A Comparative Framework for deriving True Tree Crown (TTC) from Pseudo Tree Crown (PTC)

Namya Kamboj [1], Kongwen (Frank) Zhang [1], Baoxin Hu [2]

[1] School of Computing, University of the Fraser Valley, Abbotsford, Canada – Namya.Kamboj@student.ufv.ca, Frank.Zhang@ufv.ca
[2] Department of Earth and Space Science and Engineering, York University, Toronto, Canada – baoxin@yorku.ca

Keywords: Pseudo Tree Crown (PTC), True Tree Crown (TTC), 3D Reconstruction, Remote Sensing

Abstract

Recent advances in UAV-based remote sensing have made high-resolution 2D imagery widely available, however the extraction of 3D tree structure from such data remains a primary challenge. This paper presents a novel framework for deriving True Tree Crown (TTC) geometry from Pseudo Tree Crown (PTC) representations, through a graph-based learning model. The PTC is generated from single nadir RGB images by interpreting grayscale intensity as height. This serves as an intermediate 2.5D representation that bridges the gap between conventional imagery and full 3D structure. We establish a spatial correlation between PTC and LiDAR-derived TTC meshes using geometric feature extraction and correspondence analysis. Preliminary results on synthetic data demonstrate a strong correlation between PTC and TTC height distributions, confirming that PTC encodes meaningful structural information. To learn the mapping from PTC to TTC, we propose a Graph Neural Network architecture with three GraphConv layers (64 – 128 – 256 channels), residual connections, and a composite loss function combining Chamfer distance with Laplacian and edge regularization. This framework enables the estimation of complete 3D tree crowns from single RGB images, transforming vast historical 2D image archives into valuable 3D forest data for ecological monitoring, carbon accounting, and sustainable forest management.

1. Introduction

Forests play an essential role in sustaining global ecosystems by providing wildlife habitats, promoting species diversity, storing carbon, regulating nutrient cycles, and supporting numerous ecological processes. The management of tree species databases is therefore a vital step in monitoring and assessing biodiversity health. Accurate three-dimensional (3D) characterization of forest ecosystems is fundamental to addressing global challenges such as climate change mitigation, biodiversity conservation, and sustainable resource management (Zhang et al., 2016).

Traditionally, Light Detection and Ranging (LiDAR) technology has been the primary method for collecting this vital 3D information, offering direct and precise measurements. However, the significant cost, logistic complexity, and limited temporal and spatial availability severely restrict the large-scale implementation of LiDAR. In contrast, optical remote sensing platforms, especially Unmanned Aerial Vehicles (UAVs), routinely collect vast amounts of very high-resolution two-dimensional (2D) imagery. Moreover, 2D optical datasets benefit from extensive historical archives that LiDAR lacks.

The ultimate objective of our research is to reconstruct the true three-dimensional geometry of trees, referred to as the True Tree Crown (TTC), using only widely available 2D imagery. In essence, this involves estimating the "unobserved" spatial dimensions from 2D data. To advance this goal, we introduced the concept of Pseudo Tree Crown (PTC) modelling (Miao et al., 2024; Sharma and Zhang, 2024). PTC modelling generates a fixed-view-angle pseudo-3D (2.5D) representation from a single RGB image by interpreting greyscale intensity values as height information, thereby constructing an artificial vertical profile (Figure 1) (Zhang et al., 2025). This transformation enriches the information content of the image and reduces conditional entropy by decreasing uncertainty and randomness. As a result, the PTC provides a structural profile that enhances tree species classification accuracy and establishes a conceptual bridge between 2D imagery and the TTC.

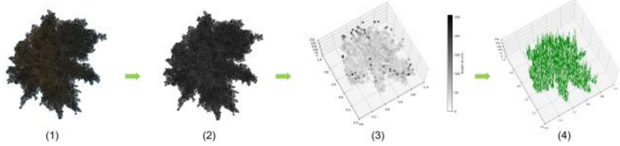

Figure 1. Generation of PTC from 2D nadir RGB image

This paper presents the framework for the current phase of our research, deriving TTC from 2D images through the intermediary of PTC, specifically by establishing the spatial correlation between PTC and TTC. A preliminary model, termed Spatial Vector Projection, was introduced by Balkenhol and Zhang (2017), in which the 3D crown structure was spatially aligned with LiDAR-derived TTC data. However, early work was limited to capturing longitudinal profiles due to computational constraints.

In this study, we explore the spatial relationship between the PTC and LiDAR-derived TTC to train advanced Machine Learning models. This will allow us to eventually generate a "reasonably representative" and "typical" 3D tree structure of a given species based solely on its 2D image input. Achieving this objective would represent a paradigm shift in remote sensing. It would unlock the potential to generate accurate 3D tree species structures for vast forested areas using existing, inexpensive 2D image archives, thereby providing unprecedented 3D forest analysis for ecologists, climate scientists, and foresters.

2. Literature Review

2.1 Individual Tree Species Classification and 3D Structural Information

Traditional forest analysis methods relied on spectral vegetation indices, for example, the normalized difference vegetation index

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLIX-B3-2026-135-2026 |

(NDVI), and empirical radiative transfer models, like SAIL at canopy level and PROSPECT at the leaf level (Zhang et al., 2025). While these approaches have proven to be valuable for broad-scale vegetation monitoring, they are limited to two-dimensional data and cannot capture the vertical structural complexity of trees.

LiDAR technology has been long considered as the gold standard for deriving detailed 3D forest structure because of its ability to capture direct range measurements with a strong precision. However, the costs and time involved in collection and storage of LiDAR data led to scarcity of data available, in terms of spatial and temporal availability. Additionally, LiDAR returns in dense forest canopies are often attenuated or deflected leading to a constraint in its effectiveness (Zhang et al., 2025). These limitations have motivated the search for alternative and cost-effective approaches to 3D forest characterization.

The novel pseudo tree crown (PTC) approach introduced by Miao et al. (2024) demonstrated an improvement in the accuracy and efficiency of urban tree classification by creating a pseudo-3D (2.5D) representation from single nadir RGB images. This study was further expanded by Zhang et al. (2025), which led to developing a pseudo-3D data representation from nadir UAV RGB imagery for individual tree classification across multiple conifer species, including pine, Douglas fir, spruce and aspen. Their results demonstrated that PTC can largely improve individual tree classification accuracy by up to 13%, reaching the high 90% range, while effectively handling challenges as snow-covered backgrounds. This proved that PTC is, in fact, a cost-effective and scalable alternative to LiDAR for classification tasks, further strengthening the approach to forest monitoring and supporting sustainable forestry.

2.2 UAV-Based 3D Reconstruction Methods

Beyond classification, significant research areas have focused on reconstructing complete 3D tree models from UAV imagery. A recent study conducted by Ghasemi et al. (2025) presented the application of geometry-based relationship constraints for point-cloud registration and unsupervised 3D reconstruction of tree structure in semi-arid forests using UAV photogrammetry. Their integrated framework focuses on combining dual-layer UAV photogrammetry, collecting data from both above and below the canopy, with an innovative geometry-based point cloud registration method. Unlike conventional approaches like Iterative Closest Point (ICP) and Random Sample Consensus (RANSAC), which tend to fail when overlap between reconstructions falls below 25-30%, their method uses spatial relationships among individual trees to align multi-view point clouds acquired under occluded and variable conditions (Ghasemi et al., 2025).

To further refine the reconstructed tree models, they developed an updated unsupervised Generative Adversarial Network (Denoise-GAN) which enable both noise reduction and structural completion without reliance on labelled training data. As a result, their models achieved high accuracy as compared to the reference data (root collar distance $R^2 = 0.93$, height $R^2 = 0.97$ and crown area $R^2 = 0.99$) (Ghasemi et al., 2025).

Similarly, researchers at the University of Sydney (*Forestry*, 2026) used aerially acquired dense point cloud datasets from airborne laser scanning to develop new 3D point cloud processing techniques for individual tree detection, segmentation, and assessment. These techniques are largely beneficial in enabling plot-level inventories of commercial plantations without requiring field crews in remote or hazardous locations and have also been expanded to analyse plantation tree health from high-resolution multi-spectral aerial imagery at a tree-by-tree level.

2.3 Deep Learning for 3D Point Cloud and Mesh Processing

Current advancements in geometric deep learning have opened new possibilities for processing 3D tree data. Xu et al. (2024) introduced a noel deep learning-based approach for rapidly generating tree models from canopy point clouds. Their methodology consists of two major steps: (1) generating 3D coordinates of tree nodes from the canopy point cloud using a designed node coordinate generation network, and (2) using a graph neural network (GNN) to predict node attributes and adjacency relationships between nodes. The discrete nodes are then connected using a minimum spanning tree algorithm along with predicted adjacency relationships to form the complete tree structure. Their method dramatically reduced the average canopy-to-tree reconstruction time from 7 minutes to less than 0.5 seconds while preserving visual authenticity and accurately matching tree canopy shapes (Xu et al., 2024).

This GNN-based approach is relevant to our work as it depicts the viability of learning-based methods for tree structure inference from point cloud data. However, unlike our approach which aims to predict full TTC geometry directly from PTC meshes, their method requires point cloud inputs.

2.4 Shadow and Illumination Handling in Forest Imagery

Irregular illumination and shadows present significant challenges for optical remote sensing in forest environments. In humid tropical regions, cloud shadows can further complicate the near-surface optical remote sensing process, thus leading to costly and repetitive surveys to maintain geographical and spectral consistency. Goncalves et al. (2024) proposed a correction method using U-net architecture trained on well-illuminated reference images augmented with artificially generated cloud shadows. Their method effectively corrected uneven illumination in near-infrared and true-colour RGB images, leading to more consistent NDVI patterns and improving classification accuracy by 2.5%. While their approach focuses on cloud shadows rather than the cast shadows within tree canopies that affect our PTC generation, it highlights the importance of illumination correction in forest remote sensing workflows.

2.5 Research Gap and Contribution

While significant progress has been made in UAV-based 3D reconstruction and deep learning for point cloud processing, several gaps remain. First, most existing methods require multi-view imagery or LiDAR data, limiting their applicability to the vast archives of single-view nadir imagery. Second, while the dual-layer UAV approach (Ghasemi et al., 2025) achieves high accuracy, it requires specialized below-canopy flights that may not be feasible in dense forests. Third, GNN-based methods (Xu et al., 2024) have been applied to point cloud data but not to PTC-derived mesh representations.

Expanding this research, our main goal is to evolve the PTC approach from being a classification tool to a bridge for 3D reconstruction. By establishing the spatial correlation between PTC and TTC and learning this mapping through graph neural networks, we aim to enable the estimation of complete 3D tree crowns from single RGB images, transforming vast historical 2D image archives into valuable 3D forest data assets.

https://doi.org/10.5194/isprs-archives-XLIX-B3-2026-135-2026

3. Proposed Methodology

Our methodology is divided into four distinct phases: 1) Data acquisition, PTC generation and species classification; 2) Matching PTC and LiDAR-derived TTC for training; 3) TTC from PTC estimation; 4) Evaluation and Validation.

3.1 Data acquisition

The dataset used in our current phase of study is the same as that described in Zhang et al. (2025). The study area is located south of Kamloops, British Columbia (50°34′19.345″N, 120°27′1.209″W), which is characterized by a typical North American boreal forest with four dominant tree species: fir, pine, spruce, and aspen.
These are the data collection specifications:

1. **Platform:** DJI M300 RTK drone with Zenmuse P1 camera
2. **Flight Altitude:** approximately 128 meters
3. **Capture date:** 17 February 2023
4. **Imagery type:** RGB (JPEG format)
5. **Coordinate system:** WGS84
6. **Dataset size:** 852 individual trees (as reported in Zhang et al., 2025)
7. **Public Availability:** CC 4.0 licensed, accessible via Roboflow Universe
8. **LiDAR companion data:** Available but only for TTC validation in this study

We had collected UAV tree data with both RGB and LiDAR available, which is shown in Figure 2.

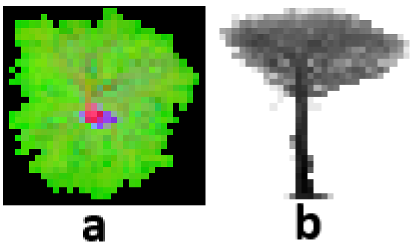

Figure 2. (a) nadir view of the tree; (b) LiDAR vertical of the same tree

The PTC was generated from the RGB image and went to classification using ResNet50 and YOLO v10 for precise individual tree species information.

3.2 Data Preprocessing and PTC Generation

For PTC generation, each individual tree image was standardized to 64x64 pixels (keeping it consistent with Miao et al., 2024). Then, the green band was extracted from the original RGB imagery to generate the grayscale basis for PTC construction (Figure 3). To remove the background noise, the threshold value is set to 180 pixels. Finally, the grayscale values were normalized to a [0,1] range to ensure consistent input scaling for the subsequent classification models.

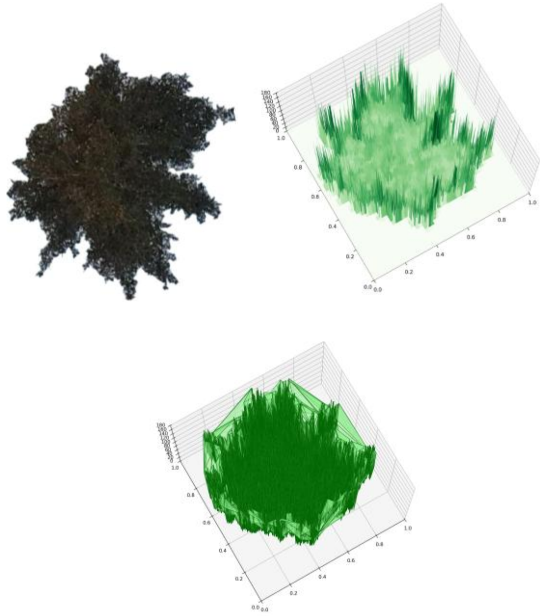

Figure 3. PTC mesh

3.3 Model Configuration and Training Strategy

This subsection details the implementation of ResNet50 and YOLOv10 for the initial PTC-based species classification, addressing reproducibility concerns.

3.3.1 ResNet50 Classification: We used a PyTorch implementation with a pre-trained ResNet50 backbone, where the final fully connected layer was adapted for four-class tree species output. The model was trained on 64x64 single-channel grayscale PTC images, following the input standardization established in earlier work (Miao et al., 2024). We used Stochastic Gradient Descent (SGD) with a momentum of 0.9 and weight decay of 5×10^{-4}, an initial learning rate of 0.001 managed by a cosine annealing scheduler, and a batch size of 32. The training proceeded for a maximum of 100 epochs with early stopping (patience of 10) based on validation loss, using Cross-Entropy as the loss function. The dataset was split into 70% for training, 15% for validation, and 15% for testing, consistent with the evaluation protocols in Zhang et al. (2025).
ResNet50 provides a well-established architecture for feature extraction from image-based representations like PTC. This results in a balance in performance and computation cost.

YOLOv10 for Tree Detection and Extraction: For object detection prior to PTC cropping, we used YOLOv10n (nano variant), due to its efficiency on UAV hardware and anchor-free design, which simplifies training and improves the inference speed (Zhang et al., 2025). YOLOv10 achieved competitive accuracy in individual tree classification tasks, along with PTC input further improving feature discrimination under challenging conditions such as overlapping canopies, or snow-covered backgrounds (Zhang et al., 2025). The model was pre-trained on the COCO dataset and fine-tuned on our annotated Kamloops dataset. Key detection parameters included:

- Input size: 640x640 RGB tiles
- Confidence threshold: 0.5
- IoU threshold: 0.45

Together, ResNet50 and YOLOv10 form a robust pipeline for species classification and individual tree extraction, using the enhanced feature representation of PTC as demonstrated in previous studies (Miao et al., 2024; Zhang et al., 2025).

3.4 PTC-TTC Spatial Correlation via Geometric Feature Mapping

This is the core analytical phase of the current work. The primary objective is to establish a quantitative spatial correlation between the Pseudo Tree Crown (PTC) and True Tree Crown (TTC) 3D representations. This process involves converting both data types into comparable mesh formats for geometric analysis.

3.4.1 Mesh Generation from PTC and TTC Data

PTC Mesh Construction: The normalized grayscale heightmap derived from UAV RGB imagery is converted into a 3D surface mesh using Delaunay triangulation (SciPy implementation). Then, the intensity value $I(x,y)$ of each pixel is interpreted as a height value, $z = I(x,y)$, creating a 2.5D heightmap. The triangulation connects adjacent pixels to form a piecewise-linear surface which represents the pseudo-crown structure of the tree.

TTC Mesh Reconstruction: LiDAR point clouds are processed to reconstruct the true 3D crown surface. Due to irregular and sparse nature of LiDAR returns in vegetation, Poisson Surface Reconstruction (vis Open3D) is applied to generate a complete surface mesh (*Surface Reconstruction - Open3D Latest (664eff5) Documentation*, n.d.). This method creates a smooth, continuous surface that preserves the architectural details of the actual tree crown while also handling occlusions and data gaps inherent in forest LiDAR scans.

3.4.2 Co-registration and Vertex Correspondence: Prior to correlation analysis, PTC and TTC meshes are aligned to a common coordinate system using centroid normalization:

$$ttc\_aligned = ttc\_vertices - mean(ttc\_vertices) + mean(ptc\_vertices)$$

For each vertex v_i in the PTC mesh, the nearest corresponding vertex v_i' in the TTC mesh is identified using KD-tree nearest neighbour search algorithm, with a maximum distance threshold of 0.1 normalized units. This creates a set of correspondence pairs $\{(v_i, v_i')\}$ for geometric comparison.

3.4.3 Geometric Feature Extraction: For each correspondence pair, we derive a feature vector capturing the spatial relationship between PTC and TTC:

$$F_i = [\Delta z_i, d_{xy,i}, r_i, \theta_i, \kappa_i]$$

where:

$\Delta z_i = z_{TTC,i} - z_{PTC,i}$ (vertical disparity)
$d_{xy,i} = \sqrt{(x_{TTC,i} - x_{PTC,i})^2 + (y_{TTC,i} - y_{PTC,i})^2}$ (horizontal distance)
$r_i = z_{PTC,i} / z_{TTC,i}$ (intensity-to-height ratio)
θ_i = angle between surface normals at v_i and v_i'
κ_i = curvature difference between correspondence regions

In addition to this, mesh-level features are extracted: convex hull values for PTC and TTC, height distribution percentiles (25th, 50th, 75th, 95th), canopy projection area in 2D, and surface area to volume ratios.

3.4.4 Correlation Metrics: To quantify the strength of the PTC-TTC relationship, we used the following metrics: Pearson correlation coefficient between PTC and TTC height distributions, Mean Absolute Error (MAE) of height predictions, and Root Mean Square Error (RMSE) of vertical disparities. Additionally, we employed Volume Ratio between convex hull approximations and Chamfer Distance between mesh surfaces.

3.5 TTC Estimation from PTC using Graph Neural Networks

Once a reliable geometric correlation is established, we train a predictive model to estimate TTC structure directly from PTC input.

3.5.1 Graph Representation of Tree Crowns: Both PTC and TTC meshes are represented as graphs $G = (V,E)$, where vertices V are the mesh points and edges E are the mesh connections. By implementing this representation, we can preserve the spatial relationships and local geometry necessary for crown structure modelling.

This phase involves training machine learning models, such as Graph Neural Networks (GNN) and diffusion model that operate precisely on mesh structures, or encoder-decoder networks, to learn the mapping function. Therefore, the model will take the PTC mesh as input and generate the prediction for full 3D structure of the corresponding TTC. The goal here is to effectively train the model to "imagine" the realistic 3D tree from its pseudo-3D representation.

For learning the mapping from PTC mesh to TTC mesh, we employ a Graph Convolutional Network (GCN) architecture (Zhang et al., 2019), designed specifically for 3D mesh data. The network processes each mesh as a graph G = (V,E) where:
V: vertices (3D coordinates + features like curvature, normals)
E: edges (mesh connectivity)

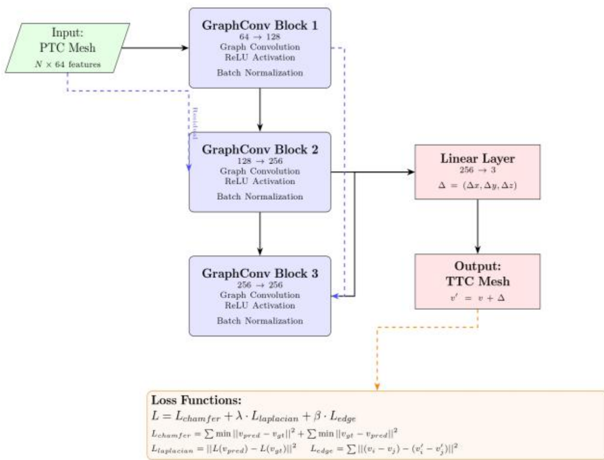

Figure 4. Proposed Graph Neural Network architecture for PTC-to-TTC prediction. The network takes a PTC mesh as input and outputs displacement vectors to transform it into a TTC mesh.

The proposed Graph Neural Network model (Figure 4) is designed to process tree crown meshes by representing each vertex with 64-dimensional feature vector. The network employs three sequential GraphConv layers, progressively increasing the channel dimensions from 64 to 128 and finally to 256. This allows the model to learn hierarchical features at multiple scales of the mesh structure. Residual connections are included between these layers to facilitate gradient flow during training and prevent degradation as the network deepens.

https://doi.org/10.5194/isprs-archives-XLIX-B3-2026-135-2026

Finally, the output layer generates three-dimensional displacement vectors (*Δx, Δy, Δz*) for each vertex, representing the transformation needed to create a mapping between the input PTC mesh and the target TTC mesh.
The model is optimized using a composite loss function that balances geometric accuracy with mesh quality. The primary component, L_chamfer, minimizes the Chamfer distance between the predicted and ground truth point sets to ensure spatial correspondence. This is augmented by a Laplacian regularization term (L_laplacian) weighted by λ. This preserves the local surface smoothness by encouraging neighbouring vertices to maintain consistent differential properties. An additional edge regularization term (L_edge) weighted by β helps maintain the original mesh connectivity and prevents the generation of unrealistic deformations or surface discontinuities. Training is conducted using the Adam optimizer with an initial learning rate of 0.001. This approach is selected for its adaptive moment estimation properties that work well with sparse gradient problems which are common in graph-based learning. A batch size of 8 meshes is used to balance memory constraints with training stability, while training proceeds up to 200 epochs with early stopping implemented to prevent overfitting. The dataset is partitioned into training (70%), validation (15%), and testing (15%) sets, making sure that the model’s performance is evaluated on unseen tree structures and maintaining sufficient data for robust learning across the four tree species present in the study area.

3.6 Evaluation and Validation

The performance of our PTC-to-TTC reconstruction pipeline will be evaluated using a combination of quantitative metrics and qualitative assessment, ensuring both geometric accuracy and structural credibility.
For the mesh-based comparisons, we use Chamfer Distance to measure the average closest-point distance between the vertices of the predicted TTC mesh and the ground truth TTC mesh. Lower score means higher levels of shape accuracy and surface similarity. The Earth Mover’s Distance (EMD) is useful to analyse the overall shape and volume distribution by calculating the effort needed to convert the point distribution of one mesh into another (lower values indicate better match). Also, we will calculate the average mesh-to-mesh distance between the predicted and ground truth meshes, which will act as a direct measure of surface deviation.
Beyond numerical scores, a visual and expert-based qualitative assessment is necessary. Visual Inspection involves a side-by-side comparison of the input PTC, the generated TTC mesh and the ground truth LiDAR model. This allows us to judge the “reasonableness” of the result, checking for branch structures and overall natural crown shape that a purely mathematical metric might miss.

4. Preliminary Results and Impact

Through the process of surface mesh comparison, we expect to establish a robust and quantitative spatial correlation between the PTC and TTC. The main outcome will be an estimated model capable of generating a “reasonably representative” 3D mesh of a tree’s structure from a 2D RGB image. The preliminary result is illustrated in Figure 5.

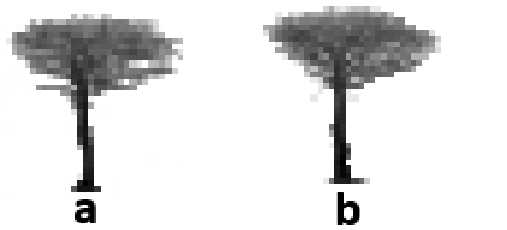

Figure 5. (a) the generated tree with full vertical profile via PTC and diffusion model; (b) the truth tree derived from LiDAR

To validate the fundamental premise of our proposed framework that “a quantifiable spatial correlation exists between PTC and TTC representations”, we conducted a proof-of-concept analysis using synthetic data designed to simulate relationship between nadir-derived pseudo crowns and LiDAR-derived true crowns. While full implementation of the GNN-based prediction model is ongoing, this preliminary investigation maintains the feasibility of the core hypothesis.

4.1 Correlation Analysis

The proof-of-concept dataset consisted of 1000 PTC vertices and 500 TTC vertices with correlated height distributions ($\rho = 0.85$ in the underlying generative model). Then using joint normalization in a shared coordinate space, correspondence analysis was performed using KD-tree nearest neighbour search with a distance threshold of 0.1 normalized units.

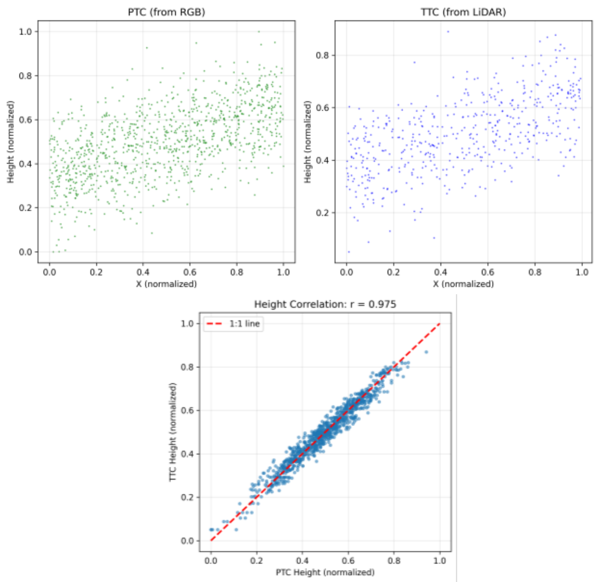

Figure 6. Height correlation between PTC and TTC for a sample tree (synthetic data) (a) PTC height distribution; (b) TTC height distribution; (c) Scatter plot showing correlation for the synthetic analysis (r=0.975)

The results, illustrated in Figure 6, show a strong correlation between PTC and TTC height values. A total of 923 corresponding point pairs were identified, yielding a Pearson correlation coefficient of $r = 0.975$ ($p < 0.001$). This high correlation confirms that the height information encoded in the PTC, derived from image intensity values, exhibits a strong linear relationship with the actual 3D height measurements from the reference TTC.
Figure 5 shows three complementary visualizations: (a) the PTC point cloud coloured by height, (b) the TTC point cloud, and (c) a scatterplot of corresponding heights with the ideal 1:1 line. The tight clustering around the diagonal confirms the strong correlation.

It is important to emphasize that these results are preliminary and based on synthetic data designed to approximate the expected relationship. Full validation on real Kamloops dataset imagery and LiDAR data is underway. However, these findings successfully demonstrate the conceptual viability of the PTC-TTC correlation framework and justify the continued development of the GNN-based prediction model.
Successful completion of this work will result in a foundational and economical pipeline for estimating 3D forest biometrics, like crown volume and structure, from inexpensive, widely accessible 2D image inputs. This will shift the focus from LiDAR-dependant methods, ultimately unlocking the potential of large-scale 3D forest monitoring, with remarkable benefits for forestry, ecology and climate science.

5. Conclusion

This research paper presents a systematic framework for establishing the spatial correlation between Pseudo Tree Crown (PTC) and True Tree Crown (TTC) representations, making a critical advancement beyond the current developments in PTC-based analysis. While previous work has demonstrated the effectiveness of PTC for specific classification tasks, we state that the ultimate value of this representation lies in its capacity to serve as a bridge between widely available 2D imagery and complete 3D structural information.
Our preliminary proof-of-concept analysis depicts the strong correlation between PTC and TTC height distributions, which confirms that the intensity-based height encoding is PTC preserves meaningful structural information. Our approach to mesh-based geometric feature extraction, correspondence analysis, and graph-based learning provides a template for similar problems where 2D-to-3D inference is required. The proposed GNN architecture, with its hierarchical GrpahConv layers, residual connections, and composite loss function, offers a principles approach to learning complex spatial transformations while preserving mesh topology and surface smoothness.
From a practical perspective, this research opens up an effective and economic pathway for large-scale 3D forest analysis. Achieving the TTC representation from simple 2D images, through the process of PTC intermediate representation, we unlock the potential of vast historical archives, collected over decades by government authorities, research institutions, and commercial operators, that currently remain underutilized for 3D analysis. Ecologists and foresters can transform these 2D archives into dynamic 2.5D or 3D forest models, enabling analytical studies of forest structure change, improved carbon stock estimation, and more accurate biodiversity assessments.
Looking forward, several directions for future work emerge from this study. Immediate priorities include validating the accuracy of our model on more real dataset imagery from different forest regions, with corresponding LiDAR data, which will provide quantitative assessment of the GNN's predictive accuracy. In the longer-term, we aim to incorporate multi-view constraints where available, develop uncertainty quantification for predicted TTC meshes, and extend the framework to handle multiple tree species simultaneously. Ultimately, we envision a fully automated pipeline capable of processing entire UAV flight campaigns to generate 3D forest inventories at unparalleled scales and resolutions.
In conclusion, this research demonstrates that PTC, originally conceived as a classification tool, can be elevated to serve as a bridge to true 3D reconstruction. By demonstrating the spatial correlation between PTC and TTC and also proposing a learning-based framework for this transformation, we lay the groundwork for a paradigm shift in forest remote sensing, one where 3D structural information becomes accessible from the vast archives of 2D imagery already in hand.

References

Balkenhol, L., Zhang, K., 2017: Identifying individual tree species structure with high-resolution hyperspectral imagery using a linear interpretation of the spectral signature. Proc. 38th Canadian Symposium on Remote Sensing (CSRS 2017), Montreal, QC, Canada.

Forestry. (2026). The University of Sydney. https://www.sydney.edu.au/engineering/our-research/research-centres-and-institutes/australian-centre-for-robotics/forestry.html

Ghasemi, M., Latifi, H., & Iranmanesh, Y. (2025). Geometry-based point cloud fusion of dual-layer UAV photogrammetry and a modified unsupervised generative adversarial network for 3D tree reconstruction in semi-arid forests. *Computers and Electronics in Agriculture*, *239*, 111024. https://doi.org/10.1016/j.compag.2025.111024

Goncalves, N. B., Nelson, B. W., Simonetti, A., and Stark, S. C.: Cloud shadow and uneven illumination detection and correction using the U-net architecture in near-surface images of complex forest canopies, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-3-2024, 191–196, https://doi.org/10.5194/isprs-archives-XLVIII-3-2024-191-2024, 2024.

Miao, S., Zhang, K., Zeng, H., Liu, J., 2024: Improving artificial-intelligence-based individual tree species classification using pseudo tree crown derived from unmanned aerial vehicle imagery. Remote Sensing, 16(11), 1849. doi.org/10.3390/rs16111849.

Sharma, R., Zhang, K., 2024: Conifer tree classification using ConvNeXt from UAV imagery data. Proc. 45th Canadian Symposium on Remote Sensing (CSRS 2024), Halifax, NS, Canada.

Surface Reconstruction — Open3D latest (664eff5) documentation. (n.d.). https://www.open3d.org/docs/latest/tutorial/Advanced/surface_reconstruction.html

Xu, J., Zhang, Z., Hu, X., & Ke, T. (2024). Accelerated forest modeling from tree canopy point clouds via deep learning. *International Journal of Applied Earth Observation and Geoinformation*, *132*, 104074. https://doi.org/10.1016/j.jag.2024.104074

Zhang, K., Robinson, J., Jing, L., 2016: Canopy vertical parameter estimation using unmanned aerial vehicle (UAV) imagery. Proc. 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China. doi.org/10.1109/IGARSS.2016.7729588.

Zhang, K., Zhang, T., Liu, J., 2025: Individual tree species classification using pseudo tree crown (PTC) in coniferous forests. Remote Sensing, 17(17), 3102. doi.org/10.3390/rs17173102.

Zhang, S., Tong, H., Xu, J. *et al.* Graph convolutional networks: a comprehensive review. *Comput Soc Netw* **6**, 11 (2019). https://doi.org/10.1186/s40649-019-0069-y