The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-G-2025
ISPRS Geospatial Week 2025 “Photogrammetry & Remote Sensing for a Better Tomorrow…”, 6–11 April 2025, Dubai, UAE

Designing the Forest of Tomorrow: Generating Virtual Trees with Adversarial Autoencoders

Hristina Hristova[1], Clemens Blattert[1], Stefano Puliti[2], Janine Schweier[1]

[1] Sustainable Forestry, Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Switzerland
(hristina.hristova, clemens.blattert, janine.schweier)@wsl.ch

[2] Norwegian Institute for Bioeconomy Research (NIBIO), Division of Forest and Forest Resources, Norway
stefano.puliti@nibio.no

Keywords: Tree species generation, Virtual forests, Forestry, Deep learning.

Abstract

This paper explores the generation of “realistic” 3D representations of individual trees to enhance visualizations of forest simulation tool outcomes. By leveraging remote sensing data, we aim to capture individual tree features and characteristics accurately, linking them to dynamic simulations of forest structures and composition. Employing a deep learning approach, we train models on existing 3D scanned data to produce diverse and realistic visual representations of specific tree species. Our method addresses the limitations of existing synthetic tree generation techniques, which often overlook species-specific characteristics. Our approach emphasizes the generation of diverse tree forms, accounting for differences in trunk shape, canopy size, and branching structures. The resulting 3D data offers potential applications for realistic future forest visualizations and improved data augmentation in tree classification models, ultimately contributing to the creation of virtual forests that represent rich species diversity.

1. Introduction

The sustainable management of forest ecosystems requires an intricate understanding of their current state and potential future development. The advancement of remote sensing technologies, such as Terrestrial Laser Scanning (TLS) and Mobile Laser Scanning (MLS), have revolutionized the analysis of forest structures through high-resolution point clouds (Kükenbrink et al., 2022, Liang et al., 2016). Complementing these technologies, photogrammetry offers a cost-effective method for reconstructing and visualizing complex forest environments. These approaches facilitate the analysis of critical features like Diameter at Breast Height (DBH) and tree position (Kükenbrink et al., 2022, Hristova et al., 2024, Mokroš et al., 2018), as well as tree species (Puliti et al., 2024b), and volume (Bornand et al., 2023). While these methods provide invaluable insights into the present conditions of forests, they do not directly relate to future forest dynamics.

Advancements in Deep Learning (DL) and computer graphics introduce opportunities to enhance our understanding of the forest’s future. Utilizing these technologies, virtual forest stands that realistically simulate the appearance and behavior of real-world forests can be produced. Recent approaches in generating virtual forests, such as (Holm and Schweier, 2024), have made strides by creating a virtual forest in the Unity game engine. Such tools allow for the visualization of long-term forest simulation outcomes using predefined tree characteristics. These visualizations may assist practitioners and foresters in understanding the potential development of future forests, which helps select appropriate management strategies.

In this context, studies exploring the integration of forest simulations with high-resolution data would be the next evolutionary step in visualizing the outcomes of various management scenarios (Neudam et al., 2023). By leveraging point cloud data filtered by tree features, like DBH and height, such visualizations provide a comprehensive framework for evaluating forest dynamics and their possible effects on forest structure and composition. Remote sensing data provides accurate information about current forest conditions and tree features, which could serve as inputs for simulations. Additionally, such data could be linked to the output of these simulations, supporting forest practitioners in their long-term decision-making.

Forest simulation tools often provide simplistic output in the form of modeled forest parameter values (*i.e.*, a list of trees comprising their dimension and species). To create a foundation for visualizing these parameters and possible future states of forest ecosystems, our work investigates the potential of generating 3D data for individual trees that accurately represent their real-world counterparts. To achieve this, we employ a DL approach and train models to produce meaningful 3D representations of specific tree species and their characteristics. Our model learns from existing 3D data of scanned individual trees, enabling the generation of a variety of tree representations for a given species from a latent space. To make the latent space more compact and avoid any gaps, we train our generator model to sample from a prior distribution.

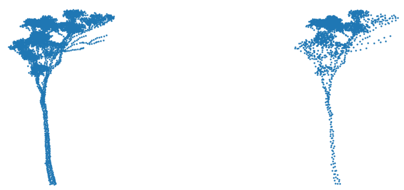
Our results indicate that the proposed generation method is a step toward effectively learning the geometric characteristics of tree species to produce new point cloud data that closely resembles real-world trees. While state-of-the-art methods also focus on generating synthetic tree data (Bryson et al., 2023, Dobbs et al., 2023), they create synthetic trees without concentrating on specific tree species. Bryson et al. (Bryson et al., 2023) focused on building an artificial dataset for training deep learning models aimed at stem and crown classification rather than generating virtual trees for visualization purposes. Thus, they do not compare the synthetic trees their model produced with real trees.

Furthermore, Dobbs et al. (Dobbs et al., 2023) propose a supervised method for tree skeletonization, which aims to learn and produce the skeleton of an input tree. Our approach, however, significantly differs from these reference methods. We focus on creating diverse representations of various tree species that exhibit distinct shapes, canopy sizes, and branch structures. Our

https://doi.org/10.5194/isprs-archives-XLVIII-G-2025-619-2025 | © Author(s) 2025. CC BY 4.0 License.

(a) Picea abies

(b) Fagus sylvatica



(c) Eucalyptus miniata

Figure 1. Each sub-figure a), b), and c) represents a pair of generated tree point clouds using our model (right tree) and corresponding real-world tree point clouds (left tree). The generated trees were produced by sampling from a prior latent distribution.

primary objective is to produce new 3D data for a given tree species without prior knowledge, making it appear as realistic as possible. This capability may enhance the realism of forest stand visualizations and assist in data augmentation for tree species classification models, ultimately facilitating the creation of virtual forests that showcase species diversity.

We summarize our main contributions as follows:

- An unsupervised model that creates life-like 3D representations of specific tree species using a DL approach, which learns from existing scanned tree data.
- Novel generative models that produce various visualizations of Europe’s most common tree species.
- A performance evaluation showcasing the benefits of our method for simulating real-world tree data. Figure 1 hints about the potential of our model.

The paper is organized as follows. First, we introduce our unsupervised method for generating 3D data of single trees and present the dataset of tree species that we use to train and test our model. This is followed by results and evaluation. Finally, we conclude the paper by discussing future work.

2. Our Method

2.1 Dataset

We used the recently published FOR-species20K dataset (Puliti et al., 2024a), which contains single-tree point clouds of 33 species obtained using TLS, MLS, and Unmanned Laser Scanning (ULS). We organized the point clouds by tree species and incorporated data from all three types of sensors. It is important to highlight that trees of a particular species can vary in appearance based on their growth stage and age. The FOR-species20K dataset offers diverse representations of Europe’s most common tree species, making it an appropriate training dataset for our model.

2.2 Input Data Preparation

The single-tree point clouds in the FOR-species20K dataset contain high-resolution data that would be too computationally demanding for our DL model. To prepare the data for input to our method, we decreased the density of the point clouds by randomly sub-sampling each point cloud to extract 4098 points using the Farthest Point Sampling (Eldar et al., 1997). This way, we preserved the overall shape of the 3D tree while decreasing the point cloud density. Furthermore, we normalized each sub-sampled tree point cloud to fit within the unit cube. The input for the DL model consists of a set of 3D points from the sub-sampled and normalized point clouds of a given tree species. Finally, we divided the dataset into training, validation, and test sets, selecting respectively 85%/5%/10% number of trees from each tree species group. The train set was used for training, the validation set was used for selecting the best training epoch, and the test set was used for computing the evaluation metrics. The FOR-species20K dataset contains species-imbalanced data, with varying quantities of samples for each tree species group. Therefore, the number of trees in the training, validation, and test sets varied across the tree species groups.

2.3 DL Model

We propose to train a model for generating the 3D shape of a given tree species in an unsupervised adversarial manner. Our model is based on 3D Adversarial Autoencoders (AAE) (Zamorski et al., 2020). However, as Zamorski et al. have trained their model for generating highly structured objects, such as chairs, tables, and cars, we aim to tackle the more complex problem of generating unstructured and diverse tree species data.

2.3.1 Adversarial Autoencoders AAE utilizes adversarial training to impose a specific prior distribution on the latent space. The encoder, denoted as `E`, encodes data into compact latent representations. The primary objective of AAE is to ensure that the encoder `E` outputs latent space values that follow a predetermined prior distribution, such as Gaussian, Beta, or Gamma distributions. The discriminator, referred to as `D`, learns to align these latent representations with the prior distribution.

This process enables the generation of new samples that are similar to the input data based on the prior distribution. By imposing a prior distribution on the latent space, we achieve a more compact and continuous latent distribution from which new data can be generated through sampling. The role of `D` is to differentiate between samples generated from the defined latent distribution and "fake" samples produced by the encoder `E`.

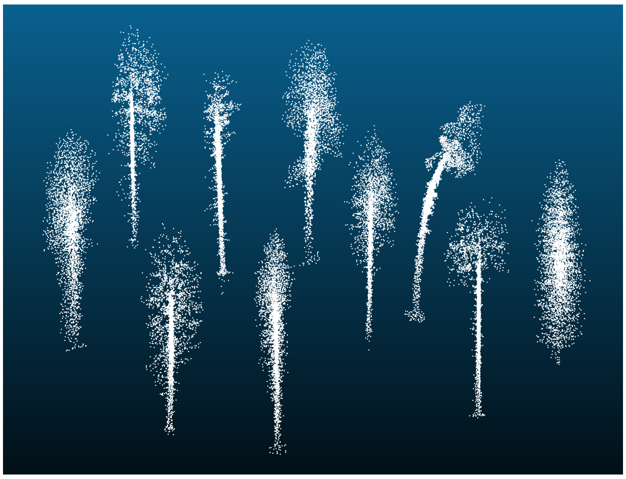
The encoder `E` is responsible for transforming the input 3D data into latent space representations while simultaneously learning to "fool" the discriminator `D`. The generator, denoted as `G`, uses samples drawn from the latent space distribution to create new data. By modeling the latent space and sampling from it, we aim to generate trees with diverse characteristics that resemble a specific tree species.

2.3.2 Loss Function The training process is guided by a loss function that consists of two main components: 1) reconstruction error, measured by the Chamfer distance between the input 3D data and the reconstructed point cloud, and 2) adversarial error, representing the negative log-likelihood of the generated encodings with respect to the discriminator. To balance these two components, we use a coefficient denoted by λ, which we have set to 10.

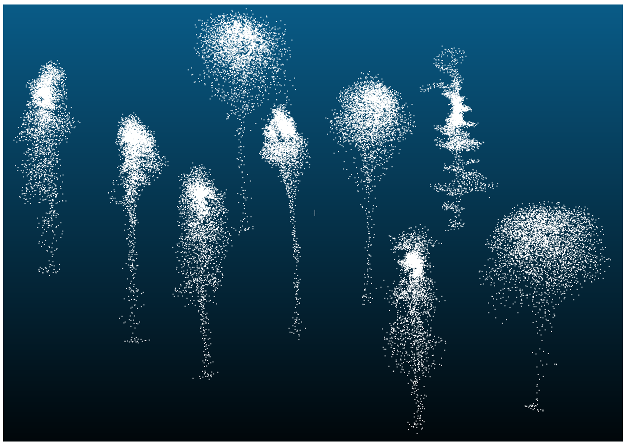
2.3.3 Validation After training, we selected the best training epoch using our validation set. For each epoch, we drew random samples from the prior distribution, each with sample size matching that of the latent space. We then used the generator `G` to create trees based on these random samples. The Jensen-Shannon Divergence (JSD) metric (Achlioptas et al., 2018), quantifying the similarity distance between two probability distributions, was employed to calculate the distance between the generated trees and the trees in our validation set. For each epoch, we computed the JSD three times using three different sets of random samples, drawn from the prior distribution. The epoch that produced the lowest mean JSD value was selected as the best epoch for the model.

2.3.4 Network Architecture The network architecture for the encoder `E` is based on PointNet++ (Qi et al., 2017). Our encoder consists of seven 1D convolution layers, one fully connected layer, and two separate fully connected layers for reparametrization. ReLU activations are applied in all layers except for the last layer used for parametrization. The generator `G` and the discriminator `D` are fully connected networks with six layers, utilizing ReLU activations in all layers except for the final one.

2.3.5 Implementation Details The network was implemented in PyTorch and is based on Zamorski et al.'s software (Zamorski et al., 2020). We trained `E`, `G`, and `D` of the AAE using the Gaussian distribution as the prior latent distribution. The model was trained for 3000 epochs with learning rates set at 0.0005 for both `E` and `G` and 0.00005 for `D`. We used a batch size of 16. The size of the latent space was 64. The training was conducted separately for each tree species in the FOR-species dataset, resulting in 33 trained models. Each model took approximately 23 hours to train on an NVIDIA TITAN V GPU with 32 GB of RAM.



(a) Picea abies



(b) Fagus sylvatica

(c) Eucalyptus miniata

Figure 2. Variety of single trees belonging to three different tree species, generated using our model by randomly sampling the latent distribution.

2.4 Evaluation Metrics

2.4.1 Reconstruction Ability We evaluated the effectiveness of our method in reconstructing real-life trees by using the JSD metric (Achlioptas et al., 2018, Zamorski et al., 2020). This metric measures the distance between two distributions and ranges from 0 to 1. We calculated the JSD metric for the reconstructed 3D data in relation to the corresponding point clouds in our test set, aiming to assess the reconstruction capability of our models quantitatively. We also computed the JSD metric on our validation set for the best training epoch.

2.4.2 Point Cloud-to-Point Cloud Metric To further evaluated the efficiency of the tree generation model, we employed the Multiscale Model-to-Model Cloud Comparison (M3C2) cloud-to-cloud metric (Lague et al., 2013). This metric meas-

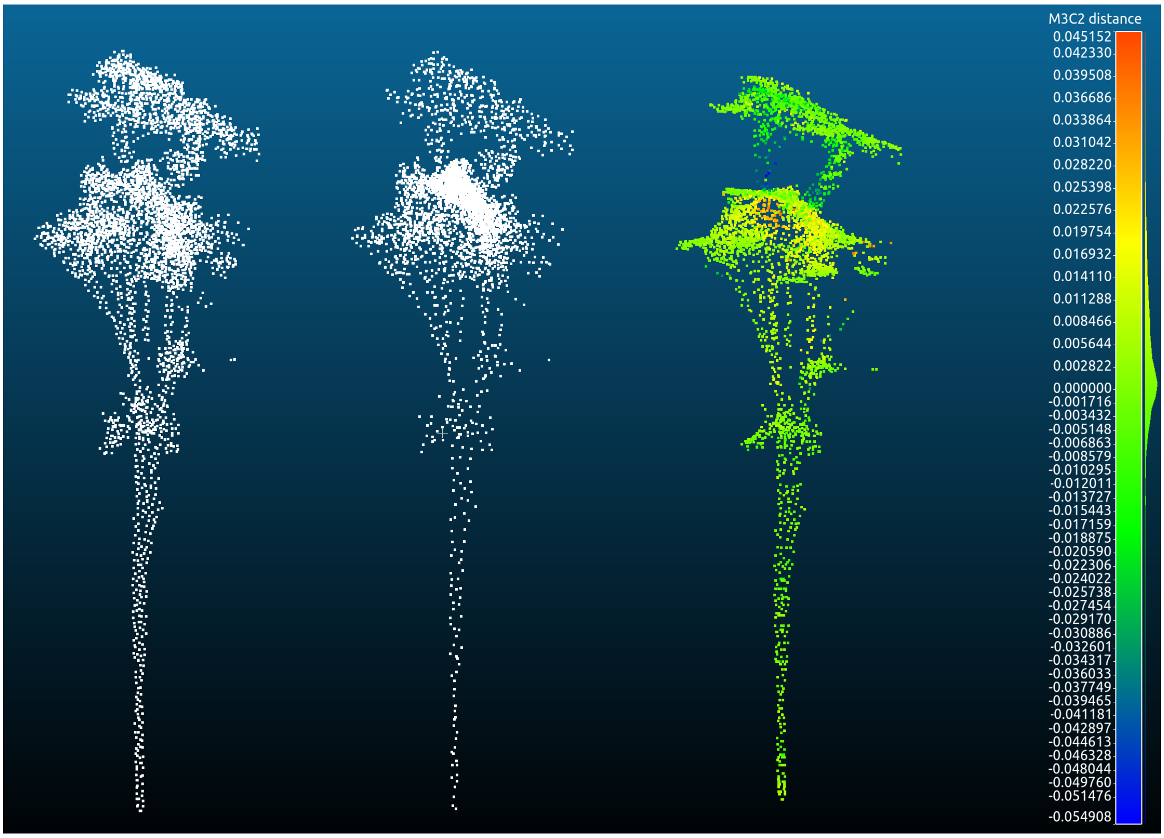


Figure 3. The figure represents a tuple of tree point clouds. The third tree shows the Multiscale Model- to-Model Cloud Comparison (M3C2) metric, computed between the real 3D data (first tree) and the generated 3D data using our method (second tree). The green and yellow colors indicate a high correlation between the generated points and the real data, while the red points highlight significant structural dissimilarities in a small portion of the canopy. The M3C2 distance demonstrates that, overall, our method effectively generates trees closely resembling the 3D tree data.

ures the structural similarities between two point clouds, focusing on local distance variations detected at reference core points. In our analysis, core points were calculated from the 3D data of real trees. We computed the distance between each core point and its corresponding neighboring points found within the generated point clouds, enabling us to evaluate the accuracy of the generated models in relation to the original data.

Moreover, we excluded core points without neighboring points, as these would not provide meaningful comparisons in our structural analysis. By focusing on valid core points with neighboring data, we ensured that our evaluation accurately reflected the model's performance in representing the tree structures.

3. Results and Discussion

3.0.1 Qualitative Assessment Figure 1 illustrates examples of trees from three tree species generated by our model alongside their real-world counterparts. To produce these trees, we encoded trees from our test set into their corresponding latent representations, which, by design, correspond to samples of our prior distribution. Then, our generator `G` transformed the latent representations into the generated trees in Figure 1. Visually, the generated trees closely resemble the real ones, demonstrating that our model effectively captures intricate branch structures. Nevertheless, the generated trees appear to contain fewer points than the original point clouds despite both having the same number of points. This discrepancy occurs because many points are clustered in specific areas, particularly in the crowns of the trees. To address this reconstruction issue, a different reconstruction loss, such as the Earth Mover's Distance (EMD) (Zamorski et al., 2020), could be adopted. However, employing the EMD may introduce a trade-off between computational efficiency and reconstruction accuracy.

Tree species	JSD validation	JSD test
Picea abies	0.176	0.189
Fagus sylvatica	0.146	0.137
Eucalyptus miniata	0.206	0.178

Table 1. The Jensen- Shannon Divergence (JSD) metric for our model, computed for three tree species groups using the validation and test sets.

By random sampling from the learned latent space, we can obtain diverse shapes of individual tree species. The input to the generator `G` is a vector sampled from the prior distribution. Specifically, Figure 2 shows trees generated for Picea abies, Fagus sylvatica, and Eucalyptus miniata, each corresponding to a different sampled vector from the prior distribution. Each reconstructed tree point cloud consists of 4098 points, as determined by the design of the generator `G`.

Quantitative Assessment Table 1 presents the JSD metric, evaluated across three tree species groups. The second column shows the JSD values computed on the validation set for the best training epoch. The third column displays the JSD values calculated on the test set. A smaller JSD value indicates bet-

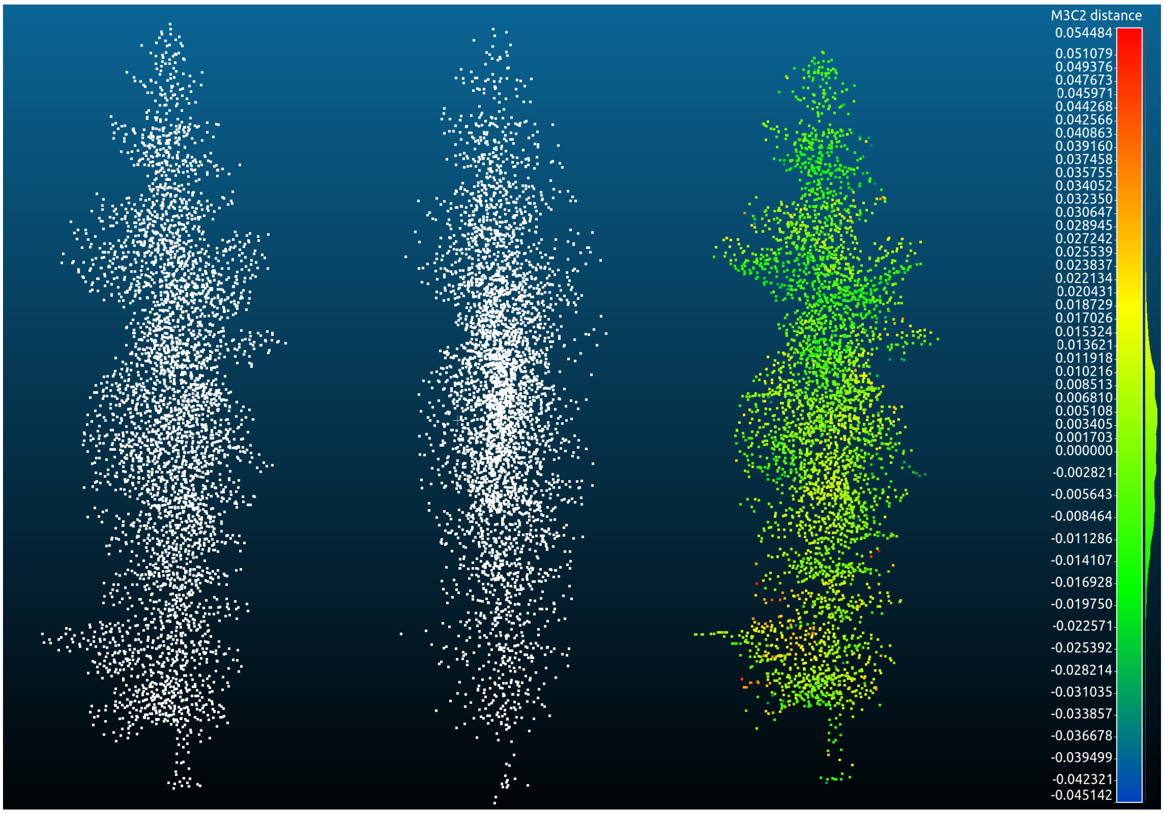


Figure 4. A tuple of tree point clouds. The third tree shows the Multiscale Model- to-Model Cloud Comparison (M3C2) metric, computed between the real 3D data (first tree) and the generated 3D data using our method (second tree). Most M3C2 points are green and yellow, revealing an overall successful reconstruction. Some red points appear close to the base of the third tree, signifying difficulties in reconstructing some of the branch structure.

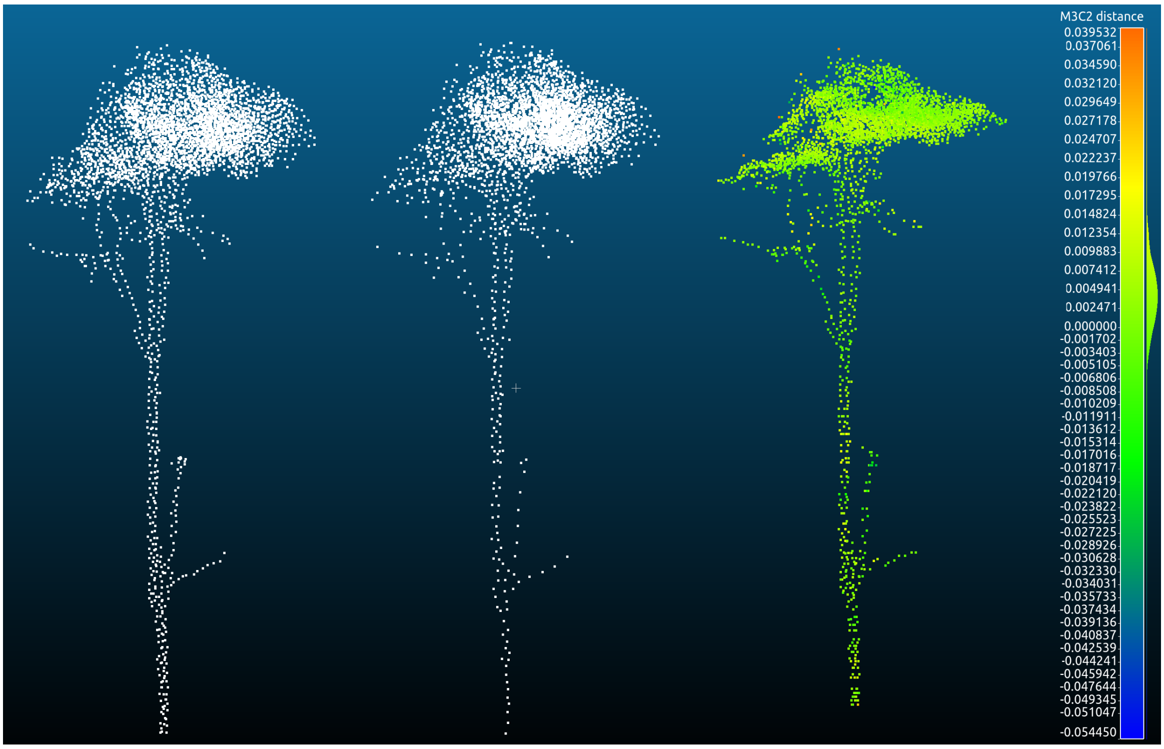


Figure 5. A tuple of tree point clouds. The third tree shows the Multiscale Model- to-Model Cloud Comparison (M3C2) metric, which is calculated between the real 3D data (represented by the first tree) and the generated 3D data produced by our method (represented by the second tree). In this visualization, green and yellow points signify a strong correlation between the generated points and the real data. Conversely, red and blue points are scarce in this particular example, indicating few significant structural differences.

https://doi.org/10.5194/isprs-archives-XLVIII-G-2025-619-2025 | © Author(s) 2025. CC BY 4.0 License.

ter model performance. Overall, the JSD metric, computed for the validation and test sets, is consistently low for the three tree species included in our evaluation.

Next, we calculated the M3C2 metric for each tree in our test set, comparing it to its corresponding generated counterpart produced by our model. Figures 3, 4 and 5 illustrate several examples of this metric. The first column in each figure presents the real tree, whereas the second column shows the generated tree. The third column displays the M3C2 metric for each point of the real tree. Insignificant M3C2 errors are marked in green and yellow, while significant deviations from the reference point cloud are indicated in red and blue.

Our analysis reveals that the M3C2 error is normally distributed around 0 for all tree species in our test set. This suggests that our model is capable of generating point clouds that capture the specific shape of individual trees from a given species.

4. Conclusion and Future Work

In this paper, we presented a novel unsupervised DL model for generating life-like 3D representations of specific tree species. By leveraging existing scanned data, our approach learned to produce diverse 3D data of trees with distinct shapes and structures. The intersection of DL and forest visualization presents a promising avenue for advancing the creation of virtual forests, which may aid the generation of digital forest twins of real forest landscapes. Digital twins provide manifold insights for scientific research as well as for decision support in forest management.

As a future work, we intend to enhance the density of the generated 3D data and further increase the network depth to improve the visualization of single trees and virtual forests. To enhance realism, the tree generator model presented in this paper can be adapted to and trained with new and bigger single-tree datasets. Finally, our current model lacks color representation for different tree species, so integrating it with photogrammetric data is a promising direction to explore. The long-term perspective would be to link the new 3D tree generator with the outcomes of forest simulation models for more realistic forest ecosystem visualizations.

5. Funding

This research was conducted as part of the Small4Good project, which received funding from the European Union Horizon Europe (HORIZON) Research and Innovation programme under Grant Agreement no. 101135517.

References

Achlioptas, P., Diamanti, O., Mitliagkas, I., Guibas, L., 2018. Learning representations and generative models for 3d point clouds. *International conference on machine learning*, PMLR, 40–49.

Bornand, A., Rehush, N., Morsdorf, F., Thürig, E., Abegg, M., 2023. Individual tree volume estimation with terrestrial laser scanning: Evaluating reconstructive and allometric approaches. *Agricultural and Forest Meteorology*, 341, 109654.

Bryson, M., Wang, F., Allworth, J., 2023. Using synthetic tree data in deep learning-based tree segmentation using LiDAR point clouds. *Remote Sensing*, 15(9), 2380.

Dobbs, H., Batchelor, O., Green, R., Atlas, J., 2023. Smart-tree: neural medial axis approximation of point clouds for 3d tree skeletonization. *Iberian Conference on Pattern Recognition and Image Analysis*, Springer, 351–362.

Eldar, Y., Lindenbaum, M., Porat, M., Zeevi, Y. Y., 1997. The farthest point strategy for progressive image sampling. *IEEE transactions on image processing*, 6(9), 1305–1315.

Holm, S., Schweier, J., 2024. Virtual forests for decision support and stakeholder communication. *Environmental Modelling & Software*, 180, 106159.

Hristova, H., Murtiyoso, A., Kükenbrink, D., Marty, M., Abegg, M., Fischer, C., Griess, V. C., Rehush, N., 2024. Viewing the Forest in 3D: How Spherical Stereo Videos Enable Low-Cost Reconstruction of Forest Plots. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.

Kükenbrink, D., Marty, M., Bösch, R., Ginzler, C., 2022. Benchmarking laser scanning and terrestrial photogrammetry to extract forest inventory parameters in a complex temperate forest. *International Journal of Applied Earth Observation and Geoinformation*, 113, 102999.

Lague, D., Brodu, N., Leroux, J., 2013. Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (NZ). *ISPRS Journal of Photogrammetry and Remote Sensing*, 82, 10–26.

Liang, X., Kankare, V., Hyyppä, J., Wang, Y., Kukko, A., Haggrén, H., Yu, X., Kaartinen, H., Jaakkola, A., Guan, F. et al., 2016. Terrestrial laser scanning in forest inventories. *ISPRS Journal of Photogrammetry and Remote Sensing*, 115, 63–77.

Mokroš, M., Liang, X., Surový, P., Valent, P., Čerňava, J., Chudý, F., Tunák, D., Saloň, Š., Merganič, J., 2018. Evaluation of close-range photogrammetry image collection methods for estimating tree diameters. *ISPRS International Journal of Geo-Information*, 7(3), 93.

Neudam, L. C., Fuchs, J. M., Mjema, E., Johannmeier, A., Ammer, C., Annighöfer, P., Paul, C., Seidel, D., 2023. Simulation of silvicultural treatments based on real 3D forest data from mobile laser scanning point clouds. *Trees, Forests and People*, 11, 100372.

Puliti, S., Lines, E., Müllerová, J., Frey, J., Schindler, Z., Straker, A., Allen, M. J., Lukas, W., Rehush, N., Hristova, H. et al., 2024a. FOR-species20K dataset.

Puliti, S., Lines, E. R., Müllerová, J., Frey, J., Schindler, Z., Straker, A., Allen, M. J., Winiwarter, L., Rehush, N., Hristova, H. et al., 2024b. Benchmarking tree species classification from proximally-sensed laser scanning data: introducing the FOR-species20K dataset. *arXiv preprint arXiv:2408.06507*.

Qi, C. R., Yi, L., Su, H., Guibas, L. J., 2017. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *Advances in neural information processing systems*, 30.

Zamorski, M., Zieba, M., Klukowski, P., Nowak, R., Kurach, K., Stokowiec, W., Trzciński, T., 2020. Adversarial autoencoders for compact representations of 3D point clouds. *Computer Vision and Image Understanding*, 193, 102921.