

Article

# Mapping of Coastal Cities Using Optimized Spectral–Spatial Features Based Multi-Scale Superpixel Classification

Aizhu Zhang<sup>1,2</sup>, Shuang Zhang<sup>1,2</sup>, Genyun Sun<sup>1,2,\*</sup>, Feng Li<sup>1,3</sup>, Hang Fu<sup>1,2</sup>, Yunhua Zhao<sup>4</sup>, Hui Huang<sup>1,2</sup>, Ji Cheng<sup>1,2</sup> and Zhenjie Wang<sup>1,2</sup> 

<sup>1</sup> School of Geosciences, China University of Petroleum (East China), Qingdao 266580, China; aizhuzhang@upc.edu.cn (A.Z.); s17010060@s.upc.edu.cn (S.Z.); B18010057@s.upc.edu.cn (F.L.); 1501060218@s.upc.edu.cn (H.F.); s17010069@upc.edu.cn (H.H.); 1501060224@s.upc.edu.cn (J.C.); sdwzj@upc.edu.cn (Z.W.)

<sup>2</sup> Laboratory for Marine Mineral Resources, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266071, China

<sup>3</sup> Shandong Provincial Climate Center, Jinan 250000, China

<sup>4</sup> Qingdao Geotechnical Investigation and Surveying Research Institute, Qingdao 266032, China; zhaoyunhua@qdkcy.com.cn

\* Correspondence: sungenyun@upc.edu.cn

Received: 3 April 2019; Accepted: 23 April 2019; Published: 26 April 2019



**Abstract:** The high interior heterogeneity of land surface covers in high-resolution image of coastal cities makes classification challenging. To meet this challenge, a Multi-Scale Superpixels-based Classification method using Optimized Spectral–Spatial features, denoted as OSS-MSSC, is proposed in this paper. In the proposed method, the multi-scale superpixels are firstly generated to capture the local spatial structures of the ground objects with various sizes. Then, the normalized difference vegetation index and extend multi-attribute profiles are introduced to extract the spectral–spatial features from the multi-spectral bands of the image. To reduce the redundancy of the spectral–spatial features, the crossover-based search algorithm is utilized for feature optimization. The pre-classification results at each single scale are, therefore, obtained based on the optimized spectral–spatial features and random forest classifier. Finally, the ultimate classification is generated via the majority voting of those pre-classification results in each scale. Experimental results on the Gaofen-2 image of Qingdao and WorldView-2 image of Hong Kong, China confirmed the effectiveness of the proposed method. The experiments verify that the OSS-MSSC method not only works effectively on the homogeneous regions, but also is able to preserve the small local spatial structures in the high-resolution remote sensing images of coastal cities.

**Keywords:** superpixels; multi-scale; feature optimization; Morphological Attribute Profiles (MAPs); coastal city; image classification

## 1. Introduction

The term “coastal cities” usually refers to cities located within 100 km from the shoreline where many physical, natural, social and economic elements intersect [1,2]. The development of coastal cities favors fast concentration of population and expansion of urban areas. The land surface covers of coastal cities are, therefore, facing rapid change. Due to changes of land surface covers having a close bearing on ecological protection and sustainable development of cities, prompt and accurate production of land-cover maps using remote sensing images play a key roles in the land-use planning and management of coastal cities [3,4].

Over the past two decades, high-resolution remote sensing images with fine spatial details of land surface covers have drawn much attention of researchers [4–6]. The related mapping techniques have evolved from the pixel-based approaches to superpixels-based methods [7–13]. This is mainly due to the complex structure of diverse urban land covers which make the pixel-based classification methods hardly applicable for high-resolution remote sensing images. To be specific, due to the abundant details in high-resolution image cause the problem of excessive heterogeneity, ground objects of the same land surface covers may reflect excessively different spectral features, thereby the spectral separation of different land surface covers is more difficult and conventional pixel-based classification methods are not efficient enough in such cases [10]. By contrast, the superpixels-based methods can simultaneously make use of the spectral and spatial information of images. Especially, the scale information of different ground objects are also considered in the superpixels-based methods, which is useful for remaining the edge information of different geometric structure [14]. For example, Csillik proposed a fast segmentation and classification method of the very high-resolution remote sensing data using the Simple Linear Iterative Clustering (SLIC) superpixels. In [15], to improve the classification accuracy, both supervised and unsupervised Fractal Net Evolution Approach (FNEA, has been embedded in the commercial software eCognition) were developed for extracting buildings from very high-resolution imagery. These methods can firstly segment the high-resolution image into a number of homogeneous regions to reduce the interior heterogeneity of different kinds of land surface covers. The efficiency of the superpixel-based methods also have been verified in the classification of highly built-up urban areas with abundant fine spatial details [16–18].

However, due to the coastal cities being located in the surface junction area of continent and sea, the urbanization of a coastal area brings out increasingly fragmented landscapes including those typical types of urban areas such as buildings, blacktop, building shadows and particular land-cover types like wharf, bathing beach, rocky islets, etc. These diverse urban land covers form complex urban morphology and structures which make the superpixels-based methods challenging. On one hand, the so-called optimal scale superpixels are always hard to generate. Once the under-segmentation problem occurs, the classification accuracy will be inferior. The fusion of multi-scale superpixels are therefore can be an effective method for accommodating the various sizes of ground objects in a coastal city image. On the other hand, the single-feature extraction method focuses on the spectral or spatial feature and cannot sufficiently describe characteristics of different land types. Although the combination of spectral–spatial features has increased the separability of land types, many of the feature extraction methods are low-utility since they require predefined image structures, whereas ground objects in the coastal cities are size various and irregularly shaped [19,20].

Recently, the Attribute Profiles (APs) have widely been used for modeling the spatial characteristics of the ground objects in remotely sensed images [21–24]. In [21], a set of profiles including the area and standard deviation attributes were used to extract the spatial information of hyperspectral images. Further, for the high-resolution images, APs can perform a multilevel analysis of the image by the sequential application of different kinds of attribute filters, which leads to the concept of Extended Multi-Attribute Profiles (EMAPs). As described in [25–27], one typical property of EMAPs is that their constructions avoid the predefining of image structures. In addition, the EMAPs can preserve the geometric characteristics of the relevant regions and effectively attenuate unimportant details [23,25,28]. This is useful for keeping the boundaries of different land surface covers and decreasing their interior heterogeneity of spectral and spatial features. Nevertheless, due to the EMAPs usually appearing as high dimensions with redundancy and relevance, feature selection or optimization is imperative.

Exhaustive algorithms are the most basic method for finding a subset of prominent features to improve predictive accuracy and to remove the redundant features [29]. However, this kind of algorithm is times consuming, that is, if an image has  $Dim$  features, an exhaustive algorithm will have to test a  $2^{Dim}$  times feature combination to search for the most informative subset features. Apart from the exhaustive algorithms, some intelligent optimization algorithms that are computationally inexpensive in terms of both memory and runtime, such as genetic algorithm (GA) and Particle Swarm

Optimizer (PSO), also have been adopted to solve the feature selection problem [30–34]. This kind of method can select the most informative feature subset from the initial feature space based on the optimization of certain criteria, like classification accuracy. The so-called Gravitational Search Algorithm (GSA) is a recently proposed intelligent optimization algorithm whose simple concept and superior performance have attracted much attention [35–37]. More recently, we have proposed a crossover-based algorithm, for the dimensionality reduction of hyperspectral images [29]. With the based method, the spectral subset with less bands and more discriminative spectral information are selected based on a combined optimization criterion constructed by the overall classification accuracy and the size of the band subset. Thus, the CROSSOVER-BASED ALGORITHM can be an effective method for the optimization of the EMAPs.

Given the background above, this paper proposes a Multi-Scale Superpixels-based Classification (MSSC) method for coastal cities using the Optimized Spectral–spatial (OSS) features (OSS-MSSC). In the proposed OSS-MSSC, the remote sensing image is firstly divided into multi-scale superpixels using the FNEA segmentation method based on three sets of scale parameters. Then, we extract the spectral and spatial features via the Normalized Difference Vegetation Index (NDVI) and EMAPs. The initial feature space is thereby constructed using the extracted spatial–spectral features and the multi-spectral bands of the image. To decrease the redundancy of these features, the CROSSOVER-BASED ALGORITHM is adopted to select the optimal combination of the spectral–spatial features. Accordingly, the pre-classification results at each single scale are generated using the Random Forest (RF) classifier and the majority vote scheme. The RF classifier is selected as the classifier mainly because of the random and rotation forest classifiers having achieved outstanding performances in classification of multi-source and multi-temporal images [38–42], such as urban sprawl modelling [43], agriculture water management [44], and mapping of the wetland park [45]. Finally, these pre-classification results are fused to obtain the ultimate land-cover map of coastal cities.

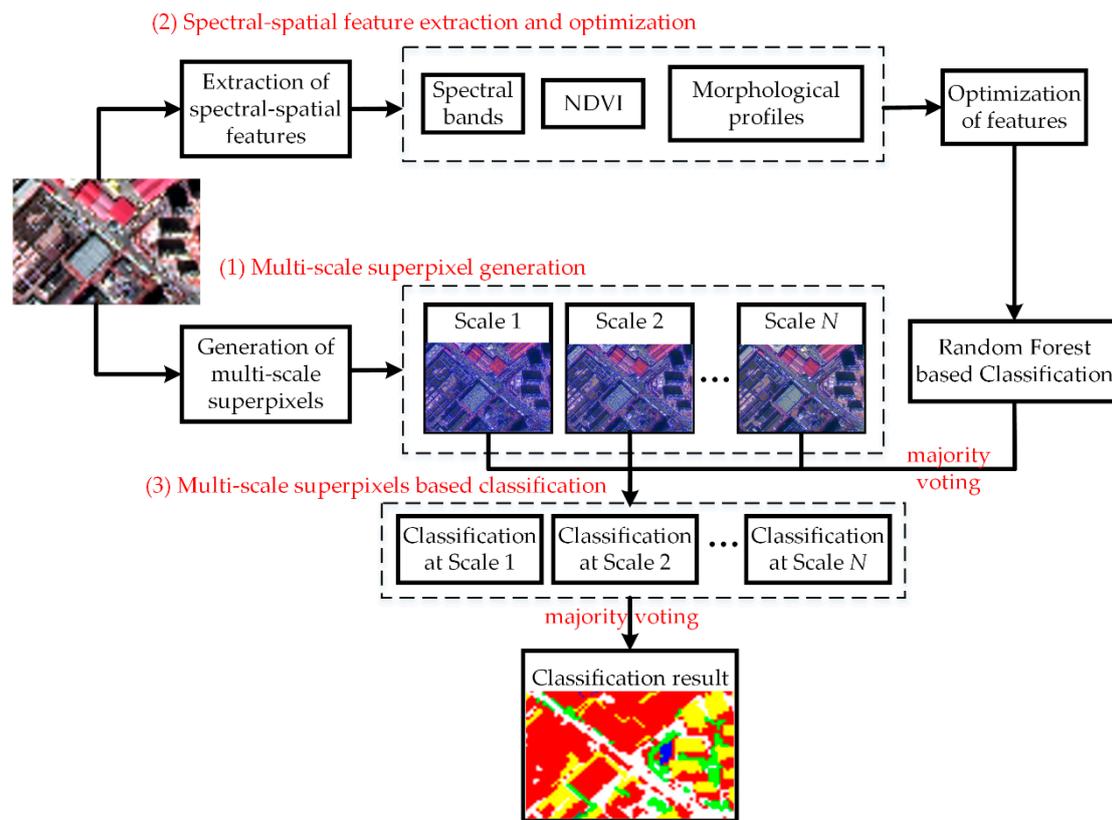
The remainder of this paper is organized as follows. Section 2 presents the details of the proposed OSS-MSSC method. Section 3 describes the experimental dataset and the corresponding comparison results. In Section 4, the detailed discussions of the proposed method is presented with some future scopes. Finally, Section 5 concludes the whole paper.

## 2. Methodology

The overall framework of the proposed OSS-MSSC is illustrated in Figure 1. As shown in Figure 1, there are three major steps within the proposed method: (1) Multi-scale superpixel generation: generating the superpixels with multiple scales to inoculate with different land surface covers; (2) Spectral–spatial feature extraction and optimization: extracting NDVI and EMAPs to construct the initial feature space along with the multi-spectral bands and optimizing the initial features to construct the optimized feature space; (3) Multi-scale superpixels-based classification: classifying the optimized spectral–spatial features first using the RF classifiers, then restricting the classification result by the generated superpixels at each single scale and majority voting scheme, and finally fusing these classification results at different scales to get the ultimate land-cover map of the coastal city via the majority voting rule.

### 2.1. Multi-Scale Superpixel Generation

Segmentation techniques are powerful tools for defining the spatial dependences of the pixels on a high-resolution remote sensing image [46]. In the segmentation techniques, adjacent pixels with similar spectral features are always gathered into superpixels. Comparison to the pixels-based methods, the superpixel-based classification methods can effectively remove the ‘salt and pepper’ effect. In this paper, the multi-resolution segmentation (MRS) algorithm of Ecognition software, i.e., the FNEA-based segmentation method is adopted to perform the hierarchical over-segmentation to generate the superpixels.



**Figure 1.** Framework of the proposed OSS-MSSC (Multi-Scale Superpixels-based Classification method using the Optimized Spectral-spatial features).

In FNEA, a set of segmentation parameters including *scale*, *shape*, and *compactness* are desired. Traditionally, users can get the proper parameters through a trial-and-error process. However, the trial-and-error process is time-consuming and operator-dependent [47]. To overcome these two problems, many researchers have carried out different experiments to identify the relationship between segmentation parameters and segmentation results to find an effective way for segmentation parameter selection [48,49]. However, in most of the cases, the solutions can only suit specific applicable conditions.

In this paper, we do not consider the selection of optimal segmentation parameters, but segment the high-resolution remote sensing image into a series of over-segmented superpixels using  $N$  set of parameters. As show in Figure 1, the superpixels at multiple scales from 1 to  $N$  can represent the local image information in an adaptive domain. For example, the small *scale* can accommodate to the small size objects like grass and trees while the large *scale* is suitable for the objects with large size such as buildings and roads. Specifically, the multi-scale superpixels can be generated following the five main steps below:

- (1) Define a scale parameter  $i \in [1, 2, \dots, N]$ . The scale parameter is the threshold for breaking the merge of different pixels and determines the size of the superpixels/objects.
- (2) Calculate the color and shape of each superpixel by:

$$h_{color} = \sum_{k=1}^D \omega_k \sigma_k \quad (1)$$

$$h_{shape} = \omega_s f_{smooth} + (1 - \omega_s) f_{compact} \quad (2)$$

where  $D$  is the number of features utilized to calculate the color heterogeneity such as spectral heterogeneity,  $\omega_k$  is the weight of the  $k$ -th feature,  $\sigma_k$  is the standard deviation of the feature values

in the  $k$ -th feature in the superpixel,  $f_{smooth}$  and  $f_{compact}$  represent the smoothness and compactness of the superpixel, respectively, and  $\omega_s$  is the weight of the smoothness.

- (3) Calculate the regional heterogeneity of each superpixel by

$$h = (1 - \omega) \times h_{color} + \omega \times h_{shape}, \quad (3)$$

where  $\omega \in [0, 1]$  is the weight of the shape.

- (4) Compare the regional heterogeneity  $h$  of each superpixel with the predefined scale parameter  $i$  and merge any of the two adjacent superpixels whose regional heterogeneity is smaller than  $i$ .
- (5) If the regional heterogeneity of the newly generated superpixel is larger than  $i$ , the merging process will be terminated.

Accordingly, as shown in Figure 1, with a series scale parameters  $i \in [1, 2, \dots, N]$ , the multi-scale superpixels are generated for the following processes.

## 2.2. Spectral–Spatial Feature Extraction and Optimization

The coastal cities scenes in study areas are complicated, multi-scaled and irregularly shaped, which leads to a high between-class homogeneity and low within-class heterogeneity. Thus, it is difficult to classify land surface covers merely depending on the spectral information. Fortunately, the high-resolution images provide fine spatial structural details. In this paper, in addition to the multi-spectral bands, both the EMAPs and NDVI are also taken to describe the spectral–spatial features of superpixels on each single scale. The NDVI is used to increase the comparability of vegetation while the EMAP is used to model the geometrical structures of different artificial facilities like buildings, roads, and bridges.

Mathematical morphology is a well-established framework which provides operators to extract high-quality spatial features [21]. The morphological profiles from these operators are particularly suitable for representing the size variability of different ground objects, but they are not sufficient to model other geometrical features. Thus, the morphological attribute filters were presented and the application of attribute filters in a multi-level way archives the extraction of the APs. Traditionally, for a grayscale image  $I$ , the APs can be generated by a sequence of thickening and thinning operations as follows [50]:

$$AP_I = \{\phi_I^K, \phi_I^{K-1}, \dots, \phi_I^1, I, \gamma_I^1, \dots, \gamma_I^{K-1}, \gamma_I^K\} \quad (4)$$

where  $\phi_I^K$  represents the filtered image generated by thickening operation  $\phi$  with threshold  $\lambda^K$ . Similarly, the  $\gamma_I^K$  is the filtered image generated by thinning operation  $\gamma$  with threshold  $\lambda^K$ . Therefore, the output  $AP_I$  is a stack of  $(2K + 1)$  images, including the original image,  $K$  filtered images from the thickening profiles, and the other  $K$  filtered images from the thinning profiles.

For the multi-spectral image, the APs can be expressed as:

$$AP_{MS} = \{AP_{B_1}, AP_{B_2}, \dots, AP_{B_i}, \dots, AP_{B_L}\} \quad (5)$$

where  $B_i$  is  $i$ -th band in the multi-spectral image and  $L$  is the number of bands.

EMAP is the extension of APs obtained using multiple types of attributes. The three morphological attributes [22,51,52], area, standard deviation, and moment of inertia, are taken into account in this paper. Thereby the EMAPs utilized in this paper can be defined as:

$$EMAP_{MS} = \{AP_{MS}^a, AP_{MS}^s, AP_{MS}^{in}\} \quad (6)$$

where  $AP_{MS}^a, AP_{MS}^s, AP_{MS}^{in}$  are the EMAPs for the multi-spectral image considering attribute area, standard deviation, and moment of inertia, respectively.

The performance of filtering operations implemented in EMAPs highly relies on the threshold value setting of each attribute [27]. Due to the irregularly shaped and multi-scale characteristics

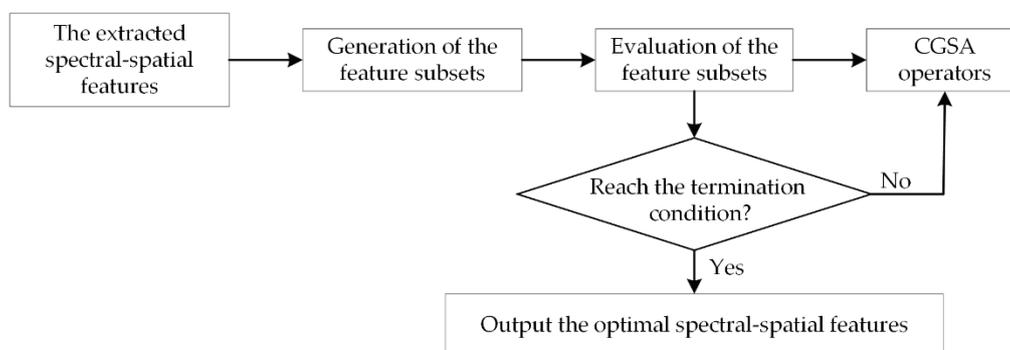
of coastal city land surface covers, adopting a proper threshold value can be a challenging task. A smaller value may introduce noise when processing complex land surface covers in a large size, while a large one could attenuate the information of small sized-objects. In order to solve this issue, an automatic threshold selection strategy [21,53] is utilized in this paper, which generates a dense sampling of threshold values from a wide range according to the characteristics of each band from the tested datasets.

The constructed EMAPs are rich in spatial information but inevitably result in increase of the computational burden and curse of dimensionality problem [27]. To reduce the information redundancy, the crossover-based method [29] is utilized to optimize the extracted spectral–spatial features in this paper. Comparing to the traditional exhaustive algorithm based feature selection method, the crossover-based method is adaptive and time-saving. In this method,  $pop$  feature subsets ( $X_i, i \in [1, 2, \dots, pop]$ ) are initialized randomly and the discriminative capability of each subset is evaluated based on the corresponding overall classification accuracy ( $OA(X_i)$ ) and size of the subset ( $D_b$ ) as illustrated in Equation (7).

$$fit(X_i) = OA(X_i) - \omega_b \times \frac{D_b}{D} \quad (7)$$

where  $\omega_b$  is a weight factor for the balancing the classification accuracy and the size of the  $i$ -th subset,  $D$  is the dimension of the spectral. Apparently, a subset with higher fitness value ( $fit$ ) represents a better discriminative capability for image classification.

Following the iterative operators of the CROSSOVER-BASED method, including update of acceleration, mass, velocity, position, and the crossover operator, the optimal spectral–spatial features can be finally acquired when reaching the termination condition as shown in Figure 2. Usually, the maximum iteration times are adopted as the termination condition.



**Figure 2.** The crossover-based search algorithm-based feature selection.

### 2.3. Multi-Scale Superpixel-based Classification

Based on the optimized features, we first perform classification of the coastal city based on the optimized spectral–spatial features using the RF classifier. However, the thickening and thinning operations in the EMAPs make the boundary of some land surface covers blurred. Moreover, the pixel-based RF classifier method always faces the ‘salt and pepper noise’ problem. In contrast, the boundaries offered by superpixels are clear and the pixels in superpixels are always more homogeneous. Therefore, the obtained classification results are then optimized by the superpixels generated at each single scale using the majority vote scheme.

Specifically, the obtained classification result is firstly mapped to the superpixels generated at each single scale to assign classification labels for all pixels. Then, the classification unit is defined as

superpixel instead of individual pixel in the post-processing and the mostly frequently appeared label in a superpixel is considered as the new label of this superpixel as shown in Equation (8).

$$SP_{reg} = \operatorname{argmax}_{n=1}^{Num} \sum \operatorname{sign}(f_{reg(i,j)} = n) \quad (8)$$

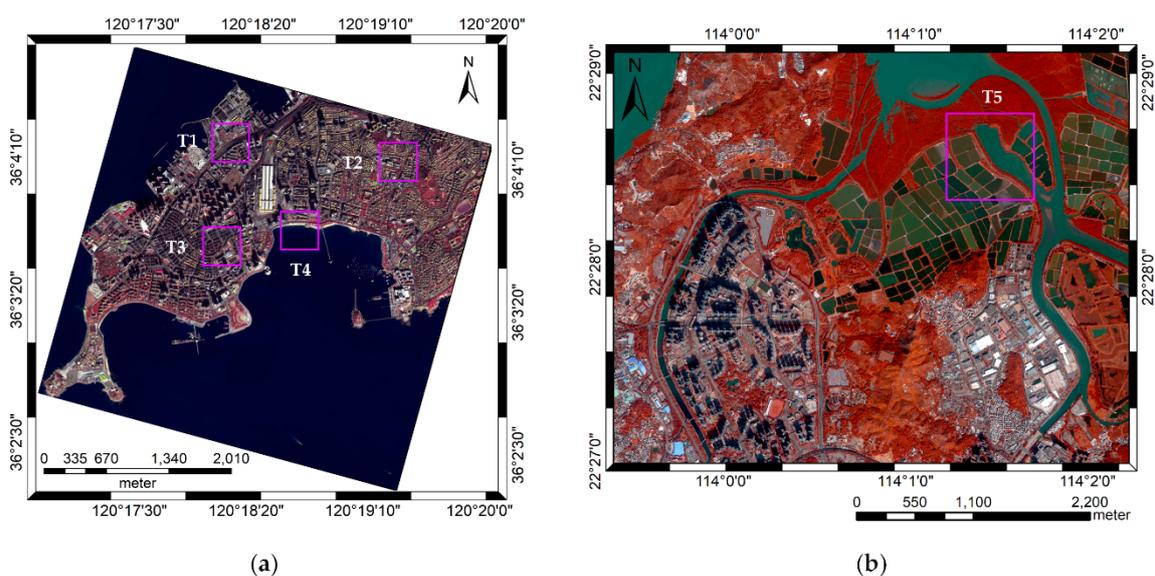
where  $Num$  is the total number of the expected classes,  $(i, j)$  is the coordinate of pixel  $reg(i, j)$ , and  $f_{reg(i,j)}$  is the label of pixel  $reg(i, j)$  in superpixel  $reg$  from the initial classification result.

Thereby, the pre-classification results at each single scale are obtained and there are  $N$  classification results for each pixel of the coastal city image. The ultimate land-cover map is obtained via the fusion of the  $N$  classification results using the majority vote scheme following Equation (8). The classification results are quantitatively evaluated using two widely used precision evaluation criteria calculated from the confusion matrix, i.e., the Overall Accuracy (OA) and kappa coefficient (kappa) [54].

### 3. Experiments and Results

#### 3.1. Introduction of Datasets

In this study, two high-resolution images from Gaofen-2 and Worldview-2 are selected as the study area as shown in Figure 3. Figure 3a shows the Gaofen-2 images of a core area of Qingdao (119°30′~121°00′E, 35°35′~37°09′N) which located on the west coast of the Yellow Sea. It is a key node city for the construction of the “21st Century Maritime Silk Road” of China with temperate monsoon climate which has undergone a rapid urbanization in the past twenty years. The Gaofen-2 data contained 4 multispectral bands spanning from visible to Near Infrared (NIR) spectral region, i.e., 3 visible spectral bands (blue: 450–520 nm defined as B1, green: 520–590 nm defined as B2, red: 630–690 nm defined as B3), 1 NIR band (770–890 nm defined as B4) with a spatial resolution of 4 m, and 1 panchromatic band (450–900 nm) with a resolution of 1 m. The used Gaofen-2 image (with a size of 4792 × 4800 pixels) was collected on 16 February 2016 under clear weather condition. The main land surface covers in this study area can be classified into 12 classes, i.e., building, open space, road, vegetation, shadow, rock, bare soil, beach, red sport ground, green sport ground, sand, and vacant. To make full use of the spectral and spatial information of the Gaofen-2 image, the tested image is fused using the Gram-Schmidt Pan Sharpening method [55] after the positive shooting correction in this paper.



**Figure 3.** The study area and the four zoomed-in representative area. (a) Gaofen-2 image of Qingdao; (b) Worldview-2 image of Hong Kong.

Figure 3b presents the Worldview-2 image of a marginal area called YuenLong of Hong Kong (113°31′–114°18′E, 22°05′–22°22′N) located on the south coast of South China Sea. The utilized Worldview-2 data contained 8 multispectral bands spanning from visible to NIR spectral region (blue: 450–510 nm defined as B1, green: 510–580 nm defined as B2, red: 630–690 nm defined as B3, NIR: 770–895 nm defined as B4, coastal: 400–450 nm defined as B5, yellow: 585–625 nm defined as B6, red edge: 705–745 nm defined as B7, and NIR 2: 860–1040 nm defined as B8) with a spatial resolution of 2 m. The used Worldview-2 image (with a size of 2636 × 2113 pixels) was collected on 14 January 2016 under clear weather condition. The main land surface covers in this study area can be classified into nine classes, i.e., building, open space, road, vegetation, shadow, river, inland water, aquafarm, and mudflat. As shown in Figure 3b, this study area contains more vegetation and nature types than that of Qingdao.

### 3.2. Experimental Settings

In this study, we conduct a series of systematical experiments to validate the effectiveness of the proposed method. The experiments contain:

- (1) Experiment merely using the raw multi-spectral features and NDVI (defined as Raw);
- (2) Experiment using the spectral–spatial features (defined as SS);
- (3) Experiment using the optimized spectral–spatial features (defined as OSS);
- (4) Experiment using the proposed method (defined as OSS-MSSC);
- (5) Experiment using a favorite image classification model using convolutional neural network with 16 weight layers (denoted as VGG (Visual Geometry Group), the source code can be download from <https://github.com/ry/tensorflow-vgg16>) because of its simplicity and accuracy;
- (6) Experiment using the multi-scale superpixel based VGG by majority voting (defined as MSS-VGG).

The first two experiments can be used to test the effects of the extracted spectral and spatial features while the third experiments can be used to verify the effectiveness of the feature selection method. The fourth experiments are used to confirm the superiority of the proposed OSS-MSSC method. The last two experiments are taken to verify the effectiveness of the OSS-MSSC compared to the advanced classification method using deep learning which have demonstrated their superiority in large-scale image recognition [56]. A typical advantage of the deep learning method is that they can abstract the spectral and spatial features in multiple levels using the repeated convolution and pooling operations. Note that the classifiers in all the first four experiments are the same RF classifier. Moreover, the utilized samples for Gaofen-2 and Worldview-2 images which contain 173,350 and 100,710 pixels, respectively, are also keep the same as listed in Table 1 (readers can ask for the tested image and samples by email). All these samples are manually selected by referring to the Google Earth images.

**Table 1.** Training and testing samples for the two study area.

Gaofen-2 Image of Qingdao				Worldview-2 Image of Hong Kong			
Classes	Training	Testing	Total	Classes	Training	Testing	Total
Building	54,704	13,676	68,380	Building	18,272	4568	22,840
Open space	8680	2170	10,850	Shadow	8552	2138	10,690
Road	28,812	7203	36,015	Open space	8220	2055	10,275
Vegetation	13,004	3251	16,255	River	6548	1637	8185
Shadow	22,128	5532	27,660	Inland water	9748	2437	12,185
Rock	3728	932	4660	Aquafarm	6660	1665	8325
Bare soil	964	241	1205	Road	8096	2024	10,120
Beach	968	242	1210	Vegetation	12,992	3248	16,240
Red sport ground	884	221	1105	Mudflat	1480	370	1850
Green sport ground	1628	407	2035				
Sand	2084	521	2605				
Vacant	1096	274	1370				

For the parameter setting, this study mainly concerns the parameters in the generation of superpixels, the extraction of EMAPs, the optimization of features using the CROSSOVER-BASED ALGORITHM, the RF classifier, and the network structures of utilized VGG.

Due to the multi-scale superpixels are generated using the FNEA method of the Ecognition Software, there are 3 parameters need to be set, including the *scale*, *weight of the smoothness*, and *weight of the shape*. To distinguish different kinds of land surface covers, we set three different scale levels in this paper as listed in Table 2.

**Table 2.** Parameter settings in the generation of superpixels.

Image	Level	Target Objects for Segmentation	Scale	Shape	Smoothness
Qingdao	1	vegetation and their shadow	10	0.2	0.8
	2	others	20	0.5	0.7
	3	road	30	0.7	0.7
Hong Kong	1	building and their shadow	10	0.7	0.8
	2	vegetation and others	15	0.2	0.5
	3	river, inland water, and aquafarm	20	0.5	0.5

For the Gaofen-2 image of Qingdao, due to the vegetation mainly displays as small patches as shown in Figure 1, the first level focuses on the vegetation their shadow and the *scale* is set to 10. Moreover, considering various texture and shape of the vegetation, the weights of smoothness and shape in the first level is set to 0.2 and 0.8, respectively. The second level mainly concerns the artificial facilities with larger size. The scale, shape, and smoothness parameters is set to 20, 0.5, and 0.7, respectively. The third level concentrates on the line objects with large scale, such as the road in the coastal cities. In this level, the scale, shape, and smoothness parameters is set to 30, 0.7, and 0.7, respectively. For the Worldview-2 image, due to the study area located in the north of Hong Kong where preserves large area of nature land surface covers, river, aquafarm, and vegetation in this area show larger scales while building and shadow display smaller scales. Thereby, the scale, shape, and smoothness parameters of building and their shadow are set to 10, 0.7, and 0.8, respectively. For the water spectral with low heterogeneity, i.e., the river, inland water, and aquafarm, the three parameters are set to 20, 0.5, and 0.5, respectively. For the other urban land covers including vegetation, the three parameters are set to 15, 0.2, and 0.5, respectively as listed in Table 2.

For the generation of the EMAPs, the thresholds were automatically computed by using the automatic framework, which had been tested on many well-known datasets with different spatial resolutions and proven to be excellent in terms of classification accuracies [21,53]. Specifically, a wide range of threshold settings for each attribute was established according to the characteristics of tested dataset as follows (each multi-spectral band):

For the area attribute,  $\lambda_a = (100/\varphi) \times \{1, 3, 5, 7, 9\}$ ;

For the standard deviation attribute,  $\lambda_s(B_i) = \mu_i \times \{\sigma_{\min}, \sigma_{\min} + \delta_s, \sigma_{\min} + 2\delta_s, \dots, \sigma_{\max}\}$ ;

For the moment of inertia attribute,  $\lambda_{in} = \{0.1, 0.3, 0.5, 0.7\}$ ;

where  $\varphi$  and  $\mu_i$  are the resolution (i.e., 1 m for Gaofen-2 image and 2 m for Worldview-2 image) and average value of the  $i$ -th band, respectively. The values of  $\sigma_{\min}$ ,  $\sigma_{\max}$ , and  $\delta_s$  are set to 2.5%, 27.5%, and 2.5%, respectively [21]. Accordingly, 164 (including 44 area, 88 standard deviation, and 32 moment of inertia attributes as shown in Table 3) and 328 (including 88 area, 176 standard deviation, and 64 moment of inertia attributes as shown in Table 3) EMAPs features for the tested Gaofen-2 and Worldview-2 images can be generated, respectively.

**Table 3.** Results of feature optimization.

Image	Multi-Spectral	NDVI	Area	Standard Deviation	Moment of Inertia			
Qingdao	4 (4)	1 (1)	19 (44)	100/ $\varphi$ $\times$ 1 (4)	$\mu \times 2.5$ (6)	0.1 (4)		
				100/ $\varphi$ $\times$ 3 (4)	$\mu \times 5$ (4)	0.3 (5)		
				100/ $\varphi$ $\times$ 5 (4)	$\mu \times 7.5$ (2)	0.5 (3)		
				100/ $\varphi$ $\times$ 7 (3)	$\mu \times 10$ (3)	0.7 (1)		
				100/ $\varphi$ $\times$ 9 (4)	$\mu \times 17.5$ (1)	$\mu \times 20$ (2)		
		1 (1)	32 (88)	100/ $\varphi$ $\times$ 1 (8)	29 (176)	$\mu \times 2.5$ (6)	32 (64)	0.1 (6)
Hong Kong	8 (8)			100/ $\varphi$ $\times$ 3 (4)	$\mu \times 5$ (6)	0.3 (12)		
				100/ $\varphi$ $\times$ 5 (4)	$\mu \times 7.5$ (8)	0.5 (10)		
				100/ $\varphi$ $\times$ 7 (4)	$\mu \times 10$ (4)	0.7 (4)		
				100/ $\varphi$ $\times$ 9 (12)	$\mu \times 15$ (5)			

For the feature optimization of using the CROSSOVER-BASED ALGORITHM, all the parameters are set to the recommended values in [29], i.e., the initial constant, the decrease coefficient, the population size, the maximum number of iterations, and the weight factor are set to 20, 100, 10, 10, and 0.6, respectively. To decrease randomness, 30 independent runs for each tested image is performed and the results with median performance is utilized for the following classification. For the RF classifier, the value of tree in each of the experiment is set to 200 for the fair comparison.

For the convolutional network in VGG with 16 layers, the recommended parameters in [56] are adopted and the patch size is set to 21. Moreover, to make fair comparison, all the algorithms are implemented using Matlab 2017b on a computer with Intel (R) Xeon (R) CPU (1.70 GHz) and 32 GB of memory.

### 3.3. Experimental Results

#### 3.3.1. Optimization Results of the Features

As introduced in Section 3.2, in the process of the feature extraction, 164 and 328 dimensional EMAPs are generated for the Qingdao and Hong Kong area, respectively. That is to say, there are 169 features (includes 5 spectral features) available for the Gaofen-2 image and 337 features (includes 9 spectral features) for the Worldview-2 image. With the crossover-based feature optimization method, 55 and 102 of the 169 and 337 spectral-spatial features are selected. The corresponding results are listed in Table 3.

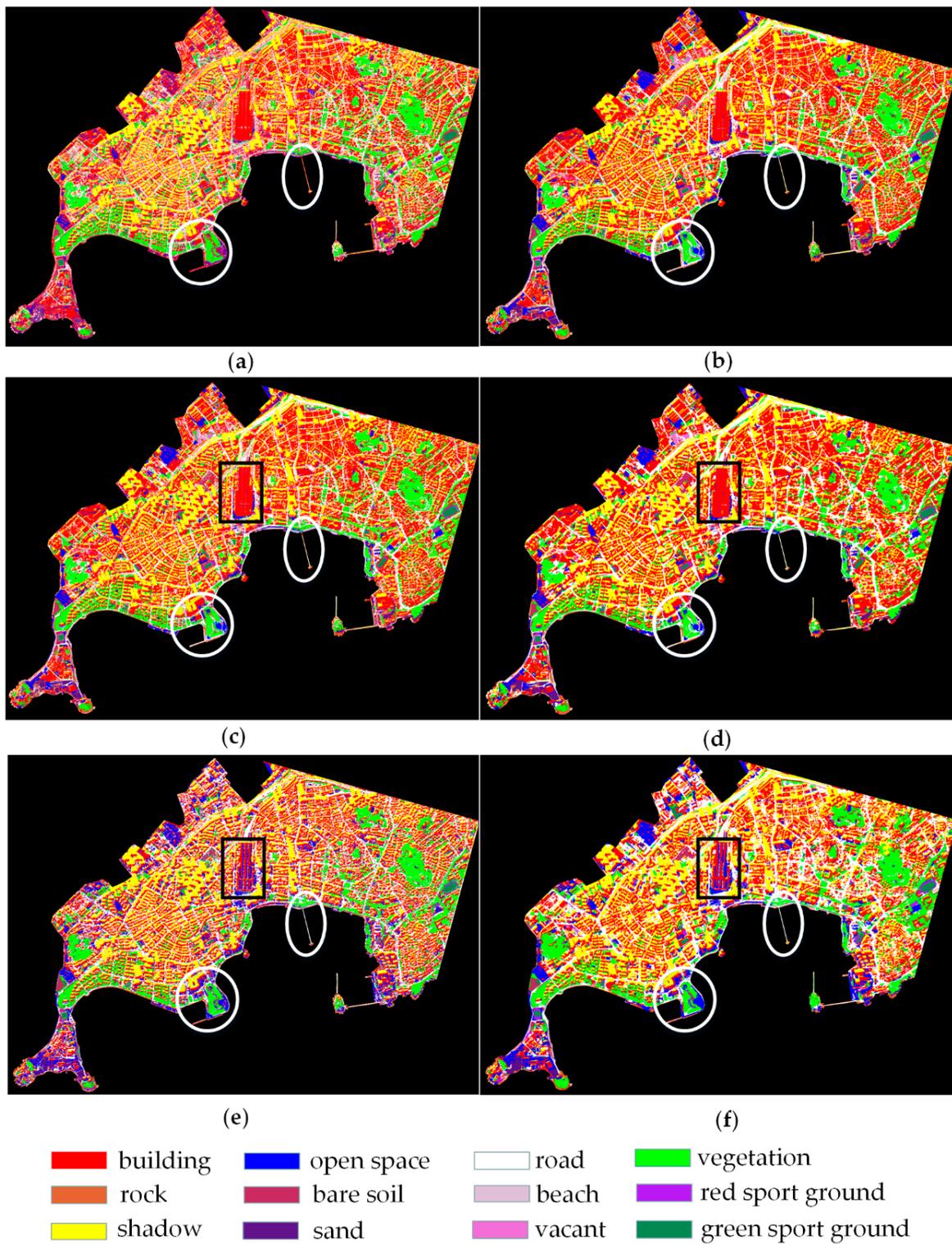
For the Gaofen-2 image of Qingdao, as shown in Table 3, with the feature selection method, the dimension of the EMAPs are sharply decreased from 164 to 50 (19 area attributes, 18 standard deviation attributes, and 13 moment of inertia attributes). Especially nearly 80% (18 of 88) of the standard deviation attributes in the EMAPs have been removed. Moreover, ten of the 18 selected standard deviation attributes come from the features at  $\mu \times 2.5$  and  $\mu \times 5$ , and the others show roughly average distribution. This indicates that the internal heterogeneities of some objects are small and that of many other objects are large. Similarly, the 19 selected area attributes present roughly average distribution as well. This confirms the high variability of scales in the complicated tested image. Due to the moment of inertia is the measure of regional compactness, the calculated value of narrow area is larger while that of the compact area is smaller. Therefore, the moment of inertia attributes contribute larger to the extraction of road than other land surface covers. This is the main reason for the number of the selected moment of inertia attributes is relatively smaller than the other two kinds of attributes. In addition, from Table 3 we can conclude that the multi-spectral bands and the NDVI feature play a key role in the mapping of the coastal cities using the Gaofen-2 image.

For the Worldview-2 image of Hong Kong, as illustrated in Table 3, the dimension of the EMAPs are decreased to 93 (32 area attributes, 29 standard deviation attributes, and 32 moment of inertia attributes) from 328. Similar to that of the Qingdao image, the standard deviation attributes are sharply decreased from 176 to 29 and their distributions are roughly average. For the area attributes, it can be seen that most (20 of 32) of the selected attributes concentrate on the small and large areas. This may come from the buildings in the Yuenlong area are basically small in size and the open spaces are scattered, while the aquafarm, river, vegetation etc. are coherent and relatively large in area. Moreover, due to the aquafarm, river, inland water, and road occupy most of the areas and show regular shape, half (32 of 64) of the moment of inertia attributes are selected and they can be used to discriminate the narrow and square objects. Besides, similar to that of the Gaofen-2 image, all the multi-spectral bands and the NDVI feature are selected as shown in Table 3.

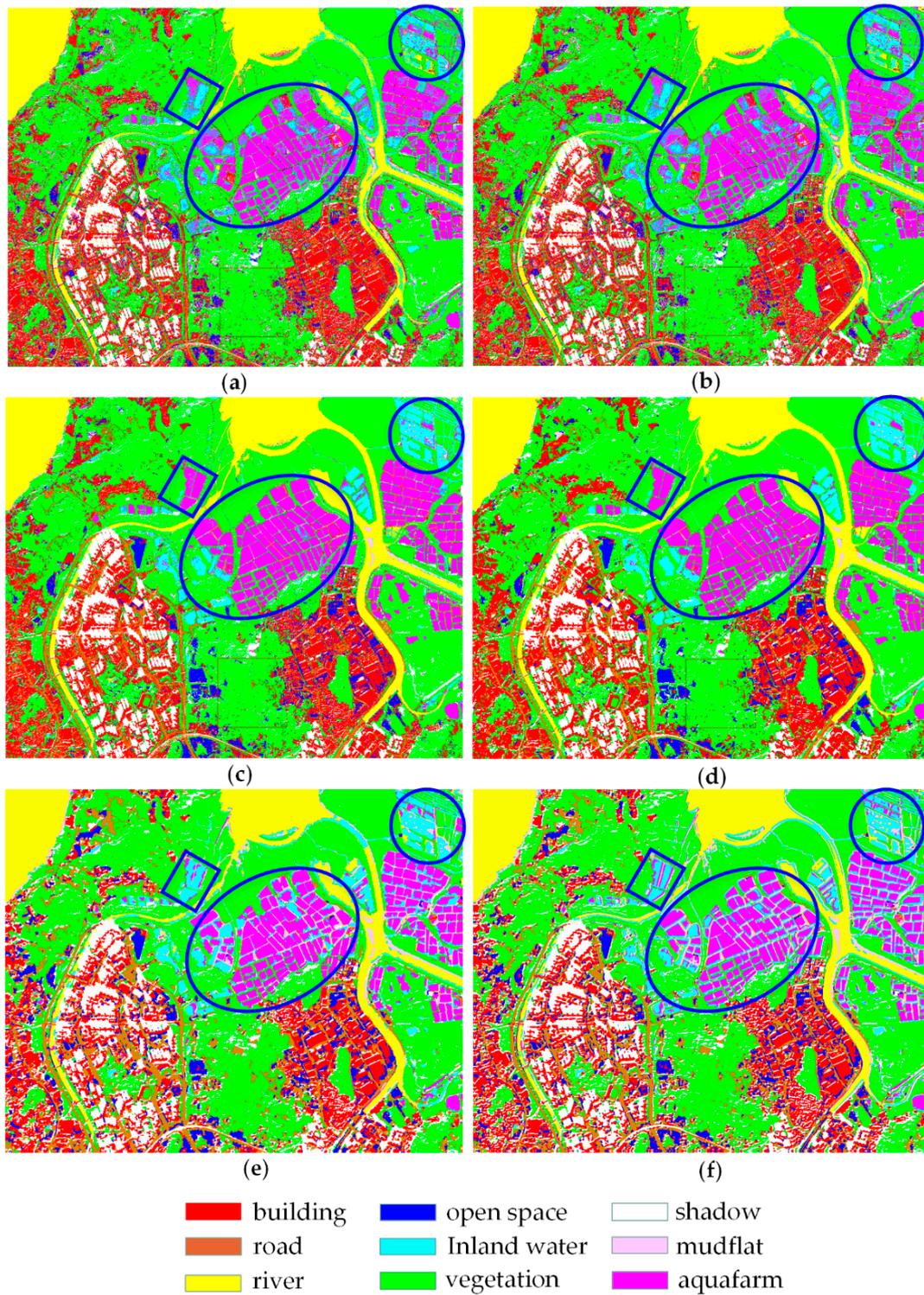
### 3.3.2. Qualitative Evaluation

The classification results of the two study area are shown in the Figures 4 and 5. For the Gaofen-2 image of Qingdao, the classification results only with the multi-spectral features (Figure 4a) is worse than the other methods based on the spectral–spatial features (Figure 4b–d), especially for the buildings and road with obvious morphological characteristics as highlighted in the white circles. For the classification results of VGG-based methods shown in Figure 4e,f, it can be found that these methods can also obtained acceptable classification results of the study area. However, some of the buildings (red) are misclassified to the open space (blue) as indicated by the black rectangles. For the Worlview-2 image of Hong Kong, as shown in Figure 5a,b, the first two methods display more complicated ‘salt and pepper’ problem. With the feature optimization strategy, the classification results of OSS show better integrality as presented in Figure 5c. Moreover, the proposed OSS-MSSC provides the best performance as shown in Figure 5d, especially for the circled water bodies. By contrast, some of the aquafarm is misclassified to inland water in the VGG-based methods as highlighted by the blue rectangles in Figure 5e,f. In addition, similar to the results of the Qingdao area, there are some buildings (red) are misclassified to the open space (blue) in the VGG-based methods. Besides, due to both the spectral and spatial information are utilized in the VGG-based method and the study areas are very large, the effectiveness of the multi-scale superpixels for the VGG-MSS method is hard to say and it is difficult to give more detailed analysis of the classification results. For further qualitative evaluation, 5 respective study areas shown in Figure 6, i.e., T1–T5 framed in Figure 3, are zoomed in and the corresponding classification results are presented in Figures 7–11.

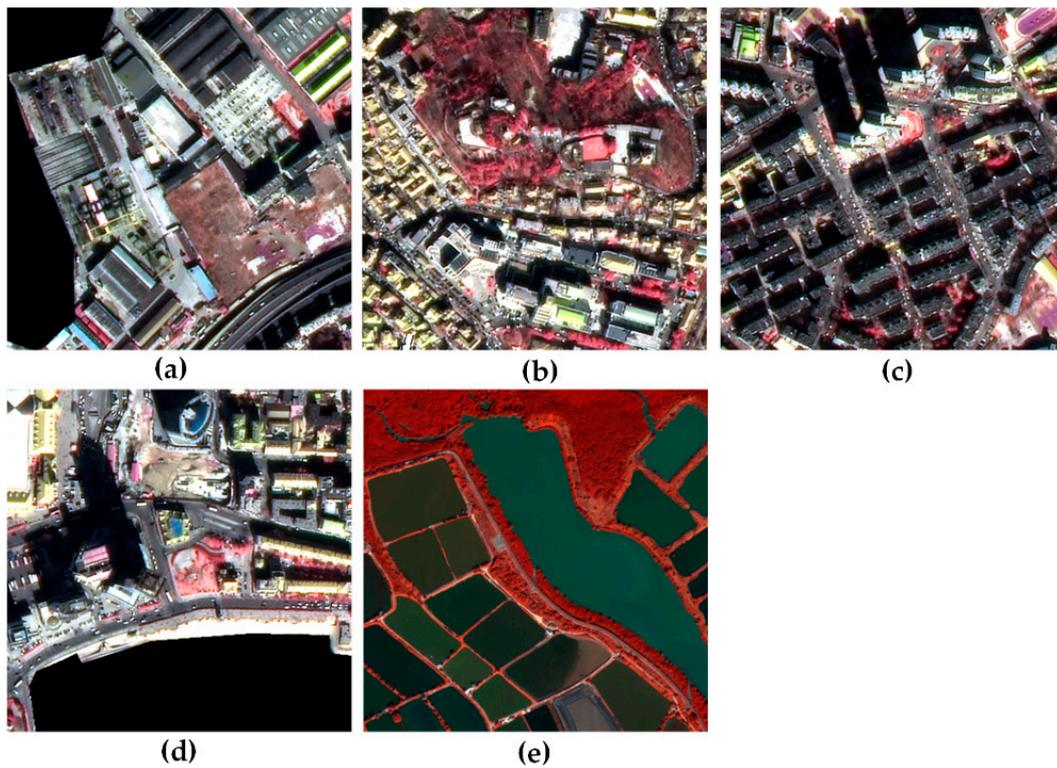
From the 5 local details of the classification results given in Figures 7–11 we can find that the combination of the spectral–spatial features effectively improves the classification results in all the tested areas. As shown in Figures 7a, 8a, 9a, 10a and 11a, without the spatial information, the classification results show serious ‘salt and pepper’ problems. Especially in the bathing beach area where the buildings (red) and the open space (blue) are heavily confused with each other as illustrated in Figures 7a and 8a and Figure 10a. Similarly, the river (yellow), inland water (cyan), and aquafarm (magenta) in aquafarm area are confused with each other as illustrated in Figure 11a. This is mainly because of the Raw multi-spectral features based method only takes use of the spectral information and cannot utilize the abundant textures information of objects, while their spectral features are similar as shown in T4 and T5 of Figure 3.



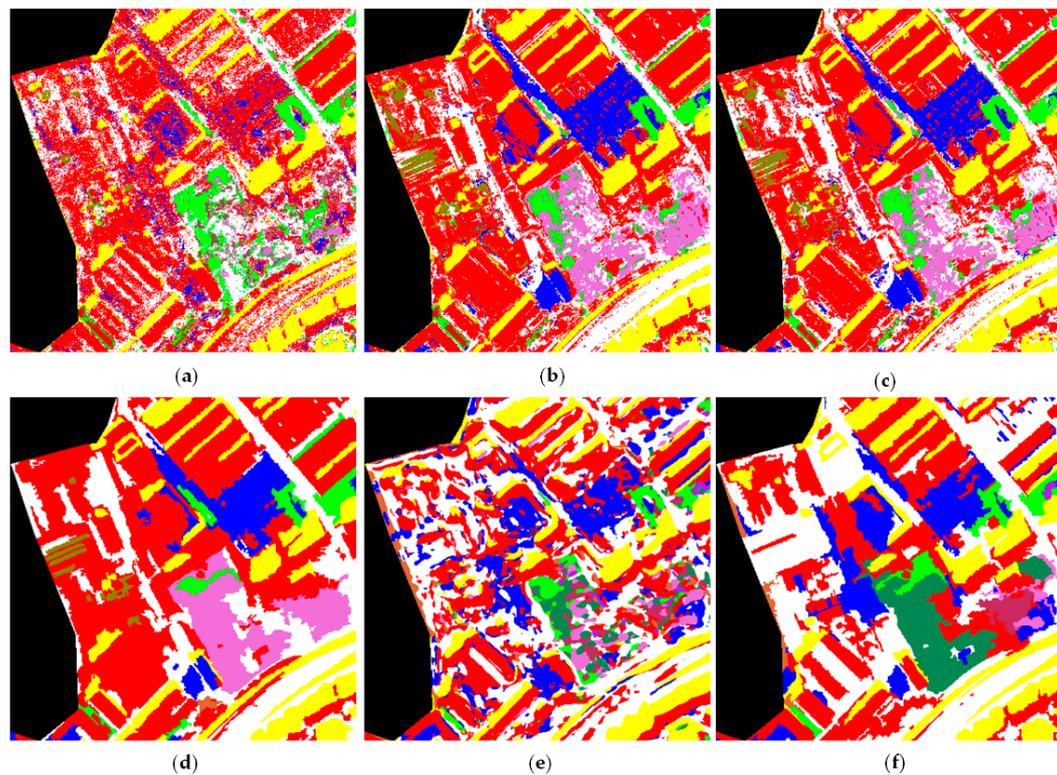
**Figure 4.** Classification results of the six methods. (a) Raw; (b) SS; (c) OSS; (d) OSS-MSSC; (e) VGG; (f) VGG-MSS.



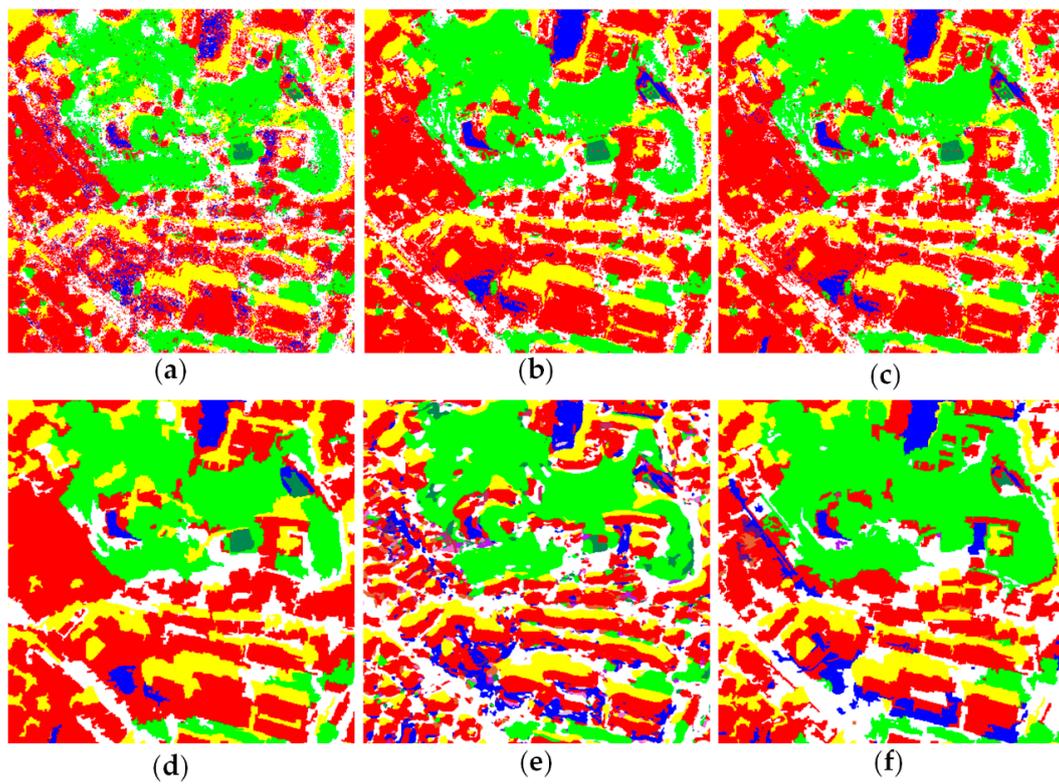
**Figure 5.** Classification results of the six methods. (a) Raw; (b) SS; (c) OSS; (d) OSS-MSSC; (e) VGG; (f) VGG-MSS.



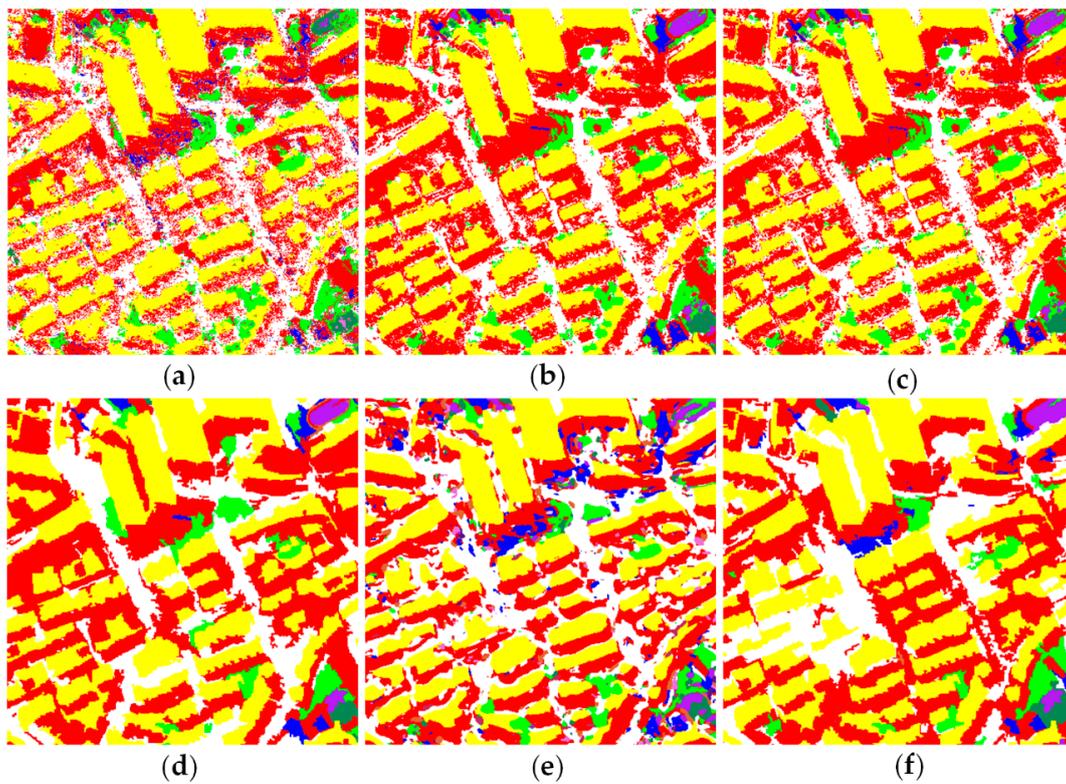
**Figure 6.** The zoomed in respective areas. (a) T1: wharf area; (b) T2: hilly city; (c) T3: residential area; (d) T4: bathing beach; (e) T5: aquafarm.



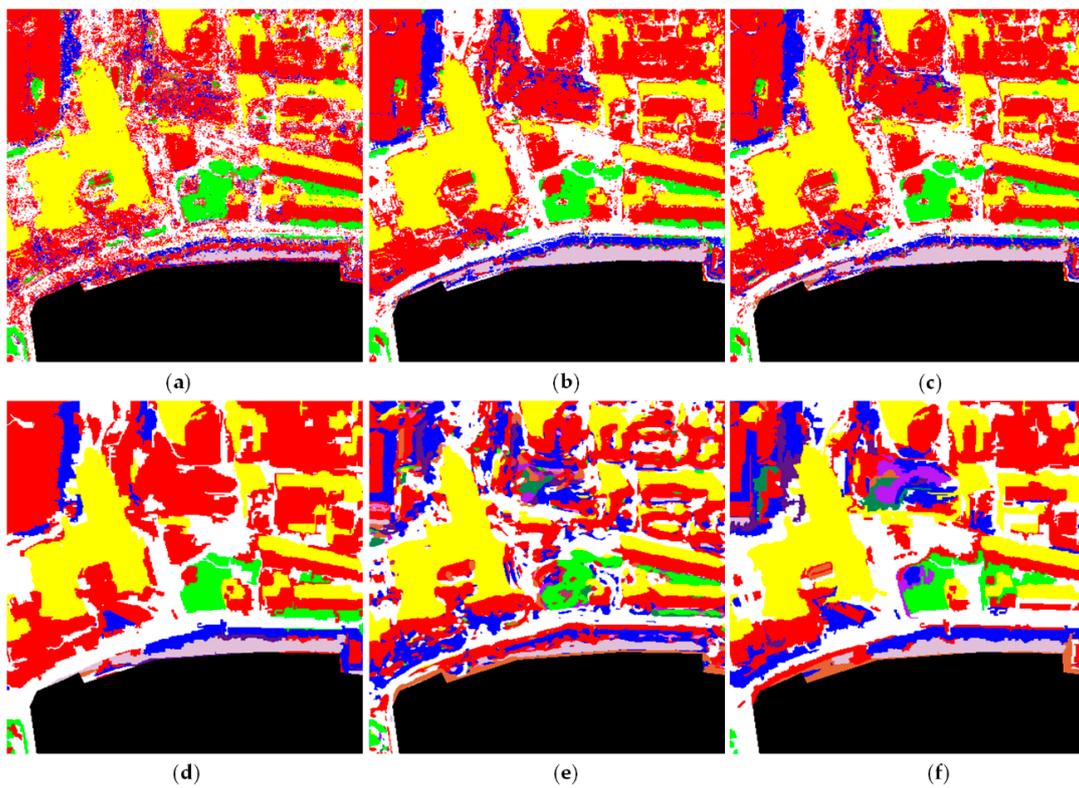
**Figure 7.** Local details of the classification results in wharf area. (a) Raw; (b) SS; (c) OSS; (d) OSS-MSSC; (e) VGG; (f) VGG-MSS.



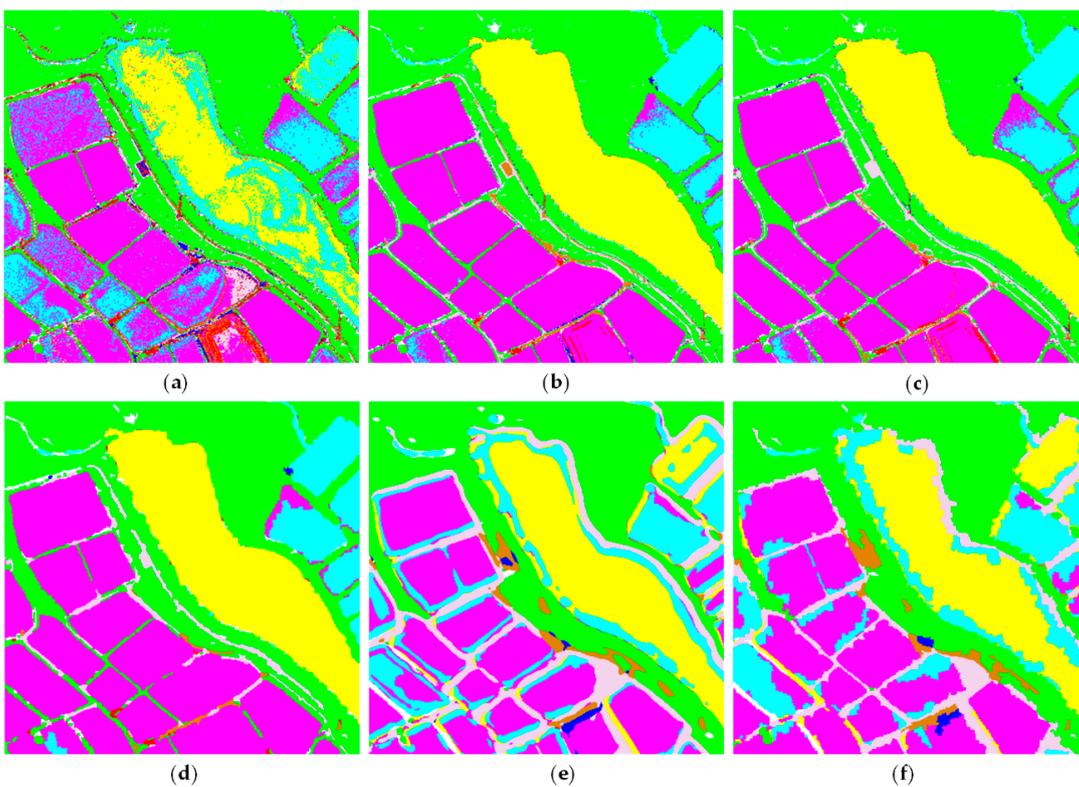
**Figure 8.** Local details of the classification results in hilly city area. (a) Raw; (b) SS; (c) OSS; (d) OSS-MSSC; (e) VGG; (f) VGG-MSS.



**Figure 9.** Local details of the classification results in residential area. (a) Raw; (b) SS; (c) OSS; (d) OSS-MSSC; (e) VGG; (f) VGG-MSS.



**Figure 10.** Local details of the classification results in bathing beach area. (a) Raw; (b) SS; (c) OSS; (d) OSS-MSSC; (e) VGG; (f) VGG-MSS.



**Figure 11.** Local details of the classification results in aquafarm area. (a) Raw; (b) SS; (c) OSS; (d) OSS-MSSC; (e) VGG; (f) VGG-MSS.

In contrast, the results shown in Figures 7b, 8b, 9b, 10b and 11b shown that the introduction of EMAPs effectively promotes the classification performance. In particular, from Figures 7b, 8b and 10b we can find that the confusion of the buildings and the open space has been largely decreased. This comes from their differences in spatial features, such as texture and roughness. For example, in the wharf area, the textures of the buildings and the open space like parking lot are very different, thus there can be separated using the SS-based method as shown in Figure 7b. For the bathing beach, the surface of the beach is very smooth while that of the building is rougher. They also can be separated with the comprehensive utilization of the spectral–spatial features as shown in Figure 10b. These results confirm that the spatial features extracted by the EMAPs describe the complicated textures of land surface covers in coastal city and this demonstrates that the spatial and geometry information is essential for the classification of the high-resolution images. Their efficiencies are also verified in the quantitative evaluation shown in Table 4.

**Table 4.** Classification Results.

Image		Raw	SS	OSS	OSS-MSSC	VGG	VGG-MSSC
Qingdao	OA	81.34	95.19	94.55	<b>97.25</b>	86.52	86.89
	Kappa	0.75	0.93	0.92	<b>0.96</b>	0.82	0.83
Hong Kong	OA	82.50	93.00	93.27	<b>95.40</b>	89.24	89.95
	Kappa	0.79	0.91	0.92	<b>0.95</b>	0.87	0.89

More importantly, from Figures 7d, 8d, 9d, 10d and 11d we can find that the classification results of the OSS-MSSC show satisfactory results in different kinds of land surface covers with various spectral and spatial characteristics. This is mainly because the multi-scale superpixels help to fully utilize the local information and overcome the “between-class homogeneity” problem in the pixel-based classification methods. Moreover, the VGG-based method faces the contour distortion problem as shown in Figures 7e, 8e, 9e, 10e and 11e. For example, the outlines of the buildings are badly smoothed and the outlines of the road are expanded as shown in Figures 7e, 8e, 9e and 10e. Even the combination of the multi-scale superpixels-based method can hardly alleviate these problem as displayed in Figures 7f, 8f, 9f and 10f. This is mainly caused by the convolution and pooling strategies of VGG. The problem of VGG-based method is more obviously revealed in the aquafarm area as shown in Figure 11. It can be found that the edges of the aquafarm (magenta) and river (yellow) are corroding inward and the surrounding areas are misclassified.

### 3.3.3. Quantitative Evaluation

From Table 4, it can be concluded that the proposed method has the best classification performance for the two study areas. For the Qingdao area, the OA reaches 97.25% with a best kappa coefficient 0.96. For the study area of Hong Kong, these evaluation criteria achieve 95.40% and 0.95, respectively. By comparison, the classification accuracy of the Raw spectral feature-based method is much lower when only spectral features are utilized as shown in Table 4. After adding the morphological profile features, the classification accuracy of the SS-based method increased greatly, the OA and kappa of Qingdao and Hong Kong images increased to 95.19%, 0.93 and 93.00%, 0.91, respectively. This demonstrate the importance of the spatial and geometry information in the mapping of the high-resolution image of the coastal cities. Moreover, for the Hong Kong image, the OSS-based method can promote the classification to a certain degree. Although the classification accuracy of OSS-based method is a little lower than that of the SS-based method for the Qingdao image, the OSS-based method greatly reduces the dimension of features and reduces the computation consuming. For the VGG and VGG-MSS based methods, owing to the deep learning of the spectral and spatial features of the high-resolution images, these methods performed superior results than the Raw-based based method. Nevertheless, due to the inaccurate boundary location problem, the classification accuracy is still lower that of the SS-, OSS-, and

OSS-MSSC-based methods. Moreover, the VGG based method usually require much more computing time than that of the proposed method.

#### 4. Discussion

A novel mapping method for coastal cities, denoted as OSS-MSSC, is proposed in this paper based on the multi-scale superpixels with optimized spectral–spatial features. This algorithm achieves effective classification of the high-resolution remote sensing image of a coastal city by using three strategies, i.e., the multi-scale superpixel generation, the spectral–spatial feature extraction and optimization, and the superpixels-based decision fusion. The contributions of these strategies are discussed as follows.

##### 4.1. The Effectiveness of Spectral–Spatial Features

In the proposed method, apart from the multi-spectral bands and the NDVI, the EMAPs are utilized to extract the geometrical structures of different ground objects which can also decrease the between-class homogeneity of different land surface covers and increase the within-class heterogeneity of one type of land surface cover. Compared to those traditional feature extraction approaches like Markov Random Field (MRF) and Gray-Level Co-occurrence Matrix (GLCM), the EMAPs avoid the need of predefined image structures and effectively preserve the geometrical characteristics of the objects with irregularly shape. From the classification results shown in Figures 4 and 5, and Figures 7c, 8c, 9c, 10c and 11c, it can be concluded that the combination of spectral and spatial features effectively improves the holistic nature of classification results and eliminates many of the salt and pepper noise. Especially for the urban land covers with large within-class heterogeneity like open space (blue), road (white), and vacant (pink). This is mainly because the spectral features of the road are similar to the buildings while their morphological characteristics are very different. In addition, the OA and kappa obtained by the SS-based method shown in Table 4 also confirms the effectiveness of the spectral–spatial features. Moreover, the classification results of the residential area of Qingdao shown in Figure 8c reveal that for some of the land types, the feature optimization strategy is effective for decreasing the ‘salt and pepper’ problem.

##### 4.2. The Effectiveness of Feature Optimization

The typical problem of high-dimensional features is their redundancy which always lead to the issue of dimensionality. Thereby, although high-dimensional spectral–spatial features have been extracted, their application needs the selection and optimization strategies. Moreover, the huge data of the extracted features requires large amount of storage space and computation times. To meet these challenges, the crossover-based feature optimization method is adopted to realize the feature selection. It has been validated that the crossover-based method outperforms many state-of-the-art intelligent algorithms in solving the dimensionality reduction problem [29]. As shown in Table 3, the operator can remarkable decrease the dimension of the EMAPs. Moreover, the multi-spectral bands and NDVI feature are demonstrated with great importance for the classification of the Gaofen-2 image. Nevertheless, one problem that cannot be ignored is that, for the Gaofen-2 image of Qingdao, the classification results after the feature selection results are little worse than that produced with all the features although the amount of data has been significantly reduced. This is because some useful information in the unselected features are lost. We will try to explore and incorporate some more excellent feature optimization algorithm in our future work.

##### 4.3. The Effectiveness of the Multi-Scale Superpixel Classification

Due to the high interior heterogeneity of different types of land surface covers, it is hard to fully utilize their spectral and spatial features to achieve high-quality image classification. To solve this problem, the multi-scale FNEA method is introduced to generated multi-scale superpixels to capture the edges of those land surface covers by the simultaneous utilization of the spectral and spatial

information. In this way, there is no need to search for the optimal scale for image segmentation and a user can set the scale parameters in a relaxed environment. Moreover, the fusion decision strategy based on the majority vote scheme can further promote the classification accuracy. The experimental results shown in Figures 5–11 and Table 4 obviously confirm the superiority of the multi-scale superpixels-based method, especially for those land surface covers with very similar spectral information and various scales like bare soil and sand in the Qingdao image. In other words, the OSS-MSSC method not only works effectively on homogeneous regions, but also is able to preserve the small local spatial structures in the high-resolution remote sensing images of coastal cities.

## 5. Conclusions

Classification of high-resolution images of a coastal city is a challenging task because of the complicated land surface covers. In this paper, the EMAPs are extracted to provide multilevel spatial description of the various land surface covers in a coastal city, and thereby decrease the inter-class variability and improve their intra-class variability. The optimized spectral–spatial features based on the crossover-based method are utilized to perform pre-classification, and the multi-scale superpixels are generated to capture the local spatial structures of the ground objects with various sizes. Experimental results on the Gaofen-2 image of Qingdao and Worldview-2 image of Hong Kong show the superiority of the proposed OSS-MSSC. The proposed method can preserve the small local spatial structures of the complicated land surface covers with irregular shapes and discriminate those land surface covers with similar spectral features. Especially, the proposed method has an outstanding capability for distinguishing the rock, bare soil, sand, and vacant areas as well as rivers, inland water, and aquafarm whose spectral features are strikingly similar to each other. However, although the crossover-based feature optimization method can effectively decrease the dimension of the spectral–spatial features, the method maybe at risk to information loss. This is mainly due to the fact that this kind of feature optimization method is based on feature reduction rather than information centralization. To meet this challenge, in the future, we will try to explore other feature optimization algorithms to promote classification accuracy.

**Author Contributions:** Conceptualization, A.Z. and S.Z.; methodology, A.Z., S.Z. and H.H.; software, A.Z., S.Z., F.L. and H.F.; validation, A.Z., Y.Z., Z.W. and J.C.; formal analysis, A.Z. and S.Z.; funding acquisition, A.Z. and G.S.

**Funding:** This work was supported by the National Natural Science Foundation of China, grant number 41801275, 41871270, the Shandong Provincial Natural Science Foundation, China, grant number ZR2018BD007, and the Fundamental Research Funds for the Central Universities, grant number 18CX02179A.

**Acknowledgments:** The authors greatly acknowledge the study material providers. The Gaofen-2 image is provided by the Ministry of Natural Resource of the People’s Republic of China and the Worldview-2 image is provided by Hongsheng Zhang of the Chinese University of Hong Kong. We also wish to thank the anonymous reviewers for their constructive comments and suggestions.

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

## References

1. Barragán, J.M.; de Andrés, M. Analysis and trends of the world’s coastal cities and agglomerations. *Ocean Coast. Manag.* **2015**, *114*, 11–20. [[CrossRef](#)]
2. de Andrés, M.; Barragán, J.M.; Sanabria, J.G. Ecosystem services and urban development in coastal social-ecological systems: The bay of cádiz case study. *Ocean Coast. Manag.* **2018**, *154*, 155–167. [[CrossRef](#)]
3. Dhiman, R.; Kalbar, P.; Inamdar, A.B. GIS coupled multiple criteria decision making approach for classifying urban coastal areas in India. *Habitat Int.* **2018**, *71*, 125–134. [[CrossRef](#)]
4. Chen, J.; Du, P.; Wu, C.; Xia, J.; Chanussot, J. Mapping urban land cover of a large area using multiple sensors multiple features. *Remote Sens.* **2018**, *10*, 872. [[CrossRef](#)]
5. Huang, X.; Weng, C.; Lu, Q.; Feng, T.; Zhang, L. Automatic labelling and selection of training samples for high-resolution remote sensing image classification over urban areas. *Remote Sens.* **2015**, *7*, 16024–16044. [[CrossRef](#)]

6. Georganos, S.; Grippa, T.; Vanhuyse, S.; Lennert, M.; Shimoni, M.; Kalogirou, S.; Wolff, E. Less is more: Optimizing classification performance through feature selection in a very-high-resolution remote sensing object-based urban application. *GISci. Remote Sens.* **2018**, *55*, 221–242. [[CrossRef](#)]
7. Fang, L.; Li, S.; Kang, X.; Benediktsson, J.A. Spectral-spatial classification of hyperspectral images with a superpixel-based discriminative sparse model. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 4186–4201. [[CrossRef](#)]
8. Vargas, J.E.; Saito, P.T.; Falcao, A.X.; De Rezende, P.J.; Dos Santos, J.A. Superpixel-based interactive classification of very high-resolution images. In Proceedings of the 27th SIBGRAPI Conference, Rio de Janeiro, Brazil, 26–30 August 2014; pp. 173–179. [[CrossRef](#)]
9. Stutz, D.; Hermans, A.; Leibe, B. Superpixels: An evaluation of the state-of-the-art. *Comput. Vis. Image Underst.* **2018**, *166*, 1–27. [[CrossRef](#)]
10. Fu, Z.; Sun, Y.; Fan, L.; Han, Y. Multiscale and multifeature segmentation of high-spatial resolution remote sensing images using superpixels with mutual optimal strategy. *Remote Sens.* **2018**, *10*, 1289. [[CrossRef](#)]
11. Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Süsstrunk, S. Slic superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal.* **2012**, *34*, 2274–2282. [[CrossRef](#)]
12. Felzenszwalb, P.F.; Huttenlocher, D.P. Efficient graph-based image segmentation. *Int. J. Comput. Vis.* **2004**, *59*, 167–181. [[CrossRef](#)]
13. Shi, J.; Malik, J. Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal.* **2000**, *22*, 888–905. [[CrossRef](#)]
14. Tu, B.; Yang, X.; Li, N.; Ou, X.; He, W. Hyperspectral image classification via superpixel correlation coefficient representation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 4113–4127. [[CrossRef](#)]
15. Belgiu, M.; Drăguț, L. Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high-resolution imagery. *ISPRS J. Photogramm. Remote Sens.* **2014**, *96*, 67–75. [[CrossRef](#)]
16. Csillik, O. Fast segmentation and classification of very high-resolution remote sensing data using slic superpixels. *Remote Sens.* **2017**, *9*, 243. [[CrossRef](#)]
17. Dingle Robertson, L.; King, D.J. Comparison of pixel-and object-based classification in land cover change mapping. *Int. J. Remote Sens.* **2011**, *32*, 1505–1529. [[CrossRef](#)]
18. Hussain, M.; Chen, D.; Cheng, A.; Wei, H.; Stanley, D. Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS J. Photogramm. Remote Sens.* **2013**, *80*, 91–106. [[CrossRef](#)]
19. Herold, M.; Liu, X.H.; Clarke, K.C. Spatial metrics and image texture for mapping urban land use. *Photogramm. Eng. Remote Sens.* **2003**, *69*, 991–1002. [[CrossRef](#)]
20. Zhang, C.; Sargent, I.; Pan, X.; Li, H.; Gardiner, A.; Hare, J.; Atkinson, P.M. An object-based convolutional neural network (oCNN) for urban land use classification. *Remote Sens. Environ.* **2018**, *216*, 57–70. [[CrossRef](#)]
21. Ghamisi, P.; Benediktsson, J.A.; Cavallaro, G.; Plaza, A. Automatic framework for spectral–spatial classification based on supervised feature extraction and morphological attribute profiles. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 2147–2160. [[CrossRef](#)]
22. Bhardwaj, K.; Patra, S. An unsupervised technique for optimal feature selection in attribute profiles for spectral-spatial classification of hyperspectral images. *ISPRS J. Photogramm. Remote Sens.* **2018**, *138*, 139–150. [[CrossRef](#)]
23. Demir, B.; Bruzzone, L. Histogram-based attribute profiles for classification of very high-resolution remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 2096–2107. [[CrossRef](#)]
24. Khodadadzadeh, M.; Li, J.; Prasad, S.; Plaza, A. Fusion of hyperspectral and lidar remote sensing data using multiple feature learning. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *8*, 2971–2983. [[CrossRef](#)]
25. Falco, N.; Mura, M.D.; Bovolo, F.; Benediktsson, J.A.; Bruzzone, L. Change detection in vhr images based on morphological attribute profiles. *IEEE Geosci. Remote Sens.* **2013**, *10*, 636–640. [[CrossRef](#)]
26. Benediktsson, J.A.; Bruzzone, L.; Chanussot, J.; Mura, M.D.; Salembier, P.; Valero, S. Hierarchical analysis of remote sensing data: Morphological attribute profiles and binary partition trees. In Proceedings of the International Conference on Mathematical Morphology and Its Applications to Image and Signal Processing, Verbania-Intra, Italy, 6–8 July 2011; pp. 306–319. [[CrossRef](#)]
27. Song, B.; Li, J.; Mura, M.D.; Li, P.; Plaza, A.; Bioucas-Dias, J.M.; Benediktsson, J.A.; Chanussot, J. Remotely sensed image classification using sparse representations of morphological attribute profiles. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 5122–5136. [[CrossRef](#)]

28. Xia, J.; Ghamisi, P.; Yokoya, N.; Iwasaki, A. Random Forest Ensembles and Extended Multiextinction Profiles for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 202–216. [[CrossRef](#)]
29. Zhang, A.; Ma, P.; Liu, S.; Sun, G.; Huang, H.; Zabalza, J.; Wang, Z.; Lin, C. Hyperspectral band selection using crossover-based gravitational search algorithm. *IET Image Process.* **2018**, *13*, 280–286. [[CrossRef](#)]
30. Li, C.; Huang, J.; Li, Z.; Wang, R. Plane-wave least-squares reverse time migration with a preconditioned stochastic conjugate gradient method. *Geophysics* **2018**, *83*, S33–S46. [[CrossRef](#)]
31. Huang, J.; Liao, W.; Li, Z. A multi-block finite difference method for seismic wave equation in auxiliary coordinate system with irregular fluid-solid interface. *Eng. Comput.* **2018**, *35*, 334–362. [[CrossRef](#)]
32. Yong, P.; Huang, J.; Li, Z.; Liao, W.; Qu, L.; Li, Q.; Liu, P. Optimized equivalent staggered-grid fd method for elastic wave modelling based on plane wave solutions. *Geophys. J. Int.* **2017**, *208*, 1157–1172. [[CrossRef](#)]
33. Huang, C.-L.; Wang, C.-J. A GA-based feature selection and parameters optimization for support vector machines. *Expert Syst. Appl.* **2006**, *31*, 231–240. [[CrossRef](#)]
34. Ghamisi, P.; Benediktsson, J.A. Feature selection based on hybridization of genetic algorithm and particle swarm optimization. *IEEE Geosci. Remote Sens.* **2015**, *12*, 309–313. [[CrossRef](#)]
35. Rashedi, E.; Nezamabadi-Pour, H.; Saryazdi, S. GSA: A gravitational search algorithm. *Inform. Sci.* **2009**, *179*, 2232–2248. [[CrossRef](#)]
36. Güvenç, U.; Sönmez, Y.; Duman, S.; Yörükeren, N. Combined economic and emission dispatch solution using gravitational search algorithm. *Sci. Iran.* **2012**, *19*, 1754–1762. [[CrossRef](#)]
37. Zhang, A.; Liu, S.; Sun, G.; Huang, H.; Ma, P.; Rong, J.; Ma, H.; Lin, C.; Wang, Z. Clustering of Remote Sensing Imagery Using a Social Recognition-Based Multi-objective Gravitational Search Algorithm. *Cogn. Comput.* **2018**, 1–10. [[CrossRef](#)]
38. Rodriguez-Galiano, V.F.; Chica-Olmo, M.; Abarca-Hernandez, F.; Atkinson, P.M.; Jeganathan, C. Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sens. Environ.* **2012**, *121*, 93–107. [[CrossRef](#)]
39. Ghosh, A.; Sharma, R.; Joshi, P.K. Random forest classification of urban landscape using Landsat archive and ancillary data: Combining seasonal maps with decision level fusion. *Appl. Geogr.* **2014**, *48*, 31–41. [[CrossRef](#)]
40. Kavzoglu, T.; Colkesen, I. An assessment of the effectiveness of a rotation forest ensemble for land-use and land-cover mapping. *Int. J. Remote Sens.* **2013**, *34*, 4224–4241. [[CrossRef](#)]
41. Kavzoglu, T. Chapter 33—Object-Oriented Random Forest for High-resolution Land Cover Mapping Using Quickbird-2 Imagery. In *Handbook of Neural Computation*; Samui, P., Sekhar, S., Balas, V.E., Eds.; Academic Press: Cambridge, MA, USA, 2017; pp. 607–619. [[CrossRef](#)]
42. Han, M.; Zhu, X.; Yao, W. Remote sensing image classification based on neural network ensemble algorithm. *Neurocomputing* **2012**, *78*, 133–138. [[CrossRef](#)]
43. Padmanaban, R.; Bhowmik, A.K.; Cabral, P.; Zamyatin, A.; Almegdadi, O.; Wang, S. Modelling Urban Sprawl Using Remotely Sensed Data: A Case Study of Chennai City, Tamilnadu. *Entropy* **2017**, *19*, 163. [[CrossRef](#)]
44. Cheema, M.J.M.; Bastiaanssen, W.G.M. Land use and land cover classification in the irrigated Indus Basin using growth phenology information from satellite data to support water management analysis. *Agric. Water Manag.* **2010**, *97*, 1541–1552. [[CrossRef](#)]
45. Bassa, Z.; Bob, U.; Szantoi, Z.; Ismail, R. Land cover and land use mapping of the iSimangaliso Wetland Park, South Africa: Comparison of oblique and orthogonal random forest algorithms. *J. Appl. Remote Sens.* **2016**, *10*, 15–17. [[CrossRef](#)]
46. Li, S.; Ni, L.; Jia, X.; Gao, L.; Zhang, B.; Peng, M. Multi-scale superpixel spectral-spatial classification of hyperspectral images. *Int. J. Remote Sens.* **2016**, *37*, 4905–4922. [[CrossRef](#)]
47. Tong, H.; Maxwell, T.; Zhang, Y.; Dey, V. A supervised and fuzzy-based approach to determine optimal multi-resolution image segmentation parameters. *Photogramm. Eng. Remote Sens.* **2012**, *78*, 1029–1044. [[CrossRef](#)]
48. Costa, G.; Feitosa, R.; Cazes, T.; Feijó, B. Genetic adaptation of segmentation parameters. In *Object-Based Image Analysis*; Springer: Heidelberg/Berlin, Germany, 2008; pp. 679–695. [[CrossRef](#)]
49. Chen, J.; Pan, D.; Mao, Z. Image-object detectable in multiscale analysis on high-resolution remotely sensed imagery. *Int. J. Remote Sens.* **2009**, *30*, 3585–3602. [[CrossRef](#)]
50. Das, A.; Bhardwaj, K.; Patra, S. Morphological complexity profile for the analysis of hyperspectral images. In Proceedings of the 2018 4th International Conference on Recent Advances in Information Technology (RAIT), Dhanbad, India, 15–17 March 2018; pp. 1–6. [[CrossRef](#)]

51. Mura, M.D.; Benediktsson, J.A.; Waske, B.; Bruzzone, L. Morphological attribute profiles for the analysis of very high-resolution images. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 3747–3762. [[CrossRef](#)]
52. Ghamisi, P.; Couceiro, M.S.; Benediktsson, J.A. A novel feature selection approach based on FODPSO and SVM. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 2935–2947. [[CrossRef](#)]
53. Marpu, P.R.; Pedernana, M.; Mura, M.D.; Benediktsson, J.A.; Bruzzone, L. Automatic generation of standard deviation attribute profiles for spectral–spatial classification of remote sensing data. *IEEE Geosci. Remote Sens.* **2012**, *10*, 293–297. [[CrossRef](#)]
54. Zhang, A.; Sun, G.; Ma, P.; Jia, X.; Ren, J.; Huang, H.; Zhang, X. Coastal Wetland Mapping with Sentinel-2 MSI Imagery Based on Gravitational Optimized Multilayer Perceptron and Morphological Attribute Profiles. *Remote Sens.* **2019**, *11*, 952. [[CrossRef](#)]
55. Maurer, T. How to pan-sharpen images using the Gram-Schmidt pan-sharpen method—a recipe. In Proceedings of the 2013 ISPRS Hannover Workshop, Hannover, Germany, 21–24 May 2013; pp. 239–244. [[CrossRef](#)]
56. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. In Proceedings of the 2015 International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015; pp. 1–14.



© 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/>).