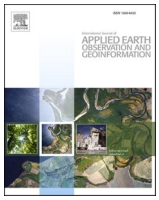
International Journal of Applied Earth Observation and Geoinformation 144 (2025) 104877

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag



Individual tree species prediction using airborne laser scanning data and derived point-cloud metrics within a dual-stream deep learning approach

Brent A. Murray [a,*], Nicholas C. Coops [a], Joanne C. White [b], Adam Dick [c], Ignacio Barbeito [a], Ahmed Ragab [d,e]

[a] Department of Forest Resources Management, University of British Columbia, 2424 Main Mall, Vancouver, British Columbia V6T 1Z4, Canada
[b] Canadian Forest Service, (Pacific Forestry Centre), Natural Resources Canada, 506 West Burnside Road, Victoria V8Z 1M5, BC, Canada
[c] Canadian Forest Service (Atlantic Forestry Centre), Natural Resources Canada, 1350 Regent Street, P.O. Box 4000, Fredericton, NB E3B 5P7, Canada
[d] CanmetENERGY, Natural Resources Canada 1615 Lionel-Boulet Blvd., P.O. Box 4800, Varennes, Québec J3X 1P7, Canada
[e] Department of Mathematics and Industrial Engineering, Polytechnique Montreal, Québec H3T 1J4, Canada

ARTICLE INFO

Keywords:
Lidar
Boreal forest
Convolutional neural network
Transformer
Attention mechanism
ITD

ABSTRACT

Accurate tree species mapping is essential for effective forest management but is often constrained by manual, labour-intensive workflows that limit scalability. While airborne laser scanning (ALS) supports large-scale forest attribute prediction, species classification remains difficult in complex, multi-species forests. To address this, we propose an automated, data-driven dual-stream deep learning framework that integrates ALS data with point-cloud metrics to identify individual tree species. Our framework incorporates an automated approach to individual tree segmentation and species labelling using existing forest inventory and field data, resulting in a dataset of 16,269 labelled individual tree point-clouds of four species across a 630,000 ha boreal mixed species forest in Ontario, Canada. Our dual-stream deep learning model integrates a Point Extractor to generate feature representations from raw ALS point-clouds and a complementary Metrics Network to process the point-cloud metrics. Results, based on the split test set of 2441 trees, showed that the inclusion of the Metrics Network improved tree species classification accuracy by approximately 11 % compared to models that rely solely on the Point Extractor. A weighted F1-score of 0.70 and area under the receiver operating characteristic curve of 0.88 was achieved using this dual-stream approach, along with enhanced predictive probabilities for all species thus improving the reliability of the predicted results. This approach reduces the manual processing bottleneck of individual tree segmentation and labelling and demonstrates the value of combining raw point-clouds and point-cloud metrics into a deep learning framework, offering a scalable and operational solution for reliable species predictions.

1. Introduction

A spatially explicit inventory of individual tree species is a vital resource for understanding and managing ecosystem services, including habitat (Prasad et al., 2020), carbon storage (Mohren et al., 2012), nutrient cycling, timber production (Rahlf et al., 2021), recreation, and cultural practices (Benner et al., 2021). Knowledge of the species and location of individual trees supports more effective forest stand management and silvicultural prescriptions, promoting sustainable forest management (Pretzsch et al., 2021).

Conventional forest inventories are spatial polygon-based products that incorporate a two-phase approach with manual aerial photo interpretation and stratified ground sampling (Leckie & Gillis, 1995). Traditionally, tree species information is recorded as leading species and the percentage of presence in the polygon, often determined by quantitative criteria, such as basal area. As such, the relative proportion of each species within the stand is known, also referred to as tree species composition. This allows identification of dominant and co-dominant species in a stand but lacks individual tree level detail. The accuracy of manually interpreted species identification is generally low (Magnussen & Russo, 2012; Thompson et al., 2007) or unassessed (Tompalski et al., 2021). Forest inventory polygons range in size, with the average stand size in Canada of 11 ha (White et al., 2025) with conventional inventories offering limited individual tree level species information required for some operational management practices.

Active and passive remote sensing technologies have shown promise

* Corresponding author.
E-mail address: brntmrry@student.ubc.ca (B.A. Murray).

https://doi.org/10.1016/j.jag.2025.104877
Received 25 June 2025; Received in revised form 17 September 2025; Accepted 19 September 2025
Available online 24 September 2025

for tree species prediction (Hermosilla et al., 2022; Queinnec et al., 2022). Airborne laser scanning (ALS) offers improved vertical profiling capabilities compared to passive sensors, capturing three-dimensional structural information that can contribute to accurate tree species prediction (Blomley et al., 2017; Ørka et al., 2012; Qian et al., 2023). ALS can be acquired efficiently over vast forested landscapes, providing critical data for enhanced forest inventories (EFI; White et al., 2025). EFIs typically follow an area-based approach (ABA; Næsset, 2002) and often have limited information on species (White et al., 2025). However, White et al. (2025) identified that integrating individual-tree based approaches into EFIs would enable targeted management practices for specific silvicultural requirements. Additionally, this information would allow for the accurate quantification of species abundance within a given area. Statistical point-cloud metrics commonly used in EFIs could highlight key structural and intensity-based characteristics relevant to species differentiation, leading to high prediction accuracies (Lin & Hyyppä, 2016; Vahrenhold et al., 2025; Z. Zhang et al., 2024).

With airborne laser scanning (ALS) becoming standard in operational forest inventories, point-based deep learning (DL) models have gained research interest for predicting forest attributes such as tree species (J. Chen et al., 2021; Hell et al., 2022; Liu et al., 2022a; Puliti et al., 2025). These models learn features from the spatial distributions of irregularly structured point-clouds and are able to capture species-specific patterns (Murray et al., 2024). Different architectures emphasize distinct patterns in the data, providing complementary views on the same input (C. Chen et al., 2024; D. Li et al., 2024; H. Zhang et al., 2023), which is particularly valuable for species classification, as they may identify different structural traits (e.g., canopy density, branching patterns). However, their performance is constrained by data quality, such as point density and noise levels, and can be limited by computational demands. Furthermore, there is often architectural and feature expansion variability which complicates direct comparisons.

Multimodal DL approaches improve performance by combining multiple data inputs or model outputs (Lahat et al., 2015; Ngiam et al., 2011). By merging learned features from different models, multimodal approaches have proven effective in remote sensing applications (Hong et al., 2021; B. Liu et al., 2023). A multimodal model combining different point-based approaches or integrating raw point-clouds with summary metrics can offer complementary insights, as explicitly including metrics may benefit learning, especially with limited data or model capacity.

A critical requirement for DL-based species prediction is a large, accurately labelled dataset, allowing for effective feature representations to be extracted. Diverse, well-annotated samples enhance model generalization, reduce overfitting, and improve performance (Barbedo, 2021). Many tree species classification studies rely on manual tree crown segmentation or species labelling (Briechle et al., 2021; Hell et al., 2022; Puliti et al., 2025), creating a bottleneck that limits dataset size and scalability. Automating detection, segmentation, and labelling would streamline data processing and enable large-scale, operational deployment.

The overall objective of this study is to develop an automated, data-driven dual-stream DL approach, one stream processing raw ALS point-clouds and the other stream processing derived point-cloud metrics, to identify and predict individual tree species within a complex, managed boreal forest. To achieve this objective, we explored three research questions:

1. How can an ALS-based automated individual tree segmentation approach be used to generate a comprehensive dataset of individual tree point-clouds with associated tree species labels for subsequent use in a dual-stream DL approach?

The use of automated, data-driven tree crown segmentation approaches for ALS point-clouds (Jing et al., 2012) enable the generation of large quantities of segmented point-clouds for use in DL models. However, a key challenge with these approaches is that crown segmentation itself does not assign species labels, and must be added in a separate step through manual annotation or classification algorithms (Dalponte et al., 2019). By leveraging existing forest inventory and field data, we aim to develop an approach that not only segments individual tree point-clouds but also assigns the corresponding species labels.

2. How do various point-based DL networks, and a combination of them, learn to extract different feature representations that are important for tree species prediction?

Point-based DL models leverage specialized layers to extract feature representations from point-clouds by capturing point-wise relationships and local geometric structures (Wang et al., 2019; Zhao et al., 2021). Through the adoption of a modular design to integrate various layers without altering the overall architecture, we aim to enable a fair and consistent evaluation of the layers for tree species classification.

3. How can statistical summary metrics derived from individual tree point-clouds be combined with features learned by a point-based DL model to improve tree species identification?

Point-based DL models can learn features directly from raw point-clouds for tree species classification (Murray et al., 2024), but may not fully capture aggregated or contextual information, especially when provided with limited data or noise associated with errors in the individual tree segmentation. Integrating summary point-cloud metrics provides complementary features that may not be captured by point-based DL (Kada & Kuramin, 2021; Lei et al., 2020). By combining features from ALS point-clouds and derived metrics, we aim to enhance input representation and predictive performance in a practical, context-aware manner.

2. Data and methods

2.1. Study area

The Romeo Malette Forest (RMF; Fig. 1) is a 630,000 ha managed forest located in Ontario, Canada on Treaty 9 territory (also known as the James Bay Treaty). The RMF is located on the traditional lands of the Ojibway, Cree, Anisininew, and the Métis Nation of Ontario. It is part of the Boreal Shield Ecozone and has a forested area of 582,430 ha, which has undergone disturbances including pest infestations (e.g., spruce budworm) and wildfire events that have affected the forest landscape. Four of the most abundant tree species within the RMF are black spruce (*Picea mariana*), jack pine (*Pinus banksiana*), trembling aspen (*Populus tremuloides),* and paper birch (*Betula papyrifera*). The production of timber and fibre procurement are the main forest products managed in the RMF (Queinnec et al., 2022).

2.2. Datasets

2.2.1. Single-photon lidar

Single-photon lidar (SPL) ALS data was acquired across the RMF in June and July of 2018 (OMNRF, 2019) using a Leica SPL100 instrument (Leica, 2024), which emits a 10 × 10 array of green ($\lambda = 532$ nm) beamlets (Mandlburger et al., 2019) and has been used in many forestry studies (Irwin et al., 2021; Murray et al., 2024; Queinnec et al., 2022). Flights were conducted at a 3800 m altitude, yielding a return density of 40 points/m^2 after processing per Gluckman (2016). Point-clouds were height-normalised using a triangular irregular network of ground classified returns, and returns below 2 m were removed, minimizing the influence of non-tree vegetation.

2.2.2. Generation of labelled individual tree point-clouds

A dataset of labelled individual tree point-clouds was generated

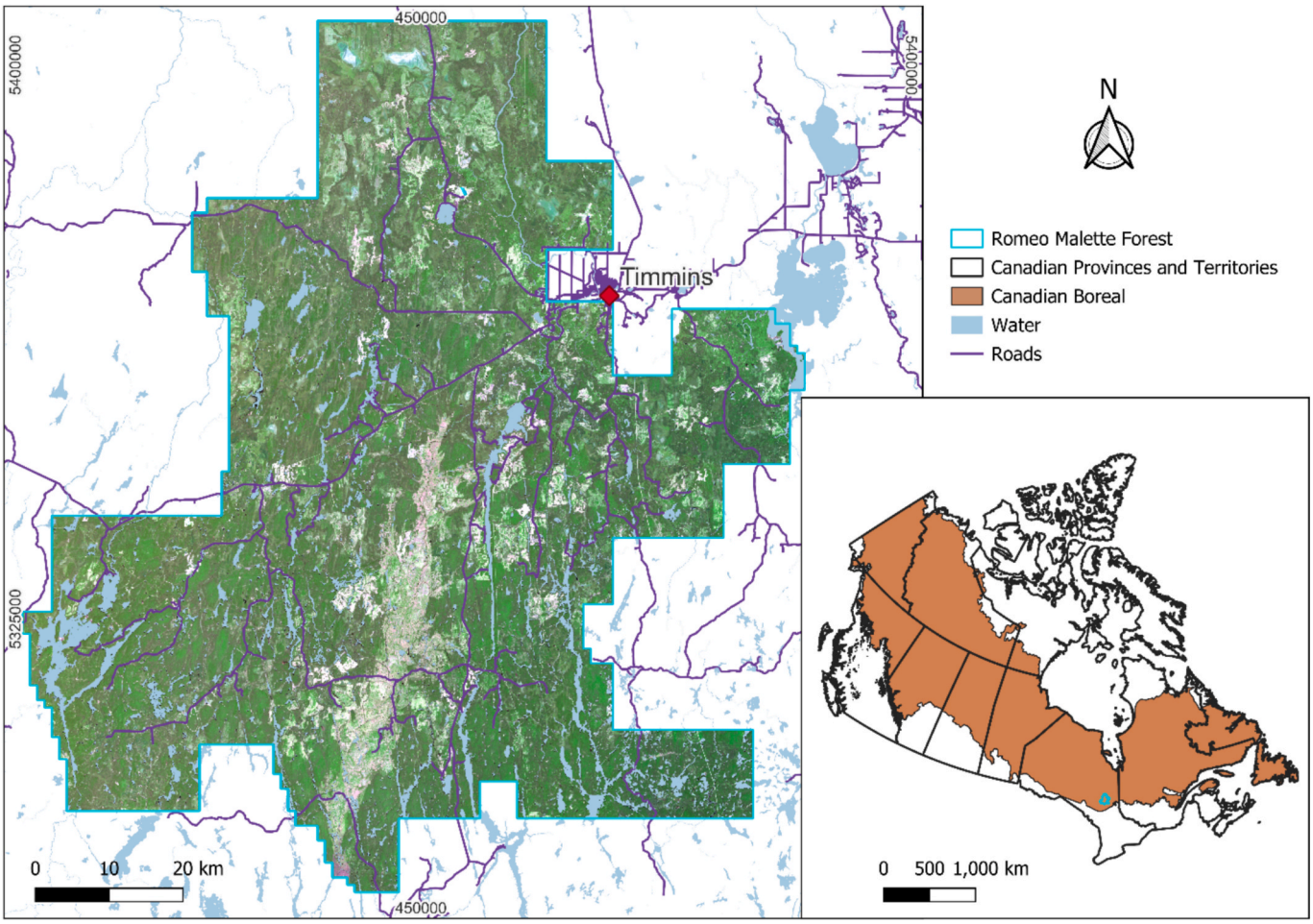


Fig. 1. Romeo Malette Forest (RMF) located in Ontario, Canada, shown with a true-colour composite. The RMF is shown within the Canadian Boreal forest and is outlined in cyan. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

using an automated individual tree detection (ITD), segmentation, and labelling workflow. This dataset generation process required no manual intervention from the ITD to the assignment of species labels, reducing the bottleneck of manual approaches. Field data from 53 circular plots (11.28 m radius or 400 m^2) were stem mapped and included tree species, height, diameter at breast height, and plot coordinates obtained from a EOS Arrow 100 GPS receiver (Queinnec et al., 2022). To supplement this dataset, pseudo-plots (11.28 m radius) were randomly generated within the forest resources inventory (FRI) polygons of the RMF (OMNRF, 1996, 2001) that were larger than 10,000 m^2, with 100 % single-species composition. The pseudo-plots were at least 50 m from polygon edges, reducing any edge effects of neighbouring forest stands. The FRI data is updated when changes occur, ensuring that the stand information used for the pseudo-plots is aligned with the ALS data. Additionally, any pseudo-plots that were identified as disturbed by harvest or fire through the National Terrestrial Ecosystem Monitoring System (Hermosilla et al., 2016) as early as 2005 were not included.

Point-clouds were then extracted from both sets of plots followed by ITD using a local maximum filter algorithm on a 25 cm resolution canopy height model (Popescu & Wynne, 2004; Queinnec et al., 2022). The detected trees were then used in the watershed segmentation algorithm (Van Der Walt et al., 2014), which was applied to the canopy height model. Point-clouds were then extracted to each segment's boundary. This approach was chosen for its simplicity, computational efficiency, and because it does not require manual segmentation of the point-clouds, which is beneficial for large areas like the RMF.

For the field plots, only trees from the ITD process within 5 m horizontally and 3 m in height of the stem-mapped trees were retained and labelled with the corresponding species. We worked under the assumption that every tree within pure stand pseudo-plots belongs to a single species and was considered a valid sample and labelled with that species. These steps (ITD, segmentation, and labelling) were integrated into a streamlined pipeline, enabling the automated generation of labelled individual tree point-clouds without manual processing. Segmented point-clouds with < 100 returns were discarded. Finally, classes with < 50 samples were removed to ensure a sufficient representation of each species, leaving four species: black spruce, jack pine, paper birch, and trembling aspen. Remaining samples were split into training, validation, and test sets (70:15:15) using a stratified random sampling approach to ensure balanced species representation. To ensure the samples were spatially disjoint, the dataset was split such that individual trees from the same field plot or pseudo-plot were assigned exclusively to either the training, validation, or test set. This provides a more realistic estimate for model generalization.

2.2.3. Statistical point-cloud metrics

Metrics related to crown height, size, and complexity, as well as signal attenuation depth and intensity, were derived for each segmented point-cloud (Grubinger et al., 2023; Queinnec et al., 2022). These metrics were selected to represent vertical distribution (e.g., Z_{q25}, Z_{skew}), canopy geometry (e.g., Rumple, Volume), and radiometric variation (e.g., I_{max}, I_{std}) of the trees. While SPL intensity differs from linear-mode lidar, it may help discriminate between needleleaf and broadleaf species (Räty et al., 2022). To ensure low co-linearity, variance inflation

factors (VIF; Perktold et al., 2024) were computed iteratively, removing the metric with the highest value until all values were below 10. Final metrics are listed in Table 1.

2.3. Model architecture

We employ a dual-stream architecture using both segmented raw point-clouds and derived point-cloud metrics (Table 1). The Point Extractor is a modular network allowing for different point-based networks to be incorporated to process the raw point-cloud to extract geometric and spatial feature representations. The Metrics Network encodes inter-metric dependences based on the point-cloud metrics. Outputs from both streams, each providing 320 feature representations, are concatenated and passed to a final classifier for tree species prediction.

2.3.1. Point-based stream

The Point Extractor ingests segmented tree point-clouds and incrementally transforms them into higher dimensional features (Fig. 2). These features are learnt representations providing information about the local geometry and contextual relationships across the ALS returns. Based on the dynamic graph convolutional neural network (DGCNN) framework (Wang et al., 2019), it applies a series of point-based layers, with layer outputs aggregated using global max pooling to yield a fixed-size global representation, which will feed into the feature fusion and classification module downstream. The key difference is its modular design enabling different point-based layers to be used without changing the overall architectural flow or feature expansion structure, allowing a direct comparison of different layers.

Three different point-based feature extraction layers were tested to better understand how variations in the feature representation affect the modeling of point-cloud data. First, EdgeConv (Wang et al., 2019) learns edge features through graph-based convolutions. Second, Point Transformer (Zhao et al., 2021), uses attention with positional encoding to dynamically weight features spatially. Finally, a multi-branched fusion layer is tested, which is a parallel ensemble of EdgeConv and Point Transformer, providing diverse perspectives on the point-cloud (Hong et al., 2021; Polikar, 2006).

2.3.2. Metrics-based stream

The Metrics Network (Fig. 3) processes the point-cloud metrics to learn informative features for predicting tree species. Inspired by TabNet (Arik & Pfister, 2020), this stream splits learned features into attention and decision branches, with each step applying two rounds of linear layers, batch normalisation, and gated linear units (GLUs) to selectively control information flow. It uses an internal guide, which is a running product of previous attention masks, that is iteratively updated at each step to regulate feature reuse. Instead of feeding attention masks directly into the decision branch, the internal guide modulates each new mask by scaling it with the cumulative influence of past masks, effectively discouraging the model from repeatedly attending the same features. This mechanism promotes a more diverse feature selection across decision steps, improving generalization and interpretability through learned feature importance weights. The decision outputs from each step

Table 1
Description of the point-cloud metrics used.

Metric	Description
N_{points}	Total number of returns
Z_{q25}	25th percentile of return heights
Z_{skew}	Skewness of return heights
Z_{kurt}	Kurtosis of return heights
$Z_{entropy}$	Evenness of the distribution of return heights (Shannon diversity index)
$Depth_{q25}$	25th percentile of signal attenuation depth of returns
$Depth_{q50}$	50th percentile of signal attenuation depth of returns
$Depth_{q75}$	75th percentile of signal attenuation depth of returns
Area	Area of the 2D convex hull of returns
Volume	Volume of the 3D convex hull of returns
Rumple	Ratio of canopy surface area to projected area
I_{max}	Max intensity of returns
I_{std}	Standard deviation of intensity of returns
I_{q50}	50th percentile of return intensity
$Apex_{mean}$	Mean of the angles (degrees) from the highest return (apex) to all other returns
$Apex_{std}$	Standard deviation of the angles (degrees) from the highest return (apex) to all other returns

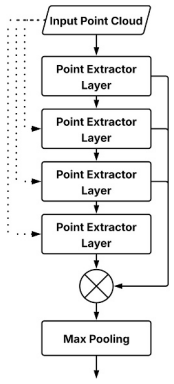


Fig. 2. Point Extractor architecture. Dotted lines represent the inclusion of the input point-cloud into the layer when the Point Transformer layer is used.

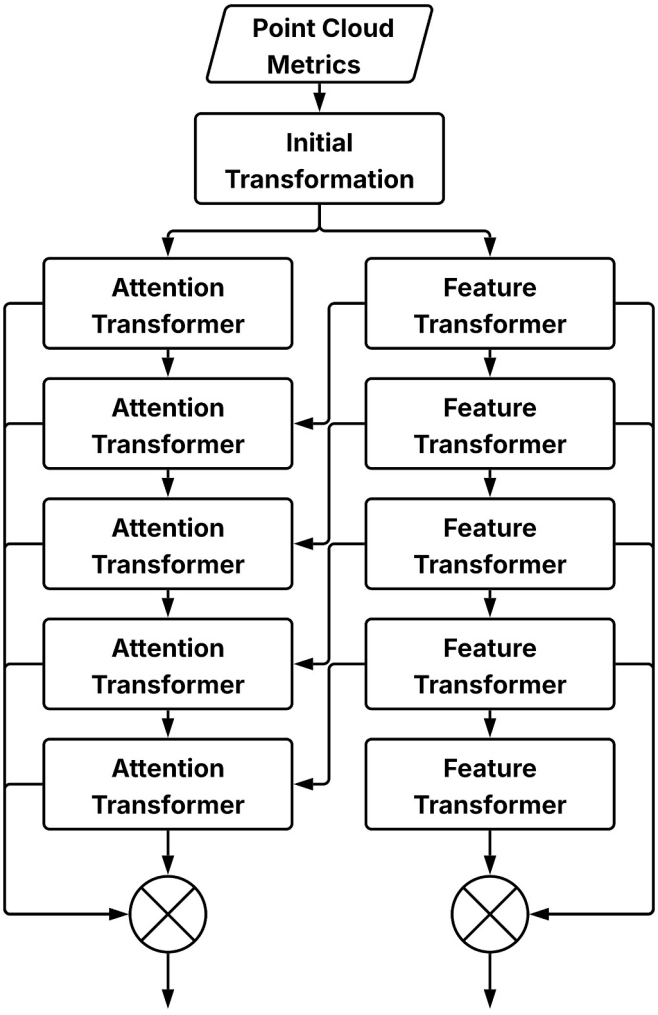


Fig. 3. Metrics Network architecture, outputting both decision features and attention features.

are concatenated to form the final feature representation that will be joined with the point-based features downstream. These aggregated representations capture both shared and step-specific information, enabling the model to learn complex patterns about the input metrics.

2.3.3. Feature fusion and classification

Features from both streams are concatenated and passed into a feedforward classifier with three sequential linear layers, rectified linear unit (ReLU) activations, and 50 % dropout to mitigate overfitting. The final softmax layer outputs a predictive probabilities vector, indicating the predicted tree species and associated confidence. This fused approach leverages both spatial structure and statistical properties of the point-cloud, enhancing model accuracy and generalization across tree species.

2.3.4. Loss function

We trained our model using a composite loss function that combines weighed cross-entropy for class predictions and centre loss for discriminative feature learning. This loss encourages accurate classification, compact intra-class feature clustering, while handling class imbalance. The composite loss is defined as:

$$L = L_{CE} + \frac{1}{2}\left(L_{\text{center}}^{(x)} + L_{\text{center}}^{(m)}\right)$$

where L_{CE} is the weighted cross-entropy loss and $L_{\text{center}}^{(x)}$ and $L_{\text{center}}^{(m)}$ are the center loss for the point-based and metric-based features, respectively.

The weighted cross-entropy is defined as:

$$L_{CE} = \frac{1}{N}\sum_{i=1}^{N} -w_{y_i} log\left(\frac{\exp\left(z_{i,y_i}\right)}{\sum_{j=1}^{C}\exp\left(z_{i,j}\right)}\right)$$

where w_{y_i} is the class-specific weight for true class y_i.

The centre loss (Wen et al., 2016) penalizes the distance between feature embeddings f_i and class centres c_{y_i}:

Here, z_{i,y_i} is the logit corresponding to the true class y_i for sample i. The centre loss, first introduced by Wen et al. (2016), encourages feature embeddings f_i to be close to their corresponding class centres c_{y_i}. With a scaling factor λ (set to 0.0001) to ensure a similar scale to the cross-entropy loss it is computed as:

$$L_{\text{center}} = \frac{\lambda}{2N}\sum_{i=1}^{N}|f_i - c_{y_i}|^2$$

This composite loss improves robustness and class separability, especially in imbalanced datasets.

2.4. Computing resources and model hyperparameters

All models were implemented using the PyTorch library (Version 2.5.1; Paszke et al., 2019) and trained on a 24 GB NVIDIA RTX A5000 GPU and 128 GB of RAM. The number of trainable parameters for each model is listed in Table 2. The batch size for each model was set to 12, and training was conducted for a maximum of 300 epochs. Early stopping was employed, stopping training if no improvements in the validation loss were observed for 20 consecutive epochs. The model weights and biases corresponding to the epoch with the lowest validation loss were saved and subsequently used on the testing dataset. We used the Adam optimizer with an initial learning rate (LR) of 0.001. A scheduler reduced the LR by 0.5 after three consecutive epochs without loss improvement, continuing until it reached 0.000012. At this point, a cosine annealing schedule (Loshchilov & Hutter, 2017) was applied, decaying the LR to 0.000008 over a 10-epoch cycle, with each subsequent cycle doubling in length. This adaptive strategy helps prevent premature convergence and improves generalization (Goodfellow et al., 2017). A visualisation of the training and validation losses for each model can be seen in Fig. S1.

Table 2
Number of trainable parameters of each model.

Model	Trainable Parameters
EdgeConv	166,020
Point Transformer	262,468
Multi-Branch	361,732
EdgeConv + Metrics Network	640,020
Point Transformer + Metrics Network	736,468
Multi-Branch + Metrics Network	835,732

2.5. Experiments and model assessment

To evaluate how different point-based DL architectures affect species prediction, we conducted an ablation study using various Point Extractor layers (Fig. 4). All models were trained and evaluated under identical conditions, using the same datasets, resources, and hyperparameters. Performance was assessed on a shared test set using overall accuracy, weighted F1-score, precision, recall, and area under the receiver operating characteristics curve (ROC AUC). We also compared pre-species accuracies and predictive probabilities across models. To examine the influence of the point-cloud metrics, we trained three additional models (one per Point Extractor layer) without the Metrics Network (Fig. 4b), adjusting the centre loss accordingly. Finally, we analyzed metric importance by comparing attention-derived rankings across models to understand how feature relevance varied with different Point Extractor layers.

3. Results

3.1. Segmented and labelled individual tree point-clouds

The process of segmenting and labelling individual tree point-clouds resulted in a total of 16,269 samples being generated (Table 3), with 11,388 for training, 2440 for validation, and 2441 for testing. 937 of those samples were extracted from the stem-mapped plots and 15,332 from the pseudo-plots. A majority of these samples were identified as either black spruce or jack pine with 6178 and 6001 segmented trees respectively, while trembling aspen and paper birch had fewer segmented trees with 3271 and 819 samples respectively. The mean number of returns per tree was highest for trembling aspen (922) and paper birch (839), followed by black spruce (635) and jack pine (633).

3.2. Ablation study evaluation

Evaluation scores for each of the trained models on the test dataset are presented in Table 4. In all cases, the inclusion of the Metrics Network in the dual-stream model improved evaluation scores. For example, the model using the EdgeConv layer had the lowest value across all scores when trained with the Point Extractor alone; however, with the addition of the Metrics Network, the model achieved the highest scores among all configurations. All model configurations that incorporated the Metrics Network had similar performance outcomes.

3.3. Per-species evaluation

Confusion matrices showing the classification percentages of the species samples for each trained model on the testing dataset are shown in Fig. 5. Overall, the inclusion of the Metrics Network stream led to an increase in the proportion of correctly classified samples for all tree species. The only exception was observed with the EdgeConv Point Extractor layer, where the proportion of correctly classified trembling aspen samples decreased from 56.12 % to 39.39 % when the Metrics Network was added. The two needleleaf species (jack pine and black spruce) showed higher classification accuracy, with correctly identified proportions ranging from 74 % to 80 %, while the broadleaf species

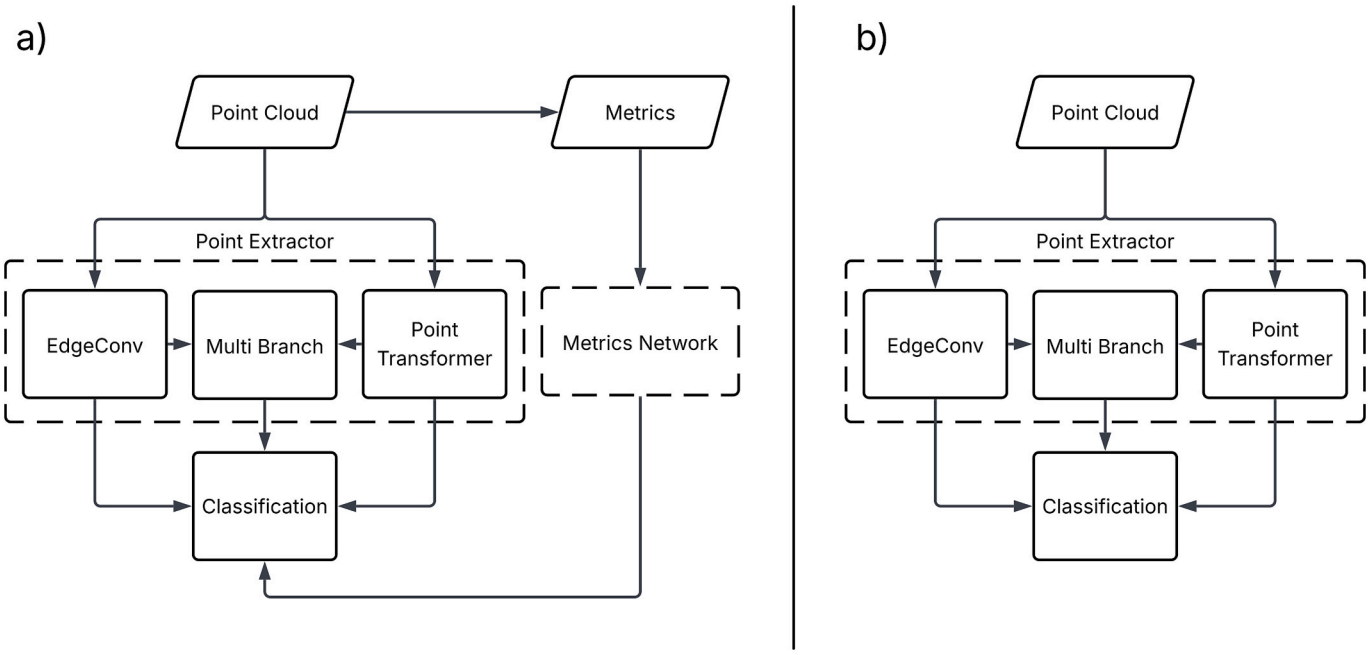


Fig. 4. Model configurations used in the ablation study to examine how different point-based layers impacted species estimation with (a)) and without (b)) the Metrics Network.

Table 3
Number of individual tree samples per species from the stem-mapped and pure stands.

Species	Stem Mapped Samples	Pure Stand Samples	Mean Number of Returns
Black Spruce	308	5870	635
Jack Pine	205	5796	633
Paper Birch	173	646	839
Trembling Aspen	251	3020	922

Table 4
Evaluation scores for the different model configurations.

Model	Overall Accuracy	F1-Score	Precision	Recall	ROC AUC
EdgeConv	0.55	0.54	0.56	0.55	0.74
Point Transformer	0.61	0.62	0.64	0.61	0.80
Multi-Branch	0.58	0.59	0.59	0.59	0.76
EdgeConv + Metrics Network	**0.69**	**0.70**	**0.73**	**0.69**	**0.88**
Point Transformer + Metrics Network	**0.69**	**0.70**	**0.73**	**0.69**	0.87
Multi-Branch + Metrics Network	**0.69**	0.69	0.72	**0.69**	0.86

(paper birch and trembling aspen) had lower accuracy, ranging from 39 % to 57 %.

Fig. 6 shows an example point-cloud for the four tree species and the predictive probabilities produced by each model. The predictive probabilities for all species improved when the Metrics Network stream was incorporated during the training. Overall, the needleleaf species exhibited the highest predictive probabilities, except when the EdgeConv layer was used in the Point Extractor; in that case, the broadleaf species consistently showed higher probabilities. The paper birch example had the lowest predictive probability when the Metrics Network stream was included, whereas trembling aspen example showed the lowest probability when it was not included.

Fig. 7 presents ROC curves, which plot the true positive rate against the false positive rate, showing how effectively each trained model distinguishes a given species from all others. Models incorporating the Metrics Network stream achieved higher AUCs for every species compared to those without it, with black spruce demonstrating excellent discrimination (AUC $\geq$ 0.90). The largest improvement in AUC was observed when EdgeConv was used in the Point Extractor, particularly for black spruce.

The calibration plots in Fig. 8 illustrate the relationship between a models' mean predicted probabilities for each species and the actual observed species labels (fraction of positives). The closer a curve is to the perfect calibration line (grey dashed line), the more accurate the model predicts that species. Curves above or below this line indicate an overprediction or underprediction, respectively. The Multi-Branch Point Extractor layer with the Metrics Network exhibited the most consistent calibration, as it more closely followed the perfect calibration line compared to the other models. Notably, paper birch consistently showed lower predicted probabilities across all models, resulting in a tendency toward underprediction, while the other species typically exhibited overprediction. Fig. 8 indicates that the addition of the Metrics Network enhances model performance without causing overfitting.

3.4. Metrics Network feature attention

Fig. 9 shows a heatmap of the ranked feature attention weights for the metrics used to train the three models which incorporated the Metrics Network. Z_{skew}, $N_{points,}$ and I_{q50} consistently showed greater feature attention across all the models, achieving the highest overall ranks. Volume and $Depth_{q25}$ continuously had low feature attention across all Point Extractor layers, ranking 15th and 12th overall, respectively. Interestingly, $Z_{entropy}$ and $Depth_{q75}$ had lower feature attention ranks when the EdgeConv and Point Transformer layers were used individually but demonstrated high feature attention ranks when combined in the Multi-Branch layer.

4. Discussion

We identified individual tree species from ALS data using an automated, dual-stream DL approach. First, we developed a method to

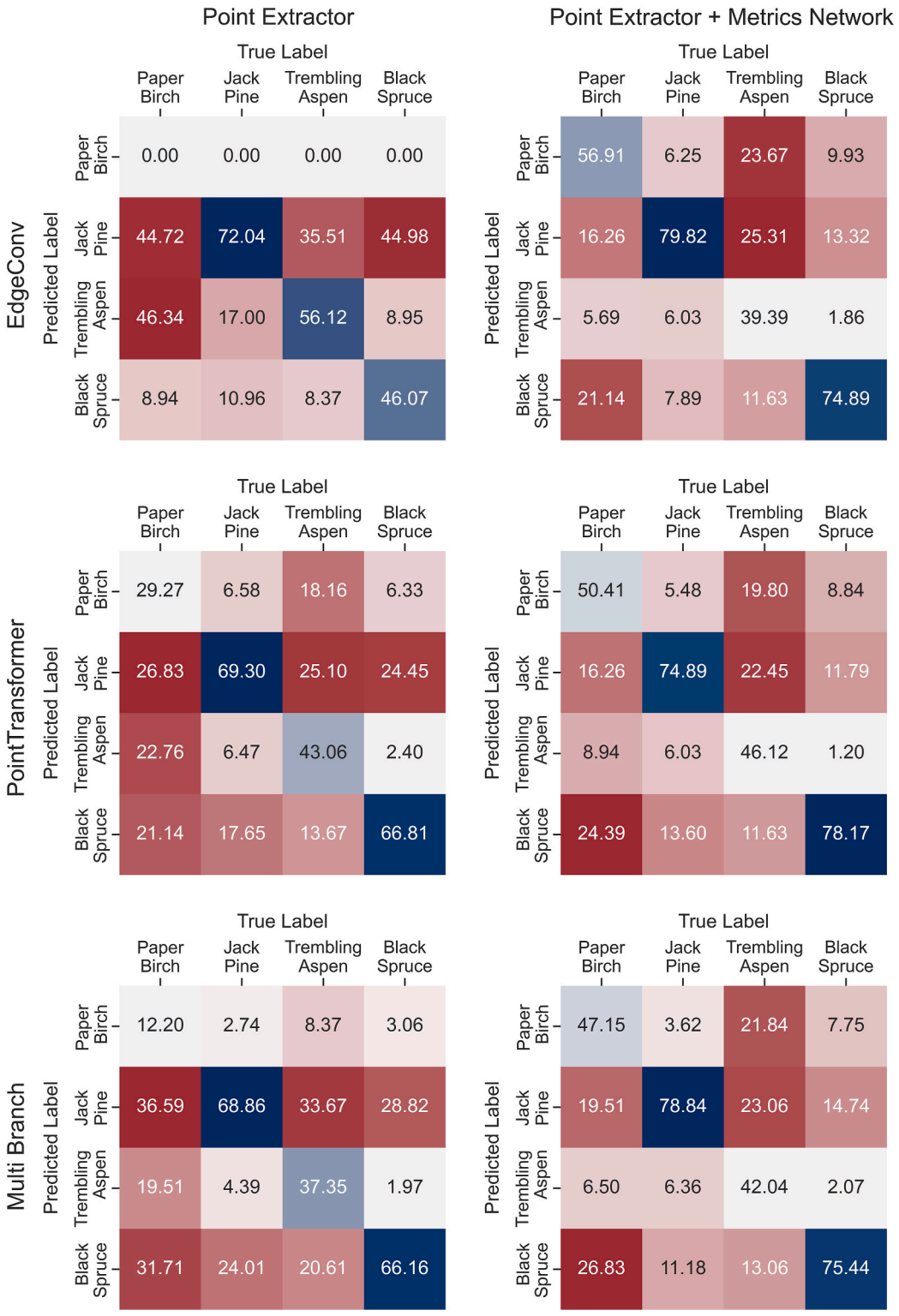


Fig. 5. Confusion matrices showing the classification percentages of species-specific test samples from the trained models of the different Point Extractor layers, with and without the Metrics Network. The darker tones represent higher values of correctly (blue) or incorrectly (red) classified samples; while lighter tones represent lower values.

generate a dataset of 16,269 labelled and segmented individual tree point-clouds using both field and inventory data (Table 3), addressing the challenge of acquiring large training datasets for DL. Second, to determine how various point-based DL layers affect species prediction, we designed the Point Extractor to maintain a consistent architecture while supporting interchangeable point-based DL layers. This enabled fair comparisons between EdgeConv, Point Transformer, and the fused multi-branch layer, identifying layer-specific performance differences. We found that the Point Transformer outperformed other layers in the evaluation scores when used alone, while EdgeConv performed best when combined with the Metrics Network (Table 4). To discern the contribution of point-cloud metrics, we built a dual-stream model

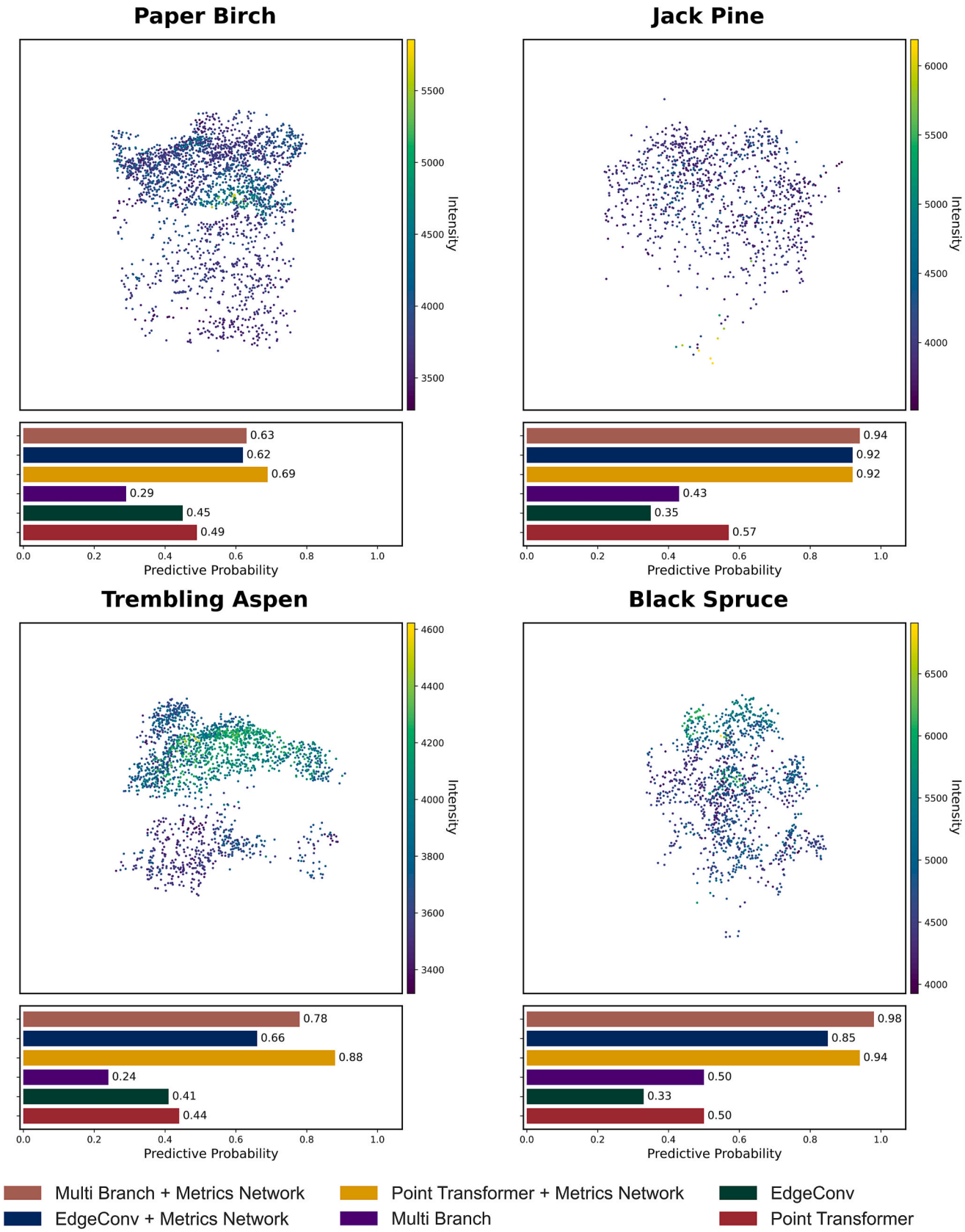


Fig. 6. Sample point-clouds of each species, coloured by intensity, with the samples predicted probability from each model.

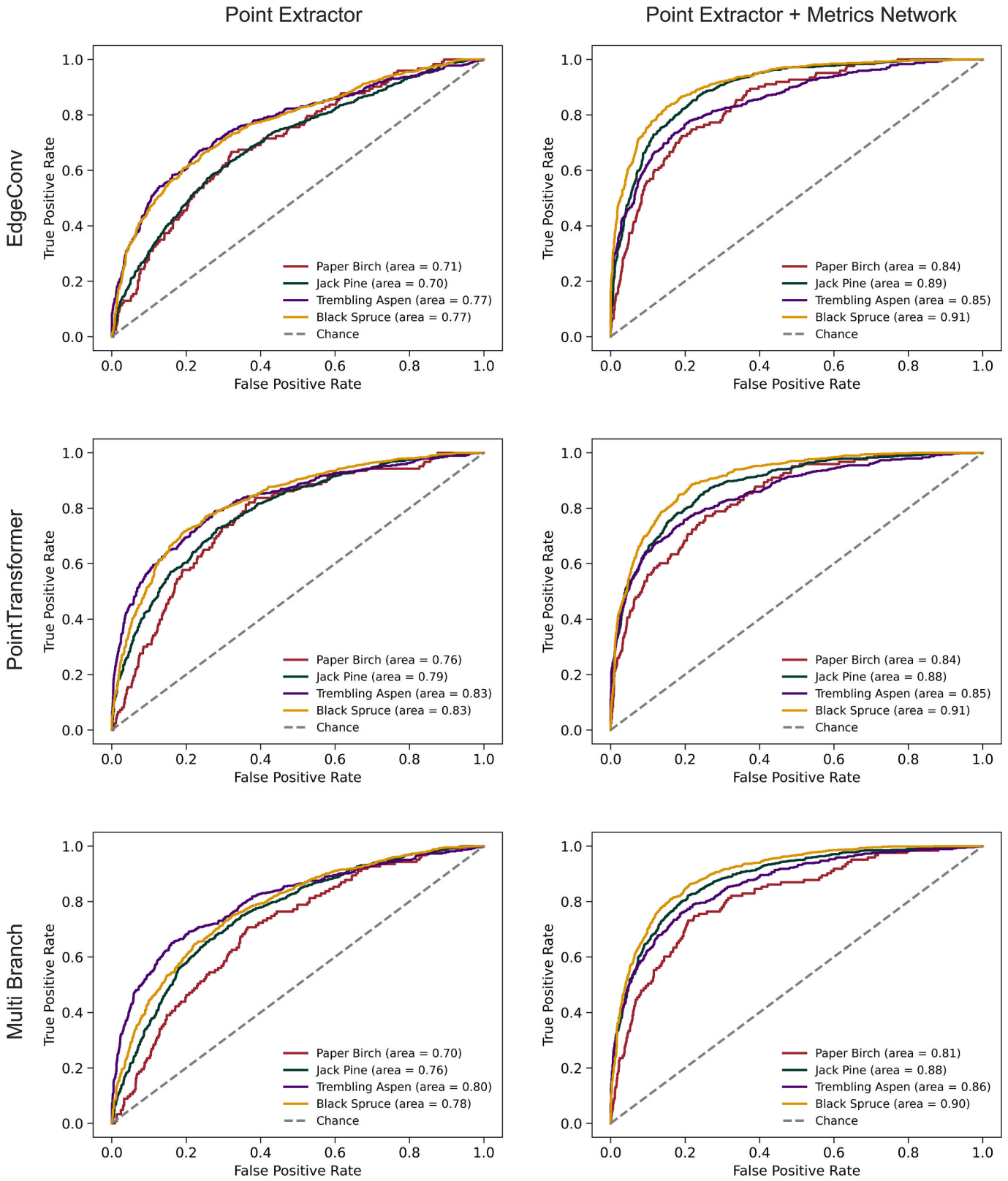


Fig. 7. Receiver operating characteristic curves for each species with associated area under the curve from the trained models using the different Point Extractor layers with and without the Metrics Network.

combining the raw point-cloud and derived metrics, processed through Point Extractor and Metrics Network, respectively. Integrating both streams improved tree species prediction accuracy by an average of 11 % (Table 4) and enhanced predictive probability across all four species (Fig. 6, Fig. 8).

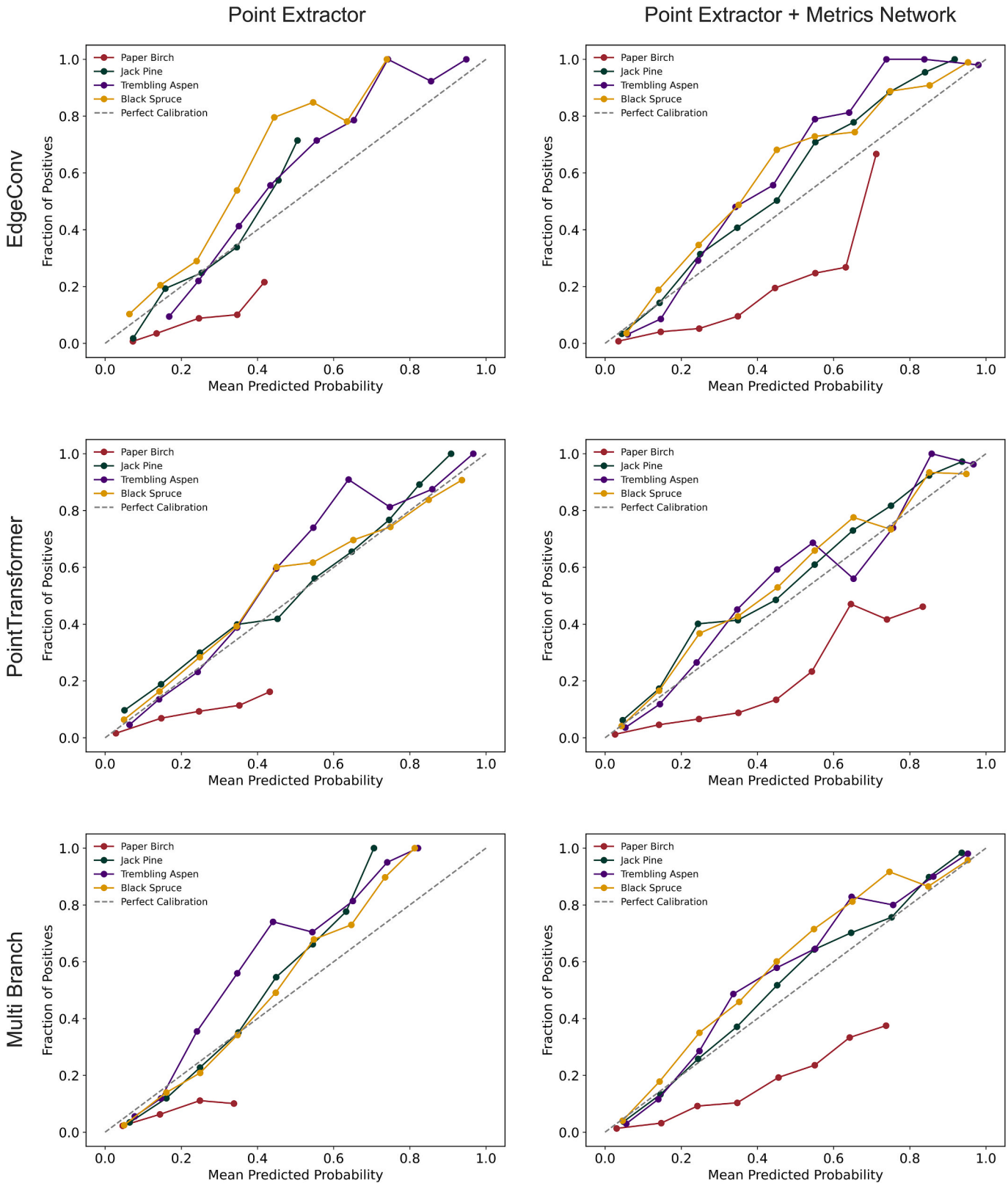


Fig. 8. Calibration plot showing the relationship between the binned mean predicted probabilities and observed fraction of positive predictions for each species from the trained models using the different Point Extractor layers with and without the Metrics Network.

4.1. Comparison and assessment of modeling approach and performance

Many studies have used point-based DL for tree species prediction (J. Chen et al., 2021; Hell et al., 2022; Puliti et al., 2025), often achieving > 90 % accuracy with high-density point-clouds from drone-based laser scanning (DLS) or terrestrial laser scanning (TLS) systems and manual segmentation. While these dense datasets have shown higher accuracy for segmentation tasks (Wielgosz et al., 2024; Xiang et al., 2024), they

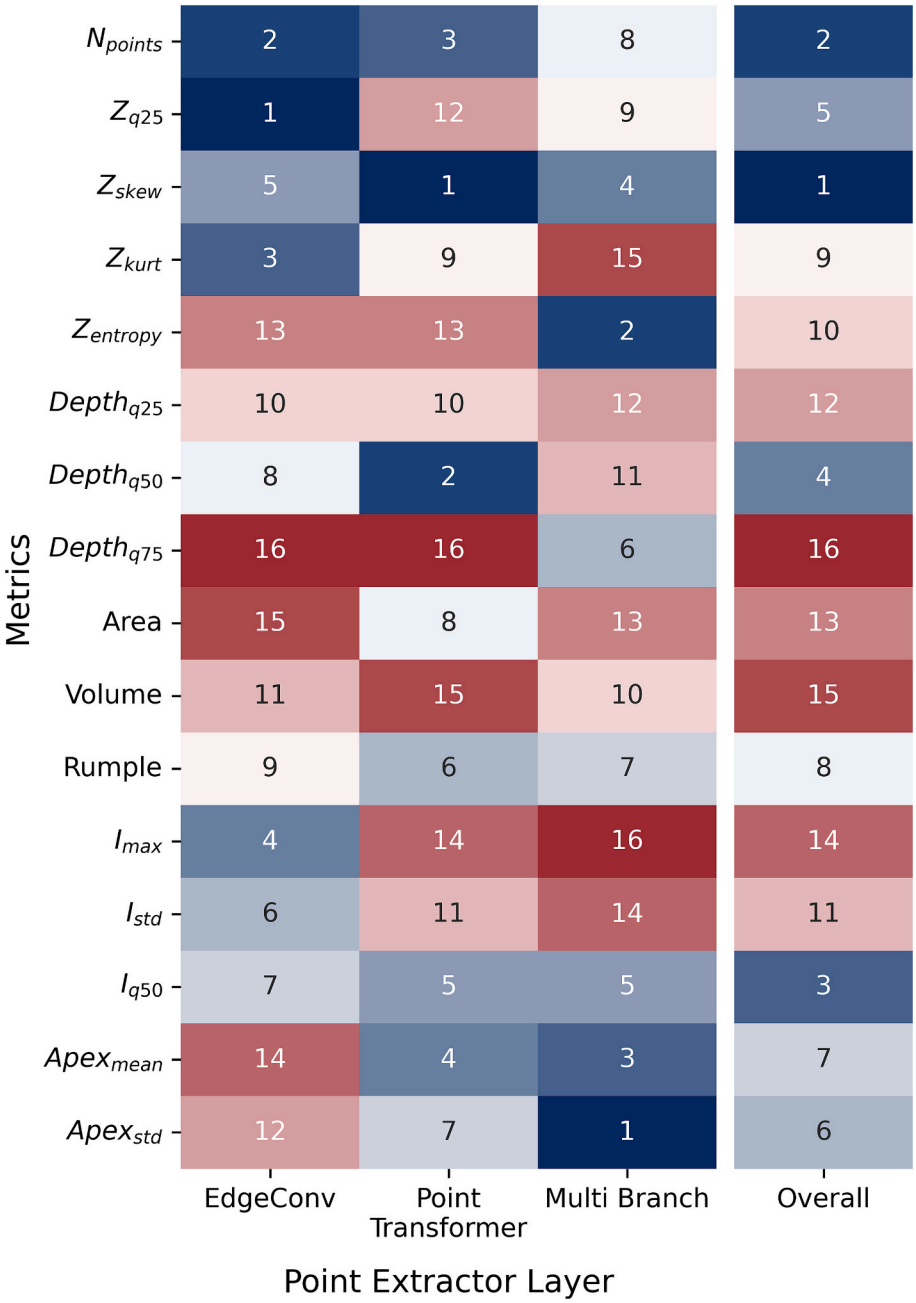


Fig. 9. Heatmap of the feature attention rank given to the point-cloud metrics used in the Metrics Network when trained with the different Point Extractor layers, as well as an overall rank. Red tones represent lower rankings while blue tones represent higher rankings. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

are impractical for large-scale inventories. ALS datasets, on the other hand, are more practical for operational and landscape-level analysis due to their ability to cover extensive areas with relatively lower costs and faster acquisition times compared to DLS/TLS acquisitions. Vahrenhold et al. (2025) used ALS data for individual tree species prediction from DL but used a manual tree segmentation approach. This manual segmentation creates a bottleneck in data processing, requiring trained experts to perform the segmentation and species labelling of the dataset, and is not ideal when considering an approach for predicting species over large landscapes. Other studies have leveraged monospecific stands for species classification (e.g., Gaydon & Roche, 2025; Oreti et al., 2021), but the integration with an automated ITD, segmentation and labelling approach remains underexplored. We wanted to focus on the use of an ALS dataset with an automated data-driven individual tree segmentation approach to demonstrate the approach's feasibility for operational use. This integration of monospecific stands within our automated data generation approach reduced labelling demands while scaling species classification to large forested areas.

4.1.1. *Assessment of point Extractor*

The major benefit of the Point Extractor is its modularity, allowing the integration of various point-based DL layers without altering the overall architecture. Our approach is similar to other modular networks (Andreas et al., 2016; Bi et al., 2025; Kirsch et al., 2018), which allow functional layers to be interchanged within a fixed architecture. This design ensures consistent feature dimensionality, maintaining comparability across experiments regardless of the point-based layer used. While Table 4 shows similar accuracies across dual-stream models, other

results (e.g., Fig. 6, Fig. 8, Fig. 9) reveal differences between Point Extractors, indicating that the point-based stream contributes complementary information alongside the Metrics Network. By interchanging point-based layers, the model adapts the feature extraction process to capture different structural characteristics of the point-cloud, enabling analysis of how each layer influences performance. Future work could expand this approach with additional point-based layers to further investigate the contribution to species prediction.

The Multi-Branch Point Extractor proved beneficial in providing a well-calibrated model, as shown in Fig. 8. Although this model showed lower values for various evaluation scores (Table 4), the calibration plots indicate that the predicted class probabilities are more reliable when combining different Point Extractor layers. This may be due to EdgeConv and Point Transformer identifying different patterns and extracting complementary features from the point-clouds (Paul et al., 2023). Additionally, $Z_{entropy}$ and $Depth_{q75}$ had lower feature attention scores when these layers were used individually, but demonstrated higher attention when combined with the Multi-Branch layer (Fig. 9), reinforcing the complementarity of this approach in identifying unique geometric and structural features of the point-cloud. This allows the model to benefit from diverse representations and may improve generalisation across datasets. A model that provides more reliable predictions also gives forest managers greater confidence in the species predictions. Future studies could explore other hybrid architectures that integrate additional geometric operations to further capture subtle structural patterns within the point-cloud.

4.1.2. Assessment of Metrics Network

The addition of the Metrics Network proved to be beneficial for the tree species prediction, which is consistent with the findings of other studies (Vahrenhold et al., 2025; Z. Zhang et al., 2024). While these metrics may introduce some bias or redundancy, they also can provide meaningful geometric or structural cues that the point-cloud DL features might not convey. The inclusion of an attention-based Metrics Network allowed for a more interpretable model, as the feature attention of the various metrics can be visualised. Since the Metrics Network and the Point Extractor were trained in parallel, the attention weights assigned to the various metrics differed based on the Point Extractor layer used. This can be observed through our experiments with different Point Extractor layers, as feature importance rank for some metrics changed more based on the layer used (Fig. 9). While this study used simple concatenation for feature fusion, future work could explore attention-based or gating mechanisms to enhance feature interaction, model performance and interpretability (Liu et al., 2022b; Pan et al., 2025).

Z_{skew}, N_{points}, and I_{q50} exhibited the highest overall feature attention ranks (Fig. 9), likely reflecting species-specific differences in vertical structure, canopy density, and reflectance. I_{q50} and other intensity metrics may rank differently with other lidar sensors (linear-mode lidar) due to inconsistent intensity values across sensors (Brown et al., 2020). Identifying and interpreting key metrics would be essential when applying this model to data from other sensors with different return densities and intensities. Volume and Area showed lower attention ranks (Fig. 9), possibly due to varying stand densities across the RMF. Retraining or fine-tuning may be necessary when applying this model to other sensors or regions to ensure accuracy.

4.2. Modeling approach and data limitations

One of the limitations of this approach is the potential propagation of errors from the watershed segmentation into the feature extraction and classification stages. Although the watershed algorithm has been widely applied in forestry studies (Edson & Wing, 2011; Kanda et al., 2004; Queinnec et al., 2022), segmentation inaccuracies, such as crown over or under segmentation, may introduce noise into the extracted features, impacting model predictions. Since the watershed segmentation uses the CHM, it will only segment trees whose crowns are found within the upper canopy, ultimately leading to additional returns being introduced to the point-cloud belonging to other understory vegetation. With new segmentation approaches using advanced DL models (Wielgosz et al., 2024) there is potential for these errors in tree segmentation to be diminished and could be incorporated into future studies using our modeling approach. While our study focused on developing and testing the automated dual-stream DL approach for predicting tree species, future research could aim to quantify segmentation errors and their influence on species prediction by using manual delineations or alternative segmentation methods.

The models trained within this study were generally more accurate in predicting needleleaf (black spruce and jack pine) compared to broadleaf (paper birch and trembling aspen) samples. Similar findings were also observed by Krzystek et al. (2020) where they classified needleleaf, broadleaf, and dead trees using full waveform ALS data and multispectral imagery. Since paper birch and trembling aspen had fewer samples in the dataset (Table 3), we incorporated a weighted cross entropy loss to penalize these classes, encouraging the model to focus more on correctly predicting them and improving overall training performance. Despite this, the models predicted these species less accurately (Fig. 5), and the mean predictive probability for paper birch remained low (Fig. 6; Fig. 8), potentially due to too few samples to capture their diverse point patterns. Additionally, the pseudo-plot sample generation strategy assumes species purity within stands, which may not always be valid in heterogeneous forests, potentially introducing systematic bias; however, this assumption was necessary to ensure sufficient training samples for effective model development.

With many segmentation tasks observing lower accuracies for broadleaf species compared to needleleaf species (Y. Li et al., 2023; Yu et al., 2024), the compounding errors in the watershed segmentation approach could have negatively impacted the predictions for paper birch and trembling aspen. Furthermore, there is often confusion between paper birch and trembling aspen samples (Fig. 5), which could be a result of similar learnt feature representations due to their structural and radiometric similarities (Dalponte et al., 2012; Kuzmin et al., 2021). This confusion may also explain the drop in trembling aspen accuracy when the Metrics Network is used with the EdgeConv layer, as paper birch was not predicted without it, eliminating the possibility of misclassifying trembling aspen as paper birch. The lower prediction accuracies observed in this study for the broadleaf species may also suggest that the geometric or structural features of the point cloud alone may not adequately capture broadleaf variability. Future studies could incorporate spectral information into the modelling approach which has shown to be beneficial especially for broadleaf species (Murray et al., 2025; Queinnec et al., 2022).

5. Conclusion

With the greater use of ALS data for forest inventories, there is a growing need to also derive tree species information from the ALS data. Species information at the individual tree level is important for deriving further information about the tree (e.g., merchantable timber, biomass) or the stand (e.g., species composition, ecosystem preferences). Through this study, we examined how an automated, data-driven dual-stream DL approach could be used to predict tree species at the individual tree level using a combination of the raw ALS point-cloud and derived metrics. Our results indicated, using this dual-stream model, which included an adaptable Point Extractor and Metrics Network, outperformed the models using only raw point-cloud data. This dual-stream model also improved the predictive probability of the four species of interest to this study (black spruce, jack pine, paper birch, and trembling aspen), providing further confidence for forest managers to make informed decisions from the predictions. These findings show that this automated, data-driven dual-stream DL approach can produce a large quantity of labeled point-clouds for use in species classification, highlighting the value of combining summary metrics with raw ALS point-clouds for

improving prediction in structurally and species-diverse forest environments.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used OpenAI's ChatGPT v 4o to improve the readability of this manuscript. After using this tool the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

CRediT authorship contribution statement

Brent A. Murray: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Nicholas C. Coops:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Joanne C. White:** Writing – review & editing. **Adam Dick:** Writing – review & editing. **Ignacio Barbeito:** Writing – review & editing. **Ahmed Ragab:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was undertaking in collaboration with Natural Resources Canada. This work was supported by the National Research Council of Canada (NRC; under project number DHGA-119-1); and partly by the Natural Sciences and Engineering Research Council of Canada (NSERC; under grant RGPIN-2018-03851). The forest inventory data of the Romeo Malette Forest used in this study was provided by Chris MacDonnell (GreenFirst Forest Products) and Grant McCartney. The single-photon lidar data was generously provided by the Ontario Ministry of Natural Resources and Forestry.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jag.2025.104877.

Data availability

The authors do not have permission to share data.

References

Andreas, J., Rohrbach, M., Darrell, T., Klein, D., 2016. Neural Module Networks. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 39–48. https://doi.org/10.1109/CVPR.2016.12.

Arik, S. O., & Pfister, T. (2020). TabNet: Attentive Interpretable Tabular Learning (No. arXiv:1908.07442). arXiv. https://doi.org/10.48550/arXiv.1908.07442.

Barbedo, J.G.A., 2021. Deep learning applied to plant pathology: the problem of data representativeness. Trop. Plant Pathol. 47 (1), 85–94. https://doi.org/10.1007/s40858-021-00459-9.

Benner, J., Nielsen, J., Lertzman, K., 2021. Using traditional ecological knowledge to understand the diversity and abundance of culturally important trees. J. Ethnobiol. 41 (2), 209–228. https://doi.org/10.2993/0278-0771-41.2.209.

Bi, S., Miao, Z., Min, Q., 2025. A modular dual learning for improving question answering and generation over knowledge graphs. IEEE Trans. Audio Speech Lang. Process. 33, 401–417. https://doi.org/10.1109/TASLPRO.2025.3527218.

Blomley, R., Hovi, A., Weinmann, M., Hinz, S., Korpela, I., Jutzi, B., 2017. Tree species classification using within crown localization of waveform LiDAR attributes. ISPRS J. Photogramm. Remote Sens. 133, 142–156. https://doi.org/10.1016/j.isprsjprs.2017.08.013.

Briechle, S., Krzystek, P., Vosselman, G., 2021. Silvi-net – a dual-CNN approach for combined classification of tree species and standing dead trees from remote sensing data. Int. J. Appl. Earth Obs. Geoinf. 98, 102292. https://doi.org/10.1016/j.jag.2020.102292.

Brown, R., Hartzell, P., Glennie, C., 2020. Evaluation of SPL100 single photon lidar data. Remote Sensing 12 (4), 722. https://doi.org/10.3390/rs12040722.

Chen, C., Wu, Y., Dai, Q., Zhou, H.-Y., Xu, M., Yang, S., Han, X., Yu, Y., 2024. A survey on graph neural networks and graph transformers in computer vision: a task-oriented perspective. IEEE Trans. Pattern Anal. Mach. Intell. 46 (12), 10297–10318. https://doi.org/10.1109/TPAMI.2024.3445463.

Chen, J., Chen, Y., Liu, Z., 2021. Classification of typical tree species in laser point cloud based on deep learning. Remote Sensing 13 (23), 4750. https://doi.org/10.3390/rs13234750.

Dalponte, M., Bruzzone, L., Gianelle, D., 2012. Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. Remote Sens. Environ. 123, 258–270. https://doi.org/10.1016/j.rse.2012.03.013.

Dalponte, M., Frizzera, L., Gianelle, D., 2019. Individual tree crown delineation and tree species classification with hyperspectral and LiDAR data. PeerJ 6, e6227. https://doi.org/10.7717/peerj.6227.

Edson, C., Wing, M.G., 2011. Airborne light detection and ranging (LiDAR) for individual tree stem location, height, and biomass measurements. Remote Sens. (Basel) 3 (11), 2494–2528. https://doi.org/10.3390/rs3112494.

Gaydon, C., Roche, F., 2025. PureForest: a large-scale aerial lidar and aerial imagery dataset for tree species classification in monospecific forests. In: 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), pp. 5895–5904. https://doi.org/10.1109/WACV61041.2025.00575.

Gluckman, J., 2016. Design of the processing chain for a high-altitude, airborne, single-photon lidar mapping instrument. Laser Radar Technology and Applications XXI 9832, 20–28. https://doi.org/10.1117/12.2219760.

Goodfellow, I., Bengio, Y., Courville, A., 2017. Optimization for training deep models. In: Deep Learning. MIT Press, pp. 274–329.

Grubinger, S., Coops, N.C., O'Neill, G.A., 2023. Picturing local adaptation: spectral and structural traits from drone remote sensing reveal clinal responses to climate transfer in common-garden trials of interior spruce (*Picea engelmannii* × *glauca*). Glob. Chang. Biol. 29 (17), 4842–4860. https://doi.org/10.1111/gcb.16855.

Hell, M., Brandmeier, M., Briechle, S., Krzystek, P., 2022. Classification of tree species and standing dead trees with lidar point clouds using two deep neural networks: PointCNN and 3DmFV-net. PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science 90 (2), 103–121. https://doi.org/10.1007/s41064-022-00200-4.

Hermosilla, T., Bastyr, A., Coops, N.C., White, J.C., Wulder, M.A., 2022. Mapping the presence and distribution of tree species in Canada's forested ecosystems. Remote Sens. Environ. 282, 113276. https://doi.org/10.1016/j.rse.2022.113276.

Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., Campbell, L.B., 2016. Mass data processing of time series Landsat imagery: pixels to data products for forest monitoring. Int. J. Digital Earth 9 (11), 1035–1054. https://doi.org/10.1080/17538947.2016.1187673.

Hong, D., Gao, L., Yokoya, N., Yao, J., Chanussot, J., Du, Q., Zhang, B., 2021. More Diverse Means Better: Multimodal Deep Learning Meets Remote-Sensing Imagery Classification. IEEE Transactions on Geoscience and Remote Sensing 59 (5), 4340–4354. https://doi.org/10.1109/TGRS.2020.3016820.

Irwin, L., Coops, N.C., Queinnec, M., McCartney, G., White, J.C., 2021. Single photon lidar signal attenuation under boreal forest conditions. Remote Sens. Lett. 12 (10), 1049–1060. https://doi.org/10.1080/2150704X.2021.1962575.

Jing, L., Hu, B., Li, J., Noland, T., 2012. Automated delineation of individual tree crowns from lidar data by multi-scale analysis and segmentation. Photogramm. Eng. Remote Sens. 78 (12), 1275–1284. https://doi.org/10.14358/PERS.78.11.1275.

Kada, M., & Kuramin, D. (2021). ALS Point Cloud Classification using PointNet++ and KPConv with Prior Knowledge. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVI-4/W4-2021, 91–96. https://doi.org/10.5194/isprs-archives-XLVI-4-W4-2021-91-2021.

Kanda, F., Kubo, M. and Muramoto, K.I., 2004, September. Watershed segmentation and classification of tree species using high resolution forest imagery. In IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium (Vol. 6, pp. 3822-3825). IEEE. https://doi.org/10.1109/IGARSS.2004.1369956.

Kirsch, L., Kunze, J., & Barber, D. (2018). Modular Networks: Learning to Decompose Neural Computation. 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, Canada.

Krzystek, P., Serebryanyk, A., Schnörr, C., Červenka, J., Heurich, M., 2020. Large-scale mapping of tree species and dead trees in šumava national park and bavarian forest national park using lidar and multispectral imagery. Remote Sens. (Basel) 12 (4), 661. https://doi.org/10.3390/rs12040661.

Kuzmin, A., Korhonen, L., Kivinen, S., Hurskainen, P., Korpelainen, P., Tanhuanpää, T., Maltamo, M., Vihervaara, P., Kumpula, T., 2021. Detection of european aspen (Populus tremula L.) based on an unmanned aerial vehicle approach in boreal forests. Remote Sens. (Basel) 13 (9), 1723. https://doi.org/10.3390/rs13091723.

Lahat, D., Adali, T., Jutten, C., 2015. Multimodal data fusion: an overview of methods, challenges, and prospects. Proc. IEEE 103 (9), 1449–1477. https://doi.org/10.1109/JPROC.2015.2460697.

Leckie, D.G., Gillis, M.D., 1995. Forest inventory in Canada with emphasis on map production. For. Chron. 71 (1), 74–88. https://doi.org/10.5558/tfc71074-1.

Lei, X., Wang, H., Wang, C., Zhao, Z., Miao, J., Tian, P., 2020. ALS point cloud classification by integrating an improved fully convolutional network into transfer learning with multi-scale and multi-view deep features. Sensors 20 (23), 6969. https://doi.org/10.3390/s20236969.

Leica. (2024). Leica SPL100 Single Photon LiDAR Sensor. https://leica-geosystems.com/products/airborne-systems/topographic-lidar-sensors/leica-spl100.

Li, D., Lu, C., Chen, Z., Guan, J., Zhao, J., Du, J., 2024. Graph neural networks in point clouds: a survey. Remote Sens. (Basel) 16 (14), 2518. https://doi.org/10.3390/rs16142518.

Li, Y., Xie, D., Wang, Y., Jin, S., Zhou, K., Zhang, Z., Li, W., Zhang, W., Mu, X., Yan, G., 2023. Individual tree segmentation of airborne and UAV LiDAR point clouds based on the watershed and optimized connection center evolution clustering. Ecol. Evol. 13 (7), e10297. https://doi.org/10.1002/ece3.10297.

Lin, Y., Hyyppä, J., 2016. A comprehensive but efficient framework of proposing and validating feature parameters from airborne LiDAR data for tree species classification. Int. J. Appl. Earth Obs. Geoinf. 46, 45–55. https://doi.org/10.1016/j.jag.2015.11.010.

Liu, B., Hao, Y., Huang, H., Chen, S., Li, Z., Chen, E., Tian, X., Ren, M., 2023. TSCMDL: multimodal deep learning framework for classifying tree species using fusion of 2-D and 3-D features. IEEE Trans. Geosci. Remote Sens. 61, 1–11. https://doi.org/10.1109/TGRS.2023.3266057.

Liu, B., Huang, H., Su, Y., Chen, S., Li, Z., Chen, E., Tian, X., 2022a. Tree species classification using ground-based LiDAR data by various point cloud deep learning methods. Remote Sens. (Basel) 14 (22), 5733. https://doi.org/10.3390/rs14225733.

Liu, Q., Kampffmeyer, M., Jenssen, R., Salberg, A.-B., 2022b. Multi-modal land cover mapping of remote sensing images using pyramid attention and gated fusion networks. Int. J. Remote Sens. 43 (9), 3509–3535. https://doi.org/10.1080/01431161.2022.2098078.

Loshchilov, I., & Hutter, F. (2017). SGDR: Stochastic Gradient Descent with Warm Restarts (No. arXiv:1608.03983). arXiv. https://doi.org/10.48550/arXiv.1608.03983.

Magnussen, S., Russo, G., 2012. Uncertainty in photo-interpreted forest inventory variables and effects on estimates of error in canada's national forest inventory. For. Chron. 88 (04), 439–447. https://doi.org/10.5558/tfc2012-080.

Mandlburger, G., Lehner, H., Pfeifer, N., 2019. A Comparison of single photon and full waveform lidar. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. IV-2/W5, 397–404. https://doi.org/10.5194/isprs-annals-IV-2-W5-397-2019.

Mohren, G., Hasenauer, H., Köhl, M., Nabuurs, G.-J., 2012. Forest inventories for carbon change assessments. Curr. Opin. Environ. Sustain. 4 (6), 686–695. https://doi.org/10.1016/j.cosust.2012.10.002.

Murray, B.A., Coops, N.C., White, J.C., Dick, A., Ragab, A., 2025. Tree species proportion prediction using airborne laser scanning and Sentinel-2 data within a deep learning based dual-stream data fusion approach. Int. J. Remote Sens. 46 (14), 5436–5464. https://doi.org/10.1080/01431161.2025.2521072.

Murray, B.A., Coops, N.C., Winiwarter, L., White, J.C., Dick, A., Barbeito, I., Ragab, A., 2024. Estimating tree species composition from airborne laser scanning data using point-based deep learning models. ISPRS J. Photogramm. Remote Sens. 207, 282–297. https://doi.org/10.1016/j.isprsjprs.2023.12.008.

Næsset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sens. Environ. 80 (1), 88–99. https://doi.org/10.1016/S0034-4257(01)00290-5.

Ngiam, J., Khosla, A., Kim, M., Nam, J., Lee, H., Ng, A.Y., 2011. Multimodal Deep Learning. In: ICML, Vol. 11, pp. 689–696.

OMNRF. (1996). Specifications for Forest Resources Inventory Photo Interpretation Standards. Ontario Ministry of Natural Resources and Forestry.

OMNRF. (2001). Forest Information Manual. Ontario Ministry of Natural Resources and Forestry.

OMNRF. (2019). Forest Resources Inventory leaf-on LiDAR [Dataset]. Ontario GeoHub. https://geohub.lio.gov.on.ca/maps/lio::forest-resources-inventory-leaf-on-lidar/about.

Oreti, L., Giuliarelli, D., Tomao, A., Barbati, A., 2021. Object oriented classification for mapping mixed and pure forest stands using very-high resolution imagery. Remote Sens. (Basel) 13 (13), 2508. https://doi.org/10.3390/rs13132508.

Ørka, H.O., Gobakken, T., Næsset, E., Ene, L., Lien, V., 2012. Simultaneously acquired airborne laser scanning and multispectral imagery for individual tree species identification. Can. J. Remote. Sens. 38 (2), 125–138. https://doi.org/10.5589/m12-021.

Pan, J., Cao, J., Xing, S., Dai, M., Liu, J., Wang, X., Zhang, Y., Huang, G., 2025. An aerial point cloud classification using point transformer via multi-feature fusion. Sci. Rep. 15 (1), 22924. https://doi.org/10.1038/s41598-025-02719-z.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., … Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. 33rd Conference on Neural Information Processing Systems (NeurIP, 2019), Vancouver, Canada.

Paul, S., Patterson, Z., Bouguila, N., 2023. DualMLP: a two-stream fusion model for 3D point cloud classification. Vis. Comput. 40 (8), 5435–5449. https://doi.org/10.1007/s00371-023-03114-3.

Perktold, J., Seabold, S., Sheppard, K., ChadFulton, Shedden, K., jbrockmendel, j-grana6, Quackenbush, P., Arel-Bundock, V., McKinney, W., Langmore, I., Baker, B., Gommers, R., yogabonito, s-scherrer, Zhurko, Y., Brett, M., Giampieri, E., yl1565, … Halchenko, Y. (2024). statsmodels/statsmodels: Release 0.14.2 (Version v0.14.2) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.10984387.

Polikar, R., 2006. Ensemble based systems in decision making. IEEE Circuits Syst. Mag. 6 (3), 21–45. https://doi.org/10.1109/MCAS.2006.1688199.

Popescu, S.C., Wynne, R.H., 2004. Seeing the trees in the forest: using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height. Photogramm. Eng. Remote Sens. 70, 589–604.

Prasad, A., Pedlar, J., Peters, M., McKenney, D., Iverson, L., Matthews, S., Adams, B., 2020. Combining US and Canadian forest inventories to assess habitat suitability and migration potential of 25 tree species under climate change. Divers. Distrib. 26 (9), 1142–1159. https://doi.org/10.1111/ddi.13078.

Pretzsch, H., Poschenrieder, W., Uhl, E., Brazaitis, G., Makrickiene, E., Calama, R., 2021. Silvicultural prescriptions for mixed-species forest stands. a european review and perspective. Eur. J. For. Res. 140 (5), 1267–1294. https://doi.org/10.1007/s10342-021-01388-7.

Puliti, S., Lines, E.R., Müllerová, J., Frey, J., Schindler, Z., Straker, A., Allen, M.J., Winiwarter, L., Rehush, N., Hristova, H., Murray, B., Calders, K., Coops, N., Höfle, B., Irwin, L., Junttila, S., Krůček, M., Krok, G., Král, K., Astrup, R., 2025. Benchmarking tree species classification from proximally sensed laser scanning data: introducing the FOR - species20K dataset. Methods Ecol. Evol. 16 (4), 801–818. https://doi.org/10.1111/2041-210X.14503.

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 (2), 406. https://doi.org/10.3390/rs15020406.

Queinnec, M., Coops, N.C., White, J.C., Griess, V.C., Schwartz, N.B., McCartney, G., 2022. Mapping dominant boreal tree species groups by combining area-based and individual tree crown LiDAR metrics with sentinel-2 Data. Can. J. Remote. Sens. 1–19. https://doi.org/10.1080/07038992.2022.2130742.

Rahlf, J., Hauglin, M., Astrup, R., Breidenbach, J., 2021. Timber volume estimation based on airborne laser scanning—comparing the use of national forest inventory and forest management inventory data. Ann. For. Sci. 78 (2), 49. https://doi.org/10.1007/s13595-021-01061-4.

Räty, J., Varvia, P., Korhonen, L., Savolainen, P., Maltamo, M., Packalen, P., 2022. A comparison of linear-mode and single-photon airborne lidar in species-specific forest inventories. IEEE Trans. Geosci. Remote Sens. 60, 1–14. https://doi.org/10.1109/TGRS.2021.3060670.

Thompson, I.D., Maher, S.C., Rouillard, D.P., Fryxell, J.M., Baker, J.A., 2007. Accuracy of forest inventory mapping: some implications for boreal forest management. For. Ecol. Manage. 252 (1), 208–221. https://doi.org/10.1016/j.foreco.2007.06.033.

Tompalski, P., White, J.C., Coops, N.C., Wulder, M.A., Leboeuf, A., Sinclair, I., Butson, C. R., Lemonde, M.-O., 2021. Quantifying the precision of forest stand height and canopy cover estimates derived from air photo interpretation. Forestry: an International Journal of Forest Research 94 (5), 611–629. https://doi.org/10.1093/forestry/cpab022.

Vahrenhold, J.R., Brandmeier, M., Müller, M.S., 2025. MMTSCNet: multimodal tree species classification network for classification of multi-source, single-tree LiDAR point clouds. Remote Sensing 17 (7), 1304. https://doi.org/10.3390/rs17071304.

Van Der Walt, S., Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., Gouillart, E., Yu, T., 2014. scikit-image: image processing in Python. PeerJ 2, e453. https://doi.org/10.7717/peerj.453.

Wang, Y., Sun, Y., Liu, Z., Sarma, S.E., Bronstein, M.M., Solomon, J.M., 2019. Dynamic graph CNN for learning on point clouds. ACM Trans. Graph. 38 (5), 1–12. https://doi.org/10.1145/3326362.

Wen, Y., Zhang, K., Li, Z., Qiao, Y., 2016. A Discriminative Feature Learning approach for deep face recognition. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (Eds.), Computer Vision – ECCV 2016, Vol. 9911. Springer International Publishing, pp. 499–515. https://doi.org/10.1007/978-3-319-46478-7_31.

White, J.C., Tompalski, P., Bater, C.W., Wulder, M.A., Fortin, M., Hennigar, C., Robere-McGugan, G., Sinclair, I., White, R., 2025. Enhanced forest inventories in Canada: implementation, status, and research needs. Can. J. For. Res. 55, 1–37. https://doi.org/10.1139/cjfr-2024-0255.

Wielgosz, M., Puliti, S., Xiang, B., Schindler, K., Astrup, R., 2024. SegmentAnyTree: a sensor and platform agnostic deep learning model for tree segmentation using laser scanning data. Remote Sens. Environ. 313, 114367. https://doi.org/10.1016/j.rse.2024.114367.

Xiang, B., Wielgosz, M., Kontogianni, T., Peters, T., Puliti, S., Astrup, R., Schindler, K., 2024. Automated forest inventory: analysis of high-density airborne LiDAR point clouds with 3D deep learning. Remote Sens. Environ. 305, 114078. https://doi.org/10.1016/j.rse.2024.114078.

Yu, J., Lei, L., Li, Z., 2024. Individual tree segmentation based on seed points detected by an adaptive crown shaped algorithm using UAV-LiDAR data. Remote Sens. (Basel) 16 (5), 825. https://doi.org/10.3390/rs16050825.

Zhang, H., Wang, C., Tian, S., Lu, B., Zhang, L., Ning, X., Bai, X., 2023. Deep learning-based 3D point cloud classification: a systematic survey and outlook. Displays 79, 102456. https://doi.org/10.1016/j.displa.2023.102456.

Zhang, Z., Wang, J., Wu, Y., Zhao, Y., Wu, B., 2024. Deeply supervised network for airborne LiDAR tree classification incorporating dual attention mechanisms. Giscience & Remote Sensing 61 (1), 2303866. https://doi.org/10.1080/15481603.2024.2303866.

Zhao, H., Jiang, L., Jia, J., Torr, P., & Koltun, V. (2021). Point Transformer (No. arXiv:2012.09164). arXiv. https://doi.org/10.48550/arXiv.2012.09164.