

Change Detection based on Stacked Generalization System with Segmentation Constraint

Kun Tan, Yusha Zhang, Qian Du, Peijun Du, Xiao Jin, and Jiayi Li

Abstract

Change detection based on a multi-classifier ensemble system can take advantage of multiple classifiers to extract change information in remote sensing images. In this paper, an efficient heterogeneous ensemble algorithm, i.e., the stacked generalization (SG) combined with image segmentation, is proposed to construct a simple multi-classifier ensemble system that can offer better detection accuracy with lower computational cost. Due to the rich spatial information in high-spatial-resolution remote sensing images, structure texture (morphological) and statistical texture features are extracted to construct the input data to the ensemble system along with spectral features. In addition, constrained analysis on segmented objects integrates the smaller heterogeneity segmentation map and pixel-wise change map to generate the final change map. The experiments were carried out on two ZY-3 and a QuickBird dataset. The results show that the proposed algorithm can integrate the advantages of both pixel-wise ensemble and object-oriented methods, and effectively improve the accuracy and stability of change detection.

Introduction

Change detection techniques based on multi-temporal remote sensing images have been widely applied to all aspects of national production and life, such as urban development monitoring, mine environmental change detection, and environmental disaster monitoring (Huang *et al.*, 2013; Jawak *et al.*, 2014; Malmir *et al.*, 2015).

Because of rich information of objects in high-resolution remote sensing images, change detection of high-resolution remote sensing images has become a popular research topic in remote sensing applications (Wen *et al.*, 2015; Huang *et al.*, 2017). According to the degree of automation, change detection algorithms for high-resolution remote sensing images can be broadly divided into unsupervised (Lv *et al.*, 2015) and supervised methods (Volpi *et al.*, 2013; Hou *et al.*, 2015). Unsupervised change detection obtains change information without additional information. The widely used unsupervised methods include change vector analysis (CVA) (Chen *et al.*, 2003; Azzouzi *et al.*, 2015), Otsu's thresholding method (Bruzzone and Prieto,

2000), and Markov random field based methods (MRF) (Moser *et al.*, 2011; Benedek *et al.*, 2015). In contrast, supervised change detection involves extracting change information through mining knowledge from prior information and mainly includes post-classification comparison and direct classification. The first one is to classify each temporal image via supervised classification and then compare classification maps to determine changes, and the other one is to directly classify the images based on selected training samples. Because of strong learning ability of individual classifiers, they are used to extract the data information in change detection. The individual classifiers that are commonly used are support vector machine (SVM) (Nemmour and Chibani, 2006; He and Laptev, 2009), k -nearest neighbor (KNN) (Guo *et al.*, 2003; Roy *et al.*, 2012), multinomial logistic regression (MLR) (Li, Bioucas *et al.*, 2012; Khodadadzhdeh *et al.*, 2014), the extreme learning machine (ELM) (Huang *et al.*, 2006; Chang *et al.*, 2010). However, no single classifier is capable of extracting all change information, and all learning algorithms have their limits. In order to improve generalization ability, ensemble learning (Nemmour and Chibani, 2006; Chel-lasamy, Ferré *et al.*, 2014; Roy *et al.*, 2014) has been introduced into change detection of high-resolution remote sensing images, including both homogeneous and heterogeneous ensemble systems. The homogeneous ensemble system means that the same classifier is used for different training samples, and the research focus is the construction of different training samples. The heterogeneous ensemble system is a combination of different classifiers, and then different fusion strategies are applied to integrate the classification results to generate the final result. The widely utilized homogeneous ensemble algorithms are bagging algorithms (Skurichina *et al.*, 2002), random subspace method (RSM) (Skurichina and Duin, 2010; Xia *et al.*, 2015), AdaBoost (Woo and Do, 2015), rotation forest (Rodriguez, Kuncheva *et al.*, 2006). The heterogeneous ensemble algorithms are mainly concerned with the choice of classifiers and fusion strategy. The fusion strategies include majority voting (MV) (Rojarath *et al.*, 2017), the Dempster-Shafer (D-S) evidential reasoning method (Peng and Zhang, 2017), the fuzzy integral (F_int) method (Nemmour and Chibani, 2013), meta-learning (Lin *et al.*, 2009). Many scholars have shown that ensemble systems can improve the accuracy of change detection. Du *et al.* (Du *et al.*, 2012) proposed a multiple classifier system (MCS) based on individual base classifiers and obtained a satisfying classification result. Roy *et al.*, 2012 (Roy *et al.*, 2012) developed a semi-supervised ensemble system based on multilayer perceptron, elliptical basis function neural network, and fuzzy KNN techniques, and showed satisfying detection performance.

Kun Tan, Yusha Zhang and Xiao Jin are with the Key Laboratory for Land Environment and Disaster Monitoring of NASG, China University of Mining and Technology, Xuzhou 221116, China (tankuncu@gmail.com).

Qian Du is with the Department of Electrical and Computer Engineering, Mississippi State University, MS 39762.

Peijun Du is with the Key Laboratory for Satellite Mapping Technology and Applications of NASG, Nanjing University, Nanjing 210023, China.

Jiayi Li is with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China.

Photogrammetric Engineering & Remote Sensing
Vol. 84, No. 11, November 2018, pp. 733–741.
0099-1112/18/733–741

© 2018 American Society for Photogrammetry
and Remote Sensing
doi: 10.14358/PERS.84.11.733

Thus, we use a stacked generalization (SG) (Ting and Witten, 2002; Hatami and Ebrahimpour, 2007; Sesmero *et al.*, 2015) for change detection of high-resolution remote sensing images. First, ELM, SVM, and KNN are utilized as the base classifiers at level-0. ELM, proposed by Huang (2005), has a fast learning speed and strong generalization ability, so we first classify most of pixels by ELM. SVM, proposed by Vapnik (Vapnik and Vladimir, 1995) in 1995, solves non-linear problems well and can avoid local minima problem. KNN is a simple algorithm which has strong robustness and classifies using analogy, and has a good classifying ability in local image. Multi-response linear regression (MRLR) is then used to construct the meta-classifier at level-1. Ting and Witten (2002) have tested four methods - C4.5, IB1, NB, and MRLR as meta-classifier in hybrid ensemble system, and found that only the MRLR ensemble achieved satisfying results. Seewald (2002) also found that MRLR can effectively solve the binary classification problem. For each sample, MRLR utilizes the predicted values of the L base classifiers to construct the input feature data but ignores the association with neighboring pixels. In this paper, considering spatial information, the weighted average of eight neighboring pixels is also taken into account when constructing the feature data for the meta-classifier. The fruit fly optimization algorithm (FOA) proposed by Pan (2012) is used to solve the parameters of MRLR.

With the increase of spatial resolution and the decrease of spectral resolution, spectral separability for similar objects is reduced. Accuracy of change detection is degraded due to “different spectrum with same objects”. Thus, object-oriented technique becomes one of the most popular methods for high-spatial-resolution images (Hao *et al.*, 2016; Peng and Zhang, 2017). However, accuracy of change detection in object-oriented methods is directly influenced by the initial image segmentation. Another effective way is constructing multi-source image features. Li *et al.* (Li, Huang *et al.*, 2017) proposed a method by integrating macro and micro-texture together and obtained high accuracy. Volpi *et al.* (2013) stacked spectral features, texture features, and morphological features to perform supervised change detection in very high-resolution images. Peng *et al.* (2017) extracted texture and spatial features by using LBP and Sobel gradient and combined them with spectral features to obtain the change information for high-resolution GF-1 image. These studies have demonstrated that the inclusion of texture and morphological features can compensate for the lack of detailed spectral information. So, in this paper, the statistical texture and structure texture features (morphological profiles) are utilized to compensate for insufficient spectral information and the initial pixel wise change map is combined with the smaller heterogeneity multi-scale segmentation map to obtain the final change map, where the influence of over-segmentation in object-oriented change detection is alleviated and the “salt and pepper noise” in pixel-wise based change detection is reduced. In experiments, two ZY-3 and QuickBird datasets are used to demonstrate the effectiveness of the proposed method.

The remainder of this paper is organized as follows. The next section introduces the proposed methodology. Then, the experimental results are discussed and analyzed. Finally, the conclusion is drawn.

Methodology

Stacked Generalization (SG)

SG is a heterogeneous ensemble algorithm, integrating different classifiers, followed by a two-level hierarchical structure (Figure 1). At level-0, each base classifier is trained by the original training set, and for each pixel, each base classifier produces a predicted value. These predicted values are used as the input data for level-1. At level-1, a trainable combiner

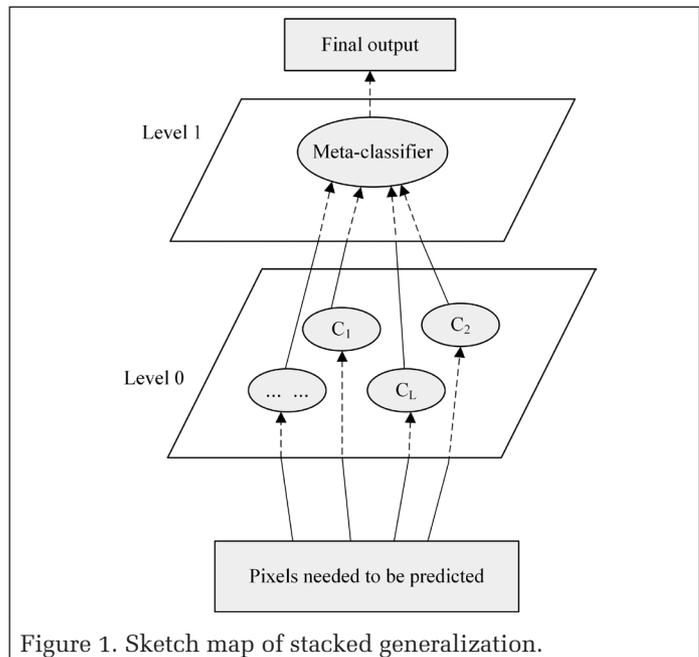


Figure 1. Sketch map of stacked generalization.

integrates the output data of each base classifier at level-0 and obtains the final prediction. This trainable combiner is also called “meta-classifier”. The overall effect of the classifier ensemble system depends on the base classifier used at level-0 and the selection of the meta-classifier at level-1.

Given a training set S , SG randomly divides the original training set into J sub-training sets of equal size and K base classifiers are used at level-0. The next steps are similar to the J -fold cross-validation process: Select a sub-training set (S_j) for training and validation of level-1, $j = 1, 2, \dots, J$. The remaining sub-training sets $S^{(-j)} = S - S_j$ are used to train the K base classifiers at level-0, and S_j is the test dataset for the K base classifiers. The base classifiers at level-0 predict all the samples in S_j . The predicted values of S_j and their training labels then form the training set for the meta-classifier at level-1. Each sample in the training set has K features (K predicted values). After training, the meta-classifier generates a fixed classification model, and for a new test dataset, the level-0 model generates an initial prediction vector as the input values of the level-1 model. The level-1 model then generates the final prediction values. The overall description of the SG algorithm is summarized as follows.

Input:
 z_train : the sub-training set ($S^{(-j)}$) of base classifiers;
 x_label : the labels of sub-training set ($S^{(-j)}$) of base classifiers;
 z_test : the test samples (S_j) of base classifiers;
 y_label : the labels of training set (S_j) of meta-classifier;
 z : the test dataset.
For all the training samples in set $S^{(-j)}$:
 Train the K base classifiers at level-0 to estimate the parameter a :

$$x_label = a * z_train$$

End for
For all the test samples in S_j :
 Get the predicted values (x_value) of base classifiers at level-0:

$$x_value = a * z_test$$

End for
 Train the meta-classifier at level-1 to estimate the parameter of weight w :

$$y_label = w * x_value$$

 Predict a new test dataset:

$$y_value = w * a * z$$

Output:
 y_value : the predicted value of the SG model.

Multi-Response Linear Regression (MRLR)

The MRLR model is an effective method for the ensemble of heterogeneous base classifiers. The advantages of using MRLR is its interpretability as it provides a method of combining the results generated by the level-0 into a final decision. The weights generated by MRLR indicate the different contributions that each base classifier makes for class prediction, which can be described as follows. Suppose the training sample set $\Phi = \{(x_i, y_i)\}_{i=1}^N$ contains N observations, where $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T$ is a p -dimensional eigenvector, y_i is a class label and $y_i \in \Gamma = \{w_1, w_2, \dots, w_m\}$. We use the training set Φ to train L different classification algorithms to obtain the integration $\zeta = \{C_1, C_2, \dots, C_L\}$ of the L base classifiers. We assume that each base classifier $C_i (i = 1, 2, \dots, L)$ predicts an observed value as a posterior probability distribution vector:

$$P_{C_i}(x) = (P_{C_i}(w_1|x), P_{C_i}(w_2|x), \dots, P_{C_i}(w_m|x))^T = (P_{C_i}^1(x), P_{C_i}^2(x), \dots, P_{C_i}^m(x))^T, i = 1, 2, \dots, L \quad (1)$$

where $P_j^i(x)$ is the possibility value of the pixels in the w_j class obtained by the i^{th} base classifier. We can therefore describe the input data of the meta-classifier as a $m \times L$ matrix $P(x)$:

$$P(x) = (P_1^T(x), P_2^T(x), \dots, P_L^T(x))^T = \begin{pmatrix} P_1^1(x), \dots, P_1^m(x), P_2^1(x), \dots, P_2^m(x), \dots, P_L^1(x), \dots, P_L^m(x) \end{pmatrix}^T \quad (2)$$

MRLR transforms the C classification problems into C regression problems. For example, for class w_j , if the sample has a class label w_j , its output value is 1; otherwise, the output value is 0. For each class w_j , MRLR chooses each base classifier's predicted x belonging to class w_j to establish a linear model, which is defined as:

$$LR_j(x) = \sum_{i=1}^L a_i^j P_i^j(x), a_i^j \geq 0, j = 1, 2, \dots, m \quad (3)$$

and the estimation of parameter $\{a_i^j\}_{j=1}^m$ usually utilizes the NNLS algorithm.

For each sample, MRLR utilizes the predicted values of the L base classifiers to construct the input feature data but ignores the association with neighboring pixels. In this paper, considering the spatial information, the weighted average of the sample's eight neighbors is also taken into account when constructing the feature data of the meta-classifier. The specific input data can be represented as:

$$Q(x) = (P_1^T(x), Q_1^T(x), \dots, P_L^T(x), Q_L^T(x))^T = \begin{pmatrix} P_1^1(x), \dots, P_1^m, Q_1^1(x), \dots, Q_1^m, \dots, P_L^1(x), \dots, P_L^m(x), Q_L^1(x), \dots, Q_L^m(x) \end{pmatrix}^T \quad (4)$$

where $Q_i^j(x)$ is the weighted average of the probabilities of the eight neighboring pixels in the w_j class obtained by the i -th base classifier.

To estimate the model parameters in MRLR, we propose the FOA and compare it with the NNLS algorithm (Li and Ngom, 2013). NNLS algorithm is the most commonly used method for parameter estimation of the MRLR model. The FOA is one of the recently developed swarm optimization algorithms, and it has global optimization ability (Iscan and Gunduz, 2015). Besides, FOA is a stable algorithm, which solves the problems fast.

Construction of Multi-Source Feature Dataset and Automatic Selection of Training Samples

As aforementioned, many studies have demonstrated the effectiveness of combination of texture, morphological, and spectral features. The gray level co-occurrence matrix (GLCM)

is a conventional method of extracting statistical texture features. In this paper, five second-moment descriptors, i.e., mean, variance, homogeneity, contrast, and dissimilarity, are applied. For the selection of window size, according to the size and distribution of various features in the image, we choose 5×5 size window and 0° direction to extract the features. The morphological features are also a type of texture features called structure texture. Two commonly used morphological operators are opening and closing. The mathematical morphology framework defines a series of operators to emphasize homogeneous spatial structures in a gray level image. The strategy of opening reconstruction is to dilate an eroded image in order to recover as much as possible of the eroded image. In contrast, closing reconstruction is to erode a dilated image in order to recover the initial shape of image structures that have been dilated. The opening-and-closing reconstruction integrates the advantages of both operations regarding their capacity to preserve original shapes of spatial structures. Therefore, these three morphological reconstruction filters are used to construct the input dataset. According to the distribution of features in images, a circular structure with a radius of 5 is chosen as the structuring element.

Despite the advantages of supervised classifiers in classification, they require training samples as labeled beforehand. Manual selection of training samples can lead to incompleteness of selected categories, and it is time-consuming. So in this paper, the training samples are selected by Change Vector Analysis (CVA), an unsupervised change detection method. CVA is very effective in combining different types of change features. The training samples are selected from the change map by using two thresholds and defined as:

$$\begin{cases} t1 = [T + k * \delta_{-c}, T + k * \delta_{-c} + \frac{1}{(a - k + 1)} * \delta_{-c}] \\ t2 = [T - l * \delta_{-nc}, T - l * \delta_{-nc} + \frac{1}{(b - l + 1)} * \delta_{-nc}] \end{cases} \quad (5)$$

where T is determined by the expectation maximization (EM) algorithm, the δ_{-c} and δ_{-nc} are the standard deviation of the changed pixels and unchanged pixels, respectively, and k and l are the adjustment coefficients as $k = 1, 2, \dots, a$, $l = 1, 2, \dots, b$. Here, $a = (x_{max} - T) / \delta_{-c}$ and $b = (T - x_{min}) / \delta_{-nc}$ with x_{max} and x_{min} being the maximum and minimum value of the CVA change map, respectively.

Pixel-wise Change Detection Based on the Stacked Generalization Hybrid Ensemble System

As mentioned earlier, ELM, SVM, and KNN are chosen to construct the base classifiers at level-0. The MRLR is utilized as the meta-classifier at level-1. In order to improve computational efficiency and ensure a high accuracy, the ELM homogeneous integration algorithm based on random subspace method (RSM) is adopted to label a large part of pixels. The remaining unlabeled pixels are then classified by the proposed SG hybrid ensemble system. The specific change detection processes are as follows.

1. Generation of the level-0 base classifier

As described in the previous section, we randomly divide the automatically acquired training samples into three sub-training sets, then we utilize two parts to train ELM, SVM, and KNN to generate the base classifiers at level-0. When training the ELM, the two sub-training sets use the RSM ensemble strategy to classify all the pixels. According to the label determination rules, a large number of pixels are labeled, and the remaining pixels are reclassified by the trained SVM and KNN. The outputs of ELM, SVM, and KNN based on the RSM homogeneous integration and the

weighted average of the eight neighboring pixels are used to construct the level-1 input feature data :

$$\left\{ \begin{aligned} Q_{SG}(x) &= (P_{ELM}^T(x), Q_{ELM}^T(x), P_{SVM}^T(x), Q_{SVM}^T(x), P_{KNN}^T(x), Q_{KNN}^T(x))^T \\ P_{ELM}(x) &= (P_{ELM}^1(x), P_{ELM}^0(x)) \\ Q_{ELM}(x) &= (Q_{ELM}^1(x), Q_{ELM}^0(x)) \\ &\dots\dots \\ P_{KNN}(x) &= (P_{KNN}^1(x), P_{KNN}^0(x)) \\ Q_{KNN}(x) &= (Q_{KNN}^1(x), Q_{KNN}^0(x)) \end{aligned} \right. \quad (6)$$

where $P_{ELM}(x)$ is the probability of the ELM output belonging to the changed class, $Q_{ELM}(x)$ is the probability of the output belonging to the unchanged class, and $Q_{SG}(x)$ is the weighted average of the eight neighboring pixels belonging to the changed class. The other variables are obtained in a similar way.

2. Construction of the level-1 meta-classifier

After generating feature data, the MRLR is used as level-1 meta-classifier to predict all the results of the base classifiers to obtain the initial change map according to (6). The optimal parameters of the level-1 classifier are obtained by using the remaining sub-training set and the two parameter optimization algorithms.

Object-based Constrained Change Detection

3. Segmentation of the stacked image

Multi-resolution segmentation is a commonly used segmentation algorithm, which is a bottom-up approach that combines adjacent pixels or small segmentation objects to ensure a minimum average heterogeneity of different objects and a maximum homogeneity of internal pixels.

4. Constraint of segmentation map

In order to make full use of spatial information in high-resolution remote sensing images, we utilize spatial information to eliminate some of the ‘‘salt-and-pepper’’ noise and reduce the omission and commission ratio. The over-segmentation of image will reduce the accuracy of object-oriented change detection method. However, the smaller scale segmentation object can reduce the omission ratio when we use the segmentation object to constrain the pixel-wise map. After obtaining the pixel-wise change map using SG ensemble system and the segmentation map segmented by multi-resolution, the ratio of change pixels in each object in segmentation map is calculated. When the ratio is less than the threshold r , the pixels in the object are determined as unchanged pixels. The flowchart of change detection based on the SG hybrid ensemble system and segmentation map is shown in Figure 2.

Experiments and Analysis

Two multi-temporal and high-resolution remote sensing datasets collected by the ZY-3 satellite with spatial resolution of 5.8 meters are used in the experiments. ZY-3 is a Chinese high-resolution imaging Earth observation satellite, launched in January 2012. Both datasets have 450×450 pixels. The first dataset (Figure 3) which is covered part of Yunlong District consists of two high-resolution images of an area of Xuzhou (Jiangsu Province, China), taken on 05 November 2012 and 04 November 2013, respectively. The two high-resolution images (Figure 4) of the second dataset, covered part of Tongshan District also located in Xuzhou, Jiangsu province, China, were acquired in November, 2013 and October, 2014, respectively. Image registration and radiometric correction are important

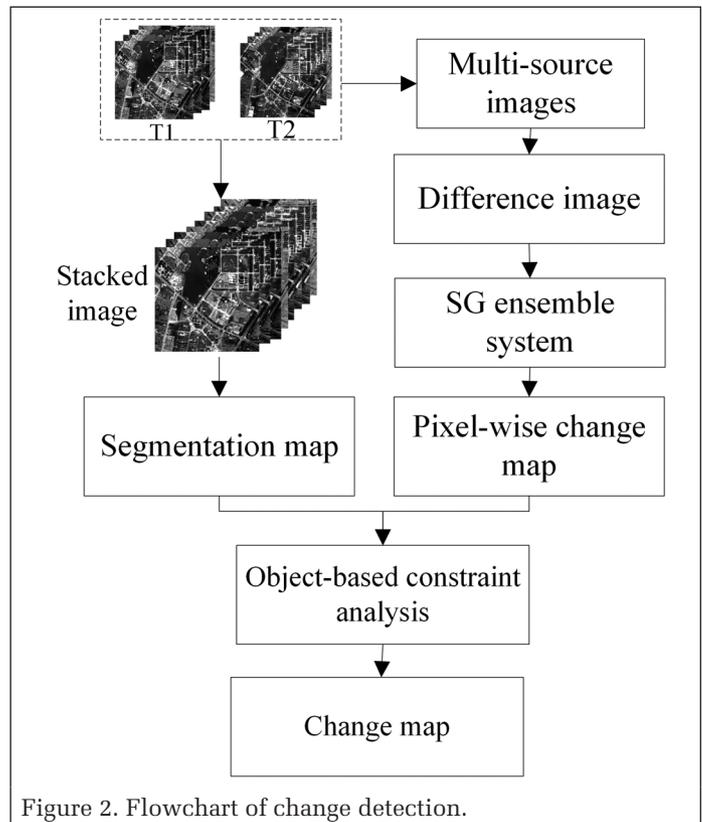


Figure 2. Flowchart of change detection.

preprocessing steps before generating the difference maps. The root mean square error of registration is less than 0.5 pixels. The relatively radiometric correction is performed by linear regression analysis (Lo and Yang, 2000).

In order to analyze the effectiveness of the proposed method, we compared our proposed method with pixel-wise change detection methods (PWCM) (i.e., ELM, KNN, SVM, MLR) and object-based change detection methods (OBCD) (i.e., CVA-OB, ELM-OB, KNN-OB, SVM-OB). In addition, the heterogeneous ensemble algorithms (HEAM) of MV, D-S, and F_{int} were used to extract the changes in both datasets. In order to demonstrate the effectiveness of the MRLR as meta-classifier, the SG-SVM and SG-MLR were also implemented. We also compared and analyzed the experimental results of the MRLR parameters based on the NNLS and FOA. Except SVM, KNN, MLR, CVA-OB, KNN-OB, and SVM-OB, all the other methods were repeated 10 times. The accuracy of these methods was the average of 10 experimental results. The overall accuracy (OA), Commission ratio and Omission ratio were used to evaluate the accuracy of change detection by comparing the detection results with the ground reference map, which are defined as

$$\begin{aligned} OA &= \frac{(N_{11} + N_{00})}{(N_{11} + N_{00} + N_{01} + N_{10})} \\ Commission\ ratio &= \frac{N_{01}}{(N_{01} + N_{11})} \\ Omission\ ratio &= \frac{N_{10}}{(N_{10} + N_{00})} \end{aligned} \quad (7)$$

where N_{11} and N_{00} are the numbers of changed pixels and unchanged pixels correctly detected, respectively, N_{10} is the number of missed changed pixels, and N_{01} is the number of



(a)



(b)

Figure 3. True color images of the first dataset.

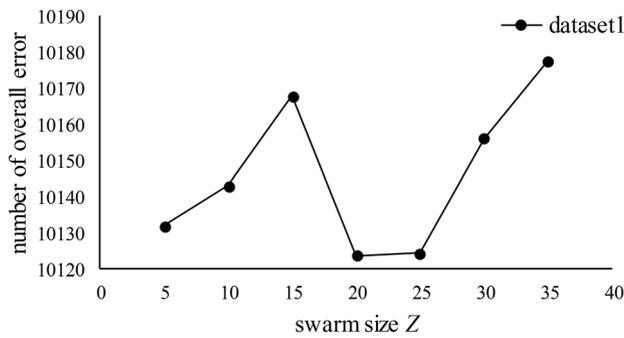


(a)

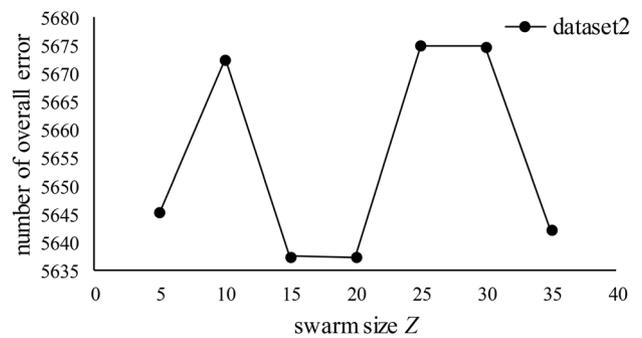


(b)

Figure 4. True color images of the second dataset.



(a)



(b)

Figure 5. The number of pixels wrongly detected by SG with different swarm sizes.

unchanged pixels in the ground reference but detected as changed in the change map.

We recorded the computing time of different methods to evaluate the computational cost. In this study, all the experiments are implemented based on Python 2.7 and ENVI 5.1.

The change detection method based on the SG hybrid ensemble system requires fewer parameters, but the key parameter is the swarm size (Z) in the FOA, which was defined in a set of comparative experiments. Another parameter is the

number of iterations (num), and it was set to 100 according to the convergence of the algorithm. In the two experiments, we set Z in the range of 5 to 35, and obtained the number of pixels wrongly detected by SG with different swarm sizes. The experimental results are shown in Figure 5. It can be observed that, the number of wrongly detected pixels fluctuates between 10,120 and 10,180 in experiment 1 and between 5,635 and 5,680 in experiment 2 when the swarm size is between 5 and 35, and the gap between them is small, which indicates



(a)



(b)

Figure 6. Overlay charts of true color images and boundaries of the segmented regions.

the robustness of the FOA. Therefore, we set $Z=20$ for the two datasets in the following experiments.

As previously mentioned, the SG ensemble system is constructed to obtain the pixel-wise change detection result, and then the change result is constrained by segmentation information to produce the final change map. The OBCD were used for comparison with the proposed method and the objects information was used to constrain the change information, so objects information was obtained with the method previously described. According to the distribution of the two regions in the experimental area, the segmentation size was set to 20, the shape size was set to 0.4, and the compactness size was 0.6 for the first dataset and 50-0.4-0.6 for the second dataset. The threshold r was experimentally set to 0.15 for the two datasets. The overlay charts of true color images and boundaries of the segmented regions for the two datasets are shown in Figure 6.

Table 1. Accuracy of the different methods for the first dataset.

| Method | OA | Commission ratio | Omission ratio | Time(s) | |
|--------|-----------------|------------------|----------------|---------------|--------------|
| PWCM | ELM | 0.9351 | 0.3983 | 0.0309 | 5.08 |
| | KNN | 0.9292 | 0.4423 | 0.0239 | 1957.15 |
| | SVM | 0.9332 | 0.4204 | 0.0245 | 11.27 |
| | MLR | 0.9317 | 0.4095 | 0.0363 | 0.305 |
| OBCD | CVA-OB | 0.9425 | 0.3540 | 0.0283 | 8.32 |
| | ELM-OB | 0.9470 | 0.2339 | 0.0415 | 5.21 |
| | KNN-OB | 0.9479 | 0.2030 | 0.0436 | 1818.92 |
| | SVM-OB | 0.9533 | 0.1771 | 0.0388 | 10.29 |
| HEAM | MV | 0.9332 | 0.4195 | 0.0252 | 1957.94 |
| | D-S | 0.9389 | 0.3441 | 0.0388 | 1966.74 |
| | F_int | 0.9328 | 0.4022 | 0.0361 | 1979.52 |
| SG | SG-LS | 0.9403 | 0.3644 | 0.0300 | 24.92 |
| | SG-ff | 0.9507 | 0.2462 | 0.0346 | 24.95 |
| | SG-SVM | 0.9422 | 0.3576 | 0.0277 | 25.31 |
| | SG-MLR | 0.9431 | 0.3461 | 0.0291 | 24.53 |
| | SG-LS-OB | 0.9547 | 0.2244 | 0.0319 | 27.62 |
| | SG-ff-OB | 0.9556 | 0.1922 | 0.0344 | 27.65 |

Table 2. Accuracy of the different methods for the second dataset.

| Method | OA | Commission ratio | Omission ratio | Time(s) | |
|--------|-----------------|------------------|----------------|---------------|--------------|
| PWCM | ELM | 0.9581 | 0.3153 | 0.0175 | 5.04 |
| | KNN | 0.9522 | 0.3884 | 0.0068 | 1925.32 |
| | SVM | 0.9614 | 0.3232 | 0.0090 | 25.58 |
| | MLR | 0.9224 | 0.5271 | 0.0229 | 0.39 |
| OBCD | CVA-OB | 0.9611 | 0.2749 | 0.0203 | 13.80 |
| | ELM-OB | 0.9628 | 0.2213 | 0.0250 | 8.11 |
| | KNN-OB | 0.9660 | 0.1665 | 0.0261 | 2144.15 |
| | SVM-OB | 0.9652 | 0.2371 | 0.0195 | 14.09 |
| HEAM | MV | 0.9610 | 0.3271 | 0.0088 | 1925.42 |
| | D-S | 0.9657 | 0.2446 | 0.0177 | 1933.75 |
| | F_int | 0.9517 | 0.3743 | 0.0146 | 1945.61 |
| SG | SG-LS | 0.9625 | 0.3108 | 0.0101 | 25.77 |
| | SG-ff | 0.9720 | 0.1852 | 0.0162 | 25.84 |
| | SG-SVM | 0.9621 | 0.3205 | 0.0084 | 26.16 |
| | SG-MLR | 0.9650 | 0.2950 | 0.0094 | 25.37 |
| | SG-LS-OB | 0.9748 | 0.1799 | 0.0131 | 28.67 |
| | SG-ff-OB | 0.9762 | 0.1195 | 0.0174 | 28.84 |

The accuracy of change detection for the two datasets are shown in Tables 1 and 2, respectively. SG-LS is the SG ensemble system with the parameters of MRLR obtained by the use of NNLS. SG-ff is the SG ensemble system with the parameters of MRLR obtained by the use of FOA. SG-SVM and SG-MLR are the SG ensemble system with the SVM and MLR as the meta-classifier, respectively. SG-LS-OB is the SG-LS algorithm refined by object information, similarly, SG-ff-OB is the refined SG-ff. All these methods, i.e., MV, D-S, F_int, are heterogeneous ensemble algorithms and consist of the same base classifiers. As can be seen from the above results, the accuracy of SG-ff-OB is the highest among the methods, with the overall accuracy of 0.9556 and 0.9762, lowest commission ratio of 0.1922 and 0.1195, because the predicted results of base classifiers and the spatial neighborhood information were utilized to

construct the feature data at level-1, and the constraint of object information after the pixel-wise change map reduced the commission ratio. It can be seen from the tables that the accuracy of OBCD is significantly higher than that of the PWCM because the PWCM is based on a single pixel, and the spectral information in a high-spatial-resolution image is insufficient resulting in more commission pixels. The detection results of different heterogeneous ensemble systems are different and the accuracy of the three heterogeneous ensemble algorithms are not all higher than that of the individual base classifiers, which demonstrates that the selection of the ensemble algorithm has impact on the performance of the multi-classifier ensemble system and that not all the ensemble methods can

improve the accuracy. Comparing the methods based on the SG hybrid ensemble system, we can observe that the overall accuracy of SG-ff are higher than that of SG-LS, SG-SVM and SG-MLR due to better optimization capability of FOA. The commission ratio of SG with the parameters of MRLR based on FOA and NNLS is improved when we utilize the object information to impose constraints on the pixel-wise change detection maps. Moreover, it can be observed that the SG ensemble system costs less time than other ensemble systems with higher accuracy. This is because in base classifiers we utilize the advantages of ELM to classify a large part of pixels and combine the series and parallel framework to classify other pixels.

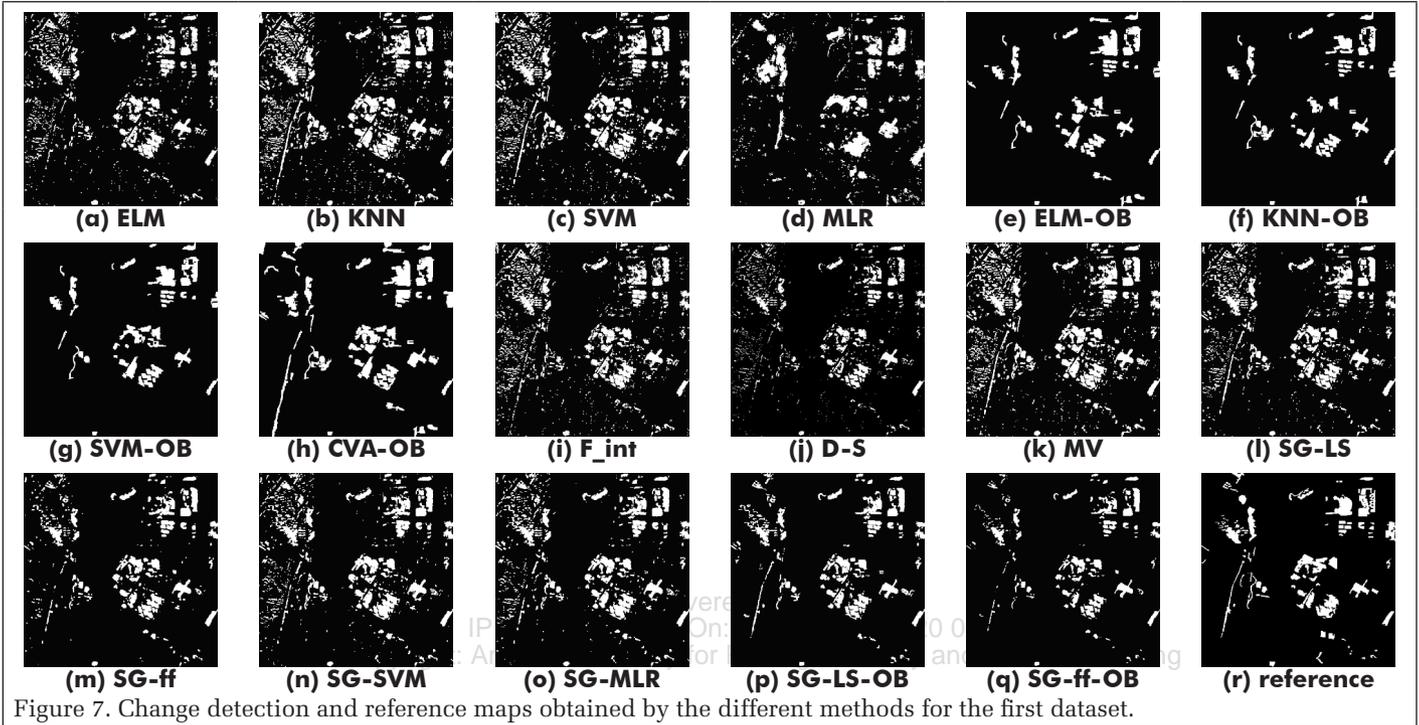


Figure 7. Change detection and reference maps obtained by the different methods for the first dataset.

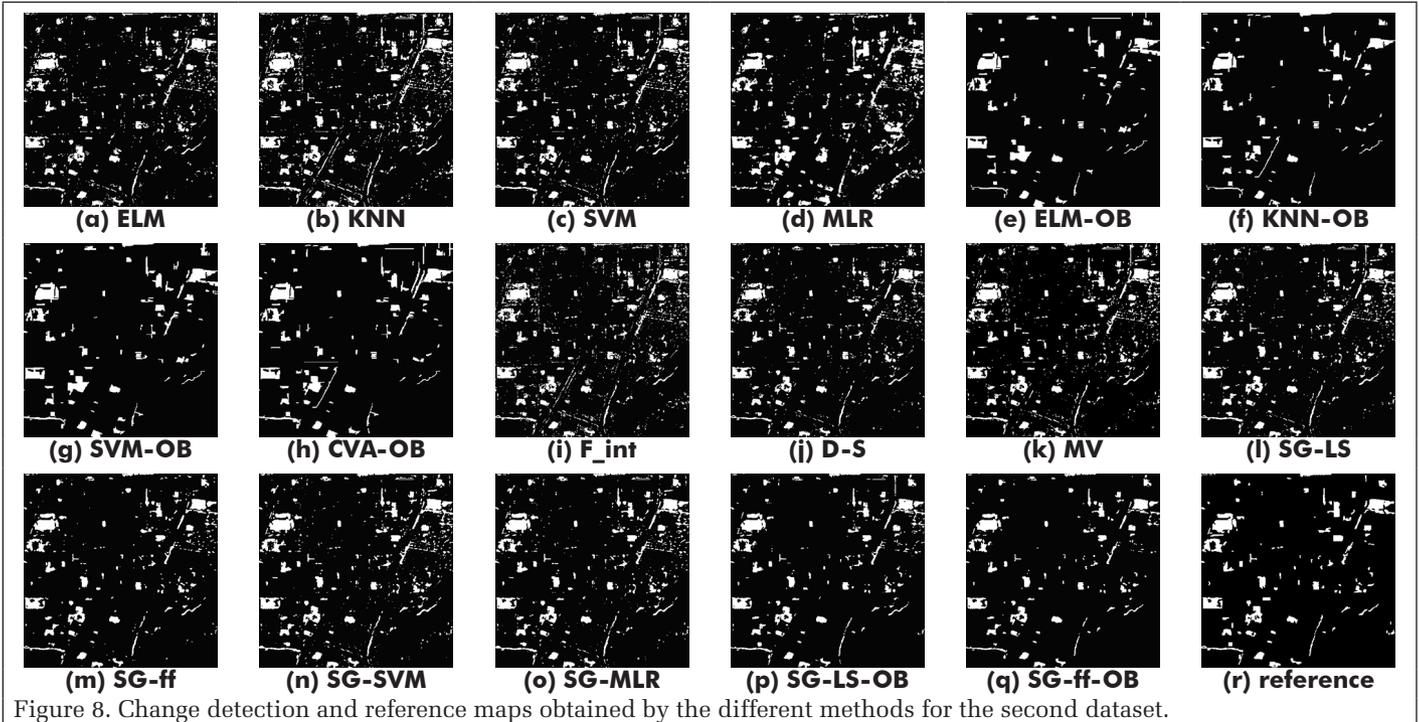


Figure 8. Change detection and reference maps obtained by the different methods for the second dataset.

Figure 7 and Figure 8 show the change maps and reference maps for the two datasets. It can be seen from the results of the change maps that the changed regions in the first dataset mainly includes the changes of vegetation and bare land to roads and buildings. Compared with the reference change map in Figure 7(r), the change detection results of SG-ff-OB shown in Figure 7(q) is more consistent with the reference change map. After utilizing the segmented object information constraint, some of the salt-and-pepper noise in the change detection result based on the SG hybrid ensemble system is suppressed. As shown in Figure 7(a)- to (d),(i) to (k), it can be observed that there are many false-alarm pixels in the left and lower part of the change maps obtained by PWCM and HEAM, whereas the pixels belonging to false detection class are significantly less in Figures 7(p) and (q) obtained by the proposed methods. That is because the areas in left part are mainly dense residential buildings where the frequency of gray scale change is large. However, the spectral information of high-spatial-resolution image is insufficient and the PWCM is based on pixel units, thus generating more noise and commission pixels. From the results of Figures 7 (a), (b) and (c), we can see that there are less false-alarm pixels in the left part but some missed pixels in the middle part of the image obtained by ELM, which is complementary with the other two methods. As can be seen from the Figures 7(e), (f), (g), and (h), some changed regions of the maps obtained by OBCDs are mis-identified as unchanged regions in the change detection maps due to the dependence of the segmentation scale of the OBCD. The SG methods yield less false-alarm pixels in change detection map, due to the consideration of spatial information when constructing the input data at level-1. The changed regions in the second dataset are mainly including the increase of buildings and roads and the decrease of vegetation and bare land. Similarly, the change maps in Figures 8(p) and (q) obtained by the SG ensemble system methods, SG-LS-OB and SG-ff-OB, are more consistent with the reference map in Figure 8 (r). As shown in Figures 8 (a), (b), (c) and (d), the results of the PWCM and Figures 8(h), (i) and (j), the results of the three HEAM, contain salt-and-pepper noise in all the maps, whereas the pixels belonging to false detection class are significantly less in the map Figure 8(q). Meanwhile, as compared with the reference map in Figure 8(r), the salt and pepper noise can be effectively suppressed in Figures 8(e), (f), (g), and (h) obtained by the OBCD, but due to the influence of the segmentation error, some changed areas are identified as unchanged regions in the lower part of the change maps, whereas the regions can be detected accurately by the proposed methods.

In order to show the generalization capability of the proposed method, an additional dataset collected by the Quick-Bird satellite is also used. The images were collected by four multispectral bands and a panchromatic band and the fusion resolution is 0.6 meters. The multi-temporal dataset contains 800 × 700 pixels. The true color images are shown in Figure 9.

For this dataset, we used the same parameter of as in the ZY-3 datasets. The spectral and texture features were utilized as the input datasets. The segmentation size was set as 40-0.5-0.5. The accuracy of different methods for the additional dataset is shown in Table 3. As can be seen, the accuracy of SG-ff-OB is also the highest among all the methods. We can see the same results as in the ZY-3 datasets, which further demonstrated the generalization of the proposed method.

Conclusions

A simple multi-classifier hybrid ensemble system for pixel-wise change detection and an object-based approach to change detection using constrained analysis were proposed in this paper. The SG hybrid ensemble system uses ELM, KNN,

Table 3. Accuracy of the different methods for the additional dataset.

| | Method | OA | Commission ratio | Omission ratio |
|------|-----------------|---------------|------------------|----------------|
| PWCM | ELM | 0.8745 | 0.3638 | 0.0405 |
| | KNN | 0.8348 | 0.4510 | 0.0269 |
| | SVM | 0.8582 | 0.4022 | 0.0387 |
| OBCD | ELM-OB | 0.8742 | 0.3713 | 0.0316 |
| | KNN-OB | 0.8121 | 0.4868 | 0.0261 |
| | SVM-OB | 0.8679 | 0.3833 | 0.0349 |
| HEAM | MV | 0.8606 | 0.4004 | 0.0328 |
| | D-S | 0.8976 | 0.2752 | 0.0563 |
| | F_int | 0.8326 | 0.4518 | 0.0699 |
| | SG-ff | 0.8829 | 0.3447 | 0.0379 |
| | SG-ff-OB | 0.9219 | 0.2087 | 0.0445 |

and SVM as the level-0 base classifiers, and MRLR as the level-1 meta-classifier by using the spectral, statistical texture and structure texture (morphological profile) features as input. In addition to the prediction results of base classifiers, the weighted average of eight neighboring pixels were also used as the input of meta-classifier. The FOA, one of the recently developed global optimization algorithms, was adopted to estimate the model parameters in MRLR. In order to utilize the advantages of high-resolution remote sensing images and decrease the direct influence of segmentation, constrained analysis on segmented objects was implemented to integrate segmentation map and pixel-wise change map for final result. The experimental results demonstrate that the proposed method performs better with lower computational cost.

Acknowledgments

This research is supported in part by Natural Science Foundation of China (No. 41471356), the Xuzhou Scientific Funds (KC16SS092), and Priority Academic Program Development of Jiangsu Higher Education Institutions.

References

- Azzouzi, S.A., A. Vidal, and H. A. Bentounes, 2015. Enhancement of the double flexible pace search threshold determination for change vector analysis, *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-7/W3(7):599-603.
- Benedek, C., M. Shadaydeh, Z. Kato, T. Szirányi, and J. Zerubia, 2015. Multilayer Markov Random Field models for change detection in optical remote sensing images, *ISPRS Journal of Photogrammetry & Remote Sensing*, 107:22-37.
- Bruzzzone, L., and D.F. Prieto, 2000. A minimum-cost thresholding technique for unsupervised change detection, *International Journal of Remote Sensing*, 21(18):3539-3544.
- Chang, N.B., M. Han, W. Yao, L.C. Chen and S.G. Xu, 2010. Change detection of land use and land cover in an urban region with SPOT-5 images and partial Lanczos extreme learning machine, *Journal of Applied Remote Sensing*, 4(1):2816-2832
- Chellasamy, M., T.P.A. Ferré, M. Humlekrog Greve, R. Larsen, and U. Chinnasamy, 2014, An ensemble classification approach for improved land use/cover change detection, *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-8(1):695-701.
- Chen, J., P. Gong, C. He, R. Pu, and P. Shi, 2003. Land-use/land-cover change detection using improved change-vector analysis, *Photogrammetric Engineering & Remote Sensing*, 69(4):369-379.
- Du, P., J. Xia, W. Zhang, K. Tan, Y. Liu, and S. Liu, 2012. Multiple classifier system for remote sensing image classification: A review, *Sensors*, 12(4): 4764-4792.

- Guo, G., H. Wang, D. Bell, Y. Bi, and K. Greer, 2003. KNN model-based approach in classification, *Lecture Notes in Computer Science*, 2888:986-996.
- Hao, M., W. Shi, K. Deng, H. Zhang, and P. He, 2016. An object-based change detection approach using uncertainty analysis for VHR Images, *Journal of Sensors*, 2016:1-17.
- Hatami, N., and R. Ebrahimpour, 2007. Combining multiple classifiers: Diversify with boosting and combining by stacking, *International Journal of Computer Science*, Vol. 7, pp. 127-131.
- He, L., and I. Laptev, 2009. Robust change detection in dense urban areas via SVM classifier, *Urban Remote Sensing Event*, pp.1-5
- Hou, B., Q. Liu, and Y. Wang, 2015. Object-based feature extraction and semi-supervised classification for urban change detection using high-resolution remote sensing images, *Proceedings of the Geoscience and Remote Sensing Symposium*, pp.1674-1677.
- Huang, G.B., Q.Y. Zhu, and C.K. Siew, 2005. Extreme learning machine: A new learning scheme of feedforward neural networks, *Proceedings of the IEEE International Joint Conference on Neural Networks*, 2:985-990.
- Huang, G.B., Q.Y. Zhu, and C.K. Siew, 2006. Extreme learning machine: Theory and applications, *Neurocomputing*, 70(1):489-501.
- Huang, X., D. Wen, J. Li, and R. Qin, 2017. Multi-level monitoring of subtle urban changes for the megacities of China using high-resolution multi-view satellite imagery, *Remote Sensing of Environment*, 56: 56-75.
- Huang, X., L. Zhang, and T. Zhu, 2013. Building change detection from multitemporal high-resolution remotely sensed images based on a morphological building index, *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 7(1):105-115.
- Iscan, H., and M. Gunduz, 2015. A survey on fruit fly optimization algorithm, *Proceedings of the International Conference on Signal-Image Technology & Internet-Based Systems*, pp. 1674-1677.
- Jawak, S.D., S.N. Panditrao, and A.J. Luis, 2014. Enhanced urban landcover classification for operational change detection study using very high resolution remote sensing data, *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-8(1):773-779.
- Khodadadzadeh, M., J. Li, A. Plaza, and J.M. Bioucas-Dias, 2014. A subspace-based multinomial logistic regression for hyperspectral image classification, *IEEE Geoscience & Remote Sensing Letters*, 11(12):2105-2109.
- Li, J., J.M. Bioucas-Dias, and A. Plaza, 2012. Spectral-spatial hyperspectral image segmentation using subspace multinomial logistic regression and Markov Random Fields, *IEEE Transactions on Geoscience & Remote Sensing*, 50(3):809-823.
- Li, Q., X. Huang, D. Wen, and H. Liu, 2017. Integrating multiple textural features for remote sensing image change detection, *Photogrammetric Engineering & Remote Sensing*, 83(2):109-121.
- Li, Y., and A. Ngom, 2013. Classification approach based on non-negative least squares, *Neurocomputing*, 118(11):41-57.
- Lin, S.C., Y.C.I. Chang, and W.N. Yang, 2009. Meta-learning for imbalanced data and classification ensemble in binary classification, *Neurocomputing*, 73(1-3):484-494.
- Lo, C.P., and X. Yang, 2000. Relative radiometric normalization performance for change detection from multi-date satellite images, *Photogrammetric Engineering & Remote Sensing*, 66(8):967-980.
- Lv, P., Y. Zhong, J. Zhao, and L. Zhang, 2015. Unsupervised change detection based on conditional random fields and texture feature for high resolution remote sensing imagery, *Proceedings of the IEEE China Summit and International Conference on Signal and Information Processing*, pp. 1081-1085
- Malmir, M., M.M. Zarkesh, S.M. Monavari, S.A. Jozi, and E. Sharifi, 2015. Urban development change detection based on multi-temporal satellite images as a fast tracking approach-A case study of Ahwaz County, southwestern Iran, *Environmental Monitoring & Assessment*, 187(3):4295.
- Moser, G., E. Angiati, and S.B. Serpico, 2011. Multiscale unsupervised change detection on optical images by Markov Random Fields and wavelets, *IEEE Geoscience & Remote Sensing Letters*, 8(4):725-729.
- Nemmour, H., and Y. Chibani, 2006. Multiple support vector machines for land cover change detection: An application for mapping urban extensions, *ISPRS Journal of Photogrammetry & Remote Sensing*, 61(2):125-133.
- Nemmour, H., and Y. Chibani, 2013. Change detector combination in remotely sensed images using fuzzy integral, *International Journal of Signal Processing*, (4):175.
- Pan, W.T., 2012. A new fruit fly optimization algorithm: Taking the financial distress model as an example, *Knowledge-Based Systems*, 26(2): 69-74.
- Peng, D., and Y. Zhang, 2017. Object-based change detection from satellite imagery by segmentation optimization and multi-features fusion, *International Journal of Remote Sensing*, 38(13):3886-3905.
- Rodriguez, J.J., L.I. Kuncheva, and C.J. Alonso, 2006. Rotation forest: A new classifier ensemble method, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 28(10):1619-1630.
- Rojarath, A., W. Songpan, and C. Pong-Inwong, 2017. Improved ensemble learning for classification techniques based on majority voting, *Proceedings of the IEEE International Conference on Software Engineering and Service Science*, pp. 107-110.
- Roy, M., S. Ghosh, and A. Ghosh, 2012. A semi-supervised change detection for remotely sensed images using ensemble classifier, *Proceedings of the International Conference on Intelligent Human Computer Interaction*, 3(1):1-5.
- Roy, M., S. Ghosh, and A. Ghosh, 2014. A novel approach for change detection of remotely sensed images using semi-supervised multiple classifier system, *Information Sciences*, 269(8):35-47.
- Seewald, A.K., 2002. Exploring the parameter state space of stacking, *Proceedings of the IEEE International Conference on Data Mining*, pp. 685.
- Sesmero, M.P., A.I. Ledezma, and A. Sanchis, 2015. Generating ensembles of heterogeneous classifiers using Stacked Generalization, *Wiley Interdisciplinary Reviews Data Mining & Knowledge Discovery*, 5(1):21-34.
- Skurichina, M., and R.P.W. Duin, 2010. The random subspace method for linear classifiers, *International Journal on Pattern Analysis and Applications*, 5(2):121-135.
- Skurichina, M., L. Kuncheva, and R.P.W. Duin, 2002. Bagging and boosting for the nearest mean classifier: Effects of Sample Size on Diversity and Accuracy, *Proceedings of the International Workshop on Multiple Classifier Systems*, 2364: 62-71.
- Ting, K.M., and I.H. Witten, 2002. Issues in stacked generalization, *Journal of Artificial Intelligence Research*, 10(1):271-289.
- Vapnik, N. and N. Vladimir, 1997. The nature of statistical learning theory, *IEEE Transactions on Neural Networks*, 38(4):409-409.
- Volpi, M., D. Tuia, F. Bovolo, M. Kanevski, and L. Bruzzone, 2013. Supervised change detection in VHR images using contextual information and support vector machines, *International Journal of Applied Earth Observation & Geoinformation*, 20(2):77-85.
- Wen, D., X. Huang, L. Zhang, and J.A. Benediktsson, 2015. A novel automatic change detection method for urban high-resolution remotely sensed imagery based on multiindex scene representation, *IEEE Transactions on Geoscience & Remote Sensing*, 54(1):609-625.
- Woo, D.-M., and V.D. Do, 2015. Post-classification change detection of high resolution satellite images using adaboost classifier, *Information Technology and Computer Science*, pp. 34-38.
- Xia, J., M.D. Mura, J. Chanussot, P. Du, and X. He, 2015. Random subspace ensembles for hyperspectral image classification with extended morphological attribute profiles, *IEEE Transactions on Geoscience & Remote Sensing*, 53(9):4768-4786.