Hyperspectral anomaly detection based on autoencoder and spatial morphology extraction

Jing Feng and **Liyan Zhang***

Capital Normal University, College of Resource Environment and Tourism,
Key Laboratory of 3D Information Acquisition and Application of Ministry, Beijing, China

Abstract. Anomaly detection (AD) is a crucial task for detecting salient objects in cluttered backgrounds. Classical AD algorithms based on statistical models or geometric models have achieved acceptable detection results. However, most of them do not hierarchically extract deep features or consider the spatial structure of the images. We propose a method that combines the reconstruction error of autoencoder (AE) and spatial morphological characteristics to estimate anomalousness. The reconstruction errors and the spatially dominant information are comprehensively considered. Specifically, given the compression capability of the AE, we use the dimensionality reduction results obtained by encoding for analyzing local pixel difference; an adaptive dual window is employed in this process. The morphological transformation commonly used in edge detection is utilized to refine the small space anomalies. Experimental results on different hyperspectral images show that the proposed AE and spatial morphology extraction-based approach significantly surpasses several traditional alternatives. © *The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.* [DOI: 10.1117/1.JRS.15.038507]

1 Introduction

In recent years, the successful launch of multiple hyperspectral satellites has dramatically improved the data acquisition capabilities. With the increased availability of data, the application of hyperspectral images (HSIs) is faced with more opportunities. Hyperspectral remote sensing can provide spectral characteristics for distinguishing objects in hundreds of contiguous bands other than two-dimensional information reflecting the spatial distribution of objects. The current HSIs processing research studies mainly concentrate on dimensionality reduction,[1,2] target detection,[3] image classification,[4,5] anomaly detection (AD),[6] and pixel unmixing.[7] Among them, target detection depends on prior knowledge in pre-investigation, and a large number of manual labeling and field survey work are often combined. As an effective supplement and improvement of target detection, AD has the advantage of detecting targets that have significant spectral differences from the background without prior knowledge. It can detect anomaly targets or anomaly situations and has substantial practical value in the military field.

Anomaly targets are generally defined as objects with a lower probability of occurrence than the background.[8] These pixels have distinctive spectral characteristics that differ from the surrounding environment, such as local fires in the forest. To realize the real-time detection of data, a series of hyperspectral AD algorithms have been proposed. For instance, the most widely used reed-xiaoli (RX) anomaly detector[9] uses the background spectrum of the entire scene to construct the covariance matrix and determines the detected pixel below the threshold according to the Mahalanobis distance (MD). As a representative based on statistical models, RX has the convenience of implementation; however, the established background model is difficult to describe complex scenes, which leads to a high false alarm rate of this algorithm. An enhancement of background selection, such as local RX (LRX) detector,[10] selects part of the pixels to

*Address all correspondence to Liyan Zhang, zhangliyan@cnu.edu.cn

calculate the background by adopting the sliding window containing the target and background pixels. The cluster-based anomaly detector assumes that background pixels within clusters can be modeled according to Gaussian distributions,[11] whereas anomalousness deviates from the distribution of the cluster. This work can detect man-made targets in the presence of natural background clutter while ensuring a relatively low false alarm rate, but the above algorithm also has a shortcoming, the background statistics always contain anomalies, which makes the dataset contaminated. Aiming at the real and complex background of the image, modified algorithms based on RX have been put forward. To deal with the high dimension and nonlinear characteristics of data, a kernel RX algorithm is presented, which maps data to higher dimensional space through a specific nonlinear mapping function.[12] In addition, the regularized RX method forces the filter coefficients to shrink through the regularization term,[13] by introducing this idea, the model can be made more stable and have better performance. Furthermore, the algorithm based on subspace RX has also achieved good detection results.[14]

Banerjee et al.[15] proposed a support vector data description (SVDD) algorithm for AD in 2006; this algorithm constrains the training sample in a hypersphere surrounded by it and utilizes the minimum closed hypersphere to separate the sample from other classes, effectively averting the problem that the assumption of Gaussian distribution is not consistent with the data distribution. Collaborative representation-based detector (CRD), which achieves advanced detection performance without assuming the background distribution, minimizes the norm of the weight vector without estimating its covariance matrix, thereby enhancing the cooperation between adjacent pixels.[16] The sparsity score estimation framework analyzes the frequency of each dictionary atom used for reconstruction based on the learned dictionary and the corresponding sparse coding.[17] Considering matrix decomposition and statistical information, the low-rank and sparse matrix decomposition-based method uses MD difference to detect anomalies.[18]

Recently, neural network-based AD has shown strong strengths in modeling and generalization of complicated data. Ma et al.[8] employed reconstruction errors generated by deep belief networks as adaptive weights combined with local Euclidean distance to determine anomalies. Xie et al.[19] proposed the structure of combining discriminator and autoencoder (AE) to form a new network. The background input to the network for training derives from clustering, and then the original HSI was tested by AE. Reference 20 adopted the fully convolutional AE and adaptive loss function for image reconstruction. The main innovation is to upgrade the network structure to adapt to the background and realize auto-detection. Both of the above two methods consider improving the network to fully learn the background, so as to separate the anomalies. One uses a semi-supervised approach, and the other is unsupervised. The main difference between this paper and the above is that the significance of the targets in the spatial domain is considered while using stacked AE obtained characteristics of the spectral domain.

Based on the AE and the connectivity between background and anomaly pixels in the spatial domain, a method for AD is proposed in this paper. The first step is to preliminarily determine the anomaly pixels by the reconstruction errors between the input data and the reconstruction results in AE. In terms of spatial morphology, the compressed results from the encoder are obtained to analyze the local pixels discrepancy using the sliding window structure. Then morphological transformations are implemented to extract the shape features in small regions to complete the refinement of the anomalies. At last, the above-mentioned processes are merged by matrix multiplication. The main characteristics of this paper are as follows: attempt to employ reconstruction error to estimate anomalies in a stacked AE; design a local spatial representation to apply information in the spatial domain; and introduce the morphological profile operations to extract anomalies. The details of the proposed method are described in the following section.

2 Related Methods

2.1 *Autoencoder*

The AE performs encoding and decoding to compress and restore the feature distribution of the original data. Since the spectral anomaly is tiny compared with the surrounding environment and is obvious in the spectral domain, the background usually occupies most areas of the HSI. In the
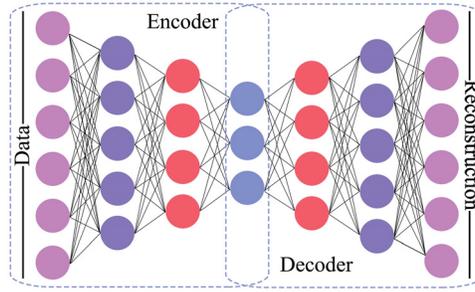
Fig. 1 Illustration of a stacked AE.

case of sufficient samples, the learning ability of the AE for background pixels is much stronger than that for anomalies. Therefore, the AE can obtain more similar reconstruction results and more minor reconstruction errors for the background pixels than the unusual pixels. AE includes encoder and decoder, which can be regarded as a self-supervised learning algorithm. Its purpose is to learn to reconstruct the original input by minimizing the reconstruction loss. The HSI (X) is input to the network pixel by pixel, and all spectral band data of a pixel are input at a time. The network structure of the AE is illustrated in Fig. 1.

The encoder of AE uses function $\sigma(\cdot): x \rightarrow h$ to project the input x into the latent layer h, and the decoder uses the function $\delta(\cdot): h \rightarrow \tilde{x}$ to reconstruct the latent layer h to ensure the output $\tilde{x}$. Mapping functions represent nonlinear activation functions, which make the neural networks possess the nonlinear ability. In this study, the Relu and sigmoid functions are selected. The output of each layer multiplied by the weight is the input of the next layer, and the output is considered to be a restored image.

Let $X \in R^{m \times n \times L}$ represent an HSI with $m \times n$ pixels and L bands in the spatial domain. X can be interpreted as a set of $m \times n$ vectors of L bands, i.e., 2D images of $X = [x_1, x_2, \cdots, x_{m \times n}]$, where $x_i \in R^{L \times 1}$. Given the input $x_i \in R^{L \times 1}$, the encoder maps x_i to the latent layer $h \in R^B$ shown in Eq. (1). B and L denote the dimensions of the layer,

$$h = \sigma(W_e x_i + b_e), \tag{1}$$

where h represents the latent layer, W_e and b_e are the weight matrix and bias of the encoder, respectively, and $\sigma(\cdot)$ is the activation function. Then decoder projects the potential layer to the output layer $\tilde{x}_i \in R^L$, as shown in Eq. (2):

$$\tilde{x}_i = \delta(W_d h + b_d), \tag{2}$$

where $\delta(\cdot)$ represents the activation, and W_d and b_d are the weight matrix and bias of the decoder, respectively. The loss function is related to the reconstruction errors between the input and output units of the AE, using the root mean square loss function:

$$\text{MSE} = \frac{1}{K} \sum_{i=1}^{K} (x_i - \tilde{x}_i)^2, \tag{3}$$

where K is the number of inputs for a given training set. During the iteration, AE learns to minimize the reconstruction error and use it to estimate anomaly score. High scores are deemed more likely to occur in anomaly pixels.

2.2 Morphological Operation

The idea of morphology is to extract corresponding shapes from the input image through specific structuring elements (SEs) for image analysis and target recognition. The dilation expands the bright area while erosion expands the dark areas.[21] They are widely used to eliminate noise, connect adjacent units, and identify obvious maximum or minimum in image areas. Dilation (Dilation $\oplus$) is defined as the maximum value in the corresponding image area of the structural element. Erosion (Erosion $\ominus$) is the minimum value in the corresponding image area of the structural element:

$$f(x, y) \oplus \text{SE} = \max\{f(x + x', y + y')|(x', y') \in D_{\text{SE}}\}, \tag{4}$$

$$f(x, y) \ominus \text{SE} = \min\{f(x + x', y + y')|(x', y') \in D_{\text{SE}}\}, \tag{5}$$

where D_{SE} is the domain of SE and f is the input image, and x' and y' are the offsets for anchor point (x, y). Opening is accomplished first by erosion and then dilation, which can eliminate the brightness of the small area. Contrary to opening operation, closing eliminates the small black holes by dilation and then erosion. The closing and opening operations in the morphology are usually used for edge detection.[22] Opening (ImO) and closing (ImC) are defined as follows:

$$ImO(f, \text{SE}) = (f \ominus \text{SE}) \oplus \text{SE}, \tag{6}$$

$$ImC(f, \text{SE}) = (f \oplus \text{SE}) \ominus \text{SE}. \tag{7}$$

It is worth noting that a specific SE is often chosen in mathematical morphology depending on the geometric structure of the original signal.[23] For example, for the image with a rectangular target area, when the square SE is selected, the result of erosion or dilation is still rectangular. If the rhombus structure slides on a rectangle, the result is an approximate octagonal shape. Additionally, the circular SE can make the processed objects smooth. Take the image with artificial objects as an example, when the experiment primarily utilizes the objects such as cars and roofs as the target, the square SE is selected according to their shape. In the subsequent experiments of this article, the square SE is used to perform morphological processing on the images.

3 Proposed Framework

This section describes the framework of the proposed autoencoder and spatial morphology extraction (AESME)-based method, as shown in Fig. 2. First, we construct an AE, including several hidden layers. The size of the input and output layer is equal to the number of channels of HSIs. There are seven neurons in the middle layer, representing encoder output and decoder input. All band ranges of a pixel in the HSI are taken as the input data; each pixel is successively inputted into the network, and reconstruction result is the same as the input image in rows, columns, and bands. The residual error is employed to determine the anomaly degree of each spectral signature; the anomaly degree is used as the evaluation:

$$r = \frac{1}{L}\sum_{i=1}^{L} \sqrt{(x_i - \tilde{x}_i)^2}, \tag{8}$$
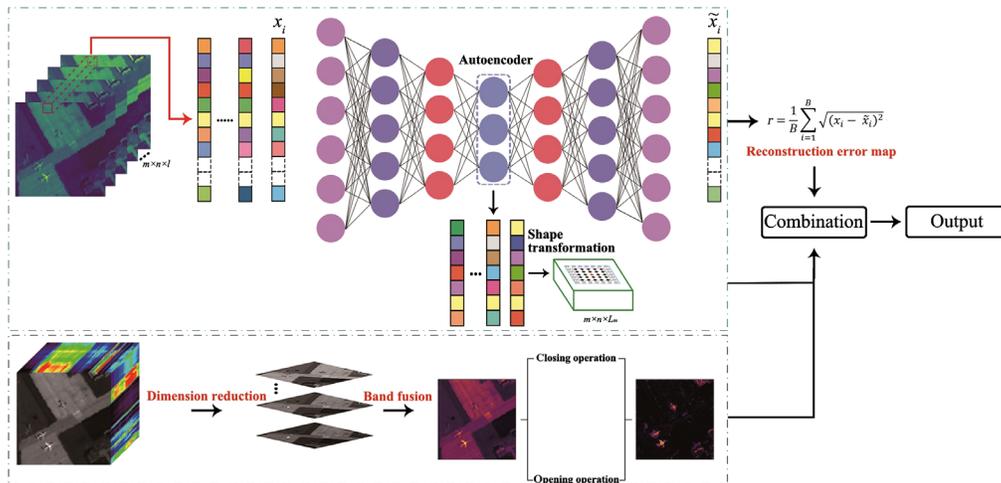

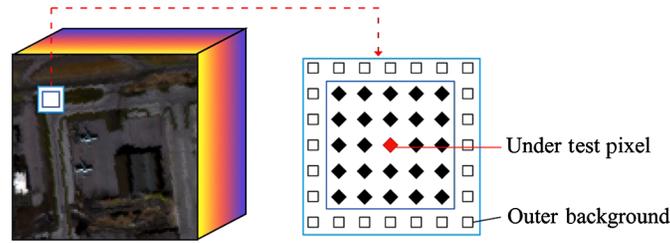**Fig. 2** Diagram of the proposed AESME AD framework.

Fig. 3 Adaptive dual-window structure for hyperspectral AD.

where r is the reconstruction error for each spectral signature, x_i is the input pixel, L is the number of spectral bands, and $\tilde{x}_i$ is the decoded output of the AE.

For the purpose of further improving the detection accuracy, the result passing through the encoder is converted into a shape-transformed image M that is equivalent to the data dimension of $m \times n \times L_m$, where L_m represents the output dimension from the encoder. A dual window with a sliding structure is established on the M, and its specific structure is shown in Figure 3; the outer neighborhoods of the window are counted as W_{out} to estimate the local background. During the sliding, the center pixel value is replaced by the difference between the local background and the corresponding center pixel; for each spectral signature, it is expressed as

$$W_c[t] = \frac{1}{L_m} \sum_{j=1}^{L_m} \left(\frac{1}{L_b} \sum_{i=1}^{L_b} |W_{\text{out}}[j][i] - W_c[j][t]| \right), \tag{9}$$

where $W_c[t]$ is the center pixel of the sliding window, $W_{\text{out}}[j][i]$ is the i'th local background pixel in the j'th band, $W_c[j][t]$ represents the center pixel in the j band, and L_b is the number of local background pixels for corresponding band. The pixel value in HSI is replaced by the average pixel discrepancy between the center pixel of the window and the background pixel. The realization of this process is shown in Fig. 4(a). It excludes an extensive range of similar backgrounds and retains areas different from the local background, mainly including pixel mutations in small areas and transition zones. When the target is retained, the regions with visual significance will be extracted, which will interfere with the anomalies extraction to some extent, so the morphological method is used to enhance the recognition of abnormal objects in space.

First, perform principal component analysis (PCA) to overcome redundant information and noise pixels. By testing the subsequent experimental data, when the optimal detection effect is obtained, the number of principal components L_n in PCA is selected variously in different images. The specific analysis is discussed in the next section. Commonly used band average fusion is employed after PCA and a new image with prominent spatial characteristics is obtained:
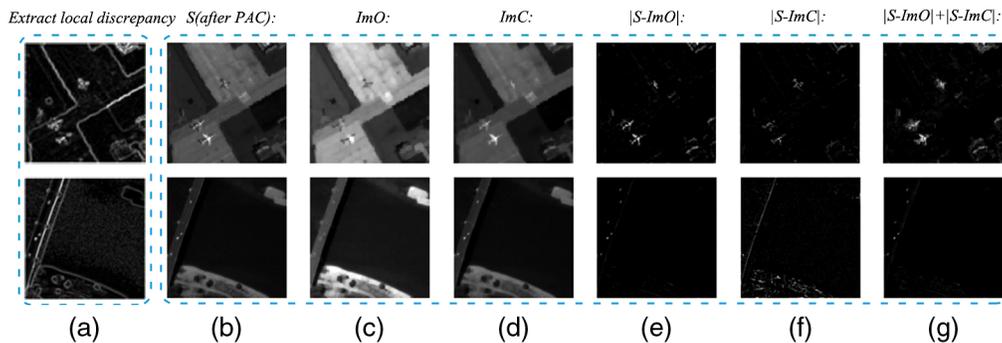

Fig. 4 The processing results in the spatial domain. (a) Spatial representation of local discrepancy. (b) Dimensionality reduction results. (c) Opening operation. (d) Closing operation. (e) The absolute value of the difference between the S and ImO. (f) The absolute value of the difference between S and ImC. (g) The final result of the morphological operation.

$$S = \frac{1}{L_n} \sum_{i=1}^{L_n} X_i, \tag{10}$$

where S represents the fused image, L_n is the total number of principal components, and X_i is the i'th principal component image. It should be noted that although the encoder can play the role of dimensionality reduction, the number of intermediate layers is fixed. The amount of information after encoding may not be the optimal result and will not be suitable for extracting information in images.

Since the morphological opening and closing operations can, respectively, remove the bright and dark connection components in a small area and will not significantly change the area of other objects, we introduced a simple operation to retain small area objects. The assumption in this section is that the actual HSI anomalies are usually manifested as small areas composed of interrelated pixels compared with the background; in addition, it also has a particularity in terms of structure attribute in the spatial domain. By fulfilling morphological operations on the fused image S, the anomalous spatial features are separated from the background. By selecting different SEs to retain or remove the connected components, the particular spatial structure of HSIs can be effectively described. Therefore, the difference map obtained is as Eq. (11) according to Ref. 24:

$$O = |S - ImO| + |S - ImC|, \tag{11}$$

where the term $|S - ImO|$ remains the bright connected components in a small region, and $|S - ImC|$ remains the dark connected components in a small region. The implementation of the opening and closing operations and their joint application results are shown in Figs. 4(b)–4(g), which can successfully erase the background.

4 Experiment

4.1 Experimental Data

The datasets employed to validate the performance of the different methods include real hyperspectral samples and their reference images. The first is San Diego dataset, which was collected by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor[25] in the San Diego airport area of California, and the noise band has been removed. Another Airport Beach City (ABU) database contains 13 sample images of 100×100 pixels and their corresponding ground reference images, which can be downloaded from its website,[26] Urban-4 data were collected by a reflection optical system imaging spectrometer (ROSIS-03) sensor,[27] all others were collected by AVIRIS sensors.

4.2 Experimental Result

To verify the proposed AESME method, four AD algorithms, including RX, LRX, SVDD, and CRD, are used for comparison. Synthetic images for experiments and detection results of different methods are indicated by Figs. 5–8. San Diego data correspond to Fig. 5, and Beach, Urban, and Airport scenes of ABU data are displayed in Figs. 6–8, respectively. In general, the RX algorithm is easy to implement and performs favorably in relatively simple scenes. LRX is
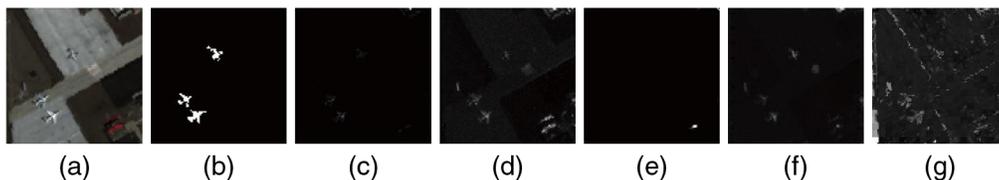

(a)   (b)   (c)   (d)   (e)   (f)   (g)

Fig. 5 Synthetic images of the (a) San Diego, (b) reference detection map; (c)–(g) are the detection maps of different methods: (c) AESME, (d) RX, (e) LRX, (f) CRD, and (g) SVDD.
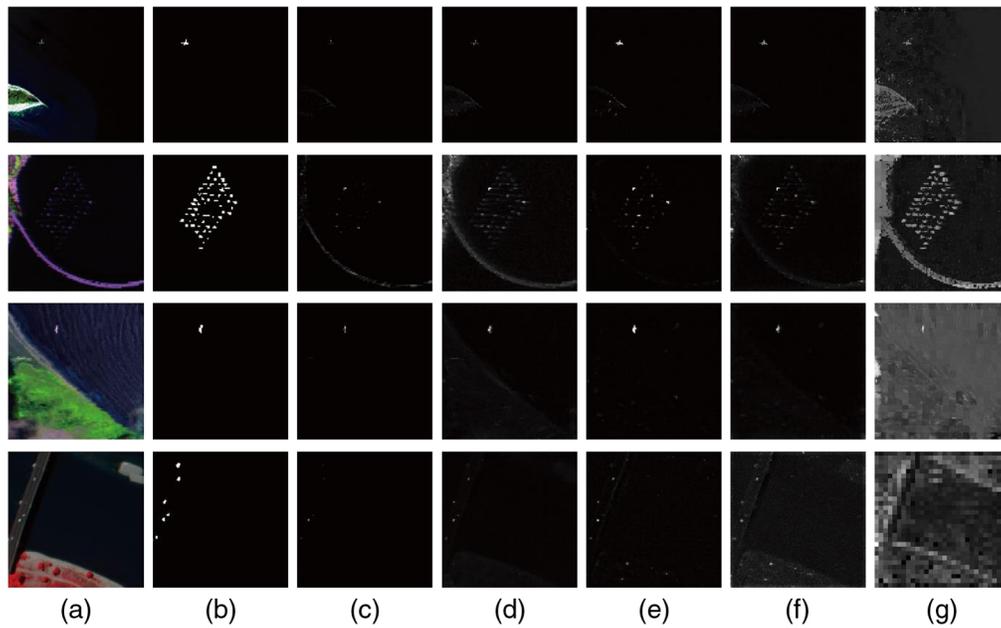
Fig. 6 Synthetic images of the (a) ABU-Beach scenes, (b) reference detection map; (c)–(g) are the detection maps of different methods: (c) AESME, (d) RX, (e) LRX, (f) CRD, and (g) SVDD.
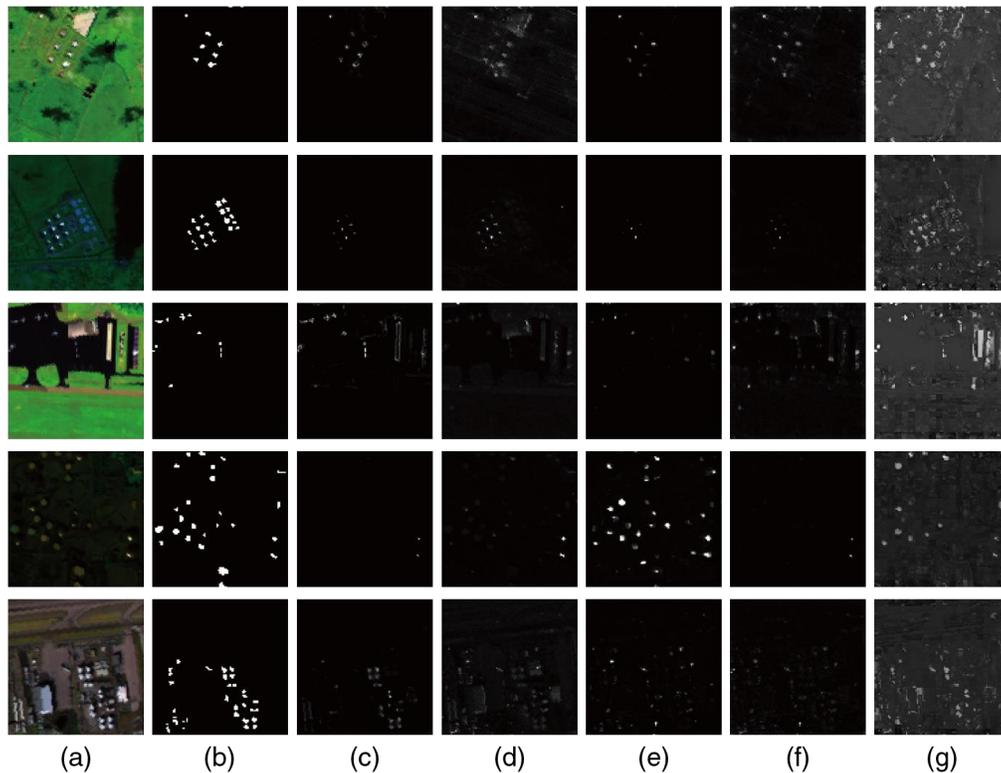

Fig. 7 Synthetic images of the (a) ABU-Urban scenes, (b) reference detection map; (c)–(g) are the detection maps of different methods: (c) AESME, (d) RX, (e) LRX, (f) CRD, and (g) SVDD.

suitable for the case of significant local spectrum variation, and the selection of the sliding window size has a crucial influence on the detection performance. The selected windows size varies according to different scenes. For example, the ABU-Beach and ABU-Urban images mostly use (11, 7) windows, and the ABU-Airport images use windows ranging from (7, 5) to (19, 17). CRD
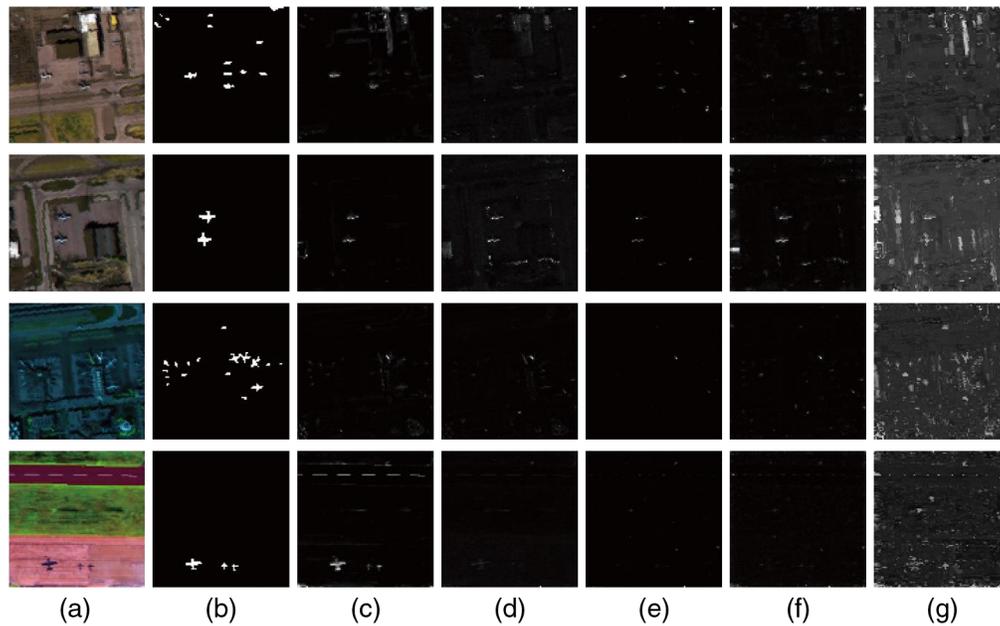
Fig. 8 Synthetic images of the (a) ABU-Airport scenes, (b) reference detection map; (c)–(g) are the detection maps of different methods: (c) AESME, (d) RX, (e) LRX, (f) CRD, and (g) SVDD.

can be properly employed in manifold scenes, and the best regularization parameter is set to 10^{-6}. The σ value of the SVDD algorithm is set between 3 and 60. Though the SVDD algorithm can be applied to various scenes, it cannot well suppress background interference. In the AESME method, the number of the AE intermediate layer is 7, and the average number of principal components is 3. In the end, the detection maps of the proposed method and the comparison methods are indicated by gray from black to white.

The receiver operating characteristic (ROC) curve is an index for intuitive evaluation of AD. The relationship between true positive rate p_d and false positive rate p_f under threshold (τ) can be effectively illustrated by drawing the ROC curves of the two ratios (p_d, p_f). The area under the ROC curve is called AUC, which is not greater than 1. A larger AUC value means the higher authenticity of the detection method. $\text{AUC} = \int_{-\infty}^{+\infty} T_{RP}(H) F_{PR}{}'(H) \mathrm{d}H$ defines the AUC value when the detection map threshold is set to H. The study uses ROC and AUC to evaluate the performance of the different methods.

The ROC curve of San Diego Airport is shown in Fig. 9(a), the curves of a scene from the Urban, Airport, and Beach dataset, respectively, are shown in Figs. 9(b)–9(d), respectively. AUC values of different HSIs are listed in Tables 1 and 2. For the visualization result of the San Diego dataset, the proposed method retains the basic shape of the airport and a relatively pure background, whereas RX, LRX, and SVDD do not eliminate the interference in the background, and the CRD method loses the shape information of the target. The AESME method obtains the highest AUC scores for the ABU-Urban dataset, which extract anomalies more completely in Urban-1, Urban-3, and Urban-5. AUC scores of Urban-2 and Urban-4 are high, but the visual results of this method do not clearly show the shapes of anomalies. The RX algorithm also shows good performance on the urban scenes except Urban-3. In Urban-4, LRX has a high false alarm rate for the background, and CRD displays the target shape incompletely; the CRD, LRX, SVDD methods cannot accurately identify the target in Urban-5. For the dataset of ABU-Airport, the AESME method has false alarm in Airport-1 and Airport-2, and LRX can obtain approximate AUC values. However, LRX and CRD have insufficient discrimination between target and background in Airport-3 and Airport-4, and the AUC values are low, whereas SVDD has a high false alarm rate. The method presented in Airport scenes is still the best in overall detection performance. For beach dataset, the proposed method shows good performance in Beach-1, Beach-3, and Beach-4, but the coastline is not excluded in Beach-2, and the LRX method avoids the false detection of linear features compared with other detection results. It is proved that this method
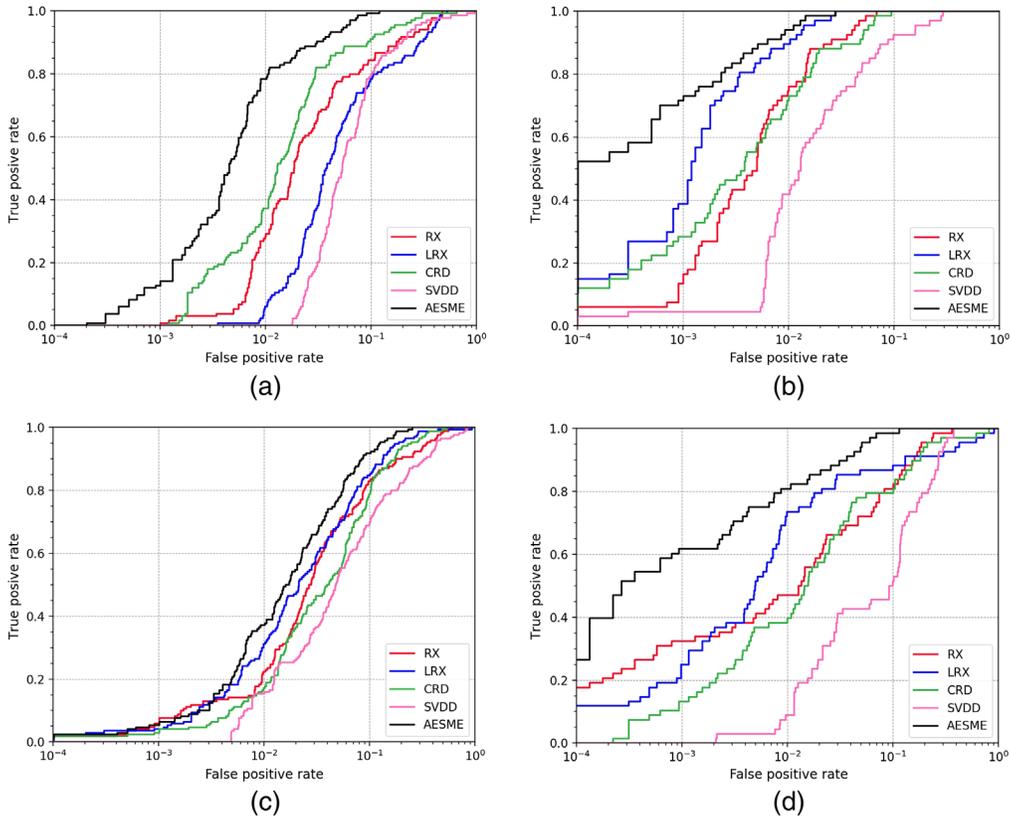
Fig. 9 ROC curves of anomaly detectors. (a) ROC curves of San Diego Airport; (b), (c), (d) are the ROC curves of Urban-1, Airport-3, and Beach-4 of ABU datasets, respectively.

Table 1 The AUC value of different AD methods for San Diego dataset.

HSIs	AESME	RX	LRX	CRD	SVDD
San Diego	0.9890	0.9403	0.9095	0.9638	0.9085

still has a slight false alarm rate for a small number of bright features in the background, which is caused by the false detection in the morphology extraction. For all experimental images, the detection maps of the proposed method are generally more visually similar to the reference images and also show a competitive performance in terms of the AUC scores of the results. Moreover, its ROC curve is the farthest from the diagonal of the graph in Figs. 9(a)–9(d), indicating its capability to distinguish anomalies. There is no significant difference in the discrimination effect of several detectors in Fig. 9(c).

4.3 *Results Discussion*

Since the overall result can be regarded as a combination of the three processes, we compared the performance when adding different components through the average AUC value of the ABU datasets. The AUC scores considering the effectiveness of different factors are shown in Table 3. In practice, when only the reconstruction error of the AE is used for AD, the average AUC score of ABU data can reach 0.951. Certainly, there is still potential for improvement. After adding the spatial filtering method, the AUC score increases by 0.024, and the detection effect is significantly improved. Considering the influence of the local spatial differences simultaneously, the average AUC score increased from 0.975 to 0.983, which achieved the ultimate effect of this method. By evaluating the detection effect of different factors, the influence of these processes

Table 2 The AUC value of different AD methods for ABU datasets.

HSIs	AESME	RX	LRX	CRD	SVDD
Urban-1	0.9980	0.9907	0.9968	0.9887	0.9643
Urban-2	0.9977	0.9946	0.9044	0.9335	0.8855
Urban-3	0.9966	0.9513	0.9800	0.9707	0.9137
Urban-4	0.9972	0.9887	0.9062	0.9728	0.8715
Urban-5	0.9814	0.9692	0.9352	0.9438	0.9062
Airport-1	0.9618	0.8221	0.9665	0.9649	0.7541
Airport-2	0.9798	0.8404	0.9847	0.9608	0.7926
Airport-3	0.9662	0.9288	0.9467	0.9313	0.8869
Airport-4	0.9920	0.9526	0.8740	0.8376	0.9401
Beach-1	0.9911	0.9807	0.9608	0.9667	0.9790
Beach-2	0.9420	0.9106	0.9616	0.9188	0.9457
Beach-3	0.9994	0.9998	0.9998	0.9994	0.9355
Beach-4	0.9910	0.9538	0.9391	0.9396	0.8904

Table 3 The average results of ABU datasets considering different factors during execution.

Factors considered in the methods	Reconstruction error only	Reconstruction error and morphological filtering	Reconstruction error, morphological filtering, and local discrepancy
AUC	0.95178	0.97560	0.98376

Table 4 The size of SE used for detection in morphological operations.

Size of SE	Data name for this size
(3, 3)	Beach-3 and Beach-2
(5, 5)	Beach-1, Beach-4, Airport-1, Airport-2, Airport-4, Urban-1, Urban-2, Urban-3, Urban-5, and San Diego
(7, 7)	Airport-3 and Urban-4

on the entire algorithm can be analyzed. In addition, Table 4 indicates the size of SE in the morphological filter when achieving the AUC values in Tables 1 and 2. The size of the targets in the image should be considered when using a specific size of SEs to extract the corresponding shape. For most datasets, the appropriate size of SE is (5, 5), and for a small portion of the data, the size of (3, 3) and (7, 7) is suitable. According to the effects to be achieved, the small size SE can better retain the objects in AD.

In addition, the influence of parameter settings on the performance of the algorithm is analyzed. The first parameter is the local window size, where the outermost neighborhoods are used as W_{out} to predict the center pixel. As shown in the left column of Fig. 10, the AUC value is used to evaluate the influence on the algorithm when the window size is set to (3, 1), (5, 3), (7, 5), (9, 7), and (11, 9). The horizontal axis in the figure is the size of the outer window, and the
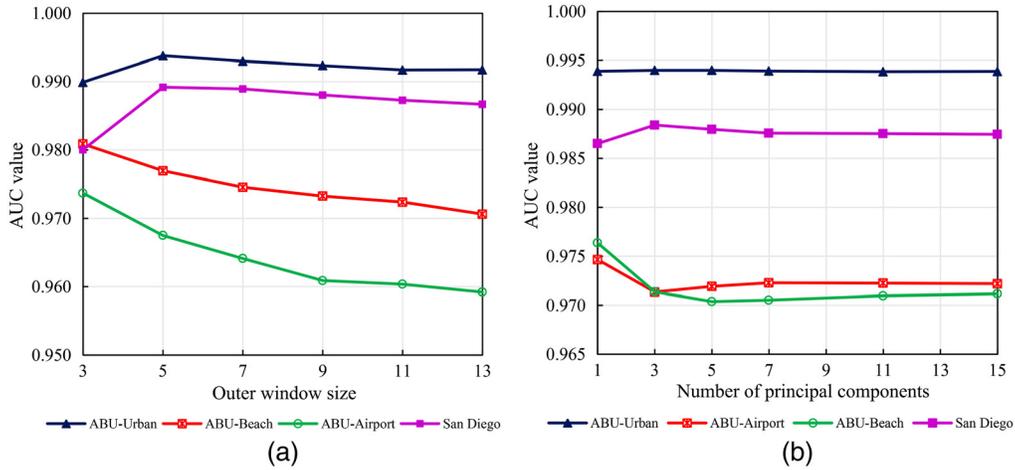
Fig. 10 The influence of parameter settings on the average AUC for different datasets.

corresponding size of the inner window is $W_{out} - 2$. It can be seen that for San Diego and ABU-Urban when the outer window is increased from 3 to 5, the detection performance improves significantly and then tends to decrease slightly. As the window increases for ABU-Airport and ABU-Beach, the AUC values present a continuous descending trend. Since the results of extracting local differences will include boundary interference in the background, this is also reflected in Fig. 4(a), where obvious boundaries with visual saliency are extracted. When the window is properly enlarged, the target can be better framed, but the extracted boundary range will also become broader, resulting in unstable or even lower AUC scores. For instance, ABU-Beach (1) and (2) have explicit boundaries, the AUC value decreases when the window size expands; when the ABU-Airport window becomes larger, the interference extraction components in the complex image also increase, so the AUC value decreases. By contrast, the background of San Diego airport and ABU-Urban dataset is simpler, so increasing the window size to fit the target has a more stable effect. The influence of the number of principal components is presented in the right column of Fig. 10. When the number of principal components is less than 3, the detection effect of different data fluctuates greatly. When the principal components continue to increase, the AUC score will tend to stabilize. Therefore, the principal components with stable detection results and small redundancy should be employed for hyperspectral data with different spectral resolutions.

5 Conclusion

A hyperspectral AD algorithm combining spatial morphological structure and reconstruction error is proposed in this paper. Our method constructs an AE network to learn the characteristics of HSI data. In this process, the reconstructed image obtained through the network that retains the typical features of the original input is adequate to calculate reconstruction errors with the original image. To identify the specific spatial morphology of anomalies, the output from the intermediate layer of the AE is transformed into a reduced-dimensional image, where a sliding window is established to refine the local spatial information. Meanwhile, the different spatial morphological filtering operations are introduced to extract the anomaly object of small targets. We combine the reconstruction errors map, the local analysis result of the sliding window, and the morphological filtering result to measure the contribution of each pixel of HSI. Experimental results reveal that the method proposed in this paper has better detection capabilities on real hyperspectral datasets in multiple scenes.

Acknowledgments

References

1. B. Rasti et al., "Noise reduction in hyperspectral imagery: overview and application," *Remote Sens.* **10**(3), 482 (2018).
2. J. Wang and C. Chang, "Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis," *IEEE Trans. Geosci. Remote Sens.* **44**(6), 1586–1600 (2006).
3. N. M. Nasrabadi, "Hyperspectral target detection: an overview of current and future challenges," *IEEE Signal Process. Mag.* **31**(1) 34–44 (2014).
4. J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geosci. Remote Sens.* **43**(3), 480–491 (2005).
5. L. He et al., "Recent advances on spectral–spatial hyperspectral image classification: an overview and new guidelines," *IEEE Trans. Geosci. Remote Sens.* **56**(3), 1579–1597 (2018).
6. S. Matteoli, M. Diani, and G. Corsini, "A tutorial overview of anomaly detection in hyperspectral images," *IEEE Aerosp. Electron. Syst. Mag.* **25**(7), 5–28 (2010).
7. J. M. Bioucas-Dias et al., "Hyperspectral unmixing overview: geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **5**(2), 354–379 (2012).
8. N. Ma et al., "An unsupervised deep hyperspectral anomaly detector," *Sensors* **18**, 693 (2018).
9. 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.* **38**(10), 1760–1770 (1990).
10. J. M. Molero et al., "Analysis and optimizations of global and local versions of the RX algorithm for anomaly detection in hyperspectral data," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **6**(2), 801–814 (2013).
11. M. J. Carlotto, "A cluster-based approach for detecting man-made objects and changes in imagery," *IEEE Trans. Geosci. Remote Sens.* **43**(2), 374–387 (2005).
12. H. Kwon and N. M. Nasrabadi, "Kernel RX-algorithm: a nonlinear anomaly detector for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.* **43**, 388–397 (2005).
13. N. M. Nasrabadi, "Regularization for spectral matched filter and RX anomaly detector," *Proc. SPIE* **6966**, 696604 (2008).
14. A. P. Schaum, "Hyperspectral anomaly detection beyond RX," *Proc. SPIE* **6565**, 656502 (2007).
15. A. Banerjee, P. Burlina, and C. Diehl, "A support vector method for anomaly detection in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.* **44**, 2282–2291 (2006).
16. W. Li and Q. Du, "Collaborative representation for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.* **53**(3), 1463–1474 (2015).
17. R. Zhao, B. Du, and L. Zhang, "Hyperspectral anomaly detection via a sparsity score estimation framework," *IEEE Trans. Geosci. Remote Sens.* **55**(6), 3208–3222 (2017).
18. Y. Zhang et al., "A low-rank and sparse matrix decomposition-based Mahalanobis distance method for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.* **54**(3), 1376–1389 (2016).
19. W. Xie et al., "Autoencoder and adversarial-learning-based semisupervised background estimation for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.* **58**(8), 5416–5427 (2020).
20. S. Wang et al., "Auto-AD: autonomous hyperspectral anomaly detection network based on fully convolutional autoencoder," *IEEE Trans. Geosci. Remote Sens.*, 1–14 (2021).
21. F. Andika, M. Rizkinia, and M. Okuda, "A hyperspectral anomaly detection algorithm based on morphological profile and attribute filter with band selection and automatic determination of maximum area," *Remote Sens.* **12**(20), 3387 (2020).
22. Y. Q. Zhao, W. H. Gui, and Z. C. Chen, "Edge detection based on multi-structure elements morphology," in *6th World Cong. Intell. Control and Autom.*, Dalian, China, pp. 9795–9798 (2006).

23. P. Salembier, "Structuring element adaptation for morphological filters," *J. Vision Commun. Image Represent.* **3**(2), 115–136 (1992).
24. J. Lei et al., "Spectral–spatial feature extraction for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.* **57**(10), 8131–8143 (2019).
25. J. Lei et al., "Discriminative reconstruction for hyperspectral anomaly detection with spectral learning," *IEEE Trans. Geosci. Remote Sens.* **58**(10), 7406–7417 (2020).
26. X. Kang, "Airport-beach-urban (ABU)," Hyperspectral anomaly detection with attribute and edge-preserving filters, http://xudongkang.weebly.com/data-sets.html (2017).
27. T. Jiang et al., "Discriminative reconstruction constrained generative adversarial network for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.* **58**(7), 4666–4679 (2020).

Jing Feng received a bachelor's degree in geographic information science from Beijing Union University in 2018 and worked in Beijing Institute of Surveying and Mapping the following year. She is now a master's student at the Capital Normal University, majoring in cartography and geographic information system. Her current research interest concerns hyperspectral remote sensing image processing.

Liyan Zhang is a lecturer at the Capital Normal University. She received her BS and MS degrees in mechanical engineering from the Shandong University of Technology and Beijing University of Technology in 2000 and 2004, respectively, and her PhD in mechanical engineering from the Beijing Institute of Technology in 2007. She finished her postdoctoral research in geography from the Capital Normal University in 2019. Her current research interests include hyperspectral imagery and its application.