

## Article

# Fusion Approaches to Individual Tree Species Classification Using Multisource Remote Sensing Data <sup>†</sup>

Qian Li <sup>1</sup>, Baoxin Hu <sup>1,\*</sup> , Jiali Shang <sup>2</sup> and Hui Li <sup>3</sup> 

<sup>1</sup> Department of Earth and Space Science and Engineering, York University, 4700 Keele Street, Toronto, ON M3J 1P3, Canada; qian2018@yorku.ca

<sup>2</sup> Ottawa Centre for Research and Development, Agriculture and Agri-Food Canada, 960 Carling Avenue, Ottawa, ON K1A 0C6, Canada; jiali.shang@canada.ca

<sup>3</sup> Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China; lihui@radi.ac.cn

\* Correspondence: baoxin@yorku.ca

<sup>†</sup> This manuscript is part of a Master's thesis by Qian Li, available online at <http://hdl.handle.net/10315/40623> (accessed on 14 December 2022).

**Abstract:** With the wide availability of remotely sensed data from various sensors, fusion-based tree species classification approaches have emerged as a prominent and ongoing research topic. However, most recent studies primarily focused on combining multisource data at the feature level, while few systematically examined their positive or negative contributions to tree species classification. This study aimed to investigate fusion approaches at the feature and decision levels deployed with support vector machine and random forest algorithms to classify five dominant tree species: Norway maple, honey locust, Austrian pine, white spruce, and blue spruce in individual crowns. Spectral, textural, and structural features derived from multispectral imagery (MSI), a very high-resolution panchromatic image (PAN), and LiDAR data were systematically exploited to assess their contributions to accurate classifications. Among the various classification schemes that were explored, both feature- and decision-level fusion approaches demonstrated significant improvements in tree species classification compared with the utilization of MSI (0.7), PAN (0.74), or LiDAR (0.8) in isolation. Notably, the decision-level fusion approach achieved the highest overall accuracies (0.86 for SVM and 0.84 for RF) and kappa coefficients (0.82 for SVM and 0.79 for RF). The misclassification analysis of fusion approaches highlighted the potential and flexibility of decision-level fusion in tree species classification.

**Keywords:** tree species classification; multisource remotely sensed data; feature selection; feature importance; decision fusion



**Citation:** Li, Q.; Hu, B.; Shang, J.; Li, H. Fusion Approaches to Individual Tree Species Classification Using Multisource Remote Sensing Data. *Forests* **2023**, *14*, 1392. <https://doi.org/10.3390/f14071392>

Academic Editor: Giorgos Mallinis

Received: 30 May 2023

Revised: 30 June 2023

Accepted: 3 July 2023

Published: 7 July 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The accurate classification of tree species in urban areas plays a crucial role in urban forest assessment, ecological management, and sustainable city planning. It allows urban planners and ecologists to effectively monitor and manage the urban ecosystem by evaluating the tree-species-specific composition and spatial distribution. This information helps with assessing the overall functionality and resilience of urban forests. It is essential to have up-to-date tree species data for making informed decisions regarding management strategies related to air and water purification, temperature regulation (especially in mitigating urban heat islands), soil erosion and flooding prevention, and climate change mitigation [1]. For instance, a study on air pollution demonstrated that certain tree species significantly influence air quality in urban environments [2]. Felton et al. [3] highlighted the importance of specific tree species by identifying potential consequences for forest biodiversity and ecosystem services resulting from changes in tree species [3]. Additionally, accurate tree

species data are integral to creating well-designed green spaces, urban parks, and infrastructure developments that enhance biodiversity and promote the overall livability and well-being of urban residents.

Remote sensing has been used as a cost-effective method in tree species classification, an alternative to conventional labor-intensive and time-consuming field surveys, and visual interpretations of aerial photographs over several decades [4]. Tree species classification using remotely sensed data has also evolved with the advancement of geospatial technologies. High-spatial-resolution remotely sensed data make it possible to derive features for individual tree crowns, thus enabling the capture of more precise spectral signatures and facilitating the analysis of textural features, such as variations in canopy structure or leaf arrangement. This approach takes into account the crown-level characteristics, resulting in a higher level of accuracy and increased discriminatory power when extracting and utilizing these features. Moreover, structural features that capture detailed 3D information about tree canopies provide valuable insights into the geometric characteristics of different tree species. Textural and structural features enable the discrimination and classification between different species with similar spectral characteristics [5–8].

Spectral features derived from multispectral imagery (MSI) play a fundamental role in tree species classification. While spectral features, including reflectance values and vegetation indices, are widely employed in tree species classification, their utility alone is limited [9–13]. The variations in spectral reflectance within individual tree crowns and the potential similarities between different species highlight the need for complementary data and features to improve the accuracy and reliability of tree species classification models. Textural features provide valuable insights into the 2D structural characteristics by quantifying the frequency and distribution of reflectance values of tree canopies and play a significant role in tree species classification [14,15]. Among the various texture analysis methods, several metrics from the statistical grey-level co-occurrence matrix (GLCM), such as homogeneity, contrast, correlation, and energy, have been commonly employed in tree species classification [16–18]. However, a comprehensive investigation of GLCM-based textural measures, including a broader range of feature types, is needed. Additionally, it is worth noting that other texture analysis techniques, such as those based on Gabor filters, have been extensively utilized for texture segmentation and classification but are rarely employed in tree species classification [19]. The inclusion of the Gabor filter technique could potentially enhance the discriminatory power and accuracy of tree species classification models.

Many studies reported successful tree species classification results using various light detection and ranging (LiDAR) data-derived features [7,8,20,21]. However, it is worth noting that Aval et al. [16] reported a marginal contribution of height statistics features derived from LiDAR data when in combination with hyperspectral and panchromatic imagery in individual tree species classification. Their study highlighted the importance of extracting the most relevant features prior to merging the complementary multisource remote sensing data. Therefore, more attention is needed to investigate the advanced 3D structural features, such as point-distribution-related features for multisource remote-sensing-assisted tree species classification, in addition to the comprehensive investigation of textural and spectral features.

Taking advantage of these multiple features extracted from the abundance of data provided by various sensors (e.g., high-resolution multispectral, hyperspectral, and LiDAR systems), researchers have recognized the potential benefits of integrating multisource data to improve the performance of tree species classification [5,11,12,22–26]. Fusion-based classification approaches have gained considerable attention in recent years and continue to be an active area of research. Extensive utilization of multisource remotely sensed data was found to be effective in improving the separation of tree species by leveraging the complementary information available from different scenarios. Most recent studies that employed fusion approaches incorporated multisource data at the feature level [11–13,22–25,27]. For example, Liu et al. [25] employed LiDAR in a combination of hyperspectral imagery to

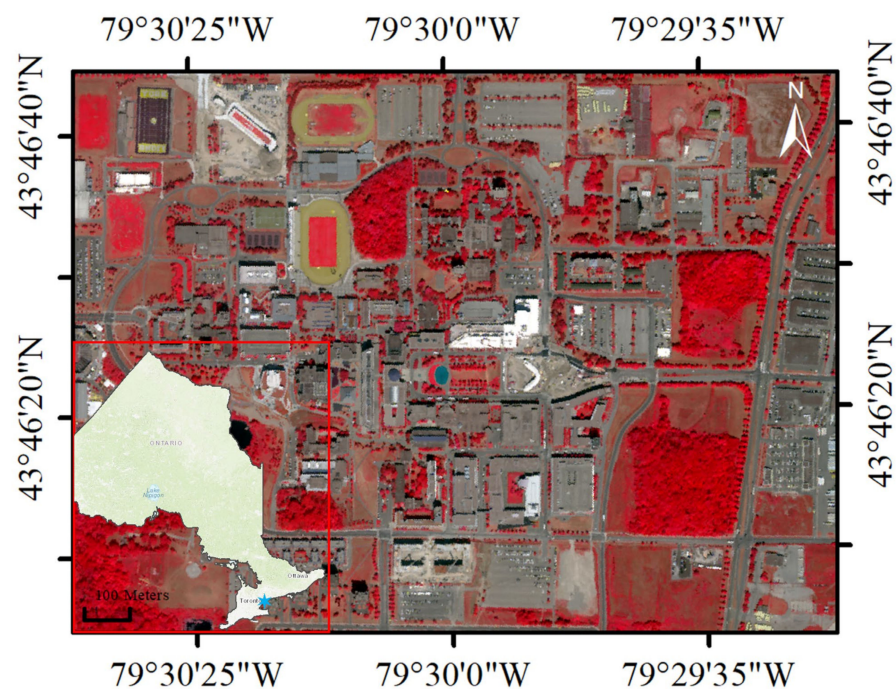
classify 15 common urban tree species using an RF classifier, resulting in an increased accuracy of 70% using feature fusion. Nevertheless, the high dimensionality in the feature space that results from feature-level fusion is likely to be a concern for applications where the size of training samples is small [28]. In addition, features derived from different data sources are treated equally by the SVM and random forest methods, even though some data sources may be more reliable than others [29]. On the other hand, several studies recently investigated fusion at the decision level for species classification, demonstrating the advantages of decision fusion over feature fusion [16,18,29]. Stavrakoudis et al. [18] mapped five forest species using the combined hyperspectral and multispectral imagery, and they showed that the decision fusion exhibited a higher degree of flexibility and accuracy than the feature fusion. Aval et al. [16] and Hu et al. [29] also highlighted that it was important but difficult to define an appropriate decision rule to merge different data sources. An in-depth analysis of decision fusion approaches is needed to realize its potential in species classification.

Although many studies illustrated the potential of fusion approaches integrating optical imagery and LiDAR data, how to effectively harness the full potential of each dataset and successfully integrate information derived from multisource data remains a significant challenge. This study aimed to address this challenge by exploring methods to extract advanced textural and structural features and investigating the optimal decision fusion approach deployed with advanced machine learning classification models. It implemented feature selection techniques and evaluated various classification schemes using spectral, textural, and structural features using feature-level fusion. This study developed a decision fusion framework based on the Dempster–Shafer theory (DST), leveraging probabilistic support vector machines (SVMs) and random forests (RFs). A misclassification analysis of the decision fusion and feature-level fusion results was also conducted to provide insights into further improvement. This study systematically described and examined the spectral, advanced textural, and structural features derived from the multispectral imagery, high-resolution PAN, and LiDAR data. Additionally, the contributions of these features toward achieving accurate classifications were investigated through feature importance analysis and feature-group-based classification evaluation to quantify the discriminative power of these features.

## 2. Study Area and Data Pre-Processing

### 2.1. Study Area

This study was conducted at the Keele Campus (43.77° N, 79.50° W) of York University, which is situated in the Greater Toronto Area, Ontario, Canada. The campus spans 457 acres of land and provides an urban setting for tree species classification (Figure 1). For this study, five dominant tree species were selected: Norway maple (*Acer platanoides*), honey locust (*Gleditsia triacanthos*), Austrian pine (*Pinus nigra*), white spruce (*Picea glauca*), and blue spruce (*Picea pungens*).



**Figure 1.** A false color composite (Band 7 printed as Red, Band 5 as Green, and Band 3 as Blue) of the WorldView-2 MSI over the study area of the Keele campus of York University (approximately indicated by the blue star overlapped on the administrative boundaries of the province of Ontario (red rectangle in the bottom-left corner).

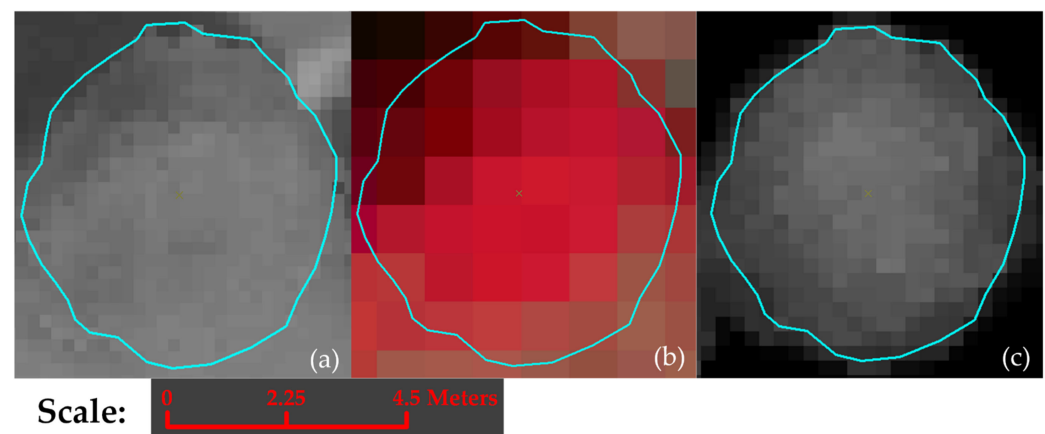
## 2.2. Data Used and Pre-Processing

To create a reference sample dataset, 751 tree crowns (Table 1) from these five dominant species were randomly chosen based on the street tree inventory conducted by Campus Services and Business Operations (CSBO) at York University in June 2015. The selected tree samples were strategically located along streets, near buildings, and other areas with high pedestrian traffic, with the aim to represent the typical distribution of trees in an urban environment. They were manually delineated on MSI and a LiDAR-derived canopy height model (CHM) using ArcMap 10.6 software. An aerial image acquired in May 2016 with a high spatial resolution of 8 cm by 8 cm, which was provided by the York University Map Library, was visually inspected to confirm the tree species. Additionally, Google Street View images were utilized to examine the ground truth of the tree species for the sample dataset. These delineated tree crowns were then superimposed onto the multisource remotely sensed data to extract target features and generate the classification sample dataset. Figure 2 illustrates an example of the delineated individual tree crowns in a PAN image, false-color MSI, and a CHM. To create training and testing subsets, tree samples of each species were randomly divided at a ratio of 7:3, resulting in 528 samples for training and 223 samples for validation.

**Table 1.** Ground reference dataset for tree species classification.

Common Name	Scientific Name	Tree Type	Samples Number	Proportion (%)
Norway maple	<i>Acer platanoides</i>	Broadleaf	188	25
Honey locust	<i>Gleditsia triacanthos</i>	Broadleaf	180	24
Austrian pine	<i>Pinus nigra</i>	Conifer	159	21
Blue spruce	<i>Picea pungens</i>	Conifer	115	15
White spruce	<i>Picea glauca</i>	Conifer	109	15
		Total	751	100

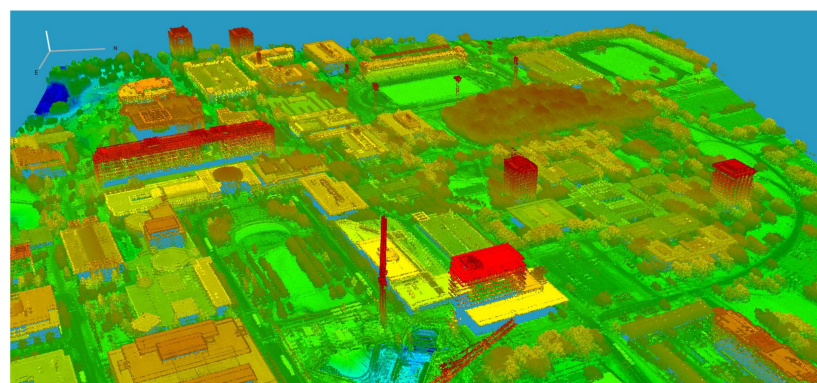




**Figure 2.** An example of a Norway maple tree crown manually delineated in (a) a PAN, (b) false-color MSI, and (c) a LiDAR-derived CHM.

This study utilized multisource remotely sensed data, including WorldView-2 (WV-2) MSI, high-resolution PAN image, and airborne LiDAR data. The WV-2 satellite sensor, which was provided by DigitalGlobe Inc. (Westminster, CO, USA), captured panchromatic imagery with a resolution of 0.4 m and eight-band MSI with a resolution of 1.6 m. The wavelength range covered by MSI was from 400 nm to 1040 nm. The data were collected in Toronto on 21 July 2016. The WV-2 MSI underwent radiometric correction to convert the digital numbers to at-sensor radiances. This correction process utilized radiometric calibration parameters and standard correction from the WV-2 image calibration file, and it was performed using MATLAB (Version 2020a, The Math Works, Inc., Natick, MA, USA, 2020). Furthermore, the atmospheric correction was conducted to obtain the surface reflectance of the WV-2 MSI. The surface reflectance values ranged from 0 to 100% and were obtained using the Atmospheric and Topographic Correction (ATCOR) model in the PCI Geomatics software (PCI Geomatics 2018).

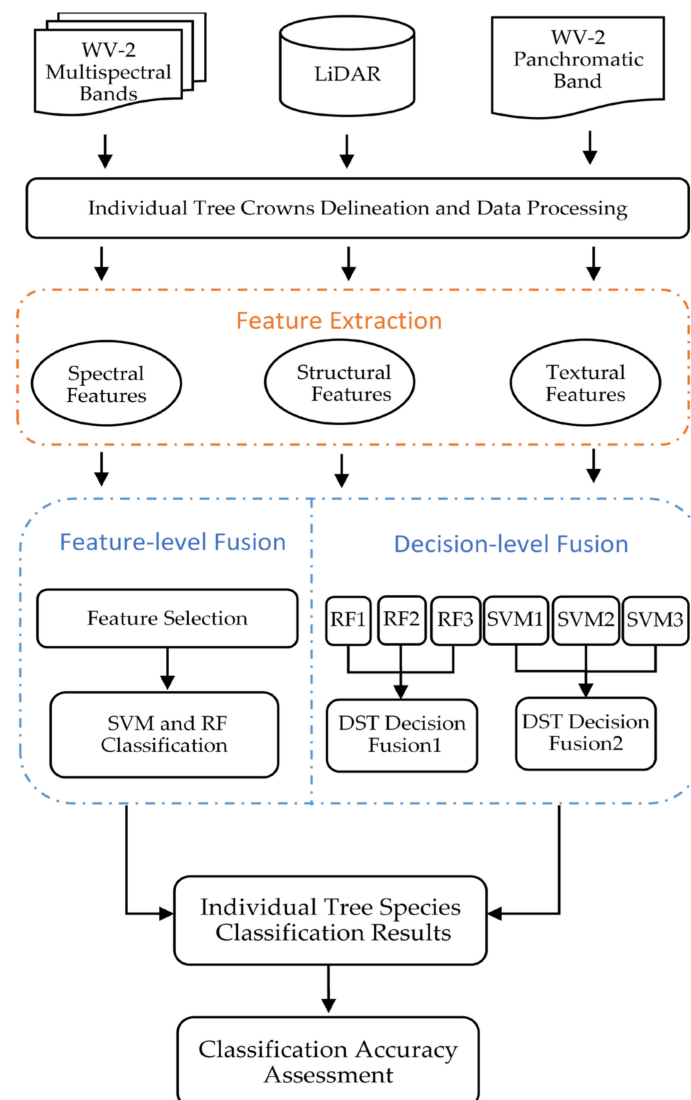
The LiDAR data provided by the York University Map Library was acquired in April 2015 during the early spring leaf-off conditions. The data set was collected using a Leica ALS70-HP discrete return LiDAR system mounted on an aircraft flying at 160 knots (82.3 m/s). The acquisition flights were undertaken at an altitude of 800 m AGL (above ground level) with the laser pulse rate set at 300,000 Hz, resulting in four discrete returns and an aggregated point density of 10 points per square meter. The horizontal and vertical accuracies of the LiDAR data were 30 cm and 10 cm, respectively. To create the CHM, also known as the normalized digital surface model (nDSM), the LiDAR data was processed by obtaining the difference between a digital elevation model (DEM) and a digital surface model (DSM) using MATLAB (Version 2020a, The Math Works, Inc., 2020). A 3D perspective view of the LiDAR point cloud is presented in Figure 3.



**Figure 3.** Three-dimensional perspective view of the LiDAR point cloud data for the study area.

### 3. Methodology

Figure 4 presents the comprehensive workflow employed in this study. Initially, individual tree crowns were manually delineated to enable species analysis at the single-tree level. Subsequently, spectral, textural, and structural features were derived from the MSI, PAN, and LiDAR data for each tree crown, respectively. The derived features were then utilized in the following investigations: (1) Classification was conducted using spectral, textural, and structural features individually to assess their discriminatory power in distinguishing the five tree species of interest. Furthermore, the contributions of advanced textual and structural features were explored. (2) Feature-level fusion was employed for classification purposes. All possible combinations of individual spectral, textural, and structural feature groups were investigated and compared to enhance our understanding of the performance of different classification schemes by utilizing the feature-level fusion approach. (3) Decision fusion classification was performed, and an in-depth analysis was carried out to evaluate the results obtained from the fusion process.




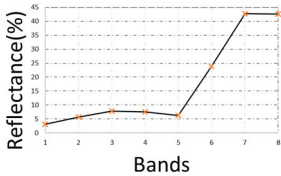


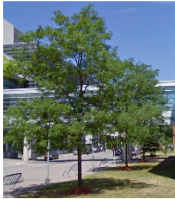
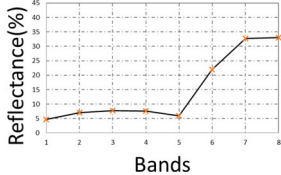

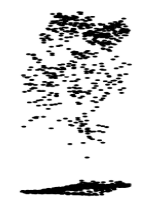
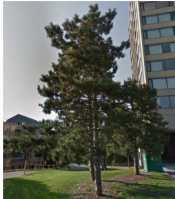
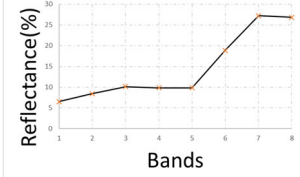
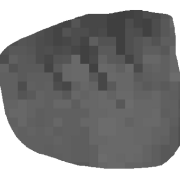
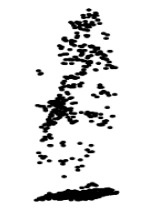

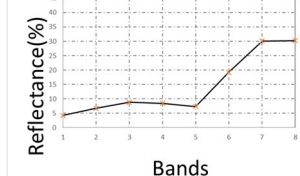
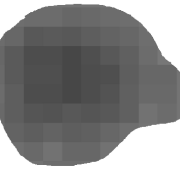


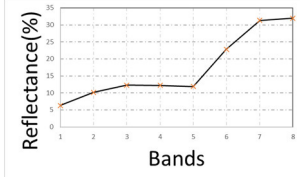
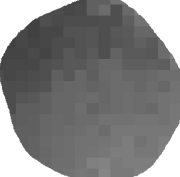
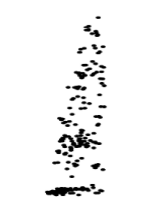
**Figure 4.** Workflow of proposed crown-based tree species classification.

#### 3.1. Feature Extraction

The identification of tree species can be facilitated by considering the physical and biophysical characteristics of their tree crowns. Integrating multisource remotely sensed data allows for the extraction of complementary features from 2D optical imagery and 3D

LiDAR points, effectively capturing the traits associated with different tree species. In this study, spectral, textural, and structural information extracted from MSI, PAN, and LiDAR data was utilized to classify the individual tree species. Table 2 presents ground photos alongside their respective spectral signatures and representations in the PAN and LiDAR point data. These examples illustrate manually delineated tree crowns of the five species of interest.

**Table 2.** Examples of tree crowns for five species and their ground photos, spectral curves, PAN image, and LiDAR point plots.

Species	Ground Photos	Spectral Curve of MSI	PAN Image	LiDAR Points
Norway maple				
Honey locust				
Austrian pine				
Blue spruce				
White spruce				

Different tree species have distinct biochemical and biophysical properties, leading to different spectral responses. This mainly drives the use of spectral signatures in tree species classification. Spectral features in Table 3 were extracted from WV-2 MSI to evaluate the performance of spectral signatures in distinguishing tree species, including the mean and standard deviation of the reflectance of individual tree crowns of each spectral band (16 in total and 11 selected) and five vegetation indices consisting of the normalized difference vegetation index (NDVI) [30], enhanced vegetation index (EVI) [31], green normalized difference vegetation index (GNDVI) and red edge normalized difference vegetation index (RENDVI) [32], and optimized soil adjusted vegetation index (OSAVI) [33,34].

**Table 3.** Spectral, textural, and structural features derived from MSI, PAN, and LiDAR data for tree species classification and selected features for classification.

Dataset	Feature Group	Representatives of Individual Tree Crowns' Characteristics	Selected Features	No.
MSI	Spectral features	Mean and standard deviation of the reflectance of tree crown using eight bands Vegetation indices combining reflectance of different bands (EVI, GNDVI, RNDVI)	Reflectance_B4,6,7,8; SD_B2,6,7,8 EVI, GNDVI, RE_NDVI	11
PAN	Textural features	Two-dimensional GLCM-based texture analysis describes the variations in the intensity of pixels belonging to tree crowns in PAN	2D-Correlation, 2D-ClusterProminence, 2D-ClusterShade, 2D-Entropy, 2D-InverseDifferenceMoment, 2D-SumVariance, 2D-MaximumProbability, 2D-DifferenceEntropy, 2D-InformationMeasureofCorrelation1, 2D-InformationMeasureofCorrelation2, 2D-InverseDifferenceNormalized	11
		Gabor-filter-based textural features provide robustness against varying brightness and contrast of pixels within the tree crown in PAN	GaborFilter-SquareEnergy 1,2,17,18,21,22,24GaborFilter-MeanAmplitude1,21	9
LiDAR point clouds	Structural features	Normalized number of points at horizontal layers using the total number of individual tree points, presenting the branch and foliage distribution at vertical profile	Density_Layer1,2,3,4,5,9	6
		Crown area and the ratio of the crown areas to the maximum crown area at horizontal layers, presenting the vertical foliage clusters at these layers	Area, Vertical_cluster1,2,5,9,10	6
		The proportion of first, second, and third returns subtracted from 1, presenting gap distributions within the tree crown opposite to foliage covers	Gap_distribution1, Gap_distribution2, Gap_distribution3	3
		Measures of the 3D spatial relationship of neighboring voxels with different LiDAR point numbers in a tree crown, characterizing the arrangement of foliage, twigs, and branch	3D-Contrast, 3D-SumMean, 3D-ClusterShade, 3D-ClusterTendency	4
CHM	Structural features	Absolute tree height statistics and the combinations with area information	Max_H/Area, Max_H*Area, SD_H/Max_H, Mean_H, (Max_H-Min_H)/Max_H, Max_H, Mean_H, (Max_H-Mean_H)/Max_H, Max_H-Mean_H, SD_H	10

The WV-2 PAN image had a significantly higher spatial resolution of 0.4 m compared with the MSI imagery (1.6 m), resulting in a significant number of pixels within individual tree crowns, allowing for a meaningful texture analysis. In this study, textural features based on a GLCM (gray-level co-occurrence matrix) texture analysis and the Gabor filter technique were derived from PAN listed in Table 3. An expansion of Haralick's original GLCM textural features was generated by calculating multiple GLCMs from each tree crown, which consisted of 22 features that were found by considering co-occurring values of neighboring pixel pairs. GLCMs were created using an array of offset parameters that defined spatial relationships in four directions (0°, 45°, 90°, and 135°), with a one-pixel distance and 64 distinct gray levels. The one-pixel distance was selected to capture detailed spatial variations within individual tree crowns, taking into account the 0.4 m spatial resolution. GLCM vectors represented each input tree crown in the four directions, which were then averaged to calculate statistical texture features. MATLAB (Version



2020a, The Math Works, Inc., 2020) was used for the computational processing of these texture statistics.

Among the 22 GLCM-based textural features, 13 were formulated by Haralick et al. [15], namely, contrast, correlation1, energy, entropy, inverse difference moment, the sum of squares: variance, sum average, sum variance, sum entropy, difference variance, difference entropy, information measure of correlation1, and information measure of correlation2. Soh et al. [35] discussed five additional features: cluster prominence, cluster shade, dissimilarity, autocorrelation, and maximum probability. Correlation2 and homogeneity were computed using the GLCM formulas in MATLAB (Version 2020a). Clausi [36] introduced two modified features: inverse difference normalized and inverse difference moment normalized.

In this study, a 2D Gabor filter [37] with five scales and six orientations was employed, which was defined as Equation (1):

$$h(x, y; f, \theta) = \frac{1}{\sigma\sqrt{\pi}} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \cdot \exp(i(f_x x + f_y y)), \quad (1)$$

where  $x$  and  $y$  are the coordinates of a given pixel in the image, the parameter  $\sigma$  is the standard deviation of the 2D Gaussian function in the  $x$  and  $y$  directions,  $f$  is the central frequency of a sinusoidal wave,  $f_x = f \cos \theta$  and  $f_y = f \sin \theta$ , and  $\theta$  is the spatial orientation of the filter. These filters were created with orientations at  $0^\circ$ ,  $30^\circ$ ,  $60^\circ$ ,  $90^\circ$ ,  $120^\circ$ , and  $150^\circ$ , and five different spatial frequencies in each direction (0.1, 0.3, 0.5, 0.7, and 0.9 cycles/pixel). The PAN image of individual tree crowns was convolved with each Gabor filter, resulting in 30 output images. The mean amplitude and square energy of the pixels in each output image were computed to capture the variation in specific frequency content in specific directions. These Gabor-filter-based textural features, mean amplitude, and square energy, derived from the PAN image of individual tree crowns, were added to a 60-feature vector (9 selected), denoted as GaborFilter-MeanAmplitude  $i$  ( $i = 1, 2, \dots, 30$ ) and GaborFilter-SquareEnergy  $j$  ( $j = 1, 2, \dots, 30$ ) in Table 3.

As shown in the last column of Table 2, the LiDAR point cloud data reflected the natural arrangements of foliage and branching patterns of individual tree crowns. This study extracted five types of 3D structural features to characterize the vertical profiles and 3D point distribution of individual tree crowns from the LiDAR point cloud data and CHM. They are summarized and described in detail in Table 3 with representatives of individual tree crowns' characteristics. Studies have employed discrete return airborne LiDAR data to derive tree crown structural features for various applications, such as the gap fraction and leaf area index [38]. Among the structural features in Table 3, the gap-distribution-related features were computed using the following Equation (2):

$$D_{gap}(j) = 1 - \frac{\sum_{z=z_i}^{z=max} R(j)}{R_{total}}, (j = 1, 2, 3, 4) \quad (2)$$

where  $R(j)$  refers to the number of first and only returns, second returns, third returns, and last returns within individual tree crowns, respectively. Based on studies [7,39], 12 3D GLCM-based statistical measures proposed by Haralick et al. [15] were computed to characterize the arrangement of tree elements, such as foliage, twigs, and branches based on the 3D point distribution inside the tree volume. First, the size of the voxels was determined as  $0.5 \text{ m}^3$  to optimally represent the internal structural properties of trees considering the point density of the LiDAR data. Second, the GLCM was calculated from 13 different directions in 3D space with a voxel distance ( $dx = dy = dz = 1 \text{ voxel}$ ) relationship between neighboring voxels, whose values were derived from the cumulative number of LiDAR points lying in each voxel. Last, this study extracted ten tree-height-related features similar to some structural features used by Alonzo et al. [22] and Aval et al. [16]. Absolute-height-related features were computed at the individual tree crown scale to capture the

general structural properties of tree species. The detailed descriptions of these features are in Table 3.

### 3.2. Feature Selection

The recursive feature elimination (RFE) algorithm [40] combined with an RF classifier was highlighted by several studies in the literature, as it can provide an unbiased feature selection and effectively increase the classification accuracy [41–43]. Specifically, Gregorutti et al. [43] recommended the use of the RFE algorithm, considering its good performance, even in the presence of correlated features. Demarchi et al. [41] reported that RFE effectively handles the fusion of complementary and diverse data sources, such as hyperspectral imagery and LiDAR data. In this study, the feature selection was implemented for the classification scheme of feature-level fusion using the “caret” package with version 6.0-90 [44] in the statistical software RStudio (Version 1.2.1335) (RStudio Team, 2018) [45]. RFE fitted the RF model with the initial feature space to rank features by importance at first and recursively eliminate the least important features afterward. RFE then re-examined the ranked features using a permutation importance measure at each backward elimination step. This process was repeated iteratively until the optimal number of feature subsets was obtained when the highest classification accuracy was produced.

### 3.3. Classification

Spectral, textural, and structural features derived from MSI, PAN, and LiDAR data were used to classify the five tree species using SVM and RF. The implementation of classification with SVM and RF algorithms was performed using the “e1071” package with version 1.7-6 [46], which is a package for R programming that provides functions for statistical and probabilistic algorithms and the “randomForest” package [47] in the statistical software RStudio (Version 1.2.1335) (RStudio Team, 2018). The same training and testing datasets were used for training and validating the SVM and RF classification models. The numeric values of the feature vectors derived from individual tree crowns were first normalized to a common scale from 0 to 1 before inputting them into the classification models.

#### 3.3.1. Classification Techniques

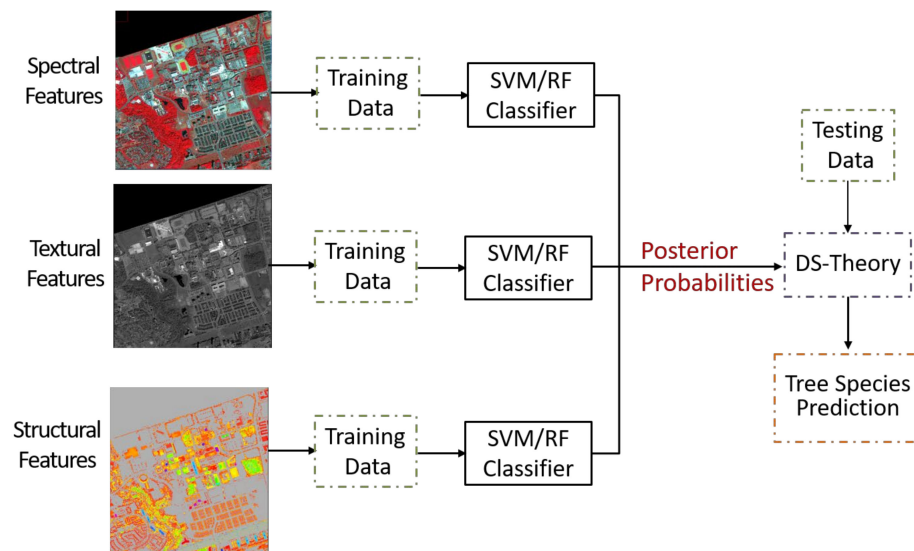
The SVM algorithm classified five tree species using the “one-against-one” approach [45], which is a pairwise classification strategy that trains multiple binary classifiers to distinguish between each pair of classes in multiclass classification problems. A total of ten binary SVM classifiers were built since five species classes were involved in this study. It is worth mentioning that four kernel types for SVM, namely, the linear kernel, polynomial kernel, radial basis function (RBF), and sigmoid kernel, were employed for the pretest and comparison. The RBF kernel was selected in the SVM classification models for the current study since it provided the best performance on both the training and testing datasets. The parameters, including the cost (C) and hyperparameter (gamma) with RBF kernel in the SVM classifier, should be optimized using the training dataset to produce the best classification performance. However, using the default parameters appeared to achieve a classification accuracy close to what was achieved through the parameter optimization process after multiple verifications of empirical tests using separate feature groups. Therefore, in order to increase the generalizability of the classification model across multiple classification schemes, the default parameters of the RBF kernel SVM classifier (cost set as one and gamma as the inverse of the number of features) were used for all the classification schemes in this study. Furthermore, the RBF-kernel-based SVM classifier computed the posterior class probabilities of each tree sample belonging to the five tree species of interest as outputs instead of crisp class labels, which were employed as the mass functions in the subsequent decision-level fusion approach.

Two parameters were set up for the RF classifier: the number of classification trees (*ntree*) and the number of features randomly sampled as candidates at each node (*mtry*). Based on the experiment by Maxwell et al. [28], a large *ntree* number of 500 was set. The

*mtry* values varied across a wide range of values from 6 to 12 for the optimization, as suggested in the documentation of the randomForest package, where the default value of *mtry* for classification was set to approximately the square root of the number of features in the training dataset. The *mtry* value that delivered the minimum out-of-bag error was chosen to build the classification model. A 10-fold cross-validation with three repeats was performed to evaluate the out-of-bag error estimate using multiple *mtry* values on the training data. In RF, the final classification result was obtained using the tree species with the majority vote. The proportion of votes predicting each sample as a specific tree species was computed as the class-specific probabilities in the RF algorithm, enabling a soft decision fusion approach. The class probabilities measuring the confidence of classification results were output as the mass functions in the subsequent decision-level fusion approach.

### 3.3.2. Fusion Approaches

The current study considered two classification cases: feature-level fusion (case A) and decision-level fusion (case B). In case A, individual feature groups without feature selection were directly utilized to build classification models using SVM and RF algorithms, respectively. The classification outputs of separate feature groups were presented in the class probabilities instead of the predicted classes. The evidence or confidence of the preliminary classification results from individual feature groups was quantified as numeric posterior probabilities using the improved Platt scaling method [48,49] in SVM and the voting of trees in RF [50]. For a tree sample, the probability indicated the percentage supporting the belief that the tree object belongs to a specific tree species. The workflow diagram in Figure 5 depicts the outline of the DST-based decision-level fusion approach deployed with SVM/RF classifiers to tree species classification. The original features from each dataset were used to compute the posterior probabilities for the decision-level fusion approach and to obtain classification accuracies for comparisons.



**Figure 5.** Workflow of the decision-level fusion approach deployed with SVM and RF.

Dempster's rule of combination (the joint mass) aggregates multiple mass functions calculated from pieces of evidence (here, preliminary classification results as posterior probabilities from the respective MSI, PAN, and LiDAR data). The frame of discernment, which is denoted by  $\theta$ , is first defined in DST, which contains all the possible classes  $A_i$  under consideration. When  $\theta = \{A_1, A_2 \dots A_i\}$ , the power set  $2^\theta$  contains all the subsets  $A_i$  of  $\theta$  (denoted as  $A_i \subseteq 2^\theta$ ), including the empty and full sets. DST uses a mass function  $m(A_i)$ , also called BPA or basic belief assignment (BBA), to represent the degree of belief in classes given the relevant and available evidence that supports the claim that the actual class belongs to  $A_i$ . In this study, the mass function  $m(A_i)$  is in the form of probabilities. For

any  $A_i \subseteq 2^\theta$ ,  $m(A_i) \in [0, 1]$ ,  $m(\emptyset) = 0$ , and  $\sum_{A_i \subseteq 2^\theta} m(A_i) = 1$ . The decision rule defined by Shafer (1976) [51] using Equation (3) was applied to combine multisource remotely sensed data for tree species classification in this study. The posterior class probabilities from individual classification schemes were employed as mass functions that were further combined through the DST combination rule for the decision-level fusion. The DST-based decision fusion approach combined probabilities for five tree species of interest computed from multisource features and arrived at a degree of confidence, considering all the available evidence for the decision-making. The most probable tree species was ultimately assigned to a given sample based on the decision criterion of the maximum probability:

$$\begin{cases} m_{1,2,\dots,i}(A) = \frac{\sum_{B_1 \cap \dots \cap B_n = A} \prod_{i=1}^n m_i(B_i)}{1-K} \\ K = \sum_{B_1 \cap \dots \cap B_n = \emptyset} \prod_{i=1}^n m_i(B_i) \end{cases} \quad (3)$$

where  $m_i(A)$  signifies the combined mass for the possible predicted tree species  $A$ .  $m_i(B_i)$  are the individual masses assigned to subsets of the frame of discernment, which are posterior probabilities from the respective MSI, PAN, and LiDAR data. The conflict parameter, which is denoted as  $K$ , measures the amount of conflict or uncertainty among the multisource-derived features. The value of  $K$  indicates the extent to which these features or data support conflicting predictions. After the combination using Equation (3), the decision criterion that the maximum probability indicated as the most credible or plausible decision was used in this study to ultimately classify a given sample into one of the interested tree species classes.

For case B, feature selection was implemented first, and classification was performed based on the selected features to calculate classification accuracies for comparison with the feature-level fusion approach. The importance of individual feature groups was analyzed based on the results of the feature selection and their discriminatory powers in the classification accuracies evaluated in case B. In particular, textural features were subdivided into GLCM- and Gabor-filter-based features, and structural features were considered as CHM and LiDAR point-cloud-derived features to further assess the discriminatory power of different feature types from the same dataset. A feature-level fusion approach integrated spectral, textural, and structural features for classification using SVM and RF algorithms. The RFE algorithm was implemented to select the most relevant features from the original features. Based on the selected features, all possible feature-level fusion schemes in terms of the types of features used were investigated and compared to enhance the understanding of the relative importance of these features and their classification performances using the feature-level fusion approach. Table 4 summarizes all the designed classification schemes in the feature-level fusion. It is worth noting that the feature selection was implemented for each combination in the feature-level fusion; however, there was no significant difference in the results from the feature selection conducted for all features combined. Sixty features selected from all features in combination are presented in Table 3.

**Table 4.** Classification schemes using the feature-level fusion approach.

Features Combination	No.
Spectral and textural features	31
Spectral and structural features	40
Textural and structural features	49
All features	60

### 3.3.3. Classification Accuracy Assessment

The performance of different classification methods was assessed by comparing the predicted results with ground truth data using independent testing data. To evaluate the classification accuracy, three evaluation metrics, namely, the overall accuracy, kappa coefficient, and F1-score, were calculated based on the confusion matrix, which is a widely accepted method for assessing classification accuracy [51]. The F1-score, which is a har-



monic mean of the user's and producer's accuracy, provides a comprehensive measure that takes into account both aspects of performance. With a range from 0 to 1, a perfect F1-score of 1 indicates optimal results for both the user's and the producer's accuracy. The user's accuracy represents the fraction of correctly identified tree samples among all the samples classified into a specific tree species, reflecting the probability that the prediction accurately represents reality. Conversely, the producer's accuracy denotes the ratio of correctly classified tree samples to the total ground truth tree samples in each species, indicating the classification quality within the testing set.

## 4. Result

### 4.1. Classification Using Individual Spectral, Textural, and Structural Features

The classification accuracies for individual feature groups in both cases are shown in Table 5, and selected features for case B are comprehensively presented in Table 6. A total of 60 features were selected, which contained 11 spectral, 20 textural, and 29 structural features (shown in Table 3).

**Table 5.** Classification results of individual feature groups using SVM and RF (entries with the star (\*) indicate the highest overall accuracy among all feature groups).

		SVM		RF	
Feature Groups		Overall Accuracy	Kappa	Overall Accuracy	Kappa
Case A	Spectral features	0.70	0.62	0.70	0.62
	Textural features	0.76	0.70	0.75	0.68
	Structural features	0.78	0.72	0.76	0.69
Case B	Selected spectral features	0.70	0.62	0.65	0.56
	Selected textural features	0.74	0.68	0.72	0.64
	Selected structural features	0.80 *	0.74 *	0.78 *	0.72 *

**Table 6.** Classification results of case B (entries with a star (\*) indicate the highest F1-score for each species and feature group with the highest overall accuracy).

	Norway Maple	Honey Locust	Austrian Pine	Blue Spruce	White Spruce	OA
<b>SVM Classification</b>						
SF	0.87	0.71	0.69	0.58	0.49	0.70
TF_GLCM	0.91	0.67	0.65	0.68	0.54	0.71
TF_GABOR	0.80	0.38	0.69	0.40	0.60 *	0.60
TF	0.89	0.70	0.74	0.72 *	0.58	0.74
STF_CHM	0.90	0.70	0.74	0.54	0.49	0.71
STF_3D	0.85	0.67	0.86	0.62	0.56	0.74
STF	0.93 *	0.80 *	0.90 *	0.67	0.53	0.80 *
<b>RF Classification</b>						
SF	0.88	0.69	0.65	0.46	0.35	0.65
TF_GLCM	0.88	0.67	0.62	0.60	0.46	0.67
TF_GABOR	0.82	0.33	0.63	0.42	0.59 *	0.58
TF	0.90	0.67	0.71	0.65 *	0.54	0.72
STF_CHM	0.91	0.73	0.72	0.57	0.48	0.72
STF_3D	0.88	0.70	0.82	0.61	0.53	0.74
STF	0.93 *	0.77 *	0.86 *	0.63	0.51	0.78 *

SF: spectral features. TF\_GLCM: textural features based on the statistical 2D GLCM. TF\_GABOR: textural features based on the Gabor filter method. TF: the combination of GLCM- and Gabor-filter-based features. STF\_CHM: structural features extracted from the LiDAR-derived CHM. STF\_3D: structural features extracted from the 3D LiDAR point cloud data. STF: combined CHM and 3D structural features.

The contribution of the advanced textural features and structural features were further demonstrated based on the classification accuracies of feature-level fusion using the testing dataset shown in Table 6, where the F1-score and overall accuracy of the selected features in the groups are detailed. The F1-score was involved in a comprehensive analysis of the classification contribution of feature groups on each tree species. GLCM- and Gabor-filter-based textural features and CHM- and 3D-LiDAR-point-cloud-based features were analyzed separately to evaluate the commonly used height-related features and rarely used 3D structural features, and the results are given in Table 6.

#### 4.2. Classification Using Feature-Level Fusion

Figure 6 depicts the importance ranking of selected features using the RFE algorithm, ordering from top to bottom as the most to least essential features. Structural features accounted for 48 percent of the selected feature subset (60 features in total). In comparison, spectral features accounted for 18 percent (11 features), and the proportion of textural features was 34% (20 features).

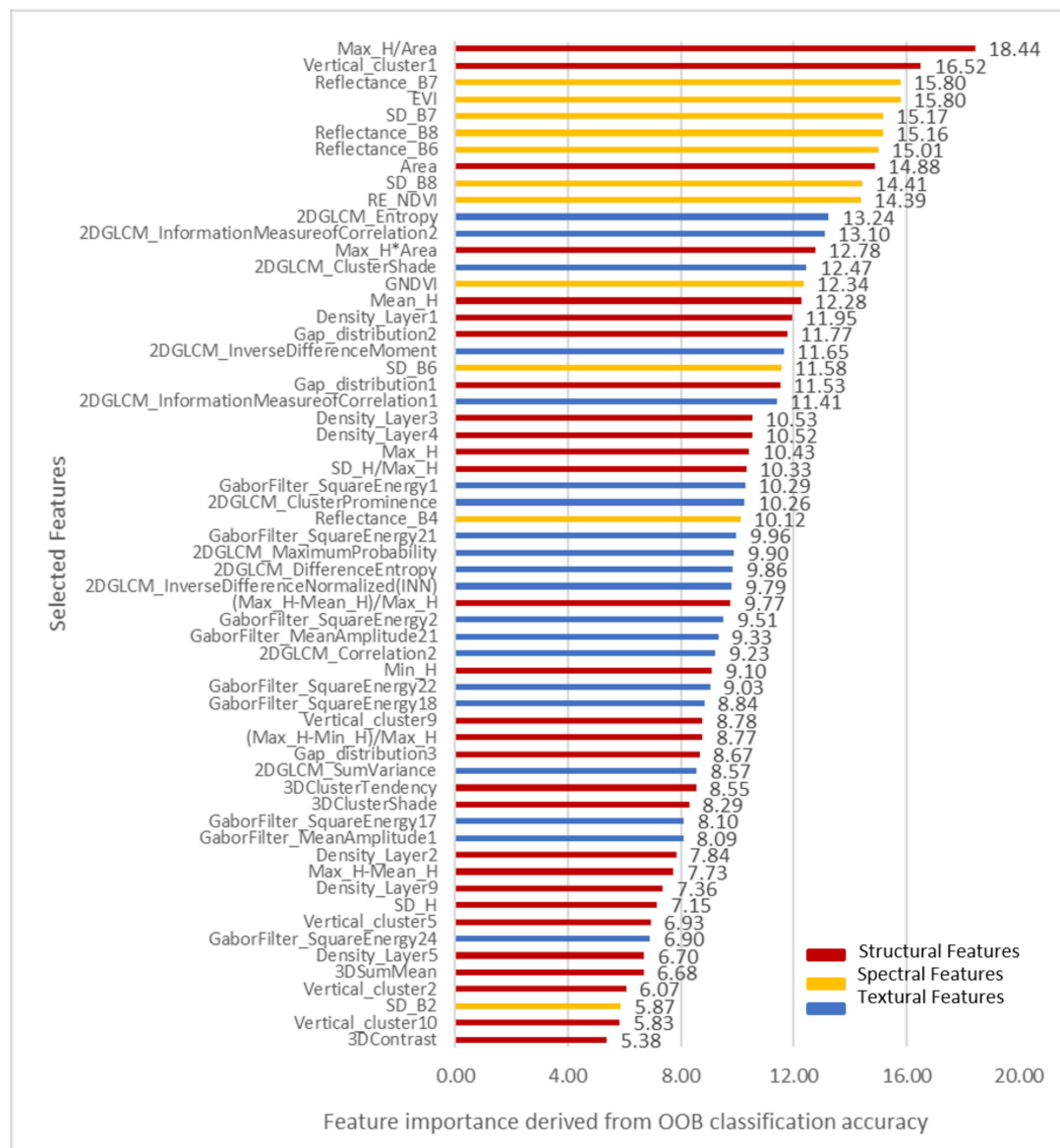


Figure 6. The importance ranking of the selected features based on the MDA in RF.

Generally, when analyzing each feature, most spectral features demonstrated a remarkable capacity to discriminate tree species, followed by structural and textural features. In addition, reflectance information at band7-NIR1 (770–895 nm), band8-NIR2 (860–1040 nm), and band6-red edge (705–745 nm) indicated greater discriminatory power than vegetation indices, except for EVI, which is consistent with related studies [4,52,53]. It is worth noting that advanced structural features held the potential for increasing the discriminatory power for remote sensing-assisted tree species classification. They accounted for 48% of the final feature subset with a wide range of importance scores. As mentioned previously, classification using different combinations of features (Table 4) was carried out to analyze the relative contribution of each type of feature group in the species classification. The results are shown in Table 7. SVM and RF delivered accordant classification results, and thus, the discussion in this section focuses on the results from SVM. Classification using the combination of spectral, textural, and structural features exhibited the best classification performance, with the highest overall accuracy of 0.85 and kappa coefficient of 0.81, which outperformed the classification accuracies by using individual feature groups (Table 5) and any other combination schemes in Table 7.

**Table 7.** Comparison of classification results at the feature-level fusion using SVM and RF (entries with a star (\*) indicate the highest classification accuracy for classification models).

Feature Combinations	SVM		RF	
	Accuracy	Kappa	Accuracy	Kappa
Spectral + textural	0.81	0.76	0.80	0.75
Spectral + structural	0.83	0.78	0.81	0.76
Textural + structural	0.82	0.77	0.81	0.76
Spectral + textural + structural	0.85 *	0.81 *	0.83 *	0.78 *

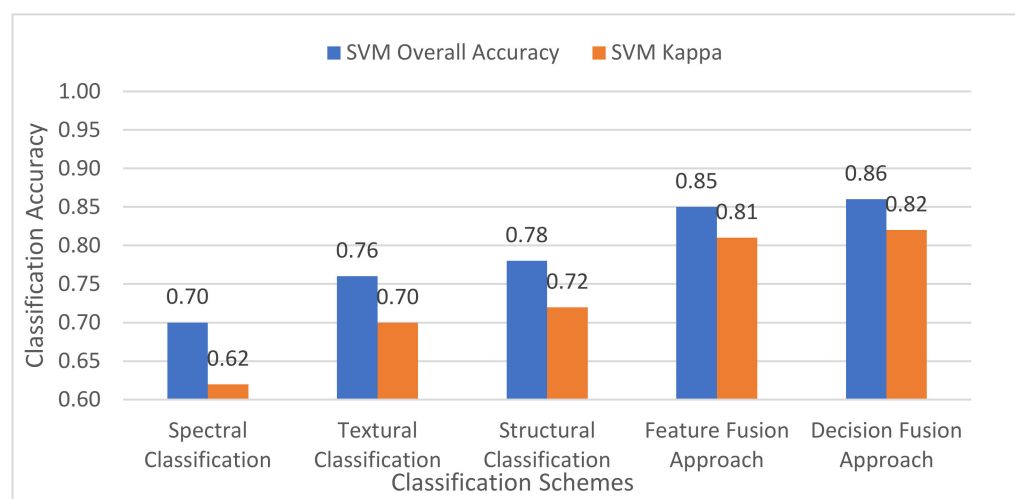
#### 4.3. Classification Using the Decision-Level Fusion Approach

The confusion matrix for the classification using the spectral, textural, and structural features based on decision-level fusion is presented in Table 8. Using the SVM algorithm, the overall accuracy and kappa coefficient were 85.65% and 0.82, respectively. RF produced slightly lower classification accuracies (83.86% and 0.79, respectively). Similarly, the classification results from SVM are discussed in detail. Austrian pine had the highest producer accuracy of 95.83% among the five tree species, showing that most Austrian pines were correctly classified. Norway maple was classified with quite a high producer accuracy (94.64%) and user accuracy (92.98%). The misclassification occurred between Norway maple and honey locust.

To clearly show the differences in classification accuracies of classification schemes investigated in this study, overall accuracies and kappa coefficients obtained by the classifications using spectral, textural, and structural features individually (Table 5) and in combination for the feature-level (Table 7) and decision-level (Table 8) fusion with SVM algorithm are compared in Figure 7. Like the feature-level fusion, the decision fusion approach significantly improved the classification accuracies obtained using individual feature groups. The accuracy of classification schemes using spectral, textural, and structural features increased by 8% to 15% when using the decision fusion method. Although the decision fusion approach achieved the highest overall accuracy and kappa coefficient, feature-level fusion produced comparable results within a narrow margin.

**Table 8.** Confusion matrix for the decision-level fusion approach using SVM and RF. User accuracy (UA) and producer accuracy (PA) were calculated as percentages.

		Actual Tree Species					
	Tree Species	Norway Maple	Honey Locust	Austrian Pine	Blue Spruce	White Spruce	UA (%)
Predicted tree species	SVM_RBF						
	Norway maple	53	4				92.98
	Honey locust	3	46	2	3		85.18
	Austrian pine		2	46	1	2	90.19
	Blue spruce				26	8	76.47
	White spruce		2		5	20	74.07
	PA (%)	94.64	85.19	95.83	74.29	66.67	
	Overall accuracy	85.65%					
	Kappa coefficient	0.82					
	RF						
	Norway maple	50	3				94.34
	Honey locust	5	46	2	1		85.19
	Austrian pine	1	3	46	1	3	85.19
	Blue spruce				27	9	75.00
	White spruce		2		6	18	69.23
	PA (%)	89.29	85.19	95.83	77.14	60.00	
	Overall accuracy	83.86%					
	Kappa coefficient	0.79					

**Figure 7.** Comparison of classification accuracies achieved by individual feature-group-based classification schemes and the feature and decision fusion approaches.

## 5. Discussion

### 5.1. Contribution of Structural and Textural Features

The contribution of structural and textural features to the accurate classifications was evaluated based on the feature importance analysis, which quantifies the discriminative power of these features and was further verified by the classification accuracies of SVM and RF using the testing dataset. The results from this study demonstrate that the structural



feature group (STF in Table 6) that combined CHM and 3D-point-cloud-based features (STF\_CHM and STF\_3D) could effectively distinguish Norway maple, honey locust, and Austrian pine better than the spectral and textural feature groups. On the other hand, the textural feature group (TF) produced the highest F1-score when separating blue spruce and white spruce. The primary reason is that these coniferous trees' natural crown cross-sectional area and crown volume significantly impacted the effective extraction of spectral features and structural features. Finally, it was worth noting that among the five tree species tested, Norway maple showed the highest classification accuracies using each feature group in Table 6. This mainly resulted from its physical characteristics, such as wide-spreading crowns and dense foliage, compared with other tree species with narrow or sparse foliage. Nevertheless, the narrow crowns and needle leaves of blue spruce and white spruce limit the extraction of expressive spectral and structural features, which resulted in severe misclassification between these two tree species under the spruce genus. In related studies, LiDAR-derived structural features were also reported to have a robust discriminatory capacity for the identification of coniferous tree species with narrow crowns, for example, features measuring the distributions of laser points along with the vertical profile [7,8], statistical analysis of height information of laser points for individual tree crowns [54], and the point density of horizontal layers at particular tree heights [20]. However, the spectral feature group (SF) contributed moderately to classifying coniferous and broadleaf tree species, as indicated by its F1-score and overall accuracy ranking in the middle.

For the contribution from advanced textural and structural features, according to the overall accuracy of classification models considering all five tree species, it could be concluded that STF\_3D contributed more than STF\_CHM, with an increase of the overall accuracy of 3%. The tree-height-related features alone were not able to effectively separate the tree species of interest in the study area. Caution should be taken since for most cities, different species might be planted at different times, and thus, the height variations might result from the age difference, though this was not the case for the study area. Except for STF, STF\_3D indicated the most significant discriminatory power in the classifications. In contrast, the Gabor filter feature group (TF\_GABOR) contributed the least to the tree species classification. The statistical GLCM feature group (TF\_GLCM) slightly improved the classification performance using spectral features that consisted of commonly used multispectral signatures and vegetation indices. Although the discrimination powers of TF\_GLCM and TF\_GABOR varied greatly regarding the classification accuracies, combining them as the textural feature group (TF) improved the overall accuracy to 0.74, which was higher than SF by 4% and comparable to STF\_3D. Furthermore, STF\_CHM exhibited the same discrimination power as the statistical TF\_GLCM, with an overall accuracy of 0.71.

The results of the feature importance ranking and classification accuracy of multisource-derived feature groups demonstrated that the combination of multisource data should be an optimal approach to improving the tree species classification accuracy. This was also verified by the feature importance ranking, which indicated that features with high importance scores for tree species classification were extracted from different feature groups and complementary. Although the LiDAR-derived structural feature group resulted in a promising overall accuracy of 0.8 when using the SVM, the potential for investigating the effective and efficient way to integrate the multiple feature groups confirmed the motivation of the current study of feature-level and decision-level fusion approaches.

## 5.2. Comparison of Feature-Level Fusion and Decision-Level Fusion

In this study, comparable classification accuracies were obtained using feature-level and decision-level fusion. This result is in accordance with recent studies on the decision fusion approach to tree species classification conducted by Aval et al. [16] and Stavrakoudis et al. [18]. On the other hand, Hu et al. [29] provided a mechanism to consider the uncertainties of multisource data in decision-level fusion and obtained better classification accuracies than feature-level fusion using SVM. Aval et al. indicated that the feature-level fusion decreased the performance of hyperspectral visible near-infrared (VNIR) image-based

classification [16]. It is worth mentioning that their study might enhance the importance of the feature selection implemented before feature-level fusion in the current study.

The misclassification analysis was conducted based on the classification results from both feature-level and decision-level fusion to show the advantages of the decision fusion approach. Forty tree samples were misclassified, considering the classification results from both the feature-level and decision-level fusion. Among them, 25 trees were misclassified by both methods, eight by feature-level fusion only and seven by decision fusion approach only. The misclassification mainly occurred among honey locusts, white spruces, and blue spruces, especially between blue and white spruces.

The margin of a tree sample is the probability of the actual class minus the maximum value of the other classes, and the size of the margin is a measure of the degree of confidence in the classification results. Among 40 misclassified samples, many misclassification results were caused by the narrow margins, indicating the uncertainties and conflicts in the classification results. Table 9 presents the classification results of a white spruce tree with the ID of 705, which was correctly classified with a marginal difference between the voting probabilities of only 0.05 compared with the probability of the blue spruce. In contrast, decision-level fusion gave rise to evidence of misclassification as blue spruce with a more significant margin of 0.36, i.e., 0.68 vs. 0.32 for the probabilities. On the other hand, Table 10 shows a positive example of the decision fusion approach. The honey locust tree with the ID of 311 was correctly classified using feature- and decision-level fusion approaches. It is worth noting that the feature-level fusion successfully classified it with an even negligible margin of 0.01 between the voting probabilities of 0.49 vs. 0.48 (Austrian pine). The decision fusion approach enhanced the evidence of the correct classification with a slightly larger margin of 0.04, i.e., 0.52 vs. 0.48 for the probabilities.

**Table 9.** Classification results indicating the posterior probabilities that a tree sample (white spruce) belonged to each of the five candidate tree species in the SVM models (entries highlighted in bold are predicted tree species based on the maximum probabilities for different classification schemes).

Tree ID: 705		SVM-Based Posterior Probabilities			
Classification Schemes	Norway Maple	Honey Locust	Austrian Pine	Blue Spruce	White Spruce
Spectral feature	0.14	<b>0.73</b>	0.01	0.07	0.05
Textural feature	0.00	0.01	0.00	<b>0.64</b>	0.35
Structural feature	0.00	0.00	0.00	0.47	<b>0.53</b>
<b>Feature-level fusion</b>	0.00	0.00	0.00	0.47	<b>0.52</b>
<b>Decision-level fusion</b>	0.00	0.00	0.00	<b>0.68</b>	0.32

**Table 10.** Classification results indicating the posterior probabilities that a tree sample (honey locust) belonged to each of the five candidate tree species in the SVM models (entries highlighted in bold are predicted tree species based on the maximum probabilities for different classification schemes).

Tree ID: 311		SVM-Based Posterior Probabilities			
Classification Schemes	Norway Maple	Honey Locust	Austrian Pine	Blue Spruce	White Spruce
Spectral feature	0.06	<b>0.51</b>	0.08	0.27	0.08
Textural feature	0.00	0.14	<b>0.85</b>	0.00	0.00
Structural feature	0.06	<b>0.47</b>	0.46	0.00	0.00
<b>Feature-level fusion</b>	0.03	<b>0.49</b>	0.48	0.00	0.00
<b>Decision-level fusion</b>	0.00	<b>0.52</b>	0.48	0.00	0.00

Even though feature-level fusion delivered successful classifications for these tree samples, the confidence in the classification accuracy was pretty weak, reducing the reliability of the classification results. The predicted tree species was ultimately assigned to a given tree sample based on the maximum probability for voting, even though the

value might not be significantly different between two different species, which made the classification method not robust, especially in the presence of noise inherently added by the sensors and image processing techniques. Furthermore, as shown in Tables 9 and 10, the information provided by individual feature groups is often imprecise and uncertain due to the inherent conflicts between remote sensing data sources. For example, the white spruce tree was misclassified as a honey locust with a probability of 0.73 and blue spruce with 0.64 using spectral and textural features, respectively. As discussed in previous sections, it was indisputable that multisource features are complementary because sensors measure different physical properties of individual tree canopies. The DST-based decision fusion approach in this study provided an effective means to combine the evidence measures from complementary multisource data and produce satisfactory results. However, the current method did not weigh the importance of different evidence pieces for identifying specific tree species. Hence, advanced decision rules for the decision fusion approach should be further investigated to represent tree crowns while adequately reducing imprecision and uncertainty.

### 5.3. Limitation of the Current Study and Future Works

The number of species types investigated was limited to five in the current study. A multitude of other tree species, such as bur oak (*Quercus macrocarpa*), sugar maple (*Acer saccharum*), white cedar (*Thuja occidentalis*), and sand basswood (*Tilia americana*), dominate or co-dominate urban forests of Toronto or other neighboring cities. Including a wide range of tree species over larger areas in future research would be of interest. Moreover, the classification results of a few trees have high uncertainties, as reflected by the posterior probabilities generated by feature and decision fusion approaches. Trees near roads and pathways are usually subject to salt, landscaping, and human interference more than distant ones, increasing the uncertainty for tree species classification. Tree species classes, including a two-species or a three-species compound class, and even an “unknown” class may reduce the confusion errors in other species classes, likely improving the classification performance.

Complementary features from multisource remote sensing data were shown to be beneficial for distinguishing tree species. The features with significant discretionary power can be applied to other detection/classification problems to improve the classification performance, such as 3D structural features and some rarely used GLCM textural measures. However, the severe misclassification among specific tree species, such as blue and white spruce, in this study indicated that additional information is needed. Very high spatial resolution imagery acquired by unmanned aerial vehicles (UAVs) and very high-density LiDAR data are capable of deriving detailed spatial and structural features to improve the classification accuracy of confusing tree species under the same genus, such as spruce.

The DST-based decision fusion approach effectively combined the measures of evidence (posterior probability masses in this study) from multisource remote sensing data. The workflow of the DST-based decision fusion approach deployed with posterior SVM and RF presented in the current study can be applied to multisource-data-based classification tasks. When significant conflicts and incompatibilities exist among the evidence, the Dempster rule combining the available evidence may result in counter-intuition decisions [29]. The current method examined the uncertainties in classification results caused by conflicts among feature groups through the misclassification analysis; however, the contribution of different pieces of evidence from multisource to specific tree species was not weighed. Based on the current study, the decision fusion approach employing alternative combination rules or comprehensive solutions can be further explored for future considerations to deal with classification uncertainties due to the conflicting information from multisource remotely sensed data.

## 6. Conclusions

This study investigated fusion approaches to improving the accuracy of tree species classification of Norway maple (*Acer platanoides*), honey locust (*Gleditsia triacanthos*), Austrian pine (*Pinus nigra*), white spruce (*Picea glauca*), and blue spruce (*Picea pungens*) using WV-2 multispectral imagery (MSI), high-resolution PAN, and LiDAR data. Advanced textural and structural features were first extracted from high-resolution PAN and LiDAR data in addition to spectral features derived from MSI. The contribution of feature groups from multisource data to the accurate classifications was comprehensively evaluated based on the feature importance analysis, which quantified the discriminative power of these features and was further verified by the classification accuracies of SVM and RF using the testing dataset. The feature selection results demonstrated the complementarity of structural, textural, and spectral features in tree species classification. Spectral and structural features were found to be more important in the feature importance ranking compared with textural features. Notably, despite their limited usage in the literature, GLCM-based textural features, such as cluster shade and inverse difference moment, showed significant importance in tree species classification. The classification results from SVM and RF further indicated the importance of structural features for identifying tree species with widespread and dense tree crowns, such as Norway maple and honey locust. On the other hand, textural features improved the classification of coniferous tree species with sparse needle leaves and small crowns, such as blue spruce and white spruce. Overall, the structural feature group demonstrated the most significant discriminatory power in classifying the five tree species, followed by the textural feature group. Although spectral features ranked high in feature importance, they yielded the lowest overall classification accuracy. Notably, GLCM-based textural features contributed more to the classification than Gabor-filter-derived features, while the 3D-point-distribution-related structural features were more important than the features derived from CHM.

A decision fusion framework based on the Dempster–Shafer theory (DST) was developed to enhance individual tree species classification. The decision fusion approach considered the uncertainty measures by quantifying the class probabilities from individual feature groups from the complementary multisource data instead of aggregating features directly with feature-level fusion. When evaluated on an independent testing dataset, both feature-level fusion and decision fusion approaches significantly improved the tree species classification compared with the uses of MSI (0.7), PAN (0.74), or LiDAR (0.8) alone. The decision fusion approach achieved the best overall accuracies (0.86 for SVM and 0.84 for RF) and kappa coefficients (0.82 for SVM and 0.79 for RF) and slightly outperformed the feature-level fusion when combining the MSI, PAN, and LiDAR data. The misclassification analysis of the decision fusion approach and feature-level fusion results was also conducted to provide insights into further improvement. The decision fusion approach based on DST provided an open perspective for individual tree species classification using multisource remotely sensed data, which holds the potential to continually improve the performance, along with comprehensive decision rules and advanced features from developed data in the future.

**Author Contributions:** Conceptualization, Q.L., B.H., J.S. and H.L.; methodology, Q.L., B.H., J.S. and H.L.; implementation, Q.L.; investigation, Q.L., B.H., J.S. and H.L.; data curation, Q.L. and B.H.; writing—first draft preparation, Q.L.; writing—review and editing, Q.L., B.H., J.S. and H.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada, grant number RGPIN-2021-03624, NSERC and Esri Canada, grant number CRDPJ 490711-2015, and Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA), grant number ND2017-3179.

**Data Availability Statement:** Data sharing is not applicable to this article.



**Acknowledgments:** The authors would like to thank York University Map Library and CSBO for the LiDAR data and tree inventory map of the study area.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Huang, X. Single Tree Urban Inventory Updates. 2019. Available online: <https://hdl.handle.net/1807/93303> (accessed on 1 January 2019).
- Fitzky, A.C.; Sandén, H.; Karl, T.; Fares, S.; Calfapietra, C.; Grote, R.; Saunier, A.; Rewald, B. The interplay between ozone and urban vegetation—BVOC emissions, ozone deposition, and tree ecophysiology. *Front. For. Glob. Chang.* **2019**, *2*, 50. [\[CrossRef\]](#)
- Felton, A.; Löfroth, T.; Angelstam, P.; Gustafsson, L.; Hjältén, J.; Felton, A.M.; Simonsson, P.; Dahlberg, A.; Lindbladh, M.; Svensson, J. Keeping pace with forestry: Multi-scale conservation in a changing production forest matrix. *Ambio* **2020**, *49*, 1050–1064. [\[CrossRef\]](#) [\[PubMed\]](#)
- 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. [\[CrossRef\]](#)
- Wang, K.; Wang, T.; Liu, X. A Review: Individual Tree Species Classification Using Integrated Airborne LiDAR and Optical Imagery with a Focus on the Urban Environment. *Forests* **2018**, *10*, 1. [\[CrossRef\]](#)
- Mojaddadi Rizeei, H.; Pradhan, B.; Saharkhiz, M.A. Urban object extraction using Dempster Shafer feature-based image analysis from worldview-3 satellite imagery. *Int. J. Remote Sens.* **2019**, *40*, 1092–1119. [\[CrossRef\]](#)
- Li, J.; Hu, B.; Noland, T.L. Classification of tree species based on structural features derived from high density LiDAR data. *Agric. For. Meteorol.* **2013**, *171–172*, 104–114. [\[CrossRef\]](#)
- Li, J.; Hu, B.; Woods, M. A Two-Level Approach for Species Identification of Coniferous Trees in Central Ontario Forests Based on Multispectral Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 1487–1497. [\[CrossRef\]](#)
- Aguilar, M.A.; Bianconi, F.; Aguilar, F.J.; Fernández, I. Object-based greenhouse classification from GeoEye-1 and WorldView-2 stereo imagery. *Remote Sens.* **2014**, *6*, 3554–3582. [\[CrossRef\]](#)
- Gillespie, T.W.; de Goede, J.; Aguilar, L.; Jenerette, G.D.; Fricker, G.A.; Avolio, M.L.; Pincetl, S.; Johnston, T.; Clarke, L.W.; Pataki, D.E. Predicting tree species richness in urban forests. *Urban Ecosyst.* **2017**, *20*, 839–849. [\[CrossRef\]](#)
- Hartling, S.; Sagan, V.; Sidike, P.; Maimaitijiang, M.; Carron, J. Urban tree species classification using a worldview-2/3 and LiDAR data fusion approach and deep learning. *Sensors* **2019**, *19*, 1284. [\[CrossRef\]](#)
- Hartling, S.; Sagan, V.; Maimaitijiang, M. Urban tree species classification using UAV-based multi-sensor data fusion and machine learning. *GIScience Remote Sens.* **2021**, *58*, 1250–1275. [\[CrossRef\]](#)
- Pu, R.; Landry, S. A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species. *Remote Sens. Environ.* **2012**, *124*, 516–533. [\[CrossRef\]](#)
- Franklin, S.E.; Maudie, A.J.; Lavigne, M.B. Using spatial co-occurrence texture to increase forest structure and species composition classification accuracy. *Photogramm. Eng. Remote Sens.* **2001**, *67*, 849–856.
- Haralick, R.M.; Shanmugam, K.; Dinstein, I.H. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *6*, 610–621. [\[CrossRef\]](#)
- Aval, J.; Fabre, S.; Zenou, E.; Sheeren, D.; Fauvel, M.; Briottet, X. Object-based fusion for urban tree species classification from hyperspectral, panchromatic and nDSM data. *Int. J. Remote Sens.* **2019**, *40*, 5339–5365. [\[CrossRef\]](#)
- Ferreira, M.P.; Wagner, F.H.; Aragão, L.E.O.C.; Shimabukuro, Y.E.; de Souza Filho, C.R. Tree species classification in tropical forests using visible to shortwave infrared WorldView-3 images and texture analysis. *ISPRS J. Photogramm. Remote Sens.* **2019**, *149*, 119–131. [\[CrossRef\]](#)
- Stavrakoudis, D.G.; Dragozi, E.; Gitas, I.Z.; Karydas, C.G. Decision fusion based on hyperspectral and multispectral satellite imagery for accurate forest species mapping. *Remote Sens.* **2014**, *6*, 6897–6928. [\[CrossRef\]](#)
- Yang, G.; Zhao, Y.; Li, B.; Ma, Y.; Li, R.; Jing, J.; Dian, Y. Tree species classification by employing multiple features acquired from integrated sensors. *J. Sens.* **2019**, *2019*, 3247946. [\[CrossRef\]](#)
- Lin, Y.; Hyypä, J. A comprehensive but efficient framework of proposing and validating feature parameters from airborne LiDAR data for tree species classification. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *46*, 45–55. [\[CrossRef\]](#)
- 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. [\[CrossRef\]](#)
- Alonzo, M.; Bookhagen, B.; Roberts, D.A. Urban tree species mapping using hyperspectral and lidar data fusion. *Remote Sens. Environ.* **2014**, *148*, 70–83. [\[CrossRef\]](#)
- Dalponte, M.; Bruzzone, L.; Gianelle, D. Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. *Remote Sens. Environ.* **2012**, *123*, 258–270. [\[CrossRef\]](#)
- Dalponte, M.; Frizzera, L.; Gianelle, D. Individual tree crown delineation and tree species classification with hyperspectral and LiDAR data. *PeerJ* **2019**, *2019*, e6227. [\[CrossRef\]](#) [\[PubMed\]](#)
- Liu, L.; Coops, N.C.; Aven, N.W.; Pang, Y. Mapping urban tree species using integrated airborne hyperspectral and LiDAR remote sensing data. *Remote Sens. Environ.* **2017**, *200*, 170–182. [\[CrossRef\]](#)

26. Shojanoori, R.; Shafri, H.Z.M. Review on the use of remote sensing for urban forest monitoring. *Arboric. Urban For.* **2016**, *42*, 400–417. [[CrossRef](#)]
27. Li, H.; Hu, B.; Li, Q.; Jing, L. CNN-based individual tree species classification using high-resolution satellite imagery and airborne LiDAR data. *Forests* **2021**, *12*, 1697. [[CrossRef](#)]
28. Maxwell, A.E.; Warner, T.A.; Fang, F. Implementation of machine-learning classification in remote sensing: An applied review. *Int. J. Remote Sens.* **2018**, *39*, 2784–2817. [[CrossRef](#)]
29. Hu, B.; Li, Q.; Hall, G.B. A decision-level fusion approach to tree species classification from multi-source remotely sensed data. *ISPRS Open J. Photogramm. Remote Sens.* **2021**, *1*, 100002. [[CrossRef](#)]
30. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring vegetation systems in the Great Plains with ERTS. In Proceedings of the Third Earth Resources Technology Satellite-1 Symposium, NASA SP-351, Greenbelt, MD, USA, 10–14 December 1974; pp. 301–317.
31. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [[CrossRef](#)]
32. Gitelson, A.A.; Merzlyak, M.N. Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. *J. Photochem. Photobiol. B Biol.* **1994**, *22*, 247–252. [[CrossRef](#)]
33. Rondeaux, G.; Steven, M.; Baret, F. Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* **1996**, *55*, 95–107. [[CrossRef](#)]
34. Wu, C.; Gonsamo, A.; Gough, C.M.; Chen, J.M.; Xu, S. Modeling growing season phenology in North American forests using seasonal mean vegetation indices from MODIS. *Remote Sens. Environ.* **2014**, *147*, 79–88. [[CrossRef](#)]
35. Soh, L.K.; Tsatsoulis, C. Texture analysis of sar sea ice imagery using gray level co-occurrence matrices. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 780–795. [[CrossRef](#)]
36. Clausi, D.A. An analysis of co-occurrence texture statistics as a function of grey level quantization. *Can. J. Remote Sens.* **2002**, *28*, 45–62. [[CrossRef](#)]
37. Zheng, D.; Zhao, Y.; Wang, J. Features extraction using a Gabor filter family. In Proceedings of the Sixth IASTED International Conference on Signal and Image Processing, Wuxi, China, 8–10 July 2004; pp. 139–144.
38. Heiskanen, J.; Korhonen, L.; Hietanen, J.; Pellikka, P.K.E. Use of airborne lidar for estimating canopy gap fraction and leaf area index of tropical montane forests. *Int. J. Remote Sens.* **2015**, *36*, 2569–2583. [[CrossRef](#)]
39. 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](#)]
40. Guyon, I.; Elisseeff, A. An introduction to variable and feature selection. *J. Mach. Learn. Res.* **2003**, *3*, 1157–1182.
41. Demarchi, L.; Kania, A.; Ciekowski, W.; Piórkowski, H.; Oświecimska-Piasko, Z.; Chormański, J. Recursive feature elimination and random forest classification of natura 2000 grasslands in lowland river valleys of poland based on airborne hyperspectral and LiDAR data fusion. *Remote Sens.* **2020**, *12*, 1842. [[CrossRef](#)]
42. Georganos, S.; Grippa, T.; Vanhuyse, S.; Lennert, M.; Shimoni, M.; Kalogirou, S.; Wolff, E. Less is more: Optimizing classification performance through feature selection in a very-high-resolution remote sensing object-based urban application. *GIScience Remote Sens.* **2018**, *55*, 221–242. [[CrossRef](#)]
43. Gregorutti, B.; Michel, B.; Saint-Pierre, P. Correlation and variable importance in random forests. *Stat. Comput.* **2017**, *27*, 659–678. [[CrossRef](#)]
44. Kuhn, M. Building Predictive Models in R Using the caret Package. *J. Stat. Softw.* **2008**, *28*, 1–26. [[CrossRef](#)]
45. RStudio. *Open Source & Professional Software for Data Science Teams—RStudio*; Version 1.2.1335; RStudio Team: Vienna, Austria, 2018.
46. Meyer, D.; Dimitriadou, E.; Hornik, K.; Weingessel, A.; Leisch, F. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien, R Package Version 1.7-3. 2019. Available online: <https://CRAN.R-project.org/package=e1071> (accessed on 2 July 2023).
47. Liaw, A.; Wiener, M. Classification and Regression by randomForest. *R News* **2002**, *2*, 18–22.
48. Platt, J. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Adv. Large Margin Classif.* **1999**, *10*, 61–74.
49. Wu, T.-F.; Lin, C.-J.; Weng, R.C. Probability estimates for multi-class classification by pairwise coupling. *J. Mach. Learn. Res.* **2004**, *5*, 975–1005.
50. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
51. Stehman, S.V. Selecting and interpreting measures of thematic classification accuracy. *Remote Sens. Environ.* **1997**, *62*, 77–89. [[CrossRef](#)]
52. Clark, M.L.; Roberts, D.A.; Clark, D.B. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sens. Environ.* **2005**, *96*, 375–398. [[CrossRef](#)]

- 
53. Waser, L.T.; Küchler, M.; Jütte, K.; Stampfer, T. Evaluating the potential of worldview-2 data to classify tree species and different levels of ash mortality. *Remote Sens.* **2014**, *6*, 4515–4545. [[CrossRef](#)]
  54. 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](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.