Deep Learning-based classification of tree species and standing dead trees using Silvi-Net

S. Briechle[1], P. Krzystek[1], G. Vosselman[2]

[1]Munich University of Applied Sciences, Munich, Germany
Email: {sebastian.briechle; peter.krzystek}@hm.edu

[2]Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, the Netherlands
Email: george.vosselman@utwente.nl

1. Introduction

Conservationists and forest managers rely on the precise mapping of individual trees from remote sensing data to efficiently derive forest attributes. In recent years, the additional quantification of deadwood in particular has attracted interest. However, tree-level approaches using segmented individual trees are still limited in their accuracy and their application is therefore mostly restricted to research studies. Furthermore, the combined classification of pre-segmented individual trees in terms of tree species and health status is important for practice, but has been insufficiently investigated so far.

In addition, the application of Deep Learning (DL)-based methods for the classification of pre-segmented individual trees based on lidar data has hardly been investigated so far. Hamraz et al. (2019) used a convolutional neural network (CNN) to classify coniferous and deciduous trees in a natural forest (330 stems/ha). By generating images from airborne laser scanning (ALS) point clouds (50 points/m²), a classification accuracy of 92% for conifers and 87% for deciduous trees was achieved. In a tropical wetland in southern China, Sun et al. (2019b) developed a patch-based classification algorithm for seven classes, including six individual tree classes (1,388 training samples, 362 test samples). Their most effective model classified image patches with an OA of 90%. In the same research area, Sun et al. (2019a) mapped 18 tree species using ALS data and high-resolution RGB imagery and achieved an overall accuracy (OA) of 73% at the individual tree level. Recently, Briechle et al. (2020) classified three tree species (pine, birch and alder) and standing dead pine trees with crowns using PointNet++ together with drone-based lidar data and multispectral (MS) imagery. In addition to 3D geometry, laser intensity values and MS features were also integrated into the classification process. Overall, their DL-based method (OA = 90%) was successful using raw 3D data and superior to a baseline method using an RF classifier and hand-crafted features (OA = 85%).

In this work, we introduce Silvi-Net, a dual CNN-based approach fusing airborne lidar data and MS images for 3D object classification.

2. Data and Methods

In our studies, we analysed the performance of Silvi-Net using data collected in two study areas, the Chernobyl Exclusion Zone (ChEZ) and the Bavarian Forest National Park (BFNP). For both study areas, the lidar point density was about 55 points/m² and the ground sampling distance values of the true orthophotos were 10 cm (ChEZ) and 20 cm (BFNP).

Using an interactive tool, single tree segments were manually labeled based on visual interpretation. In order to make our classification results independent of the segmentation quality, incorrect segments were generally not considered in the labeling process. In detail, the trees in the ChEZ were manually subdivided into the classes “pine”, “birch”, “alder”, and “dead tree”. In the BFNP, we labeled the trees with the categories “coniferous” (mostly spruce), “deciduous” (mostly beech and larch), “snag”, and “dead tree”. Here, “snag” refers to a partly or completely dead tree missing a crown or most of the smaller branches, whereas trees labeled “dead tree” are dead trees with crowns. The distinction between “snag” and “dead tree” was based on the subjective perception of three different research assistants. Finally, the labeled samples were randomly sorted into training, validation, and test datasets (Tables 1 and 2). Note that we also included class balancing for both training and validation data.

Published in: Markus Hollaus, Norbert Pfeifer (Eds.): Proceedings of the SilviLaser Conference 2021, Vienna, Austria, 28–30 September 2021. Technische Universität Wien, 2021. DOI: 10.34726/wim.1861
This paper was peer-reviewed. DOI of this paper: 10.34726/wim.1966

Table 1: Number of samples for study area ChEZ; train/val/test split: 56%/14%/30%.

Tree class	Training samples	Validation samples	Test samples
pine	93	23	51
birch	93	23	51
alder	93	23	51
dead tree	93	23	51
Σ	372	92	204

Table 2: Number of samples for study area BFNP; train/val/test split: 51%/22%/27%.

Tree class	Training samples	Validation samples	Test samples
coniferous	345	149	259
deciduous	345	149	202
snag	345	149	139
dead tree	345	149	145
Σ	1380	596	745

The methodology (Figure 1) is as follows: Initially, individual 3D trees are segmented from the lidar point cloud, and 12 silhouette-like side views are rendered and enriched with calibrated laser echo characteristics. Then, the projected outlines of the segmented trees are used to mask the MS orthomosaic and generate one image patch per tree. Next, two independent ResNet-18 networks are trained to learn relevant features from both datasets. This optimisation process is based on pre-trained CNN weights and recursive retraining of the model parameters. Subsequently, the extracted features are fused and fed to the final classification step. Here, we use a standard multi-layer perceptron that outputs 12 predictions per tree. Finally, we utilize majority voting to outvote individual misclassifications.

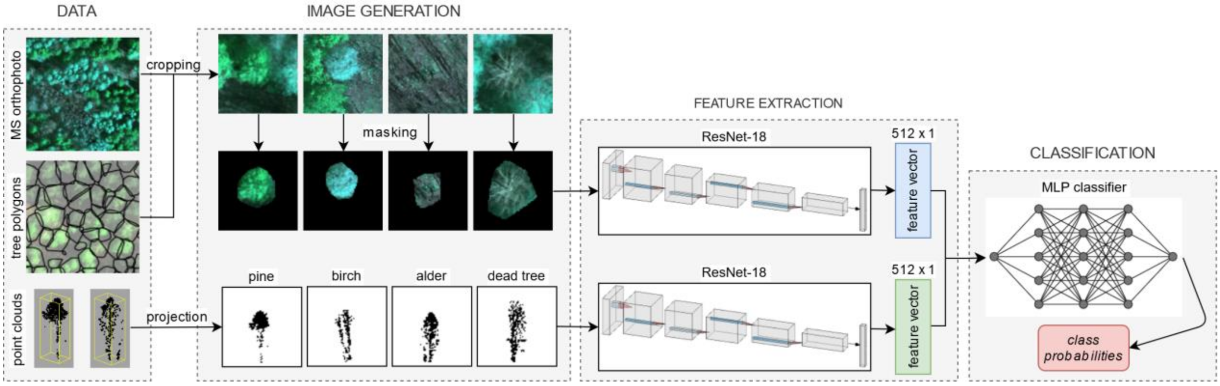

Figure 1: Outline of the proposed method, Silvi-Net.

3. Results and Discussion

In general, the trained models showed a high generalisation capacity on independent test data and achieved an OA of 96.1 % for the classification of pines, birches, alders and dead trees in the ChEZ and 91.5 % for conifers, deciduous trees, snags and dead trees in the BFNP (Figure 2). Interestingly, lidar-based imaging increased OA by 2.5 % (ChEZ) and 5.9 % (BFNP) compared to experiments using MS images only. Furthermore, Silvi-Net showed 11.3 % (ChEZ) and 2.2 % (BFNP) better OA compared to the baseline method PointNet++.

It should be noted that the datasets differ in terms of forest types and sensor models as well as geometric and spectral resolution. Both the ground resolution and the number of spectral channels of the MS images are clearly higher in the ChEZ. Thus, the MS images in this study area contain more extractable information for tree classification which mainly explains the superior results in the ChEZ.

Overall, Silvi-Net enables a convenient fusion of 2D and 3D data acquired by different sensor types. This allows information from the object geometry, laser intensity and reflection in the visible and NIR spectrum to be combined. Crucial here is the automatic extraction of meaningful features from previously generated 2D representations. The technique of transfer learning with pre-trained weights also enables fast model convergence, even with relatively small data sets.

Nevertheless, we want to make clear that a well-functioning upstream segmentation of single trees is mandatory for Silvi-Net to work well. In our study, we used almost perfectly delineated single trees. These were generated by the normalised cut segmentation algorithm by manually labelling optimal segments. This minimised the effect of under- or over-segmentation. From a practical point of view, however, many tree segmentation methods will cause problems in forests with even higher stand density and more canopy complexity.

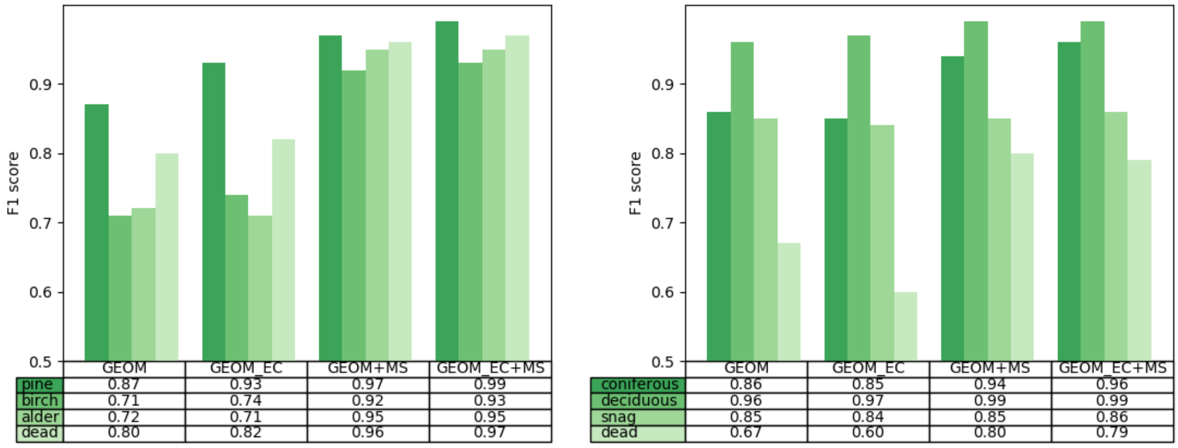

Figure 2: F_1 scores per class for Silvi-Net in the ChEZ (left) and in the BFNP (right), using different feature sets.

4. Conclusions

In this paper, we presented Silvi-Net, a dual CNN-based approach for the combined classification of pre-segmented 3D tree objects, especially with respect to tree species and deadwood. The innovative contribution of our study is the fusion of MS image patches and multiple side views rendered from 3D lidar data in a CNN-based approach. Using the transfer learning technique, Silvi-Net enables fast model convergence, even for datasets with a reduced number of samples.

The effectiveness of our approach has been demonstrated using 2D and 3D datasets from two natural forest areas (400-530 trees/ha) collected with different sensor models and different geometric and spectral resolution. Consequently, users can produce reliable maps, which are of great importance for applications such as automated inventory and monitoring projects.

In future work, the challenge will be to reliably classify ten or more individual tree species and structurally complex forests. This objective can be supported by improved optical sensors providing high-quality lidar point clouds and high-resolution multi-channel images. In addition, off-the-shelf CNNs and transfer learning can be applied to the specific task of tree species classification, even for relatively small datasets.

Acknowledgements

The authors would like to thank Prof. Dr. Marco Heurich, for providing the remote sensing data in the study area BFNP. We also highly appreciate the support from Dr. Norbert Molitor for organizing the field trips in the ChEZ. This research was funded by Federal Ministry of Education and Research (BMBF), Germany (Grant No. 03FH004IX6).

References

Briechle, S., Krzystek, P., Vosselman, G., 2020. Classification of tree species and standing dead trees by fusing UAV-based lidar data and multispectral imagery in the 3D deep neural network PointNet++. ISPRS Ann. Photogramm. Remote Sens. Spatial Informat. Sci., V-2-2020, 203–210. https://doi.org/10.5194/isprs-annals-V-2-2020-203-2020.

Hamraz, H., Jacobs, N., Contreras, M., Clark, C., 2019. Deep learning for conifer/deciduous classification of airborne lidar 3D point clouds representing individual trees. ISPRS J. Photogramm. Remote Sens. 158, 219–230. https://doi.org/10.1016/j.isprsjprs.2019.10.011.

Sun, Y., Huang, J., Ao, Z., Lao, D., Xin, Q., 2019a. Deep learning approaches for the mapping of tree species diversity in a tropical wetland using airborne lidar and high-spatial-resolution remote sensing images. Forests 10. https://doi.org/10.3390/F10111047.

Sun, Y., Xin, Q., Huang, J., Huang, B., Zhang, H., 2019b. Characterizing tree species of a tropical wetland in southern china at the individual tree level based on convolutional neural network. IEEE J. Sel. Top. Appl. Earth Obser. Remote Sens. 12, 4415–4425. https://doi.org/10.1109/JSTARS.2019.2950721.