

Article

Application of In-Segment Multiple Sampling in Object-Based Classification

Nataša Đurić ^{1,*}, Peter Pehani ² and Krištof Oštir ^{1,2}

¹ Slovenian Centre of Excellence for Space Sciences and Technologies SPACE-SI, Aškerčeva cesta 12, 1000 Ljubljana, Slovenia; E-Mail: kristof.ostir@space.si

² Research Centre of the Slovenian Academy of Sciences and Arts, Novi trg 2, 1000 Ljubljana, Slovenia; E-Mail: peter.pehani@zrc-sazu.si

* Author to whom correspondence should be addressed; E-Mail: natasa.dzuric@space.si; Tel.: +386-1-4706-502.

External Editors: Thomas Blaschke, Janet Nichol and Prasad S. Thenkabail

Received: 14 July 2014; in revised form: 11 November 2014 / Accepted: 27 November 2014 /

Published: 5 December 2014

Abstract: When object-based analysis is applied to very high-resolution imagery, pixels within the segments reveal large spectral inhomogeneity; their distribution can be considered complex rather than normal. When normality is violated, the classification methods that rely on the assumption of normally distributed data are not as successful or accurate. It is hard to detect normality violations in small samples. The segmentation process produces segments that vary highly in size; samples can be very big or very small. This paper investigates whether the complexity within the segment can be addressed using multiple random sampling of segment pixels and multiple calculations of similarity measures. In order to analyze the effect sampling has on classification results, statistics and probability value equations of non-parametric two-sample Kolmogorov-Smirnov test and parametric Student's *t*-test are selected as similarity measures in the classification process. The performance of both classifiers was assessed on a WorldView-2 image for four land cover classes (roads, buildings, grass and trees) and compared to two commonly used object-based classifiers—*k*-Nearest Neighbor (*k*-NN) and Support Vector Machine (SVM). Both proposed classifiers showed a slight improvement in the overall classification accuracies and produced more accurate classification maps when compared to the ground truth image.

Keywords: GEOBIA; in-segment analysis; sampling; urban land cover classification; Kolmogorov-Smirnov test statistics; Student's *t*-test statistics; probability value

1. Introduction

The improvements in the spatial resolution of satellite sensors that occurred over the last decade lead to the development of new image processing and classification techniques. The increasing availability and diversity of the high- and very high-resolution satellite imagery posed a challenge to researchers who not only had to deal with the abundance of available data but also with the great detail found in the image. One of the approaches that managed to overcome the aforementioned challenges was the Geographic Object-Based Image Analysis (GEOBIA). GEOBIA systems gained widespread popularity from 2000 onwards and are currently considered to be the state-of-the-art systems in both scientific and commercial thematic mapping of very high-resolution spaceborne imagery, e.g., [1–3].

The underlying concept of the two-stage GEOBIA approach is based on segmentation as the first preliminary step and classification as the second step. Generally, image segmentation is defined as a process of partitioning an image into homogenous groups called segments in which each segment is homogenous but no union of two adjacent segments is homogenous. In the classification step, regions are classified into the most appropriate classes based on their spectral, spatial and contextual information.

When compared to traditional non-contextual pixel based methods, several GEOBIA approaches presented improved results [4]. Regarding the object-based image classification as the second step of the GEOBIA approach, different methods and techniques have been investigated and employed in order to improve classification accuracy and performance and to fully exploit the additional available information provided in the image segments. However, numerous authors, e.g., [1,4–9] agree that any segmentation-based classification can only be as good as the underlying segmentation, which means that only good segmentation results can lead to object-oriented image classification out-performing pixel based classification. Li *et al.* [10] claim that the final image classification results do not depend merely on the segmentation accuracy but also on numerous other factors: the classification scheme, available images, training samples selection, data pre-processing that includes feature selection and extraction, the classification algorithm, post processing techniques, test sample collection, as well as validation methods.

In recent years the question has not been whether object-based classifiers are better than pixel based, but if and how object-based classification can gain from the classifier itself rather than from the aforementioned factors. Toure *et al.* [11] and Sridharan and Qiu [8] pointed out that the current approaches to classification of segments are based on statistical measures of central tendency and dispersion associated with normally distributed data. Due to the spectral inhomogeneity of pixels within the segments in high and very high-resolution imagery their distribution can be considered complex rather than normal [11]. Using classification methods that rely on normality assumptions such as maximum likelihood and utilizing merely the summary statistics that fail to capture the in-object heterogeneity may therefore lead to inappropriate, inaccurate and misleading results. In order to avoid normality violation, Toure *et al.* [11] proposed the use of histogram curve matching approaches. Histograms proved to be reliable features when characterizing classes. Toure *et al.* [11] also studied the

influence of wavebands on classification; the tests were performed with individual red and near infrared (NIR) bands and a combination of the two. Three different formulae for combining red and NIR bands outperformed the results obtained when using one or the other band individually.

Sridharan and Qiu [8] presented a fuzzy Kolmogorov-Smirnov based classifier that provides an object-to-object matching of the empirical distribution of the reflectance values. The Kolmogorov-Smirnov classifier has been employed as the supervised data-learning algorithm, *i.e.*, a segment is assigned to the class with which it has the most similar spectral signature relative to the training data signatures. This was tested for urban objects recognition from 8-band WorldView-2 data. Sridharan reported a minimum 10 percent increase in overall classification accuracy when compared to various popular object- and pixel-based classifiers.

Methods that assume data is derived from a particular distribution are referred to as parametric, whereas methods that do not rely on data belonging to any particular distribution are known as non-parametric. The decision as to whether to choose a parametric or non-parametric method is important when dealing with small samples. However, with small samples it may also be difficult to detect assumption violations. If the assumption deals with normal distribution, non-normality is hard to detect even when present, since small samples contain insufficient information that would enable reliable conclusions as regards the data distribution type.

In object-based image analysis, the segmentation process produces segments that vary greatly in size, *i.e.*, samples can be very big as well as very small. Various supervised classifiers tend to use spectral information within segments as it appears—all pixel values (regardless of the size of the segment) are used to compute either one summary measure or one empirical cumulative distribution function for a pair consisting of an unknown segment and a training sample. According to Sridharan and Qiu [8] empirical cumulative distributions fully characterize the analyzed segment as well as describe its inner complexity. The in-segment complexity is typical for urban areas and heterogeneous agricultural fields and should be properly addressed if we wish to obtain accurate classification results.

This study extends the classification approach based on the use of empirical cumulative distribution functions and the two-sample Kolmogorov-Smirnov test distance by Duric *et al.* [12]. The preliminary research conducted on a 3-band orthophoto image revealed the potential of an in-segment complexity analysis in regards to classification accuracy. However, due to the incomplete classification and evaluation methodology and the insufficient testing segments, it failed to provide objective conclusions as regards the benefits or disadvantages of the used classification approach.

In this paper we are going to analyze whether the complexity can be addressed using multiple random samplings of small sets of segment's pixel values and multiple calculations of similarity measures. In order to analyze the effect sampling has on classification results, statistics and probability values of non-parametric two-sample Kolmogorov-Smirnov test and the parametric Student's *t*-test are selected as similarity measures for the classification purposes. Both tests are well known in statistical hypothesis testing, however, in this analysis, hypothesis reasoning against a selected level of significance will be omitted as similarity measures will serve merely as a tool for relative comparison between different classes, namely as a degree of matching. Since the Student's *t*-test statistics computation is based on the mean and standard deviation of the segment's pixel values, it served well in the evaluation of the performance of—the hypothetically more representative—empirical cumulative distribution functions when small-sized pixel sets were sampled.

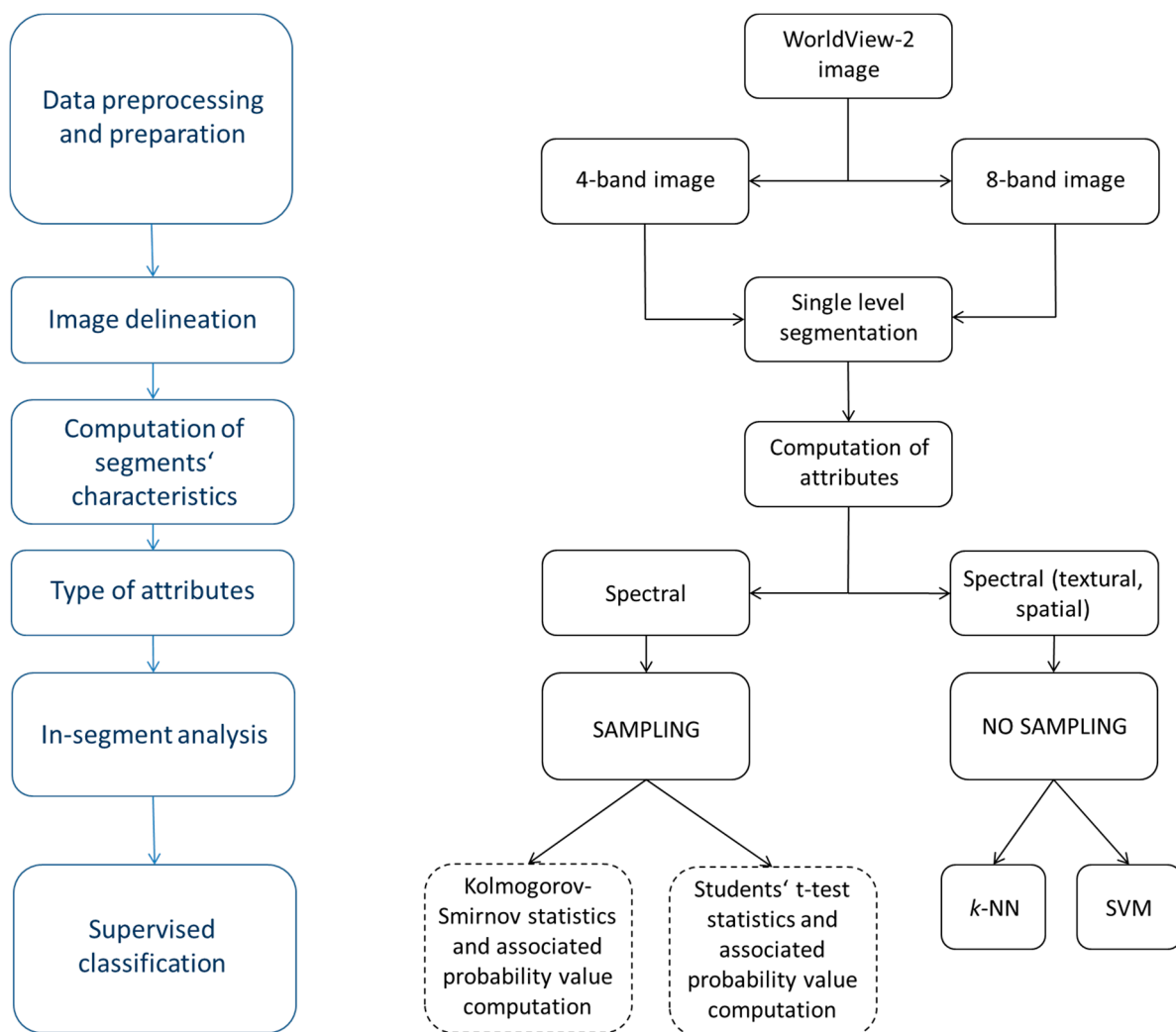
To summarize, the objectives of our study are:

- To describe in-segment pixel heterogeneity by exploiting the potential of multiple small set sampling,
- To study the effect of multiple small set sampling on normality violation with the parametric Student’s *t*-test,
- To compare the effectiveness of the Kolmogorov-Smirnov and Student’s *t*-test based classifiers, and
- To analyze the impact spectral resolution has on the classification results.

2. Data and Methodology

In order to attain the aforementioned objectives, our study applied a data processing workflow (shown in Figure 1). The analysis procedures followed an object-based supervised classification approach, which typically covers several steps: data pre-processing and preparation, segmentation, computation of the segments’ characteristics (attributes), the selection of training samples, classification and accuracy evaluation.

Figure 1. Process flow adopted in the study.



In addition to this highly used procedure, the emphasis of our study lied on the in-segment analysis, which was included in the attribute computation process. The in-segment domain was addressed using multiple pixel value samplings, followed by a multiple computation of similarity measures for each segment. In the Kolmogorov-Smirnov statistics the similarity measure is based on the computation of empirical cumulative distribution functions, while in the Student's *t*-test statistics the similarity measure is based on the computation of mean values. The performance and sensitivity of both classifiers was assessed with the use of WorldView-2 image (both in its full 8-band form and in its spectrally reduced 4-band form) and compared to two commonly used object-based classifiers—parametric *k*-Nearest Neighbor (*k*-NN) and non-parametric Support Vector Machine (SVM), in which the sampling was not adopted, *i.e.*, the segments were transferred to the classification step as they were segmented. Both, *k*-NN and SVM, classifiers are implemented in the Exelis VIS ENVI 5.0 image processing software and were used within the Feature Extraction Workflow (Example Based Feature Extraction Workflow). Sampling and two other proposed algorithms were implemented using the Interactive Data Language (IDL) within ENVI. A detailed description of the analysis is provided in the following sections.

2.1. Case Study Area and Data

In our study, the 8-band WorldView-2 satellite multispectral and panchromatic images of Ljubljana (Slovenia), acquired on 10 August 2010, were used. In order to analyze the effect spectral resolution has on the classification performance, 4- and 8-band images served as input data for the classification process. An image consisting of four bands includes standard Blue, Green, Red and Near Infrared 1 bands, whereas an 8-band image includes 4 additional bands: Coastal, Yellow, Red-Edge and Near Infrared 2. The applied pre-processing of satellite data involved precise orthorectification and pan-sharpening to a spatial resolution of 0.5 m. For the latter the Gram-Schmidt pan-sharpening algorithm in ENVI was applied.

The studied area, located west of the center of Ljubljana, the capital city of Slovenia, is highly residential, riddled with a road network. Due to the occurrence of only four most commonly observed land cover classes that were later adopted in the supervised classification—roads, buildings, grass and trees—a test site of 0.16 km² (Figure 2) was selected for the experimental purposes of this research.

The roads class included major roads and other built-up areas (e.g., courtyards). Untreated grass areas and lawns with individual trees were placed in the grass class. The trees class mainly consisted of areas covered by tree crowns and shrubbery. Despite the large variety of roofing material (black, red, light grey, dark grey and white roofs), no subclasses were considered for the buildings class. Since the Sun elevation at the time of data acquisition was high and most of the buildings are low, the shadowing effect was not severe in terms of reducing the image information. Thus, shadows were not categorized into the additional class.

Detailed national topographic vector data (scale 1:5000, from 2005) served as the reference when determining the training samples for the road and buildings classes. Due to the discrepancy between the acquisition date of the satellite imagery and the vector reference data, the latter was carefully examined in order to ensure it matched the situation on the satellite imagery. National vector data for grass and trees land cover was not available, thus, training samples were collected through the visual interpretation of very high-resolution (0.5 m) national orthophoto imagery (dated to 2011) instead of pan-sharpened

WorldView-2 image. The latter provided less clarity and sharpness on image objects, which made the interpretation harder. Although the spatial resolution of national orthophotos enabled a coarse differentiation between the trees and grass classes, this way of collecting training data is still considered to be extremely subjective.

Figure 2. The highly residential study area measuring 0.16 km² is located west of the Ljubljana center (Slovenia). It is marked with a black rectangular.



2.2. Segmentation

In GEOBIA approach segments are used as the basic units that form classified objects. Segmentation is a process in which an image is partitioned into a set of mutually disjoint regions that are more uniform within themselves than the adjacent regions [13]. The segmentation quality is important when we wish to produce an accurate classified image. Poor segmentation results can lead to high misclassification rates. Ideally, one segment should represent one object of interest and the segment and object boundaries should result in a perfect match [14]. However, in practice erroneous segments are often detected and genuine segments omitted [1].

The edge-based watershed algorithm implemented in ENVI 5.0 was used to extract land cover structure from the WorldView-2 image. With this algorithm, the appropriate level of detail and the average segment size are controlled by three parameters: the initial segmentation rate (scale level) and the merge level (both with a range between 1 and 100), and the number of selected bands that are to be included in the delineation and merging step. The scale level determines the fragmentation degree

(number of segments), whereas the merge level combines several smaller segments into a large one. Spectral resolution has also proved to be an important factor that impacts the quality of segmentation [14]. In their analysis Mesner and Ostir [14] studied the effect of spatial and spectral resolutions on the segmentation quality, which was the main focus of their research. They discovered that the poorest segmentation results were achieved when all eight spectral bands of the WorldView-2 image were used. The study area was located in North–Eastern Slovenia and the imagery used in [14] was acquired on 21 June 2011. Since segmentation is an essential part of the object-based classification approach used in this study, the findings described in [14] served merely to obtain optimal segments. The acquired segments represented an input for the detailed analysis of the classification step, which was not addressed in [14].

Through various trials and a subsequent visual inspection of the segmentation results, the scale and merge level were adopted to best delineate the image objects. Red, Blue, Green and NIR1 bands were selected as input bands in the segmentation, for both, the 4-band and 8-band, image analysis. The use of these 4 bands provided the most homogenous segmentation results in which meaningful segments with respect to the urban cover were created. However, this resulted in a few under-segmented segments, in which two land covers were located within a single segment (e.g., a tree crown and a building or a building with a grey roof and a road). Although the urban landscape was identified as a highly complex environment and therefore image objects should be addressed using multilevel segmentation, 1544 segments in our study were defined at a single segmentation level.

Attribute assignment represents the final step in the segmentation process. Attributes are computed for each segment in each band. Attributes can be spectral (variations in tone or color) or spatial (shape and spatial patterns). Spatial attributes may refer to the structure or texture of the object—understood as a tonal variation of the object of interest—or to the broader relationship between the object and its surroundings—usually referred to as context [15]. Since the computation of probability values in the two-sample Kolmogorov-Smirnov and Student's *t*-test classifiers is based solely on the spectral information (pixel values), *k*-NN and SVM classification methods were primarily conducted in ENVI with only four spectral attributes (Table 1) [16]. Thus, a more straightforward comparison of the four classification methods took place in the accuracy assessment step. In order to study the effect of the additional attributes on the classification results, the classification using *k*-NN and SVM methods was later repeated with 22 different attributes (4 spectral, 4 texture, 14 spatial) (Table 1). Spectral information is not required with 14 spatial attributes since they are computed from the polygon defining the boundary of the pixel cluster. Therefore, their value is constant in all bands.

2.3. The Selection of Training and Testing Samples

In supervised classification samples with a known identity (training samples) are necessary if we wish to construct a model capable of classifying unknown (testing) samples. Concise information as regards the number of selected training and testing samples can be found in Table 2. In our analysis, the national orthophoto image provided a basis for the training and testing sample determination. Due to our funding constraints it was impossible to perform field verification. Training samples were distributed throughout the image and we also attempted to ensure a uniform number of training samples with spectral homogeneity through careful visual inspection. However, due to the variety of roofing material in the study area, a higher number of training samples were chosen for the buildings class.

Table 1. Types of attributes computed per segment for the *k*-NN and SVM classification methods, as implemented in the ENVI 5.0 processing software.

Attribute	Type	Description
Spectral	Minimum	Minimum value of pixels comprising the region in band x.
	Maximum	Maximum value of pixels comprising the region in band x.
	Mean	Mean value of pixels comprising the region in band x.
	Standard deviation	Standard deviation value of pixels comprising the region in band x.
Texture	Range	Average data range of pixels comprising the region within the kernel.
	Mean	Average value of pixels comprising the region within the kernel.
	Variance	Average variance of pixels comprising the region within the kernel.
	Entropy	Average entropy value of pixels comprising the region within the kernel.
Spatial	Area	Total area of the polygon, minus the area of the holes.
	Length	The combined length of all polygon boundaries, including the boundaries of the holes.
	Compactness	A shape measurement that indicates the compactness of the polygon. A circle is the most compact shape with a value of $1/\pi$.
	Convexity	This attribute measures the convexity of the polygon. The convexity value for a convex polygon with no holes is 1.0, while the value for a concave polygon is below 1.0.
	Solidity	A shape measurement that compares the area of the polygon to the area of a convex hull that surrounds the polygon. The solidity value for a convex polygon with no holes is 1.0, while the value for a concave polygon is below 1.0.
	Roundness	A shape measurement that compares the area of the polygon to the square of the maximum diameter of the polygon. The roundness value of a circle is 1, while the value for a square is $4/\pi$.
	Form factor	A shape measurement that compares the area of the polygon to the square of the total perimeter. The form factor value of a circle is 1, while the value of a square is $\pi/4$.
	Elongation	A shape measurement that indicates the ratio of the major axis of the polygon to the minor axis of the polygon. The elongation value for a square is 1.0, while the value for a rectangle is greater than 1.0.
	Rectangular fit	A shape measurement that indicates how well the shape is described by a rectangle. The rectangular fit value for a rectangle is 1.0, while the value for a non-rectangular shape is below 1.0.
	Main direction	The angle subtended by the major axis of the polygon and the x-axis in degrees. The main direction value ranges between 0 and 180°. 90° is North/South, while 0 to 180° is East/West.
	Major length	The length of the major axis of an oriented bounding box that encloses the polygon.
	Minor length	The length of the minor axis of an oriented bounding box that encloses the polygon.
	Number of holes	The number of holes in the polygon.
	Hole area	The ratio of the total area of the polygon towards the area of the outer contour of the polygon. The hole-solid ratio value for a polygon with no holes is 1.0.

Table 2. The number of selected training and testing samples per individual class.

Class	Number of Selected Segments for Training Samples	Number of Selected Segments for Testing Samples
Roads	6	67
Buildings	11	257
Trees	7	176
Grass	5	
Total	29	500

In order to study the effectiveness of the four classification methods, random testing sample sets were selected—separate for roads and buildings classes (67 segments for the roads class and 257 segments for the buildings class) and combined for trees and grass classes (176 segments). Consequently, once the final layer of classified segments was obtained and intersected with the reference layer for grass and trees classes, only the presence of roads and buildings was examined and verified in the new layer. The decision to merge the testing samples for trees and grass arose from the uncertainty in determining them through the visual interpretation of the orthophoto image when a high number and variety of segments were involved.

Segments that were identified as training samples were excluded from the testing samples set. Table 2 shows that almost one third of all (1544) segments that were identified in the segmentation process were selected for the evaluation process of the classification performance.

2.4. Supervised Classification Process

The following supervised classification methods were used in our study: (a) *k*-Nearest Neighbor; (b) Support Vector Machine; (c) two-sample Kolmogorov-Smirnov classification algorithm and (d) Student's *t*-test classification algorithm. The *k*-Nearest Neighbor method computes the Euclidean distance from each segment in the segmentation image to every defined training sample. The distance is measured in an *n*-dimensional space, in which *n* is the number of attributes for the individual training sample [17]. The Support Vector Machine is a supervised classification method derived from the statistical learning theory [18]. The SVM approach seeks to find the optimal separating hyperplane between classes [19] by focusing on the points closest to the hyperplane. The closest points are called support vectors.

The following processing parameter default values were used with *k*-NN (a) and SVM (b) classifiers available in the ENVI 5.0 image processing software: (a) number of neighbors: 1; (b) kernel type: radial basis, degree of kernel polynomial: 1, bias in kernel function: 1, gamma in kernel function: 0.03, penalty parameter: 100.00. As the classification accuracies could be significantly influenced by the parameter selection, an additional analysis of the parameter values was conducted (using only the 4-band image and spectral attributes) in order to avoid favoring the proposed method. The analysis has shown that default values yielded the highest mean producer and user accuracies. Therefore, default values were considered safe to adopt in our land cover determination application.

In general, the concept of the supervised classification approach in the case of the two-sample Kolmogorov-Smirnov classification algorithm and the Student's *t*-test classification algorithm is based on the comparison of the unknown segment with the training samples (*i.e.*, all-against-one approach),

where an unknown segment is classified into a class with which a highest similarity measure was determined. The similarity analysis using two-sample Kolmogorov-Smirnov and Student's t -test measures starts with the computation of the attribute for each segment (empirical cumulative distribution function and mean value, respectively), and continues with the computation of test statistics and probability values (p -values). The p -value of the two statistics represents the input for the classification process. In order to classify an unknown segment, the p -value needs to be computed against all segments with known classes (training segments) in each band. In the following step, p -values from all bands are combined for each "unknown segment-segment with a known class" pair. These are joined in the overall similarity measure (membership grade), which is defined as the geometric mean of p -values for all bands. Sridharan and Qiu [8] consider this approach to be fuzzy. The overall similarity measure value ranges between 0 and 1, where 0 represents no similarity between the two segments and 1 represents identical segments. The final step in the classification process assigns any unknown segment to a class with the maximum overall membership grade. The theoretical background for both proposed classification algorithms is described in greater detail in Sections 2.4.1 and 2.4.2.

2.4.1. The Two-Sample Kolmogorov-Smirnov Test Statistics Based Classification Algorithm

The Kolmogorov-Smirnov test statistic is widely used to measure the closeness between two empirical cumulative distribution functions (ECDF). Given two sets of observation (e.g., two segments with pixel values), $X_{(1)}, X_{(2)}, \dots, X_{(m)}$ and $Y_{(1)}, Y_{(2)}, \dots, Y_{(n)}$, their respective empirical cumulative distribution functions, denoted as $S_{(m)}(x)$ and $S_{(n)}(x)$, are defined as (Equations (1) and (2)) [20]:

$$S_m(x) = \begin{cases} 0 & \text{if } x < X_{(1)} \\ k/m & \text{if } X_{(k)} \leq x < X_{(k+1)} \\ 1 & \text{if } x \geq X_{(m+1)} \end{cases} \quad \text{for } k = 1, 2, \dots, m-1 \quad (1)$$

and

$$S_n(x) = \begin{cases} 0 & \text{if } x < Y_{(1)} \\ k/n & \text{if } Y_{(k)} \leq x < Y_{(k+1)} \\ 1 & \text{if } x \geq Y_{(m+1)} \end{cases} \quad \text{for } k = 1, 2, \dots, n-1 \quad (2)$$

For each index k two equally sized samples will have the same empirical distribution function values, however, these values will differ for a chosen value x which makes the differentiation possible. In the two-sample case, the criterion for closeness is defined as the maximum value of the absolute difference between two empirical cumulative distribution functions (Equation (3)):

$$D_{m,n} = \max_x |S_m(x) - S_n(x)| \quad (3)$$

The Kolmogorov-Smirnov statistic $D_{(m,n)}$ is non-parametric and distribution free, therefore it has the advantage of making no assumptions as regards data distribution [21]. The statistic can assume any value between 0 and 1. The closer it is to 0, the more likely it is that the two samples were drawn from the same population. Corresponding to the statistic and size of the two samples, the probability value (p -value) is calculated in the next step. A p -value is a measure of how likely it is for the difference

between two empirical cumulative distribution functions to have occurred by chance, assuming the null hypothesis that two samples are drawn from the same distribution holds true. In our research, the p -value was calculated with an empirical determination suitable for small and medium sized samples (Equation (4)).

$$p\text{-value}_{KS} = 2 \sum_{j=1}^{\infty} (-1)^{j-1} e^{-2j^2\lambda^2} \quad (4)$$

in which $\lambda = (\sqrt{N_e} + 0.12 + 0.11/\sqrt{N_e}) \cdot D_{m,n}$ and $\sqrt{N_e} = \frac{m \cdot n}{m+n}$.

Similar to the two-sample Kolmogorov-Smirnov distance $D_{(m,n)}$, the p -value also ranges from one to zero; however, the two are inversely proportional. If the p -value is low it is highly likely that the two sets of observations were drawn from populations with different distributions. In other words, if we assume that the obtained p -value equals 0.03 random sampling from identical populations would result in 97% of the experiments with a $D_{m,n}$ difference smaller than the observed one, while in 3% of the experiments the difference would be larger than in the observed one.

2.4.2. Student's t -Test Statistics Based Classification Algorithm

Similarly to the Kolmogorov-Smirnov statistics, the separability analysis between two segments can be addressed using the Student's t -test statistics and the corresponding probability value. Unlike the Kolmogorov-Smirnov test in which the computation is based on the empirical distribution functions, this test is commonly used as a two-sample location test for the following null hypothesis: the means of two populations are equal assuming that the variances of the two populations are also equal. If the assumption as regards the variance equality is dropped, the test is known as the Welch's t -test which was employed in this research. The t statistic that tests whether the population means are different is calculated as (Equation (5)):

$$t = \frac{\overline{X}_m - \overline{Y}_n}{\sqrt{\frac{S_X^{*2}}{m} + \frac{S_Y^{*2}}{n}}} \quad (5)$$

where \overline{X}_m and \overline{Y}_n are sample means, S_X^{*2} and S_Y^{*2} are sample variances and m and n are sample sizes. The t statistics follows the ordinary Student's t distribution with ν degrees of freedom calculated as (Equation (6)):

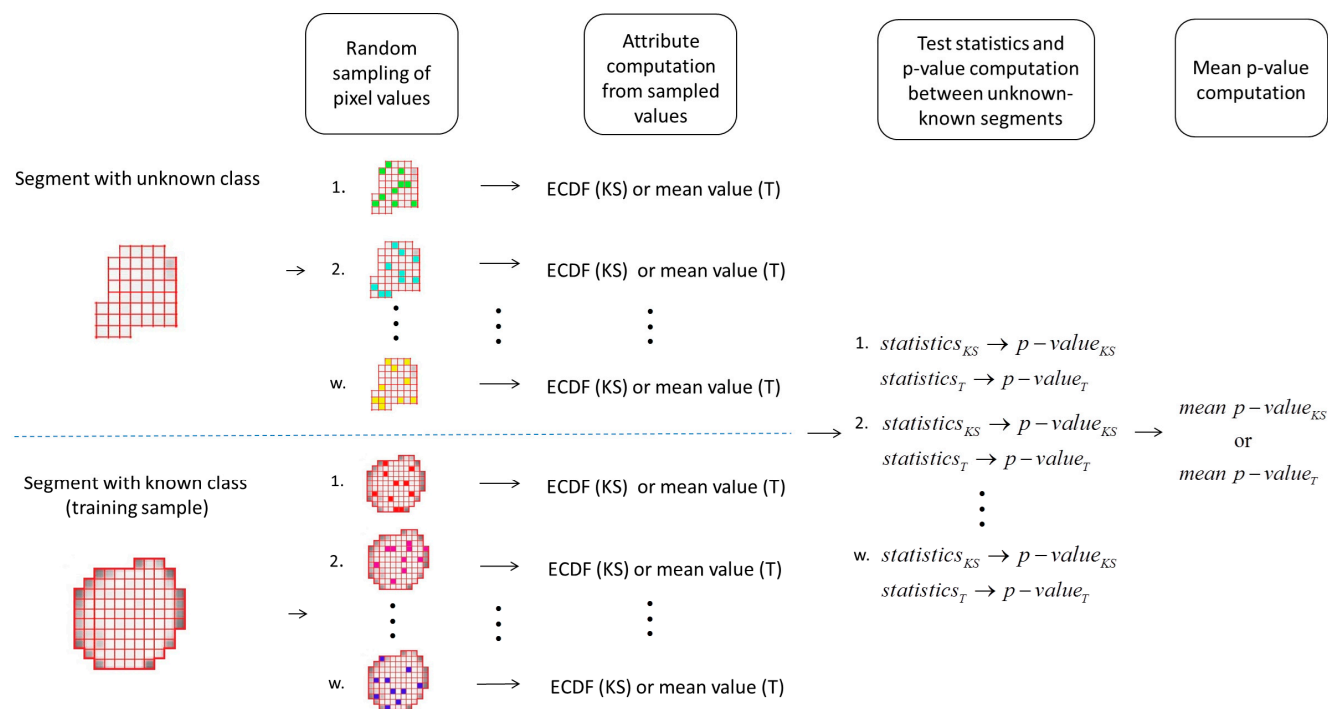
$$\nu = \frac{\left(\frac{S_X^{*2}}{m} + \frac{S_Y^{*2}}{n}\right)^2}{\frac{(S_X^{*2}/m)^2}{m-1} + \frac{(S_Y^{*2}/n)^2}{n-1}} \quad (6)$$

The p -value is computed in order to quantify how far apart the two means are. Its computation is not as straightforward as in the Kolmogorov-Smirnov statistics, since it is based on the construction of incomplete beta function for which the input variables are the t statistic and degrees of freedom.

2.4.3. Random Sampling Approach

In addition to the general supervised classification in which segments are analyzed in their full size (all pixels in the segment at once), we also implemented a random sampling approach. The idea of the random sampling approach is presented in Figure 3.

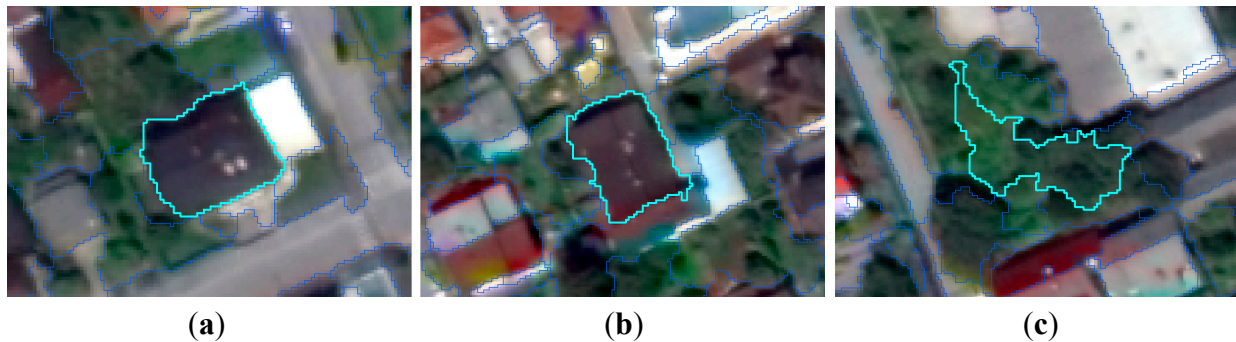
Figure 3. Random sampling approach. Empirical cumulative distribution functions and mean values are computed for each set of sampled pixel values for segments with a known and unknown class. A p -value between segments with an unknown and known class is computed after each sampling and averaged (mean p -value) in order to obtain a single value for the specific combination of the two segments. ECDF stands for empirical cumulative distribution function, KS for two-sample Kolmogorov-Smirnov test statistics and T for Student’s t -test statistics.



A selection of a random set of pixel values for the segment was used to compute multiple aforementioned variables (empirical distribution functions and mean values) that were averaged in order to obtain the final p -value used for the overall degree of matching computation. It is assumed that with multiple samplings the in-segment heterogeneity can be captured more accurately than with single computation of summary statistics, in which any normality violations associated with summary statistics can lead to misleading results. This may not be the case if small pixel sets are sampled, since non-normality is hard to detect in small samples.

We studied the effect the sample size and the number of samplings have on the p -value. In order to achieve this, we selected three homogenous segments, each containing over 1000 pixels. Two segments were similar (black roofs), while the third one differed (grass patch) (Figure 4). We selected them visually.

Figure 4. Segments from the segmentation layer selected for the sampling analysis: (a) a black roof; (b) a black roof; (c) a grass patch. Each segment contains more than 1000 pixels.



3. Results and Discussion

3.1. Sampling Analysis

Figures 5 and 6 show how the p -value between the segments with similar land covers (two black roofs, Figure 5) and segments with different land covers (a black roof and a grass patch, Figure 6) is influenced by the number of sampled pixels using the Kolmogorov-Smirnov statistics and the Student's t -test statistics. Different sized pixel sets were sampled 100 times in each segment. As shown later, with this number of samplings (100), the p -value becomes stable and does not vary significantly.

The comparison of Figures 5 and 6 shows that the p -value converged to zero as the number of sampled pixels increased. The drop is very fast when the Kolmogorov-Smirnov statistics was applied. As defined in Section 2.4., both proposed classifiers are based on the proximity or similarity analysis. The starting premise is that no two segments are identical. Instead, segments should be considered as continuous variables for which the perspective plays a significant role: the closer we look (*i.e.*, the larger the sample size), the smaller is the necessary difference that leads to the claim that the segments are not derived from the same population, despite the fact that two empirical cumulative distribution functions would appear to be the almost identical when plotted. Therefore, a sufficiently large sample would lead to the rejection of the hypothesis, if, of course, the hypothesis testing was applied. These conclusions could also reflect the equations used for the p -value computation that is better suited to small and medium sized samples. However, we also tested the asymptotic functions suitable for larger samples with the two-sample Kolmogorov-Smirnov statistic that yielded p -values of a similar magnitude.

The results shown in Figures 5 and 6 are summarized in Figure 7 that illustrates the difference in the p -value in individual bands for the two-sample Kolmogorov-Smirnov statistics (Figure 7a) and for the Student's t -test statistics (Figure 7b) in the example for 10 sampled pixels. Both classifiers clearly indicated that the largest differentiation between different land covers (a black roof and a grass patch) was seen in Red-Edge, Near Infrared 1 and Near Infrared 2 bands. In these bands, the calculated p -value was insignificant (almost equal to zero), which suggests that the two observed segments were actually different and that infrared bands could contribute useful separability information in the classification process. Different p -values in the Red band (p -value was higher for segments with different land covers) indicated a possible misclassification if only this band was used for classification purposes. The results were also summarized for cases in which a higher numbers of pixels (20, 50, 100, 200, 500,

and 1000) were sampled. As the number of the sampled pixels increased, the Coastal and Green bands became increasingly important in the differentiation with the Kolmogorov-Smirnov test statistics, as the p -values in Red-Edge, Near Infrared 1 and Near Infrared 2 bands converged towards zero very fast (with 50 sampled pixels). The Red band became unreliable for classification purposes. With the Student's t -test statistics, Yellow, Red and Red-Edge bands were the highest contributors to the separation of different segments as the number of sampled pixels increased.

Figure 5. The per-band p -value computation between segments with similar land covers (two black roofs) when different sizes of sampled pixel sets were applied and when: (a) the two-sample Kolmogorov-Smirnov statistics; (b) the Student's t -test statistics was used.

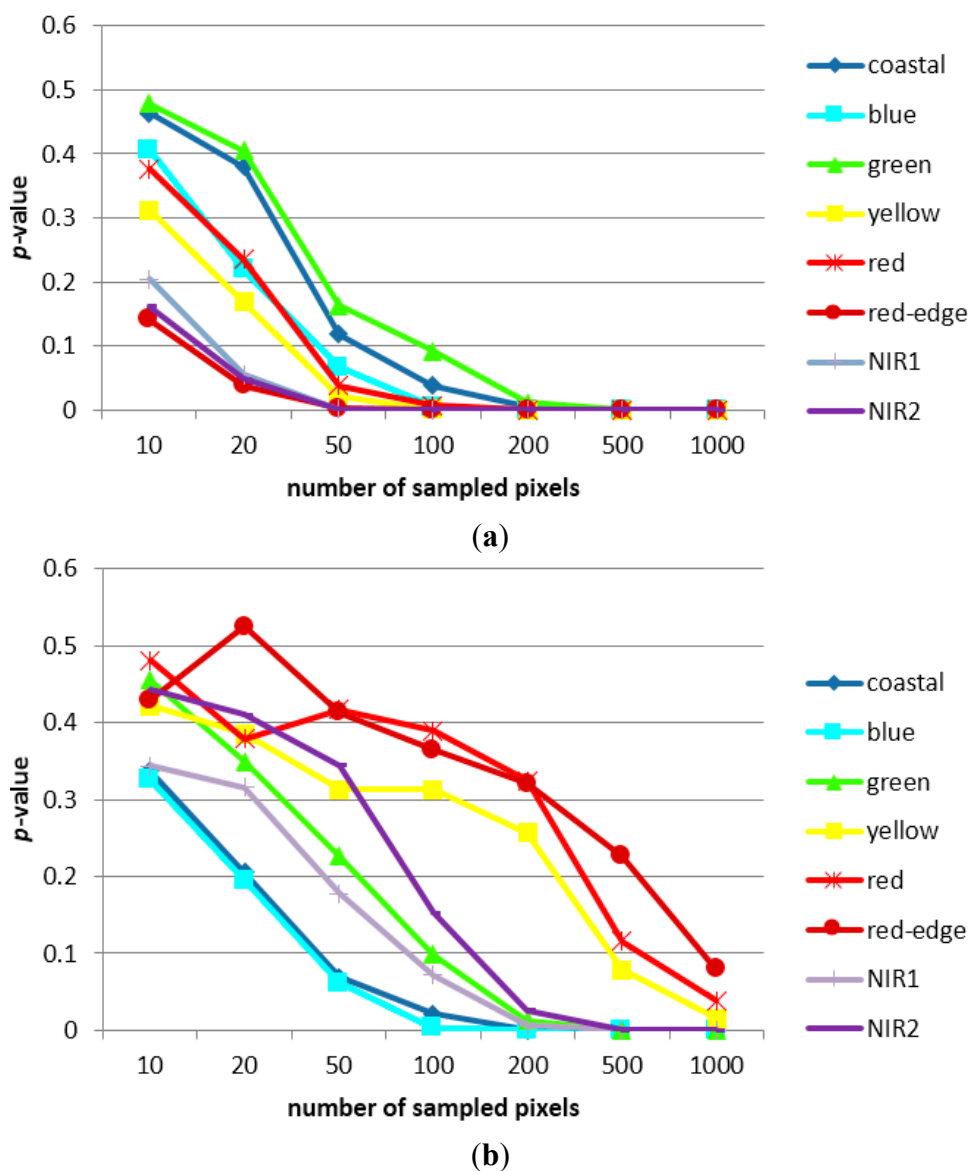


Figure 6. The per-band p -value computation between segments with different land covers (a black roof and a grass patch) when different sizes of sampled pixel sets were applied and when: (a) the two-sample Kolmogorov-Smirnov statistics; (b) the Student’s t -test statistics was used.

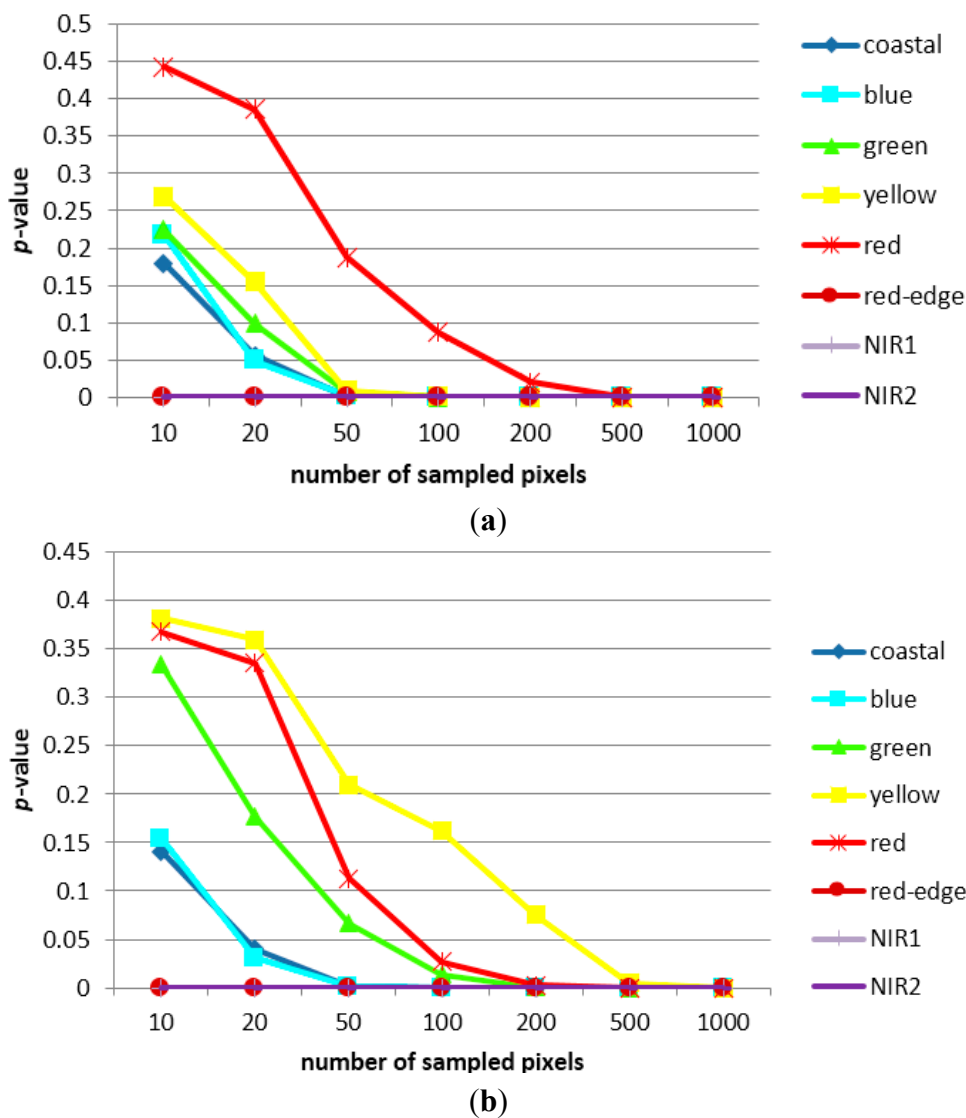


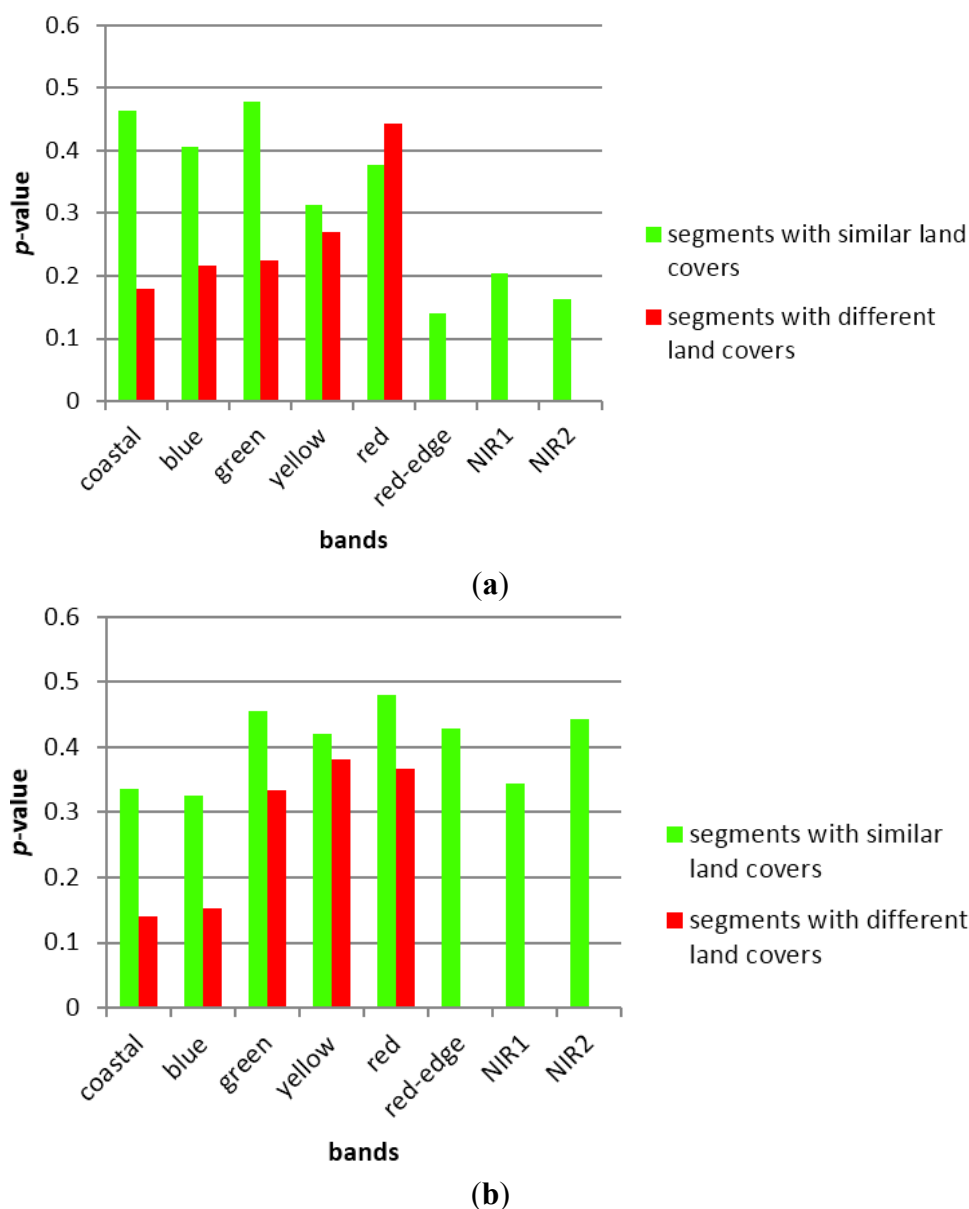
Figure 7 represents the results merely for the three analyzed segments (two black roofs and a grass patch). In order to provide more generalized and objective conclusions on a contribution of individual bands to classification results, a larger number of segments with different land covers should be selected and compared.

Finally, in order to study the effect of the number of samplings on the p -value, a sample size of 10 pixels was chosen. We selected this small sample size, because the p -value was the highest when the minimum number of pixels was sampled (see Figures 5 and 6). However, it was assumed that with a small-sized sampled pixel set, the p -value could easily be affected by outliers. The question arose as to how many times a set of pixels needs to be sampled in order for the p -value to become stable. The results are presented in Figures 8 and 9.

With both classifiers the p -value stabilized when approximately 100 samplings were performed. The sampling analysis also included the comparison of two small segments (*i.e.*, segments with less than

150 pixels) and mixed sized segments (*i.e.*, one small and one large segment). The results were consistent with the large segments.

Figure 7. The difference in the p -value per individual band for 10 sampled pixels when: (a) the Kolmogorov-Smirnov statistics (Figures 5a and 6a combined); (b) the Student's t -test statistics was applied (Figures 5b and 6b combined).



The substantial difference between the sizes of the two compared samples was the subject of the research in the field of hypothesis testing. Gordon and Klebanov [22] proved that a paradoxical situation takes place when the two-sample Kolmogorov-Smirnov test is used: one cannot use additional information contained in a very large sample if the second sample is relatively small, meaning that the two-sample Kolmogorov-Smirnov test can lose power when the sizes of the two samples differ substantially.

To summarize, 10 pixels were sampled 100 times for each segment both, with the Kolmogorov-Smirnov and Student's t -test classifiers. This assured that the two samples were always of the same size as well as enabled the inclusion of very small segments. Segments with less than 10 pixels, *i.e.*, 23 segments in

total, were left unclassified. Due to their small size they had no visual effect on the final classification image.

Figure 8. The per-band p -value computation between segments with similar land covers (two black roofs) when different numbers of 10 pixel samples were applied and when: (a) the two-sample Kolmogorov-Smirnov statistics; (b) the Student’s t -test statistics was used.

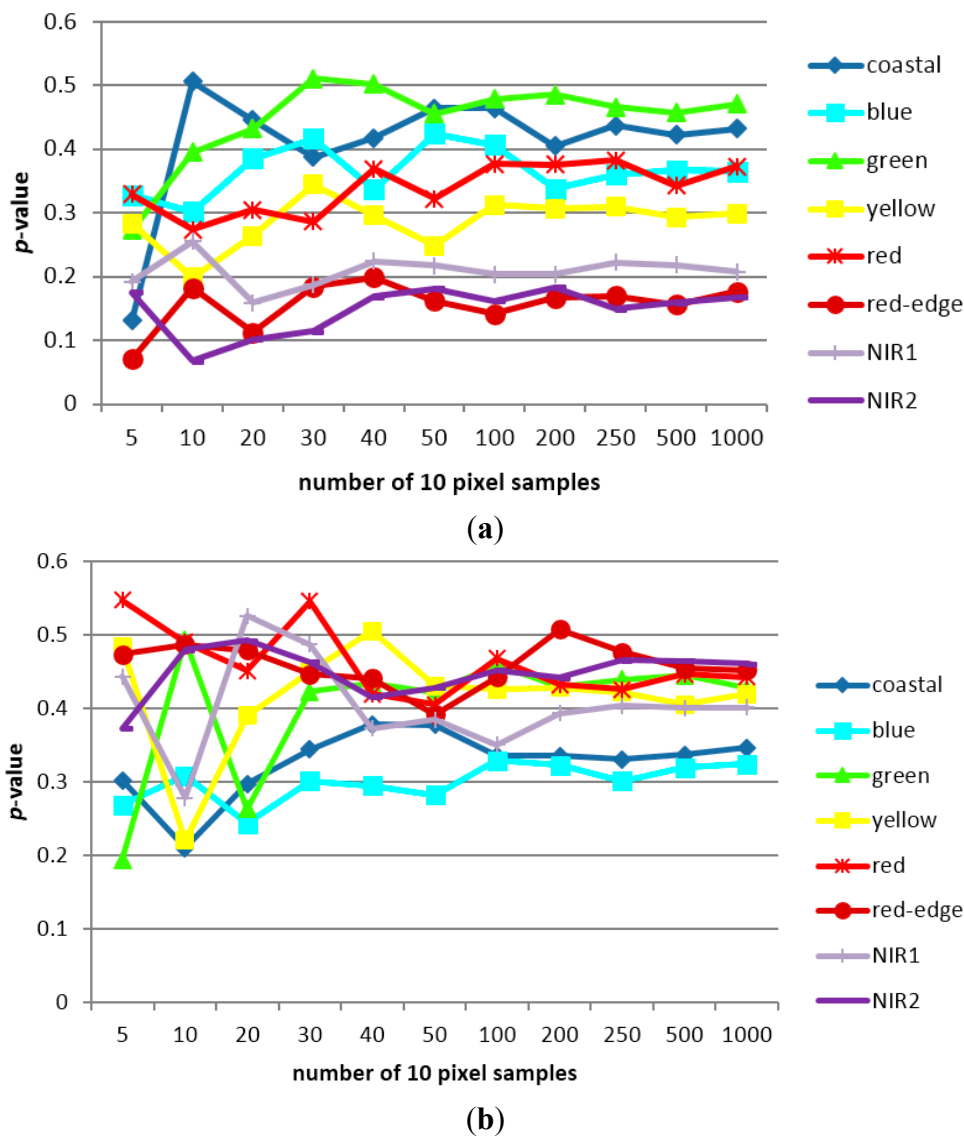
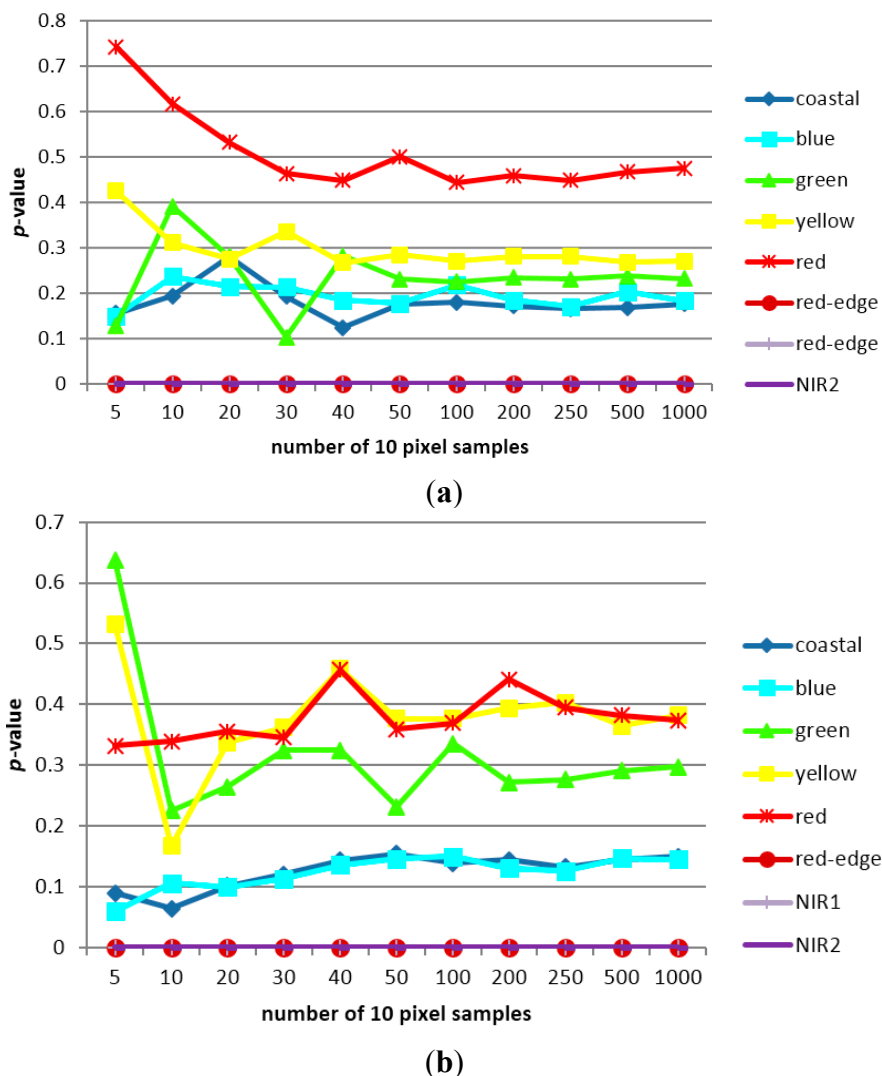


Figure 9. The per-band p -value computation between segments with different land covers (a black roof and a grass patch) when different numbers of 10 pixel samples were applied and when: (a) the two-sample Kolmogorov-Smirnov statistics; (b) the Student’s t -test statistics was used.



3.2. Classification Results and Accuracy Assessment

The final classification results and the basic accuracy assessment of the four used classifiers are summarized in Tables 3 and 4. Several accuracy measures, which give direct and basic indications of the classification quality, can be applied in the evaluation process [23]. In our study, user and producer accuracies per class and mean user and producer accuracies (in bold) were adopted and provided. Table 3 presents a confusion matrix for all four classifiers for the event in which k -NN and SVM classification methods were conducted with four spectral attributes, while Table 4 provides additional performance results for the event in which 22 attributes (4 spectral, 4 texture and 14 spatial) were used with k -NN and SVM classifiers. In the latter case, the results served to study the effect additional attributes have on the classification performance. The basic classification capability was estimated by examining the number of segments classified as the addressed class in regards to the reference data and the number of

misclassified segments for both the 4-band and the 8-band image. None of the four methods produced mean user and producer accuracies that would exceed 90%.

Moderate classification accuracies were the result of a single level segmentation, which can never ensure a clear delineation of objects and boundaries as seen in reality. Therefore, the segments appeared over- and under-segmented, which lead to a misinterpretation of their true land cover. This can be observed with all four classifiers by reading the values in the columns for trees and grass classes, where mainly buildings have been misassigned to this class. This is a result of over-segmentation, in which a building and a tree appeared within the same segment on numerous occasions.

The classification results for the 4-band image in Table 3 reveal that higher mean user and producer accuracies were achieved with Kolmogorov-Smirnov and Student's *t*-test classifiers (when the sampling approach was adopted) in overall, than with *k*-NN and SVM methods. These results show the potential of the in-segment analysis for the classification process. Amongst the four methods, the parametric Student's *t*-test classifier resulted in the best overall performance, which suggests that the normality violation is not an issue when small-sized pixel sets are sampled. The sampling of small-sized pixel sets had a greater effect on the computation of the empirical cumulative distribution functions. With a sample consisting of a mere 10 pixels, the in-segment distribution and its unique characteristics were inefficiently captured due to the large generalization, which resulted in the reduced classification capability of the two-sample Kolmogorov-Smirnov test statistics. Thus, empirical distribution functions should be approached in a different way.

Both proposed classifiers performed best for the classification of roads, grass and trees. Using merely spectral information resulted in a poor classification of objects from different classes built from the same material (*i.e.*, roads and buildings with grey roofs). Thus all four classifiers most commonly misclassified the buildings and roads. This suggests that the two proposed classifiers would also be ineffective for classifying 3-band RGB aerial images. The SVM classifier outperformed the remaining three methods when detecting buildings; however its performance was very poor for classifying roads. Six training samples were included in the roads class. The poor performance of the SVM classifier might have been caused by the large, yet spectrally relatively homogenous (no margin training samples were provided) training samples. The SVM method is otherwise known to achieve high classification results with small training sets [19].

It is insufficient to rely merely on the spectral information in supervised classification when training samples are not representative or individual classes do not show a great variety of different objects. This issue was most evident in both proposed classifiers. Their characteristic is to find the best match for the segment in the analysis related to the available training data samples (the all-against-one approach), making them highly dependable on the number and variety of training samples. If an insufficient number of samples are provided, both classifiers will provide a match and assign a class to a segment; however this class might not be the real one.

Table 3. A confusion matrix for all four classifiers for the 4- and 8-band input image in the event when k -NN and SVM classification methods were conducted with four spectral attributes. The user and producer accuracy values in bold represent the mean overall user and producer accuracy.

k-Nearest Neighbor										
Reference	4-Band Image					8-Band Image				
	Roads	Building	Trees + Grass	Total	Producer Accuracy (%)	Roads	Building	Trees + Grass	Total	Producer Accuracy (%)
Roads	43	20	4	67	59.7	48	18	1	67	57.8
Building	29	190	38	257	87.6	32	195	30	257	82.2
Trees + grass	0	7	169	176	80.1	3	24	149	176	82.8
Total	72	217	211	500	75.8	83	237	180	500	74.3
User accuracy (%)	64.2	73.9	96.0	78.0		71.6	75.9	84.6	77.4	
Support Vector Machine										
Reference	4-Band Image					8-Band Image				
	Roads	Building	Trees + Grass	Total	Producer Accuracy (%)	Roads	Building	Trees + Grass	Total	Producer Accuracy (%)
Roads	22	41	4	67	59.5	23	18	1	67	62.2
Building	15	210	32	257	83.3	13	210	34	257	90.5
Trees + grass	0	1	175	176	82.9	1	4	171	176	83.0
Total	37	252	211	500	75.2	37	232	206	500	78.6
User accuracy (%)	32.8	81.7	99.4	71.3		34.3	81.7	97.1	71.0	
Two-Sample Kolmogorov-Smirnov Test Statistics Classifier										
Reference	4-Band Image					8-Band Image				
	Roads	Building	Trees + Grass	Total	Producer Accuracy (%)	Roads	Building	Trees + Grass	Total	Producer Accuracy (%)
Roads	52	14	1	67	61.1	52	14	1	67	59.0
Building	33	179	45	257	92.7	36	181	40	257	92.8
Trees + grass	0	0	176	176	79.3	0	0	176	176	81.1
Total	85	193	222	500	77.7	88	195	217	500	77.6
User accuracy (%)	77.6	69.6	100	82.4		77.6	70.4	100	82.7	

Table 3. Cont.

Student's <i>t</i> -Test Statistics Classifier										
Reference	4-Band Image					8-Band Image				
	Roads	Building	Trees + Grass	Total	Producer Accuracy (%)	Roads	Building	Trees + Grass	Total	Producer Accuracy (%)
Roads	52	14	1	67	56.5	52	14	1	67	59.1
Building	40	194	23	257	93.3	36	196	25	257	93.3
Trees + grass	0	0	176	176	88.0	0	0	176	176	87.1
Total	92	208	200	500	79.3	88	210	202	500	79.8
User accuracy (%)	77.6	75.5	100	84.4		77.6	75.5	100	84.4	

Table 4. Additional confusion matrix for *k*-NN and SVM classifiers; conducted with 22 attributes (4 spectral, 4 texture and 14 spatial attributes).

<i>k</i> -Nearest Neighbor										
Reference	4-Band Image					8-Band Image				
	Roads	Building	Trees + Grass	Total	Producer Accuracy (%)	Roads	Building	Trees + Grass	Total	Producer Accuracy (%)
Roads	40	20	7	67	68.9	44	20	3	67	74.6
Building	16	202	39	257	83.1	14	213	30	257	81.9
Trees/grass	2	21	153	176	76.9	1	27	148	176	81.8
Total	58	243	199	500	76.3	59	260	181	500	79.4
User accuracy (%)	59.7	78.6	86.9	75.1		65.7	82.9	84.1	77.6	

Support Vector Machine										
Reference	4-Band Image					8-Band Image				
	Roads	Building	Trees + Grass	Total	Producer Accuracy (%)	Roads	Building	Trees + Grass	Total	Producer Accuracy (%)
Roads	32	25	10	67	88.9	35	30	2	67	87.5
Building	4	226	27	257	87.9	5	231	21	257	87.1
Trees/grass	0	6	170	176	82.1	0	4	172	176	88.2
Total	36	257	207	500	86.3	40	265	195	500	87.6
User accuracy (%)	47.8	87.9	96.6	77.4		52.2	89.9	97.7	79.9	

There were no major differences between the classification results when the 4- or the 8-band images were used. The overall accuracies were only slightly higher (up to 3.5%) with the 8-band image for the k -NN and SVM methods and slightly lower (up to 1%) for the two-sample Kolmogorov-Smirnov and Student's t -test classifiers. In the terms of user's accuracy, the classes that benefited the most from additional bands were: roads in the k -NN and SVM classification methods, and buildings in the two-sample Kolmogorov-Smirnov classifier. With the Student's t -test classifier, user accuracies remained unchanged. While the use of the 8-band WorldView-2 image has been recognized to improve classification results [24], this did not prove significant in our study. As seen in Figure 7, a WorldView-2 image with its additional bands (Coastal, Yellow and Red-Edge) provides high separability between different land covers. However, as suggested in [24], the low positive contribution of the extra bands to classification accuracies could be due to the inability of classifiers to handle the pairwise collinearity of bands, meaning that each of the new bands is correlated to one or more of the standard band(s).

After additional attributes were selected for the k -NN and SVM methods, the latter resulted in increased user and producer accuracy (up to 10%), while the performance of the k -NN method did not change significantly (Table 4). A SVM classifier has been reported to perform well with an increased feature dimensionality (number of attributes) [25], but it has also proven to be less sensitive to increases in feature dimensionality compared to other statistical classifiers [26].

Figure 10 represents the classification output images created with the use of the 8-band image. The classification results were not field-verified, thus the Worldview-2 satellite image served as the ground truth image. A visual inspection and comparison of the classification images and the satellite data indicated that more realistic and accurate maps were obtained with two proposed classifiers. It can be seen that with the k -NN and SVM classifiers numerous segments were classified as buildings (approximately 60% of all 1544 segments) and this resulted in a generalized output classification image. A detailed comparison of classification images is provided in Figure 11.

In addition to the presented case study, four algorithms were used to classify 3338 segments in the second highly residential study area measuring 0.25 km² and located 1.5 km west of the primary study area. This classification was conducted using a 4-band Worldview-2 image (Red, Green, Blue and NIR1 bands) and a new set of training samples for four land cover classes: roads (6), buildings (9), trees (7) and grass (9). With the k -NN and SVM classification methods only four spectral attributes were used.

The performance of the classification conducted for the new study area was assessed only visually, *i.e.*, the classification results were compared to the WorldView-2 satellite image, which was used as the reference image. The determination of the reference information for 3338 segments was omitted since it would be influenced by high subjectivity. In terms of output classification images, the classification results were consistent with the previously described results. Overall, the Student's t -test statistics based classifier yielded the most realistic classification image, despite misclassifying several buildings with grey roofs as roads. The k -NN classifier misclassified several roads and other built-up and impervious areas as buildings and some of the buildings with black roofs as trees or grass. The SVM classifier performed poorly when classifying roads, as it classified (generalized) areas of buildings and roads merely as buildings, similarly to what was described for the primary study area. Compared to the Student's t -test classifier, the two-sample Kolmogorov-Smirnov test statistics based classifier classified a higher number of buildings with grey roofs as roads. Trees and grass areas were most accurately

determined with two proposed classifiers. With the use of the two-sample Kolmogorov-Smirnov statistics based classifier and Student's t -test statistics based classifier 120 segments (3% of all segments) remained unclassified.

Figure 10. The final classification image for all four classifiers related to the 8-band classification results presented in Table 4: (a) k -NN; (b) SVM; (c) the two-sample Kolmogorov-Smirnov statistics based classifier; (d) the Student's t -test statistics based classifier.

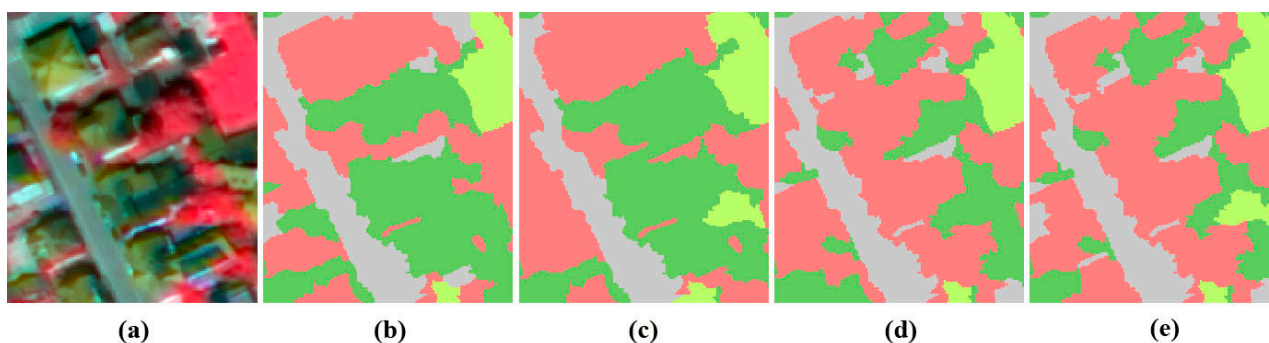


The high computing complexity and low speed performance are considered to be the major drawbacks of the proposed classification approach. On average the Student's t -test based algorithm needs 1.5 s to classify one segment with a 4-band image and 3 s to classify one segment with an 8-band image. Using the two-sample Kolmogorov-Smirnov test based algorithm, it takes an average of 5 s to classify one segment with a 4-band image and 13 s to classify one segment with an 8-band image. Both features are a result of the data structure organization of the computed variables in the IDL programming language

and could be improved significantly if the sampling approach and the computation of the empirical distribution functions were altered.

Merely a few approaches in literature have been found to deal with the analysis of the impact the in-segment investigation has on the object-based classification results. Toure *et al.* [11] reported that curve matching classifiers yielded higher classification accuracies than the nearest neighbor classifier, however, the overall accuracies of all classifiers were still moderate (between 27% and 73%). Sridharan and Qiu [8] compared the effectiveness of the Fuzzy Kolmogorov-Smirnov classifier to six other classifiers (k -NN, SVM, Spectral Angle Mapper, Maximum Likelihood classifier, Minimum Distance to Mean classifier and Parallelepiped classifier). The Fuzzy Kolmogorov-Smirnov classifier performed substantially better than the remaining classifiers. Its overall accuracy was higher by at least 10%.

Figure 11. The detailed subset of classification images represented in Figure 10 in relation to the satellite image that was used as an input for the classification process and was also considered as the ground truth image: (a) pan-sharpened Worldview-2 image (shown as a false color composite); (b) k -NN classifier; (c) SVM classifier; (d) the two-sample Kolmogorov-Smirnov statistics based classifier; (e) the Student's t -test statistics based classifier. Images reveal higher classification accuracies of the two proposed classifiers compared to the ground truth image.



The performance of the proposed classification approach was not compared to machine learning and tree-based algorithms, which are considered as advanced classification techniques and are becoming increasingly popular in the GEOBIA community. Extensive analysis conducted in [10] and [24] indicate that high classification accuracies can be achieved with the Random Forest classifier when applied to the supervised object-based approach; however, the high accuracy was to a great extent dependent on the training set representativeness and the quality of segmentation [24]. In spite of high classification accuracies, these two studies cannot serve for the straightforward comparison, since in [10], the Landsat TM data with a coarse spatial resolution was used and in [24] segments were manually delineated.

Although the k -NN and SVM methods provided less accurate results with the use of the 4-band image, this does not imply that these are inefficient classification methods. It has been suggested [27] that selecting the “best” classifier is not necessarily the wisest decision, since valuable information may be produced by classifiers that are considered to be less successful. Taking into account that a specific classification metric is better adapted to a specific situation, all classifiers can have the potential of being the “best” for a certain situation. Therefore, image understanding and active learning techniques tend to be implemented and used [1]. Image understanding is a complex cognitive process for which we may

currently lack the key concepts. Therefore, in order to ensure a strong foundation, research in this field needs to focus on the deeper understanding of the relationship between image objects and landscape objects, rather than exclusively on the development of new techniques that are tailored for specific applications [28] and rarely transferable [3].

4. Conclusions

We have presented an approach for addressing the in-segment complexity that would enhance the object-based classification performance of high-resolution satellite data. The in-segment analysis was investigated with the use of multiple small sized samples—10 pixels—with two proposed classifiers: the two-sample Kolmogorov-Smirnov test and the Student's *t*-test based statistics. The performance and capability of the approach was assessed on a WorldView-2 image for four urban land cover classes (roads, buildings, grass and trees) and compared to two commonly used object-based classifiers—*k*-Nearest Neighbor (*k*-NN) and Support Vector Machine (SVM). The analysis was conducted using a 4- as well as an 8-band image.

A highly residential area, riddled with a road network and a variety of roofing material, to the west of the Ljubljana center (Slovenia) was selected as the study region. In the study, 1544 segments were defined at a single segmentation level, out of which almost one third (500) of all segments was selected for the classification performance evaluation.

With the use of 4 spectral bands, both proposed classifiers showed improvements in the overall classification accuracy when compared to the well-established *k*-NN and SVM classifiers. The parametric Student's *t*-test yielded the highest user and producer classification accuracies in general. Buildings were most accurately classified with the use of the SVM classification method. All four classifiers were most likely to misclassify buildings and roads, which was a result of using merely spectral attributes. Mere spectral information using only 4 spectral bands is insufficient when attempting to separate objects made from similar material.

The use of empirical distribution functions showed merely a small increase in the classification accuracy. This could be due to the high generalization of the empirical distribution in the event when only 10 pixels are sampled. It is impossible for too generalized empirical distribution functions to describe the in-segment unique and complex distribution.

There were no major differences between the classification accuracies when a 4- or 8-band image was used. The low positive contribution of the extra bands in the WorldView-2 image could be a result of the classifier's inability to handle the collinearity of bands and the redundant information. As regards the output classification images, two classifiers with a proposed sampling approach yielded visually more appealing and more accurate classification results compared to the input satellite image, which was considered to be the ground truth image.

Further research will aim to investigate the various aggregation methods for multiple probability values computed within the sampling process as well as provide a different approach of empirical distribution functions analysis that would enhance the classification capabilities. The sampling approach may have a greater potential, however no *p*-value averaging should be included in the last stage of the two-sample Kolmogorov-Smirnov and Student's *t*-test statistics. In our study, the adopted averaging approach with the two proposed classifiers resulted in a similar output as the summary statistics based

methods that have been proven to fail in capturing unique segments' distributions. The idea that needs to be addressed—also in relation to improving the computation speed of the algorithms implemented in the IDL programming language—is the comparison of an equal number of subparts with two empirical distribution functions. This would mean that all pixels were used at once and no sampling would be necessary. With additional studies, which would include a larger number of more diversified study areas, the benefits and drawbacks of the sampling approach and the use of the empirical distribution functions could be more generalized and transferable within different classification applications (agriculture, land cover determination).

Acknowledgments

The Slovenian Centre of Excellence for Space Sciences and Technologies SPACE-SI is an operation partly financed by the European Union, European Regional Development Fund and Republic of Slovenia, Ministry of Higher Education, Science and Technology. The authors would like to thank all four reviewers for their thorough and constructive comments and suggestions that helped improve the quality of the manuscript.

Author Contributions

Nataša Đurić designed the methodology, conducted the experiments and wrote the paper. Peter Pehani helped by implementing the proposed classification approach in IDL programming language. Krištof Oštir supervised the research and contributed to the editing and review of the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Baraldi, A.; Boschetti, L. Operational automatic remote sensing image understanding systems: beyond Geographic Object-Based and Object-Oriented Image Analysis (GEOBIA/GEOOIA). Part 1: Introduction. *Remote Sens.* **2012**, *4*, 2694–2735.
2. Blaschke, T. Object based image analysis for remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 2–16.
3. Arvor, D.; Durieux, L.; Andrés, S.; Laporte, M-A. Advances in geographic object-based image analysis with ontologies: A review of main contributions and limitations from a remote sensing perspective. *ISPRS J. Photogramm. Remote Sens.* **2013**, *82*, 125–137.
4. Dos Santos, J.A. Semi-Automatic Classification of Remote Sensing Images. Ph.D. Thesis, Universidade Estadual de Campinas, Campinas, Brasil, 30 April 2013.
5. Qian, J.; Zhou, Q.; Hou, Q. Comparison of pixel-based and object-oriented classification methods for extracting built-up areas in Aridzone. In Proceedings of the 2007 ISPRS Workshop on Updating Geo-Spatial Databases with Imagery & The 5th ISPRS Workshop on Dynamic and Multi-Dimensional GIS, Urumchi, China, 28–29 August 2007.

6. Kressler, F.P.; Steinnocher, K.; Franzen, M. Object-oriented classification of orthophotos to support update of spatial databases. In Proceedings of the 2005 IEEE International Geoscience and Remote Sensing Symposium, Seoul, Korea, 25–29 July 2005; pp. 253–256.
7. Marpu, P.R.; Neubert, M.; Herold, H.; Niemeyer, I. Enhanced evaluation of image segmentation results. *J. Spat. Sci.* **2010**, *55*, 55–68.
8. Sridharan, H.; Qiu, F. Developing an object-based hyperspatial image classifier with a case study using WorldView-2 data. *Photogramm. Eng. Remote Sens.* **2013**, *79*, 1027–1036.
9. Lübcker, T.; Schaab, G. Optimization of parameter settings for multilevel image segmentation in Geobia. In Proceedings of the 2009 ISPRS Hannover Workshop High-Resolution Earth Imaging for Geospatial Information, Hannover, Germany, 2–5 June 2009.
10. Li, C.; Wang, J.; Wang, L.; Hu, L.; Gong, P. Comparison of classification algorithms and training sample sizes in urban land classification with Landsat thematic mapper imagery. *Remote Sens.* **2014**, *6*, 964–983.
11. Toure, S.; Stow, D.A.; Weeks, J.R.; Kumar, S. Histogram curve matching approaches for object-based image classification of land cover and land use. *Photogramm. Eng. Remote Sens.* **2013**, *79*, 433–440.
12. Đurić, N.; Pehani, P.; Oštir, K. The within and between segment separability analysis of sealed urban land use classes. In Proceedings of the 2013 Simpósio Brasileiro de Sensoriamento Remoto, Foz do Iguaçu, PR, Brasil, 13–18 April 2013; pp. 2392–2399.
13. Pal, N.R.; Pal, S.K. A review on image segmentation techniques. *Pattern Recognit.* **1993**, *26*, 1277–1294.
14. Mesner, N.; Oštir, K. Investigating the impact of spatial and spectral resolution of satellite images on segmentation quality. *J. Appl. Remote Sens.* **2014**, *8*, 083696.
15. Lizarazo, I.; Elsner, P. Improving urban land cover classification using fuzzy image segmentation. *Lect. Notes Comput. Sci.* **2009**, *5730*, 41–56.
16. List of Attributes. Available online: <http://www.exelisvis.com/docs/AttributeList.html> (accessed on 19 August 2014).
17. K Nearest Neighbor Background. Available online: <http://www.exelisvis.com/docs/BackgroundKNN.html> (accessed on 20 August 2014).
18. Support Vector Machine Background. Available online: <http://www.exelisvis.com/docs/BackgroundSVM.html> (accessed on 21 August 2014).
19. Tzotsos, A. A support vector machine approach for object based image analysis. In *Bridging Remote Sensing and GIS*, Proceedings of the 2006 International Conference on Object-based Image Analysis, Salzburg, Austria, 4–5 July 2006.
20. Gibbons, J.D.; Chakraborti, S. *Nonparametric Statistical Interference*, 4th ed.; Marcel Dekker Inc.: New York, NY, USA, 2003; p. 37.
21. Wang, H. Stochastic Modelling of the Equilibrium Speed-Density Relationship. Ph.D. Thesis, University of Massachusetts Amherst, Amherst, MA, USA, 1 September 2010.

22. Gordon, A.Y.; Klebanov, L.B. On a paradoxical property of the Kolmogorov-Smirnov two-sample test. In *Nonparametrics and Robustness in Modern Statistical Inference and Time Series Analysis: A Festschrift in Honor of Professor Jana Jurečková*; Antoch, J., Hušková, M., Sen, P.K., Eds.; Institute of Mathematical Statistics: Beachwood, OH, USA, 2010; Volume 7, pp. 70–74.
23. Arcidiacono, C.; Porto, S.M.C.; Cascone, G. Accuracy of crop-shelter thematic maps: A case study of maps obtained by spectral and textural classification of high-resolution satellite image. *J. Food Agric. Environ.* **2012**, *10*, 1071–1074.
24. 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.
25. Huang, C.; Davis, L.S.; Townshend, J.R.G. An assessment of support vector machines for land cover classification. *Int. J. Remote Sens.* **2002**, *23*, 725–749.
26. Myburgh, G.; van Niekerk, A. Effect of feature dimensionality on object-based land cover classification: A comparison of three classifiers. *S. Afr. J. Geomat.* **2013**, *2*, 13–27.
27. Maillard, P.; Clausi, A.D. Pixel based sea ice classification using the MAGSIC system. In Proceedings of the 2006 ISPRS Commission VII Symposium “Remote Sensing: From Pixels to Processes”, Enschede, The Netherlands, 8–11 May 2006.
28. Castilla, G.; Hay, J.G.; Ruiz-Gallardo, J. Size-constrained Region Merging (SCRM): An automated delineation tool for assisted photointerpretation. *Photogramm. Eng. Remote. Sens.* **2008**, *74*, 409–419.

© 2014 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 license (<http://creativecommons.org/licenses/by/4.0/>).