

SUPPLEMENTARY MATERIALS: Low-Resource White-Box Semantic Segmentation of Supporting Towers on 3D Point Clouds via Signature Shape Identification

Diogo Lavado*, Cláudia Soares*, Alessandra Micheletti[†], Giovanni Bocchi[†], Alex Coronati[‡],
Manuel Pio[‡], and Patrizio Frosini[§]

SM1. TS40K Dataset. Electrical companies are responsible for the maintenance and inspection of the transmission system. They deploy low-flying helicopters to scan rural environments, from a BEV perspective, where the electrical grid is located. The produced point clouds exhibit different data properties when compared to 3D scenes captured from other viewpoints, such as from a vehicle. Namely, they show high point density and no object occlusion, scene elements present homogeneous density and no sparsity. Then, the acquired 3D data is processed by maintenance personnel. Specifically, as the raw point clouds are dense and mainly encompass rural areas, data is sectioned into strips of land focused on the transmission system (Figure SM2). The raw data comprises several LiDAR files containing roughly 40 000 kilometers of the above land strips. 3D points therein are labeled with one out of 22 possible classes, such as power lines and their supporting towers, low and medium vegetation, rivers, railroads, human-made structures that do not belong to the transmission network, the ground, and optic cables, among others. Table SM1 depicts these classes and their density in the dataset. Rail lines and road surfaces constitute the majority of the dataset (63%), whereas classes of interest, such as power line supporting towers, make up less than 5% of the overall data.

SM1.1. Qualitative Results.. In Figure SM2, we can analyze the qualitative results on the TS40K dataset (a) of SCENE-Net (c) against a CNN is similar architecture as baseline (b). Even though the CNN achieves a higher Recall on the majority of the samples, it classifies most vertical scene elements as towers, which ultimately leads to a poor segmentation of supporting towers. In contrast, SCENE-Net segments the body of towers and rejects other vertical objects that do not exhibit the prior knowledge encoded in the model.

*NOVA School of Science and Technology, Lisbon (d.lavado@campus.fct.unl.pt, claudia.soares@fct.unl.pt).

[†]University of Milan, Milan (alessandra.micheletti@unimi.it, giovanni.bocchi1@unimi.it).

[‡]EDP NEW, Lisbon (alex.coronati@edp.pt, manuelpio.silva@edp.pt).

[§]University of Bologna, Bologna (patrizio.frosini@unibo.it).

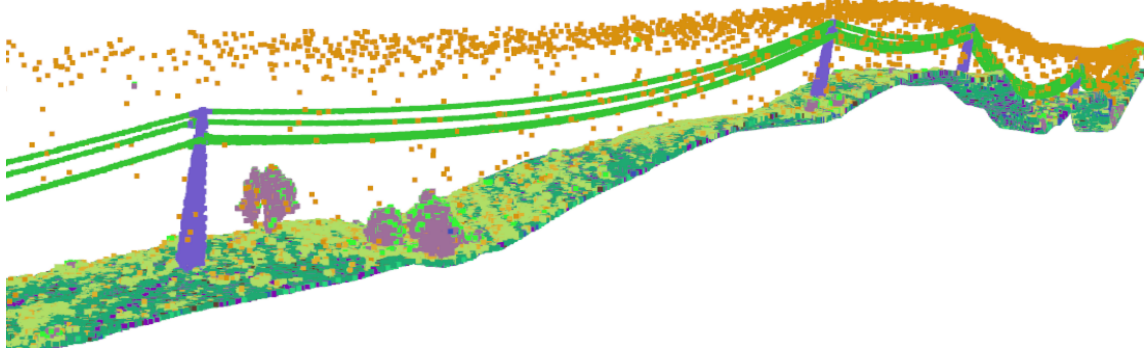


Figure SM1. Visualization of TS40K raw point cloud with colored labels.

Table SM1

Available classes in the TS40K dataset and their distribution. Rail lines and road surfaces constitute the majority of the dataset (63%). Our class of interest, power line towers, only makes up 0.52%. Moreover, around 40% of the tower points are mislabelled.

Label	Class	Density(%)	Label	Class	Density(%)
0	Created	0	11	Road surface	18.758
1	Unclassified	0.571	12	Overlap points	23.403
2	Ground	0.529	13	Medium Reliability	0
3	Low vegetation	0.681	14	Low Reliability	0
4	Medium vegetation	0.241	15	Power line support tower	0.519
5	Natural obstacle	1.069	16	Main power line	0.907
6	Human structures	0	17	Other power line	0.002
7	Low point	0.362	18	Fiber optic cable	0
8	Model key points	0	19	Not rated object to be consider	8.205
9	Water	0	20	Not rated object to be ignored	0
10	Rail	44.752	21	Incidents	0

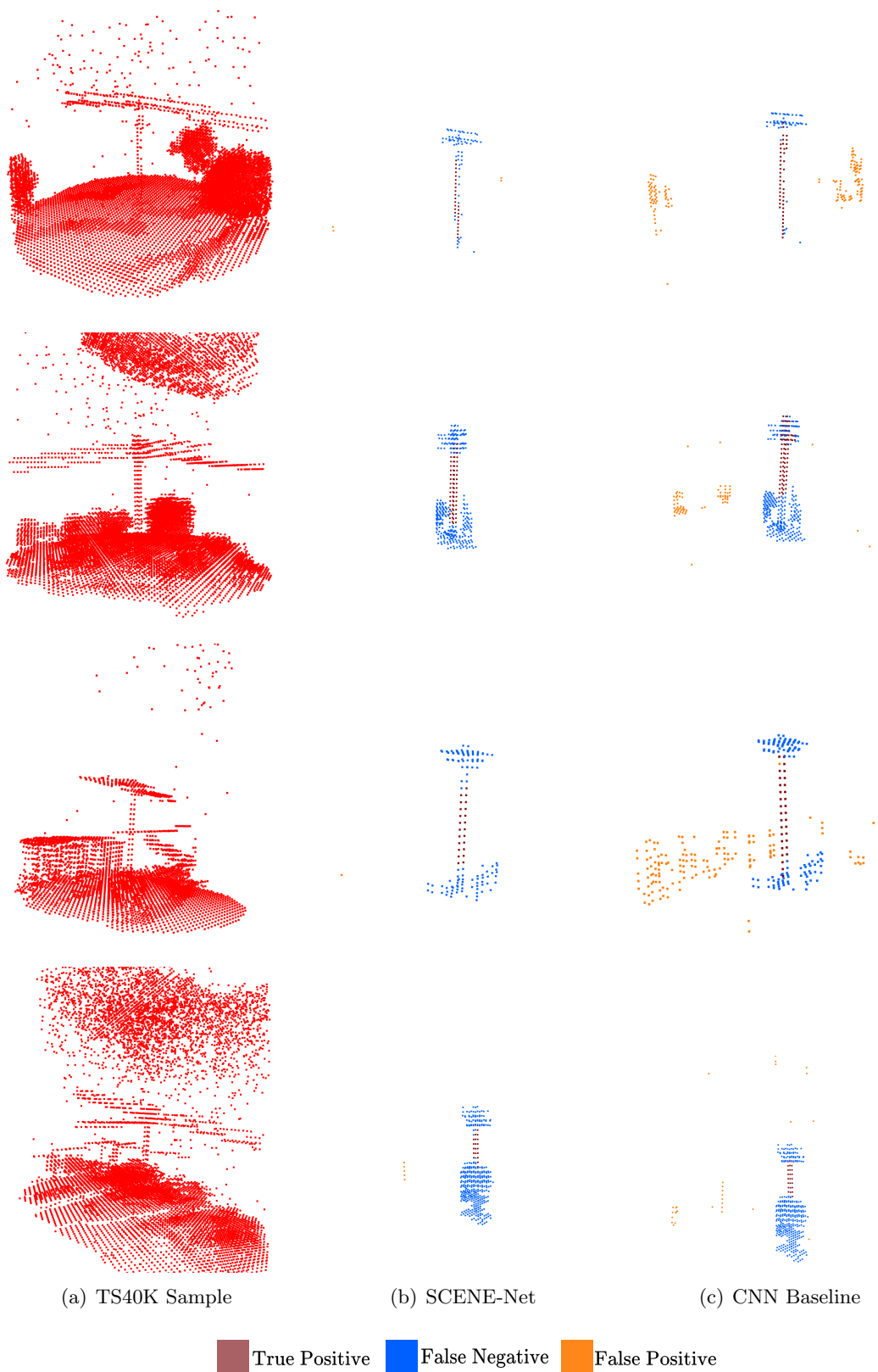


Figure SM2. Qualitative results of SCENE-Net on the testing set of TS40K, against a CNN with similar architecture. Note that, in the last example, SCENE-Net clearly identifies a second unlabeled tower, whereas CNN identifies both the second tower and vegetation as towers.