remote sensing

MDPI

Article

Dual-View Hyperspectral Anomaly Detection via Spatial Consistency and Spectral Unmixing

Jingyan Zhang [1,*], Xiangrong Zhang [2] and Licheng Jiao [2]

[1] School of Communications and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China
[2] School of Artificial Intelligence, Xidian University, Xi'an 710071, China; xrzhang@mail.xidian.edu.cn (X.Z.)
* Correspondence: jyzhang@xupt.edu.cn

Abstract: Anomaly detection is a crucial task for hyperspectral image processing. Most popular methods detect anomalies at the pixel level, while a few algorithms for anomaly detection only utilize subpixel level unmixing technology to extract features without fundamentally analyzing the anomalies. To better detect and separate the anomalies from the background, this paper proposes a dual-view hyperspectral anomaly detection method by taking account of the anomaly analysis at both levels mentioned. At the pixel level, the spectral angular distance is adopted to calculate the similarities between the central pixel and its neighbors in order to further mine the spatial consistency for anomaly detection. On the other hand, from the aspect of the subpixel level analysis, it is considered that the difference between the anomaly and the background usually arises from dissimilar endmembers, where the unmixing will be fully implemented. Finally, the detection results of both views are fused to obtain the anomalies. Overall, the proposed algorithm not only interprets and analyzes the anomalies from dual levels, but also fully employs the unmixing for anomaly detection. Additionally, the performance of multiple data sets also confirmed the effectiveness of the proposed algorithm.

Keywords: hyperspectral images; anomaly detection; spatial consistency; spectral unmixing; manifold constraint

check for updates

Citation: Zhang, J.; Zhang, X.; Jiao, L. Dual-View Hyperspectral Anomaly Detection via Spatial Consistency and Spectral Unmixing. *Remote Sens.* **2023**, *15*, 3330. https://doi.org/10.3390/rs15133330

Academic Editor: Salah Bourennane

Received: 4 April 2023
Revised: 3 June 2023
Accepted: 5 June 2023
Published: 29 June 2023

1. Introduction

In recent decades, the advent of recent remote sensing technologies has enabled various remote sensors to collect data in a more convenient way. Consequently, image processing techniques also experienced a rapid development and found widespread applications in many fields. A hyperspectral image (HSI) consists of approximately a hundred or more contiguous spectral bands, and each pixel can extract a whole complete high-resolution spectral curve [1,2]. With this feature, it contains more precise and accurate spectral information, allowing for a better characterization and identification of targets. Therefore, the HSI has attracted significant academic interest due to its competitive advantage and potential application in various industries [3–5].

In general, the most widely used processing techniques of hyperspectral remote images can be divided into several classes including mixed-pixel decomposition (unmixing) [6,7], land-cover classification [8,9], anomaly detection [10,11], target detection [12,13], and so forth. Although the hyperspectral sensor has a very high spectral resolution, its spatial resolution is limited. Moreover, the land-cover is complex, and these factors would lead to the presence of the mixed pixels. Hyperspectral unmixing (HU) aims to decompose the mixed pixel into numerous ground objects (endmembers) and determine their corresponding proportions (abundances) within this pixel [14]. The goal of anomaly detection is to identify the uncommon objects that are significantly different from the background [15].

For the anomaly detection task, the anomaly targets in the image are unknown, making it difficult to obtain sufficient prior knowledge. Therefore, the algorithms for anomaly detection can vary depending on the approach used to recognize the anomalies. Hyperspectral anomaly detection methods can generally be categorized into the following two categories based on the involvement of neural networks in the model construction phase: traditional methods and deep learning-based methods [16].

On one hand, traditional algorithms mainly include statistics-based methods, distance-based methods, collaborative representation-based methods, etc. Among these, the Reed–Xiaoli (RX) algorithm [17], which comprises the global RX detection (GRX) algorithm and the local RX detection (LRX) algorithm, is widely applied. This algorithm primarily focuses on the distribution characteristics of the anomaly points, assuming that the background follows a multivariate normal distribution. It utilizes the Mahalanobis distance to measure the differences between a test pixel and its neighbor pixels for anomaly detection [18]. Building upon the RX algorithm, several improved algorithms were developed to enhance the detection accuracy for complex backgrounds, such as in [19–21]. In addition, there are other statistics-based methods, such as single-feature anomaly detection [22], Gaussian Markov random field [23], and so on. For distance-based methods, pixels grouped based on their distance and pixels deviating from the cluster center are considered abnormal [24]. Support vector machine (SVM)-based methods are typical distance-based methods that use a one class classifier to estimate the smallest closed hypersphere, i.e., one-class SVM (OCSVM) [25]. The background samples support the hypersphere, and the pixels beyond the hypersphere model are considered anomalies [26]. Furthermore, the clustering-based method [27,28], such as the density-based methods [29], are adopted for hyperspectral anomaly detection, which is an important component of distance-based methods. These methods first cluster the original data, and then estimate the anomaly degree of the testing pixel [24]. Density-based methods, such as local outlier factor [30], connectivity-based outlier factor [31], etc., assume that the cluster density of the normal sample points is higher than that of the abnormal sample points. Moreover, the collaborative representation detection (CRD) for anomalies has received significant attention [32]. Li et al. [33] proposed a collaborative representation and a kernel version for hyperspectral anomaly detection. The algorithm is based on the concept that each background pixel can be represented approximately by its spatial neighborhood, whereas anomalies cannot. A hyperspectral anomaly detection method performed by combining the collaborative representation with the principal component analysis, suggested by Su et al. [34], turned out to achieve a better performance.

On the other hand, with the continuous development of deep learning, anomaly detection algorithms based on deep learning became a new research hotspot in recent years [35]. Depending on whether sample support is needed, deep learning-based methods can be divided into two categories, supervised and unsupervised. The typical representative of the former category is based on the convolutional neural network, and a series of related algorithms were proposed, such as the transferred deep convolutional neural network [36], the algorithm of convolutional neural network and low rank with density-based clustering [37], etc. The unsupervised deep learning-based algorithms mainly include autoencoder networks and generative adversarial networks, and lots of methods were proposed [38,39], which do not need prior samples to be suitable for hyperspectral anomaly detection.

Some scholars presented new ideas for anomaly detection that consider multiple processing technologies of the HSI. Qu et al. proposed a detection algorithm based on spectral unmixing, which utilizes the abundances as distinctive features to construct the background dictionary for HSI anomaly detection. They referred to the proposed anomaly detection algorithm as abundance- and dictionary-based low rank decomposition (ADLR) [40]. Huang et al. extracted features in the HSI at three levels, which include the subpixel level, pixel level, and super-pixel level. They performed the detection using an existing low-rank and sparse matrix decomposition algorithm, and constructed weight maps for the fusion to obtain the final result [41]. Acar et al. believed that anomalies are small, rare objects or materials with different spectral characteristics compared to their surroundings. They introduced an anomaly detection method using sparse unmixing and a Gaussian mixture model for the HSI [42]. These methods mentioned above provide a novel research direction for detecting anomalies, with many of them utilizing unmixing technology to extract features for further detection. However, they do not deeply analyze the mechanism of the anomalies on a subpixel level.

This paper proposes a dual-view anomaly detection algorithm (DVAD) that considers the anomaly characteristics at different levels by incorporating spatial consistency and spectral unmixing based on the local and global information in the HSI. On one side, similar to most algorithms, we believe that anomalies are pixels that differ from their surrounding pixels. To accurately explore the differences between the anomaly and its neighbors, we introduce the spectral angular distance (SAD) [43] to access the spatial consistency. The SAD measures the spectral similarity by calculating the angle between the spectra, and it is not affected by the spectral scale, either. Thus, we calculate the sum of the SAD values for the center pixel and its neighbors. When the sum is large, indicating a significant difference between the center pixel and its neighbors, we classify the center pixel as an anomaly pixel. On the other side, from the subpixel level view, the most essential difference between the anomaly and the background lies in the variation of the endmembers. The endmembers can be divided into anomaly endmembers and background endmembers. Anomaly endmembers only participate in the mixing of anomaly pixels, which are rare in distribution. A novel manifold-constrained sparse spectral unmixing method based on the non-negative matrix factorization (NMF) model [44] is proposed to obtain the endmembers and their corresponding abundances. For each endmember, we count the number of small values in its corresponding abundance, and when the number approaches the total number of pixels, it is considered to be the anomaly endmember for anomaly detection. Thus, by analyzing the anomaly characteristics at the subpixel level, the internal relationship between the tasks of spectral unmixing and anomaly detection were exploited. Finally, when its detection results of both two views are anomalies, the pixel will be considered as the anomaly. To achieve this goal, we fuse the detection results from both views, indicating an anomaly. For this fusion process, we adopt a strategy of multiplying the detection results from different levels. Although the fusion strategy is strict and may reduce the detection result, it effectively integrates the advantage of each level, resulting in a more reliable detection outcome.

In summary, the proposed DVAD algorithm, as depicted in Figure 1, not only performs the anomaly detection at the pixel and subpixel levels, but also applies the spectral unmixing technology into the anomaly detection task in a more meaningful manner. The effectiveness of the algorithm is demonstrated through experiments on multiple data. The main contributions of the proposed algorithm are as follows:

(1) A dual-view anomaly detection method via spatial consistency and spectral unmixing from the pixel level and subpixel level is presented, which makes full use of the local and global information in the HSI.
(2) Taking the characteristics of the HSI into account, a novel manifold-based sparse spectral unmixing algorithm is put forward.
(3) This paper proposes, for the first time, that the difference between the anomaly and background mainly comes from the difference of endmembers based on the subpixel level analysis.

The remainder of this paper is organized as follows. Section 2 presents the theoretical background. In Section 3, the proposed method is exhibited in detail. Section 4 reports the experimental results and provides a discussion. Finally, Section 5 concludes the paper and discusses future work.

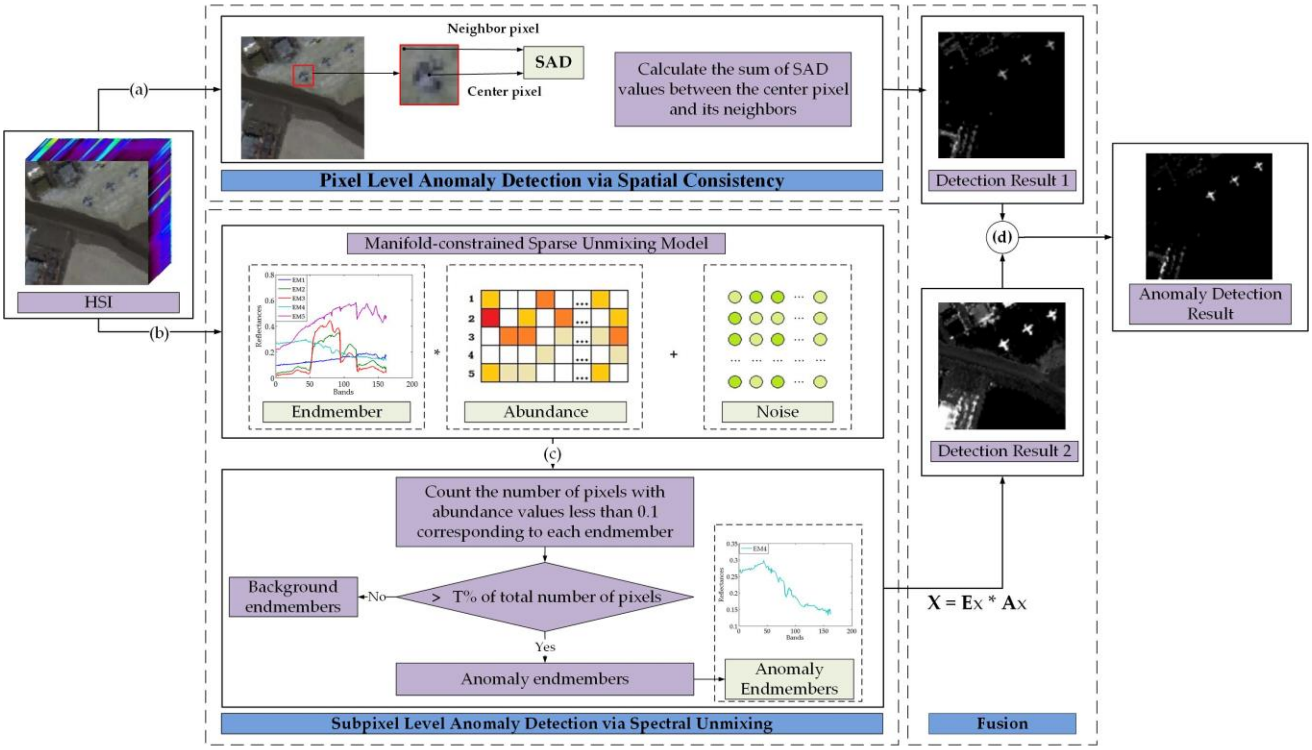


Figure 1. The flowchart of the proposed DVAD method; (**a**) pixel level detection, (**b**) spectral unmixing, (**c**) subpixel level detection, and (**d**) decision fusion.

2. Theoretical Background

The HSI data are commonly broken up into three parts, including the background, anomaly, and noise [11,15], whose model can be formulated as follows:

$$\mathbf{Y} = \mathbf{B} + \mathbf{X} + \mathbf{N} \tag{1}$$

where $\mathbf{Y} \in \mathrm{R}^{L\times P}$ is the hyperspectral matrix with L spectral bands and P pixels, $\mathbf{B} \in \mathrm{R}^{L\times P}$ and $\mathbf{X} \in \mathrm{R}^{L\times P}$ are respectively the background and the anomaly, and $\mathbf{N} \in \mathrm{R}^{L\times P}$ is the noise matrix.

The linear spectral mixture model is a popular model that is widely used in HU. The matrix form for the linear unmixing model can be written as follows:

$$\mathbf{Y} = \mathbf{EA} + \mathbf{N} \tag{2}$$

where $\mathbf{E} \in \mathrm{R}^{L\times K}$ is the endmember matrix of the K endmembers, and $\mathbf{A} \in \mathrm{R}^{K\times P}$ is the abundance matrix. The following two constraints of abundance usually need to be satisfied for unmixing process: (1) the abundance is non-negative (ANC) and (2) the sum of each column of abundance is one (ASC).

Based on the linear mixture model, the NMF is applied for HU. The basic idea of NMF model is that, for any non-negative matrix, it can find two non-negative matrices that their product is this non-negative matrice. The objective function in the view of Euclidean distance is as follows:

$$\begin{aligned} \min_{\mathbf{E},\mathbf{A}} \quad & \tfrac{1}{2}\|\mathbf{Y} - \mathbf{EA}\|_F^2 \\ s.t. \quad & \mathbf{E} \geq 0, \mathbf{A} \geq 0 \end{aligned} \tag{3}$$

where operator $\|\cdot\|_F$ represents the Frobenius norm. While there are many optimization algorithms to estimate $\mathbf{E}$ and $\mathbf{A}$, it is still difficult to obtain a globally optimal solution because of the non-convexity of Formula (3) with respect to both $\mathbf{E}$ and $\mathbf{A}$. Furthermore, the NMF is always utilized with other constraints, due to the fact that the NMF lacks a unique solution. The iterative algorithm for the minimization of the objective function [45] in Formula (3) is as follows:

$$\begin{aligned} \mathbf{E} &\leftarrow \mathbf{E}. * \mathbf{Y}\mathbf{A}^T./\left(\mathbf{E}\mathbf{A}\mathbf{A}^T\right) \\ \mathbf{A} &\leftarrow \mathbf{A}. * \mathbf{E}^T\mathbf{Y}./\left(\mathbf{E}^T\mathbf{E}\mathbf{A}\right) \end{aligned} \tag{4}$$

where $(\cdot)^T$ denotes the matrix transpose, and .* and ./ are the element-wise multiplication and division.

3. Dual-View Anomaly Hyperspectral Detection via Spatial Consistency and Spectral Unmixing

In fact, anomalies usually refer to pixels that are less distributed and not similar to the neighbor pixels [42], which is the main analysis and basic understanding of the pixel level view. Moreover, when it comes to the subpixel level, the dissimilarities between the anomaly and background are mainly reflected in different endmembers. There are anomaly endmembers and background endmembers. For instance, an anomaly endmember only participates in the mixing of the anomaly pixel. Based on the above two analyses, a method of anomaly detection that utilizes spatial consistency and spectral unmixing is proposed. In particular, this section briefly introduces the proposed algorithm in detail, including the algorithm construction, model analysis, and solution.

3.1. Pixel Level Anomaly Detection via Spatial Consistency

With regard to the pixel level, as many papers assumed, the anomalies with few distributions are considered to be the pixels that are not similar to their surrounding pixels. Firstly, we perform a sliding window to obtain the neighbors of the central pixels. Secondly, we measure the spatial consistency between the central pixel and its surrounding neighbors. Additionally, the SAD is a metric that is commonly used as a measurement for the similarities between the spectra by calculating their angles. It will not be affected by the spectral scale, either. The SAD used to measure the similarity between the pixel and its neighbors is defined as follows:

$$\mathrm{SAD} = \cos^{-1}\left(\frac{\widetilde{\mathbf{E}}^T\mathbf{E}}{\left\|\widetilde{\mathbf{E}}\right\|\left\|\mathbf{E}\right\|}\right) \tag{5}$$

where $\mathbf{E}$ refers to a spectrum, and $\widetilde{\mathbf{E}}$ is another spectrum or the estimated spectrum of $\mathbf{E}$. When the value of the SAD is small, the angle between two spectra is small, which means they are similar.

The main steps of the anomaly detection process are as follows. First of all, we calculate the SAD value between the central pixel and its neighbors. Then, we calculate the sum of the SAD value of the center pixel and its neighbors as the anomaly value of the center pixel for detection. In the experiment, the dual window shown in Figure 2 is adopted for the center pixel, and the pixels between the inner and outer windows are considered as its neighbors. Moreover, due to the different sizes of the abnormal targets in different data sets, the size of the dual window can be changed to find the exact neighbors for each pixel.

The distribution of anomalies is rare; thus, there are four circumstances to be studied, as shown in Figure 3.

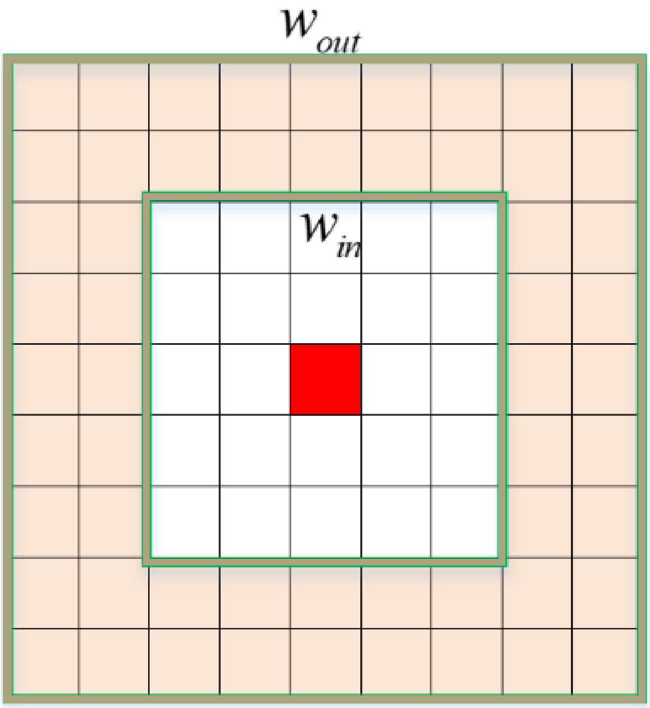


Figure 2. Dual window schematic diagram. Red represents the center pixel, and beige between green boxes represents neighbor pixels.

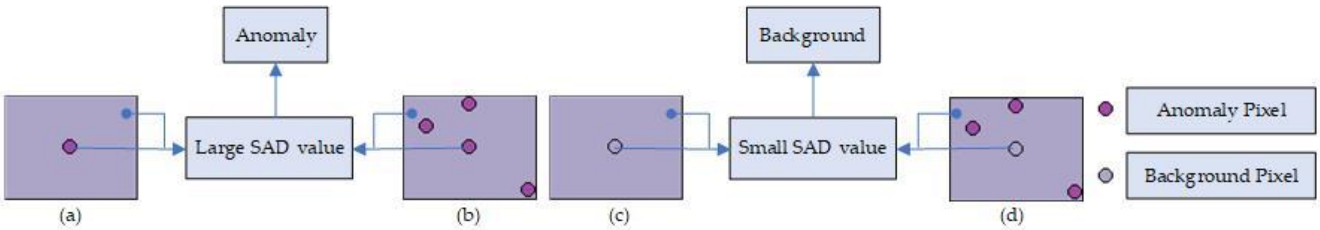


Figure 3. Four situations (**a**–**d**) during anomaly detection of pixel level.

(a) The center pixel is the anomaly, and the neighbor pixels are the background. The SAD values are all large, and the center pixel with a large sum of SAD values is deemed as the anomaly pixel.
(b) The center pixel is the anomaly, and the neighbor pixels contain a small number of anomalies. However, most neighbor pixels are still the background. Thus, most of the SAD values are large, and the center pixel with a large sum of SAD values is also detected as the anomaly.
(c) The center pixel is the background, and the neighbor pixels are the background. The SAD values are all small, which illustrates the similarity of the center pixel to its neighbors. The center pixel with a small sum of SAD values is viewed as the background.
(d) The center pixel is the background, and the neighbor pixels contain a small number of anomalies. Most of the neighbor pixels are the background, and the corresponding SAD values are small. The center pixel with a small sum of SAD values is detected as the background.

From the above-mentioned analysis, it can be seen that the proposed detection method can perform anomaly detection efficiently under various situations.

3.2. Subpixel Level Anomaly Detection via Spectral Unmixing

3.2.1. Manifold-Constrained Sparse Spectral Unmixing

In light of the subpixel level, the endmembers are largely responsible for the dissimilarities between the anomaly and the background. At this point, the endmember can be divided into the anomaly endmember and the background endmember. The anomaly endmember only participates in the mixing of the anomaly pixels, whose distribution is few. Following this, the spectral unmixing technology can be applied in the task of the anomaly detection. Then, the first step is to obtain the endmembers and abundances via spectral unmixing. The construction process of the objective function for unmixing is explained in further detail.

The proposed unmixing method is based upon the NMF model, which is a regular employed model for hyperspectral unmixing. As the objective function of the NMF is non-convex, the solution can easily fall into local optimum. Having said that, some other constraints are added to the model to address this issue. Many related studies [44,46,47] also revealed that only part of the endmembers in the mixing participate during the unmixing process. In view of the characteristics of sparsity, we added the sparse constraint in the model. The basic model of the sparse NMF algorithm works as follows:

$$\begin{array}{ll} \min\limits_{\mathbf{E},\mathbf{A}} & \frac{1}{2}\|\mathbf{Y}-\mathbf{E}\mathbf{A}\|_F^2+\alpha\|\mathbf{A}\|_{\frac{1}{2}} \\ s.t. & \mathbf{A}\geq 0,\ \mathbf{1}_K^T\mathbf{A}=\mathbf{1}_P^T \end{array} \tag{6}$$

where α represents the sparsity factor. The objective function described in Formula (6) is convex, with respect to the individual parameters $\mathbf{E}$ and $\mathbf{A}$, and the most popular algorithms for solving this NMF are the iterative ones [44]. The expected solution formula can be captured using the solution algorithm of the NMF [45], which is estimated as follows.

$$\begin{array}{c} \mathbf{E}\leftarrow\mathbf{E}.*\mathbf{Y}\mathbf{A}^T./\left(\mathbf{E}\mathbf{A}\mathbf{A}^T\right) \\ \mathbf{A}\leftarrow\mathbf{A}.*\mathbf{E}^T\mathbf{Y}./(\mathbf{E}^T\mathbf{E}\mathbf{A}+\frac{\alpha}{2}\mathbf{A}^{-\frac{1}{2}}) \end{array} \tag{7}$$

As is well known, an HSI is a kind of high-dimensional datum, and it tends to locate in the low-dimensional subspace embedded in high-dimensional space [48]. The manifold learning digs into the essence of the data, and discovers the inherent laws and potential characteristics. In addition, the low-dimensional manifold features were well studied for HSI processing. Therefore, to exploit the potential of manifold learning, we added it to the model.

The main idea is that if the two pixels $\boldsymbol{y}_i$ and $\boldsymbol{y}_j$ are very close, their representations $\boldsymbol{a}_i$ and $\boldsymbol{a}_j$ in the low-dimensional subspace, i.e., the abundance space, will be as close as possible. All points are constructed as a graph, provided that each pixel is a node. Only the nearest several points of the node, which are its most similar points, are connected with it [49]. Following this, we usually construct the connection weight between two points to effectively benefit from the characteristics of the HSI. On top of that, the weight function should be able to clearly distinguish the similarities between one point and the rest of the points by different weight values. The sigmoid function with good discrimination was proven to be a commonly used widely applied one in hyperspectral image processing. Thus, we construct the weight function by deforming the sigmoid function. If the point i is linked to the point j, the connection weight between them is estimated as follows.

$$\mathbf{W}_{i,j}=2/\left(1+\exp\left(\left\|\boldsymbol{y}_i-\boldsymbol{y}_j\right\|^2/\sigma\right)\right) \tag{8}$$

From Formula (8), it can be noted that when two points are more similar, their connection weight is closer to one. However, if they are dissimilar, the connection weight is closer to zero. That said, these two points are not connected.

It is hoped that when the two points $\boldsymbol{y}_i$ and $\boldsymbol{y}_j$ are similar in the original space, their representations $\mathbf{a}_i$ and $\mathbf{a}_j$ in the abundance space are also similar. To that end, the constraint is constructed as follows:

$$\begin{aligned}\frac{1}{2}\sum_{i,j=1}^{N}\left\|\mathbf{a}_i-\mathbf{a}_j\right\|^2\mathbf{W}_{ij} &= \sum_{i=1}^{N}\mathbf{a}_i^T\mathbf{a}_i\mathbf{D}_{ii}-\sum_{i,j=1}^{N}\mathbf{a}_i^T\mathbf{a}_j\mathbf{W}_{ij} \\ &= \mathrm{Tr}\left(\mathbf{ADA}^T\right)-\mathrm{Tr}\left(\mathbf{AWA}^T\right)=\mathrm{Tr}\left(\mathbf{ALA}^T\right)\end{aligned} \tag{9}$$

where $\mathrm{Tr}(\cdot)$ represents the trace of the matrix, $\mathbf{D}_{ii}=\sum_{j=1}^{P}\mathbf{W}_{ij}$, and $\mathbf{L}=\mathbf{D}-\mathbf{W}$. The manifold constraint on abundance is added to the sparse unmixing model to form the following objective function:

$$\begin{aligned}&\min_{\mathbf{E},\mathbf{A}} \quad \frac{1}{2}\|\mathbf{Y}-\mathbf{EA}\|_F^2+\alpha\|\mathbf{A}\|_{\frac{1}{2}}+\beta Tr\left(\mathbf{ALA}^T\right) \\ &s.t. \quad \mathbf{A}\geq 0, 1_K^T\mathbf{A}=1_P^T\end{aligned} \tag{10}$$

where β is the control parameter. The objective function in Formula (10) is also convex with respect to the individual parameters $\mathbf{E}$ and $\mathbf{A}$, and according to the iterative algorithm of NMF model, the solution formulas of the endmember and abundance are set as shown below.

$$\begin{aligned}&\mathbf{E}\leftarrow\mathbf{E}.*\mathbf{YA}^T./\left(\mathbf{EAA}^T\right) \\ &\mathbf{A}\leftarrow\mathbf{A}.*\left(\mathbf{E}^T\mathbf{Y}+\beta\mathbf{AW}\right)./\left(\mathbf{E}^T\mathbf{EA}+\frac{\alpha}{2}\mathbf{A}^{-\frac{1}{2}}+\beta\mathbf{AD}\right)\end{aligned} \tag{11}$$

Furthermore, with the purpose of making the solution of abundance to satisfy the ASC, one more row is added to the observation matrix and the endmember matrix as follows:

$$\mathbf{Y}_f=\begin{bmatrix}\mathbf{Y}\\ \varepsilon\mathbf{1}_P^T\end{bmatrix},\ \mathbf{E}_f=\begin{bmatrix}\mathbf{E}\\ \varepsilon\mathbf{1}_K^T\end{bmatrix} \tag{12}$$

where ε controls the convergence rate of the solution. Finally, taking the ASC into account, we replace $\mathbf{Y}$ and $\mathbf{E}$ with $\mathbf{Y}_f$ and $\mathbf{E}_f$. Then, the obtained iterative formula of abundance is as below.

$$\mathbf{A}\leftarrow\mathbf{A}.*\left(\mathbf{E}_f^T\mathbf{Y}_f+\beta\mathbf{AW}\right)./\left(\mathbf{E}_f^T\mathbf{E}_f\mathbf{A}+\frac{\alpha}{2}\mathbf{A}^{-\frac{1}{2}}+\beta\mathbf{AD}\right) \tag{13}$$

After this, we obtained the endmember and abundance by considering the different characteristics for the unmixing task.

3.2.2. Subpixel Level Anomaly Detection

From the subpixel level, the difference between the anomaly and the background usually lies in the difference of the endmember. In other words, the endmember can be divided into the background endmember $\mathbf{E}_B$ and the anomaly endmember $\mathbf{E}_X$. The main basic model for anomaly detection is now illustrated as follows:

$$\mathbf{Y}=\mathbf{E}_B*\mathbf{A}_B+\mathbf{E}_X*\mathbf{A}_X+\mathbf{N} \tag{14}$$

where $\mathbf{A}_B$ and $\mathbf{A}_X$ are respectively the background abundance and the anomaly abundance. It is easily observed that we can reconstruct the anomaly $\mathbf{X}$ in Formula (1) by using the second term of Formula (14).

Consequently, it is necessary to analyze the endmember and its corresponding abundance for anomaly detection. As mentioned from the previous analysis, the anomaly endmember can only participate in the mixing of the anomaly pixels, and their distribution could be very small. Hence, the distribution of each endmember is counted by analyzing its corresponding abundance. To be more specific, a small threshold is set first. As for one endmember, when its corresponding abundance value is less than this threshold at a certain pixel, this endmember is not considered to participate in the mixing of this pixel.

That is, this endmember is not distributed at this pixel. From there, we count the number of the abundance value that is smaller than the threshold for each endmember. Likewise, when the number is almost up to 90% of the total pixel number, the distribution is few, and then this endmember is regarded as the anomaly endmember.

The anomaly is reconstructed using Formula (14), and the final detection result can be achieved through the L_2 norm, which is commonly used for anomaly detection. The anomaly value of each pixel in the detection map is obtained as follows [10,15]:

$$\mathrm{R}(\mathbf{y}_i) = \|\mathbf{X}_{:,i}\|_2 = \sqrt{\sum_j (\mathbf{X}_{j,i})^2} \quad (15)$$

where when the anomaly value of a pixel is high, it is more likely to be an anomaly pixel.

3.3. Fusion for Pixel and Subpixel Levels Anomaly Detection

We analyze the characteristics of the anomaly and perform anomaly detection from the pixel and subpixel levels via spatial consistency and spectral unmixing, respectively. In order to further improve the detection results, the results based on different levels should be used to obtain the final anomaly detection. The general idea of fusion is that the final detection result will be the anomaly only when the detection results of two levels are anomalies. Here, the fusion method is to multiply the detection results of two levels. Thus, the detection result of the proposed method does not count on the result of a certain level. Although this fusion method is strict and might reduce the detection result, it ensures the reliability of the final result, which is very important for anomaly detection.

Above all, the entire DVAD algorithm was outlined in greater detail, and the whole process is reviewed in Figure 1. The DVAD algorithm not only merely analyzes the anomaly characteristics from the pixel and subpixel levels via spatial consistency and spectral unmixing, but also thoroughly combines the tasks of unmixing and anomaly detection.

4. Experiments Results

In order to start the process of verifying the effectiveness of the proposed algorithm, we designed a series of experiments on multiple data sets in this section. The subsequent arrangements are performed in three main steps. The first part mainly focuses on the performance metrics and comparison algorithms, including the traditional anomaly detection algorithms, namely, GRX, LRX, and OCSVM, as well as the related algorithms comprising the representation-based CRD and the unmixing-based ADLR. Consequently, we review the performance and analyze the effectiveness of each view and the overall algorithm. Finally, we draw conclusions in the parameter analysis part.

4.1. Data Set

The first data set of real-world anomaly detection as the test image is a subscene of the airport in San Diego, USA, which is a popular data set collected by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor [10,50,51]. It has 189 bands and 100×100 pixels, in which the bands with water absorption regions, low SNR, and poor quality are removed. The reference anomaly map, the false color image, and the ground truth are presented in Figure 4a,e,i, respectively. The buildings with different roofs, parking aprons with different materials, an airport runway, and a small quantity of vegetation are considered as the main land types. Additionally, the airplanes in the image are regarded as the anomaly target to be detected. Fifty-seven pixels were selected as anomalies, composed of full-pixel anomalies in the main body of the airplanes and subpixel targets on the edges of the airplanes [10]. Although the spectral of aircraft is largely different from that of the background, it is somewhat correlated with the spectra of roof and shadow. It leads to the difficulty for anomaly detection [52].

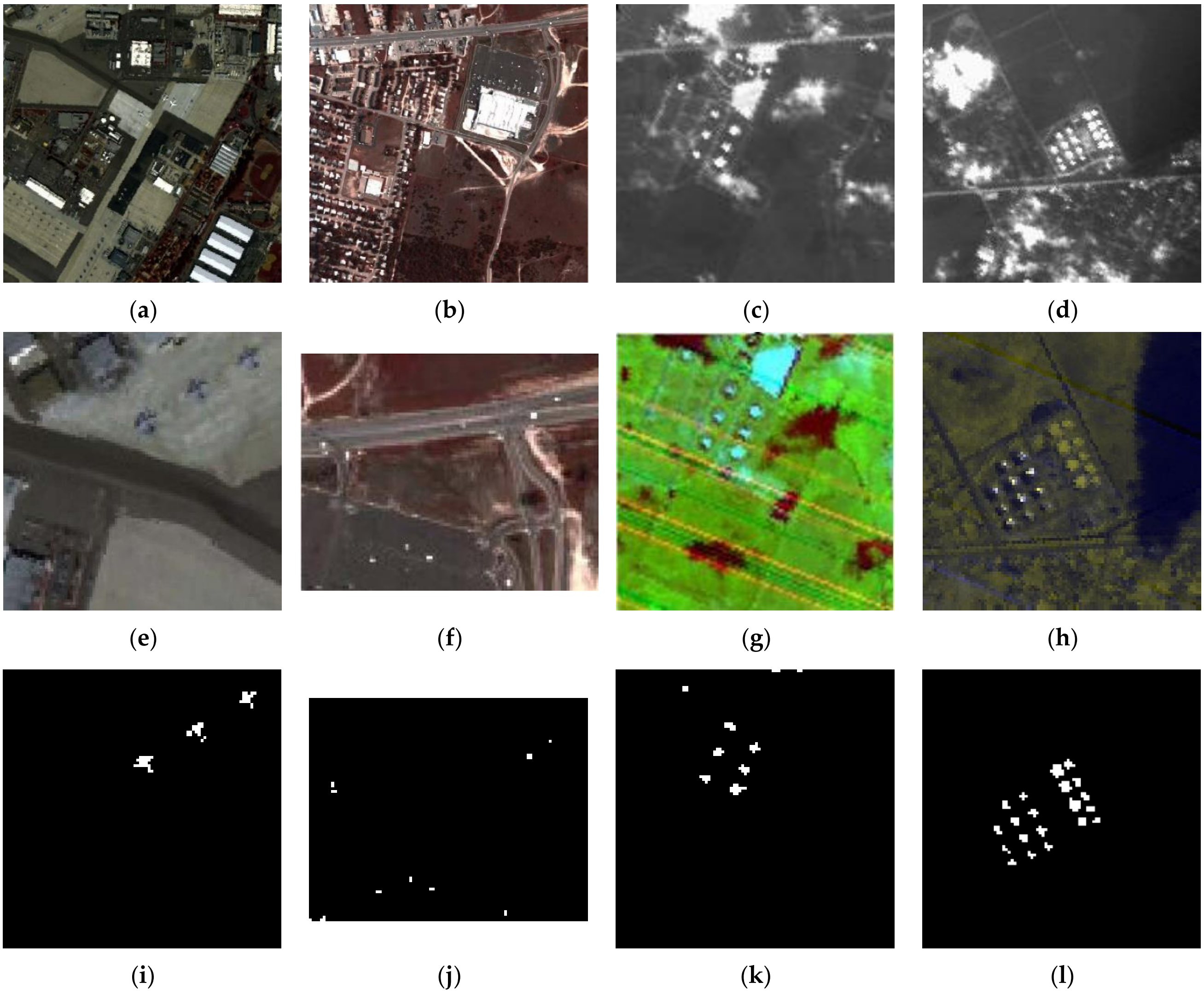


Figure 4. The illustration of data sets and their corresponding ground truths. (**a**) Reference anomaly map of San Diego data set; (**b**) reference anomaly map of HYDICE data set; (**c**) reference anomaly map of Texas data set; (**d**) reference anomaly map of Urban data set; (**e**) false color image of San Diego data set; (**f**) false color image of HYDICE data set; (**g**) false color image of Texas data set; and (**h**) false color image of Urban data set. (**i**) Reference detection map of San Diego data set; (**j**) reference detection map of HYDICE data set; (**k**) reference detection map of Texas data set; and (**l**) reference detection map of Urban data set.

The second data set in the experiment is captured by the Hyperspectral Digital Imagery Collection Experiment (HYDICE) airborne sensor covering the urban area [53], which comprises a vegetation area, a construction area, and several roads including some vehicles. Its spectral and spatial resolution are, respectively, 10 nm and 1 m. There are 175 spectral bands remaining after the removal of the water absorption bands (1–4, 76, 87, 101–111, 136–153, and 198–210). The subscene with 80 $\times$ 100 pixels is picked from an entire scene of 307 $\times$ 307 pixels. Additionally, 21 anomalous target pixels in urban scene are roof and cars of different sizes [3,33]. The ground truth defines that the anomalous targets in [10] are the cars and roofs embedded in the different backgrounds in the upper rightmost area of the scene, and the considered subscene consists of pixels covering this area. The original anomaly map, color representation, and ground truth map are illustrated in Figure 4b,f,j.

The third data set is from an open Airport–Beach–Urban data set, which was collected over the Texas Coast on 29 August 2010. The sample image with a size of 100 $\times$ 100 pixels is manually extracted from large images downloaded from the AVIRIS website [54], whose reference maps are manually labeled with the help of the Environment for Visualizing Images (ENVI) software. It contains 67 anomaly pixels, and its spatial resolution is 17.2 m per pixel with 204 bands remaining after removing the noise bands in the original images by using a recently published noise level estimation method [3]. The original anomaly map, color representation, and reference anomaly map of Texas data set are shown in Figure 4c,g,k.

The final data set, which is called the Urban data set, was also collected over the Texas Coast in 2010, and its procedure of the reference map creation is the same as the third data set. The original anomaly map acquired by the AVIRIS sensor [3,55], false color image, and reference map are displayed in Figure 4d,h,l, respectively. It has 100 × 100 pixels with 155 anomaly pixels and 207 bands remaining after the removal of noise bands, whose spatial resolution is 17.2 m per pixel. Except for the second data set that covers a spectral range of 0.4~2.5 µm, all the other data sets were collected using the AVIRIS sensor, which provides a spatial resolution of 20 m and a spectral resolution of 10 nm, covering a spectral range of 0.4~2.5 µm.

4.2. Comparison Methods and Performance Metrics

In our experiment, four representative anomaly detection methods are selected to be compared with the proposed DVAD algorithm to evaluate its performance, particularly with regard to the GRX, LRX, OCSVM, CRD, and ADLR. The GRX and LRX algorithms are widely considered as two traditional methods for anomaly detection, which belong to the classical statistic-based methods. The OCSVM method, a traditional algorithm for anomaly detection, maps the data to the feature space corresponding to the kernel, builds a hyperplane between the data and the origin, and maximizes the distance from the hyperplane to the zero point. Given the fact that the proposed pixel level detection method is similar to the method based on representation, we use the CRD algorithm as a comparison in our following experiments. The ADLR algorithm only employs the unmixing technique to extract features, and these features will be further used as the input for a low-rank sparse model to detect the anomaly. It is an unmixing-related anomaly detection algorithm based on the subpixel level analysis. Furthermore, to further illustrate its efficiency, we also make a contrast to the detection algorithms of the two levels in the DVAD algorithm.

Moreover, with the purpose of evaluating the performance of the different mentioned algorithms, we present two metrics that were extensively applied in anomaly detection. One is the receiver operating characteristic (ROC) curve, and the other is the area under ROC curve (AUC) [56], whose value is usually smaller than one. When a good performance for the algorithm is achieved, the ROC curve is closer to the upper left corner, and correspondingly, the AUC value is higher.

4.3. Anomaly Detection Performance

In this section, we objectively evaluate the detection performance of the proposed DVAD algorithm and make a comparison with four state-of-the-art detection methods, including the GRX, LRX, OCSVM, CRD, and ADLR. The size of the dual window (win_{in}, win_{out}) for the LRX and CRD is the same as (5, 7) for the HYDICE data set and the Urban data set, (9, 13) for the San Diego data set, and (3, 9) for Texas data set. The parameter λ of the CRD and ADLR is defined as 0.01. In addition, the HySime algorithm [57], a broadly used method in unmixing, is selected to estimate the endmember number. The endmember number in the ADLR algorithm is set to be 1.5–2 times larger than the number estimated by the HySime method according to [40]. The initial endmember and abundance are extracted and estimated by the VCA-FCLS algorithm [58,59].

The detection maps of the different algorithms on different data sets are provided in Figure 5. It can be easily observed that the anomaly detected by the DVAD algorithm is more obvious than the comparisons. For the San Diego data set, in view of the similarity of the anomaly target and roof, some roof pixels are detected to be anomaly pixels. However, the overall detection result turned out to be better than the other existing algorithms. The proposed DVAD algorithm benefited from the multiple level analysis for anomaly and achieved good detection efficiency based on these data sets in general, which demonstrates its effectiveness for anomaly detection.

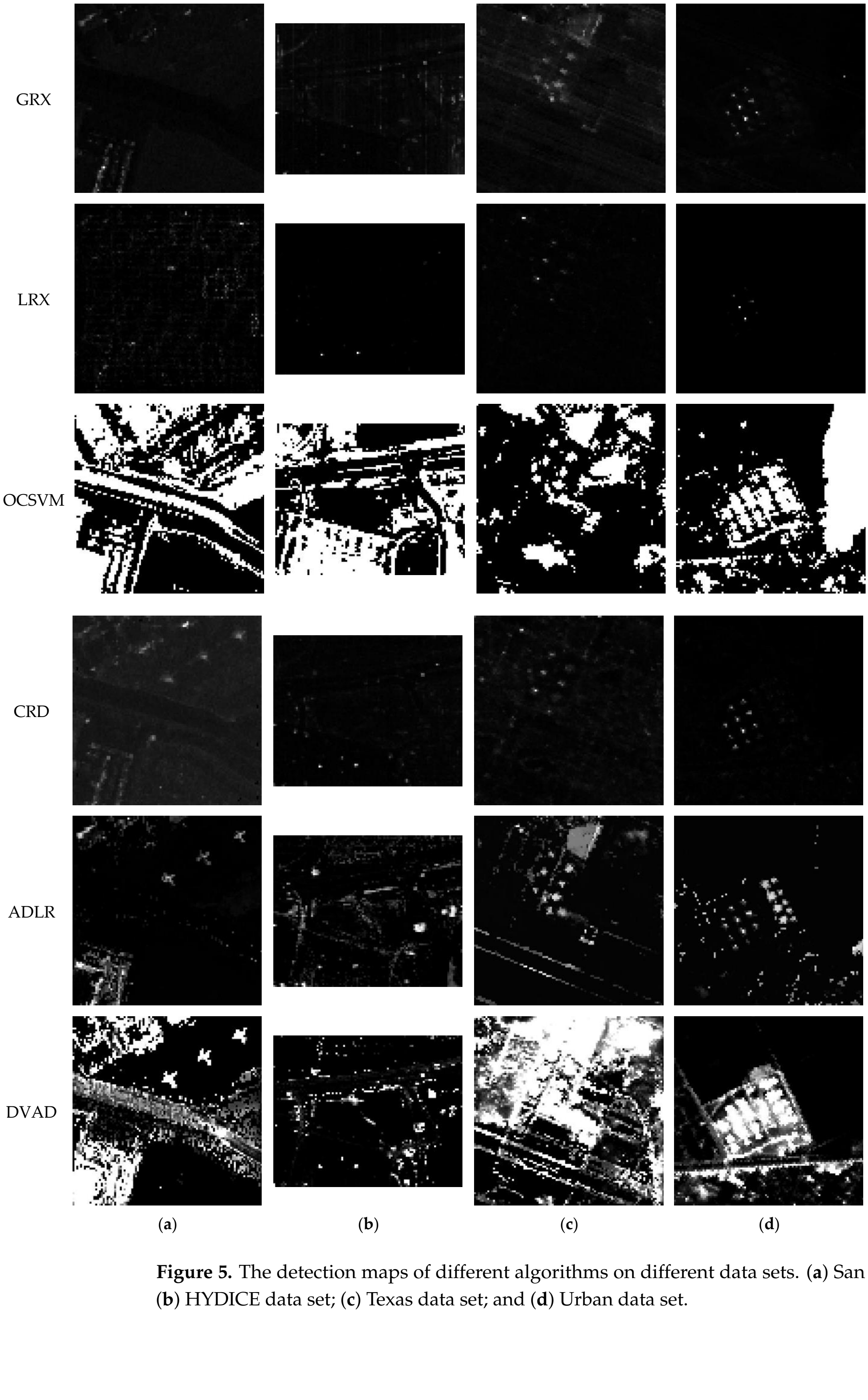
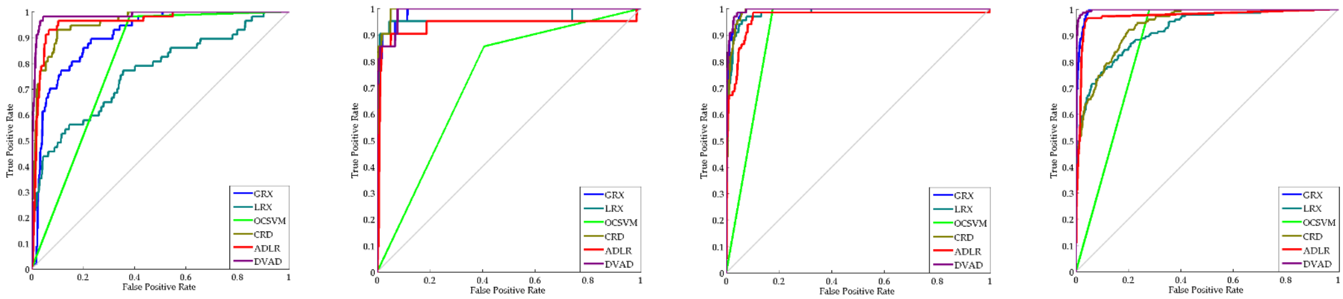


Figure 5. The detection maps of different algorithms on different data sets. (**a**) San Diego data set; (**b**) HYDICE data set; (**c**) Texas data set; and (**d**) Urban data set.

In addition, as for the quantitative analysis and comparison, the ROC curves of the different algorithms are exhibited in Figure 6. From Figure 6, we can derive that the DVAD algorithm almost produces the best results of anomaly detection. Thanks to its global statistical characteristic, the GRX detector can obtain good detection results on most data sets, except for the San Diego data set. The LRX detector and the CRD detector both adopt the dual window to obtain the local information for anomaly detection, whose detection efficiency is influenced by the window size. Their results might be low when the distribution of anomalies in the data set is dense, or when the background is cluttered. Additionally, the performance of the ADLR algorithm is not stable and it may be affected by the results of the endmember and clustering. In brief, the DVAD method does not only fully consider the local and global information via spatial consistency and spectral unmixing, but also analyzes the anomaly from the pixel and subpixel levels. Therefore, the proposed DVAD algorithm achieves the satisfactory results both quantitatively and qualitatively.



(**a**) (**b**) (**c**) (**d**)

Figure 6. The ROC curves obtained by different detection algorithms on different data sets. (**a**) San Diego data set; (**b**) HYDICE data set; (**c**) Texas data set; and (**d**) Urban data set.

Additionally, the AUC values and the time of the different detectors on the different data sets are listed in Table 1, in which the best result of each data set is shown in bold. The AUC values obtained by the proposed DVAD algorithm are almost higher than the comparisons on the different data sets, except for the HYDICE data set. Since the anomaly targets in the HYDICE data set are some independently scattered pixels that are similar to the noise in the image, the final detection result might be affected. In addition, the GRX algorithm has the shortest running time, followed by the DAVD algorithm and the LRX algorithm, with the OCSVM and ADLR algorithms taking the longest time. Therefore, from the perspective of real-time performance analysis, the proposed algorithm achieved relatively good detection results with relatively less time consumption. In general, the DVAD algorithm successfully achieves the satisfactory result for anomaly detection owing to the consideration of the spatial consistency and spectral unmixing.

Table 1. AUC and time for the detectors on different data sets with the best result in bold.

Data Set		GRX	LRX	OCSVM	CRD	ADLR	DVAD
San Diego	AUC	0.9055	0.7624	0.8888	0.9587	0.9577	**0.9847**
	Time(s)	1.24	33.59	1691.02	45.92	1667.23	47.15
HYDICE	AUC	0.9857	0.9605	0.6738	**0.9935**	0.9335	0.9880
	Time(s)	0.07	22.37	168.70	24.56	5897.29	27.50
Texas	AUC	0.9910	0.9827	0.8743	0.9890	0.9641	**0.9950**
	Time(s)	0.09	38.67	365.16	48.22	3858.96	21.62
Urban	AUC	0.9934	0.9224	0.7861	0.9387	0.9711	**0.9980**
	Time(s)	0.12	41.23	993.76	45.08	35.80	13.17

In addition, the detection maps of each level on the four data sets are expressed in Figure 7 to further exhibit the detailed result of the DVAD algorithm. From the result of the pixel level, it is not difficult to find out that there is some structure information in the detection maps on the different data sets. Given that the detection method of the pixel level is a window-based algorithm, it would obtain some local information of the image. However, the detection maps of the subpixel level can reflect the global information of the image via the manifold constraint in the unmixing model. For the San Diego data set, the detection results of some regions, which are mixed by similar materials, such as the airplane region and roof region, are more likely to be the same. Therefore, it is necessary to detect the anomaly from different views to fully integrate the advantages. Then, the fusion strategy is adopted to make the final result benefit from different level detection results. The final result would be an anomaly only when both of the detection results on two levels is an anomaly. This fusion strategy makes full use of the information of the different levels.

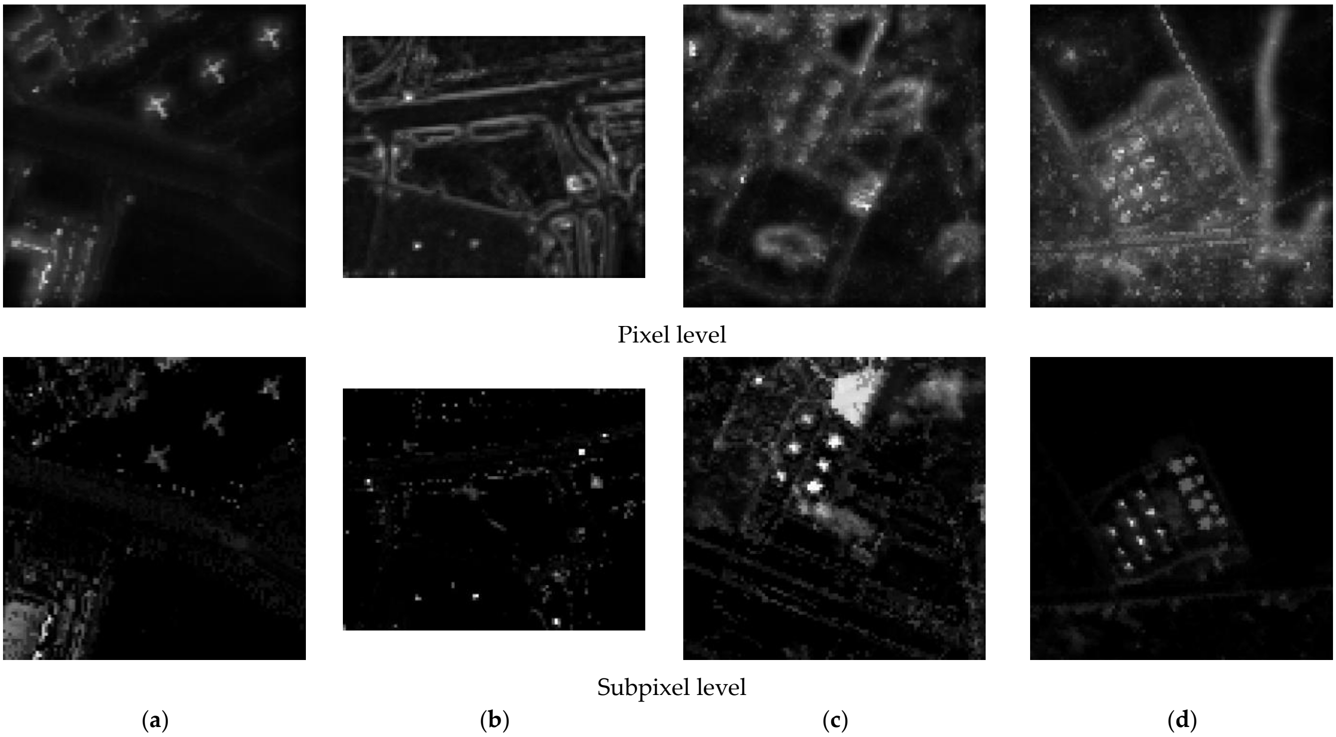


Figure 7. The detection maps of each level obtained by the proposed DVAD method on four data sets. (**a**) San Diego data set; (**b**) HYDICE data set; (**c**) Texas data set; and (**d**) Urban data set.

The endmembers extracted by the proposed spectral unmixing method on the four data sets are displayed in Figure 8, including the background endmembers and anomaly endmembers. A large difference can be found between the background endmembers and the anomaly endmembers. Due to the unknown number of endmembers, the HySime algorithm [57] is used to estimate the endmember number. As we know, the endmember number estimated by the HySime algorithm is usually higher than the real number. Nevertheless, overestimation may not have a bad effect. For example, in the experiment part of paper [40], the number of endmembers is set to be 1.5–2 times higher than the estimated number by HySime algorithm. The estimated endmember numbers of the four different data sets are, respectively, 15, 17, 15, and 4. In the experiment conducted in this paper, a small trick is adopted to further weaken the influence of overestimation. Firstly, the number of endmembers is estimated by the HySime algorithm, and the corresponding abundance is obtained after unmixing. Then, for each endmember, the number of abundance values less than 0.01 is counted. When the number almost reaches the total number of pixels, which is greater than 98% of the total number of pixels [43], this endmember is considered to be the redundant endmember.

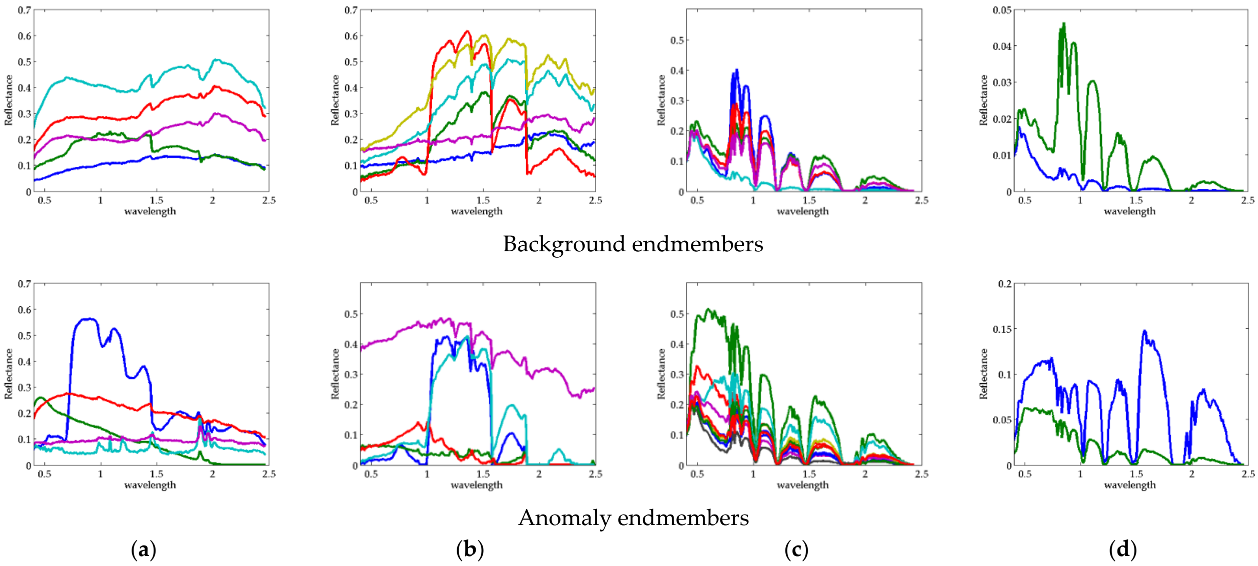


Figure 8. The endmembers with different colors extracted by the proposed spectral unmixing model on four data sets, including the background endmembers and anomaly endmembers. (**a**) San Diego data set; (**b**) HYDICE data set; (**c**) Texas data set; and (**d**) Urban data set.

Additionally, the abundance maps and anomaly abundance maps obtained from the subpixel level on the different data sets are displayed in Figure 9, which also separately marks the background endmembers and anomaly endmembers on the *y*-axis. From the abundance map of Figure 9, it can be clearly noticed that the pixels are mainly mixed by the background endmembers, and the distribution of anomaly endmembers in the whole image is very sparse. Furthermore, in the anomaly abundance map of Figure 9, the anomaly endmember mostly participates in the mixing of anomaly pixels, which also contain a small number of background endmembers.

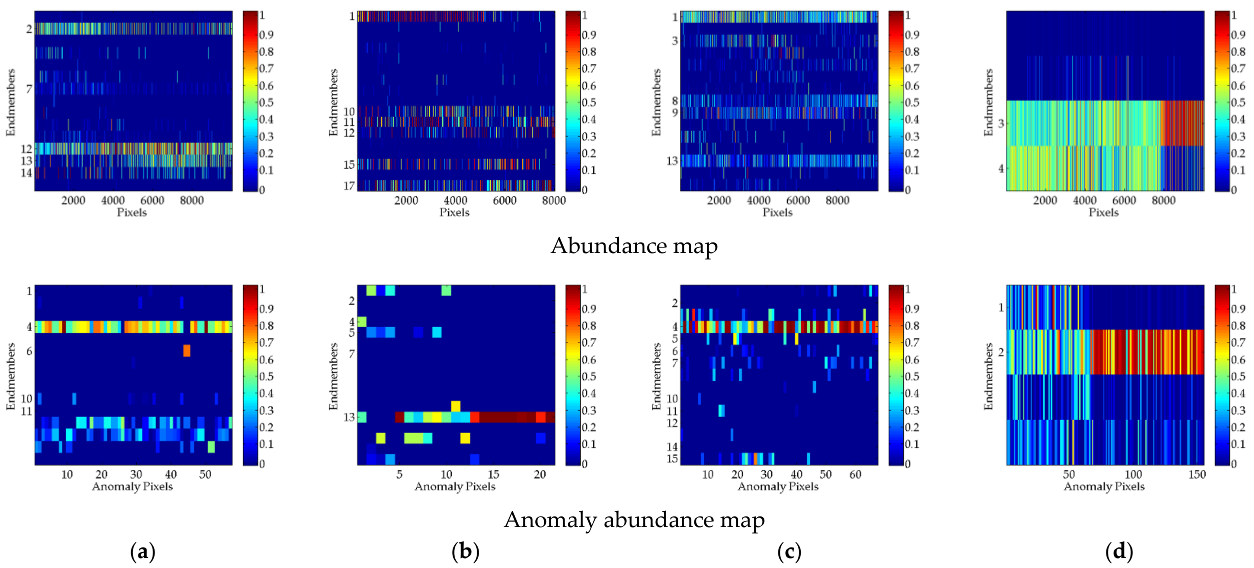


Figure 9. The abundance map and anomaly abundance map obtained by the proposed spectral unmixing model on four data sets. (**a**) San Diego data set; (**b**) HYDICE data set; (**c**) Texas data set; and (**d**) Urban data set.

4.4. Parameter Analysis

There are two level detection methods in the DVAD algorithm, and we first analyze the parameters in the pixel level. The detection algorithm based on the pixel level has two parameters to analyze, including the window size of Win and $Wout$. The AUC performance of the DVAD algorithm, with various values of Win and $Wout$ for the different data sets, is recorded in Table 2. On one hand, by comparing and analyzing the data in Table 2, we can find that the suitable window size on the different data sets is different. For the data set with a small anomaly target, such as the HYDICE data set, the Urban data set, and the Texas data set, the size of the window is small. In this case, the window size is related to the size of the anomaly target. After determining the parameters on the pixel level, its detection result is shown in Figure 7. It can be noticed that there is some local structural information in the detection map due to the sliding window.

Table 2. The AUC performance of the proposed DVAD method with varying window size.

Data Set	Wout \ Win	1	3	5
San Diego	7	0.9923	0.9920	0.9910
	9	0.9930	0.9932	0.9928
	11	0.9932	0.9933	**0.9934**
HYDICE	3	**0.9050**	-	-
	7	0.8407	0.8344	0.7893
	9	0.8184	0.8188	0.7913
Texas	7	0.9651	**0.9690**	0.9675
	9	0.9617	0.9660	0.9674
	11	0.9551	0.9587	0.9633
Urban	7	0.9044	0.9030	0.8968
	9	0.9079	**0.9121**	0.9068
	11	0.9059	0.9108	0.9127

Moreover, the parameters of the subpixel level also need to be discussed. There are two regularization parameters, α and β, in the spectral unmixing model. The parameter α is closely related to the sparsity of abundance, and it could be estimated by calculating the sparsity of the hyperspectral data recorded in [44,60,61], which is defined as follows:

$$\alpha = \frac{1}{\sqrt{L}} \sum_{l=1}^{L} \frac{\sqrt{P} - \|\boldsymbol{y}_l\|_1 / \|\boldsymbol{y}_l\|_2}{\sqrt{P} - 1} \tag{16}$$

where $\boldsymbol{y}_l$ is l-th band in the HSI. The analysis maps for parameter β on the different data sets are shown in the first line of Figure 10. The curves of parameter β in Figure 10 eventually keep stable, and the corresponding AUC values are ultimately higher than 0.98. The curve of parameter β is stable, and the difference between the maximum and minimum values is not significant. The values of β for four data sets, respectively, are 0.045, 0.1, 0.01, and 0.001.

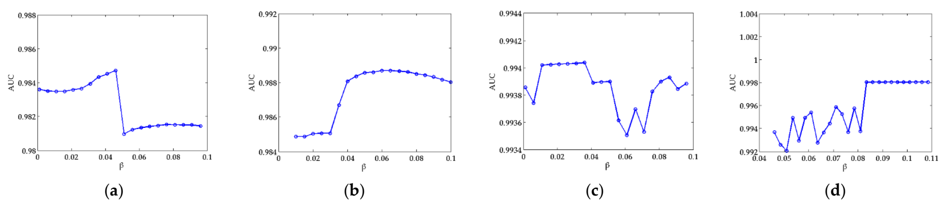


Figure 10. The analysis of parameter β on four data sets. (**a**) San Diego data set; (**b**) HYDICE data set; (**c**) Texas data set; and (**d**) Urban data set.

5. Conclusions

In this paper, we present a dual-view anomaly detection algorithm via spatial consistency and spectral unmixing aimed at detecting anomalies from different levels. In terms of the pixel level view, as many papers assumed, the anomaly mainly refers to the pixel that is dissimilar to its neighbors. From the subpixel level, the difference between the background and anomaly mainly reflects on the endmember difference. The anomaly endmember only participates in the mixing of the anomaly pixels. Therefore, with the anomaly analysis above in mind, we construct the detection model separately from different views and obtain the satisfactory detection result via fusion, which fully considers the spatial consistency and spectral unmixing by utilizing the local and global information in the HSI. The proposed algorithm not only performs the anomaly detection from different levels, but also applies the task of unmixing to anomaly detection. It outperforms the other methods and is verified its effectiveness, as seen in the results of multiple data sets.

Further research is needed to improve our algorithm, for instance, how to eliminate redundant endmembers, how to obtain the true number of endmembers, etc. In our future work, we will take into consideration the correlation between the tasks of anomaly detection and unmixing from other aspects to detect the anomaly more efficiently.

Author Contributions: Conceptualization, J.Z.; Methodology, J.Z. and X.Z.; Formal analysis, J.Z.; Writing—original draft, J.Z.; Writing—review & editing, X.Z. and L.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China, grant number 62276197.

Data Availability Statement: The data presented in this study are openly available in reference number [3,10].

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Stein, A.D.W.J.; Beaven, S.G.; Hoff, L.E.; Winter, E.M.; Schaum, A.P.; Stocker, A.D. Anomaly detection from hyperspectral imagery. *IEEE Signal Process. Mag.* **2002**, *19*, 58–69. [CrossRef]
2. Zhang, X.; Zhang, J.; Li, C.; Cheng, C.; Jiao, L.; Zhou, H. Hybrid unmixing based on adaptive region segmentation for hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 3861–3875. [CrossRef]
3. Kang, X.; Zhang, X.; Li, S.; Li, K.; Li, J.; Benediktsson, J.A. Hyperspectral anomaly detection with attribute edge-preserving filters. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 5600–5611. [CrossRef]
4. Zhang, X.; Song, Q.; Liu, R.; Wang, W.; Jiao, L. Modified co-training with spectral spatial views for semi-supervised hyperspectral image classification. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2014**, *7*, 2044–2055. [CrossRef]
5. Zhang, X.; Gao, Z.; Jiao, L.; Zhou, H. Multifeature hyperspectral image classification with local nonlocal spatial information via Markovrom field in semantic space. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 1409–1424. [CrossRef]
6. Zhang, J.; Zhang, X.; Jiao, L. Sparse nonnegative matrix factorization for hyperspectral unmixing based on endmember independence spatial weighted abundance. *Remote Sensing.* **2021**, *13*, 2348. [CrossRef]
7. Zhang, X.; Sun, Y.; Zhang, J.; Wu, P.; Jiao, L. Hyperspectral unmixing via deep convolutional neural networks. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 1755–1759. [CrossRef]
8. Li, S.; Song, W.; Fang, L.; Chen, Y.; Ghamisi, P.; Benediktsson, J.A. Deep learning for hyperspectral image classification: An overview. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 6690–6709. [CrossRef]
9. Roy, S.K.; Deria, A.; Shah, C.; Haut, J.M.; Du, Q.; Plaza, A. Spectral-spatial morphological attention transformer for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5503615. [CrossRef]
10. Xu, Y.; Wu, Z.; Li, J.; Plaza, A.; Wei, Z. Anomaly detection in hyperspectral images based on low-rank sparse representation. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 1990–2000. [CrossRef]
11. Huyan, N.; Zhang, X.; Zhou, H.; Jiao, L. Hyperspectral anomaly detection via background potential anomaly dictionaries construction. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 2263–2276. [CrossRef]
12. Shang, X.; Song, M.; Wang, Y.; Yu, C.; Yu, H.; Li, F.; Chang, C.-I. Target-constrained interference-minimized selection for hyperspectral target detection. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 6044–6064. [CrossRef]
13. Wang, Y.; Chen, X.; Wang, F.; Song, M.; Yu, C. Meta-learning based hyperspectral target detection using siamese network. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5527913. [CrossRef]
14. Zhang, J.; Zhang, X.; Tang, X.; Chen, P.; Jiao, L. Sketch-based region adaptive sparse unmixing applied to hyperspectral image. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 8840–8856. [CrossRef]

15. Cheng, T.; Wang, B. Total variation sparsity regularized decomposition model with union dictionary for hyperspectral anomaly detection. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 1472–1486. [CrossRef]
16. Lv, C.; Shen, F.; Zhang, Z.; Zhang, F. Review of image anomaly detection. *Acta Autom. Sin.* **2022**, *48*, 1402–1428.
17. Reed, I.S.; Yu, X. Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. *IEEE Trans. Acoust. Speech Signal Process.* **1990**, *38*, 1760–1770. [CrossRef]
18. Molero, J.M.; Garzón, E.M.; García, I.; Plaza, A. Analysis and optimizations of global and local versions of the RX algorithm for anomaly detection in hyperspectral data. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2013**, *6*, 801–814. [CrossRef]
19. Guo, Q.; Zhang, B.; Ran, Q.; Gao, L.; Li, J.; Plaza, A. Weighted-RXD and linear filter-based RXD: Improving background statistics estimation for anomaly detection in hyperspectral imagery. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2014**, *7*, 2351–2366. [CrossRef]
20. Kwon, H.; Nasrabadi, N.M. Kernel RX-algorithm: A nonlinear anomaly detector for hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 388–397. [CrossRef]
21. Zhou, J.; Kwan, C.; Ayhan, B.; Eismann, M.T. A novel cluster kernel RX algorithm for anomaly change detection using hyperspectral images. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6497–6504. [CrossRef]
22. Khazai, S.; Safari, A.; Mojaradi, B.; Homayouni, S. An approach for subpixel anomaly detection in hyperspectral images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 769–778. [CrossRef]
23. Schweizer, S.M.; Moura, J.M.F. Hyperspectral imagery: Clutter adaptation in anomaly detection. *IEEE Trans. Inf. Theory* **2000**, *46*, 1855–1871. [CrossRef]
24. Su, H.; Wu, Z.; Zhang, H.; Du, Q. Hyperspectral anomaly detection: A survey. *IEEE Geosci. Remote Sens. Mag.* **2022**, *10*, 64–90. [CrossRef]
25. Olkopf, B.S.; Williamson, R.; Smola, A.; Shawe-Taylor, J.; Platt, J. Support vector method for novelty detection. *Adv. Neural Inf. Process. Syst.* **2000**, *12*, 582–588.
26. Banerjee, A.; Burlina, P.; Diehl, C. A support vector method for anomaly detection in hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2282–2291. [CrossRef]
27. Carlotto, M.J. A cluster-based approach for detecting man-made objects and changes in imagery. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 374–387. [CrossRef]
28. Hytla, P.C.; Hardie, R.C.; Eismann, M.T.; Meola, J. Anomaly detection in hyperspectral imagery: Comparison of methods using diurnal and seasonal data. *J. Appl. Remote Sens.* **2009**, *3*, 033546. [CrossRef]
29. Li, J.; Zhang, H.; Zhang, L.; Ma, L. Hyperspectral anomaly detection by the use of background joint sparse representation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 2523–2533. [CrossRef]
30. Breunig, M.M.; Kriegel, H.P.; Ng, R.T.; Sander, J. LOF: Identifying density-based local outliers. In Proceedings of the ACM Sigmod International Conference on Management of Data, Dallas, TX, USA, 15–18 May 2000; ACM: New York, NY, USA, 2000.
31. Nowak-Brzezińska, A.; Horyń, C. Outliers in rules—The comparision of LOF, COF and KMEANS algorithms. *Procedia Comput. Sci.* **2020**, *176*, 1420–1429. [CrossRef]
32. Lin, S.; Zhang, M.; Cheng, X.; Zhou, K.; Zhao, S.; Wang, H. Hyperspectral anomaly detection via sparse representation and collaborative representation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2023**, *16*, 946–961. [CrossRef]
33. Li, W.; Du, Q. Collaborative representation for hyperspectral anomaly detection. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 1463–1474. [CrossRef]
34. Su, H.; Wu, Z.; Du, Q.; Du, P. Hyperspectral anomaly detection using collaborative representation with outlier removal. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2018**, *11*, 5029–5038. [CrossRef]
35. Pang, G.; Shen, C.; Cao, L.; Anton, V. Deep learning for anomaly detection: A review. *ACM Comput. Surv.* **2022**, *54*, 1–38. [CrossRef]
36. Li, W.; Wu, G.; Du, Q. Transferred deep learning for anomaly detection in hyperspectral imagery. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 597–601. [CrossRef]
37. Song, S.; Zhou, H.; Yang, Y.; Song, J. Hyperspectral anomaly detection via convolutional neural network and low rank with density-based clustering. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 3637–3649. [CrossRef]
38. Li, K.; Ling, Q.; Wang, Y.; Cai, Y.; Qin, Y.; Lin, Z.; An, W. Spectral difference guided graph attention autoencoder for hyperspectral anomaly detection. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 5001817. [CrossRef]
39. Wang, L.; Wang, X.; Vizziello, A.; Gamba, P. RSAAE: Residual self-attention-based autoencoder for hyperspectral anomaly detection. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5510614. [CrossRef]
40. Qu, Y.; Wang, W.; Guo, R.; Ayhan, B.; Kwan, C.; Vance, S.; Qi, H. Hyperspectral anomaly detection through spectral unmixing and dictionary-based low-rank decomposition. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 4391–4405. [CrossRef]
41. Huang, Z.; Fang, L.; Li, S. Subpixel-pixel-superpixel guided fusion for hyperspectral anomaly detection. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 5998–6007. [CrossRef]
42. Erdinç, A.; Aksoy, S. Anomaly detection with sparse unmixing and Gaussian mixture modeling of hyperspectral images. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 26–31 July 2015; pp. 5035–5038.
43. Zhang, X.; Li, C.; Zhang, J.; Chen, Q.; Feng, J.; Jiao, L.; Zhou, H. Hyperspectral unmixing via low-rank representation with space consistency constraint and spectral library pruning. *Remote Sens.* **2018**, *10*, 339. [CrossRef]

44. Qian, Y.; Jia, S.; Zhou, J.; Robles-Kelly, A. Hyperspectral unmixing via $L_{1/2}$ sparsity-constrained nonnegative matrix factorization. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 4282–4297. [CrossRef]
45. Lee, D.D.; Seung, H.S. Learning the parts of objects by nonnegative matrix factorization. *Nature* **1999**, *401*, 788–791. [CrossRef]
46. Fang, H.; Li, A.; Xu, H.; Wang, T. Sparsity-constrained deep nonnegative matrix factorization for hyperspectral unmixing. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 1105–1109. [CrossRef]
47. Rajabi, R.; Ghassemian, H. Spectral unmixing of hyperspectral imagery using multilayer NMF. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 38–42. [CrossRef]
48. Lu, X.; Wu, H.; Yuan, Y.; Yan, P.; Li, X. Manifold regularized sparse NMF for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 2815–2826. [CrossRef]
49. Cai, D.; He, X.; Han, J. Document clustering using locality preserving indexing. *IEEE Trans. Knowl. Data Eng.* **2005**, *17*, 1624–1637. [CrossRef]
50. Du, B.; Zhao, R.; Zhang, L.; Zhang, L. A spectral-spatial based local summation anomaly detection method for hyperspectral images. *Signal Process.* **2016**, *124*, 115–131. [CrossRef]
51. Zhang, Y.; Du, B.; Zhang, L. Regularization framework for target detection in hyperspectral imagery. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 313–317. [CrossRef]
52. Zhao, R.; Du, B.; Zhang, L. Hyperspectral anomaly detection via a sparsity score estimation framework. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 3208–3222. [CrossRef]
53. Ma, L.; Crawford, M.M.; Tian, J. Local manifold learning based k-nearest-neighbor for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 4099–4109. [CrossRef]
54. Liu, X.; Tanaka, M.; Okutomi, M. Single-image noise level estimation for blind denoising. *IEEE Trans. Image Process.* **2013**, *22*, 5226–5237. [CrossRef]
55. Xing, Y.; Gomez, R.B. Hyperspectral image analysis using ENVI (environment for visualizing images). *Proc. SPIE* **2001**, *4383*, 79–87.
56. Kerekes, J. Receiver operating characteristic curve confidence intervals and regions. *IEEE Geosci. Remote Sens. Lett.* **2008**, *5*, 251–255. [CrossRef]
57. Bioucas-Dias, J.M.; Nascimento, J.M.P. Hyperspectral subspace identification. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 2435–2445. [CrossRef]
58. Nascimento, J.M.P.; Dias, J.M.B. Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 898–910. [CrossRef]
59. Heinz, D.C.; Chang, C.-I. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 529–545. [CrossRef]
60. He, W.; Zhang, H.; Zhang, L. Total variation regularized reweighted sparse nonnegative matrix factorization for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 3909–3921. [CrossRef]
61. Tong, L.; Zhou, J.; Li, X.; Qian, Y.; Gao, Y. Region-based structure preserving nonnegative matrix factorization for hyperspectral unmixing. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2017**, *10*, 1575–1588. [CrossRef]