Saliency-Guided Collaborative-Competitive Representation for Hyperspectral Anomaly Detection

Yufan Yang, Hongjun Su, *Senior Member, IEEE*, Zhaoyue Wu, *Graduate Student Member, IEEE*, and Qian Du, *Fellow, IEEE*

***Abstract*—Hyperspectral anomaly detection based on representation learning has received much attention in recent years. Due to the lack of prior knowledge about anomalies, it is difficult for a collaborative representation (CR) model to obtain a pure dictionary in the ideal case. Some algorithms proposed to remove anomalous pixels from a dictionary, which may result in the removal of contributing background atoms. To address such a problem, this article introduces a competitive regularization constraint term into the CR model, and divides the dictionary into anomaly and background classes using an outlier searching strategy, while adding competition weights to improve the competitiveness of the background. To better reconstruct the pixels, the Jaccard similarity coefficient is combined with the distance-weighted regularization matrix to adjust the representation coefficients. In addition, to make the most of the information from the hyperspectral data, a significance mechanism is introduced to construct an anomaly saliency weight to achieve the purpose of suppressing the background and highlighting anomalies. Experiments on five real datasets show that the detection performance of the proposed method is better than other advanced algorithms.**

***Index Terms*—Anomaly saliency weight, collaborative representation (CR), competitive regularization constraint term, hyperspectral anomaly detection.**

I. Introduction

THE key difference between hyperspectral and traditional multispectral images lies in narrow-band imaging. A hyperspectral image is a three-dimensional data cube, which provides two-dimensional spatial information and a third densely sampled spectral band information [1]. This rich spectral information can reflect the basic attributes of objects and provide the potential for diagnosis and recognition of ground objects [2], [3]. Therefore, hyperspectral data has incomparable advantages in target detection [4], ground object classification [5], [6], environmental monitoring, and other fields [7].

Considering the merits of hyperspectral data and the practicability of target detection, the research of this article is focused on hyperspectral target detection. According to whether the target is a priori, hyperspectral target detection can be achieved by supervised [8] or unsupervised [9] methods. Unsupervised target detection, namely anomaly detection, does not consider specific unknown targets but finds targets with different characteristics from the surrounding neighbors or the environment [10]. Since no prior knowledge is required, hyperspectral anomaly detection (HAD) has been greatly applied in civil search and rescue [11], national defense security [12], medical inspection [13], and food health detection [14].

Among the existing algorithms for HAD, the most representative ones are those based on statistical models and those based on representation learning. In addition, anomaly detection methods based on deep learning have become a research hotspot in recent years. Reed-Xiaoli (RX) [15] algorithm is regarded as the most classical algorithm in statistical models. It assumes background pixels obey multivariate Gaussian distribution. The RX algorithm includes the global RX algorithm, which treats the entire hyperspectral image as the background, and the local RX algorithm (LRX) [16], which uses the local data in a dual-window as the background model. The key point of hyperspectral anomaly detection is the modeling of the background. Therefore, many anomaly detection methods will consider the use of global and local information. Wu et al. proposed an anomaly detection method via global and local joint modeling of the background [17]. The RX algorithm is simple and powerful, but not all background images obey the Gaussian distribution, especially scenes with complex feature distribution. Therefore, some improved variants were proposed. In [18], a rigid dual-window construction dictionary method is developed in favor of adaptive determination of the detection window for LRX using super-pixel segmentation. Kwon et al. proposed the kernel RX algorithm [19] based on the kernel theory, which solves the nonlinear anomaly-background separation problem by mapping the original data into a higher dimensional feature space. However, this algorithm has high complexity. To reduce the computational cost, an anomaly detection algorithm based

Manuscript received 24 May 2023; revised 29 June 2023; accepted 10 July 2023. Date of publication 19 July 2023; date of current version 31 July 2023. This work was supported by the National Natural Science Foundation of China under Grant 42122008 and Grant 41871220. *(Corresponding author: Hongjun Su.)*

Yufan Yang and Hongjun Su are with the School of Earth Sciences and Engineering, Hohai University, Nanjing 211100, China, and also with the Jiangsu Province Engineering Research Center of Water Resources and Environment Assessment Using Remote Sensing, Hohai University, Nanjing 211100, China (e-mail: yufanyang0623@163.com; hjsu@hhu.edu.cn).

Zhaoyue Wu is with the Department of Technology of Computer and Communication, University of Extremadura, E-10071 Cáceres, Spain (e-mail: zhaoyue_wu@163.com).

Qian Du is with the Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS 39762 USA (e-mail: du@ece.msstate.edu).

Digital Object Identifier 10.1109/JSTARS.2023.3296876

on clustering and kernel method [20] was proposed, which not only performs well in detection performance but also has high computational efficiency. An improved RX technique based on subspace (named SSRX) that can accurately detect the target and lower the false detection rate is presented in [21]. Another reliable anomaly detector that is commonly mentioned is the blocked adaptive highly computationally efficient outlier nomenclature [22], which effectively discovers anomalies via blocking approaches and adaptive parameter change. minimum covariance determinant RX [23] which uses multiple adaptively sized windows to estimate the local background covariance matrix, an algorithm with high anomaly detection performance at different signal-to-noise ratios.

Although the principle of HAD algorithm based on statistics is simple and computational cost is relatively low, it does not make full use of the deep relationship between hyperspectral pixels, and does not consider the characteristics of spectral redundancy. To avoid the drawbacks of these algorithms, representation-based approaches without any statistical assumptions have drawn wide attention in recent years [24], the representation-based learning approaches have proven to be powerful tools for HSI processing and are widely used in different HSI domains [25]. According to different constrained conditions, the algorithm of HAD based on representation can be further divided into collaborative representation (CR), sparse representation (SR), and low-rank representation (LRR). SR uses either an l_1 norm or an l_{21} norm [26], [27], [28]. In [29], anomaly detection is achieved by calculating the sparsity divergence lift of pixels in each band. In background joint sparse-representation detector [30], a background dictionary representing local background information is used to achieve robust background estimation. In [31], a maximum likelihood estimation-based joint sparse representation model is applied to the classification, which brings new thinking to sparse representation-based anomaly detection methods. The LRR model describes the global characteristics of hyperspectral images, and assumes the background has low-rank property. The background and anomalies of hyperspectral images can thus be distinguished by using the low-rank theory. Typical methods include robust principal component analysis (RPCA) [32], anomaly detection based on low-rank sparse matrix decomposition [33], [34]. Since the low-rank representation does not well consider the local structure of each pixel, Xu et al. proposed a hyperspectral anomaly detection method based on the low-rank sparse representation [35], which achieves an accurate description of the pixel structure by adding sparse induced regularization terms.

While the importance of sparsity has been emphasized in many studies, it is pointed out in [36] that the collaborative nature of the dictionary atoms is more important, and that CR imposes l_2 parametric constraints on the regularization parameters. The first application of CR to hyperspectral anomaly detection was made by Li and Du [37]. The anomaly detection algorithm based on CR considers that a background pixel can be reconstructed by the background dictionary but not anomaly pixels, and the reconstructed representation residuals are considered as the final anomaly detection results. Although the CR model is simple and fast, and does not require any prior assumptions, it suffers from the vulnerability of the background dictionary to anomaly contamination and under-utilization of features. To address these drawbacks of CR model, many scholars have proposed variants for improvements. In [38], a hyperspectral collaborative representation anomaly detection algorithm for anomaly removal is proposed, which uses statistical theorems to set upper and lower bounds for thresholds, and pixels outside the threshold bounds are considered as anomalies for removal. On top of the outlier removal strategy, Vafadar and Ghassemian [39] uses both spectral and spatial information to detect anomalies. Su et al. [40] adopt spatial domain principal components analysis (PCA) to extract the main pixel information of the background as a sample for the CR model, while removing the anomalous pixel information from the background. A density peak clustering algorithm was used in the construction of the background dictionary in [41], and atoms with high density were considered as background dictionaries. In [42], the representation coefficients are ranked based on an outlier strategy, and the few dictionary atoms with small representation coefficients are considered outliers and removed. At the same time, a saliency weight is constructed using the atoms in the inner window to weight the testing pixels, but this saliency weight does not consider the fact that for scenes with complex features, the background may also be salient and significant, resulting in a poor generalization performance of the algorithm. These algorithms focus on the construction of a complete and pure background dictionary.

As an important branch and frontier of machine learning, deep learning-based methods have been well applied in hyperspectral remote sensing classification [43], and have brought new ideas to anomaly detection. Deep learning frameworks have been divided into supervised and unsupervised ones, typical unsupervised deep learning networks such as autoencoders [44] and generative adversarial networks [45] have been successfully applied to hyperspectral anomaly detection. The representative of supervised deep learning is typically a convolutional neural network (CNN), which is used in [46] for hyperspectral anomaly detection. Moreover, Song et al. proposed a HAD method based on CNN and LRR in [47], which combines representation learning with deep learning methods.

Although the statistical model-based hyperspectral anomaly detection algorithm is computationally efficient, it is limited by the assumption of a single statistical distribution model, which can lead to serious false alarm rates in scenes with complex background distributions. There are no assumptions needed for the anomaly detection approach based on the representation model, which has a simple and effective model, but suffers from the problem that the background dictionary is susceptible to anomalous pixels and feature utilization needs to be improved. Deep learning is highly capable of learning, but most deep learning-based methods need to be retrained before they can be applied to new test scenarios, and a single deep learning network for anomaly detection is not as effective and needs to be mined for more image information. Therefore, the combination of deep learning and traditional machine learning will be a trend. This article is developed based on the method of representation learning.

We learned that while collaboration between dictionary atoms is important, such collaboration is only between background atoms, and once there is an anomaly in a dictionary atom, then it introduces competition between dictionary atoms. Moreover, the lack of prior knowledge of anomalies makes it difficult for the CR detector, and its improved algorithm for impure dictionaries to achieve the ideal situation where the background dictionaries are not contaminated. The background atoms that contribute to the testing pixel may also be removed as anomalies, affecting the process of reconstruction. In addition, in the CR model, the dictionary atoms constructed by the dual-window ignore the global information, making the model itself incomplete for feature extraction of hyperspectral images. To address the above problems, this article introduces a competitive regularization constraint term in the CR model. Meanwhile, the saliency mechanism is also introduced into anomaly detection, and anomaly saliency weight is proposed.

In summary, this article proposes a saliency-guided collaborative-competitive representation (SG-CCR) method to identify anomalous pixels. Instead of selecting purified atoms, the method divides the dictionary atoms within the dual-window into a background class and an anomaly class, introducing a competitively constrained regularization term. The competition of background pixels is also improved to better reconstruct the pixels to be measured, and the initial reconstruction error is obtained. Then anomaly saliency weight is added to the initial reconstruction error to obtain the final anomaly detection results. The following summarizes the major contributions of this article.

1) A competitively constrained regularization term is added to the CR model. Due to the lack of prior information on anomalies, it is difficult to obtain a pure background dictionary in an ideal state. Considering the competitive relationship between anomalies and background in dictionary atoms, a competitive constraint is introduced into CR model to achieve the purpose of better reconstruction of pixels.
2) A spectral Jaccard similarity coefficient is added to the distance-weighted Tikhonov regularization. As the spectral Jaccard similarity coefficient considers the similarity of the two curves from the curve trend, it can be used to adjust the weight vector to better reconstruct the testing pixels.
3) Spatial and spectral information are used to construct anomaly saliency weights. The construction of this weight takes into account the characteristics of the human visual system when recognizing targets, and therefore introduces the saliency mechanism into anomaly detection in order to highlight anomalies.

The rest of the article are arranged as follows. The related work is presented in Section II. Section III proposes the SG-CCR anomaly detection method. Section IV introduces the experimental dataset, evaluation metrics, comparison algorithms, parameter settings, and experimental results. Section V conducts the ablation study, then discusses the validity of the proposed algorithm, the parameter sensitivity analysis, and the window stability analysis. Finally, Section VI concludes this article.

II. Related Work

A. Collaborative Representation Detector

The CR model believes that the collaboration between atoms is important. It assumes that each normal background pixel can be well represented by surrounding pixels, while abnormal pixels cannot. The model can be expressed as follows:

$$\underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \|\boldsymbol{y}-\boldsymbol{X}_{\boldsymbol{s}}\boldsymbol{\alpha}\|_2^2+\lambda\|\boldsymbol{\alpha}\|_2^2 \tag{1}$$

where $\boldsymbol{X}_{\boldsymbol{s}} \in R^{b\times s}$ is the dictionary taken from dual-window centered on the testing pixel $\boldsymbol{y} \in R^{b\times 1}$, b is the number of bands, s is the number of samples, and $\boldsymbol{\alpha}$ and λ are the coefficient vector and Lagrange multiplier, respectively. The solution for $\boldsymbol{\alpha}$ is

$$\boldsymbol{\alpha}=\left(\boldsymbol{X}_{\boldsymbol{s}}^T\boldsymbol{X}_{\boldsymbol{s}}+\lambda\boldsymbol{I}\right)^{-1}\boldsymbol{X}_{\boldsymbol{s}}^T\boldsymbol{y}. \tag{2}$$

However, for the testing pixel, not every dictionary atom used for reconstructs is similar to it. Therefore, a larger reconstruction coefficient should be given to a dictionary atom that is very similar to the testing pixel, while a smaller coefficient should be given to dissimilar atoms. Taking this into account, Tikhonov regularization with distance weighting $\boldsymbol{\Gamma}_y$ is usually adopted to adjust the weight vector. Then, the CRD model expression becomes

$$\underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \|\boldsymbol{y}-\boldsymbol{X}_{\boldsymbol{s}}\widehat{\boldsymbol{\alpha}}\|_2^2+\lambda\|\boldsymbol{\Gamma}_y\widehat{\boldsymbol{\alpha}}\|_2^2 \tag{3}$$

where $\boldsymbol{\Gamma}_y$ is a diagonal regularization matrix to the dictionary atoms and the testing pixel $\boldsymbol{y}$

$$\boldsymbol{\Gamma}_y=\begin{bmatrix}\|\boldsymbol{y}-\boldsymbol{x}_1\|_2 & & 0\\ & \ddots & \\ 0 & & \|\boldsymbol{y}-\boldsymbol{x}_{\boldsymbol{s}}\|_2\end{bmatrix}. \tag{4}$$

$\boldsymbol{x}_{\boldsymbol{s}} \in R^{b\times 1}$ is the sth dictionary pixel, and the solution to $\widehat{\boldsymbol{\alpha}}$ is

$$\widehat{\boldsymbol{\alpha}}=\left(\boldsymbol{X}_{\boldsymbol{s}}^T\boldsymbol{X}_{\boldsymbol{s}}+\lambda\boldsymbol{\Gamma}_y^T\boldsymbol{\Gamma}_y\right)^{-1}\boldsymbol{X}_{\boldsymbol{s}}^T\boldsymbol{y}. \tag{5}$$

In CRD, the reconstruction error r is used to determine the anomaly

$$r=\|\boldsymbol{y}-\boldsymbol{X}_{\boldsymbol{s}}\hat{\boldsymbol{\alpha}}\|_2. \tag{6}$$

If it is larger than a certain threshold, then y is declared as an anomalous pixel

The selection of background dictionary of CRD model is easily affected by abnormal pixels, which affects the accuracy of detection.

B. RX-Detector

The RX algorithm is one of the most classical algorithms for hyperspectral anomaly detection. For hyperspectral images, the data matrix is defined as $\boldsymbol{X} \in R^{b\times n}$($n$ is the total number of pixels). Each spectral pixel consists of b spectral bands, denoted as: $x(n)=(x_1(n),x_2(n)\ldots,x_b(n))^T$. According to the principle of generalized likelihood ratio detection, H_0 is set as that the testing pixel is a background pixel, and H_1 is that the pixel is an abnormal pixel. Then, the binary hypothesis of the

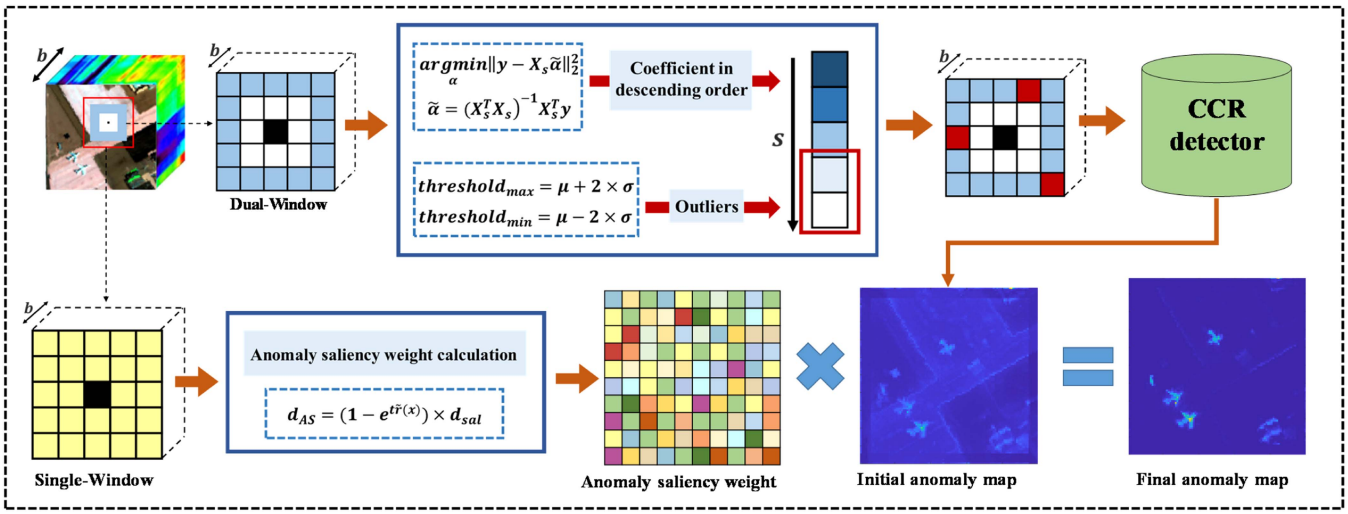

Fig. 1. Flowchart of the proposed SG-CCR method. First, the background dictionary and the anomaly dictionary are constructed by using the outlier strategy, and the obtained dictionary is put into the CCR model to obtain the initial anomaly result. Then, the anomaly significance weight is constructed, multiplied by the initial anomaly result, and the final anomaly result graph is obtained.

RX algorithm is as follows:

$$H_0 : \boldsymbol{x} = \boldsymbol{n} \ \text{(Target absent)} \tag{7}$$

$$H_1 : \boldsymbol{x} = \boldsymbol{s} + \boldsymbol{n} \ \text{(Target present)} \tag{8}$$

where $\boldsymbol{x}$ is the pixel to be measured, $\boldsymbol{n}$ is the background noise vector, and $\boldsymbol{s}$ is the target spectrum vector. It is assumed that the pixel follows multivariate normal distribution. The expression of RX algorithm is as follows:

$$RX\,(x) = (\boldsymbol{x} - \boldsymbol{u_b})^T \boldsymbol{C_b^{-1}} (\boldsymbol{x} - \boldsymbol{u_b}) \tag{9}$$

where $\boldsymbol{u_b}$ is the mean value of the background spectral vector, and $\boldsymbol{C_b}$ is the covariance matrix of the background spectral vector. Set η as the threshold. If $RX(x) > \eta$, then x is considered as an abnormal pixel; otherwise, it is a background pixel.

III. Proposed Method

To improve the reconstruction effect and achieve the purpose of suppressing background and highlighting anomalies, this article proposes a collaborative-competitive representation (CCR) hyperspectral anomaly detection method based on saliency guidance. The flowchart of the method is shown in Fig. 1, and the main steps are as follows.

1) Based on the CR model, a competitive regularization constraint term is introduced to build the CCR model.
2) Construct a dual-window dictionary for the pixel to be measured, and use the outlier strategy to classify the dictionary atoms with different contributions into two categories: background dictionary and anomaly dictionary.
3) Input the obtained dictionary into the CCR model, and obtain the initial anomaly detection residuals.
4) Construct anomaly saliency weight, multiply by initial anomaly detection residual, and obtain the final anomaly detection result.

A. Collaborative-Competitive Representation Detector

CCR is an extension of CR, which impose a competitive constraint on the CR model [48]. The CCR model is currently validated effectively in the field of object classification. In [49], the CCR model is used for anomaly detection by introducing local constraints. The most critical step in CR is the acquisition of the dictionary, but it is difficult to acquire a pure dictionary in an ideal state due to the lack of prior knowledge of anomaly detection. For anomaly detection, anomaly atoms and background atoms are antagonistic and they are competitively involved in the reconstruction. Once there is an anomaly in the dictionary, it will affect the reconstruction result. Therefore, this article considers the competition between background and anomaly in the dictionary, and introduces the competitive constraint into the CR model. Since there are usually fewer anomaly atoms, the background atoms are more competitive, meanwhile, we introduce competitive weight to improve the competition of background atoms in the reconstruction process, so the purpose of introducing the competition constraint term is to better reconstruct the test pixels and attenuate the influence of anomaly pixels on the reconstruction. The objective function of the CCR is

$$\underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{X_s}\boldsymbol{\alpha}\|_2^2 + \lambda \sum_{i=1}^{C} \|\boldsymbol{y} - \boldsymbol{X_i}\boldsymbol{\alpha_i}\|_2^2 + \beta \|\boldsymbol{\alpha}\|_2^2. \tag{10}$$

Here $\boldsymbol{X_s} \in R^{b \times s}$ is the dictionary matrix, and the dictionaries in this article are chosen from a dual-window centered on the pixel to be measured (the number of samples $s = w_{out} \times w_{out} - w_{in} \times w_{in}$, where w_{out} is the outer window size and w_{in} is the inner window atom), $\boldsymbol{X_i}$ is a dictionary of different kinds divided from $\boldsymbol{X_s}$. In (10), λ and β are regularization parameters. In anomaly detection, the dictionary includes only two classes, background and anomaly, so $C = 2$. According to the outlier strategy, the dictionary can be divided into background dictionary $\boldsymbol{X_{s_1}}$ and anomaly dictionary $\boldsymbol{X_{s_2}}$, whose corresponding

coefficient vectors are $\boldsymbol{\alpha}_{\boldsymbol{s}_1}$ and $\boldsymbol{\alpha}_{\boldsymbol{s}_2}$, respectively. Firstly, we calculate the mean value μ and standard deviation σ of dictionary $\boldsymbol{X}_{\boldsymbol{s}}$, and set the threshold to determine the number m_0 of outlier pixels. The formula of threshold construction is as follows:

$$\text{threshold}_{\max} = \mu + 2 \times \sigma \tag{11}$$

$$\text{threshold}_{\min} = \mu - 2 \times \sigma \tag{12}$$

where $\text{threshold}_{\max}$ and $\text{threshold}_{\min}$ denote the maximum and minimum intensity values respectively. Since the testing pixel can be represented linearly by the surrounding pixels, the resulting representation coefficients are the degree of contribution of the dictionary atoms to the testing pixel, so the magnitude of the representation coefficients is used as a basis for classifying background and anomalous pixels, and the vector of coefficients obtained using the least squares solution is

$$\underset{\tilde{\boldsymbol{\alpha}}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{X}_{\boldsymbol{s}}\tilde{\boldsymbol{\alpha}}\|_2^2 \tag{13}$$

$$\tilde{\boldsymbol{\alpha}} = \left(\boldsymbol{X}_{\boldsymbol{s}}^T \boldsymbol{X}_{\boldsymbol{s}}\right)^{-1} \boldsymbol{X}_{\boldsymbol{s}}^T \boldsymbol{y}. \tag{14}$$

The coefficient vectors are ranked by their contributions

$$|\tilde{\boldsymbol{\alpha}}_1| > |\tilde{\boldsymbol{\alpha}}_2| > \cdots > |\tilde{\boldsymbol{\alpha}}_s| \tag{15}$$

the m_0 outlier atoms corresponding to the representation coefficients with smaller contribution are considered as the anomaly dictionary $\boldsymbol{X}_{\boldsymbol{s}_2}$ and the rest are considered as the background dictionary $\boldsymbol{X}_{\boldsymbol{s}_1}$. Therefore, the objective function of CCR with the competitive constraint of anomaly and background is

$$\underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{X}_{\boldsymbol{s}}\boldsymbol{\alpha}\|_2^2 + \lambda \sum_{i=1}^{2} \|\boldsymbol{y} - \boldsymbol{X}_{\boldsymbol{s}_i}\boldsymbol{\alpha}_i\|_2^2 + \beta\|\boldsymbol{\alpha}\|_2^2. \tag{16}$$

Meanwhile, in order to improve the contribution of background, this article increases the competitive weight w_{-i} on the basis of CCR model inspired by [50], w_{-i} refers to the competitive weight of other training samples except ith class samples. The expression of w_{-i} is

$$w_{-i} = \exp\left(\frac{r_{\max} - r_{-i}}{\delta}\right) \tag{17}$$

where $\delta = 1$ and the expressions for the residuals r_{-1}, r_{-2} obtained using the anomaly dictionary and the background dictionary are

$$r_{-1} = \|\boldsymbol{y} - \boldsymbol{X}_{\boldsymbol{s}_2}\boldsymbol{\alpha}_{\boldsymbol{s}_2}\|_2 \tag{18}$$

$$r_{-2} = \|\boldsymbol{y} - \boldsymbol{X}_{\boldsymbol{s}_1}\boldsymbol{\alpha}_{\boldsymbol{s}_1}\|_2 \tag{19}$$

$$r_{\max} = \max\left(r_{-1},\ r_{-2}\right). \tag{20}$$

In general, the background dictionary $\boldsymbol{X}_{\boldsymbol{s}_1}$ has more contributions to the reconstruction process. Fig. 2 shows a portion of competitive weights for the testing pixels. It can be seen that w_{-2} is usually larger than w_{-1}, so adding the competitive weight w_{-i} can improve the contribution of the background dictionary. The

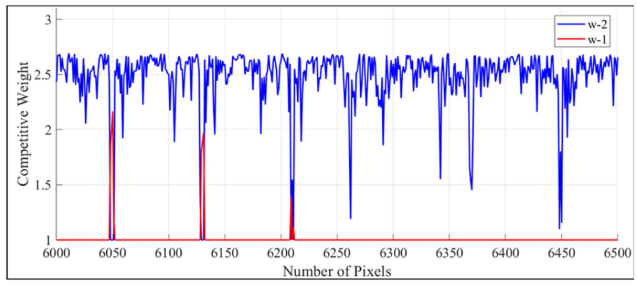

Fig. 2. Competitive weight of anomaly and background.

CCR objective function equation at this point becomes

$$\underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{X}_{\boldsymbol{s}}\boldsymbol{\alpha}\|_2^2 + \lambda \sum_{i=1}^{2} w_{-i}\|\boldsymbol{y} - \boldsymbol{X}_{\boldsymbol{s}_i}\boldsymbol{\alpha}_i\|_2^2 + \beta\|\boldsymbol{\alpha}\|_2^2. \tag{21}$$

The closed form solution of $\boldsymbol{\alpha}$ is

$$\boldsymbol{\alpha} = (1+\lambda)\left(\boldsymbol{X}_{\boldsymbol{s}}^T \boldsymbol{X}_{\boldsymbol{s}} + \lambda \boldsymbol{M} + \beta \boldsymbol{I}\right)^{-1} \boldsymbol{X}_{\boldsymbol{s}}^T \boldsymbol{y} \tag{22}$$

where $\boldsymbol{M}$ is

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{X}_{\boldsymbol{s}_1}^T \boldsymbol{X}_{\boldsymbol{s}_1} & 0 \\ 0 & \boldsymbol{X}_{\boldsymbol{s}_2}^T \boldsymbol{X}_{\boldsymbol{s}_2} \end{bmatrix}. \tag{23}$$

The anomaly and background vectors should be uncorrelated. Thus, their inner product is assigned a value of $\mathbf{0}$, which can also be reflected in the expression of $\boldsymbol{M}$.

B. Jaccard Similarity Coefficient

The CCR anomaly detector proposed above only improves the competitiveness of background pixels in the reconstruction process, but it should not be ignored that background dictionaries may be interfered by abnormal pixels. In some algorithms, different penalties are given by means of distance-weighted Tikhonov regularization parameters. We found that the trend of most of the curves for the background pixels were similar, while the trend of the abnormal curve is inconsistent with most of the background pixels. According to this characteristic, we add the Jaccard similarity coefficient to the distance-weighted regularization matrix, adjusting the contribution of each background pixel to achieve a better reconstruction of the testing pixel.

The spectral Jaccard similarity coefficient is added in this article [51], which is used to compare the degree of similarity between samples, with higher values representing greater similarity. The Jaccard coefficient is defined as follows:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}. \tag{24}$$

During most of the reconstruction process, if the pixel to be measured is the background, it is more similar to the spectral curve trend of the dictionary atom, and conversely, if the pixel to be measured is abnormal, it is not similar to the spectral curve trend of the dictionary atom. Therefore, in this article, we introduce the Jaccard index to represent the similarity between the curves and use the first order derivative to represent the

Algorithm 1. CCR Anomaly Detector

Input: Dual-Window Dictionary $\boldsymbol{X_s} \in R^{b\times s}$.

1. Use the outlier strategy of (11) and (12) to determine the number of outlier pixels in the dictionary;
2. Obtain the coefficient vector using (14) and rank them according to their contribution;
3. Divide the dictionary $\boldsymbol{X_s}$ into background class $\boldsymbol{X_{s_1}}$ and anomaly class $\boldsymbol{X_{s_2}}$ according to step 1 and 2;
4. Calculate the Jaccard similarity index using (25) to obtain the regularization parameter $\hat{\boldsymbol{\Gamma}}_y$;
5. Use (28) to obtain the representation coefficient α;
6. Calculate the reconstruction error of the testing pixel using (29) ;

Output: Initial anomaly detection result r^*

spectral trends. The Jaccard coefficient of similarity between the two spectral curves is then defined as follows:

$$JSC = \frac{B_{11} + B_{00}}{B}. \tag{25}$$

Here B_{11} is the band where the first order derivative of the two image spectral curves is greater than zero at the same time, that is, the band where the curve trend is up at the same time. B_{00} is the band where the first-order derivative of the two image spectral curves is less than or equal to zero simultaneously, that is, the band where the curve trend is down simultaneously. B represents the total number of bands. By adding the Jaccard similarity coefficient on the basis of Tikhonov regularization, the new regularization parameters can be obtained as follows:

$$\hat{\boldsymbol{\Gamma}}_y = \begin{bmatrix} \frac{1}{jsc_1}\|\boldsymbol{y} - \boldsymbol{x}_1\|_2 & & 0 \\ & \ddots & \\ 0 & & \frac{1}{jsc_s}\|\boldsymbol{y} - \boldsymbol{x_s}\|_2 \end{bmatrix} \tag{26}$$

where jsc_s is the Jaccard similarity coefficient for the dictionary pixel. We add the Jaccard similarity coefficient to the distance-weighted Tikhonov regularized CCR model to obtain the JCCR model. The objective function of the JCCR model is

$$\underset{\widehat{\boldsymbol{\alpha}}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{X_s}\widehat{\boldsymbol{\alpha}}\|_2^2 + \lambda \sum_{i=1}^{2} w_{-i}\|\boldsymbol{y} - \boldsymbol{X_{s_i}}\widehat{\boldsymbol{\alpha}}_i\|_2^2 + \beta\|\hat{\boldsymbol{\Gamma}}_y\widehat{\boldsymbol{\alpha}}\|_2^2. \tag{27}$$

The closed solution of $\widehat{\alpha}$ is

$$\widehat{\boldsymbol{\alpha}} = (1+\lambda)\left(\boldsymbol{X_s}^T\boldsymbol{X_s} + \lambda\boldsymbol{M} + \beta\hat{\boldsymbol{\Gamma}}_y^T\hat{\boldsymbol{\Gamma}}_y\right)^{-1}\boldsymbol{X_s}^T\boldsymbol{y} \tag{28}$$

After obtaining the representation coefficient, the final residual result is

$$r^* = \|\boldsymbol{y} - \boldsymbol{X_s}\widehat{\boldsymbol{\alpha}}\|_2 \tag{29}$$

where the residual r^* is the initial anomaly detection result. Algorithm 1 is the pseudocode of CCR anomaly detector proposed in this article.

C. Anomaly Saliency Weight

Inspired by papers [42] and [52], abnormal objects are regarded as significant objects in the visual system, the significance mechanism is introduced, and significance weight is represented by the ratio of spectral Angle distance and Euclidean distance between single-window pixels and the testing pixel, so as to distinguish anomalies and backgrounds. The formula of saliency weight is as follows:

$$D\ (\boldsymbol{x}_i, \boldsymbol{y}) = \frac{d_{\text{spectral}}\ (\boldsymbol{x}_i, \boldsymbol{y})}{1 + cd_{\text{position}}\ (\boldsymbol{x_i}, \boldsymbol{y})} \tag{30}$$

where d_{spectral} represents the spectral angular distance and d_{position} represents the Euclidean distance. The expression formula of spectral angular distance is

$$d_{\text{spectral}}\ (\boldsymbol{x}_i, \boldsymbol{y}) = cos^{-1}\left(\frac{\boldsymbol{x}_i^T\boldsymbol{y}}{\sqrt{\boldsymbol{x}_i^T\boldsymbol{x}_i}\sqrt{\boldsymbol{y}^T\boldsymbol{y}}}\right). \tag{31}$$

The testing pixel is marked as the center origin $(0, 0)$, and the position of $\boldsymbol{x}_i$ is marked as (p, q), then the expression of d_{position} is

$$d_{\text{position}}\ (\boldsymbol{x}_i, \boldsymbol{y}) = \sqrt{p^2 + q^2}. \tag{32}$$

Considering that hyperspectral images have multiple bands, the pixel values can be interpreted as different spectral scores for pixels between different bands, i.e., different bands generate different spectral values for the same pixel. This is not considered in simple cosine similarity, so the cosine similarity coefficient d_{spectral} in the significance weight is replaced by the adjusted cosine similarity degree d_{sim} as a way to address the shortcoming of considering only the similarity of spectral dimensions in cosine similarity without taking into account their differences. The adjusted cosine similarity is corrected by the operation of subtracting the mean value from each dimension

$$d_{\text{sim}}\ (\boldsymbol{x}_i, \boldsymbol{y}) = cos^{-1}\left(\frac{(\boldsymbol{x}_i - \bar{\boldsymbol{x}}_i)^T(\boldsymbol{y} - \bar{\boldsymbol{y}})^T}{\sqrt{(\boldsymbol{x}_i - \bar{\boldsymbol{x}}_i)^T(\boldsymbol{x}_i - \bar{\boldsymbol{x}}_i)}\sqrt{(\boldsymbol{y} - \bar{\boldsymbol{y}})^T(\boldsymbol{y} - \bar{\boldsymbol{y}})}}\right). \tag{33}$$

Therefore, the new significance weight expression is

$$\tilde{D}\ (\boldsymbol{x}_i, \boldsymbol{y}) = \frac{d_{\text{sim}}\ (\boldsymbol{x}_i, \boldsymbol{y})}{1 + cd_{\text{position}}\ (\boldsymbol{x}_i, \boldsymbol{y})} \tag{34}$$

where c usually takes the value of 1. Significance weights for test pixels are the average of all weights within a single window

$$d_{\text{sal}} = \frac{\sum_{i=1}^{s_0}\tilde{D}\ (\boldsymbol{x}_i, \boldsymbol{y})}{s_0} \tag{35}$$

where $s_0 = w_s \times w_s - 1$ and w_s is the single window size divided for the purpose of calculating significance weights.

Considering that in the image with relatively complex background, some backgrounds are also significant under visual effects, which will cause negative effects on anomaly detection. In order to highlight the anomalies and suppress the background, the anomaly saliency weight is constructed by multiplying this by the weight of the suppressed background. The expression for the anomaly saliency weight is

$$d_{AS} = (1 - e^{t\tilde{r}(x)}) \times d_{\text{sal}} \tag{36}$$

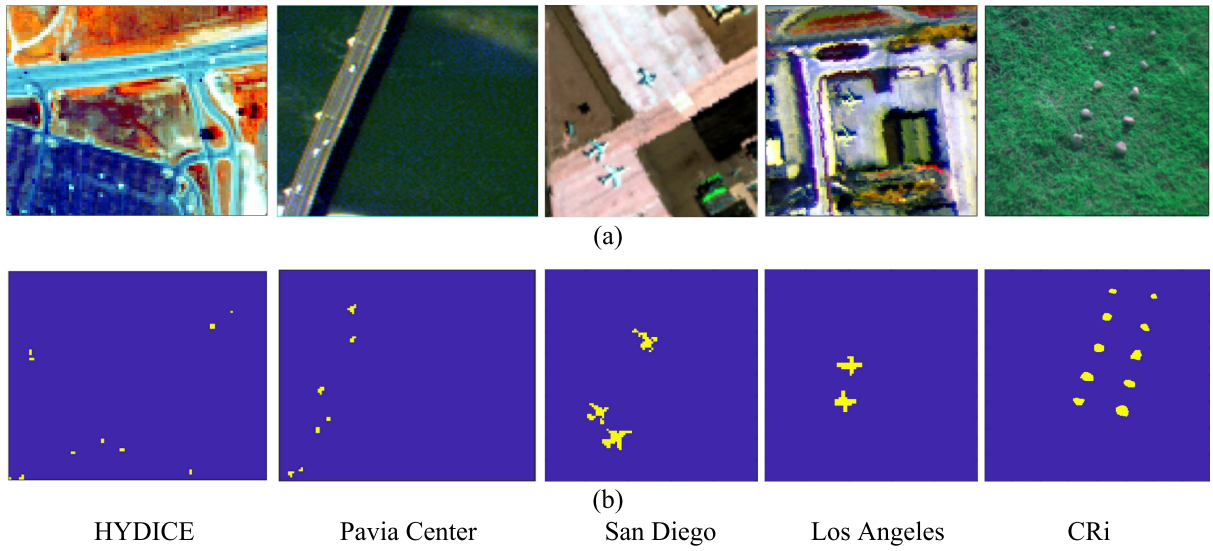

Fig. 3. (a) Pseudocolor images. (b) Reference anomaly maps.

Algorithm 2: Anomaly Saliency Weight.

Input: Single-Window Dictionary $\boldsymbol{X}_{\boldsymbol{s}_0} \in R^{b\times s_0}$.

1. Using Algorithm 1 to obtain the initial reconstruction error r^*;
2. Use (9) to obtain the RX distance;
3. Intercept M_0 pixels with the largest initial reconstruction error r^* and RX distance, and use (35) to obtain the set T.
4. Assign the RX distance of all pixels in the set T to the maximum RX distance, and leave the RX distance of the remaining pixels unchanged, and the adjusted RX distance $\tilde{r}(x)$ was obtained;
5. Use (34) to obtain the anomaly saliency weight

Output: Anomaly saliency weight d_{AS}

where the parameter t is the adjusted value of the function, and $\tilde{r}(x)$ is the adjusted global RX distance, and the adjusted RX distance is adopted because the RX distance of the anomaly may be smaller than that of the background. Therefore, it is very likely to suppress the anomaly while suppressing the background pixel, which will also bring negative effects to the anomaly detection. The specific operation is to sort the initial reconstruction error r^* and RX distance. Then the top M_0 pixels with the largest r^* and RX distances are selected to obtain the sets T_1 and T_2, and the sets T_1 and T_2 are intersected by the set T. The RX distances of the pixels in the set T are assigned to the maximum RX distance, and the RX distances of the remaining pixels remain unchanged, and the adjusted RX distance $\tilde{r}(x)$ was obtained. Anomaly saliency weight das is obtained by formula (34)

$$T = T_1 \cap T_2. \tag{37}$$

Algorithm 2 is the pseudocode for constructing abnormal significant weights. After obtaining the anomaly saliency weight, it is multiplied by the initial reconstruction error r^* to obtain the final anomaly detection result e.

$$e = r^* \times d_{AS}. \tag{38}$$

IV. Experiments

A. Experimental Datasets

Five datasets were used for this experiment, with pseudocolor images and anomaly reference maps as shown in Fig. 3.

1) The first dataset was acquired by the Hyperspectral Digital Image Collection Experiment (HYDICE) sensor [53] with a spatial resolution of 1.56 m and a spectral resolution of 10 nm. The image size is 80 × 100 pixels with 175 effective bands. The anomalous targets in the images are cars and rooftops, and these anomalous targets are made up of 21 pixels. This dataset has a complex background, with small and scattered anomaly targets that are not prominent enough from the global perspective.
2) The second dataset was acquired by the reflection optical system imaging spectrometer at the Pavia Centre with a spatial resolution of 1.3 m. The image size of 105 × 120 pixels was selected on this dataset and 102 bands were retained after removing the undesirable bands. The background consists mainly of bridges, rivers, and soil, with the anomaly being a vehicle on a bridge, consisting of 43 pixels. The background of this dataset is simple and the anomalies are located on a bridge, which affects the saliency of the anomaly target in the global view as the bridge is quite prominent.
3) The third dataset was acquired by the airborne visible/infrared imaging spectrometer (AVIRIS) sensor over the San Diego airport area, with a spatial resolution of 3.5 m and a spectral resolution of 10 nm. The image size is 100 × 100 pixels with 189 effective bands. The background consists mainly of tarmac and bare soil, and the anomaly is a scattered trio of aircraft, consisting of

134 pixels. The background of this dataset is relatively complex and the abnormal targets are large, which are more prominent in the global perspective.

4) The fourth dataset was acquired by the AVIRIS sensor over the Los Angeles Airport area in California, USA [54], with a spatial resolution of 7.1 m. The image size is 100 × 100, with 205 bands retained after removing undesirable bands. The background consists mainly of hangars and airport runways and aprons. The anomalies are the two aircraft in the center, consisting of 87 pixels. The background of this dataset is relatively complex. The anomalous targets are larger and more visible in the global view, where the two aircraft in the lower left are more prominent compared to the aircraft in the upper right.
5) The fifth dataset was acquired by the Nuance Cri sensor [34]. The spectral resolution is 10 nm. The image size is 400 × 400, and 46 bands are retained after removing undesirable bands. The background is mainly different types of grass and the anomalies are ten rocks, consisting of 2216 pixels. The background of this dataset is simple, the anomalous targets are relatively salient in the global view, especially those located in the center.

B. Evaluation Indicators

In this article, we adopt five evaluation metrics widely used in anomaly detection, which are the (ROC) curve [55], the area under the curve (AUC) [56], the box plot, the square error ratio (SER) [57], and the area error ratio (AER) [58].

ROC curve is the classical metric to measure detection performance, which describes the relationship between detection rate P_D and false alarm rate P_{FA}. AUC is the area under the ROC curve, which is used to measure the overall detection accuracy. The larger the AUC, the better the algorithm performance. The box plot depicts the degree of background-anomaly separation and background suppression. The larger the spacing between the two boxes, the greater the separation between the anomaly and the background, and if an algorithm has better background suppression effect, the shorter its background box is in the box plot.

The sum of the squares of the residuals between the actual and detected values of the anomalous and background pixels is the SER value. The anomalous pixel has a value of 1 and the background pixel has a value of 0 in the ideal case. The objective function of SER is

$$\mathrm{SER} = \frac{\sum_{i=1}^{n_{\mathrm{abomaly}}(GT)} (p_i - 1)^2 + \sum_{i=1}^{n_{\mathrm{bak}}(GT)} (p_j - 0)^2}{N_p} \times 100 \tag{39}$$

where p_i and p_j denote the detected values of anomalous and background pixels, respectively, $n_{\mathrm{abomaly}}(GT)$ and $n_{\mathrm{bak}}(GT)$ are the numbers of true anomalous pixels and true background pixels in the image respectively, and N_p is the number of total pixels. The smaller the SER, the more effective the anomaly detection.

The AER describes the separability of the anomalies from the background by analyzing the P_D and P_{FA} at different thresholds. Assuming that in a two-dimensional coordinate system, the y-axis represents the P_D and P_{FA} obtained at different thresholds, and the x -axis represents the normalized threshold. The AER index can then be described as follows:

$$\mathrm{AER} = \frac{1 - AP_{FA}}{1 - AP_D} \tag{40}$$

where AP_D and $\mathrm{A}P_{FA}$ are the areas under the P_D and P_{FA} curves, respectively. Contrary to the performance of SER on anomaly detection, the larger the AER value, the better the detection effect.

C. Comparison Methods and Parameter Settings

To verify the SG-CCR anomaly detection method proposed in this article, eight mature anomaly detection methods are used for comparison, which are the RX algorithm [15], LRX [16], low-rank and sparse matrix decomposition-based anomaly detection for hyperspectral imagery (LRaSMD) [33], CRD [37], hyperspectral anomaly detection with kernel isolation forest (KIFD) [59], multipixel anomaly detection with unknown patterns for hyperspectral imagery (2S-GLRT) [60], hyperspectral anomaly detection with robust graph autoencoders (RGAE) [61], hyperspectral anomaly detection with guided autoencoder (GAED) [62].

The parameters of the proposed anomaly detection algorithm include the outer window dictionary w_{out} the inner window dictionary w_{in}, the single-window dictionary w_s for constructing anomaly salient weights, the parameters λ and β for the CCR model. and the parameter M_0 for selecting the first M_0 pixels with the largest distance between r^* and RX for constructing anomaly saliency, and the adjustment parameter c for the function. The inner window w_{in} ranges from 3 to 17, the outer window w_{out} ranges from 5 to 25 and the single window w_s ranges from 3 to 25. The range of the parameters λ and β is set to {1e-6–1e0}. The parameters M_0 are chosen from the number of pixels from 0 to n. and the parameter $t \subseteq [1, 2, 4, 8]$. The selection of parameters M_0 and t did not have a significant effect on the results, so the parameters M_0 and t were usually fixed for the optimal results to analyze the effect of other parameters on the experiments. The dual-window range settings for LRX, CRD, and 2S-GLRT are consistent with the range settings of the algorithms proposed in this article. The parameters for LRaSMD, KIFD, RGAE, and GAED are set according to the suggested values in their articles. The best anomaly detection accuracy over the parameter variation range was chosen as the final result in both the experimental and comparison experiments.

D. Experimental Results

The anomaly detection results are shown in Fig. 4, the ROC curves and box plots are shown in Figs. 5 –9, and the AUC, AER, SER and running times for each method are shown in Table I. The optimal detection results are marked in thick type in the table and the sub-optimal detection results are underlined.

For the dataset HYDICE, the AUC of the proposed algorithm in this article is 0.9994, which is higher than all other compared algorithms, as evidenced by the ROC curve in Fig. 5(a). As can be seen from Fig. 4(a), the anomalies detected by the

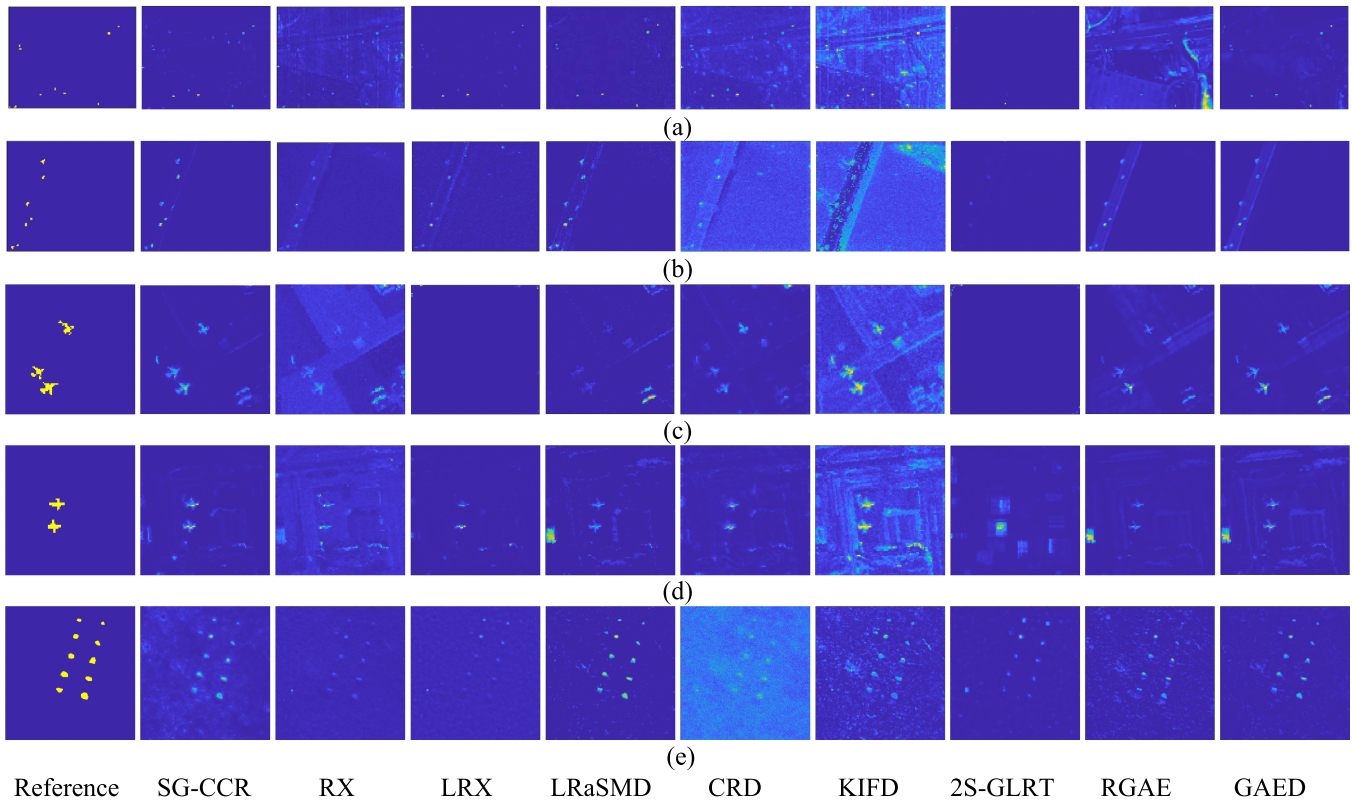

Fig. 4. Anomaly detection maps of SG-CCR, RX [14], LRX [15], LRaSMD [29], CRD [33], KIFD [50], 2S-GLRT [51], RGAE [52], and GAED [53] for the five datasets. (a) HYDICE. (b) Pavia center. (c) San Diego. (d) Los Angeles. (e) CRi.

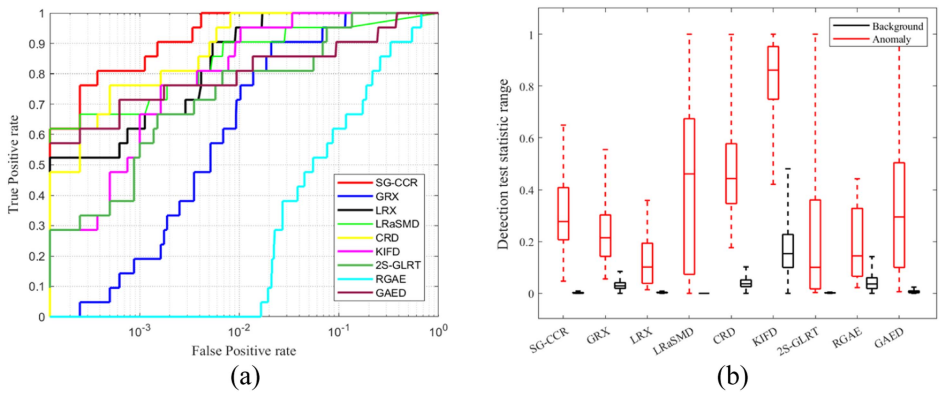

Fig. 5. ROC curves and box plots for several methods in the HYDICE dataset. (a) ROC curves. (b) Box plots.

SG-CCR are more prominent relative to those detected by other algorithms. The RX, CRD, KIFD, RGAE algorithms are less effective in suppressing the background than SG-CCR. the LRX, LRaSMD,2S-GLRT, and GAED algorithms are less effective in detecting and highlighting the anomalies than SG-CCR. The box plot in Fig. 5(b) shows the proposed SG-CCR algorithm has better background suppression, although the suppression is less effective than that of LRaSMD and 2S-GLRT, its separation of background and anomaly is much greater than that of these two algorithms, so that the anomaly is only relatively more prominent. Table I shows that the LRaSMD algorithm has the smallest SER value, SG-CCR has the next best SER value, CRD has the largest AER value, and SG-CCR is more effective in separating anomalies from the background.

For the dataset Paiva Center, it can be seen from the anomaly detection graph in Fig. 4(b) that the proposed algorithm in this article is better at highlighting anomalies than all other algorithms except LRaSMD, and the suppression of the background is the best, LRaSMD does not suppress the background as well as SG-CCR. As can be seen from Fig. 6(a), the ROC curves of the algorithms SG-CCR, RGAE, and GAED are more consistent, and their AUC values are also less different. The box plot in

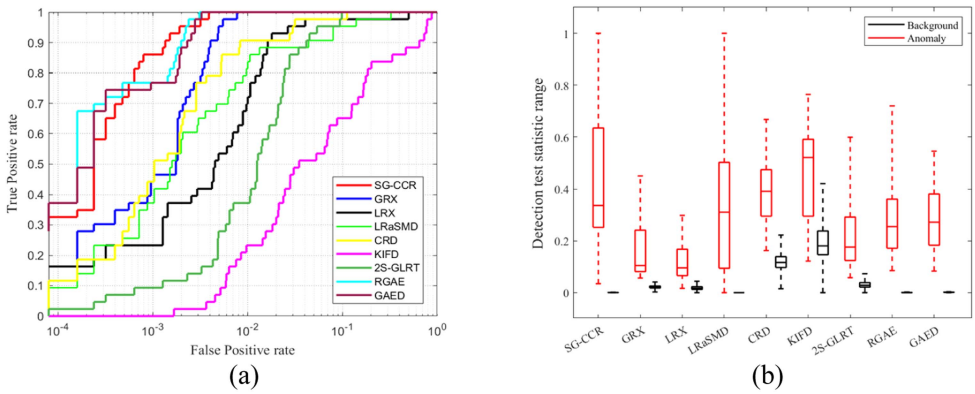

Fig. 6. ROC curves and box plots for several methods in the Pavia center dataset. (a) ROC curves. (b) Box plots.

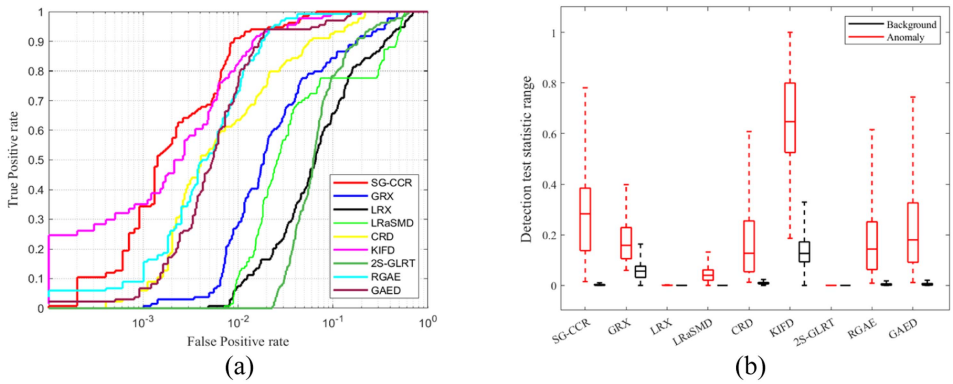

Fig. 7. ROC curves and box plots for several methods in the San Diego dataset. (a) ROC curves. (b) Box plots.

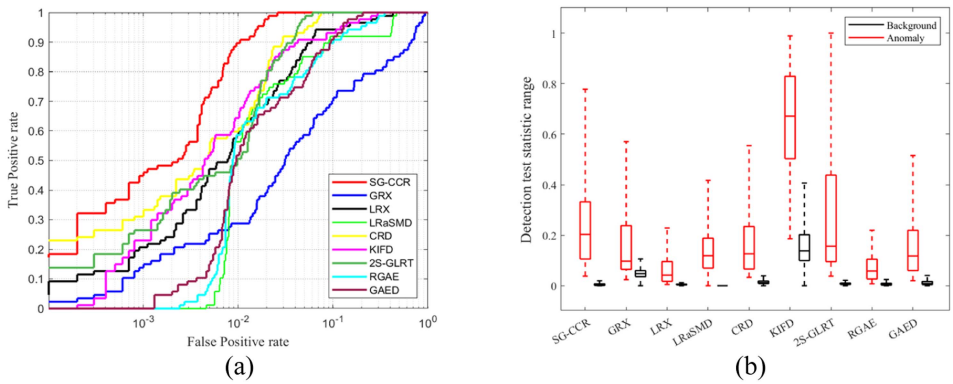

Fig. 8. ROC curves and box plots for several methods in the Los Angeles dataset. (a) ROC curves. (b) Box plots.

Fig. 6(b) shows that SG-CCR suppresses the background better and separates the background from the anomaly better than the other algorithms. Table I shows that the SER results for SG-CCR are optimal, the AER values are suboptimal, and the AER values for LRaSMD are optimal. These metrics show that the algorithm proposed in this article performs well on this dataset.

For the dataset San Diego, as shown in Fig. 4(c), the aircraft in the lower left corner of the anomaly detection graph of the SG-CCR algorithm is more conspicuous than the aircraft in the upper right corner. SG-CCR outperforms other algorithms in highlighting the anomalies and suppressing the background. And although the anomaly target of KIFD is conspicuous, the algorithm is poor at suppressing the background. The ROC curve in Fig. 7(a) shows that SG-CCR starts off less well than KIFD, but is overall optimal and has the largest AUC value. The box plot in Fig. 7(b) shows that SG-CCR is optimal for background suppression although not as good as LRX, LRaSMD, and 2S-GLRT. Although not as good as KIFD for separating the background from anomalies, SG-CCR is optimal overall. Table I shows that the SER values for SG-CCR are the smallest, the AER values for

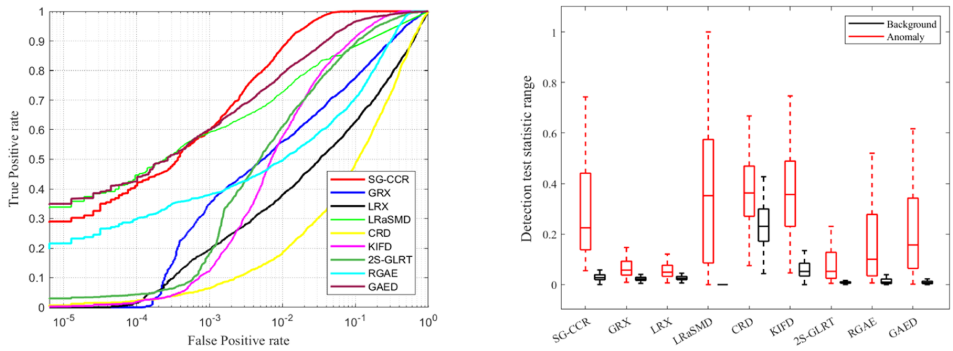

Fig. 9. ROC curves and box plots for several anomaly detection method in the CRi dataset. (a) ROC curves. (b) Box plots.

TABLE I
AUC, SER, AER, AND RUNNING TIME OF SEVERAL ALGORITHMS IN FIVE DATASETS

Datasets		SG-CCR	RX	LRX	LRaSMD	CRD	KIFD	2S-GLRT	RGAE	GAED
HYDICE	**AUC**	**0.9994**	0.9857	0.9976	0.9710	0.9987	0.9964	0.9829	0.8581	0.9645
	SER	0.1407	0.3815	0.1879	**0.1276**	0.3720	4.3373	0.1909	1.3438	0.1573
	AER	1.4017	1.2528	1.2367	1.6440	**1.7890**	1.5055	1.2323	1.1593	1.4265
	Time(s)	29.84	**0.51**	17.91	8.62	18.62	41.23	34.11	51.59	34.84
Pavia Center	**AUC**	**0.9995**	0.9982	0.9815	0.9819	0.9937	0.8451	0.9812	0.9994	0.9993
	SER	**0.1483**	0.3004	0.2960	0.2390	1.7336	6.0666	0.4963	0.2039	0.2004
	AER	1.5073	1.1895	1.2185	**1.5378**	1.4968	1.1469	1.2526	1.3814	1.4068
	Time(s)	20.64	**0.16**	37.37	4.77	15.32	80.54	14.93	97.95	49.99
San Diego	**AUC**	**0.9947**	0.9403	0.8750	0.8854	0.9769	0.9921	0.9025	0.9914	0.9876
	SER	**0.7925**	1.4744	1.3717	1.3838	1.0489	3.0082	1.3599	0.9690	0.9620
	AER	1.1441	1.1434	1.0002	1.0402	1.1802	**1.3865**	0.9996	1.2303	1.2945
	Time(s)	1121.14	**1.20**	77.90	8.14	38.40	64.90	117.46	122.27	49.90
Los Angeles	**AUC**	**0.9960**	0.8404	0.9722	0.9490	0.9881	0.9782	0.9873	0.9623	0.9690
	SER	**0.5354**	0.9803	0.7215	0.9331	0.6497	3.9840	0.6175	0.9248	0.8689
	AER	1.1200	1.1570	1.1346	1.1696	1.2201	**1.3619**	1.3607	1.0968	1.2052
	Time(s)	89.19	**0.56**	49.95	10.34	44.07	68.90	187.46	151.32	66.07
CRi	**AUC**	**0.9961**	0.9135	0.8284	0.9245	0.7867	0.9705	0.9605	0.9129	0.9831
	SER	0.8762	1.2713	1.2905	**0.7347**	7.1882	1.4476	1.1903	1.0531	0.9131
	AER	1.2472	1.0480	1.0463	**1.5196**	1.0341	1.2448	1.0869	1.1976	1.2722
	Time(s)	851.37	**0.48**	83.07	42.07	17351.83	352.21	7389.99	981.22	783.27

The optimal detection results are marked in thick type and the sub-optimal detection results are underlined.

KIFD are the largest, and the SER and AER values for RGAE and GAED do not differ significantly from the optimal results.

For the dataset Los Angeles, as can be seen in Fig. 4(d), although SG-CCR is not as effective as KIFD in highlighting anomalies, its suppression of background is much better than KIFD. So, on balance, SG-CCR has the best anomaly detection effect. The ROC curve in Fig. 8(a) shows that SG-CCR has a better ROC curve trend than other algorithms and the best AUC value. The box plot in Fig. 8(b) shows that SG-CCR is not as effective as LRaSMD in suppressing the background and not as effective as KIFD in separating the background from the anomaly, but in combination, SG-CCR outperforms the other algorithms. In Table I, the SER value of SG-CCR is the smallest and the algorithm with the largest AER value is KIFD.

Comparing datasets San Diego and Los Angeles, although their anomaly targets are similar, in the pseudocolor images, the target for dataset Los Angeles is closer to the background and the anomaly target is less significant. Therefore, the anomaly saliency weight works better for the dataset San Diego.

For dataset Cri, it can be seen from Fig. 4(e) that both SG-CCR and LRaSMD have similar suppression of background and highlighting of anomalies, but both outperform other algorithms overall. The ROC curve in Fig. 9(a) shows that the SG-CCR curve follows a similar trend to that of LRaSMD, GAED, but the further to the right axis the ROC curve of SG-CCR is closer to the upper left corner, thus the SG-CCR algorithm has the largest AUC value. The box plot in Fig. 9(b) shows that SG-CCR does not suppress the background as well as LRaSMD, does not separate the anomalies from the background as well as KIFD, and performs well although the combined effect is not optimal. From Table I, it can be obtained that the SER and AER values of LRaSMD are optimal and the

TABLE II
COMPARISON OF SG-CCR AND ITS DERIVATES ON FIVE DATASETS

Datasets		Algorithms			
		CRD	CCR	JCCR	SG-CCR
HYDICE	AUC	0.9987	0.9987	0.9987	**0.9994**
	SER	0.3720	0.4105	0.3978	**0.1407**
	AER	1.7890	**1.8392**	1.8121	1.4017
Pavia Center	AUC	0.9937	**0.9995**	**0.9995**	**0.9995**
	SER	1.7336	1.3656	1.5741	**0.1483**
	AER	1.4968	2.6080	**2.9156**	1.5073
San Diego	AUC	0.9766	0.9869	0.9887	**0.9947**
	SER	1.0489	1.2957	1.1989	**0.7925**
	AER	1.1802	1.1809	**1.2118**	1.1441
Los Angeles	AUC	0.9881	0.9905	0.9916	**0.9960**
	SER	0.6497	0.6575	0.6358	**0.5354**
	AER	1.2201	1.3082	**1.3086**	1.1200
CRi	AUC	0.7867	0.8922	0.8799	**0.9961**
	SER	7.1822	9.9066	9.8529	**0.8762**
	AER	1.0341	**1.5104**	1.4684	1.2472

The anomaly detection results of four models are obtained through ablation study. The bold value represents the optimal result of AUC, SER, and AER.

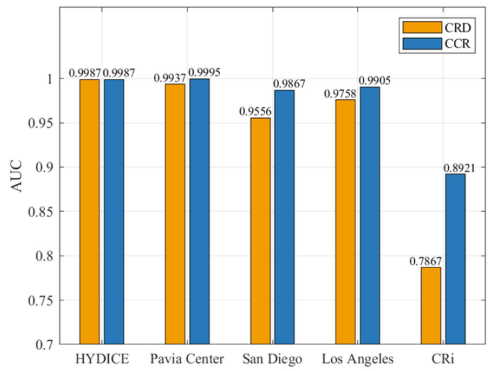

Fig. 10. AUC of CRD and CCR for five datasets.

SER value of SG-CCR is suboptimal. The combination of the above metrics shows that the SG-CCR algorithm performs well. Furthermore, for this large scene dataset, the time consumption of the SG-CCR algorithm is within an acceptable range.

V. DISCUSSION

A. Ablation Analysis

In order to verify the effectiveness of the proposed algorithm, an ablation study is conducted in this chapter. The CCR model is obtained by adding competitive constraint term to the CRD, the JCCR model is obtained by adding the Jaccard similarity coefficient to the CCR model, and the SG-CCR model is constructed by adding anomaly saliency weight to the JCCR model. Through the experiments, the accuracy of these four models on five datasets was obtained, as shown in Table II. Based on the information in the table, we used this to analyze the validity of the three contributions.

1) Validity Analysis of Competitive Constraint Term: To assess the validity of this contribution, the CRD model was compared with the CCR model. Fig. 10 shows the accuracy of CRD compared with CCR on five datasets, and it can be found that the detection accuracy of CCR is higher than that of CRD in all the remaining datasets, except in dataset HYDICE, where the accuracy of both models is the same, and the improvement in detection accuracy is particularly obvious in the CRi data. It can also be seen from Table II that after adding the competitive constraint term, the AER values of all datasets improved, and the SER values of dataset Pavia Center decreased.

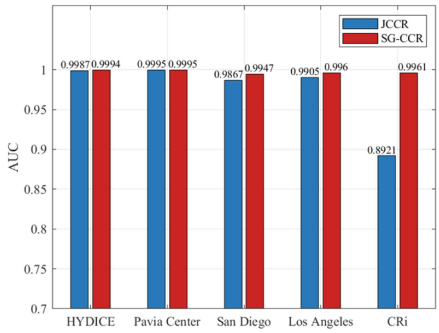

Fig. 11. AUC of JCCR and SG-CCR for five datasets.

The above analysis shows that whether it is small scene data or large scene image, the performance of anomaly detection is improved, which proves the effectiveness and stability of the competition constraint term.

2) Validity Analysis of Spectral Jaccard Coefficient: We use the JCCR model for comparison with CCR to reflect the validity of the spectral Jaccard coefficient. In Table II, we can find that while the AUC values of JCCR are only greater than CCR in the San Diego and Los Angeles datasets, the SER values decrease in both datasets HYDICE and CRi, and the AER values increase in dataset Pavia Center.

This shows that the contribution of the Jaccard similarity coefficient varies from data to data, but overall, the contribution is helping to improve the detection performance.

3) Validity Analysis of Anomaly Saliency Weight: The effectiveness of this contribution was assessed by comparing the JCCR and SG-CCR models. Fig. 11 shows a comparison of the accuracy of JCCR and SG-CCR on five datasets, and it can be found that, except in the dataset Pavia Center where the two models have the same accuracy, in the rest of the datasets, the detection accuracy of SG-CCR is higher than that of JCCR, especially in the CRi data where the accuracy improvement is obvious. Table II shows that the SER values of all datasets decreased after adding the anomalous significant weights, which indicates the robustness of the contribution of the anomalous significant weights to the suppression of background.

From the above analysis, we can find that the competitive constraint, the Jaccard similarity coefficient, and the anomaly saliency weight all contribute to anomaly detection, with the competitive constraint and the anomaly saliency weight contributing relatively more to anomaly detection. Furthermore, we find that the competitive constraint term and the saliency weight are not only contributing but also applicable in different data scenarios with robustness.

B. Parameter Analysis

The parameters of the proposed algorithm are: w_{out}, w_{in}, w_s, λ, β, M_0 and t. M_0 and t are auxiliary parameters in the

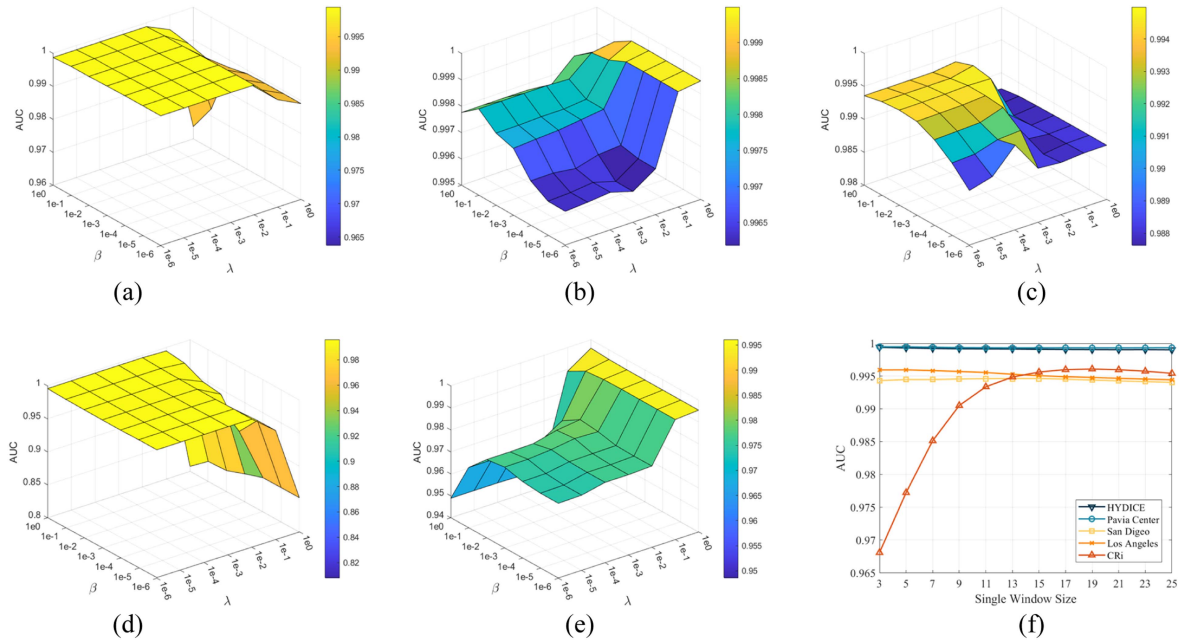

Fig. 12. Analysis of SG-CCR to λ, β, w_s. (a)–(e): Sensitivity analysis of λ and β. (a) HYDICE. (b) Pavia center. (c) San Diego. (d) Los Angeles. (e) CRi. (f) Stability analysis of single-window w_s.

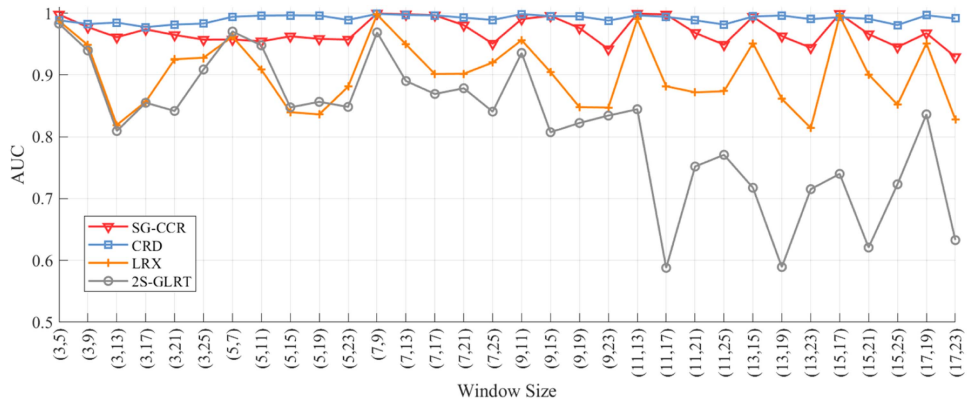

Fig. 13. Stability analysis of dual-window in HYDICE dataset.

construction of saliency weight for anomalies and have little effect on the results of anomaly detection. For dataset HYDICE, $M_0 = 55$, $t = 8$. For dataset Pavia, $M_0 = 105$, $t = 2$. For dataset San Diego, $M_0 = 4000$, $t = 8$. For dataset Los Angeles, $M_0 = 9500$, $t = 1$. For dataset Cri, M_0 takes the total number of pixels, $t = 8$. Fixed parameters M_0 and t. The effects of other parameters are analyzed and the range of values of the parameters is described in Section IV.

1) *Setting of λ and β:* Fig. 12(a)–(e) show the variation in detection accuracy of SG-CCR under different λ and β. The dataset Los Angeles is not sensitive to λ. When λ is set in the range {1e0–1e-6}and β is chosen from {1e-2–1e-6}, the accuracy of datasets HYDICE and Los Angeles is higher and the proposed algorithm is insensitive to both λ and β within this parameter setting. In the dataset Pavia Center, the anomaly detection accuracy was better when β was taken as {1e-1,1e0} and the detection accuracy of SG-CCR was essentially the same when fixing β to change λ. In the San Diego dataset, the detection accuracy of SG-CCR is high and stable when the range of λ is {1e0–1e-3} and β is {1e-3–1e-6}, i.e., the algorithm is not sensitive to λ and β in this range. The CRi dataset reaches its maximum when both λ and β are taken as 1e0, and anomaly detection accuracy changes steadily when λ and β are set in {1e-2–1e-6}.

2) *Stability of Single Window w_s*: The single-window w_s was set to highlight the significant characteristics of anomalies in the local area, fixing other parameter conditions and obtaining anomaly detection accuracy values for the five data sets within the parameters of the single-window. Fig. 12(f) shows the stability analysis of the single-window. It can be seen from the figure that the dataset HYDICE, Pavia Center, San Diego, Los Angeles has a stable change in accuracy over the course of the window change. The

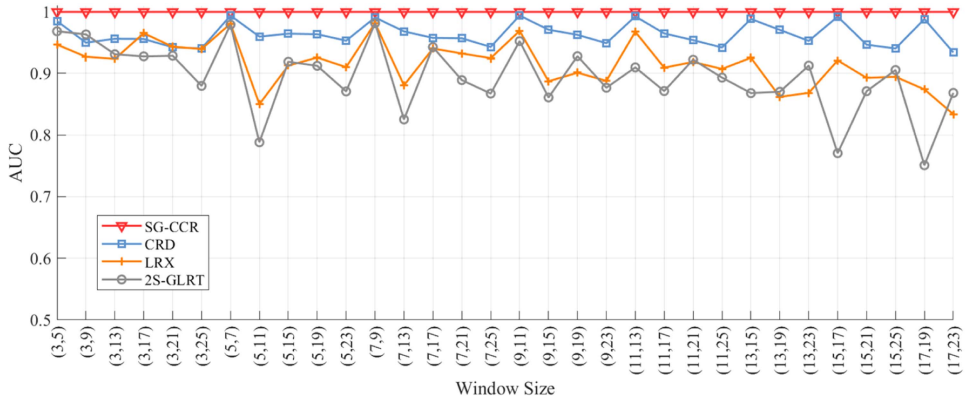

Fig. 14. Stability analysis of dual-window in Pavia center dataset.

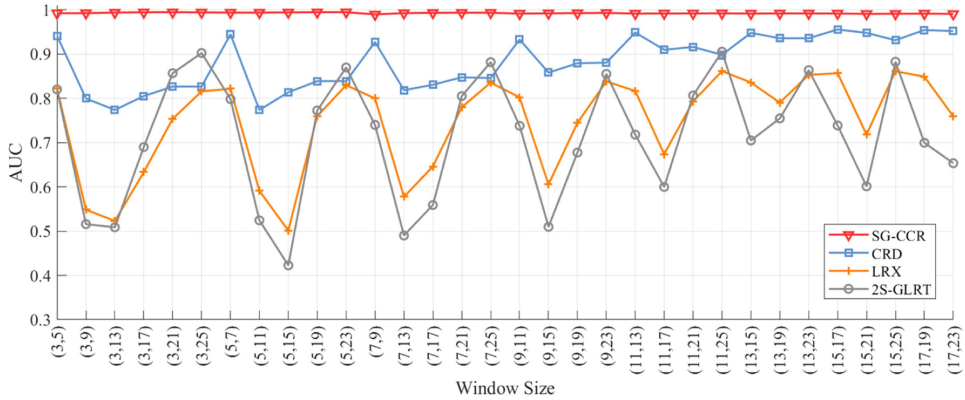

Fig. 15. Stability analysis of dual-window in San Diego dataset.

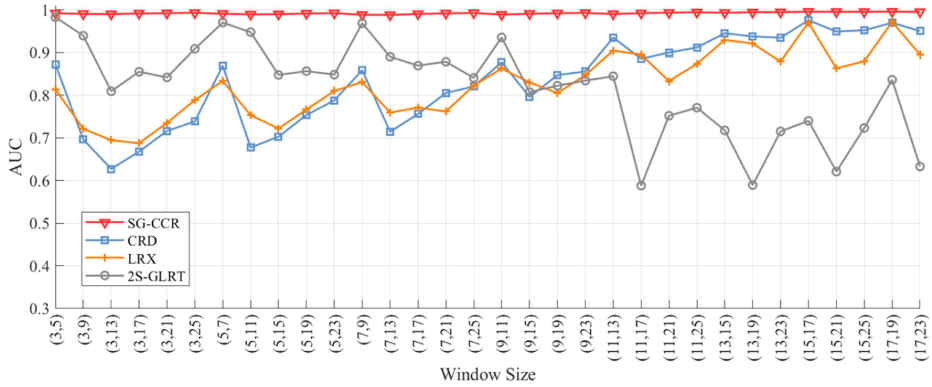

Fig. 16. Stability analysis of dual-window in Los Angeles dataset.

dataset CRi, on the other hand, increases in accuracy as the single-window becomes larger until its size is 17×17, when the accuracy also stabilizes. This may be because the anomaly targets for CRi are large and dispersed, and the smaller window may result in the anomaly not being locally significant.

3) *Stability Analysis of Internal Window w_{in} and External Window w_{out}* : In this article, 32 sets of windows were selected within the parameters of the dual-window and compared with three comparative algorithms, CRD, LRX, and 2S-GLRT, as a way to analyze the effect of window variation on the algorithm. Figs. 13–17 show the AUC values for the five datasets with different window size. Figs. 14–17 show that the curve of the SG-CCR algorithm is almost a straight line in the dataset Pavia Center, San Diego, Los Angeles, CRi, which indicates that the detection performance of SG-CCR is stable and is not affected by window size variations in these four datasets. Moreover, the average detection performance of SG-CCR outperformed the other algorithms. Fig. 13 shows the

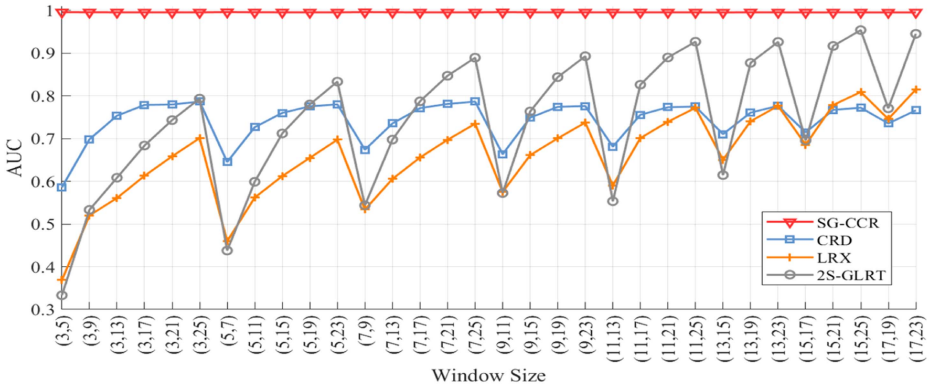

Fig. 17. Stability analysis of dual-window in CRi dataset.

variation in the accuracy of the dataset HYDICE under different datasets. The detection performance of SG-CCR is less stable than that of CRD, which is because the background of this dataset is complex, with small and few anomalies, so the collaboration between dictionary atoms is stronger under most window settings, while the algorithm proposed in this article adds competitive constraint term, so the dictionary selection under different windows may affect the collaboration between atoms effect, which affects the detection performance and leads to slightly lower stability of this algorithm than CRD, but the window-related stability of this algorithm is better compared with other algorithms. Overall, the detection performance of the proposed SG-CCR is relatively stable and has good robustness.

VI. Conclusion

In this article, a hyperspectral CCR anomaly detector with anomaly salient weights is proposed. In the CCR method, the competition between the anomaly and the background in the dictionary is considered and therefore a competition constraint is imposed. The construction of anomaly saliency weights suppresses background pixels thus highlighting anomaly targets. Experimental results on five datasets demonstrate that the proposed algorithm not only outperforms other comparative algorithms in detection but also achieves good performance in both large and small target scenarios.

Acknowledgment

The authors would like to thank Prof. W. Li, Beijing University of Chemical Technology, for providing CRD code, Dr. P. Duan, Hunan University, for sharing KIFD code, Dr. G. Fan, Wuhan University, for sharing RGAE code, Dr. P. Xiang, Xidian University, for sharing GAED code, Computational Intelligence Group of the University of the Basque Country (UPV/EHU) for sharing Pavia Center dataset.

References

[1] H. Su, Z. Wu, H. Zhang, and Q. Du, "Hyperspectral anomaly detection: A survey," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 1, pp. 64–90, Mar. 2022.

[2] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.

[3] 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.

[4] X. Shang et al., "Target-constrained interference-minimized band selection for hyperspectral target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 6044–6064, Jul. 2021.

[5] H. Su, Y. Yu, Q. Du, and P. Du, "Ensemble learning for hyperspectral image classification using tangent collaborative representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 6, pp. 3778–3790, Jun. 2020.

[6] B. Xie, Y. Zhang, S. Mei, G. Zhang, Y. Feng, and Q. Du, "Spectral variation augmented representation for hyperspectral imagery classification with few labeled samples," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5543212.

[7] 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.

[8] Z. Zou and Z. Shi, "Random access memories: A new paradigm for target detection in high resolution aerial remote sensing images," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1100–1111, Mar. 2018.

[9] J. E. Fowler and Q. Du, "Anomaly detection and reconstruction from random projections," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 184–195, Jan. 2012.

[10] C.-I. Chang, "Hyperspectral anomaly detection: A dual theory of hyperspectral target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5511720.

[11] M. T. Eismann, A. D. Stocker, and N. M. Nasrabadi, "Automated hyperspectral cueing for civilian search and rescue," *Proc. IEEE*, vol. 97, no. 6, pp. 1031–1055, Jun. 2009.

[12] I. Makki, R. Younes, C. Francis, T. Bianchi, and M. Zucchetti, "A survey of landmine detection using hyperspectral imaging," *ISPRS J. Photogrammetry Remote Sens.*, vol. 124, pp. 40–53, 2017.

[13] V. Dremin et al., "Skin complications of diabetes mellitus revealed by polarized hyperspectral imaging and machine learning," *IEEE Trans. Med. Imag.*, vol. 40, no. 4, pp. 1207–1216, Apr. 2021.

[14] P. K. Goel, S. O. Prasher, R. M. Patel, J. A. Landry, R. B. Bonnell, and A. A. Viau, "Classification of hyperspectral data by decision trees and artificial neural networks to identify weed stress and nitrogen status of corn," *Comput. Electron. Agriculture*, vol. 39, no. 2, pp. 67–93, 2003.

[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] K. Heesung, Z. D. Sandor, and M. N. Nasser, "Adaptive anomaly detection using subspace separation for hyperspectral imagery," *Opt. Eng.*, vol. 42, no. 11, pp. 3342–3351, 2003.

[17] Z. Wu, W. Zhu, J. Chanussot, Y. Xu, and S. Osher, "Hyperspectral anomaly detection via global and local joint modeling of background," *IEEE Trans. Signal Process.*, vol. 67, no. 14, pp. 3858–3869, Jul. 2019.

[18] L. Ren, L. Zhao, and Y. Wang, "A superpixel-based dual window RX for hyperspectral anomaly detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 7, pp. 1233–1237, Jul. 2020.

[19] K. Heesung 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.
[20] 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.
[21] A. Schaum, "Joint subspace detection of hyperspectral targets," in *Proc. IEEE Aerosp. Conf.*, 2004, vol. 3, pp. 1–1824.
[22] N. Billor, A. S. Hadi, and P. F. Velleman, "BACON: Blocked adaptive computationally efficient outlier nominators," *Comput. Statist. Data Anal.*, vol. 34, no. 3, pp. 279–298, 2000.
[23] M. Stefania, D. Marco, and C. Giovanni, "Improved estimation of local background covariance matrix for anomaly detection in hyperspectral images," *Opt. Eng.*, vol. 49, no. 4, 2010, Art. no. 046201.
[24] H. Su, H. Zhang, Z. Wu, and Q. Du, "Relaxed collaborative representation with low-rank and sparse matrix decomposition for hyperspectral anomaly detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 6826–6842, 2022.
[25] J. Peng et al., "Low-rank and sparse representation for hyperspectral image processing: A review," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 1, pp. 10–43, Mar. 2022.
[26] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Simultaneous joint sparsity model for target detection in hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 676–680, Jul. 2011.
[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] X. Zhang, "Hyperspectral target detection via locality-constrained group sparse representation," in *Proc. IEEE Int. Conf. Signal Process., Commun. Comput.*, 2015, pp. 1–5.
[29] Z. Yuan, H. Sun, K. Ji, Z. Li, and H. Zou, "Local sparsity divergence for hyperspectral anomaly detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 10, pp. 1697–1701, Oct. 2014.
[30] 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.
[31] J. Peng, L. Li, and Y. Y. Tang, "Maximum likelihood estimation-based joint sparse representation for the classification of hyperspectral remote sensing images," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 6, pp. 1790–1802, Jun. 2019.
[32] E. J. Candès et al., "Robust principal component analysis?," *J. ACM*, vol. 58, no. 3, pp. 1–37, 2011.
[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, 2014, Art. no. 083641.
[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] L. Zhang, M. Yang, and F. Xiangchu, "Sparse representation or collaborative representation: Which helps face recognition?," in *Proc. Int. Conf. Comput. Vis.*, 2011, pp. 471–478.
[37] 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.
[38] M. Vafadar and H. Ghassemian, "Hyperspectral anomaly detection using outlier removal from collaborative representation," in *Proc. 3rd Int. Conf. Pattern Recognit. Image Anal.*, 2017, pp. 13–19.
[39] 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.
[40] 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.
[41] G. Zhang, N. Li, B. Tu, Z. Liao, and Y. Peng, "Hyperspectral anomaly detection via dual collaborative representation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 4881–4894, 2020.
[42] Z. Hou, W. Li, R. Tao, P. Ma, and W. Shi, "Collaborative representation with background purification and saliency weight for hyperspectral anomaly detection," *Sci. China Inf. Sci.*, vol. 65, no. 1, 2021, Art. no. 112305.
[43] Y. Fang, Q. Ye, L. Sun, Y. Zheng, and Z. Wu, "Multiattention joint convolution feature representation with lightweight transformer for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5513814.
[44] K. Kayabol, E. B. Aytekin, S. Arisoy, and E. E. Kuruoglu, "Skewed t-distribution for hyperspectral anomaly detection based on autoencoder," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 5510705.
[45] Y. Li, T. Jiang, W. Xie, J. Lei, and Q. Du, "Sparse coding-inspired GAN for hyperspectral anomaly detection in weakly supervised learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5512811.
[46] L. Zhang and B. Cheng, "Transferred CNN based on tensor for hyperspectral anomaly detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 12, pp. 2115–2119, Dec. 2020.
[47] S. Song, H. Zhou, Y. Yang, and J. Song, "Hyperspectral anomaly detection via convolutional neural network and low rank with density-based clustering," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 9, pp. 3637–3649, Sep. 2019.
[48] H. Yuan, X. Li, F. Xu, Y. Wang, L. L. Lai, and Y. Y. Tang, "A collaborative-competitive representation based classifier model," *Neurocomputing*, vol. 275, pp. 627–635, 2018.
[49] C. Shah and Q. Du, "Collaborative-competitive representation with spatial regularization for hyperspectral anomaly detection," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2022, pp. 1253–1256.
[50] J. Gou et al., "Double competitive constraints-based collaborative representation for pattern classification," *Comput. Elect. Eng.*, vol. 84, 2020, Art. no. 106632.
[51] M. Rezaei and P. Fränti, "K-sets and k-swaps algorithms for clustering sets," *Pattern Recognit.*, vol. 139, 2023, Art. no. 109454.
[52] W. Liu et al., "Random selection-based adaptive saliency-weighted RXD anomaly detection for hyperspectral imagery," *Int. J. Remote Sens.*, vol. 39, no. 8, pp. 2139–2158, 2018.
[53] F. Zhu, Y. Wang, S. Xiang, B. Fan, and C. Pan, "Structured sparse method for hyperspectral unmixing," *ISPRS J. Photogrammetry Remote Sens.*, vol. 88, pp. 101–118, 2014.
[54] 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.
[55] J. Kerekes, "Receiver operating characteristic curve confidence intervals and regions," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 251–255, Apr. 2008.
[56] D. F. Williamson, R. A. Parker, and J. S. Kendrick, "The box plot: A simple visual method to interpret data," *Ann. Intern. Med.*, vol. 110, no. 11, pp. 916–921, 1989.
[57] M. Díaz, R. Guerra, P. Horstrand, S. López, and R. Sarmiento, "A line-by-line fast anomaly detector for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 8968–8982, Nov. 2019.
[58] M. Díaz, R. Guerra, S. López, and R. Sarmiento, "An algorithm for an accurate detection of anomalies in hyperspectral images with a low computational complexity," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 1159–1176, Feb. 2018.
[59] S. Li, K. Zhang, P. Duan, and X. Kang, "Hyperspectral anomaly detection with kernel isolation forest," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 319–329, Jan. 2020.
[60] J. Liu, Z. Hou, W. Li, R. Tao, D. Orlando, and H. Li, "Multipixel anomaly detection with unknown patterns for hyperspectral imagery," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 10, pp. 5557–5567, Oct. 2022.
[61] G. Fan, Y. Ma, X. Mei, F. Fan, J. Huang, and J. Ma, "Hyperspectral anomaly detection with robust graph autoencoders," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5511314.
[62] P. Xiang, S. Ali, S. K. Jung, and H. Zhou, "Hyperspectral anomaly detection with guided autoencoder," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5538818.

Yufan Yang received the B.E. degree in surveying and mapping engineering from Henan Polytechnic University, Jiaozuo, China, in 2021. She is currently working toward the M.E. degree in surveying and mapping engineering in the School of Earth Sciences and Engineering, Hohai University, Nanjing, China.

Her research interests include hyperspectral remote sensing, and hyperspectral target detection.

Hongjun Su (Senior Member, IEEE) received the Ph.D. degree in cartography and geography information system from the Key Laboratory of Virtual Geographic Environment (Ministry of Education), Nanjing Normal University, Nanjing, China, in 2011.

He is currently a Full Professor with the School of Earth Sciences and Engineering, Hohai University, Nanjing. His main research interests include hyperspectral remote sensing dimensionality reduction, classification, and spectral unmixing.

Dr. Su is an Associate Editor for IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. He was the recipient of the 2016 Best Reviewer Award from the *IEEE Geoscience and Remote Sensing Society*.

Zhaoyue Wu (Graduate Student Member, IEEE) received the B.E. degree in geodesy and geomatics engineering and the M.E. degree in photogrammetry and remote sensing from the School of Earth Sciences and Engineering, Hohai University, Nanjing, China, in 2018 and 2021, respectively. She is currently working toward the Ph.D. degree in aerospace engineering from the University of Extremadura, Cáceres, Spain.

Her research interests include hyperspectral target detection, especially anomaly detection, hyperspectral image processing, machine learning, and deep learning.

Qian Du (Fellow, IEEE) received the Ph.D. degree in electrical engineering from the University of Maryland, Baltimore, MD, USA, in 2000.

She is currently the Bobby Shackouls Professor with the Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS, USA. Her research interests include hyperspectral image analysis and applications, pattern classification, data compression, and neural networks.

Dr. Du is a Fellow of the IEEE-Institute of Electrical and Electronics Engineers, and a Fellow of the SPIE-International Society for Optics and Photonics. She was a Co-Chair of the Data Fusion Technical Committee of the *IEEE Geoscience and Remote Sensing Society* from 2009 to 2013, and the Chair of the Remote Sensing and Mapping Technical Committee of the International Association for Pattern Recognition from 2010 to 2014. She was an Associate Editor for IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, the *Journal of Applied Remote Sensing*, the *IEEE Signal Processing Letters*, IEEE Transactions of Geoscience and Remote Sensing, Pattern Recognition, and MDPI Remote Sensing. In 2016–2020, she was the Editor-in-Chief of the IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. She is a Member of *IEEE Periodicals Review and Advisory Committee*, and *SPIE Publications Committee.*