Supplementary Material
CoSMix: Compositional Semantic Mix for Domain Adaptation in 3D LiDAR Segmentation

Cristiano Saltori[1], Fabio Galasso[2], Giuseppe Fiameni[3], Nicu Sebe[1], Elisa Ricci[1,4], and Fabio Poiesi[4]

[1] University of Trento, Trento, Italy
[2] Sapienza University of Rome, Rome, Italy
[3] NVIDIA AI Technology Center
[4] Fondazione Bruno Kessler, Trento, Italy
cristiano.saltori@unitn.it

1 Introduction

We provide the supplementary material in support of our main paper. This document is organized as follows:

- Sec. 2 reports preliminary results on the real→real UDA setup of SemanticPOSS→SemanticKITTI with CoSMix.
- Sec. 3 provides qualitative examples of our mixed point clouds $\mathcal{X}^{s\to t}$ and $\mathcal{X}^{t\to s}$ from SynLiDAR [3] to SemanticKITTI [1].
- Sec. 4 reports additional qualitative results on SemanticPOSS [2] and SemanticKITTI [1].

2 Real to real adaptation performance

Domain adaptation between different real datasets requires their classes to be compatible. SemanticPOSS and SemanticKITTI are labelled neither by using the same semantic classes nor by following the same annotation protocol. We created a mapping between SemanticPOSS and SemanticKITTI classes to train the source model and to adapt it with CoSMix. Specifically, during the experiment we consider only the semantic classes *person*, *car/vehicle*, *trunk*, *plants*, *traffic-sign*, *pole*, *building*, *fence* and *ground*. Tab. 1 reports some preliminary results of this experiment.

Table 1: Preliminary results of CoSMix on SemanticPOSS→SemanticKITTI.

Method	**Source**	**CoSMix**
mIoU	22.5	26.4

3 Qualitative mixed samples

We provide qualitative examples of the mixed input point clouds $\mathcal{X}^{s\to t}$ and $\mathcal{X}^{t\to s}$ on SynLiDAR→SemanticKITTI. Each point cloud is randomly sampled during adaptation with local h and global r augmentations activated, $\zeta = 0.9$ and $\alpha = 0.5$. In Fig. 1, we report $\mathcal{X}^{s\to t}$ taken from the $s \to t$ branch while in Fig. 2 we report $\mathcal{X}^{t\to s}$ taken from the $t \to s$ branch. We provide paired source, target and mixed ($s \to t$ and $t \to s$) point clouds by reporting labels (top row) and binary masks (bottom row). Source labels are the ground-truth labels while target labels are the pseudo-labels filtered with $\zeta = 0.9$. Mixed samples show hybrid scenes with both source and target components. Although classes with points distributed over wide regions may be mixed (*e.g.* road and especially in the case of $t \to s$), mixed point clouds often include complementary elements among domains.

4 Qualitative adaptation results

Fig. 3-4 show additional qualitative examples after adaptation on SynLiDAR → SemanticPOSS while Fig. 5-6 show additional qualitative examples after adaptation on SynLiDAR→SemanticKITTI. We report results before adaptation (source), after adaptation with CoSMix (ours) and add ground-truth labels (gt) for comparison. We highlight regions with interesting results using red circles.

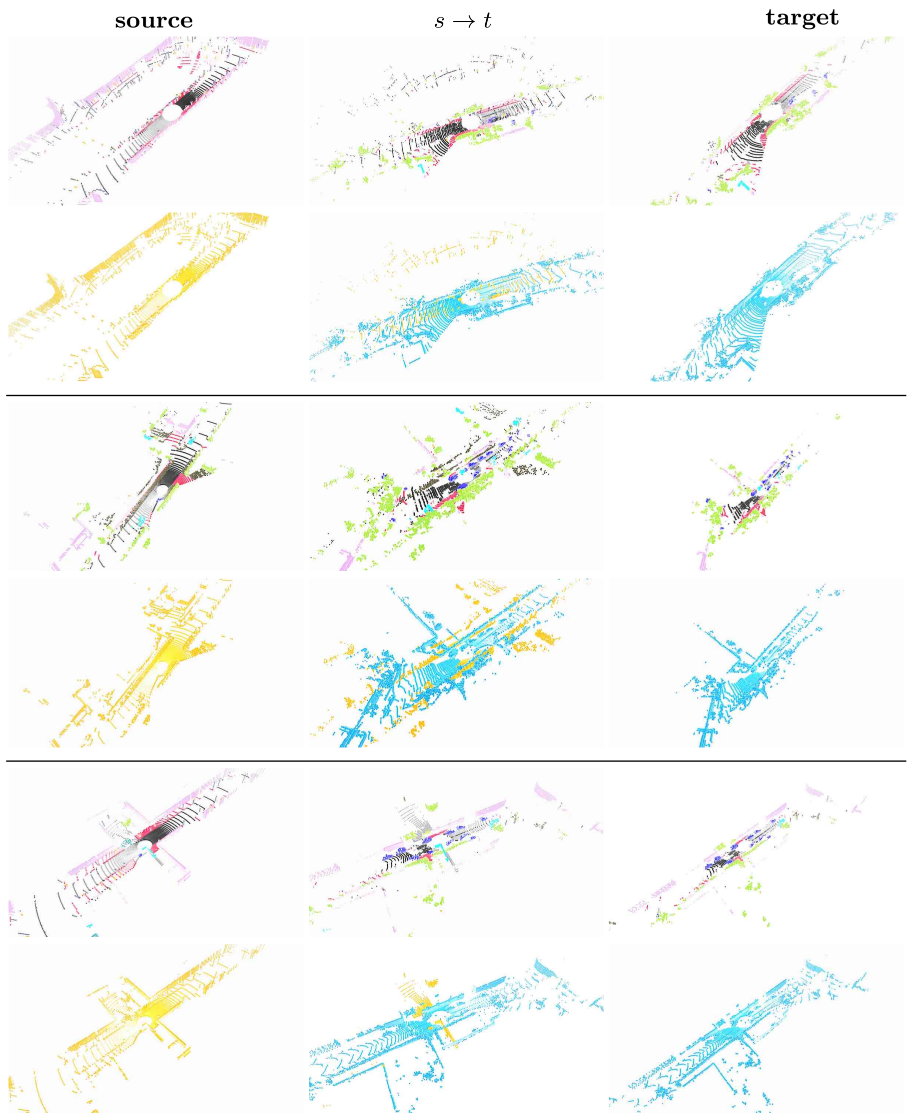

Fig. 1: Example of mixed point clouds in the s→t branch on SynLiDAR → SemanticKITTI. We report scenes with annotations (top rows) and binary masks (bottom rows).

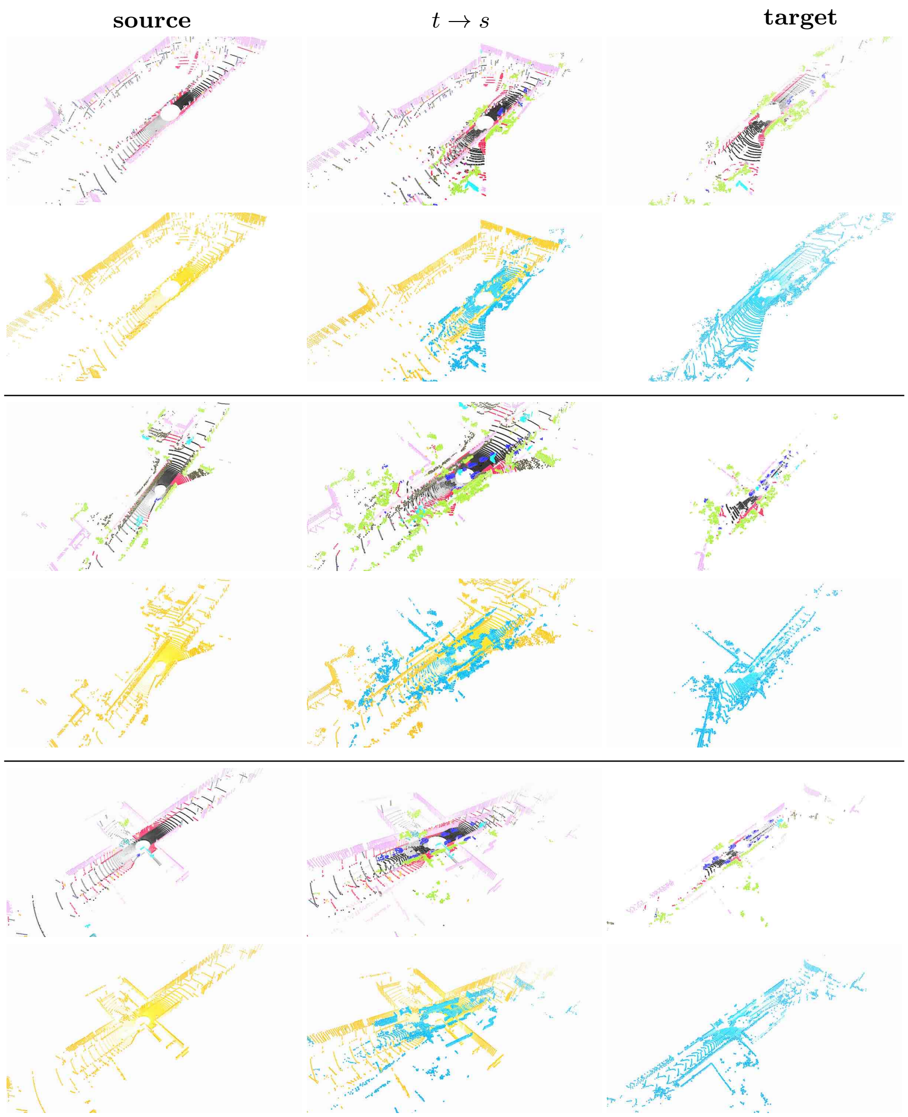

Fig. 2: Example of mixed point clouds in the t→s branch on SynLiDAR → SemanticKITTI. We report scenes with annotations (top rows) and binary masks (bottom rows).

Fig. 3: Results on SynLiDAR→SemanticPOSS. Source predictions are often wrong and mingled in the same region. After adaptation, CoSMix improves the segmentation accuracy with homogeneous predictions and correctly assigned classes. The red circles highlight regions with interesting results.

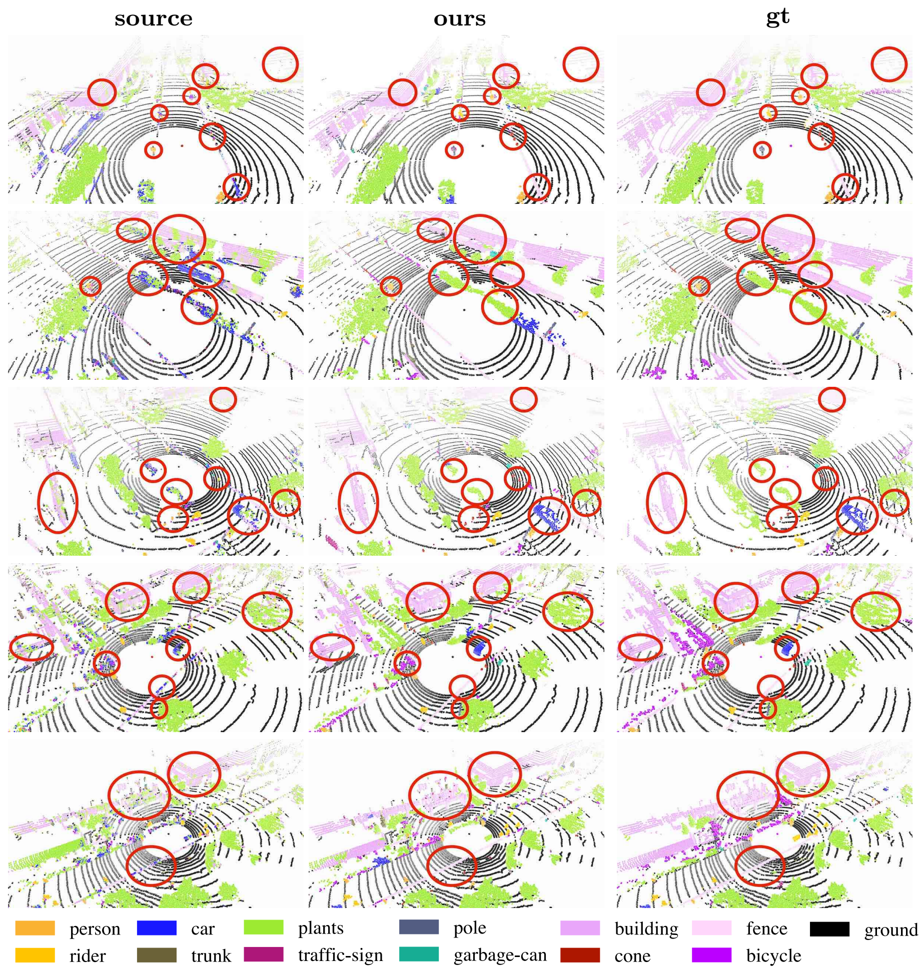

Fig. 4: Results on SynLiDAR→SemanticPOSS. Source predictions are often wrong and mingled in the same region. After adaptation, CoSMix improves the segmentation accuracy with homogeneous predictions and correctly assigned classes. The red circles highlight regions with interesting results.

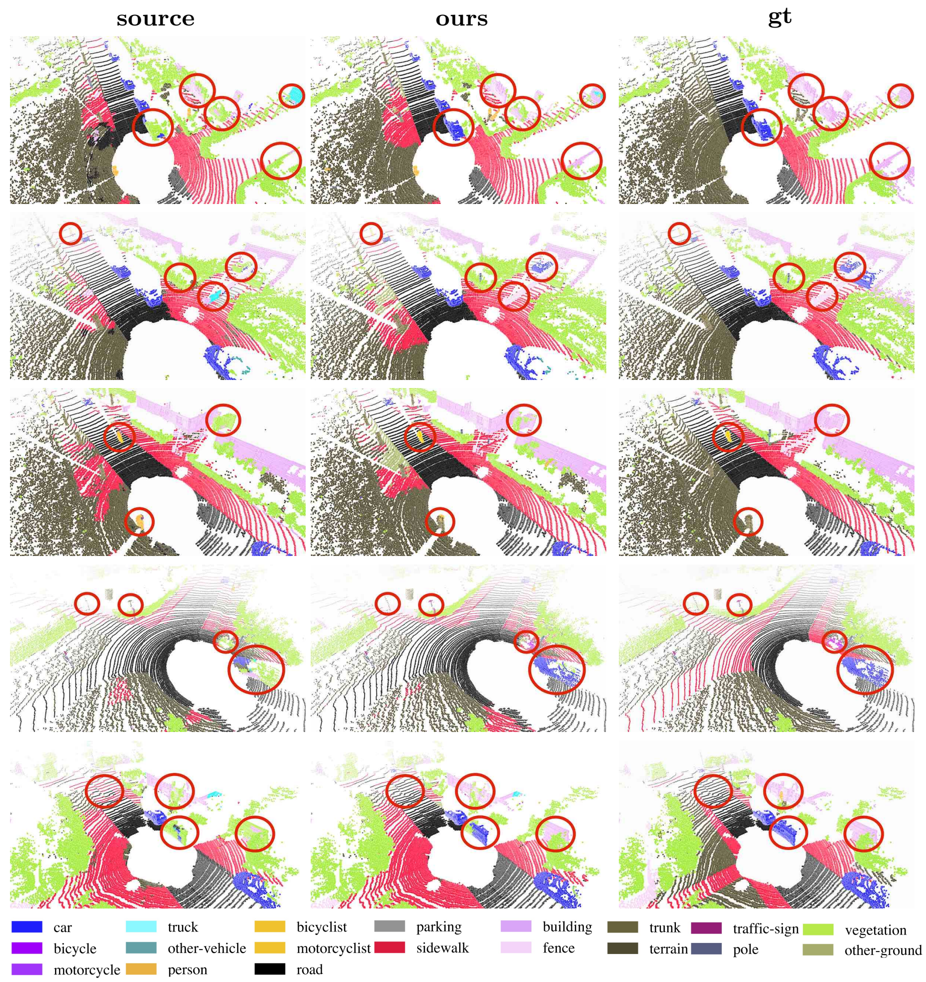

Fig. 5: Results on SynLiDAR→SemanticKITTI. Source predictions are often wrong and mingled in the same region. After adaptation, CoSMix improves the segmentation accuracy with homogeneous predictions and correctly assigned classes. The red circles highlight regions with interesting results.

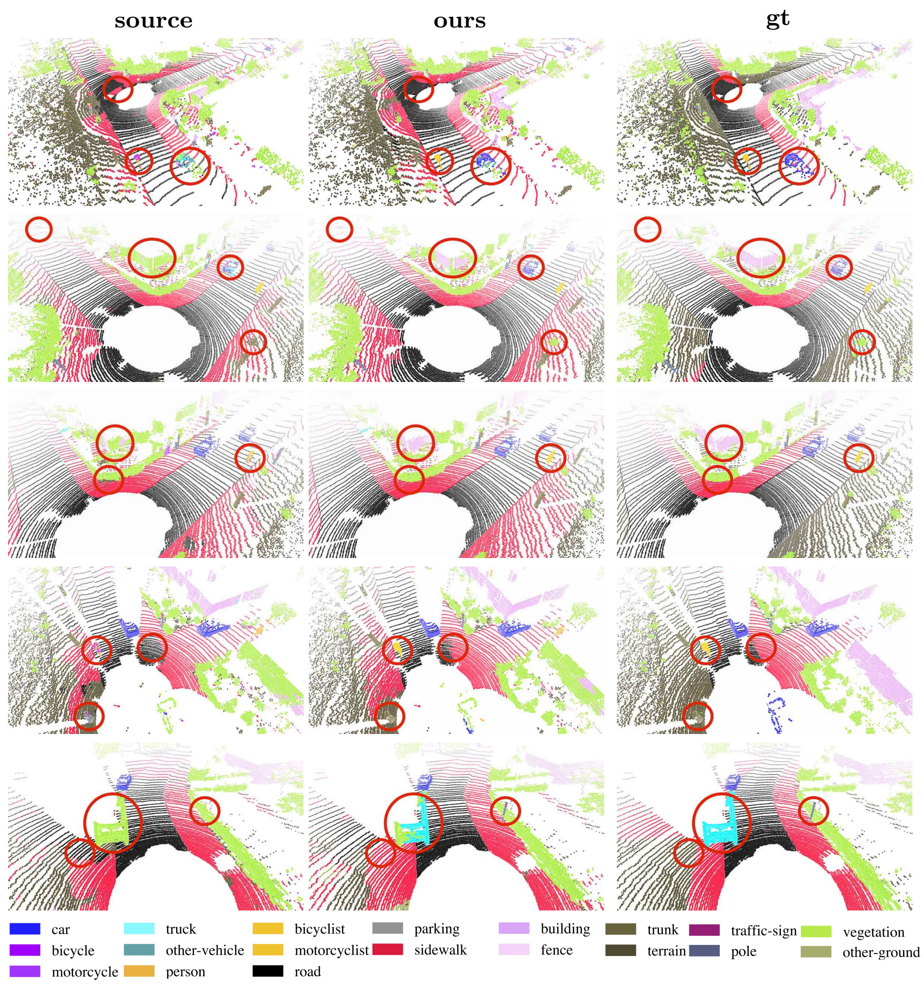

Fig. 6: Results on SynLiDAR→SemanticKITTI. Source predictions are often wrong and mingled in the same region. After adaptation, CoSMix improves the segmentation accuracy with homogeneous predictions and correctly assigned classes. The red circles highlight regions with interesting results.

References

1. Behley, J., Garbade, M., Milioto, A., Quenzel, J., Behnke, S., Stachniss, C., Gall, J.: SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences. In: ICCV (2019) 1
2. Pan, Y., Gao, B., Mei, J., Geng, S., Li, C., Zhao, H.: SemanticPOSS: A Point Cloud Dataset with Large Quantity of Dynamic Instances. arXiv (2020) 1
3. Xiao, A., Huang, J., Guan, D., Zhan, F., Lu, S.: Synlidar: Learning from synthetic lidar sequential point cloud for semantic segmentation. AAAI (2022) 1