Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 4, **Number** 1, May 2024
https://doi.org/10.47709/brilliance.v4i1.3673

Comparative Analysis of Machine Learning Algorithms for Multi-Class Tree Species Classification Using Airborne LiDAR Data

Gregorius Airlangga[1*]
[1]Atma Jaya Catholic University of Indonesia, Indonesia
[1]gregorius.airlangga@atmajaya.ac.id

open access

***Corresponding Author**

Article History:
Submitted: 29-02-2024
Accepted: 01-03-2024
Published: 08-03-2024

Keywords:
LiDAR Remote Sensing;
Tree Species Classification;
Machine Learning Algorithms;
Forest Ecology;
Airborne LiDAR Data Analysis

ABSTRACT

Forests hold vital ecological significance, and the ability to accurately classify tree species is integral to conservation and management practices. This research investigates the application of machine learning techniques to airborne Light Detection and Ranging (LiDAR) data for the multi-class classification of tree species, specifically Alder, Aspen, Birch, Fir, Pine, Spruce, and Tilia. High-density LiDAR data from varied forest landscapes were subjected to a rigorous preprocessing and noise reduction protocol, followed by feature extraction to discern structural characteristics indicative of species identity. We assessed the performance of six machine learning models: Logistic Regression, Decision Tree, Random Forest, Support Vector Classifier (SVC), k-Nearest Neighbors (KNN), and Gradient Boosting. The analysis was based on metrics of accuracy, precision, recall, and F1 score. Logistic Regression and Random Forest models outperformed others, achieving accuracies of 0.81, precision of 0.80, recall of 0.81, and an F1 score of 0.80. In contrast, the KNN algorithm had the lowest accuracy of 0.60, precision and recall of 0.60, and an F1 score of 0.59. These results demonstrate the robustness of Logistic Regression and Random Forest for classifying complex LiDAR datasets. The study underscores the potential of these models to support ecological monitoring, enhance forest management, and aid in biodiversity conservation. Future research directions include the fusion of LiDAR data with other environmental variables, application of deep learning for improved feature extraction, and validation of the models across broader species and geographical ranges. This research marks a significant step towards leveraging advanced machine learning to interpret and utilize LiDAR data for environmental and ecological applications.

INTRODUCTION

Forests play a crucial role in sustaining life on Earth, covering approximately 31% of the global land area (Mackey et al., 2020). They are central to biodiversity conservation, climate change mitigation, and the provision of livelihoods for millions of people worldwide (Muluneh, 2021). The accurate identification and classification of tree species within these ecosystems are fundamental for effective forest management, biodiversity conservation, and climate modeling efforts (Blanco, Ameztegui, & Rodr\'\iguez, 2020). Traditional approaches to species classification have predominantly relied on field-based observations, which, despite their accuracy, are labor-intensive, time-consuming, and impractical over extensive areas (Yip et al., 2024). The advent of Light Detection and Ranging (LiDAR) technology has significantly transformed ecological monitoring practices by providing high-resolution three-dimensional data on forest structures, enabling remote sensing applications such as tree species classification from both airborne and terrestrial LiDAR scans (Camarretta et al., 2020). The capability of LiDAR to capture the vertical complexity of forests offers distinct advantages over traditional remote sensing methods, facilitating the distinction between species based on morphological characteristics (Xu, Wang, Xu, Luan, & Xu, 2021). The integration of machine learning with LiDAR data analysis has emerged as a particularly promising avenue for enhancing the accuracy and efficiency of tree species classification (Gharineiat, Tarsha Kurdi, & Campbell, 2022). Various machine learning algorithms, from logistic regression to more sophisticated ensemble methods like gradient boosting and random forests, have been applied to this challenge, demonstrating the potential to significantly advance our understanding of forest biodiversity from a remote sensing perspective (Blanco et al., 2020). However, these advancements have also highlighted critical gaps in the field, particularly regarding the generalizability of models across different ecosystems, the handling of the high-dimensional nature of LiDAR data, and the specific challenge of multi-class classification among diverse tree species (Chehreh, Moutinho, & Viegas, 2023).

This research aims to address these gaps by focusing on the multi-class classification problem, aiming to distinguish between several key tree species, including Alder, Aspen, Birch, Fir, Pine, Spruce, and Tilia. Such classification is inherently complex due to the subtle differences in the structural features that LiDAR data must capture

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 4, **Number** 1, May 2024
https://doi.org/10.47709/brilliance.v4i1.3673

to accurately distinguish between species (Hastings et al., 2020). Therefore, the study explores the application and comparison of various machine learning models to identify the most effective approaches for classifying these tree species based on LiDAR data. By implementing and rigorously evaluating a range of algorithms, including logistic regression, k-nearest neighbors, support vector machines, random forests, decision trees, and gradient boosting, this study seeks to make a significant contribution to the field of remote sensing and forest ecology (Gao, Skutsch, Paneque-Gálvez, & Ghilardi, 2020). The urgency for developing robust and scalable tree species classification methodologies is underscored by the ongoing challenges of biodiversity loss, climate change, and land-use pressures (Krivoguz, 2024). Accurate and efficient classification tools are essential for informing conservation strategies, managing forest resources sustainably, and understanding the ecological dynamics of forested landscapes (Nitoslawski et al., 2021). Despite the progress made in applying machine learning techniques to LiDAR data for species classification, significant challenges remain. Many existing studies have focused on binary classification problems or a limited selection of species, with less attention given to the complexities of multi-class classification across a broader range of tree species (Rapinel & Hubert-Moy, 2021). Moreover, there is a recognized need for improved feature extraction techniques that can more effectively capture the distinguishing characteristics of different species from the data provided by LiDAR scans.

In addressing these challenges, this research not only contributes to the methodological advancements in the field but also provides practical insights that could significantly impact forest management and conservation practices. By offering a comprehensive comparison of machine learning techniques for the multi-class classification of tree species and introducing innovative approaches to feature extraction from LiDAR data, the study aims to enhance the accuracy and applicability of remote sensing methods in ecological research and practice. The remainder of the article is structured to guide the reader through the research process and findings comprehensively. The Methods section will detail the LiDAR data collection, preprocessing, feature extraction techniques, and the specifics of the machine learning models used. The Results will present the outcomes of the comparative analysis, focusing on the performance metrics such as accuracy, precision, recall, and F1 scores for each model. In the Discussion, the implications of these findings will be explored in the context of existing literature, highlighting the study's contributions to forest management and conservation strategies, as well as its limitations and potential avenues for future research. Finally, the Conclusion will summarize the key points of the study, emphasizing its contributions to the field and suggesting directions for further investigation into the classification of tree species using LiDAR data and advanced machine learning techniques. This research article, by delving into the complexities of multi-class classification and exploring a range of computational approaches, aims to provide a significant contribution to the conservation and management of forest ecosystems worldwide.

LITERATURE REVIEW

Initial approaches to tree species classification from LiDAR data primarily utilized basic statistical and geometric features extracted from the point clouds, such as height, crown diameter, and volume (Michałowska & Rapiński, 2021). These studies laid the groundwork for subsequent research by demonstrating the potential of LiDAR data to differentiate between broadleaf and coniferous trees based on structural characteristics. However, these methods often struggled with finer-scale distinctions between species within these broad categories (Iljas, 2022). The integration of machine learning techniques with LiDAR data analysis marked a significant turning point in the field. Algorithms ranging from decision trees and random forests to support vector machines (SVMs) and neural networks began to be applied to the classification problem, offering more nuanced and accurate identification of tree species (Wang, 2023). These studies capitalized on the ability of machine learning models to process and learn from the high-dimensional and complex nature of LiDAR data, achieving substantial improvements in classification accuracy (El-Omairi & El Garouani, 2023).

More recent research has focused on developing sophisticated feature extraction techniques and exploring innovative machine learning models to further enhance classification performance. For example, studies have employed deep learning algorithms, which can automatically learn hierarchical features from raw LiDAR data, showing promising results in distinguishing between species with similar structural traits (Ma, Zhao, Im, Zhao, & Zhen, 2024). Additionally, the use of multi-source data fusion, combining LiDAR with hyperspectral or RGB imagery, has been explored to leverage complementary information, enriching the feature set available for classification (Mohammadpour & Viegas, 2022). Despite these advancements, several challenges persist in the field of tree species classification from LiDAR data. First, most studies have concentrated on binary classification tasks (coniferous vs. deciduous) or a limited number of species, often within controlled experimental sites (Illarionova, Trekin, Ignatiev, & Oseledets, 2021). There is a notable gap in research addressing multi-class classification across a diverse array of tree species in natural, heterogeneous forests (Rapinel & Hubert-Moy, 2021). Second, the effectiveness of different machine learning algorithms in processing LiDAR data for species classification has not been comprehensively compared, particularly in terms of their robustness, scalability, and generalizability across different forest types and regions (Xi, Hopkinson, & Chasmer, 2024). Lastly, there is an ongoing need for innovative feature extraction methods that can more accurately capture the unique structural characteristics of individual tree species from LiDAR data (Gharineiat et al., 2022). The

E-ISSN : 2807-9035
Volume 4, **Number** 1, May 2024
https://doi.org/10.47709/brilliance.v4i1.3673

current study builds upon this body of literature by addressing these identified gaps. It contributes to the field by implementing a comparative analysis of multiple machine learning techniques for the multi-class classification of tree species, emphasizing the need for robust, scalable, and generalizable models. Furthermore, it explores advanced feature extraction methods to improve the accuracy and specificity of tree species classification from LiDAR data, aiming to enhance the ecological and conservation applications of remote sensing technologies.

METHOD

Data Acquisition

The acquisition of airborne LiDAR data was a critical first step in this study, targeting multiple forest sites characterized by a rich diversity of tree species including Alder, Aspen, Birch, Fir, Pine, Spruce, and Tilia. Utilizing a high-precision laser scanner mounted on an aerial platform, the data collection was executed at an average altitude of 800 meters above the ground. This strategic altitude facilitated the attainment of a dense point cloud, with an approximate point density of 10 points per square meter, ensuring a high level of detail in the structural representation of the forest canopy. The collected dataset comprised raw point clouds along with derived metrics such as Canopy Height Models (CHMs), Digital Elevation Models (DEMs), and Digital Surface Models (DSMs). These derived metrics play a pivotal role in subsequent analyses, offering a quantified perspective of the forest structure above ground level. The data can be downloaded in (Sentinel3734, 2023).

Preprocessing

Preprocessing the raw LiDAR point clouds was essential to refine the data for accurate analysis. Initially, noise present in the data was minimized through a statistical outlier removal process, which is crucial for eliminating anomalies that could skew the analysis. Following this, a digital terrain model (DTM) was employed to distinguish ground points from vegetation points. This distinction is vital for isolating the vegetation structure from the ground surface, allowing for a more focused analysis of the canopy. Subsequently, the data underwent a normalization process, where the DTM was subtracted from the vegetation points, yielding the height of vegetation points above ground level. Normalization ensures that the analysis concentrates on the vegetative features rather than the underlying topography, which is instrumental for accurate feature extraction and species classification.

Feature Extraction

The feature extraction process aimed to distill the structural characteristics indicative of species identity from the LiDAR data. This involved a multifaceted approach combining first-order statistical features (including the mean, standard deviation, skewness, and kurtosis of point heights), which provide a basic statistical overview of the canopy structure, with second-order texture features derived from the gray-level co-occurrence matrix (GLCM) of the CHMs. These texture features, such as entropy, contrast, and homogeneity, offer insights into the spatial variability and texture patterns within the canopy, reflecting species-specific architectural traits. Additionally, three-dimensional structural features, encompassing canopy volume, rugosity (a measure of surface roughness), and the vertical distribution of foliage, were calculated to capture the complex, species-specific structural diversity of the forest canopy. The selection of these features was informed by their proven significance in differentiating tree species in previous studies, ensuring a comprehensive set of indicators for accurate classification.

Machine Learning Models

The study evaluated an array of machine learning algorithms for their efficacy in classifying tree species from the extracted LiDAR features. This evaluation encompassed a wide range of models, from the relatively simple Logistic Regression and k-Nearest Neighbors (k-NN) to more complex algorithms like Support Vector Machines (SVM), Random Forests (RF), Decision Trees (DT), and Gradient Boosting (GB). Each model was chosen based on its ability to manage the high-dimensional nature of the data and its applicability to similar ecological and remote sensing challenges in prior research. Implementation was carried out using the Python scikit-learn library, with a rigorous hyperparameter optimization process conducted via grid search and cross-validation on a designated subset of the data to ensure the optimal configuration of each model.

Model Training and Validation

For training and validation, the dataset was strategically partitioned into a 60% training set and a 40% testing set. This division was designed to mitigate the risk of overfitting and to assess the models' generalizability. A 10-fold cross-validation method was employed during the training phase to further validate the models' effectiveness and robustness. The performance of the models was quantitatively assessed using a suite of metrics, including accuracy, precision, recall, and the F1 score, derived from the confusion matrix of observed versus predicted species classifications. These metrics provided a holistic view of each model's classification performance, enabling a comprehensive comparison and selection of the most effective algorithm for tree species classification from LiDAR data.

E-ISSN : 2807-9035
Volume 4, **Number** 1, May 2024
https://doi.org/10.47709/brilliance.v4i1.3673

RESULT

The results section of a research paper is where the outcomes of the statistical analyses are presented. Based on the table 1, which summarizes the performance of various machine learning models used for tree species classification using LiDAR data, the results section can be structured to discuss each method and its corresponding performance metrics. In assessing the performance of machine learning models for the classification of tree species from LiDAR data, six algorithms were evaluated: Logistic Regression, Decision Tree, Random Forest, Support Vector Classifier (SVC), k-Nearest Neighbors (KNN), and Gradient Boosting. The effectiveness of each model was measured using four key metrics: accuracy, precision, recall, and F1 score.

Firstly, Logistic Regression emerged as one of the top-performing models, achieving an accuracy of 0.81. This indicates that the model correctly predicted the tree species 81% of the time. The precision score of 0.80 suggests that when the model predicted a particular tree species, it was correct 80% of the time. The recall of 0.81 indicates that the model successfully identified 81% of all available instances of each species. The F1 score, which balances precision and recall, was also 0.80, reflecting a strong overall performance of the Logistic Regression model. Furthermore, Decision Tree, a model known for its simplicity and interpretability, demonstrated moderate performance with an accuracy of 0.77. This model's precision and recall were relatively balanced, with scores of 0.78 and 0.75, respectively. The F1 score for the Decision Tree was 0.77, which, while respectable, suggests room for improvement compared to more complex models. Then, Random Forest, an ensemble method that builds multiple decision trees and merges them to get more accurate and stable predictions, matched the performance of Logistic Regression with an accuracy, precision, recall, and F1 score all at 0.80. This highlights the Random Forest's capability as a robust classifier for this type of multi-class classification task.

Furthermore, Support Vector Classifier (SVC), which is often preferred for high-dimensional space classification, showed lower performance metrics in comparison to Logistic Regression and Random Forest. It achieved an accuracy of 0.75, with a matching precision and recall, suggesting consistent but not outstanding performance across different aspects of classification. The F1 score of 0.75 indicates that the balance between precision and recall is maintained.

On the other hand, k-Nearest Neighbors (KNN), a model that classifies data based on the majority vote of its neighbors, with the data being assigned to the class most common among its k nearest neighbors, displayed the weakest performance among the evaluated models with an accuracy of 0.60. Both precision and recall were also at 0.60, and the F1 score was slightly lower at 0.59. These results imply that KNN may not be as effective for the complex task of classifying multiple tree species using LiDAR data, possibly due to the curse of dimensionality affecting its performance. Lastly, Gradient Boosting, another ensemble technique that builds on weak learners and focuses on errors from previous iterations to improve its predictions, showed a good performance with an accuracy of 0.77. Precision and recall were both at 0.77, with a slightly lower F1 score of 0.76. This indicates that Gradient Boosting was reasonably effective but did not outperform the Logistic Regression and Random Forest models in this study. The results suggest that while all models provided valuable insights into the classification of tree species from LiDAR data, Logistic Regression and Random Forest were the most effective according to the metrics used in this study. However, each model has its own strengths and could be suitable for different aspects of the classification task or under varying conditions and parameter tunings.

Table 1. The Comparison Results of Machine Learning Methods

Methods	Accuracy	Precision	Recall	F1 Score
Logistic Regression	0.81	0.80	0.81	0.80
Decision Tree	0.77	0.78	0.75	0.77
Random Forest	0.81	0.80	0.81	0.80
SVC	0.75	0.75	0.75	0.75
KNN	0.60	0.60	0.60	0.59
Gradient Boosting	0.77	0.77	0.77	0.76

DISCUSSION

The purpose of this study was to evaluate the performance of several machine learning models for the classification of tree species using LiDAR data. The results revealed notable differences in the effectiveness of the models, which are indicative of their suitability for processing and interpreting the complex, high-dimensional nature of LiDAR data in the context of ecological classification tasks. The Logistic Regression and Random Forest models outperformed other algorithms in terms of accuracy, precision, recall, and F1 score. These results are consistent with previous studies that have found these models to be reliable and effective for classification tasks involving environmental data (References to prior studies). Logistic Regression's high performance may be attributed to its ability to handle linear relationships between features, which are prevalent in datasets derived from physical measurements such as those obtained from LiDAR. On the other hand, the Random Forest model's success is likely due to its ensemble approach, which reduces the risk of overfitting—a common concern in complex datasets—and increases the

E-ISSN : 2807-9035
Volume 4, **Number** 1, May 2024
https://doi.org/10.47709/brilliance.v4i1.3673

generalizability of the model's predictions.

The Decision Tree model, while not performing as strongly as Logistic Regression or Random Forest, still achieved moderate accuracy and may be favored in scenarios where model interpretability is a priority. Decision Trees provide transparent reasoning for their classifications, which can be crucial for applications requiring explainability, such as informing policy decisions or management practices in forestry. In contrast, the Support Vector Classifier (SVC) and Gradient Boosting models displayed lower performance metrics. The relatively lower performance of the SVC model could be due to its sensitivity to the choice of the kernel and hyperparameter settings, which may not have been fully optimized for the dataset at hand. Gradient Boosting's performance, while better than SVC, still did not match the top-performing models. This might be due to the model's complexity and the risk of overfitting despite efforts to optimize its parameters. The k-Nearest Neighbors (KNN) algorithm showed the weakest performance across all metrics. This outcome might stem from the model's reliance on local information, which can be a disadvantage when dealing with high-dimensional data where the notion of 'neighborhood' becomes less meaningful due to the curse of dimensionality. The dimensions in LiDAR data represent different attributes of the forest structure, which may not be effectively captured in the feature space utilized by KNN. Furthermore, KNN is sensitive to noise in the dataset, and while preprocessing aimed to minimize this, some level of noise is inevitable, potentially impacting the model's performance.

One of the study's limitations is the potential for class imbalance, which could disproportionately affect the performance metrics of the models. Future studies should ensure a balanced representation of each tree species to generalize the models effectively. Additionally, while the high dimensionality of the LiDAR data provides rich information, it also introduces challenges in model training and optimization. Selecting or engineering the right features from LiDAR data that are most informative for species classification remains an area for further research. The implications of this study are significant for ecological monitoring and forest management. The ability to accurately classify tree species over large areas using LiDAR data can aid in the conservation of biodiversity, the management of natural resources, and the assessment of ecological health. The superior performance of Logistic Regression and Random Forest models suggests that these could be applied in practical settings, providing a balance between accuracy and interpretability.

CONCLUSION

This study undertook a comparative analysis of multiple machine learning algorithms to classify tree species using airborne LiDAR data, a critical task for ecological monitoring and forest management. The investigation revealed that Logistic Regression and Random Forest models outperformed other algorithms such as Decision Trees, SVC, KNN, and Gradient Boosting in terms of accuracy, precision, recall, and F1 score. These findings suggest that, for the dataset and feature set used, Logistic Regression and Random Forest offer a robust approach for the classification of tree species from LiDAR data, balancing model complexity with predictive performance. The successful application of these models has significant implications for the fields of remote sensing and forest ecology. The ability to accurately distinguish between tree species across large areas can enhance the management of forest resources, contribute to the conservation of biodiversity, and inform climate change mitigation strategies through more accurate carbon sequestration assessments.

One of the notable contributions of this research is the demonstration of the effectiveness of machine learning in interpreting complex environmental data, an area that continues to grow in importance as we develop more advanced remote sensing technologies. However, the study is not without its limitations. The performance of the models could be further improved by addressing potential class imbalance and optimizing feature selection to better capture the unique characteristics of each species. Future research should focus on incorporating a more diverse set of environmental variables, potentially including multi-spectral or hyperspectral data, to enrich the models' inputs. Experimenting with deep learning architectures could also yield improvements by automating the feature extraction process and capturing more intricate patterns within the data. Moreover, extending the analysis to include a broader range of tree species and different geographical locations would enhance the generalizability and applicability of the findings.

REFERENCES

Blanco, J. A., Ameztegui, A., & Rodr\'\iguez, F. (2020). Modelling Forest Ecosystems: A crossroad between scales, techniques and applications. *Ecological Modelling*, Vol. 425, p. 109030. Elsevier.

Camarretta, N., Harrison, P. A., Bailey, T., Potts, B., Lucieer, A., Davidson, N., & Hunt, M. (2020). Monitoring forest structure to guide adaptive management of forest restoration: a review of remote sensing approaches. *New Forests*, *51*, 573–596.

Chehreh, B., Moutinho, A., & Viegas, C. (2023). Latest Trends on Tree Classification and Segmentation Using UAV Data—A Review of Agroforestry Applications. *Remote Sensing*, *15*(9), 2263.

El-Omairi, M. A., & El Garouani, A. (2023). A review on advancements in lithological mapping utilizing machine learning algorithms and remote sensing data. *Heliyon*.

E-ISSN : 2807-9035
Volume 4, **Number** 1, May 2024
https://doi.org/10.47709/brilliance.v4i1.3673

Gao, Y., Skutsch, M., Paneque-Gálvez, J., & Ghilardi, A. (2020). Remote sensing of forest degradation: a review. *Environmental Research Letters*, *15*(10), 103001.

Gharineiat, Z., Tarsha Kurdi, F., & Campbell, G. (2022). Review of automatic processing of topography and surface feature identification LiDAR data using machine learning techniques. *Remote Sensing*, *14*(19), 4685.

Hastings, J. H., Ollinger, S. V, Ouimette, A. P., Sanders-DeMott, R., Palace, M. W., Ducey, M. J., … Orwig, D. A. (2020). Tree species traits determine the success of LiDAR-based crown mapping in a mixed temperate forest. *Remote Sensing*, *12*(2), 309.

Iljas, T. (2022). *Using Very High Resolution Remotely Piloted Aircraft Imagery to Map Peatland Vegetation Composition and Configuration Patterns within an Elevation Gradient*. University of Waterloo.

Illarionova, S., Trekin, A., Ignatiev, V., & Oseledets, I. (2021). Tree species mapping on sentinel-2 satellite imagery with weakly supervised classification and object-wise sampling. *Forests*, *12*(10), 1413.

Krivoguz, D. (2024). Geo-spatial analysis of urbanization and environmental changes with deep neural networks: Insights from a three-decade study in Kerch peninsula. *Ecological Informatics*, 102513.

Ma, Y., Zhao, Y., Im, J., Zhao, Y., & Zhen, Z. (2024). A deep-learning-based tree species classification for natural secondary forests using unmanned aerial vehicle hyperspectral images and LiDAR. *Ecological Indicators*, *159*, 111608.

Mackey, B., Kormos, C. F., Keith, H., Moomaw, W. R., Houghton, R. A., Mittermeier, R. A., … Hugh, S. (2020). Understanding the importance of primary tropical forest protection as a mitigation strategy. *Mitigation and Adaptation Strategies for Global Change*, *25*(5), 763–787.

Michałowska, M., & Rapiński, J. (2021). A review of tree species classification based on airborne LiDAR data and applied classifiers. *Remote Sensing*, *13*(3), 353.

Mohammadpour, P., & Viegas, C. (2022). Applications of Multi-Source and Multi-Sensor Data Fusion of Remote Sensing for Forest Species Mapping. *Advances in Remote Sensing for Forest Monitoring*, 255–287.

Muluneh, M. G. (2021). Impact of climate change on biodiversity and food security: a global perspective—a review article. *Agriculture \& Food Security*, *10*(1), 1–25.

Nitoslawski, S. A., Wong-Stevens, K., Steenberg, J. W. N., Witherspoon, K., Nesbitt, L., & van den Bosch, C. C. (2021). The digital forest: Mapping a decade of knowledge on technological applications for forest ecosystems. *Earth's Future*, *9*(8), e2021EF002123.

Rapinel, S., & Hubert-Moy, L. (2021). One-class classification of natural vegetation using remote sensing: a review. *Remote Sensing*, *13*(10), 1892.

Sentinel3734. (2023). *UAV Point Clouds of Individual Trees*.

Wang, C. K. (2023). Sentiment Analysis Using Support Vector Machines, Neural Networks, and Random Forests. *2023 International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2023)*, 23–34.

Xi, Z., Hopkinson, C., & Chasmer, L. (2024). Supervised terrestrial to airborne laser scanner model calibration for 3D individual-tree attribute mapping using deep neural networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, *209*, 324–343.

Xu, D., Wang, H., Xu, W., Luan, Z., & Xu, X. (2021). LiDAR applications to estimate forest biomass at individual tree scale: Opportunities, challenges and future perspectives. *Forests*, *12*(5), 550.

Yip, K. H. A., Liu, R., Wu, J., Hau, B. C. H., Lin, Y., & Zhang, H. (2024). Community-based plant diversity monitoring of a dense-canopy and species-rich tropical forest using airborne LiDAR data. *Ecological Indicators*, *158*, 111346.