Classification of Forest Tree Species Using High-Resolution UAV Imagery and Deep Learning

Josafat-Mattias Burmeister[1] Stefan Reder Koithahn[2] David Kuska[3] Rico Richter[1] Jan-Peter Mund[2] Jürgen Döllner[1,3]

[1]University of Potsdam, Digital Engineering Faculty, Potsdam, Germany

[2]Eberswalde University for Sustainable Development, Faculty of Forest and Environment, Eberswalde, Germany

[3]Hasso Plattner Institute, Potsdam, Germany

Context and Research Objective

- UAV imagery with RGB information is well-suited for automated tree species mapping because it is cost-efficient to acquire and depicts species-specific details of the canopy structure
- This study develops a deep learning approach for tree species mapping in high-resolution UAV RGB imagery (spatial resolution < 5cm) of temperate mixed forests in Germany
- Image-based tree species mapping is modeled as a semantic segmentation task
- To obtain a tree-level species map, the semantic segmentation results are combined with point-cloud-based tree instance segmentation

Datasets

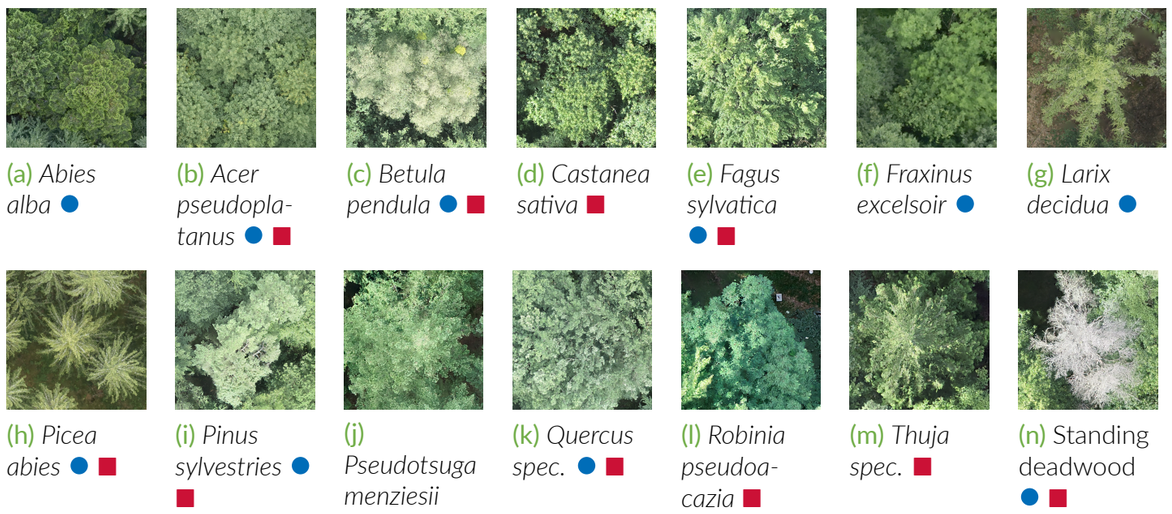
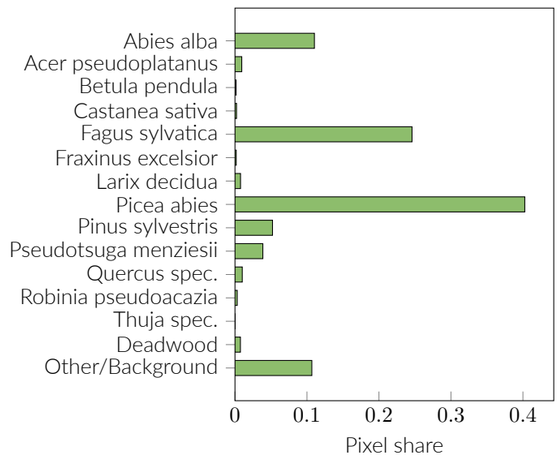


Figure 1. Examples of the species categories included in the datasets used in this study. ● indicates that a species is included in the FORTRESS dataset and ■ that it is included in the Brandenburg dataset.

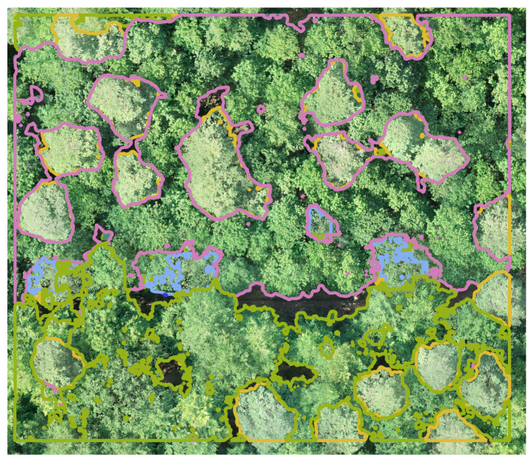
- Two datasets were combined in this study:
  - FORTRESS dataset: [4, 3]
    - Publicly available dataset collected in the Southern Black Forest, Germany
    - Covers approx. 59 ha of different forest types (mixed, deciduous, and coniferous)
    - Imagery was collected between April and October, with pixel sizes of 0.7 cm to 1.9 cm
  - Brandenburg dataset
    - Custom dataset collected in Brandenburg, Germany
    - Covers approx. 5 ha of mixed forest sites
    - Imagery was collected between May and August, with pixel sizes of 1.3 cm to 1.9 cm
    - Field inventories were conducted for selected plots and used as a reference for labeling
  - The labels of both datasets were harmonized and rare classes (*Aesculus spec.*, *Carpinus betulus*, *Fallopia spec.*, *Ilex spec.*) were aggregated into the class "Other/Background"
  - For model training and evaluation, the images were split into non-overlapping tiles (256 pixels wide), and the plots were split into training, validation, and test areas

Model Architecture and Training Settings

- Three model architectures were tested: U-Net++ [6], DeepLabv3 [2], SegFormer [5]
- Each model was trained for 50 epochs, using a class-weighted cross-entropy loss function with label smoothing of 0.1
- The training dataset was augmented using random rotations and vertical and horizontal flips



(a) Class distribution of the training data after combining the FORTRESS and the Brandenburg dataset.



(b) Example predictions of the U-Net++ model (purple = *Castanea sativa*, green = *Fagus sylvatica*, yellow = *Pinus sylvestris*, blue = *Quercus spec.*)

Model Accuracy

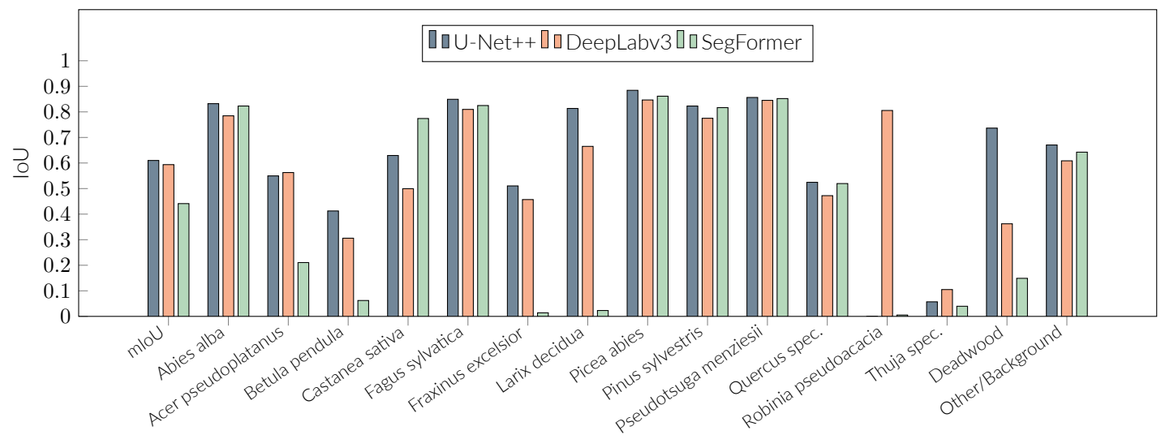


Figure 3. Intersection over Union (IoU) scores of the evaluated model architectures on the test set.

Fusion with 3D Point Cloud-Based Instance Segmentation

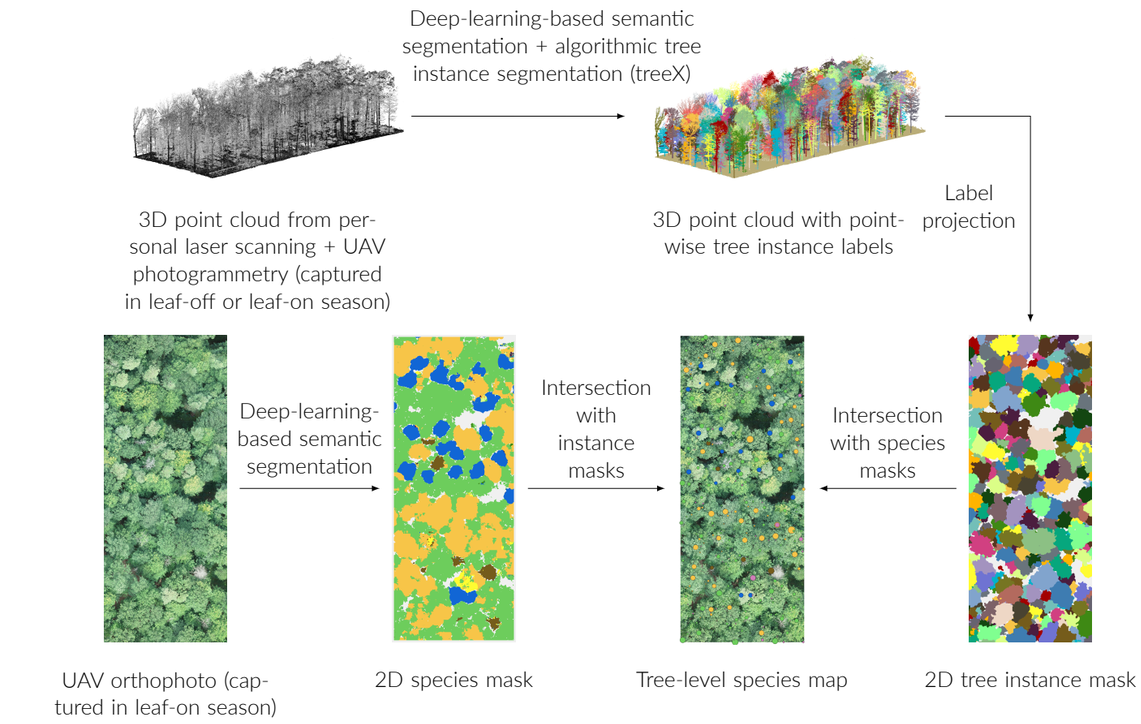


Figure 4. Approach for deriving a tree-level species map by combining image-based semantic segmentation of tree species with 3D point cloud-based tree instance segmentation. The intersection of species masks and instance masks is done using a majority voting approach.

Main Findings and Outlook

- Pixel-wise tree species maps can be automatically derived from high-resolution RGB UAV imagery
- The U-Net++ architecture outperforms the DeepLabv3 and SegFormer architectures, especially for rare classes
- Additional training data (more species, sites, and acquisition times) are required to improve generalizability
- Combining pixel-wise species masks with instance masks from 3D point clouds enables tree-level species mapping → this requires a co-registration of 3D point clouds and UAV imagery, and accurate tree instance segmentation labels
- Instance masks derived from 3D point clouds could also be used for semi-automatic generation of additional tree species labels in the future

References

[1] Josafat-Mattias Burmeister, Andreas Tockner, Stefan Reder, Markus Engel, Rico Richter, Jan-Peter Mund, and Jürgen Döllner. treex: Unsupervised tree instance segmentation in dense forest point clouds. *arXiv preprint*, 2509.03633, 2025. doi: 10.48550/arXiv.2509.03633.

[2] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Computer Vision – ECCV 2018*, pages 833–851. Springer, 2018. doi: 10.1007/978-3-030-01234-2_49.

[3] Felix Schiefer, Teja Kattenborn, Annett Frick, Julian Frey, Peter Schall, Barbara Koch, and Sebastian Schmidtlein. Mapping forest tree species in high resolution UAV-based RGB-imagery by means of convolutional neural networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, 170: 205–215, 2020. doi: 10.1016/j.isprsjprs.2020.10.015.

[4] Felix Schiefer, Julian Frey, and Teja Kattenborn. FORTRESS dataset. *Karlsruhe Institute of Technology*, 2022. doi: 10.35097/538.

[5] Enze Xie, Wenhai Wang, Zhiding Yu, Anima Anandkumar, Jose M. Alvarez, and Ping Luo. SegFormer: Simple and efficient design for semantic segmentation with transformers. In *Advances in Neural Information Processing Systems*, volume 34, pages 12077–12090. Curran Associates, Inc., 2021. URL `https://proceedings.neurips.cc/paper_files/paper/2021/file/64f1f27bf1b4ec22924fd0acb550c235-Paper.pdf`.

[6] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. UNet++: A nested U-Net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pages 3–11. Springer, 2018.

With funding from the:

Federal Ministry of Research, Technology and Space

FONA Forschung für Nachhaltigkeit

REGULUS Regionale Innovationsgruppen für eine klimaschützende Wald- und Holzwirtschaft

TREEdigitalTWINS

Universität Potsdam

HPI

Eberswalde University for Sustainable Development