Research Article

Advantage of Combining OBIA and Classifier Ensemble Method for Very High-Resolution Satellite Imagery Classification

Ruimei Han,[1,2] Pei Liu [ID],[1,2,3] Guangyan Wang,[4] Hanwei Zhang,[1,2] and Xilong Wu[1]

[1]*Key Laboratory of Spatio-Temporal Information and Ecological Restoration of Mines(MNR), Henan Polytechnic University, Jiaozuo, Henan 454003, China*
[2]*School of Surveying and Mapping Land Information Engineering, Henan Polytechnic University, Jiaozuo, 454003 Henan, China*
[3]*Collaborative Innovation Center of Aerospace Remote Sensing Information Processing and Application of Hebei Province, 065000 Langfang, China*
[4]*Jiangsu Province Engineering Investigation and Research Institute Co. Ltd., Yangzhou, Jiangsu 225001, China*

Correspondence should be addressed to Pei Liu; cumtlp@qq.com

Accurate and timely collection of urban land use and land cover information is crucial for many aspects of urban development and environment protection. Very high-resolution (VHR) remote sensing images have made it possible to detect and distinguish detailed information on the ground. While abundant texture information and limited spectral channels of VHR images will lead to the increase of intraclass variance and the decrease of the interclass variance. Substantial studies on pixel-based classification algorithms revealed that there were some limitations on land cover information extraction with VHR remote sensing imagery when applying the conventional pixel-based classifiers. Aiming at evaluating the advantages of classifier ensemble strategies and object-based image analysis (OBIA) method for VHR satellite data classification under complex urban area, we present an approach-integrated multiscale segmentation OBIA and a mature classifier ensemble method named random forest. The framework was tested on Chinese GaoFen-1 (GF-1), and GF-2 VHR remotely sensed data over the central business district (CBD) of Zhengzhou metropolitan. Process flow of the proposed framework including data fusion, multiscale image segmentation, best optimal segmentation scale evaluation, multivariance texture feature extraction, random forest ensemble learning classifier construction, accuracy assessment, and time consumption. Advantages of the proposed framework were compared and discussed with several mature state-of-art machine learning algorithms such as the k-nearest neighbor (KNN), support vector machine (SVM), and decision tree classifier (DTC). Experimental results showed that the OA of the proposed method is up to 99.29% and 98.98% for the GF-1 dataset and GF-2 dataset, respectively. And the OA is increased by 26.89%, 11.79%, 11.89%, and 4.26% compared with the traditional machine learning algorithms such as the decision tree classifier (DTC), support vector machine (SVM), k-nearest neighbor (KNN), and random forest (RF) on the test of the GF-1 dataset; OA increased by 32.31%, 13.48%, 9.77%, and 7.72% for the GF-2 dataset. In terms of time consuming, by rough statistic, OBIA-RF spends 223.55 s, SVM spends 403.57 s, KNN spends 86.93 s, and DT spends 0.61 s on average of the GF-1 and GF-2 datasets. Taking the account classification accuracy and running time, the proposed method has good ability of generalization and robustness for complex urban surface classification with high-resolution remotely sensed data.

1. Introduction

The classification accuracy of remotely sensed data and its sensitivity to classification algorithms have a critical importance for the geospatial community, as classified images provide the base layers for many applications and models [1].

Recent availability of submeter resolution imagery from advanced satellite sensors, such as WorldView-3 and Chinese GaoFen series, can provide new opportunities for detailed urban land cover mapping at the object level [2]. Applications such as environmental monitoring, natural resource management, and change detection require more accurate,

detailed, and constantly updated land cover-type mapping [3]. Detailed urban land cover information is not only essential for understanding the urban environment changes and monitoring and managing the urban ecological environment but also for supporting the government to make a decision on urban expansion, urban planning, and management [4–6]. In the last few decades, it has become an effective and convenient mean to obtain this information from remotely sensed imagery, because of its unique advantages of frequent and wide coverage, by machine learning classification technology [7].

Most of land use and land cover classification research are traditionally based on low- and medium-resolution remotely sensed imagery, such as MODIS [6, 8, 9], Landsat [10–12], and SPOT1/4 [13]. However, urban surface coverage presents high-frequency heterogeneity, resulting in a large number of mixed pixels in medium- and low-resolution images. With the rapid development of sensor technology, a large number of high-resolution remotely sensed imagery (IKONOS, Quickbird, GeoEye-1, World-View-1-4, GF-1/2, etc.) in meters or submeters are becoming more and more popular [14]. With the characteristics of high definition and abundant spatial information, high-resolution satellite image can compensate the shortcomings of mixing pixels in low- and medium-resolution images in urban land cover classification [15, 16]. And high spatial resolution images, where spatial resolution is equal or a little equal to 4 meters, could make it possible to map complex urban surface. A major challenge in using high spatial resolution for detailed urban mapping comes from the high level of intraclass spectral variability, such as building roof and road, and low level of interclass spectral variability, such as water body and shadow. In this condition, traditional pixel-based classification algorithms such as the maximum likelihood classification (MLC) can easily make missclass error and generate the salt-and-pepper effect which may reduce classification accuracy for very high-resolution imagery.

There are currently various classification algorithms, each with its own advantages and limitations [17]. And that, common mature statistical-based machine learning algorithms, such as MLC, requires hypothesis that training data follows a normal distribution, but high-resolution images cannot meet this requirement. Many previous studies revealed a bunch of machine learning algorithms such as support vector machines (SVM) [1], artificial neural networks (ANN) [18], and decision tree [19] have been popular for land cover classification. These classifiers always have limitations in practical applications in areas such as volatile and complex urban area, due to the enhanced complex of spatial relationship between pixels and the complex earth's surface phenomenon [21, 22]. Recently, ensemble methods have been introduced to integrate multiple single classifiers to improve classification performances. The combination of multisource remote sensing and geographic data is believed to offer improved accuracies in land cover classification [23]. In general, there are two steps to build the ensemble, namely, generating base learners and combining base learners. In order to obtain a good ensemble, the base learner should be as accurate as possible and as diverse as possible. Due to its high potential and superior performance, ensemble methods have been employed in a

remote sensing community. Existing theoretical and empirical studies have reported that ensemble classifiers can obtain more accuracy prediction and outperform individual classifiers [3, 17, 23–25]. The random forest (RF) classifier, as one of the more popular ensemble learning algorithms in recent years, is composed of multiple decision trees in that each tree is trained using bootstrap sampling and employing the majority vote for the final prediction [26, 27]. It has received increasing attention due to its excellent classification result, the ability to avoid overfitting, and the rapid speed to process [28–31].

In order to overcome limitations of pixel-based classification, object-based image analysis (OBIA) or geospatial object-based image analysis (GEOBIA) has been introduced to improve the quality of information extraction from high-resolution imagery. Image segmentation is a critical and important step in (geographic) object-based image analysis (GEOBIA or OBIA). The final feature extraction and classification in OBIA are highly dependent on the quality of image segmentation [32]. There are two main steps, containing segmentation and classification in OBIA. The processed object of OBIA is not a pixel, but an object composed of multiple adjacent homogenous pixels through segmentation, which containing not only the spectral information but also the textual and contextual information from imagery [32, 33]. Due to its advantages, OBIA has been more popular in the remote sensing community and successfully applied in land cover classification [34–36]. Many previous studies have showed that the OBIA method had outperformed the pixel-based classification [32, 37–40]. However, the number of input features used for classification has grown exponentially, some of which are irrelevant and redundant features, affecting the performance of the classifier, especially, when the purpose of segmentation has been changed from helping pixel labeling to object identification at present era [32].

In this paper, we verify the ability of GF-1 and GF-2 very high-resolution imagery in urban land use and land cover classification. For this purpose, we combined the random forest ensemble classifier with the OBIA method. We test the method on two selected complex urban areas of a metropolis city. And the proposed strategy was also compared with the pixel-based random forest and the state-of-art mature machine learning algorithm including the SVM, KNN, and DT classifiers from classification accuracy and operational efficiency aspects.

The rest of this paper is organized as follows: a brief introduction about the study area and dataset and preprocessing are given in Section 2. The framework and details of the proposed methodology strategy based on object-oriented analysis and random forest are drawn in Section 3. The results and discussion are shown in Section 4. Finally, the conclusions are drawn in Section 5.

2. Study Area and Data Preprocess

The study site is a central business district (CBD) of Zhengdong new district which is located in the eastern part of Zhengzhou city, capital of Henan province, and is a new urban area invested and developed by Zhengzhou Municipal Committee, municipal government in accordance with the
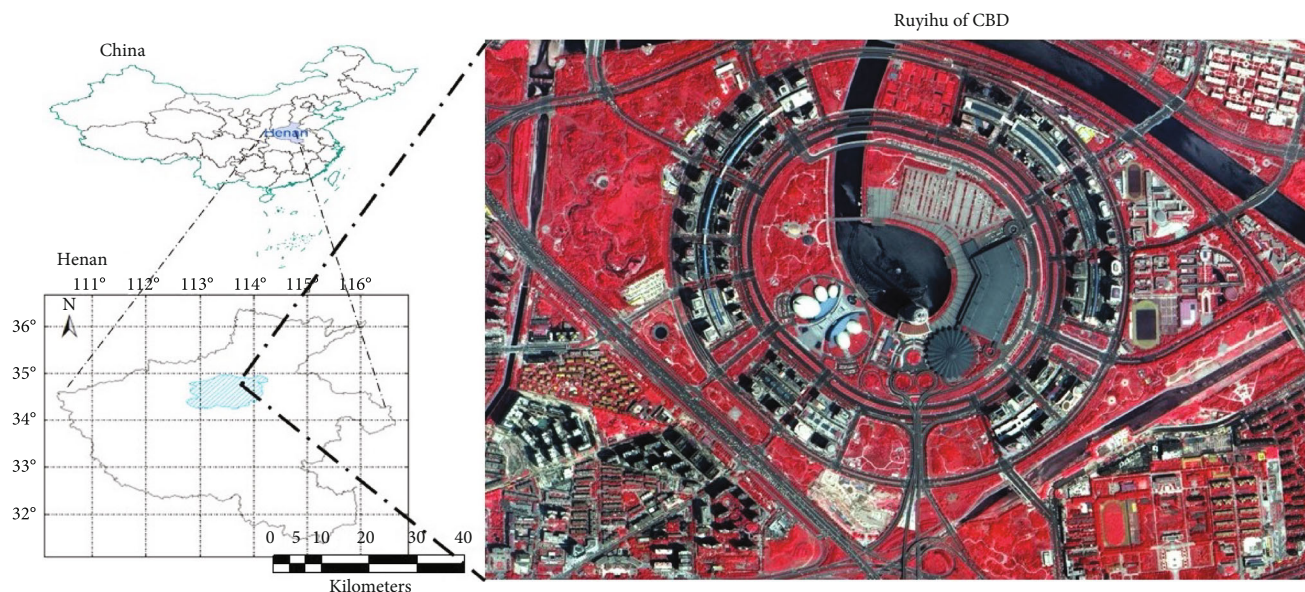
Ruyihu of CBD

Figure 1: The study area location and corresponding GF-2 imagery.

Table 1: The classification scheme of land cover in the study area.

Category	Symbols	Description
Urban building areas (UB)	1	Residential area, campus, low density building
Urban commercial areas (UC)	2	Commercial area, high density building
Urban green areas (UG)	3	Grassland, tree group, oasis in the lake, artificial grass
Urban road areas (UR)	4	Cement and asphalt pavement
Urban water area (UW)	5	River, lake
High building shadow (HS)	6	Shadow of high buildings

State Council approved the City of Zhengzhou city master plan in order to implement the megacity framework, expand the size of the city, and accelerate urbanization and urban modernization strategy. Based on the National Economic and Technological Development Zone, the original area of CBD is about $25 \, km^2$, west from 107 national road, east to Jingzhu Expressway, south to the airport highway, north to Lianhuo Expressway, and the long-term planning area of CBD is about $150 \, km^2$. The study area is focused on Ruyihu, the center of CBD (see Figure 1), which is surrounded by three landmarks of the CBD-Zhengzhou International Convention and Exhibition Center, Henan Arts Center, Zhengzhou Convention and Exhibition Hotel. The land surface is dominated by human-made material, which is a challenging task to identify different land use and land cover types. According to the planning and construction situation, the types of surface cover are mainly divided into the urban building areas (UB), urban commercial area (UC), urban green areas (UG), urban road areas (UR), urban water area (UW), and high building shadow (HS) (see Table 1). Data availability statement stated that the very high-resolution remotely sensed data used in this research is provided by Henan Data and Application Center of the High Resolution Earth Observation System through signing a contract with National Defense Science and Technology Bureau of Henan.

Unfortunately, we do not have the priority to share the high-resolution satellite remote sensed data. Anyway, we can share our code used in this research. Researchers can test algorithms using these codes with their own datasets and repeat experiments to obtain similar research conclusions. Researchers who are interested in this code can download it from hyperlink https://pan. http://baidu.com/s/19nXD7oHwq0FnpZJ7T6p5HQ, using password 6xae, or contact with the corresponding author to obtain source data to conduct secondary analysis.

2.1. Remote Sensing Data and Preprocessing. Under the background of "Chinese high-resolution earth observation" major project, a series of high-resolution satellites have been launched, involving GF-1 (2 m res. panchromatic camera/8 m res., multispectral camera/16 m res., and wide-angle multispectral camera), GF-2 (1 m res., panchromatic camera/4 m res., and multispectral camera), GF-3 (1 m res., C-band synthetic aptitude radar), GF-4 (50 m res., fixed-point camera in geostationary orbit), GF-5 (VNIR hyperspectral camera), GF-6 (2 m res., wide-angle multispectral camera), and GF-7 (stereographic cartography cameras). GF-1 satellite is the first low earth orbit remote sensing satellite of China's high-resolution earth observation system, which breaks through the key technologies of optical remote sensing for

TABLE 2: Sensor parameters of GF-1 and GF-2 satellites.

Parameters			PAN/multispectral	
			GF1	GF2
	Orbit type		Sun synchronization	Sun synchronization
Satellite	Orbit altitude		646 km	631 km
	Repeat time		41 days	69 days
	Panchromatic		0.45–0.90 μm	0.45–0.90 μm
		Blue	0.45–0.52 μm	0.45–0.52 μm
Spectral range		Green	0.52–0.59 μm	0.52–0.59 μm
	Multispectral	Red	0.63–0.69 μm	0.63–0.69 μm
		NIR	0.77–0.89 μm	0.77–0.89 μm
Spatial resolution	Panchromatic		2 m	1 m
	Multispectral		8 m	4 m
Width			60 km	45 km
Receive time			2015-07-14	2015-7-3
Orbit ID			11937	5113
Product type			Standard	Standard
Product level			Level 1A	Level 1A
(Path, row)			(3, 97)	(3,152)
(Width, height)			(4548, 4544), (18192, 18164)	(29200, 27620)
Cloud			0%	5%
Solar(azimuth, zenith)			(141.278, 73.7496)	(134.537, 21.6698)
Satellite(azimuth, zenith)			(297.354, 88.3146)	(283.342, 83.754)

high spatial resolution and multispectral and wide coverage. It can meet the needs of research data support in the fields of resources and environment, precision agriculture, and disaster measurement, which has become an important means of information services and other aspects. It is of great strategic significance to improve the level of satellite engineering in China and the self-sufficiency rate of high-resolution data.

The selected remotely sensed data is GF-1 and GF-2 very high-resolution satellite images, which was acquired on July 14, 2015 and July 30, 2015, respectively. The specific parameters are shown in Table 2.

The preprocess of selected dataset includes radiation calibration, atmospheric correction, geometric registration, orthorectification, image fusion (NNDiffuse pan-sharpening algorithm), and image resize. Radiation calibration is the process of converting DN values of image data into apparent reflectivity using atmospheric correction techniques. Equation (1) can be used to convert the channel observation DN value to equivalent brightness value.

$$L_\varepsilon(\lambda_\varepsilon) = \text{Gain} * \text{DN} + \text{Bias}, \tag{1}$$

where Gain is the calibration slope, DN is satellite observation value, Bias is calibration intercept, and these parameters can be obtained from meta file with satellite data. Then, the apparent reflectivity can be calculated based on the brightness value using

$$\rho = \frac{\pi L_\lambda d^2}{\text{ESUN}_\lambda \cos \theta}, \tag{2}$$

where ESUN is solar spectral radiation, d is solar-earth distance, and $\cos \theta$ is the zenith angle of sun.

The selected atmospheric correction is based on a 6S radiative transfer model, which is a package included in Pixel Information Export (http://www.piesat.cn/en/index.html). The purpose of atmospheric correction is to eliminate the absorption and dispersion from the sun and target.

Geometric correction includes image registration and orthorectification. The purpose of geometric correction is to correct image deformation caused by system and nonsystemic factors. In this research, the image-to-image registration method was selected to correct multispectral data based on panchromatic data of GF-1 and GF-2 sensors, respectively.

Orthorectification is the process of correcting image space and geometric distortion to generate a multicenter projection plane orthographic image. In addition to correcting geometric distortions caused by general system factors, it can also eliminate geometric distortion caused by terrain.

Image fusion is the process of generating new images under the prescribed geographical coordinate system according to a certain algorithm. This study combines multispectral data with high spatial resolution and single-band images with high spatial resolution, making the fused images have both high spatial resolution and rich spectral resolution.

3. Methodology

The proposed methodology in this research (shown in Figure 2) includes three main stages: (1) multiscale segmentation and multifeature extraction; (2) construction of the state-of-art machine learning algorithms such as decision
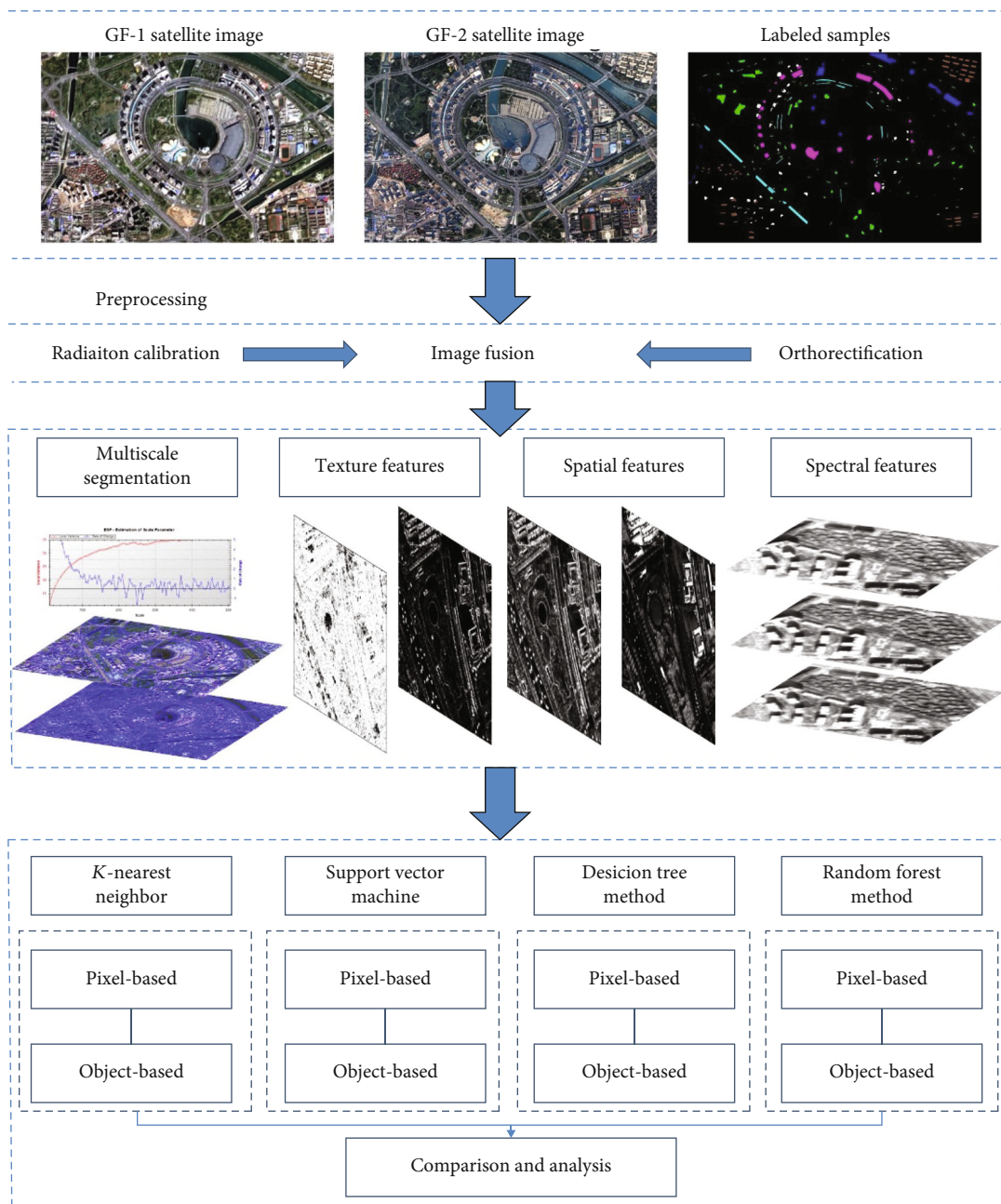
FIGURE 2: Framework of urban land cover classification using GF-1 and GF-2 remotely sensed data.

tree classifier (DTC), random forest (RF), support vector machine (SVM), k-nearest neighbor (KNN), and object-based image analysis random forest (OBIA-RF); and (3) accuracy assessment and comparison analysis and discussion; more details can be found in Figure 2.

3.1. Multiscale Segmentation and Feature Extraction. Very high-resolution (VHR) remote sensing images have a limitation in spectral information which means there are 4 spectral bands including green, blue, red, and near infrared in general. While the VHR remote sensing images are always rich in detailed characters, more specific details of land surface will be presented on the images. In order to overcome this shortcoming, we conquer the disadvantage and make full use of

advantages of these data. After data preprocessing, we performed multiscale segmentation and employed feature extraction based on the segmented results to obtain multifeature image sets as inputs of image classification models.

Quality of segmentation has a direct effect on the performance of classification, which is related to the segmentation parameters selected by an analyst. Most of the segmentation algorithms are regarded as a subjective task with the trial-and-error strategy. The multiscale segmentation algorithm, the most popular method currently, employed in this experiment is merging pixels of the original image into small object patches from bottom to top, and then merging the small patches into large patches to complete the merging of the regional objects [41–43]. Three major parameters for the

TABLE 3: Overview of extracted features based on segmentation.

Feature type	Quantity	Feature name
Spectral feature	9	Spectral bands 1-4, brightness, ratio band, standard deviation band4, NDVI, NDWI
Texture feature	8	Entropy, mean, variance, homogeneity, contrast, dissimilarity, correlation, angular second moment
Spatial feature	7	Area, border index, compactness, density, length/width, length, shape index

multiscale segmentation algorithm are scale, shape, and compactness, defining within-object homogeneity. Here, we select the appropriate scale parameters and heterogeneity standard specifications to ensure the highest homogeneity within the generated object and the heterogeneity between adjacent objects and other objects. Scale parameter is considered the most effective parameter affecting the segmentation quality [44–46]. In this study, shape parameter was set as 0.1. The optimal segmentation scale parameter of GF-1 image is quantitatively evaluated by ESP-2 (estimation of scale parameter), of which the principle is to select the optimal scales based on the rate of change (ROC) curve for the local variance (LV)of object heterogeneity at a corresponding scale [44, 47]. Peaks value of ROC-LV curves were considered the most appropriate segmentation scales at which the image can be segmented in the most optimal levels. Methodology model of ROC can be described as

$$ROC = \frac{LV_L - LV_{L-1}}{LV_{L-1}} \times 100, \tag{3}$$

where LV_L is mean standard deviation of the object in the L layer and LV_{L-1} is mean standard deviation in the next lower layer. When ROC was obtained, the optimal segmentation scale parameter is selected by visually interpreting based on segmentation result and the boundary matching effect of the actual feature [47].

On the basis of the best segmentation result, a total of 24 spectral features, texture features, and spatial geometric features were extracted. The extracted spectral, textural, and spatial features from segmented VHR remotely sensed data can be summarized as shown in Table 3.

In detail, the selected spectral features are mean value of all four bands (which means average value of all image objects). The brightness feature is that the sum of the average values of the layers containing spectral information divided by the number of layers of the image object, which can be calculated by

$$B = \frac{1}{N_{band}} \sum_{i=1}^{n} \overline{C_i}, \tag{4}$$

where B is the brightness value of the object, N_{band} is the total number of bands contained in the object, and C_i is the average gray value of the object.

In the high-resolution image, since the reflectivity of the water body in the near-infrared band is significantly lower than that of other ground objects, the shadows are very similar in many features with water body so that they are difficult to separate. In order to highlight the water, the ratio

and standard deviation of the fourth band and NDWI were additionally extracted. The ratio of the fourth band is the average gray value of all pixels in the fourth band divided by the average gray value of all pixels in all 4 bands of the image. In addition, only layers containing spectral information can be used to obtain reasonable results. The standard deviation of band 4 is calculated from all the pixel values contained in an object in band 4.

The NDWI refers to the normalized ratio index between the green band and the near-infrared band in the image. Using NDWI can better distinguish the water in the image from other features. It can be calculated by

$$NDWI = \frac{B_{green} - B_{nir}}{B_{green} + B_{nir}}, \tag{5}$$

where B_{green} is the value of object in the green band and B_{nir} is the value of object in the near-infrared band.

NDVI is a vegetation index proposed based on the reflection characteristics of vegetation in the visible and infrared bands. It is the ratio of the difference between the reflection intensity value in the visible red band and the reflection intensity value in the near-infrared band to the sum of the two. The formula can be described as

$$NDVI = \frac{B_{nir} - B_{red}}{B_{nir} + B_{red}}, \tag{6}$$

where B_{nir} is the value of object in the near-infrared band and B_{red} is the value of object in the red band.

Although limited to spectral information, high-resolution remote sensing images contain rich geometric and structural information, which can reflect the spatial distribution and geometric forms of ground objects. The selected texture features contain eight features unit extracted based on the gray level cooccurrence. The selected eight textural features extracted from GLCM include entropy, mean, variance, homogeneity, contrast, dissimilarity, correlation, and angular second moment. Gray-level cooccurrence matrix (GLCM), which is calculated based on statistic method, is also known as gray-level spatial dependence matrix considered one of the most popular techniques used for texture analysis. GLCM has strong ability to assess texture features by considering spatial relationship of pixels and its surrounding. GLCM mean value is not simply the average of all original pixel values; pixel value is weighted by its frequency of its occurrence in combination with a certain neighbor pixel value. Variance in GLCM texture performs the same task as does the common descriptive statistic called variance.

Entropy measures the complexity of a given image, which reflects the sharpness of the image and the depth of the texture; entropy can be calculated using

$$\text{Entropy} = -\log P_{i,j} \sum_{i,j=0}^{N-1} P_{i,j}, \tag{7}$$

where i and j standards position of pixels of GLCM and $P_{i,j}$ is probability of presence of pixel pairs at certain distance and angle.

Contrast measures local variations and texture of shadow depth in GLCM. The larger the contrast, the deeper groove of texture and clearer effect of the image will be shown. Contrast can be calculated by

$$\text{Contrast} = \sum\sum (i-j)^2 P_{i,j}. \tag{8}$$

Homogeneity represents values by the inverse of the contrast weight, with weights decreasing exponentially away from the diagonal, which can be calculated using

$$\text{Homogeneity} = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2}. \tag{9}$$

Correlation coefficient concludes that the degree of two variable's activities is associated and can be calculated by

$$\text{CC} = \sum_{i,j=0}^{N-1} (i-\bar{i})(j-\bar{j}) \frac{P_{i,j}}{\delta_x \delta_y}. \tag{10}$$

Angular second moment (ASM) uses $P_{i,j}$ as weight for itself; high values of ASM occurs when the window is very orderly, so it measures the homogeneousness of a given image, and ASM can be calculated using

$$\text{ASM} = \sum_{i,j=0}^{N-1} P_{i,j}^2. \tag{11}$$

In addition to spectral and texture features, combining the geometric characteristics of high-resolution remote sensing images are extremely significant for detailed land cover information extraction. In this article, the article seven spatial geometric features include area, border index, compactness, density, length, length/width, and shape index were selected. Area of an image can be obtained through multiplying number of pixels constituting the image object and the covered area of the object. The boundary index can be calculated as the ratio between the boundary length of the image object and the smallest enclosing rectangle. The tighter the image object, the smaller its border. The density describes the distribution in the pixel space of the image object, that is, how tight the image object is. Density is based on the covariance matrix, which is calculated by dividing the number of pixels constituting the image object by its approximate radius. Length-width ratio can be used as one of the features for road

extraction. It is calculated by the length and the width of the objects. In addition, the length-width ratio can be used to calculate the length of the image object. Shape index was selected to describe the smoothness of the surface of an object. The smoother the surface of the image object, the lower its shape index. The more fragmented the image object, the larger its shape index. It can be calculated by dividing the frame length of an object by the volume of the object.

3.2. Classification Algorithms. In this research, an OBIA-RF method, also known as a combination of OBIA and classifier ensemble method which can take advantage of OBIA and classifier ensemble was constructed, and four state-of-art classification algorithms named KNN, SVM, and DTC were selected for performance comparison. Performance evaluation was carried out by quantitative indicators such as overall accuracy, kappa coefficient, and execution time consumption.

3.2.1. Random Forest. Random forest algorithm is an ensemble learning method proposed by Leo Breiman in 2001 [48], and is one of the most well-known ensemble learning methodology and has advantages of, i.e., performing out-of-sample prediction rapidly, requiring only slight parameter tuning, having capable ranking of the importance of features [28]. Decision trees in RF are generated by randomly selecting sample (bootstrap sampling) subsets in the training sample set and randomly selecting the feature variables to achieve optimal splitting. The obtained decision trees do not need pruning, and the final classification result is obtained by the majority vote method from the classification results of all decision trees in the integration. Gini index which measures the impurity of a given element with respect to the result of the classes is selected as a measure for the best split selection for RF [49]. There are two key parameters in the process of constructing a random forest pattern: the number of spanning trees and the number of randomly selected features. By literature review, the number of selected features is more important than the number of how many trees are trained; especially, generally, each split number of randomly selected features is set as the square root of the number of input characters [50–52]. This parameter can be optimized based on the out-of-bag error estimate. In this research, the number of trees is set as 100 and number of random attribute selection is $\log_2 d$, where d is the total number of features. And then, the number of trees for RF was tuning from 50 to 500 with a step of 10.

3.2.2. Support Vector Machine. SVM is one of the most appealing algorithms for remotely sensed data classification due to their advantages of generalization even with limited training samples which is common in remote sensing data processing [53]. And, as a supervised nonparametric statistical learning method, SVM does not need a training set strictly conforming to the standard independent and identical distribution. The advantages of SVM come from two aspects, transforming original space training set into a very high-dimensional new space and finding a large margin linear boundary in the new space. SVM is a classifier based on theory of structural risk minimization, which tries to lower the
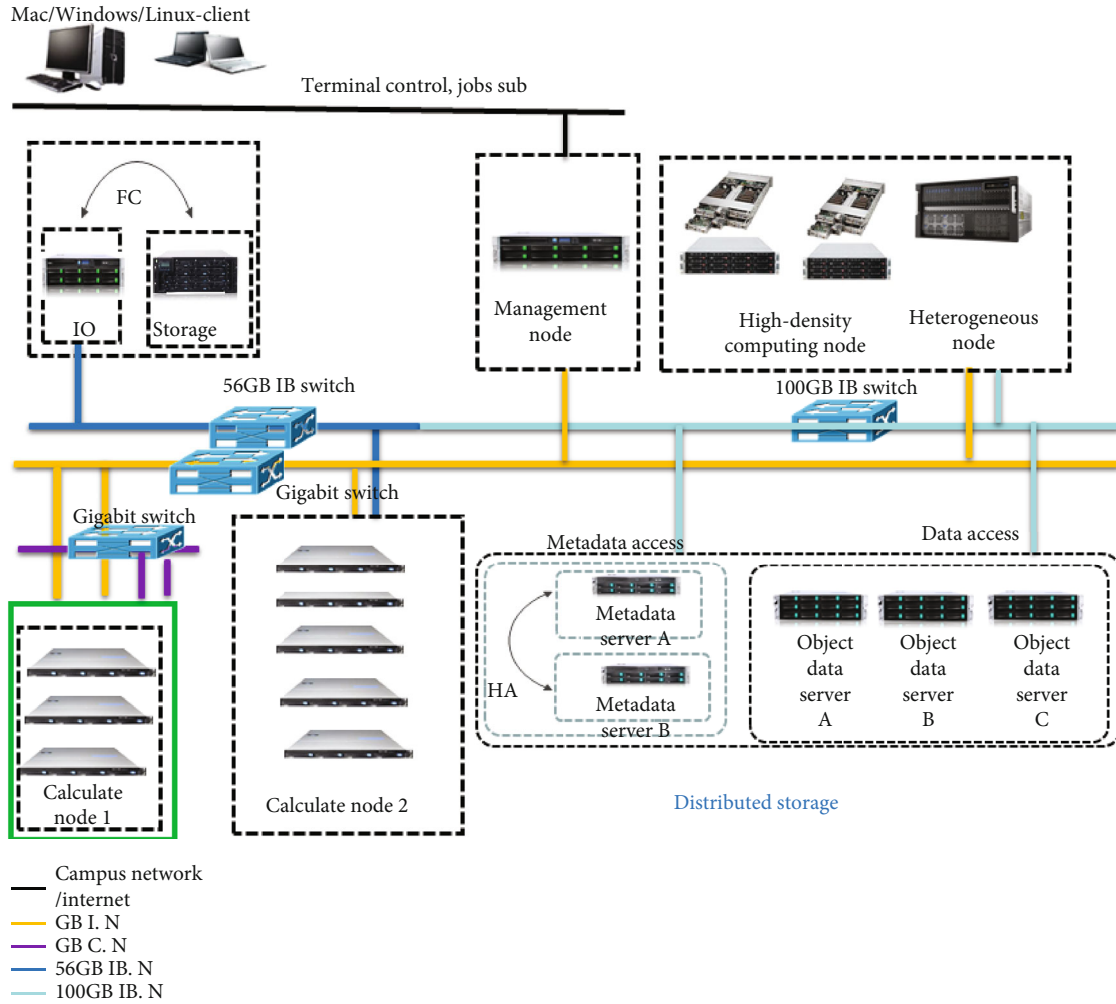
FIGURE 3: Cluster topology of high-performance computing.

generalization error by maximizing the margins on the training data. Thus, SVM looks for an ideal margin by solving optimization problem as

$$\min_{\omega,b} \frac{1}{2} ||\omega||^2,$$
$$s.t. y_i(\omega^T x_i + b) \geq 1, i = 1, 2, \cdots, m. \tag{12}$$

Furthermore, for classes that are nonseparable, the optimization can be solved by the so-called 'kernel stick.' The optimization procedure seeks to find coefficients ai and w0 in Equatuion (13), where $K(.,.)$ is kernel function. By default, the kernel function is set as the Gaussian kernel, kernel scale is 8, and box constraint is 1 standardized. And then, the scale of kernels for SVM was tuning from 0.1 to 8 with a step of 0.4

$$f(x) = \sum_{i=1}^{N} \alpha_i K(x, z_i) + \omega_0. \tag{13}$$

3.2.3. k-Nearest Neighbor. The *k*-nearest neighbor classifier (*k*-NN) is a kind of nonparametric and memory-based

learning, as well as instance-based learning or lazy learning used for classification and regression [26]. In the classification procedure, a given pixel will be classified by plurality vote of its neighbors in the feature space. The most intuitive *k*-NN classifier is 1-NN classifier; in this case, a given pixel will be assigned to the class of its closest neighbor in the feature space, which can be described as $C_n^{1NN}(x) = Y(1)$. The useful technique which can help nearer neighbors contribute more than the more distant one is assigning different weights to the neighbors. A common weighting scheme is setting each neighbor a weight of $1/d$, where d is the distance to the neighbor. Classification performance of *k*-NN can be significantly improved by metric learning, and diversity can be introduced to *k*-NN classifier by using different subsets of features, different distance metrics, different values of *k*, etc. In the experiment part, Euclidean distance is selected and number of neighbors is 100, distance weight is setting as equal to evaluate algorithm performance, and the number of neighbors for KNN was tuning from 1 to 300 with step of 10.

3.3. Comparisons and Assessment. Classification accuracy was evaluated by the confusion matrix as well as overall accuracy and the Kappa coefficient [54]. With the help of

TABLE 4: Labeled samples.

	UB (pixels)	UC (pixels)	UR (pixels)	UG (pixels)	UW (pixels)	HS (pixels)
GF-1	628	524	846	1085	427	1246
GF-2	30620	58059	34323	48874	39996	27412

TABLE 5: Overall accuracy and kappa coefficient of selected algorithms.

			DTC	SVM	KNN	RF
GF-1	P	OA (%)	72.4	87.5	87.4	94.67
		Kappa	0.68	0.85	0.85	0.93
		Time used (s)	0.032	0.57	0.05	27.13
	O	OA (%)	82.86	90.86	89.71	99.43
		Kappa	0.79	0.89	0.87	0.99
		Time used (s)	0.034	0.71	0.1	15.53
GF-2	P	OA (%)	66.67	85.5	89.21	91.26
		Kappa	0.63	0.83	0.87	0.89
		Time used (s)	0.36	589.31	1.89	251.32
	O	OA (%)	64.76	94.97	96.69	98.98
		Kappa	0.6	0.94	0.96	0.98
		Time used (s)	0.58	402.86	86.83	208.02

the confusion matrix, overall accuracy and kappa coefficient can be calculated using

$$\text{OA} = \frac{\sum_{i=1}^{q} n_{ii}}{n} \times 100\%, \tag{14}$$

$$\text{Kappa} = \frac{n\sum_{i=1}^{q} n_{ii} - \sum_{i=1}^{q} n_{i+} n_{+i}}{n^2 - \sum_{i=1}^{q} n_{i+} n_{+i}} \times 100\%, \tag{15}$$

where q is the number of classes, n represents the total number of considered pixel, n_{ii} are the diagonal elements of the confusion matrix, n_{i+} represents the marginal sum of the rows in the confusion matrix, and n_{+i} represents the marginal sum of the columns in the confusion matrix [55].

All experiments are performed on high-performance computing system using a portable bash system (PBS), as shown in Figure 3. Each algorithm is programmed as a job which can be submitted to cluster. Finally, time consumption of all algorithms is compared.

4. Results and Discussion

4.1. LULC Mapping. On the basis of segmentation results, spectral features and texture features are integrated and fused to generate multifeature images as inputs to all classifiers. In this research, the best segmentation scale is setting to 105 for GF-1 dataset and 210 for GF-2 dataset through estimation of scale parameter analysis. Land cover types in the study area include 6 categories named UB, UC, UR, UG, UW, and HS, more details can be found in Table 1. Training and testing samples are labeled by an expert of remote sensing with the

assistance of Google Earth. Labeled samples for training and testing are shown in Table 4. In order to test the sample sensitivity of the proposed process chain, limited and sufficient samples were selected from the GF-1 and GF-2 datasets, respectively. And 3-hold out validation method was selected for accuracy assessment.

During the research procedure, the same labeled training and testing samples are used as inputs of the DT, SVM, KNN, and RF classifiers. The default parameters of the constructed DT, RF, SVM, and KNN classifier models are selected for GF-1 and GF-2 image processing. Overall accuracy, kappa coefficient of all experiments, and the best classification results for GF-1 and GF-2 are shown in Table 5 and Figures 4 and 5.

It can be seen that the OBIA-RF algorithm has best classification results with an overall accuracy of 99.43% and 98.98% for GF-1 and GF-2, respectively (Table 5 and Figures 4 and 5). In general, the correct classification accuracy of all categories covered by urban land surface reached 91%. The overall accuracy of the original RF classification is 94.67% for GF-1 and 91.26% for GF-2, and the Kappa coefficient is 0.93 and 0.89, respectively. The overall accuracy of the SVM classification method is 87.5% and 85.5%, and Kappa coefficient is 0.85, 0.83 for the GF-1 and GF-2 datasets, respectively. The overall accuracy of the DTC classification is the lowest, only 72.4% and 66.67% for GF-1 and GF-2.

By analyzing the single-class accuracy of original RF and OBIA-RF classification results, the accuracy of corrected classification of UB, UR, and UC is relative lower than other classes, that is, 91.4%, 90.4%, and 92.2% for the GF-1 dataset, while the lowest single accuracy land use type of the GF-2 dataset is UR and UC for 85.4% and 86.7%, respectively. When the OBIA method is combined with RF accuracy, these difficult identified classes are improved to 91.4%, 90.4%, and 92.2% for GF-1 data and 99.2% and 99.5% for GF-2 data.

When pixel-based approach and object-based approach are compared, research results demonstrate that object-based approach outstands pixel-based approach, which is also demonstrated by other research [37, 38]. While our research further demonstrates that among the selected algorithms, OBIA is more suitable for the classifier ensemble method when compared with stand-of-art single classifier, especially for higher spatial resolution satellite data, GF-2 instead of GF-1.

4.2. General Discussion. Based on Table 5, the best accuracies are achieved by the OBIA-RF model for both GF-1 and GF-2 datasets. Obviously, these experiments showed the superiority of OBIA-RF over the selected state-of-art method in terms of classification accuracy. By statistics, the OBIA-RF method lets 4.62%, 11.89%, 11.79%, and 26.89% better accuracy than RF, KNN, SVM, and DTC for the GF-1 dataset and lets 7.72%, 9.77%, 13.48%, and 32.31% better accuracy than RF, KNN, SVM, and DTC for the GF-2 dataset.

The DTC achieved the worst overall accuracy. SVM model improves classification accuracy by 15.1% and 18.83% at pixel level and improves classification accuracy by 21.95% and 30.21% at object level for GF-1 and GF-2, respectively. The RF model led to further improvement in classification accuracy by 22.27% and 24.59% at pixel level

FIGURE 4: The classification map derived from GF-1 using the OBIA-RF method.

FIGURE 5: The classification map derived from GF-2 using the OBIA-RF method.

for GF-1 and GF-2. RF reduces the correlation between trees through random sampling of observations and features. This can demonstrate the advantage of the classifier ensemble for classification of VHR remotely sensed data. The OBIA method led to final improvement in classification accuracy by 4.62% and 7.72% in this study. This advantage is especially valuable for the relative high benchmark of random forest performance. When the OBIA method was combined with traditional machine learning model, especially classifier ensemble which takes advantage of textual, spatial structure information, and spectral information, classification results will be improved undoubtedly.

Furthermore, we investigate the sensitivity of the proposed model as well as the selected state-of-art machine learning model including DT, SVM, and KNN to parameter choice. Figures 6 and 7 plot the OA as a function of parameter for the corresponding machine learning model selected. The sensitivity of models to parameter choice for the GF-1 and GF-2 datasets (Figures 6 and 7) shows that (a) OA of the DTC model increased with the maximum number of splits, the peak value appears when number of split, GF-1 dataset equals to (290, 430) and GF-2 dataset equals to (480, 250), for pixel and OBIA training, respectively; (b) OA of SVM model increased to peak value when scale of kernels, GF-1 dataset equals to (0.5, 2.5) and GF-2 dataset equals to (0.1, 4.1), for pixel and OBIA and then significantly declines; (c) OA of KNN model appears when neighbors, GF-1 dataset equals to (291, 71) and GF-2 dataset equals to (221, 171), for pixel and OBIA training, respectively; (d) accuracy of RF fluctuates continuously with parameters changes, peak value appears when the number of trees reached, GF-1 dataset equals to (70, 240) and GF-2 dataset

equals to (230, 450), for pixel and OBIA training, respectively; and (e) the OBIA method performed better than pixel-based method for all selected model including the GF-1 dataset and GF-2 dataset.

Sensitivity test of models to parameter choice of GF-1 and GF-2 shows that rank of OA for selected model is RF, DTC, SVM, and KNN. Furthermore, we also wonder which features are most important in the procedure of prediction. With the help of out-of-bag estimation, feature importance in the RF ensemble learning are calculated and shown in Figure 8 for the GF-1 dataset and Figure 9 for the GF-2 dataset. This feature importance rank shows the contribution weight of different features for complex urban surface classification. Figure 8 demonstrates that the most important features for GF-1 remotely sensed data interpretation is texture information (mean value calculated using gray-level cooccurrence matrix(GLCM)) and the second important feature is standard deviation calculated using GLCM, and followed by the important spatial feature calculated by ratio, and then spatial feature of length/width. And, for the GF-2 dataset, the first five rank features for complex urban surface interpretation belongs to spatial feature (length, ratio) and spectral feature (NDVI, NDWI, mean2). By a comprehensive consideration of the processing results of the selected two datasets, a preliminary conclusion can be drawn as that spatial and texture features play an important role for complex urban surface classification with high- and very high-resolution remotely sensed data.

The quality of segmentation directly affects the effect of subsequent classification. In this research, the multiscale segmentation method was selected, which has three user-defined parameters: scale, shape, and compactness. And the
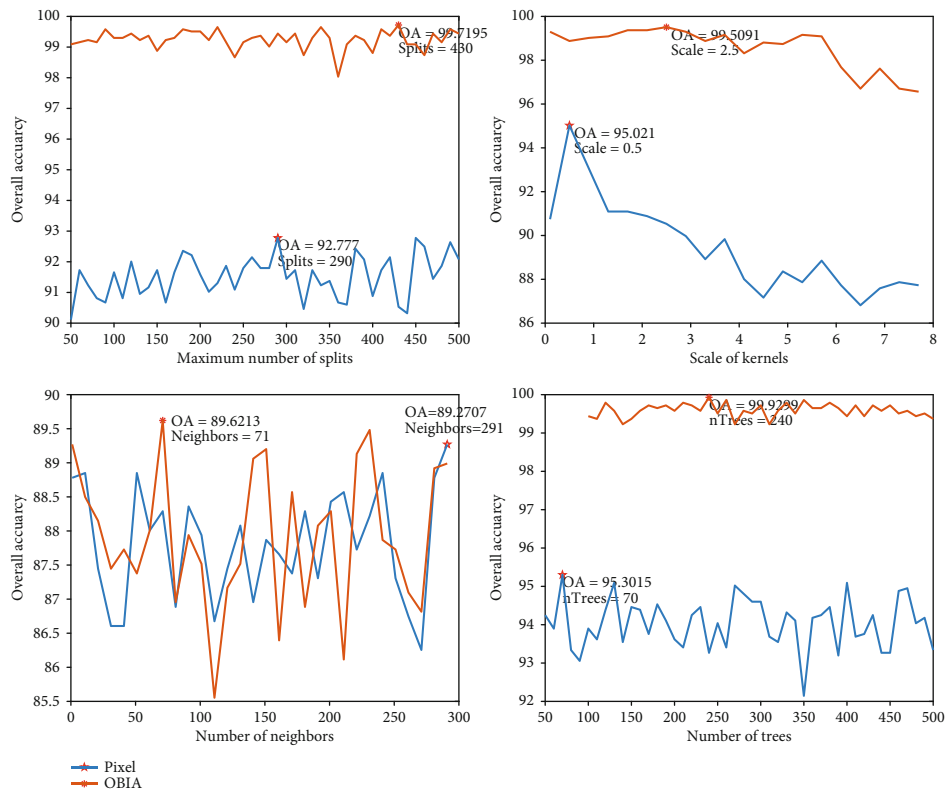
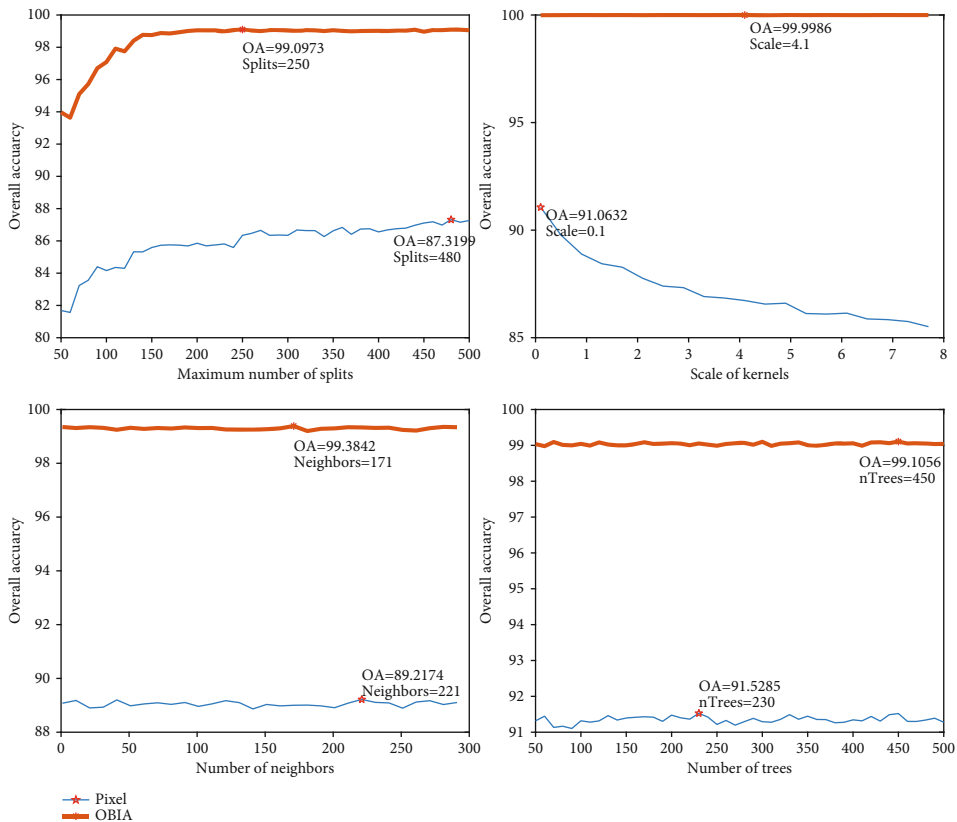FIGURE 6: Sensitivity of models to parameter choice for the GF-1 dataset.

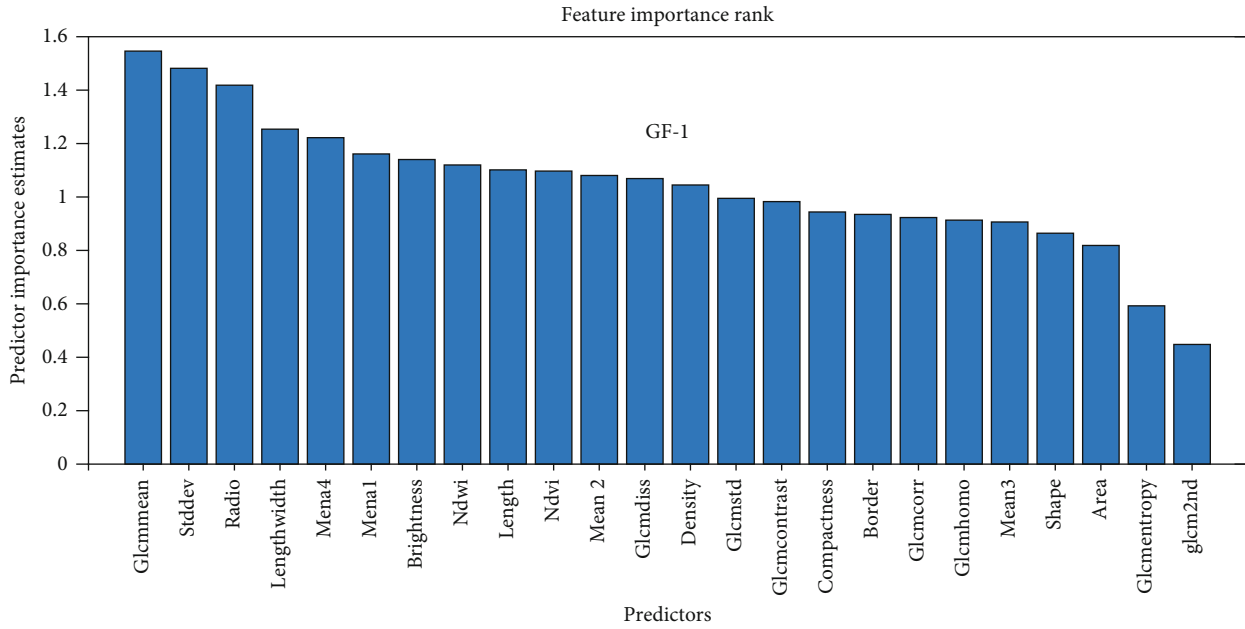FIGURE 7: Sensitivity of models to parameter choice for the GF-2 dataset.

Feature importance rank

FIGURE 8: Feature importance rank for the GF-1 dataset.
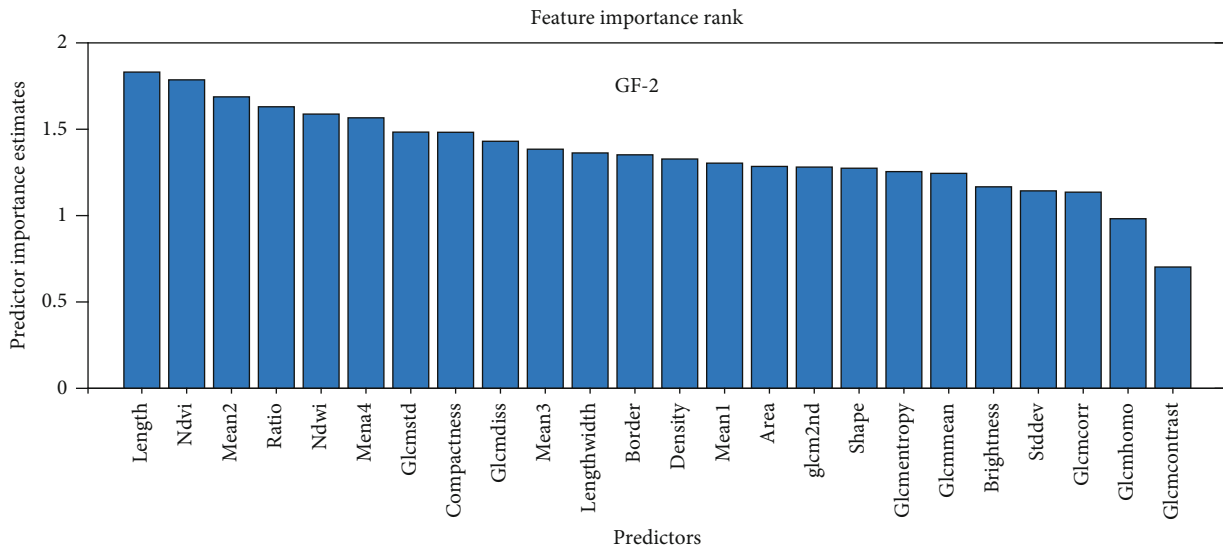
Feature importance rank

FIGURE 9: Feature importance rank for the GF-2 dataset.

scale parameter that defines the average size of the image object is considered to be the most effective parameter that affects the segmentation quality, while there is no universal rule for this scale determination. By literature review [46], the ESP2 (estimation of scale parameter) scale parameter estimation tool is introduced and combined with visual interpretation to evaluate the optimal scale value and segmentation effect of GF-1 and GF-2 remote sensing images in this research. Based on statistic results (Figure 10), the appearance of first peaks are 105 and 220 for GF-1 and GF-2, respectively, which are the optimal segmentation scales of GF-1 and GF-2 images in this study. The shape and compactness parameters have limited influence on the performance of OBIA, and they were setconstant at 0.1 and

0.5, respectively. Therefore, in this research, we chose 0.1 and 0.5 to participate in the segmentation to get the final segmentation map (Figure 10).

Finally, to more explicitly evaluate the practical speed of proposed image classification chain compared to RF, SVM, and KNN, we consider empirical run times. In terms of image performance speed test, to ensure fair comparison, all methods shared the same code. The algorithms were deployed in the Henan Polytech High Performance Computing Center server, shown in Figure 3 (here, node is 1, and thread number of per CPU core per node is 24). Through rough statistic, OBIA-RF spent 223.55 s, SVM spent 403.57 s, KNN spent 86.93 s, and DTC spent 0.64 s for the GF-1 and GF-2 datasets in average. The model with the most
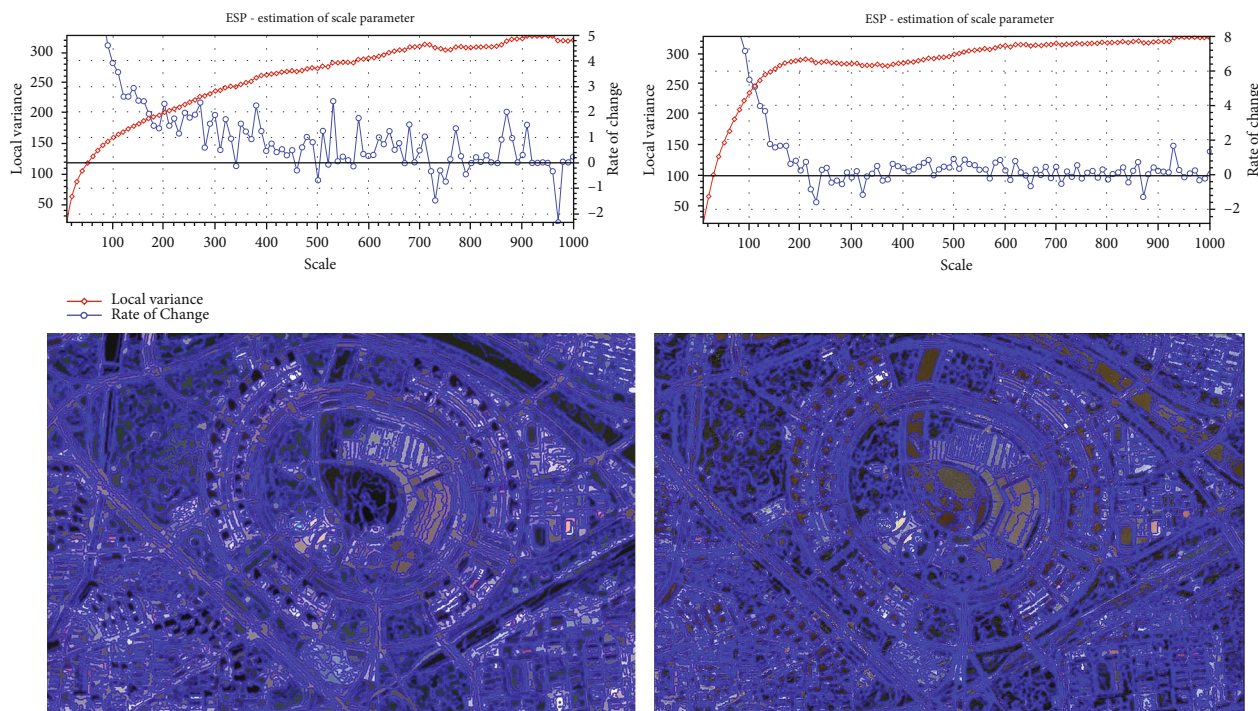
FIGURE 10: Segmentation results of GF-1 and GF-2 based on ESP.

time consumption is SVM, followed by KNN. The most time saving model is DTC, which got less than 70% classification accuracy.

5. Conclusion and Future Work

In this paper, we have proposed a novel urban mapping process chain which can take advantage of both the OBIA and classifier ensemble methods. The novelty in this paper is in the direction of successful evaluation of OBIA-RF on Chinese high-resolution satellite images GF-1 and GF-2 datasets. The performance of the OBIA-RF method has been compared with the state-of-art model such as DTC, SVM, and KNN, and performance of proposed OBIA-RF method has also been examined from urban mapping accuracy, the sensitivity to parametric selection, and time consumption. As the proposed process chain considered only spectral and GLCM texture features for image semantic, other features which might be useful to urban mapping such as local indicator of spatial association, mathematical morphology profiles, and decomposition characteristics of full polarized SAR features will be considered for future research.

Data Availability

The very high-resolution remotely sensed data used in this research is provided by Henan Data and Application Center of the High Resolution Earth Observation System through signing a contract with National Defense Science and Technology Bureau of Henan. Unfortunately, we do not have the priority to share the high-resolution satellite remote sensed data. Anyway, we can share our code used in this research. Researchers can test algorithms using these codes with their own datasets and repeat experiments to obtain similar research conclusions. Researchers who are interested in this code can download it from hyperlink https://pan.baidu.com/s/1hUyKWuoSzv2mrDx-NZ192w, using password ku0w, or contact with the corresponding author.

Conflicts of Interest

The authors declare that they have no conflict of interest.

Authors' Contributions

RM.H did the conceptualization. P.L. did the methodology. P.L. did the formal analysis. RM.H and HW. Z. wrote the manuscript. RM.H proposed the idea and wrote the original manuscript and the following revisions. P.L. provided the funding. RM.H, GY.W., and XL.W. executed all the experiments, XL.W, RM.H, and HW. Z contributed to the revisions and provided valuable comments. Ruimei Han, Pei Liu, Guangyan Wang, Hanwei Zhang, and Xilong Wu contributed equally to this work.

Acknowledgments

References

[1] M. Ustuner, F. B. Sanli, and B. Dixon, "Application of support vector machines for landuse classification using high-resolution RapidEye images: a sensitivity analysis," *EUROPEAN JOURNAL OF REMOTE SENSING*, vol. 48, no. 1, pp. 403–422, 2017.

[2] W. Zhao, S. du, and W. J. Emery, "Object-based convolutional neural network for high-resolution imagery classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 7, pp. 3386–3396, 2017.

[3] C. Vasilakos, D. Kavroudakis, and A. Georganta, "Machine learning classification ensemble of multitemporal sentinel-2 images: the case of a mixed Mediterranean ecosystem," *REMOTE SENSING*, vol. 12, no. 12, p. 2005, 2020.

[4] Y. Qian, W. Zhou, C. J. Nytch, L. Han, and Z. Li, "A new index to differentiate tree and grass based on high resolution image and object-based methods," *Urban Forestry and Urban Greening*, vol. 53, p. 126661, 2020.

[5] P. Lynch, L. Blesius, and E. Hines, "Classification of urban area using multispectral indices for urban planning," *Remote Sensing*, vol. 12, no. 15, p. 2503, 2020.

[6] Y. Xu, B. Du, L. Zhang et al., "Advanced multi-sensor optical remote sensing for urban land use and land cover classification: outcome of the 2018 IEEE GRSS data fusion contest," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 2019, no. 12, pp. 1709–1724, 2019.

[7] E. O. Yilmaz, B. Varol, R. H. Topaloglu, and E. Sertel, "Object-based classification of Izmir Metropolitan City by using Sentinel-2 images," in *2019 9th International Conference on Recent Advances in Space Technologies (RAST)*, pp. 407–412, Istanbul, Turkey, June 2019.

[8] D. Sulla-Menashe, J. M. Gray, S. P. Abercrombie, and M. A. Friedl, "Hierarchical mapping of annual global land cover 2001 to present: the MODIS collection 6 land cover product," *Remote Sensing of Environment*, vol. 222, pp. 183–194, 2019.

[9] L. H. Nguyen, D. R. Joshi, D. E. Clay, and G. M. Henebry, "Characterizing land cover land use from multiple years of Landsat and MODIS time series: a novel approach using land surface phenology modeling and random forest classifier," *Remote Sensing of Environment*, vol. 238, p. 111017, 2020.

[10] L. B. Zeferino, L. F. T. de Souza, C. H. do Amaral, E. I. F. Filho, and T. S. de Oliveira, "Does environmental data increase the accuracy of land use and land cover classification?," *International Journal of Applied Earth Observation and Geoinformation*, vol. 91, p. 102128, 2020.

[11] S. Q. Liu and N. Zhang, "Urban land use and land cover classification using multisource remote sensing images and social media data," *Remote Sensing*, vol. 11, no. 22, p. 2719, 2019.

[12] I. Colkesen and T. Kavzoglu, "Ensemble-based canonical correlation forest (CCF) for land use and land cover classification using sentinel-2 and Landsat OLI imagery," *Remote Sensing Letters*, vol. 8, no. 11, pp. 1082–1091, 2017.

[13] M. Turker and A. Ozdarici, "Field-based crop classification using SPOT4, SPOT5, IKONOS and QuickBird imagery for agricultural areas: a comparison study," *INTERNATIONAL JOURNAL OF REMOTE SENSING*, vol. 32, no. 24, pp. 9735–9768, 2011.

[14] J. A. Benediktsson, J. Chanussot, and W. M. Moon, "Advances in very-high-resolution remote sensing [Scanning the Issue]," *PROCEEDINGS OF THE IEEE*, vol. 101, no. 3, pp. 566–569, 2013.

[15] R. Momeni, P. Aplin, and D. Boyd, "Mapping complex urban land cover from spaceborne imagery: the influence of spatial resolution, spectral band set and classification approach," *REMOTE SENSING*, vol. 8, no. 2, p. 88, 2016.

[16] Q. Song, Q. Hu, Q. Zhou et al., "In-season crop mapping with GF-1/WFV data by combining object-based image analysis and random forest," *REMOTE SENSING*, vol. 9, no. 11, p. 1184, 2017.

[17] M. Amani, B. Salehi, S. Mahdavi, B. Brisco, and M. Shehata, "A multiple classifier system to improve mapping complex land covers: a case study of wetland classification using SAR data in Newfoundland, Canada," *INTERNATIONAL JOURNAL OF REMOTE SENSING*, vol. 39, no. 21, pp. 7370–7383, 2018.

[18] L. Zhou and X. Yang, "Training algorithm performance for image classification by neural networks," *PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING*, vol. 76, no. 8, pp. 945–951, 2010.

[19] G. Shukla, R. D. Garg, P. Kumar, H. S. Srivastava, and P. K. Garg, "Using multi-source data and decision tree classification in mapping vegetation diversity," *Spatial Information Research*, vol. 26, no. 5, pp. 573–585, 2018.

[20] F. Wang, Q. Wang, F. Nie, Z. Li, W. Yu, and F. Ren, "A linear multivariate binary decision tree classifier based on K-means splitting," *PATTERN RECOGNITION*, vol. 107, p. 107521, 2020.

[21] H. Shen, Y. Lin, Q. Tian, K. Xu, and J. Jiao, "A comparison of multiple classifier combinations using different voting-weights for remote sensing image classification," *INTERNATIONAL JOURNAL OF REMOTE SENSING*, vol. 39, no. 11, pp. 3705–3722, 2018.

[22] G. Lei, A. Li, J. Bian et al., "OIC-MCE: a practical land cover mapping approach for limited samples based on multiple classifier ensemble and iterative classification," *REMOTE SENSING*, vol. 12, no. 6, p. 987, 2020.

[23] G. J. Briem, J. A. Benediktsson, and J. R. Sveinsson, "Multiple classifiers applied to multisource remote sensing data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 10, pp. 2291–2299, 2002.

[24] P. Du, J. Xia, W. Zhang, K. Tan, Y. Liu, and S. Liu, "Multiple classifier system for remote sensing image classification: a review," *SENSORS*, vol. 12, no. 4, pp. 4764–4792, 2012.

[25] M. Xia, N. Tian, Y. Zhang, Y. Xu, and X. Zhang, "Dilated multi-scale cascade forest for satellite image classification," *INTERNATIONAL JOURNAL OF REMOTE SENSING*, vol. 41, no. 20, pp. 7779–7800, 2020.

[26] P. Thanh Noi and M. Kappas, "Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery," *SENSORS*, vol. 18, no. 2, p. 18, 2018.

[27] C. Pelletier, S. Valero, J. Inglada, N. Champion, and G. Dedieu, "Assessing the robustness of random forests to map land cover with high resolution satellite image time series over large areas," *REMOTE SENSING OF ENVIRONMENT*, vol. 187, pp. 156–168, 2016.

[28] M. Belgiu and L. Dragut, "Random forest in remote sensing: a review of applications and future directions," *ISPRS JOURNAL OF PHOTOGRAMMETRY AND REMOTE SENSING*, vol. 114, pp. 24–31, 2016.

[29] A. Zafari, R. Zurita-Milla, and E. Izquierdo-Verdiguier, "Evaluating the performance of a random forest kernel for land

cover classification," *REMOTE SENSING*, vol. 11, no. 5, p. 575, 2019.

[30] D. Phiri, M. Simwanda, V. Nyirenda, Y. Murayama, and M. Ranagalage, "Decision tree algorithms for developing rule-sets for object-based land cover classification," *ISPRS INTERNATIONAL JOURNAL OF GEO-INFORMATION*, vol. 9, no. 5, p. 329, 2020.

[31] A. Zafari, R. Zurita-Milla, and E. Izquierdo-Verdiguier, "A multiscale random forest kernel for land cover classification," *IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING*, vol. 13, pp. 2842–2852, 2020.

[32] M. D. Hossain and D. Chen, "Segmentation for object-based image analysis (OBIA): a review of algorithms and challenges from remote sensing perspective," *ISPRS JOURNAL OF PHOTOGRAMMETRY AND REMOTE SENSING*, vol. 150, pp. 115–134, 2019.

[33] M. Kucharczyk, G. J. Hay, S. Ghaffarian, and C. H. Hugenholtz, "Geographic object-based image analysis: a primer and future directions," *REMOTE SENSING*, vol. 12, no. 12, p. 2012, 2020.

[34] C. M. D. de Pinho, L. M. G. Fonseca, T. S. Korting, C. M. de Almeida, and H. J. H. Kux, "Land-cover classification of an intra-urban environment using high-resolution images and object-based image analysis," *International Journal of Remote Sensing*, vol. 33, no. 19, pp. 5973–5995, 2012.

[35] T. Su, T. Liu, S. Zhang, Z. Qu, and R. Li, "Machine learning-assisted region merging for remote sensing image segmentation," *ISPRS JOURNAL OF PHOTOGRAMMETRY AND REMOTE SENSING*, vol. 168, pp. 89–123, 2020.

[36] V. S. Martins, A. L. Kaleita, B. K. Gelder, H. L. F. da Silveira, and C. A. Abe, "Exploring multiscale object-based convolutional neural network (multi-OCNN) for remote sensing image classification at high spatial resolution," *ISPRS JOURNAL OF PHOTOGRAMMETRY AND REMOTE SENSING*, vol. 168, pp. 56–73, 2020.

[37] S. W. Myint, P. Gober, A. Brazel, S. Grossman-Clarke, and Q. Weng, "Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery," *Remote Sensing of Environment*, vol. 115, no. 5, pp. 1145–1161, 2011.

[38] D. C. Duro, S. E. Franklin, and M. G. Dube, "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 SENSING OF ENVIRONMENT*, vol. 118, pp. 259–272, 2012.

[39] B. Fu, Y. Wang, A. Campbell et al., "Comparison of object-based and pixel-based Random Forest algorithm for wetland vegetation mapping using high spatial resolution GF-1 and SAR data," *Ecological Indicators*, vol. 73, pp. 105–117, 2017.

[40] T. Blaschke, "Object based image analysis for remote sensing," *ISPRS JOURNAL OF PHOTOGRAMMETRY AND REMOTE SENSING*, vol. 65, no. 1, pp. 2–16, 2010.

[41] H. Grybas, L. Melendy, and R. G. Congalton, "A comparison of unsupervised segmentation parameter optimization approaches using moderate- and high-resolution imagery," *GISCIENCE & REMOTE SENSING*, vol. 54, no. 4, pp. 515–533, 2017.

[42] Y. Shen, J. Chen, L. Xiao, and D. Pan, "Optimizing multiscale segmentation with local spectral heterogeneity measure for high resolution remote sensing images," *ISPRS JOURNAL OF PHOTOGRAMMETRY AND REMOTE SENSING*, vol. 157, pp. 13–25, 2019.

[43] S. Georganos, M. Lennert, T. Grippa, S. Vanhuysse, B. Johnson, and E. Wolff, "Normalization in unsupervised segmentation parameter optimization: a solution based on local regression trend analysis," *REMOTE SENSING*, vol. 10, no. 2, p. 222, 2018.

[44] L. Dragut, D. Tiede, and S. R. Levick, "ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data," *INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE*, vol. 24, no. 6, pp. 859–871, 2010.

[45] J. Liu, M. Du, and Z. Mao, "Scale computation on high spatial resolution remotely sensed imagery multi-scale segmentation," *INTERNATIONAL JOURNAL OF REMOTE SENSING*, vol. 38, no. 18, pp. 5186–5214, 2017.

[46] J. Liu, H. Pu, S. Song, and M. Du, "An adaptive scale estimating method of multiscale image segmentation based on vector edge and spectral statistics information," *INTERNATIONAL JOURNAL OF REMOTE SENSING*, vol. 39, no. 20, pp. 6826–6845, 2018.

[47] L. Dragut, O. Csillik, C. Eisank, and D. Tiede, "Automated parameterisation for multi-scale image segmentation on multiple layers," *ISPRS JOURNAL OF PHOTOGRAMMETRY AND REMOTE SENSING*, vol. 88, no. 100, pp. 119–127, 2014.

[48] L. Breiman, "*Random forests*," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[49] C. Su, S. Ju, Y. Liu, and Z. Yu, "Improving random forest and rotation forest for highly imbalanced datasets," *INTELLIGENT DATA ANALYSIS*, vol. 19, no. 6, pp. 1409–1432, 2015.

[50] H. Deng and G. Runger, "Gene selection with guided regularized random forest," *PATTERN RECOGNITION*, vol. 46, no. 12, pp. 3483–3489, 2013.

[51] J. L. Speiser, M. E. Miller, J. Tooze, and E. Ip, "A comparison of random forest variable selection methods for classification prediction modeling," *EXPERT SYSTEMS WITH APPLICATIONS*, vol. 134, pp. 93–101, 2019.

[52] P. Lou, B. Fu, H. He et al., "An optimized object-based random forest algorithm for marsh vegetation mapping using high-spatial-resolution GF-1 and ZY-3 data," *REMOTE SENSING*, vol. 12, no. 8, p. 1270, 2020.

[53] G. Mountrakis, J. Im, and C. Ogole, "Support vector machines in remote sensing: a review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 3, pp. 247–259, 2011.

[54] R. G. Congalton, "A review of assessing the accuracy of classifications of remotely sensed data," *Remote Sensing of Environment*, vol. 37, no. 1, pp. 35–46, 1991.

[55] G. M. Foody, "Status of land cover classification accuracy assessment," *Remote Sensing of Environment*, vol. 80, no. 1, pp. 185–201, 2002.