ADVANCING LARGE-SCALE TREE CROWN DELINEATION AND CLASSIFICATION: A SEMI-SUPERVISED APPROACH LEVERAGING DEEP LEARNING AND REMOTE SENSING DATA IN THE BAVARIAN FOREST NATIONAL PARK

S. Dersch[a], B. Koch[a], A. Schöttl[e,f], P. Krzystek[e,g], J. Ebenbeck[h], M. Heurich[b,c,d,a]

[a] Steinbeis Innovation Center: System Analysis for Environment and Sustainability, 84079 Bruckberg, Germany
(sebastian.dersch, barbara.koch)@steinbeis.de

[b] Faculty of Environment and Natural Resources, University of Freiburg, 79085 Freiburg im Breisgau, Germany

[c] Bavarian Forest National Park, Dept. of Visitor Management and National Park Monitoring, 94481 Grafenau, Germany
(marco.heurich, jakob.ebenbeck)@npv-bw.bayern.de

[d] Institute of Forestry and Wildlife Management, Inland Norway University of Applied Science, NO-2480 Koppang, Norway

[e] Institute for Applications of Machine Learning and Intelligent Systems, Munich University of Applied Sciences, 80335 Munich, Germany
(peter.krzystek, alfred.schoettl)@hm.edu

[f] Dept. of Electrical Engineering and Information Technology, Munich University of Applied Sciences, 80335 Munich, Germany

[g] Dept. of Geoinformatics, Munich University of Applied Sciences, 80333 Munich, Germany

[h] The Bavarian State Institute of Forestry (LWF), 85354 Freising, Germany

ABSTRACT

Large-scale tree crown delineation and classification, including standing dead trees, provides essential information for forest management, ecological monitoring, and biodiversity assessment. While deep learning-based instance segmentation methods for tree crown delineation achieve high accuracies, their operational use is often limited by the extensive effort required for manual training data generation. This contribution addresses the challenge of scaling single-tree analysis to large forest areas while reducing the dependency on manually labelled reference data. The approach is particularly relevant for stakeholders involved in long-term forest monitoring and decision-making, including national park administrations, forest authorities, conservation agencies, and researchers working on biodiversity assessment, disturbance monitoring, and climate-change impacts on forests.

We present a semi-supervised framework that combines deep learning-based instance segmentation with clustering-based pre-labelling for single tree crown delineation. Specifically, Mask R-CNN is integrated with a normalized cut clustering approach, where tree-level training labels are automatically derived from airborne LiDAR point clouds. This methodology substantially reduces the reliance on manually labelled training data and enables the use of significantly larger remote sensing datasets for model training. In addition to tree crown delineation and classification into three tree groups - deciduous trees, coniferous trees, and standing dead trees - tree-level structural parameters such as tree height, crown volume, and crown base height are derived for each individual tree. A key component of the framework is the voxel-based analysis of airborne LiDAR point clouds, which enables the robust estimation of three-dimensional tree structure metrics at scale. This approach supports the detailed analysis of vertical tree structure, allowing structural attributes and vertical distribution patterns to be systematically derived and analyzed across different forest types.

The study was conducted and evaluated in the Bavarian Forest National Park (BFNP) in south-east Germany, which is characterized by diverse forest structures dominated by European beech (*Fagus sylvatica*), Norway spruce (*Picea abies*), and silver fir (*Abies alba*). Multispectral image data with a

ground sampling distance of 10 cm and airborne LiDAR data with an average point density of approximately 55 points/m², acquired in June 2017, formed the basis for the analysis. Model evaluation was performed on seven reference plots comprising 516 trees, using a combination of field measurements and expert visual interpretation as reference data.

The results demonstrate the effectiveness of the proposed approach. Mask R-CNN models trained on CIR imagery achieved a mean F1 score of up to 79% for tree crown delineation, outperforming the normalized cut baseline method by up to 18%. Furthermore, the overall classification accuracy for coniferous, deciduous, and standing deadwood tree groups reached 96%, representing an improvement of 6% compared with the baseline approach.

Building on these promising results, the methodology is currently being extended to encompass the entire BFNP, enabling comprehensive and consistent tree crown delineation and classification at landscape scale. This large-scale application will evaluate model performance on remote sensing datasets acquired in different years, allowing for an assessment of robustness and adaptability to temporal variations in forest structure and composition. The operational deployment of the software framework enables the automated and parallelized processing of very large remote sensing datasets and supports the efficient derivation of tree-level structural parameters for several million individual trees.

Overall, the presented framework holds significant potential for advancing remote sensing applications in forest monitoring, offering a scalable, cost-effective, and accurate tool for addressing a wide range of ecological and forest management challenges.