

Article

Context-Enabled Extraction of Large-Scale Urban Functional Zones from Very-High-Resolution Images: A Multiscale Segmentation Approach

Shouji Du, Shihong Du *, Bo Liu and Xiuyuan Zhang

Institute of Remote Sensing and GIS, Peking University, Beijing 100871, China

* Correspondence: shdu@pku.edu.cn; Tel.: +86-010-6275-0294

Received: 15 July 2019; Accepted: 10 August 2019; Published: 14 August 2019



Abstract: Urban functional-zone (UFZ) analysis has been widely used in many applications, including urban environment evaluation, and urban planning and management. How to extract UFZs' spatial units which delineates UFZs' boundaries is fundamental to urban applications, but it is still unresolved. In this study, an automatic, context-enabled multiscale image segmentation method is proposed for extracting spatial units of UFZs from very-high-resolution satellite images. First, a window independent context feature is calculated to measure context information in the form of geographic nearest-neighbor distance from a pixel to different image classes. Second, a scale-adaptive approach is proposed to determine appropriate scales for each UFZ in terms of its context information and generate the initial UFZs. Finally, the graph cuts algorithm is improved to optimize the initial UFZs. Two datasets including WorldView-2 image in Beijing and GaoFen-2 image in Nanchang are used to evaluate the proposed method. The results indicate that the proposed method can generate better results from very-high-resolution satellite images than widely used approaches like image tiles and road blocks in representing UFZs. In addition, the proposed method outperforms existing methods in both segmentation quality and running time. Therefore, the proposed method appears to be promising and practical for segmenting large-scale UFZs.

Keywords: very-high-resolution image; urban functional zone; multiresolution segmentation; GEOBIA; urban land use

1. Introduction

Urban socioeconomic activities demonstrate strong clustering patterns in space, leading to the generation of various urban functional zones (UFZs) to accommodate people's diverse needs for living, working, education, recreation, and public service [1,2]. UFZs act as an important space carrier to implement urban economic and social functions, thus they are typically employed as the basic units in urban planning and management [3–5]. There has been increasing demand for fine-grained, large-scale, accurate, and timely UFZ maps as they can assist municipal administrations in monitoring urban growths and creating sustainable urban planning during ongoing urbanization processes.

Figure 1 illustrates an urban area comprised of multiple types of UFZs, including a campus, seven residential districts, two parks, a commercial zone, a shantytown, and an industrial zone. There are two issues challenging the production of an accurate functional-zone map: how to generate the spatial units of UFZs? And how to assign the units to appropriate UFZs categories? The solution to the former issue is required before addressing the latter. While relying on human photo interpretation and field surveys to delineate UFZ boundaries has been a traditional practice, they are labor intensive and time consuming [6,7]. Recent advancement in analyzing very-high-resolution (VHR) remote sensing provides an alternative to generate the spatial units of UFZs, which takes advantage of the spectral and

spatial patterns of geographic objects that vary across various types of UFZs [8]. Figure 1 shows that there are significant differences in multiple types of UFZs at both visual features (spectrums, textures, and shapes) and spatial configuration of geographic objects. Therefore, automatically generating spatial units for UFZs from VHR images is still possible but rarely studied, thus it is selected as the purpose of this study.

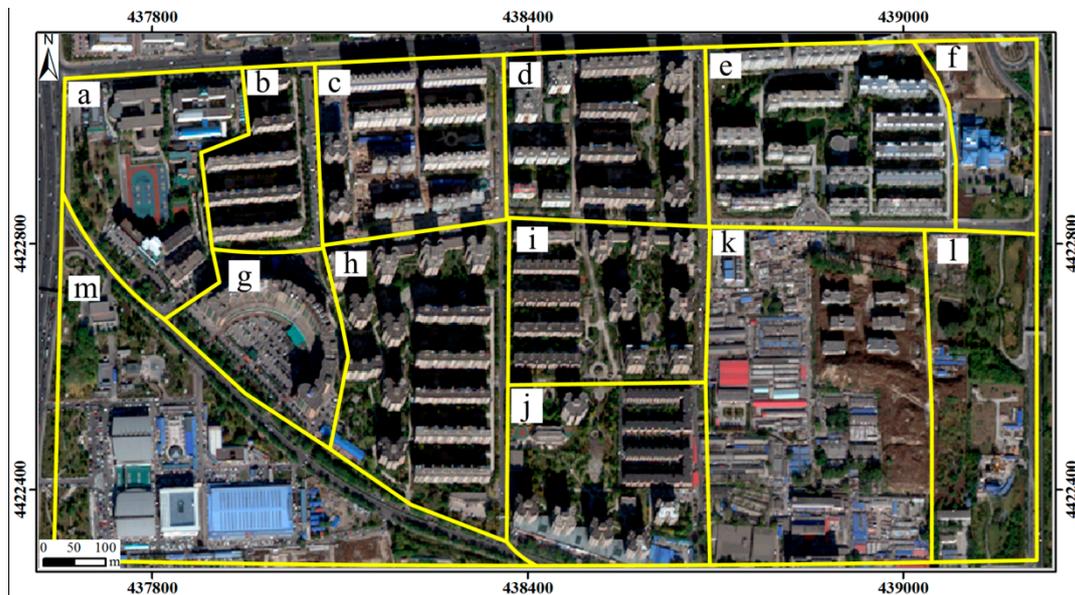


Figure 1. Urban functional zones (UFZs) outlined by yellow lines in a very-high-resolution (VHR) image. (a): A campus; (b–e,h–j): residential districts; (f,l): parks; (g): a commercial zone; (k): a shantytown; and (m): an industrial zone.

Existing methods for functional-zone analyses mainly rely on two spatial units: image tiles (e.g., UC Merced Land Use Dataset, <http://weegee.vision.ucmerced.edu/datasets/landuse.html>) and road blocks. Image tiles are rectangles and simple to use, but they cannot represent the exact UFZs with various shapes and sizes (see Figure 1) [2,9–12]. With urban road data becoming more available (e.g., OpenStreetMap), recent studies have increasingly used road vectors to divide VHR images into UFZs [1,3,5,7,13,14]. However, UFZs of this kind are likely to be under-segmented because UFZs are not necessarily separated by roads in reality. The accuracy of segmenting UFZs is highly influenced by the completeness and the quality of road vectors. Although existing polygon features (e.g., housing boundary data) can supplement image tiles and road blocks to improve UFZ segmentation, they are often not accessible or not up-to-the-date in developing and underdeveloped regions [15,16]. Therefore, there is still a lack of automatic method for generating spatial units of UFZs from VHR images.

Image segmentation techniques should be addressed to generate UFZs directly from VHR images since pixels, image tiles and road blocks are inadequate to represent UFZs [8,17–22]. Existing widely used segmentation methods include object segmentation [23–28] and semantic segmentation [29–32]. Object segmentation is designed to obtain homogenous image objects like buildings and roads rather than heterogeneous patches. On the other hand, semantic segmentation methods were presented to obtain more accurate and complete geographic objects than those from object segmentation in the form of per-pixel classification. However, neither object segmentation nor semantic segmentation can generate exact UFZs. First, their purposes are to obtain homogeneous objects with consistent visual cues (spectrums, textures, and shapes) while UFZs are heterogeneous patches and composed of diverse geographic objects with substantial discontinuities in visual cues. Second, for UFZs, spatial patterns are also essential for generating UFZs, while both object segmentation and semantic segmentation only consider visual features or deep features, but ignore spatial patterns.

Some related studies have focused on complex patterns and land use extraction from VHR images. The method, named window independent context segmentation, was presented to generate different urban categories (such as city center and industry) [33–35]. This method first extracts context features to characterize spatial relations among different categories of objects, and then clusters the pixels with similar context features into an urban category. However, the method is only designed to identify different urban area categories rather than segmenting UFZs. Zhang et al. [8] presented a multi-level aggregation approach considering both object features and spatial pattern features to achieve geoscene segmentation. However, this approach requires conducting image object segmentation and classification. As a result, both segmentation and classification not only require much more computation time and resources and greatly rely on prior information, but also their accuracies lead to substantial influences on UFZ segmentation results. In one word, an efficient and automatic method, with less computation time and resources and prior information, for segmenting UFZs from VHR images is still open and needs to be resolved urgently.

As shown in Figure 1, UFZs are often composed of diverse objects with different spectral features, thus spectral features are hardly able to characterize UFZs alone. A UFZ holds a specific social and economic activity, and meanwhile, it typically has a specific spatial configuration of geographic objects. For example, Figure 1 shows that buildings are neatly arranged with similar sizes and structures in residential districts, although they are small and tightly connected in shantytowns. A school campus usually contains different building structures and one or more sports fields. A park may contain large areas of green space, while a commercial zone typically has big buildings with oversized open spaces for parking. Accordingly, UFZs tend to be defined by their context information which reflects the spatial configurations between geographic objects. The window independent context (WIC) feature reported in the work of Nielsen [33] refers to the nearest-neighbor distance from a specific pixel to pixels of different classes providing the context information for each individual pixel. Accordingly, it is useful to define context information for different UFZs. Furthermore, segmentation scale controls the segmentation quality and bridges the gap between pixels of VHR images and UFZs. The heterogeneity is strong and varies over different urban areas, requiring that segmentation scales of UFZs must be variant in different areas.

In summary, a multiscale image segmentation method is proposed in this work to segment UFZs from VHR images. Four contributions are made in this study: (1) a context feature is explored for UFZ segmentation and its effectiveness is discussed in detail; (2) a scale-adaptive method is proposed to determine segmentation scales over different urban areas for best delineating UFZs; (3) a graph cuts algorithm is improved to optimize the UFZ segmentation results; and (4) to our best knowledge, this is the first method for large-scale, efficient, and automated UFZ segmentation in this field.

2. Methodology

The input of the proposed UFZ segmentation method is only the remote sensing image. The workflow (Figure 2) includes the following four steps.

(1) WIC feature calculation. Context features are explored for segmenting UFZs as they can measure the spatial patterns of diverse objects in urban areas. For extracting WIC features, the VHR image is first clustered into groups of pixels by ISODATA [36], and then WIC features are computed for each pixel in the form of geographic nearest-neighbor distance from a pixel to different groups of pixels.

(2) Image objects segmentation. Geographic objects with homogeneous WIC features will make up a UFZ, thus the original VHR image is segmented into objects as the initial units of UFZ segmentation using multiresolution segmentation (MRS) [23].

(3) UFZ segmentation. Image objects are further merged to produce UFZs with homogeneous WIC features, i.e., spatial patterns of diverse objects. To do so, a scale-adaptive method is presented to determine the appropriate scales and obtain the initial UFZs by merging image objects in terms of WIC feature.

(4) UFZ optimization. The graph cuts algorithm is adopted to achieve globally optimal results considering the similarities between adjacent objects.

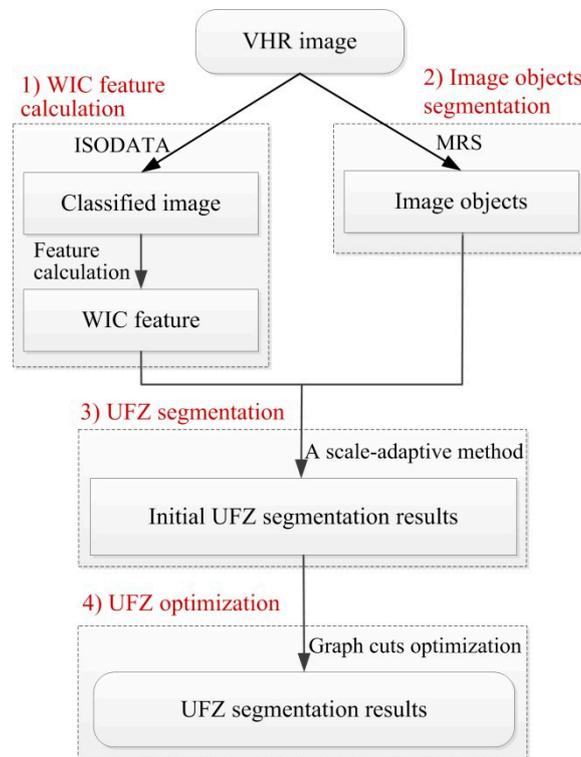


Figure 2. Workflow of the proposed UFZ segmentation method.

2.1. WIC Feature Calculation

The WIC feature refers to the nearest-neighbor distances from a specific pixel to pixels of different classes and provides the context information for each individual pixel. Accordingly, it can be used to define spatial contexts for different UFZs. Figure 3 shows the diagram of the WIC feature calculation, including the following two steps.

(1) The original image is first classified by ISODATA. A total of 20 classes are generated in this study because it is enough not only to ensure the computational efficiency to perform the analysis, but also to retain enough variance in the input data.

(2) For each individual pixel, the nearest-neighbor distances to all the 20 classes are calculated and recorded. Therefore, the WIC feature of each pixel is a vector with 20 elements, each of which represents a nearest distance to one of the classes. The WIC vector can be expressed as:

$$F_{p(x,y)} = \left\{ \begin{array}{c} d_{c_1} \\ d_{c_2} \\ \vdots \\ d_{c_i} \\ \vdots \\ d_{c_{20}} \end{array} \right\}, \quad (1)$$

$$d_{c_i} = \min(\text{dist}(p(x,y), c_i)), \quad (2)$$

where d_{c_i} , \min and dist respectively refer to the nearest-neighbor distance, the minimum distance, and the Euclidian distance between pixel $p(x,y)$ and a pixel belonging to a specific class c_i .

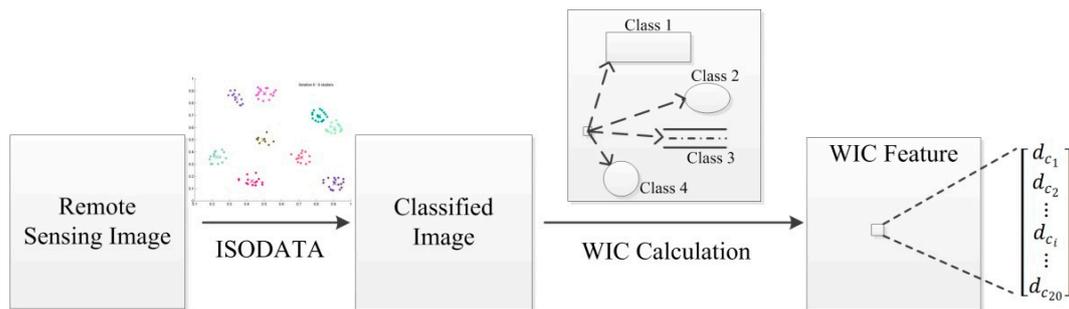


Figure 3. The diagram of window independent context (WIC) feature calculation.

2.2. UFZ Segmentation

UFZs tend to be defined by their context information which reflects the spatial configurations between geographic objects. Accordingly, geographic objects with homogeneous context features can make up a UFZ. Therefore, the original image is segmented into objects using MRS and then objects will be further merged to generate UFZs in terms of WIC features. For merging image objects, two issues need to be determined: the heterogeneity between two image objects and the merging strategy, i.e., in the merging procedure, how the two objects are merged when their heterogeneity meets the merging criterion.

2.2.1. UFZ Heterogeneity

The heterogeneity between two image objects is determined by the heterogeneity increase, which refers to the changes of heterogeneity before and after merging two objects. The WIC standard deviation and area (i.e., pixel number) of two adjacent objects are denoted by $\sigma_{1,i}$, $\sigma_{2,i}$, n_1 , and n_2 , and WIC standard deviation and area of the merged object are denoted as $\sigma_{\text{merg},i}$ and n_m . Then WIC heterogeneity increase can be defined as:

$$h_{\text{WIC}} = \sum_i w_i (n_m \sigma_{\text{merg},i} - (n_1 \sigma_{1,i} + n_2 \sigma_{2,i})), \quad (3)$$

where i refers to i -th dimension of WIC feature, while w_i refers to the weight of i -th dimension of WIC feature.

Shape heterogeneity increase of two objects is also considered. It can be represented as two shape indices: smoothness and compactness which can be defined as $\frac{1}{b}$ and $\frac{1}{\sqrt{n}}$, where l is the perimeter of the object, b is the perimeter of the smallest bounded rectangle of the object, and n is the area of the object. Similar to WIC heterogeneity increase, the smoothness and compactness heterogeneity increases are defined as:

$$h_{\text{smooth}} = n_{\text{merg}} \frac{l_{\text{merg}}}{b_{\text{merg}}} - \left(n_1 \frac{l_1}{b_1} + n_2 \frac{l_2}{b_2} \right), \quad (4)$$

$$h_{\text{com}} = n_{\text{merg}} \frac{l_{\text{merg}}}{\sqrt{n_{\text{merg}}}} - \left(n_1 \frac{l_1}{\sqrt{n_1}} + n_2 \frac{l_2}{\sqrt{n_2}} \right), \quad (5)$$

and then shape heterogeneity increase can be obtained:

$$h_{\text{shape}} = w_{\text{smooth}} \times h_{\text{smooth}} + (1 - w_{\text{smooth}}) \times h_{\text{com}}, \quad (6)$$

where w_{smooth} is the weight of smoothness in shape heterogeneity increase.

After calculating h_{WIC} and h_{shape} , the heterogeneity increase is defined as:

$$f = w_{\text{WIC}} \times h_{\text{WIC}} + (1 - w_{\text{WIC}}) \times h_{\text{shape}}, \quad (7)$$

where w_{WIC} is the weight of WIC feature in the heterogeneity increase. f is used as the index for whether or not to merge two objects.

2.2.2. UFZ Segmentation: A Scale-Adaptive Method

Image objects merging is achieved in iterative procedures. In an iteration procedure, each object will be judged whether it can be merged with an object around it (if one object has been merged, it does not need to be judged). For one object, if the minimal heterogeneity increase between it and its neighboring objects meets the merging criterion (i.e., the segmentation scale), this object will be merged with the neighboring object with minimal heterogeneity increase to form a new object; otherwise, nothing is done for this object and another object will be chosen for further processing. After all objects are processed, this iteration is completed and next iteration will be executed, in which objects that were not merged in the previous iteration and all the new objects will be processed, until no objects are merged.

The key issue is to determine the UFZ segmentation scale, which represents the largest tolerant heterogeneity of segmented UFZs. Since the WIC heterogeneity of UFZs varies over different urban areas, segmentation scales should also vary in terms of the WIC heterogeneity. Actually, it is difficult to automatically determine the segmentation scales. For UFZ segmentation, a good solution is to use different scales in different urban areas (e.g., residential districts and parks) because different types of UFZs usually have different landscape components and structures. However, this will depend on additional classification information. Therefore, a scale-adaptive method is proposed in this study based on WIC feature.

According to the definition of WIC feature, the larger a UFZ will, the larger its WIC feature. The reason is that the pixels, especially central ones, will be far from other class pixels, leading to a large heterogeneity of WIC feature. Accordingly, this fact is helpful to determine appropriate scales for different urban areas. The idea is that the areas with large WIC values should be segmented by a large scale. The proposed adaptive scale is defined as follows:

$$S = \begin{cases} S_{set} \times \frac{d_{ij}}{d_m}, & \text{if}(d_i > d_{uq} \& d_j > d_{uq}) \\ S_{set}, & \text{otherwise} \end{cases}, \quad (8)$$

where S_{set} is the initial segmentation scale specified by users; d_i and d_j are the mean WIC values of current objects i and j respectively; d_{ij} is the mean WIC value of the merged object from objects i and j ; d_m is the median of WIC distribution in the whole image; and d_{uq} is the upper quartile of WIC distribution. Since the distribution of WIC values is positively skewed, the median and upper quartiles of the numerical distribution are adopted.

For the proposed scale-adaptive method, the scale parameter S will be adjusted according to whether the mean WIC values of current objects (i.e., d_i and d_j) are larger than the given threshold (i.e., d_{uq}) or not. Namely, if they are smaller than the given threshold of the WIC feature, the initial segmentation scale S_{set} will be used; otherwise, the scale will be enlarged proportionally to d_m . Accordingly, the larger d_{ij} is, the larger the segmentation scale will be. For example, for an object to be merged, its neighboring object with minimal heterogeneity increase is first found, then the adaptive scale used for these two objects are determined by Equation (8). If their heterogeneity increase is less than the calculated scale S , they will be merged; otherwise, nothing is done for this object and another object will be chosen for further processing. Consequently, the initial UFZ segmentation results will be generated.

2.3. UFZ Optimization

Two issues are still not resolved in the initial UFZ segmentation results. First, the above iteration process uses locally optimal strategy and ignores the global information, leading to locally optimal results in some areas. Second, objects, that are significantly different in WIC feature from their

neighboring objects will be segmented into one individual UFZ. Therefore, post-processing is needed to optimize the initial UFZ segmentation results. In our method, the graph cuts algorithm is adopted because its objective is to achieve globally optimal results considering similarities between adjacent objects [37–39].

Given a set of nodes P and a finite set of labels L , the goal of graph cuts is to assign each node $p \in P$ a label $l_p \in L$ by minimizing the energy function as defined as follows:

$$E(l) = \sum_{p \in P} D_p(l_p) + \lambda \sum_{\{p,q\} \in N} w_{(p,q)} \times V_{(p,q)}(l_p, l_q), \tag{9}$$

where l is the collection of all label assignments; $\sum_{p \in P} D_p(l_p)$ and $\sum_{\{p,q\} \in N} w_{(p,q)} \times V_{(p,q)}(l_p, l_q)$ denote unary and pairwise terms respectively, balanced by λ . $D_p(l_p)$ measures how well label l_p fits node p ; $w_{(p,q)}$ defines the similarity between p and q ; N is the collection of neighboring node pairs. $V_{(p,q)}(l_p, l_q)$ is defined as:

$$V_{(p,q)}(l_p, l_q) = \begin{cases} 0, & \text{if } (l_p = l_q) \\ 1, & \text{if } (l_p \neq l_q) \end{cases} \tag{10}$$

To construct energy function, nodes, labels and their connection relationships need to be determined. In this study, image objects segmented by step (2) (the blue segments in Figure 4) and initial UFZ units segmented by step (3) (the purple segments in Figure 4) are regarded as nodes and labels respectively. The connection relationships between nodes and labels can be obtained in terms of the inclusion relations between image objects and initial UFZs. In our method, all the initial UFZs will grow in space in order to make image objects located at the boundaries of UFZs correspond to several labels. Therefore, for one initial UFZ, its label will be assigned to the image objects that belong to this UFZ and those which are two-order neighborhoods of this UFZ. For example, in Figure 4, image objects I_0, I_1 are the two-order neighborhood of initial UFZ U_1 , thus the label of U_1 will also be assigned to I_0 and I_1 . After graph cuts optimization, the label of each image object will be redistributed, and all the adjacent image objects with the same label will make up a final UFZ (the yellow segments in Figure 4). Accordingly, the graph cuts optimization can be briefly illustrated in Figure 4e (supposing only three UFZs).

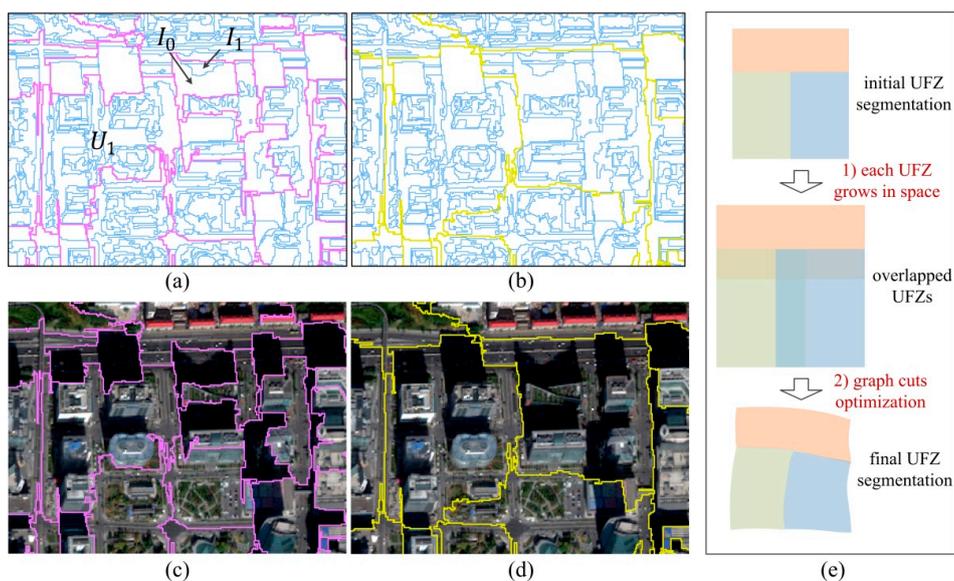


Figure 4. The diagram of energy function construction (the blue polygons refer to the image objects segmented by step (2); the purple polygons refer to the initial UFZs segmented by step (3), while the yellow polygons refer to the final UFZs). (a) Image objects and initial UFZs; (b) image objects and final UFZs; (c) initial UFZs; (d) final UFZs; and (e) brief illustration of graph cuts optimization.

The unary term, indicating the corresponding relationships between the nodes and the labels, can be obtained according to the above explanation. Let $S(l_i)$ denote the set of image objects covered by the initial UFZ U_i (containing image objects which are two-order neighborhood of U_i , such as I_0, I_1 for U_1 in Figure 4), then the unary term is defined as:

$$D_p(l_i) = \begin{cases} 1, & \text{if}(p \in S(l_i)) \\ \infty, & \text{otherwise} \end{cases} \quad (11)$$

The pairwise term denotes the similarity between objects p and q , and it is defined as:

$$w_{(p,q)} = \exp\left(-\frac{f^2}{\text{dist}(p,q) \times 2\sigma^2}\right), \quad (12)$$

where $f = w_{WIC} \times h_{WIC} + (1 - w_{WIC}) \times h_{\text{shape}}$ refers to the heterogeneity increase after merging p and q , $\text{dist}(p, q)$ is the Euclidean distance between objects p and q , and σ is used to adjust the value range of the pairwise term. In our method, heterogeneity increase after merging p and q is used in the pairwise term considering both WIC and shape heterogeneity increase, which will help to keep good edge information of UFZs.

After constructing the energy function, α – expansion algorithm is employed to minimize the energy function [37,38,40]. Finally, each image object will obtain a new label, then all the objects with the same label will make up a UFZ.

2.4. Segmentation Evaluation

To quantitatively evaluate image segmentation results, supervised [41,42] and unsupervised methods [43,44] are proposed. The supervised methods are more objective for evaluating the image segmentation quality [41]. In this study, object-level consistency error (OCE) is employed to evaluate the segmentation quality [41]. OCE takes into account the existence, size, shape, and position of each object. Compared with existing error measures, OCE works at object level and considers both over- and under-segmentation [41]. In addition, it retains the properties of being normalized, i.e., $0 \leq \text{OCE}(I_s, I_g) \leq 1$ and $\text{OCE}(I_s, I_g) = 0$ only if $I_s = I_g$, where I_s and I_g are the segmented and the ground truth images. A smaller OCE indicates a better segmentation result. According to the experiments in the work of Polak et al. [41], segmentation results with an OCE value smaller than 0.55 can be typically considered satisfactory. In our evaluation procedure, each UFZ is regarded as one object.

3. Results

3.1. Experimental Data

To evaluate the effectiveness of the proposed UFZ segmentation method, a study area located in Beijing was used to perform the experiments (Figure 5) and it covered approximately 36 km². Beijing, the capital of China, is a rapidly developing city at a high urbanization level and contains diverse archaic and modern zones with highly disparate architectural styles, making it difficult but valuable to study UFZs. The WorldView-2 image with multispectral bands was employed in the experiments. The image resolution was 2-meters and the size was 3000 × 3000 pixels.

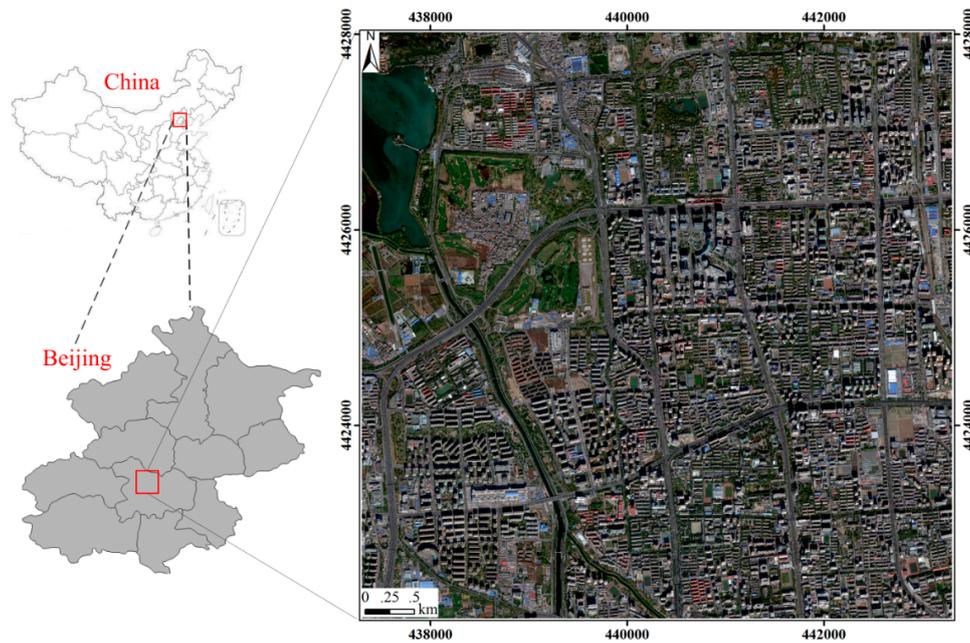


Figure 5. The study area in Beijing.

3.2. Results of UFZ Segmentation

The ISODATA classified image and WIC feature image (R, G, B assigned with feature bands 1, 2, and 3, i.e., the nearest-neighbor distance to classes 1, 2 and 3) are shown in Figure 6.

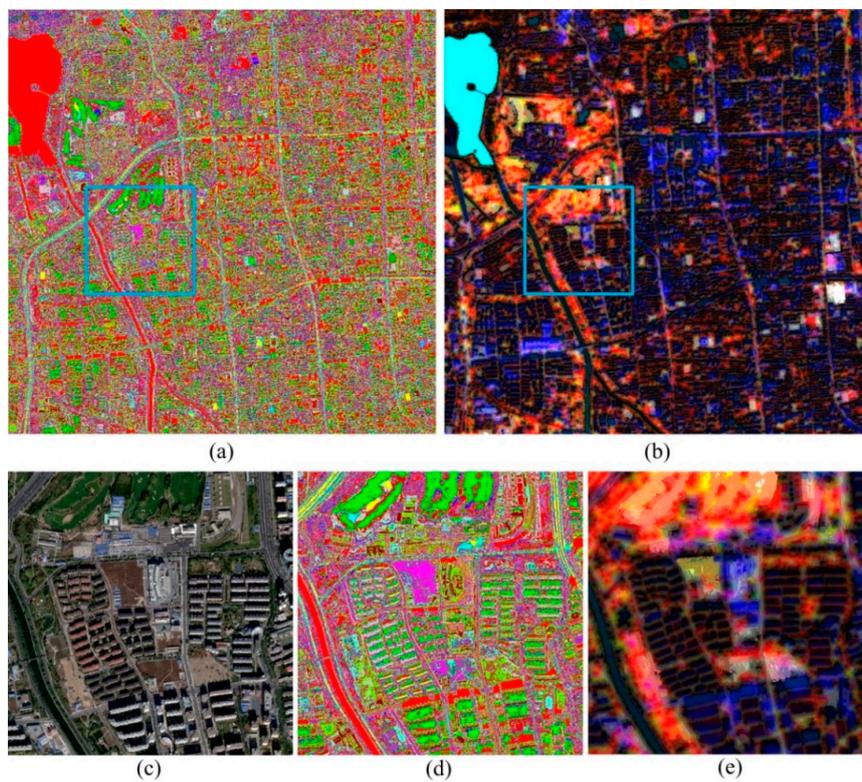


Figure 6. The ISODATA classified image and WIC feature image. (a) The ISODATA classified image; (b) the WIC feature image (R, G, and B assigned with feature bands 1, 2, and 3 respectively); (c) amplified view of original image of blue rectangles in (a,b); (d) amplified view of ISODATA classified image; (e) amplified view of WIC feature image.

Considering that the purpose of ISODATA is to reduce the dimensions of the image to a manageable size instead of generating an accurate land-cover map [35], the clustering results are enough for WIC feature calculation. The WIC feature contains spatial patterns of ground objects especially for buildings. Moreover, different UFZs are characterized with different types of WIC features, thus WIC features can describe the spatial configurations of different UFZs. The UFZ segmentation results are shown in Figure 7.

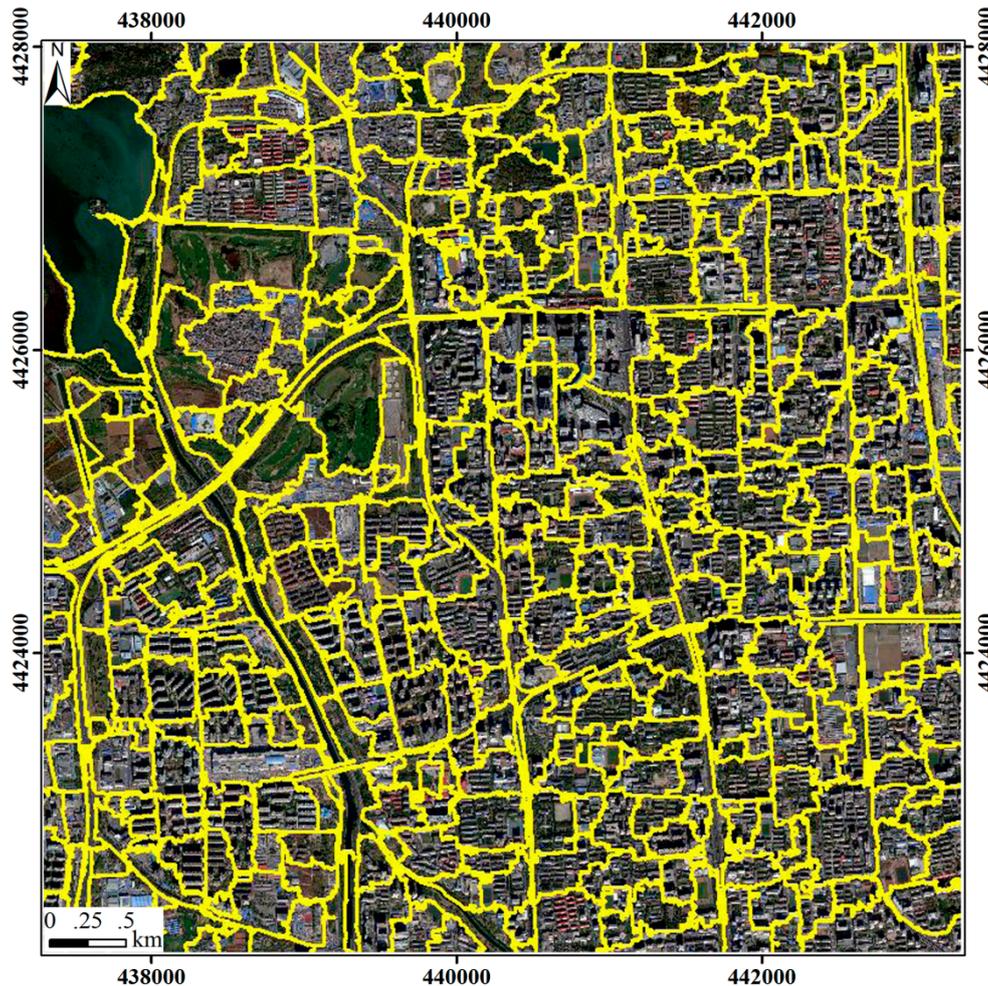


Figure 7. UFZ segmentation results.

As shown in Figure 7, most of the UFZs are segmented appropriately in visual check, demonstrating the effectiveness of the proposed method. In the results, the segmented UFZs, especially those surrounded by roads, reserve good boundary information. Moreover, UFZs with significant spatial configurations such as residential districts are segmented better. There is a park with a large area in the upper left corner of the study area, in which a large segmentation scale is automatically selected by the scale-adaptive method.

As the ground truth of UFZs is not completely available, a number of representative referenced UFZs are manually delineated in the VHR image assisted by Baidu Map (<https://map.baidu.com>, because POI (point of interest) can help us to recognize UFZs) and field investigation (Figure 8). These referenced UFZs have different categories and sizes including 23 residential districts, 10 commercial zones, 3 campuses, 6 parks, 6 industrial zones, and 2 shantytowns. Therefore, there are a total of 50 references. The OCE between the reference data and segmentation results is 0.58, which is relatively satisfactory [41]. According to the visual check on the results and reference data, most UFZs are segmented well. In Figure 8, five different areas are selected to demonstrate the segmentation quality. Except for the

shantytowns which are over-segmented in Figure 8b, the other four areas are segmented appropriately. Large spectral heterogeneity in shantytown leads to more ISODATA clustered classes, making the WIC values of this area to be small; as a result, the segmentation scale in this area has not been enlarged.

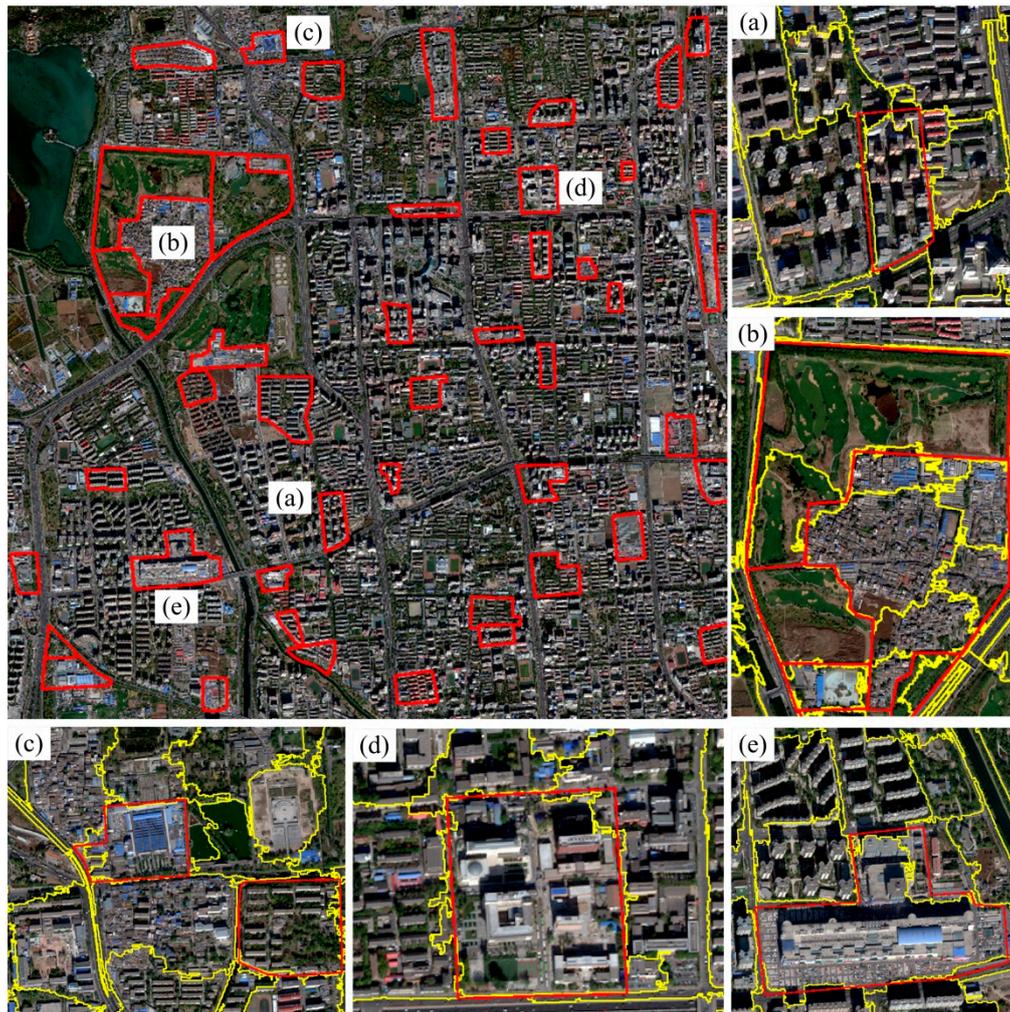
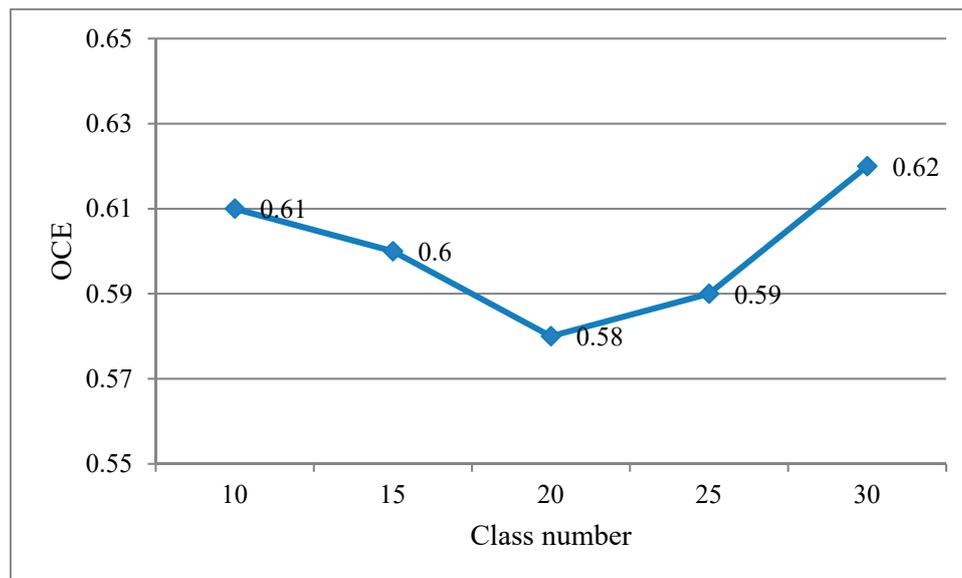


Figure 8. The reference data and the comparison with segmentation results (the red lines are reference data and the yellow lines are segmentation results). (a) Residential districts; (b) shantytowns and parks; (c) industrial zones and residential districts; (d) campuses; and (e) commercial zones.

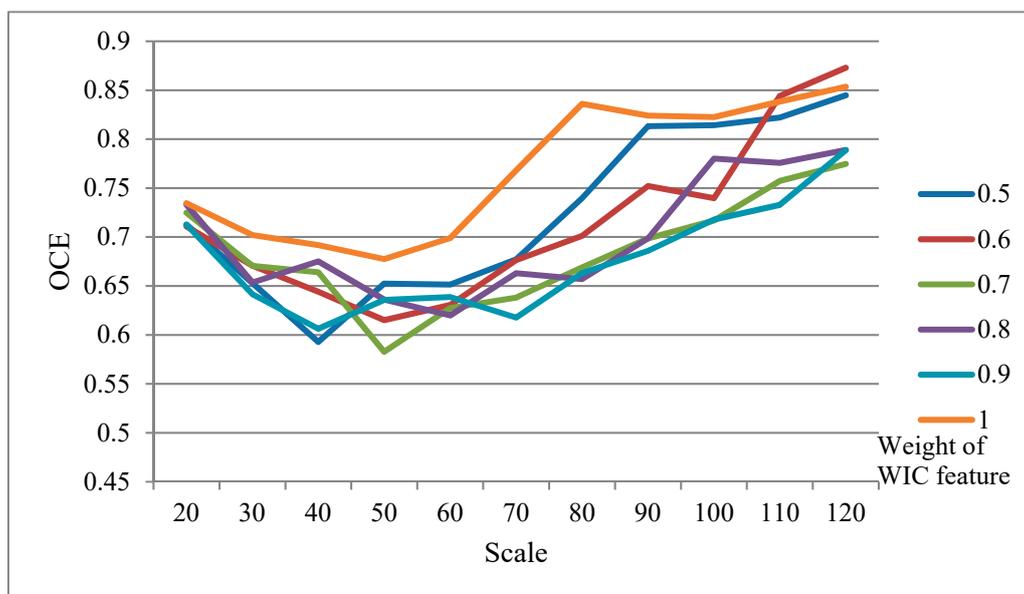
Efficiency is significantly important as it decides the practicability of an algorithm. Our experiments were performed on a desktop computer with Intel Core i7-3770 at a 3.40 GHz CPU with 8.0 GB memory. The running time was about 285 s, 52 s, 195 s, and 66 s for WIC feature calculation, image object segmentation, initial UFZ segmentation, and graph cuts optimization, respectively, thus the whole time was about 10 min. It can be seen that the proposed method behaves efficiently and has a certain potential for practical applications.

3.3. Parameter Analyses

Three parameters are analyzed including the number of ISODATA class, the initial UFZ segmentation scale (in Equation (8)) and the weights of WIC feature and shape in calculating heterogeneity increase (in Equation (7)). Different parameter settings are evaluated and the calculated OCEs are shown in Figure 9.



(a)



(b)

Figure 9. Parameter analyses. (a) Analysis of the number of ISODATA class; and (b) analyses of the segmentation scale and the weight of the WIC feature.

Experiments about the influence of the number of ISODATA class have been performed using Beijing data and the results are shown in Figure 9a. It can be concluded that when the number of ISODATA class ranges from 10 to 30, the OCEs range from 0.58 to 0.62, while the minimal OCE 0.58 is achieved by setting 20. The purpose of the unsupervised classification is mainly to reduce the multidimensional input data to a manageable size, instead of producing an accurate land cover classification. Therefore, 20 classes are enough not only to ensure the computational efficiency, but also to retain enough variance in the input data.

Compared with the number of ISODATA class, both the scale and the weight of WIC feature can remarkably affect segmentation results. A better OCE can be obtained at scales from 40 to 60; as the scale increases, the OCE becomes larger. It can be seen that if only WIC feature is considered in calculating heterogeneity increase, the worst result is obtained; in the meanwhile, when the weight

of WIC feature is 0.5, the result is almost equally poor. Therefore, both the WIC feature and shape heterogeneity increases should be considered, with WIC feature dominating. The minimum OCE is obtained when the scale is 50 and the weight of WIC feature is 0.7. In fact, if the scale is between 40 to 60 and the weight of WIC feature ranges from 0.7 to 0.9 are used, better results can be expected.

The weights of compactness and smoothness (in Equation (6)) are set to 0.5 and 0.5 respectively, because they have little impacts on segmentation results. In graph cuts optimization, we set parameter $\lambda = 1$ (in Equation (9)) and $\sigma = 500$ (in Equation (12)) empirically.

3.4. Experiments on GF-2 Image of Nanchang, China

To demonstrate the adaptability, the proposed method was applied to another city with different urban structures from that of Beijing. Nanchang, the capital of Jiangxi Province in central China, is a developing city at a low urbanization level. This city is undergoing rapid modernization and urban construction. The study area of Beijing is well developed in urban construction while Nanchang is undergoing large-scale urbanization. The two study areas have different urban structures and are located at different urbanization levels, leading to significant differences in their UFZs, thus they can better demonstrate the effectiveness of our method. In order to segment UFZs in Nanchang, GF-2 image with 3.2 meter resolution covering about 164 km² was employed in the experiments (Figure 10).

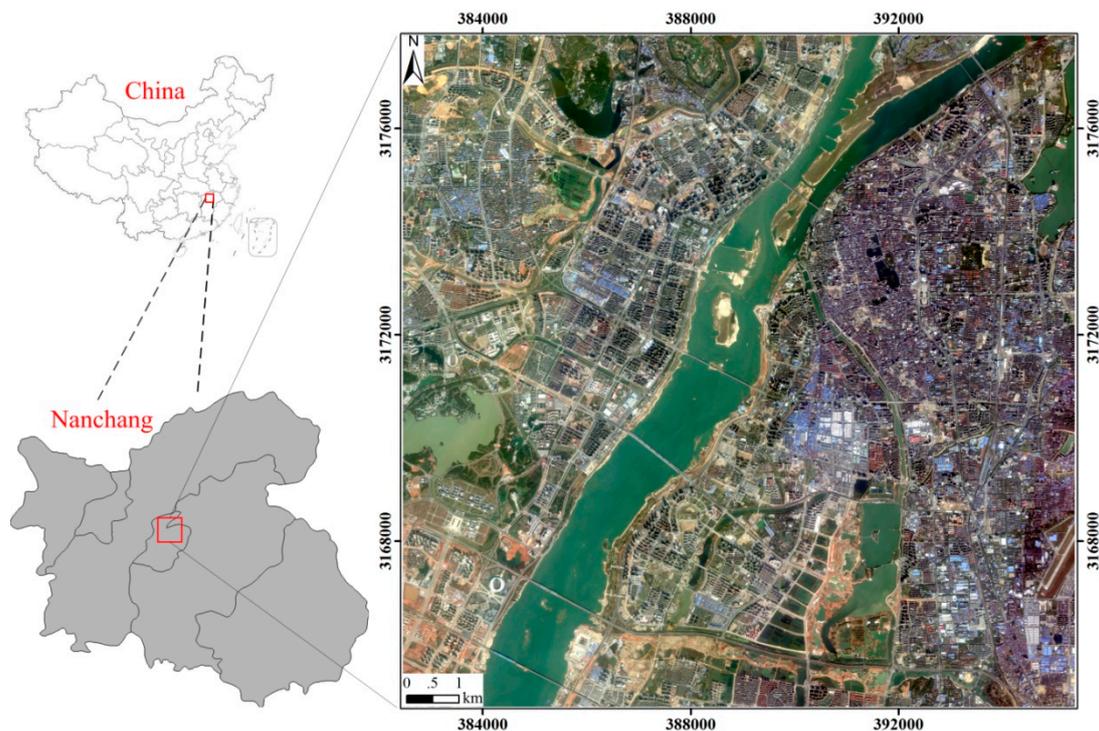


Figure 10. The study area in Nanchang.

As shown in Figure 10, the east part of this area is an old town with lots of dense shantytowns; on the contrary, the south and west parts are newly developed urban areas with broad roads and neat residence and commercial zones. For this area, except for the segmentation scale which is set to 60, the other parameters are the same as those of Beijing area. The segmentation results are shown in Figure 11 and 50 UFZs are delineated manually as reference data, six of which are shown in Figure 11.



Figure 11. UFZ segmentation results of Nanchang (red lines refer to the reference data, while yellow lines represent the segmentation results).

The calculated OCE is 0.52 in this area. From Figure 11, we can see that the UFZ segmentation results are satisfactory in most areas. For the old town, it is difficult to segment UFZs as different UFZs are more likely to be mixed together. The segmentation results in Figure 11c,f that are related to old towns are visually acceptable. For the newly developed area, the buildings show clearer spatial configurations due to neater arrangement, meanwhile, the roads are wider, providing efficient demarcations for different UFZs. Therefore, more satisfactory segmentation results can be obtained in these areas such as shown in Figure 11a,b,d.

4. Discussion

In this section, the effectiveness of context feature and our proposed segmentation method are discussed using Beijing data. Moreover, the proposed method is compared with existing methods for generating UFZs, including the multi-level aggregation approach [8], image tiles, and road blocks. Finally, the contributions of this study are discussed in detail.

4.1. The Effectiveness of Context Feature

The WIC feature measures context information in the form of geographic nearest-neighbor distance from a specific pixel to different image classes, thus it can well characterize the context for each pixel. To demonstrate the effectiveness of the WIC feature, the following three experiments were conducted: (1) visual comparisons of WIC and spectral features; (2) using the traditional MRS method

to directly segment the original image and the WIC feature image; and (3) explaining the reason of the segmentation differences between the original image and WIC feature image.

First, to visually demonstrate WIC feature, five pixels were selected to compare their WIC and spectral features (Figure 12). As shown in Figure 12, P1, P2, and P3 represent building, shadow, and vegetation pixels within the same UFZ, while both P4 and P5 are building pixels from another UFZ. Figure 12b,c are the characteristic curves of the WIC feature (20 bands) and spectral feature (8 bands) for P1, P2, and P3. The WIC value ranges of P1, P2, and P3 are (0, 14), (0, 25), and (0, 15), while their spectral value ranges are (330, 530), (100, 280), and (170, 590), respectively. Compared with spectral feature, P1, P2, and P3 have relatively similar WIC value ranges. Besides, the WIC curves of P1, P2, and P3 are intertwined, while the spectral curves are clearly separated. That is, the UFZ is greatly heterogeneous in spectrum because different ground objects have different spectral curves. However, the WIC feature excels in providing context information for each pixel, thus different categories of pixels in the same UFZ will have similar WIC value ranges. The comparisons of building pixels P1, P4, and P5 are shown in Figure 12d,e. It can be obtained that although they are both building pixels with similar spectral curves, P4 and P5 have almost consistent curves of WIC feature, while P1 is quite different, indicating that WIC feature has the ability to differentiate the same objects from different UFZs. Therefore, it can be concluded that WIC outperforms spectrum in obtaining good UFZ segmentation results.

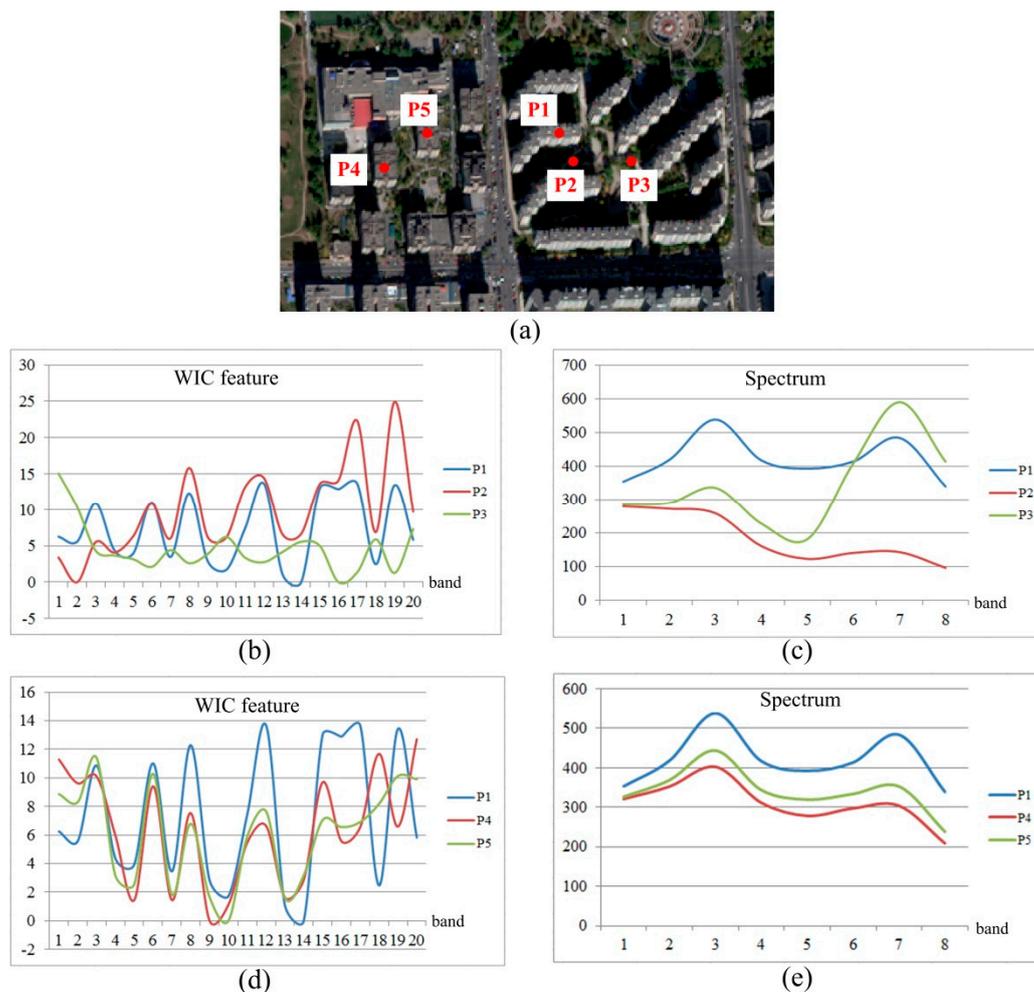


Figure 12. Comparisons of WIC and spectral features. (a) The selected five pixels (P1, P4, and P5 are building pixels; P2 and P3 are shadow and vegetation pixel respectively); (b,c) WIC and spectral feature curves of P1, P2, and P3; and (d,e) WIC and spectral feature curves of P1, P4, and P5.

Second, to further validate the advantage of the WIC feature, we also used the traditional MRS in e-Cognition software to directly segment the original image and the WIC feature image, and the results at different scales are shown in Figure 13. For the original image segmentation, a small scale will lead to a very broken segmentation result, especially for the shadow areas (scale = 300). While for a large scale, a large number of diverse objects will be merged together (scale = 600). On the contrary, the performance of the WIC feature image is better than that of the original image because the objects with similar context information are segmented into one patch. In one word, the WIC feature is more effective for UFZ segmentation while spectrum is almost impossible to achieve.

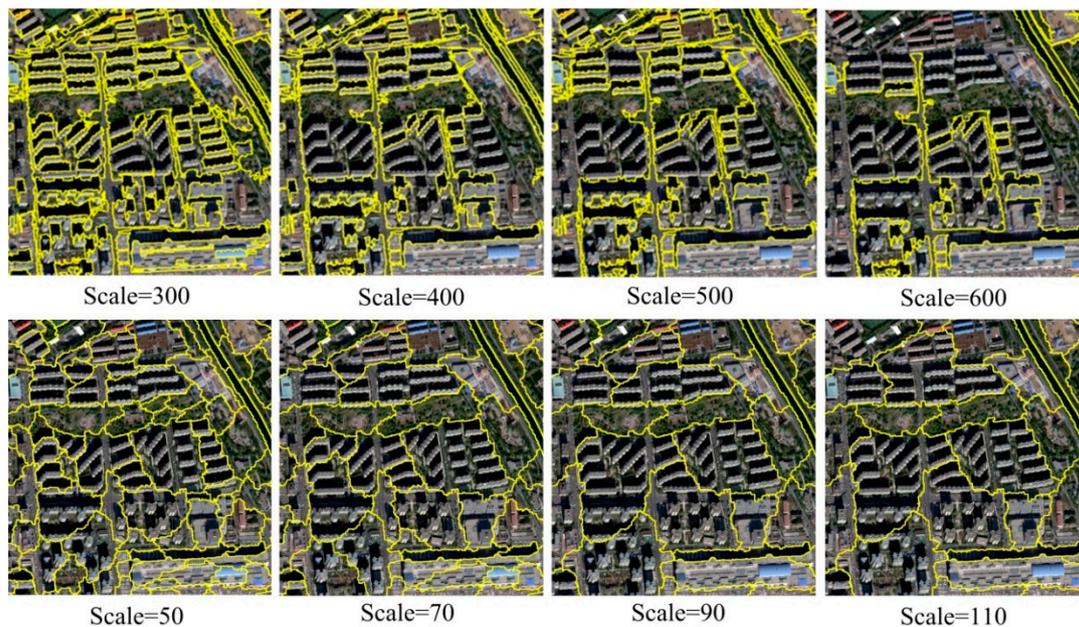


Figure 13. The segmentation results using traditional multiresolution segmentation (MRS) in e-Cognition software (the top images are the results of the original image, while the bottom images are the results of the WIC feature image).

Finally, in order to explain the reason of the segmentation differences between the original and WIC feature images in Figure 13, we calculated the heterogeneity increase caused by the object merging in MRS using WIC and spectral features respectively. The results are shown in Figure 14, where a total of 64 objects crossed by an orange line are selected (Figure 14a), which represent soil, residence-1, park, residence-2, residence-3, and commercial zone in turn. Thereby, 63 heterogeneity increase values are calculated (Figure 14b). By comparing the two different curves, two conclusions can be delivered. First, for each UFZ, the WIC feature has smoother curve than the spectral feature, which is beneficial to make each segmented UFZ more complete. Second, there are more obvious demarcation peaks for different UFZs in WIC feature curves, such as merger 7, 18, 24, and 62 (horizontal axis in Figure 14b), which contribute to the separation between different UFZs. Therefore, the WIC feature measuring context information is much more effective for UFZ segmentation than the spectral feature.

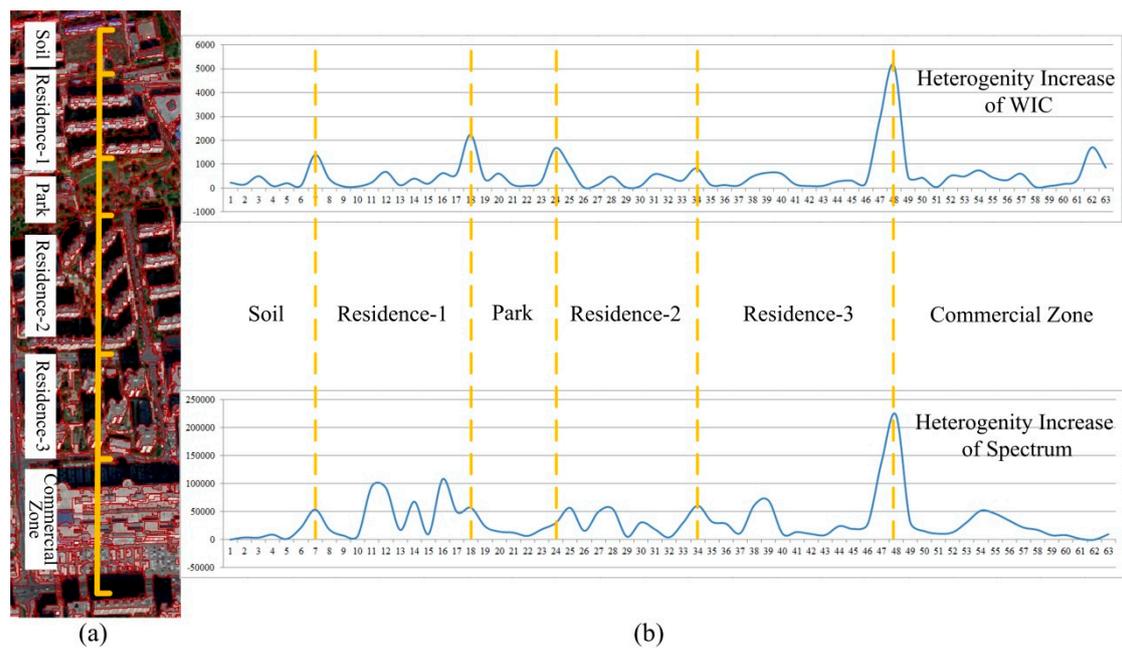


Figure 14. The heterogeneity increase of the object merging in MRS using the WIC and spectral features. (a) The objects crossed by the orange line are selected; and (b) the heterogeneity increase of the WIC and spectral features.

4.2. The Effectiveness of the Proposed Method

In this section, three experiments are conducted to demonstrate the effectiveness of our method including: (1) directly using the traditional MRS in e-Cognition software to segment the WIC feature image and comparing the results with our method; (2) comparing the differences of results before and after graph cuts optimization; and (3) comparing the results of the fixed- and adaptive-scale methods. The calculated OCEs are listed in Table 1.

Table 1. The effectiveness of the proposed segmentation method.

Experiment	(1)	(2)	(3)	The Proposed Method
OCE	0.79	0.67	0.74	0.58

First, the OCE of 0.79 was obtained by the traditional MRS (Figure 15), while the OCE of 0.58 was achieved by our method. According to Figure 15, two issues exist in traditional MRS compared with our method (Figure 7). First, large UFZs are seriously over-segmented (Figure 15a). The reason is that the WIC values are large in big UFZs, which leads to great WIC heterogeneity. Second, since the segmentation is directly performed on the WIC feature image, the objects' boundaries are missed, leading to many UFZs crossing the roads (Figure 15b).

Second, the OCEs before and after graph cuts optimization are 0.67 and 0.58 respectively, indicating that graph cuts can significantly improve the UFZ segmentation results. The results of three local regions are shown in Figure 16 and we can draw the following conclusions: first, these initial segmented UFZs with similar context features are merged together (such as region M in Figure 16); second, objects with significant spectral differences from their surrounding environments are merged into neighboring UFZs (such as region P in Figure 16); third, in the results of graph cuts optimization, the UFZs retain good edge information, especially those surrounded by roads (such as region B in Figure 16). In sum, the graph cuts algorithm aims to achieve the global optimum so that better UFZ segmentation results can be obtained.

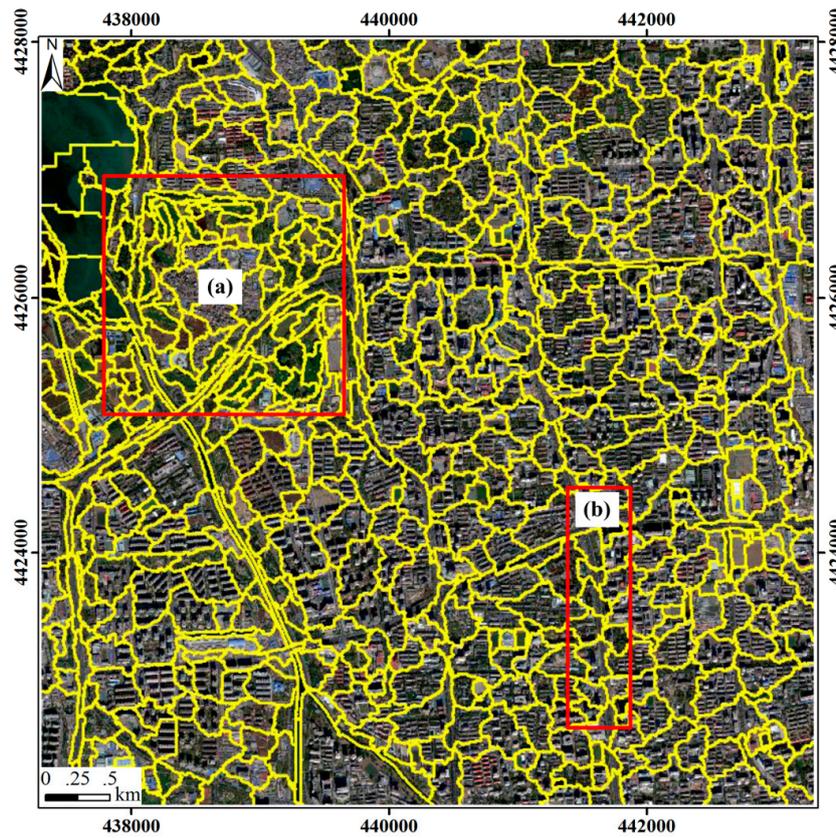


Figure 15. Segmentation results of WIC feature image using the traditional MRS in e-Cognition software with scale = 90. (a) Large UFZs are over-segmented seriously; and (b) many UFZs cross roads.

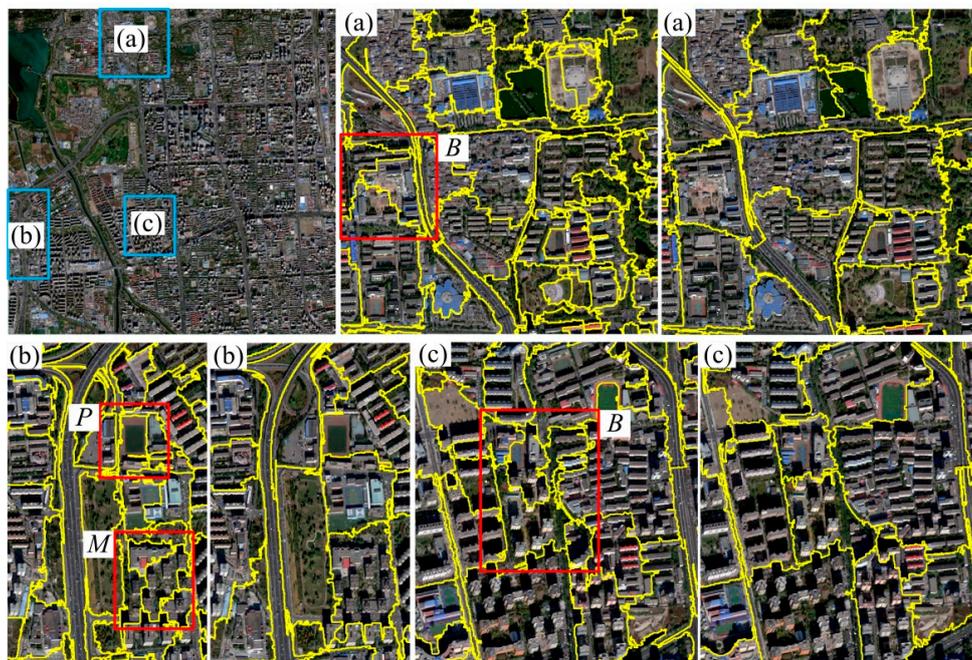


Figure 16. The effectiveness of graph cuts optimization (For each sub-figure, the left image is the initial segmentation result, while the right one is the result of graph cuts optimization).

A scale-adaptive method based on WIC values is proposed in this study. Figure 17 shows the distribution of WIC values (each pixel is assigned with the mean of 20 dimension WIC values). It can be seen that large UFZs have large WIC values, especially for the parks, soil, and large commercial

zones (Figure 17b). The distribution of WIC values is positive skew distribution (Figure 17c), thus the median and upper quartile of the distribution are adopted. The OCE of the scale-adaptive method is 0.58, and when compared with 0.74 of the fixed-scale method, it is clear that significant improvement has been achieved. Figure 18 shows the results of the fixed-scale method and adaptive-scale method in two local regions, demonstrating the effectiveness of the latter. The shantytowns in the center of Figure 18a,b are over-segmented because the segmentation scales in this area are not enlarged, which is caused by the WIC values that are not large enough (see Figure 17b).

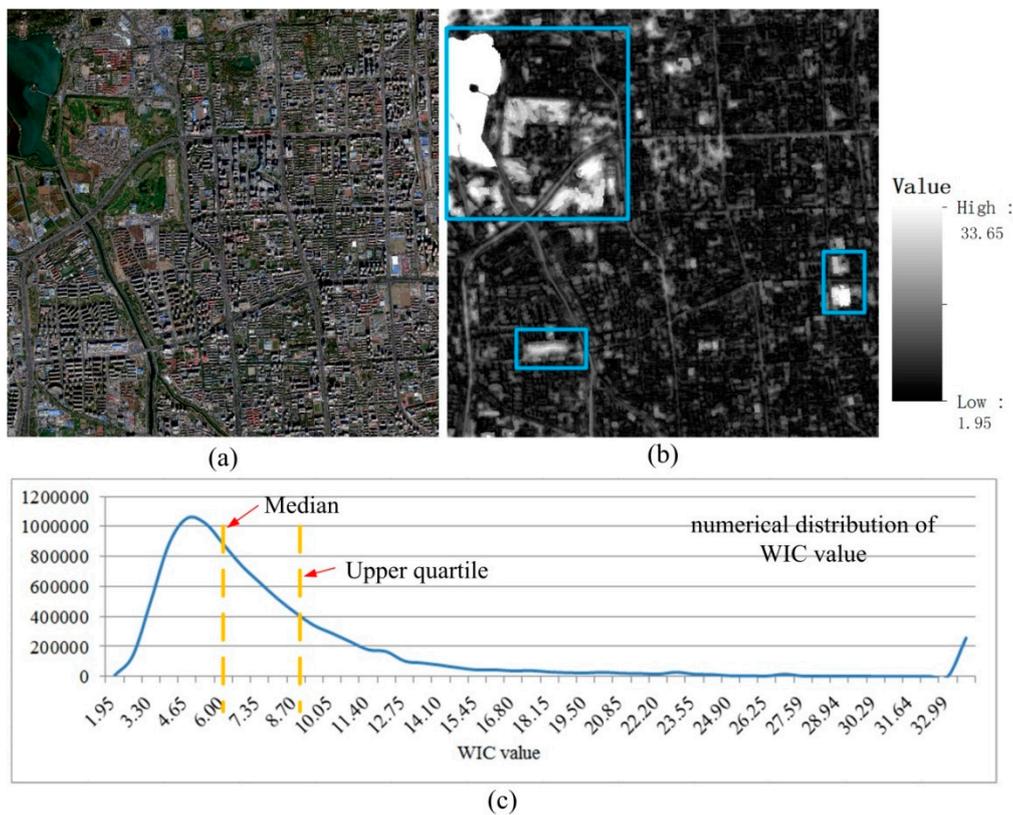


Figure 17. The distribution of WIC values. (a) The original image; (b) the corresponding WIC values (each pixel is assigned with the mean of 20 dimension WIC values); and (c) the distribution of WIC values.

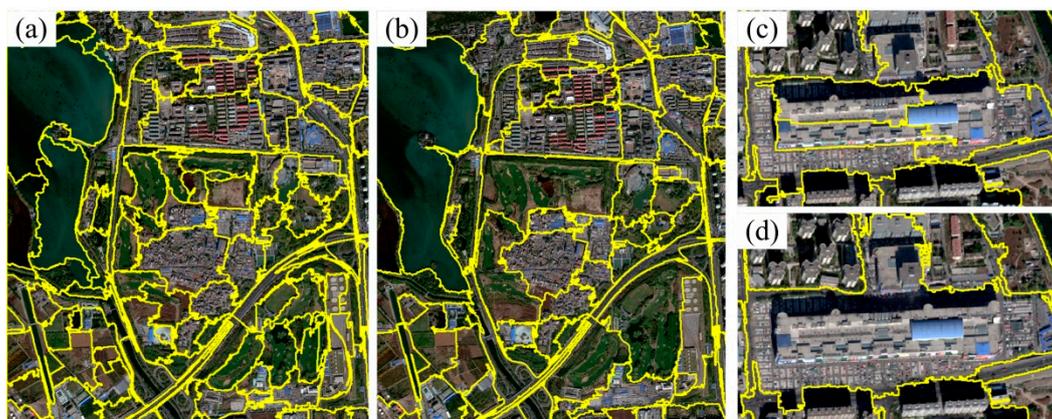


Figure 18. The effectiveness of scale-adaptive method. (a,c) are the results of fixed-scale method; (b,d) are the results of adaptive-scale method.

4.3. Comparing with Existing Methods

In this section, our method is compared with existing methods for generating UFZs, including the multi-level aggregation approach [8], image tiles, and road blocks. First, the comparison with the multi-level aggregation approach is performed in terms of segmentation quality and running time. Figure 19 shows the segmented UFZs by the multi-level aggregation approach, while Table 2 illustrates the calculated OCEs and running time of the two methods.

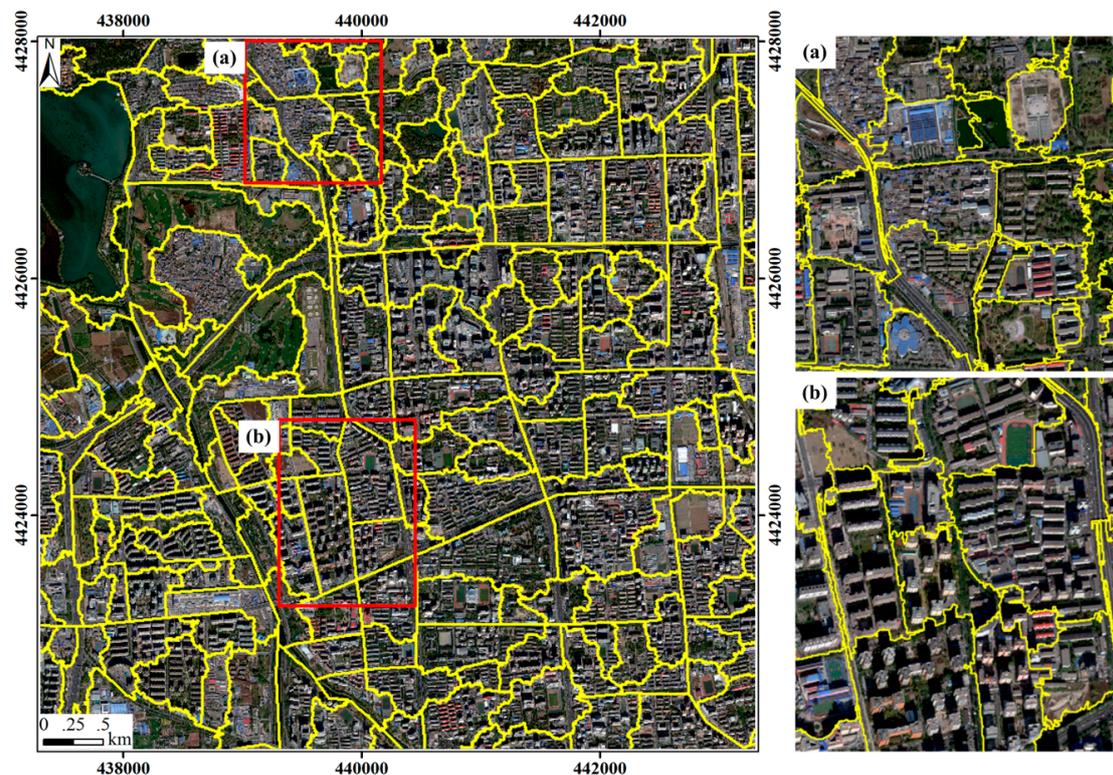


Figure 19. UFZs generated by multi-level aggregation. Right sub-images (a,b) are the results generated by our method.

Table 2. Segmentation quality and running time of the two methods.

Method	The Proposed Method	The Multi-Level Aggregation
OCE	0.58	0.75
Running time	10 min	1.4 h

From Figure 7, Figure 19, and Table 2, it can be concluded that the segmentation results of our method are better than the multi-level aggregation approach in both segmentation quality and running time because the former achieved a smaller OCE value (0.58 vs. 0.75) and much less running time than the latter (10 min vs. 1.4 h). The results of multi-level aggregation are visually acceptable, but there are three issues: (1) it needs the constraints of road vectors; (2) object segmentation and classification must be conducted, during which process low accuracies may influence segmentation results by reducing the robustness of spatial pattern features used in aggregation procedure; and (3) the complex procedure makes the method time-consuming and unsuitable for large-scale UFZ segmentation. Moreover, the segmentation scale of multi-level aggregation is fixed during one segmentation process while the scales used in our method are adaptive for different UFZs. This demonstrates the effectiveness of the proposed segmentation method.

As discussed in Section 1, UFZs are usually represented by image tiles or road blocks in existing UFZ analyses. However, these two units are weak when representing UFZs with arbitrary shapes

and sizes; moreover, UFZs should be analyzed at different scales [45]. Fortunately, our methods can represent and generate UFZs well at multiple scales. The comparisons are shown in Figure 20.

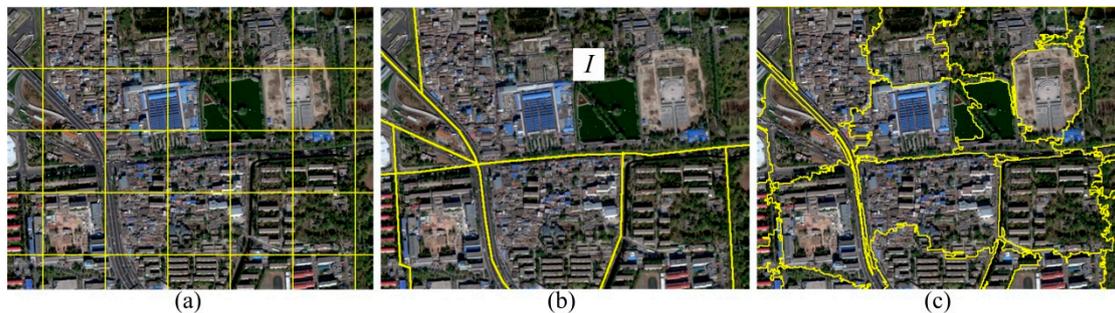


Figure 20. Comparisons of (a) image tiles (100×100 pixels), (b) road blocks, and (c) UFZs produced by our method.

It is easy to tell from Figure 20a that image tiles cannot represent real UFZs in shapes and sizes because they are all rectangles. For the road blocks, they are seriously under-segmented and a block usually contains different kinds of UFZs (Figure 20b). For example, the region I in Figure 20b consists of multiple UFZs in Figure 20c, including shantytowns, industrial zones, and parks. Two reasons are related to this issue. First, a number of UFZs are not separated by roads in reality; second, the road data are often incomplete. The rough road blocks can cause various issues that cannot benefit the UFZ analyses [7]. As to the multiscale image segmentation method proposed in this study, it can produce multiscale UFZs for satisfying the various application demands. Besides, it is automatic, fast, and only remote sensing images are required as the input. Overall, the proposed method is more appropriate to represent and generate UFZs.

4.4. Limitations of the Proposed Method

Although satisfactory results are obtained by the proposed method, two problems still exist in the results. First, as city main roads are often wider, they may be segmented into individual UFZs (Figure 21a). With urban road data becoming more available (e.g., OpenStreetMap), a feasible solution is to introduce road vectors to restrict the segmentation procedure. Second, some UFZs such as campuses and residential districts are easy to be mixed as they are usually similar in spectral features and spatial configurations (Figure 21b). The introduction of POI can resolve this problem effectively as POI has rich semantic information.

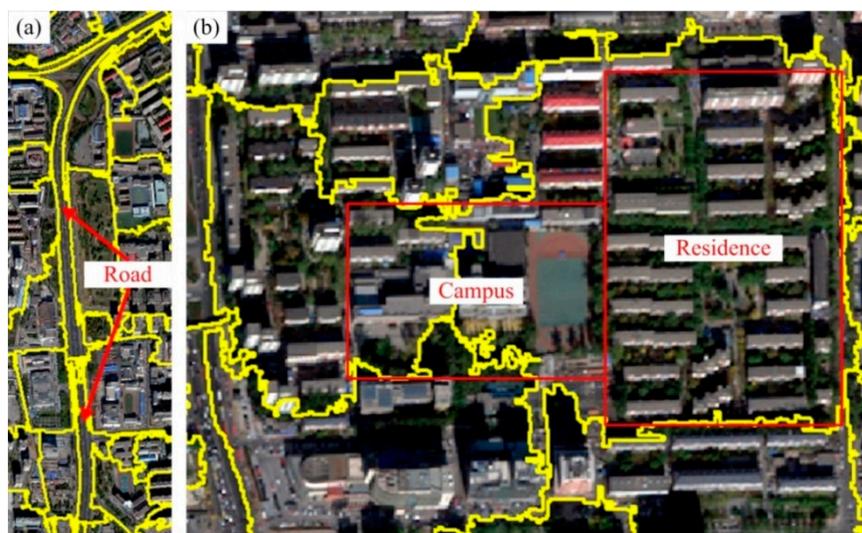


Figure 21. Road segments (a) and confusions between campuses and residential districts (b).

4.5. Contributions of this Study

UFZs are mainly influenced by two factors: government macro policy-planning and human micro socioeconomic activities. UFZ mapping results can reflect the real status of urban forms and functions, the differences in UFZs between the real status and the government planning, and the possible reasons for the differences. Regardless of the real status or the differences, it can help the government to formulate more accurate and detailed planning and management strategies in the future. In order to generate spatial units for UFZ mapping, this study proposes a context-enabled multiscale image segmentation method.

For UFZ analyzes, whether based on VHR images or social sensing data, the commonly used spatial units are image tiles and road blocks, which cannot exactly represent UFZs [2,7,13,14]. Zhang et al. [7] indicated that the spatial units have vast influences on analyzing UFZs. Since different UFZs can be described based on spatial configurations between geographic objects, context information is explored for segmenting UFZs in our study. Therefore, units produced by our method are homogeneous in semantics of urban functions. Accordingly, the proposed image segmentation method committed to solving the generation of spatial units for UFZs is an important complement to the existing UFZ analyzes. Moreover, segmentation results by our method can be used not only for UFZ mapping and planning, but also as units for urban change analysis [46], thermal environment evaluation [47], and population estimation [48]. For example, compared with the pixel, grid, or local-climate-zone based research on urban heat islands [47], using the segmentation results by our method can better investigate the effects of landscape configuration on the urban thermal environment.

Spatial configuration of landscape components is an important content of landscape pattern [49]. Since the WIC features function as descriptions of the spatial context surrounding each individual pixel, the segmented units are consistent in spatial patterns. Therefore, these multiscale segmented units can be used to represent landscape patches, which are required to be delineated at different scales considering spatial patterns, both including man-made and natural landscapes [8,50].

To our best knowledge, it is the first method for large-scale, efficient, and automated UFZ segmentation in this field. Accordingly, this method is an important complement to the existing UFZ analyzes. Since the spatial units are produced, physical features derived from VHR images or social attributes of social sensing data can be applied to identify their categories to generate UFZ maps [2,7,13,14], which are important for policy-makers and professionals in the field of urban planning because they can help to understand the relationship between physical and socioeconomic structures within urban areas [1]. Specifically, the spatial configuration of UFZs can influence the efficiency of urban life. For example, the sharp separation of residences and workplaces causes urban problems such as traffic jams and air pollution.

5. Conclusions

UFZs are the basis of urban planning and play an important role in monitoring urbanization. UFZ segmentation aiming at delineating UFZ boundaries is the prerequisite to urban applications, but it is still unresolved. This study presents a context-enabled multiscale image segmentation method to generate the spatial units of UFZs from VHR images. WorldView-2 image in Beijing and GaoFen-2 image in Nanchang are used to validate the proposed method. The segmentation results demonstrate that our method is promising and practical for segmenting UFZs. The experimental analyses indicate that:

- (1) Context features can be effectively applied to UFZ segmentation.
- (2) Graph cuts algorithm aims to achieve global optimum so that it is effective to optimize the segmentation results. The presented scale-adaptive method based on WIC values can adaptively determine the segmentation scales in different urban areas.
- (3) The proposed segmentation method is more appropriate for UFZ delineation than traditionally used image tiles and road blocks.

Our future work includes two aspects. First, road vectors will be introduced to improve the segmentation quality. Second, since the spatial units of UFZs are generated in this study, function classification by integrating VHR images and social sensing data for producing UFZ maps will be another aspect in our future work.

Author Contributions: Funding acquisition, S.D. (Shihong Du); Investigation, S.D. (Shouji Du) and B.L.; Methodology, S.D. (Shouji Du), S.D. (Shihong Du), B.L., and X.Z.; Supervision, S.D. (Shihong Du); Visualization, S.D. (Shouji Du); Writing—original draft, S.D. (Shouji Du); Writing—review and editing, S.D. (Shihong Du) and X.Z.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 41871372.

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

References

1. Yuan, N.J.; Zheng, Y.; Xie, X.; Wang, Y.; Zheng, K.; Xiong, H. Discovering urban functional zones using latent activity trajectories. *IEEE Trans. Knowl. Data Eng.* **2015**, *27*, 712–725. [[CrossRef](#)]
2. Tu, W.; Hu, Z.; Li, L.; Cao, J.; Jiang, J.; Li, Q.; Li, Q. Portraying urban functional zones by coupling remote sensing imagery and human sensing data. *Remote Sens.* **2018**, *10*, 141. [[CrossRef](#)]
3. Banzhaf, E.; Hofer, R. Monitoring urban structure types as spatial indicators with CIR aerial photographs for a more effective urban environmental management. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2008**, *1*, 129–138. [[CrossRef](#)]
4. Matsuoka, R.H.; Kaplan, R. People needs in the urban landscape: Analysis of landscape and urban planning contributions. *Landsc. Urban Plan.* **2008**, *84*, 7–19. [[CrossRef](#)]
5. Heiden, U.; Heldens, W.; Roessner, S.; Segl, K.; Esch, T.; Müller, A. Urban structure type characterization using hyperspectral remote sensing and height information. *Landsc. Urban Plan.* **2012**, *105*, 361–375. [[CrossRef](#)]
6. Herold, M.; Liu, X.; Clarke, K.C. Spatial metrics and image texture for mapping urban land use. *Photogramm. Eng. Remote Sens.* **2003**, *69*, 991–1001. [[CrossRef](#)]
7. Zhang, Y.; Li, Q.; Huang, H.; Wu, W.; Du, X.; Wang, H. The combined use of remote sensing and social sensing data in fine-grained urban land use mapping: A case study in Beijing, China. *Remote Sens.* **2017**, *9*, 865. [[CrossRef](#)]
8. Zhang, X.; Du, S.; Wang, Q.; Zhou, W. Multiscale geoscene segmentation for extracting urban functional zones from VHR satellite images. *Remote Sens.* **2018**, *10*, 281. [[CrossRef](#)]
9. Zhao, W.; Du, S. Scene classification using multi-scale deeply described visual words. *Int. J. Remote Sens.* **2016**, *37*, 4119–4131. [[CrossRef](#)]
10. Nogueira, K.; Penatti, O.A.; dos Santos, J.A. Towards better exploiting convolutional neural networks for remote sensing scene classification. *Pattern Recognit.* **2017**, *61*, 539–556. [[CrossRef](#)]
11. Cheng, G.; Li, Z.; Yao, X.; Guo, L.; Wei, Z. Remote sensing image scene classification using bag of convolutional features. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1735–1739. [[CrossRef](#)]
12. Li, A.; Lu, Z.; Wang, L.; Xiang, T.; Wen, J.R. Zero-shot scene classification for high spatial resolution remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 4157–4167. [[CrossRef](#)]
13. Zhang, X.; Du, S.; Wang, Y.C. Semantic classification of heterogeneous urban scenes using intrascene feature similarity and interscene semantic dependency. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 2005–2014. [[CrossRef](#)]
14. Hu, T.; Yang, J.; Li, X.; Gong, P. Mapping urban land use by using landsat images and open social data. *Remote Sens.* **2016**, *8*, 151. [[CrossRef](#)]
15. Wu, S.-S.; Qiu, X.; Usery, E.L.; Wang, L. Using geometrical, textural, and contextual information of land parcels for classification of detailed urban land use. *Ann. Assoc. Am. Geogr.* **2009**, *99*, 76–98. [[CrossRef](#)]
16. Hu, S.; Wang, L. Automated urban land-use classification with remote sensing. *Int. J. Remote Sens.* **2013**, *34*, 790–803. [[CrossRef](#)]
17. Hay, G.J.; Castilla, G. Geographic object-based image analysis (geobia): A new name for a new discipline. In *Object-Based Image Analysis*; Blaschke, T., Lang, S., Hay, G., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 75–89.

18. Blaschke, T. Object based image analysis for remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 2–16. [[CrossRef](#)]
19. 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](#)]
20. Tehrany, M.S.; Pradhan, B.; Jebuv, M.N. A comparative assessment between object and pixel-based classification approaches for land use/land cover mapping using SPOT 5 imagery. *Geocarto Int.* **2014**, *29*, 351–369. [[CrossRef](#)]
21. Blaschke, T.; Hay, G.J.; Kelly, M.; Lang, S.; Hofmann, P.; Addink, E.; Feitosa, R.Q.; van der Meer, F.; van der Werff, H.; van Coillie, F. Geographic object-based image analysis—towards a new paradigm. *ISPRS J. Photogramm. Remote Sens.* **2014**, *87*, 180–191. [[CrossRef](#)]
22. Pertuz, S.; Garcia, M.A.; Puig, D. Focus-aided scene segmentation. *Comput. Vis. Image Underst.* **2015**, *133*, 66–75. [[CrossRef](#)]
23. Baatz, M.; Schäpe, A. Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation. In *Angewandte Geographische Informationsverarbeitung XII: Beiträge zum AGIT-Symposium Salzburg 2000 (German Edition)*; Strobl, J., Blaschke, T., Griesebner, G., Eds.; Wichmann-Verlag: Heidelberg, Germany, 2000; pp. 12–23.
24. Sandeep, V.M.; Kulkarni, S.; Kohir, V. Level set issues for efficient image segmentation. *Int. J. Image Data Fusion* **2011**, *2*, 75–92. [[CrossRef](#)]
25. Zhang, X.; Xiao, P.; Feng, X.; Wang, J.; Wang, Z. Hybrid region merging method for segmentation of high-resolution remote sensing images. *ISPRS J. Photogramm. Remote Sens.* **2014**, *98*, 19–28. [[CrossRef](#)]
26. Zhang, X.; Du, S. Learning selfhood scales for urban land cover mapping with very-high-resolution satellite images. *Remote Sens. Environ.* **2016**, *178*, 172–190. [[CrossRef](#)]
27. Yang, J.; He, Y.; Caspersen, J. Region merging using local spectral angle thresholds: A more accurate method for hybrid segmentation of remote sensing images. *Remote Sens. Environ.* **2017**, *190*, 137–148. [[CrossRef](#)]
28. Chen, G.; Weng, Q.; Hay, G.J.; He, Y. Geographic object-based image analysis (GEOBIA): Emerging trends and future opportunities. *GISci. Remote Sens.* **2018**, *55*, 159–182. [[CrossRef](#)]
29. Arbeláez, P.; Hariharan, B.; Gu, C.; Gupta, S.; Bourdev, L.; Malik, J. Semantic segmentation using regions and parts. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 3378–3385.
30. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
31. Zheng, C.; Zhang, Y.; Wang, L. Semantic segmentation of remote sensing imagery using an object-based markov random field model with auxiliary label fields. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 3015–3028. [[CrossRef](#)]
32. Maggiori, E.; Tarabalka, Y.; Charpiat, G.; Alliez, P. Convolutional neural networks for large-scale remote-sensing image classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 645–657. [[CrossRef](#)]
33. Nielsen, M.M. Extraction of different urban area categories from satellite images using window independent context segmentation. In Proceedings of the Joint Urban Remote Sensing Event, Munich, Germany, 11–13 April 2011; pp. 101–104.
34. Nielsen, M.M. Remote sensing for urban planning and management: The use of window-independent context segmentation to extract urban features in Stockholm. *Comput. Environ. Urban Syst.* **2015**, *52*, 1–9. [[CrossRef](#)]
35. Nielsen, M.M.; Ahlqvist, O. Classification of different urban categories corresponding to the strategic spatial level of urban planning and management using a SPOT4 scene. *J. Spat. Sci.* **2015**, *60*, 99–117. [[CrossRef](#)]
36. Memarsadeghi, N.; Mount, D.M.; Netanyahu, N.S.; Moigne, J.L. A fast implementation of the ISODATA clustering algorithm. *Int. J. Comput. Geom. Appl.* **2007**, *17*, 71–103. [[CrossRef](#)]
37. Boykov, Y.; Kolmogorov, V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans. Pattern Anal. Mach. Intell.* **2004**, *26*, 1124–1137. [[CrossRef](#)]
38. Boykov, Y.; Veksler, O.; Zabih, R. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.* **2001**, *23*, 1222–1239. [[CrossRef](#)]

39. Veksler, O.; Boykov, Y.; Mehrani, P. Superpixels and supervoxels in an energy optimization framework. In Proceedings of the European Conference on Computer Vision, Crete, Greece, 5–11 September 2010; pp. 211–224.
40. Kolmogorov, V.; Zabini, R. What energy functions can be minimized via graph cuts? *IEEE Trans. Pattern Anal. Mach. Intell.* **2004**, *26*, 147–159. [[CrossRef](#)] [[PubMed](#)]
41. Polak, M.; Zhang, H.; Pi, M. An evaluation metric for image segmentation of multiple objects. *Image Vis. Comput.* **2009**, *27*, 1223–1227. [[CrossRef](#)]
42. Yang, J.; He, Y.; Caspersen, J.; Jones, T. A discrepancy measure for segmentation evaluation from the perspective of object recognition. *ISPRS J. Photogramm. Remote Sens.* **2015**, *101*, 186–192. [[CrossRef](#)]
43. Johnson, B.; Xie, Z. Unsupervised image segmentation evaluation and refinement using a multi-scale approach. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 473–483. [[CrossRef](#)]
44. Georganos, S.; Lennert, M.; Grippa, T.; Vanhuyse, S.; Johnson, B.; Wolff, E. Normalization in unsupervised segmentation parameter optimization: A solution based on local regression trend analysis. *Remote Sens.* **2018**, *10*, 222. [[CrossRef](#)]
45. Wu, J.; Hobbs, R.J. (Eds.) *Key Topics in Landscape Ecology*; Cambridge University Press: New York, NY, USA, 2007.
46. Du, B.; Wang, Y.; Wu, C.; Zhang, L. Unsupervised scene change detection via latent dirichlet allocation and multivariate alteration detection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 4676–4689. [[CrossRef](#)]
47. Zhou, D.; Xiao, J.; Bonafoni, S.; Berger, C.; Deilami, K.; Zhou, Y.; Frolking, S.; Yao, R.; Qiao, Z.; Sobrino, J. Satellite remote sensing of surface urban heat islands: Progress, challenges, and perspectives. *Remote Sens.* **2019**, *11*, 48. [[CrossRef](#)]
48. Grippa, T.; Linard, C.; Lennert, M.; Georganos, S.; Mboga, N.; Vanhuyse, S.; Gadiaga, A.; Wolff, E. Improving urban population distribution models with very-high resolution satellite information. *Data* **2019**, *4*, 13. [[CrossRef](#)]
49. Peng, J.; Xie, P.; Liu, Y.; Ma, J. Urban thermal environment dynamics and associated landscape pattern factors: A case study in the Beijing metropolitan region. *Remote Sens. Environ.* **2016**, *173*, 145–155. [[CrossRef](#)]
50. Kirchhoff, T.; Trepl, L.; Vicenzotti, V. What is landscape ecology? An analysis and 833 evaluation of six different conceptions. *Landsc. Res.* **2013**, *38*, 33–51. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).