Detection of Urban Built-Up Area Change From Sentinel-2 Images Using Multiband Temporal Texture and One-Class Random Forest

Xiaoxue Feng, Peijun Li, *Senior Member, IEEE*, and Tao Cheng, *Senior Member, IEEE*

***Abstract*—Detection of urban land expansion is important for understanding the urbanization process and improving urban planning. Spatio-temporal contextual information derived from multitemporal high-resolution imagery is useful for highlighting urban land cover changes. This article proposes a new method for detecting urban built-up area change from multitemporal high spatial resolution imagery by combining spectral and spatio-temporal features. A multiband temporal texture measured using pseudo cross multivariate variogram (PCMV) is adopted to quantify the local spatio-temporal dependence between bitemporal multispectral images. The PCMV textures at multiple scales, bitemporal spectral features, and normalized difference vegetation indices are together input to an improved one-class random forest classifier for urban built-up area change mapping. The proposed method is evaluated in urban built-up area change detection using multitemporal Sentinel-2 images of Tianjin area acquired from 2015 to 2019. It is also compared with three feature combinations and an existing postclassification comparison method based on one-class support vector machine. Experimental results demonstrate that the proposed method outperformed the traditional ones, with increases of 2.15%–7.38%, 2.07%–5.45%, 1.93%–6.76%, and 5.98%–13.11% in overall accuracy. Moreover, the proposed method also achieves the best performance using the bitemporal Sentinel-2 images over the east of Beijing area. The proposed method is promising as a simple and reliable way to detect urban built-up area change with multitemporal Sentinel-2 imagery.**

***Index Terms*—Built-up area change, improved one-class random forest (iOCRF), multiband temporal texture, multitemporal data, pseudo cross multivariate variogram (PCMV).**

I. Introduction

URBANIZATION process, including population growth and urban land expansion, has been given much focus in recent years [1]. To meet the needs of human life and production, urban areas are now expanding at twice their rates of population growth [2]. Urban land expansion is closely related to urban morphology, land utilization efficiency, economic development, and ecological environment [3]–[5]. The impact of urban land expansion on biochemical, hydrological, and climate changes at different scales has been a widespread concern [6], [7]. Therefore, timely and accurate information about urban land expansion is widely required by the scientific community, government sectors, and the public.

The increasing availability of remote sensing data of different resolutions has made it possible to detect urban land expansion on different scales [8], [9]. Low- and medium-resolution satellite images, such as MODIS and Landsat images, have been exploited in detecting and monitoring urban land changes across large areas during the last three decades [1], [10]–[13]. However, the limitation in spatial resolution of these images makes these data only suitable for detecting urban land expansion in large scales, and are not able to capture urban land changes at fine scales [14]–[16]. Thus, higher spatial resolution data, such as SPOT-5 and Sentinel-2 images with 10–20 m resolutions, are used to monitor urban land expansion at detailed levels [16]–[19]. In particular, Sentinel-2 satellites launched by the European Space Agency collect multispectral images with high spectral (13 bands) and spatial resolutions (10, 20, and 60 m) [15]. Two satellites (Sentinel-2A and 2B) are in orbit with a time revisit of five days [20]. The Sentinel-2 is found to provide the improved data to map and monitor human settlement [20], urban land cover [21], and imperviousness [22], with respect to Landsat derived product. The increased spatial details, temporal resolution, and thematic contents offered by Sentinel-2 data have the potential for detecting urban land expansion at a fine scale.

To detect urban land expansion at fine scales using high spatial resolution data, it is important to include spatial and temporal information, in addition to multitemporal spectral data [17], [23]. Guo *et al.* [24] proposed a joint framework for change detection by combining convolutional sparse analysis to achieve spatial utilization and temporal spectral unmixing. The addition of spatio-temporal information from multitemporal data is conducive to improve the accuracy of urban land change detection. For example, some studies showed that spatio-temporal contextual information, which represents temporal dependence of spectral signature in a local neighborhood, obtained from multitemporal data provides additional discriminative capacity

Manuscript received April 10, 2021; revised June 6, 2021; accepted June 20, 2021. Date of publication June 24, 2021; date of current version July 22, 2021. This work was supported by the National Science Foundation of China under Grant 42071307. (*Corresponding author: Peijun Li.*)

Xiaoxue Feng and Peijun Li are with the Institute of Remote Sensing and GIS, School of Earth and Space Sciences, Peking University, Beijing 100871, China, and also with the Beijing Key Lab of Spatial Information Integration and Its Applications, Peking University, Beijing 100871, China (e-mail: fengxx@pku.edu.cn; pjli@pku.edu.cn).

Tao Cheng is with the National Engineering and Technology Center for Information Agriculture, MARA Key Laboratory of Crop System Analysis and Decision Making, Jiangsu Key Laboratory for Information Agriculture, Nanjing Agricultural University, Nanjing 210095, China (e-mail: tcheng@njau.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2021.3092064

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

for detecting land cover changes [25]–[29]. Papadomanolaki *et al.* [17] presented a deep learning framework for urban land expansion from a multitemporal Sentinel-2 dataset by integrating spatial features learned using a U-Net architecture and temporal change pattern learned using recurrent networks. Im and Jensen [25] developed a change detection method based on temporal contextual information between two-date images using correlation, slope and intercept in a small neighborhood and decision tree classification. Jin *et al.* [27], [28] used multitemporal texture to quantify local temporal correlation between bitemporal images of the same bands. Moya *et al.* [29] presented a three-dimensional gray-level co-occurrence matrix (GLCM) texture from a pair of images in the identification of collapsed buildings. Gueguen *et al.* [30] used the local mixed information based change detector to analyze the temporal co-occurrences and adopted a region-based approach for consistent urbanization detection of large scenes.

In the literature, most existing methods quantify spatio-temporal dependence between bitemporal images of the same spectral band [27], [28]. In this case, an appropriate spectral band needs to be selected for generating the texture and structural images. As a result, band selection may have an impact on detection results. On the other hand, the correlation across different bands is discarded [31]. However, the spatio-temporal contextual information derived from multiple bands at two dates (e.g., bitemporal multispectral images) may be more useful to highlight different land cover changes. A spatio-temporal feature named multiband temporal texture measured by pseudo cross multivariate variogram (PCMV) was proposed in our previous study [32], [33]. The multiband temporal texture was preliminarily evaluated in the detection of forest burned area using bitemporal Landsat TM images [32]. Besides, multiband temporal texture derived from bitemporal high spatial resolution data, e.g., SPOT-5 and Sentinel-2 images, was evaluated in the detection of urban land change [33].

To take advantage of the spectral and spatio-temporal contextual information from high spatial resolution images for urban land change detection, an appropriate change detection method is required [34]. There are two problems to be considered. First, the combination of spatio-temporal features and spectral features may produce a high-dimensional dataset and affect the classification performance, i.e., "dimensionality disaster" [35]. In this case, apart from using additional feature reduction methods to reduce the data dimensionality [29], land cover classification or extraction methods insensitive to highly dimensional data are needed. Second, since built-up area change is the only target class of interest in this study, a one-class classification method that focuses on a specific class is needed, such as one-class support vector machine (OCSVM) [36] and one-class random forest (OCRF) [37]. The OCRF uses the probability distribution of target class samples to simulate an artificial outlier (i.e., samples of nontarget class), and then transforms the one-class classification task into a two-class classification [37], [38]. However, the OCRF may misestimate the probability distribution of nontarget class samples for high-dimensional data [39]. Therefore, a one-class classification method, which is insensitive to data dimensionality, would be appropriate and efficient in urban land change detection.

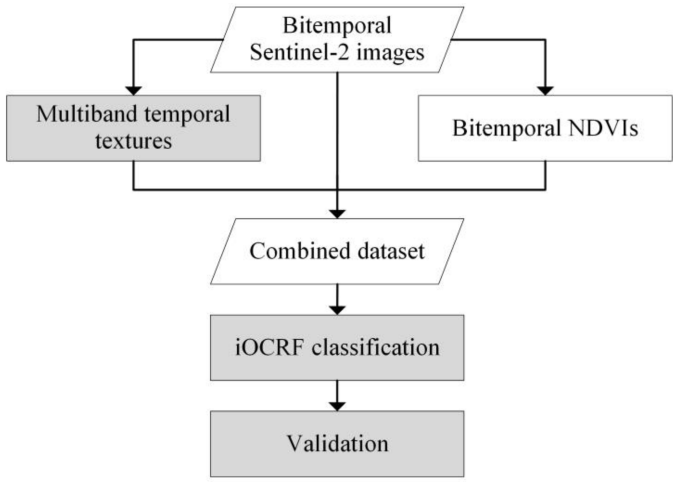

Fig. 1. Flowchart of the proposed urban built-up area change detection method. The modules in the gray boxes are described in detail in the full text.

To address the aforementioned problems, we proposed a method for the detection of urban built-up area change using the multiband temporal texture measured by PCMV [32] and one-class classification [39]. The ability of multiband temporal texture in quantifying spatio-temporal information and the effectiveness of one-class classification in the detection of urban built-up change were evaluated.

II. Methods

In this study, a method of detecting urban built-up area change from multitemporal high spatial resolution imagery by including the PCMV multiband temporal texture was proposed. First, the PCMV texture [40] was extracted from bitemporal Sentinel-2 imagery. Then, the PCMV textures obtained were combined with bitemporal spectral features and normalized difference vegetation indices (NDVIs) in the detection of urban built-up area change. The built-up area change was detected by classifying the combined dataset (i.e., bitemporal spectral features, bitemporal NDVIs, and multiband temporal textures) using a one-class classification method. Specifically, an improved one-class random forest (iOCRF), a recently proposed one-class classifier, [39] was used. The flowchart of the proposed method is shown in Fig. 1 and is described in detail in the following sections.

A. Multiband Temporal Texture by PCMV

Remote sensing images acquired at different dates in the same geographical area usually show a certain spatio-temporal dependence (or correlation) [26]–[29]. An appropriate method of quantifying spatio-temporal correlation between bitemporal images is thus required. In the existing studies, pseudo cross-variogram (PCV) $\gamma_{12}^{\mathrm{PCV}}(\boldsymbol{h})$ was originally proposed to quantify the spatial cross dependence between two spatial variables [41], [42], which is defined as

$$\gamma_{12}^{\mathrm{PCV}}(\boldsymbol{h}) = \frac{1}{2}E[(z_1(\boldsymbol{x}) - z_2(\boldsymbol{x}+\boldsymbol{h})]^2 \tag{1}$$

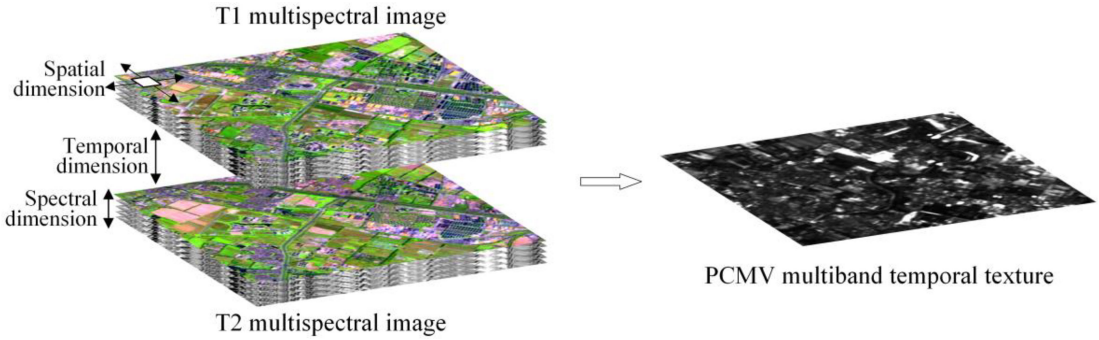

Fig. 2. PCMV multiband temporal texture calculated from multispectral images acquired on two different dates, using a moving local window w with a lag distance $\boldsymbol{h}$.

where $\boldsymbol{h}$ is a distance vector, and z_1 and z_2 denote two spatial variables. For remote sensing applications, these two spatial variables represent the digital numbers or spectral reflectance values from two spectral bands of an image. The PCV quantifies spatial cross dependence between two different bands. However, if the same spectral band acquired at two different dates is used in PCV, the PCV expresses temporal dependence of the same spectral band [27].

It is worth noting that variation across different bands is discarded when using a single spectral band of bitemporal images in PCV. Therefore, the PCV is extended to PCMV [32], which represents the semivariance of the cross-increments between two-date multispectral images as

$$\gamma_{12}^{\mathrm{PCMV}}(\boldsymbol{h}) = \frac{1}{2}E[(\boldsymbol{z}_1(\boldsymbol{x}) - \boldsymbol{z}_2(\boldsymbol{x}+\boldsymbol{h}))\boldsymbol{M}(\boldsymbol{z}_1(\boldsymbol{x}) - \boldsymbol{z}_2(\boldsymbol{x}+\boldsymbol{h}))^T] \quad (2)$$

where $\gamma_{12}^{\mathrm{PCMV}}(\boldsymbol{h})$ represents the PCMV at a lag distance vector $\boldsymbol{h}$. $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ are pixel vectors from two multispectral images acquired on dates t_1 and t_2, respectively. $(.)^T$ is the transpose of the matrix. $\boldsymbol{M}$ is a symmetric positive definite matrix used as metrics in the calculation of dissimilarities. Such metrics can be: the identity matrix, the inverse of the variance-covariance matrix [40].

The experimental PCMV is estimated by averaging the multivariate distance squared

$$\gamma_{12\_\mathrm{exp}}^{\mathrm{PCMV}}(\boldsymbol{h}) = \frac{1}{2N(\boldsymbol{h})}\sum_{i=1}^{N(\boldsymbol{h})}[(\boldsymbol{z}_1(\boldsymbol{x}_i) - \boldsymbol{z}_2(\boldsymbol{x}_i+\boldsymbol{h}))\boldsymbol{M}(\boldsymbol{z}_1(\boldsymbol{x}_i) - \boldsymbol{z}_2(\boldsymbol{x}_i+\boldsymbol{h}))^T] \quad (3)$$

where $N(\boldsymbol{h})$ is the number of pixel pairs with a certain lag distance vector $\boldsymbol{h}$, and x_i is the location of pixel.

As mentioned previously, different metrics of $\boldsymbol{M}$ could be used in the calculation of PCMV in (3). For simplicity, $\boldsymbol{M} = \boldsymbol{I}$ ($\boldsymbol{I}$ is the identity matrix) was used in the existing study [32], [33]. The experimental PCMV using $\boldsymbol{M} = \boldsymbol{I}$ is the sum of the experimental PCVs for each pair of bands on two different dates. However, different spectral bands in the multispectral images have different importance values or contributions to PCMV. It would be more reasonable to assign different weights to different bands. For example, the variables with high variability receive less weight than variables with low variability [43]. Therefore, in this study, $\boldsymbol{M} = \boldsymbol{C}^{-1}$ (the inverse of the variance-covariance matrix $\boldsymbol{C}$) is used in the calculation of the PCMV [44]. For two multispectral images in this study, $\boldsymbol{C}$ is estimated by the pooled covariance matrix [45]. The pooled covariance matrix is the weighted average, such as mean value, of the covariance matrices of two multispectral images.

The multiband temporal texture measured using PCMV is calculated in a local neighborhood (see Fig. 2). Specifically, a PCMV texture is calculated in a local moving window for a specific lag distance $\boldsymbol{h}$ (including size and direction).

The value of PCMV obtained within a local window for a specific lag distance $\boldsymbol{h}$ is assigned to the central pixel of the window. Thus, window size w and lag distance $\boldsymbol{h}$ (including size and direction) are two key parameters for the calculation of PCMV texture. The appropriate window size for texture calculation is determined by trial and error to maximize the accuracy of classification [46] or by the semivariograms method [47]. The lag distance $\boldsymbol{h}$ can be measured with one or more sizes in different directions. The average of texture values from all directions of $\boldsymbol{h}$, such as four directions (NS, EW, NE-SW, NW-SE), is commonly used as the omnidirectional texture. However, the textures from an individual direction or the omnidirectional texture usually show a significant edge effect in texture measures [48]. In contrast, the minimum of texture values from different directions of the lag distance vector is less sensitive to the edge effect [49]. Thus, the minimum of PCMV values from these four directions of $\boldsymbol{h}$ is used as the multiband temporal texture value in this study.

By analyzing (3), the multiband temporal texture value measured using PCMV quantifies spatio-temporal correlation between bitemporal multispectral images. Specifically, if the spectral signature of a local area appears to remain the same over time, there is a strong temporal correlation between the bitemporal multispectral images for the area. Thus, a low PCMV value for this area will be obtained using (3). In contrast, if the spectral signature of a local area changes significantly over time, because of either change in the land cover type or phenological variation of vegetation, there is a weak temporal correlation between multispectral images in the local area. In this case, a high PCMV value for this area will be obtained using (3).

The PCMV texture uses the contextual variations in pixel values within a moving window to reflect local spatial information like most existing texture measures, such as GLCM. The GLCM texture could only be derived from a single image at one time and two GLCM texture bands are needed for characterizing the changes between bitemporal images. However, the PCMV texture quantifies the local spatio-temporal variation between bitemporal multispectral images at one time and highlights the change intensity directly.

B. Change Detection With iOCRF

In this study, the PCMV multiband temporal texture is used in the detection of urban built-up area change. The obtained PCMV textures with different window sizes and lag distances are combined with bitemporal spectral features and NDVIs in the detection of built-up area change. Different from the existing studies that only use a single-scale texture, the PCMV textures at multiple scales, i.e., multiple window sizes and lag distances are adopted, considering that multiscale textures may provide more information to discriminate urban land change from other change and nonchange types. In addition, bitemporal NDVIs are used to reduce the impact of vegetation changes on the detection of built-up area change. Since all changes in the study area are highlighted by the multiband temporal texture, in particular, it is found that the built-up area change and vegetation change all have high texture values. The inclusion of bitemporal NDVIs will reduce the confusion between built-up area change and vegetation change, thus improving the detection accuracy. However, the combination of multiple-scale PCMV textures, spectral features, and NDVIs may result in the high dimensionality of data.

Since the urban built-up area change is the only target class of interest in this study, a one-class classification method is selected, as in the existing studies [50]. Only samples of the target class (i.e., urban built-up area change) are used in training of one-class classification, instead of using samples of all classes (both urban built-up change and other classes) in the conventional multiclass classification. In addition, considering that the high dimensionality of data combining different features may affect the classification performance, and the spectral and texture features generally have different data ranges, iOCRF is adopted to classify the combined dataset (i.e., bitemporal spectral features, NDVIs, and multiband temporal textures) in the detection of urban built-up area change. This is because the iOCRF is insensitive to high-dimensionality data and normalization in image stacking is not required.

The iOCRF is a modified version of OCRF [37] and can be considered as a special case of random forest (RF), where the labeled samples of target class and the samples of nontarget class estimated from unlabeled samples are used in RF classification [39]. A key component of iOCRF is to properly estimate training samples of nontarget class from unlabeled samples. Given that the RF classification has strong robustness for training samples containing some false class labels, a small number of mislabeled samples have a limited impact on classification result [51]. Thus, the samples randomly selected from unlabeled samples, which include some samples of target class, are initially considered as samples of nontarget class in iOCRF. These randomly selected samples of nontarget class and the labeled samples of target class are used in the initial classification using RF. After the initial classification is obtained, those highly reliable pixels of nontarget class (e.g., with posterior probabilities higher than 90%) are selected as new training samples of nontarget class. These newly selected samples of nontarget class and the training samples of target class manually selected are then used in iOCRF for final classification.

In addition, to analyze the contributions of different features (i.e., spectral feature, NDVI, and multiband temporal texture) to the classification of built-up area change, feature importance of different features in the iOCRF classification are computed by out of bag (OOB) accuracy using RF-based feature importance measure [52]. The OOB data are some observations that are not included in the bootstrap sample for training a tree in RF. Their predicted values from the tree of each feature are permuted in the OOB data. Then, the feature importance was calculated by the difference of OOB errors before and after the permutation and obtained by averaging over all trees [52]. The larger the difference, the more important the feature.

C. Evaluation

The proposed method was evaluated using accuracy measures produced from confusion matrix to assess the change detection results. The producer's accuracy (PA), user's accuracy (UA), the overall accuracy (OA), and the $F1$ score were used for accuracy assessment [53]. The $F1$ score is usually used as an accuracy measure of a dichotomous model [54] and is suitable for one-class classification method. The $F1$ score is the harmonic mean of precision (i.e., UA) and recall (i.e., PA), and its value increases with greater precision and recall, which is derived as follows:

$$F1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}. \tag{4}$$

To fully validate the effectiveness of the proposed method, three different feature combinations were used for comparison: 1) bitemporal spectral features; 2) bitemporal spectral features and bitemporal NDVIs; and 3) bitemporal spectral features, NDVIs, and single-band temporal texture features measured by PCV [27]. For the calculation of single-band temporal texture, an appropriate band should be selected. In this study, principal component analysis was conducted on the multispectral image of each date and the first principal component was used in the calculation of single-band temporal texture. For a comprehensive comparison, the OCSVM was used for urban built-up area change detection by the postclassification comparison method [55]. The built-up areas of each year were extracted using the OCSVM classifier, and then the change from non-built-up areas to built-up areas was detected from bitemporal images.

III. Study Area and Data

A. Study Area

Two study areas were selected as shown in Fig. 3(a). One area is in Tianjin, the third largest megacity in China. Tianjin has 16

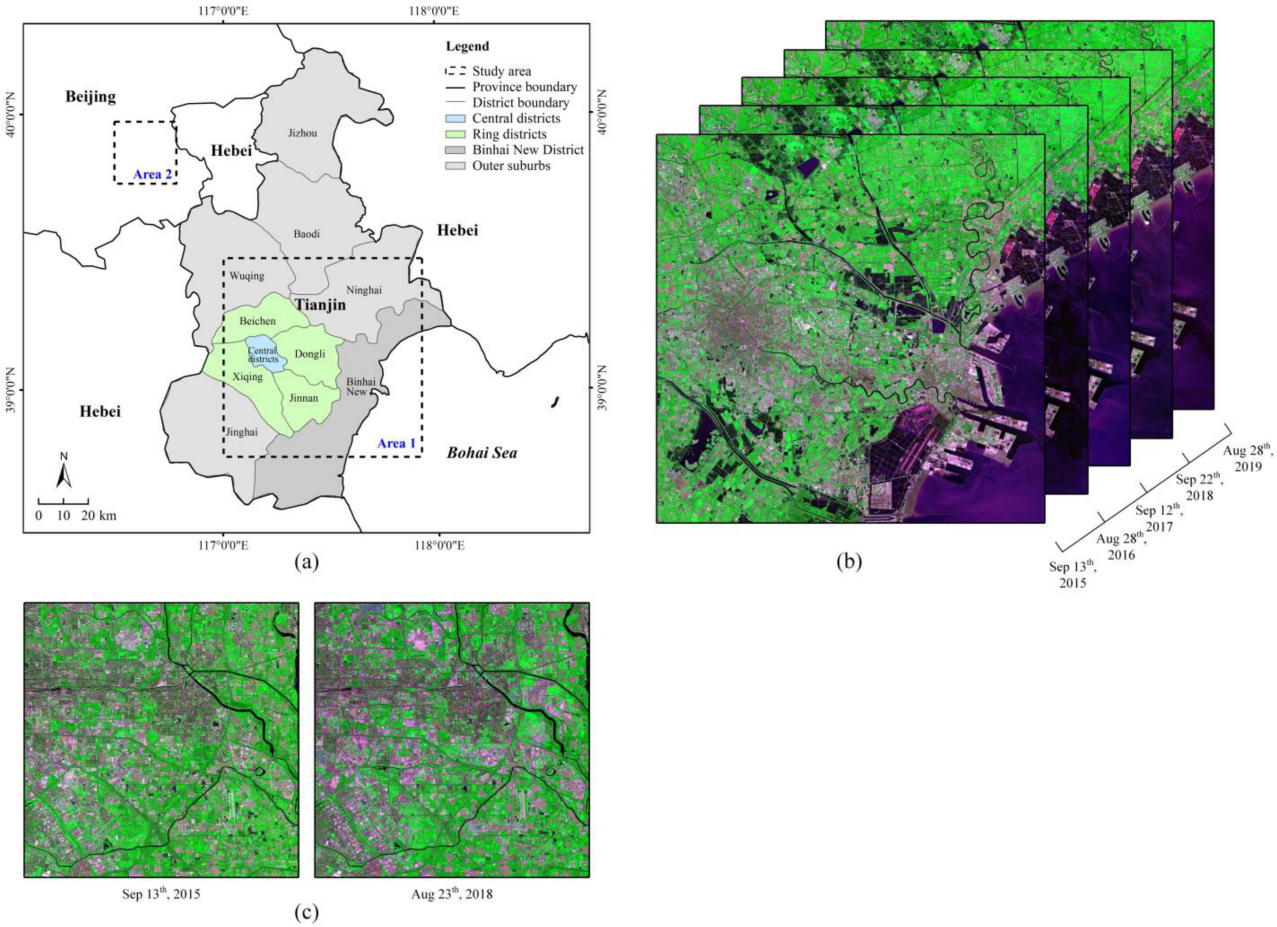

Fig. 3. (a) Geographic location of two study areas. (b) False color composite of multitemporal Sentinel-2 multispectral images dataset in Area 1. (c) False color composite of bitemporal Sentinel-2 multispectral images in Area 2 (Bands red, near infrared, and green as R, G, B color composite).

districts, including six central districts, four ring districts, Binhai New District, and five outer suburbs. The area covers the central and ring districts, and parts of the Binhai New District and outer suburbs. The main land cover types in the area are built-up area, vegetation, bare soil, and water body. Along with rapid economic development, Tianjin urban area has been dramatically expanding in the past three decades. The second area is in the east of Beijing, which mainly covers Beijing's administrative subcenter, Tongzhou District and develops rapidly these years.

B. Data

Multitemporal Sentinel-2 multispectral images from 2015 to 2019 were used in Area 1 [see Fig. 3(b)]. And bitemporal Sentinel-2 multispectral images acquired in 2015 and 2018 were used in Area 2 [see Fig. 3(c)]. A cloud-free image was selected for each year. The images used were acquired in August or September, when the vegetation flourished and bare soil area was less. The use of images in summer could reduce the confusion between built-up area and bare soil. Ten spectral bands of Sentinel-2 images containing bands 2–8, 8a, 11, and 12, with spatial resolutions of 10 and 20 m, were used.

The Sentinel-2 images used are Level-1C data after geometric correction. These Level-1C data were preprocessed to generate the bottom-of-atmosphere corrected reflectance, i.e., Level-2A data, after atmospheric correction. Sen2Cor, a Level-2A processor in ESA'S sentinel application platform (SNAP), was used to correct for the atmosphere effects on Level-1C products [56] (http://step.esa.int/main/snap-supported-plugins/sen2cor/). Four spectral bands of 20 m were then resampled to 10 m using Sen2Res model implemented in the SNAP toolbox [57] (http://step.esa.int/main/snap-supported-plugins/sen2res/). The Sentinel-2 images acquired in 2016–2019 in Area 1 and the 2018 image in Area 2 were all coregistered to their corresponding 2015 images. The root mean square errors of registration are less than 0.1 pixels. The image size of the Sentinel-2 images is 8000×8000 pixels in Area 1 and 2500×2500 pixels in Area 2.

Both the training samples of urban built-up area change and the test samples for two classes were selected by visual interpretation of bitemporal images (see Table I). Since urban built-up area change is the only target class, training samples for urban built-up area change were selected. The test samples for both the target class samples and the nontarget class samples were randomly selected. Since urban built-up area change (target class) occurred in a small portion of urban area, the nontarget class samples were selected as roughly twice of the target class samples.

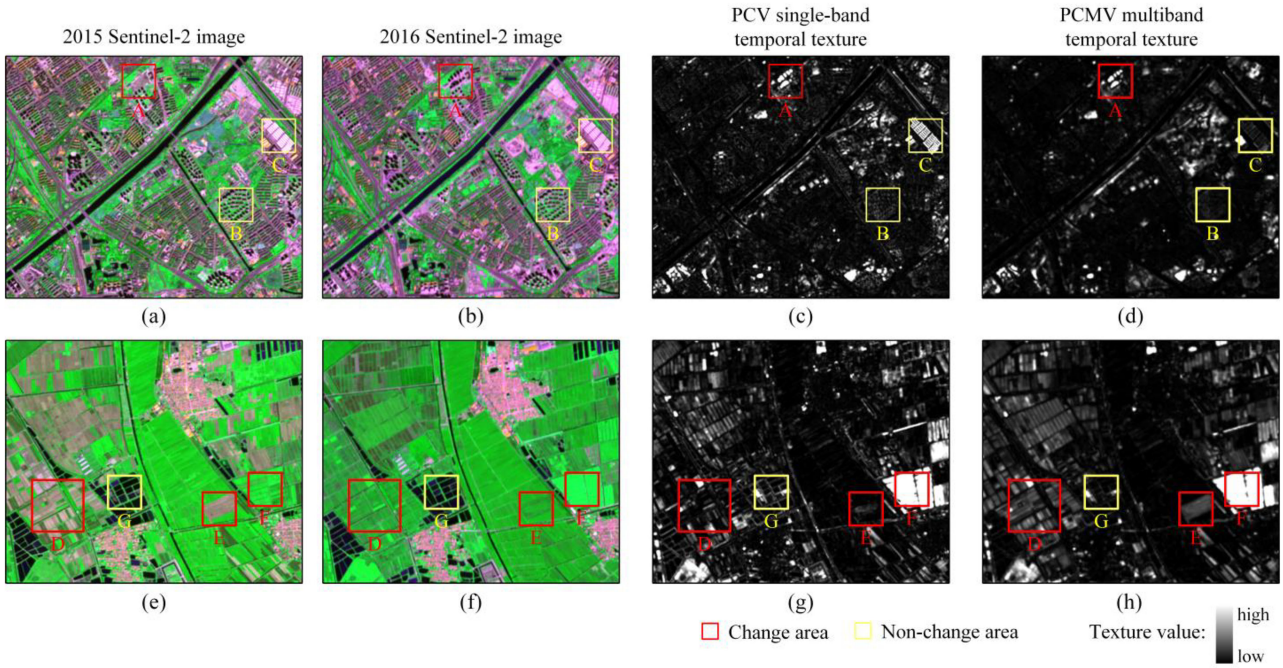

Fig. 4. Two portions of Sentinel-2 images of 2015 [(a) and (e)] and 2016 [(b) and (f)], PCV single-band temporal texture images [(c) and (g)], and PCMV multiband temporal texture images [(d) and (h)] generated from these bitemporal images ($\boldsymbol{M} = \boldsymbol{C}^{-1}$, $w = 5\times5$, $\boldsymbol{h} = 0$) in an urban area and a suburban area. Parts A–G are seven selected details of bitemporal images and temporal texture images in these areas.

TABLE I
PIXEL NUMBERS OF THE TRAINING AND TEST SAMPLES FOR FOUR BITEMPORAL IMAGES

Study area	Bi-temporal images	Training samples (Target class)	Test samples	
			Target class	Non-target class
Area 1	2015 and 2016 images	2169	46819	96747
	2016 and 2017 images	1951	40063	81824
	2017 and 2018 images	2088	37884	76065
	2018 and 2019 images	2015	59881	121765
Area 2	2015 and 2018 images	1698	15555	31027

IV. RESULTS AND DISCUSSION

A. Analysis of Multiband Temporal Texture

The multiband temporal textures measured by PCMV were computed from Sentinel-2 multispectral images between every two years. To understand how the PCMV texture highlights land cover changes, two portions of PCMV texture images from images of 2015 and 2016 in Area 1 were selected for analysis (see Fig. 4). For comparison, the PCV single-band temporal texture images of the same areas were also selected and shown in Fig. 4. These two portions represent multiband temporal textures in two different scenarios. One portion is in an urban area containing buildings and urban vegetation [see Fig. 4(a)], whereas the other portion is in a suburban area that is mainly covered by vegetation, bare soil, and water body [see Fig. 4(e)].

From bitemporal spectral images shown in Fig. 4(a) and (b), most of the change areas in this portion are urban land change. It is found that all urban land expansions in this portion are highlighted by high values in two different temporal texture images [see Fig. 4(c) and (d)]. However, nonchange areas are shown differently in these two temporal texture images. For example, Part A is changed from bare land to built-up area. This area is highlighted by high values in both single-band and multiband temporal texture images. Parts B and C are nonchange areas. The stable urban built-up area in Part B has slightly lower values in the multiband temporal texture [Fig. 4(d)] than that of the single-band temporal texture [Fig. 4(c)]. Another unchanged area in Part C also shows very low values in the multiband temporal texture, whereas it is significantly highlighted by high values in the single-band temporal texture, which is confused with the change area.

In the suburban area, the change areas mainly include changes between non-built-up classes, i.e., vegetation, bare land, and water [see Fig. 4(e) and (f)]. Three local areas of cropland were selected in temporal texture images (Parts D, E, and F), which show different texture values. The areas in Parts D and E are changed from bare soil to vegetation. These change areas are highlighted more uniform and complete in the multiband temporal texture image [Fig. 4(h)] than in the single-band temporal texture image [Fig. 4(g)]. Part F is vegetation change with different brightness. This area is significantly highlighted in both temporal texture images. To distinguish built-up area change and vegetation change, it is necessary to make a distinction of vegetation change by combining bitemporal NDVIs in the detection of built-up area change. Part G is a stable water body area. The boundary of water shows high values in the single-band

TABLE II
ACCURACY RESULTS OF BUILT-UP AREA CHANGE DETECTION WITH FOUR BITEMPORAL IMAGES IN AREA 1 USING FOUR DIFFERENT FEATURE COMBINATIONS

Bi-temporal image pairs	Feature combinations	OA (%)	F1 score	Built-up change area		Other change and non-change areas	
				PA (%)	UA (%)	PA (%)	UA (%)
2015 and 2016 images	Spectral	87.55	0.8088	80.77	81.00	90.83	90.71
	Spectral + NDVI	89.48	0.8368	82.71	84.68	92.76	91.72
	Spectral + NDVI + PCV	92.81	0.8871	86.62	90.90	95.80	93.67
	Spectral + NDVI + PCMV	**94.93**	**0.9222**	**92.19**	**92.26**	**96.26**	**96.22**
2016 and 2017 images	Spectral	91.19	0.8596	82.06	90.25	95.66	91.59
	Spectral + NDVI	91.81	0.8635	78.79	95.51	**98.19**	90.44
	Spectral + NDVI + PCV	93.05	0.8867	82.72	**95.54**	98.11	92.06
	Spectral + NDVI + PCMV	**94.98**	**0.9219**	**90.19**	94.27	97.32	**95.30**
2017 and 2018 images	Spectral	90.30	0.8406	76.93	92.65	96.96	89.41
	Spectral + NDVI	90.38	0.8374	74.46	**95.66**	**98.32**	88.54
	Spectral + NDVI + PCV	88.57	0.8110	73.76	90.05	95.94	88.01
	Spectral + NDVI + PCMV	**92.45**	**0.8796**	**82.98**	93.58	97.17	**91.98**
2018 and 2019 images	Spectral	90.72	0.8546	82.70	88.40	94.66	91.76
	Spectral + NDVI	90.67	0.8498	80.10	90.49	95.86	90.74
	Spectral + NDVI + PCV	86.78	0.7908	75.81	82.65	92.17	88.57
	Spectral + NDVI + PCMV	**93.54**	**0.8983**	**86.53**	**93.39**	**96.99**	**93.61**

Numbers in bold represent the highest accuracy values for each group. OA: overall accuracy; PA: producer's accuracy; UA: user's accuracy.

temporal texture [Fig. 4(g)], whereas it shows low values in the multiband temporal texture [Fig. 4(h)]. In general, the PCMV texture is more effective in highlighting land cover change than in the single-band temporal texture.

B. Results of Built-Up Area Change Detection

Bitemporal Sentinel-2 images pairs were selected to evaluate the effectiveness of the proposed method in the detection of urban built-up area change in two study areas. The bitemporal image pairs in Area 1 consisted of bitemporal images of two successive years from 2015 to 2019. Specifically, the multiband temporal textures of multiple scales were adopted. Five window sizes (i.e., 3, 5, 7, 9, 11) and two lag distances $\boldsymbol{h} = 0$ and 1 were used to compute the texture values in this study. Thus, 32 features, including ten multiband temporal textures, 20 spectral bands of two dates, and two NDVIs, were stacked into a combined dataset to classify urban built-up area change. In addition, the other three feature combinations were used for comparison. The first comparative feature combination included 20 spectral bands. The second comparative feature combination included 22 features, i.e., 20 spectral bands and two NDVIs. The third comparative feature combination included 32 features, i.e., 20 spectral bands of two dates, two NDVIs, and ten PCV single-band temporal textures. The ten single-band temporal textures of multiple scales were computed with five window sizes (i.e., 3, 5, 7, 9, 11) and two lag distances $\boldsymbol{h} = 0$ and 1. The parameter determination and nontarget sample selection for the iOCRF classification were followed to Shi *et al.* [39].

The detection accuracies of the proposed method and three comparative feature combinations in Area 1 are listed in Table II. First, the OAs and $F1$ scores using bitemporal spectral features and NDVIs are similar or slightly higher than those using spectral features alone, whereas the accuracies using the proposed method are higher than those two feature combinations. Although the inclusion of bitemporal NDVIs has little effect on the detection of built-up area change, the bitemporal NDVIs are helpful to reduce the impact of vegetation changes highlighted by the PCMV texture on built-up area change.

From Table II, the proposed method with the inclusion of multiband temporal texture performs better than three comparative feature combinations for all these four pairs of bitemporal images in OAs and $F1$ scores. For these four bitemporal image pairs, the OAs using the proposed method are 2.15%–7.38% higher than using the bitemporal spectral features, 2.07%–5.45% higher than using the bitemporal spectral features and NDVIs, and 1.93%–6.76% higher than those of including PCV single-band temporal texture, respectively. In addition, the computation of the PCV texture is highly affected by the band selected. Not all image pairs have obvious improvement with the inclusion of PCV texture. The accuracies of classification with the incorporation of multiband temporal texture (92.45%–94.98% in OAs) are more robust than the single-band temporal texture (86.78%–93.05% in OAs).

For the PAs and UAs of target class (urban built-up area change) (see Table II), it is found that the PAs of target class using the proposed method are all the highest for these four bitemporal image pairs, which vary from 82.98% to 92.19%. There are some differences in the UAs of target class, although the UAs (92.26%–94.27%) are not far apart for four bitemporal image pairs. Specifically, the PAs and UAs for 2015–2016 and 2018–2019 image pairs are similar. The PAs and UAs of the target class using the proposed method are about 5.6%–13.5% and 1.4%–10.7% higher than using bitemporal spectral features and NDVIs and with inclusion of single-band temporal texture. However, for 2016–2017 and 2017–2018 image pairs, although the PAs of the target class using the proposed method are higher than those two feature combinations (7.5%–11.4%), some of the UAs of the target class using the proposed method are slightly lower (1.2%–2.1%). Besides, the PAs and UAs of the target class using the proposed method are more balanced than those using the comparative feature combinations. Taking the first two image pairs as an example, there are 0.1% and 4.1% differences between PAs and UAs for the proposed method, whereas the differences are 4.3% and 12.8% for the method with inclusion of single-band temporal texture and 2% and 16.7% when using bitemporal spectral features and NDVIs.

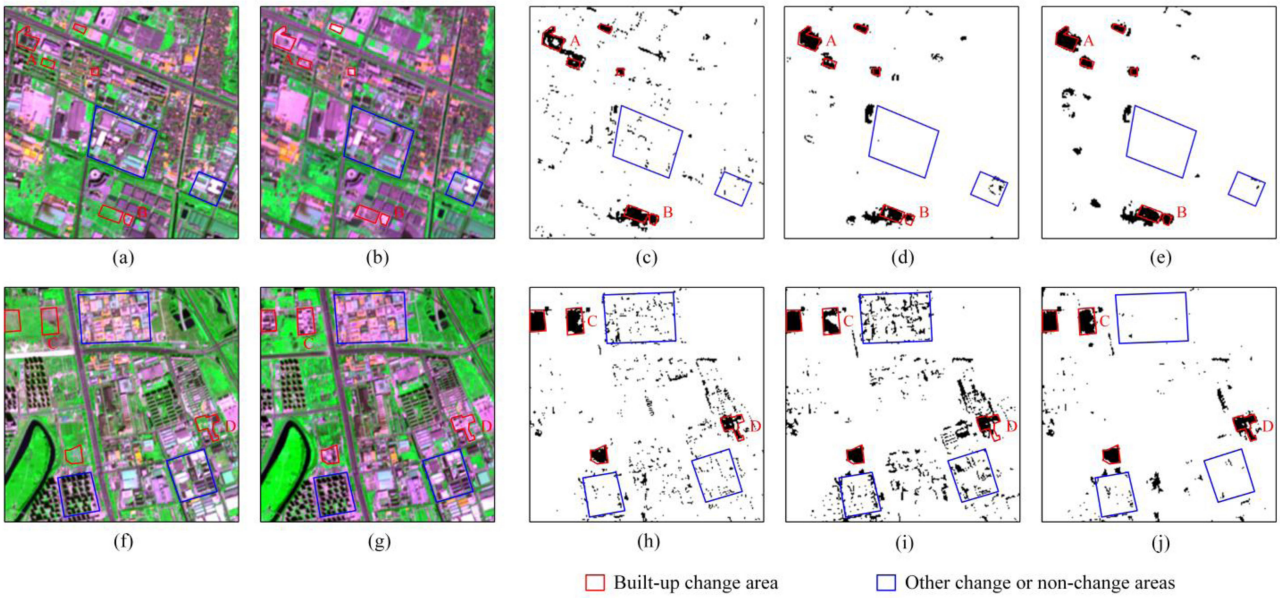

Fig. 5. Portions of Sentinel-2 images of (a) 2015, (b) 2016, (f) 2018, and (g) 2019, and their corresponding detection results of urban built-up area change using three different feature combinations: (c) and (h): the "Spectral + NDVI" combined dataset in iOCRF; (d) and (i): the "Spectral + NDVI + PCV" combined dataset in iOCRF; (e) and (j): the proposed "Spectral + NDVI + PCMV" combined dataset in iOCRF. Black denotes detection results of built-up change area and white denotes detection results of the other change and nonchange areas. Parts A–D are four selected details of built-up change areas for comparison.

Two portions of the detection results of 2015–2016 and 2018–2019 image pairs in Area 1 were selected as examples to visually compare three different feature combinations (see Fig. 5). For the portion of detection results using different feature combinations shown in Fig. 5(c)–(e), the classification results with the incorporation of two different temporal textures in the iOCRF classification perform better than using bitemporal spectral features and NDVIs. For example, for the built-up change area in Part A, some built-up change areas are underestimated using bitemporal spectral features and NDVIs (e.g., white areas in red polygon) [see Fig. 5(c)]. There is also more overestimation (e.g., black areas in blue polygon) in classification using bitemporal spectral features and NDVIs [see Fig. 5(c)]. Although the inclusions of two temporal textures (single-band and multiband) in iOCRF classification produce better results [see Fig. 5(d) and (e)], overestimation and underestimation are reduced more including multiband temporal texture. For example, the built-up change area in Part B is more complete in classification result with the inclusion of multiband temporal texture than the single-band temporal texture, as well as less overestimation in the blue polygon.

For the portion of detection results shown in Fig. 5(h)–(j), the result with inclusion of multiband temporal texture performs better than using bitemporal spectral features and NDVIs and the inclusion of single-band temporal texture. The results of this portion [see Fig. 5(f) and (g)] are different from those of the portion in Fig. 5(a) and (b) that the additions of two different temporal textures all perform better. For example, for the built-up change area in Part C, the detection result with the inclusion of multiband temporal texture [Fig. 5(j)] has less underestimation (e.g., white areas in red polygon) than that of the single-band temporal texture [Fig. 5(i)]. Another built-up change area in Part D is also extracted more completely using the proposed method [Fig. 5(j)], compared with those from two comparative feature combinations. Besides, the overestimation in the blue polygons is significantly reduced in classification with the inclusion of multiband temporal texture than those from the other two feature combinations.

The comparison of detection results from different feature combinations demonstrated that the proposed method using the PCMV texture and iOCRF classification produced more accurate built-up area change results with less noise [see Fig. 5(e) and (j)]. However, the result with the incorporation of single-band temporal texture for the 2018–2019 image pair [Fig. 5(i)] does not show the same improvement compared with that for the 2015–2016 image pair [Fig. 5(d)]. Thus, the inclusion of multiband temporal texture in iOCRF classification performed more robustly than the single-band temporal texture in different image pairs. These results are consistent with the accuracies listed in Table II.

The detection accuracies of different feature combinations for bitemporal Sentinel-2 images in Area 2 are listed in Table III. The proposed method with the inclusion of PCMV texture performs better than three comparative feature combinations. The OA using the proposed method increases 3.28%, 3.12%, and 2.65% over those of three comparative feature combinations, respectively. The PA and UA are also the highest for the proposed method. These results of the proposed method in this study area achieved the best performance, which is the same as that in Area 1. This indicated that the proposed method has a generalization ability in different study areas.

For the comparative method using postclassification comparison based on OCSVM classifier, the same test samples (as presented in Table I) were used for validation. To reduce the impact of training samples on the extraction of built-up areas in different

TABLE III
ACCURACY RESULTS OF BUILT-UP AREA CHANGE DETECTION WITH BITEMPORAL IMAGES IN AREA 2 USING FOUR DIFFERENT FEATURE COMBINATIONS

Bi-temporal image pairs	Feature combinations	OA (%)	F1 score	Built-up change area		Other change and non-change areas	
				PA (%)	UA (%)	PA (%)	UA (%)
2015 and 2018 images	Spectral	93.74	0.9031	87.35	93.48	96.94	93.86
	Spectral + NDVI	94.37	0.9133	88.79	94.01	97.16	94.53
	Spectral + NDVI + PCV	93.90	0.9054	87.45	93.85	97.13	93.92
	Spectral + NDVI + PCMV	**97.02**	**0.9557**	**96.23**	**94.91**	**97.41**	**98.10**

Numbers in bold represent the highest accuracy values for each group. OA: overall accuracy; PA: producer's accuracy; UA: user's accuracy.

TABLE IV
ACCURACY RESULTS OF BUILT-UP AREA CHANGE DETECTION WITH BITEMPORAL IMAGES USING THE OCSVM-BASED POSTCLASSIFICATION COMPARISON

Study area	Bi-temporal image pairs	OA (%)	F1 score	Built-up change area		Other change and non-change areas	
				PA (%)	UA (%)	PA (%)	UA (%)
Area 1	2015 and 2016 images	84.26	0.7214	62.50	85.31	94.79	83.93
	2016 and 2017 images	89.00	0.8326	83.22	83.31	91.84	91.79
	2017 and 2018 images	79.34	0.7031	73.58	67.32	82.21	86.20
	2018 and 2019 images	81.14	0.6831	61.68	76.54	90.70	82.80
Area 2	2015 and 2018 images	88.17	0.8163	78.70	84.78	92.92	89.69

OA: overall accuracy; PA: producer's accuracy; UA: user's accuracy.

years, training samples for each year were randomly selected in the invariant built-up area from multitemporal images. In the classification, ten spectral bands and NDVI of each image were stacked into a combined dataset to classify the built-up area. The detection accuracies of the comparative method are listed in Table IV. For the four bitemporal image pairs in Area 1, the OAs using the proposed method (see Table II) are 5.98%–13.11% higher than those using this comparative method. And the OA using the proposed method (see Table III) increases 8.85% for the bitemporal images in Area 2. The PA and UA are also higher than this comparative method. The proposed method based on multitemporal image classification extracts built-up area change directly and is not affected by the built-up area extraction of each image. These results demonstrated the effectiveness of the PCMV texture and iOCRF classification in the detection of urban built-up area change.

To understand the contributions of different features to the iOCRF classification in the proposed method, i.e., spectral features, NDVIs, and PCMV textures of multiple scales, feature importance of different features was calculated (see Fig. 6). In general, all three types of features show varying degrees of importance in urban built-up change classification. For the PCMV textures, feature importance scores for the textures with $\boldsymbol{h} = 0$ are higher than those with $\boldsymbol{h} = 1$. The textures with larger window sizes (i.e., 7, 9, and 11) have higher importance scores in classification than the textures with smaller window sizes (i.e., 3 and 5). In addition, feature importance scores of spectral features and NDVI from T2 image are higher than those from T1 image. The RGB bands, the first vegetation red edge band, the shortwave infrared bands and NDVI from T2 image and the shortwave infrared bands from T1 image have high feature importance scores in the classification.

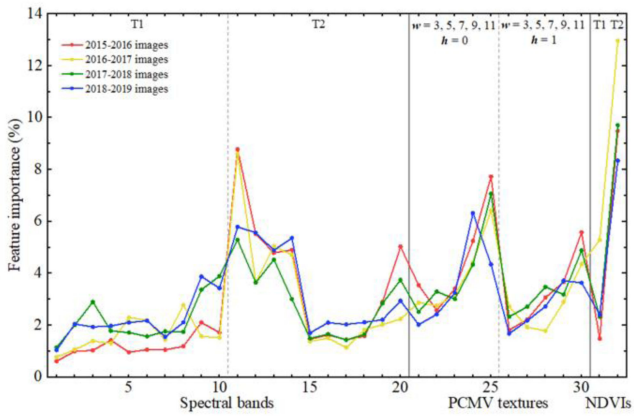

Fig. 6. Feature importance of the proposed "Spectral + NDVI + PCMV" combined dataset in iOCRF classification.

C. Analysis of Built-Up Area Changes Over Multiple Years

To further validate the proposed method and to analyze the urbanization, urban built-up area change detection results from five-year Sentinel-2 images in Area 1 using the proposed method were combined to generate an annual urban land change result. It should be noted that most built-up area changes were detected in a specific year. But in some pixel locations, urban built-up area changes were extracted for multiple times, e.g., two, three, or four years. Since the period of study is relatively short (i.e., five years), it is not possible to have multiple urban changes.

To analyze these pixels detected as urban built-up area change for multiple years, an example of multiple urban land changes detected is shown in Fig. 7. The detection results for multiple years were stacked and represented in different colors for analysis. By observing the multitemporal Sentinel-2 images of this area from 2015 to 2017 and the very high spatial resolution images collected from Google Earth, it was found that the area had been under construction in the past two years. The image in 2016 shows the status of construction. The buildings under construction are easily confused with built-up area in classification. Therefore, using multitemporal images of two successive

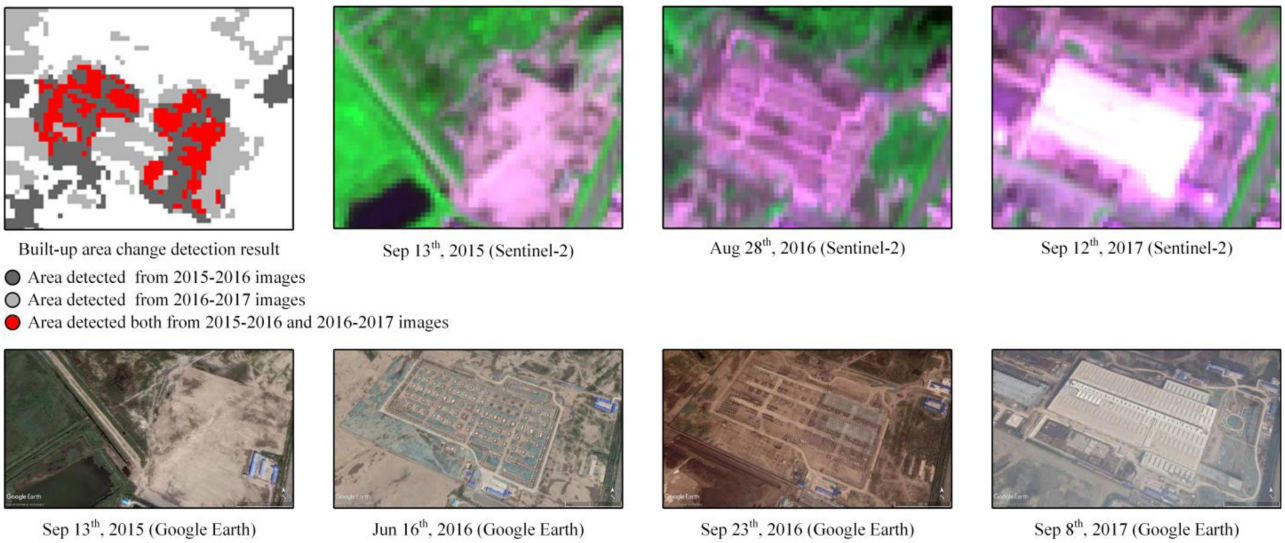

Fig. 7. Built-up area change detection result of an example of multiple urban land changes, the corresponding Sentinel-2 images from 2015 to 2017, and very high spatial resolution images acquired from Google Earth.

years may lead to misclassification and multiple extractions before the construction of buildings is completed. In this case, since built-up area change is assumed to occur in a certain year, the detection result from 2015–2016 Sentinel-2 images is not accurate. The time of urban built-up area change detected both from 2015–2016 and 2016–2017 images needs to be calibrated to the latest year of extraction, i.e., built-up area change from 2016 to 2017.

Therefore, to generate the annual mapping result of built-up area changes over these years in the study area, a postprocessing step was carried out to correct the built-up area changes extracted in multiple years. Since these areas were mainly multiyear changes of built-up area, changes extracted at an earlier time might actually only be the intermediate process of becoming built-up area. Therefore, the change time for these areas extracted in multiple years was calibrated to the latest detection year. In addition, there are some small connected components as noises in the detection results. It was found that most of these small areas are smaller than ten pixels by visual inspection. Thus, the areas less than ten pixels were considered as noises and masked out in mapping. The mapping result over the study area after the above postprocessing steps is shown in Fig. 8.

From Fig. 8, it is found that the built-up area changes mainly occurred in the central districts, four ring districts, and the Binhai New District. The central districts have relatively fewer built-up area changes, which are evenly distributed across the region. There are more urban land changes occurred in the four ring districts outside the boundary of the central districts than inside the central districts. Besides, with the development pattern of two main urban zones, i.e., the central districts and the Binhai New District, there are more built-up area changes detected in the Binhai New District than the other outer suburbs, especially in the northeastern coastal area of Binhai New District. Another area with significant urban land changes is a newly built Beijing-Tianjin cooperation demonstration zone, located in the northwest of Binhai New District. Moreover, the surrounding regions of these two main urban zones are developing in the east-west direction. The non-built-up areas in the ring districts between these two zones are changing rapidly to built-up areas. In general, about 5.2% of the total areas have been detected as built-up area changes in the study area over these years, and most areas are not affected by urban land expansion.

D. Discussion

In this study, a multiband temporal texture measured using PCMV was proposed and evaluated in the detection of urban built-up area change using multitemporal Sentinel-2 multispectral images. The proposed method combines the multiband temporal textures with bitemporal spectral features and NDVIs in the iOCRF classification. The proposed method was compared with three comparative feature combinations and a postclassification comparison method based on OCSVM for validation. The experimental results demonstrated that the urban built-up area change was extracted more accurately using the proposed method. The proposed method achieved better performance than the use of bitemporal spectral features and NDVIs, which demonstrated that the addition of multiband temporal texture in the classification improved the accuracies of detection results. The proposed method performed better and more robustly than the incorporation of PCV single-band temporal texture by experimenting in different image pairs. The proposed method also outperformed the OCSVM-based postclassification comparison method, which is affected by the built-up area extraction results of each image.

The advantages of the proposed method can be summarized as follows. First, unlike general textures that only reflect spatial information of a single image, the multiband temporal texture in this article quantifies local spatio-temporal dependence from bitemporal multispectral images. The multiband temporal texture is derived simultaneously from all available bands. Any number of input bands from bitemporal images could

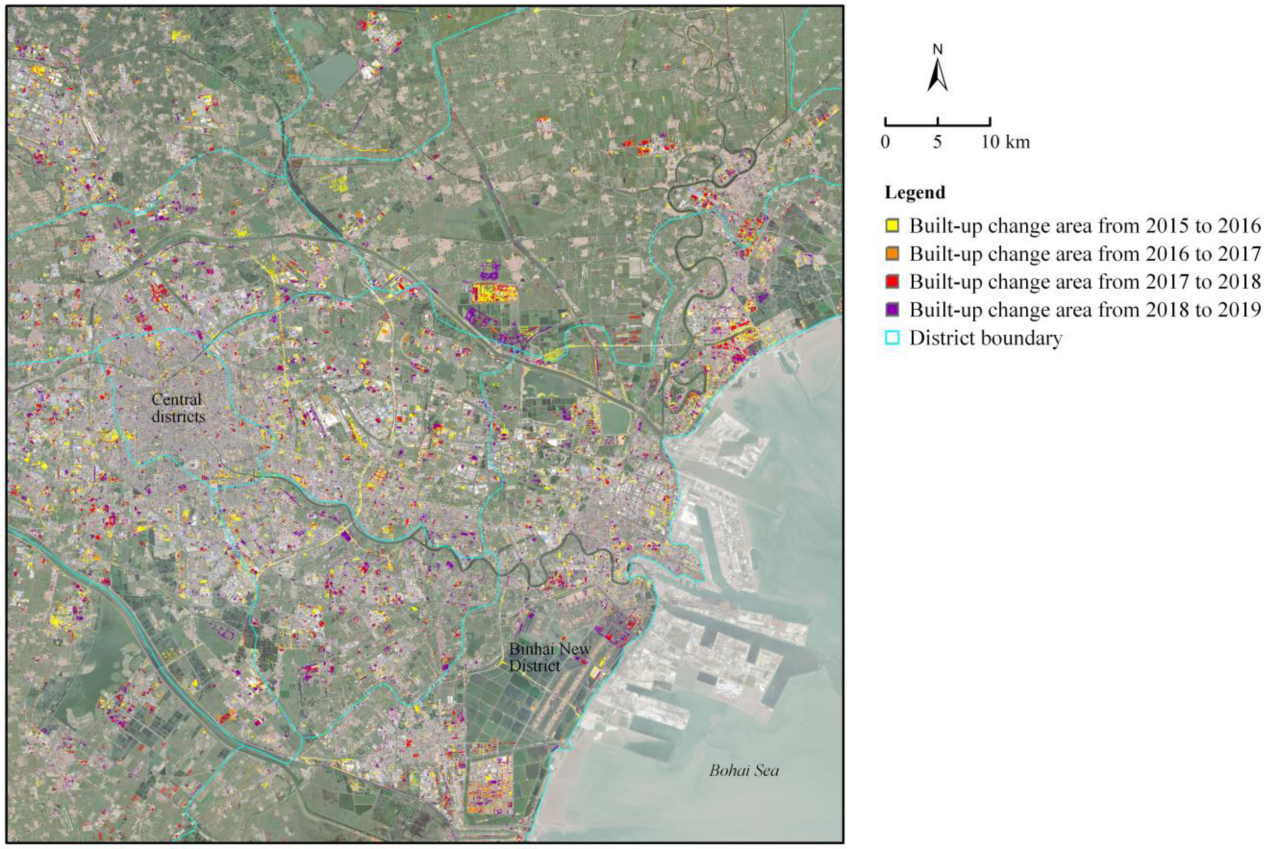

Fig. 8. Mapping result of urban built-up area changes using the proposed method with four bitemporal Sentinel-2 image pairs from 2015 to 2019 in Area 1. The 2015 image is shown as background.

be employed to extract the PCMV texture, in particular when the optimal spectral band to detect the changes of interest is not known. The experiments demonstrated the PCMV texture highlights land cover changes with different texture values and benefits for built-up area change detection than the PCV single-band temporal texture [27]. In addition, the multiband temporal textures with different window sizes and lag distances exhibit different spatio-temporal information of land cover changes at multiple scales. The combination of these PCMV textures at multiple scales provides complementary information. Therefore, instead of selecting a texture at an appropriate scale with the best performance, multiple-scale PCMV textures were used in this study.

Second, the proposed method based on multitemporal image classification is appropriate for including the PCMV texture and can detect built-up area change directly. The experiments demonstrated that the joint use of spectral features and PCMV textures improved the built-up change detection. The postclassification comparison method can only use multiple features obtained from each image for classification and cannot use any multiband temporal texture [55]. The classification of built-up areas in prechange and postchange images will affect the change detection [58].

Third, since the built-up area change is the only target class of interest, the one-class image classification in the proposed method, i.e., iOCRF, is easier to be implemented than the conventional multiclass classification methods. It was also found that the OCSVM classifier cannot be integrated into the proposed method. The iOCRF classifier is insensitive to high-dimensionality data, which is especially suitable for the classification of multitemporal multispectral images and the combination of multiple features [39]. Moreover, the iOCRF classifier has the ability to quantify feature importance [52]. From these experimental results, these three types of features (spectral feature, NDVI, and multiband temporal texture) all played an important role in the change classification.

Although the proposed method produced very promising results, there are still some problems to be solved in the future. The multiband temporal texture highlights all kinds of land cover changes, but not just the target built-up area change. Different land cover changes with similar texture values exist and are easily confused. For example, the use of images containing buildings under construction, especially for areas under construction in the past several years, can easily lead to misclassification and multiple extractions. Bare soil and built-up area with similar spectral signatures are likely to be confused in the classification. In addition, by comparing the detection results obtained from different image pairs, the differences in data quality, training samples selected for classification could also lead to different detection results. Therefore, the methods of detecting built-up area change using multiband temporal

texture and the postprocessing method for analyzing built-up area changes in multiple years will be studied in the future.

Although the multiband temporal texture was evaluated in the detection of urban built-up area change, the proposed method will be further evaluated using more high spatial resolution images over a longer period or in other urban areas. Moreover, the multiband temporal texture will be also evaluated in the situations that other land cover changes are of interest.

V. Conclusion

In this article, a simple and reliable method for detection of urban built-up change using the PCMV multiband temporal textures and iOCRF classification was proposed. The PCMV texture quantifies local spatio-temporal dependence between bitemporal multispectral images. The urban built-up area change is detected by classifying the combined dataset of spectral features, NDVIs, and multiband temporal textures obtained using the iOCRF classifier. The proposed multitemporal image classification method based on iOCRF focuses on the target class of built-up area change, and is suitable for including the multiband temporal texture and insensitive to the data dimensionality. The experimental results demonstrated that the proposed method outperformed three comparative feature combinations and the OCSVM based postclassification comparison method for all the bitemporal Sentinel-2 image pairs in two study areas. The detection results in five successive years were processed to generate an annual urban land change result and were analyzed. More studies are needed to further evaluate the performance of the proposed method in other study areas and other changes of interest.

References

[1] C. M. Mertes *et al.*, "Detecting change in urban areas at continental scales with MODIS data," *Remote Sens. Environ.*, vol. 158, pp. 331–347, 2015.

[2] S. Angel *et al.*, "The dimensions of global urban expansion: Estimates and projections for all countries, 2000–2050," *Prog. Planning*, vol. 75, no. 2, pp. 53–107, 2011.

[3] A. G. O. Yeh and X. Li, "Measurement and monitoring of urban sprawl in a rapidly growing region using entropy," *Photogramm. Eng. Remote Sens.*, vol. 67, no. 1, pp. 83–90, 2001.

[4] A. E. J. Morris, *History of Urban Form Before the Industrial Revolution*. London, U.K.: Routledge, 2013.

[5] Z. Zhang *et al.*, "Urban expansion in China based on remote sensing technology: A review," *Chin. Geo-Graphical Sci.*, vol. 28, no. 5, pp. 727–743, 2018.

[6] A. O. Onyango, "Global and regional trends of urbanization: A critical review of the environmental and economic imprints," *World Environ.*, vol. 8, no. 2, pp. 47–62, 2018.

[7] K. C. Seto, B. Güneralp, and L. R. Hutyra, "Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools," *Proc. Nat. Acad. Sci. USA*, vol. 109, no. 40, pp. 16083–16088, 2012.

[8] B. Maimaiti *et al.*, "Characterizing urban expansion of Korla city and its spatial-temporal patterns using remote sensing and GIS methods," *J. Arid Land*, vol. 9, pp. 458–470, 2017.

[9] K. Dhanaraj and D. P. Angadi, "Land use land cover mapping and monitoring urban growth using remote sensing and GIS techniques in Mangaluru, India," *GeoJournal*, 2020, doi: 10.1007/s10708-020-10302-4.

[10] F. Yuan *et al.*, "Land cover classification and change analysis of the twin cities (Minnesota) metropolitan area by multitemporal Landsat remote sensing," *Remote Sens. Environ.*, vol. 98, no. 2/3, pp. 317–328, 2005.

[11] B. Chai and K. C. Seto, "Conceptualizing and characterizing micro-urbanization: A new perspective applied to Africa," *Landscape Urban Planning*, vol. 190, 2019, Art. no. 103595.

[12] C. R. Fichera, G. Modica, and M. Pollino, "Land cover classification and change-detection analysis using multi-temporal remote sensed imagery and landscape metrics," *Eur. J. Remote Sens.*, vol. 45, no. 1, pp. 1–18, 2012.

[13] D. Lu, E. Moran, and S. Hetrick, "Detection of impervious surface change with multitemporal Landsat images in an urban–rural frontier," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 3, pp. 298–306, 2011.

[14] Y. G. Qian *et al.*, "Quantifying spatiotemporal pattern of urban greenspace: New insights from high resolution data," *Landscape Ecol.*, vol. 30, pp. 1165–1173, 2015.

[15] J. Deng *et al.*, "A methodology to monitor urban expansion and green space change using a time series of multi-sensor SPOT and Sentinel-2A images," *Remote Sens.*, vol. 11, no. 10, 2019, Art. no. 1230.

[16] Y. Cai, M. Zhang, and H. Lin, "Estimating the urban fractional vegetation cover using an object-based mixture analysis method and Sentinel-2 MSI imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 341–350, Jan. 2020, doi: 10.1109/JSTARS.2019.2962550.

[17] M. Papadomanolaki, S. Verma, M. Vakalopoulou, S. Gupta, and K. Karantzalos, "Detecting urban changes with recurrent neural networks from multitemporal Sentinel-2 data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 214–217.

[18] D. Jiang, J. Xu, and Y. Gao, "Urban change detection of Pingtan city based on bi-temporal remote sensing images," *IOP Conf. Ser., Earth Environ. Sci.*, vol. 57, 2017, Art. no. 012056.

[19] A. Alghamdi and A. R. Cummings, "Assessing Riyadh's urban change utilizing high-resolution imagery," *Land*, vol. 8, no. 12, 2019, Art. no. 193.

[20] M. Pesaresi *et al.*, "Assessment of the added-value of Sentinel-2 for detecting built-up areas," *Remote Sens.*, vol. 8, no. 4, 2016, Art. no. 299.

[21] C. Qiu *et al.*, "Local climate zone-based urban land cover classification from multi-seasonal Sentinel-2 images with a recurrent residual network," *ISPRS J. Photogramm. Remote Sens.*, vol. 154, pp. 151–162, 2019.

[22] A. Lefebvre, C. Sannier, and T. Corpetti, "Monitoring urban areas with Sentinel-2A data: Application to the update of the Copernicus high resolution layer imperviousness degree," *Remote Sens.*, vol. 8, no. 7, pp. 1–21, 2016.

[23] Y. Zhou, Y. Song, S. Cui, H. Zhu, J. Sun, and W. Qin, "A novel change detection framework in urban area using multilevel matching feature and automatic sample extraction strategy," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 3967–3987, Mar. 2021, doi: 10.1109/JSTARS.2021.3064311.

[24] Q. Guo, J. Zhang, C. Zhong, and Y. Zhang, "Change detection for hyperspectral images via convolutional sparse analysis and temporal spectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 4417–4426, Apr. 2021, doi: 10.1109/JSTARS.2021.3074538.

[25] J. Im and J. R. Jensen, "A change detection model based on neighborhood correlation image analysis and decision tree classification," *Remote Sens. Environ.*, vol. 99, no. 3, pp. 326–340, 2005.

[26] S. Li, P. Li, and T. Cheng, "Remote sensing change detection by inclusion of multitemporal texture," (in Chinese with English abstract), *Remote Sens. Land Resour.*, vol. 21, no. 3, pp. 35–40, 2009.

[27] H. Jin *et al.*, "Land cover classification using CHRIS/PROBA images and multi-temporal texture," *Int. J. Remote Sens.*, vol. 33, no. 1, pp. 101–119, 2012.

[28] J. Song, X. Wang, and P. Li, "Urban building damage detection from VHR imagery by including temporal and spatial texture features," *J. Remote Sens.*, vol. 16, no. 6, pp. 1233–1245, 2012.

[29] L. Moya *et al.*, "3D gray level co-occurrence matrix and its application to identifying collapsed buildings," *ISPRS J. Photogram. Remote Sens.*, vol. 149, pp. 14–28, 2019.

[30] L. Gueguen, M. Pesaresi, D. Ehrlich, L. Lu, and H. Guo, "Urbanization detection by a region based mixed information change analysis between built-up indicators," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 6, pp. 2410–2420, Dec. 2013.

[31] A. Safia and D. He, "Multiband compact texture unit descriptor for intra-band and inter-band texture analysis," *ISPRS J. Photogramm. Remote Sens.*, vol. 105, pp. 169–185, 2015.

[32] P. Li, T. Cheng, G. Moser, S. B. Serpico, and D. Ma, "Multitemporal change detection by spectral and multivariate texture information," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2007, pp. 1922–1925.

[33] X. Feng and P. Li, "Urban built-up area change detection using multi-band temporal texture and one-class random forest," in *Proc. 10th Int. Workshop Anal. Multitemporal Remote Sens. Images*, Shanghai, China, Aug. 5–7, 2019, pp. 1–4.

[34] H. Luo *et al.*, "Urban change detection based on Dempster–Shafer theory for multitemporal very high-resolution imagery," *Remote Sens.*, vol. 10, no. 7, 2018, Art. no. 980.
[35] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Trans. Inf. Theory*, vol. IT-14, no. 1, pp. 55–63, Jan. 1968.
[36] Y. Chen, X. S. Zhou, and T. S. Huang, "One-class SVM for learning in image retrieval," in *Proc . IEEE Int. Conf. Image Process.*, 2001, vol. 1, pp. 34–37.
[37] C. Désir *et al.*, "One class random forests," *Pattern Recognit.*, vol. 46, pp. 3490–3506, 2013.
[38] C. Desir *et al.*, "A new random forest method for one-class classification," in *Proc. Joint IAPR Int. Workshop Struct. Syntactic Pattern Recognit.*, 2012, vol. 7626, pp. 282–290.
[39] Z. Shi *et al.*, "A method for extraction of newly-built buildings in road region using morphological attribute profiles and one-class random forest," (in Chinese with English abstract), *Acta Scientiarum Naturalium Universitatis Pekinensis*, vol. 54, no. 1, pp. 105–114, 2018.
[40] G. Bourgault and D. Marcotte, "Multivariable variogram and its application to the linear model of coregionalization," *Math. Geol.*, vol. 23, no. 7, pp. 899–928, 1991.
[41] I. Clark, K. L. Basinger, and W. V. Harper, "MUCK—A novel approach to co-kriging," in *Geostatistical, Sensitivity, and Uncertainty Methods for Ground-Water Flow and Radionuclide Transport Modeling*. Columbus, OH, USA: Battelle, 1989, pp. 473–493.
[42] D. E. Meyers, "Pseudo cross-variograms, positive-definiteness, and cokriging," *Math. Geol.*, vol. 23, no. 6, pp. 805–816, 1991.
[43] R. De Maesschalck, D. Jouan-Rimbaud, and D. L. Massart, "The Mahalanobis distance," *Chemometrics Intell. Lab. Syst.*, vol. 50, no. 1, pp. 1–18, 2000.
[44] G. Bourgault, D. Marcotte, and P. Legendre, "The multivariate (co) variogram as a spatial weighting function in classification methods," *Math. Geol.*, vol. 24, no. 5, pp. 463–478, 1992.
[45] G. J. McLachlan, "Mahalanobis distance," *Resonance*, vol. 4, no. 6, pp. 20–26, 1999.
[46] S. Berberoglu *et al.*, "The integration of spectral and textural information using neural networks for land cover mapping in the Mediterranean," *Comput. Geosci.*, vol. 26, no. 4, pp. 385–396, 2000.
[47] A. O. Onojeghuo and G. A. Blackburn, "Mapping reedbed habitats using texture-based classification of quickbird imagery," *Int. J. Remote Sens.*, vol. 32, no. 23, pp. 8121–8138, 2011.
[48] C. J. S. Ferro and T. A. Warner, "Scale and texture in digital image classification," *Photogramm. Eng. Remote Sens.*, vol. 68, no. 1, pp. 51–63, 2002.
[49] M. Pesaresi, A. Gerhardinger, and F. Kayitakire, "A robust built-up area presence index by anisotropic rotation-invariant textural measure," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 1, no. 3, pp. 180–192, Sep. 2008.
[50] P. Li and H. Xu, "Land-cover change detection using one-class support vector machine," *Photogramm. Eng. Remote Sens.*, vol. 76, no. 3, pp. 255–263, 2010.
[51] A. Mellor *et al.*, "Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin," *ISPRS J. Photogramm. Remote Sens.*, vol. 105, pp. 155–168, 2015.
[52] A. Cutler, D. R. Cutler, and J. R. Stevens, "Random forests," in *Ensemble Machine Learning*. Boston, MA, USA: Springer, 2012, pp. 157–175.
[53] R. G. Congalton and K. Green, *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. Boca Raton, FL, USA: CRC Press, 2019.
[54] Y. Koga, H. Miyazaki, and R. Shibasaki, "A CNN-based method of vehicle detection from aerial images using hard example mining," *Remote Sens.*, vol. 10, no. 1, 2018, Art. no. 124.
[55] B. Schölkopf *et al.*, "Estimating the support of a high-dimensional distribution," *Neural Comput.*, vol. 13, pp. 1443–1471, 2001.
[56] R. Richter, J. Louis, and U. Müller-Wilm, "Sentinel-2 MSI—Level 2A products algorithm theoretical basis document," *Eur. Space Agency (Special Publ.)*, vol. 49, pp. 1–72, 2012.
[57] N. Brodu, "Super-resolving multiresolution images with band-independent geometry of multispectral pixels," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4610–4617, Aug. 2017.
[58] H. Liu and Q. Zhou, "Accuracy analysis of remote sensing change detection by rule-based rationality evaluation with post-classification comparison," *Int. J. Remote Sens.*, vol. 25, no. 5, pp. 1037–1050, 2004.

Xiaoxue Feng received the B.S. degree in engineering of surveying and mapping from Tongji University, Shanghai, China, in 2017. She is currently working toward the Ph.D. degree in photogrammetry and remote sensing with the Institute of Remote Sensing and GIS, Peking University, Beijing, China.

Her research interests include land cover classification and change detection, urban remote sensing analysis, and very high resolution image analysis.

Peijun Li (Senior Member, IEEE) received the Ph.D. degree in geology from the Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China, in 1995.

He is currently a Professor with Peking University, Beijing, China. He was a Visiting Researcher/Professor with several institutions, including Seoul National University, Seoul, South Korea; University of Bonn, Bonn, Germany; and Columbia University, New York, NY, USA. His research interests include land cover classification and change detection, very high resolution image analysis, urban remote sensing analysis, and geologic remote sensing.

Dr. Li was a Co-Chair of ISPRS Working Group VII/4 "Methods for Land Cover Classification" from 2008 to 2016. He is currently a member of Editorial Board of *ISPRS Journal of Photogrammetry and Remote Sensing*.

Tao Cheng (Senior Member, IEEE) received the B.S. degree in geographic information system from Lanzhou University, Lanzhou, China, in 2003, the M.E. degree in photogrammetry and remote sensing from Peking University, Beijing, China, in 2006, and the Ph.D. degree in earth and atmospheric sciences from the University of Alberta, Edmonton, AB, Canada, in 2010.

From January 2011 to December 2013, he was a Postdoctoral Scholar with the Department of Land, Air and Water Resources, University of California, Davis, CA, USA. Since December 2013, he has been a Professor with the National Engineering and Technology Center for Information and College of Agriculture, Nanjing Agricultural University, Nanjing, China. He is the author of more than 50 papers on international and national journals, one book chapter, and holds two patents. He is the Principal Investigator of several national level projects funded by the National Key R&D Program of China and the National Natural Science Foundation of China. His research interests include crop monitoring, crop type and management mapping, plant disease detection, crop phenotyping, imaging and nonimaging spectroscopy, and remote sensing for vegetation and agricultural applications.

Dr. Cheng was a Guest Editor of Special Issues in *Remote Sensing* and the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, and has been serving on the editorial board of the *ISPRS International Journal of Geo-Information* since 2015. He was the Chair of IEEE Geoscience and Remote Sensing Society Nanjing Chapter from 2016 to 2020. He was the recipient of the Youth Science and Technology Award from the Crop Science Society of China in 2019.