



Exploring Tree Species Classification from UAV Orthophotos Using Foundation Models and Transfer Learning

Abstract

Monitoring tree species at fine spatial resolutions is essential for biodiversity assessment, ecological research, and forest management. The growing availability of open-access UAV datasets has facilitated the development of deep learning methods for fine-grained forest monitoring, particularly in tasks such as tree crown detection and species classification [1].

In recent years, AI-driven approaches, including deep learning models, have been increasingly used for tree species classification, though challenges remain, such as high intra-class variability, spectral similarity among species, limited annotated data, and complex canopy structures.

Numerous studies have leveraged machine learning techniques for AI-based tree species classification, but many models still face limitations in generalization across different forest types and environmental conditions [2].

The Quebec Trees Dataset has become a valuable benchmark for tree species classification, providing a high-resolution, UAV-based orthomosaic dataset from temperate forests [3]. It comprises 21 UAV-derived orthomosaics acquired in 2021 during different phenological stages, with a ground sampling distance of approximately 2 cm. More than 23,000 manually annotated tree crowns are included across 14 species or genera. Each orthomosaic was generated using Structure-from-Motion photogrammetry and supported by accurate field-collected reference data.

Despite its high annotation quality and temporal richness, recent studies using this dataset have reported moderate classification accuracies, particularly for rare species, indicating room for improvement in model robustness and generalization.

To address these challenges, we explore the use of foundation models [4]—large pretrained neural networks originally developed for general computer vision tasks—and apply transfer learning to the UAV-based tree species classification task. By fine-tuning these models on the Quebec Trees Dataset, we aim to improve classification accuracy, particularly under conditions of class imbalance and complex canopy structures.

In this study, we implement a U-Net architecture with a ResNet-50 encoder, comparing models trained from scratch to those initialized with pretrained weights. Preliminary results indicate that while the overall improvement remains limited, pretrained models offer more stable training behavior and perform better on underrepresented species.

This work contributes to the advancement of AI-based ecological monitoring using UAV data and highlights the potential of foundation models for improving tree species classification. We also provide a perspective on the evolving research landscape surrounding the Quebec Trees Dataset and its growing use in remote sensing and forest informatics.

References

- [1] Zhong, L., Dai, Z., Fang, P., Cao, Y., & Wang, L. (2024). A Review: Tree Species Classification Based on Remote Sensing Data and Classic Deep Learning-Based Methods. *Forests*, 15(5), 852. <https://doi.org/10.3390/f15050852>
- [2] Lu, S., et al. (2024). A deep-learning-based tree species classification for natural secondary forests using unmanned aerial vehicle hyperspectral images and LiDAR. *Ecological Indicators*, 160, 113456. <https://doi.org/10.1016/j.ecolind.2024.113456>
- [3] NIAID Data Discovery Portal. (2024). Quebec Trees Dataset. https://data.niaid.nih.gov/resources?id=zenodo_8148478
- [4] Bountos, N. I., Ouaknine, A., Papoutsis, I., & Rolnick, D. (2025). FoMo: Multi-Modal, Multi-Scale and Multi-Task Remote Sensing Foundation Models for Forest Monitoring. *Proceedings of the 39th AAAI Conference on Artificial Intelligence*. <https://arxiv.org/abs/2312.10114>

Primary authors: Mr MUTREJA, Guneet (German Aerospace Center (DLR), Remote Sensing Technology Institute); Dr TIAN, Jiaojiao (German Aerospace Center (DLR), Remote Sensing Technology Institute); BEN GHORBEL, Youcef (German Aerospace Center (DLR), Remote Sensing Technology Institute)

