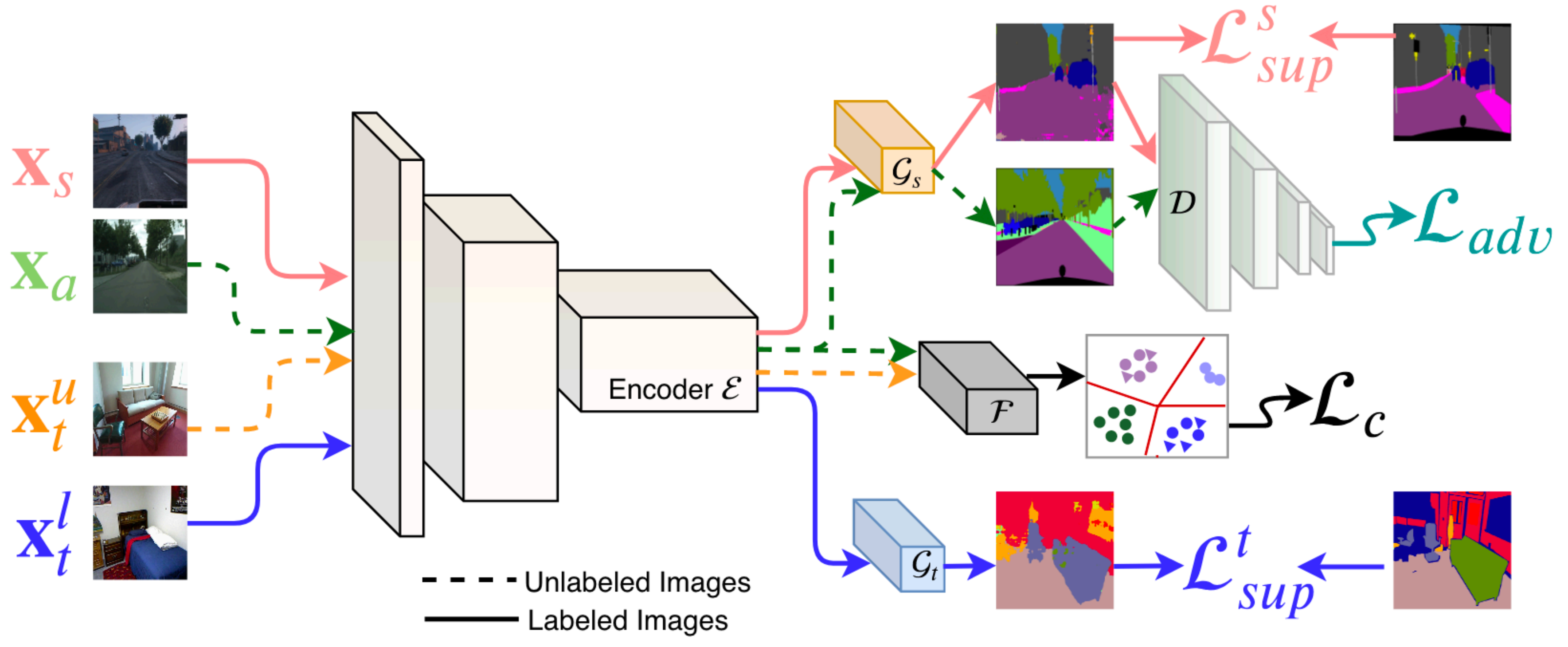


- Global alignment based may lead to **negative transfer** where different categories from source and target align post adaptation.

Using Unlabeled Bridges to Ease the Adaptation



Joint Training Using Constrained Clustering Objective



- Clustering loss groups the shared properties from source and target features into distinct groups, while preventing negative alignment.
- $p(\mu_k|v_j)$ is the probability that a feature v_j belongs to a cluster with center μ_k.

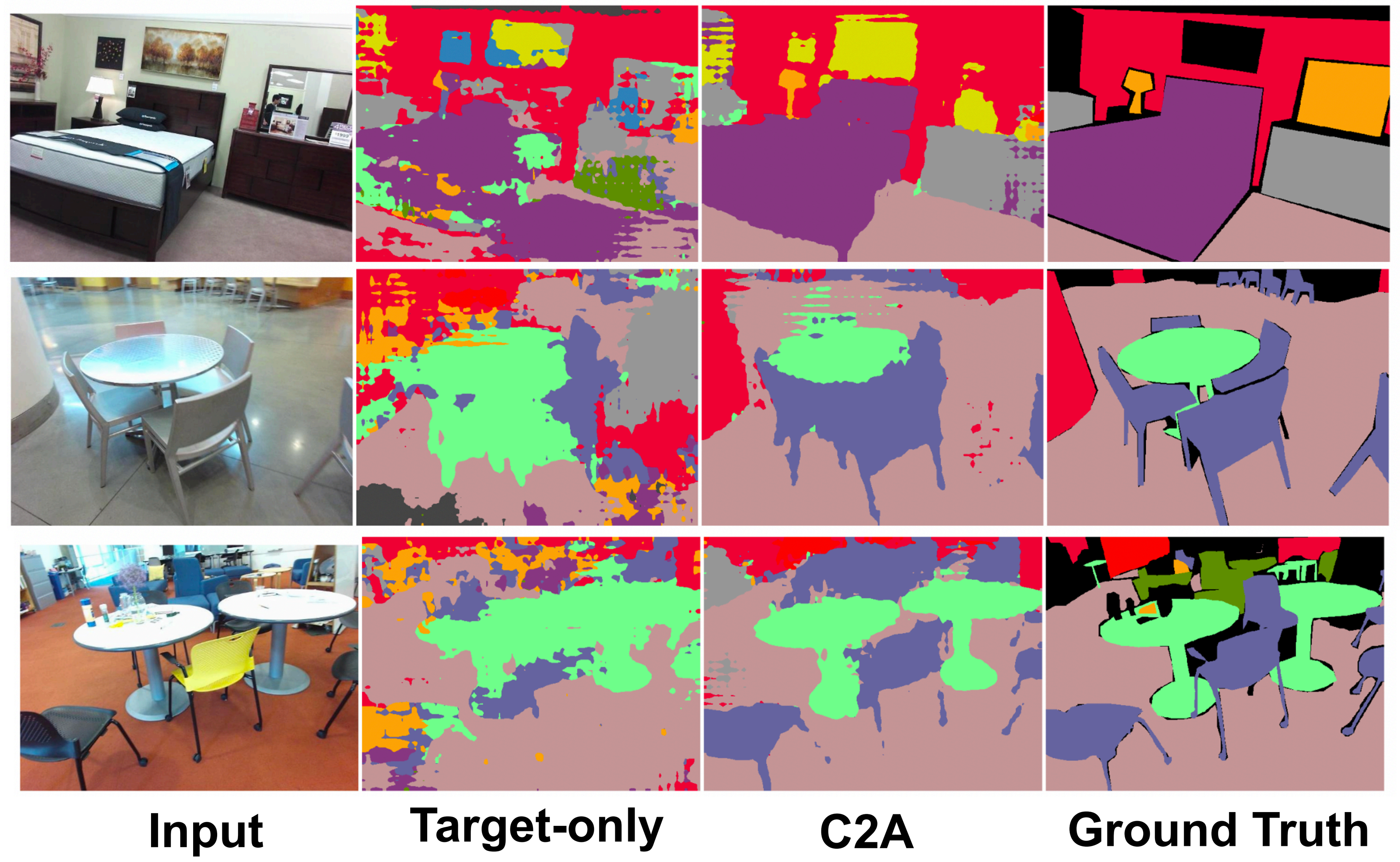
$$\mathcal{L}_c = \sum_{x\in\{\mathbb{D}_a,\mathbb{D}_t^u\}} \sum_{v_j\in\mathcal{F}(\mathcal{E}(x))} -\log(\max_k p(\mu_k|v_j))$$

$$p(\mu_k|v_j) \propto \exp\left(\frac{v_j\cdot\mu_k}{||v_j||_2||\mu_k||_2}\right)$$

Total Loss:

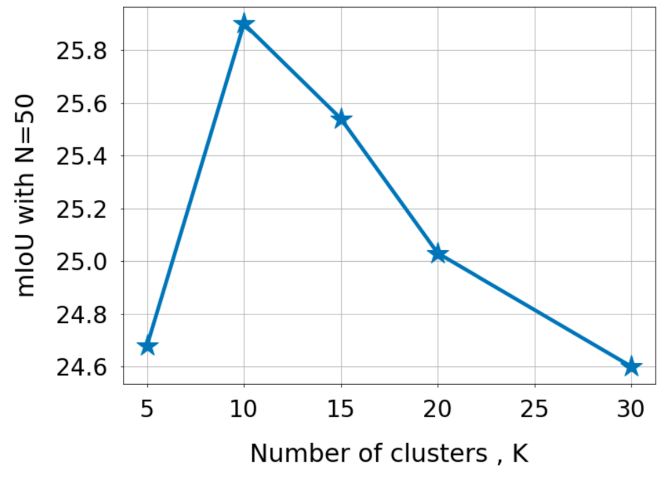
$$\arg\min_{\mathcal{E},\mathcal{G}_s,\mathcal{G}_t,\mathcal{F}} \mathcal{L}_{sup} + \lambda_{adv}\mathcal{L}_{adv} + \lambda_c(\mathcal{L}_c+\mathcal{L}_{kl})$$

Results on Few Shot Adaptation



- Few-shot training does not capture object boundaries well.
- C2A leverages larger scale synthetic data to improve segmentation.

	$\sigma = 0.01$ N = 50	$\sigma = 0.04$ N = 200	$\sigma = 0.1$ N = 500	$\sigma = 0.3$ N = 1500
Target only	22.62	30.43	36.62	43.17
Adapt SegNet	25.20	32.51	36.90	43.83
LET	25.19	32.44	35.87	42.96
UnivSeg	22.21	31.32	36.08	42.10
Adv SemiSeg	24.72	33.22	**38.46**	**45.10**
C2A [Ours]	**25.98**	**33.37**	37.41	43.16



Limitations of our method

- The gains from C2A when enough target supervision is available are limited.
- The clustering is very noisy, and the classes aligned are related only sometimes.