Urban Tree Classification from Multispectral Airborne LiDAR Using PointNet, DGCNN & RandLA-Net

Nada Hamdani[1,4], Imane Abouhat[2], Kenza Ait El Kadi[1,2], Saloua Bensiali[1,3] & Imane Sebari[1,2]

[1]Research Unit of Geospatial Technologies for a Smart Decision, Hassan II Institute of Agronomy and Veterinary Medicine. 10101 Rabat, Morocco - n.hamdani@iav.ac.ma; abouhat.imane@gmail.com; k.aitelkadi@iav.ac.ma; i.sebari@iav.ac.ma
[2]Department of Photogrammetry and Cartography, School of Geomatics and Surveying Engineering. Hassan II Institute of Agronomy and Veterinary Medicine. 10101 Rabat, Morocco
[3]Department of Applied Statistics and Computer Science - s.bensiali@iav.ac.ma
[4]Société Topographie Informatique, 91000 Evry Courcouronnes, France - nada.hamdani@sti-topographie.fr

Keywords: Tree Species Classification, Multispectral Airborne LiDAR, Deep learning, PointNet, DGCNN, RandLA-Net.

Abstract

Intelligent management of urban trees is a key issue for Smart Cities, contributing to environmental sustainability and urban well-being. Geospatial technologies and artificial intelligence are increasingly being integrated into smart cites to improve resource management and urban planning. This study provides an in-depth comparison of three deep learning methods: PointNet, DGCNN & RandLA-Net, applied to classification of seven urban trees species (Pine, Spruce, Birch, Maple, Aspen, Rowan, Linden) from the open-source Airborne Multispectral LiDAR dataset (MS-ALS-SPECIES). Each model was trained on a common training set, validated during training and evaluated on a separate test set, allowing a systematic evaluation of their classification performance. The comparison focuses on overall accuracy, F1-score, mean per-class accuracy and recall. The results demonstrate that PointNet achieves the best overall accuracy of test dataset of 82.07% and a mean per-class accuracy of 70.32%, with competitive performance on Pine (94.37%), Spruce (84.25%), and Maple (86.67%). DGCNN improves the capture of local structures, with 79.15% accuracy in validation, and 67.98% in testing, reflecting slight overfitting. RandLA-Net, although less accurate overall (56.03%), achieves the best inter-species homogeneity (62.32%), and high recall on minority species (Aspen: 86.36%, Linden 85.71%). These results demonstrate the potential of 3D deep learning combine with multispectral airborne lidar for automated urban tree species classification and their integration into the geospatial systems of smart cities for intelligent management of green spaces.

1. Introduction

Smart cities use advanced technologies, particularly Remote Sensing and Artificial Intelligence, to improve urban management, environmental sustainability and the quality of life (Rahman et al., 2023; Velasquez-Camacho et al., 2021).
In this context, identifying tree species with precision is a major challenge for sustainable urban planning. Trees plays an essential role in mitigating the urban heat island effect, sequestration carbon and improving of air quality(Alonzo et al., 2025).

Precise tree species classification is key to green management and smart urban planning.
Advances in Multispectral Airborne LiDAR and Deep Learning techniques, offer new perspectives on the automated analysis of urban vegetation. Exploiting these two techniques together enables the construction of intelligent systems capable of mapping and identifying urban trees on a large scale (Zhong et al., 2024).

This study proposes a comparison of three deep learning architectures for identifying tree species using an open-source multispectral airborne LiDAR dataset (MS-ALS-SPECIES)(Taher et al., 2025). The main aim is to evaluate their performance and efficacy in complex urban environments, with a view to developing decision-making tools for intelligent urban forest management.

2. Related works

Urban trees are important for city resilience. They mitigate heat islands and support biodiversity (Rao et al., 2025). Precise tree species classification is crucial for urban management, smart cities and development of digital twins with a high detail level (LoD4), exceeds limits of current urban models (Ambarwari et al., 2024).

Current studies exploit more and more airborne point clouds, especially those generated by multispectral airborne laser scanning (MS ALS), which provide detailed 3D structural information and spectral intensity information, whatever illumination conditions (Mustafić et al., 2025).

In addition, several researchers proved that the use of multispectral LiDAR for tree species classification improves tree species classification (Hakula et al., 2023), especially for species delineation with similar 3D structures but different spectral signatures (Gong et al., 2023; Qian et al., 2023; Taher et al., 2025);

Latest works show that deep learning methods applied directly to unstructured point clouds are most appropriate for classification tree species (Mustafić et al., 2025) and outperform traditional approaches bases on features extraction (Bayrak et al., 2023; Pierdicca et al., 2023).

However, few studies have benchmarked state of the art deep learning applied to point cloud, using multispectral airborne

LiDAR, particularly in urban environments, where species present similar structure (Taher et al., 2025).
This gap motivates the present study, which compare the performance of three deep learning models: PointNet, DGCNN and RandLA-Net, for classification seven urban tree species using the open-source multispectral airborne LiDAR dataset (MS-ALS-SPECIES).

3. Materials and Methods

This section presents the materials and methods that we used in our study. After selecting the dataset, we performed data preprocessing, trained three deep learning models (PointNet, DGCNN & RandLA-NET) and evaluated performance for each model. The figure 1 below resumes our workflow:

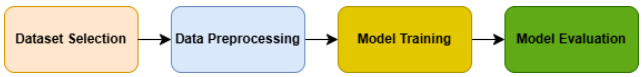

Figure 1. Workflow overview of the proposed method.

3.1 Study Area and Dataset

This study based at open-source dataset "MS-ALS-SPECIES" acquired by Optech Titan Multispectral LiDAR, collected in a suburban area of Espoonlahti in Finland. This benchmark dataset is designed to evaluate the classification of tree species from Multispectral Airborne LiDAR point clouds. It includes 6326 individual tree samples, representing nine species listed in the table 1 below. The point clouds were captured using a three-wavelength airborne LiDAR sensor: 1550 nm (channel 1), 1064 nm (channel 2), and 532 nm (channel 3). The system was mounted on a fixed-wing aircraft at an altitude of approximately 700 meters above ground (AGL), achieving point density of 35 points/m². Each individual tree is represented by several point clouds. Each point P_i in point cloud is represented by seven-dimensional vector containing the spatial coordinate (X, Y, Z), the intensities from the three scanners, and the echo return number.

Species code	Species name
1	Pine (*Pinus sylvestris*)
2	Spruce (*Picea sp.*)
3	Birch (*Betula sp.*)
4	Maple (*Acer platanoides*)
5	Aspen (*Populus tremula*)
6	Rowan (*Sorbus sp.*)
7	Oak (*Quercus robur*)
8	Linden (*Tilia sp.*)
9	Alder (*Alnus sp.*)

Table 1. Species and their respective code present in open-source dataset "MS-ALS-SPECIES"

3.2 Data Preprocessing

Data preprocessing regroups all steps involves transforming raw data into clean, consistent format that can be used by a deep learning model. It is a crucial phase in the deep learning pipeline, as the quality of the input data affects the performance of the model.

3.2.1 Filtering & Conversion:

Initially, the "MS-ALS-SPECIES" dataset contained 6326 individual tree point clouds in las format and an associated csv file containing the labels.
Firstly, we excluded samples without a corresponding species label, retaining 6171 trees.
Secondly, to ensure that each sample represented a distinct tree, only samples classified as "single tree" (profile_category=1 and profile_category_name = "Single tree") were selected. This filtering step yielded 4050 individual trees.
Thirdly, to mitigate class imbalance and ensure statistically reliable training, only species represented by at least 100 individual trees were retained.
After filtering, the final dataset comprised 3935 trees belonging to multiple species. The dataset was then randomly divided into training, validation, and testing subsets. Specifically, 75% of the samples (approximately 2,948 trees) were allocated to the training and validation sets, while the remaining 25% (approximately 987 trees) were reserved for testing. The training validation subset was split further into 80% for training (2,358 trees) and 20% for validation (590 trees), enabling robust model tuning and performance assessment. Figure 1 below illustrates the distribution of tree species included in this study.

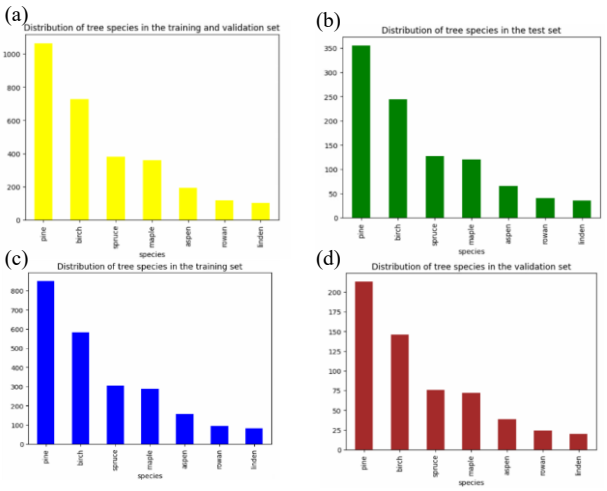

Figure 2. Distribution of tree species in (a) the training and validation set, (b) in the test set, (c) in the training set only and (d) in validation set only.

In order to improve computational efficiency and enable direct model input, all of the selected point clouds were converted into the NumPy (.npy) format.

3.2.2 Sampling:

A fixed number of points (npoints=2048) is randomly selected from the point clouds for each tree, to ensure that all model inputs are the same size. This sampling step is essential for architectures, which require uniform input dimensions. Random sampling reduces computational cost by avoiding the processing of all points in large data sets and improves model robustness by exposing it to different subsets of points during training. For this study, a batch size of 16 and an epoch of 100 training were employed to strike a balance between training time and model performance

3.2.3 Normalization:

Normalization ensures that the geometry and intensity of the point clouds are homogeneous, while preserving the relative relationships between points.
The spatial coordinates (X,Y,Z) of each point were centred by subtracting the center of mass of the cloud, thus returning the cloud to the origin of the coordinate system. This centering ensures that the model's learning is not influenced by the cloud's absolute position. Next, coordinate scaling was applied to control the spatial extent of the cloud. Each coordinate was

divided by the maximum distance from a point to the origin to ensure that the entire cloud fit within a unit-radius sphere.
For LiDAR intensity, each channel was normalized by a constant factor determined from the intensity distribution in the training set. This normalization brings the values back into a comparable range (approximately [0, 1]), thereby improving the stability and efficiency of the neural network's learning.

3.2.4 Data augmentation:

Several data augmentation techniques were applied to the input point clouds to enhance the robustness and generalization capability of the model. These techniques included:

- Random translation of the entire point cloud to simulate spatial positioning variations.
- Gaussian jitter was added to the coordinates of individual points to account for sensor noise and positional perturbations.
- Random scaling of coordinates to simulate trees of varying sizes and improve scale invariance;
- Random rotation around the vertical axis to increase invariance to orientation.

The diagram below, illustrates the adopted methodology applied in this study:

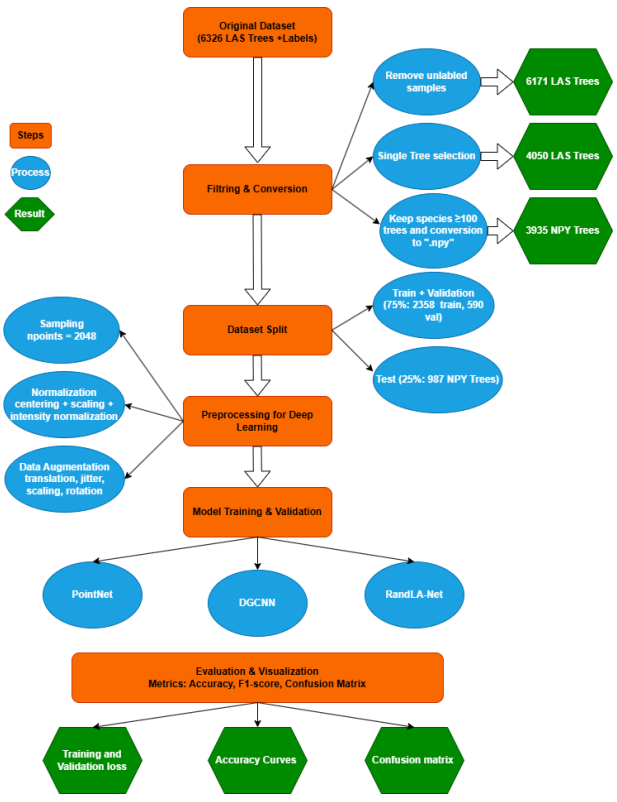

Figure 3. Adopted Methodology Applied for Urban Tree Species Classification from Multispectral Airborne LiDAR.

3.3 Deep Learning Architectures:

3.3.1 PointNet: is a deep learning model designed specifically to process raw, unordered point clouds for 3D classification and segmentation. Its architecture consists of three main components. First, shared multilayer perceptrons (MLPs) extract point-wise features, enabling the network to independently learn the local characteristics of each point. Second, a symmetric aggregation function implemented as max pooling ensures permutation invariance and generates a global shape descriptor. Third, alignment networks (T-Nets) are used to make the model invariant to geometric transformations, such as rotation and translation. The global descriptor is subsequently used for classification tasks. For segmentation, it is concatenated with local point features to predict labels for each point. Overall, PointNet provides an efficient, unified framework for deep learning on raw 3D point clouds that preserves geometric structure and is robust to permutation and geometric transformations (Qi et al., 2017).

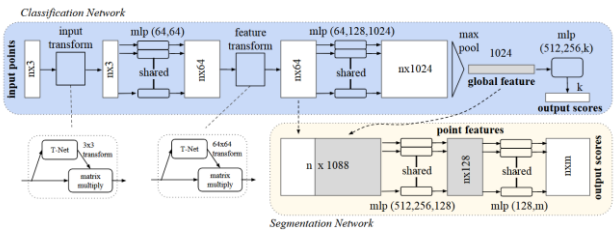

Figure 4. PointNet Architecture (Qi et al., 2017).

3.3.2 DGCNN: The Dynamic Graph Convolutional Neural Network (DGCNN) is a deep learning model introduces dynamic neighbourhood graphs (EdgeConv) designed to classify 3D point clouds. It dynamically constructs a k-nearest neighbour graph for each point at every layer to capture local geometric relationship. Local features are extracted using shared multilayer perceptrons on edge features, following by max-pooling over neighbors to obtain permutation-invariant point representations. Then, features from multiple layers are concatenated and passed through a final one-dimensional (1D) convolution and global pooling to form a global embedding. This embedding is then processed by fully connected layers to predict class scores. This architecture effectively preserves local and global structures while remaining robust to point permutations and geometric transformations (Wang et al., 2019).

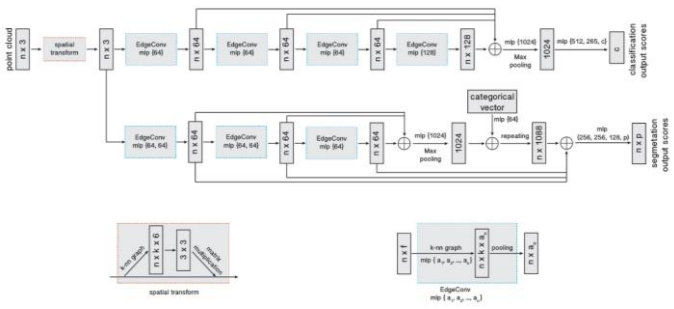

Figure 5. Dynamic Graph CNN's architecture (Wang et al., 2019).

3.3.3 RandLA-Net: Random Local Aggregation Network (RandLA-Net) is a deep learning model designed to efficiently process large-scale, 3D points clouds. It combines random point sampling with local feature aggregation to reduce computational cost while preserving local geometric details. For each point, neighbouring points are identified using the k-nearest neighbour method, and the relative position encodings are combined with the point features through shared multilayer perceptrons (MLPs) and attentive pooling. The encoder progressively extracts hierarchical, multi-scale features, and the decoder up samples and fuses these features to restore the original point resolution. For classification, point-wise features are aggregated into a global descriptor and processed through fully connected layers. RandLA-Net is fast, memory-efficient and scalable (Hu et al., 2020).

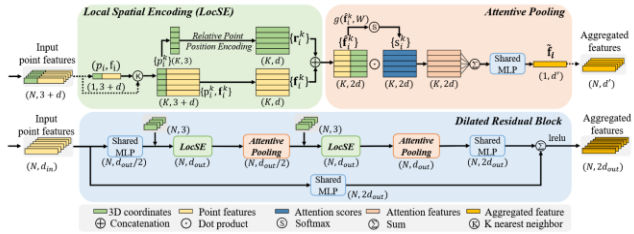

Figure 6. Local feature aggregation module of RandLA-Net (Hu et al., 2020).

3.4 Training and Evaluation

3.4.1 PointNet:
The architecture is based on a series of 1D convolutional layers applied independently to each point, allowing local features to be extracted from geometric coordinates (X, Y, Z) and spectral attributes (intensity, color, height, etc.). The network consists of five successive layers:
Conv1D (number of features → 64);
Conv1 D (64 → 64);
Conv1 D (64 → 64);
Conv1 D (64 → 128);
Conv1 D (128 → 1,024).

Each of these layers is followed by batch normalization and a ReLU activation function, which ensures gradient stability and better convergence. These convolution operations allow each point to learn a high-level representation independent of the order of the points in the cloud.

To aggregate local information into a global representation that is invariant to permutation, the model applies an adaptive max pooling operation that condenses the features from all points into a 1024-dimensional encoding vector. This vector represents the tree in a compact latent space.

Finally, this representation is passed to two fully connected layers:
Linear (1024 → 512).
Linear (512 → number of output channels).

With dropout for regularization and BatchNorm1D to improve generalization. The last output layer (7 output channels) corresponds to the number of tree species to be classified.

3.4.2 DGCNN: The process begins with the construction of a k-nearest neighbors (k-NN) graph for each point in the cloud. From these neighbourhoods, local features are extracted by calculating the difference between each central point and its neighbors, thus forming graph features rich in contextual information. These features are then processed by a series of 2D convolution layers (with dimensions 64, 64, 128, and 256), followed by LeakyReLU activation functions and batch normalization, which promotes learning stability.

The intermediate representations from each layer are concatenated to form a global vector, which is transformed by a 1D convolution into a 1,024-dimensional encoding space. Two global pooling operations (max and average) are then applied to aggregate the global information from the point cloud. Finally, three successive linear layers (512, 256, and number of output channels) perform the final classification using a LeakyReLU activation function and a dropout rate of 0.5 for regularization.

In this study, the model was parameterized with k = 20, a dropout rate of 0.5, and an embedding dimension of 1,024 in order to balance model complexity and convergence stability. This architecture allows both local and global features to be extracted, thus offering better generalization during multi-species tree classification based on three-dimensional LiDAR data.

The intermediate representations from each layer are concatenated to form a global vector, which is transformed by a 1D convolution into a 1,024-dimensional encoding space. Two global pooling operations (max and average) are then applied to aggregate the global information from the point cloud. Finally, three successive linear layers (512, 256, and number of output channels) perform the final classification using a LeakyReLU activation function and a dropout rate of 0.5 for regularization.

3.4.3 RandLA-Net:
RandLA-Net combines random sampling with careful local aggregation, which preserves essential geometric and spectral features while significantly reducing computational cost.

The network is based on an encoder-decoder structure.
In the encoding phase, each layer applies a Local Feature Aggregation (LFA) block, combining:

- a search for the K nearest neighbors (K = 32);
- relative position encoding;
- attention pooling that weights the influence of each neighbour.

These blocks progressively extract local features (16 → 64 → 128 → 256 channels) while reducing the density of the cloud by a decimation factor of 2.

In the decoder, the features are reconstructed via spatial interpolation (Nearest Neighbour Interpolation) and merged with the intermediate representations.

Finally, the aggregated features are condensed by adaptive global pooling and transmitted to an MLP classification head (256 → 128 → n_classes).

3.5 Results & Discussion:

The three deep learning architectures—PointNet, DGCNN, and RandLA-Net—were trained and evaluated on the multispectral LiDAR dataset for tree species classification. These models

differ in how they capture spatial and spectral relationships between points, resulting in varying performance across classes.

3.5.1 PointNet: The PointNet model showed rapid and stable convergence, with a final loss of 0.5107 and a validation accuracy rate of 82.20% at the hundredth epoch. The loss curve decreases steadily, without major oscillations, reflecting good learning stability and a limited risk of overfitting.

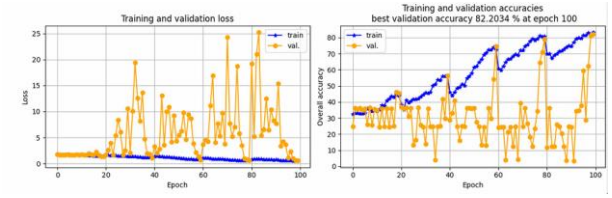

Figure 7. Training and validation loss and accuracy curves of the PointNet model.

On the test set, PointNet achieved an overall accuracy of 82.07% and an average accuracy per class of 70.32%, confirming its robustness. The confusion matrix reveals a very pronounced main diagonal, particularly for the Pine (94.37%), Spruce (84.25%), Birch (78.69%), and Maple (86.67%) classes. These classes are therefore perfectly recognized by the model.

The main confusions concern Aspen (57.58%), Rowan (45.00%), and Linden (45.71%), which are often poorly differentiated from each other, probably due to spectral and structural similarities between these deciduous species.

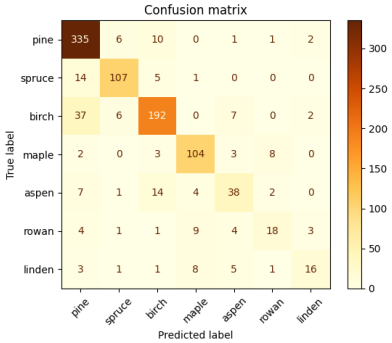

Figure 8. Confusion matrix of the PointNet model on the test dataset.

The macro F1-score of 0.73 and the weighted F1-score of 0.82 indicate that the model has strong generalization ability while maintaining a good balance between precision and recall. These performances confirm the effectiveness of PointNet for the global classification of structured point clouds, despite the absence of explicit exploitation of local relationships between points.

3.5.2 DGCNN :

The DGCNN model was trained for 100 epochs, achieving a training accuracy of 77.47%, a validation accuracy of 79.15%, and a final loss of 0.6535. The learning curves indicate correct convergence, but with a slightly higher loss than PointNet, reflecting greater optimization complexity due to dynamic graph construction.

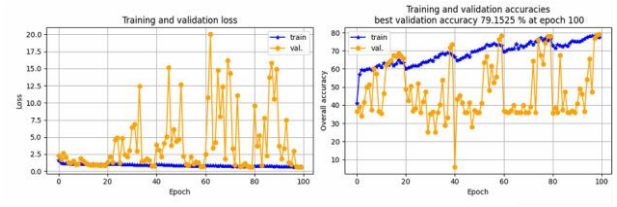

Figure 9. Training and validation loss and accuracy curves of the DGCNN model.

On the test set, DGCNN achieves an overall accuracy of 67.98% and a mean per-class accuracy of 52.53%, with a macro F1-score of 0.53.

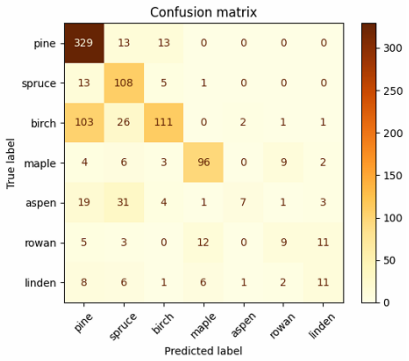

Figure 10. Confusion matrix of the DGCNN model on the test dataset.

The confusion matrix reveals high accuracy for the dominant species: Pine (92.68%), Spruce (85.04%), and Maple (80.00%), confirming the model's ability to capture local geometric structures using its dynamic neighborhood graphs.

On the other hand, very poor performance is observed for Aspen (10.61%), Rowan (22.50%), and Linden (31.43%). These results reflect a high sensitivity to class imbalance and a tendency for the model to overfit the majority classes.

The higher loss (0.63) and the divergence between training accuracy (77.5%) and test accuracy (68%) illustrate slight overfitting. Nevertheless, DGCNN remains effective in modeling fine spatial relationships, which could be leveraged on more balanced or larger datasets.

3.5.3 RandLA-Net:

The RandLA-Net model stood out for its remarkable stability during training, with a maximum validation accuracy of 84.58% at epoch 58 and a minimum loss of 0.4238.

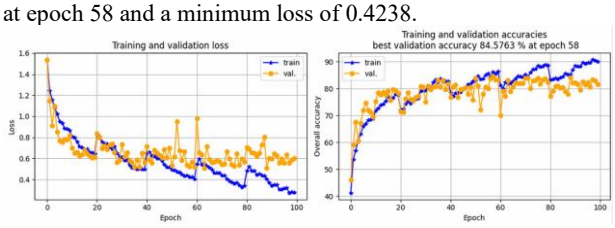

Figure 11. Training and validation loss and accuracy curves of the RandLA-Net model.

On the test set, performance drops to an overall accuracy of 56.03% and a mean per-class accuracy of 62.32%, suggesting high inter-species variability. The confusion matrix (Figure 10) reveals behaviour that is the opposite of PointNet: RandLA-Net overperforms for minority classes (Aspen: 86.36%, Linden: 85.71%) but fails to distinguish certain dominant species (Pine: 33.80%, Rowan: 20.00%).

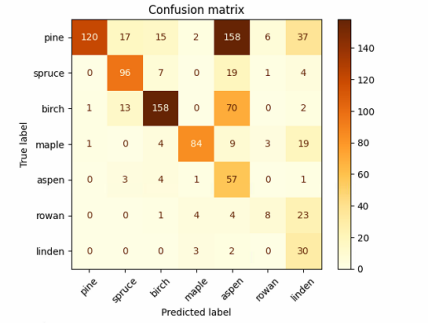

Figure 12. Confusion matrix of the RandLA-Net model on the test dataset.

This inversion stems from the random decimation mechanism used in the network, which tends to reduce point density and therefore spatial accuracy, but promotes better regularization and inter-species homogeneity.

The average F1-score of 0.53 confirms a balanced but overall inferior performance in terms of overall accuracy. RandLA-Net nevertheless remains the most robust model in the face of class imbalance, and the most suitable for dense and heterogeneous point clouds.

3.5.4 Global Analysis comparative

The loss and validation curves (Figures 5,7&9) confirm that PointNet and RandLA-Net converge quickly and stably, while DGCNN converges more slowly and exhibits oscillations due to the complexity of the graph.

The confusion matrices (Figures 6,8 & 10) illustrate the complementarity of the models:

- PointNet: achieved strong results for the dominant classes;
- DGCNN: showed efficient capture of local structure;
- RandLA-Net: demonstrated higher performance on minority classes.

The table 2 below regroups the main result od comparison of three architectures.

Model	Final Loss	Accuracy (%)	Mean Class Accuracy (%)	Macro F1	Strength points	Main limitations
PointNet	0.5107	82.07	70.32	0.73	Strong convergence and overall accuracy	Less sensitive to local relationships
DGCNN	0.6535	67.98	52.53	0.53	Good capture of local structures	Sensitive to overfitting
RandLA-Net	0.4238	56.03	62.32	0.53	Good balance between classes, robust in rare classes	Loss of spatial information

Table 2. Comparative Evaluation of PointNet, DGCNN and RandLA-Net on Urban Tree Species Classification

The results confirm that all three models can extract discriminating features from multispectral LiDAR data, but using different strategies.

PointNet stands out due to its superior overall accuracy (82.07%) and high learning stability, which makes it well-suited to homogeneous environments.

Despite its lower overall performance, DGCNN shows high potential for modelling local spatial structures in fine detail.

Finally, RandLA-Net offers better regularisation and inter-species balance, making it a promising model for classification in unbalanced or large datasets.

Joint analysis of loss functions, convergence curves and confusion matrices thus highlights the trade-off between accuracy, stability and generalisation.

Prospective developments include the adoption of hybrid architectures (e.g. PointNet++ or Point Transformer), which combine the simplicity of PointNet with the relational capabilities of DGCNN. Additionally, data augmentation and adaptive loss weighting techniques are being explored to improve the detection of underrepresented species.

3.6 Conclusion:

This study demonstrates the potential of deep learning approaches combined with Multispectral Airborne LiDAR data for accurately identifying seven tree species. By comparing three architectures: PointNet, DGCNN and RandLA-Net using the open-source Multispectral LiDAR dataset (MS-ALS-SPECIES). The research highlights the capacity of neural networks to capture the geometric and spectral information that is crucial for distinguishing between species on a city-wide scale. Experimental results confirm the specific strengths of

each model: PointNet demonstrated strong performance on dominant classes, DGCNN effectively captured local geometric structures, and RandLA-Net outperformed the other models on minority classes.

These findings emphasise the importance of advanced, Deep learning-based methods in supporting sustainable urban management and smart city strategies. While this study was limited to seven species and a single dataset, the methodology can be extended to more diverse datasets, paving the way for automated urban trees classification. Future work will focus on improving the models' ability to generalise across different urban contexts, incorporating temporal data to monitor tree health, and developing operational tools for large-scale urban forest mapping and decision support systems.

Acknowledgements

The authors would like to express their sincere gratitude to the "Communauté d'Agglomération Grand Paris Sud (CA GPS)", particularly the GIS division of the Department of Foresight, Territorial Observation and GIS (Direction de la Prospective, de l'Observation Territoriale et du SIG), for their valuable support and active involvement throughout this research.

We gratefully acknowledge for making the open-source Multispectral airborne laser scanning for species classification in Espoonlahti (MS-ALS-SPECIES) dataset publicly available (https://zenodo.org/records/17077256), which was fundamental to the completion of this work.

References

Alonzo, M., Ibsen, P.C., Locke, D.H., 2025. Urban Trees and Cooling: A Review of the Recent Literature (2018 to 2024). isa jauf.2025.023. https://doi.org/10.48044/jauf.2025.023

Ambarwari, A., Suwardhi, D., Rani, M.S., Husni, E., Junaidy, D.W., Agirachman, F.A., Murtiyoso, A., Griess, V.C., 2024. Conceptual Model of Graph-based Individual Tree and Its Utilization in Digital Twin and Metaverse of Urban Forest. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLVIII-4–2024, 7–12. https://doi.org/10.5194/isprs-archives-XLVIII-4-2024-7-2024

Bayrak, O.C., Erdem, F., Uzar, M., 2023. deep learning based aerial imagery classification for tree species identification. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLVIII-M-1–2023, 471–476. https://doi.org/10.5194/isprs-archives-XLVIII-M-1-2023-471-2023

Gong, Y., Li, X., Du, H., Zhou, G., Mao, F., Zhou, L., Zhang, B., Xuan, J., Zhu, D., 2023. Tree Species Classifications of Urban Forests Using UAV-LiDAR Intensity Frequency Data. Remote Sensing 15, 110. https://doi.org/10.3390/rs15010110

Hakula, A., Ruoppa, L., Lehtomäki, M., Yu, X., Kukko, A., Kaartinen, H., Taher, J., Matikainen, L., Hyyppä, E., Luoma, V., Holopainen, M., Kankare, V., Hyyppä, J., 2023. Individual tree segmentation and species classification using high-density close-range multispectral laser scanning data. ISPRS Open Journal of Photogrammetry and Remote Sensing 9, 100039. https://doi.org/10.1016/j.ophoto.2023.100039

Hu, Q., Yang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., Trigoni, N., Markham, A., 2020. RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds. https://doi.org/10.48550/arXiv.1911.11236

Mustafić, S., Schardt, M., Perko, R., 2025. Tree Type Classification from ALS Data: A Comparative Analysis of 1D, 2D, and 3D Representations Using ML and DL Models. Remote Sensing 17, 2847. https://doi.org/10.3390/rs17162847

Pierdicca, R., Nepi, L., Mancini, A., Malinverni, E.S., Balestra, m., 2023. uav4tree: deep learning-based system for automatic classification of tree species using rgb optical images obtained by an unmanned aerial vehicle. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences X-1-W1-2023, 1089–1096. https://doi.org/10.5194/isprs-annals-X-1-W1-2023-1089-2023

Qi, C.R., Su, H., Mo, K., Guibas, L.J., 2017. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. https://doi.org/10.48550/arXiv.1612.00593

Qian, C., Yao, C., Ma, H., Xu, J., Wang, J., 2023. Tree Species Classification Using Airborne LiDAR Data Based on Individual Tree Segmentation and Shape Fitting. Remote Sensing 15, 406. https://doi.org/10.3390/rs15020406

Rahman, A.U., Hoskere, V., 2024. Instance Segmentation of Reinforced Concrete Bridges with Synthetic Point Clouds. https://doi.org/10.2139/ssrn.4944550

Rao, P., Torreggiani, D., Tassinari, P., Rötzer, T., Pauleit, S., Rahman, M.A., 2025. Do urban green spaces cool cities differently across latitudes? Spatial variability and climatic drivers of vegetation-induced cooling. Sustainable Cities and Society 130, 106513. https://doi.org/10.1016/j.scs.2025.106513

Taher, J., Hyyppä, E., Hyyppä, M., Salolahti, K., Yu, X., Matikainen, L., Kukko, A., Lehtomäki, M., Kaartinen, H., Thurachen, S., Litkey, P., Luoma, V., Holopainen, M., Kong, G., Fan, H., Rönnholm, P., Polvivaara, A., Junttila, S., Vastaranta, M., Puliti, S., Astrup, R., Kostensalo, J., Myllymäki, M., Kulicki, M., Stereńczak, K., Pires, R. de P., Valbuena, R., Carbonell-Rivera, J.P., Torralba, J., Chen, Y.-C., Winiwarter, L., Hollaus, M., Mandlburger, G., Takhtkeshha, N., Remondino, F., Lisiewicz, M., Kraszewski, B., Liang, X., Chen, J., Ahokas, E., Karila, K., Vezeteu, E., Manninen, P., Näsi, R., Hyyti, H., Pyykkönen, S., Hu, P., Hyyppä, J., 2025. Multispectral airborne laser scanning for tree species classification: a benchmark of machine learning and deep learning algorithms. https://doi.org/10.48550/arXiv.2504.14337

Velasquez-Camacho, L., Cardil, A., Mohan, M., Etxegarai, M., Anzaldi, G., de-Miguel, S., 2021. Remotely Sensed Tree Characterization in Urban Areas: A Review. Remote Sensing 13, 4889. https://doi.org/10.3390/rs13234889

Wang, Y., Sun, Y., Liu, Z., Sarma, S.E., Bronstein, M.M., Solomon, J.M., 2019. Dynamic Graph CNN for Learning on Point Clouds. https://doi.org/10.48550/arXiv.1801.07829

Zhong, L., Dai, Z., Fang, P., Cao, Y., Wang, L., 2024. A Review: Tree Species Classification Based on Remote Sensing Data and Classic Deep Learning-Based Methods. Forests 15. https://doi.org/10.3390/f15050852