

Article

# Object-Based Urban Tree Species Classification Using Bi-Temporal WorldView-2 and WorldView-3 Images

Dan Li, Yinghai Ke \*, Huili Gong and Xiaojuan Li

Received: 16 October 2015; Accepted: 10 December 2015; Published: 15 December 2015

Academic Editors: Randolph H. Wynne and Prasad S. Thenkabail

Beijing Laboratory of Water Resource Security, State Key Laboratory Incubation Base of Urban Environmental Processes and Digital Simulation, Capital Normal University, 105 West Third Ring Road, Haidian District, Beijing 100048, China; eternalDan@163.com (D.L.); gonghl@263.net (H.G.); xiaojuanli@vip.sina.com (X.L.)

\* Correspondence: yke@cnu.edu.cn; Tel.: +86-10-6898-0798; Fax: + 86-10-6890-2217

**Abstract:** Urban tree species mapping is an important prerequisite to understanding the value of urban vegetation in ecological services. In this study, we explored the potential of bi-temporal WorldView-2 (WV2, acquired on 14 September 2012) and WorldView-3 images (WV3, acquired on 18 October 2014) for identifying five dominant urban tree species with the object-based Support Vector Machine (SVM) and Random Forest (RF) methods. Two study areas in Beijing, China, Capital Normal University (CNU) and Beijing Normal University (BNU), representing the typical urban environment, were evaluated. Three classification schemes—classification based solely on WV2; WV3; and bi-temporal WV2 and WV3 images—were examined. Our study showed that the single-date image did not produce satisfying classification results as both producer and user accuracies of tree species were relatively low (44.7%–82.5%), whereas those derived from bi-temporal images were on average 10.7% higher. In addition, the overall accuracy increased substantially (9.7%–20.2% for the CNU area and 4.7%–12% for BNU). A thorough analysis concluded that near-infrared 2, red-edge and green bands are always more important than the other bands to classification, and spectral features always contribute more than textural features. Our results also showed that the scattered distribution of trees and a more complex surrounding environment reduced classification accuracy. Comparisons between SVM and RF classifiers suggested that SVM is more effective for urban tree species classification as it outperforms RF when working with a smaller amount and imbalanced distribution of samples.

**Keywords:** urban tree species classification; bi-temporal images; object-based method; support vector machine; random forest

## 1. Introduction

All types of urban vegetation, especially trees, play an important role in the urban ecosystem. Trees have valuable eco-service functions such as above-ground carbon storage, urban temperature mediation, air quality improvement and urban flood risk alleviation [1–4]. Acquiring timely and detailed information on spatial distribution and structural characteristics of trees within urban areas is critical for a better understanding of their eco-service values, and subsequently for developing strategies for sustainable urban development. Traditional methods for urban tree species mapping involves random sampling in various urban districts and field investigation of tree species within each sample plot. An alternative method is visual interpretation of aerial photographs. Both are time and labor consuming. Remote sensing techniques, especially high spatial resolution satellite imagery, provide a great opportunity for timely tree species mapping at a considerably lower cost. High spatial resolution satellite images, such as those acquired by IKONOS, QuickBird, WorldView-2 (WV2) or WorldView-3 (WV3) satellites, have been widely used for tree species identification in

forested areas [5–10]. Compared to the traditional four-band IKONOS and QuickBird, the WV2 satellite (DigitalGlobe Inc.) launched in 2009 has better spectral (eight bands) and spatial (0.5–2 m) resolution. The four additional bands (coastal, yellow, red-edge and near-infrared2 bands) are considered to be more capable of detailed vegetation identification [6,7,11]. The WV3 satellite was launched in August 2014. In addition to the eight bands it shares with WV2, it also has a 16-band mode which could provide an additional eight short-wave infrared (SWIR) bands that may further benefit vegetation analysis.

Most of the existing tree species identification applications have focused on forested areas, and only a few studies have evaluated the capability of high spatial resolution imagery for urban tree species classification. One of the representative studies was conducted by Pu *et al.* [12], which compared the capability of IKONOS and WV2 the tree species identification of seven types of trees in the city of Tampa, Florida, USA. The seven tree species included sand live oak (*Quercus geminata*), laurel oak (*Q. laurifolia*), live oak (*Q. virginiana*), pine (species group), palm (species group), camphor (*Cinnamomum camphora*), and magnolia (*Magnolia grandiflora*). Both the Linear Discriminant Analysis (LDA) algorithm and the Classification and Regression Tree (CART) methods were examined. The results showed that WV2 achieved better overall classification accuracy (around 55% when using CART) than IKONOS. Compared to forested areas where tree crowns are usually densely distributed and the surrounding environment is relatively homogeneous, urban areas feature complex and heterogeneous land covers and thus face specific challenges in tree species classification. First, shadows casted by high-rise objects lead to reduction or total loss of spectral information of shaded tree species so that they are difficult to be classified or interpreted. Li *et al.* [13] reported that shadows considerably affected the capability of tree species discrimination, even when the shadow area in the image was recovered using the linear-correlation correction method (the overall accuracy decreased by over 5% when trees under shadows were considered). Second, trees in an urban area are distributed in a scattered fashion, and thus spectral characteristics of tree crowns are easily affected by surrounding non-tree objects. Third, various background materials under tree crowns—such as soil, asphalt or cement—could influence spectral information of crowns observed by a high resolution sensor. These challenges may affect species classification accuracy. As reported by Pu *et al.* [12], the average producer accuracy and user accuracy of the seven tree species classification were only around 67% and 52%, respectively, which indicates lower classification accuracy in forested areas [6,7].

The aforementioned studies indicated that single-date imagery may not suffice for urban tree species classification. Multi-temporal images that represent different vegetation phenologies may assist in better classification. Tigges *et al.* [14] explored the identification of eight tree genera in an urban forest in Berlin using RapidEye imagery (6.5 m) acquired in spring, summer and autumn. The eight tree genera are: pine (*Pinus sylvestris*), chestnut (*Aesculus hypocastanum*), plane (*Platanus hispanica*), lime (*Tilia cordata*, *Tilia × vulgaris*, *Tilia platyphyllos*), maple (*Acer campestre*, *Acer platanoides*, *Acer* sp.), poplar (*Populus nigra*, *Populus alba*), beech (*Fagus sylvatica*) and oak (*Quercus robur*, *Quercus rubra*, *Quercus* sp.) In Berlin, urban vegetation covers over 40% of the total urban area, and more than 290 km<sup>2</sup> are urban forest. This study utilized pixel-based classification based on the Support Vector Machine (SVM) approach and found that a series of multi-temporal images (spring, summer, autumn) were necessary for tree genera classification; the overall accuracy reached 85.5%.

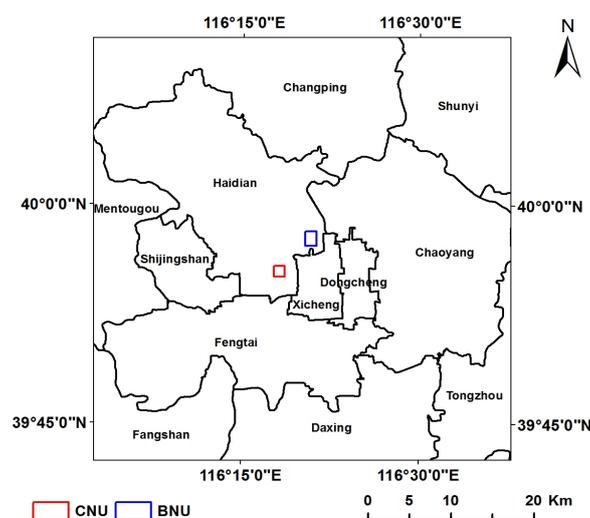
In this study, we aimed to evaluate bi-temporal high spatial resolution imagery for tree species classification in a more complex urban environment, exemplified by two study areas in Beijing, China which represent typical tree distribution in big cities. Unlike the study area in Berlin, urban tree coverage in inner-city Beijing makes up only 25.2% [15]. Trees are mainly isolated and distributed along roadsides, as well as around residential areas, school campuses, or within parks; trees are either individually distributed or clustered in groups of several trees. Tree crown sizes vary and the crown diameter is normally within 3–8 meters. In these cases, the spatial resolution of RapidEye imagery

is likely too coarse to identify crown spectral and texture characteristics. Higher spatial resolution imagery needs to be utilized to conduct tree classification at the species level. Moreover, a pixel-based classification method such as that used for RapidEye imagery cannot be used for species classification. In this study, we examined higher spatial resolution imagery, *i.e.*, WV2 and WV3 images, for bi-temporal analysis with an object-based method. At each of the two study sites, three classification schemes, including classification based on late summer WV2 images, high autumn WV3 images and both WV2 and WV3 images, were conducted to examine the effects of bi-temporal imagery on urban tree species classification. Two machine learning algorithms, SVM and Random Forest (RF) were used for object-based classification. Comparisons between the two study sites were also performed in order to analyze the impact of a complex urban environment on tree species discrimination.

## 2. Study Area and Datasets

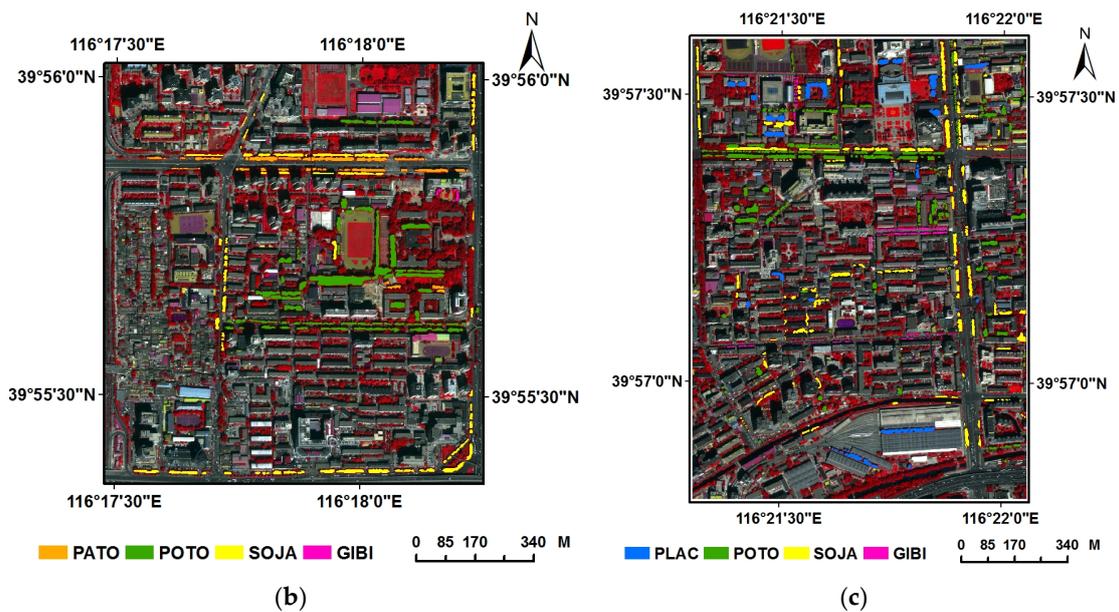
### 2.1. Study Area

Two study sites located in the Haidian district of Beijing, China were explored in this study. One site covers the campus of Capital Normal University (CNU) and residential areas around the campus. The other site is located around Beijing Normal University (BNU). The areas of the study sites are 1.4 km<sup>2</sup> and 1.6 km<sup>2</sup>, respectively (Figure 1). Dominant tree species at CNU include the empress tree (*Paulownia tomentosa*, PATO), Chinese white poplar (*Populus tomentosa* Carrière, POTO), Chinese scholar tree (*Sophora japonica*, SOJA), and ginkgo (*Ginkgo biloba*, GIBI). At the BNU site, the dominant tree species are London plane tree (*Platanus acerifolia*, PLAC), POTO, SOJA, and GIBI. At both study sites, the dominant tree species account for over 94% of all tree stems. In addition, these five species are the most common hardwood species in Beijing [16,17]. Minor tree species include cedar (*Cedrus deodara*) and Chinese toon trees (*Toona sinensis*) in the CNU area, and Sabina tibetica (*Sabina tibetica* Kom.) in the BNU area. Trees at both study sites are mainly located within the campus and residential areas and along streets, thereby representing the typical distribution of trees in an urban area. Despite similarities in tree species and location of trees, the two study sites differ in average tree crown size and spatial heterogeneity of tree species distribution. Tree crown diameter varies from 3 m to 8 m at both study sites, while tree crowns at the CNU site are generally larger than those at BNU, with an average crown diameter of 6.7 m compared to 5.5 m at BNU. In addition, same species trees are distributed in a more clustered manner at CNU (Figure 1, Section 2.3).



(a)

Figure 1. Cont.



**Figure 1.** (a) Location of CNU and BNU study sites in Beijing city; (b) False color WV2 image over CNU study site with reference polygons; (c) False color composite WV2 image over BNU study site with reference polygons.

### 2.2. WorldView-2/3 Imagery

Cloud-free WV2 and WV3 images acquired on 14 September 2012 and 18 October 2014, respectively (Table 1), were used in this study. Image acquisition dates were selected based on both data availability and the vegetation phenology period in Beijing. Data for both were required to be collected in leaf-on seasons, from a similar view angle, during different phenological phases, and able to provide cloud-free images over the study area. Although such a pair of images was not available within the same year, it is reasonably certain that the WV2 and WV3 images were acquired in different tree growth phases. In Beijing, mid-September belongs to late summer when trees reach maturity and develop a fully green canopy, while mid to late October is high autumn when most hardwood trees begin leaf coloring and senescence [18,19]. In addition, the temperature in early October 2014 dropped significantly, thus accelerating the process of pigment change and leaves falling. Both WV2 and WV3 datasets are composed of one panchromatic band (450–800 nm) with Ground Sampling Distance (GSD) of 0.5 m, and eight multispectral bands with 2 m GSD including coastal (400–450 nm), blue (450–510 nm), green (510–580 nm), yellow (585–625 nm), red (630–690 nm), red edge (705–745 nm), NIR1 (770–895 nm) and NIR2 (860–1040 nm). Although the WV3 sensor has 16 multispectral bands, the image available over the study areas only provides eightband mode. Both images were geometrically corrected and projected to WGS-84 UTM Zone 50N system.

**Table 1.** WV2 /WV3 imagery and relevant phenological information for the study areas.

Image	Date	Max off Nadir Angle	Phenology Season
WV2	14 September 2012	13.68°	Late summer (29 August to 2 October); Leaves fully developed
WV3	18 October 2014	16.03°	High autumn (18 October to 1 November); Most species start yellowing phase and the leaves then start falling

### 2.3. Reference Data

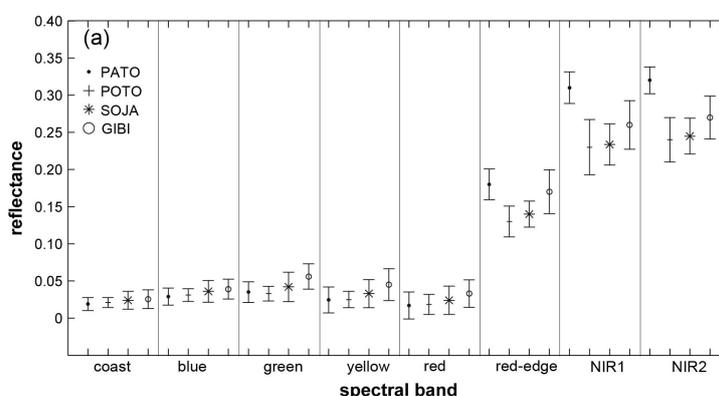
Field investigations were conducted from July to October 2014. Hardcopies of WV2 and WV3 false color composite images were brought to the field to locate tree crowns and identify tree species

in the study areas. A total of 187 polygons with a total area of 0.06 km<sup>2</sup> were manually outlined in the CNU study area, and 564 polygons with a total area of 0.08 km<sup>2</sup> were outlined in the BNU study area (Table 2, Figure 1). Each polygon may consist of either an individual tree crown or a group of adjacent trees of the same species; if a group of trees of the same species cannot be visually separated on the image, one polygon may cover several crowns. Due to more uniformly distributed species at CNU, the number of polygons drawn is less than for BNU. In both areas, POTO is the most dominant species and GIBI has the smallest coverage. Because acquisition dates of WV2 and WV3 images were two years apart, several trees were cut down during city construction and were thus unseen on the WV3 images. A few trees also have slightly larger crown size in 2014. Nevertheless, over 95% of the samples can be identified at the same positions on both WV2 and WV3 images. In order to investigate the distribution of tree species, buffer areas with a 100 m radius were outlined around the centroid of each reference polygon, and the number of tree species within the buffer area was counted. The average number of tree species within the buffer areas is 2.3 at CNU and 3.5 at BNU, meaning that same-species trees at CNU have more clustered distribution than those at BNU.

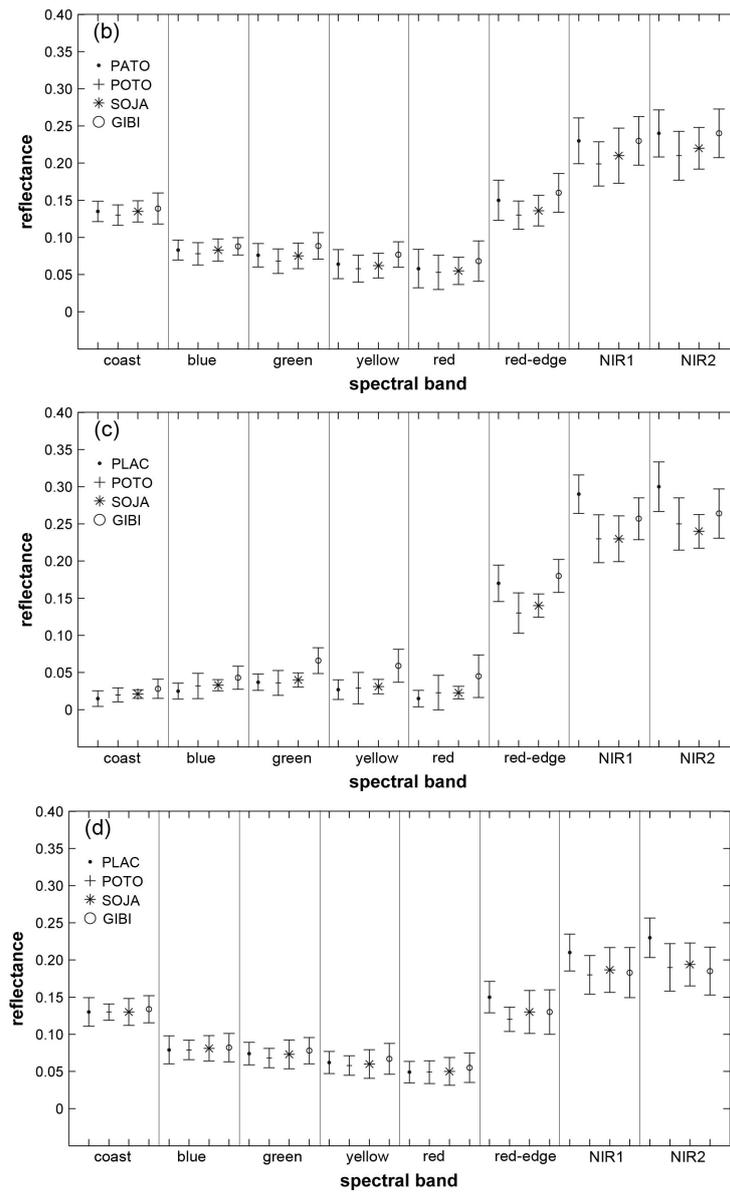
**Table 2.** Distribution and percentage of samples delineated in both areas.

Tree Species	Polygons	Pixel	Area (m <sup>2</sup> )	Area Percentage (%)
<i>(a) CNU Area</i>				
PATO	31	35,752	8938	16.23
POTO	55	111,213	27,803.25	50.49
SOJA	62	70,136	17,534.00	31.84
GIBI	39	3173	793.25	1.44
Sum	187	220,274	55,068.50	100
<i>(b) BNU Area</i>				
PLAC	33	57,443	14,360.75	17.06
POTO	163	132,424	33,106.00	39.33
SOJA	147	125,890	31,472.50	37.39
GIBI	221	20,928	5232.00	6.22
Sum	564	336,685	84,171.25	100

Figure 2 illustrates mean spectral reflectance of the dominant tree species within the delineated reference polygons on WV2 and WV3 images. It is obvious that PATO at CNU and PLAC at BNU have very different spectral characteristics compared to the other species at the red-edge, NIR1 and NIR2 bands, especially on the WV2 image during late summer (Figure 3a,c). GIBI has different reflectance values compared to other species in green, yellow, red, red-edge, NIR1 and NIR2 on both WV2 and WV3 images. Spectral characteristics of POTO and SOJA are very close at all eight bands. From late summer to high autumn, all species have decreasing reflectance in red-edge, NIR1 and NIR2 bands, and increasing reflectance in coastal, blue, green, yellow and red bands. The phenology variation represented in surface reflectance change is expected to help identify tree species.



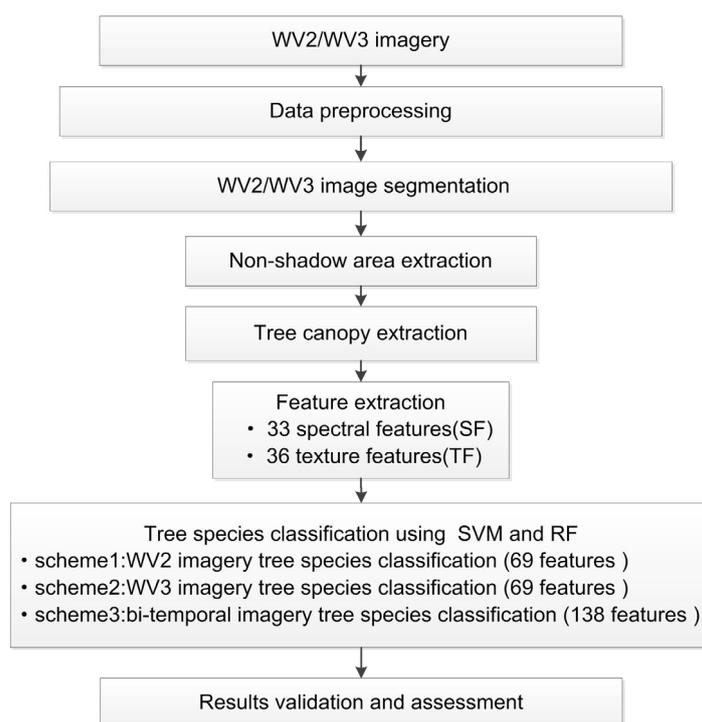
**Figure 2.** Cont.



**Figure 2.** Mean spectral characteristic of the dominant tree species on (a) WV2 image over CNU study area; (b) WV3 image over CNU study area; (c) WV2 over BNU study area; and (d) WV3 over BNU study area.

### 3. Methods

Figure 3 presents the flowchart of the urban tree species classification procedure using an object-based method and machine learning algorithms. The procedure consists of five steps: (1) data preprocessing; (2) images' object generation by image segmentation; (3) tree crown area extraction; (4) feature extraction; and (5) tree species classification using machine learning algorithms SVM and RF. In order to evaluate the effect of bi-temporal images on urban tree species identification, for each study site three classification schemes were tested: classification based solely on the WV2 image, WV3 image, and a combination of WV2 and WV3 images. Accuracy assessment was then performed for each classification scheme. A tree species classification map was finally produced based on the best results.



**Figure 3.** Flowchart of the tree species classification procedure.

### 3.1. Data Preprocessing

Each of the WV2 and WV3 panchromatic and multispectral Digital Number (DN) images was converted to Top of Atmosphere (TOA) radiance based on radiometric calibration parameters [20] and standard correction formula [21]. For each band, surface reflectance was generated using Second Simulation of a Satellite Signal in the Solar Spectrum Vector (6SV) [22] radiative transfer algorithm based on a sensor spectral response function and specified atmospheric condition. A mid-latitude summer climate model was used to specify water vapor and ozone content. Aerosol optical thickness was obtained from MODIS Aerosol product (MOD04\_L2) on the same day. Then the 2 m multispectral WV2 (or WV3) surface reflectance image was fused with 0.5 m panchromatic WV2 (or WV3) surface reflectance image to generate a pan-sharpened 0.5 m WV2 (or WV3) image using the Gramm-Schmidt Spectral Sharpening (GSPS) method with nearest-neighbor resampling [23]. Studies have proven that this pan-sharpening method could reflect the synergic effectiveness of both multispectral and high resolution panchromatic images. It has also proved that this method is spectrally stronger than other sharpening techniques for the fusion of WV2 multispectral bands with the panchromatic band [24]. The GSPS method has been widely applied in land cover classification [25], forest tree species classification [26], urban tree species classification [12], and detection of mineral alteration in a marly limestone formation [27]. In our study, the pan-sharpened image provides clear boundaries of tree crowns, and thus was used for further processing including segmentation and classification.

Both WV2 and WV3 pan-sharpened images were co-registered to eliminate small location displacement of tree crowns caused by differences in image acquisition time and satellite observation angle. The co-registration error, *i.e.*, Root Mean Square Error, was within 0.5 pixel size (0.28 pixel at BNU study area and 0.17 pixel at CNU study area).

### 3.2. Image Segmentation

Multi-resolution segmentation algorithm in Trimble eCognition™ Developer 8.7 software was used to generate image objects from each of the pan-sharpened WV2 and WV3 images.

The multi-resolution segmentation algorithm is a bottom-up region merging technique. Starting from one-pixel objects, larger objects are generated by merging smaller ones with a series of iterative steps [28]. Parameters required as input for the segmentation algorithm include: (1) weight of each input layer; (2) scale parameter; (3) color/shape weight; and (4) compactness/smoothness weight. In this study, all of the eight spectral bands of WV2 or WV3 pan-sharpened imagery were used as input. Previous research has shown that the eight WV2 bands are equally important for urban land cover classification [29]. Furthermore, as is shown in Figure 2, the spectral reflectance of the species are different at each of the eight bands, and it is hard to tell which bands are more important in distinguishing these species. Therefore, all of the eight spectral band layers were assigned the same weight in the segmentation process. The weights of color and compactness were set as 0.8 and 0.5, respectively, in order to balance the difference of spectral/shape heterogeneity between tree species and buildings. For bi-temporal classification, segmented polygons from each of the single-date images were intersected to generate new image objects.

For object-based classification, scale parameter is a key factor because it is closely related to the resultant image object size [30,31]. Under- or over-segmentation both decreases classification accuracy, although significant under-segmentation tends to produce much worse results than over-segmentation [9,32,33]. In this study, we adopted the Bhattacharyya Distance (BD) index method presented by Xun *et al.* [25] and Wang *et al.* [26] in order to determine the best segmentation scale parameter. In statistics, BD is used to measure the similarity between two discrete or continuous probability distributions and the amount of overlap between two statistical samples. In remote sensing, a greater BD value corresponds to greater spectral separation between two distinct classes. The BD index method assumed that the best scale parameter leads to the maximum separation of classes; thus, when the pair-wise BD values reach the highest, the corresponding scale parameter was selected as optimum. Six scale parameters, 100, 110, 120, 130, 140 and 150, were tested (Table 3). For each segmented image, BD index values between every two tree species were calculated based on mean spectral reflectance of each band within the segmented polygons. The scale parameter that results in the overall maximum BD values was selected as optimal scale parameter. BD index equations are as follows:

$$BD(i,j) = 2 \times \left[ 1 - e^{-a(i,j)} \right] \quad (1)$$

$$a(i,j) = \frac{1}{8} [M(i) - M(j)]^T \times A(i,j)^{-1} \times [M(i) - M(j)] + \frac{1}{2} \ln \left\{ \frac{\det A(i,j)}{\sqrt{\det S(i) * \det S(j)}} \right\} \quad (2)$$

$$A(i,j) = \frac{1}{2} (S(i) + S(j)) \quad (3)$$

where  $i, j$  represent class  $i$  and class  $j$ ,  $BD(i,j)$  is the Bhattacharyya Distance between tree species class  $i$  and  $j$ ,  $M(i)$  and  $M(j)$  are the matrices composed of mean reflectance values of all polygons at each of the eight spectral bands.  $S(i)$  and  $S(j)$  are the covariance matrix of  $M(i)$  and  $M(j)$ , respectively;  $A(i,j)$  is half of the sum of  $S(i)$  and  $S(j)$ . The result is a value in a range of 0 to 2, with greater BD values representing greater separability.

**Table 3.** Parameters for pan-sharpened WV2/WV3 image segmentation.

Band Number	Input Layers	Weight	Scale Parameters	Color	Compactness
1	coastal	1			
2	blue	1	100		
3	green	1	110		
4	yellow	1	120		
5	red	1	130	0.8	0.5
6	red-edge	1	140		
7	NIR1	1	150		
8	NIR2	1			

### 3.3. Tree Canopy Extraction

Hierarchical classification strategy was utilized to extract tree crowns in non-shadow areas. Tree canopy under shadows casted by buildings were not considered in this study as it is difficult to recover accurate spectral information of tree crowns. First, shadow and non-shadow areas were separated with NIR1 threshold. Studies have demonstrated that urban land covers usually have higher reflectivity at the NIR spectrum than the visual spectrum, and the reflectance in the shadow area drops more significantly at the NIR band than the non-shadow area because of the occlusion of sunlight [34]. The threshold value was determined using a bimodal histogram splitting method, which has been successfully used for shadow detection [34–36]. Image objects with mean NIR1 reflectance values higher than the threshold were extracted as the non-shadow area. Next, vegetation was extracted from the non-shadow area with a NDVI threshold. Because there was an overlap between NDVI values of blue roofs (0.29–0.60 at CNU and 0.22–0.64 at BNU) and those of vegetation (0.35–0.98 at CNU and 0.43–0.99 at BNU), we further used the blue band threshold to remove these buildings because blue roofs have significantly higher blue band reflectance than vegetation. Both NDVI and blue band thresholds were determined with a stepwise approximation method [12], which searches an initial threshold values in the histogram and then identifies the optimal threshold value as the one that results in the best match with reference polygons. Finally, tree canopy objects were separated from the other vegetated areas such as grass and shrub. New metrics were first calculated by multiplication of a textural feature such as Grey Level Co-occurrence Matrix Entropy (GLCME) of NIR1 band ( $GLCME_{NIR1}$ ) or Grey Level Difference Vector Angular Second Moment (GLDVA) of NIR1 band ( $GLDVA_{NIR1}$ ) and a color feature such as hue or intensity. Both texture and color features were considered because of the observations that tree canopy normally had a higher  $GLCME_{NIR1}$  value or hue values than grass/shrub, and lower  $GLDVA_{NIR1}$  or intensity values. All combinations of  $GLCME_{NIR1}$ ,  $GLDVA_{NIR1}$ , hue and intensity values were evaluated. Similar as the vegetation/non-vegetation classification, threshold values were determined using the stepwise approximation method. After trial and error, threshold values of  $GLCME_{NIR1} \times$  hue metrics and  $GLDVA_{NIR1} \times$  intensity metrics were chosen for the CNU and BNU areas, respectively (Table 4). The accuracy of tree crown extraction was approximately 95% for both sites. Table 4 summarized all threshold values for each classification step.

**Table 4.** Threshold values for hierarchical classification steps.

	Non-Shaded Area	Vegetation	Tree Canopy
CNU_WV2	$NIR1 \geq 0.083$	$NDVI > 0.35$ and $blue < 0.075$	$GLCME_{NIR1} \times hue > 2.45$
CNU_WV3	$NIR1 \geq 0.094$	$NDVI > 0.35$ and $blue < 0.12$	$GLCME_{NIR1} \times hue > 2.6$
BNU_WV2	$NIR1 \geq 0.063$	$NDVI > 0.43$ and $blue < 0.077$	$GLDVA_{NIR1} \times intensity < 0.0019$
BNU_WV3	$NIR1 \geq 0.085$	$NDVI > 0.33$ and $blue < 0.12$	$GLDVA_{NIR1} \times intensity < 0.0013$

### 3.4. Tree Canopy Feature Extraction

For tree canopy objects on each of the WV2 and WV3 images, a total of 69 features including 33 spectral features and 36 textural features were extracted from the pan-sharpened eight-band image in addition to the first principal component layer derived by Principal Component Analysis (PCA). As listed in Table 5, spectral features consisted of means and standard deviations of surface reflectance of each band, ratios calculated by mean spectral value of each band divided by the sum of spectral values of all eight bands, NDVIs derived from red, NIR1 bands and four additional bands, brightness values derived from the traditional bands (band 2, 3, 5, 7) and additional bands (band 1, 4, 6, 8), and mean and standard deviation of the first principal component image. Textural features included 24 GLCM and 12 GLDV features [9,37]. GLCM is a tabulation of the frequency of different combinations of grey levels at a specified distance and orientation in an image object, and GLDV is the sum of the diagonals of the GLCM within the image object [9,28]. The spectral

indices were considered because they reflect spectral discrimination between tree species. Texture indices were used because tree crowns with different species have different crown structures and distribution of branches or twigs. All these spectral and textural indices are potentially useful for forest or urban tree species classification [9,12]. The first principle component image accounts for most of the variance in the eight bands. Previous research [38] has shown that PCA analysis helped discrimination of trees, grass, water and impervious surfaces in urban areas. The 69 features were used for each single-date classification scheme. For the bi-temporal classification scheme, a total of 138 features, including 69 features from each of the WV2 and WV3 images, were used. Note that object geometric features such as size, shape index, or length/width ratio were not incorporated because an image segment may consist of several connected trees, and could therefore not represent shape and size characteristics of each tree crown.

**Table 5.** Description of object features derived from WV2/WV3 images.

Feature Name	Description
Mean1-8	Mean of bands 1–8
SD1-8	Standard deviations of individual bands 1–8
Ratio1-8	<i>i</i> th band mean divided by sum of band 1 through band 8 means
BTRA	Brightness derived from traditional bands 2, 3, 5, 7
BADD	Brightness derived from additional bands 1, 4, 6, 8
NDVI75	$(\text{band7} - \text{band5}) / (\text{band7} + \text{band5})$
NDVI86	$(\text{band8} - \text{band6}) / (\text{band8} + \text{band6})$
NDVI84	$(\text{band8} - \text{band4}) / (\text{band8} + \text{band4})$
NDVI61	$(\text{band6} - \text{band1}) / (\text{band6} + \text{band1})$
NDVI65	$(\text{band6} - \text{band5}) / (\text{band6} + \text{band5})$
GLCMH	GLCM homogeneity from bands 3, 6, 7, 8
GLCMCON	GLCM contrast from bands 3, 6, 7, 8
GLCMD	GLCM dissimilarity from bands 6, 7, 8
GLCMM	GLCM mean from bands 3, 6
GLCME	GLCM entropy from bands 3, 6, 7, 8
GLCMSD	GLCM standard deviation from bands 3, 6, 7, 8
GLCMCOR	GLCM correlation from bands 6, 7, 8
GLDVA	GLDV angular second moment from bands 6, 7, 8
GLDVE	GLDV entropy from bands 6, 7, 8
GLDVC	GLDV contrast from bands 3, 6, 7, 8
GLDVM	GLDV mean from bands 3, 6
PCAM	Mean of the first principal component from Principal Component Analysis
PCASD	Standard deviation of first component from Principal Component Analysis

### 3.5. Tree Species Classification

Image objects that intersected with reference polygons were used as reference samples. The number of samples for each tree species was summarized in Table 6. For each classification scheme at each study area, two machine learning algorithms, Support Vector Machine (SVM) and Random Forest (RF), were used to mode the classifiers.

**Table 6.** Number of sample objects for each study area.

Study Area	Tree Species	WV2	WV3	Bi-Temporal
CNU	PATO	98	98	260
	POTO	269	214	620
	SOJA	205	157	396
	GIBI	74	50	87
	Sum	646	519	1363
BNU	PLAC	168	115	397
	POTO	448	303	726
	SOJA	459	335	814
	GIBI	241	187	259
	Sum	1316	940	2196

Support Vector Machine (SVM) was developed by Cortes and Vapnik [39]. It attempts to find the optimal hyper plane in the high-dimensional feature space to maximize the margin between classes with a kernel function in polynomial, radial basis, or sigmoid form. SVM has proved to perform well in handling small numbers of training samples with high-dimensional space [14,40,41]. In this study, we used Radial Basis Function Kernel (RBF) as the kernel function because it works well in many classification tasks [42–44]. Optimal classifier parameters including RBF kernel parameter  $g$  and penalty factor  $c$  were determined by the grid search method in the LibSVM software package [45].

Random Forest (RF) [46] is constituted by many Classification and Regression Trees (CARTs) and suitable for high-dimensional dataset classification [40,47,48]. Each decision tree in RF is constructed by extracting an individual bootstrap sample (sampling with replacement) from the original dataset. The splitting variables used at each node are determined based on the Gini Index [7,40,49–51]. Then, each tree assigns the single vote to the most frequent class for the input data. To finish, the class gaining the majority vote is classified into the corresponding category. RF employed out-of-bag (OOB) samples that are not in the bootstrap samples to estimate the prediction performance. During classification, two important parameters [7,49] are necessary: (1)  $ntree$ , *i.e.*, the number of decision trees executing classification; and (2)  $mtry$ , *i.e.*, the number of input variables used at each node. In this study, we tried the parameter  $ntree$  ranging from 100 to 1000 and  $mtry$  as 1, 8 or 11 (the square root of the number of features) to determine the best group setting. After trial and error, the validation results indicated that the accuracy reached its maximum with  $ntree = 500$  and  $mtry = 8$  for single-date image classification, and  $ntree = 500$  and  $mtry = 11$  for bi-temporal image classification. These parameters were thus used for RF classification in this study.

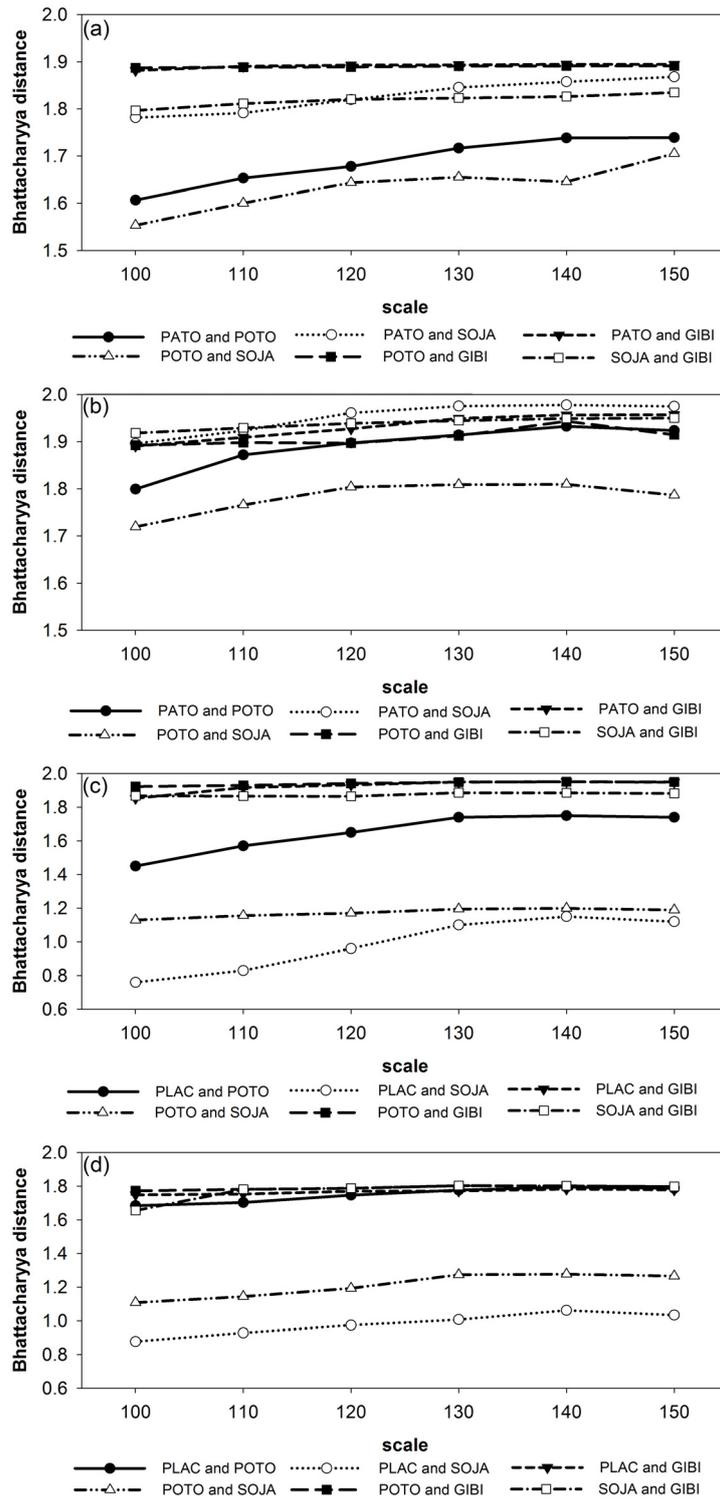
Image objects with over 10% of the area intersecting with the reference samples were used as sample objects. Ten-fold cross-validation with fixed subsets of training/validation sample objects was conducted to for each classifier and classification scheme. Each subset of the training/validation samples were selected based on a stratified random sampling process in order to ensure that the sample selected had a proportional number of each class. The average overall accuracy, kappa value, user and producer accuracy from validation datasets were used to assess the classification performance.

## 4. Results

### 4.1. Selection of Optimal Image Segmentation Scale Parameter

Pair-wise BD values were calculated for each image at each study area. Figure 4 shows that at both areas, BD values of all six pairs of species increase gradually with a scale parameter ranging from 100 to 130. For the WV2 image at CNU (Figure 4a), all BD values except that between POTO and SOJA reach their maximum at the scale parameter of 140 and become stable from 140 to 150. The maximum BD between POTO and SOJA is obtained at a scale parameter of 150. Thus, 150 was selected as the best scale parameter. For the WV3 image at CNU, all BD values reach a plateau at a scale parameter of 140, thus 140 was selected as best scale parameter (Figure 4b). Similarly, 140 was selected as best scale parameters for both images at BNU (Figure 4c,d). In both study areas, the number of tree stems per segment varies from one to eight depending on the tree crown layout. For individual tree crowns, each segment only covers one tree crown. For connected tree crowns (for example, Chinese white poplars on the CNU campus are planted densely along the avenue), one segment may cover as many as eight stems.

At both study sites, GIBI and any other species always have higher BD values, thereby indicating that GIBI has better spectral separability from other species, and thus better classification accuracy is expected. POTO and SOJA species always have smaller BD values, meaning that the two species are less spectrally separable. PLAC and SOJA, PATO and POTO species are also less separable. Overall, BD values at BNU are lower than those at CNU; lower overall classification accuracy is therefore expected.



**Figure 4.** BD values between six pairs of tree species at image segmentation scales from 100 to 150. (a) BD values from WV2 image at CNU; (b) BD values from WV3 image at CNU; (c) BD values from WV2 image at BNU; (d) BD values from WV3 image at BNU.

4.2. Classification Results

Table 7 lists the average overall accuracy (OA) and kappa values from all classification schemes at both study sites. The OAs range from 70.0% to 92.4% at the CNU study site, and from 71.0% to

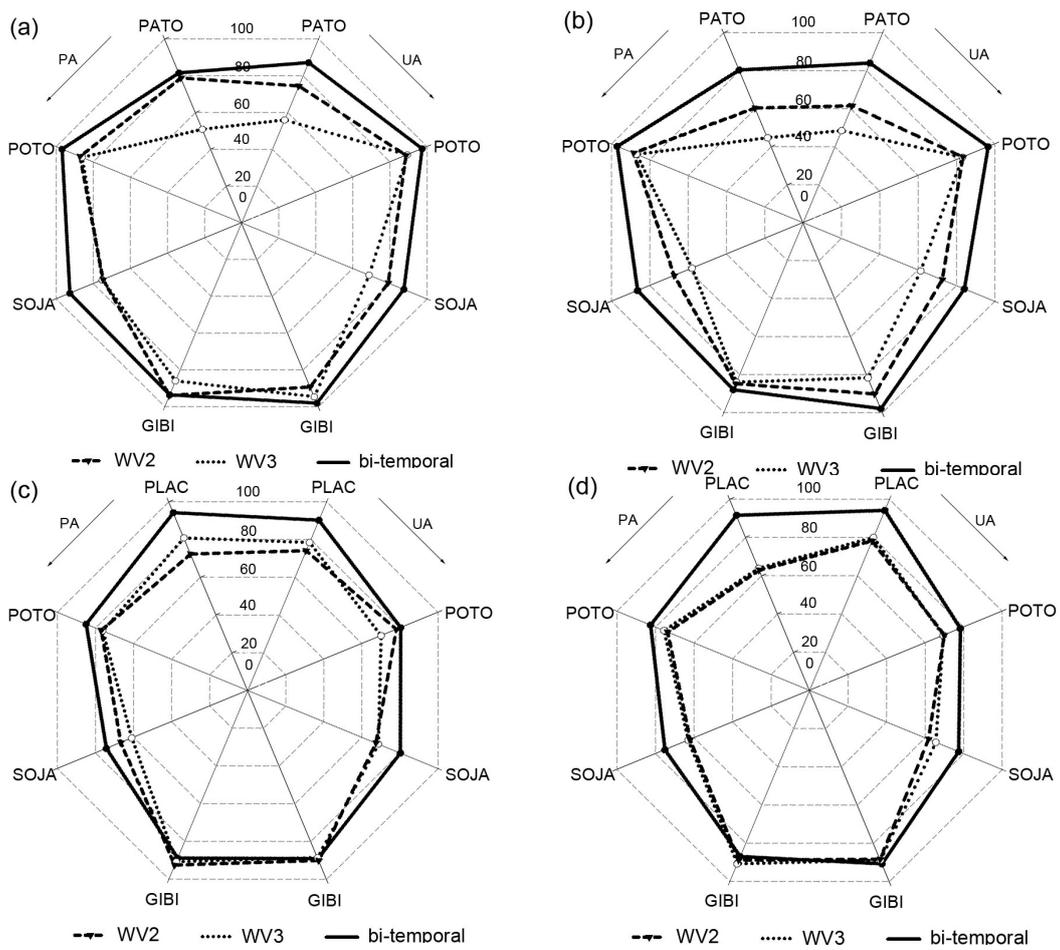
83.0% at the BNU study site. Using either the SVM or RF algorithm, tree species classification based on bi-temporal WV2 and WV3 images produces considerably higher accuracies than those based on each image alone at both study sites, with an average increase in OA of 11.5%. At CNU, the OA is around 9.7%–20.2% higher than when using a single-date image, regardless of the classification method used. Kappa values based on bi-temporal images are higher than 0.85, while those based on single-date images can be as low as 0.56. At BNU, OAs from bi-temporal image classification increase around 5%–12% compared to single-date image classification, and kappa values increase from less than 0.60 to over 0.75.

**Table 7.** Classification results with single-date images and bi-temporal images using the SVM and RF methods in the CNU and BNU areas. OA: overall accuracy.

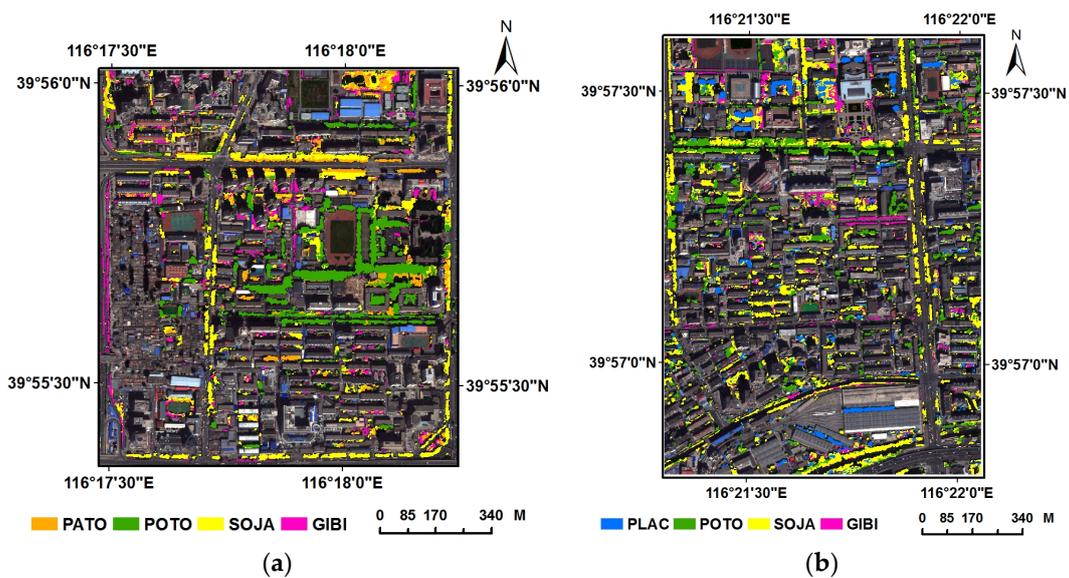
Classification Schemes	CNU				BNU			
	SVM		RF		SVM		RF	
	OA(%)	Kappa	OA(%)	Kappa	OA(%)	Kappa	OA(%)	Kappa
WV2	82.7	0.75	77.2	0.67	75.6	0.66	71.0	0.59
WV3	76.3	0.66	70.0	0.56	74.2	0.64	72.7	0.62
bi-temporal	92.4	0.89	90.2	0.85	80.3	0.76	83.0	0.76

Radar charts in Figure 5 show that for almost all tree species, both producer accuracies (PAs) and user accuracies (UAs) derived from bi-temporal image classification are notably higher than those from single-date image classification; the increases of PAs and UAs are on average 10.7%. At CNU, PAs of SOJA species increase by over 15% regardless of classification algorithm used, and UAs increase by over 11%; PATO has low PA and UA (44.7%–56.0%) based on the WV3 image alone, which may be caused by the small spectral reflectance gap between PATO and other species during high autumn (Figure 2b), while the addition of the WV2 image acquired in late summer helps increase the accuracy to 81.5%–87.1%. At BNU, PA of PLAC increase from 62.9% to 91.2% and UA increase from 78.2% to 94.2%. GIBI has relatively higher PAs and UAs than other species at both study sites using either WV2 or WV3 images. Especially at BNU, both PA and UA of GIBI are over 85%, while PA and UA of other species are lower than 80%. PATO, PLAC and SOJA had lower PA and UA when single-date images were used in the classification in each area. However, when using a combination of WV2 and WV3 images, PA and UA of PATO and SOJA at CNU, and PLAC at BNU increase substantially and are comparable with those of GIBI species (Figure 5). It is evident that bi-temporal classification not only produces higher but also more balanced PAs and UAs among tree species. The standard deviation of PAs of all species decreases from 11.7% using a single-date image to 6.2% using a bi-temporal image, and that of UAs decreases from 13.6% to 6.4%.

Table 7 demonstrates that OAs at the CNU site are consistently higher than those from BNU, regardless of the classification methods or schemes. For example, the overall accuracy generated from WV2 images using the SVM classifier at CNU is about 7% higher than that at BNU. For the same dominant species at both sites, *i.e.*, POTO, SOJA and GIBI, both PAs and UAs at CNU are higher than those at BNU (Figure 5). Comparisons between SVM and RF classification results show that SVM is superior to RF at both sites (Table 7 and Figure 5). OAs from SVM were around 1.5%–6.3% higher than those from RF except that from the bi-temporal classification scheme at BNU. The SVM classifier was thus used for tree species mapping. Figure 6 illustrates the resultant classification maps for dominant tree species using the SVM method based on bi-temporal images over both areas. Minor tree species were classified as one of the dominant classes. It is obvious that the classification map at BNU is more fragmented than that at CNU.



**Figure 5.** Radar charts of three classification schemes with SVM and RF methods in two study areas. (a) PAs and UAs using SVM at CNU; (b) PAs and UAs using RF at CNU; (c) PAs and UAs using SVM at BNU; (d) PAs and UAs using RF at BNU. PA: producer accuracy; UA: user accuracy.



**Figure 6.** Tree species map at (a) CNU area and (b) BNU area using bi-temporal images and SVM method.

### 4.3. Feature Importance

Tables 8 and 9 list the first 20 important metrics ranked by SVM and RF classifiers in the bi-temporal classification scheme. F-score and Mean Decrease Accuracy (MDA) were used to calculate feature importance in SVM and RF, respectively. F-score is a tool embedded in LibSVM software and measures the discrimination of two sets of real numbers [52]. MDA is generally used to measure the contribution of each feature to prediction accuracy in the RF model and can be calculated by permuting the  $m$ th features of each tree for the out-of-bag data [53]. Higher F-score or MDA values indicate higher importance of the feature in classification.

Among the 138 features used for bi-temporal classification, the top 20 important features identified by the SVM and RF classifier are mostly dominated by spectral characteristics. In both the CNU and BNU areas, less than five textural features are listed in the top 20 features by either the SVM or RF classifier, indicating that the spectral features make a more significant contribution to the species classification than the textural features. In both study areas, both the F-score and MDA identify the red-edge band (Band 6), the new NIR band (NIR2, Band 8) and green band (Band 3) as the most important bands since the features based on these bands, such as WV2\_NDVI86, WV3\_SD6, WV2\_Mean3 and WV3\_Ratio6, are consistently listed among the top-ranking features. Compared to the traditional four bands of high resolution satellite sensors, the new bands, especially the red-edge and NIR2 band designed for WV2 and WV3, make more of a contribution to urban tree species identification. In the two study areas, both WV2 and WV3 features are listed as important. Removing features derived from either WV2 or WV3 result in a decrease in accuracy, thereby emphasizing the role of bi-temporal spectral information in urban tree species classification.

**Table 8.** F-score weight of each feature in SVM for bi-temporal classification.

Rank	CNU		Rank	BNU	
	Features	F-Score		Features	F-Score
1	WV2_Mean3	0.696	1	WV2_Ratio6	0.479
2	WV3_Mean3	0.511	2	WV2_NDVI86	0.409
3	WV2_Mean4	0.511	3	WV3_Ratio6	0.403
4	WV2_NDVI86	0.455	4	WV2_Mean3	0.376
5	WV3_SD6	0.446	5	WV2_Mean4	0.331
6	WV3_SD3	0.353	6	WV3_SD6	0.309
7	WV3_SD5	0.344	7	WV2_Mean6	0.288
8	WV3_PCASD	0.343	8	WV2_SD6	0.249
9	WV3_PCAM	0.343	9	WV2_Mean5	0.236
10	WV3_Mean4	0.337	10	WV3_Mean6	0.229
11	WV3_SD2	0.322	11	WV3_PCASD	0.207
12	WV3_SD4	0.322	12	WV2_BADD	0.196
13	WV2_Ratio6	0.312	13	WV2_PCASD	0.189
14	WV3_Ratio6	0.310	14	WV2_BTRA	0.187
15	WV2_Mean5	0.304	15	WV2_PCAM	0.185
16	WV2_Mean2	0.291	16	WV3_NDVI61	0.185
17	WV2_Mean1	0.262	17	WV3_Mean3	0.184
18	WV3_SD1	0.262	18	WV2_GLDVE7	0.169
19	WV3_SD8	0.261	19	WV2_GLDVA8	0.167
20	WV3_Mean2	0.254	20	WV2_GLDVE8	0.165

**Table 9.** MDA of each feature in RF for bi-temporal classification.

Rank	CNU		Rank	BNU	
	Features	MDA		Features	MDA
1	WV2_NDVI86	20.3	1	WV2_Ratio6	24.7
2	WV2_Ratio6	19.8	2	WV2_NDVI86	20.0
3	WV2_Mean3	16.1	3	WV3_Ratio6	16.8
4	WV3_SD6	15.1	4	WV3_SD6	16.6
5	WV3_PCASD	13.6	5	WV3_NDVI86	15.8
6	WV3_Ratio6	13.5	6	WV2_Mean3	15.7
7	WV3_Mean5	13.3	7	WV2_Ratio3	14.4
8	WV3_Mean3	13.2	8	WV2_GLCMSD6	13.9
9	WV3_SD8	12.7	9	WV2_Ratio5	13.8
10	WV3_NDVI86	12.6	10	WV2_SD6	13.7
11	WV2_GLCMM3	12.2	11	WV3_Mean3	13.7
12	WV2_Mean6	12.1	12	WV2_NDVI65	13.6
13	WV3_SD7	11.9	13	WV3_NDVI65	13.2
14	WV3_PCAM	11.8	14	WV3_GLCMM6	13.2
15	WV2_SD6	11.7	15	WV2_GLCMH8	13.0
16	WV2_BTRA	11.7	16	WV2_GLDVA8	12.9
17	WV2_Mean8	11.7	17	WV3_Mean1	12.8
18	WV2_GLDVC3	11.7	18	WV3_SD7	12.7
19	WV2_GLCMCON3	11.7	19	WV2_NDVI57	12.6
20	WV2_PCAM	11.7	20	WV3_GLDVA6	12.4

## 5. Discussion

### 5.1. Effect of Bi-Temporal Images on Urban Tree Species Classification

In this study, we applied an object-based approach with SVM and RF classification algorithms for urban tree species classification at two study areas in Beijing, China. Our results show that using bi-temporal WV2 and WV3 images consistently improve urban tree species mapping accuracy regardless of study area or classifier used. The increasing accuracy in bi-temporal classification is mainly attributed to phenology variation that is represented in WV2 images acquired in late summer and WV3 images acquired in high autumn. From summer to autumn, the content of chlorophyll in tree leaves decreases gradually while the lutein increases. This results in a slight increase of blue and red band reflectance and a significant increase of costal and yellow band reflectance, while the reflectance of NIR1 and NIR2 bands declines due to change of porous thin-walled cells and tissues of the plant leaves. Variations in phenological patterns among tree species represented in the bi-temporal images enhance the spectral heterogeneity of crowns, thus helping to improve the classification accuracy of each tree species. It should be noted that the unique phenological characteristics of each tree species is also influenced by the climate conditions at the time. As already mentioned, temperature in early October 2014 dropped significantly. This possibly accelerated the process of pigment change and leaves falling, especially for the GIBI species whose leaf coloring and senescence occur earlier than other species. Our results suggest that local phenological phases during a certain year need to be considered when selecting bi-temporal or multi-temporal images for tree species classification.

Compared to previous research considering three or more phenological phases within one year [14,54], we used only two seasons in different years in late summer and high autumn in order to identify tree species and also gained promising accuracy. Hill *et al.* [55] reported that the highest classification accuracy was obtained when combining spring, summer and autumn image, while an autumn image with an image from both the green-up and full-leaf phases were sufficient for forest tree species classification. Tigges *et al.* [13] pointed out that spring, summer and autumn images were needed to achieve high class separability in an urban forest in Berlin using five-band RapidEye images. Our results confirmed these findings by combining late summer images with high

autumn images. Although the two images were acquired in different years, there was not much change in species distribution, rendering them sufficient for bi-temporal analysis. The resultant OAs from bi-temporal image classification of 80.3% to 92.4% (CNU area) and 80.3% (BNU area), and Kappa values of 0.89 (CNU area) and 0.76 (BNU area) indicate high consistency with the actual species distribution. In addition to higher overall classification accuracy from bi-temporal image classification that was reported similarly in previous research, we found that the PAs and UAs of tree species were more balanced compared to the single-date image scenario. Smaller differences of PAs (UAs) among tree species suggest that each individual species can be identified with similar accuracy, and the overall species distribution map is more reliable.

Our study extends the existing research [13] in that we explored bi-temporal images with higher spatial resolution for urban tree species mapping. In urban areas where trees have various crown sizes and scattered distribution, resolution of either IKONOS or RapidEye cannot provide spatial details of tree crowns, and thus does not support species-level mapping. Our study suggests that satellite sensors should provide spatial and spectral resolution comparable with WV2 or WV3, and images acquired on at least two phenological phases are needed for the purpose of urban tree species mapping.

Our analysis on feature importance rank further confirms the effect of bi-temporal images on urban tree species identification, as both WV2- and WV3-derived features are listed among the most important features ranked by the F-score used for SVM and the mean decrease accuracy (MDA) used for RF. Removing features from any date reduces the accuracy. It is worth noting that spectral features always have a greater contribution than texture features in species discrimination, which varies from the studies in forested areas [9,50]. This can be explained by a much lower density of trees in an urban area. As trees are isolated, differences in texture patterns of tree crowns are difficult to uncover. This further highlights the importance of spectral features, especially those associated with chlorophyll or leaf cellular structure such as red-edge and NIR reflectance. Therefore, the utilization of bi-temporal or multi-temporal images is critical, because the temporal change of spectral features enhances the discrimination of tree species.

### *5.2. Effect of Complex Urban Environment on Tree Species Classification*

Two study areas were selected in this study in order to explore the capability of WV2/WV3 in urban tree species classification. Although both study sites represent a typical urban environment in big cities in China and tree species are similar at both sites, they are different in crown spatial distribution and represent different urban environment heterogeneities. As previously mentioned, compared to the BNU site, trees with same species in the CNU area are more clustered and evenly distributed than those with scattered distribution; our field investigation reported that the average number of tree species within buffer areas with the same size is 2.3 at CNU and 3.5 at BNU. On the other hand, BNU has a greater number of tall buildings and they are in general taller than those at CNU. Previous research showed that species distribution influences the classification results in a forested area [56]. Uniform distribution of species tends to reduce the effect of overlapping crown from different adjacent species and thus yields higher accuracy. More abundant species within a small region increases variations of spectral information, thus reducing spectral separability. In our study, we found a similar pattern in the urban environment. The overall accuracies at CNU with more uniform species distribution were consistently higher than those at the BNU site except that of the WV3 classification scheme using RF. Urban tree spectral characteristics are also susceptible to the surrounding artificial materials such as buildings and backgrounds. In addition to tree distribution, the scattering and diffuse reflectance from tall buildings may also affect the spectral information of tree crowns. Pure spectral information of a single tree species next to the highrise is difficult to observe, thus potentially leading to large within-species spectral variation and lower classification accuracy at BNU.

In our study, we used two machine learning algorithms, SVM and RF, in order to ensure that the improvement of bi-temporal images in urban tree species classification is independent of the classification approaches used. Interestingly, we found that SVM consistently outperforms RF at both study sites for all classification schemes except bi-temporal classification at the BNU site (Table 7). Many studies have compared SVM and RF in various classification applications and showed different comparison results [40,47,57]. Barrett *et al.* [41] applied SVM, RF and the Extremely Randomized Tree method for grassland classification in Ireland, and the results indicated that the accuracies of SVM and RF were very close. Duro *et al.* [51] reported that the classification accuracies yielded by SVM were slightly higher than those produced by RF when classifying the agricultural landscapes using the object-based image analysis approach; however, the difference was not statistically significant ( $p > 0.05$ ). Li *et al.* [52] compared several machine learning algorithms for urban land classification and reported that RF was always better than SVM. Malahlela *et al.* [53] reported higher accuracy of RF for canopy gap mapping based on WV2 data. Previous studies have suggested that the comparison between RF and SVM is affected by the amount of samples and the distribution of samples among each class [49]. SVM tends to outperform RF in the case of a small amount and uneven distribution of samples. Investigation of the reference sample data in our study shows that the number of samples of the four species is imbalanced. SOJA and POTO species have much greater coverage than PATO, PLAC and GIBI at both sites, and thus a greater amount of samples were collected. The disparity among the number of samples might explain the higher overall accuracies obtained with the SVM algorithm. In an urban area, tree species are more scattered and unevenly distributed than in a forested area. Our results, along with previous research [43], imply that SVM may be more suitable for urban tree species classification.

Higher overall accuracies may be due to the fact that the selected study areas are relatively small and that there are only four tree species at each study area compared to six or more species in previous studies [12,13]. Lower accuracy is expected for bigger study areas and for a more complex distribution of tree species. However, this study provides a solid example of the impact of different urban environments on tree species classification, as both study areas have similar tree species and the same datasets were examined. Nevertheless, district- or city-wide species-level classification still needs to be conducted for operational investigation of urban trees.

## 6. Conclusions

This study evaluated bi-temporal WV2 and WV3 imagery for tree species identification in complex urban environments. An object-based classification method based on SVM and RF machine learning algorithms was conducted using late summer WV2 images and/or high autumn WV3 images. Five dominant tree species, the empress tree (*Paulownia tomentosa*), Chinese white poplar (*Populus tomentosa* Carrière), Chinese scholar tree (*Sophora Japonica*), ginkgo (*Ginkgo biloba*) and London plane tree (*Platanus acerifolia*), were examined at two study sites in urban Beijing, China.

Results showed that phenology variations presented in the bi-temporal imagery helped enhance the species identification capability. The overall accuracies based on both WV2 and WV3 imagery reached 92.4% and were significantly higher than those based on single-date image alone. The PAs and UAs produced by a single-date image were relatively low (44.7%–82.5%) and could hardly satisfy requirements for most urban ecological assessment applications, whereas those derived from bi-temporal images were on average 10.7% higher, and the PAs and UAs more balanced. The feature importance analysis revealed that spectral characteristics were more important than texture features. The new wavebands designed for WV2 and WV3 sensors such as red-edge and NIR2 bands contributed more to the classification than the traditional four bands. Both WV2- and WV3-derived features are listed among the most important features. This highlights the necessity of bi-temporal or even multi-temporal high resolution images for urban tree species classification. Comparison between two study sites showed that urban environment heterogeneity and distribution pattern of trees in the study area influenced classification accuracy. Clustered and evenly distributed

tree species were more easily identified than trees with scattered distribution. Our study also suggests that SVM is superior to RF in urban tree species classification as the former algorithm is more suitable for imbalanced sample distribution.

This study only identified the species in non-shadowed areas and did not consider the shadow effect on species classification. Our future research will focus on the development of approaches for spectral information restoration under shadowed areas in order to perform better and more complete tree species inventory. Because of the rounded shape of tree crowns, the uneven sun illumination within tree crowns may affect the species classification, especially for trees individually distributed. Impact of the uneven illumination on tree species classification will be incorporated in our future research. As the WV3 satellite could provide images in the 16-band mode, future research will also include evaluation of the eight shortwave infrared spectral bands in urban tree species classification.

**Acknowledgments:** The authors acknowledge the support from the National Natural Science Foundation of China (Grant: 41401493; 41130744), 2015 Beijing Nova Program (xx2015B060) and Specialized Research Fund for the Doctoral Program of Higher Education (20131108120006).

**Author Contributions:** Yinghai Ke and Xiaojuan Li conceived and designed the experiments; Dan Li performed the experiments; Dan Li and Yinghai Ke analyzed the data; Huili Gong contributed data; Dan Li and Yinghai Ke wrote the paper.

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

## References

1. Nowak, D.J.; Hirabayashi, S.; Bodine, A.; Hoehn, R. Modeled PM<sub>2.5</sub> removal by trees in ten U.S. cities and associated health effects. *Environ. Pollut.* **2013**, *178*, 395–402. [[CrossRef](#)]
2. Davies, Z.G.; Edmondson, J.L.; Heinemeyer, A.; Leake, J.R.; Gaston, K.J. Mapping an urban ecosystem service: Quantifying above-ground carbon storage at a city-wide scale. *J. Appl. Ecol.* **2011**, *48*, 1125–1134. [[CrossRef](#)]
3. Nowak, D.J. Institutionalizing urban forestry as a “biotechnology” to improve environmental quality. *Urban For. Urban Green.* **2006**, *5*, 93–100. [[CrossRef](#)]
4. Escobedo, F.J.; Kroeger, T.; Wagner, J.E. Urban forests and pollution mitigation: Analyzing ecosystem services and disservices. *Environ. Pollut.* **2011**, *159*, 2078–2087. [[CrossRef](#)]
5. Mustafa, Y.T.; Habeeb, H.N. Object based technique for delineation and mapping 15 tree species using VHR WorldView-2 (WV-2) imagery. *SPIE Remote Sens. Int. Soc. Opt. Photonics* **2014**. [[CrossRef](#)]
6. Waser, L.; 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](#)]
7. Immitzer, M.; Atzberger, C.; Koukal, T. Tree species classification with random forest using very high spatial resolution 8-band Worldview-2 satellite data. *Remote Sens.* **2012**, *4*, 2661–2693. [[CrossRef](#)]
8. Kim, S.; Lee, W.; Kwak, D.; Biging, G.; Gong, P.; Lee, J.; Cho, H. Forest cover classification by optimal segmentation of high resolution satellite imagery. *Sensors* **2011**, *11*, 1943–1958. [[CrossRef](#)] [[PubMed](#)]
9. Ke, Y.; Quackenbush, L.J.; Im, J. Synergistic use of QuickBird multispectral imagery and LiDAR data for object-based forest species classification. *Remote Sens. Environ.* **2010**, *114*, 1141–1154. [[CrossRef](#)]
10. Ke, Y.; Quackenbush, L.J. Forest species classification and tree crown delineation using QuickBird imagery. In Proceedings of the ASPRS 2007 Annual Conference, Tampa, FL, USA, 7 May 2007; pp. 7–11.
11. Cho, M.A.; Malahlela, O.; Ramoelo, A. Assessing the utility WorldView-2 imagery for tree species mapping in South African subtropical humid forest and the conservation implications: Dukuduku forest patch as case study. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *38*, 349–357. [[CrossRef](#)]
12. 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](#)]
13. Li, D.; Ke, Y.; Gong, H.; Chen, B.; Zhu, L. Tree species classification based on WorldView-2 imagery in complex urban environment. *IEEE Proc.* **2014**. [[CrossRef](#)]
14. Tigges, J.; Lakes, T.; Hostert, P. Urban vegetation classification: Benefits of multitemporal RapidEye satellite data. *Remote Sens. Environ.* **2013**, *136*, 66–75. [[CrossRef](#)]

15. Chen, H.; Li, W. Analysis on forest distribution and structure in Beijing. *For. Res. Manag.* **2011**, *2*, 32–35. (In Chinese) [[CrossRef](#)]
16. Zhao, J.; OuYang, Z.; Zheng, H.; Xu, W.; Wang, X. Species composition and spatial structure of plants in urban parks of Beijing. *Chin. J. Appl. Ecol.* **2009**, *20*, 298–305.
17. Meng, X.; OuYang, Z.; Cui, G.; Li, W.; Zheng, H. Composition of plant species and their distribution patterns in Beijing urban ecosystem. *Acta Ecol. Sin.* **2004**, *24*, 2200–2206.
18. Zhong, S.; Ge, Q.; Zheng, J.; Dai, J.; Wang, H. Changes of main phenophases of natural calendar and phenological seasons in Beijing for the last 30 years. *Chin. J. Plant Ecol.* **2012**, *36*, 1217–1225. [[CrossRef](#)]
19. Zhang, M.; Du, Y.; Ren, Y. Phenological season of Shishahai Area in Beijing. *J. Cap. Norm. Univ.* **2007**, *28*, 78–99.
20. Tarantino, C.; Lovergine, F.; Pasquariello, G.; Adamo, M.; Blonda, P.; Tomaselli, V. 8-Band image data processing of the WorldView-2 Satellite in a wide area of applications. Available online: <http://www.intechopen.com/books/earth-observation/8-band-image-data-processing-of-the-worldview-2-satellite-in-a-wide-area-of-applications> (accessed on 30 Oct 2015).
21. Pu, R.; Landry, S.; Zhang, J. Evaluation of atmospheric correction methods in identifying urban tree species with WorldView-2 imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 1886–1897. [[CrossRef](#)]
22. Vermote, E.; Tanré, D.; Deuzé, J.L.; Herman, M.; Morcrette, J.J.; Kotchenova, S.Y. Second Simulation of A Satellite Signal in the Solar Spectrum-Vector (6SV); 6S User Guide Version 3. Available online: [http://6s.ltdri.org/files/tutorial/6S\\_Manual\\_Part\\_1.pdf](http://6s.ltdri.org/files/tutorial/6S_Manual_Part_1.pdf) (accessed on 30 October 2015).
23. ITT ENVI Version 4.8. Available online: <http://envi-ex.software.informer.com/4.8/> (accessed on 30 October 2015).
24. Padwick, C.; Deskevich, M.; Pacifici, F.; Smallwood, S. WorldView-2 pan-sharpening. In Proceedings of the ASPRS 2010 Annual Conference, San Diego, CA, USA, 26–30 April 2010.
25. Jain, S.; Jain, R.K. A remote sensing approach to establish relationships among different land covers at the micro level. *Int. J. Remote Sens.* **2006**, *27*, 2667–2682. [[CrossRef](#)]
26. Kosaka, N.; Akiyama, T.; Tsai, B.; Kojima, T. Forest type classification using data fusion of multispectral and panchromatic high-resolution satellite imageries. *Int. Geosci. Remote Sens. Symp.* **2005**, *4*, 2980–2980.
27. Salati, S.; van Ruitenbeek, F.; van der Meer, F.; Naimi, B. Detection of alteration induced by onshore gas seeps from ASTER and WorldView-2 Data. *Remote Sens.* **2014**, *6*, 3188–3209. [[CrossRef](#)]
28. Baatz, M.; Benz, U.; Dehghani, S.; Heynen, M.; Höltje, A.; Hofmann, P.; Lingenfelder, I.; Mimler, M.; Sohlbach, M.; Weber, M. *eCognition User Guide 4*; Definiens Imaging: Munich, Germany, 2004; pp. 133–138.
29. Novack, T.; Esch, T.; Kux, H.; Stilla, U. Machine learning comparison between WorldView-2 and QuickBird-2-simulated imagery regarding object-based urban land cover classification. *Remote Sens.* **2011**, *3*, 2263–2282. [[CrossRef](#)]
30. Xun, L.; Wang, L. An object-based SVM method incorporating optimal segmentation scale estimation using Bhattacharyya distance for mapping salt cedar (*Tamarisk spp.*) with QuickBird imagery. *GISci. Remote Sens.* **2015**, *52*, 257–273. [[CrossRef](#)]
31. Wang, L.; Sousa, W.P.; Gong, P. Integration of object-based and pixel-based classification for mapping mangroves with IKONOS imagery. *Int. J. Remote Sens.* **2004**, *25*, 5655–5668. [[CrossRef](#)]
32. Im, J.; Jensen, J.R.; Hodgson, M.E. Object-based land cover classification using high-posting-density LiDAR data. *GISci. Remote Sens.* **2008**, *45*, 209–228. [[CrossRef](#)]
33. Mathieu, R.; Aryal, J. Object-based classification of Ikonos imagery for mapping large-scale vegetation communities in urban areas. *Sensors* **2007**, *7*, 2860–2880. [[CrossRef](#)]
34. Song, H.; Huang, B.; Zhang, K. Shadow detection and reconstruction in high-resolution satellite images via morphological filtering and example-based learning. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 2545–2554. [[CrossRef](#)]
35. Zhou, W.; Huang, G.; Troy, A.; Cadenasso, M.L. Object-based land cover classification of shaded areas in high spatial resolution imagery of urban areas: A comparison study. *Remote Sens. Environ.* **2009**, *113*, 1769–1777. [[CrossRef](#)]
36. Dare, P.M. Shadow analysis in high-resolution satellite imagery of urban areas. *Photogramm. Eng. Remote Sens.* **2005**, *71*, 169–177. [[CrossRef](#)]
37. Haralick, R.M. Statistical image texture analysis. *Handb. Pattern Recognit. Image Proc.* **1986**, *86*, 247–279.

38. Myint, S.W.; Gober, P.; Brazel, A.; Grossman-Clarke, S.; Weng, Q. Per-pixel *vs.* object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sens. Environ.* **2011**, *115*, 1145–1161. [[CrossRef](#)]
39. Cortes, C.; Vapnik, V. Support vector network. *Mach. Learn.* **1995**, *3*, 273–297. [[CrossRef](#)]
40. Waske, B.; Benediktsson, J.A.; Árnason, K.; Sveinsson, J.R. Mapping of hyperspectral AVIRIS data using machine-learning algorithms. *Can. J. Remote Sens.* **2009**, *35*, S106–S116. [[CrossRef](#)]
41. Van der Linden, S. Classifying segmented hyperspectral data from a heterogeneous urban environment using support vector machines. *J. Appl. Remote Sens.* **2007**, *1*, 013543. [[CrossRef](#)]
42. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science and Business Media: New York, NY, USA, 1999; p. 333.
43. Heumann, B.W. An object-based classification of mangroves using a hybrid decision tree—Support vector machine approach. *Remote Sens.* **2011**, *3*, 2440–2460. [[CrossRef](#)]
44. Kavzoglu, T.; Colkesen, I. A kernel functions analysis for support vector machines for land cover classification. *Int. J. Appl. Earth Obs. Geoinf.* **2009**, *11*, 352–359. [[CrossRef](#)]
45. Chang, C.; Lin, C. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 1–27. [[CrossRef](#)]
46. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
47. Barrett, B.; Nitze, I.; Green, S.; Cawkwell, F. Assessment of multi-temporal, multi-sensor radar and ancillary spatial data for grasslands monitoring in Ireland using machine learning approaches. *Remote Sens. Environ.* **2014**, *152*, 109–124. [[CrossRef](#)]
48. Rodriguez-Galiano, V.F.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sanchez, J.P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm.* **2012**, *67*, 93–104. [[CrossRef](#)]
49. Adelabu, S.; Mutanga, O.; Adam, E.; Cho, M.A. Exploiting machine learning algorithms for tree species classification in a semiarid woodland using RapidEye image. *J. Appl. Remote Sens.* **2013**, *7*, 073480. [[CrossRef](#)]
50. Pal, M. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* **2005**, *26*, 217–222. [[CrossRef](#)]
51. Liaw, A.; Wiener, M. Classification and regression by random forest. *R News* **2002**, *2*, 18–22.
52. Chen, Y.; Lin, C. Combining SVMs with various feature selection strategies. In *Feature Extraction*; Springer: Berlin, Germany, 2006; pp. 315–324.
53. Xu, S.; Huang, X.; Xu, H.; Zhang, C. Improved prediction of coreceptor usage and phenotype of HIV-1 based on combined features of V3 loop sequence using random forest. *J. Microbiol.* **2007**, *45*, 441–446.
54. Key, T.; Warner, T.A.; McGraw, J.B.; Fajvan, M.A. A comparison of multispectral and multitemporal information in high spatial resolution imagery for classification of individual tree species in a temperate hardwood forest. *Remote Sens. Environ.* **2001**, *75*, 100–112. [[CrossRef](#)]
55. Hill, R.A.; Wilson, A.K.; George, M.; Hinsley, S.A. Mapping tree species in temperate deciduous woodland using time-series multi-spectral data. *Appl. Veg. Sci.* **2010**, *13*, 86–99. [[CrossRef](#)]
56. Chubey, M.S.; Franklin, S.E.; Wulder, M.A. Object-based analysis of Ikonos-2 imagery for extraction of forest inventory parameters. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 383–394. [[CrossRef](#)]
57. Duro, D.C.; Franklin, S.E.; Dubé, M.G. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sens. Environ.* **2012**, *118*, 259–272. [[CrossRef](#)]



© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).