


Hyperspectral Simultaneous Anomaly Detection and Denoising: Insights From Integrative Perspective

Minghua Wang , *Member, IEEE*, Lianru Gao , *Senior Member, IEEE*, Longfei Ren ,
Xian Sun , *Senior Member, IEEE*, and Jocelyn Chanussot , *Fellow, IEEE*

Abstract—In data acquisition and transmission, hyperspectral images are inevitably corrupted by additive noises, making it challenging to accurately observe and recognize the materials on the surface of the Earth. However, scholars have been addicted to developing numerous complex methods for separable two-stage denoising and anomaly detection (AD) tasks over the past years, rarely paying attention to the real effect of noises for subsequent intelligent interpretation. To this date, we propose a hierarchical integration framework for hyperspectral simultaneous AD and denoising (HyADD). Joint AD and denoising are mutually integrated and their outputs in each iteration stimulate each other, breaking through the respective performance bottlenecks of the separable two-stage scheme. Inspired by spatial–spectral gradient domain-based constraint, HyADD removes additive noises and preserves advantageous image smoothness information to improve intermediate detection performances in the iteration loop. Conversely, with the assistance of the antinoise dictionary conduction and the subspace domain-based low-rankness, the identification of anomalies with different features from the background can provide effective feedback to the denoising process. The proposed algorithm is efficiently solved by the well-designed linearized alternating direction method of multipliers with an adaptive penalty. A comparison with the existing well-known AD algorithms via simulated and real-world experiments establishes the competitiveness of the proposed HyADD with state-of-the-art methods.

Index Terms—Anomaly detection (AD), denoising, feedback, integration framework, linearized alternating direction method of multipliers.

Manuscript received 28 April 2024; revised 10 July 2024; accepted 27 July 2024. Date of publication 2 August 2024; date of current version 15 August 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 42325104 and Grant 62201552, in part by the China Postdoctoral Science Foundation under Grant 2023T160656 and Grant 2023M733578, and in part by the MIAI@Grenoble Alpes under Grant ANR-19-P3IA-0003 and Grant AXA Research Fund. (*Corresponding author: Lianru Gao.*)

Minghua Wang is with the Institute of Robotics and Automatic Information System (IRAIS), College of Artificial Intelligence, Tianjin Key Laboratory of Intelligent Robotics (tjKLIR), Nankai University, Tianjin 300350, China (e-mail: wangminghua@nankai.edu.cn).

Lianru Gao and Longfei Ren are with the Key Laboratory of Computational Optical Imaging Technology, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China (e-mail: gaolr@aircas.ac.cn; renlongfei0131@163.com).

Xian Sun is with the Key Laboratory of Network Information System Technology (NIST), Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100190, China, and also with the School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100190, China (e-mail: sunxian@aircas.ac.cn).

Jocelyn Chanussot is with the CNRS, Grenoble INP, GIPSA-Lab, Université Grenoble Alpes, 38000 Grenoble, France, and also with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China (e-mail: jocelyn@hi.is).

The codes and real dataset will be available at <https://github.com/MinghuaWang123/HyADD>.

Digital Object Identifier 10.1109/JSTARS.2024.3437460

I. INTRODUCTION

RECENT advances in hyperspectral (HS) imaging techniques have provided users with abundant spatial and spectral information on Earth’s surface, deriving a wide range of applications, such as superresolution [1], [2], classification [3], [4], unmixing [5], [6], and target detection [7], [8]. Due to abominable environmental change and unfavourable equipment conditions, real-world noise usually exists in all the HS imaging systems and degrades the performances of results and analysis [9]. Therefore, HS image denoising is considered as a crucial preprocessing step for the subsequent refinement of spectral information [10]. As a branch of HS target location, anomaly detection (AD) aims at the separation of abnormal objects from backgrounds. For instance, the sudden appearance of unknown aircraft or vessels in village or wharf scenes can be marked as anomalies or outliers. While the precise definition of HS anomalous objects is often context-dependent, they are typically characterized by low occurrence probabilities and a lack of prior information [11], [12], [13]. These anomalies usually manifest as several pixels or small objects with distinct shapes within an observation scene, and their spectral signatures significantly differ from the surrounding background environment.

Without the available prior information on anomalies or background, existing methodologies usually focus on the different feature extraction for accurate distinction. As a typical statistical AD benchmark, the classical Reed–Xiaoli (RX) algorithm was proposed early in 1990 [14]. Based on a generalized likelihood ratio, RX is a constant false alarm rate (FAR) detector. With the background assumption obeying a Gaussian distribution, the mean vector and covariance matrix of all HS samples are calculated to estimate the Mahalanobis distance between a testing pixel and the background clutter. GRX considers all the pixels from one whole HS image, whereas local RX takes a local region as the background [15], [16]. However, the conception of a multivariate Gaussian distribution deteriorates the detection performances of basic RX algorithms in practice. Subsequently, many modified RX approaches have sprung up over the last two decades. A nonlinear kernel RX (KRX) method was proposed to map HS images in the feature spaces [17]. The principal component analysis (PCA) of the background covariance matrix was used to gain subspace RX [18]. Segment RX was proposed via the critical factors that affect the estimated results of local background covariance matrices [19]. A modified KRX was

proposed by using a Gaussian background purification and an inverse-of matrix-free model [20].

Nowadays, with the advent of machine learning and compressive sensing theory, representation-based HS AD methods have gradually become a research hotspot, specifically including sparse representation (SR), collaborative representation (CR), and low-rank representation (LRR). Ling et al. [21] advocated the concept that the background is approximately represented as a sparse linear combination of its spatial surrounding pixels, but anomalies cannot when they are separated from its neighborhood. Motivated by homogeneous region-based joint SR, Yang et al. [22] presented a robust background feature extraction method. Vafadar and Ghassemian [23] developed a modified CR-based with an outlier removal anomaly algorithm. The proposed detector obtained the background estimation by a sliding dual-window adaptively. Wang et al. [24] combined a suitable weight learning strategy with CR to achieve a joint objective function. Nevertheless, the above-mentioned SR and CR approaches have a limitation of pixel-pixel detection [25].

Liu et al. [26] proved LRR enables exact recovery of potential subspace structures even when the data are corrupted by outliers. Without any statistical hypothesis in advance, many scholars masterly explore the low rankness of the background and the sparsity of anomalies. The correlation of HS bands leads to the low rank (LR) characteristic of an unfolding matrix whose columns are composed of the vectorized background part of each band. The sparsity originates from a small proportion of anomalies and a low probability of occurrence. The robust principal component analysis (RPCA) [27] algorithm was employed to separate an HS image into the LR and SR parts [28]. In [29] and [30], the conception of LR and sparse matrix decomposition (LRaSMD) has been introduced for HS AD. An LRaSMD-based method [31] adopted the go decomposition algorithm to solve the AD optimization with the Mahalanobis distance estimation. Xu et al. [32] explored the background information with a new antinnoise dictionary learning strategy and fused a sparsity-inducing regularization into an LRR framework. Li et al. [33] designed a global detector via LRaSMD with the assumption of the mixture of Gaussian. Cheng and Wang [34] injected the graph and TV regularization into the LRR formulation (GTVLRR) to preserve the local geometrical information. With the wide application of deep learning, Xie et al. [35] formulated the background in a weakly supervised manner. An LRR-based dictionary was constructed by observed training data and hidden learned data. Wang et al. [36] proposed a novel tensor LR and sparse representation (LRASR) method with a PCA preprocessing step to exploit the 3-D inherent structure of HS images. This was the first time to expand the concept of tensor LR representation in HS AD.

In practical HS AD applications, observed images are unavoidably caused by a variety of noises during data acquisition and transmission [37], [38]. A natural and intuitive pipeline to address this issue is to remove the noise first with the subsequent AD procedure. When the slight denoising preprocessing step may eliminate some of the noise, the sparse residual will confuse the identification of anomalies. On the contrary, the crude denoising processing destroys the image details, such as

the geometrical structure and edges of anomalies, leading to the loss of detection accuracy. This two-stage scheme cuts off the inherent relationship between HS denoising and AD. In addition, existing AD models rarely analyze the explicit influence of noise on detection performances. In Cheng and Wang's [34] work, many detectors are compared under different levels of just zero-mean Gaussian noise. As the noise level rises, the detection performances of all the algorithms gradually deteriorate. Moreover, due to the spatial TV regularization, GTVLRR appears the most stable among all the competing detectors. However, due to the independent denoising optimization ignored in the iterative loop, GTVLRR hardly copes with additive noise (i.e., Gaussian noise, salt-and-pepper noise, deadlines, and stripes). In this article, we mainly attend to HS AD in the presence of additive noise. We analyze the impact of additive noise on AD and take a coupled denoising and AD task into account in the field of HS remote sensing (RS). This optimization policy is able to overcome the irrelevance of two-stage denoising and AD. In the integration iteration loop, the accuracy of anomaly locations is improved by denoising, and simultaneously, denoising is developed by the LRR-based AD model. The denoised HS image and the anomaly locations are produced by an integration framework. We summarize the main contributions of this study as follows.

- 1) We blaze a trail for AD under additive noise in the field of practical HS RS applications. AD is specifically analyzed in the presence of additive noise and a coupled AD-denoising task is mathematically formulated in a novel integration manner.
- 2) We exploit spatial-spectral gradient domain-based smoothness and subspace domain-based low-rankness to promote denoising and AD each other. Noise removal contributes to improving the detection performances of anomaly locations. Assisted by a novel antinnoise dictionary construction, the LRR-based AD model provides positive feedback for additive noise removal.
- 3) The linearized alternating direction method with adaptive penalty (LADMAP) is employed to solve the resulting model more effectively. The proposed method outperforms various two-step denoising and AD algorithms through simulated and real-data experiments, which also include our homemade GF5-MouMan dataset.

The rest of this article is organized as follows. Section II elaborates on the motivation and study of the integration denoising and AD. Then, the experimental results and analysis are presented in Section III. Finally, Section IV concludes this article.

II. PROPOSED METHOD

A. Motivation

Initially, we investigate the impact of additive mixed noise on anomaly locations and explore different AD approaches, including classic global RX (GRX) [14], LRASR [32], and LSMAD [31]. Additive Gaussian noise is first introduced into a benchmark AIR-HI dataset. Each band of this HS image is corrupted by the salt-and-pepper noise at a small proportion of 0.01. The detected outcomes from three AD approaches under

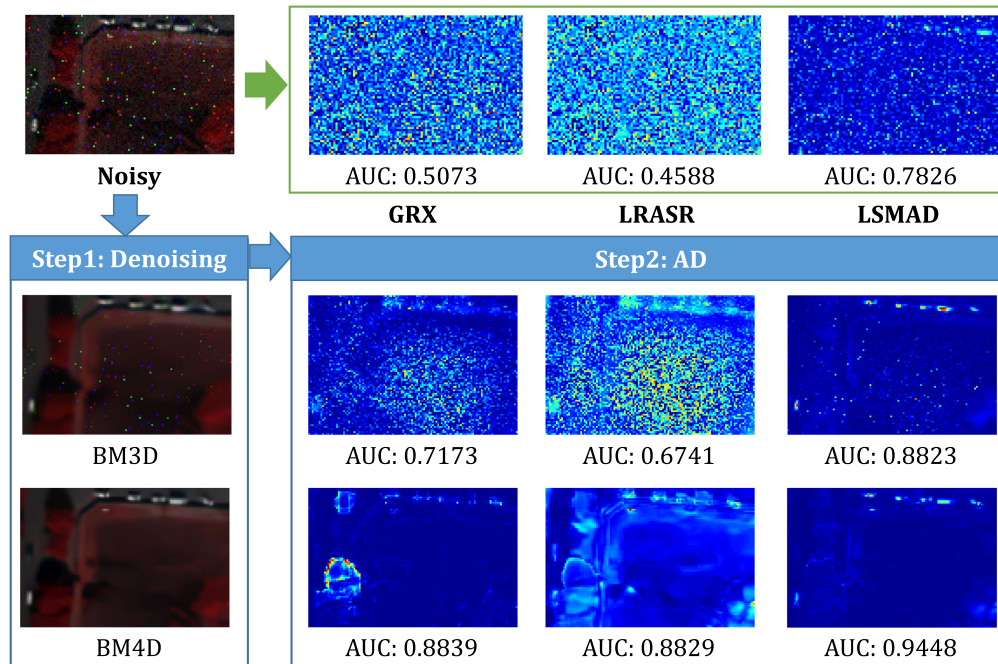


Fig. 1. Detection performances by GRX, LRASR, and LSMAD with an additive noise or denoising cases.

the additive noise case are visually compared in the first row of Fig. 1, revealing how additive noise affects different AD techniques. GRX, as a Mahalanobis distance-based method, assumes a multivariate Gaussian distribution for the background and fails to detect anomalous objects. LSMAD, which models the observed data corrupted by Gaussian noise, yields better results with less noise than GRX and LRASR. The undesirable influence of noise on AD is hardly ignored since the original geometrical structures of the HS image background and anomalies are contaminated by noise, ineluctably worsening the postprocessing interpretation task.

Traditionally, denoising and AD are conducted separately in the real HS image processing system. The second and third rows of Fig. 1 depict the conventional two-stage HS denoising and AD processing flowchart, where noise is first removed using different algorithms. However, this approach has its drawbacks, as seen with the residual sparse noise after denoising with BM3D, which inevitably increases FARs of AD detectors. On the other hand, denoising with BM4D removes most of the noise but also leads to loss of image detail information and a decrease in edge detection rates of anomaly objects. This highlights the importance of selecting appropriate denoisers and detectors based on the specific scene and noisy conditions in a two-step scheme.

To address these limitations, we propose an integrated framework that couples denoising and AD procedures. In this framework, denoising and AD mutually constrain each other in each iteration. AD can benefit from noise removal, while denoising can be assisted by intermediate AD results. Specifically, in the AD scheme, the HS image is decomposed into background and anomalies, with anomalies having distinct spectral signatures and statistical distributions from the surrounding background.

The identification of anomalies provides positive feedback to enhance spatial and spectral smoothness during denoising. This joint framework allows for the derivation of advanced models with effective and rational constraints for denoising and AD.

B. Simultaneous AD and Denoising Model

Fig. 2 presents the schematic diagram of the proposed HyADD model. The input is a noisy HS image with anomalies. The spatial-spectral TV (SSTV) of an HS tensor $\mathcal{X} \in \mathbb{R}^{m \times n \times z}$ is developed to remove additive noise and preserve the smoothness in the spatial and spectral domains. The denoised result of each iteration provides beneficial support for the AD procedure. Assisted by the specific antinoise dictionary, LRR can explore the multiple underlying subspaces and detect the outliers as well. The exploration of subspace domain-based low-rankness and the identification of outliers offer effective feedback to the denoising process. Finally, the denoised result and anomaly locations are produced together for further analysis and interpretation of ground objects.

When HS images are often contaminated by noise, the simultaneous AD and denoising task in third unfolding forms can be formulated as follows:

$$\begin{cases} \mathbf{Y} = \mathbf{X} + \mathbf{N} \\ \mathbf{X} = \mathbf{D}\mathbf{A} + \mathbf{S} \end{cases} \quad (1)$$

where \mathbf{Y} , \mathbf{X} , and \mathbf{N} , $\mathbf{S} \in \mathbb{R}^{z \times mn}$, $\mathbf{D} \in \mathbb{R}^{z \times l}$, and $\mathbf{A} \in \mathbb{R}^{l \times mn}$ are the noisy HS image, the latent clean image, additive noise, sparse part, the background dictionary, and representation coefficients, respectively. The key to AD under additive noise is to estimate the background and anomalies with the elimination

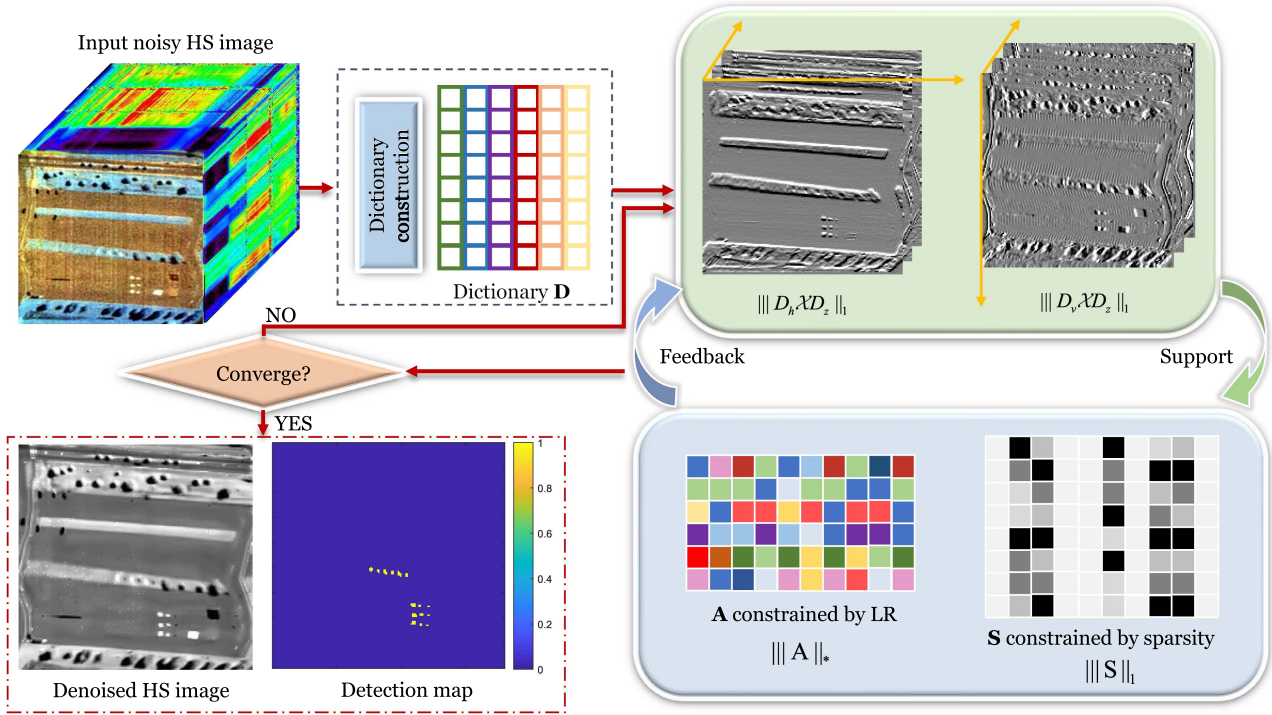


Fig. 2. Schematic diagram of the proposed HyADD integration model.

of additive noise. Differ from the LRSMD models [31], [32], [34] ignoring the independent fidelity term of noise removal, the proposed joint method is solved by minimizing an objective function as follows:

$$\min_{\mathbf{X}, \mathbf{A}, \mathbf{S}} \frac{1}{2} \|\mathbf{Y} - \mathbf{X}\|_F^2 + \frac{1}{2} \|\mathbf{X} - \mathbf{DA} - \mathbf{S}\|_F^2. \quad (2)$$

In general, the distribution of ground objects varies gently in one scene. Pixels of most HS images often have similar spatial characteristics to their neighboring pixels. High correlations also exist between adjacent spectral bands. Therefore, relatively smoothing characteristics of HS images usually occur in the spatial and spectral domains. A TV method originated from Rudin et al.'s [39] work to remove the noise of gray-level images due to the ability to preserve edge information and enforce piecewise smoothness. Inspired by this implementation of gray-level images, the 2-D spatial TV norm is first used to remove Gaussian noise of HS images in a band-by-band manner [40]. Some AD studies also adapted the band-by-band TV model to constrain the spatial gradients [34], [41], [42], unfortunately ignoring the correlation of nearby bands. To remove additive noise and promote the spatial-spectral piecewise smoothness, we incorporate an SSTV regularization [43] of $\mathcal{X} \in \mathbb{R}^{m \times n \times z}$ into (2)

$$\min_{\mathbf{X}, \mathbf{A}, \mathbf{S}} \frac{1}{2} \|\mathbf{Y} - \mathbf{X}\|_F^2 + \frac{1}{2} \|\mathbf{X} - \mathbf{DA} - \mathbf{S}\|_F^2 + \mu \|\mathcal{X}\|_{\text{SSTV}} \quad (3)$$

where $\|\mathcal{X}\|_{\text{SSTV}} = \|D_h \mathbf{X} D_z\|_1 + \|D_v \mathbf{X} D_z\|_1$ with two differential operators D_h and D_v along the horizontal and vertical

directions and one 1-D finite differencing D_z on the spectral signature of each pixel.

In HS imagery, each pure endmember mainly corresponds to one material. However, many additive pixels exist in one image due to the mixture caused by heterogeneity and stratified distribution of objects. Based on the LRR theory for the representation of additive pixels, the underlying HS data \mathbf{X} tends to be drawn from a union of subspaces. The goal of this representation is to identify the LR linear representation of \mathbf{X} by a dictionary \mathbf{D} . The optimization problem (2) can be rewritten as

$$\min_{\mathbf{X}, \mathbf{A}, \mathbf{S}} \frac{1}{2} \|\mathbf{Y} - \mathbf{X}\|_F^2 + \frac{1}{2} \|\mathbf{X} - \mathbf{DA} - \mathbf{S}\|_F^2 + \mu \|\mathcal{X}\|_{\text{SSTV}} + \lambda \|\mathbf{S}\|_1 + \alpha \|\mathbf{A}\|_*. \quad (4)$$

In (4), the AD and noise removal fidelity terms are constrained individually. The sparse term \mathbf{S} is utilized to identify anomaly objects within the background of \mathbf{X} . To further ensure the separation of noise from anomalies, an SSTV constraint is applied to \mathbf{X} . This constraint is particularly effective in isolating sparse noise from \mathbf{X} , leveraging both spatial and spectral information to enhance the separation process. This approach is advantageous because images contaminated with sparse noise rarely satisfy the gradient sparsity condition, which the SSTV constraint addresses effectively. In addition, choosing an appropriate dictionary is vital to facilitate the LRR of the background and additive noise removal. A new strategy for dictionary construction: A median filter is first applied to \mathbf{X} to get rid of the impact of sparse noise on the background. After filtering, pixels are divided into k clusters by k -means with the Euclidean distance. The

basic RX algorithm is used to find the more likely background pixels, whose detection results are lower than the others in each cluster. Anomalies occur in the same areas of each band and appear dense in the spectral domain. Thus, the $l_{2,1}$ norm often constrains this spectral–spatial sparsity of anomalies. However, sparse noise, such as salt-and-pepper noise and dead pixels, tends to exhibit stronger group sparsity compared to anomaly objects. This characteristic makes the $l_{2,1}$ norm more sensitive to sparse noise, as it is designed to handle group sparsity. In contrast, the l_1 norm is more robust in such scenarios because it focuses on individual element sparsity, reducing the influence of sparse noise and dead pixels. The specific experimental comparison between HyADD with the l_1 norm and the $l_{2,1}$ norm is provided in Section III-A5.

An antinoise dictionary construction for HyADD: As a vital role, an antinoise dictionary \mathbf{D} should be designed before the iteration loop. The simplest way to construct the unknown dictionary is to directly employ the original data \mathbf{Y} . However, the naturally large size of \mathbf{Y} will bring a heavy computation burden due to the need for the singular value decomposition operation. Another efficient approach is to select some pixels to constitute the dictionary in a certain manner [32]. These pixels should correspond to the material with more representation samples, whose probability of belonging to the background material will be higher. Meanwhile, in the mixed noise case, salt-and-pepper noise arises with a point distribution when some of the imaging sensors are saturated or fail to sample [44], [45]. The process starts with a traditional median filter method to weaken the negative influence of these sparse noises on the performance of the background dictionary. The filtered image \mathbf{Y}_{med} can be expressed as

$$\mathbf{Y}_{\text{med}} = \text{med}(\mathbf{Y}) \quad (5)$$

where med represents a median filtering directly applied to \mathbf{Y} band by band. All pixels are divided by a K-means algorithm into C clusters $\mathbf{Y}_{\text{med}} = \{\mathbf{y}_{\text{med},1}, \mathbf{y}_{\text{med},2}, \dots, \mathbf{y}_{\text{med},W}\}$, where W should be larger than the number of material classes. After that, the classical RX detector is assigned to obtain the Mahalanobis distance of each test pixel and choose the background pixels with the higher probability as the antinoise dictionary \mathbf{D} .

C. Optimization Procedure

Based on the alternating direction method, the existing LRR optimization often needs to introduce two auxiliary variables [26]. The matrix–matrix multiplications and matrix inversions of this solver cause $O(n^3)$ computation complexity and further a slow convergence, where $n \times n$ is the matrix size. To overcome this defect, Lin et al. [46] proposed LADMAP with $O(rn^2)$ complexity of solving LRR, where r is the rank of \mathbf{A} . LADMAP is an effective approach to solve (4) [47]. By

introducing auxiliary variables \mathbf{P} and \mathbf{Q} , (4) is reformulated as

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{A}, \mathbf{S}} \quad & \frac{1}{2} \|\mathbf{Y} - \mathbf{X}\|_F^2 + \frac{1}{2} \|\mathbf{X} - \mathbf{DA} - \mathbf{S}\|_F^2 + \mu(\|\mathcal{P}\|_1 \\ & + \|\mathcal{Q}\|_1) + \lambda \|\mathbf{S}\|_1 + \alpha \|\mathbf{A}\|_* \\ \text{s.t.} \quad & \mathcal{P} = D_h \mathcal{X} D_z, \mathcal{Q} = D_v \mathcal{X} D_z. \end{aligned} \quad (6)$$

Equation (6) is efficiently solved by LADMAP, and the corresponding unconstrained augmented Lagrangian function is expressed as follows:

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{A}, \mathbf{S}} \quad & \frac{1}{2} \|\mathbf{Y} - \mathbf{X}\|_F^2 + \frac{1}{2} \|\mathbf{X} - \mathbf{DA} - \mathbf{S}\|_F^2 + \mu(\|\mathcal{P}\|_1 + \|\mathcal{Q}\|_1) \\ & + \lambda \|\mathbf{S}\|_1 + \alpha \|\mathbf{A}\|_* + \frac{\beta}{2} \|\mathcal{P} - D_h \mathcal{X} D_z - \mathcal{B}_1\|_F^2 \\ & + \frac{\beta}{2} \|\mathcal{Q} - D_v \mathcal{X} D_z - \mathcal{B}_2\|_F^2 \end{aligned} \quad (7)$$

where \mathcal{B}_1 and \mathcal{B}_2 are two Lagrangian multipliers. The optimization procedure of (7) is achieved by updating one variable while fixing the others. The global optimized solution is guaranteed since (7) is convex. Suppose that optimization is in the k th iteration, the specific subproblems are given as follows.

1) Update \mathcal{P} and \mathcal{Q} : Exact all their related items from (7)

$$\min_{\mathcal{P}} \mu \|\mathcal{P}\|_1 + \frac{\beta}{2} \|\mathcal{P} - D_h \mathcal{X} D_z - \mathcal{B}_1\|_F^2 \quad (8)$$

$$\min_{\mathcal{Q}} \mu \|\mathcal{Q}\|_1 + \frac{\beta}{2} \|\mathcal{Q} - D_v \mathcal{X} D_z - \mathcal{B}_2\|_F^2. \quad (9)$$

The operator $\Delta_\varepsilon[x] = \text{sgn}(x) * \max(|x| - \varepsilon, 0)$ is used to obtain the solutions

$$\mathcal{P}_{k+1} = \Delta_{\mu/\beta}[D_h \mathcal{X}_k D_z + \mathcal{B}_{1,k}] \quad (10)$$

$$\mathcal{Q}_{k+1} = \Delta_{\mu/\beta}[D_v \mathcal{X}_k D_z + \mathcal{B}_{2,k}]. \quad (11)$$

2) Update \mathbf{A} : The subproblem of \mathbf{A} is equivalent to the following constrained minimization problem:

$$\min_{\mathbf{A}} \alpha \|\mathbf{A}\|_* + \frac{1}{2} \|\mathbf{X} - \mathbf{DA} - \mathbf{S}\|_F^2. \quad (12)$$

By linearizing the quadratic term f in (12) at previous iteration \mathbf{A}_k and adding a proximal term, the iteration of \mathbf{A} goes as follows:

$$\begin{aligned} \mathbf{A}_{k+1} &= \underset{\mathbf{A}}{\text{argmin}} \alpha \|\mathbf{A}\|_* + \langle \nabla_{\mathbf{A}} f(\mathbf{A}_k, \mathbf{X}_k, \mathbf{S}_k, \beta_k), \mathbf{A} - \mathbf{A}_k \rangle \\ &+ \frac{\gamma}{2} \|\mathbf{A} - \mathbf{A}_k\|_F^2 \\ &= \underset{\mathbf{A}}{\text{argmin}} \alpha \|\mathbf{A}\|_* \\ &+ \frac{\gamma}{2} \|\mathbf{A} - \mathbf{A}_k + [-\mathbf{D}^T(\mathbf{X}_k - \mathbf{DA}_k - \mathbf{S}_k)]/\gamma\|_F^2 \end{aligned} \quad (13)$$

where the quadratic term $\nabla_{\mathbf{A}} f$ is the partial differential of f with respect to \mathbf{A} and $\gamma = \|\mathbf{D}\|_F^2$.

3) Update \mathbf{S} : The \mathbf{S} -related subproblem is rewritten as

$$\min_{\mathbf{S}} \lambda \|\mathbf{S}\|_1 + \frac{1}{2} \|\mathbf{X} - \mathbf{DA} - \mathbf{S}\|_F^2. \quad (14)$$

Algorithm 1: HyADD Algorithm.

Input: Input matrix with noise $\mathbf{Y}_0 = \mathbf{Y}$, parameter $\beta, \mu, \lambda, \alpha, K, \beta_{max} = 10^{10}, \rho_0 = 1.1$,

- 1: **Ditionaryconstruction** ::
 - 1.1 a median filter is applied to \mathbf{Y} : $\mathbf{Y}_{med} = \text{med}(\mathbf{Y})$,
 - 1.2 \mathbf{Y}_{med} is divided into W clusters using K-means,
 - 1.3 the background pixels are selected by the RX algorithm,
- 2: **while** $0 \leq k \leq K$ or meet the convergence conditions **do**
- 3: update \mathcal{P}_{k+1} and \mathcal{Q}_{k+1} by solving (10) and (11), respectively:

$$\mathcal{P}_{k+1} = \Delta_{\mu/\beta}[D_h \mathcal{X}_k D_z + \mathcal{B}_{1,k}],$$

$$\mathcal{Q}_{k+1} = \Delta_{\mu/\beta}[D_v \mathcal{X}_k D_z + \mathcal{B}_{2,k}],$$
- 4: update \mathbf{A}_{k+1} by solving (13):

$$\mathbf{A}_{k+1} = \underset{\mathbf{A}}{\text{argmin}} \alpha \|\mathbf{A}\|_* + \frac{\gamma}{2} \|\mathbf{A} - \mathbf{A}_k + [-\mathbf{D}^T(\mathbf{X}_k - \mathbf{D}\mathbf{A}_k - \mathbf{S}_k)]/\gamma\|_F^2,$$
- 5: update \mathbf{S}_{k+1} by solving (15):

$$\mathbf{S}_{k+1} = \Delta_\lambda[\mathbf{X}_k - \mathbf{D}\mathbf{A}_{k+1}],$$
- 6: update \mathbf{x}_{k+1} by solving (17) :
- 7: $(\beta D_1^T D_1 + \beta D_2^T D_2 + 2\mathbf{I})\mathbf{x}_{k+1} = \mathbf{y}_k + \text{vec}(\mathbf{D}\mathbf{A}_{k+1}) + \mathbf{s}_{k+1} + \beta D_1^T(\mathbf{p}_{k+1} - \mathbf{b}_{1,k+1}) + \beta D_2^T(\mathbf{q} - \mathbf{b}_{2,k+1})$,
- 8: update $\mathbf{X}_{k+1} = \text{rearrange}(\mathbf{x}_{k+1})$,
- 9: update the multipliers $\mathbf{B}_{1,k+1}$ and $\mathbf{B}_{2,k+1}$:

$$\mathbf{B}_{1,k+1} = \mathbf{B}_{1,k} + D_1 \mathbf{X}_{k+1} - \mathbf{P}_{k+1},$$

$$\mathbf{B}_{2,k+1} = \mathbf{B}_{2,k} + D_2 \mathbf{X}_{k+1} - \mathbf{Q}_{k+1},$$
- 10: Update β as: $\beta = \min(\beta_{max}, \rho\beta)$, where
- 11: $\rho =$

$$\begin{cases} \rho_0, & \text{if } \beta \max(\sqrt{\eta} \|\mathbf{L}_{t+1} - \mathbf{L}_t\|_F, \|\mathbf{S}_{t+1} - \mathbf{S}_t\|_F, \\ & \|\mathbf{X}_{t+1} - \mathbf{X}_t\|_F) / \|\mathbf{X}\|_F \leq \varepsilon_2 \\ 1, & \text{otherwise,} \end{cases}$$
- 12: **end while**
- 13: $\mathbf{Y}_{k+1} = \Theta \mathbf{X}_{k+1} + (1 - \Theta) \mathbf{Y}$,
- 14: $k = k + 1$,

Output: AD map and denoised result \mathbf{X} .

The solution of (14) is obtained as

$$\mathbf{S}_{k+1} = \Delta_\lambda[\mathbf{X}_k - \mathbf{D}\mathbf{A}_{k+1}]. \quad (15)$$

4) Update \mathbf{X} by minimizing the vector form of \mathbf{X} -related subproblem as follows:

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_F^2 + \frac{1}{2} \|\mathbf{x} - \text{vec}(\mathbf{D}\mathbf{A}) - \mathbf{s}\|_F^2 + \frac{\beta}{2} \|\mathbf{p} - D_1 \mathbf{x} - \mathbf{b}_1\|_F^2 + \frac{\beta}{2} \|\mathbf{q} - D_2 \mathbf{x} - \mathbf{b}_2\|_F^2, \quad (16)$$

where $D_1 = D_h \otimes D_z$, $D_2 = D_v \otimes D_z$, and \otimes denotes the Kronecker product. By setting the derivative of (16) with respect to \mathbf{X} to zero, we obtain the following equation:

$$(\beta D_1^T D_1 + \beta D_2^T D_2 + 2\mathbf{I})\mathbf{x}_{k+1} = \mathbf{y}_k + \text{vec}(\mathbf{D}\mathbf{A}_{k+1}) + \mathbf{s}_{k+1} + \beta D_1^T(\mathbf{p}_{k+1} - \mathbf{b}_{1,k+1}) + \beta D_2^T(\mathbf{q} - \mathbf{b}_{2,k+1}). \quad (17)$$

The algorithm for sparse linear equations and least squares (LSQR) [48] can be used to solve (17). The solution \mathbf{x}_{k+1} is rearranged into the recovered result \mathbf{X}_{k+1} . Due to AD being regarded as the final goal of HyADD, we pay more attention to variables related to AD and define the convergence condition as follows:

$$\max \left\{ \begin{array}{l} \|\mathbf{S}_k - \mathbf{S}_{k-1}\|_\infty, \\ \|\mathbf{A}_k - \mathbf{A}_{k-1}\|_\infty, \\ \|\mathbf{X}_k - \mathbf{X}_{k-1}\|_\infty, \\ \|\mathbf{X}_k - \mathbf{D} * \mathbf{A}_k - \mathbf{S}_k\|_\infty \end{array} \right\} \leq \epsilon. \quad (18)$$

The observation \mathbf{Y} in each iteration with Θ ($0 \leq \Theta \leq 1$) is updated as

$$\mathbf{Y}_{k+1} = \Theta \mathbf{X} + (1 - \Theta) \mathbf{Y} \quad (19)$$

which is used to boost the denoising performances [49], [50], [51].

The numerical iterations are outlined in Algorithm 1.

Computational complexity: Here, we briefly discuss the computational complexity of HyADD. The following parts are consisted of each step.

- 1) The computation of solving both (10) and (11) depends on the SSTV operator, which requires about $2(zmn)^2 + 2zmn$.
- 2) Different from the expensive computation of the conventional LRR solver [28], solving (13) with \mathbf{A} of size $l \times z$ costs $2lzm + 2zmn + 2lmn + l(mn)^2 + l^3$.
- 3) Updating \mathbf{S} requires about $zlmn + 2zmn$ in each iteration.
- 4) LSQR algorithm is used to update \mathbf{x} and its computation costs $2(zmn)^3 + zmn + zlmn + 2(zmn)^2$.

The overall computational complexity of Algorithm 1 in each iteration is $2(zmn)^3 + 7zmn + 2lmn + 4zlmn + 4(zmn)^2 + l(mn)^2 + l^3$. Usually, $z, l \ll mn$, the order of complexity for our HyADD is $O((zmn)^3)$.

Remark: Lin et al. [46] generalized ADMM by linearizing the quadratic penalty term and has proven the convergence of LADMAP for convex optimization with two blocks. Note that the separable and independent variable is treated as one block. However, as for the three-block convex function (4) of HyADD, the extended version of the ADMM is not necessarily convergent even with the strong convex assumption [52][Theorem 4.1]. Hence, we have systematically discussed the convergence performance of the HyADD framework. Section III-A3 gives a further illustration of the convergence analysis.

III. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, experiments on two simulated and one real-world HS data are carried out to verify the performance of HyADD for two tasks. First, the denoised performance of the proposed HyADD is analyzed with existing methods, such as BM3D [53], BM4D [54], and TDL [55]. Second, different HS AD approaches chosen for comparison consist of classic GRX [14], discrete wavelet transform RX (DRX) [56], RPCA [28], LRASR [32], LSMAD [31], and graph and TV regularized LRR (GTVLRR) [34].

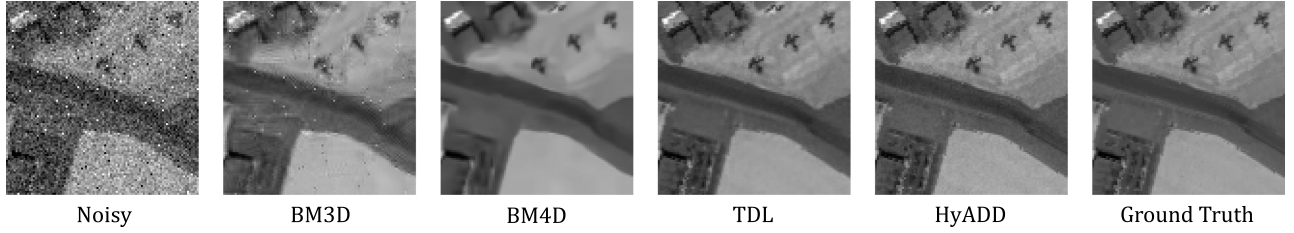


Fig. 3. Denoising results of all compared methods on the San Diego dataset under noise Case 2.

Except for the visual observation of denoised results on simulated experiments, we adopt two quality assessments including the mean peak signal to noise ratio (MPSNR) and mean structural similarity index (MSSIM). MPSNR and MSSIM are obtained by taking the averages of the PSNR and SSIM values between the ground truth and denoised ones, which depends on the denoising precision along the spatial domain

$$\text{MPSNR} = \frac{1}{z} \sum_{k=1}^z \text{PSNR}(\mathcal{X}(:, :, k), \bar{\mathcal{X}}(:, :, k)) \quad (20)$$

$$\text{MSSIM} = \frac{1}{z} \sum_{k=1}^z \text{SSIM}(\mathcal{X}(:, :, k), \bar{\mathcal{X}}(:, :, k)) \quad (21)$$

where $\mathcal{X}(:, :, k)$ and $\bar{\mathcal{X}}(:, :, k)$ denote the i th band of the reference and denoised images, and $\mathbf{X}_{(3)}(k, :)$ and $\bar{\mathbf{X}}_{(3)}(k, :)$ are the i th spectral signatures of two images.

Color detection maps are given in all experimental results and analysis. The responses of pixels are reflected by the colors ranging from blue to red. For qualitative comparisons, we introduce the receiver operating characteristic (ROC) curve [57] with pointwise confidence intervals, which reflects the compromise between the probability of detection (PD) and FAR. The area under the ROC curve is denoted as AUC [58] and its ideal value is 1.

A. Simulated HS Data Experiments

HS images acquired by sensors are often plagued by various noise sources, such as atmospheric haze and instrument noise, which significantly degrade image quality and impact subsequent HS applications. Thermally induced dark current, fluctuations in power supply and atmospheric absorption can introduce Gaussian noise [43], [59]. Salt-and-pepper noise typically occurs when some sensors become saturated or fail to sample properly [44], [45]. For HS scanners using whiskbroom technology (e.g., the airborne HyMap [60]) or pushbroom technology (e.g., the spaceborne EO-1 Hyperion [61]), stripes can appear due to miscalibration [62]. More critically, if entire rows or columns are missing due to physical damage to the scanner, the resulting HSIs will exhibit prominent deadline noise [63]. To simulate different noisy HS images, we add Gaussian noise, salt-and-pepper noise, deadlines, and stripes into two datasets, including San Diego and Urban. The details of noise cases are reported in Table I.

1) *Denoised and Detection Performances. San Diego:* The San Diego dataset was captured by the airborne visible/infrared

TABLE I
ADDED ADDITIVE NOISE IN FOUR CASES

Noise case	Additive noise
Case 1	Gaussian (Zero mean and noise variance: 0.05) +salt and pepper (0.01) in all bands
Case 2	Gaussian (Zero mean and noise variance: 0.10) +salt and pepper (0.01) in all bands
Case 3	Gaussian (Zero mean and noise variance: 0.10) +salt and pepper (0.01) in all bands + dead-lines in 91~130 bands (with the number of stripes randomly selected from 2 to 5 and the width of the dead-lines randomly generated from 1 to 2)
Case 4	Gaussian (Zero mean and noise variance: 0.10) +salt and pepper (0.01) in all bands + dead-lines in 91~130 bands (with the number of stripes randomly selected from 2 to 5 and the width of the dead-lines randomly generated from 1 to 2) + stripes in 161~170 bands (the number of stripes randomly selected from 3 to 5.)

imaging spectrometer with a spatial resolution of 3.5 m. The image size of the whole dataset is 400×400 . Herein, we select one widely used subimage, which has a size of $100 \times 100 \times 189$. In this scene, the background refers to classes of grass, roofs, roads, etc. On the contrary, three manual airplanes are regarded as anomalies.

Figs. 3 and 4 present the denoising and detection outcomes for the San Diego image under noise Case 2, respectively. Some residual sparse noise persists in the result denoised by BM3D, causing increased FARs of GRX, DRX, and RPCA. Conversely, LRASR, LSMAD, and GTVLRR demonstrate robustness against noise in the denoising results shown in the first row of Fig. 4. In Fig. 3, BM4D outperforms BM3D by eliminating all residual noise but exhibits an over-smoothing effect. Compared to BM3D, the FARs of the background objects are reduced in the second row of Fig. 4. However, the edges of the three flights appear somewhat blurred due to the oversmoothness during the denoising process of BM4D. TDL exhibits superior performance compared to BM3D and BM4D. GTVLRR, when subjected to denoising by TDL, achieves the cleanest detection map with the most prominent edges of airplanes among all compared algorithms.

In Fig. 5, the ROC curves compare HyADD with other competing approaches under noise Case 2. Among these, conventional LR-based methods consistently demonstrate higher PD

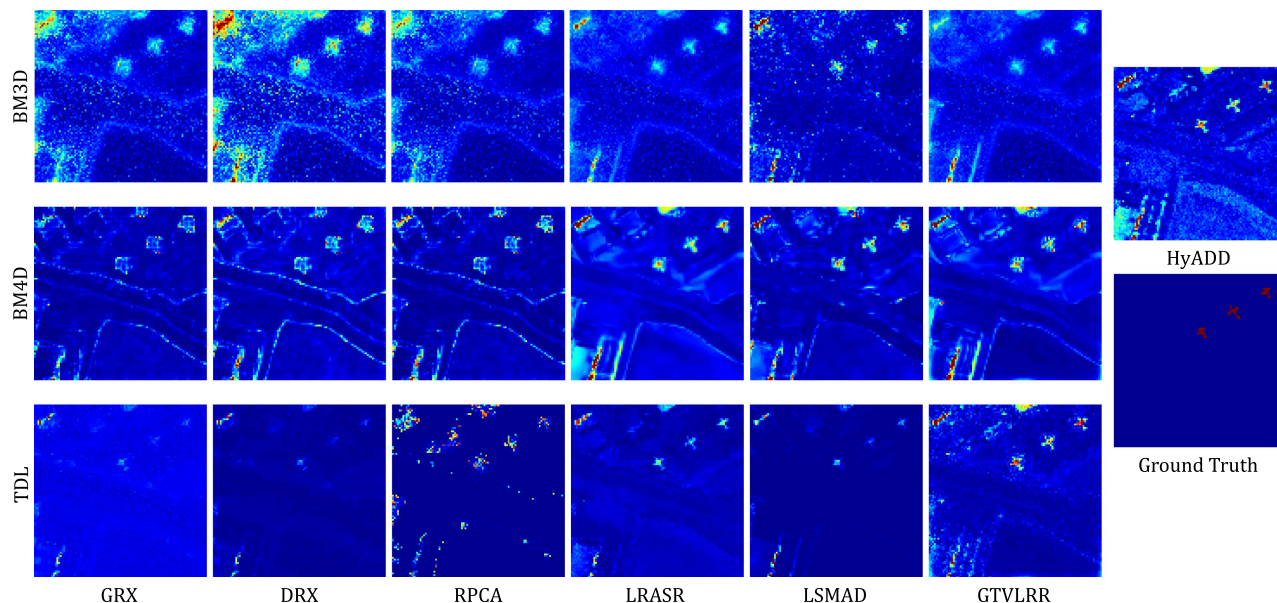


Fig. 4. Detection results of all compared methods on the San Diego dataset under noise Case 2.

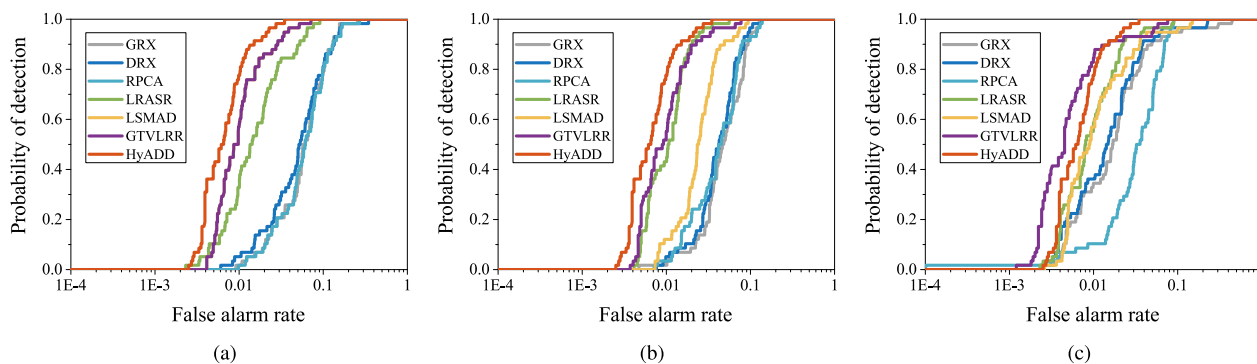


Fig. 5. ROC curves obtained by different denoising and AD methods for the San Diego dataset under noise Case 2. (a) BM3D. (b) BM4D. (c) TDL.

values compared to RX-based methods and RPCA, irrespective of the denoising technique employed. Following denoising with BM3D and BM4D, GRX, DRX, and RPCA exhibit similar curve trends. Notably, RPCA performs the poorest when encountering TDL among all compared approaches in Fig. 5(c). Before reaching a FAR of 0.01, the curve of GTVLRR surpasses that of HyADD. However, HyADD achieves a PD of 1.0 even at a FAR of 0.03, while most others remain below 0.9.

Table II lists the AUC values of different anomaly detectors after applying three denoisers under four noise cases. In addition, Table III presents the denoising performance of various denoisers under these noise scenarios. Notably, TDL achieves the second-highest MPSNR and MSSIM among all competitors. However, GTVLRR obtains higher AUC scores than the former methods with the support of BM4D under noise Case 4. It is essential to note that excellent denoising performance may not always guarantee favorable detection outcomes within the two-step scheme. Despite achieving a lower MPSNR value,

HyADD outperforms TDL in terms of AUC score, highlighting its effectiveness in the detection task under noise Case 1.

Urban: This dataset is obtained by the HYDICE airborne sensor with a 10-nm spectral solution and a 1-m spatial solution over an urban area in the USA. A subimage of 80×100 pixels with 175 spectral bands is used for evaluation. The scene consists of vast stretches of vegetation area, a construction area, and several roads, which are treated as the background. Anomalies refer to vehicles and roofs, whose precise positions are given in the ground truth of Fig. 7.

In Fig. 6(a), band 170 of the Urban dataset is depicted, showcasing various types of noise corruption. Each anomaly occupies only a few pixels, presenting a challenge for accurate detection. BM3D and BM4D fail to eliminate stripes and deadlines, resulting in high FAR in the results of GRX, DRX, RPCA, and LRASR under noise Case 2, as shown in Fig. 7. Upon applying TDL, RPCA exhibits a stronger response to outlines than the other detectors. Consistent with the visual

TABLE II
AUC VALUES OF ALL COMPARED ALGORITHMS FOR TWO DATASETS UNDER FOUR NOISE CASES

Data sets	Denoisers	Noise case	GRX [14]	DRX [56]	RPCA [28]	LRASR [32]	LSMAD [31]	GTVLRR [34]	HyADD
San Diego	BM3D	Case 1	0.9233	0.9291	0.9226	0.9837	0.9791	0.9886	0.9911 <u>0.9907</u> 0.9900 0.9892
		Case 2	0.9303	0.9340	0.9288	0.9804	0.9550	0.9834	
		Case 3	0.7981	0.8529	0.7909	0.8687	0.8559	0.9430	
		Case 4	0.7720	0.8295	0.7659	0.8583	0.8422	0.9306	
	BM4D	Case 1	0.9395	0.9467	0.9495	0.9852	0.9898	0.9903	
		Case 2	0.9455	0.9624	0.9503	0.9839	0.9717	0.9874	
		Case 3	0.9265	0.9568	0.9525	0.9714	0.9709	0.9814	
		Case 4	0.9117	0.9410	0.9401	0.9497	0.9448	<u>0.9475</u>	
	TDL	Case 1	0.9719	0.9062	0.9296	0.9903	0.9896	<u>0.9906</u>	
		Case 2	0.9669	0.9753	0.9606	0.9859	0.9833	0.9912	
		Case 3	0.9002	0.8795	0.9082	0.9567	0.9593	<u>0.9737</u>	
		Case 4	0.8870	0.8745	0.8572	0.8897	0.9013	0.9323	
Urban	BM3D	Case 1	0.7776	0.7749	0.7900	0.8594	0.7846	0.9026	0.9852 0.9756 0.9748 0.9637
		Case 2	0.8103	0.8337	0.8202	0.8640	0.7396	0.8930	
		Case 3	0.7018	0.7149	0.7061	0.7387	0.7037	0.7876	
		Case 4	0.7267	0.7302	0.7345	0.7330	0.7672	0.8274	
	BM4D	Case 1	0.8963	0.9379	0.9392	0.9185	0.9405	0.9583	
		Case 2	0.8521	0.8988	0.9041	0.9239	0.9292	<u>0.9492</u>	
		Case 3	0.8075	0.8818	0.8578	0.9016	0.8727	0.9132	
		Case 4	0.7757	0.8360	0.8364	0.8753	<u>0.9017</u>	0.8901	
	TDL	Case 1	0.9410	0.9348	<u>0.9637</u>	0.9554	0.9577	0.9602	
		Case 2	0.9190	0.8979	0.8711	0.8856	0.8514	0.8968	
		Case 3	0.8937	0.8856	<u>0.9332</u>	0.8539	0.8568	0.8879	
		Case 4	0.7420	0.8597	0.8982	0.8472	0.8582	0.8914	

The best results are indicated in bold, and the second ones are underlined.

TABLE III
DENOISING PERFORMANCES OF DIFFERENT METHODS FOR TWO DATASETS UNDER FOUR NOISE CASES

HSI datasets	Noise case	Index	Noisy	BM3D [53]	BM4D [54]	TDL [55]	HyADD
San Diego	Case 1	MPSNR	22.73dB	29.03dB	32.92dB	38.54dB	<u>38.02dB</u>
		MSSIM	0.5077	0.7974	0.8912	0.9602	<u>0.9547</u>
	Case 2	MPSNR	19.00dB	28.28dB	31.88dB	35.54dB	<u>35.15dB</u>
		MSSIM	0.2995	0.7802	0.8698	<u>0.9205</u>	0.9334
	Case 3	MPSNR	18.53dB	25.36dB	30.68dB	<u>32.25dB</u>	33.91dB
		MSSIM	0.2919	0.6667	0.8395	<u>0.8720</u>	0.9233
	Case 4	MPSNR	18.48dB	24.51dB	30.61dB	<u>31.97dB</u>	33.80dB
		MSSIM	0.2918	0.6341	0.8611	<u>0.8706</u>	0.9114
Urban	Case 1	MPSNR	22.48dB	26.96dB	33.38dB	36.12dB	<u>35.16dB</u>
		MSSIM	0.6439	0.7601	0.9232	0.9603	<u>0.9466</u>
	Case 2	MPSNR	19.02dB	26.15dB	31.31dB	<u>33.67dB</u>	35.54dB
		MSSIM	0.4450	0.7353	0.8961	0.9447	<u>0.9209</u>
	Case 3	MPSNR	18.67dB	24.81dB	29.53dB	<u>31.43dB</u>	34.18dB
		MSSIM	0.4349	0.7209	0.8755	<u>0.9161</u>	0.9165
	Case 4	MPSNR	18.64dB	24.75dB	29.47dB	<u>31.23dB</u>	33.90dB
		MSSIM	0.4344	0.7213	0.8737	<u>0.9107</u>	0.9148

The best results are indicated in bold, and the second ones are underlined.

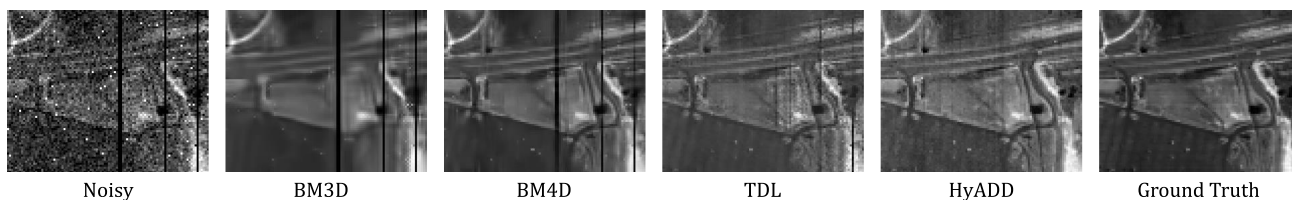


Fig. 6. Denoising results of all compared methods on the Urban dataset under noise Case 2.

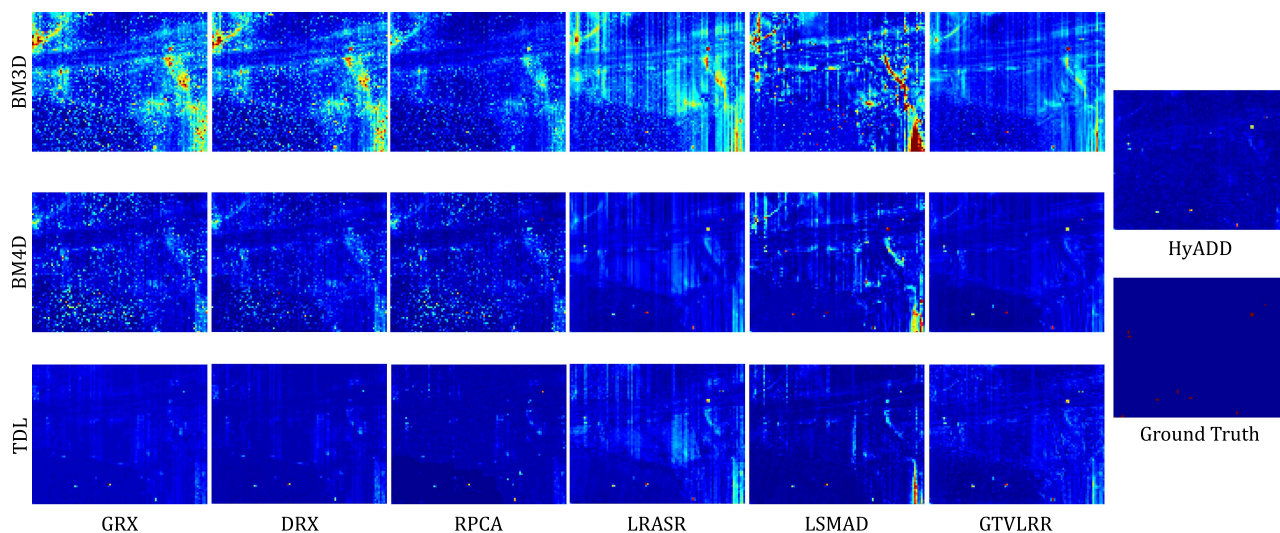


Fig. 7. Detection results of all compared methods on the Urban dataset under noise Case 2.

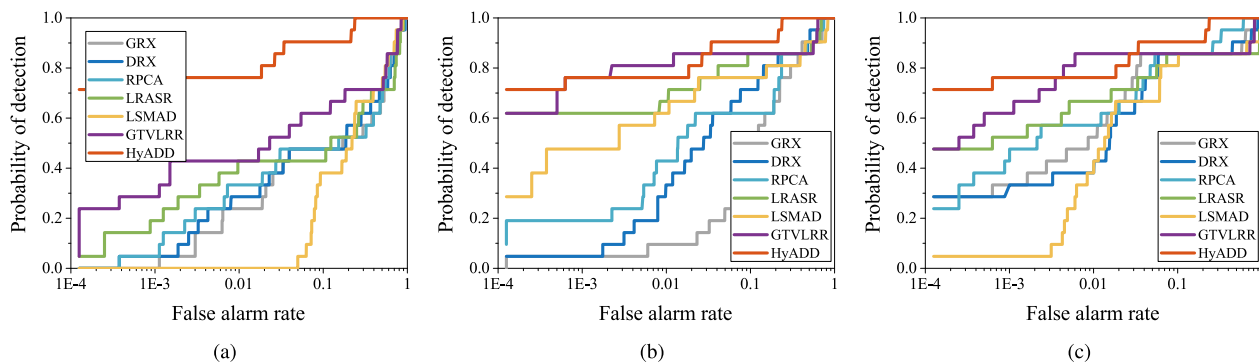


Fig. 8. ROC curves obtained by different denoising and AD methods for the Urban dataset under noise Case 2. (a) BM3D. (b) BM4D. (c) TDL.

observations, RPCA obtains the second-highest AUC in Table II under noise Cases 1 and 3. Therefore, different denoisers and detectors should be selected for different scenes and noisy conditions for the two-scheme step. Remarkably, our HyADD achieves appealing denoising and detection performance. Fig. 8 reports the ROC curves, which also validate the effectiveness of the proposed algorithm. Especially, the PD value of HyADD is always the highest and reaches 1.0 fastest among all ROC curves.

In summary, while achieving superior denoising performance is a critical aspect, it may not consistently translate into favorable detection outcomes within a two-step framework. Despite

BM4D achieving a lower MPSNR value than TDL, GTVLRR after denoising by BM4D obtains the highest AUC score on the Urban dataset under noise Case 2 among the two-step schemes. The selection of appropriate denoisers and detectors is paramount, contingent upon the specific scene characteristics and prevailing noise conditions within a two-step configuration. Notably, our HyADD exhibits commendable denoising and detection efficacy, thus highlighting its proficiency in mitigating noise-related challenges across diverse noise scenarios.

2) *Parameter Analysis*: In the proposed HyADD solver, four parameters β , α , μ , and λ need to be determined. All parameter discussion is conducted for the San Diego dataset under noise

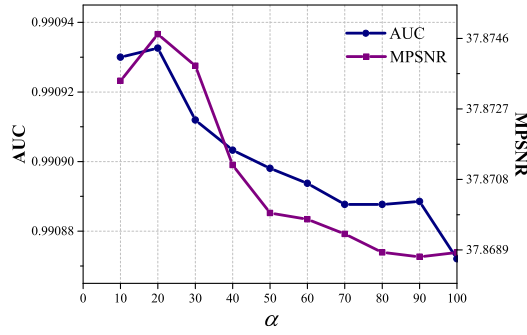


Fig. 9. AUC and MPSNR value curves with respect to α for the San Diego dataset.

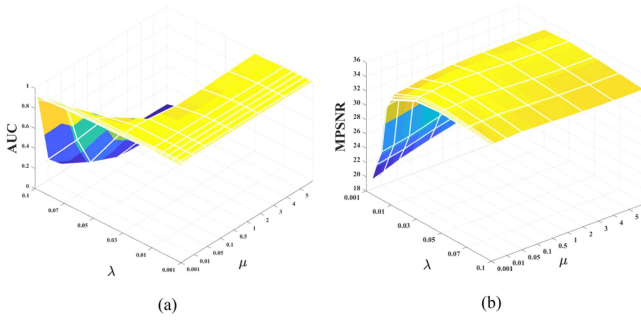


Fig. 10. Parameter sensitivity analysis in terms of λ and μ for the San Diego dataset. (a) AUC values. (b) MPSNR values.

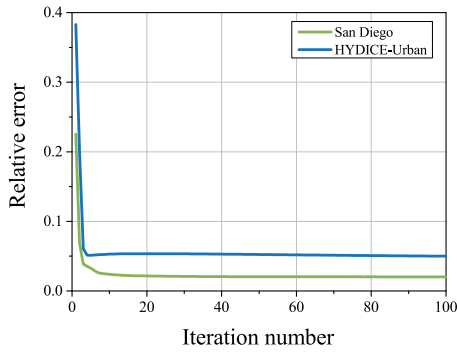


Fig. 11. Relative errors versus iterations on three datasets.

Case 2. Penalty parameter β is just related to the convergence speed in ADMM iterations [47], [64]. β can be selected from the range of [1,100] with an interval of 10. Parameter α controls the strength of the LR property for the background. Fig. 9 shows the sensitivity of two parameters to AUC and MPSNR values. Two curves have the same changing tendency with α ranging from 10 to 100.

Parameter λ plays an important role in controlling the strength of the sparsity term. Parameter μ restricts the impact of SSTV regularization on denoising effectiveness. To investigate the AD and denoising performance gains, these parameters λ and μ are simultaneously taken into consideration in Fig. 10. λ lies in the range of [0.001, 0.003, 0.005, 0.01, 0.03, 0.05, 0.07, 0.1, 0.2, 0.3] and μ is chosen from [0.001, 0.01, 0.05, 0.1, 0.5, 1, 2,

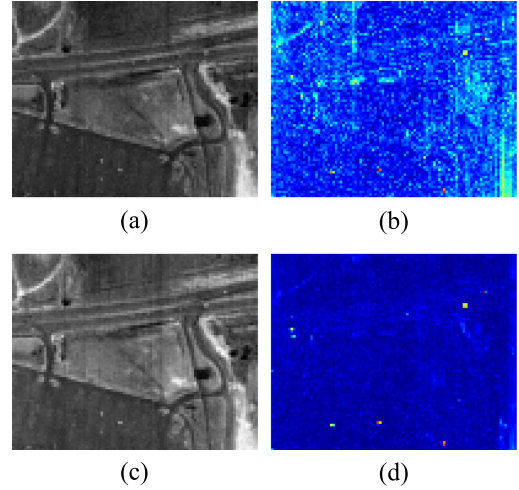


Fig. 12. Experimental results between HyADD with the $L_{2,1}$ norm and the L_1 norm. $L_{2,1}$. (a) Denoised result (PSNR: 30.55 dB). (b) Anomaly map (AUC: 0.9429); L_1 . (c) Denoised result (PSNR: **34.18 dB**). (d) Anomaly map (AUC: **0.9748**).

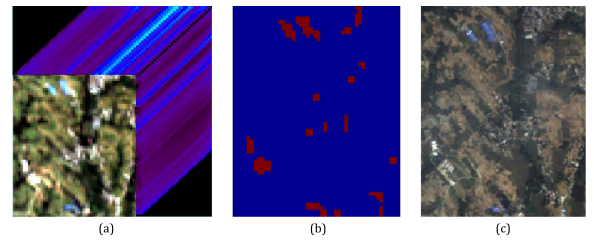


Fig. 13. GF5-MouMan dataset. (a) HS image. (b) Reference map. (c) Responding multispectral image.

3, 4, 5]. When λ increases from 0.001 to 0.05 and μ is varied among [0.1, 0.2, 0.3], the AUC and MPSNR curves of the two figs show upward trends. After $\lambda = 0.2$, the AD performances fall off gradually and the MPSNR curves stay almost stable. In summary, the parameter tuning of HyADD is reported as follows.

- 1) Choose β as an integer from the scope of [1,100] with 10 as step size.
- 2) The range of α from 10 to 20.
- 3) λ and μ is given within [0.05, 0.2] and [0.1, 0.3], respectively.

3) *Convergence Analysis*: We investigate the numerical convergence on three datasets. Fig. 11 describes the relative errors of (18) versus iterations. With the increasing of iteration numbers, the relative errors monotonically decrease first and then keep stable, which numerically proves the good convergence of HyADD.

4) *Running Time*: Table IV reported the running times of different approaches. All of the experiments are conducted in Matlab 2018a on the same PC with an Intel i9 CPU at 3.00 GHz and 128 GB of memory. GRX is the fastest among the state-of-the-art methods, but it rarely shows a satisfactory detection performance under mixed noise. DWT RX has better computing times and detection results than RPCA. The running times of

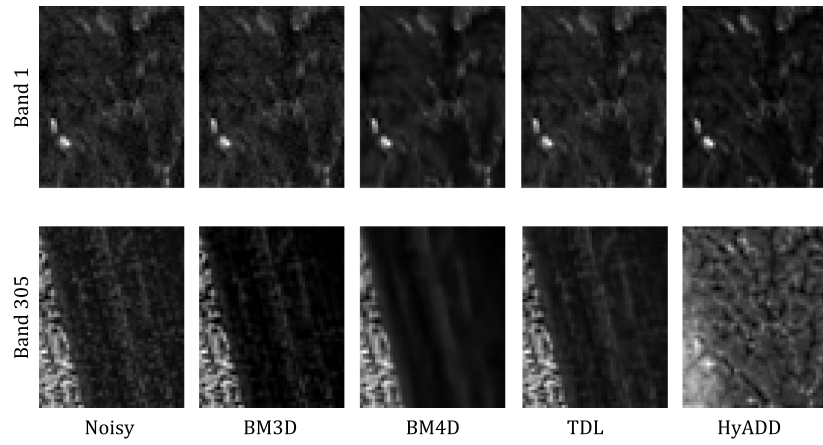


Fig. 14. Denoising results of all compared methods on the GF5-MouMan dataset.

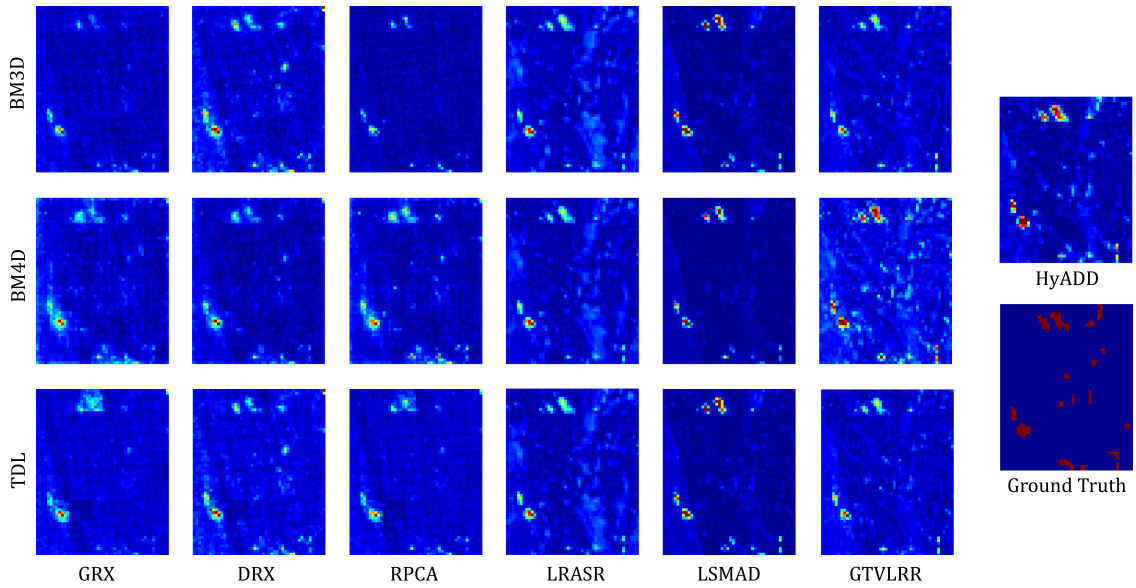


Fig. 15. Detection results of all compared methods on the GF5-MouMan dataset.

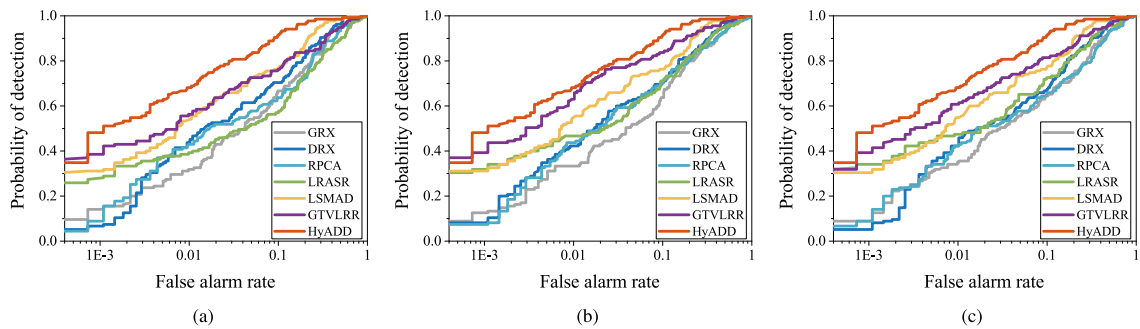


Fig. 16. ROC curves obtained by different denoising and AD methods for the GF5-MouMan dataset. (a) BM3D. (b) BM4D. (c) TDL.

TABLE IV
RUNNING TIME COMPARISON OF DIFFERENT METHODS ON DIFFERENT DATASETS

HSI data sets	GRX	DWT RX	RPCA	LRASR	LSMAD	GTVLRR	HyADD
San Diego	0.08 s	4.75 s	2.23 s	17.93 s	6.57 s	277.03 s	520.65 s
HYDICE-Urban	0.05 s	4.92 s	2.62 s	17.73 s	7.04 s	255.59 s	510.02 s

TABLE V
AUC VALUES OF ALL COMPARED ALGORITHMS FOR THE REAL-WORLD GF5-MOUMAN DATASET

Denoisers	GRX	DRX	RPCA	LRASR	LSMAD	GTVLRR	HyADD
BM3D	0.8799	0.9052	0.8705	0.8484	0.9362	0.9160	0.9679
BM4D	0.8739	0.9019	0.8866	0.9041	0.9364	<u>0.9448</u>	
TDL	0.8659	0.8957	0.8621	0.9001	0.9370	0.9292	

The best results are indicated in bold, and the second ones are underlined.

GTVLRR and HyADD are more than LRASR and LSMAD due to the incorporation of the TV terms. In particular, joint AD and noise removal tasks are unified into one intelligent framework. The detected objects and denoised results are simultaneously obtained by the proposed HyADD.

5) *Comparison Between HyADD With the $L_{2,1}$ Norm and the L_1 Norm*: We experimentally compare HyADD with different sparsity constraints to illustrate the effectiveness of the L_1 norm for the noisy case. Fig. 12 shows the experimental results obtained by HyADD with the $L_{2,1}$ norm and the L_1 norm. Residual salt-and-pepper noise and deadlines are detected as anomalies by HyADD with the $L_{2,1}$ norm, resulting in lower PSNR and AUC values than the original HyADD. In noisy cases, sparse noise can significantly affect the performance of $L_{2,1}$ -based methods. The L_1 norm mitigates this issue by penalizing the absolute values of the coefficients, making it less influenced by extreme values caused by noise.

B. Real-World HS Data Experiments

The real-world GF5 data acquired by the GF5-Advanced HS imager on 31 July 2019 was used to verify the effectiveness of the proposed method, ranging from 390 to 2500 nm with a spatial resolution of 30 m. The popular environment for visualizing images is utilized to make pre-processing steps and select one subimage of $60 \times 48 \times 305$ for the experiment, as shown in Fig. 13(a). In this scene, the background includes a mountain, trees, and roads, while the manmade buildings refer to the anomaly objects. Therefore, this subimage is named the GF5-MouMan dataset. The reference detection map is displayed in Fig. 13(b), whose number of anomalies is 135, with a 4.69% proportion. These anomalies are manually labeled according to the high-resolution GF6 multispectral image [shown in Fig. 13(c)] and four common characteristics of anomalies [12], [65]: no prior knowledge; low probability of occurrence; small spatial size; and insignificance in spectral statistics.

The first column of Fig. 14 showcases two noisy HS bands, with the first corrupted by Gaussian noise and the second affected by both Gaussian noise and stripes. BM4D and TDL demonstrate superior performance over BM3D in terms of accurate Gaussian noise removal. However, BM4D exhibits more oversmoothness, leading to a loss of detailed information below the stripes. In contrast, HyADD effectively removes most of the Gaussian noise and stripes while preserving clear surface features. With the assistance of BM4D, GTVLRR exhibits the strongest response to anomaly objects among all the two-step schemes, as depicted in Fig. 15. Notably, the background-anomaly separation achieved by HyADD is notably superior to that of GTVLRR.

In Fig. 16, although GTVLRR and LSMAD deliver remarkable performances, their PDs reach the highest value slowly. Conversely, the PDs of HyADD reach 1.0 when FAR = 0.4, outpacing the other algorithms. Regarding quantitative quality, Table V presents the AUC values of seven detectors for the GF5-MouMan dataset. Our method surpasses GTVLRR by 1.31%, 3.87%, and 5.19% after denoising by BM4D, TDL, and BM3D, respectively, respectively. GTVLRR (after BM4D) and LSMAD (after TDL) achieve the second-highest and third-highest AUCs among the competing algorithms.

IV. CONCLUSION AND FURTHER ANALYSIS

In this article, we have pointed out the drawbacks of the two-step AD-denoising scheme and the advantages of the AD-denoising integration framework. Noise removal benefits the exploration of subspace domain-based LR property to locate anomalies. With the assistance of an antinoise dictionary design, anomaly location can offer effective feedback for denoising, especially preserving the piecewise spatial-spectral smoothness. The HyADD algorithm is designed and solved by LADMAP to output AD maps and denoised results simultaneously. Experimental results on simulated and real-world HS datasets verify the superior performances of our proposed method.

REFERENCES

- [1] J. Li, K. Zheng, Z. Li, L. Gao, and X. Jia, "X-shaped interactive autoencoders with cross-modality mutual learning for unsupervised hyperspectral image super-resolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5518317.
- [2] J. Li, K. Zheng, L. Gao, L. Ni, M. Huang, and J. Chanussot, "Model-informed multi-stage unsupervised network for hyperspectral image super-resolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5516117.
- [3] M. Bi, M. Wang, Z. Li, and D. Hong, "Vision transformer with contrastive learning for remote sensing image scene classification," *IEEE J. Select. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 738–749, Dec. 2023, doi: 10.1109/JSTARS.2023.3230835.
- [4] K. Bernard, Y. Tarabalka, J. Angulo, J. Chanussot, and J. A. Benediktsson, "Spectral-spatial classification of hyperspectral data based on a stochastic minimum spanning forest approach," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2008–2021, Apr. 2012.
- [5] L. Ren, L. Gao, M. Wang, X. Sun, and J. Chanussot, "HADGSM: A unified nonconvex framework for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5503415.
- [6] L. Ren, Z. Ma, F. Bovolo, and L. Bruzzone, "A nonconvex framework for sparse unmixing incorporating the group structure of the spectral library," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5506719.
- [7] B. Tu, X. Yang, W. He, J. Li, and A. Plaza, "Hyperspectral anomaly detection using reconstruction fusion of quaternion frequency domain analysis," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 6, pp. 8358–8372, Jun. 2024.
- [8] M. Wang, Q. Wang, D. Hong, S. K. Roy, and J. Chanussot, "Learning tensor low-rank representation for hyperspectral anomaly detection," *IEEE Trans. Cybern.*, vol. 53, no. 1, pp. 679–691, Jan. 2023.

- [9] B. Zhang et al., "Progress and challenges in intelligent remote sensing satellite systems," *IEEE J. Select. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 1814–1822, Feb. 2022, doi: [10.1109/JSTARS.2022.3148139](https://doi.org/10.1109/JSTARS.2022.3148139).
- [10] L. Zhuang, M. K. Ng, L. Gao, and Z. Wang, "Eigen-CNN: Eigenimages plus eigennoise level maps guided network for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5512018.
- [11] B. Tu, Z. Wang, H. Ouyang, X. Yang, J. Li, and A. Plaza, "Hyperspectral anomaly detection using the spectral–spatial graph," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5542814.
- [12] S. Wang, X. Wang, L. Zhang, and Y. Zhong, "Auto-AD: Autonomous hyperspectral anomaly detection network based on fully convolutional autoencoder," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5503314.
- [13] B. Tu, X. Yang, X. Ou, G. Zhang, J. Li, and A. Plaza, "Ensemble entropy metric for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5513617.
- [14] I. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 38, no. 10, pp. 1760–1770, Oct. 1990.
- [15] J. M. Molero, E. M. Garzón, I. García, and A. Plaza, "Analysis and optimizations of global and local versions of the RX algorithm for anomaly detection in hyperspectral data," *IEEE J. Select. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 801–814, Apr. 2013.
- [16] Y. Taitano, B. Geier, and K. Bauer, "A locally adaptable iterative RX detector," *EURASIP J. Adv. Signal Process.*, vol. 2010, 2010, Art. no. 341908.
- [17] H. Kwon and N. Nasrabadi, "Kernel RX-algorithm: A nonlinear anomaly detector for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 2, pp. 388–397, Feb. 2005.
- [18] A. Schaum, "Joint subspace detection of hyperspectral targets," in *Proc. IEEE Aerosp. Conf.*, 2004, pp. 1818–1824.
- [19] S. Matteoli, M. Diani, and G. Corsini, "Improved estimation of local background covariance matrix for anomaly detection in hyperspectral images," *Opt. Eng.*, vol. 49, no. 4, pp. 1–16, 2010.
- [20] C. Zhao, X.-F. Yao, and Y. Yan, "Modified kernel RX algorithm based on background purification and inverse-of-matrix-free calculation," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 4, pp. 544–548, Apr. 2017.
- [21] Q. Ling, Y. Guo, Z. Lin, and W. An, "A constrained sparse representation model for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2358–2371, Apr. 2019.
- [22] Y. Yang, S. Song, D. Liu, J. Zhang, and T. C.-W. Chan, "Robust background feature extraction through homogeneous region-based joint sparse representation for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 10, pp. 8723–8737, Oct. 2021.
- [23] M. Vafadar and H. Ghassemian, "Anomaly detection of hyperspectral imagery using modified collaborative representation," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 4, pp. 577–581, Apr. 2018.
- [24] R. Wang, H. Hu, F. He, F. Nie, S. Cai, and Z. Ming, "Self-weighted collaborative representation for hyperspectral anomaly detection," *Signal Process.*, vol. 177, 2020, Art. no. 107718.
- [25] H. Su, Z. Wu, H. Zhang, and Q. Du, "Hyperspectral anomaly detection: A survey," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 1, pp. 64–90, Mar. 2022.
- [26] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 171–184, Jan. 2013.
- [27] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?," *J. ACM*, vol. 58, no. 3, pp. 1–37, 2011.
- [28] S.-Y. Chen, S. Yang, K. Kalpakis, and C.-I. Chang, *Low-Rank Decomposition-Based Anomaly Detection*. Bellingham, WA, USA: SPIE, 2013, pp. 1–24.
- [29] M. Farrell and R. Mersereau, "On the impact of covariance contamination for adaptive detection in hyperspectral imaging," *IEEE Signal Process. Lett.*, vol. 12, no. 9, pp. 649–652, Sep. 2005.
- [30] N. Billor, A. S. Hadi, and P. F. Velleman, "Bacon: Blocked adaptive computationally efficient outlier nominators," *Comput. Statist. Data Anal.*, vol. 34, no. 3, pp. 279–298, 2000.
- [31] Y. Zhang, B. Du, L. Zhang, and S. Wang, "A low-rank and sparse matrix decomposition-based Mahalanobis distance method for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1376–1389, Mar. 2016.
- [32] Y. Xu, Z. Wu, J. Li, A. Plaza, and Z. Wei, "Anomaly detection in hyperspectral images based on low-rank and sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 1990–2000, Apr. 2016.
- [33] L. Li, W. Li, Q. Du, and R. Tao, "Low-rank and sparse decomposition with mixture of Gaussian for hyperspectral anomaly detection," *IEEE Trans. Cybern.*, vol. 51, no. 9, pp. 4363–4372, Sep. 2021.
- [34] T. Cheng and B. Wang, "Graph and total variation regularized low-rank representation for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 391–406, Jan. 2020.
- [35] W. Xie, X. Zhang, Y. Li, J. Lei, J. Li, and Q. Du, "Weakly supervised low-rank representation for hyperspectral anomaly detection," *IEEE Trans. Cybern.*, vol. 51, no. 8, pp. 3889–3900, Aug. 2021.
- [36] M. Wang, Q. Wang, D. Hong, S. Roy, and J. Chanussot, "Learning tensor low-rank representation for hyperspectral anomaly detection," *IEEE Trans. Cybern.*, vol. 53, no. 1, pp. 679–691, Jan. 2023.
- [37] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. M. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.
- [38] M. Wang, Q. Wang, J. Chanussot, and D. Li, "Hyperspectral image mixed noise removal based on multidirectional low-rank modeling and spatial-spectral total variation," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 1, pp. 488–507, Jan. 2021.
- [39] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D: Nonlinear Phenom.*, vol. 60, no. 1/4, pp. 259–268, 1992.
- [40] W. He, H. Zhang, L. Zhang, and H. Shen, "Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 178–188, Jan. 2016.
- [41] T. Cheng and B. Wang, "Total variation and sparsity regularized decomposition model with union dictionary for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1472–1486, Feb. 2021.
- [42] S. Yang and Z. Shi, "Hyperspectral image target detection improvement based on total variation," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 2249–2258, May 2016.
- [43] H. K. Aggarwal and A. Majumdar, "Hyperspectral image denoising using spatio-spectral total variation," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 3, pp. 442–446, Mar. 2016.
- [44] A. Majumdar, N. Ansari, H. Aggarwal, and P. Biyani, "Impulse denoising for hyper-spectral images: A blind compressed sensing approach," *Signal Process.*, vol. 119, pp. 136–141, 2016.
- [45] H. K. Aggarwal and A. Majumdar, "Exploiting spatio-spectral correlation for impulse denoising in hyperspectral images," *J. Electron. Imag.*, vol. 24, no. 1, 2015, Art. no. 013027.
- [46] Z. Lin, R. Liu, and Z. Su, "Linearized alternating direction method with adaptive penalty for low-rank representation," in *Proceed. Int. Conf. Neural Inf. Process. Syst.*, 2011, pp. 612–620.
- [47] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.
- [48] C. C. Paige and M. A. Saunders, "LSQR: An algorithm for sparse linear equations and sparse least squares," *ACM Trans. Math. Softw.*, vol. 8, no. 1, pp. 43–71, 1982.
- [49] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 2862–2869.
- [50] Y. Chang, L. Yan, and S. Zhong, "Hyper-Laplacian regularized unidirectional low-rank tensor recovery for multispectral image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 5901–5909.
- [51] W. He, Q. Yao, C. Li, N. Yokoya, and Q. Zhao, "Non-local meets global: An integrated paradigm for hyperspectral denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 6861–6870.
- [52] Y. Ye, C. Chen, B. He, and X. Yuan, "The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent," *Math. Prog.*, vol. 155, pp. 57–79, 2016.
- [53] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
- [54] M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi, "Nonlocal transform-domain filter for volumetric data denoising and reconstruction," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 119–133, Jan. 2013.
- [55] Y. Peng, D. Meng, Z. Xu, C. Gao, Y. Yang, and B. Zhang, "Decomposable nonlocal tensor dictionary learning for multispectral image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 2949–2956.
- [56] Y. Y. Tang, Y. Lu, and H. Yuan, "Hyperspectral image classification based on three-dimensional scattering wavelet transform," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2467–2480, May 2015.
- [57] J. Kerekes, "Receiver operating characteristic curve confidence intervals and regions," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 251–255, Apr. 2008.

- [58] S. Khazai, S. Homayouni, A. Safari, and B. Mojaradi, "Anomaly detection in hyperspectral images based on an adaptive support vector method," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 646–650, Jul. 2011.
- [59] C. Rogass et al., "Reduction of uncorrelated striping noise—applications for hyperspectral pushbroom acquisitions," *Remote Sens.*, vol. 6, no. 11, pp. 11082–11106, 2014.
- [60] T. D. Cocks, R. Jenssen, A. E. Stewart, I. Wilson, and T. Shields, "1 the hymap TM airborne hyperspectral sensor: The system, calibration and performance," in *Proc. EARSEL Workshop Imag. Spectro.*, 1998, pp. 37–42.
- [61] J. Pearlman, S. Carman, C. Segal, P. Jarecke, P. Clancy, and W. Browne, "Overview of the hyperion imaging spectrometer for the NASA EO-1 mission," in *Proc. IEEE Geosci. Remote Sens. Symp.*, 2001, pp. 3036–3038.
- [62] C. Rogasß et al., "Reduction of radiometric miscalibration—applications to pushbroom sensors," *Sensors*, vol. 11, no. 6, pp. 6370–6395, 2011.
- [63] H. Zhang, W. He, L. Zhang, H. Shen, and Q. Yuan, "Hyperspectral image restoration using low-rank matrix recovery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4729–4743, Aug. 2014.
- [64] M. Wang et al., "Tensor decompositions for hyperspectral data processing in remote sensing: A comprehensive review," *IEEE Geosci. Remote Sens. Mag.*, vol. 11, no. 1, pp. 26–72, Mar. 2023.
- [65] X. Yang, B. Tu, Q. Li, J. Li, and A. Plaza, "Graph evolution-based vertex extraction for hyperspectral anomaly detection," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Aug. 5, 2023, doi: [10.1109/TNNLS.2023.3303273](https://doi.org/10.1109/TNNLS.2023.3303273).



Minghua Wang (Member, IEEE) received the B.S. degree from the School of Automation, Harbin Institute of Technology, Harbin, China, in 2016, and the Ph.D. degree in control science and engineering from the Harbin Institute of Technology, in 2021.

She is currently an Associate Professor with the College of Artificial Intelligence, Nankai University, Tianjin, China. From 2019 to 2020, she was a Visiting Ph.D. Student with University of Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, Grenoble, France. From 2021 to 2023, she was with the Aerospace

Information Research Institute, Chinese Academy of Sciences, Beijing, China. Her research interests include remote sensing image processing, noise removal, anomaly detection, machine learning, and deep learning.

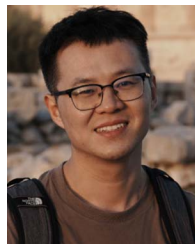


Lianru Gao (Senior Member, IEEE) received the B.S. degree in civil engineering from Tsinghua University, Beijing, China, in 2002, and the Ph.D. degree in cartography and geographic information system from Institute of Remote Sensing Applications, Chinese Academy of Sciences (CAS), Beijing, in 2007.

He is currently a Professor with the Key Laboratory of Computational Optical Imaging Technology, Aerospace Information Research Institute, CAS. In 2014, he was a Visiting Scholar with the University of Extremadura, Cáceres, Spain, and with Mississippi

State University (MSU), Starkville, MS, USA, in 2016. In the last ten years, he was the Principal Investigator of ten scientific research projects at national and ministerial levels, including projects by the National Natural Science Foundation of China from 2010 to 2012, 2016 to 2019, and 2018 to 2020, the National Key R&D Program of China (2021–2025), and so on. He has authored or coauthored more than 240 peer-reviewed papers, and there are more than 150 journal papers included by Science Citation Index (SCI). He has coauthored three academic books including "Hyperspectral Image Information Extraction." He obtained 30 National Invention Patents in China. His research interests include hyperspectral image processing and information extraction.

Dr. Gao was the recipient of the Outstanding Science and Technology Achievement Prize of the CAS in 2016, won the Second Prize of The State Scientific and Technological Progress Award in 2018, and supported by the National Science Foundation for Distinguished Young Scholars of China in 2023. He was also the recipient of the 2021 Outstanding Paper Award at the IEEE Workshop on Hyperspectral Image Processing: Evolution in Remote Sensing (WHISPERS). He is an Associate Editor for the IEEE TGRS and IET Image Processing. He is a Fellow of the Institution of Engineering and Technology.



Longfei Ren received the Ph.D. degree from the School of Information Science and Technology, Southwest Jiaotong University, Chengdu, China, in 2021.

He is currently with the Key Laboratory of Computational Optical Imaging Technology, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China. From 2019 to 2020, he was a Visiting Ph.D. Student with Remote Sensing Laboratory, Department of Information and Communication Technologies, University of Trento, Trento, Italy, and

the Remote Sensing for Digital Earth Unit, Fondazione Bruno Kessler, Trento, Italy. His research interests include hyperspectral unmixing, optimization, and image processing.



Xian Sun (Senior Member, IEEE) received the B.Sc. degree from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 2004, and the M.Sc. and Ph.D. degrees from the Institute of Electronics, Chinese Academy of Sciences, Beijing, China, in 2009.

He is currently a Professor with Aerospace Information Research Institute, Chinese Academy of Sciences. His research interests include computer vision, geospatial data mining, and remote sensing image understanding.



Jocelyn Chanussot (Fellow, IEEE) received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (Grenoble INP), Grenoble, France, in 1995, and the Ph.D. degree from the Université de Savoie, Annecy, France, in 1998.

Since 1999, he has been with Grenoble INP, where he is currently a Professor of signal and image processing. He has been a Visiting Scholar with Stanford University, Stanford, CA, USA; KTH (Sweden), Stockholm, Sweden; and National University of Singapore, Singapore. Since 2013, he has been an Adjunct Professor with the University of Iceland, Reykjavik, Iceland. From 2015 to 2017, he was a Visiting Professor with the University of California, Los Angeles, CA, USA. He holds the AXA chair in remote sensing and is an Adjunct Professor with the Chinese Academy of Sciences, Aerospace Information research Institute, Beijing, China. His research interests include image analysis, hyperspectral remote sensing, data fusion, machine learning, and artificial intelligence.

Dr. Chanussot was the recipient of multiple outstanding paper awards. He is the founding President of IEEE Geoscience and Remote Sensing French chapter (2007–2010) which received the 2010 IEEE GRS-S Chapter Excellence Award. He was the Vice-President of the IEEE Geoscience and Remote Sensing Society, in charge of meetings and symposia (2017–2019). He was the General Chair of the first IEEE GRSS Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote sensing (WHISPERS). He was the Chair (2009–2011) and Co-chair of the GRS Data Fusion Technical Committee (2005–2008). He was a Member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society (2006–2008) and the Program Chair of the IEEE International Workshop on Machine Learning for Signal Processing (2009). He is an Associate Editor for IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and the Proceedings of the IEEE. He was the Editor-in-Chief of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (2011–2015) and an Associate Editor for IEEE TRANSACTIONS ON IMAGE PROCESSING. In 2014, he was a Guest Editor for the IEEE Signal Processing Magazine. He is a Member of the Institut Universitaire de France (2012–2017) and a Highly Cited Researcher (Clarivate Analytics/Thomson Reuters).