

# Binary Change Guided Hyperspectral Multiclass Change Detection

Meiqi Hu, *Graduate Student Member, IEEE*, Chen Wu<sup>ID</sup>, *Member, IEEE*,  
 Bo Du<sup>ID</sup>, *Senior Member, IEEE*, and Liangpei Zhang<sup>ID</sup>, *Fellow, IEEE*

**Abstract**—Characterized by tremendous spectral information, hyperspectral image is able to detect subtle changes and discriminate various change classes for change detection. The recent research works dominated by hyperspectral binary change detection, however, cannot provide fine change classes information. And most methods incorporating spectral unmixing for hyperspectral multiclass change detection (HMCD), yet suffer from the neglection of temporal correlation and error accumulation. In this study, we proposed an unsupervised Binary Change Guided hyperspectral multiclass change detection Network (BCG-Net) for HMCD, which aims at boosting the multiclass change detection result and unmixing result with the mature binary change detection approaches. In BCG-Net, a novel partial-siamese united-unmixing module is designed for multi-temporal spectral unmixing, and a groundbreaking temporal correlation constraint directed by the pseudo-labels of binary change detection result is developed to guide the unmixing process from the perspective of change detection, encouraging the abundance of the unchanged pixels more coherent and that of the changed pixels more accurate. Moreover, an innovative binary change detection rule is put forward to deal with the problem that traditional rule is susceptible to numerical values. The iterative optimization of the spectral unmixing process and the change detection process is proposed to eliminate the accumulated errors and bias from unmixing result to change detection result. The experimental results demonstrate that our proposed BCG-Net could achieve comparative or even outstanding performance of multiclass change detection among the state-of-the-art approaches and gain better spectral unmixing results at the same time.

**Index Terms**—Hyperspectral multiclass change detection, multi-temporal unmixing, temporal correlation constraint, unsupervised learning, deep neural network.

## I. INTRODUCTION

**K**OWN for the high spectral resolution, hyperspectral image (HSI) provides tremendous spectral information for object discrimination [1], [2], [3] and has made advanced

Manuscript received 10 December 2021; revised 19 September 2022 and 14 November 2022; accepted 10 December 2022. Date of publication 6 January 2023; date of current version 13 January 2023. This work was supported in part by the National Key Research and Development Program of China under Grant 2022YFB3903302, in part by the National Natural Science Foundation of China under Grant T2120014 and Grant 61971317, in part by the Natural Science Foundation of Hubei Province under Grant 2020CFB594, and in part by the Fundamental Research Funds for the Central Universities. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Leyuan Fang. (*Corresponding author: Chen Wu*)

Meiqi Hu, Chen Wu, and Liangpei Zhang are with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China (e-mail: meiqi.hu@whu.edu.cn; chen.wu@whu.edu.cn; zlp62@whu.edu.cn).

Bo Du is with the School of Computer Science, Wuhan University, Wuhan 430072, China (e-mail: gunspace@163.com).

Digital Object Identifier 10.1109/TIP.2022.3233187

development in hyperspectral classification [4], [5], [6], target detection [7], [8], anomaly detection [9], [10], denoising [11], [12], spectral unmixing (SU) [13], [14], super-resolution [15], [16], image fusion [17], [18], and change detection [19], [20]. Among various researches of HSIs, change detection has been one of the hottest remote sensing application topics in the past decades, aiming to detect the change information of two remote sensing images acquired in the same geographical location on different times [21], [22], and has been widely applied in land cover and land usage change detection [23], [24], disaster emergency and assessment [25], [26], dynamic urban monitoring and management [27], [28], etc.

In general, hyperspectral change detection is categorized as hyperspectral binary change detection (HBCD) [29] and hyperspectral multiclass change detection (HMCD) [30]. HMCD has long been a challenging research work. Compared with HBCD, which requires to detect the changed area, HMCD not only needs to detect the change area and the unchanged area, but also demands to identify the number of changes and distinguish different kinds of change transformation associated with variant land-cover transitions [31], [32]. By contrast, HMCD offers more detailed change information.

The most straightforward solution to HMCD, named post classification comparison (PCC) [33], is to classify the HSIs separately and then compare the multi-temporal classification maps to get the multiclass change detection (MCD) result. However, abundant training samples are desired for accurate classification. Then Du et al. [34] proposed an unsupervised or semi-supervised HMCD method incorporating the SU to detect multiclass changes, called post unmixing comparison (PUC) [35]. SU is the process of decomposing the spectral signature of a mixed pixel into a set of pure materials (called endmember) and their corresponding proportions (called abundance) [36], [37], and is a powerful tool to the mixed pixel problem due to coarse spatial resolution of HSI. Generally, the independent unmixing is conducted on each image and corresponding abundance vectors of each pixel are then compared to acquire binary and multiclass changes results for PUC. Limited by the separate unmixing, the abundance maps of multi-temporal images lack the temporal correlation and tend to yield different unmixing results on the same area. But the support of sub-pixel information provided by the abundance still sparks much interest in the utilization of unmixing for HMCD. Liu et al. proposed a novel multitemporal spectral unmixing (MSU) [38] approach which extracted the multi-temporal endmembers (MT-EMs) from the stacked

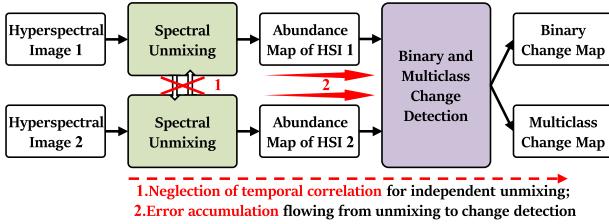


Fig. 1. Traditional framework of incorporating spectral unmixing for hyperspectral multiclass change detection. Limited by the independent unmixing, the abundance maps of multi-temporal images lack the temporal correlation. And the tiny numerical abundance variance and either unmixing errors accumulate and are magnified when the binary and multiclass change detection results are directly computed from the abundance maps.

HSIs, and introduced a change analysis strategy to distinguish change MT-EMs from no-change MT-EMs. Jafarzadeh and Hasanlou put forward a new rule of detecting binary and multiclass changes by identifying the pure changed pixels and mixed changed pixels hierarchically from the abundance vectors pairs [39]. Guo and Zhong [40] represented an original temporal spectral unmixing model by using the current temporal HSI and the endmembers of the previous to obtain the endmember of the current temporal HSI, then the binary and multiclass changes were detected from the abundance maps. These methods, however, have a challenging issue of detecting accurate change maps without error accumulation when the binary and multiclass change detection results are directly computed from the abundance maps (presented in Fig. 1). The tiny numerical abundance variance and either unmixing errors of the multi-temporal abundance maps might be magnified by the abundance comparison. Consequently, the accumulated error significantly impacts the unchanged pixel-matching and change classes transformation, producing noise and spots of false positive changes and wrong shifted multiple changes.

We found that the binary change detection result computed by SU result reflects the performance of unmixing to some extent and contains the error accumulation and bias between unmixing and change detection. Can we start with improving the binary change detection result to further accomplish collaborative optimization between unmixing and change detection? In this paper, we propose a binary change guided multiclass change detection approach to solve the problem of error accumulation, as shown in Fig. 2. We design a temporal correlation constraint directed by the pseudo-labels of BCD result to guide the unmixing process from the perspective of change detection. Actually, the BCD result of multi-temporal images is a token of temporal correlation. That is, if the pixel at the same location of two phases has not changed, the abundance of the two pixels are likely to be similar; if the pixel has changed, the abundance of them are likely to be quite different. Specifically, given the reliable pseudo-labels from BCD result, the gaps between the BCD result calculated from the SU results and that of pseudo-labels indicate the accumulated error and bias passing from the unmixing process to change detection. Then the gaps are minimized through back propagation, which acts over the unmixing process to boost the abundance of the unchanged pixels more similar and that of the changed pixels more accurate. Driven by the iterative optimization of unmixing result and BCD results, the error accumulation from the unmixing to change detection is eliminated, and high-accuracy binary and multiclass change detection results are accomplished. Meanwhile, more accurate multi-temporal SU results can be obtained.

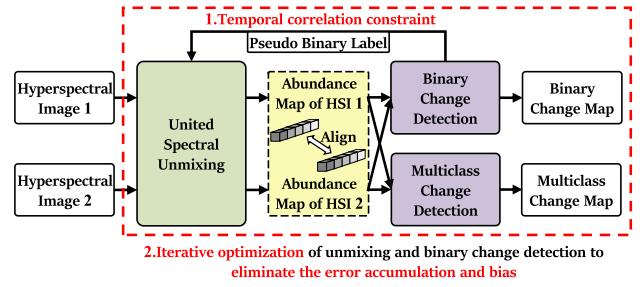


Fig. 2. Proposed binary change guided multiclass change detection framework with temporal correlation constraint. A temporal correlation constraint directed by the pseudo-labels of binary change detection result is tailored to guide the unmixing process from the perspective of change detection, where the abundance maps of the unchanged pixels are boosted to be more similar and that of the changed pixels more accurate. And the unmixing and change detection are iteratively trained together by alternate optimization to clear the impact of error accumulation and bias.

Inspired by this idea, an unsupervised multi-task learning neural network, Binary Change Guided hyperspectral multi-class change detection Network (BCG-Net), is put forward for HMCD. Firstly, we developed a novel partial-siamese united-unmixing module, where the scaled spectral attention block is designed to dig diverse spectral features, especially for the emphasis on the distinguished spectral characteristics of different ground objects. The sum-to-one constraint and non-negative constraint are integrated into the neural network perfectly, making it convenient to conduct unmixing in an unsupervised way. The bi-temporal HSIs are both fed into the network to acquire the abundance maps simultaneously. Secondly, the multi-temporal abundance vectors are input into the designed temporal correlation module to acquire the binary change information, testifying the performance of unmixing from the point of view of change detection. Then the unmixing and change detection are iteratively optimized to eliminate the bias and the accumulated error, yielding greater MCD result.

The main contributions are summarized as follows:

1) We propose an innovative constraint for spectral unmixing from the point view of change detection, integrating the temporal correlation into the spectral unmixing, where the abundance maps of the unchanged pixels are boosted to be more similar and that of the changed pixels more accurate.

2) To further effectively clear the impact of error accumulation and bias, we designed the alterative optimization between the unmixing and change detection, contributing to greater binary and multiclass change detection results.

3) Based on the proposed thought, we put forward an outstanding unsupervised multi-task learning network, named BCG-Net, where the designed novel united-unmixing module and a groundbreaking temporal correlation module are tailored and integrated for binary and multiclass change detection.

4) The experimental results show that our method could gain comparable or even outstanding performance among the state-of-the-art methods, demonstrating the effectiveness of BCG-Net. And the discussions about the effect of temporal correlation constraint indicate that the temporal correlation constraint really works on both the temporal unmixing result and the binary and multiclass change detection results.

The rest of the article is organized as follows. The detailed description of the proposed BCG-Net will be provided in Section II. And the data description and experiment results will be represented in Section III. Section IV will discuss the effect of temporal correlation constraint on SU and change

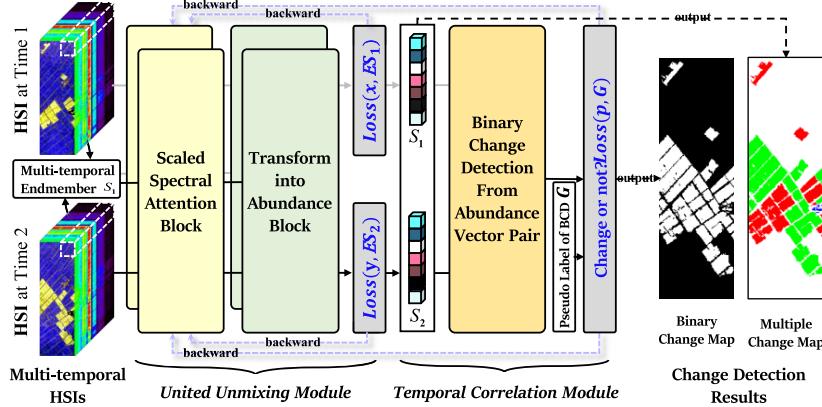


Fig. 3. Architecture of the proposed BCG-Net. The network consists of UU-Module and TC-Module. The TC-Module verifies the performance of the unmixing result in the view of change detection, placing a temporal correlation constraint on the UU-Module.

detection result. The conclusion of this paper will be given in Section V.

## II. METHODOLOGY

### A. General Framework

As shown in Fig. 3, the BCG-Net consists of two parts, namely, the united unmixing module (UU-Module) and temporal correlation module (TC-Module). A pair of patch blocks centered on a certain pixel of the bi-temporal images are first fed into the UU-Module, which is tailored to get the corresponding abundance vectors. The outputs of the UU-Module are then fed into the TC-Module, which is designed to acquire the binary change information of the pixel. And the result is then compared with the given binary pseudo-label. The united unmixing result is qualified if the binary change result is consistent with the given binary label. Otherwise the unmixing result would be optimized further and the binary change result would be further compared with the given label. During the back propagation, the TC-Module takes a role of temporal correlation constraint on the UU-Module in the view of change detection, where the unmixing process is optimized further upon boosting the coherence of the abundance vectors of those unchanged objects to decrease the false change alarms, and more accurate abundance vectors for those changed objects. Consequently, better binary change result is outputted from the network and the multiclass change result is produced by the optimized united unmixing result.

### B. United Unmixing Module

The UU-Module aims at extracting the corresponding abundance from bi-temporal HSIs by united spectral unmixing. The UU-Module composes of two blocks, where Scaled Spectral Attention Block (SSAB) is designed to extract the spatial and multi-scaled spectral features from the input and the Transform into Abundance Block (TAB) to convert the extracted features into abundance.

For a single HSI  $X$ , the details of the single-branch UU-Module are shown Fig. 4 (a). The hyperspectral patch  $X_{\text{patch}} \in \mathbb{R}^{m \times m \times C}$  centered on a pixel  $x \in \mathbb{R}^{C \times 1}$  is first processed by the  $3 \times 3 \times 1$  three-dimensional convolutional layer (3Dconv) [41] and  $3 \times 3 \times 3$  3Dconv separately, followed by two  $3 \times 3 \times 3$  3Dconvs respectively, where  $m$  is the patch size and  $C$  is the number of the spectral band. The 3Dconv is able to extract spatial and spectral features simultaneously. The smaller spectral kernel size is introduced to extract the band-by-band spectral features, while larger

one is beneficial to dig the spectral feature of contiguous spectral channels. Furthermore, the efficient channel attention (ECA) [42] block (shown in Fig. 4 (b)) is introduced to obtain the spectral attention after the two 3Dconvs, aiming at capturing the cross-channel interaction and stressing the discriminated spectral features. ECA block firstly adopts a global average pooling (GAP) to convert each channel of feature map into a representative statistical value. The one-dimensional convolutional layer designed in the ECA captures local cross-channel interaction by considering every channel and its  $k$  neighbors, where  $k = \lceil [\log_2(C)]/2 + 1 \rceil_{\text{odd}}$  according to [34] and  $|t|_{\text{odd}}$  refers to the nearest odd number of  $t$ . The weights of each channels are computed by the gate mechanism, where the sigmoid activation function  $\sigma$  converts the extracted channel features into a channel attention map ranging from 0 to 1. The bigger the weight is, the more attractive the spectral information of the input feature is. The weights are adaptively adjusted according to the inputted features and discriminative spectral characteristics are highlighted for various ground objects. The one-dimensional convolutional layer designed in the ECA facilitates faster and more efficient channel attention representation than fully connection layer of the Squeeze-and-Excitation Network [43] does without dimensionality reduction. The two feature extraction branches focus on different scaled features, which are beneficial to distinguish the landmark spectral features of different ground objects in SU. Another  $3 \times 3 \times 1$  3Dconv is adopted after that the pointwise concatenation of two branches. Additionally, no max pooling layer is used in the SSAB. The TAB converts the spatial and spectral features into abundance vector of the center pixel using  $1 \times 1$  two-dimensional convolutional layer (2Dconv). It is noted that ReLU [44] activation function is adopted to satisfy the non-negative constraint of abundance as the last layer. The length of the output abundance vector is equivalent to the number of the endmembers of multi-temporal HSIs. And the sum-to-one function  $F$  is conducted after the ReLU activation, designed to meet the sum to one constraint of abundance vector. Given a vector  $S = [s_1, s_2, \dots, s_K]^T \in \mathbb{R}^{K \times 1}$ ,  $F(s_i)$  is defined as:

$$F(s_i) = \frac{s_i}{\sum_{i=1}^K s_i + \varepsilon}, \quad i = 1, 2, \dots, K \quad (1)$$

where  $\varepsilon$  is a very small value to avoid the computation disorder. The output abundance vector  $S_1 \in \mathbb{R}^{K \times 1}$  of the first

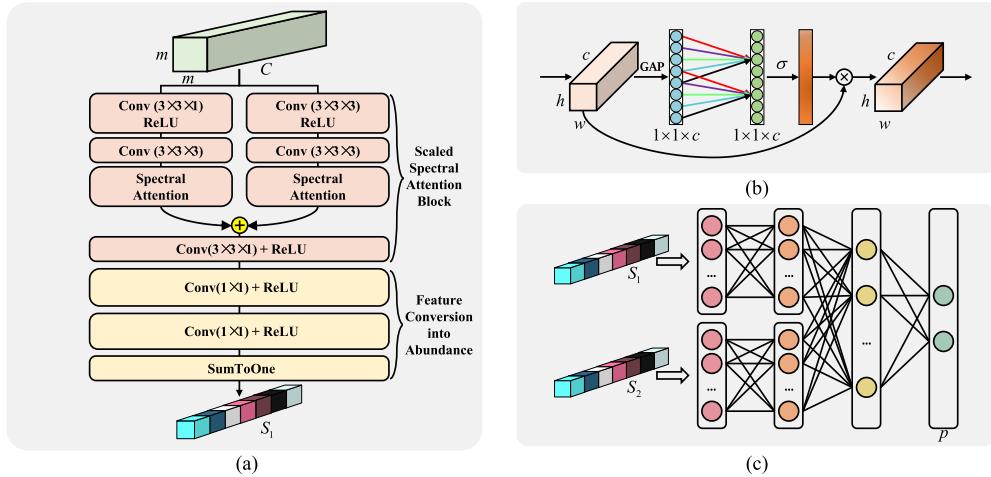


Fig. 4. (a) The detailed architecture of a single-branch UU-Module; (b) The Efficient Channel Attention block; (c) The detailed architecture of the proposed TC-Module, which learns the rules of detecting binary change information from two temporal abundance vectors automatically and adaptively.

single-branch UU-Module can be expressed as:

$$\begin{aligned} S_1 &= U_1(X_{\text{patch}}) = F(C_5(C_4(C_3(\text{Feature}_X)))) \\ \text{Feature}_X &= \text{ECA}_1(C_{12}(C_{11}(X_{\text{patch}}))) \\ &\quad + \text{ECA}_2(C_{22}(C_{21}(X_{\text{patch}}))) \end{aligned} \quad (2)$$

where  $K$  depicts the total number of all multi-temporal end-members; the  $U_1$  refers to the single branch of UU-Module with input as HSI  $X$ ;  $C_{11}$ ,  $C_{12}$ ,  $C_{21}$ ,  $C_{22}$  and  $C_3$  are all 3Dconvs of SSAB; ECA<sub>1</sub> and ECA<sub>2</sub> represent the ECA block, while  $C_4$  and  $C_5$  are the 2Dconv of TAB. Then the predictive spectrum of the center pixel  $\hat{x} \in \mathbb{R}^{C \times 1}$  can be obtained according to the linear mixed model [45], as follows:

$$\hat{x} = E \cdot S_1 \quad (3)$$

where the multi-temporal endmember matrix  $E \in \mathbb{R}^{C \times K}$  is available according to the following steps. Concretely, the two HSIs  $X \in \mathbb{R}^{H \times W \times C}$  and  $Y \in \mathbb{R}^{\tilde{H} \times \tilde{W} \times C}$  are firstly concatenated along the direction of image width to get the concatenation image  $Z \in \mathbb{R}^{H \times (2 \cdot W) \times C}$ . Next, the Hyperspectral Signal Identification by Minimum Error [46] is applied to compute the total number of all endmembers of two HSIs. Finally, we employ the Vertex Component Analysis [47] to calculate the endmember matrix  $E$  from the concatenation image  $Z$ . We chose cosine similarity function (named as COS\_SIM) to measure the difference between the predicted spectrum and the input center pixel spectrum. COS\_SIM emphasizes more the direction difference between two vectors than the absolute values. The closer the cosine similarity is to 1, the more parallel the two vectors are in direction, and vice versa. And the designed loss function can be described as follows:

$$\begin{aligned} L_{\cos}(x, \hat{x}) &= 1 - \text{COS\_SIM}(x, \hat{x}) \\ &= 1 - \frac{\sum_{i=1}^C x_i \cdot \hat{x}_i}{\sqrt{\sum_{i=1}^C x_i^2} \cdot \sqrt{\sum_{i=1}^C \hat{x}_i^2}} \end{aligned} \quad (4)$$

For the bi-temporal HSIs, the UU-Module is actually a special partial-siamese network. The SSABs of two HSIs share the same weights for the purpose of balancing the similarity of bi-temporal inputs, and gain the same abundance for the unchanged pixels of the two HSIs. However, the weights in TABs are different and are updated independently to acquire

more accurate united unmixing results in the case of two inputs with violent change.

### C. Temporal Correlation Module

The TC-Module represents a new binary change analysis rule based on the bi-temporal abundance vectors. It is noted that most of the unsupervised SU networks are trained by minimizing the distance between the reconstructive spectrum and original spectrum to obtain the abundance [48], [49]. The shorter distance, however, does not mean better abundance when there is noise or spectral disturbance of imaging condition in the original HSI. And the proposed TC-Module also places temporal correlation constraint on the UU-Module in the view of change detection.

As is shown in Fig. 4 (c), the TC-Module comprises four fully connected layers. The pair of abundance vectors  $S_1$  and  $S_2$  are fed into the TC-Module to extract features from the abundance vectors separately, then are concatenated to extract the inter-temporal change features, finally are transformed into binary change information. The number of units in the output layer is set as two. Considering the imbalanced samples between the changed and unchanged, focal loss [50] is a terrific choice to address the imbalanced problem and prevent the change detector or binary classifier from being overwhelmed by the huge number of easy samples. Given a pair of abundance vectors  $S_1$  and  $S_2$ , the output of the new change detection rule  $Q(S_1, S_2)$  from TC-Module is expressed as follows:

$$\begin{aligned} p &= Q(S_1, S_2) \\ &= \text{SOFTMAX}(W_3 \cdot g(W_2 \cdot \text{Feature}_{XY} + B_2) + B_3) \\ \text{Feature}_{XY} &= \text{CONCAT}(g(W_{11} \cdot S_1 + B_{11}), \\ &\quad g(W_{12} \cdot S_1 + B_{12})) \end{aligned} \quad (5)$$

where  $W_{11}$  and  $W_{12}$  are the weight matrices between the input layer and the first hidden layer of two input temporal abundance vectors;  $B_{11}$  and  $B_{12}$  represent the corresponding bias vectors;  $W_2$  and  $B_2$  are the weight matrix and bias vector between the first hidden layer and the second hidden layer, while  $W_3$  and  $B_3$  refer to that between the second hidden layer and the output layer. In addition,  $g$  represents the ReLU activation function; CONCAT denotes the operation of concatenation, and SOFTMAX indicates the softmax operator

that turns the extracted features into probability. The loss function we used is defined as follows:

$$L_{fc}(p, G) = -\alpha_t \cdot (1 - p_t)^\gamma \cdot \log(p_t) \quad (6)$$

$$p_t = \begin{cases} p & \text{if } G = 1 \\ 1 - p & \text{if } G = 0. \end{cases} \quad (7)$$

$$\alpha_t = \begin{cases} \alpha & \text{if } G = 1 \\ 1 - \alpha & \text{if } G = 0. \end{cases} \quad (8)$$

where  $G$  refers to the given pseudo binary class label. And there is a weighting factor  $\alpha \in [0, 1]$  which is used to address the imbalanced samples. For those easy samples,  $p_t$  tends to be close to one. In such case, the bigger  $\gamma$  is, the smaller the modulating factor  $(1 - p_t)^\gamma$  is, thus decreasing the rate of the loss of the easy samples more greatly.

#### D. The Scheme of Proposed BCG-Net

The proposed BCG-Net integrated SU and change detection in an unsupervised way for binary and multiclass change detection. The UU-Module and TC-Module adopt warmup strategy firstly to gain relatively stable unmixing and binary change detection performance. Next, the two modules are alternately optimized to ensure the stepwise enhancement of capability of unmixing and binary change detection of two modules. The TC-Module is kept constant during the UU-Module training phase, where the unmixing loss and the focal loss between the given pseudo binary labels and the binary change result generated from the multi-temporal abundance maps are both computed for updating the UU-Module. Similarly, the UU-Module is kept constant during the TC-Module training phase. Only the focal loss is back forward to update the TC-Module. The samples for the warmup of the UU-Module are not supervised training samples but the pixels randomly selected from the whole images. We adopted simple and widely used Change Vector Analysis (CVA) [51] and Expectation Maximization (EM) [52] to obtain pseudo binary labels for warmup of the TC-Module and alternate optimization of the whole training. We can get change probability of all pixels from the output after training. The binary change map is obtained from the output of the TC-Module, where a threshold value of 0.5 is employed to segment the change probability map.

For those changed pixels, each pixel is tagged as the class owning the maximum abundance value of the abundance vector. MCD are then acquired from the bi-temporal class comparison.

The scheme of BCG-Net is depicted as Algorithm 1. The detailed architecture of BCG-Net is exhibited in Table I.

### III. EXPERIMENTAL RESULTS AND ANALYSIS

#### A. Datasets Description

The first dataset is the simulative Urban dataset designed for evaluating the proposed temporal correlation constraint on the performance of united unmixing and change detection. As the first row of Fig 5 shows, (a) is the original Urban dataset; (b) is the variant of (a) with simulative temporal changes; (c) and (d) refer to binary and multiclass reference change maps, respectively. The original Urban data was collected on Copperas Cove, USA, in 1995. The image size is  $307 \times 307$ , with 162 spectral bands. There are a great deal of trees, grasses, and artificial architectures in the scene and four endmember materials were used, including Asphalt, Grass,

---

#### Algorithm 1 Process of Training and Generating Binary and Multiclass Change Detection Result for BCG-Net

---

##### Input:

HSI  $X$ , HSI  $Y$ , patch size  $m$ ;

##### Output:

The binary and multiclass change detection map  $I_b$  and  $I_m$ ;

- 1: Calculate the multi-temporal endmember  $E$ ;
- 2: Generate the pixels samples from bi-temporal HSIs,  $\{x^{(j)}, j=1, 2, \dots, M\}$  from HSI  $X$ ,  $\{y^{(j)}, j=1, 2, \dots, M\}$  from HSI  $Y$  for warmup of UU-Module, and pseudo binary labels  $\{G^{(j)}, j=1, 2, \dots, N\}$  from pre-detection binary change map;
- 3: Warm up the UU-Module and TC-Module;
- 4: **while**  $epoch < max\_epochs$  **do**
- 5: Update the TC-Module by descending its stochastic gradient:

$$\nabla_{\theta_t} \sum_{j=1}^N \left[ \frac{1}{N} \cdot L_{fc} \left( Q(S_1^{(j)}, S_2^{(j)}), G^{(j)} \right) \right]$$

- 6: Update the UU-Module by descending its stochastic gradient:

$$\nabla_{\theta_U} \sum_{j=1}^N \left[ L_{cos} \left( x^{(j)}, E \cdot U_1 \left( X_{patch}^{(j)} \right) \right) + L_{cos} \left( y^{(j)}, E \cdot U_2 \left( Y_{patch}^{(j)} \right) \right) + \frac{1}{N} \cdot \omega \cdot L_{fc} \left( Q(U_1(X_{patch}^{(j)}), U_2(Y_{patch}^{(j)})), G^{(j)} \right) \right]$$

where  $\omega$  is the weight factor to leverage the losses of two modules.

- 7:  $epoch++$ ;
  - 8: **end while**
  - 9: Calculate the predicted binary change probability and get the binary change map  $I_b$  with a threshold as 0.5;
  - 10: Calculate the abundance maps of the two HSIs and acquire the multiclass change map  $I_m$ ;
  - 11: **return**  $I_b$  and  $I_m$ ;
- 

Tree, and Roof, separately. Based on the original Urban data, several blocks were selected and inserted into another area to create changes. Additionally, we also added white Gaussian noise with different levels of SNR values (i.e., SNR = 20, 30, 40, 50 dB) to model the spectral variance of imaging conditions. The details of multiclass change information are depicted as the second column of Table II with seven changes.

The second hyperspectral dataset is a real-world dataset, named as USA dataset. As is shown in the second row of Fig. 5, (a) is acquired on May 1, 2004, and (b) on May 8, 2007 by Hyperion in Hermiston city, USA in an irrigated agricultural field; (c) and (d) are the binary and multiclass reference change maps, separately. The land cover types involve irrigated fields, soil, grassland, river, cultivated land, and building. The image covers a size of  $307 \times 241$ , with 154 spectral bands. Table II (the third column) displays the detailed change information.

The third one, China dataset, is the other real dataset. Fig 5 (a) (b) of the third row present the bi-temporal hyperspectral datasets shot on May 3, 2006 and April 23, 2007 by Hyperion, in the city of Yuncheng, Jiangsu, China over farmland area. (c) and (d) are the binary and multiclass reference change maps, individually. The image comprises a size of  $450 \times 140$ , and 155 spectral bands. There are only three change classes in this dataset, as Table II (the fourth column) shows.

#### B. Experimental Setting

We implemented our method by Pytorch and conducted experiments on a single NVIDIA RTX 2080 TI GPU. And

TABLE I  
THE ARCHITECTURE DETAILS FOR BCG-NET

Module	Layers	Type	Kernel and Padding Size	Channel
United Unmixing Module	3Dconv11	3Dconv+ReLU	Kernel (3×3×1), Padding (0,0,0)	C
	3Dconv12	3Dconv	Kernel (3×3×3), Padding (0,0,1)	C
	Spectral Attention1	AvgPooling+1Dconv+Sigmoid	Kernel= $\lceil \log_2(C) \rceil / 2 + 1 \rceil_{odd}$	C
	3Dconv21	3Dconv+ReLU	Kernel (3×3×3), Padding (0,0,1)	C
	3Dconv21	3Dconv	Kernel (3×3×3), Padding (0,0,1)	C
	Spectral Attention2	AvgPooling+1Dconv+Sigmoid	Kernel= $\lceil \log_2(C) \rceil / 2 + 1 \rceil_{odd}$	C
	3Dconv3	3Dconv+ReLU	Kernel (3×3×1), Padding (0,0,0)	C
	3Dconv4	2Dconv+ReLU	Kernel (1×1)	C/2
Feature Conversion into Abundance1	3Dconv5	2Dconv+ReLU+sumToOne	Kernel (1×1)	K
	3Dconv6	2Dconv+ReLU	Kernel (1×1)	C/2
	3Dconv7	2Dconv+ReLU+sumToOne	Kernel (1×1)	K
	FC11	Fully Connected+ReLU	—	K
Temporal Correlation Module	FC12	Fully Connected+ReLU	—	K
	FC2	Fully Connected+ReLU	—	K
	FC3	Fully Connected+softmax	—	2

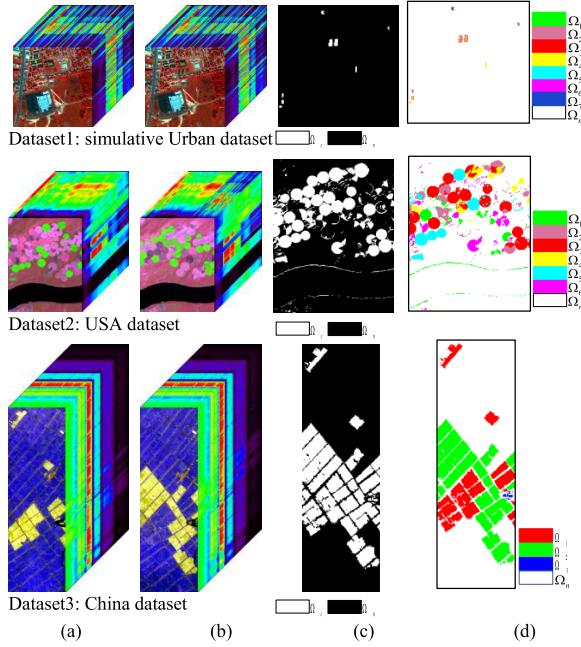


Fig. 5. From top to down are the simulative Urban dataset, USA dataset and China dataset. From left to right are (a) HSI  $X$ ; (b) HSI  $Y$ ; (c) Binary reference change map,  $\Omega_n$  refers to the unchanged, where  $\Omega_c$  refers to the changed; (d) Multiclass reference change map, where  $\Omega_n$  is the unchanged and  $\{\Omega_i\}_{i=1,2,\dots,7}$  represent different change classes.

TABLE II  
DETAILS OF THE MULTICLASS CHANGE INFORMATION

NO. of change class	Number of change pixels		
	Urban	USA	China
1	29	1850	6559
2	169	2905	11613
3	119	5896	105
4	182	1524	
5	4	2744	
6	38	1808	
7	49		

He-normal [53] way was selected as initialization of the network, and Adam [54] as optimizer with L2 regularization,

TABLE III  
DETAILS OF THE HYPERPARAMETERS OF BCG-NET

Hyperparameter	Urban	USA	China
$\alpha, \gamma$	0.25, 2	0.5, 2	0.25, 1
$\omega$	1	1	20
Learning rate	0.001	0.001	0.001
Weight decay rate	64	64	64
Batch size	7	7	7
Patch size $m$	200	200	200
Epoch (whole training)			

with the learning rate and weight decay rate both fixed as 0.001. The batch size is set as 64 and the patch size is equal to 7. The experiments are randomly repeated 3 times. The detailed hyperparameters settings are presented in Table III. The warmup samples for UU-Module are opted from the whole images at random, the number of which is all set 16384 for three datasets. For the pseudo binary labels, concretely, 2048 samples (2.17% of total pixel number, 1648 unchanged, 400 changed) from pre-detection result of EM are chose for the simulative Urban dataset, and 9216 samples (12.26% of whole pixel number, 6144 unchanged, 3072 changed) from pre-detection result of EM for the USA dataset, and 12288 samples (19.50% of total pixel number, 8192 unchanged, 4096 changed) from pre-detection result of CVA for the China dataset.

Overall accuracy (OA) and the Kappa coefficient were calculated for quantitative evaluation. To validate the effectiveness of proposed method, ten classical and state-of-the-art methods were conducted for comparison, with five binary change detection methods, including CVA [51], Iterative Slow Feature Analysis (ISFA) [55], [56], BIT [57], GETNET (without unmixing), as well as GETNET (with unmixing) [58], and another five multiclass change detection methods, Compressed Change Vector Analysis ( $C^2VA$ ) [59], KPCA-Mnet [60], MSU [38], PUC [34], as well as (Re3FCN) [61] included. The multi-temporal principle component analysis is first applied to the HSIs, and the main components are used for BCD for ISFA. BIT introduces transformer for change detection of remote sensing images to effectively model contexts

information within the bitemporal image. GETNET proposes an end-to-end 2-D convolutional neural network framework for binary change detection, constructing a novel mixed affinity matrix for cross channel features learning. GETNET (with unmixing) integrates the pixel spectral information and sub-pixel abundance information to build the mixed affinity matrix and extract multi-source information. Noted that the number of changed class is provided for the Re3FCN as well as C<sup>2</sup>VA as priori knowledge. KPCA-MNet is a novel unsupervised model, where kernel principal component analysis (KPCA) convolution is proposed to extract high-level spatial-spectral feature maps for both binary and multiple change detection. Re3FCN is also an end-to-end method combined 3D convolution and ConvLSTM for detecting of binary and multiple changes. Among all comparative methods, GETNET (with unmixing), MSU, PUC, Re3FCN all introduce unmixing or endmembers for change detection, which are opted specially for comparison. And GETNET (with unmixing), PUC and proposed method share the same endmember matrix extracted from the hyperspectral data for the sake of justice. All methods use the optimal hyperparameter setting recommended in their original references. The parameter settings for the comparative methods are presented in Table IV.

### C. Change Detection Results on Simulative Urban Datasets

Fig. 6 depicts the binary change detection results on the simulative Urban datasets. As SNR decreases, the amount of noise in HSI grows. From the binary change detection results with SNR as 20db, all methods can detect most of changes except CVA, MSU and Re3FCN. And there is a great deal of noise in the result of PUC, which reduces greatly when the SNR of dataset increases. Large false alarms can be found in the result of BIT (Fig. 6 (c)) due to the loss of local spatial details. More changes are detected in the results of CVA and MSU and less false alarm is found in the result of Re3FCN with higher SNR value. Among all the results of different comparative methods under variant SNR values, our proposed BCG-Net acquires the best binary change maps and is immune to noise.

The multiclass change detection maps of proposed BCG-Net with regard to C<sup>2</sup>VA, MSU, PUC, and Re3FCN are shown in Fig. 7. Different kinds of change classes are displayed by different colors, areas of no change are in white, and those change classes that cannot match with the reference are in black. Noted that there are seven change classes, which brings huge challenge for multiclass change detection. As shown in Fig. 7 (a) (b), it is obvious that the result of C<sup>2</sup>VA and KPCA-MNet under the four SNR levels are not consistent with the reference. Noted that the multiple changes of C<sup>2</sup>VA and KPCA-MNet are gained by clustering according to the change strength and angles. The results of PUC (Fig. 7 (d)) are severely impacted by the noise. And Re3FCN does not work very well on this dataset due to small number of samples and numerous changes classes. Compared with these comparative methods, the multiclass change map of proposed BCG-Net is the most similar with the reference, performing well even when SNR is as low as 20db and not affected by the small number of pixels of different changes class. By contrast, the multiple change results of BCG-Net acquired from the comparison of the multi-temporal abundance maps, incorporate more rich semantic information than C<sup>2</sup>VA and KPCA-MNet do. Notably, UU-Module extracts both spectral and spatial features from the input and the TC-Module places temporal correlation constraint on the SU from the point of

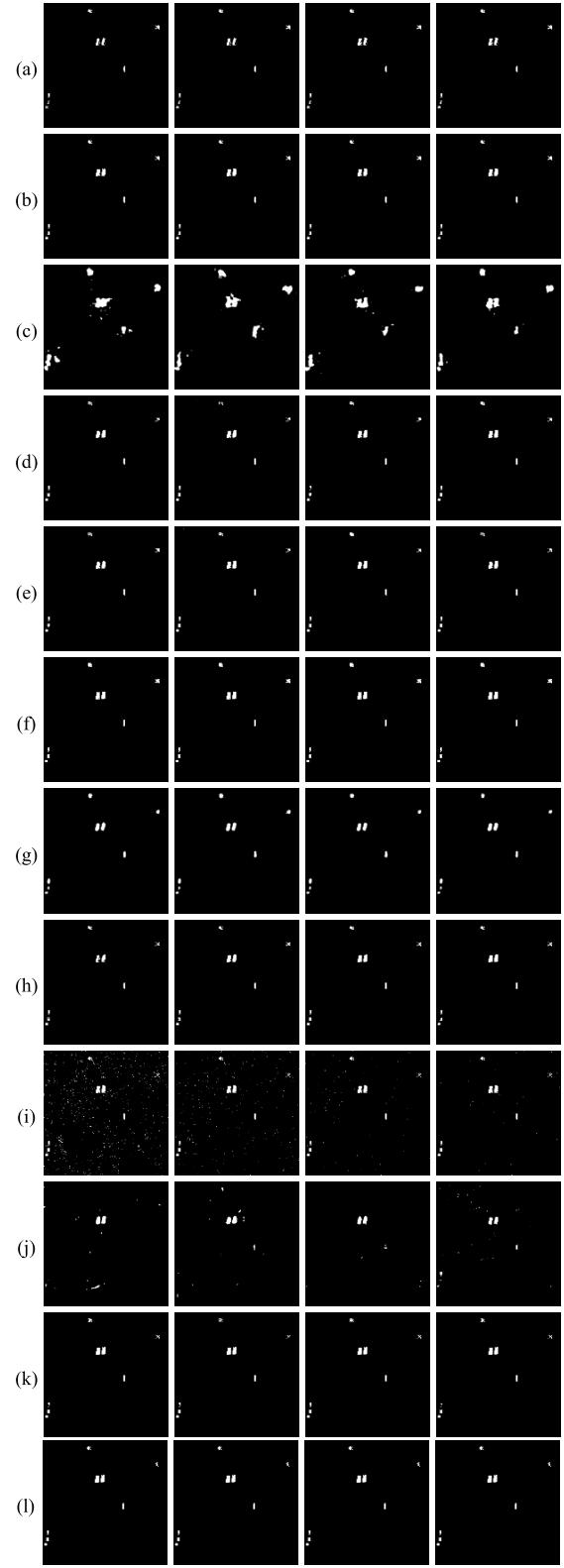


Fig. 6. The binary change detection results of Urban dataset. (a) CVA, (b) ISFA, (c) BIT, (d) GETNET (without unmixing), (e) GETNET (with unmixing), (f) C<sup>2</sup>VA, (g) KPCA-MNet, (h) MSU, (i) PUC, (j) Re3FCN, (k) BCG-Net, (l) Reference of binary change detection.

change detection, which contribute to the noise-free change detection result.

Table V summarizes the quantitative assessment of binary and multiclass change detection results. The maximum is

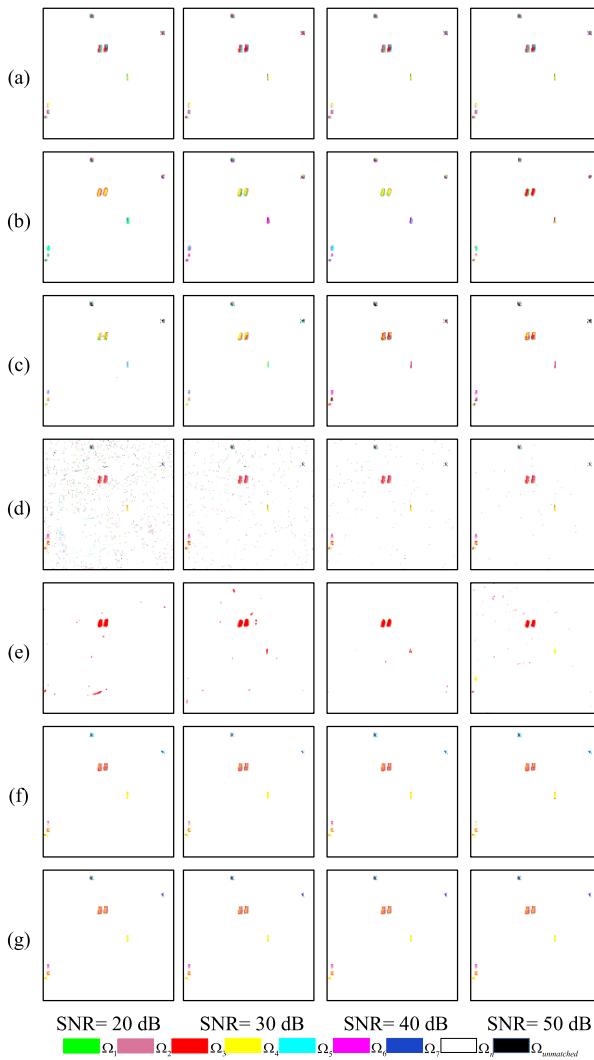


Fig. 7. The multiclass change detection results of simulative Urban dataset. (a) C<sup>2</sup>VA, (b) KPCA-MNet, (c) MSU, (d) PUC, (e) Re3FCN, (f) BCG-Net, (g) Reference of multiclass change detection.

marked in bold, and the second-best value is underlined. Noted that Kappa is more important for assessment of change detection task given the seriousness of the imbalance between the changed and the unchanged in the Urban dataset. It is found that the proposed BCG-Net acquires the best OA and Kappa coefficient of both binary and multiclass change detection result under all SNR values. For binary change detection, the ISFA and C<sup>2</sup>VA get the second-best OA and Kappa coefficient with different SNR values, respectively. As for multiclass change detection, the MSU obtains second-best OA and Kappa coefficient when SNR equals to 20db, 30db and 40db, separately. The PUC, however, obtains second-best OA and Kappa coefficient with SNR equal to 50db. The high false alarm rate leads to poor performance of Kappa for binary and multiple change detection results of BIT. KPCA-MNet and Re3FCN get low multiple change accuracy due to mistaken detected changes under the situation of seven changed class. It is worthy to mention that the second-best Kappa coefficient is much lower than the top one acquired by BCG-Net for multiclass change detection, confirming the superiority of proposed method.

#### D. Change Detection Results on USA Dataset

The binary change detection results of proposed BCG-Net on USA dataset as well as ten comparison methods are presented in Fig. 8. Compared with the binary reference change map (Fig. 8 (1)), it is obvious that there are larger white areas in the binary maps of MSU and PUC, revealing plenty of false positive changes. And there are some omitted changed areas in result of CVA, ISFA and Re3FCN. By contrast, GETNET, C<sup>2</sup>VA, KPCA-MNet, and proposed BCG-Net could detect most of changes accurately.

Fig. 9 shows the multiclass change detection maps. Lots of unmatched black areas can be found in the result of PUC (Fig. 9 (d)), where the before and after abundance comparison brings about far more change indications than the reference. By contrast, no black area is found in the multiclass change maps of C<sup>2</sup>VA and Re3FCN due to the provided prior information of the number of change classes, avoiding the redundant unmatched change classes. But the detected multiple changes in the black box of KPCA-MNet (Fig. 9 (b)) are inconsistent with the reference. As observed, most of the green area in the result of MSU is unchanged in effect. It is found that BCG-Net can detect most change classes correctly, despite of a piece of unmatched area. The constraint of temporal correlation supported by pseudo binary labels is able to optimize the united unmixing to encourage the coherence of the abundances of those unchanged objects and improve the accuracy of the abundances of those change objects.

Table VI reports the quantitative assessment of the results on USA dataset. It is observed that C<sup>2</sup>VA achieves the largest OA and Kappa coefficient as 0.9564 and 0.8749 for binary and as 0.8666 and 0.6586 for multiclass change detection results, respectively; BCG-Net gains second-best quantitative assessment performance as OA 0.9546 and Kappa 0.8662 for binary change detection and 0.5940 as Kappa for multiclass change detection, and the Re3FCN acquires the second-best OA as 0.8592 for the multiclass change detection result. It is worth noting that the no prior information is provided for BCG-Net while the C<sup>2</sup>VA and Re3FCN receive the number of change classes for multiple change detection. And BCG-Net gains comparative result in a totally unsupervised way. The performance of PUC is not optimistic, far behind the best OA and Kappa coefficient. The mission of changed in the result of KPCA-MNet, which may result from the loss of information during the kernel Principle component analysis, yields low Kappa values. Although BCG-Net gains the second-best performance, BCG-Net is an unsupervised method for both binary and multiclass change detection, acquiring a relatively nice effect for both binary and multiclass change maps.

#### E. Change Detection Results on China Dataset

Fig. 10 and Fig. 11 are the binary and multiclass change detection results on China dataset. As shown in Fig. 10, GETNET, KPCA-MNet, and proposed BCG-Net acquire nice performance on the binary change results, with most changes detected and little noise. It is observed that the binary change result of BCG-Net is not affected by the noise in the result of pre-detection algorithm CVA. Many stripes misclassified as changes can be found in the change maps of BIT, C<sup>2</sup>VA, MSU, and PUC, weakening the performance of multiclass change results further.

For multiclass change maps, many unmatched areas in black in Fig. 11 (c) and (d) indicate redundant change classes

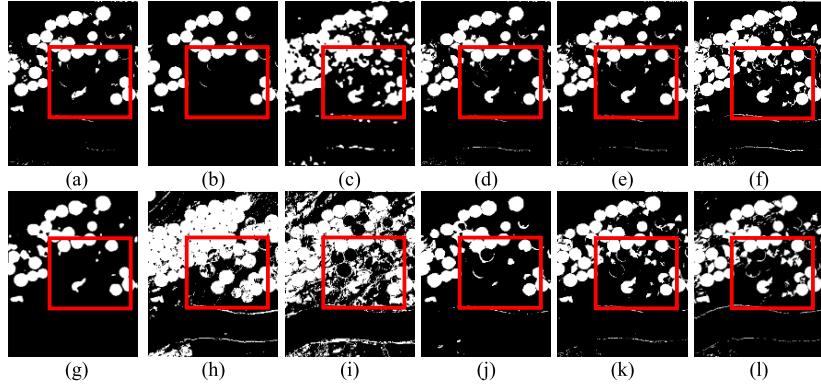


Fig. 8. The binary change detection results of USA dataset. (a) CVA, (b) ISFA, (c) BIT, (d) GETNET (without unmixing), (e) GETNET (with unmixing), (f) C<sup>2</sup>VA, (g) KPCA-MNet, (h) MSU, (i) PUC, (j) Re3FCN, (k) BCG-Net, (l) Reference of binary change detection.

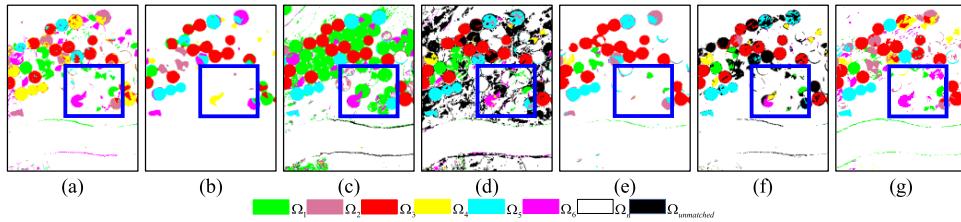


Fig. 9. The multiple change detection results of USA dataset. (a) C<sup>2</sup>VA, (b) KPCA-MNet, (c) MSU, (d) PUC, (e) Re3FCN, (f) BCG-Net, (g) Reference of multiclass change detection.

TABLE IV  
THE HYPERPARAMETER SETTINGS OF ALL COMPARATIVE METHODS

Method	Hyperparameter Setting	Urban(20db)	Urban(30db)	Urban(40db)	Urban(50db)	USA	China
CVA							
ISFA	Number of components of the multi-temporal Principle Component Analysis	7	7	7	7	9	10
BIT	Token length	4	4	4	4	4	4
GETNET (without unmixing)	Batch size	96	96	96	96	96	96
GETNET (with unmixing)	Batch size	96	96	96	96	96	96
C <sup>2</sup> VA	Alpha	0.4	0.4	0.4	0.4	0.5	0.5
KPCA-MNet	Kernel size	3	3	3	3	3	3
	Number of kernel size	8	8	8	8	8	8
MSU	Number of training patches	300	300	300	300	300	500
	Number of segmented patches	8	4	4	4	8	1
PUC							
	Batch size	64	64	64	64	64	64
RE3FCN	Patch size	9	9	9	9	9	9

detected by MSU and PUC. A little bit false alarms can be observed in the multiple change map of KPCA-MNet (Fig. 11 (b)). The change class in blue shown in Fig. 11 (a) of C<sup>2</sup>VA is not matched with the reference. By contrast, there is little noise and black unmatched area in the multiclass change map of BCG-Net.

For quantitative evaluation, Table VII lists the OA and Kappa coefficient on binary and multiclass change results. Concretely, for the assessment of binary change detection, the OA and Kappa coefficient of BCG-Net rank first, equal to 0.9664 and 0.9184, respectively. GETNET (with unmixing) wins second-best as OA 0.9626 and Kappa coefficient 0.9076, separately. And the result of MSU gets the worst performance, resulting from voluminous false alarm. For the assessment of multiclass change detection, the top of OA and Kappa coefficient are 0.9660 and 0.9249 respectively, attained by BCG-Net. KPCA-MNet wins second-best OA and Kappa coefficient as 0.9551 and 0.9018, separately. Some noise sparkle in the

binary change map of BIT hampers the performance of Kappa. To conclude, BCG-Net gains best OA and Kappa coefficient of both binary and multiclass change results, exhibiting the effectiveness of proposed method.

#### F. Parameter Analysis

The hyperparameter analysis of proposed BCG-Net is exhibited in this part. The hyperparameters involved in BCG-Net include the  $\alpha$ ,  $\gamma$  of the focal loss, and the  $\omega$ , the weight of focal loss in the final loss. The weighting factor  $\alpha$  put weight  $\alpha$  for the changed class and  $1 - \alpha$  for the unchanged class. The bigger  $\gamma$  is, the more the rate of the loss of the easy samples are decreased greatly. Concretely, the hyperparameter analyses of the  $\alpha$  and  $\gamma$ , where  $\alpha \in [0.25, 0.50, 0.75]$  and  $\gamma \in [1, 2]$  according to [50], are implemented on the warmup of the TC-Module. And the analysis of the  $\omega$  is conducted on the alternate optimization of the whole network, where  $\omega \in [1, 5, 10, 20, 30, 40]$ . The Urban dataset with SNR as

TABLE V  
QUANTITATIVE ASSESSMENT ON THE SIMULATIVE URBAN DATASET UNDER DIFFERENT SNR VALUES

SNR	20db				30db			
	Two-class		Multiclass		Two-class		Multiclass	
Method	OA	Kappa	OA	Kappa	OA	Kappa	OA	Kappa
CVA	0.9976	0.7855			0.9978	0.8100		
ISFA	<u>0.9990</u>	0.9159			<u>0.9990</u>	0.9160		
BIT	0.9837	0.4438			0.9831	0.4211		
GETNET (without unmixing)	0.9986	0.8890			0.9985	0.8811		
GETNET (with unmixing)	0.9985	0.8789			0.9989	0.9133		
C <sup>2</sup> VA	0.9989	<u>0.9205</u>	0.9947	0.6012	<u>0.9990</u>	<u>0.9221</u>	0.9947	0.6075
KPCA-MNet	0.9972	0.7848	0.9936	0.5151	0.9971	0.7824	0.9936	0.5184
MSU	0.9985	0.8698	<u>0.9964</u>	<u>0.6876</u>	0.9989	0.9087	<u>0.9959</u>	<u>0.6678</u>
PUC	0.9852	0.4307	0.9834	0.3626	0.9942	0.6607	0.9924	0.5545
Re3FCN	0.9951	0.5777	0.9932	0.4240	0.9952	0.5851	0.9933	0.4212
BCG-Net (ours)	<b>0.9992</b>	<b>0.9324</b>	<b>0.9986</b>	<b>0.8901</b>	<b>0.9991</b>	<b>0.9263</b>	<b>0.9987</b>	<b>0.8909</b>
SNR	40db				50db			
	Two-class		Multiclass		Two-class		Multiclass	
Method	OA	Kappa	OA	Kappa	OA	Kappa	OA	Kappa
CVA	0.9982	0.8491			0.9982	0.8512		
ISFA	<u>0.9990</u>	0.9163			<u>0.9990</u>	0.9163		
BIT	0.9867	0.4815			0.9884	0.5188		
GETNET (without unmixing)	0.9988	0.9017			<u>0.9990</u>	0.9205		
GETNET (with unmixing)	0.9989	0.9152			0.9988	0.9017		
C <sup>2</sup> VA	<u>0.9990</u>	<u>0.9222</u>	0.9946	0.5994	0.9989	<u>0.9215</u>	0.9946	0.5989
KPCA-MNet	0.9972	0.7866	0.9935	0.5138	0.9969	0.7920	0.9935	0.5173
MSU	0.9986	0.8856	<u>0.9964</u>	<u>0.6934</u>	0.9989	0.9109	0.9963	0.7008
PUC	0.9975	0.8183	0.9957	0.6868	0.9986	0.8893	<u>0.9968</u>	<u>0.7459</u>
Re3FCN	0.9961	0.6204	0.9943	0.4374	0.9960	0.6524	0.9937	0.4463
BCG-Net (ours)	<b>0.9993</b>	<b>0.9463</b>	<b>0.9989</b>	<b>0.9098</b>	<b>0.9991</b>	<b>0.9316</b>	<b>0.9987</b>	<b>0.8973</b>

TABLE VI  
QUANTITATIVE ASSESSMENT ON THE USA DATASET

Method	Two-class		Multiclass	
	OA	Kappa	OA	Kappa
CVA	0.9272	0.7670		
ISFA	0.9023	0.6716		
BIT	0.9259	0.7938		
GETNET (without unmixing)	0.9289	0.7764		
GETNET (with unmixing)	0.9332	0.7897		
C <sup>2</sup> VA	<b>0.9564</b>	<b>0.8749</b>	<b>0.8666</b>	<b>0.6586</b>
KPCA-MNet	0.9241	0.7551	0.8418	0.5323
MSU	0.8132	0.5752	0.7140	0.4424
PUC	0.7825	0.4975	0.6945	0.4008
Re3FCN	0.9131	0.7255	<u>0.8592</u>	0.5921
BCG-Net (ours)	<u>0.9546</u>	<u>0.8662</u>	0.8456	<u>0.5940</u>

TABLE VII  
QUANTITATIVE ASSESSMENT ON CHINA DATASET

Method	Two-class		Multiclass	
	OA	Kappa	OA	Kappa
CVA	0.9548	0.8926		
ISFA	0.9575	0.8996		
BIT	0.9405	0.8588		
GETNET (without unmixing)	0.9584	0.8974		
GETNET (with unmixing)	<u>0.9626</u>	<u>0.9076</u>		
C <sup>2</sup> VA	0.9364	0.8534	0.8627	0.7238
KPCA-MNet	0.9551	0.8919	<u>0.9551</u>	<u>0.9018</u>
MSU	0.8432	0.6713	0.8407	0.7121
PUC	0.8650	0.6997	0.8612	0.7280
Re3FCN	0.9306	0.8381	0.9305	0.8536
BCG-Net (ours)	<b>0.9664</b>	<b>0.9184</b>	<b>0.9660</b>	<b>0.9249</b>

20db is selected as the representative of all four Urban datasets with different SNR values.

As shown in Fig. 12, we find that Urban dataset (SNR = 20db) and USA dataset work well when  $\alpha \in [0.25, 0.50]$  and  $\gamma \in [1, 2]$ . And it achieves the best OA and Kappa for both two-class and multiclass change detection results at  $\alpha = 0.25, \gamma = 2$  on Urban dataset (SNR=20db) (Fig. 12(a)) and  $\alpha = 0.5, \gamma = 2$  for USA dataset (Fig. 12 (c)), respectively. For China dataset (shown in Fig. 12 (e)), it is obvious that  $\alpha = 0.25, \gamma = 1$  gains the highest result. It is found that all three datasets behave badly when  $\alpha = 0.75$ , owing to too much stress on the changed class and little concern on the unchanged class. Consequently, lots of unchanged area are detected as the changed, bringing about high false alarm rate and leading to the drop of the Kappa performance on both binary and multiple change results. Fig. 12 (b) (d) (f) present the hyperparameter analysis of the  $\omega$ . It is observed that the performance of the three datasets initially increase and then drop as the increase of the  $\omega$ . And the Urban(SNR=20db) and USA datasets obtain the best effect when  $\omega = 1$ . The China dataset gets the highest accuracy under  $\omega = 20$ . Noted that the change detection results of the final alternate optimization outperform the result of warmup of the TC-Module.

#### IV. DISCUSSIONS

On this section, the effect of temporal correlation constraint on the SU is firstly discussed. Then the discussion on the effect of TC-Module on change detection is provided. And the running time cost is analyzed finally.

##### A. The Effect of Temporal Correlation Constraint on Spectral Unmixing

Here, the unmixing performance comparison on the simulative Urban dataset is represented to test the effect of the

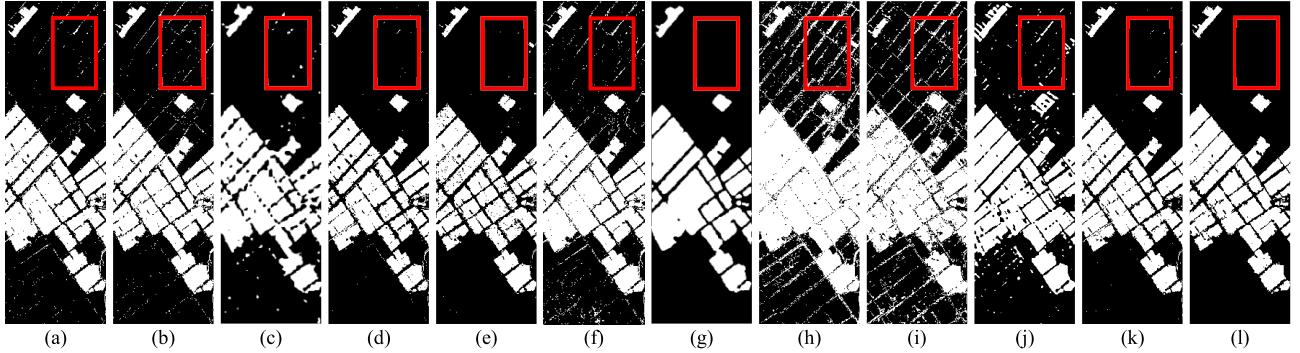


Fig. 10. The binary change detection of China dataset. (a) CVA, (b) ISFA, (c) BIT, (d) GETNET (without unmixing), (e) GETNET (with unmixing), (f) C<sup>2</sup>VA, (g) KPCA, (h) MSU, (i) PUC, (j) Re3FCN, (k) BCG-Net, (l) Reference of binary change detection.

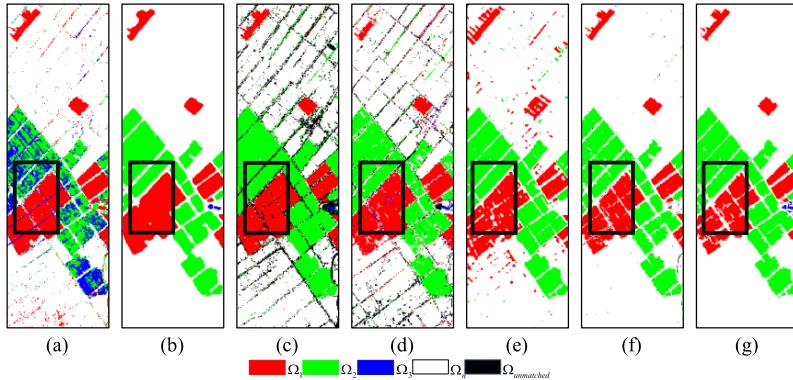


Fig. 11. The multiclass change detection of China dataset. (a) C<sup>2</sup>VA, (b) KPCA-MNet, (c) MSU, (d) PUC, (e) Re3FCN, (f) BCG-Net, (g) Reference of multiclass change detection.

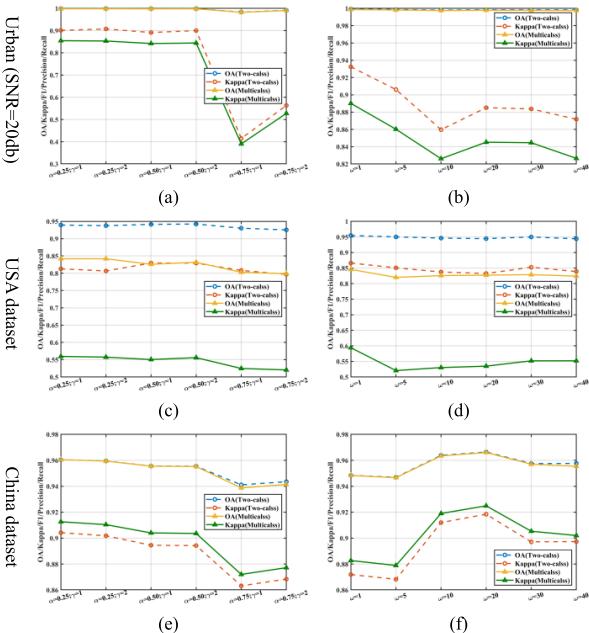


Fig. 12. The hyperparameter analysis of BCG-Net. The first column is the hyperparameter analyses of  $\alpha$  and  $\gamma$  for the focal loss for (a) Urban dataset ( $\text{SNR}=20\text{db}$ ), (c) USA dataset, and (e) China dataset. The second column is the hyperparameter analysis of weight  $\omega$  in the final loss for (b) Urban dataset ( $\text{SNR}=20\text{db}$ ), (d) USA dataset, and (f) China dataset.

proposed temporal correlation constraint. For sake of fairness, we use the result of UU-Module as comparison without temporal correlation. And we also test the abundance map

obtained by fully constrained least square (FCLS) [62]. The metric to evaluate the abundance maps is depicted as:

$$\text{MSE} = \frac{1}{K} \sum_{i=1}^K \|S_j^{(i)} - \hat{S}_j^{(i)}\|^2, \quad j = 1, 2 \quad (9)$$

where  $K$  is number of multi-temporal endmembers;  $S_1^{(i)}$  and  $S_2^{(i)}$  refer to abundance map of the endmember  $i$  of HSI  $X$  and HSI  $Y$ ;  $\hat{S}_1^{(i)}$  and  $\hat{S}_2^{(i)}$  are the corresponding reference abundance maps.

The MSE comparison result is represented in Table VIII. The best results are in bold. As can be seen, the MSE of UU-Module and BCG-Net are far lower than that of FCLS for both HSI  $X$  and  $Y$  under four different SNR values. It might be contributed to the powerful feature extraction ability of the designed united unmixing network, where spectral and spatial information are taken into consideration at the same time. Moreover, BCG-Net obtains lower abundance MSE than UU-Module does for most of the endmembers of two HSIs at different SNR values, which confirms that the temporal correlation constraint really works to improve the performance of SU. For BCG-Net, the designed TC-Module is able to constrain the unmixing result from the view of binary change detection. Generally, for the unchanged pixels, the abundances of them at two temporal HSIs are probably similar with each other. Likewise, the abundances of the changed pixels at two temporal HSIs are probably different with each other. With the stimulation of the temporal correlation constraint, the abundances of those unchanged pixels are encouraged to be consistent and that of changed pixels be more accurate, contributing to greater multi-temporal unmixing results. Besides,

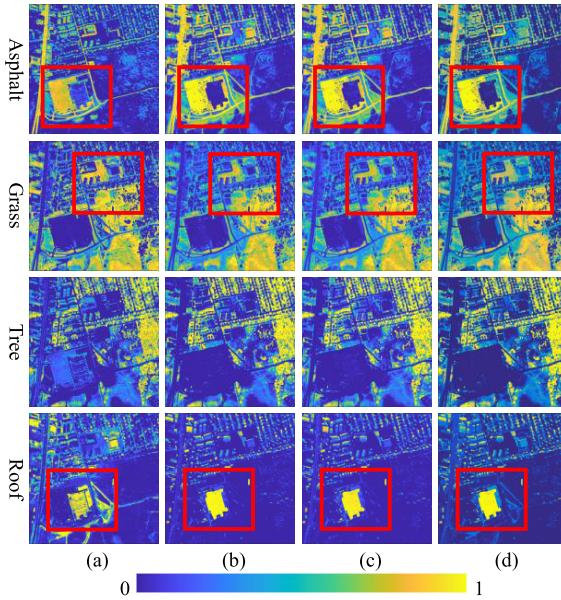


Fig. 13. The visual comparison of estimated abundance maps on simulative Urban dataset with SNR equals to 20db. From left to right are the estimated abundance maps of HSI  $Y$  on the (a) FCLS, (b) UU-Module, (c) BCG-Net, and (d) ground truth.

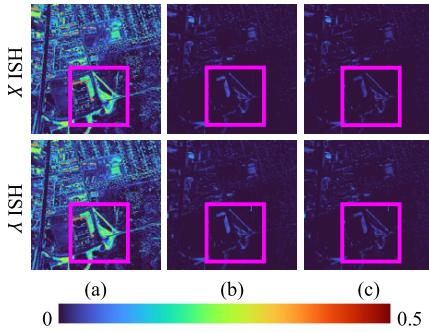


Fig. 14. The visual comparison of residual error maps of estimated abundance maps on simulative Urban dataset with SNR as 20db. From left to right are the residual error maps of (a) FCLS, (b) UU-Module and (c) BCG-Net.

the abundance MSE of HSI  $Y$  is higher than the one of HSI  $X$  when SNR is equivalent to 20db, 30db. This is because the massive noise resulting from low SNR undermines the quality of HSIs and further influences the performance of unmixing. And the abundance MSE of all unmixing result improves with higher SNR value.

Fig. 13 shows the visual comparison of abundance maps on simulative Urban dataset under SNR as 20db. It is obvious that the abundance maps of UU-Module (Fig. 13(b)) and BCG-Net (Fig. 13(c)) are closely similar with the reference (Fig. 13(d)). FCLS (Fig. 13(a)), however, performs not well especially in the abundance maps on the Asphalt (Fig. 13(a), the first row) and the Roof (Fig. 13(a), the fourth row), compared with the reference. For better visualization of the difference between the abundance results of UU-Module and BCG-Net, the residual error between the estimated abundance and the reference is computed according to:

$$R_{MSE} = \frac{1}{K} \sum_{i=1}^K (S_j^{(i)} - \hat{S}_j^{(i)})^2, \quad j = 1, 2 \quad (10)$$

Fig. 14 represents the residual error maps of estimated abundance maps from FCLS, UU-Module and BCG-Net on

simulative Urban dataset with SNR equals to 20db. From the marked pinked frame, it is observed that the error of BCG-Net is lower than that of UU-Module, both of which are largely lower than the residual error acquired by FCLS. All in all, it is concluded that the proposed BCG-Net achieves more accurate abundance maps than UU-Module and FCLS does, demonstrating the validity of proposed temporal correlation constraint on the SU.

### B. The Effect of Temporal Correlation Constraint on Change Detection

To test the effect of temporal correlation constraint on change detection, comparison experiment of proposed BCG-Net and the traditional post unmixing comparison is conducted on the three datasets. To be fair, the SU method here both employs the proposed UU-Module, named as PUC (UU-Module). Additionally, the PUC with FCLS as SU method is also taken for comparison, named as PUC(FCLS).

Table IX represents the quantitative comparison of binary and multiclass change results of these three methods. The greatest result is in bold. For all datasets, BCG-Net acquires much better OA and Kappa coefficient in both binary and multiclass change detection result than PUC(UU-Module) and PUC(FCLS) do. Compared with PUC(UU-Module), BCG-Net obtains more 16% around rise on the OA of both binary and multiclass change results for USA dataset, and 3% around growth on that of China dataset. Besides, for the Kappa coefficient of both binary and multiclass result, BCG-Net acquires a more 22% to 58% increase than PUC (UU-Module) does especially for simulative Urban dataset and USA dataset, and 2% and 5% growth on the China dataset. Compared with PUC(FCLS), BCG-Net obtains a more 10% to 17% rise on the OA of both binary and multiclass results for USA and China datasets, and more 19% to 52% increase on the Kappa coefficient of that for all the three datasets.

It is noted that PUC(FCLS) obtains better performance than PUC(UU-Module) does for simulative Urban dataset and USA dataset, which indicates that the traditional PUC method does not always gain better change detection result with better unmixing result. The reason can be summarized as follows. Traditional PUC method usually adopts such a rule that each pixel is firstly assigned the class holding the maximum of the abundance vector and then the bi-temporal class labels are compared to acquire binary change information. That is to say, the class with the maximum abundance value is regarded as the main component of the pixel. If the main components remain unchanged, the pixel is considered to be unchanged. Otherwise, the pixel is thought to be changed. However, this brief binary change detection rule is susceptible to subtle differences of abundance values. For example, a mixed pixel is composed of only two endmembers with equal abundances and keeps unchanged during two phases. The abundance components of these two endmembers, however, are not equal resulting from numerical computation of unmixing process, leading to the pixel altered. And the phenomenon is more common in the datasets with complicated background, such as simulative Urban dataset and USA dataset.

For the proposed BCG-Net, apart from casting a temporal correlation constraint on the UU-Module, TC-Module also plays a role of powerful binary change detector, seeking the exact relationship between the multi-temporal abundance vectors pair and the binary change information and bringing

TABLE VIII

THE MSE COMPARISON OF ABUNDANCE MAP FOR SPECTRAL UNMIXING WITH AND WITHOUT PROPOSED TEMPORAL CORRELATION CONSTRAINT

SNR	Method	20db		30db		40db		50db	
		FCLS	UU-Module	BCG-Net	FCLS	UU-Module	BCG-Net	FCLS	UU-Module
X	Asphalt	0.0595	0.0067	<b>0.0064</b>	0.0595	0.0071	<b>0.0064</b>	0.0595	0.0071
	Grass	0.0266	<b>0.0043</b>	<b>0.0043</b>	0.0266	0.0050	<b>0.0044</b>	0.0266	0.0049
	Tree	0.0219	<b>0.0019</b>	<b>0.0019</b>	0.0219	0.0031	<b>0.0019</b>	0.0219	0.0026
	Roof	0.0793	0.0022	<b>0.0020</b>	0.0793	0.0028	<b>0.0021</b>	0.0793	0.0027
	Average	0.0468	0.0038	<b>0.0036</b>	0.0468	0.0045	<b>0.0037</b>	0.0468	0.0043
Y	Asphalt	0.0600	0.0074	<b>0.0073</b>	0.0598	0.0075	<b>0.0070</b>	0.0598	0.0068
	Grass	0.0269	0.0054	<b>0.0052</b>	0.0266	0.0054	<b>0.0048</b>	0.0266	0.0050
	Tree	0.0222	<b>0.0023</b>	0.0024	0.0220	0.0031	<b>0.0021</b>	0.0220	0.0022
	Roof	0.0801	<b>0.0025</b>	0.0026	0.0800	0.0029	<b>0.0022</b>	0.0800	0.0026
	Average	0.0473	<b>0.0044</b>	<b>0.0044</b>	0.0471	0.0047	<b>0.0040</b>	0.0471	0.0042

TABLE IX

QUANTITATIVE COMPARISON ON THE CHANGE DETECTION PERFORMANCE BETWEEN THE PUC(UU-MODULE) WITHOUT TEMPORAL CONSTRAINT AND BCG-NET WITH TEMPORAL CONSTRAINT

Dataset	Model	Two-class		Multiclass	
		OA	Kappa	OA	Kappa
Urban (20db)	PUC(UU-Module)	0.9779	0.3522	0.9774	0.3409
	PUC(FCLS)	0.9852	0.4307	0.9834	0.3626
	BCG-Net	<b>0.9992</b>	<b>0.9324</b>	<b>0.9986</b>	<b>0.8901</b>
Urban (30db)	PUC(UU-Module)	0.9902	0.5523	0.9896	0.5274
	PUC(FCLS)	0.9942	0.6607	0.9924	0.5545
	BCG-Net	<b>0.9991</b>	<b>0.9263</b>	<b>0.9987</b>	<b>0.8909</b>
Urban (40db)	PUC(UU-Module)	0.9911	0.5889	0.9906	0.5647
	PUC(FCLS)	0.9975	0.8183	0.9957	0.6868
	BCG-Net	<b>0.9993</b>	<b>0.9463</b>	<b>0.9989</b>	<b>0.9098</b>
Urban (50db)	PUC(UU-Module)	0.9916	0.5919	0.9911	0.5707
	PUC(FCLS)	0.9986	0.8893	0.9968	0.7459
	BCG-Net	<b>0.9991</b>	<b>0.9316</b>	<b>0.9987</b>	<b>0.8973</b>
USA	PUC(UU-Module)	0.7926	0.5237	0.6769	0.3703
	PUC(FCLS)	0.7825	0.4975	0.6945	0.4008
	BCG-Net	<b>0.9546</b>	<b>0.8662</b>	<b>0.8456</b>	<b>0.5940</b>
China	PUC(UU-Module)	0.9393	0.8589	0.9383	0.8710
	PUC(FCLS)	0.8650	0.6997	0.8612	0.7280
	BCG-Net	<b>0.9664</b>	<b>0.9184</b>	<b>0.9660</b>	<b>0.9249</b>

no false detection problem. Besides, the UU-Module and TC-Module are optimized alternately, eliminating the bias and accumulated errors from abundance result to change detection result.

Fig. 15 presents the visual comparison of binary and multiclass change maps of the traditional PUC (UU-Module) method and BCG-Net on simulative Urban dataset. Massive noise can be observed in the binary and multiclass change maps of PUC (UU-Module), as shown in Fig. 15 (a), (c), whereas most of them disappear in the result of BCG-Net (Fig. 15 (b), (d)). Fig. 16 is another example of the effect of the TC-Module on the USA dataset. Compared with the result of PUC (UU-Module) in Fig. 16(a), BCG-Net (shown

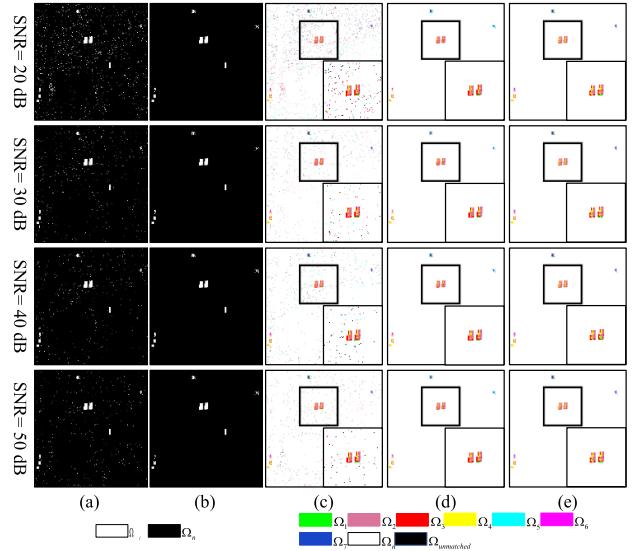


Fig. 15. The visual comparison of change detection result on the simulative Urban dataset under SNR as 20db, 30db, 40db and 50db, respectively (from left to right). From top to down are binary change maps of (a) traditional PUC(UU-Module), (b) proposed BCG-Net, and multiclass change maps of (c) traditional PUC(UU-Module), (d) proposed BCG-Net, and (e) multiclass ground truth.

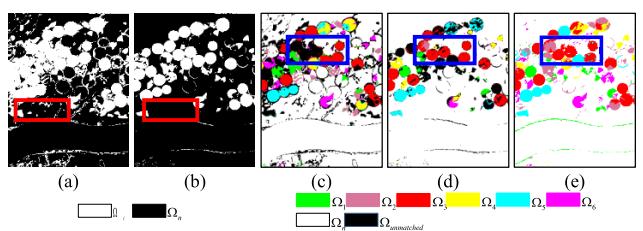


Fig. 16. The visual comparison of change detection result on the USA dataset. From left to right are binary change maps of (a) traditional PUC (UU-Module), (b) proposed BCG-Net, and multiclass change maps of (c) traditional PUC(UU-Module), (d) proposed BCG-Net, and (e) multiclass ground truth.

TABLE X  
THE COMPARISON OF THE RUNTIME COST OF ALL METHODS ON THREE DATASETS

Method	Urban (20db)	Urban (30db)	Urban (40db)	Urban (50db)	USA	China
CVA	4.5111	4.4784	4.3969	4.3628	1.6825	0.6106
ISFA	5.3724	5.0764	5.0702	4.9500	3.4796	2.1950
BIT	17.5265	20.4259	18.6879	18.7523	12.6461	11.6416
GETNET (without unmixing)	267.2014	164.6019	195.8644	195.7137	191.4209	195.7750
GETNET (with unmixing)	267.1340	182.7625	255.5028	237.0207	227.4805	193.0283
C <sup>2</sup> VA	7.8085	6.7920	6.6122	6.6742	5.5430	5.6023
KPCA-MNet	17.4711	17.7046	18.0086	18.0461	11.4562	11.2171
MSU	67.4801	66.9681	79.4838	81.2448	45.6035	37.5106
PUC	16.8674	16.8743	16.9566	17.4751	14.1700	7.2490
RE3FCN	98.8500	375.1560	198.4764	204.9609	<b>2139.4033</b>	<b>1812.8435</b>
BCG-Net (ours)	<b>921.2013</b>	<b>1421.7168</b>	<b>1418.9848</b>	<b>1259.5052</b>	1920.7794	1664.4297

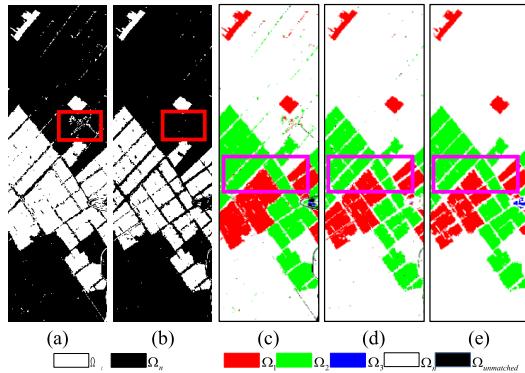


Fig. 17. The visual comparison of change detection result on the China dataset. From left to right are binary change maps of (a) traditional PUC (UU-Module), (b) proposed BCG-Net, and multiclass change maps of (c) traditional PUC(UU-Module), (d) proposed BCG-Net, (e) multiclass ground truth.

in Fig. 16(b)) can detect more accurate change area and less false alarms on the binary change maps. And for the multiclass change maps, as the blue frame marks, BCG-Net (shown in Fig. 16 (d)) acquires more sound multiclass change detection results than PUC (UU-Module) (shown in Fig. 16(c)) does. This is because that the TC-Module encourages the UU-Module to acquire better aligned unmixing result, which contributes to more accurate multiclass change detection result. From the comparison result of China dataset shown in Fig. 17, the noise in the red box of the binary change map of PUC (UU-Module) (Fig. 17 (a)) is eliminated to some extent in the result of BCG-Net (Fig. 17 (b)). And the multiclass change map of BCG-Net (Fig. 17 (d)) is more similar with the reference than that of PUC (UU-Module) (Fig. 17 (c)). In summary, with TC-Module, BCG-Net shows better performance on reducing the false alarm of binary change map and boosting the multiclass detection result.

### C. Runtime Cost Analysis

To test the runtime efficiency of the proposed BCG-Net and all other comparative method, the running time cost experiment have been conducted on all datasets. As shown in Table X, the longest running time is in bold and the shortest running time is in italic. It is observed that the proposed BCG-Net takes the longest time for the Urban dataset and the Re3FCN spends the longest time for the USA and China dataset. The traditional CVA gains the shortest running time for all datasets. Noted that it really takes a long time to solve

the spectral unmixing of the UU-Module to obtain the multi-temporal abundance maps. Besides, the iterative optimization of the UU-Module and TC-Module is indeed time-consuming. But more importantly, among all methods, the proposed BCG-Net achieves the best binary and multiple change detection results in totally unsupervised way, and greater unmixing performance are discovered in the discussions. Under the present circumstances, it is suggested to call on more attention to multiple class change detection research to acquire a good balance between the efficiency and the performance.

### V. CONCLUSION

In this article, we propose an unsupervised hyperspectral multiclass change detection method named as BCG-Net to solve the problem of error accumulation and neglection of temporal correlation encountered by traditional methods. Instead of obtaining the binary and multiclass changes directly from the unmixing result like most previous methods, a novel temporal correlation constraint directed by pseudo binary labels is designed to boost the SU process from the point of view of change detection, where the abundance of the unchanged pixels is encouraged to be more consistent and that of the changed pixels more accurate. Besides, we put forward a new rule based on neural network to build an effective relationship between the abundance vectors pairs and the change information, bypassing the abundant false alarms traditional rule suffers from. The represented innovative strategy of alternately optimizing the unmixing process and the change detection process provides a terrific solution to eliminate the accumulated error and bias from unmixing result to changed detection result. The qualitative and quantitative evaluation are conducted on three hyperspectral datasets to test the effectiveness of proposed method. In summary, the proposed binary change guided hyperspectral multiclass change detection network achieves competitive or even superior result on all tested datasets, demonstrating the validity of temporal correlation constraint on the binary and multiclass change detection results, as well as the multi-temporal spectral unmixing result.

### REFERENCES

- [1] J. Yue, L. Fang, and M. He, "Spectral-spatial latent reconstruction for open-set hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 31, pp. 5227–5241, 2022, doi: [10.1109/TIP.2022.3193747](https://doi.org/10.1109/TIP.2022.3193747).
- [2] J. W. P. Sun, H. Li, W. Li, X. Meng, C. Ge, and Q. Du, "Low-rank and sparse representation for hyperspectral image processing: A review," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 1, pp. 10–43, Jun. 2021, doi: [10.1109/MGRS.2021.3075491](https://doi.org/10.1109/MGRS.2021.3075491).

- [3] S. Zhou, Z. Xue, and P. Du, "Semisupervised stacked autoencoder with cotraining for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 6, pp. 3813–3826, Jun. 2019, doi: [10.1109/TGRS.2018.2888485](https://doi.org/10.1109/TGRS.2018.2888485).
- [4] Y. Xu, L. Zhang, B. Du, and F. Zhang, "Spectral–spatial unified networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 5893–5909, Oct. 2018, doi: [10.1109/TGRS.2018.2827407](https://doi.org/10.1109/TGRS.2018.2827407).
- [5] W. Sun et al., "A simple and effective spectral–spatial method for mapping large-scale coastal wetlands using China ZY1–02D satellite hyperspectral images," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 104, Dec. 2021, Art. no. 102572, doi: <https://doi.org/10.1016/j.jag.2021.102572>.
- [6] Z. Xue, Y. Zhou, and P. Du, "S3Net: Spectral–spatial Siamese network for few-shot hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–19, 2022, doi: [10.1109/TGRS.2022.3181501](https://doi.org/10.1109/TGRS.2022.3181501).
- [7] D. Zhu, B. Du, and L. Zhang, "Target dictionary construction-based sparse representation hyperspectral target detection methods," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 4, pp. 1254–1264, Apr. 2019, doi: [10.1109/JSTARS.2019.2902430](https://doi.org/10.1109/JSTARS.2019.2902430).
- [8] A. W. Bitar, L.-F. Cheong, and J.-P. Ovarlez, "Sparse and low-rank matrix decomposition for automatic target detection in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5239–5251, Aug. 2019, doi: [10.1109/TGRS.2019.2897635](https://doi.org/10.1109/TGRS.2019.2897635).
- [9] S. Chang, B. Du, and L. Zhang, "A subspace selection-based discriminative forest method for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 6, pp. 4033–4046, Jun. 2020, doi: [10.1109/TGRS.2019.2960391](https://doi.org/10.1109/TGRS.2019.2960391).
- [10] H. Su, Z. Wu, A.-X. Zhu, and Q. Du, "Low rank and collaborative representation for hyperspectral anomaly detection via robust dictionary construction," *ISPRS J. Photogramm. Remote Sens.*, vol. 169, pp. 195–211, Nov. 2020.
- [11] H. Zhang, H. Chen, G. Yang, and L. Zhang, "LR-Net: low-rank spatial–spectral network for hyperspectral image denoising," *IEEE Trans. Image Process.*, vol. 30, pp. 8743–8758, 2021, doi: [10.1109/TIP.2021.3120037](https://doi.org/10.1109/TIP.2021.3120037).
- [12] B. Lin, X. Tao, and J. Lu, "Hyperspectral image denoising via matrix factorization and deep prior regularization," *IEEE Trans. Image Process.*, vol. 29, pp. 565–578, 2020, doi: [10.1109/TIP.2019.2928627](https://doi.org/10.1109/TIP.2019.2928627).
- [13] Y. Su, X. Xu, J. Li, H. Qi, P. Gamba, and A. Plaza, "Deep autoencoders with multitask learning for bilinear hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 10, pp. 8615–8629, Oct. 2021, doi: [10.1109/TGRS.2020.3041157](https://doi.org/10.1109/TGRS.2020.3041157).
- [14] Y. Zhong, X. Wang, L. Zhao, R. Feng, L. Zhang, and Y. Xu, "Blind spectral unmixing based on sparse component analysis for hyperspectral remote sensing imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 119, pp. 49–63, Sep. 2016.
- [15] R. Dian, S. Li, and L. Fang, "Learning a low tensor-train rank representation for hyperspectral image super-resolution," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 9, pp. 2672–2683, Sep. 2019, doi: [10.1109/TNNLS.2018.2885616](https://doi.org/10.1109/TNNLS.2018.2885616).
- [16] R. Dian, L. Fang, and S. Li, "Hyperspectral image super-resolution via non-local sparse tensor factorization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5344–5353, doi: [10.1109/CVPR.2017.411](https://doi.org/10.1109/CVPR.2017.411).
- [17] K. Ren, W. Sun, X. Meng, G. Yang, J. Peng, and J. Huang, "A locally optimized model for hyperspectral and multispectral images fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–15, 2022, doi: [10.1109/TGRS.2021.3133670](https://doi.org/10.1109/TGRS.2021.3133670).
- [18] W. Sun, K. Ren, X. Meng, C. Xiao, G. Yang, and J. Peng, "A band divide-and-conquer multispectral and hyperspectral image fusion method," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–13, 2022, doi: [10.1109/TGRS.2020.3046321](https://doi.org/10.1109/TGRS.2020.3046321).
- [19] C. Wu, B. Du, and L. Zhang, "Hyperspectral anomalous change detection based on joint sparse representation," *ISPRS J. Photogramm. Remote Sens.*, vol. 146, pp. 137–150, Dec. 2018.
- [20] A. A. Nielsen, "The regularized iteratively reweighted mad method for change detection in multi- and hyperspectral data," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 463–478, Feb. 2007.
- [21] P. Ghamisi et al., "Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 37–78, Dec. 2017, doi: [10.1109/MGRS.2017.2762087](https://doi.org/10.1109/MGRS.2017.2762087).
- [22] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: A systematic survey," *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 294–307, Mar. 2005, doi: [10.1109/TIP.2004.838698](https://doi.org/10.1109/TIP.2004.838698).
- [23] Z. Zhu, "Change detection using Landsat time series: A review of frequencies, preprocessing, algorithms, and applications," *ISPRS J. Photogramm. Remote Sens.*, vol. 130, pp. 370–384, Aug. 2017, doi: [10.1016/j.isprsjprs.2017.06.013](https://doi.org/10.1016/j.isprsjprs.2017.06.013).
- [24] P. Fu and Q. Weng, "A time series analysis of urbanization induced land use and land cover change and its impact on land surface temperature with Landsat imagery," *Remote Sens. Environ.*, vol. 175, pp. 205–214, Mar. 2016.
- [25] Z. Zheng, Y. Zhong, J. Wang, A. Ma, and L. Zhang, "Building damage assessment for rapid disaster response with a deep object-based semantic change detection framework: From natural disasters to man-made disasters," *Remote Sens. Environ.*, vol. 265, Nov. 2021, Art. no. 112636.
- [26] F. Bovolo and L. Bruzzone, "A split-based approach to unsupervised change detection in large-size multitemporal images: Application to tsunami-damage assessment," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1658–1670, Jun. 2007.
- [27] X. Li, P. Gong, and L. Liang, "A 30-year (1984–2013) record of annual urban dynamics of Beijing city derived from Landsat data," *Remote Sens. Environ.*, vol. 166, pp. 78–90, Sep. 2015.
- [28] J. Deng et al., "A methodology to monitor urban expansion and green space change using a time series of multi-sensor SPOT and Sentinel-2A images," *Remote Sens.*, vol. 11, no. 10, p. 1230, 2019.
- [29] Z. Hou, W. Li, L. Li, R. Tao, and Q. Du, "Hyperspectral change detection based on multiple morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2021, Art. no. 5507312.
- [30] D. Marinelli, F. Bovolo, and L. Bruzzone, "A novel change detection method for multitemporal hyperspectral images based on binary hyperspectral change vectors," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4913–4928, Jul. 2019.
- [31] S. Liu, D. Marinelli, L. Bruzzone, and F. Bovolo, "A review of change detection in multitemporal hyperspectral images: Current techniques, applications, and challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 2, pp. 140–158, Jun. 2019, doi: [10.1109/MGRS.2019.2898520](https://doi.org/10.1109/MGRS.2019.2898520).
- [32] L. Bruzzone, S. Liu, F. Bovolo, and P. Du, "Change detection in multitemporal hyperspectral images," in *Multitemporal Remote Sensing: Methods and Applications*, Y. Ban, Ed. Cham, Switzerland: Springer, 2016, pp. 63–88, doi: [10.1007/978-3-319-47037-5\\_4](https://doi.org/10.1007/978-3-319-47037-5_4).
- [33] A. Singh, "Review article digital change detection techniques using remotely-sensed data," *Int. J. Remote Sens.*, vol. 10, no. 6, pp. 989–1003, Jun. 1989.
- [34] Q. Du, L. Wasson, and R. King, "Unsupervised linear unmixing for change detection in multitemporal airborne hyperspectral imagery," in *Proc. Int. Workshop Anal. Multi-Temporal Remote Sens. Images*, 2005, pp. 136–140.
- [35] A. Ertürk, M.-D. Iordache, and A. Plaza, "Sparse unmixing-based change detection for multitemporal hyperspectral images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 2, pp. 708–719, Feb. 2016.
- [36] C. Shi and L. Wang, "Incorporating spatial information in spectral unmixing: A review," *Remote Sens. Environ.*, vol. 149, pp. 70–87, Jun. 2014.
- [37] T. Uezato, R. J. Murphy, A. Melkumyan, and A. Chilingaryan, "Incorporating spatial information and endmember variability into unmixing analyses to improve abundance estimates," *IEEE Trans. Image Process.*, vol. 25, no. 12, pp. 5563–5575, Dec. 2016.
- [38] S. Liu, L. Bruzzone, F. Bovolo, and P. Du, "Unsupervised multitemporal spectral unmixing for detecting multiple changes in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 2733–2748, May 2016, doi: [10.1109/TGRS.2015.2505183](https://doi.org/10.1109/TGRS.2015.2505183).
- [39] H. Jafarzadeh and M. Hasanolou, "An unsupervised binary and multiple change detection approach for hyperspectral imagery based on spectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 12, pp. 4888–4906, Dec. 2019.
- [40] Q. Guo and C. Zhong, "Change detection for hyperspectral images via convolutional sparse analysis and temporal spectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 4417–4426, 2021.
- [41] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, "Learning spatiotemporal features with 3D convolutional networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 4489–4497.
- [42] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient channel attention for deep convolutional neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 1–9.
- [43] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141.

- [44] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, vol. 25, Dec. 2012, pp. 1097–1105.
- [45] J. M. Bioucas-Dias et al., "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.
- [46] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.
- [47] J. M. P. Nascimento and J. M. Bioucas-Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.
- [48] Y. Su, J. Li, A. Plaza, A. Marinoni, P. Gamba, and S. Chakravortty, "DAEN: Deep autoencoder networks for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4309–4321, Jul. 2019.
- [49] B. Palsson, M. O. Ulfarsson, and J. R. Sveinsson, "Convolutional autoencoder for spectral-spatial hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 1, pp. 535–549, Jan. 2020.
- [50] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 318–327.
- [51] R. D. Johnson and E. S. Kasischke, "Change vector analysis: A technique for the multispectral monitoring of land cover and condition," *Int. J. Remote Sens.*, vol. 19, no. 3, pp. 411–426, 1998.
- [52] L. Bruzzone and D. F. Prieto, "Automatic analysis of the difference image for unsupervised change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1171–1182, May 2000.
- [53] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Santiago, Chile, Dec. 2015, pp. 1026–1034, doi: [10.1109/ICCV.2015.123](https://doi.org/10.1109/ICCV.2015.123).
- [54] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent.*, 2014, pp. 1–15.
- [55] L. Zhang, C. Wu, and B. Du, "Automatic radiometric normalization for multitemporal remote sensing imagery with iterative slow feature analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6141–6155, Oct. 2014.
- [56] C. Wu, B. Du, and L. Zhang, "Slow feature analysis for hyperspectral change detection," in *Proc. 6th Workshop Hyperspectral Image Signal Process., Evol. Remote Sens. (WHISPERS)*, Jun. 2014, pp. 1–4.
- [57] H. Chen, Z. Qi, and Z. Shi, "Remote sensing image change detection with transformers," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–14, 2022, doi: [10.1109/TGRS.2021.3095166](https://doi.org/10.1109/TGRS.2021.3095166).
- [58] Q. Wang, Z. Yuan, Q. Du, and X. Li, "GETNET: A general end-to-end 2-D CNN framework for hyperspectral image change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 3–13, Jan. 2019.
- [59] F. Bovolo, S. Marchesi, and L. Bruzzone, "A framework for automatic and unsupervised detection of multiple changes in multitemporal images," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 6, pp. 2196–2212, Jun. 2012.
- [60] C. Wu, H. Chen, B. Du, and L. Zhang, "Unsupervised change detection in multitemporal VHR images based on deep kernel PCA convolutional mapping network," *IEEE Trans. Cybern.*, vol. 52, no. 11, pp. 12084–12098, Nov. 2021, doi: [10.1109/TCYB.2021.3086884](https://doi.org/10.1109/TCYB.2021.3086884).
- [61] A. Song, J. Choi, Y. Han, and Y. Kim, "Change detection in hyperspectral images using recurrent 3D fully convolutional networks," *Remote Sens.*, vol. 10, no. 11, p. 1827, 2018.
- [62] D. C. Heinz and C.-I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.



**Meiqi Hu** (Graduate Student Member, IEEE) received the B.S. degree in surveying and mapping engineering from the School of Geoscience and Info-Physics, Central South University, Changsha, China, in 2019. She is currently pursuing the Ph.D. degree with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China.

Her research interests include deep learning and multitemporal remote sensing image change detection.



Chen Wu (Member, IEEE) received the B.S. degree in surveying and mapping engineering from Southeast University, Nanjing, China, in 2010, and the Ph.D. degree in photogrammetry and remote sensing from the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China, in 2015.

He is currently an Associate Professor with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing. His research interests include multitemporal remote sensing image change detection and analysis in multispectral and hyperspectral images.



**Bo Du** (Senior Member, IEEE) received the Ph.D. degree in photogrammetry and remote sensing from the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China, in 2010.

He is currently a Professor with the School of Computer Science, Institute of Artificial Intelligence, Wuhan University. He is also the Director of the National Engineering Research Center for Multimedia Software, Wuhan University. He has published more than 80 research articles in the IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON CYBERNETICS, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE JOURNAL OF SELECTED TOPICS IN EARTH OBSERVATIONS AND APPLIED REMOTE SENSING, and IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, 13 of them are ESI hot papers or highly cited papers. His major research interests include pattern recognition, hyperspectral image processing, and signal processing.

Dr. Du regularly serves as a Senior PC Member for IJCAI and AAAI. He served as an Area Chair for ICPR. He won the Highly Cited Researcher 2019 by the Web of Science Group. He also won the International Joint Conferences on Artificial Intelligence (IJCAI) Distinguished Paper Prize, the IEEE Data Fusion Contest Champion, and the IEEE Workshop on Hyperspectral Image and Signal Processing Best paper Award, in 2018. He serves as a Reviewer for 20 Science Citation Index (SCI) magazines, including IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, IEEE TRANSACTIONS ON CYBERNETICS, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE JOURNAL OF SELECTED TOPICS IN EARTH OBSERVATIONS AND APPLIED REMOTE SENSING, and IEEE GEOSCIENCE AND REMOTE SENSING LETTERS. He also serves as an Associate Editor for *Neural Networks, Pattern Recognition, and Neurocomputing*.



**Liangpei Zhang** (Fellow, IEEE) received the B.S. degree in physics from Hunan Normal University, Changsha, China, in 1982, the M.S. degree in optics from the Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, China, in 1988, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1998.

He is the Chair Professor at the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University. He was a Principal Scientist of the China State Key Basic Research Project (2011–2016) appointed by the Ministry of National Science and Technology of China to lead the Remote Sensing Program in China. He has published more than 700 research papers and five books. He is the Institute for Scientific Information (ISI) highly cited author. He is the holder of 30 patents. His research interests include hyperspectral remote sensing, high-resolution remote sensing, image processing, and artificial intelligence.

Dr. Zhang is a fellow of the Institution of Engineering and Technology (IET). He was a recipient of the 2010 Best Paper Boeing Award, the 2013 Best Paper ERDAS Award from the American Society of Photogrammetry and Remote Sensing (ASPRS), and the 2016 Best Paper Theoretical Innovation Award from the International Society for Optics and Photonics (SPIE). His research teams have won the top three prizes of the IEEE GRSS 2014 Data Fusion Contest, and his students have been selected as the winners or finalists of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS) Student Paper Contest in recent years. He is the Founding Chair of the IEEE Geoscience and Remote Sensing Society (GRSS) Wuhan Chapter. He serves as an associate editor or the editor for more than ten international journals. He also serves as an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.