

Change Detection From Very-High-Spatial-Resolution Optical Remote Sensing Images

Methods, applications, and future directions

DAWEI WEN, XIN HUANG, FRANCESCA BOVOLO, JIAYI LI, XINLI KE, ANLU ZHANG, AND JÓN ATLI BENEDIKTSSON

XXXXX

Change detection is a vibrant area of research in remote sensing. Thanks to increases in the spatial resolution of remote sensing images, subtle changes at a finer geometrical scale can now be effectively detected. However, change detection from very-high-spatial-resolution (VHR) (≤ 5 m) remote sensing images is challenging due to limited spectral information, spectral variability, geometric distortion, and information loss. To address these challenges, many change detection algorithms have been developed. However, a comprehensive review of change detection in VHR images is lacking in the existing literature. This review aims to fill the gap and mainly includes three aspects: methods, applications, and future directions.

BACKGROUND

Change detection is a vibrant area of research with wide-ranging applicability, including damage assessment, land management, and environment monitoring. Due to the revisit property of Earth observation sensors, multitemporal remote sensing images at a large geographical scale can be acquired easily and conveniently. Due to their extensive availability, optical images become the main data sources for change detection [1]. Since these satellite sensors are able to acquire images with meter and submeter spatial resolutions, ground objects in fine spatial detail can be investigated [2]. Subtle change detection using these VHR images has drawn great interest in both the academic and industrial communities. However, multitemporal VHR

images exhibit unique properties, such as limited spectral information, intrinsic spectral variability, spatial displacement, and information loss, that limit the usefulness of traditional change detection methods. Therefore, a great number of studies have been carried out on VHR change detection, and a series of new research topics has emerged along with advances in remote sensing technology and data computing methods. In this regard, a timely overview of VHR change detection is required to summarize the new techniques and applications.

Although a number of reviews about change detection using remote sensing data [3]–[10] exist in the literature, the publications discuss general change detection methods and do not focus on high-spatial-resolution images. Only a few available works involve VHR images, e.g., the reviews in [6] and [7]. However, those two works concern object-based change detection methods for VHR data, neglecting other aspects, e.g., recent technological advances in deep learning and multiview and 3D change detection. Moreover, specific applications of VHR change detection have rarely been summarized and discussed in the currently available literature. Therefore, a comprehensive review of change detection from VHR remote sensing images, including methods, applications, and future directions, is presented (Figure 1).

ISSUES RELATED TO VHR IMAGES AND THEIR CHANGE DETECTION

With the ongoing development of remote sensing imaging techniques, an increasing number of VHR sensors are available, and many new sensors are being planned and launched [11]. New platforms, such as unmanned aerial

Digital Object Identifier 10.1109/MGRS.2021.3063465
Date of current version: 5 April 2021

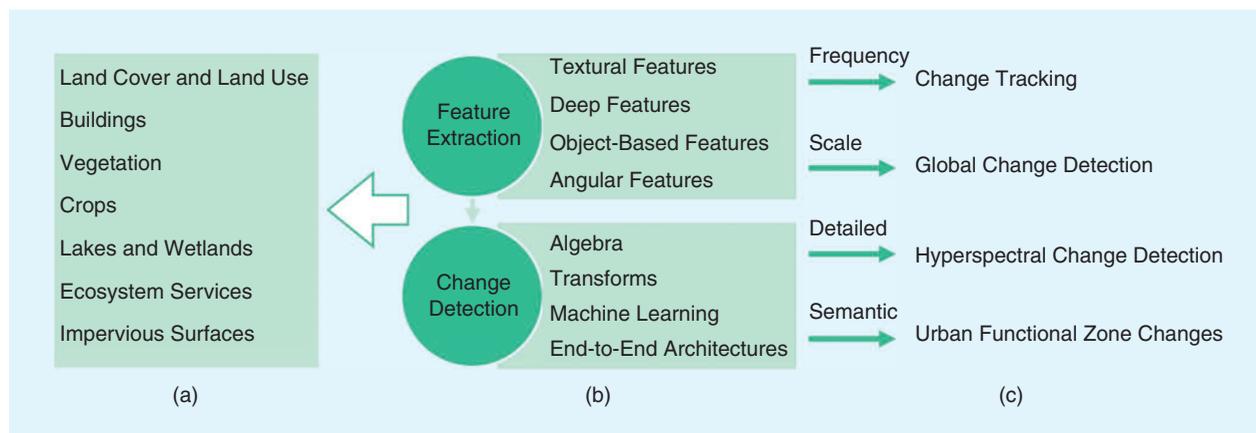


FIGURE 1. An outline of this review, including (a) applications, (b) methods, and (c) future directions.

vehicles (UAVs) and remotely piloted aircraft systems, have grown in popularity [12] and are now providing a large amount of VHR remote sensing data. As seen in Table 1, the imaging capabilities of VHR platforms and sensors are continually being improved with higher spatial resolutions, more spectral bands, and higher temporal revisit frequencies. In addition, most VHR sensors provide an along-track and across-track pair for stereo capture [12], [13]. With the improved capability of VHR remote sensing equipment, it is now becoming possible to achieve subtle, detailed, and frequent 3D change detection. Although change detection using VHR images is advantageous, from a technological point of view, it remains a challenge due to 1) limited spectral information, 2) intrinsic spectral variability, 3) spatial displacement, and 4) information loss, as discussed in the following.

- 1) *Limited spectral information:* Compared to coarse- and medium-resolution sensors, images captured by VHR sensors usually provide a smaller number of bands. Although WorldView-3, one of the most advanced VHR sensors, can provide images with 16 spectral bands, most VHR images, e.g., from IKONOS, QuickBird, WorldView-2, and Ziyuan-3, cover only four bands (blue, green, red, and near-infrared) [14]. With limited spectral information, it is difficult to separate classes that have similar spectral signatures because of the low between-class variance [15]–[18]. Researchers have also pointed out that it is difficult to achieve high-accuracy change detection with the limited spectral information [5], [15], [19]–[21] of VHR images. This may inhibit the direct use of traditional spectral-based change detection methods, e.g., change vector analysis (CVA) [22]. Therefore, other categories of features are often adopted to augment the spectral information for VHR change detection.
- 2) *Spectral variability:* There exists a high degree of spectral variability in VHR images. Buildings, for example, have complicated appearances, with various roof superstructures, such as chimneys, water tanks, and pipelines; this leads to significantly heterogeneous

TABLE 1. THE MAIN PARAMETERS OF SOME VHR SENSORS.

SENSOR	SPATIAL RESOLUTION (M)	NUMBER OF BANDS	REVISIT TIME (DAYS)	LAUNCH YEAR
IKONOS	1	Four	One to three	1991
QuickBird	0.61	Four	1.5–2.5	2001
SPOT-5	2.5	Four	26	2002
OrbView-3	1	Four	Three	2003
Cartosat-2	0.8	One	Four	2007
WorldView-1	0.5	One	1.7	2007
GeoEye-1	0.41	Four	Fewer than three	2008
WorldView-2	0.46	Four	1.1	2009
KOMPSAT-3	0.7	Four	Three	2012
Ziyuan-3	2.1	Four	Four to five	2012
SPOT-6/7	2	Four	One	2012/2014
Gaofen-1	2	Four	Fewer than four	2013
Gaofen-2	0.8	Four	Four	2014
Planet Labs	3	Four	One or two	2014
Deimos-2	1	Four	One or two	2014
WorldView-3	0.31	16	Fewer than one	2014
DMC-3	1	Four	One	2015
WorldView-4	0.31	Four	Fewer than one	2016

SPOT: Satellite Pour l’Observation de la Terre; KOMPSAT: Korean Multipurpose Satellite; DMC: Disaster Monitoring Constellation.

spectral characteristics in VHR images [23], [24]. High spectral variability within geographic objects increases the within-class variance, which inevitably leads to the uncertainty of spectral-based image interpretation methods. External factors, such as atmospheric conditions, phenological stages, sun angles, soil moisture, tidal stages, and water turbidity, may make unchanged objects temporally variant in their spectral features and hence result in them being incorrectly identified as changed ones [25], [26]. In addition, temporary objects, such as cars on a road, visible in VHR images

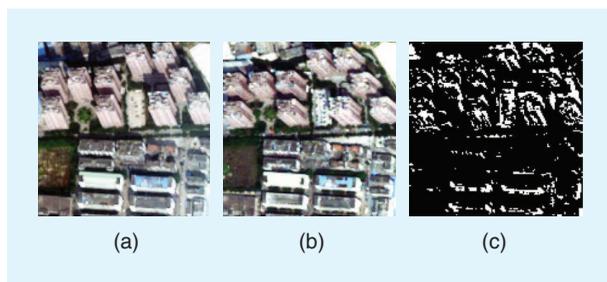


FIGURE 2. The spatial displacement in multispectral data acquired with different viewing geometries in an unchanged urban scene [21]: (a) Image (t_1), with a satellite angle zenith of 153° , and (b) image (t_2), with a satellite angle zenith of $129^\circ 12'$. (c) The result of traditional spectral-based CVA shows a high number of false alarms (black and white indicate unchanged and changed areas, respectively) [31].

can also affect the performance of traditional spectral-based change detection methods using VHR images.

- 3) *Spatial displacement*: The VHR imaging systems on optical satellites are highly agile platforms and can operate as constellations [27] that can support rapid retargeting, high revisit times (for instance, <1 day for WorldView-3 and WorldView-4), and stereoscopic coverage for rapid disaster response and 3D change detection [28]. However, this imaging mode makes it extremely difficult to acquire multitemporal images with the same or close viewing angles for accurate change detection [29], [30]. As such, multitemporal VHR images may suffer from apparent spatial displacement due to the parallax distortion of land cover objects, especially for high-rise buildings [31]. Specifically, a building may display distinct spatial morphologies (e.g., roofs and facades) in multitemporal VHR images due to different viewing angles (Figure 2). This may lead to a large number of commission errors if traditional spectral and pixel-based change detection methods are adopted. To solve such a problem, precise orthorectification using VHR digital surface models (DSMs) is a feasible solution. In particular, sensors equipped with multiview imaging systems, for instance, the three-line array of Ziyuan-3 and the two cameras of Cartosat-2, that can nearly simultaneously collect multiview images are preferred in similar atmospheric conditions for their stereo pairs and convenient collection of multitemporal data.
- 4) *Information loss*: VHR images suffer from serious information loss owing to the presence of clouds/haze, cloud shadows, and shadows cast by terrain, buildings, and trees. The problem of cloud and cloud shadow contamination can be avoided by selecting cloud-free observations [32]. However, shadows cast by terrain, buildings, and trees seem unavoidable in VHR imagery, especially in urban areas [33]. Although shadow information is useful in building detection and height estimation [34]–[36], it becomes a problem for change detection in wider

areas [37]. Since the direction and length of shadows are dependent on the sun's azimuth and elevation angle at the time of image acquisition, shadow-affected areas are different in multitemporal images. Besides, in the case of occlusions by vertical structures (e.g., high-rise buildings and trees), the problem of information loss can be more complicated. With different viewing geometries in multitemporal images, the size and direction of the tilting effect can vary, as shown in Figure 2. Overall, the regions affected by shadow and occlusions may become invisible and different in multitemporal VHR images.

METHODS

Change detection methods for VHR images are commonly based on two steps: 1) feature extraction and 2) change detection (see Figure 1).

FEATURE EXTRACTION

Change detection methods rely on effective multitemporal feature representation to indicate whether and what changes have occurred. It has been agreed that spectral-based methods become ineffective in dealing with the challenges facing VHR change detection. During the past decades, a large number of image features have been extracted, which can compensate for the limited spectral information contained in VHR images and improve the discriminative capability of image change information. In this review, image features designed for VHR change detection are divided into the following categories: textural, deep, object based, and angular (Figure 3). These are potentially useful for dealing with the challenges of limited spectral information and intrinsic spectral variability. A summary of the major features used for VHR change detection, including categories, subcategories, descriptions, characteristics, most-used sensors, and corresponding references, is presented in Table 2.

TEXTURAL FEATURES

Textural features depict contextual and structural information by using a moving window or kernel, where the parameters of size, direction, and distance must be appropriately determined [5], [38]. Textural features for VHR change detection can be categorized as statistical, structural, model based, and transform based. Statistical textures describe the relationships between the gray levels of local windows, e.g., the gray-level cooccurrence matrix (GLCM); local binary patterns (LBPs); and pixel shape index (PSI). The GLCM, the most popular statistical texture, measures the contrast (e.g., dissimilarity and homogeneity), orderliness (e.g., the angular second moment and entropy), and statistical (e.g., the mean, variance, and correlation) attributes within local windows [39], [40]. The LBP, an ordered set of binary comparisons of pixel values between the central pixel and its neighboring ones, is invariant to monotonic grayscale change [41]. The PSI aims to measure the length of direction lines, which are extended based on gray-level similarity along a series of directions [42]. Some representative

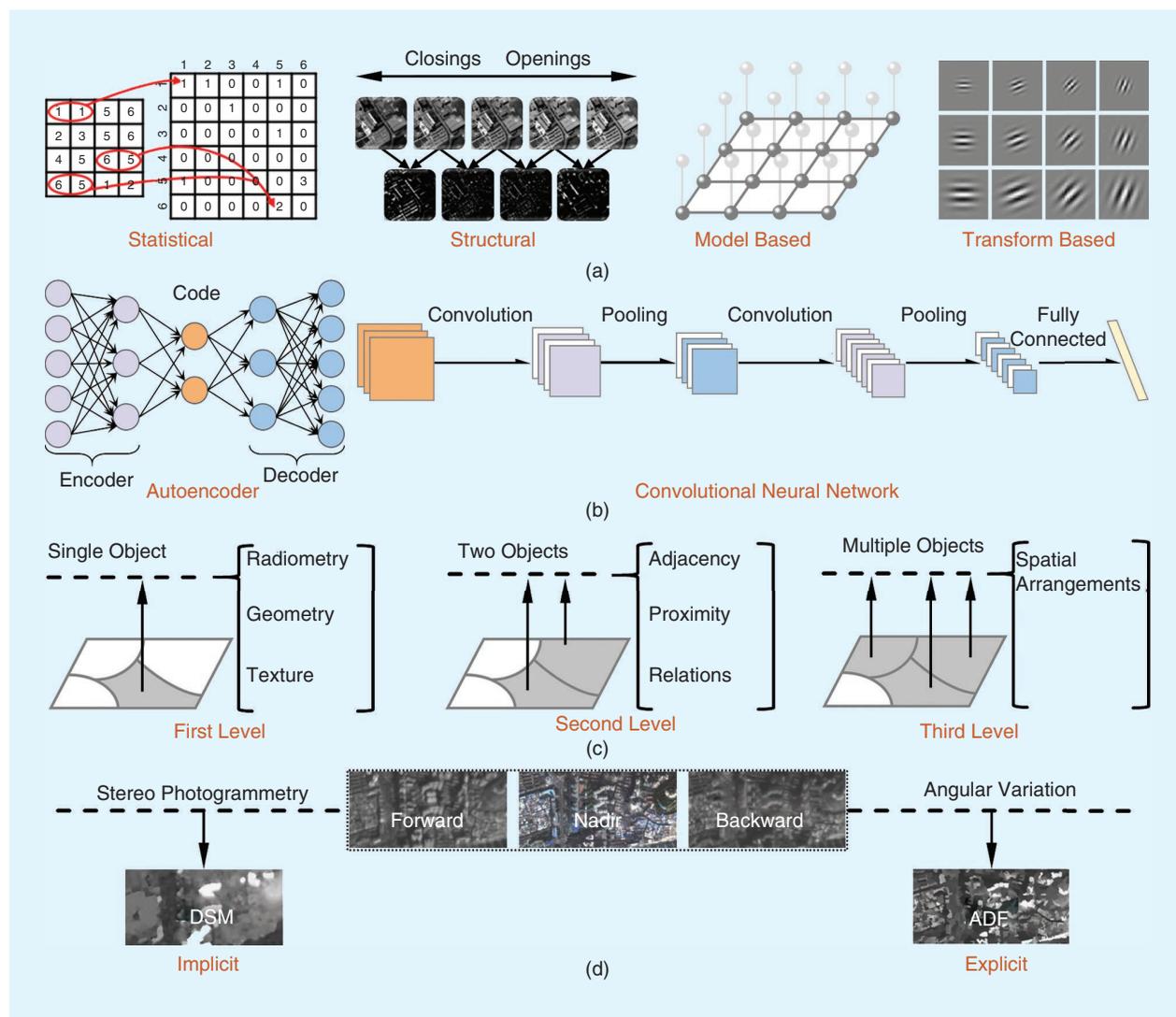


FIGURE 3. Features for change detection using high-spatial-resolution remote sensing images. (a) Textural features. (b) Deep features. (c) Object-based features. (d) Angular features. ADF: angular difference feature.

examples for VHR change detection using statistical textures are briefly introduced in the following.

Tan et al. [43] adopted the GLCM in an automatic change detection method to consider the variation information of direction, distance, and amplitude in images. Li et al. [44] applied the local similarity of GLCM textures to detect changes and demonstrated that this kind of feature was robust against both noise and spectral similarity. Peng and Zhang [45] used the LBP for change detection from *Gaofen-1* imagery, and both qualitative and quantitative analyses demonstrated the effectiveness of the proposed approach. Zhang et al. [46] identified building change types, i.e., new construction, demolition, and reconstruction, by using LBP features and obtained satisfactory change detection results with a high detection accuracy and precise structure boundaries. Liu et al. [47] proposed a line-constrained shape feature, a modified version of the PSI, for building change detection, and the results showed the approach's advantage in individual building change detection in a lightly populated region.

Structural textures, e.g., morphological profiles (MPs) and attribute profiles (APs), facilitate the investigation of the geometries, shapes, and edges of regions, with the convex and concave components being erased so that the geometric information of relevant structures is preserved and unimportant details are attenuated [48], [49]. MPs and APs have proved to be effective in VHR change detection since they can simplify results and reduce noise components (e.g., spectral variations) [48], [49]. For instance, Liu et al. [50] took the geometrical structure of change targets into account using MPs. In addition, the morphological building index (MBI) [36], which is defined as differential MPs with linear structural elements, has been extensively used in VHR change detection in urban areas since it can highlight bright and high-contrast structures, mostly consisting of buildings, in remote sensing images. For example, Huang et al. [51] proposed an automatic building change detection framework based on the MBI. Experimental results showed that the proposed method

TABLE 2. A SUMMARY OF THE FEATURES USED FOR VHR IMAGE CHANGE DETECTION.

CATEGORY	SUBCATEGORY	DESCRIPTION	CHARACTERISTICS	SENSOR	REFERENCES
Textural features	Statistical	Describe the relationships among the gray levels of local windows	Edge effect, difficulty of identifying parameters	QuickBird [48]–[53]	[43]–[47]
	Structural	Investigate the geometry, shapes, and edges of regions			[48]–[53]
	Model based	Obtain coefficients from the model describing the relationships among the local image neighborhood			[56]–[61]
Deep features	Transform based	Capture local structures in a transformed space	Complex training and parameter tuning, “black-box” nature, high computational burden, overfitting, and so on	Gaofen-2 [66], [77] and Google Earth images [66], [76]	[63], [64]
	Autoencoders	Learn efficient encoding through the optimization of a series of criteria			[70]–[73]
Object-based features	Convolutional neural networks	Extract mid- and high-level abstract features by interleaving convolutional and pooling layers	Determination of appropriate segmentation parameters and uncertainties of the segmentation results	QuickBird [88], [89]	[66], [67], [75]–[78]
	First level	Radiometry, geometry, and texture for each image object			[85], [88], [89]
	Second level	Relationships between two image objects, e.g., adjacency and proximity, and relationships with neighboring objects			[91], [92]
Angular features	Third level	Spatial arrangements of multiple objects	Availability of multiangle images	Ziyuan-3 [2], [21]	[95]
	Implicit	Orthographic images and DSMs			[21], [98], [99]
	Explicit	Quantify the differences contained in multiangle images, such as angular difference features			[2]

outperformed supervised classification via a support vector machine (SVM). In addition, point and line features, for instance, Harris [52] and scale-invariant feature transforms (SIFTs) [53], can improve the discriminability of man-made objects, such as buildings, roads, and cars, by describing corners and edges, therefore improving results.

Model-based textures, e.g., Markov random fields (MRFs) and fractal models, aim to represent textures through stochastic processes [54]. MRF models present spatial context through a graph-based image representation, where the nodes and edges of the graph express pixels and their relationship with connected nodes, respectively. Fractal models can depict texture roughness and complexity by capturing self-similar and self-affine patterns [55]. A number of MRF-based methods have been proposed to deal with VHR image change detection [56]–[60] because of their ability to describe local spatial relationships. Specifically, Bruzzone and Prieto [57] introduced a change detection method based on an MRF to model prior class probabilities by interpixel dependence, which increased the accuracy and reliability of the change detection results. In [60], spatial constraints between neighboring samples were formulated using an MRF in an active learning process for change detection. Multifractal features were applied to change detection in [61], and experiments on a complex landscape that included urban areas, agricultural fields, trees, and an unregulated river indicated that the features were tolerant to some degree to multitemporal differences caused by the viewing geometry and illumination angles.

Transform-based textures, e.g., Gabor, wavelets, and contourlets (CTs), aim to convert images into a new space

to capture local structures corresponding to scale, localization, and orientation [62]. For example, Li et al. [63] used a Gabor-based approach to improve the change detection performance since the technique can capture contextual information at different scales and orientations. Wei et al. [64] introduced wavelet pyramid decomposition features to VHR change detection. Thus, in VHR images, the complexity of homogeneous regions can be reduced in low-scale features, and details and edge information can be retained in high-scale ones [64]. In a comparative study conducted by Li et al. [65], a number of representative textural features were selected for change detection using VHR images, and it was shown that texture-based change detection methods can obtain better performance than spectral-based pixel ones. Texture change detection results are demonstrated in Figure 4, and it can be seen that, compared to using individual textures, combining multiple textures can improve change detection accuracy.

DEEP FEATURES

Deep feature representation based on the layer-wise learning of image patterns is a very promising research direction for change detection in VHR images [66], [67]. Differing from traditional handcrafted features, higher-level abstractions (both linear and nonlinear features) can be automatically extracted and optimized by multilayer neural networks, which can retain crucial variations and discard uncorrelated differences for change detection tasks [68]. In recent years, many deep learning methods have been developed, such as autoencoder (AE) models and convolutional neural networks (CNNs), for deep feature extraction in change detection with VHR images.

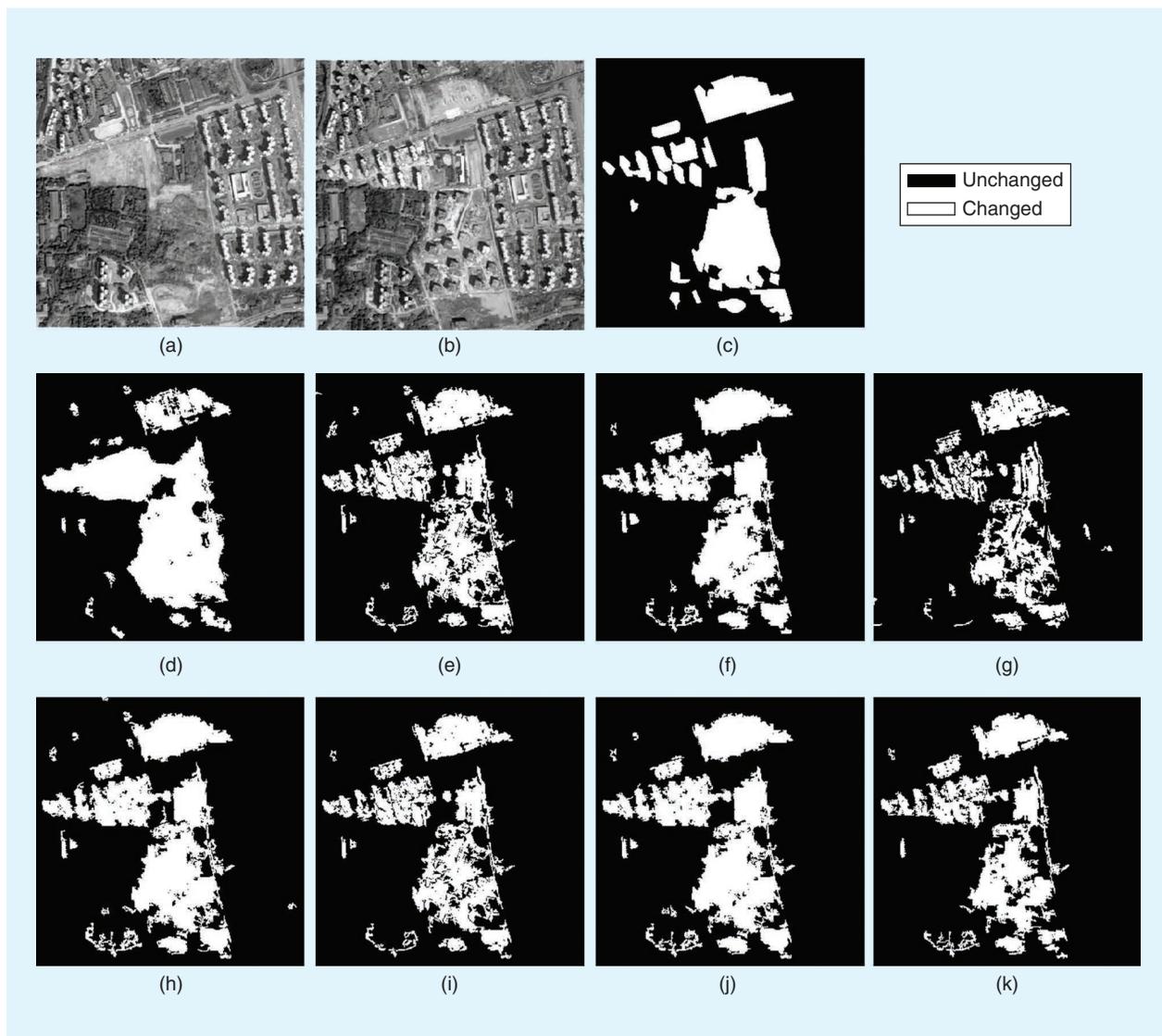


FIGURE 4. Change detection results based on textures: (a) image (t_1), (b) image (t_2), (c) the reference change map, (d) the GLCM, (e) APs, (f) a 2D wavelet transform (WT), (g) a fractal, (h) a fuzzy set (APs plus a 2D WT plus a 3D WT), (i) a fuzzy set (all textures), (j) a random forest (APs plus a 2D WT plus a 3D WT), and (k) a random forest (all textures) [65].

The AE is an unsupervised feature learning model that is constructed by minimizing the reconstruction error. However, it may learn a useless feature representation, such as a simple copy of the input [69]. To overcome that issue, variant models, e.g., the denoising AE (DAE) [70], sparse AE (SAE) [71], and Fisher AE (FAE) [72], have been employed for VHR change detection, with denoising, sparsity, and Fisher discriminant criteria, respectively. Specifically, a stacked DAE was used to learn high-level features from the local neighborhood [70]. In [70], it was found that the filters learned by a stacked DAE have a stronger representation capability than existing explicit ones. Based on the SAE, Su et al. [71] transformed a difference image into a suitable feature space for suppressing noise and extracting key change information in the change detection framework. Liu et al. [72] used the FAE for unsupervised layer-wise feature learning and showed that the model can generate more discriminative features than

the original AE. In addition to unsupervised feature learning through the optimization of certain criterions, AE-based models can learn effective features in a supervised way by considering label consistency, e.g., the contractive AE [73].

It is well recognized that CNNs are effective in extracting mid- and high-level abstract features by interleaving convolutional and pooling layers [74]. According to the feature learning strategy, CNNs can be categorized as unsupervised [67], [75], [76], supervised [77], fine-tuning [66], and transfer learning based [78]. For example, Zhan et al. [75] used a pretrained CNN to automatically extract deep spatial-spectral features for change detection in VHR satellite images. Saha et al. [67] developed unsupervised deep CVA for change detection, and a network trained on remote sensing aerial images for semantic labeling by Volpi and Tuia [79] was adopted for deep feature extraction. As detailed in Figure 5, the experimental results demonstrated

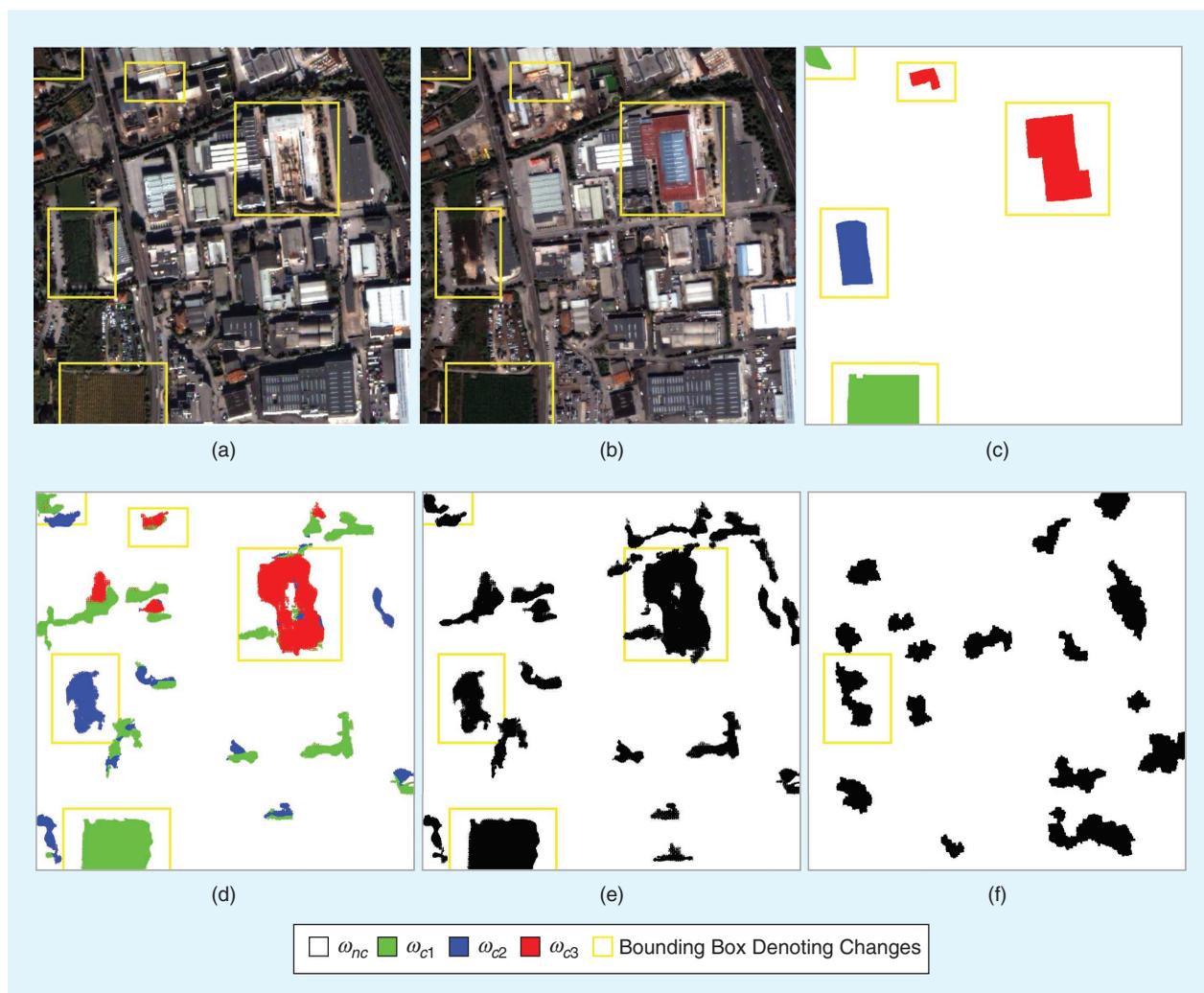


FIGURE 5. Change detection results for *QuickBird* bi-temporal images: (a) image (t_1), (b) image (t_2), (c) the reference change map, (d) multi-class deep CVA, (e) binary change deep CVA, and (f) object-based CVA [67].

that, compared to object-based methods, deep features are effective for capturing change information and are promising for distinguishing multiclass change information. Wang et al. [77] trained a model through manually selected samples, where the parameters of the shared convolutional layers were initialized by the pretrained ResNet-50 model, and the others were randomly initialized. Hou et al. [66] chose to extract CNN-based deep features through a fine-tuned Visual Geometry Group (VGG)-16 by transferring a model pretrained on large-scale natural images to the remote sensing domain via an aerial image data set. Liu et al. [78] proposed a CNN-based transfer learning method for change detection. In particular, the loss function was designed by combining high-level features extracted from a pretrained model (i.e., the U-net model trained on an open source data set) and semantic information contained in change detection data sets.

Notably, deep learning methods depend on an enormous amount of training data, which may not be available for multitemporal VHR remote sensing imagery [74].

Meanwhile, great differences in spectral properties and image contexts among natural red–green–blue (RGB) images and remote sensing data result in deep features extracted by fine-tuned models that do not fully represent the essential characteristics of remote sensing images. As a result, the contrast between a small number of remote sensing data sets and a large number of natural images during model learning may hamper the further improvement of VHR change detection using deep features. In recent years, large multitemporal data sets have been released, such as 86 image pairs from the DigitalGlobe satellite constellation (i.e., *QuickBird*, *WorldView-1*, *WorldView-2*, and *GeoEye-1*) [80], 291 pairs of multitemporal aerial images [81], and more than 700,000 labeled instances for building damage assessment [82]. It can be anticipated that more and larger multitemporal VHR remote sensing data sets with diverse image characteristics and various acquisition conditions will appear in the near future. In this case, the essential change features for VHR remote sensing images can be effectively extracted by a deep network specialized for multitemporal remote sensing data.

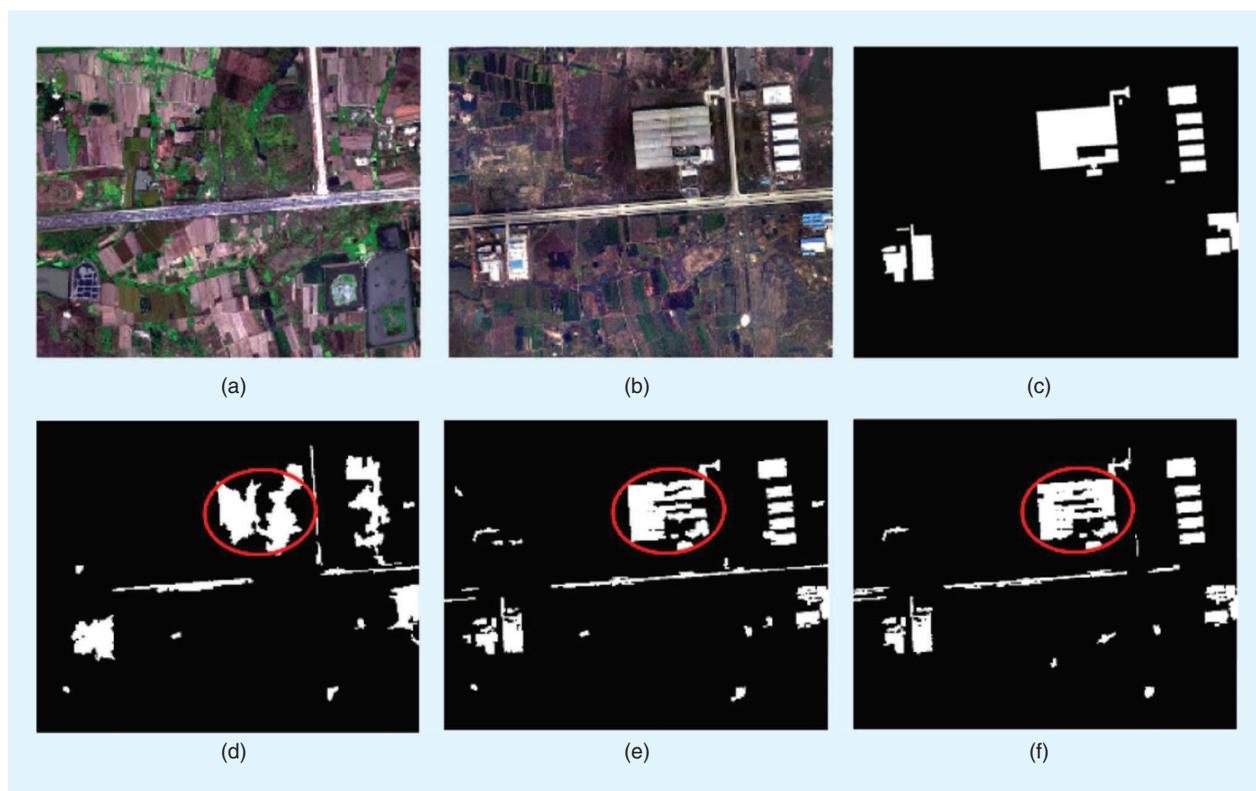


FIGURE 6. Object-based CVA results from different multitemporal segmentation strategies: (a) image (t_1), (b) image (t_2), (c) the reference change map, (d) the segmentation of image(t_1), (e) the segmentation of stacked multitemporal images, and (f) the separate segmentation of each monotemporal image [85].

OBJECT-BASED FEATURES

Object-based features refer to spectral, geometry, texture, extent, and contextual information at the object scale rather than single pixels and groups of pixels within a kernel filter/moving window. In this way, an image object is viewed as the processing unit for change detection. An object is a set of spatially adjacent pixels that are spectrally similar and that can be extracted through image segmentation. Overall, object-based features are effective in VHR change detection since they mitigate radiometric differences, spectral variability, and misregistration errors [38], [83]. However, appropriate segmentation parameters, which are often dependent on subjective and laborious trial-and-error experiments, need to be determined [84]. Furthermore, shortcomings and problems in different multitemporal image segmentation strategies, e.g., 1) the segmentation of only one monotemporal image, 2) the segmentation of stacked multitemporal images, and 3) the independent segmentation of multitemporal images, should be carefully considered and tackled [5], [85]. Specifically, geometric changes (e.g., the size and shape) cannot be captured by 1) and 2) [85]. Moreover, strategy 2) may also result in “sliver objects” caused by image misregistration. As for strategy 3), spatial correspondence between multitemporal objects needs to be established. Object-based CVA results [85] derived from different multitemporal segmentation strategies are presented in Figure 6, where it can be observed

that different multitemporal segmentation strategies can significantly affect change detection results.

Generally speaking, three levels of object-based features can be used for change detection [86]. In the first, the object-based features include the radiometry, geometry, and texture for each image object [87]. For instance, in [88], key points of each object are extracted in change detection, which was successfully applied in three landslide scenes and one view that examined land use changes. Bovolo [89] computed the mean values of texture measures in separate parcels for change detection, and better accuracy with high fidelity in the homogeneous and border regions was achieved by the object-based method than with the pixel-based one. However, in these studies, texture is still extracted in a pixel-based manner and depends on the size of a moving window (or kernel). More importantly, kernel- and window-based texture can create between-class texture, leading to an edge effect [87]. Therefore, object-oriented texture computed within the boundary of an object is recommended, such as object-wise GLCM texture measures [87] and object-based MPs [90].

The second-level object-based features exploit relationships between two image objects, e.g., adjacency, proximity, and relations between neighboring objects [87]. For example, Liang et al. [91] considered the relations of neighboring objects in feature extraction for object-oriented change detection. Yu et al. [92] combined a relative border with a

“forest with no change” and the normalized difference vegetation index (NDVI) to identify the category of “change from forest to developed land.” The third-level features refer to spatial arrangements among multiple objects [87]. Third-level object-based features have been used in image classification, such as urban functional zone extraction [93] and urban village detection [94]. Nevertheless, such features have rarely been used in VHR image change detection. In [95], spatial dependency and sharing boundaries among multiple objects are considered to reduce spurious errors caused by shadow in urban vegetation change detection.

ANGULAR FEATURES

Multiangle satellite images can be acquired by *WorldView-2*, *IKONOS*, *Cartosat-1*, and *Ziyuan-3* through across-track and along-track stereoscopy [96]. Spatial and spectral variations encoded in multiangle images can be extracted as new information sources for change detection. To be specific, multiangle observations can capture information about bidirectional reflectance signatures and vertical structures (e.g., trees and buildings) and hence complement conventional spectral and spatial features [27]. In this article, angular features are categorized as 1) implicit ones that are generated by stereo photogrammetry, such as orthographic images and DSMs, and 2) explicit ones that capture angular variations, such as angular difference features [97].

Most existing change detection studies based on multiangle VHR imagery adopt implicit angular features. For example, Chaabouni-Chouayakh et al. [98] presented a fully automatic change detection method for urban monitoring using *IKONOS* stereo data, and their experimental results verified the effectiveness of the joint use of multispectral and DSM features. Tian et al. [99] investigated building and forest change detection using panchromatic *Cartosat-1* stereo imagery, and they found that extracted height values from DSMs can greatly improve change detection accuracy. Huang et al. [21] used photogrammetrically derived orthographic images from multiangle *Ziyuan-3* data to monitor subtle changes across urban areas, and it was shown that the use of orthographic images can minimize the influence of spatial inconsistency among multitemporal data, e.g., misregistration and parallax distortion for high-rise buildings.

On the other hand, explicit angular features aim at describing the differences contained in multiangle images, e.g., the angular difference feature [100], multangular built-up index (MABI) [101], multiangle spectral variation feature [27], stacked multiangle spectral feature [102], and bidirectional reflectance distribution function-based index [103]. Benefiting from these explicit angular features, detailed urban and vegetation classifications were achieved using multiangle VHR images. Nevertheless, in the current literature, the previously mentioned explicit angular features have seldom been employed for change detection. One exception is a recent study presented in [2]. In it, the MABI, which indicates spectral and structural variations in multiview images, was used. Specifically, Huang et al. [2] integrated planar (i.e.,

MBI, Harris, and PanTex) and vertical [multispectral image (MSI), normalized DSM (nDSM), and MABI] features to detect newly constructed buildings and identify their change timing by using time-series, multiview *Ziyuan-3* imagery. Figure 7 gives an example of change results from different feature combinations. It shows that the joint use of planar and vertical features can generate more accurate results in terms of change extents and timings.

To better evaluate the different kinds of features, we create a *Ziyuan-3* multiview change detection (MVCD) data set, which is available at http://irsip.whu.edu.cn/resources/resources_en_v2.php. It includes both urban and rural scenes with diverse and complex change types, and, moreover, it considers seasonal and illumination influences. These characteristics enable the MVCD to function as a challenging change detection data set. A comparative analysis between different attributes, including the GLCM [39], AP [49], CT [62], MABI, [101], object-wise GLCM (GLCM-Obj) [87], and deep features [67], has been carried out. Specifically, the change intensity map was obtained by CVA, and the threshold for each feature was determined based on receiver operating characteristic curves to achieve a balance between commission and omission errors [65]. Qualitative and quantitative experimental results are provided in Figure 8 and Table 3, respectively. The spectral feature fails to detect changes between spectrally similar classes (e.g., bare soil and buildings), and unchanged objects with spectral variation are incorrectly detected as changed ones. The GLCM, AP, and CT can depict textural changes, e.g., the spatial distribution of the gray value, geometry, and local details. Among them, the CT gives more complete changed regions, and the AP produces more false alarms. The MABI emphasizes building changes, but it is not sensitive to other variations (e.g., soil, vegetation, and roads), which therefore leads to a large omission error. The GLCM-Obj generates smoother results with smaller omissions but larger commission errors than its pixel-wise version. Deep CVA outperforms the other methods, but false alarms caused by shadows and seasonal effects can be still observed.

CHANGE DETECTORS

VHR change detectors can be categorized as algebra-, transform-, and machine learning-based indicators. CVA is one of the most widely used algebraic approaches, and it is carried out by measuring the difference among bi-temporal multifeature vectors to derive a change vector for VHR images [67], [104], [105]. Transform-based methods, such as principal component analysis [106] and multivariate alteration detection [107], attempt to suppress no-change areas and emphasize change information in the transformed feature space. In the machine learning community, change detection is often viewed as a classification problem. In conventional classification-based VHR change detection, spectral-spatial feature extraction and detectors (e.g., SVMs [108] and the random forest [65]) are separately implemented. The recent hot spot, i.e., deep learning, can integrate these two

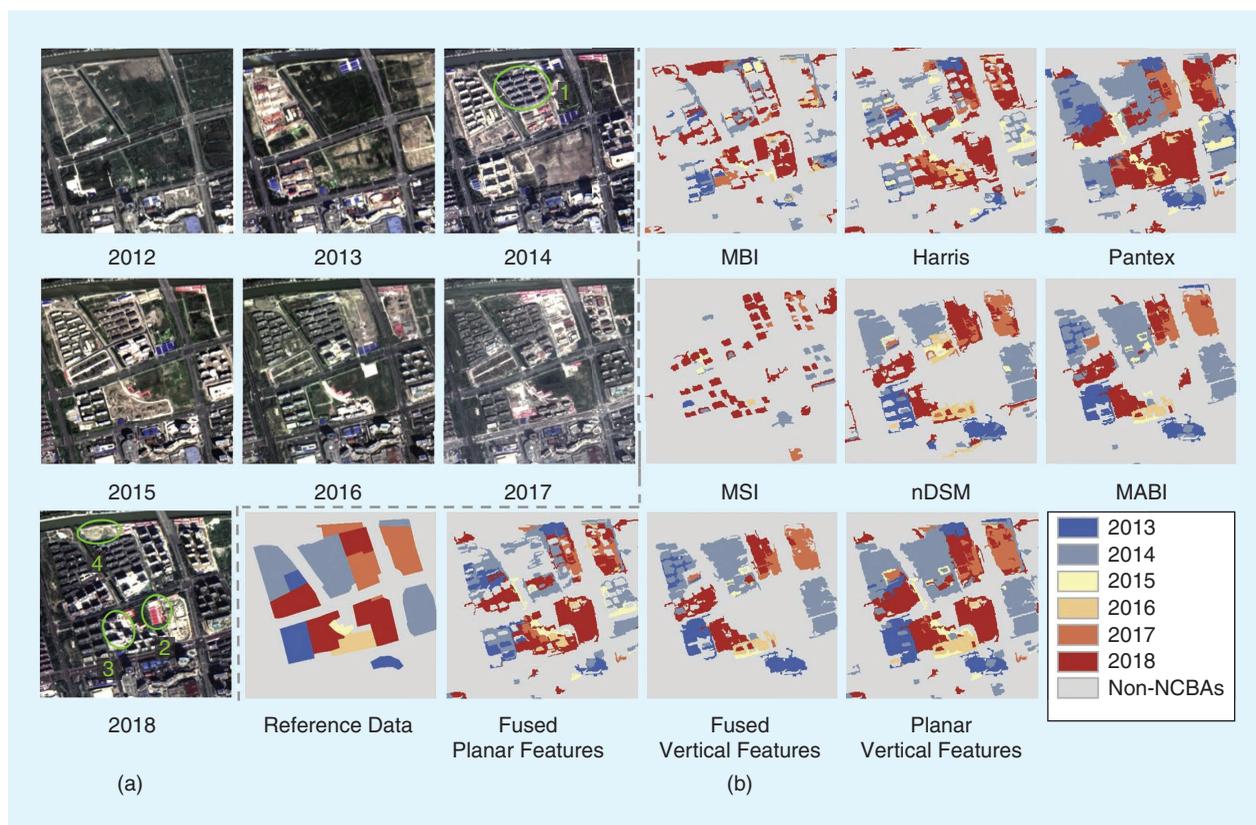


FIGURE 7. Experimental results for the automatic monitoring of newly constructed building areas (NCBAs) using planar (i.e., MBI, Harris, and Pantex) and vertical (MSI, nDSM, and MABI) features [2]. (a) Multitemporal *Ziyuan-3* images. (b) NCBAs and their change timing.

operations in a joint learning framework, which is therefore very promising for VHR change detection [109], [110].

Deep learning-based change detectors can be grouped in terms of different criteria, including learning and fusion strategies, network models, and processing units (Table 4). We first discuss learning strategies. On the basis of a large amount of annotated data, supervised deep learning methods can capture semantic changes, and hence they are sensitive to actual variations of interest and tolerant to “pseudo changes” (such as geometric deformation and radiation distortions caused by spatial displacement and phenology variation, respectively) [110]–[116]. However, it is difficult to learn a deep model only from the training samples of a study area since the proportion of the change area is usually very small. To tackle this problem, on the one hand, transfer learning [117] and meta-learning [118] are considered to leverage knowledge from other data sources. Transfer learning strategies focus on fine-tuning pretrained models that are designed for different but related tasks. Meta-learning can learn from data, and it can learn how to learn by utilizing previous experiences [119]. Regarding the huge difference between VHR remote sensing images and data from other fields (e.g., natural RGB images) in terms of the image modality, spectral bands, spatial resolution, viewing angle, and so on, large amounts of publicly available multitemporal VHR remote sensing data are required to construct a robust VHR deep change detector. On the other hand,

semisupervised deep learning methods, with the consideration of unlabeled samples [120], can relieve the burdensome labeling process, although the effects of unlabeled samples as well as the complexity of the semisupervised model should be further investigated.

With regard to the fusion strategy, according to how bi-temporal images are dealt with, deep learning-based change detectors can be classified as early fusion and late fusion. Early fusion methods concatenate multitemporal images as a whole input into a deep network [110]. Early fusion is able to capture the hierarchical difference representation, i.e., from low-level grayscale differences in shallow layers to high-level semantic changes in deep layers, while grayscale differences that are not relevant to semantic changes, e.g., spatial misalignment and the internal variability of objects, may propagate to deeper layers and therefore lead to false alarms [113]. In contrast to early fusion, late fusion methods separately learn monotemporal features and concatenate them later as an input to the change detection layers [121]. This kind of network architecture may lead to insufficient learning, e.g., during network training. Gradients in high layers are difficult to flow backward to lower ones [122] and hence affect the change detection performance. Thus, as an attempt in [113], early and late fusion networks were combined to complement one another.

As for network models, AE [123], [124], deep belief networks [125], CNNs [110], [112], [113], [115], [120], [126],

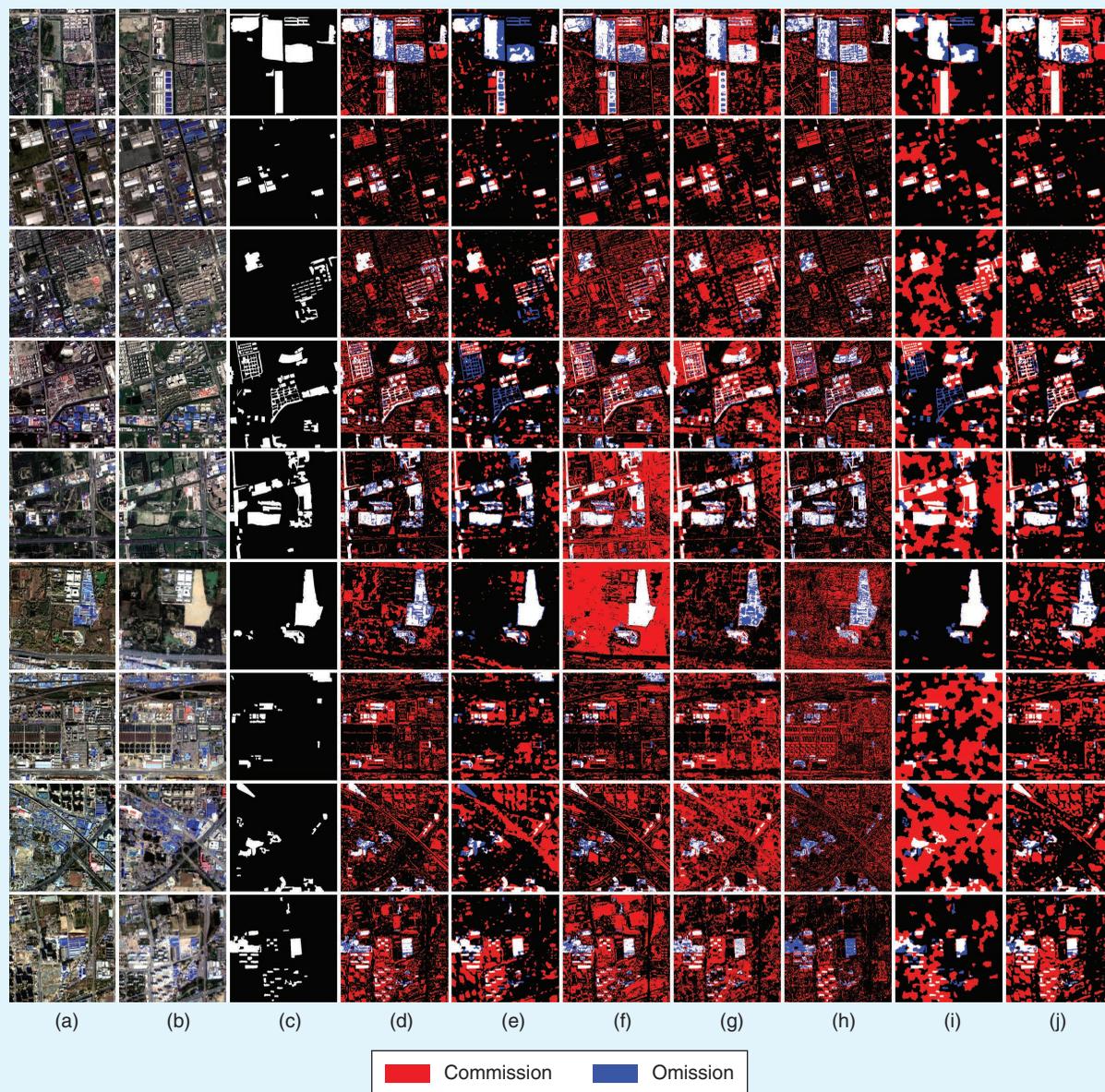


FIGURE 8. A comparison of different features for the MVCD data set: (a) image (t_1), (b) image (t_2), (c) the ground reference, (d) spectral features, (e) the GLCM, (f) APs, (g) CTs, (h) the MABI, (i) the GLCM-Obj, and (j) deep features.

TABLE 3. THE CONSIDERED METHODS' CHANGE DETECTION ACCURACY WHEN USING THE MVCD DATA (%).

METHOD	CORRECTNESS	COMMISSION ERROR	OMISSION ERROR	OVERALL ERROR
Spectral	70.88	24.51	29.12	26.62
GLCM	65.05	16.1	34.95	22.04
AP	71.75	40.2	28.25	33.18
CT	75.13	33.84	24.87	28.67
MABI	57.87	28.76	42.13	34.18
GLCM-Obj	74.51	30.47	25.49	27.76
Deep	79.98	25.46	20.02	22.41

recurrent neural networks (RNNs) [127]–[129], generative adversarial networks (GANs) [130], [131], and graph neural networks [132] have been adopted for end-to-end change detection. The CNN is one of the most widely used methods, and mainstream CNN architectures, such as AlexNet [133], VGGNet [134], GoogleNet [135], ResNet [136], and DenseNet [137] as well as their variants, have been considered [138]. RNNs with modules, such long short-term memory and gated recurrent units as well as their variants, are also widely employed to model the phenological process of multitemporal VHR images, due to the superiority of recurrent layers in processing sequential data and modeling time-series dependence. In addition, the U-net and

its variants, which are composed of an encoder to hierarchically extract semantic information and a counterpart decoder to delineate spatial details, can be viewed as AE architectures for VHR change detection. They receive much attention due to their ability to maintain change object spatial details.

Recently, some studies proposed hybrid models, such as those in [111] and [127]. For instance, as illustrated in Figure 9, a CNN and an RNN are combined in one end-to-end network to extract joint spectral-spatial-temporal features [111]. In [139], difference-based methods using edge-based level set evolution (ELSE), region-based level set evolution (RLSE), MRFs, and fully convolutional networks (FCNs) as well as postclassification-based methods with SVMs, CNNs, GANs, Siamese convolutional networks (SCNs), and end-to-end GAN-based Siamese frameworks (GSFs) are compared for landslide detection (Figure 10). Since observing landslides separately from unchanged and other changed regions is required, this kind of change detection is challenging. As can be seen, the four difference-based methods lead to more false alarms. As for the five postclassification methods, deep learning techniques generally outperformed SVMs, due to their explorative capabilities in representing related changes and suppressing irrelevant variations.

According to the processing unit, deep learning-based detectors are divided into patch- [116], [130], pixel- [110], [117], and object-based [127], [140] varieties. For a patch-based change detection task, a sliding window with a fixed size is used to divide the study area into a series of patches, and each patch is assigned a label by the detector. In this way, each pixel in the patch is assigned the same label. Consequently, rough location—not fine-grained—boundary-of-change information is obtained. However, patch-based change detection can reduce the influence of spatial misalignment to some extent in VHR change detection. Since patch-based deep learning networks view each patch as the change detection unit and encode each patch as a set of feature maps with coarser spatial resolutions, the spatial misalignment of these feature maps becomes smaller, and some errors of spatial alignment are therefore avoided in a change detection task. In other words, when regarding a patch as the change analysis unit, only a very large misalignment can cause an unchanged image patch to be identified as a changed one, and a small misalignment can be tolerated. Several important issues should be noticed for the patch-based method, such as the oversmoothing of results and the selection of the patch size.

The multiscale strategy [135] may be appropriate for addressing these issues, but it inevitably leads to larger computation burdens. Pixel-based methods usually employ semantic segmentation architectures to predict pixel-wise change detection results [33]. Specifically, in semantic segmentation architectures, after extracting abstract semantic information through multilayer encoding (e.g., convolution layers), a series of operations, e.g., interpolation,

TABLE 4. A SUMMARY OF DEEP LEARNING-BASED CHANGE DETECTORS.

CRITERIA	CATEGORY	DESCRIPTION	REFERENCES
Learning strategy	Supervised	Based on a large number of labeled samples	[110]–[116], [121], [124]–[130], [139], [140]
	Transfer learning	Fine-tunes pretrained models that are designed for different but related tasks	[117], [131]
	Meta-learning	Learns from little labeled data and learns how to learn	[118]
Fusion strategy	Semisupervised	Joint use of labeled and unlabeled data	[120], [132]
	Early fusion	Uses concatenated multi-temporal images as input	[110], [114], [115], [125]–[129], [131], [132], [139]
Network model	Late fusion	Learns monotemporal features separately and then concatenates them as a whole input	[111], [112], [116], [117], [121], [130], [140]
	CNN	Stacked convolutional, pooling, and fully connected layers	[110], [112]–[116], [120], [126]
	Recurrent neural network	Models with a recurrent hidden state, e.g., gated recurrent units and long short-term memory	[127]–[129]
Processing unit	AE	Reconstructs the input with an encoder–decoder structure	[123], [124]
	Deep belief network	Composed of layer-wise restricted Boltzmann machine	[125]
	Graph neural network	Learns graph structure, e.g., relationships between features of pixels/objects	[132]
Processing unit	Generative adversarial network	Generator and discriminator that are adversarially trained	[130], [131], [139]
	Patch	Assigns a label to each patch	[111], [115]–[117], [120], [121], [128]–[130]
	Pixel	Predicts change labels for each pixel	[110], [113], [114], [126], [131], [139]
Processing unit	Object	Incorporation of segments/superpixels	[124], [125], [127], [132], [140]

deconvolution, and upsampling, is used to progressively decode semantic information into feature maps that have the same spatial resolution as the input images. Unlike traditional pixel-based change detectors that suffer from misregistration, viewing angle differences, and occlusions, deep learning methods can predict pixel-wise change detection with a highly semantic abstraction of the spatial context. However, object boundaries are often blurred in the change detection results, as up-sampling layers reconstruct the appearance but not the shape of objects. To cope with this issue, better networks are designed. UNet++, for example, combines nested features to preserve change region boundaries, considering that shallow layers are better able to capture spatial details [110].

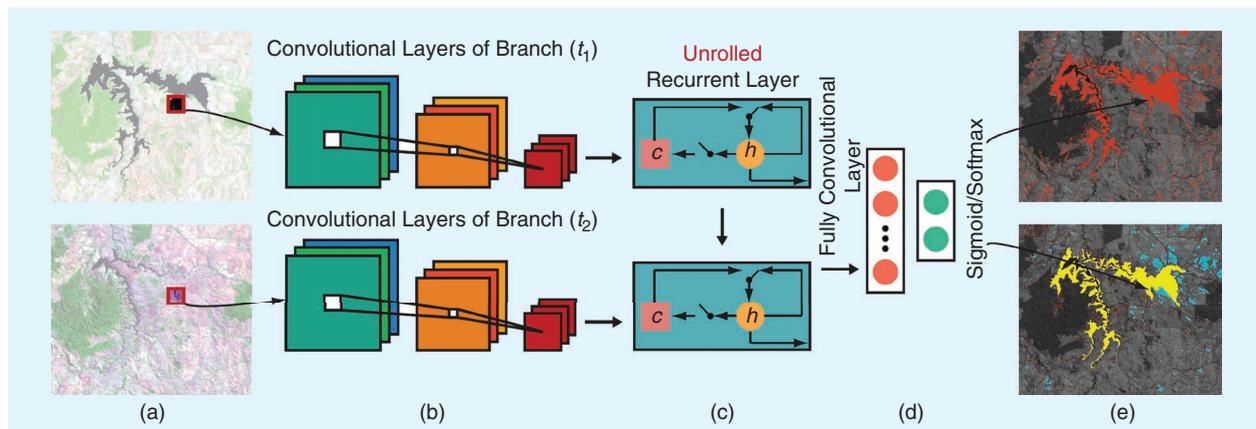


FIGURE 9. An end-to-end architecture composed of a CNN, RNN, and fully connected network for change detection [111]. (a) Image (t_1) (top) and image (t_2). (b) The convolutional subnetwork. (c) The recurrent subnetwork. (d) The fully convolutional layers. (e) The binary change detection (top) and multiclass change detection.

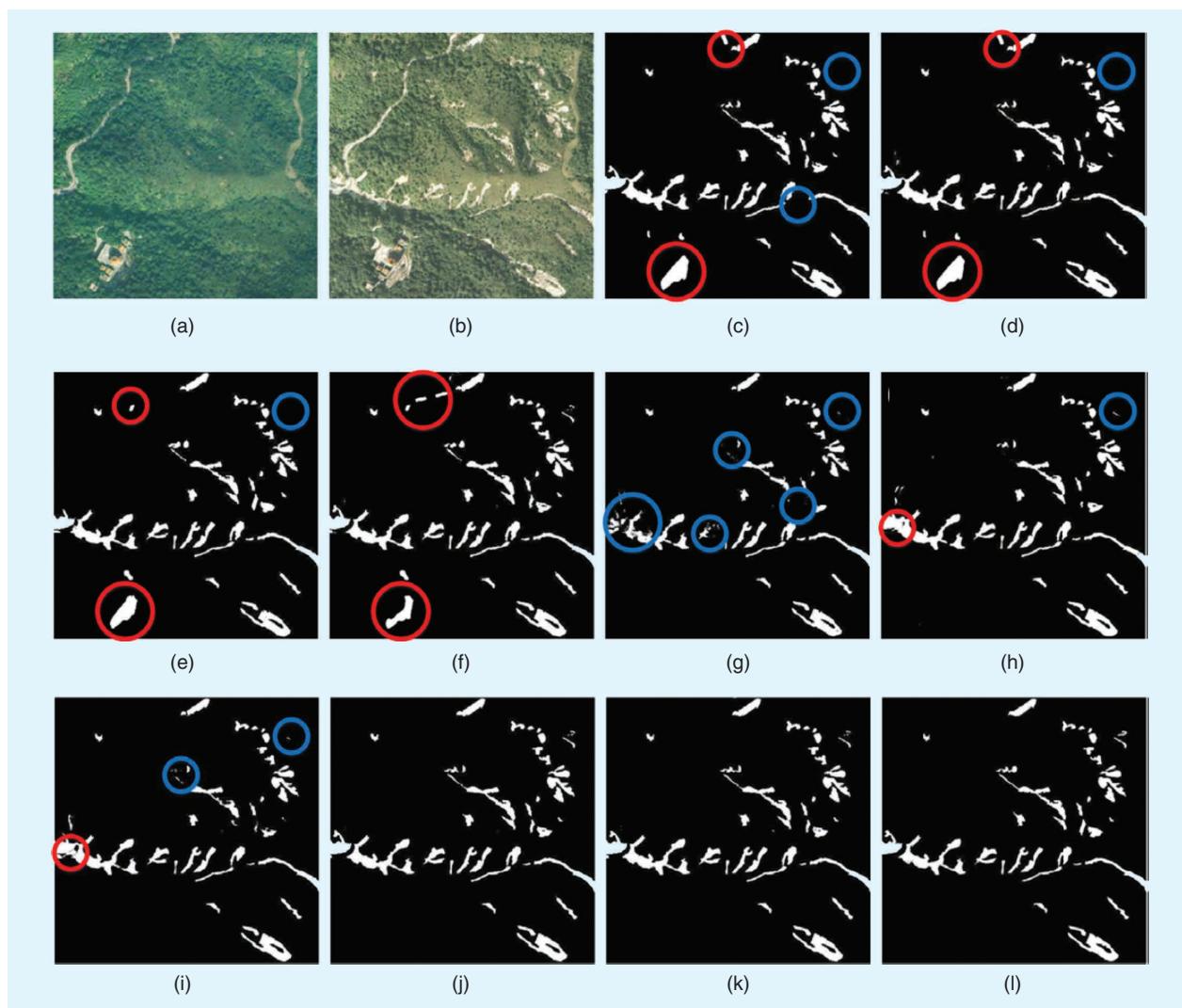


FIGURE 10. Landslide detection results from different methods: (a) image (t_1), (b) image (t_2), (c) ELSE, (d) RLSE, (e) an MRF, (f) an FCN, (g) an SVM, (h) a CNN, (i) a GAN, (j) an SCN, (k) a GSF, and (l) the ground truth. White and black indicate areas where landslides are detected and not detected, respectively. Red and blue circles represent landslide pixels that are wrongly detected and omitted [139].

Object-based deep learning methods are also considered for change detection [127], [140]. A simple approach is to adopt object-based segmentation in the pre/postprocessing step, as shown in [140]. On the other hand, object information can be also considered during the training process by adding object-wise loss terms [127]. However, issues related to conventional multitemporal image segmentation, such as oversegmentation, undersegmentation, and “sliver objects” caused by misregistration, remain unsolved. In the future, object-based detectors need to generate semantic segments and establish spatial correspondence between multitemporal segments.

The types of characteristics most often used for each criterion (i.e., the learning strategy, fusion strategy, network model, and processing unit) in VHR change detection are summarized in the following:

- 1) For the learning strategy, supervised learning is the most widely used method for VHR change detection. However, the great amount of labor required to collect a large number of training samples becomes a bottleneck, especially for deep network models, which leads to increasing attention for other learning strategies.
- 2) Late and early fusion strategies have their own strengths and weaknesses in representing multitemporal features and their differences, and hence hybrid fusion is sometimes chosen.
- 3) Among various network models, CNNs are the most commonly considered, and they are coupled with other networks, i.e., hybrid models, for instance, CNN-RNNs [111].
- 4) As for the processing unit, most studies consider patch- and pixel-level models. Patch-level detectors are more tolerant to spatial misalignment, but pixel-based ones are more appropriate for identifying fine-grained changes.

APPLICATIONS OF VHR CHANGE DETECTION

VHR image change detection is widely used in a large number of practical scenarios. A series of representative applications is the focus of this review, including the monitoring and change detection of 1) land cover and land use, 2) buildings, 3) vegetation, 4) crops, 5) lakes and wetlands, 6) ecosystem services, and 7) impervious surfaces.

LAND COVER AND LAND USE CHANGE DETECTION

Compared to coarse- and medium-resolution images, VHR images can reveal detailed and subtle intraurban change information [141]. Specifically, urban change detection by combining multiple features (e.g., object-based spectral, shape, and texture attributes) was presented in [142], where changes to detailed urban objects, e.g., buildings, roads, and playgrounds, can be detected. Huang et al. [21] identified pixel-level change transitions in 2012–2013 using *Ziyuan-3* orthographic images, and the experimental result is presented in Figure 11. It can be seen that, even in the one-year period, small-scale changes extensively occurred in the urban area of Wuhan, China. For instance, fine-scale urban land cover transitions caused by pond infilling, building

demolitions, building construction, weed growth, and site preparation can be observed. In [143], changes in detailed land cover classes, including bright roofs, gray roofs, tile roofs, brown fields, dark asphalt, light asphalt, and so on, were analyzed using *IKONOS* and *GeoEye-1* images.

As for land use change detection, Wu et al. [108] interpreted change transitions, e.g., from sparse housing to industrial areas, by combining spectral and SIFT features. In [144], land use maps of Shenzhen (a highly dynamic and developed megacity in China) were generated in 2005 and 2017 based on VHR satellite data. As demonstrated in Figure 12, detailed land use categories, including residential, commercial, industrial, infrastructure, grassland, farmland, woodland, water, breeding surfaces, and unused land, were monitored. In addition, the performance of different features, i.e., color histograms (CHs), LBPs, SIFTs, and deep features, were compared, and the best accuracies of 96.9% and 97.1% were obtained by the deep learning method [Figure 12(b)].

BUILDING CHANGE DETECTION

Buildings are one of the most dynamic artificial structures, and building change detection is important for urban development monitoring (e.g., building demolition and construction) and disaster management (e.g., building damage caused by natural hazards). Numerous methods for building change detection have been proposed [19], [51]–[53], [85], [145]–[157]. Some studies focus on multitemporal building observation and subsequent change analysis, where descriptors for building detection in VHR images are a critical issue. The descriptors can be categorized as template matching (e.g., the snake model) [158], knowledge based (e.g., shadow evidence and the MBI) [36], [159], and machine learning [148], [160]. For example, in [52], the MBI and the Harris detector were used to identify building areas, and then building change detection was conducted through interest point matching. Other types of methods directly explore changes in shapes, colors, and textural properties that are highly related to characteristics of buildings. For example, in [51], multitemporal variations in the MBI and spectral information were used to identify altered buildings. Likewise, in [85], the change feature generated by the MBI and spectral features was considered the indicator of building change. In [161], building changes were detected through the aggregation of spectral and textural features.

Figure 13 provides building change detection results from different methods, including SVMs based on MBI features (MBI-SVM), building interest point detection using the MBI and the Harris detector, MBI-based CVA (MBI-CVA), the fusion of the MBI and spectral and shape features, CVA using morphological features, and object-based CVA. It can be seen that automatic methods can achieve performance comparable to or better than supervised ones, i.e., the MBI-SVM [Figure 13(d)]. Meanwhile, the results of the MBI-CVA [Figure 13(g)] show more small false alarms. The fusion of the MBI and other

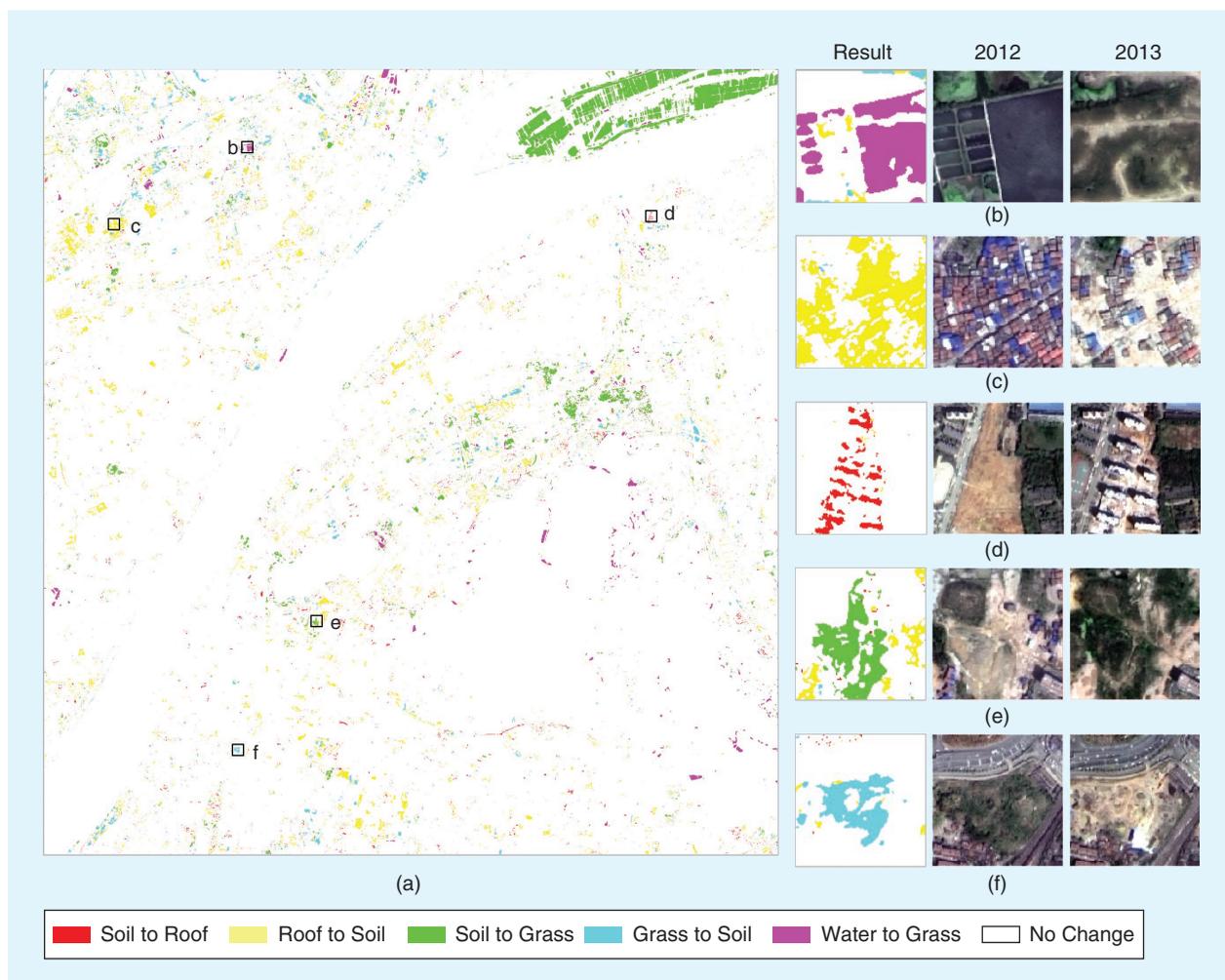


FIGURE 11. Land cover change detection using *Ziyuan-3* satellite imagery from 2012 and 2013. (a) The change detection result of the study area in Wuhan. (b)–(f) Five example cases of the change detection result and corresponding bi-temporal images [21].

features, e.g., the Harris detector [Figure 13(f)] and spectral and shape features [Figure 13(h)], can reduce these errors. These results illustrate that effective feature representation is the key to achieving good performance for VHR change detection.

Apart from 2D characteristics, 3D information has been exploited for building change detection in recent studies. With easier access to 3D data, such as multiview images, 3D information indicated by angular features can be conveniently used. More importantly, misregistration caused by spatial displacement is minimized [162]. Turker and Cetinkaya [163] detected damaged buildings by calculating the difference between digital elevation maps derived from pre- and postearthquake stereo images. In [157], multichannel indicators, such as height differences and texture similarities, are fused to monitor building changes. The incorporation of angular features is effective in improving the performance of building change detection, and it has potential for quantifying 3D dynamic processes in urban renewal and development. However, due to the relatively high cost of 3D data

acquisition, such as lidar and multiview UAV images, only a few studies investigate detailed building change processes in 3D space. Benefiting from time-series, multiview satellite imagery, Wen et al. [155] analyzed 3D annual building changes in inner city areas of four Chinese megacities (Beijing, Shanghai, Xi'an, and Wuhan). Their results characterized changes in the horizontal direction, such as construction and demolition, and quantified changes in the vertical direction, i.e., height and volume (Figure 14).

It should be noted that uncertainty and the cost of 3D data can present a bottleneck for the development and application of 3D building change detection. Specifically, on the one hand, lidar data are relatively accurate but not recurrently acquired. On the other hand, photogrammetrically derived 3D data from multiview images are a sufficiently cost-effective alternative to lidar, but their 3D reconstruction qualities depend on metaparameters of stereo pairs (e.g., intersections, off-nadir angles, sun elevations, azimuth angles, completeness, and time differences) [164]. Therefore, successful

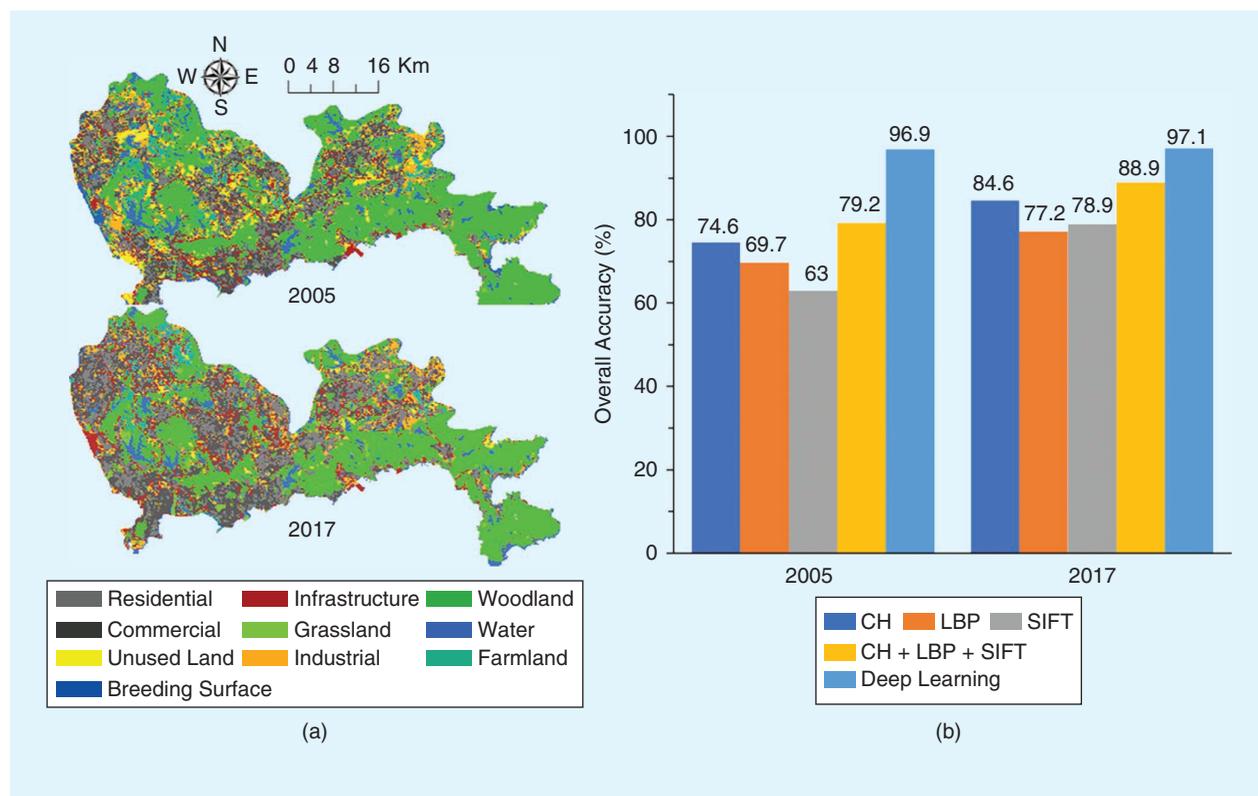


FIGURE 12. Land use change detection in the city of Shenzhen using high-spatial-resolution satellite imagery from 2005 to 2017, including (a) land use maps and (b) an accuracy assessment with different features [144].

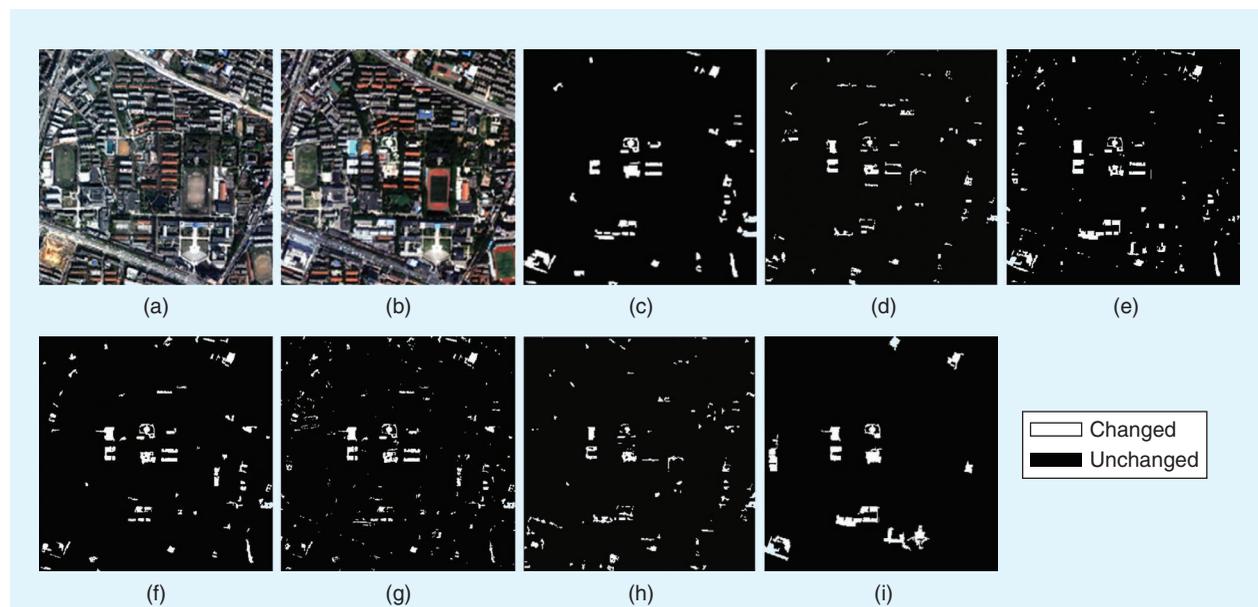


FIGURE 13. Building change detection maps obtained by different algorithms: (a) image (t_1), (b) image (t_2), (c) the reference change map, (d) the MBI-SVM, (e) object-based CVA, (f) the MBI and the Harris detector, (g) the MBI-CVA, (h) the fusion of the MBI with spectral and shape features, and (i) CVA using morphological features [51], [52].

3D building change detection relies on more advanced models that can produce accurate multitemporal 3D data in an economical and effective way. Very recently, deep learning has been explored for 3D reconstruction from multiview images. For example, a CNN-based method

was proposed for dense image matching in [165]. This novel technique may provide a new research orientation for 3D urban change detection when vertical and height information can be accurately derived from multiview satellite images.

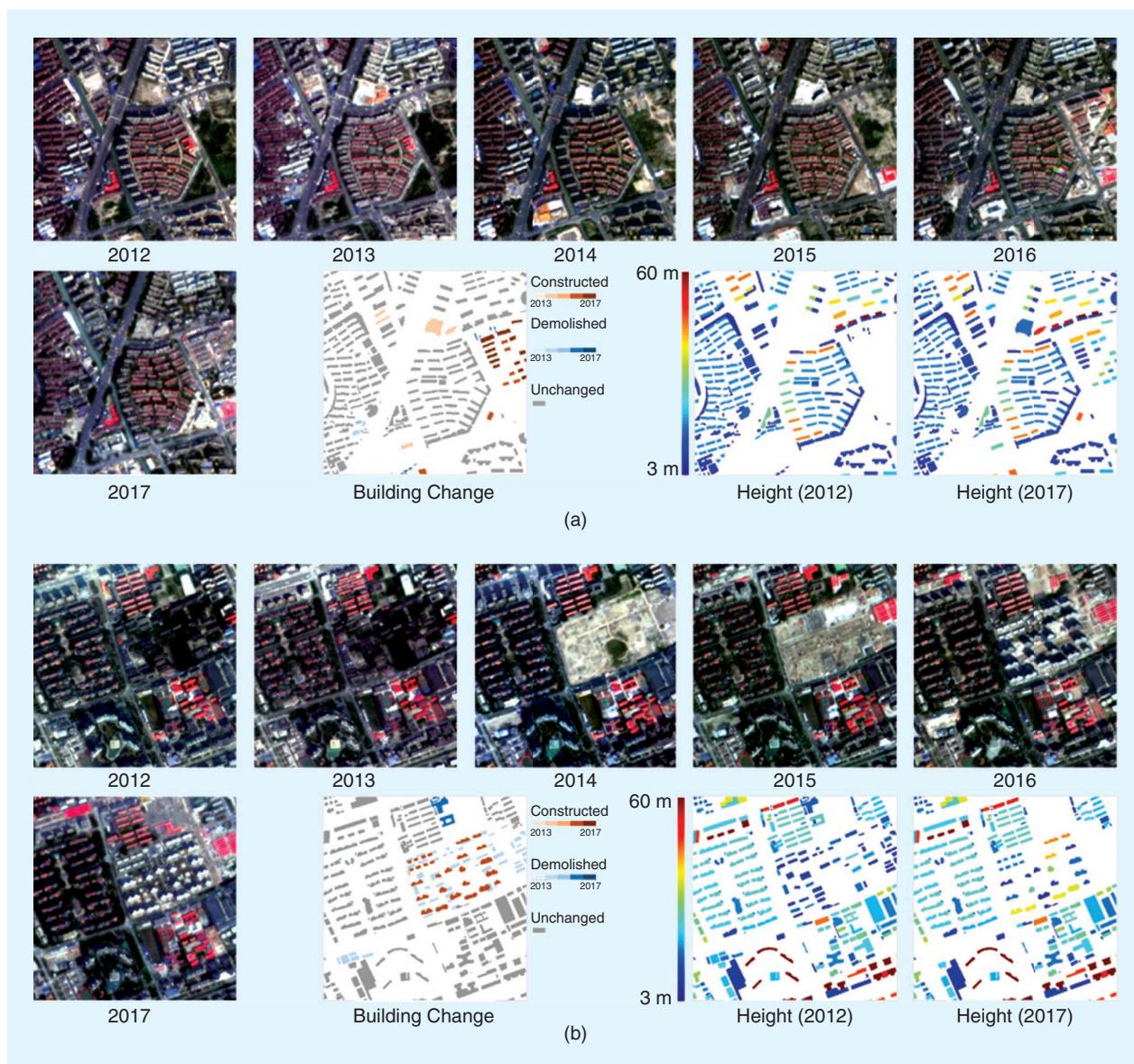


FIGURE 14. The annual 3D building change in subset areas of Shanghai that was achieved using multiview satellite imagery. (a) Subset area 1. (b) Subset area 2 [155].

VEGETATION CHANGE DETECTION

Analysis of vegetation change is important to understanding ecological transitions [166]. Using VHR imagery, vegetation change can be investigated at a much finer scale, e.g., from forest stands to individual trees. In general, there are three types of vegetation changes: 1) seasonal, caused by plant phenology; 2) gradual, caused by interannual climate variability, land management, and land degradation; and 3) abrupt, caused by disturbances, e.g., urbanization, deforestation, and fires [167]. In [168], to assess seasonal changes, both spectral and textural information extracted from multiseasonal *Pléiades* imagery (2 m) was used for multiseasonal leaf area index (LAI) mapping. The results showed that the highest LAI occurred in midsummer, followed by late spring, autumn, and winter, and the observed seasonal change trend was similar to that based on the in

situ measured LAI. Seasonal changes in the crown scale in an Amazon tropical evergreen forest were assessed by Wang et al. [169] using *Planet* constellation imagery with a spatial resolution of 3 m. The crown scale fraction of nonphotosynthetic vegetation showed large seasonal trend variability from June to November.

As for gradual changes, Gärtner et al. [170] used *QuickBird* and *WorldView-2* imagery to quantify tree crown diameter changes in a degraded riparian tugai forest in northwestern China, and their results indicated that the diameter increased by 1.14 m, on average, during 2005–2011. Tian et al. [171] explored DSMs from satellite stereo sensors to monitor vertical tree growth and found that periodic annual increments at the study sites were in the range of 0.3–0.5 m. In the case of abrupt change, Dalagnol et al. [172] quantified tree canopy loss and gap recovery in

tropical forests where there was low-intensity logging by using *WorldView-2* and *GeoEye-1* images. Their study showed that VHR satellite imagery has potential for tracking small-scale human disturbances. Ardila et al. [173] identified bi-temporal tree crown elliptical objects through the iterative surface fitting of a Gaussian model to crown membership in two urban residential areas in The Netherlands using *QuickBird* and aerial images. A detection rate of 77% was reported for both removed and planted trees.

In addition to coverage, tree crown diameters, and canopy heights, species types are an essential parameter of vegetation community structures. In particular, VHR imagery is able to identify small and highly mixed species. Since different vegetation types exhibit similar spectral characteristics, textures are often used to identify various species. For instance, Lu and He [174] investigated seasonal species variations in a tall grassland in Ontario, Canada, during the growing season (from April to December) in 2015 using UAV images. The reflectance value, vegetation indices, and GLCM textures were used in the classification, and temporal change analysis revealed the growing process and succession of different species. Notably, some advanced methods, e.g., deep features [175], photogrammetric-derived DSMs from stereo images [176], phenological characteristics [177], and data fusion (e.g., lidar and airborne hyperspectral images) [178], have been considered for the change analysis of vegetation species. Moreover, some researchers attempted to discriminate vegetation function types, e.g., park, roadside, and residential-industrial trees in urban areas [179]. Likewise, vegetation function-type change monitoring is of great significance but has not been addressed in the current research.

MONITORING CROP CHANGES

Information about agricultural land changes, crop type conversions, and crop growth, critical for precision agriculture, can be effectively captured using VHR images. In [180], land cover data for Guanlin, Yixing City, China, in 2006, 2009, 2012, and 2015 were generated using *QuickBird* images, and they showed a decrease followed by an increase in the agricultural land area that was observed. Malinverni et al. [181] quantified the temporal variation of main crop rotations on the Capitanata plain of Southern Italy using *WorldView-2* images, and the textural features (e.g., the GLCM and the Gabor wavelet) were employed to improve the classification accuracy. The study suggests that multi-temporal classification is preferred in crop mapping, due to its rich phenological characteristics. Furthermore, frequent crop growth monitoring is extremely important for timely decision making in precision agriculture. Therefore, time-series data are recommended, although dense time series of VHR images are relatively difficult to acquire.

Recently, new generation micro-/nanosatellites (e.g., *Planet*) and UAV systems have become available and are able to obtain time-series VHR images, which has potential for agricultural applications. For example, Sadeh et al.

[182] detected sowing dates using dense time-series *Planet* CubeSat data with an interval of two days. As shown in Figure 15, a partly sown field was successfully detected, implying that detailed processes on a near daily basis can be monitored by dense time series of VHR data. Likewise, Bendig et al. [183] monitored plant growth based on crop surface models using stereo UAV images. Notably, height differences between cultivars and their increased trend during the growing season can be observed.

Crop change caused by disease and insect damage can also be located. VHR images are able to identify small-extent disease and insect damage, which is beneficial for controlling problems at early stages. Generally, diseases and insects can result in various kinds of harm to crop canopies, such as the removal of leaves, skeletonizing of leaf tissue, and discoloration of leaves, and these effects vary depending on the type of disease, insect, and crop [184]. Therefore, different damage shows various spectral and structural characteristics in remote sensing images, which makes the identification of disease and insect problems via VHR images a challenging task. One of the successful applications was presented by Johansen et al. [185], where *GeoEye-1* images acquired in 2012, 2013, and 2014 were used to detect canegrub damage in sugarcane fields. In the study, objects with low NDVI values and rough textures were identified as likely to be damaged, and they were further classified as low, medium, and high likelihood. Franke and Menz [186] observed different levels of disease severity in a plot of winter wheat using multitemporal *QuickBird* images acquired in April, May, and June.

The experimental results show that VHR multispectral data are only moderately suitable for damage detection at an early growth stage, a fact attributed to the subtle spectrum and texture differences between damaged and healthy crops [187], [188]. However, VHR hyperspectral sensors seem to have potential to address this issue. For example, in [189], spectral and spatial features were extracted by a CNN from UAV hyperspectral images for the detection of yellow rust across a whole crop cycle of winter wheat. Satisfactory accuracy was achieved through all growing stages, due to the detailed spectral information and rich spatial details in VHR hyperspectral images.

MONITORING LAKES AND WETLANDS

Lakes and wetlands, which play a critical role in biodiversity, ecosystems, hydrology, and climate regulation, are highly dynamic due to various natural and anthropogenic factors, such as climate change, farming, urbanization, floods, and hydrological interventions [190]. Therefore, accurate and timely monitoring of lakes and wetlands is important for management, restoration, and protection. Many studies have used remote sensing data for monitoring lakes, from a local to a global scale. They include lake changes between 1975 and 2015 across the Yangtze floodplain in China via *Landsat* images [191], water clarity changes in lakes and reservoirs across China that were observed using

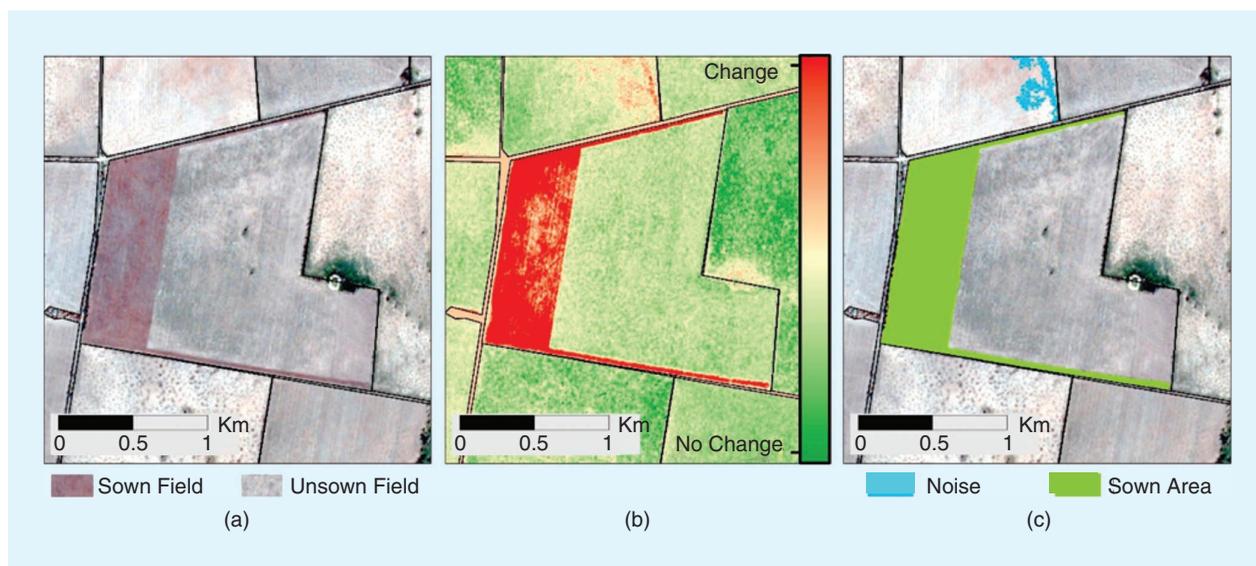


FIGURE 15. A sowing detection result obtained using time-series *Planet* CubeSat images [182]. (a) RGB satellite imagery. (b) The change result. (c) The sowing detection result.

Moderate Resolution Imaging Spectroradiometer (MODIS) data [192] from 2000 to 2017, and global surface water changes between 1984 and 2015 acquired through *Landsat* images [193]. In these studies, which were subject to relatively low spatial resolution, lakes with large areas were targeted. However, more than 303.6 million of the 304 million lakes at the global scale are smaller than 1 km² [194]. Therefore, VHR remote sensing images are required for observing them. To our knowledge, however, only a few studies have focused on lake monitoring using VHR images.

Cooley et al. [195] tracked water changes in the 470 lakes (0.0025–1.23 km²) in the Yukon Flats of north-central Alaska during mid to late summer (23 June to 1 October) in 2016, using *Planet* CubeSat images with a spatial resolution of 3 m. A time-series analysis revealed that the area of 83% of the studied lakes had decreased and that 22% of the lakes had lost more than half their surface. Notably, more applications of advanced methods of water detection through VHR images, e.g., deep learning [196] and physical approaches [197], are needed. Furthermore, information about black and odorous water [198] and water types (e.g., rivers, lakes, canals, and ponds) [199] is of increasing interest, and multitemporal monitoring is imperative.

In addition to lakes, VHR images have potential for monitoring detailed changes in wetland ecosystems. In [200], the results of five-level mangrove features, including vegetation boundaries, mangrove stands, mangrove zonation, individual tree crowns, and species communities, using different data sets [*Landsat* (30 m), *Advanced Land Observing Satellite Advanced Visible and Near-Infrared Radiometer 2* (10 m), pan-sharpened *WorldView-2* (0.5 m), and lidar] were generated and compared. As described in Figure 16, the *Landsat* image cannot accurately discriminate the mangrove extent, due to the mixed-pixel problem [Figure 16(e)], and more fine-scale mangrove features, i.e., tree-crown-level

species, can be captured only by pan-sharpened *WorldView-2* imagery [Figure 16(l)–(p)]. By summarizing the current literature, it can be found that most studies focus on detecting the extent of wetland change but ignore species change. For instance, Hu et al. [201] monitored land cover changes in the Hangzhou Xixi wetland from 2000 to 2013 using *IKONOS*, *QuickBird*, and *WorldView-2* images. It was shown that the nonwetland area increased by approximately 100%, mostly in the form of herbaceous zones, followed by forests, ponds, cropland, marshes, and rivers. Wu et al. [202] integrated lidar data and multitemporal aerial imagery (1 m) to map wetland inundation dynamics in the Prairie Pothole region of North America, which is characterized by millions of small depressional wetlands.

The difficulties of species change detection in wetlands lie in the following aspects. On the one hand, tidal and phenological changes make different plant species highly dynamic on daily and seasonal frequencies, respectively. On the other hand, many species have a similar spectral reflectance during the peak biomass in complex wetland landscapes [203], and the spectral signature of the same species can be influenced by many complex factors, such as the off-nadir angle, sun-viewing geometry, crown porosity, leaf clumping, and ground surface scattering [204]. For instance, in [200], mangrove species were categorized from *WorldView-2* images using the nearest-neighbor classifier to extract object-based spectral and textural features within tree crowns, but a low overall accuracy of around 54% was reported. As demonstrated in Figure 16(p), misclassified open scrub *Avicennia marina* can be clearly observed. To improve the discriminative power among various species, the potential of VHR hyperspectral images, dense time-series data, and vertical information for characterizing detailed spectral, phenological, and height attributes needs to be explored.

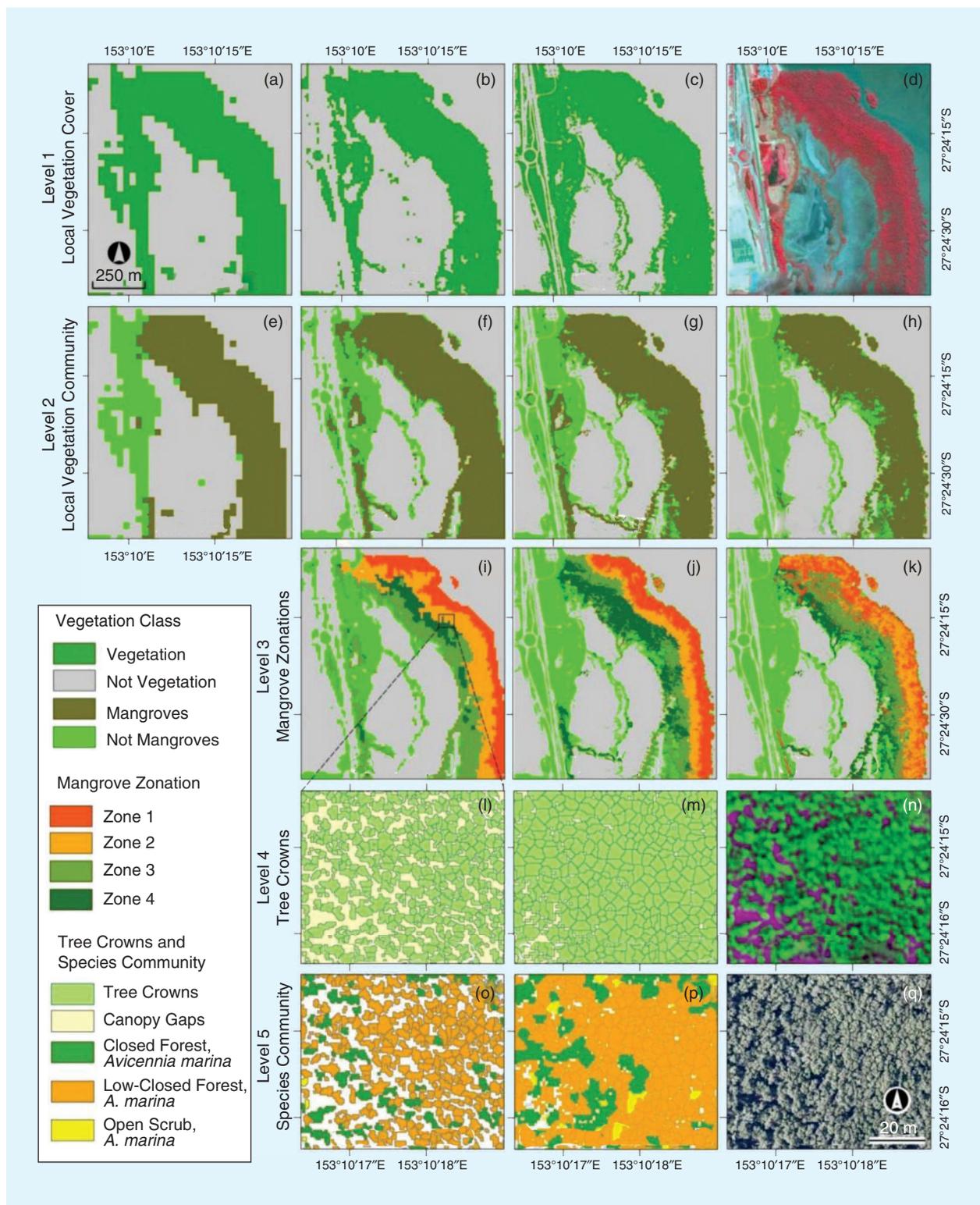


FIGURE 16. Five-level mangrove features generated using different data sets [200]. (a) Level 1 TM, (b) level 1 AVNIR-2, (c) level 1 WorldView-2, (d) WorldView-2 RGB image, (e) level 2 TM, (f) level 2 AVNIR-2, (g) level 2 WorldView-2, (h) level 2 WorldView-2+LiDAR, (i) level 3 AVNIR-2, (j) level 3 WorldView-2, (k) level 3 WorldView-2+LiDAR, (l) level 4 pan-sharpened WorldView-2, (m) level 4 pan-sharpened WorldView-2+LiDAR, (n) WorldView-2 PC1,2,1, (o) level 5 pan-sharpened WorldView-2, (p) level 5 pan-sharpened WorldView-2+LiDAR, and (q) aerial photograph.

ECOSYSTEM SERVICES MONITORING

Ecosystem services link ecosystems to human welfare by regarding nature as a stock providing a flow of services (e.g., local climate regulation and water purification) [205]. Monitoring urban ecosystem services is of great value for investigating ecological function changes and can help improve the understanding of urbanization impacts on local ecological benefits. VHR satellite data can monitor spatially explicit ecosystem services at fine scales. Generally speaking, there are two categories of methods to derive ecosystem services: 1) statistical regression and radiative transfer models and 2) land use/cover-based methods [206]. Since in situ observations are not always available and the validity of statistical regression and radiative transfer models is affected by time inconsistencies between ground and remotely sensed measurements, land use/cover-based methods are often preferred. For example, in [207], land use/cover maps of Shanghai's urban core from 2000 to 2009 were classified using *IKONOS* and *GeoEye-1* images, and the classes were then transformed into ecosystem service supply and demand budgets, including regulating, provisioning and cultural services, and ecological integrity. An increase of at least 20% in ecosystem service supply budgets was observed, which was mainly attributed to the replacement of continuous urban fabric and industrial areas by high-rise commercial/residential areas despite a slight increase in urban green sites.

Huang et al. [144] assessed ecosystem service change in Shenzhen from 2005 to 2017 using *Gaofen-2* (4-m) and *QuickBird* (2.4-m) images. In the study, multitemporal land use maps were generated by a transferred deep CNN (as shown in Figure 12), based on which ecosystem service supply and demand values were estimated. It was found that supply capacity had decreased by 13.7% due to a reduction in woodlands, water, farmland, and so on, but, on the other hand, demand values had grown by 23.5% because urban expansion and redevelopment had increased the amount of residential, commercial, and infrastructure land. The results clearly demonstrated the ecosystem degradation of Shenzhen during the previous 10 years. Ren et al. [208] evaluated the ecosystem services of Guyuan City in 2003, 2009, and 2014 via VHR satellite imagery (e.g., *QuickBird* and *Gaofen-1*) and showed that VHR images were advantageous in the dynamic, quantitative, and visual examination of ecological changes. With VHR remote sensing images, fine-scale ecosystem services within urban areas can be effectively quantified. However, most of the current works focus on urban areas and ignore the ecosystem services of natural scenes, such as forests and wetlands. Moreover, these works present only case studies, and large-scale examinations are still lacking.

IMPERVIOUS-SURFACE CHANGE DETECTION

The change detection of impervious surfaces is important in monitoring and understanding urban development and has been extensively studied in the remote sensing

literature. However, most of the existing studies monitor the change of impervious surfaces based on coarse- and medium-spatial-resolution satellite imagery, such as *MODIS* and *Landsat* [209], [210], which, on the other hand, have difficulty dealing with areas that have low impervious-surface intensities and mixed pixels [211]. During recent decades, images with high spatial resolution have provided new opportunities for subtle impervious-surface monitoring at very fine scales. However, impervious-surface monitoring using VHR imagery is a challenging task. VHR multitemporal images exhibit a large number of details (e.g., buildings, roads, driveways, and sidewalks), greater spatial heterogeneity (e.g., different viewing geometries), and occlusion by urban trees, shadow, and vertical structure layover [212]. To address the problem caused by shadow, Li et al. [213] extracted multiscale object features and further classified shaded areas to extract impervious surfaces using *QuickBird* and *IKONOS* imagery. More recently, Zhang and Huang [214] developed a two-stage object-based classification method based on multilevel features (i.e., spectral, textural, shape, and class related) for time-series impervious-surface change detection in Shenzhen in 2003–2017, including the impervious-surface mapping of both nonshaded and shaded areas. As can be seen in Figure 17, in addition to single changes across the studied period (i.e., cases 1 and 2), some regions (e.g., case 3) experienced multiple changes.

SUMMARY OF VHR CHANGE DETECTION DIMENSIONS

As suggested in [10], remote sensing change detection can be categorized according to different dimensions, e.g., input data, temporal resolutions, change categories, targets, and analysis units. Since this research focuses on VHR optical images, the input data are discussed in terms of spatial resolutions. Therefore, we divide VHR change detection studies by considering the following five categorization schemes:

- 1) spatial resolution: HR (2–5 m), VHR (1–2 m), and ultra-HR (UHR) (<1 m)
- 2) temporal resolution: bi-temporal and multitemporal
- 3) analysis unit: pixel, object, and patch
- 4) change category: binary change (BC), multiple change (MC), and directional change (DC) categories
- 5) targets.

In terms of the previously mentioned categorization schemes, a distribution of the literature reviewed in this study appears in Figure 18. Most articles use only bi-temporal images (78.12%) and concern binary change (66.32%). With regard to spatial resolution, 43.75% of the papers use UHR images, followed by VHR (33.33%) and HR (22.92%) images. As for analysis units, pixels and objects have almost the same number of articles, but patch-based change detection is rarely reported. Of the studies reviewed in this research, more than half involve land cover and land use change detection with multiple targets considered, followed by a series of specific targets, including buildings (20%), vegetation (10.53%), crops (8.42%), lakes and

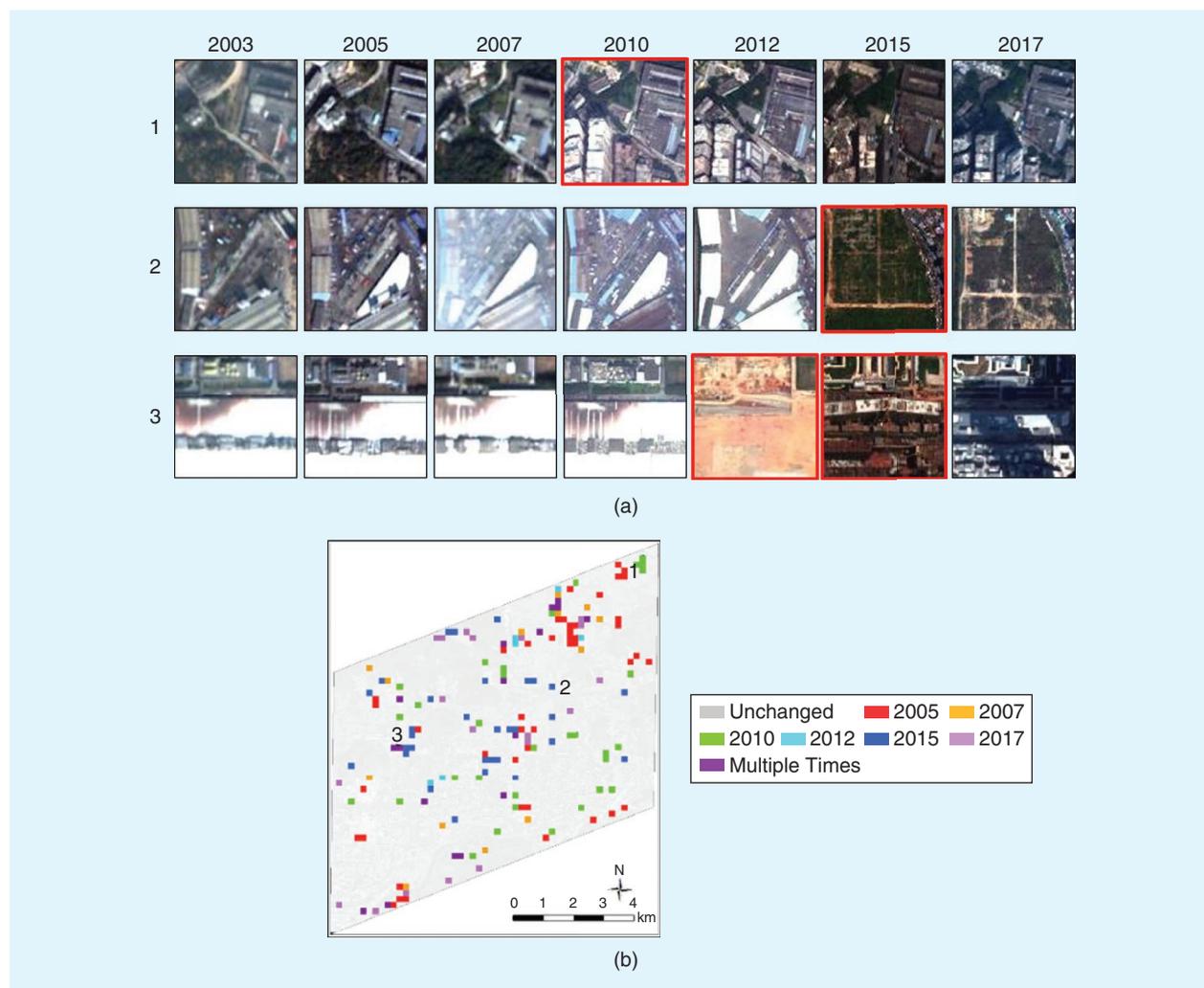


FIGURE 17. Impervious-surface monitoring results from Shenzhen during 2003–2017. (a) Some typical cases of change profiles and (b) change detection results [214]. Red borders represent corresponding change times.

wetlands (5.26%), ecosystem services (3.16%), and impervious surfaces (2.1%).

RECOMMENDATIONS FOR FUTURE WORK

FROM CHANGE DETECTION TO TRACKING

Most VHR change detection studies focus on bi-temporal images and multiple time series. However, change events, such as phenology and urban development, cannot be well characterized by coarse temporal observations. Frequent HR monitoring of both human and natural activities deserves much attention, especially when small satellite constellation (e.g., *Planet*) images become available. With time series VHR images, change detection is advanced from simply locating variations via bi-temporal data to dense time-series monitoring [215]. There have been attempts at time-series monitoring using VHR images of buildings [155], crops [216], water [195], impervious surfaces [214], newly constructed building areas [2], forests [217], and landslides [218]. However, most of these methods are merely an extension of bi-temporal techniques by multiple

pair comparisons, which is not sufficient to capture the temporal context and semantics and to support time series analysis.

Recently, VHR videos acquired by *SkySat-1*, *Jinlin-1*, and the *UrtheCast Iris* camera have shown great potential for near-real-time target tracking from space. Most of the current change detection studies have focused on the appearance/disappearance and shape changes of objects, but studies related to tracking moving objects (e.g., ships, planes, trains, and vehicles) in VHR sequential videos are limited. In [219], the automatic detection and tracking of moving ships using satellite video was achieved based on multiscale saliency and surrounding contrast analysis. Wang et al. [220] presented a UAV-based vehicle detecting and tracking system, which jointly considered edges, optical flows, and local feature points. The first-ranked team at the 2016 IEEE Geoscience and Remote Sensing Society Data Fusion Contest designed an innovative deep neural network with an MSI and spaceborne video as input, and object activity was analyzed using the Kanade–Lucas–Tomasi key point tracker [221], [222]. During the coming

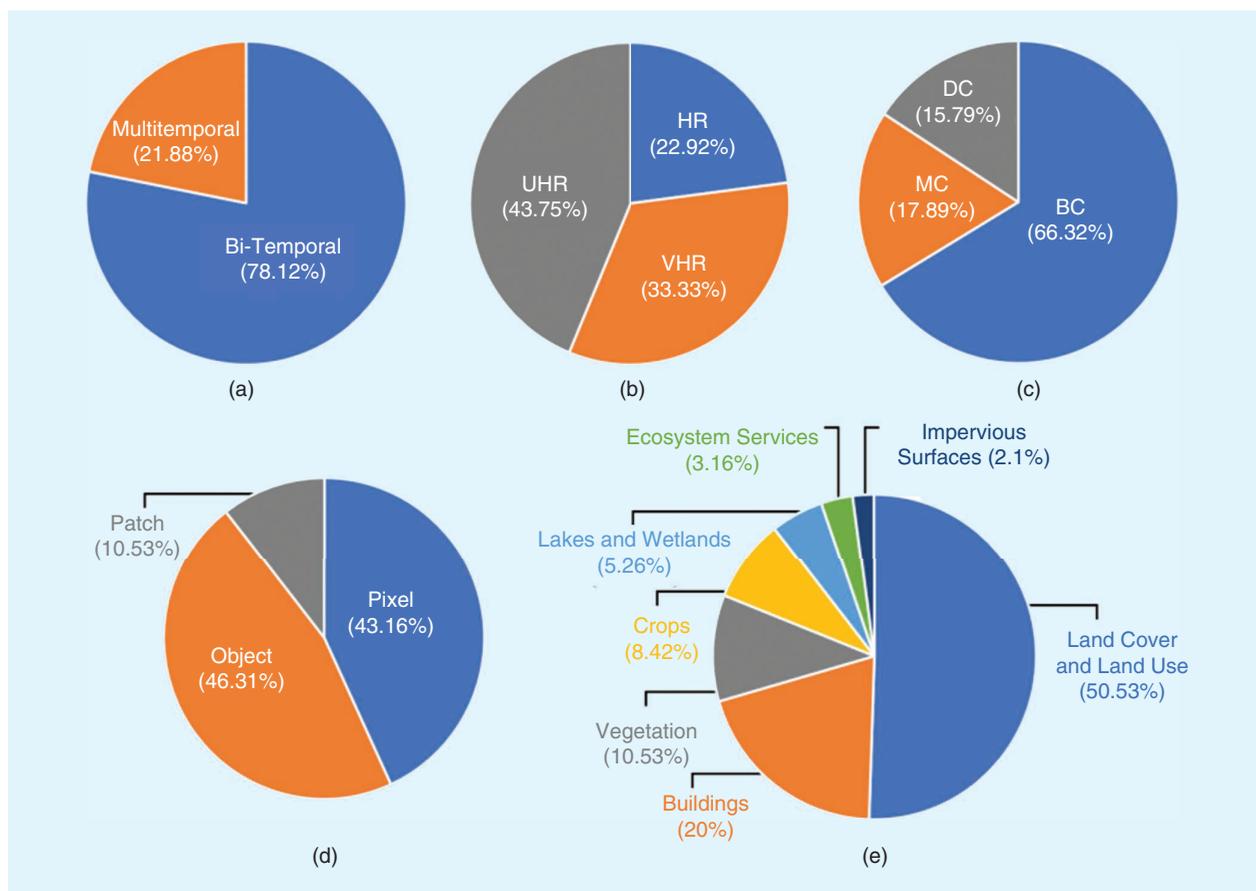


FIGURE 18. The distribution of different dimensions for the studies reviewed in this research: (a) temporal resolution, (b) spatial resolution, (c) change categories, (d) analysis units, and (e) targets.

years, space videos are likely to be a very important data source for Earth monitoring, and more promising studies based on VHR sequential videos can be expected, while a new era in VHR change detection that shifts from conventional multitemporal change detection to video sequential tracking may dawn. Despite the preceding attempts, change tracking using VHR videos is still in its early stage and needs to be further explored. Notably, unlike conventional videos, challenges related to satellite video processing may include the small size of moving objects (e.g., vehicles), complex backgrounds (e.g., building relief displacement in urban scenes), camera movements, and low frame rates.

HR GLOBAL CHANGE DETECTION

Remote sensing imagery has long been considered an effective data source for global change detection, due to its large coverage area, convenient access, and frequent revisits. Previous multitemporal global maps of land cover and thematic change detection are often generated at a relatively coarse resolution (i.e., >300 m), e.g., 8-km-resolution global forest change based on *Advanced Very High Resolution Radiometer* data for 1982–1999 [223], 500-m resolution mapping of the global urban extent from *MODIS* data from 2005 and 2009 [224], [225], and

the 300-m resolution annual Climate Change Initiative Land Cover maps from 1992 to 2015 [226]. More recently, global-scale change detection with fine spatial resolution (around 30 m) has been attempted with open source *Landsat* imagery. Notable examples include the Global Forest Cover database [227], GlobeLand30 global land cover product [228], Global Artificial Impervious Area annual maps [229], Global Surface Water data sets by the European Commission Joint Research Center [230], and Global Human Settlement Layer framework [231]. Please note that 30 m is not a high spatial resolution in a common sense, but it should be regarded as high in the case of intercontinental and global mapping. Recently, Gong et al. [232] developed a 10-m resolution global land cover map through *Sentinel-2* images acquired in 2017.

It is a trend that global products are being developed in finer spatial and temporal resolutions that can characterize heterogeneous and mixed areas more accurately. For instance, the *Planet* CubeSats are able to acquire images at a 3–5-m spatial resolution with near-real-time daily global coverage [233], which has potential for VHR global change detection in the future. In addition, cloud computing platforms, such as Google Earth Engine and Amazon Web Services, can facilitate the processing of large volumes of

satellite images and speed the development of VHR global mapping [234].

HYPERSPECTRAL CHANGE DETECTION

Hyperspectral data can distinguish more detailed land cover types due to their rich spectral information. For a long time, the data availability of hyperspectral images seemingly limited real applications in precise change detection. Recently, however, the development of hyperspectral satellites with a relatively fine spatial resolution, e.g., *Gaofen-5* (30 m, with 330 spectral bands), *Tiangong-1* (10 m, with 128 spectral bands), and *Zhuhai-1* (10 m, with 32 spectral bands), and airborne hyperspectral sensors, e.g., HyMap (3 m, with 126 spectral bands) and the *Reflective Optics System Imaging Spectrometer (ROSIS)* (1.3 m, with 115 spectral bands), has significantly increased the availability of multitemporal hyperspectral images. However, studies related to VHR hyperspectral change detection are very limited, and even the existing methodologies were developed based on synthetic data [235]. Moreover, advances in hyperspectral image classification benefit from a set of widely used public benchmark data sets, e.g., the ROSIS Pavia University and Airborne Visible/Infrared Imaging Spectrometer Salinas data sets [236]. Therefore, there is an urgent need for public hyperspectral change detection data sets to promote the development of the related research fields.

URBAN FUNCTIONAL ZONE CHANGE DETECTION

Currently, the classification of urban functional zones is one of the important research areas in interpreting VHR remote sensing images, as the urban functional zones can bridge the semantic gap between land cover and human socioeconomic activities. Current urban functional zone mapping not only involves various image features, e.g., deep [237], [238], angular [97], object based [239], and textural [240], but it also refers to multisource geographic information, such as points of interest (POIs) [241], social media [242], and mobile phone positioning [100]. In rapidly urbanizing regions, the timely and accurate monitoring of urban functional zones is crucial for planning and management. However, studies for change detection in urban functional zones are lacking. Frankly, urban functional zone change detection is a difficult task since land cover change does not necessarily signify the conversion of a functional zone type. Meanwhile, multisource geographic data, e.g., POIs, are widely used for functional zone classification [230], but these data do not provide a time tag, which hampers the dynamic monitoring of urban functional zones. These issues should be overcome to effectively monitor changes in cities.

CONCLUSIONS

With the increasing availability of VHR remote sensing images, precise, frequent, and stereo change detection becomes possible. To the best of our knowledge, a comprehensive review of VHR change detection is lacking in the

current literature. Therefore, this article aimed to summarize recent advances in VHR remote sensing image change detection, including methods and applications. The review of methods focused on feature extraction and change detectors for multitemporal VHR images. Applications including change detection for land cover and land use, impervious surfaces, buildings, crops, vegetation, lakes and wetlands, and ecosystem services were reviewed. Finally, some future directions were suggested and discussed for this important research area. Recommendations for future work include focusing on change tracking, global change detection, hyperspectral change detection, and urban functional zone change detection to generate frequent and detailed semantic change information on a global scale.

ACKNOWLEDGMENTS

The authors are grateful to the editor-in-chief, associate editor, and reviewers for their insightful comments and suggestions. This research was supported by the National Natural Science Foundation of China, under grants 41901279, 41771360, and 41971295, and the Chinese Academy of Sciences Interdisciplinary Innovation Team, under grant JCTD-2019-04. (Corresponding author: Xin Huang.)

AUTHOR INFORMATION

Dawei Wen (daweiwen@mail.hzau.edu.cn) received the B.E. degree in surveying and mapping and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2013 and 2018, respectively. She is a postdoctoral researcher in the College of Public Administration, Huazhong Agricultural University, Wuhan, 430070, China. Her research interests include the change analysis of multitemporal remote sensing images and remote sensing applications.

Xin Huang (xhuang@whu.edu.cn) received the Ph.D. degree in photogrammetry and remote sensing in 2009 from Wuhan University, Wuhan, China. He is a Luojia Distinguished Professor at Wuhan University, Wuhan, 430079, China, where he teaches remote sensing, photogrammetry, and image interpretation. He is the founder and director of the Institute of Remote Sensing Information Processing, School of Remote Sensing and Information Engineering, Wuhan University. He has published more than 150 peer-reviewed articles (Science Citation Index papers) in international journals. His research interests include remote sensing image processing methods and applications. He was the recipient of the Boeing Award for the Best Paper in Image Analysis and Interpretation from the American Society for Photogrammetry and Remote Sensing (ASPRS) in 2010, the second-place recipient of the John I. Davidson President's Award from ASPRS in 2018, and the winner of the IEEE Geoscience and Remote Sensing Society 2014 Data Fusion Contest. He was an associate editor of *Photogrammetric Engineering and Remote Sensing* (2016–2019) and of *IEEE Geoscience and Remote Sensing Letters* (2014–2020), and he now serves as an associate editor of *IEEE Journal of*

Selected Topics in Applied Earth Observations and Remote Sensing (since 2018). He is also an editorial board member of *Remote Sensing of Environment* (since 2019), *Science of Remote Sensing* (since 2020), and *Remote Sensing* (since 2018). He is a Senior Member of IEEE.

Francesca Bovolo (bovolo@fbk.eu) received the B.S. and M.S. degrees in telecommunication engineering (summa cum laude) and the Ph.D. degree in communication and information technologies from the University of Trento, Italy, in 2001, 2003, and 2006, respectively, where she remained as a research fellow until June 2013. She is the founder and head of the Remote Sensing for Digital Earth unit at Fondazione Bruno Kessler, Trento, 38123, Italy, and a member of the Remote Sensing Laboratory, Trento. Her research interests include multitemporal remote sensing image analysis; change detection in multispectral, hyperspectral, and synthetic aperture radar images and VHR images; time series analysis; content-based time series retrieval; domain adaptation; and lidar and radar sounders. She was the publication chair for the 2015 IEEE International Geoscience and Remote Sensing Symposium. She is the cochair of the Society of Photographic Instrumentation Engineers International Conference on Signal and Image Processing for Remote Sensing. She is a Senior Member of IEEE.

Jiayi Li (zjjerica@whu.edu.cn) received the B.S. degree from Central South University, Changsha, China, in 2011 and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2016. She is currently an assistant professor in the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, 430079, China. She has authored more than 30 peer-reviewed articles (Science Citation Index papers) in international journals. Her research interests include hyperspectral imagery, sparse representation, computation vision and pattern recognition, and remote sensing images. She is a reviewer for more than 10 international journals, including *IEEE Transactions on Geoscience and Remote Sensing*, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *IEEE Geoscience and Remote Sensing Letters*, *IEEE Signal Processing Letters*, and *International Journal of Remote Sensing*. She is the guest editor of the special issue "Change Detection Using Multisource Remotely Sensed Imagery" of *Remote Sensing* (an open-access journal of the Multidisciplinary Digital Publishing Institute). She is a Member of IEEE.

Xinli Ke (kexl@mail.hzau.edu.cn) received the B.S. degree in land planning and utilization from Huazhong Agricultural University, Wuhan, China, in 2001 and the M.S. degree in cartography and geographical information systems and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2006 and 2009, respectively. He is a professor in the College of Public Administration, Huazhong Agricultural University, Wuhan, 430070, China.

Anlu Zhang (zhanglanlu@mail.hzau.edu.cn) received the Ph.D. degree in 1999 from Huazhong Agricultural University, Wuhan, China. He has been a professor at Huazhong

Agricultural University, Wuhan, 430070, China, since 2000. He is an executive director of the China Land Society; deputy director of the Academic Committee, China Land Society; deputy director of the Youth Working Committee, China Land Society; and a member of the Expert Committee, Land Remediation Center, Ministry of Land and Resources.

Jón Atli Benediktsson (benedikt@hi.is) received the Cand.Sci. degree in electrical engineering from the University of Iceland, Reykjavik, Iceland, in 1984, and the M.S.E.E. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, Indiana, USA, in 1987 and 1990, respectively. He is with the Faculty of Electrical and Computer Engineering, University of Iceland, Reykjavik, IS 107, Iceland. From 2009 to 2015, he was the prorector of science and academic affairs and a professor of electrical and computer engineering at the University of Iceland. In 2015, he was the rector of the University of Iceland. He is a cofounder of Oxymap, Reykjavik, a biomedical start-up company. He has authored and coauthored extensively in his fields of interest. His research interests include remote sensing, image analysis, pattern recognition, biomedical analysis of signals, and signal processing. He is a Fellow of IEEE.

REFERENCES

- [1] S. Liu, Q. Du, X. Tong, A. Samat, and L. Bruzzone, "Unsupervised change detection in multispectral remote sensing images via spectral-spatial band expansion," *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 12, no. 9, pp. 3578–3587, 2019. doi: 10.1109/JSTARS.2019.2929514.
- [2] X. Huang, Y. Cao, and J. Li, "An automatic change detection method for monitoring newly constructed building areas using time-series multi-view high-resolution optical satellite images," *Remote Sensing Environment*, vol. 244, p. 111,802, 2020. doi: 10.1016/j.rse.2020.111802.
- [3] I. J. P. Coppin, K. Nackaerts, B. Muys, and E. Lambin, "Digital change detection methods in ecosystem monitoring: A review," *Int. J. Remote Sens.*, vol. 25, no. 9, pp. 1565–1596, 2004. doi: 10.1080/0143116031000101675.
- [4] D. Lu, P. Mausel, E. Brondizio, and E. Moran, "Change detection techniques," *Int. J. Remote Sens.*, vol. 25, no. 12, pp. 2365–2401, 2004. doi: 10.1080/0143116031000139863.
- [5] A. P. Tewkesbury, A. J. Comber, N. J. Tate, A. Lamb, and P. F. Fisher, "A critical synthesis of remotely sensed optical image change detection techniques," *Remote Sens. Environ.*, vol. 160, pp. 1–14, 2015. doi: 10.1016/j.rse.2015.01.006.
- [6] M. Hussain, D. Chen, A. Cheng, H. Wei, and D. Stanley, "Change detection from remotely sensed images: From pixel-based to object-based approaches," *ISPRS J. Photogrammetry Remote Sens.*, vol. 80, pp. 91–106, June 2013. doi: 10.1016/j.isprsjprs.2013.03.006.
- [7] G. Chen, G. J. Hay, L. M. T. Carvalho, and M. A. Wulder, "Object-based change detection," *Int. J. Remote Sens.*, vol. 33, no. 14, pp. 4434–4457, 2012. doi: 10.1080/01431161.2011.648285.
- [8] S. Liu, D. Marinelli, L. Bruzzone, and F. Bovolo, "A review of change detection in multitemporal hyperspectral images: Current techniques, applications, and challenges," *IEEE Geosci. Re-*

- ote Sens. Mag. (replaces Newslett.)*, vol. 7, no. 2, pp. 140–158, 2019. doi: 10.1109/MGRS.2019.2898520.
- [9] F. Bovolo and L. Bruzzone, "The time variable in data fusion: A change detection perspective," *IEEE Geosci. Remote Sens. Mag. (replaces Newslett.)*, vol. 3, no. 3, pp. 8–26, 2015. doi: 10.1109/MGRS.2015.2443494.
- [10] H. Si Salah, S. E. Goldin, A. Rezugui, B. Nour El Islam, and S. Ait-Aoudia, "What is a remote sensing change detection technique? Towards a conceptual framework," *Int. J. Remote Sens.*, vol. 41, no. 5, pp. 1788–1812, 2020. doi: 10.1080/01431161.2019.1674463.
- [11] H. Han et al., "A mixed property-based automatic shadow detection approach for VHR multispectral remote sensing images," *Appl. Sci.*, vol. 8, no. 10, p. 1883, 2018. doi: 10.3390/app8101883.
- [12] C. Toth and G. Józków, "Remote sensing platforms and sensors: A survey," *ISPRS J. Photogrammetry Remote Sens.*, vol. 115, pp. 22–36, May 2016. doi: 10.1016/j.isprsjrs.2015.10.004.
- [13] D. Poli and T. Toutin, "Review of developments in geometric modelling for high resolution satellite pushbroom sensors," *Photogrammetric Rec.*, vol. 27, no. 137, pp. 58–73, 2012. doi: 10.1111/j.1477-9730.2011.00665.x.
- [14] M. Dalla Mura, S. Prasad, F. Pacifici, P. Gamba, J. Chanussot, and J. A. Benediktsson, "Challenges and opportunities of multimodality and data fusion in remote sensing," *Proc. IEEE*, vol. 103, no. 9, pp. 1585–1601, 2015. doi: 10.1109/JPROC.2015.2462751.
- [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 Sens.*, vol. 8, no. 2, p. 88, 2016. doi: 10.3390/rs8020088.
- [16] M. Volpi, D. Tuia, F. Bovolo, M. Kanevski, and L. Bruzzone, "Supervised change detection in VHR images using contextual information and support vector machines," *Int. J. Appl. Earth Observat. Geoinf.*, vol. 20, pp. 77–85, Feb. 2013. doi: 10.1016/j.jag.2011.10.013.
- [17] J. P. Ardila, W. Bijker, V. A. Tolpekin, and A. Stein, "Multitemporal change detection of urban trees using localized region-based active contours in VHR images," *Remote Sensing Environ.*, vol. 124, pp. 413–426, 2012. doi: 10.1016/j.rse.2012.05.027.
- [18] J. Gong, C. Liu, and X. Huang, "Advances in urban information extraction from high-resolution remote sensing imagery," *Sci. China Earth Sci.*, vol. 63, no. 4, pp. 463–475, 2020. doi: 10.1007/s11430-019-9547-x.
- [19] R. Qin, X. Huang, A. Gruen, and G. Schmitt, "Object-based 3-D building change detection on multitemporal stereo images," *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 8, no. 5, pp. 2125–2137, 2015. doi: 10.1109/JSTARS.2015.2424275.
- [20] D. Liu et al., "Integration of historical map and aerial imagery to characterize long-term land-use change and landscape dynamics: An object-based analysis via Random Forests," *Ecol. Indicators*, vol. 95, pp. 595–605, Dec. 2018. doi: 10.1016/j.ecolind.2018.08.004.
- [21] X. Huang, D. Wen, J. Li, and R. Qin, "Multi-level monitoring of subtle urban changes for the megacities of China using high-resolution multi-view satellite imagery," *Remote Sens. Environ.*, vol. 196, pp. 56–75, July 2017. doi: 10.1016/j.rse.2017.05.001.
- [22] G. Xian and C. Homer, "Updating the 2001 National Land Cover Database impervious surface products to 2006 using Landsat imagery change detection methods," *Remote Sens. Environ.*, vol. 114, no. 8, pp. 1676–1686, 2010. doi: 10.1016/j.rse.2010.02.018.
- [23] M. Pesaresi et al., "A global human settlement layer from optical HR/VHR RS data: Concept and first results," *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 6, no. 5, pp. 2102–2131, 2013. doi: 10.1109/JSTARS.2013.2271445.
- [24] L. Bruzzone and F. Bovolo, "A novel framework for the design of change-detection systems for very-high-resolution remote sensing images," *Proc. IEEE*, vol. 101, no. 3, pp. 609–630, 2012. doi: 10.1109/JPROC.2012.2197169.
- [25] M. Lu, J. Chen, H. Tang, Y. Rao, P. Yang, and W. Wu, "Land cover change detection by integrating object-based data blending model of Landsat and MODIS," *Remote Sens. Environ.*, vol. 184, pp. 374–386, Oct. 2016. doi: 10.1016/j.rse.2016.07.028.
- [26] S. Ye, D. Chen, and J. Yu, "A targeted change-detection procedure by combining change vector analysis and post-classification approach," *ISPRS J. Photogrammetry Remote Sens.*, vol. 114, pp. 115–124, Apr. 2016. doi: 10.1016/j.isprsjrs.2016.01.018.
- [27] N. Longbotham, C. Chaapel, L. Bleiler, C. Padwick, W. J. Emery, and F. Pacifici, "Very high resolution multiangle urban classification analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 4, pp. 1155–1170, 2012. doi: 10.1109/TGRS.2011.2165548.
- [28] D. Poli, F. Remondino, E. Angiuli, and G. Agugiaro, "Radiometric and geometric evaluation of GeoEye-1, WorldView-2 and Pléiades-1A stereo images for 3D information extraction," *ISPRS J. Photogrammetry Remote Sens.*, vol. 100, pp. 35–47, 2015/02/01/, 2015. doi: 10.1016/j.isprsjrs.2014.04.007.
- [29] F. Pacifici, N. Longbotham, and W. J. Emery, "The importance of physical quantities for the analysis of multitemporal and multiangular optical very high spatial resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6241–6256, 2014. doi: 10.1109/TGRS.2013.2295819.
- [30] K. Jacobsen, "High resolution satellite imaging systems-an overview," *Photogrammetrie Fernerkundung Geoinf.*, vol. 2005, pp. 487–496, Jan. 2005.
- [31] D. Wen, X. Huang, L. Zhang, and J. A. Benediktsson, "A novel automatic change detection method for urban high-resolution remotely sensed imagery based on multiindex scene representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 609–625, 2016. doi: 10.1109/TGRS.2015.2463075.
- [32] N. Tatar, M. Saadateseresh, H. Arefi, and A. Hadavand, "A robust object-based shadow detection method for cloud-free high resolution satellite images over urban areas and water bodies," *Adv. Space Res.*, vol. 61, no. 11, pp. 2787–2800, 2018. doi: 10.1016/j.asr.2018.03.011.
- [33] A. Movia, A. Beinat, and F. Crosilla, "Shadow detection and removal in RGB VHR images for land use unsupervised classification," *ISPRS J. Photogrammetry Remote Sens.*, vol. 119, pp. 485–495, Sept. 2016. doi: 10.1016/j.isprsjrs.2016.05.004.
- [34] G. Liasis and S. Stavrou, "Satellite images analysis for shadow detection and building height estimation," *ISPRS J. Photogrammetry Remote Sens.*, vol. 119, pp. 437–450, Sept. 2016. doi: 10.1016/j.isprsjrs.2016.07.006.

- [35] N. Kadhim and M. Mourshed, "A shadow-overlapping algorithm for estimating building heights from VHR satellite images," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 1, pp. 8–12, 2018. doi: 10.1109/LGRS.2017.2762424.
- [36] X. Huang and L. Zhang, "A multidirectional and multiscale morphological index for automatic building extraction from multispectral GeoEye-1 imagery," *Photogrammetric Eng. Remote Sens.*, vol. 77, no. 7, pp. 721–732, 2011. doi: 10.14358/PERS.77.7.721.
- [37] H. Song, B. Huang, and K. Zhang, "Shadow detection and reconstruction in high-resolution satellite images via morphological filtering and example-based learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2545–2554, 2014. doi: 10.1109/TGRS.2013.2262722.
- [38] T. Blaschke et al., "Geographic object-based image analysis – Towards a new paradigm," *ISPRS J. Photogrammetry Remote Sens.*, vol. 87, pp. 180–191, Jan. 2014. doi: 10.1016/j.isprsjprs.2013.09.014.
- [39] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *IEEE Trans. on systems, man, and cybernetics*, vol. SMC-3, no. 6, pp. 610–621, 1973. doi: 10.1109/TSMC.1973.4309314.
- [40] M. Hall-Beyer, "GLCM texture: A tutorial, version v3.0," Univ. of Calgary, 2007. [Online]. Available: <http://www.fp.ucalgary.ca/mhallbey/tutorial.htm>
- [41] S. Yao, S. Pan, T. Wang, C. Zheng, W. Shen, and Y. Chong, "A new pedestrian detection method based on combined HOG and LSS features," *Neurocomputing*, vol. 151, pp. 1006–1014, Mar. 2015. doi: 10.1016/j.neucom.2014.08.080.
- [42] L. Zhang, X. Huang, B. Huang, and P. Li, "A pixel shape index coupled with spectral information for classification of high spatial resolution remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2950–2961, 2006.
- [43] K. Tan, X. Jin, A. Plaza, X. Wang, L. Xiao, and P. Du, "Automatic change detection in high-resolution remote sensing images by using a multiple classifier system and spectral-spatial features," *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 9, no. 8, pp. 3439–3451, 2016. doi: 10.1109/JSTARS.2016.2541678.
- [44] Z. Li, W. Shi, M. Hao, and H. Zhang, "Unsupervised change detection using spectral features and a texture difference measure for VHR remote-sensing images," *Int. J. Remote Sens.*, vol. 38, no. 23, pp. 7302–7315, 2017. doi: 10.1080/01431161.2017.1375616.
- [45] D. Peng and Y. Zhang, "Object-based change detection from satellite imagery by segmentation optimization and multi-features fusion," *Int. J. Remote Sens.*, vol. 38, no. 13, pp. 3886–3905, 2017. doi: 10.1080/01431161.2017.1308033.
- [46] L. Zhang, B. Zhong, and A. Yang, "Building change detection using object-oriented LBP feature map in very high spatial resolution imagery," in *Proc. 10th Int. Workshop on the Anal. Multitemporal Remote Sens. Images (MultiTemp)*, 2019, pp. 1–4. doi: 10.1109/Multi-Temp.2019.8866919.
- [47] H. Liu, M. Yang, J. Chen, J. Hou, and M. Deng, "Line-constrained shape feature for building change detection in VHR remote sensing imagery," *ISPRS Int. J. Geo-Inform.*, vol. 7, no. 10, p. 410, 2018. doi: 10.3390/ijgi7100410.
- [48] M. Dalla Mura, J. A. Benediktsson, F. Bovolo, and L. Bruzzone, "An unsupervised technique based on morphological filters for change detection in very high resolution images," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 3, pp. 433–437, 2008. doi: 10.1109/LGRS.2008.917726.
- [49] N. Falco, M. Dalla Mura, F. Bovolo, J. A. Benediktsson, and L. Bruzzone, "Change detection in VHR images based on morphological attribute profiles," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 3, pp. 636–640, 2013. doi: 10.1109/LGRS.2012.2222340.
- [50] S. Liu, Q. Du, X. Tong, A. Samat, L. Bruzzone, and F. Bovolo, "Multiscale morphological compressed change vector analysis for unsupervised multiple change detection," *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 10, no. 9, pp. 4124–4137, 2017. doi: /10.1109/JSTARS.2017.2712119.
- [51] X. Huang, L. Zhang, and T. Zhu, "Building change detection from multitemporal high-resolution remotely sensed images based on a morphological building index," *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 7, no. 1, pp. 105–115, 2014.
- [52] Y. Tang, X. Huang, and L. Zhang, "Fault-tolerant building change detection from urban high-resolution remote sensing imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 5, pp. 1060–1064, 2013.
- [53] X. Huang, T. Zhu, L. Zhang, and Y. Tang, "A novel building change index for automatic building change detection from high-resolution remote sensing imagery," *Remote sensing letters*, vol. 5, no. 8, pp. 713–722, 2014. doi: 10.1080/2150704X.2014.963732.
- [54] G. R. Cross and A. K. Jain, "Markov random field texture models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-5, no. 1, pp. 25–39, 1983. doi: 10.1109/TPAMI.1983.4767341.
- [55] T. Xu, I. D. Moore, and J. C. Gallant, "Fractals, fractal dimensions and landscapes—A review," *Geomorphology*, vol. 8, no. 4, pp. 245–262, 1993. doi: 10.1016/0169-555X(93)90022-T.
- [56] C. Benedek, M. Shadaydeh, Z. Kato, T. Szirányi, and J. Zerubia, "Multilayer Markov Random Field models for change detection in optical remote sensing images," *ISPRS J. Photogrammetry Remote Sensing*, vol. 107, pp. 22–37, Sept. 2015. doi: 10.1016/j.isprsjprs.2015.02.006.
- [57] L. Bruzzone and D. F. Prieto, "An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images," *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 452–466, 2002. doi: 10.1109/TIP.2002.999678.
- [58] A. Ghosh, B. N. Subudhi, and L. Bruzzone, "Integration of Gibbs Markov random field and Hopfield-type neural networks for unsupervised change detection in remotely sensed multitemporal images," *IEEE Trans. Image Process.*, vol. 22, no. 8, pp. 3087–3096, 2013. doi: 10.1109/TIP.2013.2259833.
- [59] B. N. Subudhi, F. Bovolo, A. Ghosh, and L. Bruzzone, "Spatio-contextual fuzzy clustering with Markov random field model for change detection in remotely sensed images," *Optics Laser Technol.*, vol. 57, pp. 284–292, Apr. 2014. doi: 10.1016/j.optlastec.2013.10.003.
- [60] H. Yu, W. Yang, G. Hua, H. Ru, and P. Huang, "Change detection using high resolution remote sensing images based on active learning and Markov random fields," *Remote Sensing*, vol. 9, no. 12, p. 1233, 2017. doi: 10.3390/rs9121233.

- [61] S. Aleksandrowicz, A. Wawrzaszek, W. Drzewiecki, and M. Krupiński, "Change detection using global and local multifractal description," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 8, pp. 1183–1187, 2016. doi: 10.1109/LGRS.2016.2574940.
- [62] S. Luan, C. Chen, B. Zhang, J. Han, and J. Liu, "Gabor convolutional networks," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4357–4366, 2017.
- [63] Z. Li, W. Shi, H. Zhang, and M. Hao, "Change detection based on Gabor wavelet features for very high resolution remote sensing images," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 5, pp. 783–787, 2017. doi: 10.1109/LGRS.2017.2681198.
- [64] C. Wei, P. Zhao, X. Li, Y. Wang, and F. Liu, "Unsupervised change detection of VHR remote sensing images based on multi-resolution Markov Random Field in wavelet domain," *Int. J. Remote Sens.*, vol. 40, no. 20, pp. 7750–7766, 2019. doi: 10.1080/01431161.2019.1602792.
- [65] Q. Li, X. Huang, D. Wen, and H. Liu, "Integrating multiple textural features for remote sensing image change detection," *Photogrammetric Eng. Remote Sens.*, vol. 83, no. 2, pp. 109–121, 2017. doi: 10.14358/PERS.83.2.109.
- [66] B. Hou, Y. Wang, and Q. Liu, "Change detection based on deep features and low rank," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2418–2422, 2017. doi: 10.1109/LGRS.2017.2766840.
- [67] S. Saha, F. Bovolo, and L. Bruzzone, "Unsupervised deep change vector analysis for multiple-change detection in VHR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 6, pp. 3677–3693, 2019. doi: 10.1109/TGRS.2018.2886643.
- [68] T. Zhan, M. Gong, J. Liu, and P. Zhang, "Iterative feature mapping network for detecting multiple changes in multi-source remote sensing images," *ISPRS J. Photogrammetry Remote Sens.*, vol. 146, pp. 38–51, Dec. 2018. doi: 10.1016/j.isprsjprs.2018.09.002.
- [69] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *J. Mach. Learning Res.*, vol. 11, pp. 3371–3408, Dec. 2010.
- [70] P. Zhang, M. Gong, L. Su, J. Liu, and Z. Li, "Change detection based on deep feature representation and mapping transformation for multi-spatial-resolution remote sensing images," *ISPRS J. Photogrammetry Remote Sens.*, vol. 116, pp. 24–41, June 2016. doi: 10.1016/j.isprsjprs.2016.02.013.
- [71] L. Su, M. Gong, P. Zhang, M. Zhang, J. Liu, and H. Yang, "Deep learning and mapping based ternary change detection for information unbalanced images," *Pattern Recogn.*, vol. 66, pp. 213–228, June 2017. doi: 10.1016/j.patcog.2017.01.002.
- [72] G. Liu, L. Li, L. Jiao, Y. Dong, and X. Li, "Stacked Fisher autoencoder for SAR change detection," *Pattern Recogn.*, vol. 96, p. 106,971, Dec. 2019. doi: 10.1016/j.patcog.2019.106971.
- [73] N. Lv, C. Chen, T. Qiu, and A. K. Sangaiah, "Deep learning and superpixel feature extraction based on contractive autoencoder for change detection in SAR images," *IEEE Trans. Ind. Inf.*, vol. 14, no. 12, pp. 5530–5538, 2018. doi: 10.1109/TII.2018.2873492.
- [74] X. X. Zhu et al., "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geosci. Remote Sens. Mag. (replaces Newstlett.)*, vol. 5, no. 4, pp. 8–36, 2017. doi: 10.1109/MGRS.2017.2762307.
- [75] T. Zhan, M. Gong, X. Jiang, and M. Zhang, "Unsupervised scale-driven change detection with deep spatial-spectral features for VHR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 8, pp. 1–13, 2020. doi: 10.1109/TGRS.2020.2968098.
- [76] S. Saha, L. Mou, C. Qiu, X. X. Zhu, F. Bovolo, and L. Bruzzone, "Unsupervised deep joint segmentation of multitemporal high-resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 12, pp. 1–13, 2020. doi: 10.1109/TGRS.2020.2990640.
- [77] Q. Wang, X. Zhang, G. Chen, F. Dai, Y. Gong, and K. Zhu, "Change detection based on Faster R-CNN for high-resolution remote sensing images," *Remote Sensing Letters*, vol. 9, no. 10, pp. 923–932, 2018. doi: 10.1080/2150704X.2018.1492172.
- [78] J. Liu et al., "Convolutional neural network-based transfer learning for optical aerial images change detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 1, pp. 127–131, 2019. doi: 10.1109/LGRS.2019.2916601.
- [79] M. Volpi and D. Tuia, "Dense semantic labeling of subdecimeter resolution images with convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 881–893, 2016. doi: 10.1109/TGRS.2016.2616585.
- [80] L. Gueguen and R. Hamid, "Toward a generalizable image representation for large-scale change detection: Application to generic damage analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3378–3387, 2016. doi: 10.1109/TGRS.2016.2516402.
- [81] R. Caye Daudt, B. Le Saux, A. Boulch, and Y. Gousseau, "Multi-task learning for large-scale semantic change detection," *Comput. Vision Image Understanding*, vol. 187, p. 102783, 2019. doi: 10.1016/j.cviu.2019.07.003.
- [82] R. Gupta et al., "Creating xBD: A dataset for assessing building damage from satellite imagery," in *Proc. IEEE Conf. Comput. Vision and Pattern Recogn. Workshops*, 2019, pp. 10–17.
- [83] J. Zhu, Y. Su, Q. Guo, and T. C. Harmon, "Unsupervised object-based differencing for land-cover change detection," *Photogrammetric Eng. Remote Sens.*, vol. 83, no. 3, pp. 225–236, 2017. doi: 10.14358/PERS.83.3.225.
- [84] D. Ming, J. Li, J. Wang, and M. Zhang, "Scale parameter selection by spatial statistics for GeOBIA: Using mean-shift based multi-scale segmentation as an example," *ISPRS J. Photogrammetry Remote Sens.*, vol. 106, pp. 28–41, Aug. 2015. doi: 10.1016/j.isprsjprs.2015.04.010.
- [85] P. Xiao, M. Yuan, X. Zhang, X. Feng, and Y. Guo, "Cosegmentation for object-based building change detection from high-resolution remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 3, pp. 1587–1603, 2017. doi: 10.1109/TGRS.2016.2627638.
- [86] Y. Liu, Q. Guo, and M. Kelly, "A framework of region-based spatial relations for non-overlapping features and its application in object based image analysis," *ISPRS J. Photogrammetry Remote Sens.*, vol. 63, no. 4, pp. 461–475, 2008. doi: 10.1016/j.isprsjprs.2008.01.007.
- [87] M. Kim and M. Madden, Xu, Bo, "GEOBIA vegetation mapping in great smoky mountains national park with spectral and non-spectral ancillary information," *Photogrammetric Eng. Remote Sensing*, vol. 76, no. 2, pp. 137–149, 2010. doi: 10.14358/PERS.76.2.137.
- [88] Z. Lv, T. Liu, and J. A. Benediktsson, "Object-oriented key point vector distance for binary land cover change detection

- using VHR remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 9, pp. 6524–6533, 2020. doi: 10.1109/TGRS.2020.2977248.
- [89] F. Bovolo, "A multilevel parcel-based approach to change detection in very high resolution multitemporal images," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 1, pp. 33–37, 2009. doi: 10.1109/LGRS.2008.2007429.
- [90] C. Geiß, M. Klotz, A. Schmitt, and H. Taubenböck, "Object-based morphological profiles for classification of remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 5952–5963, 2016. doi: 10.1109/TGRS.2016.2576978.
- [91] J. Liang et al., "A comparison of two object-oriented methods for land-use/cover change detection with SPOT 5 imagery," *Sensor Lett.*, vol. 10, no. 1, pp. 415–424, 2012. doi: 10.1166/sl.2012.1865.
- [92] W. Yu, W. Zhou, Y. Qian, and J. Yan, "A new approach for land cover classification and change analysis: Integrating backdating and an object-based method," *Remote Sensing Environment*, vol. 177, pp. 37–47, May 2016. doi: 10.1016/j.rse.2016.02.030.
- [93] X. Zhang, S. Du, Q. Wang, and W. Zhou, "Multiscale geoscene segmentation for extracting urban functional zones from VHR satellite images," *Remote Sens.*, vol. 10, no. 2, p. 281, 2018. doi: 10.3390/rs10020281.
- [94] H. Liu, X. Huang, D. Wen, and J. Li, "The use of landscape metrics and transfer learning to explore urban villages in China," *Remote Sens.*, vol. 9, no. 4, p. 365, 2017. doi: 10.3390/rs9040365.
- [95] J. Zhou, B. Yu, and J. Qin, "Multi-level spatial analysis for change detection of urban vegetation at individual tree scale," *Remote Sens.*, vol. 6, no. 9, pp. 9086–9103, 2014. doi: 10.3390/rs6099086.
- [96] M. A. Aguilar, M. D. M. Saldana, and F. J. Aguilar, "Generation and quality assessment of stereo-extracted DSM from GeoEye-1 and WorldView-2 imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 2, pp. 1259–1271, 2013. doi: 10.1109/TGRS.2013.2249521.
- [97] X. Huang, H. Chen, and J. Gong, "Angular difference feature extraction for urban scene classification using ZY-3 multi-angle high-resolution satellite imagery," *ISPRS J. Photogrammetry Remote Sens.*, vol. 135, pp. 127–141, Jan. 2018. doi: 10.1016/j.isprsjprs.2017.11.017.
- [98] H. Chaabouni-Chouayakh, I. R. Arnau, and P. Reinartz, "Towards automatic 3-D change detection through multispectral and digital elevation model information fusion," *Int. J. Image Data Fusion*, vol. 4, no. 1, pp. 89–101, 2013. doi: 10.1080/19479832.2012.739577.
- [99] J. Tian, P. Reinartz, P. d'Angelo, and M. Ehlers, "Region-based automatic building and forest change detection on Cartosat-1 stereo imagery," *ISPRS J. Photogrammetry Remote Sens.*, vol. 79, pp. 226–239, May 2013. doi: 10.1016/j.isprsjprs.2013.02.017.
- [100] W. Tu et al., "Portraying urban functional zones by coupling remote sensing imagery and human sensing data," *Remote Sens.*, vol. 10, no. 1, p. 141, 2018. doi: 10.3390/rs10010141.
- [101] C. Liu, X. Huang, Z. Zhu, H. Chen, X. Tang, and J. Gong, "Automatic extraction of built-up area from ZY3 multi-view satellite imagery: Analysis of 45 global cities," *Remote Sens. Environ.*, vol. 226, pp. 51–73, June 2019. doi: 10.1016/j.rse.2019.03.033.
- [102] R. Duca and F. D. Frate, "Hyperspectral and multiangle CHRIS-PROBA images for the generation of land cover maps," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 2857–2866, 2008. doi: 10.1109/TGRS.2008.2000741.
- [103] Y. Yan, L. Deng, X. Liu, and L. Zhu, "Application of UAV-based multi-angle hyperspectral remote sensing in fine vegetation classification," *Remote Sens.*, vol. 11, no. 23, p. 2753, 2019. doi: 10.3390/rs11232753.
- [104] M. Zanetti and L. Bruzzone, "A theoretical framework for change detection based on a compound multiclass statistical model of the difference image," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 1129–1143, 2017. doi: 10.1109/TGRS.2017.2759663.
- [105] Y. T. Solano-Correa, F. Bovolo, and L. Bruzzone, "An approach to multiple change detection in VHR optical images based on iterative clustering and adaptive thresholding," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 8, pp. 1–5, 2019. doi: 10.1109/LGRS.2019.2896385.
- [106] J. S. Deng, K. Wang, Y. H. Deng, and G. J. Qi, "PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data," *Int. J. Remote Sens.*, vol. 29, no. 16, pp. 4823–4838, 2008. doi: 10.1080/01431160801950162.
- [107] A. Tahraoui, R. Kheddami, A. Bouakache, and A. Belhadj-Aissa, "Land change detection using multivariate alteration detection and Chi squared test thresholding," in *Proc. 4th Int. Conf. Adv. Technol. Signal and Image Process. (ATSIP)*, 2018, pp. 1–6. doi: 10.1109/ATSIP.2018.8364501.
- [108] C. Wu, L. Zhang, and L. Zhang, "A scene change detection framework for multi-temporal very high resolution remote sensing images," *Signal Process.*, vol. 124, pp. 184–197, July 2016. doi: 10.1016/j.sigpro.2015.09.020.
- [109] X. Zhang, R. Fan, L. Ma, X. Liao, and X. Chen, "Change detection in very high-resolution images based on ensemble CNNs," *Int. J. Remote Sens.*, vol. 41, no. 12, pp. 4757–4779, 2020. doi: 10.1080/01431161.2020.1723818.
- [110] D. Peng, Y. Zhang, and H. Guan, "End-to-end change detection for high resolution satellite images using improved UNet++," *Remote Sens.*, vol. 11, no. 11, p. 1382, 2019. doi: 10.3390/rs11111382.
- [111] L. Mou, L. Bruzzone, and X. X. Zhu, "Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 924–935, 2019. doi: 10.1109/TGRS.2018.2863224.
- [112] T. Bao, C. Fu, T. Fang, and H. Huo, "PPCNET: A combined patch-level and pixel-level end-to-end deep network for high-resolution remote sensing image change detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 10, pp. 1–5, 2020. doi: 10.1109/LGRS.2019.2955309.
- [113] C. Zhang et al., "A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images," *ISPRS J. Photogrammetry Remote Sensing*, vol. 166, pp. 183–200, Aug. 2020. doi: 10.1016/j.isprsjprs.2020.06.003.
- [114] T. Lei, Y. Zhang, Z. Lv, S. Li, S. Liu, and A. K. Nandi, "Landslide inventory mapping from bitemporal images using deep convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 6, pp. 982–986, 2019. doi: 10.1109/LGRS.2018.2889307.

- [115] W. Wiratama, J. Lee, S.-E. Park, and D. Sim, "Dual-dense convolution network for change detection of high-resolution panchromatic imagery," *Appl. Sci.*, vol. 8, no. 10, p. 1785, 2018. doi: 10.3390/app8101785.
- [116] W. Zhang and X. Lu, "The spectral-spatial joint learning for change detection in multispectral imagery," *Remote Sens.*, vol. 11, no. 3, p. 240, 2019. doi: 10.3390/rs11030240.
- [117] A. Song and J. Choi, "Fully convolutional networks with multiscale 3D filters and transfer learning for change detection in high spatial resolution satellite images," *Remote Sens.*, vol. 12, no. 5, p. 799, 2020. doi: 10.3390/rs12050799.
- [118] M. Zhai, H. Liu, and F. Sun, "Lifelong learning for scene recognition in remote sensing images," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 9, pp. 1472–1476, 2019. doi: 10.1109/LGRS.2019.2897652.
- [119] M. Rußwurm, S. Wang, M. Korner, and D. Lobell, "Meta-learning for few-shot land cover classification," in *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recogn. Workshops*, 2020, pp. 200–201.
- [120] R. Hedjam, A. Abdesselam, and F. Melgani, "Change detection in unlabeled optical remote sensing data using Siamese CNN," *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 13, pp. 4178–4187, July 2020. doi: 10.1109/JSTARS.2020.3009116.
- [121] H. Chen, C. Wu, B. Du, L. Zhang, and L. Wang, "Change detection in multisource VHR images via deep siamese convolutional multiple-layers recurrent neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2848–2864, 2020. doi: 10.1109/TGRS.2019.2956756.
- [122] J. Liu, M. Gong, A. K. Qin, and K. C. Tan, "Bipartite differential neural network for unsupervised image change detection," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 3, pp. 876–890, 2020. doi: 10.1109/TNNLS.2019.2910571.
- [123] X. Junfeng, Z. Baoming, G. Haitao, L. Jun, and L. Yuzhun, "Combining iterative slow feature analysis and deep feature learning for change detection in high-resolution remote sensing images," *J. Appl. Remote Sens.*, vol. 13, no. 2, pp. 1–16, 2019. doi: 10.1117/1.JRS.13.024506.
- [124] J. Fan, K. Lin, and M. Han, "A novel joint change detection approach based on weight-clustering sparse autoencoders," *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 12, no. 2, pp. 685–699, 2019. doi: 10.1109/JSTARS.2019.2892951.
- [125] A. Argyridis and D. P. Argialas, "Building change detection through multi-scale GEOBIA approach by integrating deep belief networks with fuzzy ontologies," *Int. J. Image Data Fusion*, vol. 7, no. 2, pp. 148–171, 2016.
- [126] P. F. Alcantarilla, S. Stent, G. Ros, R. Arroyo, and R. Gherardi, "Street-view change detection with deconvolutional networks," *Autonom. Robots*, vol. 42, no. 7, pp. 1301–1322, 2018. doi: 10.1007/s10514-018-9734-5.
- [127] R. Jing et al., "Object-based change detection for VHR remote sensing images based on a Trisiamese-LSTM," *Int. J. Remote Sens.*, vol. 41, no. 16, pp. 6209–6231, 2020. doi: 10.1080/01431161.2020.1734253.
- [128] J. Geng, J. Fan, H. Wang, and X. Ma, "Change detection of marine reclamation using multispectral images via patch-based recurrent neural network," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, 2017, pp. 612–615. doi: 10.1109/IGARSS.2017.8127028.
- [129] H. Lyu, H. Lu, and L. Mou, "Learning a transferable change rule from a recurrent neural network for land cover change detection," *Remote Sens.*, vol. 8, no. 6, p. 506, 2016. doi: 10.3390/rs8060506.
- [130] M. Gong, X. Niu, P. Zhang, and Z. Li, "Generative adversarial networks for change detection in multispectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2310–2314, 2017. doi: 10.1109/LGRS.2017.2762694.
- [131] M. Gong, Y. Yang, T. Zhan, X. Niu, and S. Li, "A generative discriminatory classified network for change detection in multispectral imagery," *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 12, no. 1, pp. 321–333, 2019. doi: 10.1109/JSTARS.2018.2887108.
- [132] S. Saha, L. Mou, X. X. Zhu, F. Bovolo, and L. Bruzzone, "Semisupervised change detection using graph convolutional network," *IEEE Geosci. Remote Sens. Lett.*, 2020.
- [133] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Information Process. Syst.*, 2012, pp. 1097–1105.
- [134] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, arXiv:1409.1556.
- [135] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," 2016, arXiv:1602.07261.
- [136] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vision Pattern Recogn.*, 2016, pp. 770–778.
- [137] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vision Pattern Recogn.*, 2017, pp. 4700–4708.
- [138] Y. Wu, Z. Bai, Q. Miao, W. Ma, Y. Yang, and M. Gong, "A classified adversarial network for multi-spectral remote sensing image change detection," *Remote Sensing*, vol. 12, no. 13, p. 2098, 2020. doi: 10.3390/rs12132098.
- [139] B. Fang, G. Chen, L. Pan, R. Kou, and L. Wang, "GAN-based Siamese framework for landslide inventory mapping using bi-temporal optical remote sensing images," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 11, pp. 1–5, 2020. doi: 10.3390/rs11111292.
- [140] A. Song, Y. Kim, and Y. Han, "Uncertainty analysis for object-based change detection in very high-resolution satellite images using deep learning network," *Remote Sensing*, vol. 12, no. 15, p. 2345, 2020. doi: 10.3390/rs12152345.
- [141] S. I. Toure, D. A. Stow, H.-c. Shih, J. Weeks, and D. Lopez-Carr, "Land cover and land use change analysis using multi-spatial resolution data and object-based image analysis," *Remote Sens. Environ.*, vol. 210, pp. 259–268, June 2018. doi: 10.1016/j.rse.2018.03.023.
- [142] X. Wang, S. Liu, P. Du, H. Liang, J. Xia, and Y. Li, "Object-based change detection in urban areas from high spatial resolution images based on multiple features and ensemble learning," *Remote Sens.*, vol. 10, no. 2, 2018. doi: 10.3390/rs10020276.
- [143] M. Chini, C. Bignami, A. Chiancone, and S. Stramondo, "Classification of VHR optical data for land use change analysis by scale object selection (SOS) algorithm," in *Proc. IEEE Geosci. Remote Sens. Symp.*, 2014, pp. 2834–2837.

- [144] X. Huang, X. Han, S. Ma, T. Lin, and J. Gong, "Monitoring ecosystem service change in the City of Shenzhen by the use of high-resolution remotely sensed imagery and deep learning," *Land Degradation Develop.*, vol. 30, no. 12, 2019. doi: 10.1002/ldr.3337.
- [145] G. Doxani, K. Karantzas, and M. Tsakiri-Strati, "Monitoring urban changes based on scale-space filtering and object-oriented classification," *Int. J. Appl. Earth Observat. Geoinf.*, vol. 15, pp. 38–48, Apr. 2012. doi: 10.1016/j.jag.2011.07.002.
- [146] Z. Guo and S. Du, "Mining parameter information for building extraction and change detection with very high-resolution imagery and GIS data," *GISci. Remote Sens.*, vol. 54, no. 1, pp. 38–63, 2017. doi: 10.1080/15481603.2016.1250328.
- [147] B. Hou, Y. Wang, and Q. Liu, "A saliency guided semi-supervised building change detection method for high resolution remote sensing images," *Sensors*, vol. 16, no. 9, p. 1377, 2016. doi: 10.3390/s16091377.
- [148] X. Huang, H. Liu, and L. Zhang, "Spatiotemporal detection and analysis of urban villages in mega city regions of China using high-resolution remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3639–3657, 2015. doi: 10.1109/TGRS.2014.2380779.
- [149] M. Janalipour and A. Mohammadzadeh, "Building damage detection using object-based image analysis and ANFIS from high-resolution image (case study: BAM earthquake, Iran)," *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 9, no. 5, pp. 1937–1945, 2016. doi: 10.1109/JSTARS.2015.2458582.
- [150] T. Leichtle, C. Geiß, M. Wurm, T. Lakes, and H. Taubenböck, "Unsupervised change detection in VHR remote sensing imagery—an object-based clustering approach in a dynamic urban environment," *Int. J. Appl. Earth Observat. Geoinf.*, vol. 54, pp. 15–27, 2017. doi: 10.1016/j.jag.2016.08.010.
- [151] Y. Li, X. Huang, and H. Liu, "Unsupervised deep feature learning for urban village detection from high-resolution remote sensing images," *Photogrammetric Eng. Remote Sensing*, vol. 83, no. 8, pp. 567–579, 2017. doi: 10.14358/PERS.83.8.567.
- [152] S. Radhika, Y. Tamura, and M. Matsui, "Cyclone damage detection on building structures from pre-and post-satellite images using wavelet based pattern recognition," *J. Wind Eng. Ind. Aerodynamics*, vol. 136, pp. 23–33, 2015. doi: 10.1016/j.jweia.2014.10.018.
- [153] N. Sofina and M. Ehlers, "Building change detection using high resolution remotely sensed data and GIS," *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 9, no. 8, pp. 3430–3438, 2016. doi: 10.1109/JSTARS.2016.2542074.
- [154] X. Tong et al., "Use of shadows for detection of earthquake-induced collapsed buildings in high-resolution satellite imagery," *ISPRS J. Photogrammetry Remote Sensing*, vol. 79, pp. 53–67, 2013. doi: 10.1016/j.isprsjprs.2013.01.012.
- [155] D. Wen, X. Huang, A. Zhang, and X. Ke, "Monitoring 3D building change and urban redevelopment patterns in inner city areas of Chinese megacities using multi-view satellite imagery," *Remote Sens.*, vol. 11, no. 7, p. 763, 2019. doi: 10.3390/rs11070763.
- [156] J. Tian, S. Cui, and P. Reinartz, "Building change detection based on satellite stereo imagery and digital surface models," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 406–417, 2014. doi: 10.1109/TGRS.2013.2240692.
- [157] R. Qin, "Change detection on LOD 2 building models with very high resolution spaceborne stereo imagery," *ISPRS J. Photogrammetry Remote Sens.*, vol. 96, pp. 179–192, Oct. 2014. doi: 10.1016/j.isprsjprs.2014.07.007.
- [158] A. Kovacs and T. Sziranyi, "Orientation based building outline extraction in aerial images," *ISPRS Ann. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, vol. I-7, pp. 141–146, July 2012. doi: 10.5194/isprsannals-I-7-141-2012.
- [159] A. O. Ok, "Automated detection of buildings from single VHR multispectral images using shadow information and graph cuts," *ISPRS J. Photogrammetry Remote Sensing*, vol. 86, pp. 21–40, Dec. 2013. doi: 10.1016/j.isprsjprs.2013.09.004.
- [160] M. Vakalopoulou, K. Karantzas, N. Komodakis, and N. Paragios, "Building detection in very high resolution multispectral data with deep learning features," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, 2015, pp. 1873–1876.
- [161] M. Janalipour and M. Taleai, "Building change detection after earthquake using multi-criteria decision analysis based on extracted information from high spatial resolution satellite images," *Int. J. Remote Sens.*, vol. 38, no. 1, pp. 82–99, 2017. doi: 10.1080/01431161.2016.1259673.
- [162] X. Huang and Y. Wang, "Investigating the effects of 3D urban morphology on the surface urban heat island effect in urban functional zones by using high-resolution remote sensing data: A case study of Wuhan, Central China," *ISPRS J. Photogrammetry Remote Sens.*, vol. 152, pp. 119–131, June 2019. doi: 10.1016/j.isprsjprs.2019.04.010.
- [163] M. Turker and B. Cetinkaya, "Automatic detection of earthquake-damaged buildings using DEMs created from pre- and post-earthquake stereo aerial photographs," *Int. J. Remote Sens.*, vol. 26, no. 4, pp. 823–832, 2005. doi: 10.1080/01431160512331316810.
- [164] R. Qin, "A critical analysis of satellite stereo pairs for digital surface model generation and a matching quality prediction model," *ISPRS J. Photogrammetry Remote Sensing*, vol. 154, pp. 139–150, Aug. 2019. doi: 10.1016/j.isprsjprs.2019.06.005.
- [165] S. Ji, J. Liu, and M. Lu, "CNN-based dense image matching for aerial remote sensing images," *Photogrammetric Eng. Remote Sens.*, vol. 85, no. 6, pp. 415–424, 2019. doi: 10.14358/PERS.85.6.415.
- [166] Y. Lü et al., "Recent ecological transitions in China: Greening, browning and influential factors," *Sci. Rep.*, vol. 5, no. 1, p. 8732, 2015. doi: 10.1038/srep08732.
- [167] J. Verbesselt, R. Hyndman, G. Newnham, and D. Culvenor, "Detecting trend and seasonal changes in satellite image time series," *Remote Sens. Environ.*, vol. 114, no. 1, pp. 106–115, 2010. doi: 10.1016/j.rse.2009.08.014.
- [168] R. Pu and S. Landry, "Evaluating seasonal effect on forest leaf area index mapping using multi-seasonal high resolution satellite pléiades imagery," *Int. J. Appl. Earth Observat. Geoinf.*, vol. 80, pp. 268–279, Aug. 2019. doi: 10.1016/j.jag.2019.04.020.
- [169] J. Wang, D. Yang, M. Detto, B. W. Nelson, M. Chen, K. Guan, et al. "Multi-scale integration of satellite remote sensing improves characterization of dry-season green-up in an Amazon tropical evergreen forest," *Remote Sens. Environ.*, vol. 246, p. 111,865, 2020. doi: 10.1016/j.rse.2020.111865.

- [170] P. Gärtner, M. Förster, A. Kurban, and B. Kleinschmit, "Object based change detection of Central Asian Tugai vegetation with very high spatial resolution satellite imagery," *Int. J. Appl. Earth Observat. Geoinf.*, vol. 31, pp. 110–121, Sept. 2014. doi: 10.1016/j.jag.2014.03.004.
- [171] J. Tian, T. Schneider, C. Straub, F. Kugler, and P. Reinartz, "Exploring digital surface models from nine different sensors for forest monitoring and change detection," *Remote Sens.*, vol. 9, no. 3, p. 287, 2017. doi: 10.3390/rs9030287.
- [172] R. Dalagnol et al., "Quantifying canopy tree loss and gap recovery in tropical forests under low-intensity logging using VHR satellite imagery and airborne LiDAR," *Remote Sensing*, vol. 11, no. 7, p. 817, 2019. doi: 10.3390/rs11070817.
- [173] J. P. Ardila, W. Bijker, V. A. Tolpekin, and A. Stein, "Quantification of crown changes and change uncertainty of trees in an urban environment," *ISPRS J. Photogrammetry Remote Sens.*, vol. 74, pp. 41–55, 2012. doi: 10.1016/j.isprsjprs.2012.08.007.
- [174] B. Lu and Y. He, "Species classification using Unmanned Aerial Vehicle (UAV)-acquired high spatial resolution imagery in a heterogeneous grassland," *ISPRS J. Photogrammetry Remote Sens.*, vol. 128, pp. 73–85, 2017. doi: 10.1016/j.isprsjprs.2017.03.011.
- [175] Y. Sun, Q. Xin, J. Huang, B. Huang, and H. Zhang, "Characterizing tree species of a tropical wetland in southern China at the individual tree level based on convolutional neural network," *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 12, no. 11, pp. 4415–4425, 2019. doi: 10.1109/JSTARS.2019.2950721.
- [176] Z. Xie, Y. Chen, D. Lu, G. Li, and E. Chen, "Classification of land cover, forest, and tree species classes with ZiYuan-3 multispectral and stereo data," *Remote Sens.*, vol. 11, no. 2, p. 164, 2019. doi: 10.3390/rs11020164.
- [177] R. Pu, S. Landry, and Q. Yu, "Assessing the potential of multi-seasonal high resolution Pleiades satellite imagery for mapping urban tree species," *Int. J. Appl. Earth Observat. Geoinf.*, vol. 71, pp. 144–158, Sept. 2018. doi: 10.1016/j.jag.2018.05.005.
- [178] S. Hartling, V. Sagan, P. Sidike, M. Maimaitijiang, and J. Carron, "Urban tree species classification using a WorldView-2/3 and LiDAR data fusion approach and deep learning," *Sensors*, vol. 19, no. 6, p. 1284, 2019. doi: 10.3390/s19061284.
- [179] D. Wen, X. Huang, H. Liu, W. Liao, and L. Zhang, "Semantic classification of urban trees using very high resolution satellite imagery," *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 10, no. 4, pp. 1413–1424, 2017. doi: 10.1109/JSTARS.2016.2645798.
- [180] M. Xia, Y. Zhang, Z. Zhang, J. Liu, W. Ou, and W. Zou, "Modeling agricultural land use change in a rapid urbanizing town: Linking the decisions of government, peasant households and enterprises," *Land Use Policy*, vol. 90, pp. 104266, 2020. doi: 10.1016/j.landusepol.2019.104266.
- [181] E. S. Malinverni, M. Rinaldi, and S. Ruggieri, "Agricultural crop change detection by means of hybrid classification and high resolution images," *EARSeL eProc.*, vol. 11, no. 2, pp. 132–154, 2012.
- [182] Y. Sadeh, X. Zhu, K. Chenu, and D. Dunkerley, "Sowing date detection at the field scale using CubeSats remote sensing," *Comput. Electron. Agriculture*, vol. 157, pp. 568–580, 2019. doi: 10.1016/j.compag.2019.01.042.
- [183] J. Bendig, A. Bolten, and G. Bareth, "UAV-based imaging for multi-temporal, very high resolution crop surface models to monitor crop growth variability monitoring des Pflanzenwachstums mit Hilfe multitemporaler und hoch auflösender Oberflächenmodelle von Getreidebeständen auf Basis von Bildern aus UAV-Befliegungen," *Photogrammetrie-Fermerkundung-Geoinf.*, vol. 2013, pp. 551–562, Dec. 2013.
- [184] P. L. Hatfield and P. J. Pinter, "Remote sensing for crop protection," *Crop Protection*, vol. 12, no. 6, pp. 403–413, 1993. doi: 10.1016/0261-2194(93)90001-Y.
- [185] K. Johansen et al., "Using GeoEye-1 imagery for multi-temporal object-based detection of canegrub damage in sugarcane fields in Queensland, Australia," *GISci. Remote Sensing*, vol. 55, no. 2, pp. 285–305, 2018. doi: 10.1080/15481603.2017.1417691.
- [186] J. Franke and G. Menz, "Multi-temporal wheat disease detection by multi-spectral remote sensing," *Precision Agriculture*, vol. 8, no. 3, pp. 161–172, 2007. doi: 10.1007/s11119-007-9036-y.
- [187] L. Yuan, Y. Huang, R. W. Loraamm, C. Nie, J. Wang, and J. Zhang, "Spectral analysis of winter wheat leaves for detection and differentiation of diseases and insects," *Field Crops Res.*, vol. 156, pp. 199–207, 2014. doi: 10.1016/j.fcr.2013.11.012.
- [188] A. M. Mouazen et al., "Chapter 2—Monitoring," in *Agricultural Internet of Things and Decision Support for Precision Smart Farming*, A. Castrignanò, G. Buttafuoco, R. Khosla, A. M. Mouazen, D. Moshou, and O. Naud, Eds. New York: Academic Press, 2020, pp. 35–138.
- [189] X. Zhang et al., "A deep learning-based approach for automated yellow rust disease detection from high-resolution hyperspectral UAV images," *Remote Sens.*, vol. 11, no. 13, p. 1554, 2019. doi: 10.3390/rs11131554.
- [190] Y. Wang and H. Yésou, "Remote sensing of floodpath lakes and wetlands: A challenging frontier in the monitoring of changing environments," *Remote Sens.*, vol. 10, no. 12, p. 1955, 2018. doi: 10.3390/rs10121955.
- [191] C. Xie, X. Huang, H. Mu, and W. Yin, "Impacts of land-use changes on the lakes across the Yangtze floodplain in China," *Environ. Sci. Technol.*, vol. 51, no. 7, pp. 3669–3677, 2017. doi: 10.1021/acs.est.6b04260.
- [192] S. Wang et al., "Changes of water clarity in large lakes and reservoirs across China observed from long-term MODIS," *Remote Sens. Environ.*, vol. 247, pp. 111949, 2020. doi: 10.1016/j.rse.2020.111949.
- [193] J.-F. Pekel, A. Cottam, N. Gorelick, and A. S. Belward, "High-resolution mapping of global surface water and its long-term changes," *Nature*, vol. 540, no. 7633, pp. 418–422, 2016. doi: 10.1038/nature20584.
- [194] J. A. Downing et al., "The global abundance and size distribution of lakes, ponds, and impoundments," *Limnol. Oceanogr.*, vol. 51, no. 5, pp. 2388–2397, 2006. doi: 10.4319/lo.2006.51.5.2388.
- [195] S. W. Cooley, L. C. Smith, L. Stepan, and J. Mascaro, "Tracking dynamic northern surface water changes with high-frequency planet CubeSat imagery," *Remote Sens.*, vol. 9, no. 12, p. 1306, 2017. doi: 10.3390/rs9121306.
- [196] W. Feng, H. Sui, W. Huang, C. Xu, and K. An, "Water body extraction from very high-resolution remote sensing imagery

- using deep U-Net and a superpixel-based conditional random field model," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 4, pp. 618–622, 2019. doi: 10.1109/LGRS.2018.2879492.
- [197] F. Chen, X. Chen, T. Van de Voorde, D. Roberts, H. Jiang, and W. Xu, "Open water detection in urban environments using high spatial resolution remote sensing imagery," *Remote Sens. Environ.*, vol. 242, p. 11, 1706, June 2020.
- [198] Q. Shen et al., "A CIE color purity algorithm to detect black and odorless water in urban rivers using high-resolution multispectral remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6577–6590, 2019. doi: 10.1109/TGRS.2019.2907283.
- [199] X. Huang, C. Xie, X. Fang, and L. Zhang, "Combining pixel- and object-based machine learning for identification of water-body types from urban high-resolution remote-sensing imagery," *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 8, no. 5, pp. 2097–2110, 2015. doi: 10.1109/JSTARS.2015.2420713.
- [200] M. Kamal, S. Phinn, and K. Johansen, "Object-based approach for multi-scale mangrove composition mapping using multi-resolution image datasets," *Remote Sens.*, vol. 7, no. 4, pp. 4753–4783, 2015. doi: 10.3390/rs70404753.
- [201] T. Hu, J. Liu, G. Zheng, Y. Li, and B. Xie, "Quantitative assessment of urban wetland dynamics using high spatial resolution satellite imagery between 2000 and 2013," *Sci. Rep.*, vol. 8, no. 1, p. 7409, 2018. doi: 10.1038/s41598-018-25823-9.
- [202] Q. Wu et al., "Integrating LiDAR data and multi-temporal aerial imagery to map wetland inundation dynamics using Google Earth Engine," *Remote Sens. Environ.*, vol. 228, pp. 1–13, July 2019. doi: 10.1016/j.rse.2019.04.015.
- [203] K. S. Schmidt and A. K. Skidmore, "Spectral discrimination of vegetation types in a coastal wetland," *Remote Sens. Environ.*, vol. 85, no. 1, pp. 92–108, 2003. doi: 10.1016/S0034-4257(02)00196-7.
- [204] G. Viennois, C. Proisy, J. Féret, J. Prosperi, F. Sidik, Suhardjono, et al. "Multitemporal analysis of high-spatial-resolution optical satellite imagery for mangrove species mapping in Bali, Indonesia," *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 9, pp. 3680–3686, 2016. doi: 10.1109/JSTARS.2016.2553170.
- [205] R. B. Norgaard, "Ecosystem services: From eye-opening metaphor to complexity blinder," *Ecol. Econ.*, vol. 69, no. 6, pp. 1219–1227, 2010. doi: 10.1016/j.ecolecon.2009.11.009.
- [206] Y. Z. Ayanu, C. Conrad, T. Nauss, M. Wegmann, and T. Koellner, "Quantifying and mapping ecosystem services supplies and demands: A review of remote sensing applications," *Environ. Sci. Technol.*, vol. 46, no. 16, pp. 8529–8541, 2012. doi: 10.1021/es300157u.
- [207] J. Haas and Y. Ban, "Mapping and monitoring urban ecosystem services using multitemporal high-resolution satellite data," *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 10, no. 2, pp. 669–680, 2016. doi: 10.1109/JSTARS.2016.2586582.
- [208] X. Ren, X. Chen, and Q. Ma, "Urban spatial ecological performance based on the data of remote sensing of Guyuan," *Int. Arch. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, vol. 42, p. 3, Apr. 2018.
- [209] C. R. Hakkenberg, M. P. Dannenberg, C. Song, and G. Vinci, "Automated continuous fields prediction from landsat time series: application to fractional impervious cover," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 1, pp. 132–136, 2019. doi: 10.1109/LGRS.2019.2915320.
- [210] L. Zhang, Q. Weng, and Z. Shao, "An evaluation of monthly impervious surface dynamics by fusing Landsat and MODIS time series in the Pearl River Delta, China, from 2000 to 2015," *Remote Sens. Environ.*, vol. 201, pp. 99–114, Nov. 2017. doi: 10.1016/j.rse.2017.08.036.
- [211] G. Xian, H. Shi, J. Dewitz, and Z. Wu, "Performances of WorldView 3, Sentinel 2, and Landsat 8 data in mapping impervious surface," *Remote Sens. Appl., Soc. Environ.*, vol. 15, p. 100, 246, 2019. doi: 10.1016/j.rsase.2019.100246.
- [212] W. Zhou, G. Huang, A. Troy, and M. L. Cadenasso, "Object-based land cover classification of shaded areas in high spatial resolution imagery of urban areas: A comparison study," *Remote Sens. Environ.*, vol. 113, no. 8, pp. 1769–1777, 2009. doi: 10.1016/j.rse.2009.04.007.
- [213] P. Li, J. Guo, B. Song, and X. Xiao, "A multilevel hierarchical image segmentation method for urban impervious surface mapping using very high resolution imagery," *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 4, no. 1, pp. 103–116, 2011. doi: 10.1109/JSTARS.2010.2074186.
- [214] T. Zhang and X. Huang, "Monitoring of urban impervious surfaces using time series of high-resolution remote sensing images in rapidly urbanized areas: A case study of Shenzhen," *IEEE J. of Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 11, no. 8, pp. 2692–2708, 2018. doi: 10.1109/JSTARS.2018.2804440.
- [215] C. E. Woodcock, T. R. Loveland, M. Herold, and M. E. Bauer, "Transitioning from change detection to monitoring with remote sensing: A paradigm shift," *Remote Sens. Environ.*, vol. 238, p. 111, 558, 2020. doi: 10.1016/j.rse.2019.111558.
- [216] D. Helman et al., "Using time series of high-resolution planet satellite images to monitor grapevine stem water potential in commercial vineyards," *Remote Sens.*, vol. 10, no. 10, p. 1615, 2018. doi: 10.3390/rs10101615.
- [217] M. A. Wulder, J. C. White, N. C. Coops, and C. R. Butson, "Multi-temporal analysis of high spatial resolution imagery for disturbance monitoring," *Remote Sens. Environ.*, vol. 112, no. 6, pp. 2729–2740, 2008. doi: 10.1016/j.rse.2008.01.010.
- [218] D. Turner, A. Lucieer, and S. M. De Jong, "Time series analysis of landslide dynamics using an Unmanned Aerial Vehicle (UAV)," *Remote Sens.*, vol. 7, no. 2, pp. 1736–1757, 2015. doi: 10.3390/rs70201736.
- [219] H. Li, L. Chen, F. Li, and M. Huang, "Ship detection and tracking method for satellite video based on multiscale saliency and surrounding contrast analysis," *J. Appl. Remote Sens.*, vol. 13, no. 2, p. 026511, 2019. doi: 10.1117/1.JRS.13.026511.
- [220] L. Wang, F. Chen, and H. Yin, "Detecting and tracking vehicles in traffic by unmanned aerial vehicles," *Automat. Construct.*, vol. 72, pp. 294–308, Dec. 2016. doi: 10.1016/j.autcon.2016.05.008.
- [221] L. Mou et al., "Multitemporal very high resolution from space: Outcome of the 2016 IEEE GRSS data fusion contest," *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 10, no. 8, pp. 3435–3447, 2017. doi: 10.1109/JSTARS.2017.2696823.
- [222] L. Mou and X. X. Zhu, "Spatiotemporal scene interpretation of space videos via deep neural network and tracklet analysis,"

- in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, 2016, pp. 1823–1826.
- [223] M. C. Hansen and R. S. DeFries, “Detecting long-term global forest change using continuous fields of tree-cover maps from 8-km advanced very high resolution radiometer (AVHRR) data for the years 1982–99,” *Ecosystems*, vol. 7, no. 7, pp. 695–716, 2004. doi: 10.1007/s10021-004-0243-3.
- [224] A. Schneider, M. A. Friedl, and D. Potere, “A new map of global urban extent from MODIS satellite data,” *Environ. Res. Lett.*, vol. 4, no. 4, p. 044003, 2009. doi: 10.1088/1748-9326/4/4/044003.
- [225] M. A. Friedl et al., “MODIS collection 5 global land cover: Algorithm refinements and characterization of new datasets,” *Remote Sens. Environ.*, vol. 114, no. 1, pp. 168–182, 2010. doi: 10.1016/j.rse.2009.08.016.
- [226] ESA. CCI-LC Product User Guide v2.4 [Online]. Available: [Http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-PUG-v2.4.pdf](http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-PUG-v2.4.pdf)
- [227] M. C. Hansen et al., “High-resolution global maps of 21st-century forest cover change,” *Science*, vol. 342, no. 6160, pp. 850–853, 2013. doi: 10.1126/science.1244693.
- [228] J. Chen et al., “Global land cover mapping at 30m resolution: A POK-based operational approach,” *ISPRS J. Photogrammetry Remote Sens.*, vol. 103, pp. 7–27, May 2015. doi: 10.1016/j.isprsjprs.2014.09.002.
- [229] P. Gong et al., “Annual maps of global artificial impervious area (GAIA) between 1985 and 2018,” *Remote Sens. Environ.*, vol. 236, p. 111,510, Jan. 2020. doi: 10.1016/j.rse.2019.111510.
- [230] P. Gong et al., “Mapping essential urban land use categories in China (EULUC-China): Preliminary results for 2018,” *Sci. Bull.*, vol. 65, no. 3, pp. 182–187, 2020. doi: 10.1016/j.scib.2019.12.007.
- [231] M. Pesaresi, D. Ehrlich, A. J. Florczyk, S. Freire, A. Julea, T. Kemper, et al. “GHS built-up grid, derived from Landsat, multitemporal (1975, 1990, 2000, 2014),” *European Commission, Joint Res. Centre, JRC Data Catalogue*, 2015.
- [232] P. Gong, H. Liu, M. Zhang, C. Li, J. Wang, H. Huang, et al. “Stable classification with limited sample: Transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017,” *Sci. Bull.*, vol. 64, no. 6, pp. 370–373, 2019. doi: 10.1016/j.scib.2019.03.002.
- [233] R. Houborg and M. McCabe, “High-resolution NDVI from Planet’s constellation of earth observing nano-satellites: A new data source for precision agriculture,” *Remote Sens.*, vol. 8, no. 9, p. 768, 2016. doi: 10.3390/rs8090768.
- [234] L. Wang, M. Jia, D. Yin, and J. Tian, “A review of remote sensing for mangrove forests: 1956–2018,” *Remote Sens. Environ.*, vol. 231, p. 111,223, 2019. doi: 10.1016/j.rse.2019.111223.
- [235] A. Ertürk, M. Iordache, and A. Plaza, “Sparse unmixing with dictionary pruning for hyperspectral change detection,” *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 10, no. 1, pp. 321–330, 2017. doi: 10.1109/JSTARS.2016.2606514.
- [236] M. V. M. Graña, B. Ayerdi. “Hyperspectral remote sensing scenes.” Grupo de Inteligencia Computacional (GIC). http://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes&redirect=no (accessed 2012).
- [237] K. Nogueira, O. A. B. Penatti, and J. A. dos Santos, “Towards better exploiting convolutional neural networks for remote sensing scene classification,” *Pattern Recognition*, vol. 61, pp. 539–556, Jan. 2017. doi: 10.1016/j.patcog.2016.07.001.
- [238] W. Zhou, D. Ming, X. Lv, K. Zhou, H. Bao, and Z. Hong, “SO-CNN based urban functional zone fine division with VHR remote sensing image,” *Remote Sens. Environ.*, vol. 236, p. 111,458, 2020. doi: 10.1016/j.rse.2019.111458.
- [239] M. Li, K. M. d Beurs, A. Stein, and W. Bijker, “Incorporating open source data for Bayesian classification of urban land use from VHR stereo images,” *IEEE J. Select. Topics Appl. Earth Observat. Remote Sens.*, vol. 10, no. 11, pp. 4930–4943, 2017. doi: 10.1109/JSTARS.2017.2737702.
- [240] S. Du, S. Du, B. Liu, and X. Zhang, “Context-enabled extraction of large-scale urban functional zones from very-high-resolution images: A multiscale segmentation approach,” *Remote Sens.*, vol. 11, no. 16, p. 1902, 2019. doi: 10.3390/rs11161902.
- [241] X. Zhang, S. Du, and Q. Wang, “Hierarchical semantic cognition for urban functional zones with VHR satellite images and POI data,” *ISPRS J. Photogrammetry Remote Sens.*, vol. 132, pp. 170–184, Oct. 2017. doi: 10.1016/j.isprsjprs.2017.09.007.
- [242] X. Liu et al., “Classifying urban land use by integrating remote sensing and social media data,” *Int. J. Geographical Inform. Sci.*, vol. 31, no. 8, pp. 1675–1696, 2017. doi: 10.1080/13658816.2017.1324976.

GRS