Classification of Tree Species as Well as Standing Dead Trees Using Triple Wavelength ALS in a Temperate Forest

Nina Amiri [1], Peter Krzystek [2,3,*] , Marco Heurich [4] and Andrew Skidmore [1,2]

[1] Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands; n.amiri@utwente.nl (N.A.); a.k.skidmore@utwente.nl (A.S.)

[2] Department of Earth and Environmental Science, Macquarie University, Sydney, NSW 2109, Australia

[3] Faculty of Geoinformatics, Munich University of Applied Sciences, Karlstr. 6, D-80333 Munich, Germany

[4] Faculty of Environment and Natural Resources, University of Freiburg, 79085 Freiburg im Breisgau, Germany; marco.heurich@wildlife.uni-freiburg.de

* Correspondence: krzystek@hm.edu

check for updates

Abstract: Knowledge about forest structures, particularly of deadwood, is fundamental for understanding, protecting, and conserving forest biodiversity. While individual tree-based approaches using single wavelength airborne laserscanning (ALS) can successfully distinguish broadleaf and coniferous trees, they still perform multiple tree species classifications with limited accuracy. Moreover, the mapping of standing dead trees is becoming increasingly important for damage calculation after pest infestation or biodiversity assessment. Recent advances in sensor technology have led to the development of new ALS systems that provide up to three different wavelengths. In this study, we present a novel method which classifies three tree species (Norway spruce, European beech, Silver fir), and dead spruce trees with crowns using full waveform ALS data acquired from three different sensors (wavelengths 532 nm, 1064 nm, 1550 nm). The ALS data were acquired in the Bavarian Forest National Park (Germany) under leaf-on conditions with a maximum point density of 200 points/m^2. To avoid overfitting of the classifier and to find the most prominent features, we embed a forward feature selection method. We tested our classification procedure using 20 sample plots with 586 measured reference trees. Using single wavelength datasets, the highest accuracy achieved was 74% (wavelength = 1064 nm), followed by 69% (wavelength = 1550 nm) and 65% (wavelength = 532 nm). An improvement of 8–17% over single wavelength datasets was achieved when the multi wavelength data were used. Overall, the contribution of the waveform-based features to the classification accuracy was higher than that of the geometric features by approximately 10%. Our results show that the features derived from a multi wavelength ALS point cloud significantly improve the detailed mapping of tree species and standing dead trees.

Keywords: segmentation; forest structure analysis; dead wood; ALS point cloud; feature importance analysis

1. Introduction

Remote sensing can provide valuable information for ecosystem structure and function over large areas [1] that influence biodiversity [2]. Aside from area-based approaches [3], a large number of studies have mentioned the fundamental role of species identification at the single tree level in much forest inventory and management [4–6] as well as biodiversity [7]. Therefore, in order to maintain

up-to-date information, effective methods and techniques need to be developed that accurately classify tree species.

Single tree species can initially be detected using individual tree segmentation approaches and later mapped by a classification strategy. Recent innovative methods for single tree detection have utilized a 3D approach instead of using the canopy height model (CHM) alone to reduce the over/under-segmentation problems [8]. The detection rates for single trees can be improved significantly by applying the spectral clustering normalized cut method (*NCut*) to a (super) voxel forest structure [9,10], and introducing a classifier-based adaptive stopping criterion [11]. Moreover, to segment individual trees Strîmbu and Strîmbu [12] proposed an approach that captures the topological structure of the forest in hierarchical data structures and quantifies the tree crown component relationships in a weighted graph. Overall, accurate single tree segmentation is an important step for high quality species determinations at the individual tree level.

Over decades, optical imagery that can remotely measure the spectral reflectance of an object has been used as a standard source to discriminate tree species [6,13]. Optical aerial and spaceborne instruments can record the spectral signatures of tree species not only in the visible spectral range (RGB), but also in the near-infrared (NIR), short-wave infrared (SWIR), and even thermal infrared. Depending on the radiometric resolution, the radiation can be measured in multiple bands. Multispectral sensors typically provide up to 10 spectral bands, whereas hyperspectral sensors have hundreds of bands [14]. Recently, spaceborne optical sensors with high spatial and temporal resolution have also been successfully applied to tree species classification. Moreover, dense matching has become a mature technique used to reconstruct objects from a series of highly overlapping images on the pixel level with excellent subpixel accuracy [15]. Regarding forestry applications, this novel computer vision method enables a dense point cloud to be generated from canopy surfaces that later can be used for tree species classification either on the tree level, or using an area-based approach [16]. Moreover, multispectral and hyperspectral sensors can be combined with ALS to enrich the limited radiometric information of ALS. The authors of Ullah et al. [17] demonstrated that estimating forest structural parameters at the stand and forest compartment level can be improved by using point clouds generated from aerial imagery. The authors of Nevalainen et al. [18] combined an RGB camera and a frame format hyperspectral camera mounted on a multicopter drone. They showed that the single tree detection rate is strongly dependent on the forest stand characteristics and ranges between 40% and 95%. Moreover, the tree species classification accuracy (four different tree species) achieved an overall accuracy of 95% and an F-score of 0.93. In this regard, Maschler et al. [19] demonstrated the feasibility of classification of 13 tree species (8 broadleaf, 5 coniferous) with an overall accuracy of 89.4% by combining ALS-based tree segments with hyperspectral data acquired by a Hyspex VNIR 1600 push broom instrument. Finally, the study of Grabska et al. [20] reported a mapping of nine tree species using Sentinel-2 time series. The usage of only two images from different seasons resulted in an overall accuracy higher than 90%. However, the main drawback of these passive sensors for forest applications is the limited forest canopy surface penetration, hence, the forest structure beneath the canopy cannot not be fully captured in 3D.

Over the past decade, ALS point clouds have become an important data source for classifying tree species. Several studies have proposed using structural features from ALS point clouds, such as crown shapes, height distribution percentiles, and proportions of first/single returns for distinguishing between tree species [21–24]. Separating trees by height is important for single tree classification, especially in forests where tree height distributions differ between species [21,25]. After the advent of single full waveform ALS systems, several studies reported accuracy improvements by applying waveform features that use detailed backscattered pulse information, such as the intensity and pulse width [26,27]. The authors of Höfle et al. [28] used calibrated waveforms from the ALS data to distinguish between European larch (*Larix decidua*), English oak (*Quercus robur*), durmast oak (*Quercus petraea*), and European beech (*Fagus sylvatica*), and found that echo width could separate larch from broadleaf trees. However, the responses of the oak and beech represented by the backscatter

cross-section and echo width were similar. Reitberger et al. [9] found that radiometric information derived from full waveform ALS, such as the intensity and pulse width, provide a strong basis for distinguishing between broadleaf and coniferous trees. The authors of Heinzel and Koch [29] explored a set of waveform-based features for classifying four groups of tree species in a mixed temperate forest with an overall accuracy of 78%. The authors of Yao et al. [30] found that single wavelength ALS data (1550 nm) could be advantageously used to classify coniferous and broadleaf trees in the Bavarian Forest National Park with a maximum overall accuracy of 90%. However, Shi et al. [31] found that the classification accuracy decreases by 30% if the detailed tree species mapping (six different tree species) is attempted for the same study area. Further, Hovi et al. [27] focused on systematically analyzing the identification potential of ALS point cloud features by investigating the sources of the within-species variations. They achieved an overall accuracy of 75% for the identification of three main tree species in Finland using ALS waveform features. Overall, considering the limitations of optical imagery, single wavelength ALS point clouds (full waveform) are superior data sources for classifying tree species [32]. However, due to the lack of spectral information, detailed tree species identifications have yet to reach sufficiently high accuracies (up to 90%) [33,34].

The aforementioned results suggest that the intensity of single wavelength ALS (full waveform) is useful for classifying tree species. Moreover, the reflectivity of each tree is dependent on the laser wavelength. For instance, according to [35], by using single wavelength ALS (1064 nm) under leaf-on conditions, the average intensity values of broadleaf trees are higher than those of most coniferous tree species. This difference is mainly due to differences in the tree structures. Broadleaf trees have larger single leaves, while coniferous trees have needles with a non-continuous leaf surface [36]. Recently, Shi et al. [31] verified that the intensity features make a more significant contribution to tree species classification in mixed temperate forests than the geometric features. Moreover, the bidirectional reflectance and geometry of the volumetric target surfaces significantly influence the intensity values recorded by a ALS system [6].

Recently introduced multispectral ALS technology is promising for improving forest mapping as it can provide a denser point cloud and higher spectral information. A few studies have focused on the potential of using multispectral ALS point clouds for classifying tree species [24,34,37–40]. The authors of Lindberg et al. [24] generated multispectral ALS data using three different instruments during different flights with a point density of 20 points/m^2 to characterize tree species. They employed visual interpretation to show that, if both spectral and geometric information from multi wavelength ALS data are used, the accuracy of tree species classification is better than that obtained when information from single wavelength ALS data are used. The authors of St-Onge and Budei [37] used the intensity-based features extracted from three spectral channels of a Titan multispectral ALS system [41] to classify broadleaf vs. Needle-leaf trees in a Canadian boreal forest and achieved a classification accuracy over 90%. The authors of Yu et al. [34] used the same sensor and achieved an overall tree species classification accuracy of 85.6% for three different tree species in southern Finland using intensity-based features. The authors of Hopkinson et al. [38] compared terrain and forest canopy attributes extracted from each wavelength of two multispectral ALS datasets (multisensor and single-sensor). They achieved an overall accuracy of 78% for the classification of land surface and vegetation (8 classes) by integrating spectral and structural information. The authors of Axelsson et al. [39] used the Optech Titan *X* System to investigate ten tree species in a boreal forest, and achieved a cross-validated accuracy of 76.5% using the height and intensity distribution of features from the tree segments. The authors of Budei et al. [40] classified 10 tree species using the Optech Titan system with an overall accuracy of 75%. So far, combinations of various features from multi spectral ALS point clouds have been mainly used to examine tree species classification in boreal forests.

Thus far, the detection of individual standing dead trees from ALS point clouds has been of minor research interest. The authors of Yao et al. [10] tackled for the first time the detection of dead trees with crowns using only the ALS intensity (single wavelength 1550 nm) and geometric features such as the crown shape and point height distribution. The study reports a classification accuracy

between 71% to 73%. Recently, Polewski [42] presented a method that uses single tree 3D segments in combination with multispectral aerial imagery. Based on features generated from the covariance matrix of the three image channels, a two-class classification (dead tree and non-dead tree) led to an accuracy of around 88%. The same author reported in the study Polewski et al. [43] on the detection of standing dead trees without crowns (snags). The method used free shape contexts to generate salient features suitable to describe the single snags in a sparse point cloud. After optimizing the highly dimensional feature space with a genetic algorithm, the new approach classified 285 objects with a classification accuracy of 84.2%.

In summary, the classification of coniferous and broadleaf trees with single wavelength ALS data (full waveform) is possible with high accuracy. The classification of multiple tree species has not yet reached a practical performance. Multispectral ALS has the potential to improve the tree species classification accuracy. The combined classification of tree species and standing dead trees with crowns has been of minor research interest. Moreover, due to the expected high dimensional feature space, techniques are mandatory to reduce the huge number of predictive variables to the most prominent ones.

Therefore, the main objectives of this study are to evaluate the accuracy of classification tree species and standing dead trees in a temperate forest located in southeastern Germany using (i) triple wavelength ALS (1550 nm, 1064 nm, and 532 nm), (ii) single wavelength ALS, and (iii) to identify the most important features based on a specific feature selection approach. Thereby, we demonstrate the current potentials and limitations of triple wavelength ALS without fusing optical imagery. We simulate a multispectral ALS sensor by compiling the data from three different ALS sensors which have been carried by two aircrafts on the same day.

2. Materials and Methods

2.1. Study Area

Located in the Bavarian Forest National Park (49°12′N, 12°58′E), our study area is a temperate and complex forest situated in the south-eastern part of Germany along the Czech Republic border covering an area of 24,250 ha. The forest is dominated by Norway spruce (*Picea abies*) with European beech (*Fagus sylvatica*) and Silver fir (*Abies alba*). Rare tree species are also present in the park, such as white birch (*Betula pendula*), sycamore maple (*Acer pseudoplatanus*), and common rowan (*Sorbus aucuparia*) [44]. Figure 1 shows in yellow the sample plots site (=transect) where the experiments were carried out on a color infrared orthophotograph of the Bavarian Forest National Park. A color infrared image of the study area that presents the difference between the trees (living vegetation) and standing dead spruce trees with crowns by color, is shown in Figure 2.

The multi wavelength ALS data were acquired on 18 August 2016 during leaf-on conditions using three RIEGL scanners LMS-680i, LMS-Q780, and VQ-880-G. The VQ-880-G and LMS-Q780 sensors were flown in an one aircraft, while the LMS-680i instrument was carried in a helicopter. In both campaigns, several parallel flight strips were flown with varying side lap between 50% and 100%. Therefore, the resulting point densities amounted in the center of the transect to 80 points/m^2 (LMS-680i) and 60 points/m^2 (LMS-Q780 and VQ-880-G). To this end, the combined 3D point cloud containing three different spectral channels (1550 nm (Ch1), 1064 nm (Ch2), and 532 nm (Ch3)) showed an average point density of around 200 points/m^2. Table 1 illustrates the details of the acquisition flight. Note that the weather conditions remained constant during data acquisition.
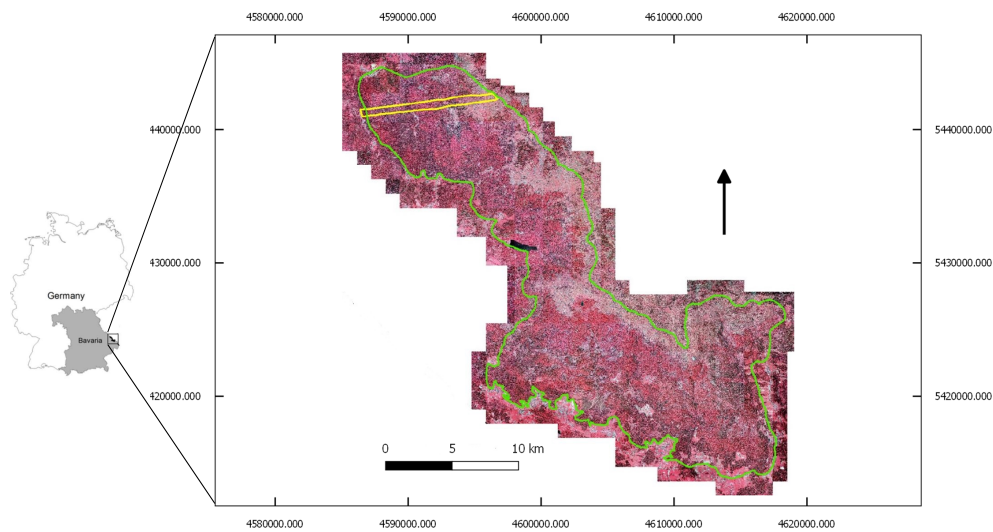
Figure 1. Color infrared orthophoto of the Bavarian Forest National Park (park border in green, transect border in yellow.
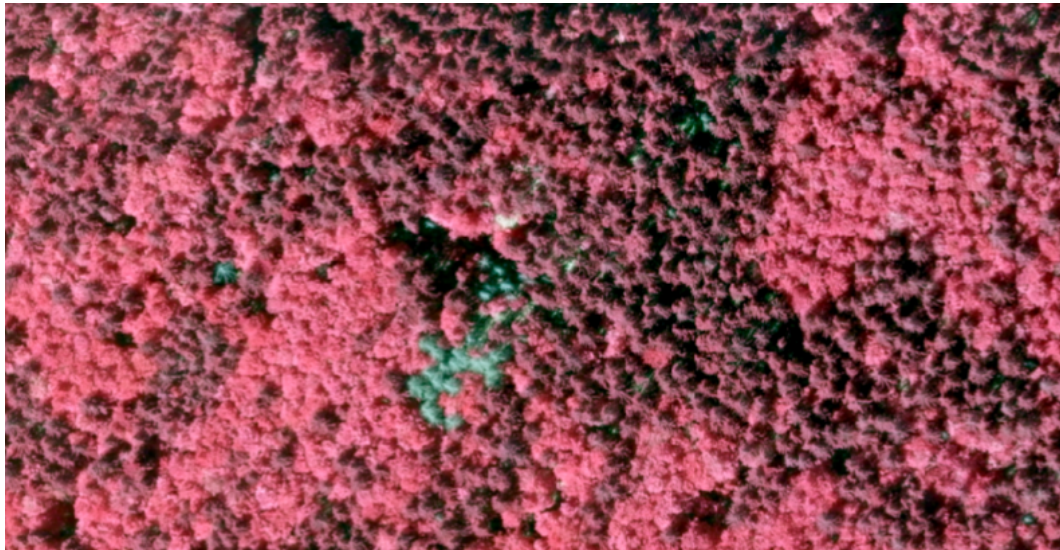

Figure 2. A forest scene in the color infrared aerial image with a ground pixel resolution of 20 cm. Standing dead spruce trees with crowns appear in gray, while the living trees are shown in red (broadleaf and coniferous trees).

Table 1. Flight campaigns of multi wavelength airborne laserscanning (ALS) data.

RIEGL Scanner Type	LMS-680i	LMS-Q780	VQ-880-G
Platform type	Helicopter: D-HALL/AS350	Aircraft: Tecnam P2006T	Aircraft: Tecnam P2006T
Spectral wavelength (nm)	1550 (=Ch1)	1064 (=Ch2)	532 (=Ch3)
Beam divergence (mrad)	0.5	0.25	0.2
Measurement rate (kHz)	266	400	550
Flight speed (kts)	50	90–100	90–100
Av. flight height (m)	400	550	550
Side lap (%)	50–80	50–100	50–100
Swath width (m)	350–450	500	500
# flight strips	5	30	10
Max. point density (pts/m^2)	80	62	60
Footprint size (mm)	150	137	110

The raw data of spectral Ch2 and Ch3 channels were horizontally and vertically shifted to reference data set Ch1. The alignment was visually checked on appropriate objects (e.g., standing dead trees without crowns; tree trunks; small flat areas). The reference channel had been geometrically calibrated in advance based on vertical and planimetric objects, such as enclosed building polygons and flat areas, respectively.

2.2. Reference Data

The ground truth data for the sample plots were acquired by field measurements and included 586 single trees that were measured from 20 circular sample plots with an area of 500 m^2 for each plot. The reference trees positions (all trees with diameter at breast height (DBH) > 7 cm) were measured from the center of the plot with a vertex for the distance and a compass for the angle using the Leica GS GPS system (see Bässler et al. [45]). Moreover, the DBH values for all the trees and heights of all trees were measured. The measurement campaign was conducted during the summer of 2017. Based on the available ground truth data and, due to the lack of reference data for some of the rare species, we selected three species for the study: Norway spruce (*Picea abies*), European beech (*Fagus sylvatica*), and Silver fir (*Abies alba*). Additionally, we added standing dead spruce trees with crowns to the list of trees to be classified. The distribution of three tree species and standing dead trees in the sample plots for the classification were as follows: Norway spruce (43%), European beech (30%), Silver fir (10%), and standing dead trees with crowns (16%).

2.3. Pre-Processing of ALS Data

A full waveform ALS system provides the reflected digitized (typically 15 cm) waveform at regular intervals and includes information regarding the reflected intensity and pulse width. An appropriate waveform decomposition is required to obtain these parameters. By using superimposed Gaussian functions, the 3D coordinates (x_v, y_v, z_v) of each reflecting object v hit by the laser pulse were obtained in combination with the intensity I_v and pulse width PW_v as physical properties [9]. Overall, this decomposition generated a point cloud for the forest area represented by the vector $P_n(x_n, y_n, z_n, I_n, PW_n), n = 1, ..., N$ (N is the total number of points in the point cloud) [46]. The intensity I_n can be interpreted as the pulse energy and is equal to the area of a single Gaussian function. This value is dependent on the traveling distance r_n (in one direction) and must be normalized with respect to a reference distance r_{ref} according to Equation (1) [47–49].

$$I_n^{corr} = I_n \left(\frac{r_n}{r_{ref}} \right)^m \tag{1}$$

The m parameter could be estimated from ALS data acquired in a special calibration flight [50]. In this study, we used the theoretical value $m = 2$. Commercial software packages, such as RIEGL RiAnalyze© can provide the amplitude a_n and/or the reflectance ρ_n. The latter is also often referred to as an intensity and refers to the fraction of incident optical power reflected by a target at a certain wavelength. The amplitude a_n is defined by each hardware manufacturer (see [51]). In RIEGL scanners, the amplitude a_n is defined as a linear measure for the pulse energy. Extending Equation (1) and following the suggestions of Höfle and Pfeifer [52], the range-independent reflectance ρ_n can be approximately converted from the amplitude a_n by the following formula ([51], personal communication [53]):

$$\rho_n = \rho_{ref} \left(\frac{a_n}{a_{ref}} \right) \left(\frac{r_n}{r_{ref}} \right)^2 \eta_{atm} \tag{2}$$

Equation (2), refers to an instrument whose emitted laser beam perpendicularly hits a target area of 100% reflectance (=ρ_{ref}) and measures at a distance of r_{ref} and an amplitude of a_{ref}. The term $\eta_{atm} = exp(0.0000978 \times 2 \times (r_n - r_{ref}))$ describes the atmospheric attenuation (the loss of energy through the scattering and absorption of photons from the laser beam (1550 nm) in the atmosphere)

is assumed to cover a visual range of 23 km [50]. Note that the reflectance ρ_n values are also affected by the angle of incidence (the angle between the emitted laser beam and the target surface normal) [54]. However, we assumed that the incident angle was unknown and thus this effect has been neglected.

The ρ_n reflectance values were calculated from the LMS-Q780 and VQ-880-G scanner data using the RIEGL RiAnalyze© software. By default, these values were corrected considering the traveling distance r_n. However, for the LMS-680i scanner data, the RIEGL RiAnalyze© software could only provide the amplitude values. Therefore, the amplitude a_n for each reflection point P_n was approximately converted to the reflectance ρ_n using Equation (2). In order to obtain reasonable values for the parameters a_{ref} and r_{ref}, a sample of the amplitude $a_n = 165$ in Equation (2) was taken as the mean value from four small concrete areas (4 × 4 m) located on an airfield (see Figure 3). Each area was selected in the nadir of four different ALS strips that had approximately the same flying height $r_n = 420$ m. Finally, we defined that the $a_{ref} = 250$ and $r_{ref} = 600$ m parameters should refer to a reflectance of $\rho_{ref} = 100\%$. This in turn means that according to Equation (2), the amplitude $a_n = 165$ is equivalent to a reflectance of $\rho_{ref} = 31\%$. After the pre-processing step, a final visual inspection of the overlaying ALS strip areas showed there were no tiling effects.

Figure 3. The four selected small target areas located on the airfield that were used for the amplitude calibration a_{ref}.

2.4. Outline of Method

In our approach, we classify tree species in combination with standing dead trees using triple wavelength ALS. A huge feature set is generated comprising geometric features, waveform features, and BoW-features. We reduce the highly dimensional feature space to the most salient features which

contribute at most to the classification. To this end, we avoid the typical overfitting effect and optimize the classification result. The method used in our study is as follows. We segment the ALS point cloud into 3D segments representing single trees using the normalised cut algorithm [9,55] (see Section 2.5) after performing the above mentioned pre-processing step (see Section 2.3). The procedure recursively partitions the ALS point cloud until the level representing single trees is achieved. The aim of the segmentation step is to divide points in the cloud with similar attributes into homogeneous 3D segments [11]. Because the classification step is our main focus, we regard this phase as an external procedure. We then extract features in the 3D point cloud spaces. The forward stepwise selection method is applied to extract the most significant features for classification. The final selected features are used in a multinomial logistic regression for classification. A schematic overview of the entire processing procedure is presented in Figure 4. This method is explained in detail by the following subsections.
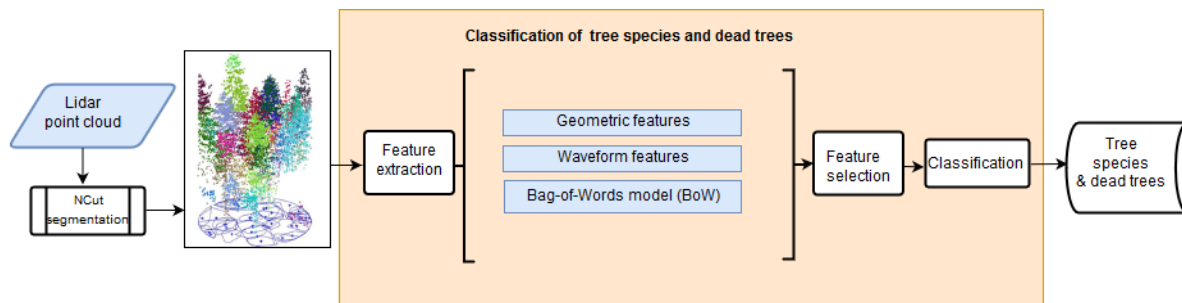

Figure 4. Overview of the approach for classification of tree species and standing dead trees.

2.5. Normalized Cut Segmentation

The normalised cut algorithm [55] is a top-down method for segmenting objects over a discrete graph structure $G = (V, E)$. The vertices V represent the individual objects and the edges E correspond to the neighborhood topology. The input 3D ALS point clouds is split up into disjointed segments where the normalised cut criterion is minimized. A recursive bisection of the graph's vertices V into disjointed segments A and B maximizes the intra-segment similarity of the objects and minimizes their inter-segment dissimilarity (see Reitberger et al. [9]). The normalised cut to be minimized is:

$$NCut(A, B) = \frac{Cut(A, B)}{Assoc(A, V)} + \frac{Cut(A, B)}{Assoc(B, V)} \tag{3}$$

where $Cut(A, B) = \sum_{i \in A, j \in B} w_{ij}$ is defined as the total sum of the weights between the A and B segments, while $Assoc(A, V) = \sum_{i \in A, j \in V} w_{ij}$ is the sum of the weights of all the edges ending in segment A.

The recursive splitting of graph G into disjoint subgraphs must normally be terminated by a threshold value $NCut_{max}$. In our study, we used the parameter recommended by Reitberger et al. [9]. The presented tree segmentation algorithm was implemented in C++ and in MATLAB R2017a.

2.6. Feature Extraction

From the multi wavelength ALS point cloud, we calculate for each tree segment the following features. The features can be divided into three main groups. The different feature sets are presented in the Table 2.

(i)　*Geometric features:*

This feature set includes: Percentiles of the point height distribution in a tree segment S_h (referred to as the height dependent variables) at 10% intervals from ($h = 10\%$–100%); the two axis lengths of a paraboloid fitted to a tree crown summarized in feature S_g; the percentage of points per height layer of a tree S_d (referred to as the density dependent variables) at

10 intervals ($d = 1–10$); and the point count ratios by reflection type (single S_n_single, first S_n_first, middle S_n_middle) (see Reitberger et al. [9]). Additionally, the crown polygon area A_p and the minimum enclosing circle of the projected polygon ec_min are extracted.

(ii) *Waveform features:*

The waveform decomposition described in Section 2.3 provides for each laser point reflectance and the pulse width. Based on these laser points attributes, the mean reflectance of single S_I_{single} and first S_I_{first} reflections, histograms from the intensity $S\_I_j^H$, and the pulse width $S\_PW_j^H$ ($j = 1, ..., 10$) are extracted for each spectral channel [9]. Furthermore, triple wavelength data offer an unique opportunity to use the band variances and covariances that were impossible for approaches only using single wavelength ALS data. Therefore, six independent variance and features S_cov_uk (u = 1, ..., 3; k = u, ..., 3) from the covariance matrix are obtained that corresponded to the upper triangle of the band covariance matrix.

(iii) *Bag-of-Words model:*

This model is an established method for representing high-level characteristics used for classification purposes [56]. The main aim of the bag-of-words (BoW) model is to approximate the feature vectors using a vector quantization algorithm with a set of prototypes. As proposed by Weinmann et al. [57], our BoW model $S\_BoW.[C]_b^{nd}$ compiles eight geometric features generated from the local covariance matrix. In the BoW model, C stands for the linearity, planarity, sphericity, omnivariance, anisotropy, eigenentropy, sum of eigenvalues, and change of curvature that are commonly used in ALS data processing [58,59]. The frequency histograms for each feature are then constructed using $b = 2–14$ bins. This is conducted for 16 spherical point neighborhood sizes with radii of $nd = 0.2–1.6$ m (0.1). The range of the neighborhood radii nd are defined using the tree crown diameters from the reference data.

Table 2. Description of the generated features for each tree segment.

Features	Definition	Features	Definition
S_g	Axis lengths of a fitted paraboloid	A_p	Crown polygon area
S_h	Percentiles of point height distribution	ec_min	Minimum enclosing circle
S_d	Percentage of points per height layer	S_I_{single}	Mean intensity (single reflections)
S_n_single	Point counts ratio by single reflectio	S_I_{first}	Mean intensity (first reflections)
S_n_first	Point counts ratio by first reflection	$S\_I_j^H$	Intensity histograms
S_n_middle	Point counts ratio by middle reflection	$S\_PW_j^H$	Pulse width histograms
S_cov_uk	Elements of covariance matrix	$S\_BoW.[C]_b^{nd}$	BoW model features

2.7. Feature Selection

In this study, we use a large number of features to classify tree species and dead trees. All the features presented in Section 2.6 sum up to a long list with around 3600 elements. However, only a small number of the features are meaningful and suited for the classification. The high number of features raises the methodical problem that the large hyper-dimensional feature space faces a spare number of samples [6]. Therefore, a feature selection method needs to be applied that identifies and removes the irrelevant and redundant attributes from the data that do not contribute to the accuracy of the classification model. The authors of Weinmann et al. [57], Guyon and Elisseeff [60], Liu et al. [61] report on feature selection techniques for finding the most robust subsets of the relevant features to optimize the classification accuracy and to improve the computational efficiency. They are subdivided into wrapper methods, filter methods, and embedded methods. Wrapper-based methods apply either a sequential backward elimination or forward selection in combination with a certain classifier to rank the feature set. The process works iteratively until the best learning performance is obtained or the certain number of selected features is achieved. Due to the high search space, these methods are

computationally fairly complex. Filter-based methods rank in a first step the features according to a score function and filter lowly ranked features out in the second step [62]. Finally, embedded methods interact with the learning procedure but do not evaluate the feature set iteratively, thereby gaining a significantly higher computational performance. A typical example is the feature importance of the Random Forest classifier which is based on the permutation importance of a feature variable.

Here, we propose a wrapper method referred to as stepwise forward selection that begins with a small feature set randomly selected from the full set and then proceeds in an iterative fashion, selecting one additional feature in each step. A single iteration inspects every available feature by adding it to the active feature set and obtaining an estimate of the classification error rate on the augmented data through cross-validation [63]. The feature that enter the active set with the lowest error rate are incorporated into the result set, and the iterations proceeds. The process is terminated when the inclusion of additional features ceases to decrease the classification error rate. In order to avoid any randomness effects, the selection procedure is repeated five times for each scenario (see Section 4.2) and the results are similar for all the iterations. Further, this wrapper method uses the multinomial logistic regression of the Section 2.8 as the predictive model. The final result is a list of features organized in ascending order according to the error rate generated in the classification. In this study, the first members on the list are interpreted as the most important features with the highest contribution to the classification result.

2.8. Classification of Tree Species and Dead Trees

To classify the trees based on the features extracted and discussed in Section 2.6, we apply a multinomial logistic regression. Logistic regression models represent the probability distribution of the class label y as follows:

$$p(y_i = k|x_i; \Theta) = exp(\theta_k^T x_i) / [1 + \sum_l \exp(\theta_l^T x_i)] \tag{4}$$

where $x_i \in X$, $i = 1, ..., M$ denotes the M feature vectors of the training examples and y_i represents their corresponding labels. Training the model involves minimizing the penalized joint negative log-likelihood of the training examples with respect to the $\beta > 0$ term, as shown in Equation (5):

$$\min_{\Theta} \sum_{i=1}^{M} -log\, p(y_i|x_i; \Theta) + \beta||\Theta||_1 \tag{5}$$

where $||\Theta||_1$ is the regularization term that increases the sparsity of the coefficient vector Θ. After feature selection, a few features remain and the sparsity effect is not very significant. Therefore, a penalized logistic regression can be applied using the following model:

$$\frac{1}{M} \sum_{i=1}^{M} L(\beta, X, y) - \lambda[\frac{(1-\alpha)||\beta||_2^2}{2} + \alpha||\beta||_1] \tag{6}$$

where α is a balance coefficient between penalties $L_1(lasso)$ and $L_2(ridge)$, and the optimal value for λ can be selected using cross-validation. Equation (6) is defined as the 'elastic net penalty', that is a regularized-generalized model and is implemented in the *glmnet* function as an *R* source package.

3. Experimental Setup

3.1. Control Parameters for Normalized Cut Segmentation

The normalised cut segmentation is based on control parameters whose values were optimized experimentally. The main control parameters are summarized in the Table 3.

Table 3. Main control parameters for the single tree segmentation using the normalised cut method.

Parameters	Symbols	Values
Normalized cut threshold	$NCut_{max}$	0.16
Minimum number of points in a segment	Min_{num}	10

3.2. Tree Selection for the Classification

The segmented single trees were subdivided into three canopy layers with respect to the top tree height h_{top} in the plot that was defined as the average height of the 100 tallest trees per ha [64,65]. We focused on the upper canopy layer that contained trees with heights that were at least 80% of the h_{top} and could be identified in the reference data using a matching strategy. The reference data found in the upper canopy layer were linked to the correctly segmented single trees by conducting the following procedure [50]. The single tree positions from the normalized cut segmentation were matched with the reference trees if, (a) the distance from the center of a segment to the reference tree was smaller than 60% of the mean tree distance of the plot, and (b) the height difference between h_{top} and the height of the reference tree was smaller than 20% of h_{top}. Note that the threshold values in (a) is an approximation of the average crown radius (=50%) of a plot plus 10% [65]. In the second criterion (b), we consider the accuracy of tree height estimation from ALS (=10%), which is dependent on the parameter h_{top} [66]. Due to the low accuracy of field measurement, we added 10% to this value. If a reference tree was assigned to more than one tree position detected by the normalized cut segmentation, the tree position with the minimum distance to the reference tree was selected, otherwise, it was removed from further analysis. The matched tree segments in the upper canopy layer were extracted and assigned to the corresponding reference trees for all sample plots. The final number of the matched tree segments with the reference data in the upper canopy layer that were used for further classification analyses are summarized in Table 4.

Table 4. The number of matched trees in the upper canopy layer with the overall percentage of tree species and dead trees in the plots.

	Tree Species and Dead Trees (%)	Number of Trees
Norway spruce	35	161
European beech	30	138
Silver fir	16	67
Dead tree	19	82

3.3. Classifier Training

The reference data were unbalanced and dominated by beech and spruce trees. Therefore, the reference data were balanced for the classifier and the experiment was run 20 times by randomly selecting a subset of the dominant class each time so that the reference data contained no more than 32% of any class. This resulted in the significant reduction in the number of matched trees mentioned in the Table 4 for the classification.

The classification accuracy was estimated and evaluated based on the results of the matching step between the reference data and the 3D segmented trees in the upper canopy layer (see Section 3.2) (with a proportion of 60%–40% for training and testing the classifier) with respect to the distribution of tree species and dead trees. For the training, we used a 15-fold cross-validation to obtain the overall classification accuracy, correctness, and completeness as a compromise between the computational efficiency and reducing the effects of randomness.

Finally, we defined two test cases for the evaluation of the classification. In scenarios 1 to 4, we highlight the potential of the multi wavelength lidar. For comparison, we used in scenarios 5 to 7 individual wavelengths Ch1, Ch2, and Ch3.

4. Results

4.1. Main Outcomes

In general, our experiments demonstrate that the multi wavelength ALS outperforms single wavelength ALS by at least 8%. The best classification result of 82% is achieved using the full feature set (geometric features, intensity-based features, BoW-features). Norway spruce and European beech are classified with fairly high completeness of 93% and 87%, respectively (Table 6). The completeness for silver firs and standing dead trees drops down to 59% and 73% (Table 6). The embedded feature selection provides a clear overview on the most prominent features as well. The wavelenghth 1064 nm (=Ch2) turns out as the best discriminative channel. In addition to the intensity-based features, which appear as the most salient features, the covariance between channels Ch1 and Ch2, the geometric shape of the tree crown, and the sphericity of the BoW model are ranked best in the list of the top five features. The feature selection optimizes the classification results by 6% (Table 5).

Table 5. Results for classification of tree species and dead trees using the test data set. The best result of each test case is highlighted in bold.

Scenario Number	Feature Set	Overall Accuracy (%)	Kappa
1	Multi wavelength ALS : Geometric	68	0.59
2	Multi wavelength ALS : Geometric + Waveform	79	0.69
3	Multi wavelength ALS : BoW	54	0.46
4	Multi wavelength ALS : Geometric + Waveform + BoW	**82**	0.75
5	Single wavelength ALS (Ch1) : Geometric + Waveform + BoW	69	0.61
6	Single wavelength ALS (Ch2) : Geometric + Waveform + BoW	**74**	0.65
7	Single wavelength ALS (Ch3) : Geometric + Waveform + BoW	65	0.55

4.2. Classification of Tree Species and Dead Trees

The quantitative evaluation of the classification using the test data set is presented in Table 5. The accuracy values refer to the average overall classification accuracy, and the best result for each test case is highlighted in bold.

First, we start with an analysis of the results generated from the multi wavelength ALS data (scenarios 1 to 4). The results of scenario 2 demonstrates an improvement of the classification accuracy by 11% by adding the waveform features to the geometric features (scenario 1) improves. Among the multi wavelength ALS features, the BoW model features alone provide the lowest accuracy of 54% (scenario 3). The scenario 4 turns out as the best showing an classification result of 82%. Analyzing the results collected from single wavelength ALS (scenarios 5 to 7) indicate a worse overall accuracy when compared to the best multi wavelength ALS scenario 4. As expected, the channels Ch1 and Ch2 perform better than the channel Ch3. In summary, we conclude an improvement of the classification via the multi wavelength approach by at least 8% when compared to a single wavelength scenario.

We now focus on the detailed mapping of the three tree species Norway spruce, European beech, silver fir, and standing dead trees. The confusion matrices showing the correctness and completeness are presented in Tables 6 and 7 based on the best scenarios 4 and 6. For scenario 4, we achieve for Norway spruce high values of 95% and 93% for correctness and completeness. The correctness and completeness for European beech are also fairly good with values of 83% and 87%, respectively. Silver fir trees, proportionally less represented in the study area than spruces, are only classified with 59% completeness and correctness. Finally, standing dead trees are detected with 73% completeness and 76% correctness. The average rate of false negatives for the silver fir trees and standing dead trees in scenario 4 are 41% and 27%, respectively.

Table 6. Confusion matrix of scenario 4 using features extracted from multi wavelength ALS point cloud from the test data set.

		Reference				
		Spruce	**Beech**	**Fir**	**Dead Tree**	**Correctness (%)**
	Spruce	76	0	0	4	95
	Beech	0	69	11	3	83
Predicted	Fir	1	7	22	7	59
	Dead tree	5	3	4	39	76
	Completeness (%)	93	87	59	73	
Overall accuracy: 82.1%	Kappa value: 0.75					

Table 7. Confusion matrix of scenario 6 using features extracted from single wavelength (1064 nm) ALS point cloud from test data set.

		Reference				
		Spruce	**Beech**	**Fir**	**Dead Tree**	**Correctness (%)**
	Spruce	70	1	3	6	87
	Beech	3	64	12	4	77
Predicted	Fir	1	9	19	8	51
	Dead tree	6	5	6	34	67
	Completeness (%)	87	81	47	65	
Overall accuracy: 74.5%	Kappa value: 0.65					

4.3. Feature Relevance Assessment

In Table 8, we focus on analyzing the important features listed for the scenarios 4 and 6 that represent the best classification results. Note that these scenarios are representative for their different dataset combinations (Multi wavelength ALS, single wavelength ALS). After the feature selection step, from 3600 features identified, 30 features were selected as the most important features for the classification. Following Section 2.7, the error rate, which each additional feature generates in the classification, represents the individual feature importance. For both scenarios, we present the top five features.

Table 8. List of the top five features for two classification scenarios 4 and 6 (each scenario is defined in Table 5). Each scenario has 30 features selected via the feature selection step. The below mentioned feature abbreviations are explained in Section 2.6.

Scenarios	Top Five Features
Scenario 4: Multi wavelength	$S\_I_6^H(Ch2), S\_I_{single}(Ch2), S\_cov\_21, S\_g, S\_BoW.[sphericity]_7^{0.8}$
Scenario 6: Single wavelength (Ch2)	$S\_I_6^H, S\_I_8^H, S\_I_{single}, S\_g, S\_BoW.[sphericity]_7^{0.8}$

In the multi wavelength case (=scenario 4), channel Ch2 is the most important wavelength. In detail, intensity histogram feature $S\_I_6^H$ and mean intensity for single reflections S_I_{single} generated from this channel exhibit the highest contribution to the classification. These top two features are followed by radiometric feature S_cov_21 and geometric feature S_g. The first represents the covariance between channels Ch1 and Ch2, whereas the latter describes the outer geometric shape of the tree crown. Interestingly, the BoW-based feature sphericity generated in neighborhood 0.8 m ranks as the next important feature. Noteworthy, the mean intensity of the single reflections S_I_{single} for Ch2 performs better than those from channels Ch1 and Ch3.

Finally, we present in Figure 5 the learning curves that the feature selection provided using the error rates of the classification. In the case of scenario 4, the error rate remains constant after around seven feature. The error rate of scenario 6 is significantly improved by the first 10 features and continues almost constant by adding the next best features.
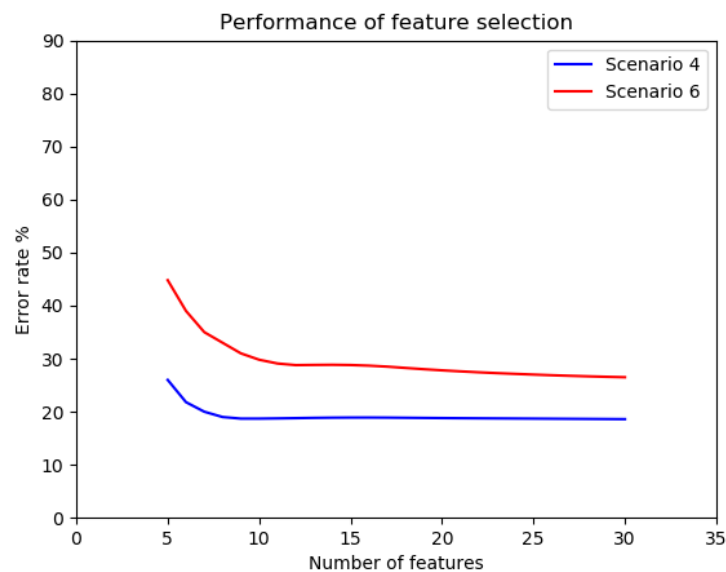
Figure 5. Performance of feature selection for scenarios 4 and 6.

5. Discussion

5.1. Main Goals of the Study

The main goals of this study were to classify three tree species as well as standing dead trees using triple wavelength ALS data. The fairly large dimension of the highly correlated feature space was successfully reduced by a special feature selection method that provided the top five features of the tree species classification.

5.2. Classification of Tree Species and Dead Trees

In Tables 6 and 7, we notice a fairly high confusion between firs and beeches. The results for the firs might be due to the spectral similarities between firs and beeches. Structural crown correlation between dead trees and living spruces might be the reason for dead trees whose crowns still have in the early infestation stage a similar geometrical structure. In general, the completeness and correctness for the three tree species and standing dead trees are on average better by around 7% for the multi wavelength scenario 4. Finally, the kappa value of 0.75 also indicates a higher accuracy for the multi wavelength scenario 4.

5.3. Feature Assessment

If we compare the findings of the feature relevance assessment for the multi wavelength case (=scenario 4) and the single wavelength case (=scenario 6 using Ch2), we notice an almost equivalent feature composition. Radiometric feature S_cov_21 in scenario 4 is effectively replaced by feature $S\_I_8^H$, which is an element in the intensity histogram. Finally, the analysis of the long feature list summarized in Table 2 shows that features describing the penetration of the laser points in the tree segments (S_h, S_d), the type of reflections (S_n_single, S_n_first, S_n_middle), and the pulse width distribution ($S\_PW_j^H$) are irrelevant for the classification.

The important feature selection step was applied to extract robust feature subsets that optimize the classification accuracy. It is important to note that the accuracy is on average reduced by nearly 6% when all the generated features (without feature selection step) are used. Recent feature selection approaches primarily applied wrapper methods, filter methods, or embedded methods [60]. In general, the wrapper methods (such as forward feature selection) performed better than the other mentioned

techniques in terms of classification accuracy [67]. However, our feature selection method requires as a wrapper method high computational effort because the classifier model needs retraining as part of the cross-validation used for accuracy assessment purposes.

5.4. Related Work

5.4.1. Tree Species Classification and Feature Assessment

The results of this study should be compared to those of earlier studies [31,68] conducted on the Bavarian Forest National Park that classified the six tree species (Spruce, Beech, Fir, Maple, Rowan, and Birch). The accuracy of our study was 20% higher under leaf-on conditions than the findings of Shi et al. [31], who used only features extracted from single wavelength ALS data for multiple tree species classification.

Our results demonstrated that the multi wavelength ALS features increased the accuracy of the classification of the three tree species and standing dead trees by approximately 8 to 17% compared to the single wavelength approach. Our study verified the findings of Yu et al. [34] (overall accuracy: 85.9%, kappa: 0.75) and Axelsson et al. [39] (overall accuracy: 75.5%), who stated that intensity-based features extracted from multispectral ALS data, specifically those from channel Ch2, were the most important for classifying tree species. We showed that in both scenarios 4 and 6 (see Table 8) the intensity histogram-based feature $S_6^{I^H}(Ch2)$ played a more important role in the species identification than any other geometric features. As far as the classification using single wavelengths is concerned, our findings confirmed the results of Shi et al. [31] who concluded the importance of the intensity-based features (within the top seven features) when employing an embedded feature ranking strategy for the Random Forest classifier. Moreover, Budei et al. [40] showed that the intensity-based features of infrared and green channels of a three-wavelength ALS system using the Random Forest classifier could improve the detailed species identification accuracy in a temperate forest compared to single channel ALS systems. Concluding, our analysis was based on sample plots with relatively limited variability between the tree species (see Section 2.1). Therefore, the applicability of our identified important features needs to be tested further in other study sites.

5.4.2. Dead Tree Detection

Finally, we compare our dead tree classification with other approaches obtained from the same forest site. The findings of the study Yao et al. [30] are similar to ours with a classification accuracy between 71% and 73%. The method of Polewski [42] classified standing dead trees with an accuracy of around 88%. However, in this study, the reference data were generated from the tree point cloud by visual inspection using orthophotos superimposed on the tree crown polygons. Very likely, the combination of ALS data and multispectral optical imagery in a two-class classification is more efficient for the detection of standing dead trees with crowns. The authors of Polewski [42] successfully detected the crowns of infested spruces solely using features from the covariance matrix calculated in the projected tree crown polygon. Apparently, the crowns of standing dead trees have a characteristic spectral appearance in multispectral images compared to the other trees.

6. Conclusions

In summary, our experiment illustrated that multi wavelength ALS point clouds improves the characterization of tree species in approaches that work at the single tree level. However, both firs and standing dead trees with crowns could only be classified with reduced completeness of 59% and 73%, respectively. The use of the feature selection step showed that mainly the intensity-related features from spectral channel Ch2 (1064 nm) notably improved the classification rate, achieving an overall accuracy of 82%.

This study performed one of the first experiments examining the applicability of multi wavelength ALS data for classification of tree species and dead trees on a temperate forest. Instead of using one

single instrument, we simulated multispectral ALS data by combining sensor data acquired from two different instruments on the same day with stable weather conditions. The flying height and the atmospheric attenuation was the same for the two sensors and almost the same for the third one. Of course, this system configuration has different scan angles and different sensors as well. All in all, we believe that this setup provided radiometric data comparable to those captured by a single instrument consisting of three non-collinear ALS units, e.g., the Titan sensor from Optech. Clearly, the use of a single multispectral ALS instrument has advantages over multiple sensors. Besides the data processing, the data from a single instrument refer to the same flying height and are consistently calibrated. However, both instrument approaches have the drawback that non-collinear laser beams are used meaning that the backscattered pulses do not necessarily result from the same part of the object.

Furthermore, as expected, the feature selection step considerably reduced the high dimensional feature space used to optimize the classification accuracy. Based on the prominent features list, the feature selection procedure was able to identify the most discriminative features. Interestingly, the classification deteriorated by about 6% if all the extracted features without the feature selection step had been used.

From a practical point of view, the accuracy level is not yet optimal and should be at least 90%. Apparently, the radiometric information from the non-collinear multi wavelength ALS data provided at three distinct wavelengths limited the classification of the four forest objects (three tree species and standing dead trees with crowns) to this accuracy level in a temperate forest. Further research needs to be conducted to investigate the relationships between the structural tree crown characteristics and the multi wavelength ALS point cloud features to determine their impact on the classification. Furthermore, a collinear multispectral ALS system, whose emitted laser beams strike the same target simultaneously, or the fusion of ALS with hyperspectral imagery may improve the detailed classification of tree species and dead trees.

References

1. Baldeck, C.A.; Asner, G.P.; Martin, R.E.; Anderson, C.B.; Knapp, D.E.; Kellner, J.R.; Wright, S.J. Operational tree species mapping in a diverse tropical forest with airborne imaging spectroscopy. *PLoS ONE* **2015**, *10*, e0118403. [CrossRef]

2. Turner, W.; Spector, S.; Gardiner, N.; Fladeland, M.; Sterling, E.; Steininger, M. Remote sensing for biodiversity science and conservation. *Trends Ecol. Evol.* **2003**, *18*, 306–314. [CrossRef]

3. Korpela, I.; Koskinen, M.; Vasander, H.; Holopainen, M.; Minkkinen, K. Airborne small-footprint discrete-return LiDAR data in the assessment of boreal mire surface patterns, vegetation, and habitats. *Forest Ecol. Manag.* **2009**, *258*, 1549–1566. [CrossRef]

4. Treitz, P.; Howarth, P. High spatial resolution remote sensing data for forest ecosystem classification: An examination of spatial scale. *Remote Sens. Environ.* **2000**, *72*, 268–289. [CrossRef]

5. Heinzel, J.; Koch, B. Investigating multiple data sources for tree species classification in temperate forest and use for single tree delineation. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *18*, 101–110. [CrossRef]

6. Fassnacht, F.E.; Latifi, H.; Stereńczak, K.; Modzelewska, A.; Lefsky, M.; Waser, L.T.; Straub, C.; Ghosh, A. Review of studies on tree species classification from remotely sensed data. *Remote Sens. Environ.* **2016**, *186*, 64–87. [CrossRef]

7.   Skidmore, A.K.; Pettorelli, N.; Coops, N.C.; Geller, G.N.; Hansen, M.; Lucas, R.; Mücher, C.A.; O'Connor, B.; Paganini, M.; Pereira, H.M.; et al. Environmental science: Agree on biodiversity metrics to track from space. *Nature* **2015**, *523*, 403–405. [CrossRef]

8.   Wang, Y.; Weinacker, H.; Koch, B.; Sterenczak, K. Lidar point cloud based fully automatic 3D single tree modelling in forest and evaluations of the procedure. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2008**, *37*, 45–51.

9.   Reitberger, J.; Schnörr, C.; Krzystek, P.; Stilla, U. 3D segmentation of single trees exploiting full waveform LIDAR data. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 561–574. [CrossRef]

10.   Yao, W.; Krzystek, P.; Heurich, M. Enhanced detection of 3D individual trees in forested areas using airborne full-waveform LiDAR data by combining normalized cuts with spatial density clustering. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2013**, *1*, 349–354. [CrossRef]

11.   Amiri, N.; Polewski, P.; Heurich, M.; Krzystek, P.; Skidmore, A.K. Adaptive stopping criterion for top-down segmentation of ALS point clouds in temperate coniferous forests. *ISPRS J. Photogramm. Remote Sens.* **2018**, *141*, 265–274. [CrossRef]

12.   Strîmbu, V.F.; Strîmbu, B.M. A graph-based segmentation algorithm for tree crown extraction using airborne LiDAR data. *ISPRS J. Photogramm. Remote Sens.* **2015**, *104*, 30–43. [CrossRef]

13.   Erikson, M. Species classification of individually segmented tree crowns in high-resolution aerial images using radiometric and morphologic image measures. *Remote Sens. Environ.* **2004**, *91*, 469–477. [CrossRef]

14.   Adão, T.; Hruška, J.; Pádua, L.; Bessa, J.; Peres, E.; Morais, R.; Sousa, J. Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry. *Remote Sens.* **2017**, *9*, 1110. [CrossRef]

15.   Hirschmüller, H. Stereo processing by semiglobal matching and mutual information. *IEEE Trans. Pattern Anal. Mach. Intell.* **2008**, *30*, 328–341. [CrossRef]

16.   White, J.C.; Coops, N.C.; Wulder, M.A.; Vastaranta, M.; Hilker, T.; Tompalski, P. Remote sensing technologies for enhancing forest inventories: A review. *Can. J. Remote Sens.* **2016**, *42*, 619–641. [CrossRef]

17.   Ullah, S.; Dees, M.; Datta, P.; Adler, P.; Koch, B. Comparing airborne laser scanning, and image-based point clouds by semi-global matching and enhanced automatic terrain extraction to estimate forest timber volume. *Forests* **2017**, *8*, 215. [CrossRef]

18.   Nevalainen, O.; Honkavaara, E.; Tuominen, S.; Viljanen, N.; Hakala, T.; Yu, X.; Hyyppä, J.; Saari, H.; Pölönen, I.; Imai, N.N.; et al. Individual tree detection and classification with UAV-based photogrammetric point clouds and hyperspectral imaging. *Remote Sens.* **2017**, *9*, 185. [CrossRef]

19.   Maschler, J.; Atzberger, C.; Immitzer, M. Individual tree crown segmentation and classification of 13 tree species using airborne hyperspectral data. *Remote Sens.* **2018**, *10*, 1218. [CrossRef]

20.   Grabska, E.; Hostert, P.; Pflugmacher, D.; Ostapowicz, K. Forest stand species mapping using the Sentinel-2 time series. *Remote Sens.* **2019**, *11*, 1197. [CrossRef]

21.   Brandtberg, T. Classifying individual tree species under leaf-off and leaf-on conditions using airborne lidar. *ISPRS J. Photogramm. Remote Sens.* **2007**, *61*, 325–340. [CrossRef]

22.   Holmgren, J.; Persson, Å.; Söderman, U. Species identification of individual trees by combining high resolution LiDAR data with multi-spectral images. *Int. J. Remote Sens.* **2008**, *29*, 1537–1552. [CrossRef]

23.   Ørka, H.O.; Næsset, E.; Bollandsås, O.M. Classifying species of individual trees by intensity and structure features derived from airborne laser scanner data. *Remote Sens. Environ.* **2009**, *113*, 1163–1174.

24.   Lindberg, E.; Briese, C.; Doneus, M.; Hollaus, M.; Schroiff, A.; Pfeifer, N. Multi-wavelength airborne laser scanning for characterization of tree species. In Proceedings of the SilviLaser, La Grande Motte, France, 28–30 September 2015; pp. 271–273.

25.   Holmgren, J.; Persson, Å. Identifying species of individual trees using airborne laser scanner. *Remote Sens. Environ.* **2004**, *90*, 415–423. [CrossRef]

26.   Shang, X.; Chazette, P. Interest of a full-waveform flown UV lidar to derive forest vertical structures and aboveground carbon. *Forests* **2014**, *5*, 1454–1480. [CrossRef]

27.   Hovi, A.; Korhonen, L.; Vauhkonen, J.; Korpela, I. LiDAR waveform features for tree species classification and their sensitivity to tree-and acquisition related parameters. *Remote Sens. Environ.* **2016**, *173*, 224–237. [CrossRef]

28. Höfle, B.; Hollaus, M.; Lehner, H.; Pfeifer, N.; Wagner, W. Area-based parameterization of forest structure using full-waveform airborne laser scanning data. SilviLaser 2008. In Proceedings of the 8th International Conference on LiDAR Applications in Forest Assessment and Inventory, Edinburgh, UK, 17–19 September 2008; pp. 17–19.

29. Heinzel, J.; Koch, B. Exploring full-waveform LiDAR parameters for tree species classification. *Int. J. Appl. Earth Obs. Geoinf.* **2011**, *13*, 152–160. [CrossRef]

30. Yao, W.; Krzystek, P.; Heurich, M. Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. *Remote Sens. Environ.* **2012**, *123*, 368–380. [CrossRef]

31. Shi, Y.; Wang, T.; Skidmore, A.K.; Heurich, M. Important LiDAR metrics for discriminating forest tree species in Central Europe. *ISPRS J. Photogramm. Remote Sens.* **2018**, *137*, 163–174. [CrossRef]

32. Korpela, I.; Ørka, H.O.; Maltamo, M.; Tokola, T.; Hyyppä, J. Tree species classification using airborne LiDAR–effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type. *Silva Fennica* **2010**, *44*, 319–339. [CrossRef]

33. Leckie, D.G.; Gougeon, F.A.; Walsworth, N.; Paradine, D. Stand delineation and composition estimation using semi-automated individual tree crown analysis. *Remote Sens. Environ.* **2003**, *85*, 355–369. [CrossRef]

34. Yu, X.; Hyyppä, J.; Litkey, P.; Kaartinen, H.; Vastaranta, M.; Holopainen, M. Single-sensor solution to tree species classification using multispectral airborne laser scanning. *Remote Sens.* **2017**, *9*, 108–124. [CrossRef]

35. Kim, S. Individual Tree Species Identification Using LIDAR-Derived Crown Structures and Intensity Data. Ph.D. Thesis, University of Washington, Seattle, WA, USA, 2008.

36. Kim, S.; McGaughey, R.J.; Andersen, H.E.; Schreuder, G. Tree species differentiation using intensity data derived from leaf-on and leaf-off airborne laser scanner data. *Remote Sens. Environ.* **2009**, *113*, 1575–1586. [CrossRef]

37. St-Onge, B.; Budei, B.C. Individual tree species identification using the multispectral return intensities of the Optech Titan LiDAR system. In Proceedings of the SilviLaser, La Grande Motte, France, 28–30 September 2015; pp. 71–73.

38. Hopkinson, C.; Chasmer, L.; Gynan, C.; Mahoney, C.; Sitar, M. Multisensor and multispectral lidar characterization and classification of a forest environment. *Can. J. Remote Sens.* **2016**, *42*, 501–520. [CrossRef]

39. Axelsson, A.; Lindberg, E.; Olsson, H. Exploring multispectral ALS data for tree species classification. *Remote Sens.* **2018**, *10*, 183. [CrossRef]

40. Budei, B.C.; St-Onge, B.; Hopkinson, C.; Audet, F.A. Identifying the genus or species of individual trees using a three-wavelength airborne lidar system. *Remote Sens. Environ.* **2018**, *204*, 632–647. [CrossRef]

41. Optech. Optech Titan. 2015. Available online: http://www.teledyneoptech.com/index.php/product/titan/ (accessed on 28 April 2018).

42. Polewski, P.P. Reconstruction of Standing and Fallen Single Dead Trees in Forested Areas from LiDAR Data and Aerial Imagery. Ph.D. Thesis, Technische Universität München, Munich, Germany, 2017.

43. Polewski, P.; Yao, W.; Heurich, M.; Krzystek, P.; Stilla, U. Free shape context descriptors optimized with genetic algorithm for the detection of dead tree trunks in ALS point clouds. In Proceedings of the Isprs Geospatial Week 2015, La Grande Motte, France, 28 September–3 October 2015; pp. 41–48.

44. Cailleret, M.; Heurich, M.; Bugmann, H. Reduction in browsing intensity may not compensate climate change effects on tree species composition in the Bavarian Forest National Park. *For. Ecol. Manag.* **2014**, *328*, 179–192. [CrossRef]

45. Bässler, C.; Förster, B.; Moning, C.; Müller, J. The BIOKLIM-Project: Biodiversity research between climate change and wilding in a temperate montane forest–the conceptual framework. *Waldökologie Landschaftsforschung und Naturschutz* **2008**, *7*, 21–33.

46. Amiri, N.; Yao, W.; Heurich, M.; Krzystek, P.; Skidmore, A.K. Estimation of regeneration coverage in a temperate forest by 3D segmentation using airborne laser scanning data. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 252–262. [CrossRef]

47. Hopkinson, C. The influence of flying altitude, beam divergence, and pulse repetition frequency on laser pulse return intensity and canopy frequency distribution. *Can. J. Remote Sens.* **2007**, *33*, 312–324. [CrossRef]

48. Briese, C.; Höfle, B.; Lehner, H.; Wagner, W.; Pfennigbauer, M.; Ullrich, A. Calibration of full-waveform airborne laser scanning data for object classification. In *Proceedings of SPIE Laser Radar Technology and Applications XIII*; Turner, M.D., Kamerman, G.W., Eds.; Volume 6950, pp. 0H1–0H8. Available online: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/6950/69500H/Calibration-of-full-waveform-airborne-laser-scanning-data-for-object/10.1117/12.781086.short?SSO=1 (accessed on 7 November 2019).

49. Reitberger, J.; Krzystek, P.; Stilla, U. Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees. *Int. J. Remote Sens.* **2008**, *29*, 1407–1431. [CrossRef]

50. Reitberger, J. 3D-Segmentierung von Einzelbäumen und Baumartenklassifikation aus Daten flugzeuggetragener Full Waveform Laserscanner. Ph.D. Thesis, Technische Universität München, Munich, Germany, 2010.

51. RIEGL. *LAS Extrabytes Implementation in RIEGL Software*; Technical Report; 2012. Available online: http://www.riegl.com/uploads/tx_pxpriegldownloads/Whitepaper_LASextrabytes_implementation_in-RIEGLSoftware_2017-12-04.pdf (accessed on 7 November 2019).

52. Höfle, B.; Pfeifer, N. Correction of laser scanning intensity data: Data and model-driven approaches. *ISPRS J. Photogramm. Remote Sens.* **2007**, *62*, 415–433. [CrossRef]

53. RIEGL. (Horn, Austria). Personal communication, November 2017.

54. Kashani, A.G.; Olsen, M.J.; Parrish, C.E.; Wilson, N. A review of LiDAR radiometric processing: From ad hoc intensity correction to rigorous radiometric calibration. *Sensors* **2015**, *15*, 28099–28128. [CrossRef]

55. Shi, J.; Malik, J. Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 888–905.

56. Toldo, R.; Castellani, U.; Fusiello, A. The bag of words approach for retrieval and categorization of 3D objects. *Visual Comput.* **2010**, *26*, 1257–1268. [CrossRef]

57. Weinmann, M.; Jutzi, B.; Hinz, S.; Mallet, C. Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. *ISPRS J. Photogramm. Remote Sens.* **2015**, *105*, 286–304. [CrossRef]

58. Weinmann, M.; Urban, S.; Hinz, S.; Jutzi, B.; Mallet, C. Distinctive 2D and 3D features for automated large-scale scene analysis in urban areas. *Comput. Graph.* **2015**, *49*, 47–57. [CrossRef]

59. Hackel, T.; Wegner, J.D.; Schindler, K. Fast semantic segmenation of 3D point clouds with strongly varying density. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* **2016**, *3*, 177–184. [CrossRef]

60. Guyon, I.; Elisseeff, A. An introduction to variable and feature selection. *J. Mach. Learn. Res.* **2003**, *3*, 1157–1182.

61. Liu, H.; Motoda, H.; Setiono, R.; Zhao, Z. (Eds.) Feature selection: An ever evolving frontier in data mining. In *Proceedings of the The Fourth Workshop on Feature Selection in Data Mining*; Proceedings of Machine Learning Research; PMLR: Hyderabad, India, 2010; Volume 10, pp. 4–13.

62. Li, J.; Cheng, K.; Wang, S.; Morstatter, F.; Trevino, R.P.; Tang, J.; Liu, H. Feature Selection: A Data Perspective. *ACM Comput. Surv.* **2017**, *50*, 94:1–94:45. [CrossRef]

63. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning*; Springer New York Inc.: New York, NY, USA, 2001; pp. 57–60.

64. Van Laar, A.; Akça, A. *Forest Mensuration*; Springer Science & Business Media: Amsterdam, The Netherlands, 2007; Volume 13.

65. Heurich, M. Automatic recognition and measurement of single trees based on data from airborne laser scanning over the richly structured natural forests of the Bavarian Forest National Park. *For. Ecol. Manag.* **2008**, *255*, 2416–2433. [CrossRef]

66. Pretzsch. Bestandsstruktur. Quantifizierung und Analyse. In *Grundlagen der Waldwachstumsforschung*; Springer Spektrum: Berlin/Heidelberg, Germany, 2019.

67. Sikora, R.; Piramuthu, S. Framework for efficient feature selection in genetic algorithm based data mining. *Eur. J. Oper. Res.* **2007**, *180*, 723–737. [CrossRef]

68. Shi, Y.; Skidmore, A.K.; Wang, T.; Holzwarth, S.; Heiden, U.; Pinnel, N.; Zhu, X.; Heurich, M. Tree species classification using plant functional traits from LiDAR and hyperspectral data. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *73*, 207–219. [CrossRef]