A baseline for semi-supervised learning of efficient semantic segmentation models: supplementary material

Ivan Grubišić*, Marin Oršić*, Siniša Šegvić
University of Zagreb, Faculty of Electrical Engineering and Computing
ivan.grubisic@fer.hr, marin.orsic@gmail.com, sinisa.segvic@fer.hr

1 Qualitative results

Figure 1 illustrates qualitative results of semi-supervised and supervised training evaluated on half-resolution Cityscapes val. The left column displays clean and perturbed versions of images from Cityscapes val. The second column displays corresponding fround truth. The last two columns display predictions of supervised training and semi-supervised MT-PhTPS training. MT-PhTPS results in less errors on clean images and much more robustness to T_{τ} perturbations.
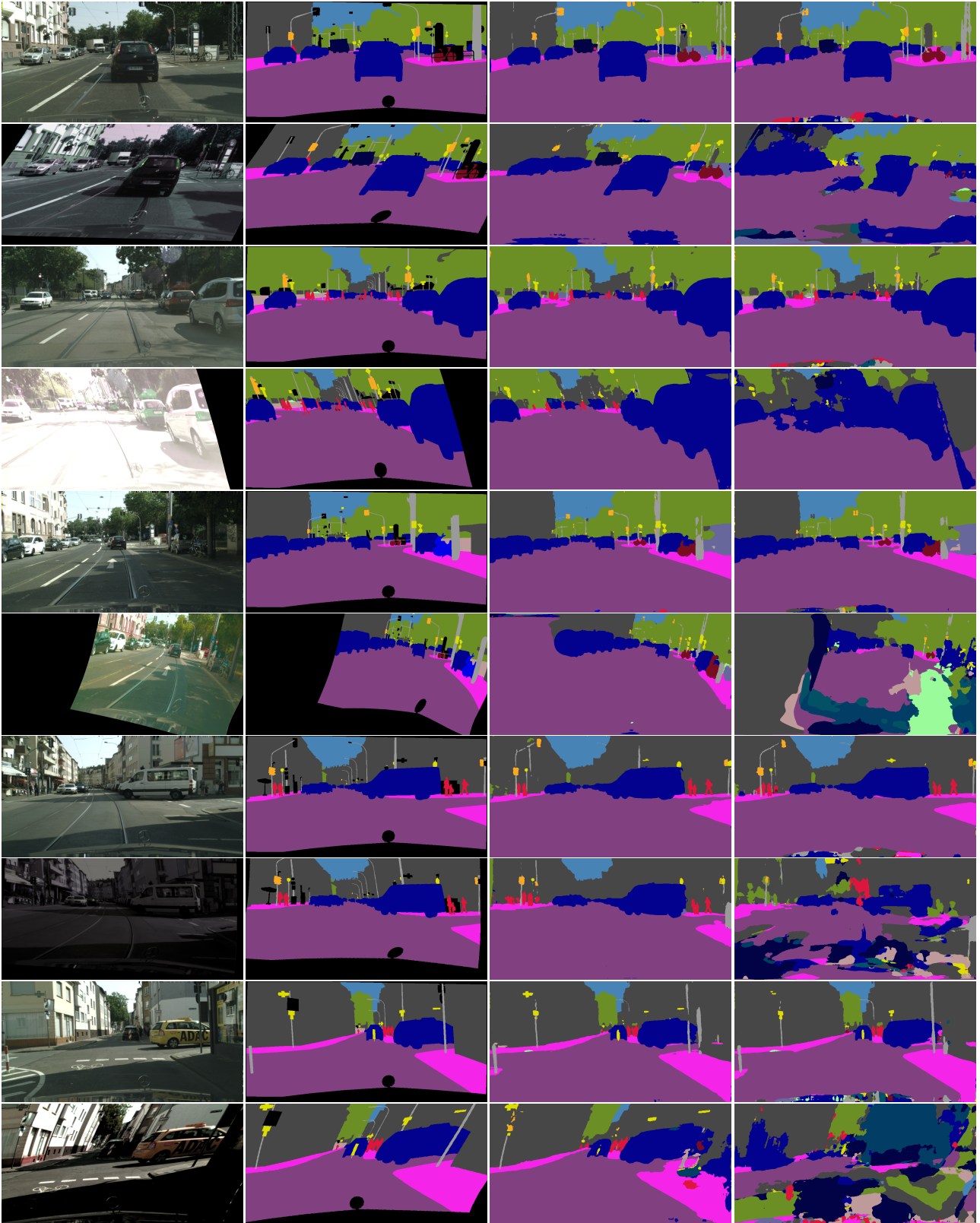
---

*Equal contribution.

Figure 1. Qualitative results on images from half-resolution Cityscapes val with SwiftNet-RN18 trained on half-resolution Cityscapes train with $1/4$ of the labels. The colums are as follows: (1) input image (perturbed with $T_{\boldsymbol{\tau}}$ in odd rows), (2) corresponding ground truth label (warped with $T_{\boldsymbol{\gamma}}^{\mathrm{G}}$ in even rows), (3) segmentation predicted when trained with MT-PhTPS, and (4) segmentation predicted when trained only on labeled images (supervised). Models trained with semi-supervised consistency produce better segmentations in clean images and are much more robust to perturbations used during training.