Article

Multi-Dimensional Low-Rank with Weighted Schatten p-Norm Minimization for Hyperspectral Anomaly Detection

Xi'ai Chen [1,2,*], Zhen Wang [1,2], Kaidong Wang [3], Huidi Jia [1,2,4], Zhi Han [1,2] and Yandong Tang [1,2]

1   State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China; wangzhen2@sia.cn (Z.W.); jiahuidi@sia.cn (H.J.); hanzhi@sia.cn (Z.H.); ytang@sia.cn (Y.T.)
2   Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110016, China
3   Center for Intelligent Decision-Making and Machine Learning, School of Management, Xi'an Jiaotong University, Xi'an 710049, China; wangkd13@gmail.com
4   University of Chinese Academy of Sciences, Beijing 100049, China
*   Correspondence: chenxiai@sia.cn

Abstract: Hyperspectral anomaly detection is an important unsupervised binary classification problem that aims to effectively distinguish between background and anomalies in hyperspectral images (HSIs). In recent years, methods based on low-rank tensor representations have been proposed to decompose HSIs into low-rank background and sparse anomaly tensors. However, current methods neglect the low-rank information in the spatial dimension and rely heavily on the background information contained in the dictionary. Furthermore, these algorithms show limited robustness when the dictionary information is missing or corrupted by high level noise. To address these problems, we propose a novel method called multi-dimensional low-rank (MDLR) for HSI anomaly detection. It first reconstructs three background tensors separately from three directional slices of the background tensor. Then, weighted schatten p-norm minimization is employed to enforce the low-rank constraint on the background tensor, and $L_{F,1}$-norm regularization is used to describe the sparsity in the anomaly tensor. Finally, a well-designed alternating direction method of multipliers (ADMM) is employed to effectively solve the optimization problem. Extensive experiments on four real-world datasets show that our approach outperforms existing anomaly detection methods in terms of accuracy.

Keywords: anomaly detection; multi-dimensional; low-rank

1. Introduction

Compared with conventional images such as RGB images, multispectral images, SAS images [1], and delay-Doppler images [2], hyperspectral images (HSIs) offer the advantage of capturing hundreds of contiguous spectral bands of the same scene. This unique characteristic of HSI proves to be beneficial for target detection and finds wide applications in various fields such as land cover classification [3–5], mineral survey [6–8], environmental protection [9–11], and other applications [12–18]. In hyperspectral target detection, when the target information is unknown, the unsupervised processing of the target detection is called anomaly detection. However, in practical applications, it is often difficult to obtain the prior information of the target, so hyperspectral anomaly detection is more suitable. In essence, hyperspectral anomaly detection can be viewed as an unsupervised binary classification problem that separates an image into background and anomalies, where anomalies typically represent rare targets that occupy only a small number of pixels.

Over the past two decades, there has been a growing interest in hyperspectral anomaly detection, leading to the development of numerous detection algorithms. The Reed–Xiaoli (RX) algorithm [19] is a classical statistical modelling method for anomaly detection, assuming that the background follows a multivariate Gaussian distribution. The main objective of the RX algorithm is to compute the Mahalanobis distance between the measured pixel and the background [20], which involves estimating the mean vector and the covariance

matrix of the background. Two commonly studied extended versions of the RX algorithm are the global RX (GRX) [21] and the local RX (LRX) [22], where the former calculates the distance between the measured pixel and all background pixels, and the latter calculates the distance between the measured pixel and the surrounding background pixels. However, in hyperspectral applications, it is crude to describe the background with a single Gaussian distribution, and the mean vector and covariance matrix of the background are susceptible to the noisy pixels and anomalies.

In general, the HSI can be represented as a three-order tensor with two spatial dimensions and one spectral dimension. Taking into account the similarity between spectral bands, the HSI can be transformed into a matrix along the spectral dimension, which inspires the matrix-based anomaly detection methods. Anomalies are assumed to be randomly distributed in the background and to have sparse properties. By formulating a constrained convex optimization problem that incorporates the characteristics of both the background and the anomalies, successful separation of the anomalies from the background can be achieved. Consequently, the low-rank and sparse matrix decomposition (LRaSMD) algorithms [23–25] have been used to separate the HSI data into low-rank background and sparse anomalies and have demonstrated their effectiveness in previous studies [26–28]. According to the LRaSMD approach, the spectral response of a pixel $y_i (i \in \{1, \ldots, N\})$ in d bands of the HSI can be represented as a spectral vector $\mathbf{y}_i \in \mathbb{R}^d$ with the decomposition.

$$\mathbf{y}_i = \mathbf{x}_i + \mathbf{s}_i, \quad \begin{cases} \mathbf{s}_i = 0, & \text{if } \mathbf{y}_i \text{ is part of background,} \\ \mathbf{s}_i \neq 0, & \text{if } \mathbf{y}_i \text{ is part of anomalies,} \end{cases} \tag{1}$$

which can further be written in matrix form as:

$$\mathbf{Y} = \mathbf{X} + \mathbf{S}, \tag{2}$$

where $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N]^T$ and $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_N]^T \in \mathbb{R}^{N \times d}$ represent the background and anomaly components of the HSI matrix $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N]^T \in \mathbb{R}^{N \times d}$, where N represents the number of pixels in the HSI, and d represents the number of spectral bands. Furthermore, to address the attention imbalance between anomalies and the background observed in LRaSMD, Zhang et al. [29] proposed the LRaSMD-based Mahalanobis distance (LSMAD) method. Xu et al. [30] integrated cooperative representation and Euclidean distance into the LRaSMD framework. Li et al. [31] investigated LRaSMD under the assumption of a mixture-of-Gaussian (MoG) distribution and developed a global detector based on the Manhattan distance. To further exploit the intrinsic information of the background, low-rank representation (LRR) [32,33] was proposed, which maps the HSI to multiple linear subspaces using a dictionary. Xu et al. [34] proposed a new anomaly detection method called low-rank and sparse representation (LRASR), which employed a dictionary construction strategy and a sparsity-inducing regularization term to reconstruct the background matrix. To preserve the local geometric structure and spatial relationships of the background, the graph and total variation regularized low-rank representation (GTVLRR) [35] method was introduced for HSI anomaly detection. Fu et al. [36] used convolutional neural network (CNN) denoisers [37] as priors for the coefficients of the dictionary.

The aforementioned matrix-based anomaly detection methods tend to destroy the spatial structure of HSI and fail to effectively exploit the inherent spatial information [38,39]. In recent years, tensor-based methods have emerged as a promising approach to HSI anomaly detection, allowing the decomposition of HSI data into low-rank and sparse components. Sun et al. [23] used Tucker decomposition to obtain the low-rank background, using an unmixing method to extract the spectral features of the anomaly. Li et al. [40] embedded priors into the dimensions of a tensor with different regularizations. Song et al. [41] proposed a dictionary construction strategy based on Tucker decomposition, which improved the inclusion of spectral segment information in the dictionary. Shang et al. [42] found a new prior that describes the sparsity of the core tensor of a

gradient map (GCS) under Tucker decomposition. However, Tucker decomposition has inherent limitations in terms of rank. To address this issue, Wang et al. [43] extended the concept of LRR from matrix to tensor, taking into account the three-dimensional structure of HSI. Sun et al. [44] represented the background tensor as the product of a transformed tensor and a low-rank matrix. However, these methods pay primary attention to the low-rank of the spectral dimension of the tensor, neglecting the low-rank information in the spatial dimensions. The dictionary, which maps the HSI into multiple linear subspaces, plays a crucial role in the reconstruction of the background component. To achieve an effective separation of background and anomalies, the dictionary should primarily contain background information. Although some methods choose the original data themselves as the dictionary, they may still contain anomalies that can adversely affect the background reconstruction process.

To address these issues, we propose a multi-dimensional low-rank (MDLR) strategy for HSI anomaly detection. Unlike the existing tensor-based methods that construct one background tensor, our approach constructs three background tensors, two capturing the spatial dimension and one representing the spectral dimension. Using the tensor singular value decomposition (t-SVD) technique, we obtain the f-diagonal tensor $\mathcal{S}$, characterizing the background. To enforce low-rankness in the background tensor, we apply the weighted Schatten p-norm minimization (WSNM) to the slices of $\mathcal{S}$. Finally, the three background tensors are merged into a single background tensor. In addition, anomalies in the HSI tend to occur at consistent spatial locations across all spectral bands and exhibit a slight spectral density. To capture this property, we impose a joint spectral–spatial sparsity on the anomaly tensor using the $L_{F,1}$ norm. The main contributions of this work can be summarized as follows:

1. Low-rankness along three dimensions in the frequency domain is exploited. Through the low-rank property analysis of the tensor along different dimensions, we found that it is not sufficient to measure the low-rankness along only one dimension. Therefore, multi-dimensional low-rankness is embedded into different tensors with t-SVD along different slices. These tensors are then fused to form a background tensor that captures the low-rank characteristics across all three dimensions and enables the MDLR method to effectively explore more comprehensive background information.
2. To enforce low-rank in the background tensor, WSNM is applied to the frontal slices of the f-diagonal tensor, which enhances the preservation of the low-rank structure in the background tensor.

The rest of this paper is organized as follows. In Section 2, notations and preliminaries are introduced. The proposed multidimensional low-rank model is presented in detail in Section 3. The experimental results are demonstrated in Section 4. The conclusion is given in Section 5.

2. Notations and Preliminaries

In this section, we introduce the notations and preliminaries used in this paper. The column vectors are represented by lowercase letters, e.g., $\mathbf{x}$. The matrix is represented by bold capital letters, e.g., $\mathbf{X}$. An HSI with w rows, h columns, and d spectral bands can be naturally represented as a third-order tensor, $\mathcal{X} \in \mathbb{R}^{w \times h \times d}$. The discrete Fourier transform (DFT) of $\mathcal{X}$ along the spectral dimension can be written as $\hat{\mathcal{X}} = \text{fft}(\mathcal{X}, [], 3)$. The inverse DFT of $\hat{\mathcal{X}}$ is written as $\mathcal{X} = \text{ifft}(\hat{\mathcal{X}}, [], 3)$. $\mathcal{X}^*$ represents the conjugate transpose of $\mathcal{X}$. $\mathcal{X}^{(i)}$ is the i-th frontal slice of $\mathcal{X}$. The block circulant matrix $\text{bcirc}(\mathcal{N})$ of $\mathcal{N} \in \mathbb{R}^{w \times h \times d}$ is defined as follows:

$$\text{bcirc}(\mathcal{N}) = \begin{bmatrix} \mathcal{N}^{(1)} & \mathcal{N}^{(d)} & \mathcal{N}^{(d-1)} & \cdots & \mathcal{N}^{(2)} \\ \mathcal{N}^{(2)} & \mathcal{N}^{(1)} & \mathcal{N}^{(d)} & \cdots & \mathcal{N}^{(3)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathcal{N}^{(d)} & \mathcal{N}^{(d-1)} & \ddots & \mathcal{N}^{(2)} & \mathcal{N}^{(1)} \end{bmatrix}.$$

The block vectorization operation bvec$(\cdot)$ of $\mathcal{N}$ and its inverse operation bvfold$(\cdot)$ are denoted as:

$$\mathrm{bvec}(\mathcal{N}) = \begin{bmatrix} \mathcal{N}^{(1)} \\ \mathcal{N}^{(2)} \\ \vdots \\ \mathcal{N}^{(d)} \end{bmatrix}, \quad \mathrm{bvfold}(\mathrm{bvec}(\mathcal{N})) = \mathcal{N}.$$

Definition 1 (Tensor product). *The product of three-order tensor $\mathcal{N} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $\mathcal{M} \in \mathbb{R}^{n_2 \times n_4 \times n_3}$ is $\mathcal{A} \in \mathbb{R}^{n_1 \times n_4 \times n_3}$ defined as follows:*

$$\mathcal{A} = \mathcal{N} * \mathcal{M} = bvfold(bcirc(\mathcal{N}) * bvec(\mathcal{M})). \tag{3}$$

Definition 2 (Slices of Tensor). *There are three types of slices in a tensor: that is, horizontal slices $\mathcal{X}_{i::}$, lateral slices $\mathcal{X}_{:j:}$, and frontal slices $\mathcal{X}_{::k}$.*

Definition 3 (Identity Tensor). *The identity tensor $\mathcal{I} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is defined by $\mathcal{I}(:,:,1) = eye(n_1, n_2)$, $\mathcal{I}(:,:,2:n_3) = 0$, where eye (n_1, n_2) is an identity matrix $(n_1 \times n_2)$.*

Definition 4 (Conjugate Transpose). *The conjugate transpose of a tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is denoted as $\mathcal{X}^*$ with*

$$\widehat{\mathcal{X}^*}^{(i)} = (\hat{\mathcal{X}}^{(i)})^T, i = 1, 2, ..., n_3. \tag{4}$$

Definition 5 (Orthogonal Tensor). *The orthogonal tensor $\mathcal{D}$ satisfies $\mathcal{D}^* * \mathcal{D} = \mathcal{D} * \mathcal{D}^* = \mathcal{I}$.*

Definition 6 (t-SVD). *The singular value decomposition of a tensor $\mathcal{X} \in \mathbb{R}^{w \times h \times d}$ can be decomposed into the product of three three-order tensors.*

$$\mathcal{X} = \mathcal{U} * \mathcal{S} * \mathcal{V}^*, \tag{5}$$

where $\mathcal{U} \in \mathbb{R}^{w \times w \times d}$ and $\mathcal{V} \in \mathbb{R}^{h \times h \times d}$ are orthogonal tensors and $\mathcal{S} \in \mathbb{R}^{w \times h \times d}$ is an f-diagonal tensor. The procedure of t-SVD is described in Algorithm 1.

Definition 7 (Tensor Tubal Rank). *For a tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ with t-SVD $\mathcal{X} = \mathcal{U} * \mathcal{S} * \mathcal{V}^*$, its tubal rank is the number of non-zero tubes of $\mathcal{S}$:*

$$rank_t(\mathcal{X}) = \#\{k : \mathcal{S}(k, k, :) \neq 0\}. \tag{6}$$

Definition 8 (Tensor Nuclear Norm(TNN)). *The TNN of a tensor $X \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is the sum of singular values of all front slices of $\hat{\mathcal{X}}$, that is,*

$$\|\mathcal{X}\|_* := \sum_{k=1}^{n_3} \|\hat{\mathcal{X}}^{(k)}\|_*. \tag{7}$$

3. Proposed Method

An illustration of the proposed model is shown in Figure 1. Figure 1a illustrates the different dimension low-rank property of the HSI in the frequency domain. To exploit the low-rankness along different dimensions, we combine these three different dimensional tensors to form the background tensor and apply tensor low-rank and sparse decomposition to extract the sparse anomaly object from the low-rank background. Final detection map M can be obtained via the sparse $\mathcal{S}$ by computing $\sqrt{\sum_{k=1}^{d} |\mathcal{S}(i,j,k)|^2}$. We will introduce each part in detail in the following subsections.
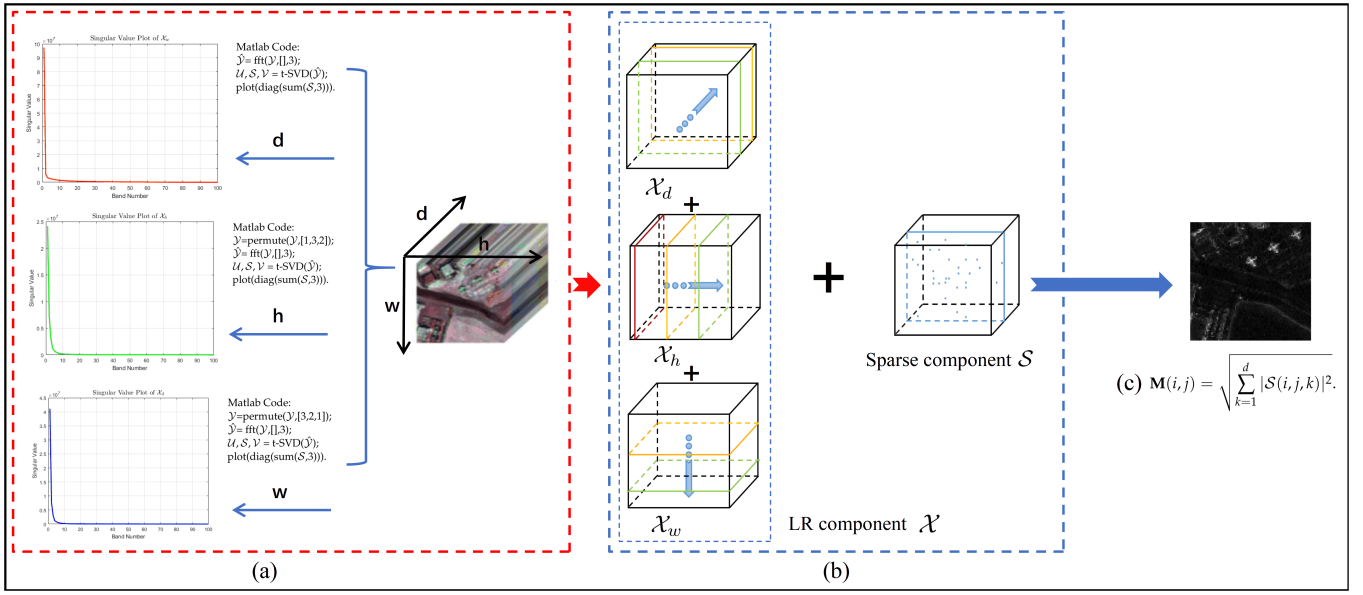
Figure 1. Illustration of the proposed model for HSI anomaly detection. (**a**) Multi-dimensional low-rank in frequency domain. (**b**) Tensor low-rank and sparse decomposition. (**c**) Detection map.

3.1. Tensor Low-Rank Linear Representation

LRR uses a dictionary to explore low-rank linear representations of HSI, but the matrix-based approach breaks the tensor structure inherent in HSI. To overcome this limitation, tensor LRR is proposed, which incorporates the t-product to preserve the spatial structure of the tensor. Given a tensor $\mathcal{Y} \in \mathbb{R}^{w \times h \times d}$, it can be decomposed using the tensor LRR formulation as follows:

$$\mathcal{Y} = \mathcal{A} * \mathcal{X} + \mathcal{S}, \tag{8}$$

where $\mathcal{X}$ is the low-rank background tensor, $\mathcal{S}$ is the sparse anomaly tensor, and $\mathcal{A}$ is the dictionary. Equation (8) aims to construct the low-rank and sparse components exactly and efficiently by dictionary $\mathcal{A}$ from HSI data.

$$\min_{\mathcal{X}, \mathcal{S}} \quad \text{rank}_t(\mathcal{X}) + \lambda \text{sparse}(\mathcal{S})$$

$$s.t. \quad \mathcal{Y} = \mathcal{A} * \mathcal{X} + \mathcal{S}, \tag{9}$$

where $\text{rank}_t(\mathcal{X})$ denotes the tensor tubal rank function [45], λ is a regularization parameter of the $\mathcal{S}$, $\text{sparse}(\mathcal{S})$ is the sparse norm.

3.2. Weighted Schatten p-Norm Minimization

The problem of determining the $\text{rank}_t(\mathcal{X})$ is known to be NP-hard. To approximate the rank of a matrix, a commonly used method is nuclear norm minimization (NNM), which calculates the sum of the singular values of the matrix $\mathcal{X}$. NNM is typically solved using a singular value thresholding algorithm. However, to obtain a more accurate low-rank approximation, other methods [46–48] have been developed. These methods treat different singular values individually rather than uniformly as in NNM, resulting in improved performance. In WSNM, each singular value is assigned a specific weight, and the optimization problem aims to minimize the weighted Schatten p-norm of the matrix $\mathbf{X} \in \mathbb{R}^{h \times w}$.

$$\|\mathbf{X}\|_{w,S_p} = \left(\sum_{i=1}^{\min\{n,m\}} w_i \sigma_i^p \right)^{\frac{1}{p}}, \tag{10}$$

where σ_i is the *i*-th singular value of $\mathbf{X}$, w_i is the weight of σ_i, $\mathbf{w}=[w_1,\ldots,w_{min(n,m)}]$ is a non-negative vector to constrain the single value of $\mathbf{X}$. The weighted Schatten *p* norm minimization problem can be effectively solved by the generalized soft thresholds. Given *p* and w_i, the specific threshold can be obtained by:

$$GST(w_i, p) = (2w_i(1-p))^{\frac{1}{2-p}} + w_i p(2w_i(1-p))^{\frac{p-1}{2-p}}. \tag{11}$$

The main procedures of this approach are shown in Algorithm 1. In this work, the low-rank problem of HSI is solved in tensor form and the nuclear norm of the matrix is converted to the tensor nuclear norm (TNN). The WSNM is applied to the forward slices of $\mathcal{S}$.

3.3. Mutil-Dimensional Tensor Low-Rank Norm

According to the tensor LRR, tensor $\mathcal{X}$ can be expressed as the linear combination of the tensor dictionary $\mathcal{A}$. The choice of the dictionary plays a crucial role in the background tensor reconstruction. Conventional dictionary construction methods are often sensitive to noise and require separate construction for different datasets, making the anomaly detection process complicated. When the dictionary is an identity tensor, the tensor LRR is converted to tensor robust principal component analysis (TRPCA) [49]. By combining WSNM and TRPCA, we have the following:

$$\min_{\mathcal{X},\mathcal{S}} \quad \|\mathcal{X}\|_{w,S_p} + \lambda\|\mathcal{S}\|_{F,1}$$

$$s.t. \quad \mathcal{Y} = \mathcal{X} + \mathcal{S}. \tag{12}$$

In the field of HSI unmixing, a latent low-rank representation theory (LatLRR) has been proposed [50]. LatLRR treats itself as a dictionary and learns its own rows and columns separately to obtain two different background representations while incorporating low-rank constraints. Motivated by this concept, we aim to explore the background tensor and reorganize it from different directions of slices. To achieve this, we introduce three background tensors: $\mathcal{X}_w, \mathcal{X}_h, \mathcal{X}_d \in \mathbb{R}^{w \times h \times d}$. We run WSNM separately on these three tensors along different dimensional frontal slices. The proposed tensor-based method, called multi-dimensional low-rank (MDLR), can be expressed as follows:

$$\|\mathcal{X}\|_{msp,*} = \mu_w\|\mathcal{X}_w\|_{w,S_p} + \mu_h\|\mathcal{X}_h\|_{w,S_p} + \mu_d\|\mathcal{X}_d\|_{w,S_p}, \tag{13}$$

where $0 \leq \mu_w \leq 1$, $0 \leq \mu_h \leq 1$, and $\mu_d = 1 - \mu_w - \mu_h$ balance the contributions of $\mathcal{X}_w, \mathcal{X}_h$ and $\mathcal{X}_d$. We call the reconstruction of the background tensor $\mathcal{X}_w$, the reconstruction of the background from the *w* dimension, $\mathcal{X}_h$, the reconstruction of the background from the *h* dimension, $\mathcal{X}_d$ the reconstruction of the background from the *d* dimension. Finally, our model formulation can be written as:

$$\min_{\mathcal{X},\mathcal{S}} \quad \|\mathcal{X}\|_{msp,*} + \lambda\|\mathcal{S}\|_{F,1}$$

$$s.t. \quad \mathcal{Y} = \mathcal{X} + \mathcal{S}. \tag{14}$$

3.4. Optimization Procedure

By introducing auxiliary $\mathcal{X}_w, \mathcal{X}_h, \mathcal{X}_d$, Equation (12) can be written as the following equivalent problem:

$$\min_{\mathcal{X}_w,\mathcal{X}_h,\mathcal{X}_d,\mathcal{S}} \mu_w\|\mathcal{X}_w\|_{w,S_p} + \mu_h\|\mathcal{X}_h\|_{w,S_p} + \mu_d\|\mathcal{X}_d\|_{w,S_p} + \lambda\|\mathcal{S}\|_{F,1}$$

$$s.t. \; \mathcal{X} = \mathcal{X}_w, \; \mathcal{X} = \mathcal{X}_h, \; \mathcal{X} = \mathcal{X}_d, \; \mathcal{Y} = \mathcal{X} + \mathcal{S}. \tag{15}$$

The Lagrange multipliers $\mathcal{E}$, $\mathcal{Q}_{1,2,3}$ are introduced and we use the ADMM to solve the augmented Lagrange function. The optimization problem above is written as follows:

$$\min_{\mathcal{X}_w, \mathcal{X}_h, \mathcal{X}_d, \mathcal{S}, \mathcal{E}} \mu_w \|\mathcal{X}_w\|_{w, S_p} + \mu_h \|\mathcal{X}_h\|_{w, S_p} + \mu_d \|\mathcal{X}_d\|_{w, S_p} + \lambda \|\mathcal{S}\|_{F, 1}$$

$$+ \frac{\alpha}{2} \|\mathcal{X} - \mathcal{X}_w + \frac{\mathcal{Q}_1}{\alpha}\|_F^2 + \frac{\alpha}{2} \|\mathcal{X} - \mathcal{X}_h + \frac{\mathcal{Q}_2}{\alpha}\|_F^2$$

$$+ \frac{\alpha}{2} \|\mathcal{X} - \mathcal{X}_d + \frac{\mathcal{Q}_3}{\alpha}\|_F^2 + \frac{\alpha}{2} \|\mathcal{Y} - \mathcal{X} - \mathcal{S} + \frac{\mathcal{E}}{\alpha}\|_F^2 \tag{16}$$

(1) **Update** $\mathcal{X}$

$$\mathcal{X} = \underset{\mathcal{X}}{\arg\min} \frac{\alpha}{2} \|\mathcal{X} - \mathcal{X}_w + \frac{\mathcal{Q}_1}{\alpha}\|_F^2 + \frac{\alpha}{2} \|\mathcal{X} - \mathcal{X}_h + \frac{\mathcal{Q}_2}{\alpha}\|_F^2$$

$$+ \frac{\alpha}{2} \|\mathcal{X} - \mathcal{X}_d + \frac{\mathcal{Q}_3}{\alpha}\|_F^2 + \frac{\alpha}{2} \|\mathcal{Y} - \mathcal{X} - \mathcal{S} + \frac{\mathcal{E}}{\alpha}\|_F^2. \tag{17}$$

The closed-form solution of $\mathcal{X}$ can be obtained by taking the derivative of the above objective function and setting it to zero, as follows:

$$\mathcal{X} = (\mathcal{Y} - \mathcal{S} + \frac{\mathcal{E}}{\alpha} + \mathcal{X}_w + \mathcal{X}_h + \mathcal{X}_d - \sum_{i=1}^{3} \frac{\mathcal{Q}_i}{\alpha}) / 4 \tag{18}$$

(2) **Update** $\mathcal{X}_w$

$$\mathcal{X}_w = \underset{\mathcal{X}_w}{\arg\min} \quad \mu_w \|\mathcal{X}_w\|_{w, s_p} + \frac{\alpha}{2} \|\mathcal{X} - \mathcal{X}_w + \frac{\mathcal{Q}_1}{\alpha}\|_F^2 \tag{19}$$

(3) **Update** $\mathcal{X}_h$

$$\mathcal{X}_h = \underset{\mathcal{X}_h}{\arg\min} \quad \mu_h \|\mathcal{X}_h\|_{w, s_p} + \frac{\alpha}{2} \|\mathcal{X} - \mathcal{X}_h + \frac{\mathcal{Q}_3}{\alpha}\|_F^2 \tag{20}$$

(4) **Update** $\mathcal{X}_d$

$$\mathcal{X}_d = \underset{\mathcal{X}_d}{\arg\min} \quad \mu_d \|\mathcal{X}_d\|_{w, s_p} + \frac{\alpha}{2} \|\mathcal{X} - \mathcal{X}_d + \frac{\mathcal{Q}_3}{\alpha}\|_F^2. \tag{21}$$

The subproblem $\mathcal{X}_{w,h,d}$ can be solved using generalized soft-thresholding as shown in Algorithm 1. Before applying the Algorithm 1, $\mathcal{X}_w$ should be converted to $\mathcal{X}_w \in \mathbb{R}^{d \times h \times w}$, and then it must be reshaped again as $\mathcal{X}_w \in \mathbb{R}^{w \times h \times d}$ after Algorithm 1. Similarly, $\mathcal{X}_h$ should be converted to $\mathcal{X}_w \in \mathbb{R}^{w \times d \times h}$ before Algorithm 1 and back to $\mathcal{X}_h \in \mathbb{R}^{w \times h \times d}$ after Algorithm 1.

(5) **Update** $\mathcal{S}$

$$\mathcal{S} = \underset{\mathcal{S}}{\arg\min} \lambda \|\mathcal{S}\|_{F,1} + \frac{\alpha}{2} \|\mathcal{Y} - \mathcal{X} - \mathcal{S} + \frac{\mathcal{E}}{\alpha}\|_F^2. \tag{22}$$

Then, we have the following closed solution:

$$\mathcal{S}(:,:,k) = \begin{cases} \frac{\|\mathcal{M}(:,:,k)\|_F + \lambda}{\mathcal{M}(:,:,k)\|_F} \mathcal{M}(:,:,k), & \lambda < \|\mathcal{M}(:,:,k)\|_F \\ 0, & \text{otherwise} \\ , \end{cases}$$

$$\tag{23}$$

where $\mathcal{M} = \mathcal{Y} - \mathcal{X} + \frac{\mathcal{E}}{\alpha}$.

(6) **Update** Lagrange multiplier $\mathcal{E}$ and $\mathcal{Q}_{1,2,3}$

$$\mathcal{E} = \mathcal{E} + \alpha * (\mathcal{Y} - \mathcal{X} - \mathcal{S})$$

$$
\begin{aligned}
\mathcal{Q}_1 &= \mathcal{Q}_1 + \alpha * (\mathcal{X} - \mathcal{X}_w) \\
\mathcal{Q}_2 &= \mathcal{Q}_2 + \alpha * (\mathcal{X} - \mathcal{X}_h) \\
\mathcal{Q}_3 &= \mathcal{Q}_3 + \alpha * (\mathcal{X} - \mathcal{X}_d).
\end{aligned}
\tag{24}
$$

The overall process of the proposed method is concluded in Algorithm 2. When the optimization process is complete, the anomaly detection map **M** of HSI data can obtain by the sparse anomaly tensor S as follows:

$$
\mathbf{M}(i,j) = \sqrt{\sum_{k=1}^{d} |\mathcal{S}(i,j,k)|^2}.
\tag{25}
$$

Due to the WSNM regularization, the solution process of the problem in Equations (19)–(21) is actually not a convex optimization problem. Nevertheless, Xie et al. [51] prove that WSNM is not convex, and if the weights satisfy $0 \le w_1 \le w_2 \le \ldots w_i$, at least one accumulation point satisfies (26). A convergence analysis can be found in Theorem 3 of WSNM.

$$
\lim_{k \to \infty} \quad \|\mathcal{X}_{k+1} - \mathcal{X}_k\|_F^2 + \|\mathcal{S}_{k+1} - \mathcal{S}_k\|_F^2.
\tag{26}
$$

Algorithm 1 WSNM based on t-SVD.

Input: $\mathcal{X}, \mathcal{Q}, p, \alpha, \tau$
1: $\mathcal{P} = \mathcal{X} + \frac{\mathcal{Q}}{\alpha}$
2: $\hat{\mathcal{P}} = \mathrm{fft}(\mathcal{P}, [], 3)$
3: **for** i = 0,1,..., $[\frac{d+1}{2}]$
4:     $[\hat{\mathcal{U}}(:,:,i), \hat{\mathcal{S}}(:,:,i), \hat{\mathcal{V}}(:,:,i)] = \mathrm{SVD}(\hat{\mathcal{P}}(:,:,i))$;
5:     $\hat{\mathcal{S}} = \hat{\mathcal{S}}(:,:,i)$;
6:     **for** $j = 1 : \mathrm{size}(\mathrm{diag}(\hat{\mathcal{S}}))$
7:         $w_j = \tau(2\sqrt{2}((\frac{1}{\alpha^2})\sqrt{wh})/(\mathrm{diag}(\hat{\mathcal{S}}(j))^{1/p} + 1^{-6})$;
        get t by calculating Equation (11);
8:         **if** $|\mathrm{diag}(\hat{\mathcal{S}}(j))| \le t$, **then**
9:            $\mathrm{diag}(\hat{\mathcal{S}}(j)) = 0$;
10:         **else**
11:            $k = 0, \mu_k = |\mathrm{diag}(\hat{\mathcal{S}}(j))|$
12:            **for** $k = 0, 1, ..., J$ **do**
13:                $\mu_{k+1} = |\mathrm{diag}(\hat{\mathcal{S}}(j))| - w_j p(\mu_k)^{p-1}$;
14:                $k = k + 1$;
15:            **end**
16:            $\mathrm{diag}(\hat{\mathcal{S}}(j)) = \mathrm{sgn}(\mathrm{diag}(\hat{\mathcal{S}}(j))\mu_k$;
17:            $\hat{S}_{new}(:,:,i) = \mathrm{diag}(\hat{\mathcal{S}}(j))$;
18:         **end**
19:     **end**
20: **end**
21: **for** $i = [\frac{z+1}{2}] + 1, ..., z$
22:     $\hat{\mathcal{U}}(:,:,i) = \mathrm{conj}(\hat{\mathcal{U}}(:,:,z-i+2))$
23:     $\hat{S_{new}}(:,:,i) = \mathrm{conj}(\hat{S_{new}}(:,:,i))$
24:     $\hat{\mathcal{V}}(:,:,i) = \mathrm{conj}(\hat{\mathcal{V}}(:,:,z-i+2))$
25: $\mathcal{U} = \mathrm{ifft}(\hat{\mathcal{U}}, [], 3), \mathcal{S}_{new} = \mathrm{ifft}(\hat{\mathcal{S}}_{new}, [], 3), \mathcal{V} = \mathrm{ifft}(\hat{\mathcal{V}}, [], 3)$
26: **end**
Output: $\mathcal{Z} = \mathcal{U} * \mathcal{S}_{new} * \mathcal{V}^*$;

Algorithm 2 MDLR for HSI anomaly detection.

Input: HSI tensor $\mathcal{Y}$, $\mu_{w,h,d}$, λ, α
Initializtion: $\mathcal{X}$, $\mathcal{S}$, $\mathcal{X}_{w,h,d} = 0$, $\mathcal{Q}_{1,2,3}$, $\mathcal{E} = 0$, $i = 0$.
If i<maxiter or satisfy Equation (11).
    Update $\mathcal{X}$ by Equation (18);
    Update $\mathcal{X}_{w,j,d}$ by algorithm (2);
    Update $\mathcal{S}$ by Equation (23);
    Update $\mathcal{E}$ and $\mathcal{Q}_{1,2,3}$ by Equation (24);
End
compute the anomaly detection map **M** by Equation (25);
Output: anomaly detection map **M**.

3.5. Computational Complexity

Given a tensor $\mathcal{X} \in \mathbb{R}^{w \times h \times d}$, the computational complexity of our model mainly consists of the following two parts: (1) solving the subproblem of Equations (19)–(21) depends on the t-SVD, and the complexity is approximately $O(h^3d + d^3h + h^3w)$; (2) for Equation (22), $O(wh^2d + h(w + h)d\log(d))$ is required. Therefore, the total main cost of the proposed model is $O(h^3d + d^3h + h^3w + wh^2d + h(w + d)d\log(d))$. The computational complexity is studied by comparing their running time on San Diego data as shown in Table 1.

Table 1. Running time of all compared algorithms on San Diego.

HSI Data	RX	LSMAD	LRASR	GTVLRR	DeCNN-AD	PTA	PCA-TLRSR	MDLR
San Diego	2.054	38.46	56.394	214.343	256.589	34.344	8.312	132.46

4. Experimental Results

In this section, we verify the effectiveness of our method on an extensive dataset compared with the SOTA methods. Standard metrics, such as the 2D receiver operating characteristic (ROC) curve [52] and the area under the curve (AUC) metric [53], are used to quantify the results quantitatively. The ROC curve plots the probability of detection (PD) against the false alarm rate (FAR) for all possible thresholds. The AUC is calculated by integrating the area under the ROC curve. To effectively evaluate the detection performance, 3D ROCs [54] generated from 2D ROC and separability maps are also used for quantitative comparison. All the experimental algorithms are performed in MATLAB 2020a on a computer with Core i9-11900KF 3.50GHz CPU and 32-GB of RAM in Windows 11.

4.1. HSI Datasets

San Diego: The dataset is part of a collection captured by the AVIRIS sensor [55], which measures $100 \times 100 \times 189$ and consists mainly of roofs, shadows, and grass, of which aircrafts are considered anomalies to be detected.

HYDICE-Urban: The dataset is collected by a hyperspectral digital imagery collection experiment (HYDICE) sensor [56] for an urban area, including one vegetation area, one built-up area, and several built-up areas of roads and some vehicles. Its spatial resolution is 1 meter. The size of the crop plus the water vapor removal is $80 \times 100 \times 175$. The 21 pixels occupied by vehicles and roofs of different sizes are used as anomalies.

Airport 1–4: The dataset consists of four images of 100 by 100 pixels in 205 bands taken by the airborne visible/infrared imaging spectrometer (AVIRIS) sensor [57]. As above, they include surface vegetation, roads, and buildings as background. Aircraft flying at different altitudes are treated as anomalies.

Urban 1–4: This dataset of four urban scenes is obtained from a class of sensors as with the airport dataset, with pixels of 100×100 and a band number between 190 and 210.

4.2. Compared Methods and Parameter Setting

In this section, we briefly introduce the compared anomaly detection methods and their parameter settings. The parameter values of the compared methods in our experiments are tuned according to the corresponding references.

- **RX [19]:** The classical anomaly detection algorithm calculates the Mahalanobis distance between the pixel under test and the background pixels. The parameter λ of RX is set to $1/\min(w,h)$.
- **LSMAD [29]:** A method based on low-rank sparse matrix decomposition (LRaSAM) with Mahalanobis distance. We set $r = 3$, $k = 0.8$.
- **LRASR [34]:** Learn low-rank linear representation (LRR) of backgrounds by constructing dictionaries. The parameters λ and β of LRASR are set to 0.1 and 0.1 in LRASR.
- **GTVLRR [35]:** Adding total variation (TV) and graph regularization to the restructuring of the background in the LRR-based method, we set $\lambda = 0.5$, $\beta = 0.2$, and $\omega = 0.05$ according to the GTVLRR.
- **PTA [40]:** According to the properties of the spatial and spectral dimensions of the HSI, PTA adds TV into spatial dimensions and low-rank into spectral dimensions. The parameters α, τ, β of PTA are set to 1, 1, and 0.01 separately.
- **DeCNN-AD [36]:** Using convolutional neural network (CNN)-based denoisers as the prior for the dictionary representation coefficients, the cluster number of DeCNN-AD is set to 8 and λ, β are set to 0.01.
- **PCA-TLRSR [43]:** The first method extends LRR to tensor LRR for HSI anomaly detection. The reduced dimensions of PCA are tuned according to PCA-TLRSR and parameter λ is set to 0.4.

4.3. Detection Performance

4.3.1. San Diego

The false-color image, ground-truth map, and detection maps of all compared methods are shown in Figure 2. In the San Diego detection maps, methods such as RX, LSMAD, and LRASR fail to accurately detect the three aircrafts in the upper right corner of the HSI data. DeCNN-AD and GTVLRR have difficulty clearly identifying the outline of the aircrafts. PTA can observe the aircrafts, but it contains some background information, such as road buildings. PCA-TLRSR obtains a relatively good performance by recovering the outline of the aircrafts with less background information. In addition, our method can also capture more detailed features of the aircrafts. Figure 3 shows the anomaly detection evaluation metrics of different anomaly detectors for the San Diego dataset, including 3D ROC curves, 2D ROC curves, and separability maps. The proposed method has a slightly higher detection probability than PCA-TLRSR in 3D and 2D ROC curves. The gap between the background box and the anomaly box shows the degree of separation between background and anomaly on the separability maps. The separability maps on the San Diego dataset are shown in Figure 3. The proposed MDLR obtains a larger gap between the background and the anomaly box over all the other methods compared, which indicates that it has a better ability to separate the background and anomaly. The AUC values in the second row of Table 2 provide further evidence that our method achieves the highest performance on the San Diego dataset.

4.3.2. HYDICE-Ubran

The false-color image, ground-truth map, and detection maps of the competitive methods are visually shown in Figure 4. In the detection maps of all compared methods on the HYDICE-Urban dataset, RX and LSMAD have difficulty in clearly recovering the anomaly information. PTA and PCA-TLRSR can observe the anomaly information, but they also contain a significant amount of background information, such as roads. DeCNN-AD, LRASR, GTVLRR, and our proposed method can clearly identify the anomaly information. However, both LRASR and GTVLRR struggle to detect the anomaly in the lower left corner

of the image. Our proposed method shows an improvement in terms of visual quality and achieves a higher AUC value compared to other methods, as shown in Table 2.

Table 2. AUC values of all compared algorithms on different datasets.

HSI Datasets	RX	LSMAD	LRASR	GTVLRR	DeCNN-AD	PTA	PCA-TLRSR	MDLR
San Diego	0.8885	0.9773	0.9853	0.9795	0.9901	0.9946	<u>0.9957</u>	**0.9976**
HYDICE-Urban	0.9856	0.9901	0.9918	0.9856	0.9935	<u>0.9953</u>	0.9941	**0.9975**
Airport-1	0.8220	0.8334	0.7854	0.9013	0.8503	0.9207	<u>0.9478</u>	**0.9538**
Airport-2	0.8403	0.9189	0.8657	0.8695	0.9204	0.9428	<u>0.9697</u>	**0.9738**
Airport-3	0.9228	0.9401	0.9408	0.9295	0.9434	0.9355	<u>0.9574</u>	**0.9590**
Airport-4	0.9526	0.9862	0.9723	0.9875	0.9897	0.9875	<u>0.9943</u>	**0.9953**
Urban-1	**0.9907**	0.9829	0.9797	0.9605	0.9820	0.9826	<u>0.9902</u>	0.9835
Urban-2	0.9946	0.9836	0.9628	0.8539	<u>0.9973</u>	0.9970	0.9941	**0.9980**
Urban-3	0.9513	0.9636	0.9415	0.9385	0.9394	0.9578	**0.9833**	<u>0.9812</u>
Urban-4	0.9887	0.9809	0.9575	0.9205	0.9868	0.9907	<u>0.9869</u>	**0.9966**
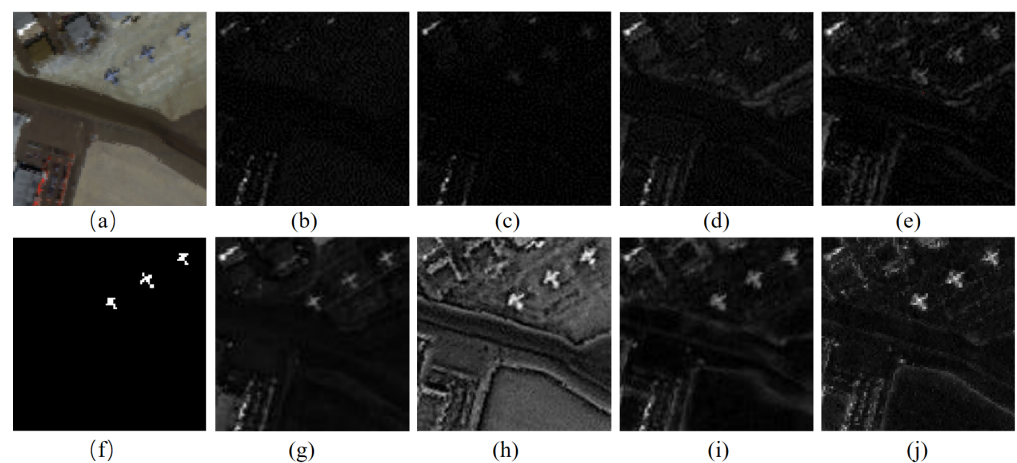
Figure 2. Detection maps obtained by all compared methods on San Diego dataset. (**a**) HSI. (**b**) RX. (**c**) LSMAD. (**d**) LRASR. (**e**) GTVLRR. (**f**) Ground-truth. (**g**) DeCNN-AD. (**h**) PTA. (**i**) PCA-TLRSR. (**j**) Ours.
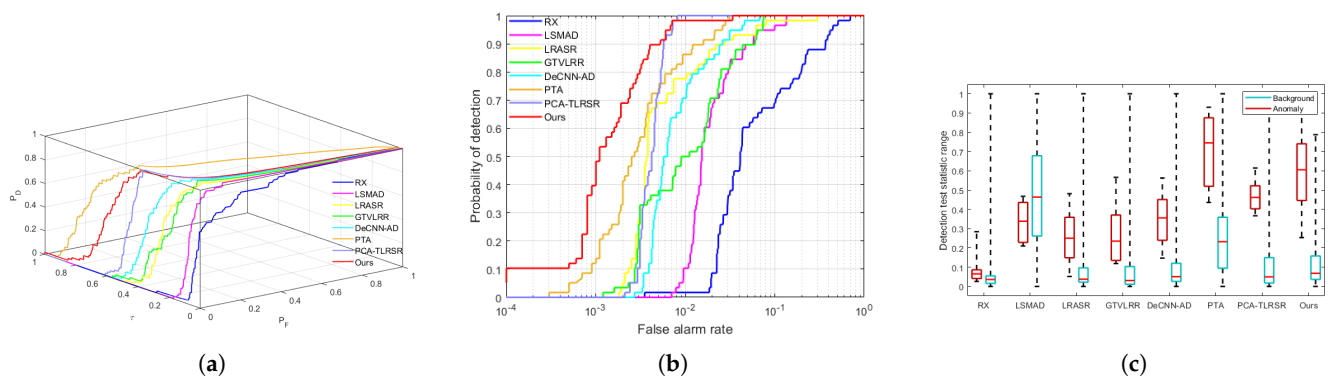

Figure 3. Anomaly detection evaluation metrics obtained by different methods on the San Diego dataset. (**a**) Three-dimensional (3D) ROC curves, (**b**) 2D ROC curves, (**c**) separability map.

4.3.3. Airport 1–4

The AUC values of four airport datasets are provided in Table 2. Our method has achieved the highest values. The false-color images, ground-truth maps, and detection maps of four airport dates are demonstrated in Figure 5. In the detection maps of Airport-1, RX, LSMAD, LRASR, and DeCNN-AD datasets, it is difficult to distinguish the aircrafts.

GTVLRR, PTA, PCA-TLRSR, and ours can distinguish the aircraft in the middle, but they contain a lot of roof information and the aircraft in the upper left corner is not visible. In the detection maps of Airport-2, our method can clearly observe the middle plane of the figure with less background information compared to other methods, but does not fully preserve its edge information due to the effect of some mixed pixels. Compared to the ground truth of Airport-3, the detection maps of all comparison methods can barely detect the outline of an aircraft. This indicates that the existing methods are not sensitive to detecting dense small targets and are easily contaminated by background information. In the detection maps of the Airport-4 dataset, our detection result shows a clear outline of the aircraft compared to other methods. There is no interference from road information compared to LSMAD, LRASR, DeCNN-AD, GTVLRR, PTA, and PCA-TLRSR. The first row of Figures 6 and 7 show the 2D and 3D ROC curves of different anomaly detectors for Airport 1–4. They demonstrate that our method produces detection maps with relatively little interference from background information. The separability maps on the Airport dataset are shown in the first row of Figure 8. The compared methods fail to effectively separate the background boxes and anomaly boxes, while the proposed MDLR achieves a bigger gap.
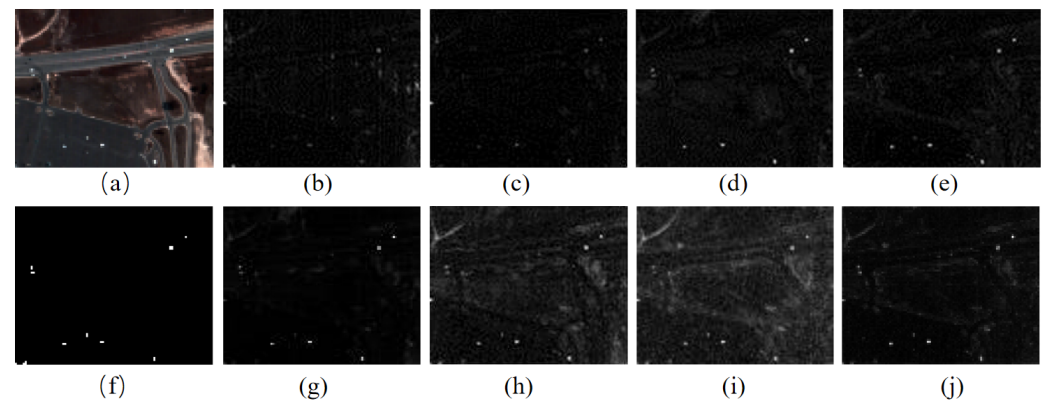
Figure 4. Detection maps on HYDICE-Urban dataset obtained by all compared methods. (**a**) HSI. (**b**) RX. (**c**) LSMAD. (**d**) LRASR. (**e**) GTVLRR. (**f**) Ground-truth. (**g**) DeCNN-AD. (**h**) PTA. (**i**) PCA-TLRSR. (**j**) Ours.
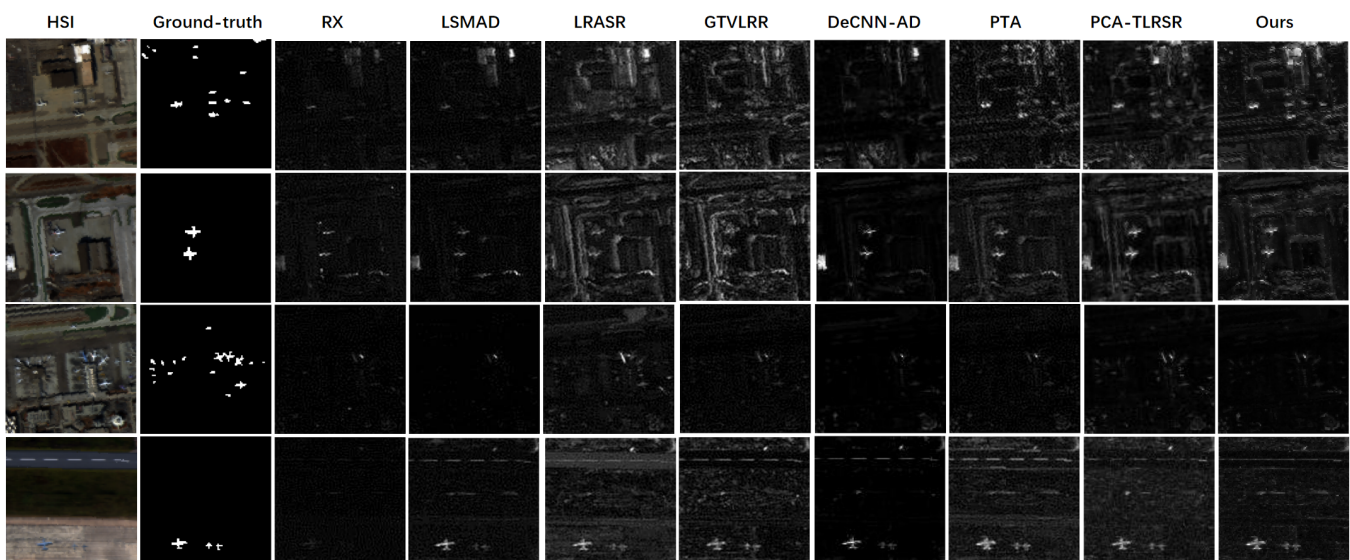

Figure 5. Detection maps obtained by all compared methods on Airport-1 (first line), Airport-2 (second line), Airport-3 (third line), and Airport-4 (forth line) datasets.
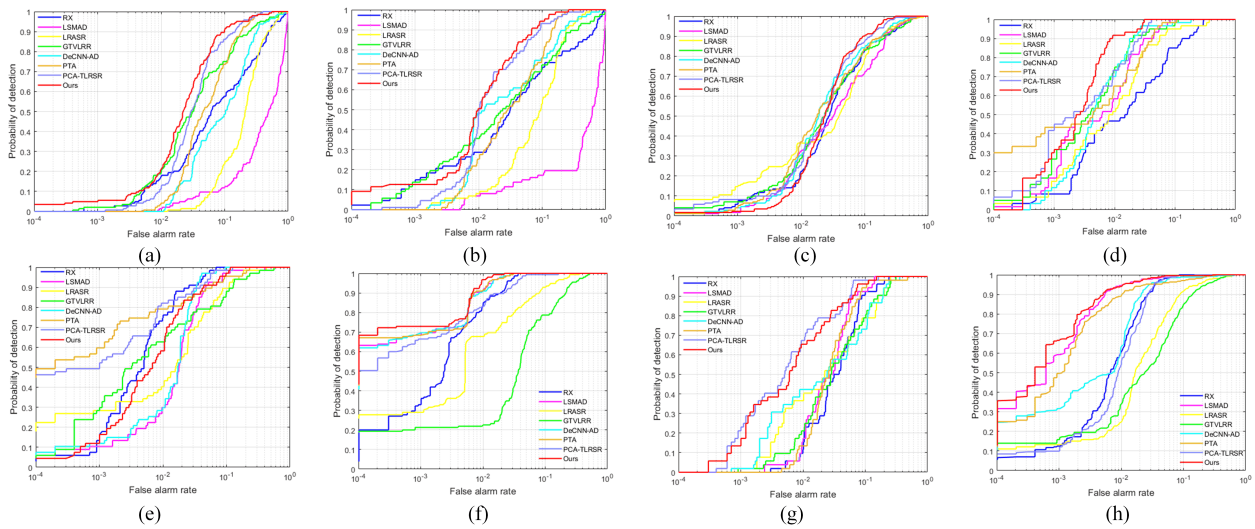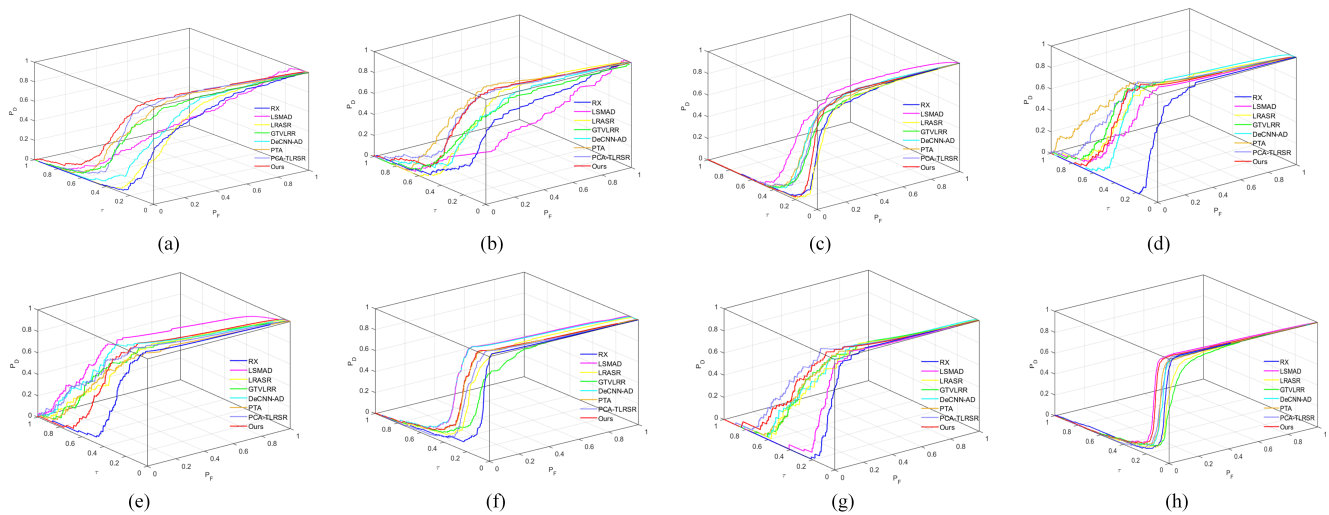
Figure 6. Two-dimensional (2D) ROC curves obtained by all compared methods. (**a**) Airport-1, (**b**) Airport-2, (**c**) Airport-3, (**d**) Airport-4, (**e**) Urban-1, (**f**) Urban-2, (**g**) Urban-3, (**h**) Urban-4.

Figure 7. Three-dimensional (3D) ROC curves obtained by all compared methods. (**a**) Airport-1, (**b**) Airport-2, (**c**) Airport-3, (**d**) Airport-4, (**e**) Urban-1, (**f**) Urban-2, (**g**) Urban-3, (**h**) Urban-4.

4.3.4. Urban 1–4

The AUC values in Table 2 are optimal for all datasets except Urban-1 and Urban-3. For the Urban-1 dataset, from the detection maps shown in Figure 9 we can observe clear lines running through the maps in LSMAD, LRASR, GTVLRR, DeCNN-AD, PTA, and our method. However, the detection map of PCA-TLRSR is difficult to interpret. PCA-TLRSR uses PCA for dimensionality reduction, which aims to reduce noise in the image. The presence of noise in a particular image may have hindered the achievement of optimal results. In the false-color image of Urban-3, there are many large and obvious targets in the background. In the third row of Figure 9, the detection maps of RX and LSMAD barely show the anomaly targets. Other detection algorithms can detect the anomaly targets but retain most of the background contour information. In the detection maps of Urban-2 and Urban-4, compared to the detection maps obtained by other algorithms and the ROC curves in the second row of Figures 6 and 7 obtained by other methods, our method obtains a clearer and more accurate observation of the anomalies. In the separability maps of the Urban dataset in Figure 8, it can be seen that the background boxes and the anomaly

boxes of proposed MDLR are obviously separated, which also proves that our method can achieve effective separation of background and anomaly.
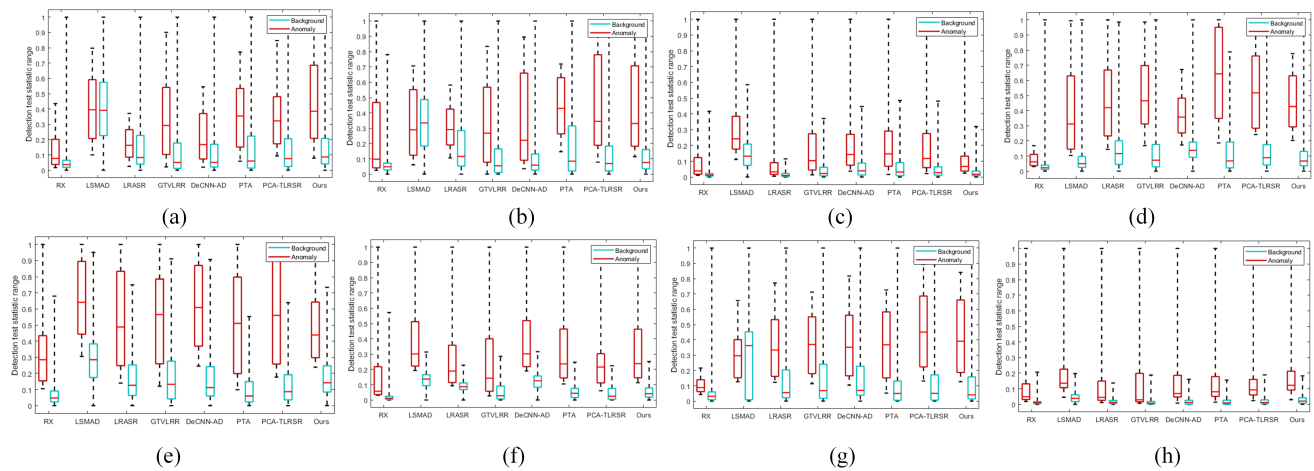

Figure 8. Separability maps obtained by all compared methods. (**a**) Airport-1, (**b**) Airport-2, (**c**) Airport-3, (**d**) Airport-4, (**e**) Urban-1, (**f**) Urban-2, (**g**) Urban-3, (**h**) Urban-4.
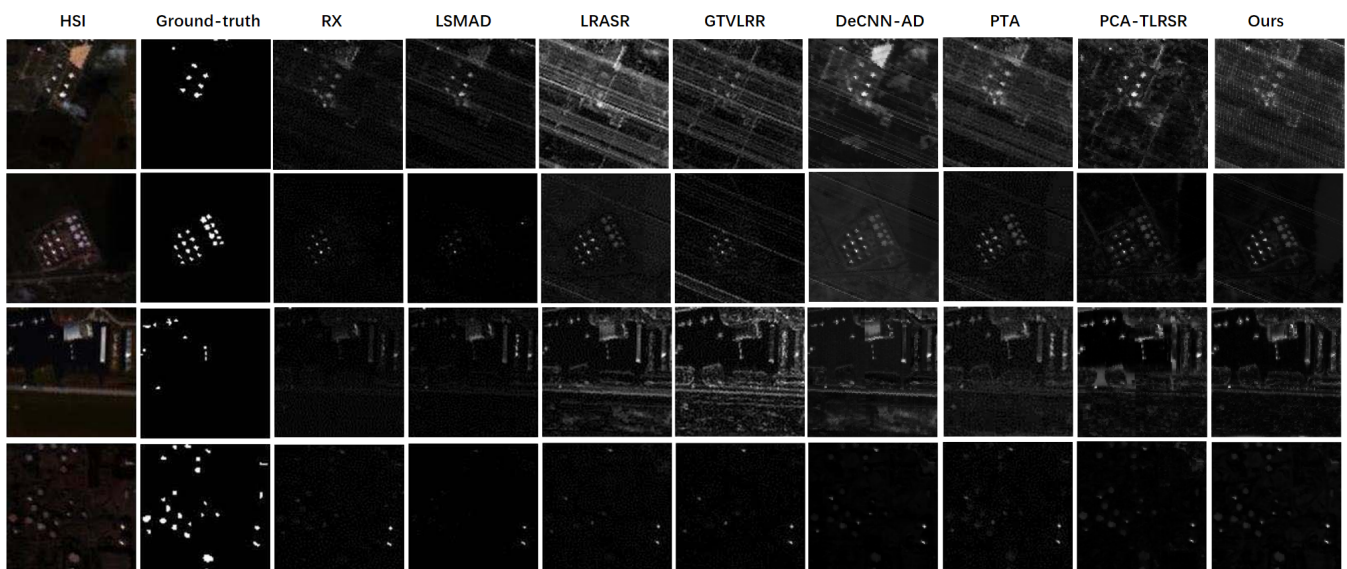
Figure 9. Detection maps obtained by all compared methods on Urban-1 (first line), Urban-2 (second line), Urban-3 (third line), and Urban-4 (forth line) datasets.

4.4. Discussion of Multi-Dimensional Low-Rank

In this section, we analyze the necessity of reconstructing the background with multi-dimensional low-rank.

The discussion on single-dimensional and three-dimensional low-rank: The results presented in Table 3 show that reconstructing the background along a multi-dimensional (w, h, d dimension) gives significantly higher AUC values compared to reconstructing along a single-dimensional (d dimension). This improvement is evident across all datasets, with notable increases in AUC values for the HYDICE-Urban and Airport-1 HSI datasets. Specifically, the AUC values for anomaly detection increased by 4 percent and 6 percent for HYDICE-Urban and Airport-1 datasets, respectively, compared to the one-dimensional reconstruction. These improvements demonstrate the benefit of using multi-dimensional information to effectively separate the background from anomalies in the HSI data. By considering the data from multiple dimensions simultaneously, the proposed method is able to capture more comprehensive and discriminative information about the background,

leading to an improved detection performance. This highlights that multi-dimensional has an advantage over single-dimensional in separating background and anomaly.

Table 3. AUC values of single-dimensional and multi-dimensional low-rank.

HSI dataset	San Diego	Airport-1	Airport-2	Airport-3	Airport-4
S-dimensional	0.9966	0.8957	0.9655	0.9345	0.9921
M-dimensional	**0.9976**	**0.9538**	**0.9738**	**0.9590**	**0.9953**

HSI dataset	HYDIE-Urban	Urban-1	Urban-2	Urban-3	Urban-4
S-dimensional	0.9546	0.9619	0.9928	0.9527	**0.9966**
M-dimensional	**0.9975**	**0.9835**	**0.9980**	**0.9812**	**0.9966**

The discussion on two-dimensional and three-dimensional low-rank: From the results shown in Figure 10, it can be seen that the AUC values obtained by reconstructing the background with different combinations of two dimensions are different. In the case of Airport-1 in Figure 10a,b, reconstructing the background with $\mathcal{X}_d$ and $\mathcal{X}_w$ achieves higher AUC values compared to reconstructing with $\mathcal{X}_d$ and $\mathcal{X}_h$. On the other hand, for Airport-3 in Figure 10c,d, reconstructing the background with $\mathcal{X}_h$ and $\mathcal{X}_w$ yields higher AUC values compared to reconstructing with $\mathcal{X}_d$ and $\mathcal{X}_h$. These results indicate that there is no fixed combination of two dimensions that consistently gives the best performance for background tensor reconstruction. The optimal combination may vary depending on the specific dataset and the characteristics of the HSI data. Therefore, it is important to explore and analyze the relationship between different dimensional background tensors to achieve the best reconstruction results.
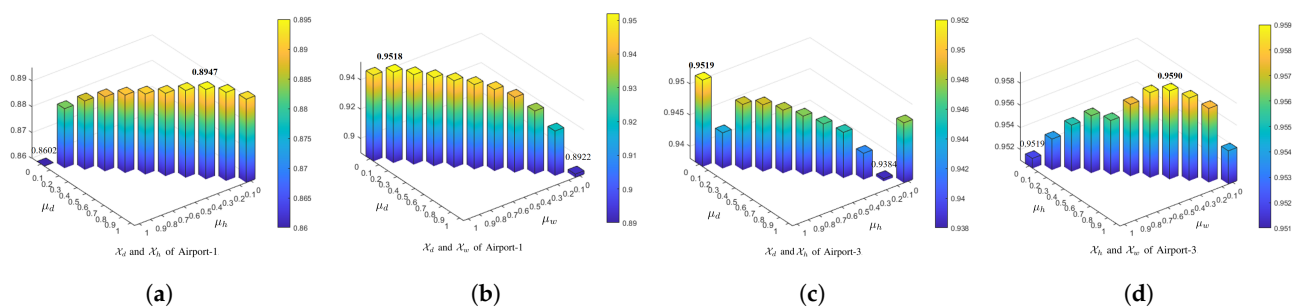

Figure 10. AUC value bars obtained from coefficients μ via the reconstruction of the background with two different dimensions. (**a**) Airport-1: μ_w and μ_h, (**b**) Airport-1: μ_w and μ_d, (**c**) Airport-3: μ_w and μ_h, (**d**) Airport-3: μ_h and μ_d.

In the analysis of the reconstruction of the background along two different dimensions of the HSI, the focus is on investigating the relationship between these dimensions. For this purpose, the experimental datasets Airport-1 and Airport-3 are selected. Two dimensions of the data are chosen to reconstruct the background, resulting in four different comparison experiments: (a) Reconstruction using $\mathcal{X}_d$ and $\mathcal{X}_h$ from Airport-1; (b) Reconstruction using $\mathcal{X}_d$ and $\mathcal{X}_w$ from Airport-1; (c) Reconstruction using $\mathcal{X}_d$ and $\mathcal{X}_h$ from Airport-3; (d) Reconstruction using $\mathcal{X}_h$ and $\mathcal{X}_w$ from Airport-3. The aim of these experiments is to investigate the performance and effectiveness of background reconstruction when using different combinations of two dimensions.

Effects of coefficient μ in two reconstructed background tensors: This study investigates the relationship between the reconstructed background tensors from different dimensions, which aims to better understand their impact on anomaly detection performance. The coefficients between two dimensions are varied in the range [0:0. 1:1] so that the results can be observed when each dimension acts alone and when two dimensions work together. The AUC values of the comparison experiments are visualized in Figure 10. In Figure 10a, the individual reconstruction of $\mathcal{X}_d$ in Airport-1 achieves an accu-

racy of 0.8922, while the individual reconstruction of $\mathcal{X}_h$ only achieves an accuracy of 0.8602. However, when $\mathcal{X}_d$ and $\mathcal{X}_h$ are combined in the reconstruction, they achieve a maximum AUC value of 0.8947. Similarly, in Figure 10b, the individual reconstruction of $\mathcal{X}_d$ in Airport-1 gives an AUC of 0.8922, while the individual reconstruction of $\mathcal{X}_w$ gives an AUC of 0.9487. However, their combination leads to a maximum AUC value of 0.9518. The same trend can be seen in Figure 10c,d for Airport-3. From Figure 10a,c, it is clear that both Airport-1 and Airport-3 benefit from the background reconstruction using $\mathcal{X}_d$ and $\mathcal{X}_h$. Interestingly, the coefficients of the reconstructed background from the same dimensions differ between the two datasets, indicating that the relationship between the reconstructed background tensors can vary for different datasets. The final AUC values in Table 2 also support the effectiveness of reconstructing the background along multiple dimensions. Although the AUC values of Airport-1 and Airport-3 in Figure 10 are slightly lower, they still demonstrate the validity of the multi-dimensional reconstruction approach in improving the anomaly detection performance.

4.5. Parameter Tuning

In this section, we focus on analyzing the effect of the values of λ and p on the AUC results.

(1) **Effects of parameter λ:** The influence of parameter λ on model performance was analysed on four HSI datasets. The parameter λ was selected from the set [0.001, 0.005, 0.01, 0.05, 0.1, 1, 2] while keeping the other parameters fixed. AUC value curves with respect to λ on four datasets are shown in Figure 11a. The AUC values of the San Diego, HYIDE-Urban and Airport-4 datasets reach their maximum value when λ is equal to 1. Airport-1 and Airport-2 datasets both reach their maximum value when λ is 2. Airport-3 has a downward trend when λ is 1. For the experiment as a whole, when λ is 0.1 or 1, the AUC values are relatively stable. So for San Diego, HYDICE-Urban, and Airport-4 datasets, λ is set to 1. λ is set to 0.1 on Airport 1-3.

(2) **Effects of parameter p:** The AUC value curves for the parameter p are shown in Figure 11b. The parameter p has been chosen in the range [0.1, 1]. As the value of p increases, the AUC values on different HSI datasets start to improve. The growing trend of the AUC values for San Diego, HYDICE-Urban, Airport-2, Airport-3, and Airport-4 datasets tends to level off when p reaches 0.6. However, the AUC value of Airport-1 continues to increase as p increases. Based on the AUC value curves, the following choices of p are made. For the San Diego, HYDICE-Urban, Airport-1, and Airport-4 datasets, p is set to 1. For the Airport-2 dataset, p is set to 0.9. For the Airport-3 dataset, p is set to 0.6.
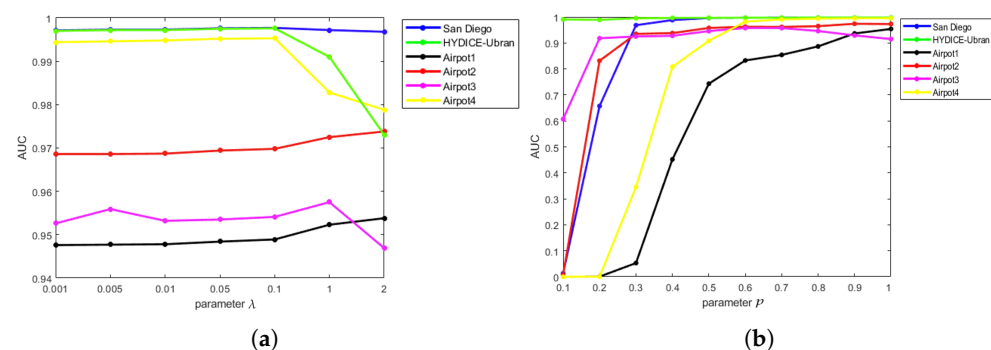

| (a) | (b) |

Figure 11. The effect of parameter tuning on AUC values. (**a**) AUC value curves with respect to λ on four datasets; (**b**) AUC value curves with respect to p on three datasets.

5. Conclusions

In this paper, a novel multi-dimensional low-rank (MDLR) method is proposed for HSI anomaly detection. The MDLR method exploits the low-rank properties of HSI from three dimensions, namely the spatial and spectral dimensions. Multi-dimensional background tensors are reconstructed. Weighted Schatten p-norm minimization is used to enforce the

low-rank constraints. In addition, the $L_{F,1}$ norm is used to penalize the anomaly tensor to promote joint spectral–spatial sparsity. The optimization problem is solved using ADMM. Experimental results on real HSI datasets demonstrate its effectiveness compared with SOTA in terms of anomaly detection. However, one of the major limitations of MDLR is the computational complexity introduced by the t-SVD operation, especially when dealing with large spectral bands. In future work, we would like to try to incorporate dimensionality reduction preprocessing techniques, which is a promising direction to take in order to address this computational challenge.

References

1. Zhang, X.; Wu, H.; Sun, H.; Ying, W. Multireceiver SAS Imagery Based on Monostatic Conversion. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 10835–10853. [CrossRef]
2. Zhu, J.; Song, Y.; Jiang, N.; Xie, Z.; Fan, C.; Huang, X. Enhanced Doppler Resolution and Sidelobe Suppression Performance for Golay Complementary Waveforms. *Remote Sens.* **2023**, *15*, 2452. [CrossRef]
3. Hong, D.; Gao, L.; Yokoya, N.; Yao, J.; Chanussot, J.; Du, Q.; Zhang, B. More Diverse Means Better: Multimodal Deep Learning Meets Remote-Sensing Imagery Classification. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 4340–4354. [CrossRef]
4. Liu, F.; Wang, Q. A sparse tensor-based classification method of hyperspectral image. *Signal Process.* **2020**, *168*, 107361. [CrossRef]
5. An, W.; Zhang, X.; Wu, H.; Zhang, W.; Du, Y.; Sun, J. LPIN: A Lightweight Progressive Inpainting Network for Improving the Robustness of Remote Sensing Images Scene Classification. *Remote Sens.* **2021**, *14*, 53. [CrossRef]
6. Tan, K.; Wu, F.; Du, Q.; Du, P.; Chen, Y. A Parallel Gaussian–Bernoulli Restricted Boltzmann Machine for Mining Area Classification With Hyperspectral Imagery. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* **2019**, *12*, 627–636. [CrossRef]
7. Ren, Z.; Sun, L.; Zhai, Q.; Liu, X. Mineral Mapping with Hyperspectral Image Based on an Improved K-Means Clustering Algorithm. In Proceedings of the IGARSS 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 2989–2992.
8. Rukhovich, D.I.; Koroleva, P.V.; Rukhovich, D.D.; Rukhovich, A.D. Recognition of the Bare Soil Using Deep Machine Learning Methods to Create Maps of Arable Soil Degradation Based on the Analysis of Multi-Temporal Remote Sensing Data. *Remote Sens.* **2022**, *14*, 2224. [CrossRef]
9. Wang, Q.; Li, J.; Shen, Q.; Wu, C.; Yu, J. Retrieval of water quality from China's first satellite-based Hyperspectral Imager (HJ-1A HSI) data. In Proceedings of the 2010 IEEE International Geoscience and Remote Sensing Symposium, Honolulu, Hawaii, USA, 25–30 July 2010; pp. 371–373.
10. Zhang, X.; Han, L.; Dong, Y.; Shi, Y.; Huang, W.; Han, L.; González-Moreno, P.; Ma, H.; Ye, H.; Sobeih, T. A Deep Learning-Based Approach for Automated Yellow Rust Disease Detection from High-Resolution Hyperspectral UAV Images. *Remote Sens.* **2019**, *11*, 1554. [CrossRef]
11. Wan, Y.; Hu, X.; Zhong, Y.; Ma, A.; Wei, L.; Zhang, L. Tailings Reservoir Disaster and Environmental Monitoring Using the UAV-ground Hyperspectral Joint Observation and Processing: A Case of Study in Xinjiang, the Belt and Road. In Proceedings of the IGARSS 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 9713–9716.

12. Farrar, M.B.; Wallace, H.M.; Brooks, P.R.; Yule, C.M.; Tahmasbian, I.; Dunn, P.K.; Bai, S.H. A Performance Evaluation of Vis/NIR Hyperspectral Imaging to Predict Curcumin Concentration in Fresh Turmeric Rhizomes. *Remote Sens.* **2021**, *13*, 1807. [CrossRef]

13. Légaré, B.; Bélanger, S.; Singh, R.K.; Bernatchez, P.; Cusson, M. Remote Sensing of Coastal Vegetation Phenology in a Cold Temperate Intertidal System: Implications for Classification of Coastal Habitats. *Remote Sens.* **2022**, *14*, 3000. [CrossRef]

14. Cen, Y.; Huang, Y.H.; Hu, S.; Zhang, L.; Zhang, J. Early Detection of Bacterial Wilt in Tomato with Portable Hyperspectral Spectrometer. *Remote Sens.* **2022**, *14*, 2882. [CrossRef]

15. Yin, C.; Lv, X.; Zhang, L.; Ma, L.; Wang, H.; Zhang, L.; Zhang, Z. Hyperspectral UAV Images at Different Altitudes for Monitoring the Leaf Nitrogen Content in Cotton Crops. *Remote Sens.* **2022**, *14*, 2576. [CrossRef]

16. Thornley, R.H.; Verhoef, A.; Gerard, F.F.; White, K. The Feasibility of Leaf Reflectance-Based Taxonomic Inventories and Diversity Assessments of Species-Rich Grasslands: A Cross-Seasonal Evaluation Using Waveband Selection. *Remote. Sens.* **2022**, *14*, 2310. [CrossRef]

17. Pang, D.; Ma, P.; Shan, T.; Li, W.; Tao, R.; Ma, Y.; Wang, T. STTM-SFR: Spatial–Temporal Tensor Modeling With Saliency Filter Regularization for Infrared Small Target Detection. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–18. [CrossRef]

18. Gao, Q.; Zhang, P.; Xia, W.; Xie, D.; Gao, X.; Tao, D. Enhanced Tensor RPCA and its Application. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *43*, 2133–2140. [CrossRef] [PubMed]

19. 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]

20. Manolakis, D.G.; Shaw, G.A. Detection algorithms for hyperspectral imaging applications. *IEEE Signal Process. Mag.* **2002**, *19*, 29–43. [CrossRef]

21. Molero, J.M.; Garzón, E.M.; García, I.; Plaza, A.J. Analysis and Optimizations of Global and Local Versions of the RX Algorithm for Anomaly Detection in Hyperspectral Data. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* **2013**, *6*, 801–814. [CrossRef]

22. Taitano, Y.P.; Geier, B.A.; Bauer, K.W. A Locally Adaptable Iterative RX Detector. *Signal Process.* **2010**, *2010*, 1–10. [CrossRef]

23. Sun, W.; Liu, C.; Li, J.; Lai, Y.M.; Li, W. low-rank and sparse matrix decomposition-based anomaly detection for hyperspectral imagery. *J. Appl. Remote Sens.* **2014**, *8*, 15823048. [CrossRef]

24. Farrell, M.D.; Mersereau, R.M. On the impact of covariance contamination for adaptive detection in hyperspectral imaging. *IEEE Signal Process. Lett.* **2005**, *12*, 649–652. [CrossRef]

25. Billor, N.; Hadi, A.S.; Velleman, P.F. BACON: blocked adaptive computationally efficient outlier nominators. *Comput. Stat. Data Anal.* **2000**, *34*, 279–298. [CrossRef]

26. Sun, W.; Tian, L.; Xu, Y.; Du, B.; Du, Q. A Randomized Subspace Learning Based Anomaly Detector for Hyperspectral Imagery. *Remote Sens.* **2018**, *10*, 417. [CrossRef]

27. Sun, W.; Yang, G.; Li, J.; Zhang, D. Randomized subspace-based robust principal component analysis for hyperspectral anomaly detection. *J. Appl. Remote Sens.* **2018**, *12*, 015015. [CrossRef]

28. Qu, Y.; Wang, W.; Guo, R.; Ayhan, B.; Kwan, C.; Vance, S.D.; Qi, H. Hyperspectral Anomaly Detection Through Spectral Unmixing and Dictionary-Based low-rank Decomposition. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 4391–4405. [CrossRef]

29. Zhang, Y.; Du, B.; Zhang, L.; Wang, S. A low-rank and Sparse Matrix Decomposition-Based Mahalanobis Distance Method for Hyperspectral Anomaly Detection. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 1376–1389. [CrossRef]

30. Xu, Y.; Du, B.; Zhang, L.; Chang, S. A low-rank and Sparse Matrix Decomposition- Based Dictionary Reconstruction and Anomaly Extraction Framework for Hyperspectral Anomaly Detection. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 1248–1252. [CrossRef]

31. Li, L.; Li, W.; Du, Q.; Tao, R. low-rank and Sparse Decomposition With Mixture of Gaussian for Hyperspectral Anomaly Detection. *IEEE Trans. Cybern.* **2020**, *51*, 4363–4372. [CrossRef]

32. Liu, G.; Lin, Z.; Yu, Y. Robust Subspace Segmentation by low-rank Representation. In Proceedings of the International Conference on Machine Learning, Haifa, Israel, 21–24 June 2010.

33. Liu, G.; Lin, Z.; Yan, S.; Sun, J.; Yu, Y.; Ma, Y. Robust Recovery of Subspace Structures by low-rank Representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *35*, 171–184. [CrossRef]

34. Xu, Y.; Wu, Z.; Li, J.; Plaza, A.J.; Wei, Z. Anomaly Detection in Hyperspectral Images Based on low-rank and Sparse Representation. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 1990–2000. [CrossRef]

35. Cheng, T.; Wang, B. Graph and Total Variation Regularized low-rank Representation for Hyperspectral Anomaly Detection. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 391–406. [CrossRef]

36. Fu, X.; Jia, S.; Zhuang, L.; Xu, M.; Zhou, J.; Li, Q. Hyperspectral Anomaly Detection via Deep Plug-and-Play Denoising CNN Regularization. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 9553–9568. [CrossRef]

37. Zhang, K.; Zuo, W.; Zhang, L. FFDNet: Toward a Fast and Flexible Solution for CNN-Based Image Denoising. *IEEE Trans. Image Process.* **2017**, *27*, 4608–4622. [CrossRef] [PubMed]

38. Guo, S.; Chen, X.; Jia, H.; Han, Z.; Duan, Z.; Tang, Y. Fusing Hyperspectral and Multispectral Images via low-rank Hankel Tensor Representation. *Remote Sens.* **2022**, *14*, 4470. [CrossRef]

39. Zhang, Z.; Ding, C.; Gao, Z.; Xie, C. ANLPT: Self-Adaptive and Non-Local Patch-Tensor Model for Infrared Small Target Detection. *Remote Sens.* **2023**, *15*, 1021. [CrossRef]

40. Li, L.; Li, W.; Qu, Y.; Zhao, C.; Tao, R.; Du, Q. Prior-Based Tensor Approximation for Anomaly Detection in Hyperspectral Imagery. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *33*, 1037–1050. [CrossRef] [PubMed]

41. Song, S.; Zhou, H.; Gu, L.; Yang, Y.; Yang, Y. Hyperspectral Anomaly Detection via Tensor- Based Endmember Extraction and low-rank Decomposition. *IEEE Geosci. Remote. Sens. Lett.* **2020**, *17*, 1772–1776. [CrossRef]
42. Shang, W.; Peng, J.; Wu, Z.; Xu, Y.; Jouni, M.; Mura, M.D.; Wei, Z. Hyperspectral Anomaly Detection via Sparsity of Core Tensor Under Gradient Domain. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 1–16. [CrossRef]
43. Wang, M.; Wang, Q.; Hong, D.; Roy, S.K.; Chanussot, J. Learning Tensor low-rank Representation for Hyperspectral Anomaly Detection. *IEEE Trans. Cybern.* **2022**, *53*, 679–691. [CrossRef]
44. Sun, S.; Liu, J.; Zhang, Z.; Li, W. Hyperspectral Anomaly Detection Based on Adaptive low-rank Transformed Tensor. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, 1–13. [CrossRef]
45. Lu, C.; Feng, J.; Chen, Y.; Liu, W.; Lin, Z.; Yan, S. Tensor Robust Principal Component Analysis: Exact Recovery of Corrupted low-rank Tensors via Convex Optimization. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26 June–1 July 2016; pp. 5249–5257.
46. Zhang, D.; Hu, Y.; Ye, J.; Li, X.; He, X. Matrix completion by Truncated Nuclear Norm Regularization. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 2192–2199.
47. Oh, T.H.; Kim, H.; Tai, Y.W.; Bazin, J.C.; Kweon, I.S. Partial Sum Minimization of Singular Values in RPCA for Low-Level Vision. In Proceedings of the 2013 IEEE International Conference on Computer Vision, Sydney, Australia, 1–8 December 2013; pp. 145–152.
48. Gu, S.; Zhang, L.; Zuo, W.; Feng, X. Weighted Nuclear Norm Minimization with Application to Image Denoising. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 2862–2869.
49. Lu, C.; Feng, J.; Chen, Y.; Liu, W.; Lin, Z.; Yan, S. Tensor Robust Principal Component Analysis with a New Tensor Nuclear Norm. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *42*, 925–938. [CrossRef] [PubMed]
50. Liu, G.; Yan, S. Latent low-rank Representation for subspace segmentation and feature extraction. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 1615–1622.
51. Xie, Y.; Gu, S.; Liu, Y.; Zuo, W.; Zhang, W.; Zhang, L. Weighted Schatten p-Norm Minimization for Image Denoising and Background Subtraction. *IEEE Trans. Image Process.* **2015**, *25*, 4842–4857. [CrossRef]
52. Kerekes, J.P. Receiver Operating Characteristic Curve Confidence Intervals and Regions. *IEEE Geosci.* **2008**, *5*, 251–255. [CrossRef]
53. Khazai, S.; Homayouni, S.; Safari, A.; Mojaradi, B. Anomaly Detection in Hyperspectral Images Based on an Adaptive Support Vector Method. *IEEE Geosci.* **2011**, *8*, 646–650. [CrossRef]
54. Chang, C.I. Comprehensive Analysis of Receiver Operating Characteristic (ROC) Curves for Hyperspectral Anomaly Detection. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–24. [CrossRef]
55. Li, W.; Du, Q. A survey on representation-based classification and detection in hyperspectral remote sensing imagery. *Pattern Recognit. Lett.* **2016**, *83*, 115–123. [CrossRef]
56. 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]
57. Kang, X.; Zhang, X.; Li, S.; Li, K.; Li, J.Y.; Benediktsson, J.A. Hyperspectral Anomaly Detection with Attribute and Edge-Preserving Filters. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 5600–5611. [CrossRef]