Ensemble and Random Collaborative Representation-Based Anomaly Detector for Hyperspectral Imagery

Rong Wang, Wei Feng, Qianrong Zhang, Feiping Nie, Zhen Wang, and Xuelong Li, *Fellow, IEEE*

Abstract—**In recent years, hyperspectral anomaly detection (HAD) has become an active topic and plays a significant role in military and civilian fields. As a classic HAD method, the collaboration representation-based detector (CRD) has attracted extensive attention and in-depth research. Despite the good performance of CRD method, its computational cost is too high for the widely demanded real-time applications. To alleviate this problem, a novel ensemble and random collaborative representation-based detector (ERCRD) is proposed for HAD. This approach comprises two main steps. Firstly, we propose a random background modeling to replace the sliding dual window strategy used in the original CRD method. Secondly, we can obtain multiple detection results through multiple random background modeling, and these results are further refined to final detection result through ensemble learning. Experiments on four real hyperspectral datasets exhibit the accuracy and efficiency of this proposed ERCRD method compared with ten state-of-the-art HAD methods.**

Index Terms—**Hyperspectral imagery (HSI), hyperspectral anomaly detection (HAD), collaborative representation, random background modeling, ensemble learning.**

I. INTRODUCTION

Hyperspectral imagery (HSI) offers plentiful useful spectral and spatial information to monitor the earth's surface for the fine identification of various land cover materials [1]–[5]. Owing to the powerful advantage of HSI, it has been widely applied to various remote sensing fields, such as scene classification [3], [5], unmixing [6], clustering [7], [8], change detection [9] and target or anomaly detection [2], [10]–[12]. Hyperspectral anomaly detection (HAD) has also attracted much interest and in-depth research for its widespread application in areas like military reconnaissance, civilian search and rescue, environmental monitoring and mineral exploration [13], [14].

In essence, HAD is an unsupervised binary classification problem, which detects anomalies against background without any prior knowledge of this scene. In general, anomalies usually refer to pixels with two significant features, namely, low occurrence probability and distinct spectral-spatial characteristic different from the background. For example, tanks in the forest background, warships in the sea background and aircraft in the airport background.

In the last three decades, various methods have been developed to detect anomalies in HSI. Generally speaking, there are two main kinds of existing HAD methods: statistics modeling HAD and representation-based HAD. Statistics modeling HAD assume a multivariate normal (Gaussian) background distribution. See [15] for a well-known statistics modeling method called the Reed-Xiaoli (RX) detector, which identifies anomalies based on the Mahalanobis distance between pixel and the background, with an assumption of multivariate Gaussian background distribution. The RX detector is now the benchmark in HSI anomaly detection and comes in two variants: the global RX (GRX) detector which uses the entire image to model the background, and the local RX (LRX) detector which uses the local dual window image. Afterwards, numerous variants of the RX detector have been proposed [16]–[25]. For example, the kernel RX (KRX) [16] method, considered as the nonlinear version of the RX detector, uses the kernel theory to characterize the non-Gaussian distributions in high-dimensional feature space. The cluster-based anomaly detection (CBAD) [17] method first grouped the entire image into several clusters, then the above RX detector was applied on each cluster to achieve better detection performance. The subspace RX (SSRX) [18] detector assumes that the high-variance principal components and low-variance principal components are associated with background subspace and anomaly subspace, respectively. The SSRX detector is obtained by applying the RX detector to the low-variance principal components. The weighted RX (WRX) [21] detector assigned specific weight to every pixel for background modeling. To speed up the KRX detector, the cluster KRX (CKRX) [25] detector applied the fast eigenvalue decomposition on clusters, which are all obtained by clustering the entire image. Chang *et al.* [26] designed a background-anomaly component projection and separation optimized detector (BASO) to detect anomalies from the perspective of optimization theory.

Besides, many representation-based HAD methods, including sparse representation [27]–[32], low-rank representation [33]–[37] and collaborative representation [38]–[43], have also attracted substantial attention and in-depth research. For example, Chen *et al.* [27] came up with the sparse representation-based HAD method, which assumes the back-

R. Wang is with the School of Cybersecurity and Center for OPTical IMagery Analysis and Learning (OPTIMAL), Northwestern Polytechnical University, Xi'an 710072, P. R. China. E-mail: wangrong07@tsinghua.org.cn.

W. Feng, Q. Zhang, F. Nie, and X. Li are with the School of Computer Science and Center for OPTical IMagery Analysis and Learning (OPTIMAL), Northwestern Polytechnical University, Xi'an 710072, P. R. China. E-mail: feipingnie@gmail.com.

Z. Wang is with the Center for OPTical IMagery Analysis and Learning (OPTIMAL) and School of Mechanical Engineering, Northwestern Polytechnical University, Xi'an 710072, China.

ground can be approximated by several atoms in the dictionary, while the anomalies cannot. The low-rank and sparse matrix decomposition (LRaSMD) detector designed by Sun *et al.* [33] separates the anomalies by calculating the Euclidean distance between each pixel and the mean vector on the sparse part. Zhang *et al.* [34] came up with the low-rank and sparse matrix decomposition-based Mahalanobis distance (LSMAD) method, which exert the low-rank constraint on the background and the sparse constraint on anomalies. Under the assumption that each background pixel can be approximated by its spatial neighborhood pixels in a sliding dual window centered at this pixel, while the anomalies cannot, Li and Du proposed the collaborative representation-based detector (CRD) for HSI in [38]. On the basis of the CRD, many variants have been proposed recently. To make better use of the spatial features from HSI, a morphology-based collaborative representation detector (MCRD) [39] was presented. Principal component analysis was applied in CRD method to remove outliers (PCAro-CRD) [41] first uses PCA to extract the main background pixel information, and then predicts the anomalies with collaborative representation. It comes in two variants: Global PCAroCRD and Local PCAroCRD. In order to decrease the complexity of computing, a recursive CRD (RCRD) [42] algorithm was proposed based on the matrix inversion lemma. In addition to the above two categories of methods, many scholars have also explored HAD methods on the strength of support vector data description (SVDD) [44], [45], morphological and attribute filters [46]–[48], tensor decomposition [49]–[52], and deep convolutional neural networks [53]–[55], etc., which shows that the in-depth study of HAD method has become a popular topic.

Here we are focusing on the family of collaborative representation-based HAD methods. These methods are built based on an assumption: each background pixel can be approximated by its spatial neighborhood pixels in a dual window centered at this pixel, while the anomalies cannot. Most traditional collaborative representation-based HAD methods, however, focus on detection accuracy while ignoring the underlying computational complexity, which is of great importance for the anomaly detection of a large dataset. The complexity mainly arises from the sliding dual window strategy. Therefore, we come up with a novel ensemble and random collaborative representation-based detector (ERCRD) with random background modeling and ensemble learning. Firstly, we propose a random background modeling to replace the sliding dual window strategy and use the same background pixels for each pixel in our model. Then, we can obtain multiple detection results through multiple random background modeling, and these results are further refined to final detection result through ensemble learning. The proposed ERCRD shows its advantages over many existing HAD methods, such as GRX [15], LRX [15], SSRX [18], CBAD [17], and LSMAD [34]. In comparison with the CRD [38] and its variants [39], [41], [42], we also validate that our ERCRD method is able to attain considerable or better detection accuracy with even much less implementation time.

The rest of this paper is arranged as follows. In Section II, we briefly review the traditional CRD and the recently

proposed PCAroCRD. The proposed approach ERCRD is described in Section III. In Section IV, we conduct empirical studies on four real datasets to validate the accuracy and efficiency of our approach. Finally, we summarize this paper in Section V.

II. RELATED WORK

Let $\mathbb{X} \in \mathbb{R}^{d \times h \times w}$ denotes the hyperspectral imagery, where d is the number of spectral bands, and h and w are the height and width of the background respectively. The three-dimensional (3-D) hyperspectral imagery $\mathbb{X}$ is transformed into a two-dimensional (2-D) matrix $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n] \in \mathbb{R}^{d \times n}$, where $n = h \times w$ is the total number of pixels.
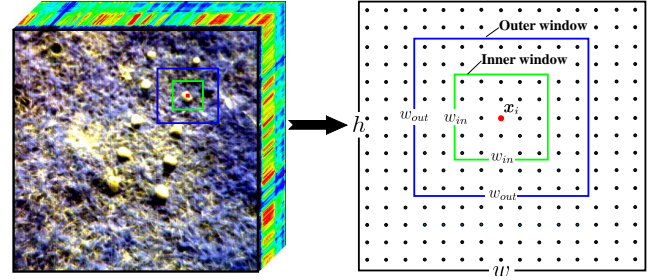

Fig. 1: The sliding dual window of the CRD.

A. CRD

In this subsection, we review the recently proposed collaborative representation-based detector (CRD) [38]. The CRD assumed that the background pixel can be well approximated by its spatial neighborhood pixels, whereas the anomalies cannot. For the pixel $\boldsymbol{x}_i \in \mathbb{R}^{d \times 1}$, its spatial neighborhood pixels are selected by a sliding dual window, as can be seen from Fig. 1. The spatial neighborhood pixels specifically refer to pixels outside the inner window and inside the outer window. The size of the outer window and the inner window are represented as w_{out} and w_{in}, respectively. Thus, the spatial neighborhood pixels are resized into a two-dimensional matrix $\boldsymbol{X}_s = [\bar{\boldsymbol{x}}_1, \bar{\boldsymbol{x}}_2, \cdots, \bar{\boldsymbol{x}}_s] \in \mathbb{R}^{d \times s}$, where s denotes the number of spatial neighborhood pixels and $s = w_{out} \times w_{out} - w_{in} \times w_{in}$. The objective function of the CRD is defined as [38]

$$\min_{\boldsymbol{\alpha}_i} \|\boldsymbol{x}_i - \boldsymbol{X}_s \boldsymbol{\alpha}_i\|_2^2 + \lambda \|\boldsymbol{\alpha}_i\|_2^2, \tag{1}$$

where $\boldsymbol{\alpha}_i \in \mathbb{R}^{s \times 1}$ denotes the weight vector and λ is the regularization parameter. Taking the derivative w.r.t $\boldsymbol{\alpha}_i$ and setting it to zero, so we have

$$\hat{\boldsymbol{\alpha}}_i = (\boldsymbol{X}_s^T \boldsymbol{X}_s + \lambda \boldsymbol{I})^{-1} \boldsymbol{X}_s^T \boldsymbol{x}_i, \tag{2}$$

where $\boldsymbol{I}$ is an identity matrix. The reconstruction error r_i for the pixel $\boldsymbol{x}_i$ is regarded as the anomaly score and can be computed by

$$r_i = \|\boldsymbol{x}_i - \boldsymbol{X}_s \hat{\boldsymbol{\alpha}}_i\|_2. \tag{3}$$

If r_i is greater than a threshold, then the pixel $\boldsymbol{x}_i$ is called anomalous pixel.

The computational complexity of this step is $O(ds^2 + s^3 + ds + d)$. Each pixel $\boldsymbol{x}_i$ needs to get the corresponding surrounding pixel matrix $\boldsymbol{X}_s$ on its own sliding double window. Thus, the weight vector $\boldsymbol{\alpha}_i$ and the anomaly score r_i need to calculate n times, thereby the total computational complexity of the CRD is $O(nds^2 + ns^3 + nds + nd)$. This repeated process needs a high computational burden, reduces the speed of the anomaly detection, and limits the application of the CRD in real-time tasks.

B. PCAroCRD

The PCAroCRD method was recently proposed by Su *et al.* [41], which applied PCA to CRD method for removing outliers. It has two versions: Global PCAroCRD and Local PCAroCRD.

The Global PCAroCRD first obtains the projection matrix $\boldsymbol{W}_g \in \mathbb{R}^{n \times n}$ by solving the standard PCA model:

$$\max_{\boldsymbol{W}_g^T \boldsymbol{W}_g = \boldsymbol{I}} \operatorname{tr}(\boldsymbol{W}_g^T \boldsymbol{X}^T \boldsymbol{X} \boldsymbol{W}_g), \qquad (4)$$

then the spatial-domain PCA transformation is represented as

$$\hat{\boldsymbol{X}} = \boldsymbol{X} \boldsymbol{W}_g \qquad (5)$$

where $\hat{\boldsymbol{X}} = [\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{x}}_2, \cdots, \hat{\boldsymbol{x}}_n] \in \mathbb{R}^{d \times n}$ denotes the transformed data matrix. For the pixel $\boldsymbol{x}_i$, the objective function of the Global PCAroCRD can be written as follows [41]

$$\min_{\boldsymbol{\alpha}_i} \|\boldsymbol{x}_i - \boldsymbol{X}_m \boldsymbol{\alpha}_i\|_2^2 + \lambda \|\boldsymbol{\Gamma}_i \boldsymbol{\alpha}_i\|_2^2. \qquad (6)$$

where $\boldsymbol{X}_m = [\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{x}}_2, \cdots, \hat{\boldsymbol{x}}_m] \in \mathbb{R}^{d \times m}$ denotes the first m principal components of $\hat{\boldsymbol{X}}$ and contains the most information of $\boldsymbol{X}$ in the spatial domain. $\boldsymbol{\Gamma}_i^g$ denotes the Tikhonov regularization matrix and is defined as:

$$\boldsymbol{\Gamma}_i^g = \begin{bmatrix} \|\boldsymbol{x}_i - \hat{\boldsymbol{x}}_1\|_2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \|\boldsymbol{x}_i - \hat{\boldsymbol{x}}_m\|_2 \end{bmatrix}. \qquad (7)$$

Similarly, taking the derivative w.r.t $\boldsymbol{\alpha}_i$ and setting the derivative to zero, we have

$$\hat{\boldsymbol{\alpha}}_i = (\boldsymbol{X}_m^T \boldsymbol{X}_m + \lambda \boldsymbol{\Gamma}_i^{gT} \boldsymbol{\Gamma}_i^g)^{-1} \boldsymbol{X}_m^T \boldsymbol{x}_i. \qquad (8)$$

The reconstruction error r_i can be computed by

$$r_i = \|\boldsymbol{x}_i - \boldsymbol{X}_m \hat{\boldsymbol{\alpha}}_i\|_2. \qquad (9)$$

If r_i is greater than a threshold, the pixel $\boldsymbol{x}_i$ is referred to as an anomaly. The total computational complexity of the Global PCAroCRD algorithm is $O(n^2 + ndm^2 + nm^3 + ndm + nd)$.

The Local PCAroCRD first selects the surrounding pixels for each pixel $\boldsymbol{x}_i$ by a sliding dual window, which is exactly the same as the CRD. Then, the spatial-domain PCA transformation is performed on the surrounding pixels $\boldsymbol{X}_s \in \mathbb{R}^{d \times s}$ in the spatial domain:

$$\bar{\boldsymbol{X}} = \boldsymbol{X}_s \boldsymbol{W}_l \qquad (10)$$

where $\check{\boldsymbol{X}} = [\check{\boldsymbol{x}}_1, \check{\boldsymbol{x}}_2, \cdots, \check{\boldsymbol{x}}_s] \in \mathbb{R}^{d \times s}$ denotes the transformed data matrix. The projection matrix $\boldsymbol{W}_l \in \mathbb{R}^{s \times s}$ is obtained by solving the PCA model:

$$\max_{\boldsymbol{W}_l^T \boldsymbol{W}_l = \boldsymbol{I}} \operatorname{tr}(\boldsymbol{W}_l^T \boldsymbol{X}_s^T \boldsymbol{X}_s \boldsymbol{W}_l). \qquad (11)$$

For the pixel $\boldsymbol{x}_i$, the objective function of the Local PCAroCRD can be written as follows [41]

$$\min_{\boldsymbol{\alpha}_i} \|\boldsymbol{x}_i - \boldsymbol{X}_k \boldsymbol{\alpha}_i\|_2^2 + \lambda \|\boldsymbol{\Gamma}_i^l \boldsymbol{\alpha}_i\|_2^2. \qquad (12)$$

where $\boldsymbol{X}_k = [\check{\boldsymbol{x}}_1, \check{\boldsymbol{x}}_2, \cdots, \check{\boldsymbol{x}}_k] \in \mathbb{R}^{d \times k}$ denotes the first k principal components of $\bar{\boldsymbol{X}}$ and contains the most information of $\boldsymbol{X}_s$ in the spatial domain. $\boldsymbol{\Gamma}_i^l$ denotes the Tikhonov regularization matrix and is defined as:

$$\boldsymbol{\Gamma}_i^l = \begin{bmatrix} \|\boldsymbol{x}_i - \bar{\boldsymbol{x}}_1\|_2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \|\boldsymbol{x}_i - \bar{\boldsymbol{x}}_k\|_2 \end{bmatrix}. \qquad (13)$$

Similarly, taking the derivative w.r.t $\boldsymbol{\alpha}_i$ and setting the derivative to zero, we have

$$\hat{\boldsymbol{\alpha}}_i = (\boldsymbol{X}_k^T \boldsymbol{X}_k + \lambda \boldsymbol{\Gamma}_i^{lT} \boldsymbol{\Gamma}_i^l)^{-1} \boldsymbol{X}_k^T \boldsymbol{x}_i. \qquad (14)$$

The reconstruction error r_i can be computed by

$$r_i = \|\boldsymbol{x}_i - \boldsymbol{X}_k \hat{\boldsymbol{\alpha}}_i\|_2. \qquad (15)$$

If r_i is greater than a threshold, then the pixel $\boldsymbol{x}_i$ is referred to as an anomaly. The total computational complexity of the Local PCAroCRD algorithm is $O(ns^2 + ndk^2 + nk^3 + ndk + nd)$.

III. THE PROPOSED METHOD

For the HAD applications, computational efficiency is an important factor to assess the detection performance of the detector. To reduce the computational complexity of CRD, a novel collaborative representation-based detector using the random background modeling and ensemble learning is proposed in this section.
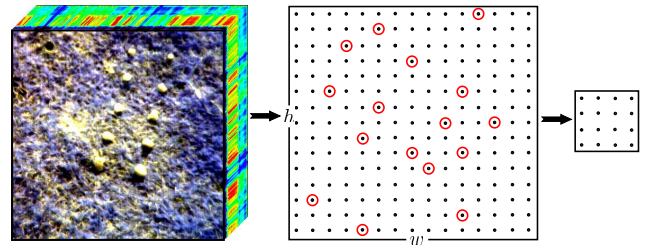

Fig. 2: Random background modeling.

A. Random CRD

The main idea of the CRD hypothesize that the background pixel can be approximated by a linear combination of the background pixels, but the anomalous pixel cannot. In the CRD, the background pixels for each pixel is represented by the surrounding pixels, which are selected by a sliding dual window centered at this pixel.

Different from the sliding dual window in the CRD, we select the background pixels for each pixel by the random background modeling, which is completed by random sampling from the whole hyperspectral image scene, as shown in Fig. 2.

Then, the background pixels obtained by the random sampling are resized into a matrix $\boldsymbol{X}_r = [\tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_2, \cdots, \tilde{\boldsymbol{x}}_r] \in \mathbb{R}^{d \times r}$. In the CRD, every pixel has its own surrounding pixels, and the surrounding pixels of each pixel are different. Unlike the CRD, we use the same matrix $\boldsymbol{X}_r$ for each pixel in our model. Therefore, the objective function of the proposed Random CRD is written as

$$\min_{\boldsymbol{A}} \|\boldsymbol{X} - \boldsymbol{X}_r\boldsymbol{A}\|_F^2 + \lambda\|\boldsymbol{A}\|_F^2, \tag{16}$$

where $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ denotes the weight matrix and λ denotes the regularization parameter. Taking the derivative w.r.t $\boldsymbol{A}$ and setting the derivative to zero, we have the following equation:

$$\boldsymbol{A} = (\boldsymbol{X}_r^T\boldsymbol{X}_r + \lambda\boldsymbol{I})^{-1}\boldsymbol{X}_r^T\boldsymbol{X}. \tag{17}$$

Thus, the matrix $\boldsymbol{X}$ can be reconstructed by the matrix $\boldsymbol{X}_r\boldsymbol{A}$. The reconstruction error for the pixel $\boldsymbol{x}_i$ is regarded as anomaly score and can be obtained by

$$\delta_i = \|\boldsymbol{x}_i - \boldsymbol{X}_r\boldsymbol{a}_i\|_2. \tag{18}$$

where $\boldsymbol{x}_i$ and $\boldsymbol{a}_i$ denote the i column of $\boldsymbol{X}$ and $\boldsymbol{A}$, respectively.

If δ_i is larger than a threshold, then the pixel $\boldsymbol{x}_i$ is called an anomaly. The computational complexity of the proposed Random CRD is $O(ndr + nd + dr^2 + r^3)$. The detailed process can be found in Algorithm 1.

Algorithm 1 Random CRD

Input: The two-dimensional HSI matrix $\boldsymbol{X}$, the number of random sampling r.
1. Randomly select r pixels from the matrix $\boldsymbol{X}$ and resize these pixels into the matrix $\boldsymbol{X}_r$.
2. Calculate the weight matrix $\boldsymbol{A}$ by Eq. (17).
3. Obtain the anomaly score δ_i for the pixel $\boldsymbol{x}_i$ by Eq. (18).
Output: The anomaly scores for all pixels.

B. Ensemble and Random CRD

As an anomaly detector ensemble that employs Random CRDs, the proposed Ensemble and Random CRD (ERCRD) has multiple Random CRDs acting as 'experts' to detect different anomalies. Through multiple Random CRDs, the anomaly score for the pixel x_i is obtained by

$$\gamma_i = \sum_{t=1}^{T} \delta_i^t, \tag{19}$$

where T denotes the ensemble size of the Random CRDs. Therefore, the total computational complexity of the proposed ERCRD is $O(ndrT + ndT + dr^2T + r^3T)$. The details of the ERCRD can be found in Algorithm 2.

Algorithm 2 Ensemble and Random CRD (ERCRD)

Input: The two-dimensional HSI matrix $\boldsymbol{X}$, the number of random sampling r and the ensemble size of the Random CRDs T.
1. Repeat the Random CRD (Algorithm 1) T times.
2. Obtain the anomaly score γ_i for the pixel $\boldsymbol{x}_i$ by Eq. (19).

Output: The anomaly scores for all pixels.

IV. EXPERIMENTAL RESULTS

To explore the detection performance of our ERCRD method, we conduct several experiments on a PC with E5-2680 v4 @2.40GHz and 256GB RAM, MATLAB 2016b. We use four hyperspectral datasets obtained from different scenes, which are described as follows:

1) AVIRIS-I Dataset: This dataset was acquired by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) from San Diego, CA, USA, with a spatial resolution of 3.5 m per pixel and a spectral resolution of 10 nm. This dataset has 224 spectral bands with wavelengths ranging from 370 to 2,510 nm. After removing the bands with water absorption, low signal-to-noise ratio, and poor-quality ($1 - 6$, $33 - 35$, 97, $107 - 113$, $153 - 166$, and $221 - 224$), 189 bands are retained in this experiment. The size of the entire image scene is 400×400 pixels, from which we select a 120×120 pixels area in the upper left corner to test and mark it as AVIRIS-I. The three airplanes in the image are considered to be anomalies, which consist of 58 pixels and should be detected.

2) AVIRIS-II Dataset: This dataset was derived from [46]. Same as the above dataset, we select a 100×100 pixels area at the center of the San Diego image to test and mark it as AVIRIS-II. The three airplanes consist of 134 pixels in the image are considered to be anomalies.

3) AVIRIS-III Dataset: This dataset was obtained from [56]. Again, we select a 200×240 pixels area in the upper left of the San Diego image to test and mark it as AVIRIS-III. The six airplanes consist of 90 pixels in the image are considered to be anomalies.

4) Cri Dataset: This dataset was derived from [34] and collected by the Nuance Cri hyperspectral sensor, with a spectral resolution of 10 nm. This dataset has a size of 400×400 pixels and 46 spectral bands with wavelengths ranging from 650 to 1,100 nm. The ten rocks consist of 2,216 pixels in the image are considered to be anomalies.

Note that Fig. 3a and 3b respectively present the false color image and the corresponding ground truth map of the AVIRIS-I dataset. In the same way, Fig. 3c and 3d correspond to the AVIRIS-II dataset, Fig. 3e and 3f correspond to the AVIRIS-III dataset, and Fig. 3g and 3h correspond to the Cri dataset.

In addition, the color detection map is used as the qualitative evaluation metric in our experiments. The receiver operating characteristic (ROC) curve [57], the area value under the ROC curve (AUC), the normalized background-anomaly separation map and the running time are used as quantitatively evalua-

tion metrics in our experiments. The ROC curve reflects the relations between the detection probability (DP) and the false alarm rate (FAR) at the thresholds ranging from 0 to 1 on the strength of ground truth. An excellent detector usually has a high DP value under the same FAR value, which leads to the phenomena that the corresponding ROC curve located close to the upper left corner, making the area under the curve larger. The value of the area enclosed by the ROC curve and the false alarm rate axis is called AUC. The normalized background-anomaly separation map describes the normalized anomaly score distributions of the background, and anomalous pixels are represented by box plot. Generally speaking, a good method should have a high AUC value and a distinct gap between the background box and the anomaly box.

A. Detection Performance

In this subsection, we carry out two experiments to verify the performance of our ERCRD method. Firstly, our ERCRD method is compared with five state-of-the-art methods. Since the proposed ERCRD method is a variant of the CRD, the CRD and four representative variants of the CRD are subsequently compared.

In the first experiment, in order to evaluate detection performance, we make a comparison between the proposed ERCRD method and five state-of-the-art methods: GRX [15], LRX [15], SSRX [18], CBAD [17], and LSMAD [34]. Among them, GRX is known as the benchmark anomaly detector for HSI. The LRX, SSRX and CBAD are three representative improved versions of RX. LSMAD is a typical low-rank and sparse matrix decomposition-based detector with remarkable detection performance. We choose the inner window size w_{in} ranging from 3 to 23 and the outer window size w_{out} ranging from 5 to 25 for the reason that the detection performance of LRX is sensitive to them. Moreover, we set the parameters of SSRX, CBAD and LSMAD to be accordant with earlier work [17], [18], [34].

The color detection maps of different methods based on AVIRIS-I dataset, AVIRIS-II dataset, AVIRIS-III dataset and Cri dataset are presented in Fig. 4, Fig. 6, Fig. 8, Fig. 10 respectively.

As for AVIRIS-I dataset, our ERCRD method is able to identify the locations of three airplanes, but fail to precisely picture the shapes of them. The GRX, LRX, SSRX, CBAD and LRaSMD methods not only fail to detect the anomalies, but also misidentify several normal background pixels as anomalies. Moreover, the ROC curves, the corresponding AUC values and the normalized background-anomaly separation maps are displayed in Fig. 5. It can be observed that the curve of the ERCRD method is closer to the upper left corner than the others and its AUC value is 0.9870, which is larger than others. We can see that the separation gap for the proposed ERCRD method is larger than those for the other methods. This indicates that the ERCRD method achieves the best separation result. Moreover, the LSMAD, SSRX and CBAD methods obtain relatively better separation capacity, while the GRX and LRX methods perform unsatisfactorily separation capacity.
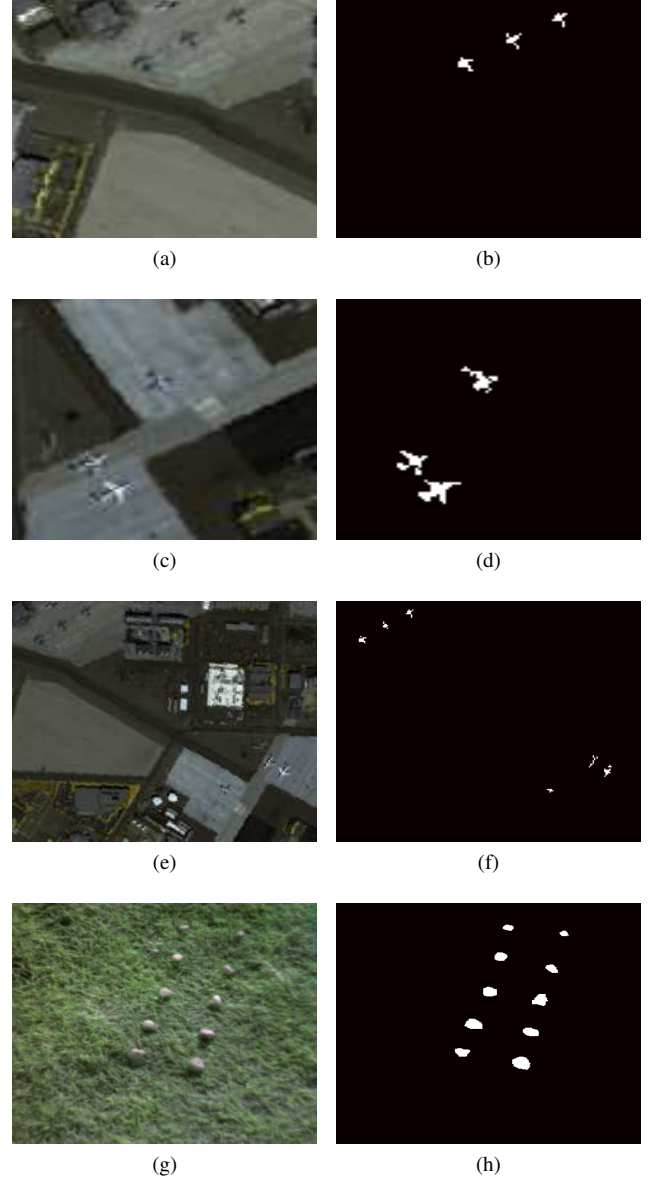

Fig. 3: Image scene descriptions. (a) False color image of AVIRIS-I dataset. (b) Ground truth map of AVIRIS-I dataset. (c) False color image of AVIRIS-II dataset. (d) Ground truth map of AVIRIS-II dataset. (e) False color image of AVIRIS-III dataset. (f) Ground truth map of AVIRIS-III dataset. (g) False color image of Cri dataset. (h) Ground truth map of Cri dataset.

As for AVIRIS-II dataset, the LRX and CBAD methods still can not separate the anomalies from the background and even misidentifies some normal pixels as anomalies. The GRX, SSRX, LSMAD and ERCRD methods can identify the locations of three airplanes, but the shapes of them are fuzzy and some false anomalies are also detected. The ROC curves, corresponding AUC values and normalized background-anomaly separation maps of different methods are given in Fig. 7. It can be concluded that the ROC curve of the ERCRD method is closer to the upper left corner than the others, and its
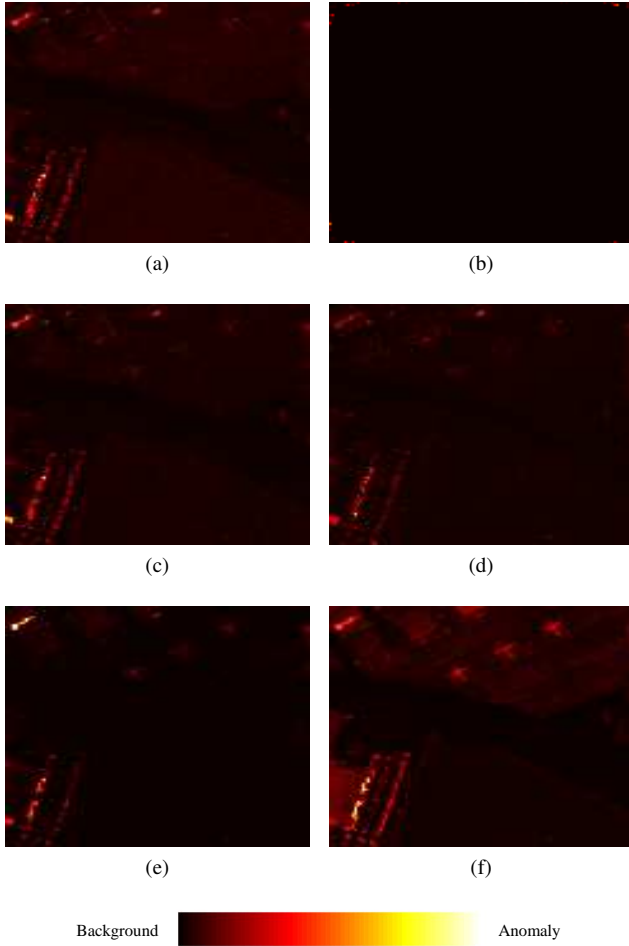
Fig. 4: Color detection maps obtained by different algorithms for AVIRIS-I dataset. (a) GRX. (b) LRX. (c) SSRX. (d) CBAD. (e) LSMAD. (f) ERCRD.

AUC value is 0.9793, which is also larger than others. Here, the proposed ERCRD method still achieves larger separation gaps while the separation capabilities of the GRX, SSRX, CBAD and LSMAD methods are slightly poorer. Compared with the above methods, the LRX method performs relatively unsatisfactorily.

As for AVIRIS-III dataset, the proposed ERCRD method is able to identify the locations of six airplanes but some anomalous pixels are missing and several normal pixels are misidentified. Unfortunately, other methods cannot detect anomalies effectively. In Fig. 9a, the ROC curves indicate that the proposed ERCRD method obtains a higher detection probability than others. The AUC values of all methods are illustrated in Fig. 9b; these values indicate that the proposed ERCRD method can achieve the best detection results among all the compared methods. Fig. 9c presents the separation maps for this dataset. Here, the proposed ERCRD method still achieves larger separation gaps. Moreover, the separation abilities of the other methods are greatly poorer, since their separation gaps are narrower than that of the ERCRD method.

As for Cri dataset, the SSRX, LSMAD and proposed ERCRD methods can effectively detect the locations and clear
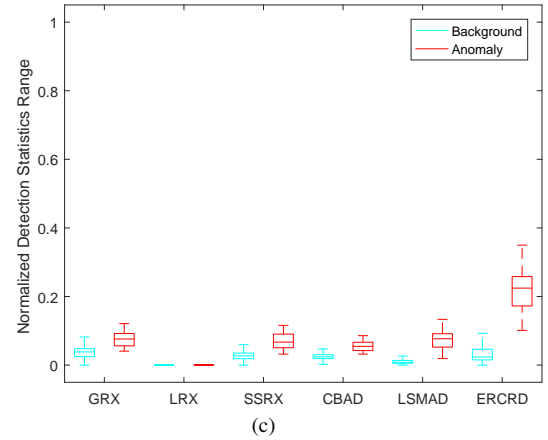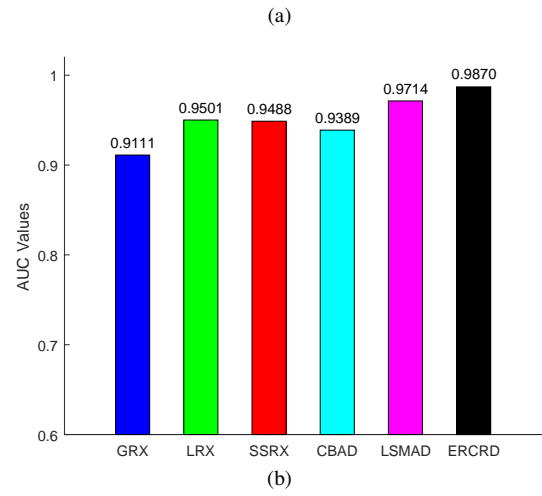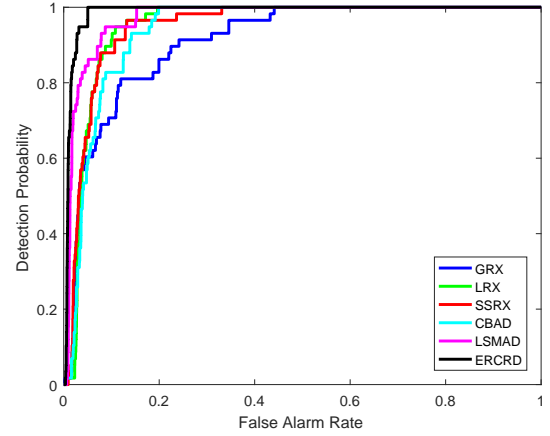

Fig. 5: Detection accuracy evaluation for AVIRIS-I dataset. (a) ROC curves. (b) AUC values. (c) Normalized background-anomaly separation map.

shapes of ten rocks, while the GRX, LRX and CBAD methods can only detect the positions of several anomalous pixels but the shapes of some are missing. It can be seen from Fig. 11 that the proposed ERCRD method performs high detection abilities with a low false alarm rate, high AUC value and achieves larger separation gaps than the other methods.

In addition, the running times for the four datasets are displayed in Table I. The GRX and SSRX methods are very

(a)   (b)

(c)   (d)
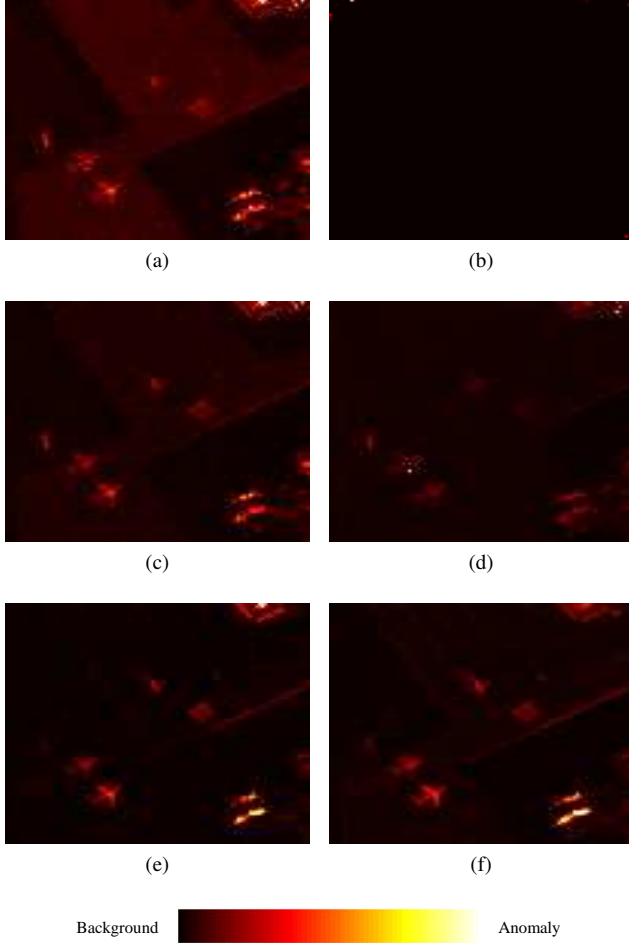
(e)   (f)

Background   Anomaly

Fig. 6: Color detection maps obtained by different algorithms for AVIRIS-II dataset. (a) GRX. (b) LRX. (c) SSRX. (d) CBAD. (e) LSMAD. (f) ERCRD.

fast compared to other methods. It is noteworthy that the running time of the proposed ERCRD and CBAD methods are similar to that of the GRX and SSRX methods; meanwhile, the proposed ERCRD method also achieves excellent detection performance. Thus, the proposed ERCRD method can be easily utilized in real-time applications. Furthermore, the LRX and LSMAD methods are more time-consuming than the other methods.

TABLE I: Running time (seconds)

Dataset	GRX	LRX	SSRX	CBAD	LSMAD	ERCRD
AVIRIS-I	0.27	140.75	0.21	0.41	20.28	0.77
AVIRIS-II	0.15	98.04	0.14	0.23	14.50	0.79
AVIRIS-III	0.52	443.55	0.67	3.56	73.99	2.45
Cri	0.79	192.86	0.98	2.76	55.01	1.82

In the second experiment, the detection performance of ER-CRD method is assessed and compared with the CRD [38] and four representative variants of CRD: Global PCAroCRD [41], Local PCAroCRD [41], MCRD [39], and RCRD [42]. It can be seen that the detection performance of the CRD, Local PCAroCRD, MCRD and RCRD methods are sensitive to the
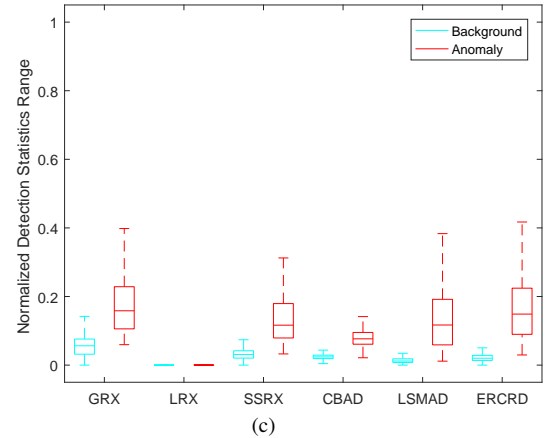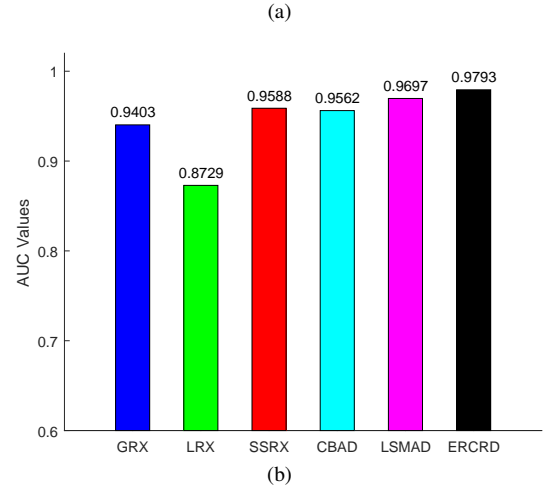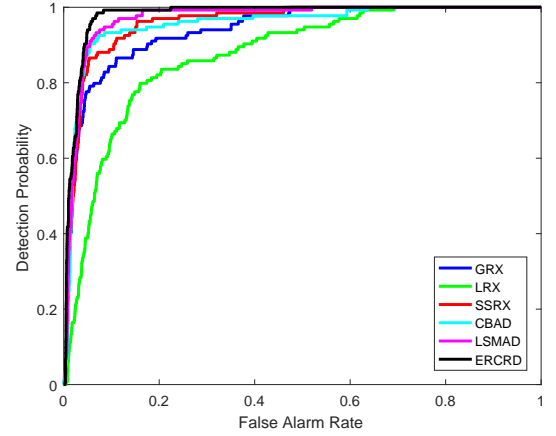

(a)

(b)

(c)

Fig. 7: Detection accuracy evaluation for AVIRIS-II dataset. (a) ROC curves. (b) AUC values. (c) Normalized background-anomaly separation map.

inner window size w_{in} and the outer window size w_{out}. Thus, we employed four window sizes: $(5, 9)$, $(7, 11)$, $(9, 13)$ and $(11, 15)$. The regularization parameter λ of these six methods is set to 10^{-6}. The AUC values and the corresponding running times of these six methods are displayed in Table II and Table III, respectively.

As for AVIRIS-I dataset, the AUC value of ERCRD method is 0.9870, only smaller than that of MCRD with window sizes
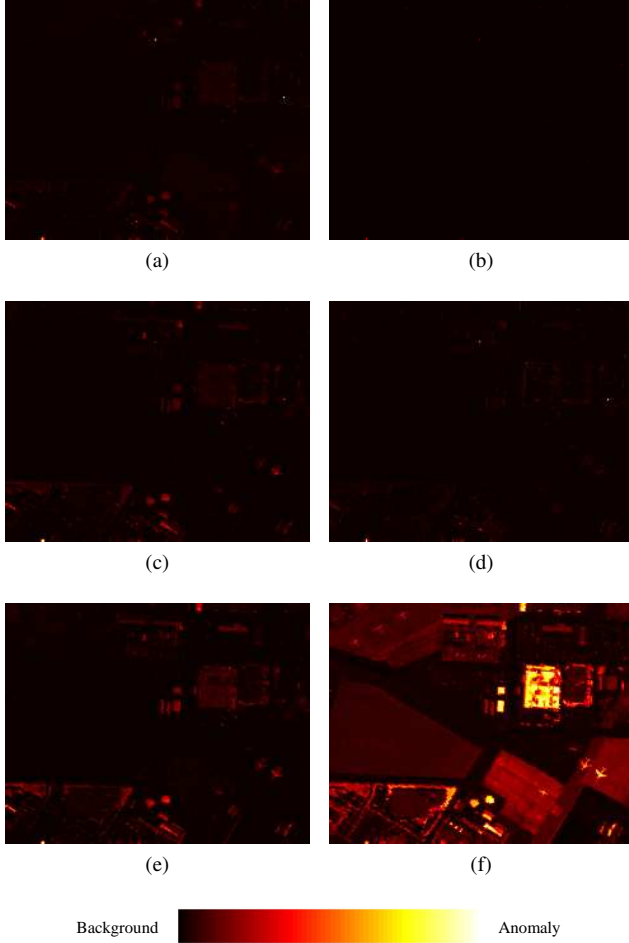
Fig. 8: Color detection maps obtained by different algorithms for AVIRIS-III dataset. (a) GRX. (b) LRX. (c) SSRX. (d) CBAD. (e) LSMAD. (f) ERCRD.

TABLE II: AUC values

Dataset		AVIRIS-I	AVIRIS-II	AVIRIS-III	Cri
CRD	$(5, 9)$	0.7116	0.9025	0.8393	0.8888
	$(7, 11)$	0.7835	0.8970	0.8880	0.9355
	$(9, 13)$	0.8842	0.9035	0.9388	0.9539
	$(11, 15)$	0.9493	0.9179	0.9659	0.9619
Local PCAroCRD	$(5, 9)$	0.8773	0.9517	0.8790	0.6106
	$(7, 11)$	0.9033	0.9455	0.9115	0.6416
	$(9, 13)$	0.9146	0.9468	0.9318	0.6726
	$(11, 15)$	0.9515	0.9590	0.9500	0.6990
MCRD	$(5, 9)$	0.8460	0.9068	0.8206	0.5561
	$(7, 11)$	0.9138	0.9162	0.8964	0.5463
	$(9, 13)$	0.9603	0.9262	0.9415	0.5393
	$(11, 15)$	0.9904	0.9495	0.9696	0.5396
RCRD	$(5, 9)$	0.7030	0.9016	0.8324	0.6361
	$(7, 11)$	0.7846	0.8963	0.8824	0.6261
	$(9, 13)$	0.8804	0.9051	0.9341	0.6113
	$(11, 15)$	0.9446	0.9197	0.9588	0.6322
Global PCAroCRD		0.9403	0.9259	0.9180	0.8183
ERCRD		0.9870	0.9793	0.9385	0.9844

Fig. 9: Detection accuracy evaluation for AVIRIS-III dataset. (a) ROC curves. (b) AUC values. (c) Normalized background-anomaly separation map.

$(11, 15)$ and larger than that of other methods. The running time of ERCRD is 0.77s, which is much lower than that of the other methods. The running time of MCRD with window sizes $(11, 15)$ is 167.34s, which is much higher than that of
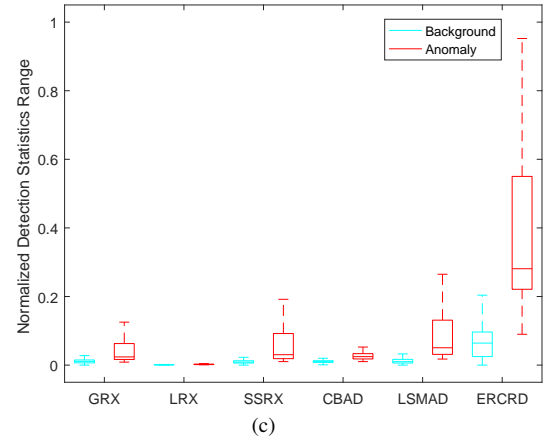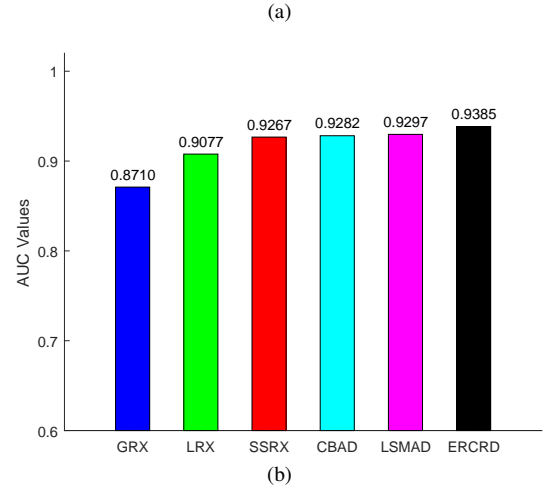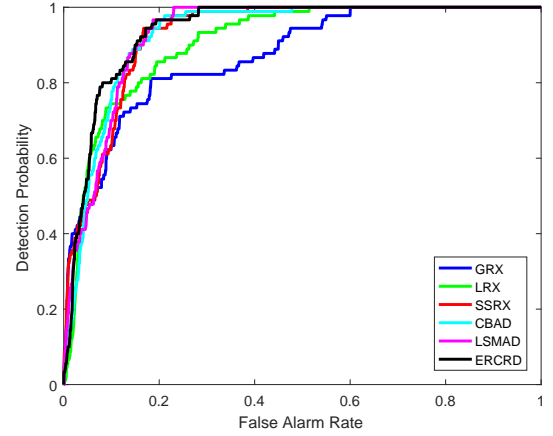
the ERCRD method. As for AVIRIS-II dataset, the AUC value of ERCRD is 0.9793, higher than that of other methods. The running time of ERCRD is 0.79s, which is much lower than that of the other methods. As for AVIRIS-III dataset, the AUC value of ERCRD is 0.9385, smaller than that of CRD with window sizes $(9, 13)$ and $(11, 15)$, Local PCAroCRD with window sizes $(11, 15)$, MCRD with window sizes $(9, 13)$ and $(11, 15)$, RCRD with window sizes $(11, 15)$. Note that the

(a)

(b)
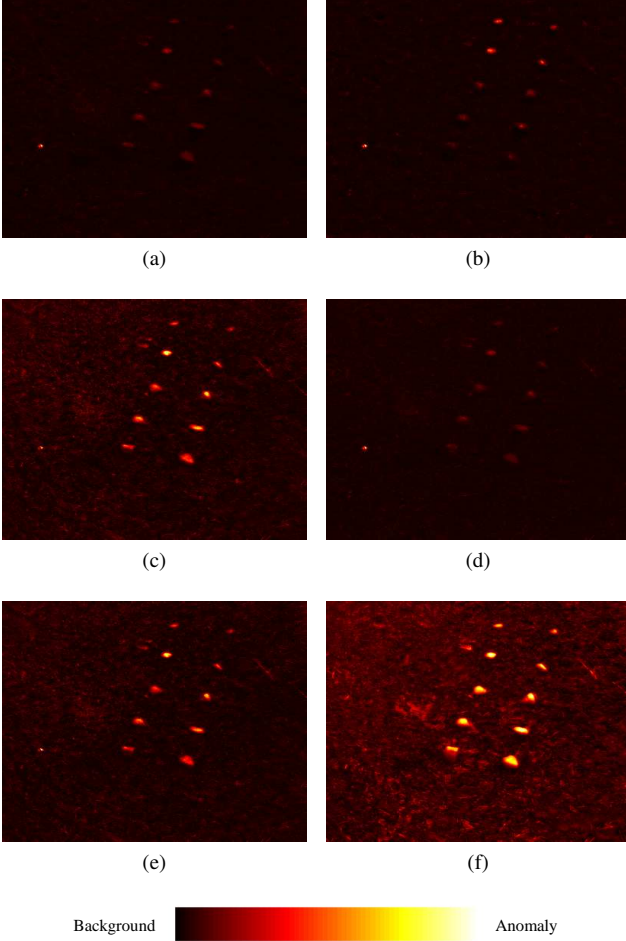

(c)

(d)

(e)

(f)

Background   Anomaly

Fig. 10: Color detection maps obtained by different algorithms for Cri dataset. (a) GRX. (b) LRX. (c) SSRX. (d) CBAD. (e) LSMAD. (f) ERCRD.

TABLE III: Running time (seconds)

Dataset		AVIRIS-I	AVIRIS-II	AVIRIS-III	Cri
CRD	$(5, 9)$	19.31	13.11	63.92	186.66
	$(7, 11)$	25.12	18.53	88.05	264.28
	$(9, 13)$	33.21	23.43	119.41	360.93
	$(11, 15)$	44.66	31.01	153.33	434.43
Local PCAroCRD	$(5, 9)$	11.39	7.72	39.79	101.31
	$(7, 11)$	15.61	10.48	52.32	131.17
	$(9, 13)$	20.97	13.89	70.89	186.62
	$(11, 15)$	26.82	18.12	91.24	214.92
MCRD	$(5, 9)$	58.45	52.14	249.23	816.61
	$(7, 11)$	82.64	72.11	311.20	1195.25
	$(9, 13)$	115.99	91.41	463.19	1667.27
	$(11, 15)$	167.34	114.62	568.85	2051.54
RCRD	$(5, 9)$	12.55	8.43	42.65	99.89
	$(7, 11)$	15.88	10.56	51.18	123.59
	$(9, 13)$	17.47	11.65	56.84	145.64
	$(11, 15)$	23.72	16.15	79.23	224.16
Global PCAroCRD		52.90	37.37	115.89	809.39
ERCRD		0.77	0.79	2.45	1.82

running time of ERCRD method is 2.45s, which is much lower than others. As for Cri dataset, the AUC value of ERCRD method is 0.9844, higher than that of other methods. The running time of ERCRD method is 1.82s, which is much lower
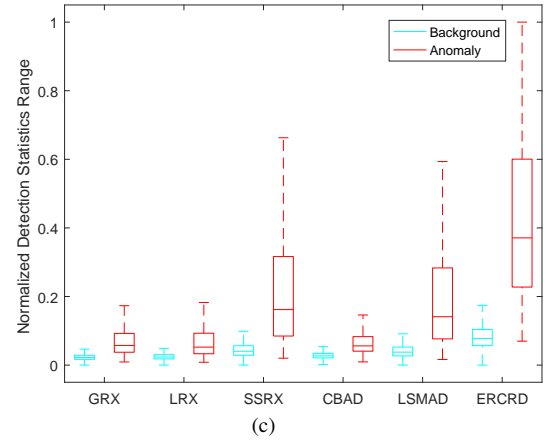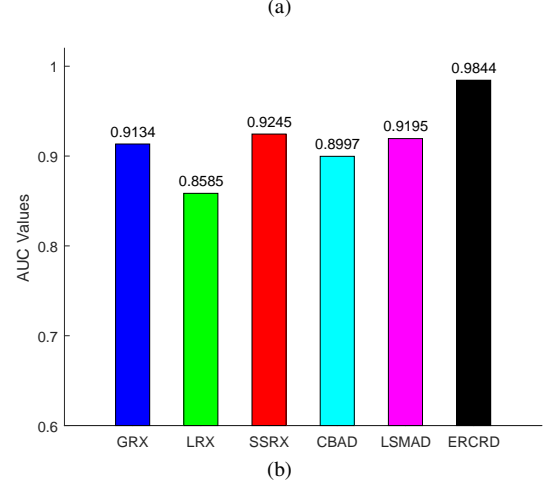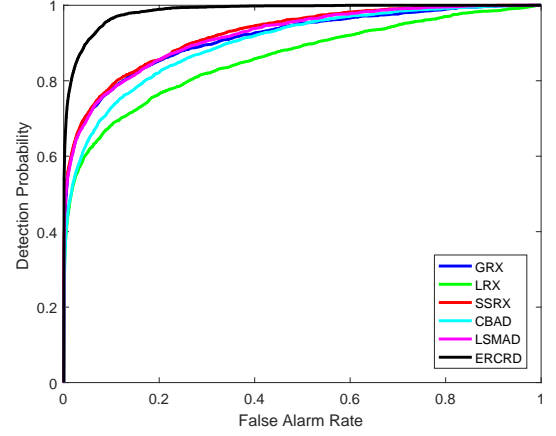

(a)

(b)

(c)

Fig. 11: Detection accuracy evaluation for Cri dataset. (a) ROC curves. (b) AUC values. (c) Normalized background-anomaly separation map.

than that of the other methods.

B. Parameter Analysis and Discussion

There are two parameters in the proposed ERCRD method: the number of random sampling r and the ensemble size T. We will conduct parameter analysis by comparing the AUC value and running time on four datasets.

Fig. 12 presents the impact of different random sampling number r on the detection performance and running time for each dataset. The value of the random sampling number r is varied in the range $[1, 2, \cdots, 20]$ and the ensemble size T is set to 10. It can be seen from Fig. 12a that the AUC values on AVIRIS-I and AVIRIS-II datasets fluctuate in a small range from 0.97 to 0.99, while on AVIRIS-III dataset it fluctuate in a slightly larger range from 0.91 to 0.95, and within a relatively larger range from 0.93 to 0.99 in Cri dataset. It is known from Fig. 12b that the running time on each dataset is nearly stable while the random sampling number r increasing.

Then the impact of different ensemble size T on the four datasets is analyzed. The value of ensemble size T is changed within the range of $[5, 10, \cdots, 50]$ and the number of random sampling r is set to 5. In Fig. 13a, we see that the AUC values on AVIRIS-I and AVIRIS-II datasets are nearly stable, while on AVIRIS-III and Cri datasets are increasing at first, and then fluctuate on a small scale. Fig. 13b illustrates that with the ensemble size T increasing, the running time of each dataset increases almost linearly.

V. CONCLUSION

In this paper, we came up with a novel ensemble and random collaborative representation-based detector (ERCRD) for hyperspectral imagery by exploring random background modeling and ensemble learning. Fistly, the random CRD is proposed by the random background modeling. The computational complexity of random CRD is greatly reduced compared to that of the original CRD. Then, multiple random CRDs act as 'experts' to detect different anomalies by using ensemble learning. Experimental results on four real hyperspectral datasets validate that our method outperforms its counterparts in the aspect of detection accuracy and running time.

REFERENCES

[1] G. Shaw and D. Manolakis, "Signal processing for hyperspectral image exploitation," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 12–16, Jan. 2002.

[2] D. W. J. Stein, S. G. Beaven, L. E. Hoff, E. M. Winter, A. P. Schaum, and A. D. Stocker, "Anomaly detection from hyperspectral imagery," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 58–69, Jan. 2002.

[3] C.-I. Chang and S.-S. Chiang, "Anomaly detection and classification for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 6, pp. 1314–1325, Jun. 2002.

[4] L. Zhang, Q. Zhang, B. Du, X. Huang, Y. Y. Tang, and D. Tao, "Simultaneous spectral-spatial feature selection and extraction for hyperspectral images," *IEEE Trans. Cybern.*, vol. 48, no. 1, pp. 16–28, 2018.

[5] F. Luo, B. Du, L. Zhang, L. Zhang, and D. Tao, "Feature learning using spatial-spectral hypergraph discriminant analysis for hyperspectral image," *IEEE Trans. Cybern.*, vol. 49, no. 7, pp. 2406–2419, 2019.

[6] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "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.

[7] R. Wang, F. Nie, and W. Yu, "Fast spectral clustering with anchor graph for large hyperspectral images," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 11, pp. 2003–2007, Nov. 2017.

[8] R. Wang, F. Nie, Z. Wang, F. He, and X. Li, "Scalable graph-based clustering with nonnegative relaxation for large hyperspectral image," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 7352–7364, Oct. 2019.

[9] 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.

[10] Y. Yuan, D. Ma, and Q. Wang, "Hyperspectral anomaly detection by graph pixel selection," *IEEE Trans. Cybern.*, vol. 46, no. 12, pp. 3123–3134, 2016.

[11] L. Li, W. Li, Q. Du, and R. Tao, "Low-rank and sparse decomposition with mixture of gaussian for hyperspectral anomaly detection," *IEEE Trans. Cybern.*, pp. 1–10, 2020.

[12] N. M. Nasrabadi, "Hyperspectral target detection: An overview of current and future challenges," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 34–44, Jan. 2014.

[13] M. Huber-Lerner, O. Hadar, S. R. Rotman, and R. Huber-Shalem, "Hyperspectral band selection for anomaly detection: The role of data gaussianity," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 2, pp. 732–743, Feb. 2016.

[14] L. Wang, C. Chang, L. Lee, Y. Wang, B. Xue, M. Song, C. Yu, and S. Li, "Band subset selection for anomaly detection in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 4887–4898, Sep. 2017.

[15] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1760–1770, Oct. 1990.

[16] H. Kwon and N. M. Nasrabadi, "Kernel RX-algorithm: A nonlinear anomaly detector for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 2, pp. 388–397, Feb. 2005.

[17] M. J. Carlotto, "A cluster-based approach for detecting man-made objects and changes in imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 2, pp. 374–387, Feb. 2005.

[18] A. P. Schaum, "Hyperspectral anomaly detection beyond RX," in *Proc. SPIE*, vol. 6565, May 2007.

[19] J. M. Molero, E. M. Garzn, I. Garcła, and A. Plaza, "Analysis and optimizations of global and local versions of the RX algorithm for anomaly detection in hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 801–814, Apr. 2013.

[20] W. Liu and C. Chang, "Multiple-window anomaly detection for hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 644–658, Apr. 2013.

[21] Q. Guo, B. Zhang, Q. Ran, L. Gao, J. Li, and A. Plaza, "Weighted-RXD and linear filter-based RXD: Improving background statistics estimation for anomaly detection in hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2351–2366, Jun. 2014.

[22] M. Imani, "RX anomaly detector with rectified background," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 8, pp. 1313–1317, Aug. 2017.

[23] C. Zhao, X. Yao, and Y. Yan, "Modified kernel RX algorithm based on background purification and inverse-of-matrix-free calculation," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 4, pp. 544–548, Apr. 2017.

[24] C. Zhao and Y. Xi-Feng, "Fast real-time kernel RX algorithm based on cholesky decomposition," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 11, pp. 1760–1764, Nov. 2018.

[25] J. Zhou, C. Kwan, B. Ayhan, and M. T. Eismann, "A novel cluster kernel RX algorithm for anomaly and change detection using hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 11, pp. 6497–6504, Nov. 2016.

[26] S. Chang, B. Du, and L. Zhang, "BASO: A background-anomaly component projection and separation optimized filter for anomaly detection in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 7, pp. 3747–3761, Jul. 2018.

[27] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Sparse representation for target detection in hyperspectral imagery," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 629–640, Jun. 2011.

[28] J. Li, H. Zhang, L. Zhang, and L. Ma, "Hyperspectral anomaly detection by the use of background joint sparse representation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2523–2533, Jun. 2015.

[29] Y. Zhang, B. Du, Y. Zhang, and L. Zhang, "Spatially adaptive sparse representation for target detection in hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 11, pp. 1923–1927, Nov. 2017.

[30] R. Zhao, B. Du, and L. Zhang, "Hyperspectral anomaly detection via a sparsity score estimation framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 6, pp. 3208–3222, Jun. 2017.

[31] F. Li, X. Zhang, L. Zhang, D. Jiang, and Y. Zhang, "Exploiting structured sparsity for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 7, pp. 4050–4064, Jul. 2018.

[32] Q. Ling, Y. Guo, Z. Lin, and W. An, "A constrained sparse representation model for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2358–2371, Apr. 2019.
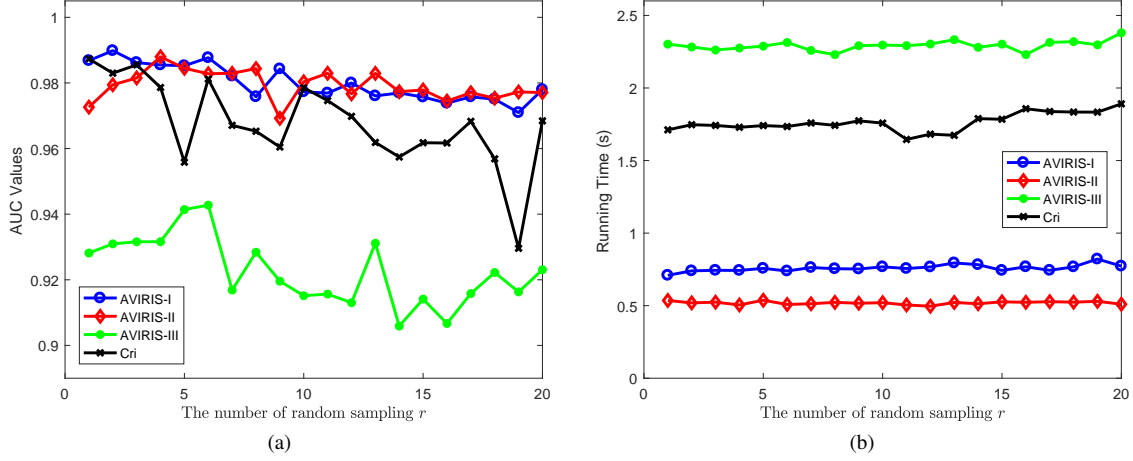
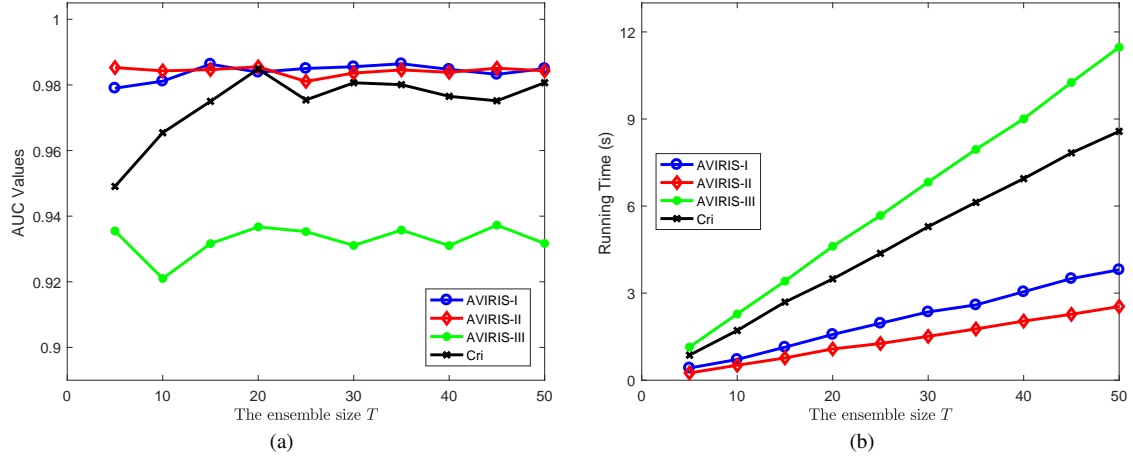Fig. 12: Effect of the number of random sampling r on each dataset. (a) AUC values. (b) Running time.

Fig. 13: Effect of the ensemble size T on each dataset. (a) AUC values. (b) Running time.

[33] W. Sun, C. Liu, J. Li, Y. M. Lai, and W. Li, "Low-rank and sparse matrix decomposition-based anomaly detection for hyperspectral imagery," *J. Appl. Remote Sens.*, vol. 8, no. 1, pp. 1–18, 2014.

[34] Y. Zhang, B. Du, L. Zhang, and S. Wang, "A low-rank and sparse matrix decomposition-based mahalanobis distance method for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1376–1389, Mar. 2016.

[35] Y. Xu, Z. Wu, J. Li, A. Plaza, and Z. Wei, "Anomaly detection in hyperspectral images based on low-rank and sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 1990–2000, Apr. 2016.

[36] Y. Qu, W. Wang, R. Guo, B. Ayhan, C. Kwan, S. Vance, and H. Qi, "Hyperspectral anomaly detection through spectral unmixing and dictionary-based low-rank decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4391–4405, Aug. 2018.

[37] B. Madathil and S. N. George, "Simultaneous reconstruction and anomaly detection of subsampled hyperspectral images using $l_{1/2}$ regularized joint sparse and low-rank recovery," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 5190–5197, Jul. 2019.

[38] W. Li and Q. Du, "Collaborative representation for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1463–1474, Mar. 2015.

[39] M. Imani, "Anomaly detection using morphology-based collaborative representation in hyperspectral imagery," *European Journal of Remote Sensing*, vol. 51, no. 1, pp. 457–471, 2018.

[40] M. Vafadar and H. Ghassemian, "Anomaly detection of hyperspectral imagery using modified collaborative representation," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 4, pp. 577–581, Apr. 2018.

[41] H. Su, Z. Wu, Q. Du, and P. Du, "Hyperspectral anomaly detection using collaborative representation with outlier removal," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 12, pp. 5029–5038, Dec. 2018.

[42] N. Ma, Y. Peng, and S. Wang, "A fast recursive collaboration representation anomaly detector for hyperspectral image," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 4, pp. 588–592, Apr. 2019.

[43] T. Cheng and B. Wang, "Graph and total variation regularized low-rank representation for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 391–406, Jan. 2020.

[44] A. Banerjee, P. Burlina, and C. Diehl, "A support vector method for anomaly detection in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 8, pp. 2282–2291, Aug. 2006.

[45] W. Sakla, A. Chan, J. Ji, and A. Sakla, "An SVDD-based algorithm for target detection in hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 2, pp. 384–388, Mar. 2011.

[46] X. Kang, X. Zhang, S. Li, K. Li, J. Li, and J. A. Benediktsson, "Hyperspectral anomaly detection with attribute and edge-preserving filters," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 10, pp. 5600–5611, Oct. 2017.

[47] S. Li, K. Zhang, Q. Hao, P. Duan, and X. Kang, "Hyperspectral anomaly detection with multiscale attribute and edge-preserving filters," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 10, pp. 1605–1609, Oct. 2018.

[48] A. Taghipour and H. Ghassemian, "Hyperspectral anomaly detection using attribute profiles," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 7, pp. 1136–1140, Jul. 2017.

[49] X. Zhang, G. Wen, and W. Dai, "A tensor decomposition-based anomaly detection algorithm for hyperspectral image," *IEEE Trans. Geosci.*

Remote Sens., vol. 54, no. 10, pp. 5801–5820, Oct. 2016.

[50] Y. Xu, Z. Wu, J. Chanussot, and Z. Wei, "Joint reconstruction and anomaly detection from compressive hyperspectral images using mahalanobis distance-regularized tensor rpca," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 5, pp. 2919–2930, May. 2018.

[51] X. Zhang and G. Wen, "A fast and adaptive method for determining k_1, k_2, and k_3 in the tensor decomposition-based anomaly detection algorithm," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 1, pp. 3–7, Jan. 2018.

[52] W. Xie, T. Jiang, Y. Li, X. Jia, and J. Lei, "Structure tensor and guided filtering-based algorithm for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4218–4230, Jul. 2019.

[53] W. Li, G. Wu, and Q. Du, "Transferred deep learning for anomaly detection in hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 5, pp. 597–601, May. 2017.

[54] C. Zhao, X. Li, and H. Zhu, "Hyperspectral anomaly detection based on stacked denoising autoencoders," *J. Appl. Remote Sens.*, vol. 11, no. 4, p. 042605, 2017.

[55] N. Ma, Y. Peng, S. Wang, and L. Phw, "An unsupervised deep hyperspectral anomaly detector," *Sensors*, vol. 18, no. 3, p. 693, 2018.

[56] Y. Zhang, B. Du, and L. Zhang, "A sparse representation-based binary hypothesis model for target detection in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1346–1354, Mar. 2015.

[57] J. Kerekes, "Receiver operating characteristic curve confidence intervals and regions," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 251–255, Apr. 2008.