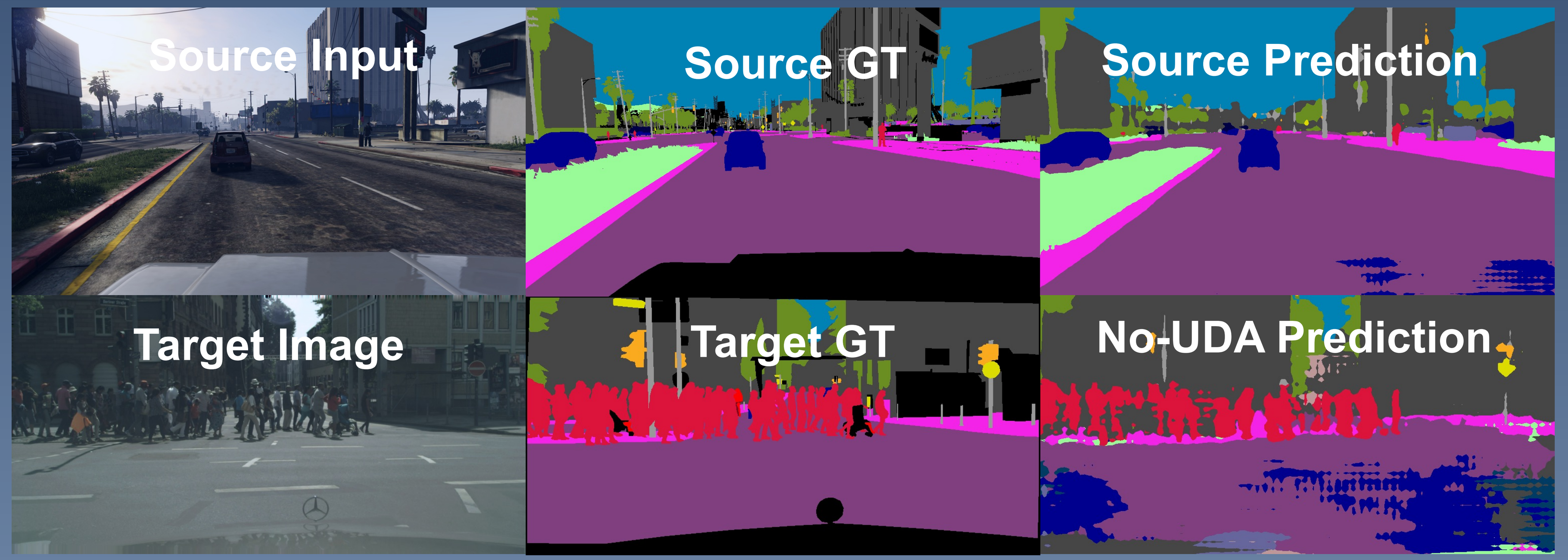
LSR: Latent Space Regularization for Unsupervised Domain Adaptation in Semantic Segmentation

Francesco Barbato, Marco Toldo, Umberto Michieli and Pietro Zanuttigh – University of Padova

CVPR VIRTUAL JUNE 19-25

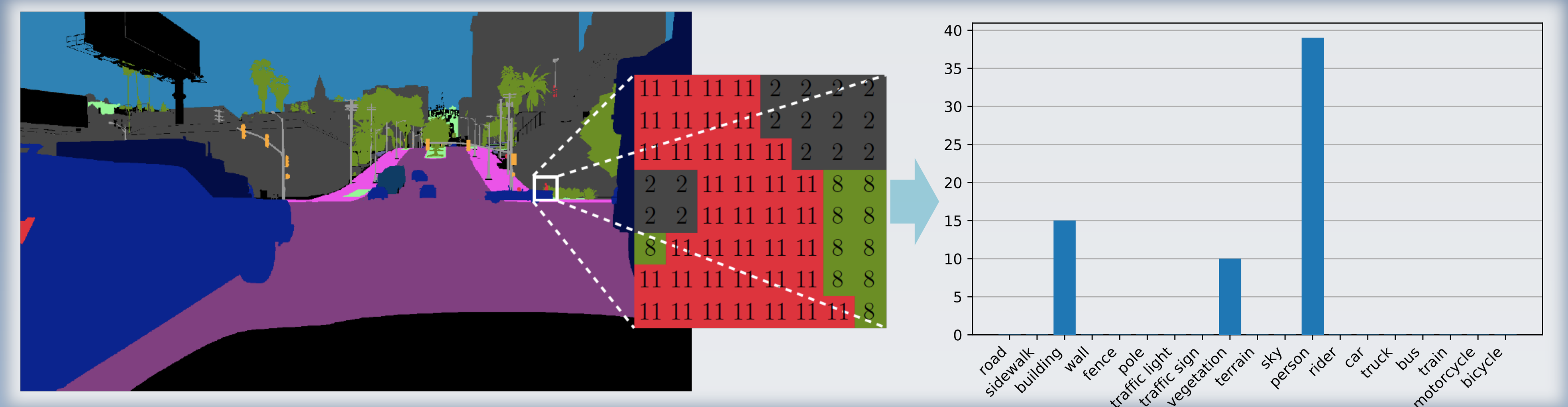
Problem Setup

Deep convolutional neural networks for semantic segmentation **do not generalize well** to distributions slightly different from the one of the training data and they **require a huge amount of labelled data** for their optimization. We introduce feature-level space-shaping regularization strategies to reduce the domain discrepancy in such scenario. Jointly enforcing a **clustering objective**, a **perpendicularity constraint** and a **norm alignment** goal on the feature vectors corresponding to source and target samples. We verify the effectiveness of our methods in the **autonomous driving setting** achieving **state-of-the-art results** in multiple synthetic-to-real road scenes benchmarks.



Feature-Level Labels

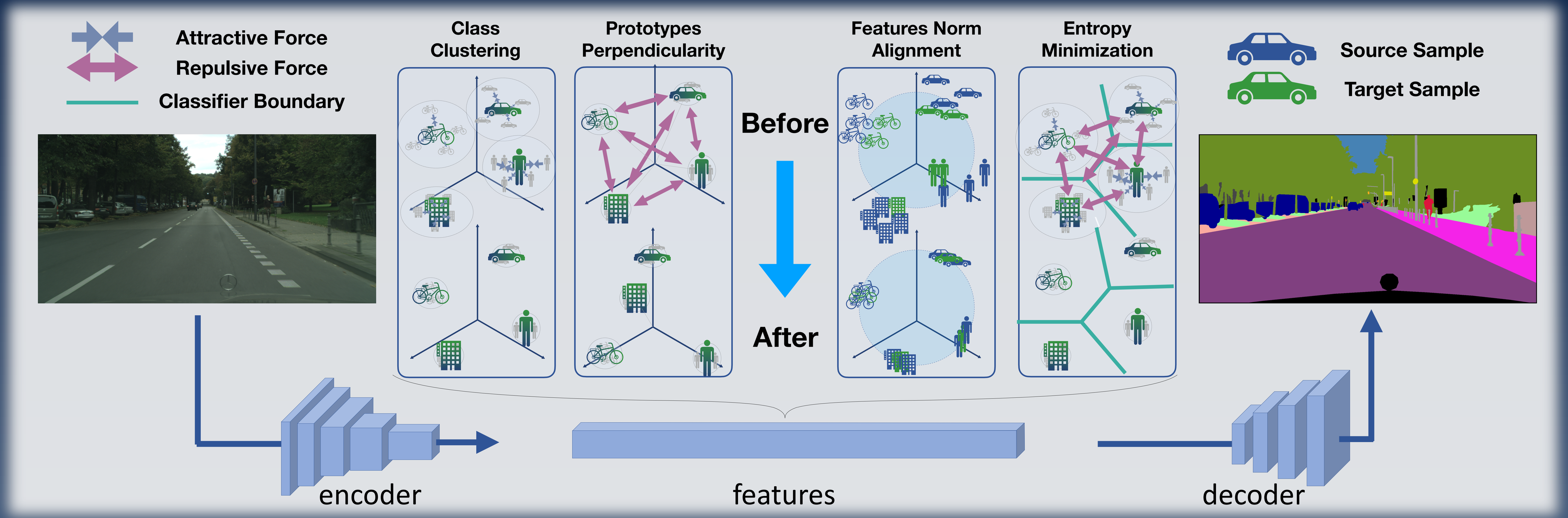
To employ **class-discriminative objectives** on the feature vectors one **needs** to **downsample** the **labels** in a **semantically-aware** manner.



Idea: Use SIFT-like **histogram filtering**, keep **only relevant classes**.

Architecture

The key **insight** of our strategy is the use of multiple **space-shaping strategies** to **enhance** the **semantic content** of the **latent space**.



To further **reduce estimation errors** in the computation of class-discriminative losses we exploit **exponentially smoothed prototypes**.

$$\mathbf{p}_c[i] = \frac{1}{|\mathcal{F}_c^s|}\sum_{\mathbf{f}\in\mathcal{F}_c^s}\mathbf{f}[i] \qquad \hat{\mathbf{p}}_c = \eta\hat{\mathbf{p}}_c' + (1-\eta)\mathbf{p}_c$$

Space-Shaping Objectives

1) Class Clustering: Align source and target feature vectors

$$\mathcal{L}_C = \frac{1}{|\mathcal{C}|}\sum_{c\in\mathcal{C}}\frac{1}{|\mathcal{F}_c|}\sum_{\mathbf{f}\in\mathcal{F}_c}||\hat{\mathbf{p}}_c - \mathbf{f}||^2$$

2) Prototype Perpendicularity: Different classes should have different activations

$$\mathcal{L}_P^s = \frac{1}{|\mathcal{C}|(|\mathcal{C}|-1)}\sum_{c_i,c_j\in\mathcal{C},i\neq j}\frac{\mathbf{p}_{c_i}}{||\mathbf{p}_{c_i}||}\cdot\frac{\mathbf{p}_{c_j}}{||\mathbf{p}_{c_j}||}$$

3) Norm Alignment: Target samples have smaller feature norms than source ones

$$\mathcal{L}_N^s = \frac{1}{|\mathcal{F}^s|}\sum_{\mathbf{f}\in\mathcal{F}^s}\Big|(\bar{f}_s+\Delta_f) - ||\mathbf{f}||\Big| \qquad \mathcal{L}_N^t = \frac{1}{|\mathcal{F}^t|}\sum_{\mathbf{f}\in\mathcal{F}^t}\max(0,(\bar{f}_s+\Delta_f) - ||\mathbf{f}||)$$

4) Entropy Minimization: Mimic source samples confidence of network in target ones

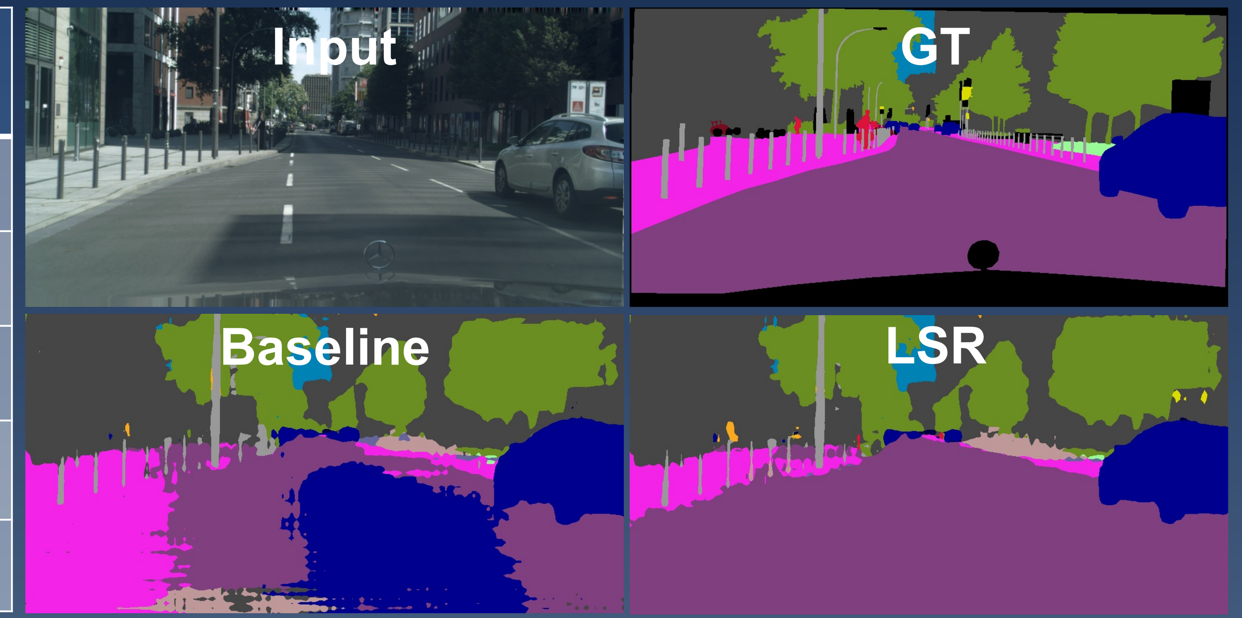
$$\mathcal{L}_{EM} = -\frac{1}{2N}\sum_{n,c}(p_t^{n,c})^2$$

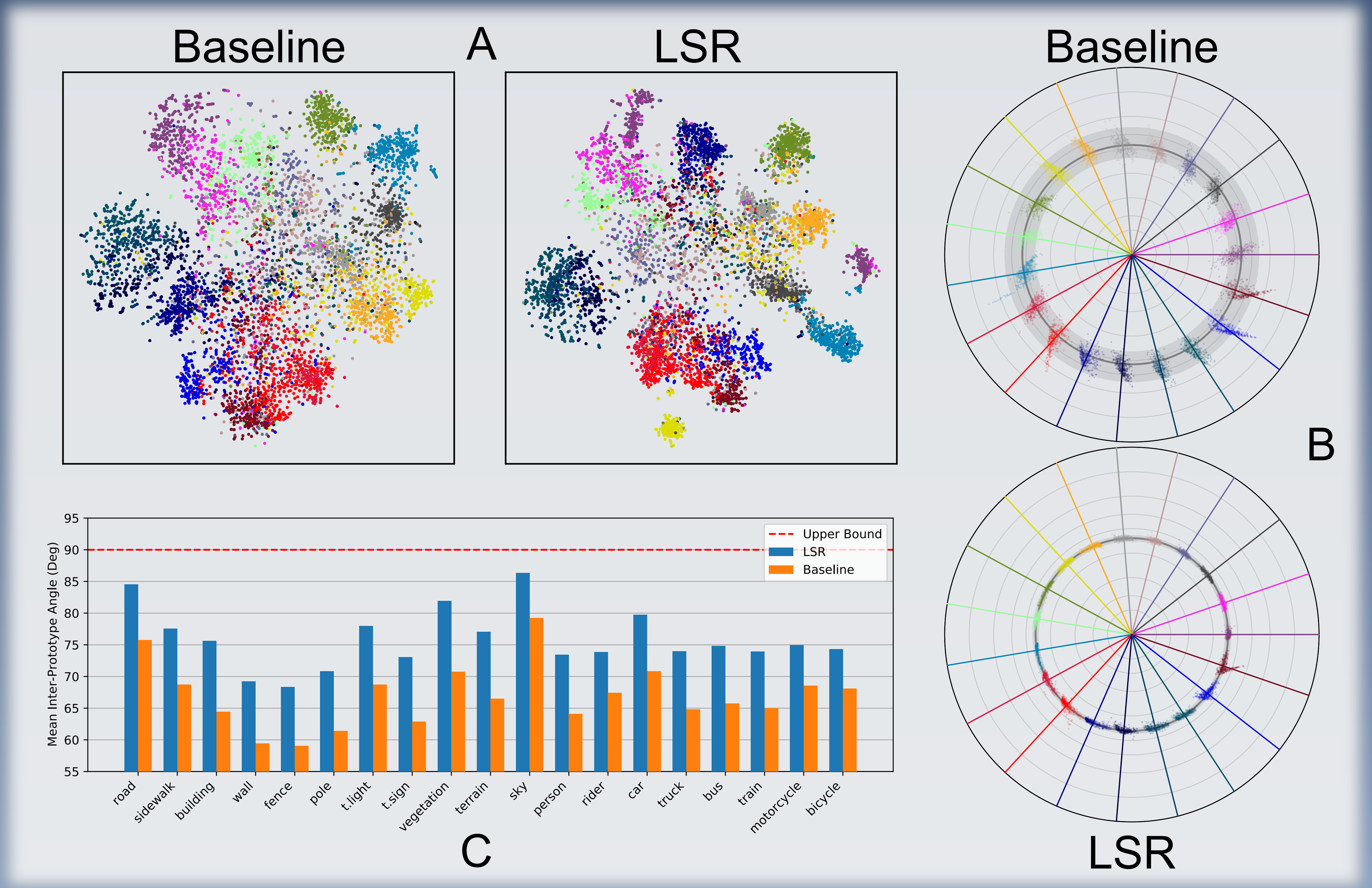
Chen, M. et al., ICCV, 2019.

Results

Our end-to-end **strategy**, despite its simplicity, achieves **state-of-the-art results** in the common **synthetic-to-real** autonomous driving **benchmark GTAV-to-CityScapes.**

Strategy	mIoU
Baseline	36.9
ASN (feats)	39.0
SAPNet	43.2
MaxSquareIW	45.5
LSR (Ours)	**46.0**





A. T-SNE plot of normalized feature vectors.
B. Star plot: Norm and intra-class angle. More details in **section 6.5** of the article.
C. Bar Plot: inter-prototype angles.