Supplementary Material:

TreeON: Reconstructing 3D Tree Point Clouds from Orthophotos and Heightmaps

Angeliki Grammatikaki[1], Johannes Escher[1], Pedro Hermosilla[1], Oscar Argudo[2], Manuela Waldner[1]

[1]TU Wien, Institute of Visual Computing and Human-Centered Technology, Vienna, Austria
[2]Universitat Politècnica de Catalunya, Department of Computer Science, Barcelona, Spain

1 Qualitative Results on Austrian Landmark Trees

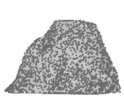
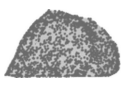
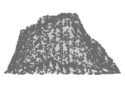
Table 1 presents qualitative reconstruction results on the Austrian landmark tree dataset [Gra24]. These results complement the qualitative evaluation presented in Table 3 of the main paper by providing the full set of Austrian landmark trees used in our real-world experiments.

For each tree, we show the DSM and orthophoto inputs used by the model, together with a real-world photograph and the reconstructed 3D tree rendered as a point cloud. Although no ground-truth 3D geometry is available for this dataset, the results demonstrate that our method recovers key morphological characteristics, such as tree height, crown extent, and overall silhouette. The reconstructions exhibit substantial variation across trees, indicating that the model generalizes beyond the synthetic training data and captures natural structural diversity without relying on species labels or multi-view supervision.

We observed only a small number of category-level confusions. Specifically, four out of 72 trees were reconstructed with an incorrect coarse category: trees 8 and 12 are coniferous but were reconstructed with a deciduous appearance, while trees 43 and 46 is deciduous but was reconstructed with a coniferous appearance. In addition, a limited number of intra-class confusions were observed among visually similar conifer species: trees 64 and 65 (pine) were reconstructed with a spruce-like appearance, and tree 35 (spruce) with a pine-like appearance. Overall, these errors are rare and occur primarily between visually similar species, reflecting the inherent ambiguity of species inference from top-down orthophotos and DSMs rather than systematic reconstruction failures.

Overall, this corresponds to a coarse conifer/deciduous correctness of 95.8% for TreeON on the 72 landmark trees. Specifically, two coniferous trees and two deciduous trees were assigned to the incorrect coarse category, yielding class-conditional correctness of 97.22% for both conifers and deciduous trees. For comparison, Grammatikaki et al. [GEL*25] report probabilities of 93.3% for conifers and 88.8% for deciduous trees being correctly classified on the same dataset. While [GEL*25] explicitly optimizes for species classification as a separate learning stage, TreeON does not use semantic supervision and instead performs end-to-end geometric reconstruction. The high coarse-category consistency observed here therefore suggests that TreeON implicitly captures meaningful species-related structure from sparse top-down inputs, despite not being trained for explicit species recognition.

Table 1: Qualitative reconstruction results on Austrian landmark trees [Gra24]. For each tree, we show a real-world photograph, the DSM and orthophoto inputs, and the reconstructed tree produced by our method. No ground-truth 3D geometry is available for this dataset; the results illustrate generalization to real-world landmarks and structural consistency across diverse tree shapes.

ID	DSM	OrthophotoTarget		Output	ID	DSM	OrthophotoTarget		Output
1					2				
3					4				
5					6				

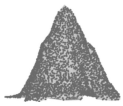
7




8

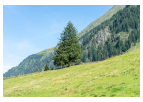



9

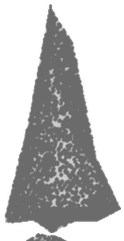



10



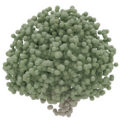



11

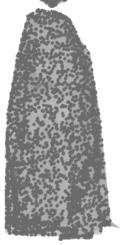



12



13




14




15




16




17

18

19

20

21

22

23

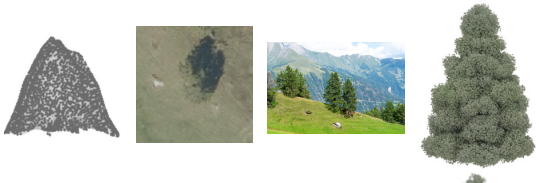
24

25

26

27 

28 

29 

30 

31 

32 

33 

34 

35 

36 

37 

38 

39 

40 

41 

42 

43 

44 

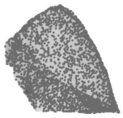
45 

46 

47 

48 

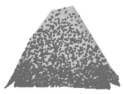
49



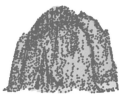

50




51



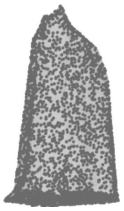

52

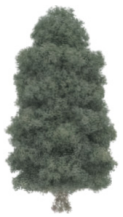



53

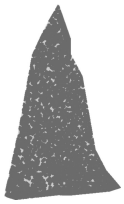



54



55



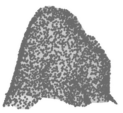

56




57




58




59

60

61

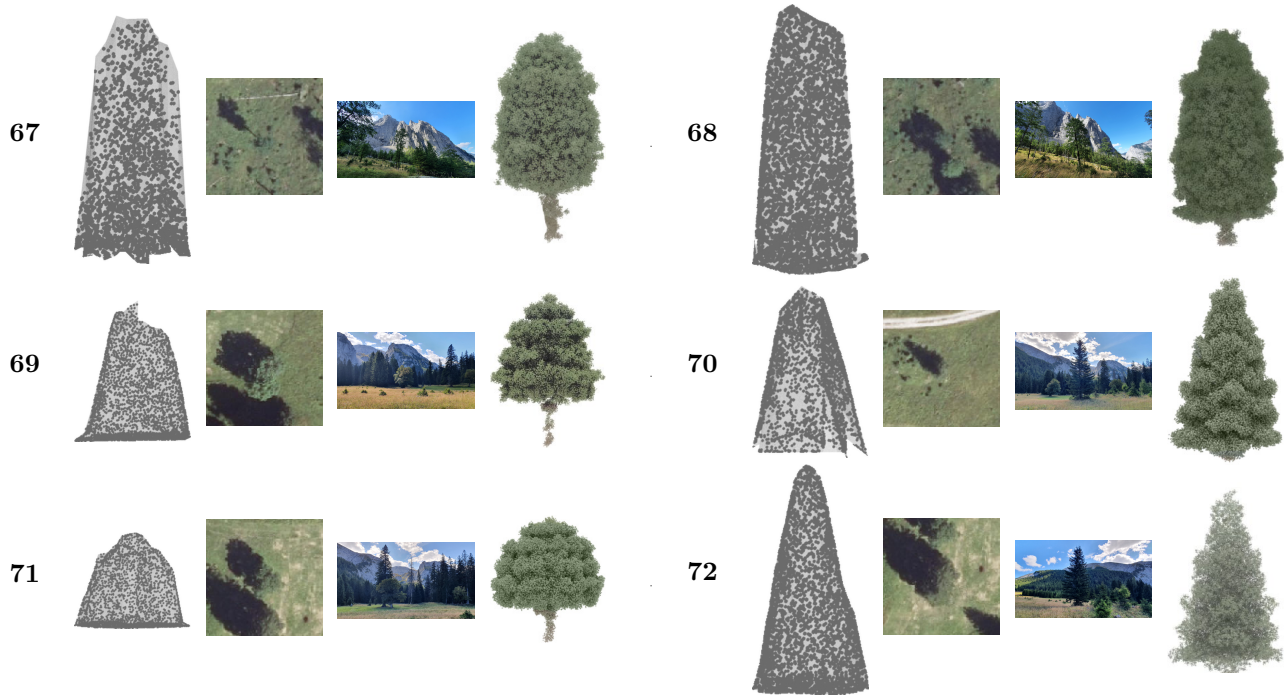
62

63

64

65

66



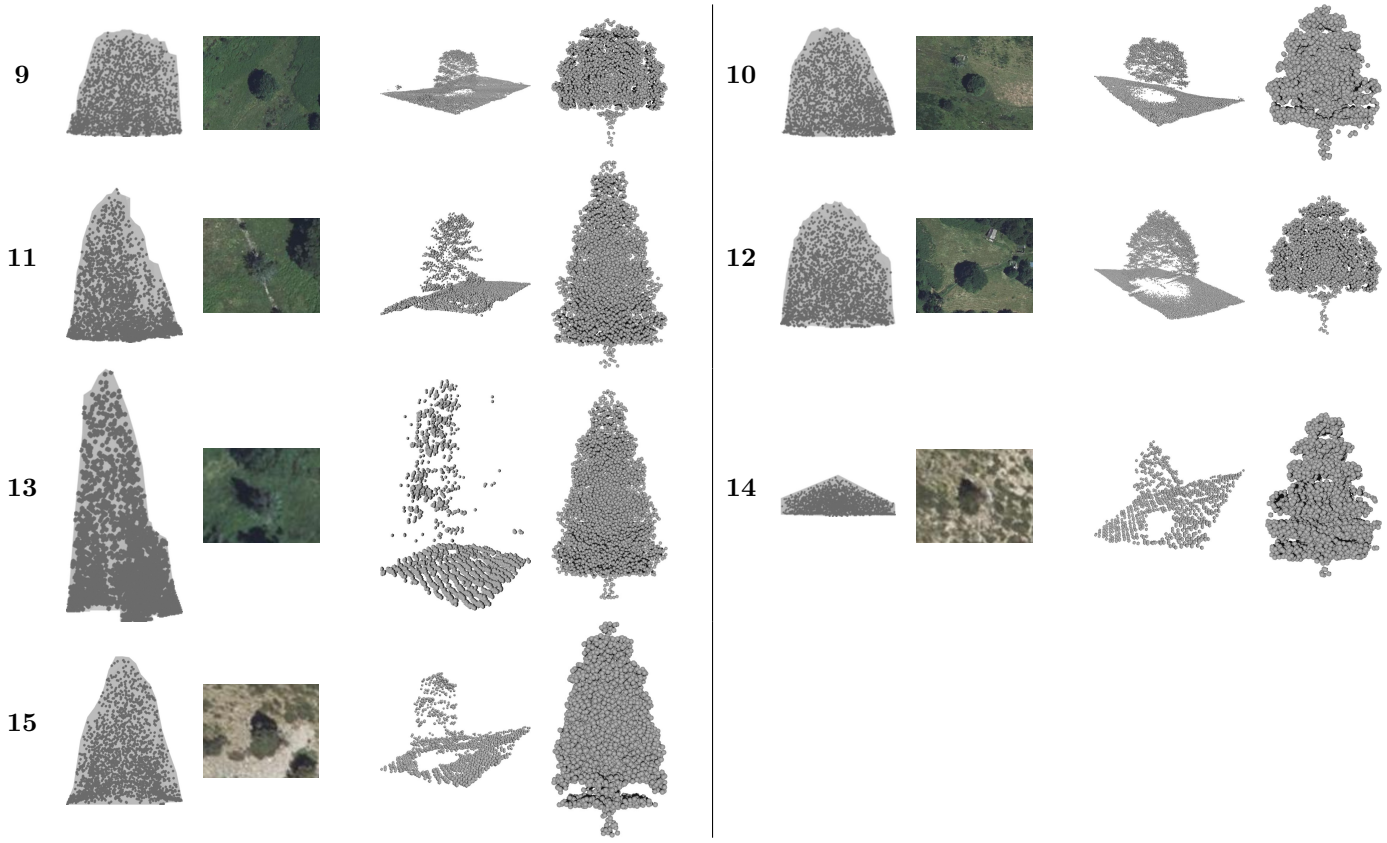
2 Qualitative Comparison with LiDAR Data (France)

Table 2 shows qualitative comparisons between our reconstructed trees and high-resolution airborne LiDAR data from the French IGN dataset. This comparison corresponds to the comparison with high-resolution LiDAR data discussed in Section 5.4 of the main paper and provides an external geometric reference where airborne LiDAR data are available.

For each example, we present the DSM and orthophoto inputs, the LiDAR-derived reference point cloud, and our reconstruction. Our reconstructions align well with the LiDAR reference in terms of overall height and crown extent, despite differences in sensing modality. While fine-scale branching varies, our method consistently recovers realistic tree geometry from sparse geodata and preserves a clearly visible trunk structure.

Table 2: Qualitative comparison with airborne LiDAR data on the French IGN dataset.

ID	DSM	Orthophoto	LiDAR	Ours	ID	DSM	Orthophoto	LiDAR	Ours
1					2				
3					4				
5					6				
7					8				



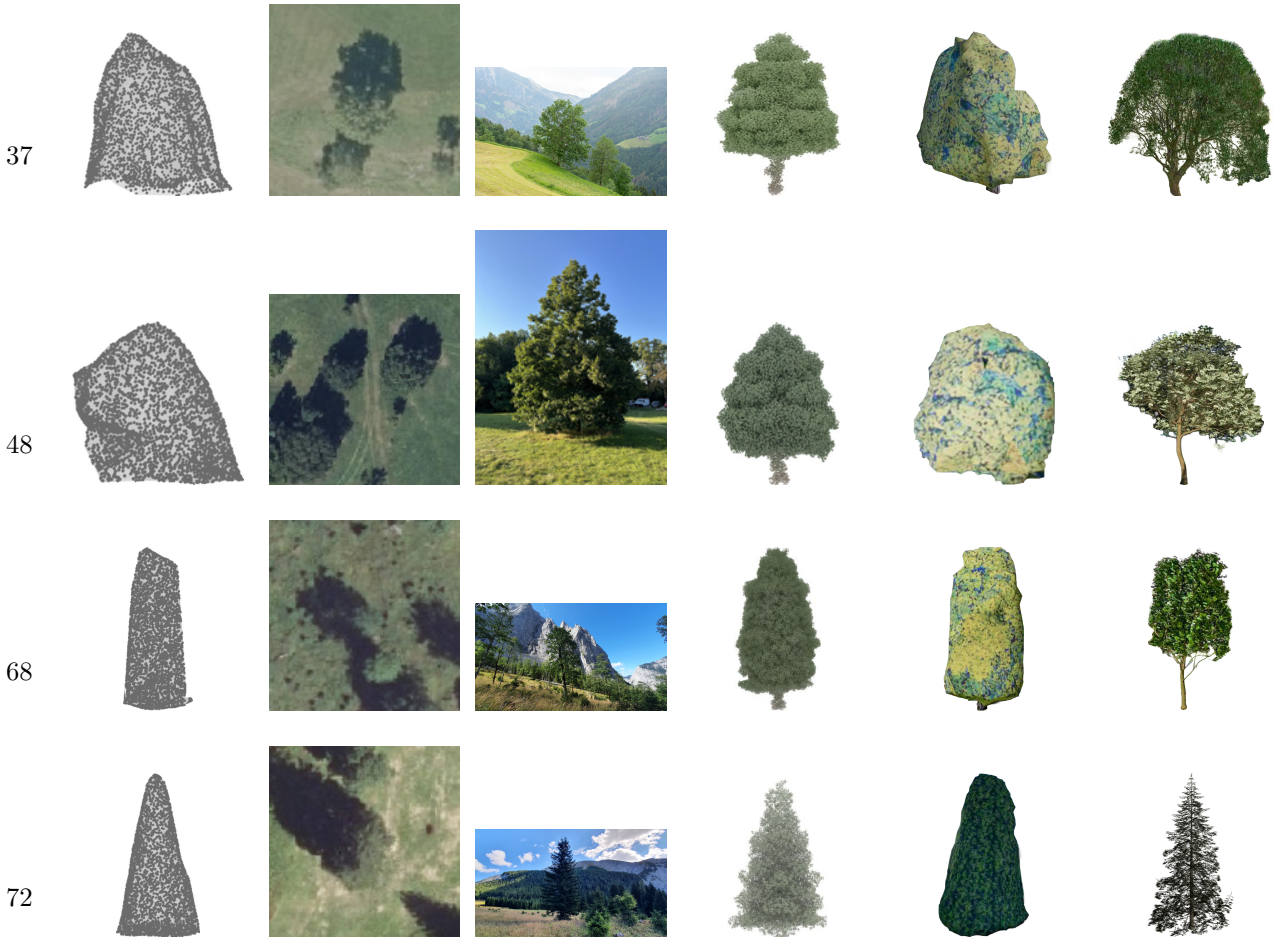
3 Extended Qualitative Comparison with Baseline Reconstructions

Table 3 provides an extended qualitative comparison between our reconstructions and alternative reconstruction variants based on the method of Grammatikaki et al. [GEL*25]. This extended comparison supplements the quantitative and runtime analysis reported in Section 5.5 of the main paper. For each tree, we show the input data alongside our reconstruction and the corresponding low-detail and high-detail models produced by their pipeline.

Whereas the pipeline of Grammatikaki et al. requires expensive procedural growth steps, with reported runtimes of up to ∼15 s per tree and storage requirements in the tens of megabytes for high-detail meshes, our method predicts a compact point-cloud representation in approximately 0.3 s per tree, with a memory footprint of about 0.2 MB. This represents an order-of-magnitude improvement in both runtime and storage, while retaining the structural and perceptual detail required for large-scale geospatial visualization.

Table 3: Qualitative comparison with baseline reconstructions based on Grammatikaki et al. [GEL*25]. For each tree, we show the input data, our reconstruction, and the corresponding low- and high-detail baseline models. Our lightweight point-cloud reconstructions preserve salient tree structure while enabling substantially lower runtime and memory usage than the mesh-based baseline.

ID	DSM	Orthophoto	Target	Ours	Low-Detail [GEL*25]	High-Detail [GEL*25]
13						
23						



4 Qualitative Comparison with Google Earth™

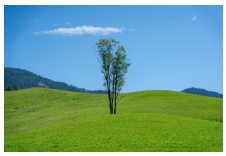
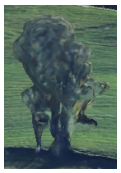
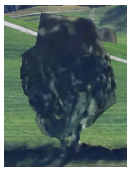
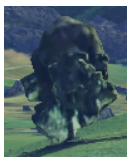
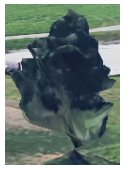
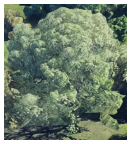
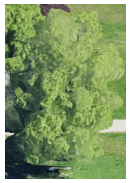
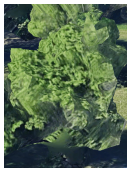
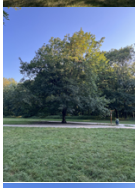
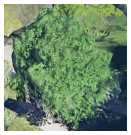
Figure 1 provides a qualitative comparison between our reconstruction and a corresponding tree visualized in Google Earth™ (tree 7 in Table 1). Table 4 extends this comparison to the 30 Austrian landmark trees for which Google Earth™ 3D data is available, out of the 72 trees listed in Section 1 of the Supplementary Material. Google Earth™ images were obtained by cropping the 3D scene at the corresponding landmark locations defined in the dataset [Gra24]. These include nine trees located in rural areas outside Vienna, with the remaining trees situated within the city. The results show that our point-cloud reconstructions reproduce consistent crown structure and visually plausible shape and density across diverse real-world tree forms.

(a) Photographic reference of landmark tree.

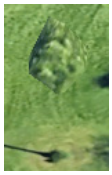
(b) LiDAR-based comparison. **Left:** Google Earth reference. **Right:** Our reconstruction.

Figure 1: A side-by-side comparison of the same scene as rendered with our approach with point cloud representations and Google Earth™.

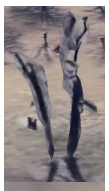
Table 4: Qualitative comparison between target photographs, Google Earth™ 3D visualizations, and our reconstructed trees. Google Earth images are shown for qualitative reference only and are not part of the dataset or evaluation.

ID	Target	Google Earth™	Ours	ID	Target	Google Earth™	Ours
1				2			
3				4			
5				6			
7				8			
9				39			
40				41			
42				43			
44				45			
46				47			
48				49			

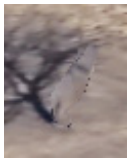
50



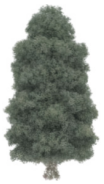
52



53



54



55

56

57

58

59

60

References

[GEL*25] Grammatikaki A., Eschner J., Ledermann F., Argudo O., Waldner M.: How to represent landmark trees in digital 3d maps? an automated workflow and user study. *Cartography and Geographic Information Science 0*, 0 (2025), 1–18. doi:10.1080/15230406.2025.2489543.

[Gra24] Grammatikaki A.: Landmark trees in austria, Mar. 2024. URL: https://doi.org/10.48436/nsj20-6ka24, doi:10.48436/nsj20-6ka24.