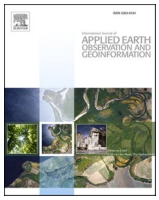
International Journal of Applied Earth Observations and Geoinformation 101 (2021) 102367

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

journal homepage: www.elsevier.com/locate/jag



A recurrent curve matching classification method integrating within-object spectral variability and between-object spatial association

Yunwei Tang [a], Fang Qiu [b,*], Linhai Jing [a], Fan Shi [c], Xiao Li [b]

[a] Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, No. 9 South Road, Beijing 100094, China
[b] Geospatial Information Sciences, The University of Texas at Dallas, 800 West Campbell Road, Richardson, TX 75080, USA
[c] College of Information Science and Engineering, Henan University of Technology, No. 100 Lianhua Street, Zhengzhou High-Tech Development Zone, Henan 450001, China

ARTICLE INFO

Keywords:
Object-based image classification
Recurrent curve matching method
Spatial association
High spatial resolution

ABSTRACT

Object-based image analysis (OBIA), which has been commonly used for land cover and land use classification, may encounter challenges when satellite images' spatial resolution achieves at the sub-meter level. An image object may exhibit spectral heterogeneity, causing traditional object-level statistical measures such as mean values of the pixels in an object not suited to represent the feature of the object. Additionally, an image object may have strong spatial association with its surroundings. Traditional OBIA only considers spatial features of individual object, but ignoring spatial arrangement or spatial association between objects. This paper proposes a new OBIA method by integrating within-object spectral variability and between-object spatial association. The within-object spectral variability is captured by the histograms of the pixels in an object across multispectral bands to reflect the heterogeneity of their pixel values. Based on this, the initial classification result is obtained using non-parametric curve matching methods. Then, the between-object spatial association is represented by curves derived from the frequency of pairwise classes in four main directions, also in the form of curves. The curves now composed of both the histograms of spectral feature and the class pair frequency of spatial feature are then fused for another curve matching based classification. This recurrent process is repeated and the spatial association is recaptured from the previous classification result at each iteration until a stopping criterion is satisfied. The curve matching classification method based on histograms of spectral feature is superior to traditional OBIA based on only object-level statistical measures since it fully characterizes spectral variability in the objects. The between-object spatial association works as a spatial filter that considers spatial arrangement of classes in a neighborhood. The developed method is especially suitable for classifying high spatial resolution (HSR) images with land cover/land use classes in typical urban areas.

1. Introduction

High spatial resolution (HSR) data provide large amount of detailed geospatial information. Developing methods of land cover/land use classification for HSR images has been one of the most active research areas in remote sensing. The increasing intra-class heterogeneity and decreasing inter-class variability in the spectral feature space often lead to misclassification of spectrally similar classes (Bruzzone and Carlin, 2006; Lv et al., 2016), causing the classification of HSR image inaccurate. The well-known object-based image analysis (OBIA) has advantages in processing HSR images. OBIA first generates homogeneous image segments that correspond to real-world objects of interest by applying image segmentation, which groups spatially connected pixels with similar characteristics into segments (i.e., objects) (Blaschke et al., 2014). Object-level statistical measures, such as the mean value and standard deviation of the pixels in an object are commonly used as the object's features. Then classification methods are applied to classify image objects based on these object-level statistical measures.

Traditional OBIA has encountered challenges when satellite images with a sub-meter spatial resolution became widely available. An object with multiple parts, which is common in reality, may exhibit increased spectral heterogeneity at finer spatial resolution (Tang et al., 2020). A typical example is a building can include several parts with very different materials such as concrete, glasses, tiles, etc. Traditional OBIA

* Corresponding author.
E-mail address: ffqiu@utdallas.edu (F. Qiu).

https://doi.org/10.1016/j.jag.2021.102367
Received 18 February 2021; Received in revised form 13 April 2021; Accepted 13 May 2021
Available online 23 May 2021

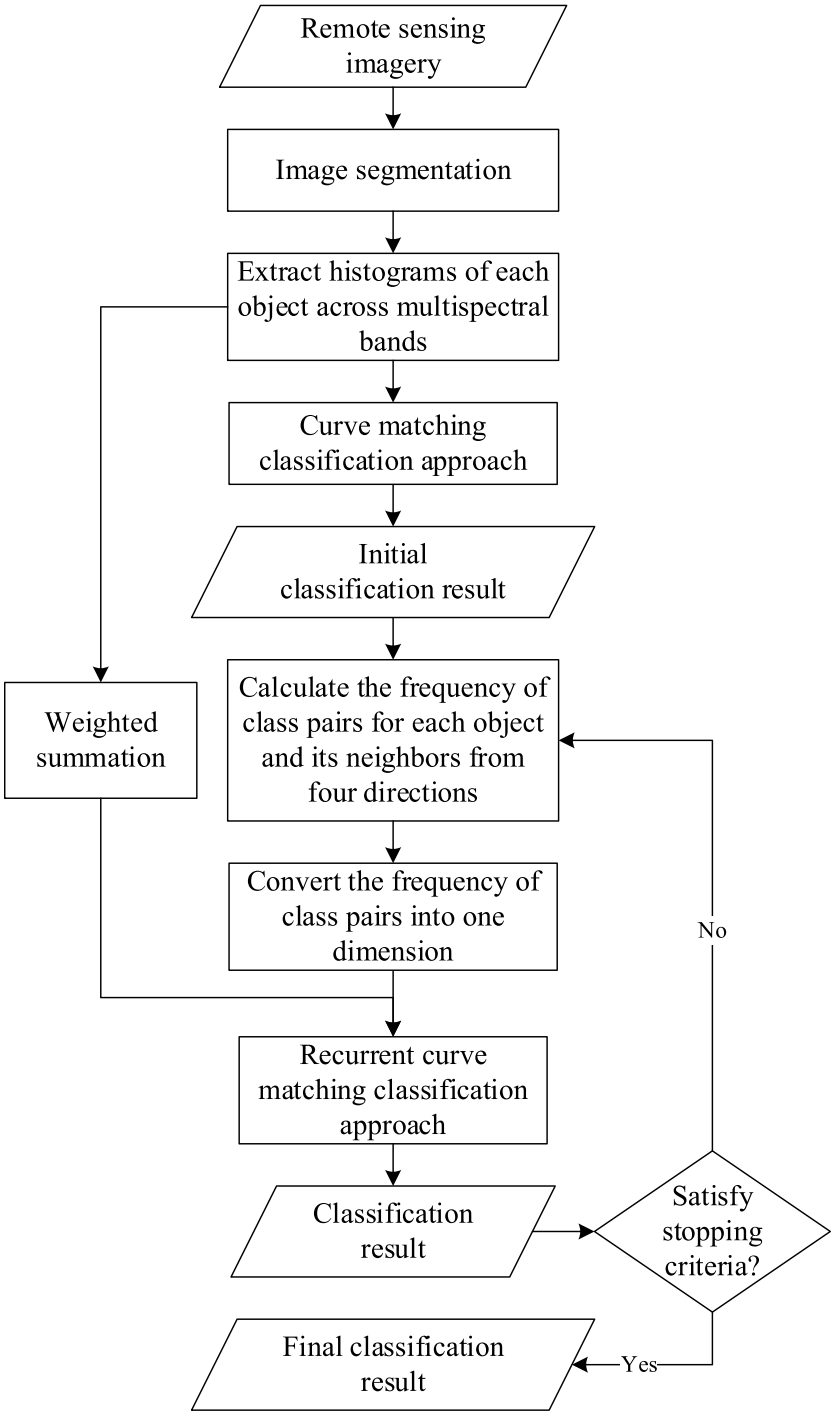


Fig. 1. The process of the RCCM classification method.

based on statistical measures may not be suited when an object appears large spectral heterogeneity. It is because its pixel values do not follow a normal distribution for an object with multiple parts. Especially, when the distribution of pixel values in the object has multiple peaks, statistical measures cannot characterize this object (Wang et al., 2019). Additionally, the traditional classification methods based only on the object-level statistical measures could not make the best use of the detailed information lies in the pixels in HSR images.

The curve matching method compares two curves based on some functions and returns a measure to indicate the divergence between these two curves. Recently, this method has been extended for object-based classification (Stow et al., 2012; Toure et al., 2013; Sridharan and Qiu, 2013; Tang et al., 2020). In these studies, the spectral histogram or the derivative of the histogram of the pixel values in an object are extracted as the feature curve, which can fully explore the spectral variability of HSR images. The functions such as the curve angle mapper (CAM) and the Kolmogorov-Smirnov (KS) distance are then employed to calculate the divergences between curves to perform the non-parametric classification. The previous research reported the curve matching methods can achieve higher accuracy than the traditional OBIA based on the statistical measures. Additionally, it was found that the curve matching methods facilitate to incorporate disparate data sources such as LiDAR to further improve classification performance because they are all fundamentally curves (Zhou and Qiu, 2015; Wan et al., 2021).

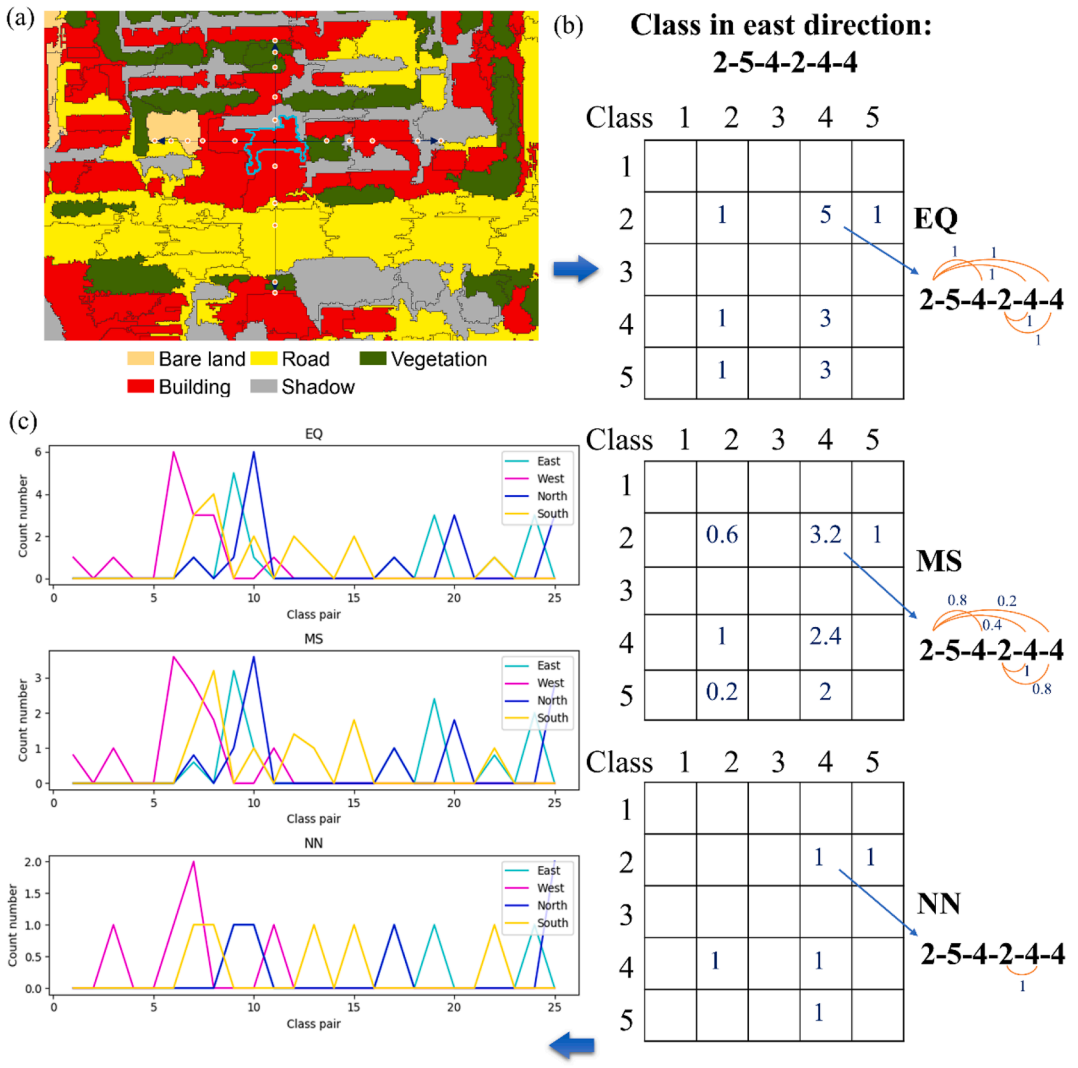


Fig. 2. An example of capturing spatial association from neighborhood for an object (class code: 1-bare land, 2-building, 3-road, 4-shadow, 5-vegetation). (a) An object and its neighborhood in four directions, (b) the class codes in the east direction and the weighted frequency of class pairs using three weighting schemes, (c) the curves of weighted frequency versus class pairs using three weighting schemes in four directions.

However, not many studies have further explored the curve matching methods in OBIA, especially the ability of curve matching methods for incorporating spatial features.

Spatial features, specifically, the spatial arrangement or spatial association can be introduced into classification to utilize the connection between objects in order to reduce the risk of misclassification (Zhong et al., 2014; Bai et al., 2020). Except for the case of an object including multiple parts, another common case in a HSR image is that an object may encompass other smaller objects of a different class. For example, vehicles on the roads and parking lots are usually not land cover/land use class of interest and therefore are expected to be segmented as part of the roads or parking lots. However, due to their spectral uniqueness, vehicles are often separated from the roads and parking lots they are on after segmentation and thus misclassified as other classes. The spatial association is helpful to capture the spatial arrangement of the road or the parking lot object and its surroundings, and may prevent this type of misclassification. The object-level spatial association can also help to identify land use types or function zones. For example, a railway station can be identified from the characteristic of the objects that comprise it (e.g., long thin platforms) and the set of objects that surround it (e.g., railway track, parking lots and multiple roads) (Tang et al., 2016a). Currently, however, most studies only focus on the pixel-level spatial association such as the classification based on the Markov random field model (Solberg et al., 1996; Tarabalka et al., 2010; Sun et al., 2015). Some researchers proposed measures to establish object-level association for some specific classes. For example, Hay et al. (1996) developed an image texture primitive neighborhood method to contextually assess the characteristics of forest objects of individual tree crowns. Huang and Zhang (2011, 2012) proposed a morphological index based on the building/shadow association to extract buildings. Recently, the geostatistical and multiple-point geostatistical weightings have been proposed for the *K* nearest neighbor (KNN) classification at the object level (Tang et al., 2016a, 2016b, 2018), in which the between-object association was modelled from the spatial covariance and multiple-point probability for each class. However, these methods rely on the performance of the KNN classifier, thus their improvements are limited. Also, these methods only establish the spatial association for the same class, without exploring the spatial association between different classes. The deep learning technique has been widely used for scene classification (Yu et al., 2017; Chaib et al., 2017) and object detection (Ding et al., 2018) due to its ability of extracting high-level features. Recently, the convolutional neural networks (CNN) have been combined with the OBIA for land cover and land use classification (Zhang et al., 2018; Zhang, Harrison, et al., 2020; Zhang, Yue, et al., 2020). Particularly, the pixel-wise fully convolutional networks (FCN) show better performance than the patch-wise CNN for OBIA (Fu et al, 2017; Liu et al., 2018).

However, the CNN method commonly requires a large number of samples to achieve a desirable accuracy.

To address the aforementioned problems, a recurrent curve matching method is developed to combine spectral and spatial features. The spectral features fully explore the variability of pixel spectral values within objects and the spatial features establish the class association between neighboring objects. The developed method is compared with curve matching methods that consider within-spectral features only and several advanced classification methods in three different study areas.

2. Methods

To test the developed method at the object level, image segmentation is first applied. Then the histogram is extracted for each object across multispectral bands to provide spectral features. The initial classification result is obtained based on the histogram-based spectral features using a curve matching method. The spatial association is captured from the initial classification result for each object in four main directions. Spatial association is defined as the frequency of pairwise classes of the object under consideration and its neighboring objects, which when recorded in the form of curves, can feed back into the curve matching method as spatial features along with spectral features to improve the classification result. A new round of spatial association can be then captured from the second classification result. This recurrent process is repeated until a stopping criterion is satisfied. The spatial association is recurrently captured from the previous classification result at each iteration, thus we designate this method as the RCCM (ReCurrent Curve Matching) method. In comparison, the curve matching method based on the histogram-based spectral features only is designated as the CM method. The process of the RCCM method is shown in Fig. 1.

2.1. Feature extraction

The fractal net evolution algorithm (FNEA) method is firstly applied to produce the segmentation result (Benz et al., 2004). The histogram of the pixels in the objects is used as the within-object spectral feature. The between-object spatial association is recurrently captured from the previous classification result for each object. To do so, the center of an object is first extracted. Starting from this point, the classes of neighboring objects in a pre-defined range r along four main directions (east, west, south and north) are detected for this object. The class pair is recorded when an object and its neighboring object are both within the image boundary. The frequency of a class pair is extracted for the combination of any two classes. If an image is classified to M classes, the number of total class pairs is M by M. The frequency of pairwise classes can generally reflect the spatial arrangement of neighboring objects. Since spatial distance may have different effects on spatial association, different weighting schemes can be applied according to the distance between the central object and its neighbors in the range r. Three weighting schemes are considered in the study: 1) Equal weight (EQ): assigning equal weighting to all the neighbors; 2) Multiscale weight (MS): giving larger weights to nearer neighbors and smaller weights to further neighbors; 3) Nearest neighboring weight (NN): considering only the adjacent neighbors.

An example of extracting spatial association is shown in Fig. 2. The central object is classified as the building class. Starting from the center of this object, the spatial association is detected in a range r, which equals 5 objects. The sequential classes along the east direction are (including the central object): building, vegetation, shadow, building, shadow and shadow. The corresponding class codes are 2, 5, 4, 2, 4 and 4. In the EQ weight scheme, the frequency of each class pair is counted, regardless of the distance between objects. For example, the weighted frequency of class pair 2–4 is 5. In the MS weight scheme, we use the inverse of the range value (i.e., $1/r$) as the smallest weight assigned to the furthest neighbor, and the value of 1 as the largest weight assigned to the nearest neighbor, the weights in between are evenly distributed. Therefore, the weights are 1.0, 0.8, 0.6, 0.4 and 0.2 when r equals 1, 2, 3, 4, and 5, respectively. The five distances between objects for class pair 2–4 are 2, 4, 5, 1 and 2. Applying MS weight, the weighted frequency of class pair 2–4 equals $0.8 + 0.4 + 0.2 + 1.0 + 0.8 = 3.2$. In the NN scheme, only the adjacent class pair is considered, thus the weighted frequency of class pair 2–4 is 1. The two-dimensional class pairs $C1$-$C2$ can be converted to a one-dimensional variable by simply applying $(C1-1) \times M + C2$, where $C1$-$C2$ represents the class pair and M is the total number of classes. For example, the class pair 2–4 corresponds to $(2-1) \times 5 + 4 = 9$ in the one-dimensional vector. The spatial association of class pair is plotted as curves with the one-dimensional class pair as x-axis, and the weighted frequency of the corresponding class pair as y-axis. The weighted frequency of spatial curves is standardized to values between 0 and 1 to keep the same value range with the spectral histogram.

2.2. The RCCM method

The classification is performed at the object level, in which the within-object spectral histogram and between-object spatial association are used as feature curves. Curve matching methods are adopted to integrate spectral and spatial features for classification. As mentioned previously, the curve matching method compares two curves by calculating their divergence. In this study, three curve matching methods are conducted: Kullback-Leibler (KL) divergence (Kullback, 1987), curve angle mapper (CAM) (Kruse et al., 1993), and root sum squared differential area (RSSDA) (Stow, 2012).

For a pair of objects, the extracted feature curves are expressed as P_1, P_2. The divergence D calculated by the KL, CAM and RSSDA methods is as follows.

$$D_{\mathrm{KL}} = \frac{1}{2}\left[\sum_{i=1}^{n} P_1(i) log\frac{P_1(i)}{P_2(i)} + \sum_{i=1}^{n} P_2(i) log\frac{P_2(i)}{P_1(i)}\right] \quad (1)$$

$$D_{\mathrm{CAM}} = arccos\left[\frac{\sum_{i=1}^{n} P_1(i)P_2(i)}{\sqrt{\sum_{i=1}^{n} P_1(i)^2 \sum_{i=1}^{n} P_2(i)^2}}\right] \quad (2)$$

$$D_{\mathrm{RSSDA}} = \sqrt{\sum_{i=1}^{n} [P_1(i) - P_2(i)]^2} \quad (3)$$

where n is the total number of curve intervals. If P represents the histogram (i.e., the spectral feature), n is the number of histogram bins. If P represents the curve of weighted frequency of class pairs (i.e., the spatial feature), n is the number of class pairs.

The histograms across multispectral bands are first obtained for each object as the within-object spectral features. The spectral features are used alone to derive the initial classification result. Since curve matching methods are non-parametric, the histogram curves of each object to be classified are compared with the histogram curves of each training object. The divergences are calculated from Eqs. (1)–(3) for each pair of the training object and the object to be classified. The minimum divergence indicates the feature curve of the object to be classified is closest to that of the training object, thus its associated class is assigned to this object. Then the recurrent process of extracting between-object spatial feature starts. In addition to the spectral divergences estimated from K multispectral bands, each pair of objects also has divergences derived from spatial association in four directions. The divergence of the RCCM method is the weighted sums of spectral and spatial divergences, as given in Eq. (4):

$$D_{\mathrm{CM\ -\ spec\ +\ spat}} = w \sum_{k=1}^{K} D(k)_{\mathrm{spectral}} + (1-w) \sum_{m=1}^{4} D(m)_{\mathrm{spatial}} \quad (4)$$

where D_{spectral} is the spectral divergence, summed across K multispectral

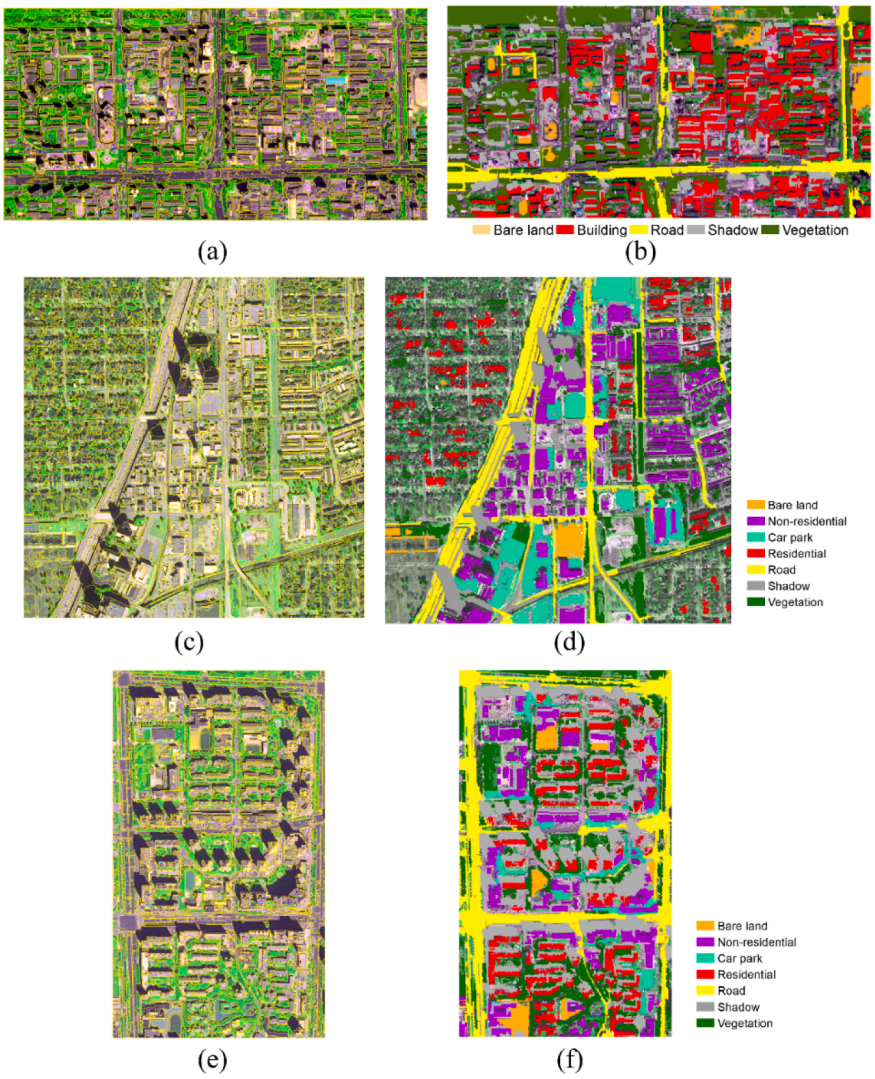


Fig. 3. The segmentation results and the sample objects of three images used for the experiments: (a) segmentation result of IKONOS image, (b) samples of IKONOS image, (c) segmentation result of WV-2 image, (d) samples of WV-2 image, (e) segmentation result of WV-3 image, (f) samples of WV-3 image.

bands, whereas $D_{spatial}$ is the spatial divergence, derived from four directions. These two divergences are calculated by one of the methods from (1)–(3). w is the weight that balances the proportion of spectral and spatial features, ranging from 0 to 1.

The curve matching method is then performed based on the combined divergences. This process is repeated until the stopping criterion is satisfied, i.e., the change of classification accuracy is less than a predefined threshold compared with that of the previous result, or the pre-defined maximum number of recurrent iterations is achieved.

3. Experiments

3.1. Experimental datasets and study sites

Identifying land cover and land use classes from HSR imagery in urban areas is a difficult objective (Johnson and Xie, 2013). The developed RCCM method was performed on three different images in typical urban environments. The IKONOS, WorldView-2 (WV-2) and WorldView-3 (WV-3) images were used in the study. The IKONOS image acquired in May 2000 is located in Beijing, China (39°57′57″– 39°58′30″ N, 116°23′19″– 116°24′46″ E). With the spatial resolution of 4 m, five land cover classes can be identified in the IKONOS image, including bare land, building, road, shadow and vegetation. The bare land, the building and the road classes are difficult to be classified from their spectral signatures only. The WV-2 image was obtained in January 2010 over downtown Dallas, Texas (32°50′18″– 32°51′21″ N, 96°45′49″– 96°47′8″ W). With a spatial resolution of 2 m, seven land use classes can be distinguished, including bare land, car park, residential, non-residential, road, shadow and vegetation. The WV-3 image acquired in September 2014 is also located in Beijing (39°59′52″–40°0′31″ N, 116°23′37″–116°24′7″ E). The same seven land use classes are identified in the WV-3 image at a spatial resolution of 1.6 m. The non-residential class includes public, business, commercial and industry. It is not easy to classify residential and non-residential since they are essentially buildings made of various materials. The road and the car park classes also have very similar spectral signatures and are difficult to be separated.

For the FNEA segmentation, the scale parameter was selected using the automated estimation of scale parameter (ESP) tool. The samples were selected at the object level from the segmentation results. The segmentation results and sample objects for the three study areas are shown in Fig. 3. Most classes can be manual interpreted from enhanced images using the Gram-Schmidt pansharpening method. Some of the residential and the non-residential classes in the WV-3 image were identified by site visitation and referencing the Gaode data (https://lbs.amap.com/). The developed method and all the benchmark methods

Table 1
The number of sample objects and the total objects in the study areas.

IKONOS		WV-2		WV-3	
Class	Number of objects	Class	Number of objects	Class	Number of objects
Bare land	22	Bare land	26	Bare land	21
Building	233	Non-residential	399	Non-residential	198
Road	25	Car park	147	Car park	58
Shadow	142	Residential	104	Residential	220
Vegetation	150	Road	131	Road	175
		Shadow	74	Shadow	94
		Vegetation	104	Vegetation	167
Total samples	**572**	**Total samples**	**985**	**Total samples**	**933**
Total objects	**1188**	**Total objects**	**3033**	**Total objects**	**1911**

were repeated ten times on different samples for a robust test (Stehman, 2009). Each time the ratio of training and testing samples is 1:2 using a random sampling scheme, selected from the sample objects in Fig. 3. The information about the sample objects is provided in Table 1.

3.2. Classification

For all the study areas, the three curve matching methods (i.e., KL, CAM and RSSDA) were applied to derive the initial classification results. These results were also used for the comparison purpose. Then the between-object spatial association was extracted, with three weighting schemes (i.e., EQ, MS, NN) applied to improve the initial classification results. For the parameter n, the number of bins for the spectral histogram was set to 100, and the numbers of class pair equal 25, 49, and 49 for the IKONOS, WV-2 and WV-3 images, respectively. To select the number of iterations and the weight w in each iteration, a 5-fold cross-validation was performed by randomly selecting 20% validation samples from training samples. The weight w was set to change from 0 to 1 with an increment of 0.1. The maximum number of iterations was set to 10, and the threshold of change in validation accuracy between two consecutive iterations was set to 0.1%, below which the iteration will be stopped. The number of iterations that achieves the highest validation accuracy was chosen, and their associated weight in each iteration was recorded. Then these parameters were applied to classify the whole image, and the testing samples were used to estimate the classification accuracy.

To test the developed RCCM method, eight advanced benchmark classification methods were employed at the object level, including i) random forest (RF) classification, ii) support vector machine (SVM), iii) extreme gradient boosting (XGBoost) (Chen and Guestrin, 2016), iv) geostatistically weighted KNN (gKNN) (Atkinson, 2004; Atkinson and Naser, 2010), v) multiple-point weighted KNN (MPKNN) (Tang et al., 2016a), vi) fully convolutional networks (FCN), vii) relearning method based on histogram (relearn-Hist) (Huang et al., 2014), and viii) relearning method based on the primitive co-occurrence matrix (relearn-PCM) (Huang et al., 2014). The RF, SVM and XGBoost are common machine learning methods. Seven spectral features, including mean, median, standard deviation, skewness, kurtosis, first and third quartile, were used in the RF, SVM and XGBoost methods. The gKNN method incorporates the spatial covariance into the KNN classifier, whereas the MPKNN method enhances gKNN by further considering spatial structure in an irregular data template of an object and its nearest neighbors in the feature space. Both the gKNN and MPKNN methods have been used at the object level for classification. In the FCN, the objects were resized to image patches of the same size and used as the input to the networks. The optimal input image patch size was set to 32×32 through trial-and-error approach. The FCN requires all the pixels in an image patch to be labeled, for which the pixels outside the object were labeled as background. Data augmentation, including flipping, rotating, noise adding and images blurring, was applied to derive more robust training data. The number of FCN layers was set to four to balance the network complexity and robustness (Chen S. et al., 2016; Chen Y. et al., 2016). The structure includes convolutional layers, pooling layers and deconvolutional layers. Thirty-two filters with the size of 3×3 were used in the convolutional layers. To train the network, the images were randomly shuffled and fed into the network in batches, with each batch containing 40 images. The learning rate was set to 0.001 and the number of epochs was set to 500. Each pixel was classified with a label for an output image patch. The majority label of the pixels within the object was assigned to this object for each output image patch. The original relearn-Hist and relearn-PCM methods were proposed at the pixel level, extracting spatial association from a moving template for each pixel. The relearn-Hist uses a spatial feature of class frequency histogram, and the relearn-PCM adds one more spatial feature of spatial arrangement for classification. We extended these two methods to the object level to make a fair comparison. These two methods are based on the SVM classifier, where the spectral features are the same seven statistical measures used with the SVM, RF and XGBoost methods. For the spatial features, the same neighborhood objects used in our method were also adopted to derive class frequency histogram and spatial arrangement. Similar to our method, a 5-fold cross-validation was applied to select optimal parameters from a set of the candidate parameters for the benchmark classification methods. The candidate parameters used for cross-validation in the classification methods are listed in Table 2.

Table 2
The candidate parameters used for cross-validation in the classification methods.

Method	Parameter	Candidate values
RF	Number of trees	10, 15, 20
	The function to measure the quality of a split	gini, entropy
	The minimum number of samples at a leaf node	2, 4, 6
SVM	Kernel type	linear, radial basis function
	Regularization parameter	1, 10, 100, 1000
	Kernel coefficient	10^{-4}, 10^{-3}
XGBoost	The maximum depth of the tree	4, 5, 6, 7
	Learning rate	From 0.03 to 0.3 with an increment of 0.03
	Number of trees	100, 200
gKNN	Number of neighbors	1, 3, 5, 7
	Proportional weight	From 0 to 1 with an increment of 0.1
MPKNN	Number of neighbors	1, 3, 5, 7
	Weight for spectral covariance	From 0 to 1 with an increment of 0.1
	Weight for multiple-point probability	From 0 to 1 with an increment of 0.1
Relearn-Hist and relearn-PCM	Number of iteration times	An integer not exceeding 10
RCCM	Number of iteration times	An integer not exceeding 10
	Weight for spectral feature in each iteration	From 0 to 1 with an increment of 0.1

3.3. Results and analysis

3.3.1. Classification results

Since there are three curve matching methods, and each has three weighting schemes, we use the abbreviations to distinguish the nine RCCM methods. For example, the RCCM methods based on the KL classifier with the EQ weighting scheme is written as KL-EQ. In total, three CM methods (KL, CAM, RSSDA), nine RCCM methods (KL-EQ, KL-

Table 3
The average classification accuracies (%) and numbers of significant improvements at 95% confidence level out of ten samplings (in parentheses).

Methods		IKONOS	WV-2	WV-3
Advanced classification methods	RF	92.62 (2)	68.59 (10)	61.93 (10)
	SVM	92.00 (2)	67.79 (10)	60.11 (10)
	XGBoost	92.60 (2)	69.16 (10)	61.64 (10)
	gKNN	92.49 (1)	64.67 (10)	61.64 (10)
	MPKNN	92.60 (0)	64.84 (10)	61.66 (10)
	FCN	92.65 (0)	72.45 (10)	61.56 (10)
	Relearn-Hist	92.39 (1)	72.19 (10)	69.33 (9)
	Relearn-PCM	**92.84** (0)	**74.81** (9)	**72.11** (10)
CM and RCCM methods	KL	89.37 (10, 9, 9)	64.12 (10, 10, 10)	58.46 (10, 10, 10)
	CAM	89.92 (6, 6, 6)	66.88 (10, 10, 10)	60.96 (10, 10, 10)
	RSSDA	89.97 (4, 4, 5)	66.96 (10, 10, 10)	59.89 (10, 10, 10)
	KL-EQ	93.07	80.97	73.20
	CAM-EQ	92.76	82.50	75.15
	RSSDA-EQ	92.10	80.08	74.44
	KL-MS	93.62	81.89	74.84
	CAM-MS	93.02	84.29	77.53
	RSSDA-MS	91.92	81.93	77.27
	KL-NN	**93.86**	82.09	76.96
	CAM-NN	92.78	**84.67**	**79.68**
	RSSDA-NN	92.26	82.91	78.44

In the significance test, all the advanced methods were compared only with the RCCM method with the highest accuracy. The CM methods were compared with their corresponding RCCM methods using three weighting schemes in a sequence of EQ, MS, and NN.

MS, KL-NN, CAM-EQ, CAM-MS, CAM-NN, RSSDA-EQ, RSSDA-MS, RSSDA-NN), and eight advanced methods (SVM, RF, XGBoost, gKNN, MPKNN, FCN, relearn-Hist, relearn-PCM) were performed on three study areas, each method was performed ten times with different samplings.

The average classification accuracies of the ten samplings for all the methods are reported in Table 3. The highest accuracy for each image is highlighted in bold. The McNemar test was performed to test if the increases in classification accuracy of the RCCM method over other methods are significant at the 95% confidence interval for each sampling. The number of improvements that are significant out of the ten samplings is indicated in parentheses. For the significance test, all the advanced classification methods were compared only with the RCCM method with the highest accuracy. The CM methods were each compared with the corresponding RCCM method using three weighting schemes (EQ, MS, and NN) with their respective number of significant improvement also included in parentheses. The classification results of the testing samples from one of the ten samplings and the corresponding testing samples are shown in Figs. 4–6. For simplicity only the results using the SVM, XGBoost, MPKNN, FCN, relearn-PCM, one of RCCM methods with the highest accuracy (KL-NN for the IKONOS image, CAM-NN for the WV-2 and WV-3 images) and the corresponding CM methods (KL for the IKONOS image, CAM for the WV-2 and WV-3 images) are displayed. The overall accuracies from one sampling (corresponding to Figs. 4–6) are listed in Tables 4–6, along with the F1-score of each class.

As can be seen in Table 3, the RF and XGBoost results have similar accuracies, and the accuracy of SVM is lower. The gKNN and MPKNN methods are based on improvement over the KNN classifier, with the latter performs slightly better than the former. The FCN method does not show a great advantage over other methods, although it performs well for the IKONOS and WV-2 images. One of reasons could be that FCN may not be able to establish a stable architecture with the available sample data, even though data augmentation is applied. It could also be that FCN may be more suited to classify image patches with a larger and regular extent rather than image objects with irregular shapes, since the resizing process of FCN is quite arbitrary and may cause some useful information to be lost. The two relearning methods based on the histogram and PCM show better accuracies than the other advanced classification algorithms, except for the relearn-Hist in the IKONOS image. The accuracy of the relearn-PCM method is constantly higher than that of the relearn-Hist method since the latter considers both class frequency and spatial arrangement. These two SVM-based relearning methods improves the accuracies of the SVM method up to 12% by the relearning process.

Among three CM methods, the RSSDA method has the highest accuracy for the IKONOS and WV-2 images, and the CAM method has the highest accuracy for the WV-3 image. All nine RCCM methods greatly improve the accuracies of their corresponding CM methods, with an increase in accuracy ranging from 1.95% to 18.72%. The RCCM methods demonstrate great advantages in classifying the HSR WV-2 and WV-3 images with more detailed classes. The nearest weighting scheme (NN) always leads to the highest accuracy, such as the KL-NN method for the IKONOS image and the CAM-NN for the WV-2 and WV-3 images. Comparing with the relearn-Hist and relearn-PCM methods, the accuracies of the nine RCCM methods are similar for the IKONOS image but all higher for the WV-2 and WV-3 images. Overall, the developed RCCM methods are superior to the advanced classification and the CM methods considering spectral features only.

For the IKONOS image with simpler classes, the accuracies of the advanced classification methods are already rather high. The KL-NN method has outperformed all other approaches, and it demonstrates significant improvements over the RF, SVM and XGBoost methods in two samplings, and over the MPKNN and relearn-PCM methods in one sampling of the total ten samplings. For the WV-2 image with more complex classes, the accuracies of the FCN and two relearning methods are over 72%, and the accuracies of all other methods are below 70%. In comparison, the accuracy of the CAM-NN method is 84.67%, resulting in significant improvement over the relearn-PCM in nine samplings and over all other methods in all the ten samplings. Similarly, for the WV-3 image, the accuracy of the CAM-NN method is significantly better than all other advanced classification methods in all the ten samplings, except for only the relearn-Hist method, over which the CAM-NN method demonstrates significant improvement in nine samplings. When compared with the CM methods, all the nine RCCM methods demonstrate significant improvement over CM counterparts in four to ten samplings for the IKONOS image and all the ten samplings for the WV-2 and WV-3 images.

The F1-score for each class shows that the bare land class in the IKONOS image has the greatest improvement (Table 4). Some advanced benchmark methods even misclassify all the objects of this class due to a small number of samples and lack of unique features. The F1-score of the KL method is 0.53 for the bare land class, and the KL-NN method improves the F1-score to 0.8. The bare land class shows similar trend in the WV-2 and WV-3 images (Tables 5 and 6). Most RCCM methods have also outperformed the advanced benchmark methods for the car park, the residential and the road classes in the WV-2 and WV-3 images. It demonstrates the effectiveness of the spatial association in classifying complex land use classes.

In the IKONOS image, the SVM, MPKNN and relearn-PCM results

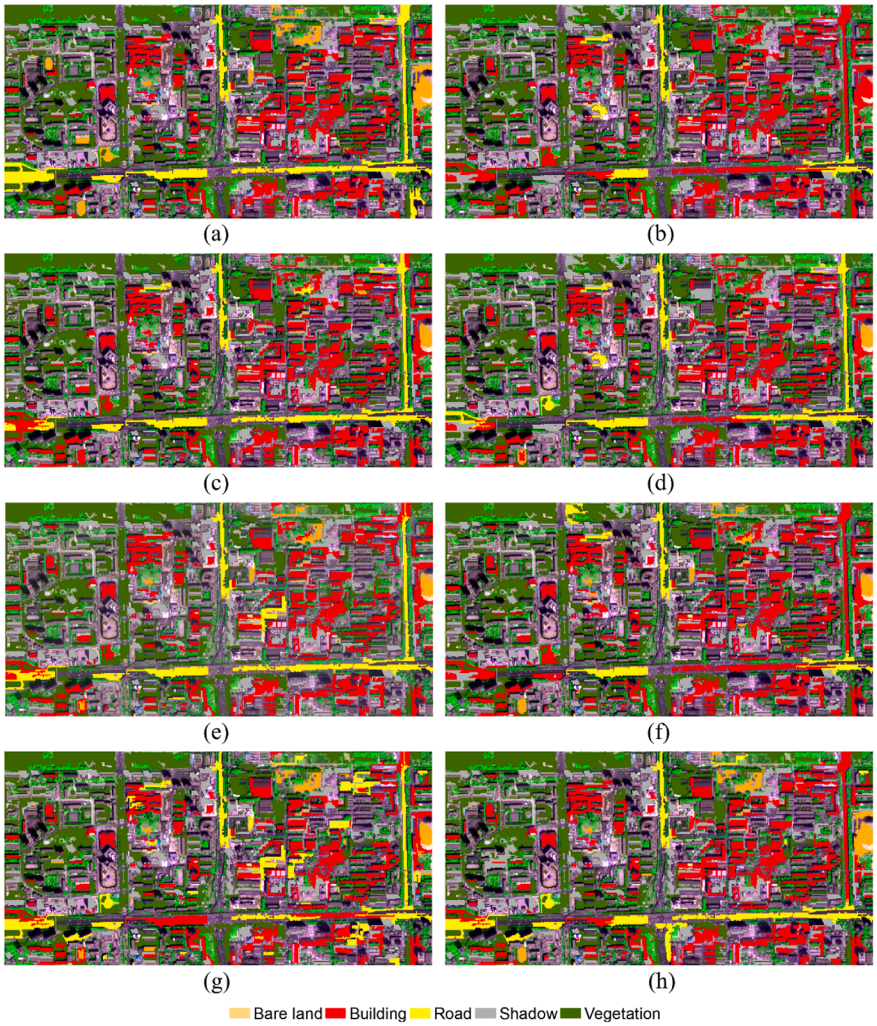


Fig. 4. The testing samples and classification results of the IKONOS image. (a) testing samples, (b) SVM, (c) XGBoost, (d) MPKNN, (e) FCN, (f) relearn-PCM, (g) KL, (h) KL-NN.

(Fig. 4b, d and f) all suffer from the misclassification of roads as buildings. The XGBoost and FCN methods (Fig. 4c and e) do better job in classifying the road. The KL method (Fig. 4g) misclassifies the road in the east-west direction as buildings and a few buildings as the roads. These misclassifications are corrected in the KL-NN result (Fig. 4h), but the road in the south-north direction is misclassified. The FCN, KL and KL-NN methods correctly classify the bare land class at most places, while all other methods fail to do so.

The upper-left part of the WV-2 image mainly includes the residential, vegetation and shadow classes. Since the reflectance of most residential buildings is dark, the advanced classification and the CAM methods (Fig. 5b-g) misclassify the residential class to the non-residential and the shadow classes in the upper-left part. The CAM-NN result (Fig. 5h) has the best performance in classifying the residential class. The road class and the car park class are also easily to be misclassified. The car park class wrongly appears in the middle of the road class in the SVM, XGBoost, MPKNN, relearn-PCM and CAM results (Fig. 5b-d, f, g). The FCN method (Fig. 5e) has the best performance in classifying the main road than the other advanced classification methods, but some narrow and curvilinear roads are still misclassified. Only the CAM-NN method (Fig. 5h) correctly classifies most of the road and the car park classes. The non-residential class is somewhat confused with the residential and the other classes in the advanced classification and CAM results (Fig. 5b-g). Again the CAM-NN method provides the best result in classifying the non-residential class.

Unlike in the WV-2 image, the residential and the non-residential classes tend to be mixed up in the WV-3 image. Specifically, the non-residential class is likely to appear near the road class, and the residential class is likely to be distributed in clusters in the WV-3 image. The relearn-PCM and the CAM-NN methods (Fig. 6g and h) show better performance in differentiate these two classes due to the consideration of spatial association. It also occurs that the road objects are often misclassified as other classes, such as in the SVM, XGBoost, MPKNN, FCN and relearn-PCM results (Fig. 6b-f). Only the CAM-NN method (Fig. 6h) shows a satisfied result for the road class.

3.3.2. RCCM on each class

To compare the RCCM methods with their CM counterparts on each land cover/land use class, Fig. 7 shows the changes in object numbers and the accuracy on the testing samples of the RCCM method over the CM method.

In the IKONOS image, the number for buildings increases sharply

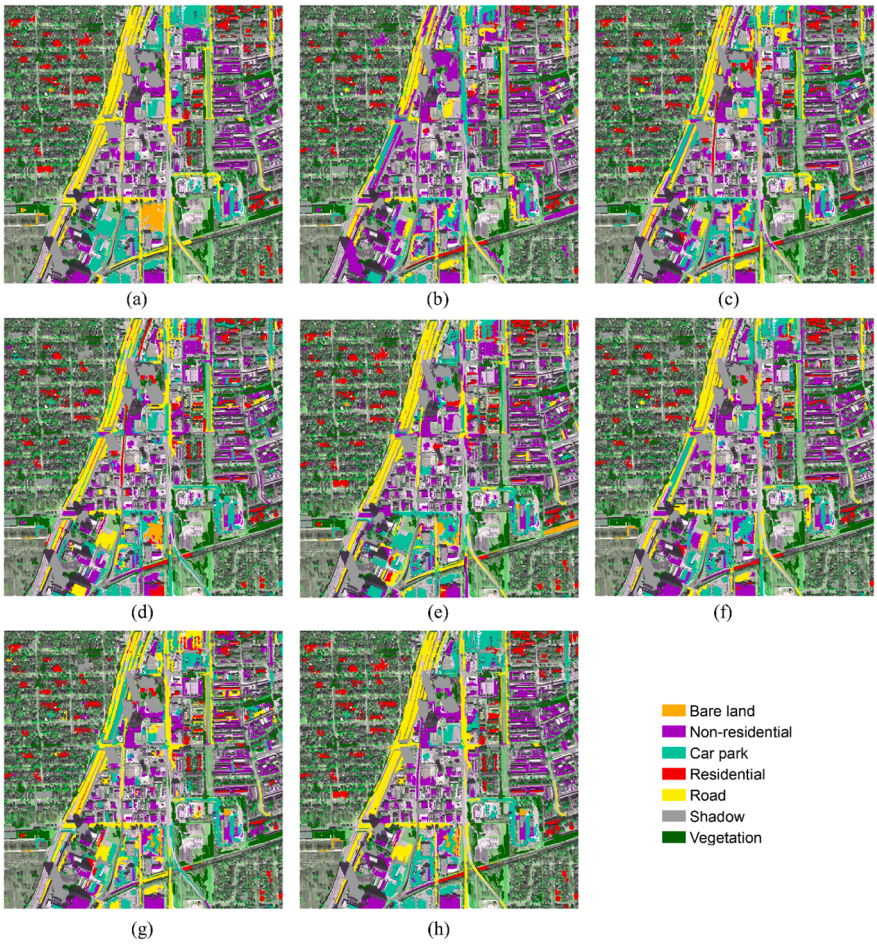


Fig. 5. The testing samples and classification results of the WV-2 image. (a) testing samples, (b) SVM, (c) XGBoost, (d) MPKNN, (e) FCN, (f) relearn-PCM, (g) CAM, (h) CAM-NN.

whereas the numbers for the road class decreases (Fig. 7a). The number for the bare land class either increases slightly or remains unchanged. Most of these changes are correct, leading to improvements in accuracy for all three classes, especially for the road and the bare land classes (Fig. 7b). The vegetation and the shadow classes barely change either in object numbers or accuracies.

In the WV-2 image, the number for the non-residential class increases the most, whereas the number for the road and the car park classes decreases the most (Fig. 7c). The residential class has different changes in object numbers for three classifiers: the CAM-NN and RSSDA-NN methods lead to an increase, the KL-NN method leads to a decrease. The bare land, vegetation and shadow classes barely change in numbers. The accuracies for all the seven classes gain improvements (Fig. 7d), with the car park and the residential classes have the greatest improvements.

In the WV-3 image, the numbers for the residential class and the non-residential class increase, whereas the number for the car park and the non-residential classes decreases (Fig. 7e). Unlike those in the WV-2 image, the number for the bare land and the vegetation classes also varies. Again, the accuracies of all the seven classes increase, with the improvement in the bare land class is the greatest (Fig. 7f). The car park, the non-residential, the residential, and the road classes gain similar improvements.

This test shows that the complex land use classes that tend to be mixed with other classes, such as the residential, non-residential, road, car park and bare land classes, benefit the most from the RCCM method. Since CM methods have already performed well on the vegetation and the shadow classes, these two classes do not change much in accuracy for the RCCM methods.

3.3.3. *Effects of spatial association*

To illustrate how the spatial association introduced in the RCCM methods impacts on the classification, an example is shown for the WV-2 image (Fig. 8). An object to be classified (indicated by a triangle in the object's center in Fig. 8a) belongs to the car park class, but is misclassified using the CAM method. It has been corrected to the car park class using the CAM-NN method after one iteration (Fig. 8a). The curve of this object and the sample curves with the five smallest spectral divergences (Fig. 8b) and spatial divergences (Fig. 8c) are shown. As can be seen, this object is misclassified as the road class using the CAM method, and three out of five most similar spectral sample curves are also of the road class. Not only this object is misclassified in the CAM result, but also one of its neighbors belonging to the car park is misclassified as residential. For the curves of spatial association, the most similar sample curve belongs to the car park class. Therefore, with the introduction of the spatial association, this object has now been correctly classified as a car park. Additionally, its misclassified neighbor has also been corrected after one iteration.

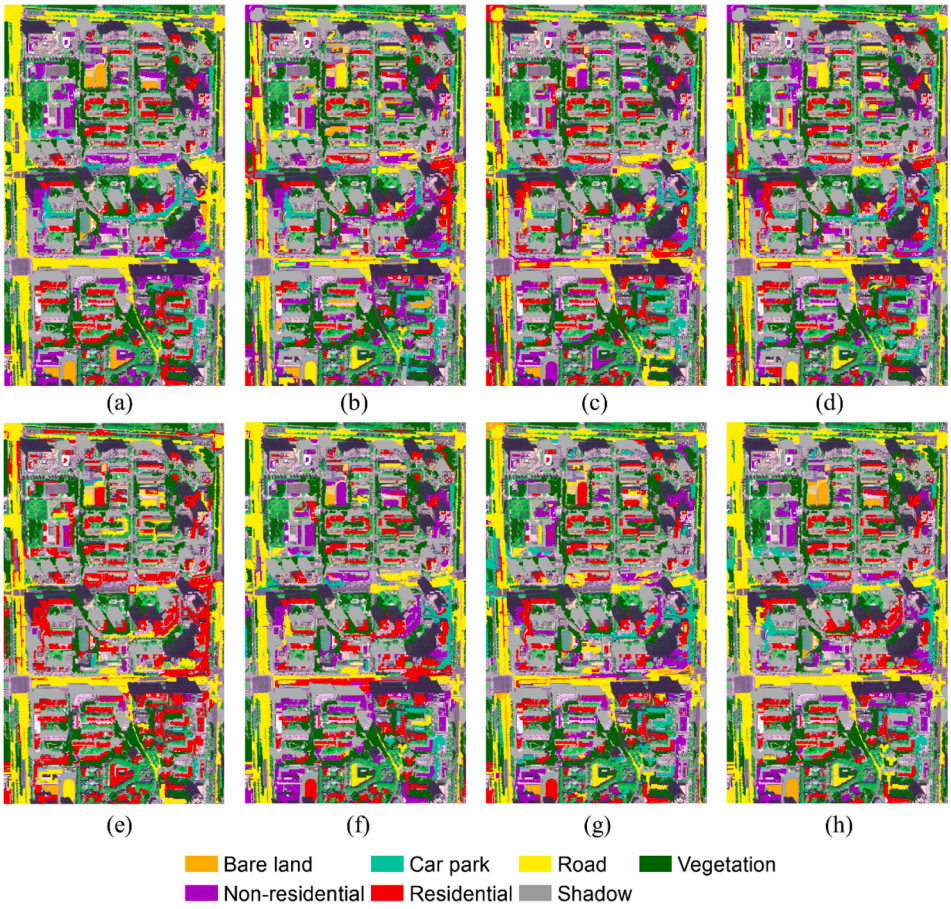


Fig. 6. The testing samples and classification results of the WV-3 image. (a) testing samples, (b) SVM, (c) XGBoost, (d) MPKNN, (e) FCN, (f) relearn-PCM, (g) CAM, (h) CAM-NN.

Table 4
The overall accuracy (%) and the F1-score per class of classification results from one sampling for the IKONOS image.

Method	Total	Bare land	Building	Road	Shadow	Vegetation
RF	92.39	0	0.92	0.48	1	0.99
SVM	90.81	0	0.91	0.36	0.97	0.98
XGBoost	93.70	0.11	0.93	0.80	0.99	0.99
gKNN	91.86	0.20	0.93	0.56	0.97	0.98
MPKNN	92.65	0.21	0.94	0.64	0.97	0.98
FCN	93.96	0.58	0.95	0.76	0.97	0.97
Relearn-Hist	91.60	0.35	0.91	0.42	0.98	0.99
Relearn-PCM	92.65	0.59	0.92	0.43	0.98	0.98
KL	88.71	0.53	0.87	0.37	1	1
CAM	88.45	0.43	0.87	0.31	0.99	0.99
RSSDA	87.40	0.35	0.87	0.30	0.98	0.98
KL-EQ	94.23	0.67	0.94	0.65	1	1
CAM-EQ	92.39	0.62	0.92	0.58	0.99	0.98
RSSDA-EQ	91.34	0.57	0.92	0.45	0.97	0.97
KL-MS	94.49	0.69	0.94	0.59	1	1
CAM-MS	92.91	0.62	0.92	0.57	0.99	0.99
RSSDA-MS	92.13	0.59	0.93	0.55	0.97	0.96
KL-NN	95.28	0.80	0.96	0.62	0.99	0.99
CAM-NN	92.65	0.59	0.93	0.56	0.98	0.98
RSSDA-NN	90.81	0.50	0.91	0.41	0.98	0.98

3.3.4. The recurrent process

It is also of interest to examine how the recurrent process in the RCCM methods influences the classification results. The improved classification results by the recurrent process from one of the RCCM methods with the highest accuracy over the original CM methods are shown in Fig. 9 using one of the ten samplings. The KL-NN has run three iterations to reach convergence for the IKONOS image. The CAM-NN has run six and four iterations to reach convergence for the WV-2 image and the WV-3 image, respectively. Only two intermediate and the final results are shown for these two images.

All three images show certain smoothing effects with the number of iterations increases. The most obvious change for the IKONOS image is the increase of the building class and the decrease of the road class, most of which are correct in the areas where buildings are densely distributed (Fig. 9a). On the other hand, some road objects misclassified as buildings using the KL method are now correctly classified by the recurrent process of the KL-NN method (e.g., in the circles).

In the WV-2 image, an obvious phenomenon is that objects belonging to the same class tend to group together in the recurrent process (Fig. 9b). The residential class and the non-residential class have been mixed up with each other and with many other classes in the initial classification (in the left and right sides of the image). They are being correctly separated from each other and from other classes in the third iteration. However, some narrow roads within the residential and the non-residential areas are over-smoothed in the recurrent process. Luckily, the roads in the south-north direction have strong spatial association, and therefore the CAM-NN method is able to correct some of these roads that were misclassified as car parks in the initial CAM result (e.g., in the circles).

The general class arrangement in the WV-3 image (Fig. 9c) is of typical single-family residential community, with the vegetation, residential and shadow classes generally appear in sequence in the south-north direction. The objects have been misclassified as the non-residential class in the initial CAM result, and are correctly classified as the residential class by the recurrent process (e.g., the bottom circles).

Table 5
The overall accuracy (%) and the F1-score per class of classification results from one sampling for the WV-2 image.

Method	Total	Bare land	Non-residential	Car park	Residential	Road	Shadow	Vegetation
RF	68.49	0.07	0.78	0.45	0.50	0.54	0.84	0.96
SVM	67.43	0.10	0.75	0.44	0.54	0.49	0.78	0.96
XGBoost	69.10	0.15	0.77	0.49	0.57	0.52	0.91	0.95
gKNN	64.84	0.44	0.76	0.33	0.50	0.46	0.87	0.99
MPKNN	64.84	0.44	0.76	0.33	0.50	0.46	0.87	0.99
FCN	72.45	0.31	0.78	0.43	0.67	0.61	0.96	0.99
Relearn-Hist	73.36	0	0.81	0.55	0.58	0.65	0.86	0.97
Relearn-PCM	76.10	0.15	0.82	0.61	0.62	0.71	0.88	0.94
KL	65.14	0.18	0.76	0.35	0.43	0.57	0.85	1
CAM	68.04	0.31	0.79	0.39	0.49	0.54	0.90	1
RSSDA	67.58	0.17	0.80	0.40	0.49	0.48	0.93	0.98
KL-EQ	79.76	0.54	0.89	0.66	0.69	0.72	0.84	0.89
CAM-EQ	84.93	0.75	0.90	0.70	0.88	0.76	0.90	0.94
RSSDA-EQ	78.84	0.34	0.86	0.62	0.71	0.68	0.92	0.98
KL-MS	82.50	0.79	0.88	0.72	0.76	0.78	0.81	0.92
CAM-MS	86.91	0.91	0.92	0.78	0.91	0.80	0.85	0.86
RSSDA-MS	84.17	0.30	0.89	0.75	0.81	0.74	0.94	0.98
KL-NN	79.45	0.52	0.86	0.71	0.75	0.77	0.80	0.81
CAM-NN	88.74	0.42	0.92	0.79	0.93	0.84	0.94	0.97
RSSDA-NN	82.19	0.53	0.88	0.75	0.84	0.71	0.81	0.89

Table 6
The overall accuracy (%) and the F1-score per class of classification results from one sampling for the WV-3 image.

Method	Total	Bare land	Non-residential	Car park	Residential	Road	Shadow	Vegetation
RF	64.79	0	0.40	0.47	0.55	0.60	1	0.98
SVM	61.74	0.06	0.43	0.36	0.50	0.59	0.95	0.98
XGBoost	64.95	0.12	0.44	0.40	0.54	0.63	0.97	0.98
gKNN	63.83	0	0.42	0.37	0.53	0.62	0.97	0.98
MPKNN	63.34	0	0.42	0.37	0.53	0.62	0.97	0.97
FCN	62.06	0.12	0.17	0.18	0.55	0.64	0.98	0.96
Relearn-Hist	71.22	0	0.49	0.52	0.63	0.73	1	0.98
Relearn-PCM	76.37	0.14	0.61	0.57	0.67	0.81	1	0.98
KL	58.52	0.26	0.39	0.19	0.42	0.56	1	0.97
CAM	63.02	0.24	0.41	0.17	0.54	0.63	0.99	0.98
RSSDA	60.61	0.17	0.36	0.20	0.53	0.60	0.99	0.98
KL-EQ	74.76	0.62	0.66	0.42	0.70	0.86	0.83	0.85
CAM-EQ	78.14	0.64	0.69	0.48	0.78	0.83	0.87	0.88
RSSDA-EQ	69.13	0.57	0.62	0.51	0.60	0.79	0.82	0.80
KL-MS	76.05	0.52	0.61	0.37	0.69	0.81	0.98	0.95
CAM-MS	80.55	0.72	0.69	0.46	0.77	0.84	0.99	0.94
RSSDA-MS	77.49	0.70	0.65	0.52	0.72	0.82	0.97	0.94
KL-NN	78.78	0.64	0.66	0.49	0.74	0.82	0.96	0.94
CAM-NN	85.05	0.83	0.77	0.56	0.83	0.89	0.99	0.93
RSSDA-NN	78.94	0.40	0.61	0.51	0.75	0.85	1	0.97

Additionally, most misclassified road objects in the initial CAM result have been filtered out after the first iteration, whereas the correctly classified vegetation objects on the road are well kept in the CAM-NN results (e.g., the left circle).

4. Discussions

4.1. Proportional weight

The weight w in the RCCM method decides the influence of the spectral features. It therefore also determines the influence of the spatial features, which equals 1–w. This parameter was chosen automatically based on the highest validation accuracy for each recurrent iteration in a 5-fold cross-validation. Fig. 10 shows the chosen weight w at each iteration using all nine RCCM methods.

As can be seen, the KL based RCCM methods generally prefer large spectral weights. Most of the final weights are above 0.5 for all three weighting schemes, although the intermediate weight sometimes goes below 0.5 during the recurrent process. The CAM based RCCM methods favor a larger spatial weight for the WV-2 and WV-3 images, but a larger spectral weight for the IKONOS image. The final spectral weights for the RSSDA based RCCM methods in WV-2 and WV-3 images range from 0.2 to 0.6, suggesting the spatial association plays an equally important role during recurrent process. This sensitivity analysis indicates that the RCCM methods favor different weights for different curve matching methods: the KL based methods prefer a larger spectral weight, whereas the CAM and the RSSDA based methods prefer a slightly larger spatial weight.

4.2. Iteration of RCCM

The irregular shapes of the FNEA segmentation results sometimes cause the neighborhood association inaccurate. For example, in Fig. 2a, the object belonging to the bare land class is counted twice in the west direction when the neighborhood association is being detected. Spatial association based on the frequency of class pairs can also reduce the

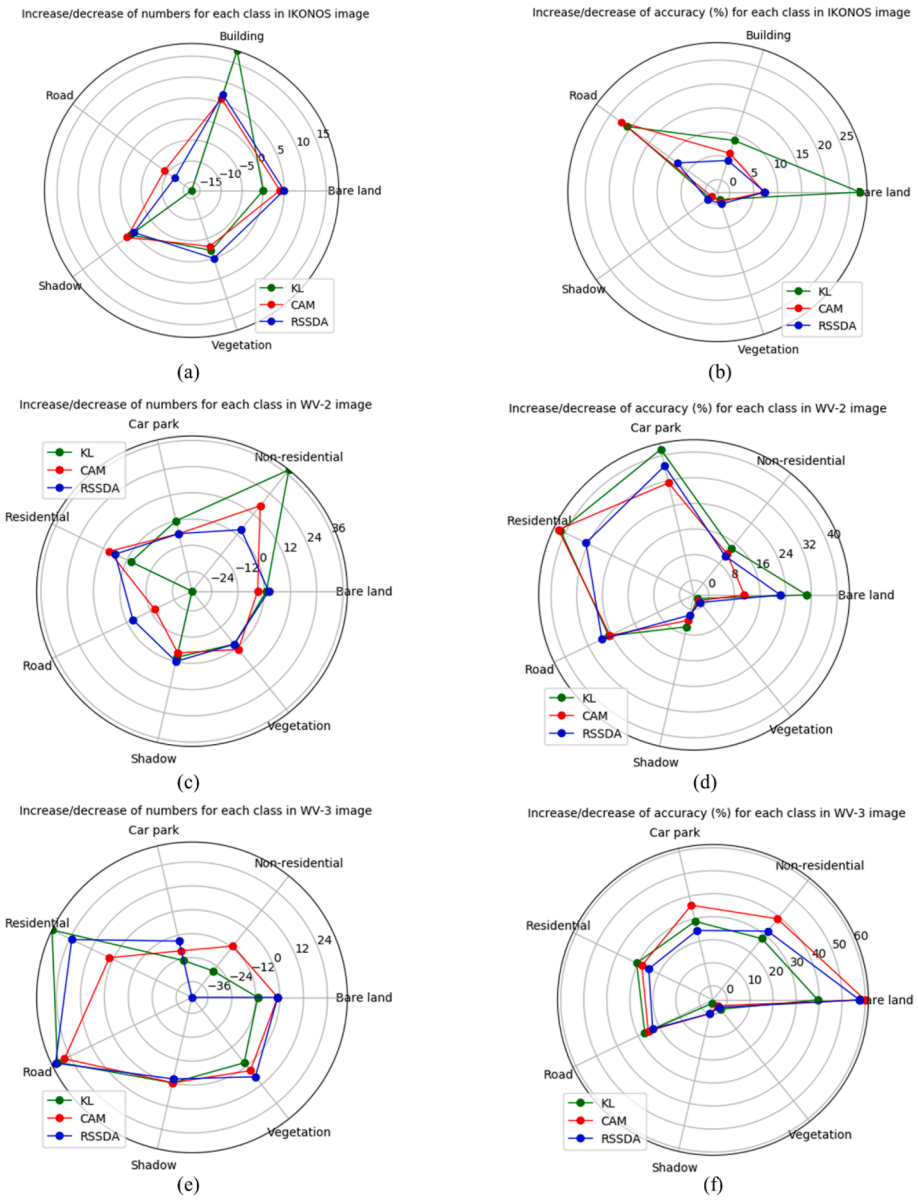


Fig. 7. The changes of numbers and accuracy on testing samples based on the RCCM methods with the NN weighting scheme compared with the CM methods.

effect of this inaccurate neighborhood association, because the most frequent class pairs can reflect the dominant spatial association and the recurrent process further reinforces the dominant spatial association and suppresses the atypical spatial association.

Usually more iterations lead to a more smoothing result, but there are exceptions since a larger spectral weight could suppress the smoothing effects of the recurrent process. The maximum number of iterations was set to 10 in this study. In general, the iterations will be stopped way before this maximum is reached. For example, the accuracy of the IKONOS image with five classes reaches its stability after 2 to 4 iterations, whereas those of the WV-2 and WV-3 images with seven classes reach their stability after 5 to 8 iterations. Some of the RCCM methods reach this maximum when applied to the WV-2 image. As a result, certain places are somewhat over-smoothed (Fig. 9b), thus a value larger than 10 is not desired.

As a non-parametric method, the efficiency of the CM method is not high. Testing on a laptop with 1.8-GHz Intel Core-i7 CPU and 16-GB memory, the CM method costs 87, 501, and 144 s on average for the IKONOS, WV-2 and WV-3 images, respectively. The computational cost of the RCCM method is about 1.8 to 3.8 times longer than the CM method depending on the number of iterations. The generalization of the curve matching method and the recurrent process can be further explored to improve the classification efficiency.

Spatial association of an object is decided by the detected directions and the range distance. In this study, the neighborhood range was set to 6 in four main directions to capture spatial association within a large distance, which allows the exploration of higher-order spatial associations. Traditional OBIA seldom considers higher-order association, because the rook and queen connectivity is commonly used. The three weighting schemes do not show significant difference in classification results, even though the nearest neighboring scheme always results in the highest accuracy.

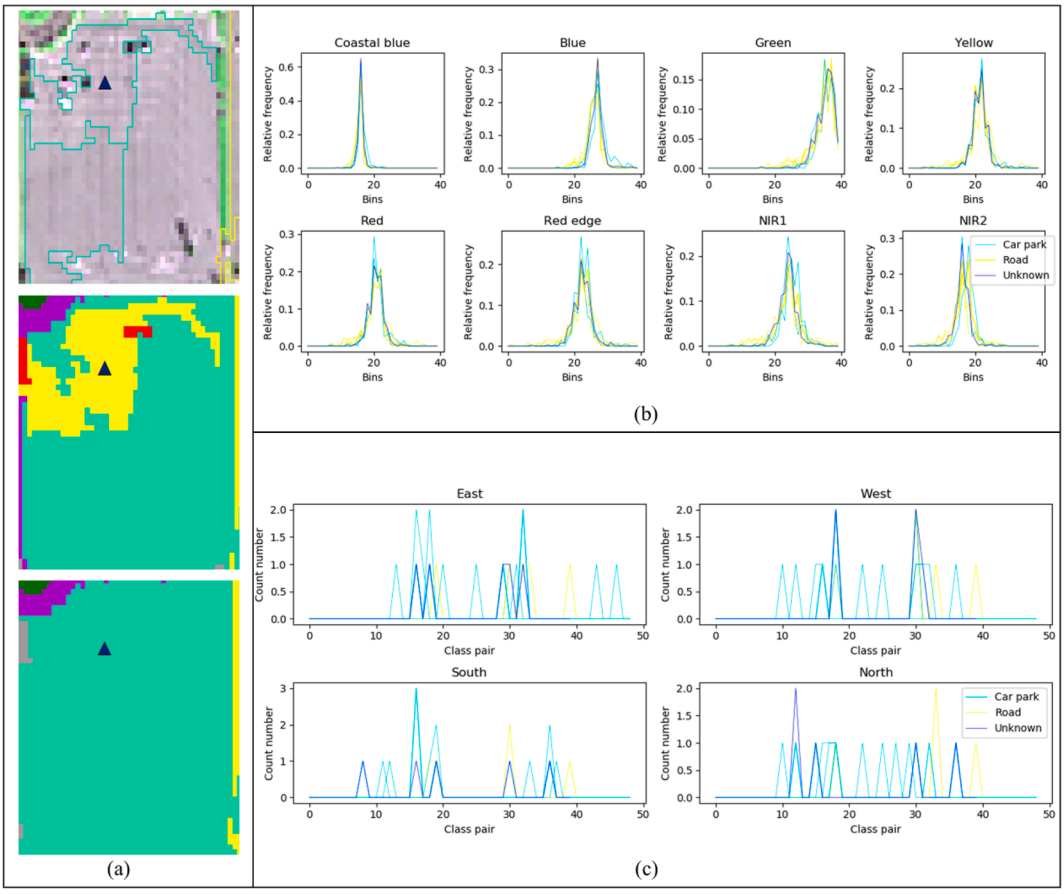


Fig. 8. An example of the RCCM method correctly classifies an object that is misclassified by the CM method. (a) An object belongs to the car park class (marked by a triangle) classified as a road using CAM method, and correctly classified using the CAM-NN method after one iteration, (b) the curves of histograms of this object across multispectral bands and its five most similar sample curves, (c) the curves of spatial association of the object based on the NN weighting scheme and its five most similar sample curves (class code: 1-bare land, 2-non-residential, 3-car park, 4-residential, 5-road, 6-shadow, 7-vegetation).

4.3. Neighborhood and range

To further explore the effects of neighborhood and range distances on the classification accuracy, Fig. 11 compares the classification accuracies of the RCCM methods using the spatial association based on eight directions (east, west, south, north, south-east, south-west, north-east, north-west) with those of the results that were based on four directions. Fig. 12 shows the changes of classification accuracies using the RCCM methods (only the method with the highest accuracy was used for each study area) with the range distance varying from 1 to 10 neighbors.

The accuracy is higher when using four neighbors than that using eight neighbors most of the time (Fig. 11). Due to the irregular shapes of the segmentation results, it is common that the neighboring object in the south direction, for example, is the same with that in the south-east and south-west directions. It is therefore unnecessary to detect the neighbors in all the directions. The accuracy shows an increase and then a decrease when the range distance increases for all three study areas (Fig. 12). It demonstrates that the higher-order spatial association is preferred, and the range of 6 neighbors is sufficient to achieve a satisfied classification accuracy.

4.4. Pros and cons

The curve matching method is capable of integrating large amount of object features without applying feature optimization, which is a very common process in traditional object-based classifications (Pu and Landry, 2012; Löw et al., 2013; Gil-Yepes et al., 2016). Unlike histograms, the *x*-axis of the spatial curve representing class pairs are categorical variables. Different class pair coding can lead different curves, but the results of the three curve matching methods used in the experiments are not affected by the order of the coding, because they are all based on the total divergence between two curves. Some of other curve matching methods, such as the Kolmogorov-Smirnov (KS) test, are based on cumulative curves. Theoretically these methods can also be used for the RCCM methods. However, the cumulative curve is affected by the order of the class code, which means different coding of class pairs would lead different classification results. Therefore, we do not use methods based on cumulative curves in this study.

The extraction of spatial feature in the RCCM methods is realized by simply deriving the weighted frequency of class pairs. In other studies, detecting spatial association is not as straightforward as this one. The well-known variograms or spatial covariance can be used to quantify spatial dependence between objects such as in the gKNN and MPKNN methods, and a modelling process is usually required. However, the type of model and some of its parameters need to be predetermined, which can be subjective and problematic. Additionally, the modelling process in OBIA may suffer from the so-called change-of-support problem (COSP) (Cressie, 1996), because the spatial association may be modelled

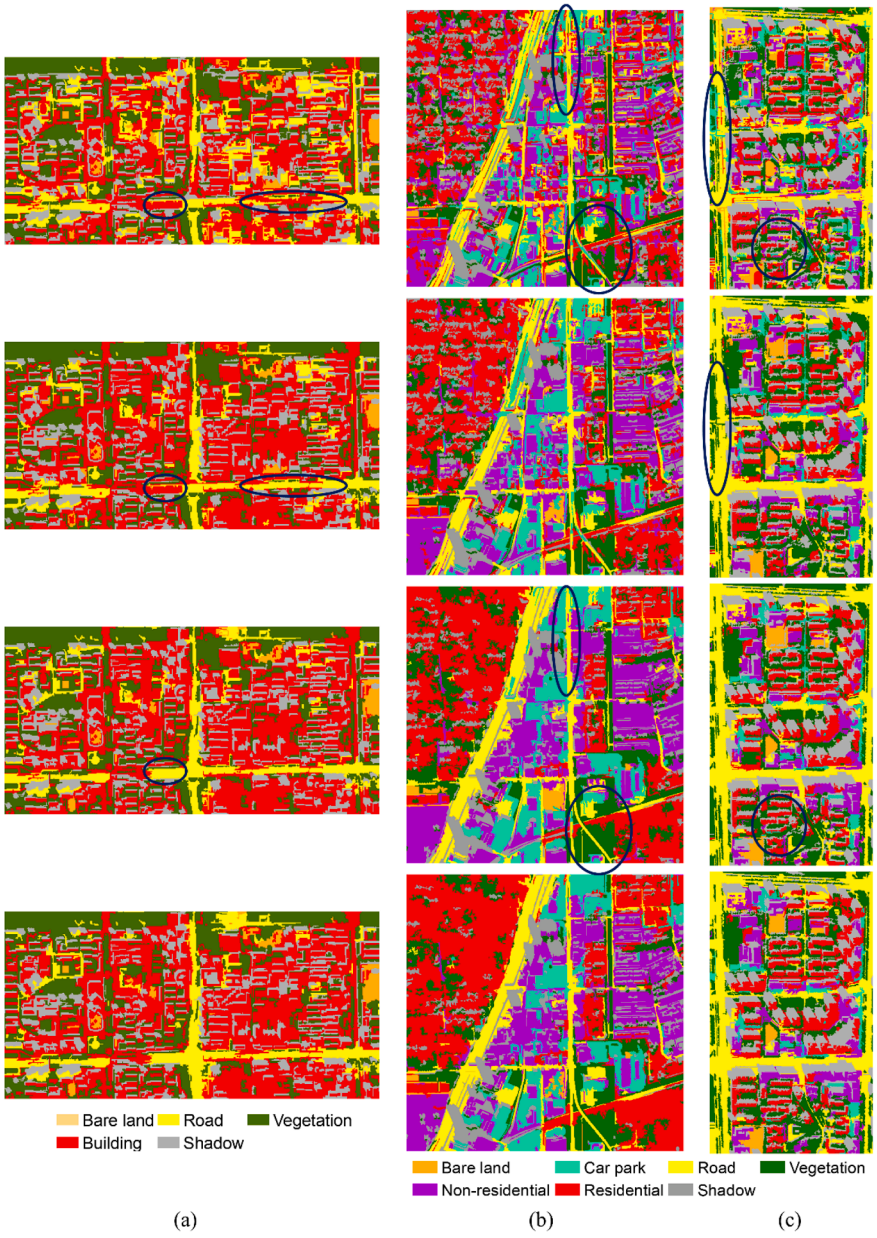


Fig. 9. The original classification, the improved intermediate and final results of the RCCM (from top to bottom): (a) KL, and KL-NN after 1, 2 and 3 iterations for the IKONOS image, (b) CAM, and CAM-NN after 1, 3 and 6 iterations for the WV-2 image, (c) CAM, and CAM-NN after 1, 2 and 4 iterations for the WV-3 image.

inaccurately without considering different geometries of objects (Tang et al., 2018). The current method does not include a modelling process and thus can avoid these problems.

The spatial association is detected only between two classes in this study. Spatial association between more than two classes may be helpful to better capture spatial details of sophisticate curvilinear objects in the classification results and will be explored in the future. The RCCM methods are designed to capture spatial association through introducing the spatial features and recapture the association through the recurrent process. A deep learning based on recurrent neural network (RNN) may achieve a similar goal if sufficiently large number of training samples are available. Albeit tested only in complex urban environments, the RCCM methods should theoretically work well in other environments, which will be also verified in future studies.

5. Conclusion

The developed RCCM method is able to integrate within-object spectral variability and between-object spatial association. The curve matching method provides an easy way to fuse these disparate features in the simple form of curves. The curves derived from histogram of the pixels within an object characterize spectral variability across multi-spectral bands. The curves derived from the weighted frequency of class pairs quantify the spatial association with neighborhood objects. The method can either filter out isolated objects based on the neighborhood association of the same class, or correct misclassification based on spatial arrangements of different classes. The recurrent process reinforces dominant patterns and suppresses atypical patterns by recapturing spatial association from previous classification result to improve

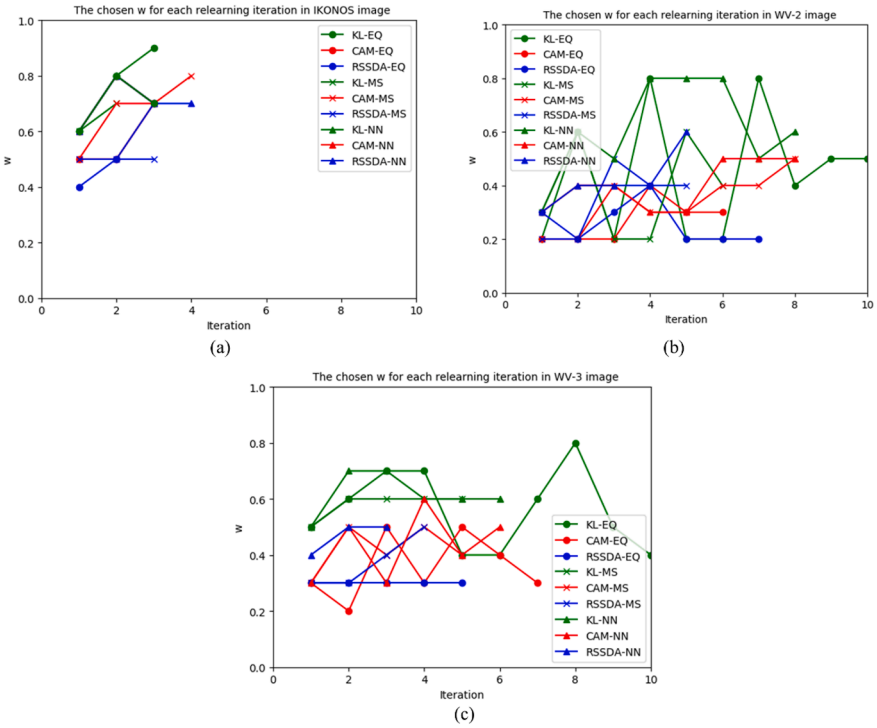


Fig. 10. Plots of iterations against weight *w* for nine RCCM methods: (a) IKONOS image, (b) WV-2 image, and (c) WV-3 image.

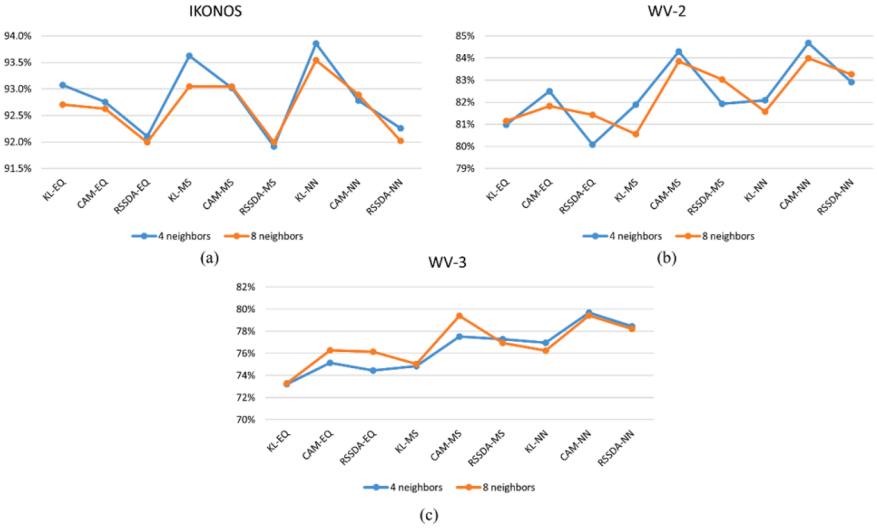


Fig. 11. The changes of classification accuracy with the spatial association detected along four and eight directions for nine RCCM methods: (a) IKONOS image, (b) WV-2 image, and (c) WV-3 image.

classification results. The experiments on the three HSR images demonstrate that the developed method constantly achieves higher accuracy than the curve matching methods using only spectral features and some advanced OBIA classification methods. The method is suited to classify complex urban land use types such as the residential, non-residential, car park, road and bare land classes.

CRediT authorship contribution statement

Yunwei Tang: Methodology, Writing - original draft. **Fang Qiu:** Supervision, Conceptualization, Writing - review & editing. **Linhai Jing:** Resources, Validation. **Fan Shi:** Software, Investigation. **Xiao Li:** Data curation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

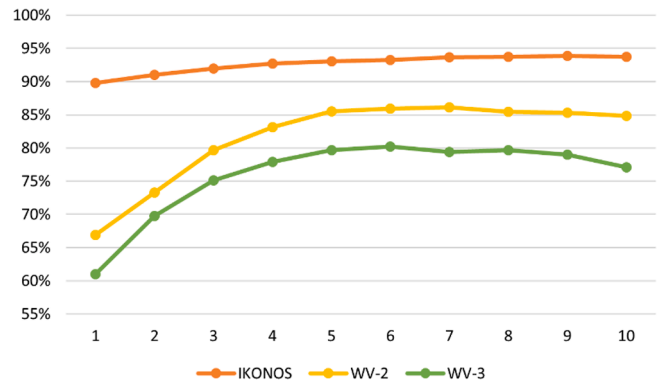


Fig. 12. The changes of classification accuracy with different neighborhood range distances for the RCCM method.

Acknowledgements

This study was supported by the National Key R&D Program of China [grant number 2019YFC1520800], the National Science Foundation (NSF) [grant number BCS-1826839], the key project of the Aerospace Information Research Institute, CAS [grant number Y951150Z2F], and the National Natural Science Foundation of China [grant numbers 41501489, 42071312]. We thank two anonymous reviewers for providing valuable comments and suggestions.

References

Atkinson, P.M., 2004. Spatially weighted supervised classification for remote sensing. Int. J. Appl. Earth Obs. Geoinf. 5, 277–291.
Atkinson, P.M., Naser, D.K., 2010. A geostatistically weighted K-NN classifier for remotely sensed imagery. Geographical Analysis 42, 204–225.
Bai, H., Cao, F., Atkinson, P.M., Chen, Q., Wang, J., Ge, Y., 2020. Incorporating spatial association into statistical classifiers: local pattern-based prior tuning. Int. J. Geographical Inform. Sci. 34 (10), 2077–2114.
Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., Heynen, M., 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS J. Photogramm. Remote Sens. 58, 239–258.
Blaschke, T., Hay, G.J., Kelly, M., Lang, S., Hofmann, P., Addink, E.A., Queiroz Feitosa, R., van der Meer, F., van der Werff, H., van Coillie, F., Tiede, D., 2014. Geographic object-based image analysis - towards a new paradigm. ISPRS J. Photogramm. Remote Sens. 87, 180–191.
Bruzzone, L., Carlin, L., 2006. A multilevel context-based system for classification of very high spatial resolution images. IEEE Trans. Geosci. Remote Sens. 44, 2587–2600.
Chaib, S., Liu, H., Gu, Y., Yao, H., 2017. Deep feature fusion for VHR remote sensing scene classification. IEEE Trans. Geosci. Remote Sens. 55 (8), 4775–4784.
Chen, S., Member, S., Wang, H., Xu, F., Member, S., 2016a. Target classification using the deep convolutional networks for SAR images. IEEE Trans. Geosci. Remote Sens. 54, 4806–4817.
Chen, T., Guestrin, C., 2016. XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13-17 August 2016; p. 785.
Chen, Y., Jiang, H., Li, C., Jia, X., Member, S., 2016b. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. IEEE Trans. Geosci. Remote Sens. 54, 6232–6251.
Cressie, N., 1996. Change of support and the modifiable areal unit problem. Geographical Systems 3, 159–180.
Ding, P., Zhang, Y., Deng, W.-J., Jia, P., Kuijper, A., 2018. A light and faster regional convolutional neural network for object detection in optical remote sensing. ISPRS J. Photogramm. Remote Sens. 141, 208–218.
Fu, G., Liu, C., Zhou, R., Sun, T., Zhang, Q., 2017. Classification for high resolution remote sensing imagery using a fully convolutional network. Remote Sensing 9, 498.
Gil-Yepes, J.L., Ruiz, L.A., Recio, J.A., Balaguer-Beser, Á., Hermosilla, T., 2016. Description and validation of a new set of object-based temporal geostatistical features for land-use/land-cover change detection. ISPRS J. Photogramm. Remote Sens. 121, 77–91.
Hay, G.J., Niemann, K.O., McLean, G., 1996. An object-specific image-texture analysis of H-resolution forest imagery. Remote Sens. Environ. 55, 108–122.
Huang, X., Lu, Q., Zhang, L., Plaza, A., 2014. New postprocessing methods for remote sensing image classification: A systematic study. IEEE Trans. Geosci. Remote Sens. 52, 7140–7159.
Huang, X., Zhang, L., 2011. A multidirectional and multiscale morphological index for automatic building extraction from multispectral GeoEye-1 imagery. Photogramm. Eng. Remote Sens. 7 (7), 721–732.
Huang, X., Zhang, L., 2012. Morphological building/shadow index for building extraction from high-resolution imagery over urban areas. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 5 (1), 161–172.
Johnson, B., Xie, Z., 2013. Classifying a high resolution image of an urban area using super-object information. ISPRS J. Photogramm. Remote Sens. 83, 40–49.
Kullback, S., 1987. The Kullback-Leibler Distance. The American Statistician 41 (4), 340–341.
Kruse, F.A., Lefkoff, A.B., Boardman, J.B., Heidebrecht, K.B., Shapiro, A.T., Barloon, P.J., Goetz, A.F.H., 1993. The Spectral Image Processing System (SIPS) - interactive visualization and analysis of imaging spectrometer data. Remote Sens. Environ. 44, 145–163.
Liu, T., Abd-Elranhman, A., Morton, J., Wilhelm, V.L., 2018. Comparing fully convolutional networks, random forest, support vector machine, and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system. GIScience & Remote Sensing 55 (2), 243–264.
Löw, F., Michel, U., Dech, S., Conrad, C., 2013. Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using support vector machines. ISPRS J. Photogramm. Remote Sens. 85, 102–119.
Lv, Z., Shi, W., Benediktsson, J.A., Ning, X., 2016. Novel object-based filter for improving land-cover classification of aerial imagery with very high spatial resolution. Remote Sensing 8, 1023.
Pu, R., Landry, S., 2012. A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species. Remote Sens. Environ. 124, 516–533.
Solberg, A.H.S., Taxt, T., Jain, A.K., 1996. A Markov random field model for classification of multisource satellite imagery. IEEE Trans. Geosci. Remote Sens. 34, 100–113.
Sridharan, H., Qiu, F., 2013. Developing an object-based hyperspatial image classifier with a case study using WorldView-2 data. Photogramm. Eng. Remote Sens. 79 (11), 1027–1036.
Stehman, S.V., 2009. Sampling designs for accuracy assessment of land cover. Int. J. Remote Sens. 30 (20), 5243–5272.
Stow, D.A., Toure, S.I., Lippitt, C.D., Lippitt, C.L., Lee, C., 2012. Frequency distribution signatures and classification of within-object pixels. Int. J. Appl. Earth Obs. Geoinf. 15 (1), 49–56.
Sun, L., Wu, Z., Liu, J., Xiao, L., Wei, Z., 2015. Supervised spectral-spatial hyperspectral image classification with weighted Markov random fields. IEEE Trans. Geosci. Remote Sens. 53, 1490–1503.
Tarabalka, Y., Fauvel, M., Chanussot, J., Benediktsson, J.A., 2010. SVM-and MRF-based method for accurate classification of hyperspectral images. IEEE Geosci. Remote Sens. Lett. 7, 736–740.
Tang, Y., Jing, L., Li, H., Atkinson, P.M., 2016a. A multiple-point spatially weighted k-NN method for object-based classification. Int. J. Appl. Earth Obs. Geoinf. 52, 263–274.
Tang, Y., Jing, L., Li, H., Liu, Q., Yan, Q., Li, X., 2016b. Bamboo classification using WorldView-2 imagery of giant panda habitat in a large shaded area in Wolong, Sichuan Province. China. Sensors 16 (11), 1957.
Tang, Y., Qiu, F., Jing, L., Shi, F., Li, X., 2020. Integrating spectral variability and spatial distribution for object-based image analysis using curve matching approaches. ISPRS J. Photogramm. Remote Sens. 169, 320–336.
Tang, Y., Zhang, J., Jing, L., Gao, H., 2018. Geostatistical modelling of spatial dependence in area-class occurrences for improved object-based classifications of remote-sensing images. ISPRS J. Photogramm. Remote Sens. 141, 219–236.
Toure, S.I., Stow, D.A., Weeks, J.R., Kumar, S., 2013. Histogram curve matching approaches for object-based image classification of land cover and land use. Photogramm. Eng. Remote Sens. 79 (5), 433–440.
Wan, H., Tang, Y., Jing, L., Li, H., Qiu, F., Wu, W., 2021. Tree species classification of forest stands using multisource remote sensing data. Remote Sensing 13, 144.
Wang, L., Dong, Q., Yang, L., Gao, J., Liu, J., 2019. Crop classification based on a novel feature filtering and enhancement method. Remote Sensing 11 (4), 455.
Yu, X., Wu, X., Luo, C., Ren, P., 2017. Deep learning in remote sensing scene classification: a data augmentation enhanced convolutional neural network framework. GIScience & Remote Sensing 54 (5), 741–758.
Zhang, C., Harrison, P.A., Pan, X., Li, H., Sargent, I., Atkinson, P.M., 2020a. Scale Sequence Joint Deep Learning (SS-JDL) for land use and land cover classification. Remote Sens. Environ. 237, 111593.
Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J., Atkinson, P.M., 2018. An object-based convolutional neural network (OCNN) for urban land use classification. Remote Sens. Environ. 216, 57–70.
Zhang, C., Yue, P., Tapete, D., Shangguan, B., Wang, M., Wu, Z., 2020b. A multi-level context-guided classification method with object-based convolutional neural network for land cover classification using very high resolution remote sensing images. Int. J. Appl. Earth Obs. Geoinf. 88, 102086.
Zhong, Y., Zhao, B., Zhang, L., 2014. Multiagent object-based classifier for high spatial resolution imagery. IEEE Trans. Geosci. Remote Sens. 52 (2), 841–857.
Zhou, Y., Qiu, F., 2015. Fusion of high spatial resolution WorldView-2 imagery and LiDAR pseudo-waveform for object-based image analysis. ISPRS J. Photogramm. Remote Sens. 101, 221–232.