Hyperspectral Image Denoising: From Model-Driven, Data-Driven, to Model-Data-Driven

Qiang Zhang, *Member, IEEE*, Yaming Zheng, Qiangqiang Yuan, *Member, IEEE*, Meiping Song, *Member, IEEE*, Haoyang Yu, *Member, IEEE*, and Yi Xiao

***Abstract*— Mixed noise pollution in HSI severely disturbs subsequent interpretations and applications. In this technical review, we first give the noise analysis in different noisy HSIs and conclude crucial points for programming HSI denoising algorithms. Then, a general HSI restoration model is formulated for optimization. Later, we comprehensively review existing HSI denoising methods, from model-driven strategy (nonlocal mean, total variation, sparse representation, low-rank matrix approximation, and low-rank tensor factorization), data-driven strategy [2-D convolutional neural network (CNN), 3-D CNN, hybrid, and unsupervised networks], to model-data-driven strategy. The advantages and disadvantages of each strategy for HSI denoising are summarized and contrasted. Behind this, we present an evaluation of the HSI denoising methods for various noisy HSIs in simulated and real experiments. The classification results of denoised HSIs and execution efficiency are depicted through these HSI denoising methods. Finally, prospects of future HSI denoising methods are listed in this technical review to guide the ongoing road for HSI denoising. The HSI denoising dataset could be found at https://qzhang95.github.io.**

***Index Terms*— Data-driven, denoising, hyperspectral image, model-data-driven, model-driven, technical review.**

I. Introduction

HYPERSPECTRAL image (HSI) contains various spectra from ultraviolet bands, visible bands, to infrared bands via the imaging spectrometer sensors [1]. The spectral interval of HSI is usually narrow at the nanometer scale. HSI could jointly receive the spatial and spectral information of observed objects. Through these broad and dense spectral features, HSI can better distinguish the subtle physical differences between different surface materials, compared with natural RGB or gray images. In recent years, HSI has been widely put into practice, such as classification, object detection, material identification, anomaly detection [2], [3], [4], [5], and so on.

Manuscript received 23 August 2022; revised 10 January 2023 and 11 April 2023; accepted 18 May 2023. Date of publication 6 June 2023; date of current version 8 October 2024. This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant 3132023262; and in part by the National Natural Science Foundation of China under Grant 41922008, Grant 42101350, Grant 61971082, and Grant 61971319. *(Corresponding author: Qiangqiang Yuan.)*

Qiang Zhang, Yaming Zheng, Meiping Song, and Haoyang Yu are with the Center of Hyperspectral Imaging in Remote Sensing, Information Science and Technology College, Dalian Maritime University, Dalian 116026, China (e-mail: whuqzhang@gmail.com; zym1505@dlmu.edu.cn; smping@163.com; yuhy@dlmu.edu.cn).

Qiangqiang Yuan and Yi Xiao are with the School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China (e-mail: yqiang86@gmail.com; xiao_yi@whu.edu.cn).

Digital Object Identifier 10.1109/TNNLS.2023.3278866

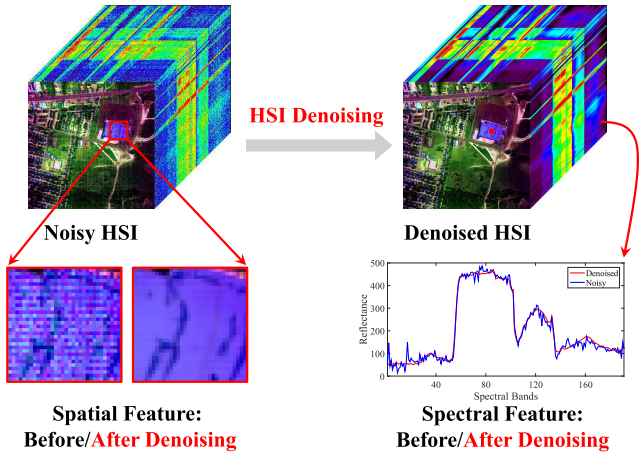


Fig. 1. Spatial and spectral features in noisy and denoised HSI: before/after HSI denoising (lower left: spatial feature and lower right: spectral feature).

However, everything has two sides, including HSI. Due to the atmospheric interference and signal response, almost all of the acquired HSIs are inevitably polluted by various noises to different degrees [6]. The noise pollution destroys both initial spectral and spatial features in HSI, as depicted in Fig. 1.

On the one hand, if we discard all these noisy HSIs in subsequent applications, it will lead to severe data waste [7]. On the other hand, if we ignore these noise pollution problems and directly utilize them for subsequent applications, it will also disturb the interpretation algorithms [8], [9], [10], [11]. For instance, many works manifest that the noise pollution problem results in reducing HSI classification accuracy. However, it is hard to suppress the mixed noise in HSI through hardware improvement [12], [13].

Therefore, how to remove noise in HSI via denoising algorithms is indispensable before HSI interpretation and applications [14], [15]. As shown in Fig. 2, the HSI denoising task has gradually become a hotspot in recent two decades.

Up to now, a few literatures have concentrated on HSI denoising tasks from a signal processing perspective. Therefore, we devote ourselves to technically reviewing existing HSI denoising methods, to further exploit improved ideas and directions for prospective HSI denoising algorithms. Current HSI denoising methods are classified into three types: 1) model-driven strategy; 2) data-driven strategy; and

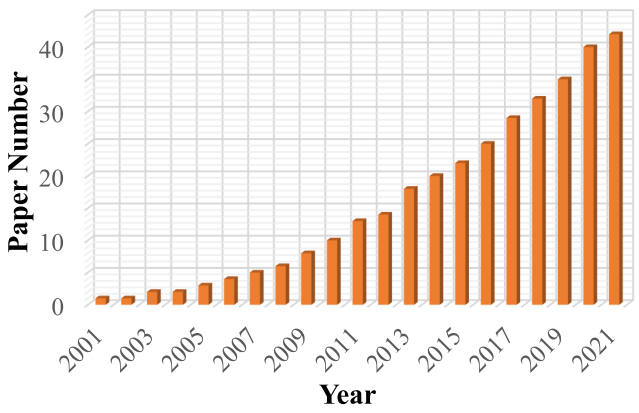


Fig. 2. Paper number which searched in Google Scholar on the topic of HSI denoising in recent twenty years.

3) model-data-driven strategy. Later, related experiments and discussions are given to manifest the effectiveness of these noise-remove algorithms for HSI. Finally, we generalize the conclusion and future directions for HSI denoising. Overall, the primary contributions of this work are listed in the following.

1) We specifically give the noise analysis in different noisy HSIs and conclude five crucial points for programming HSI denoising algorithms. Then, the general HSI restoration model is formulated for subsequent optimization.
2) We technically review existing HSI denoising methods, from model-driven, data-driven, to model-data-driven strategy. The advantages and disadvantages of each strategy for HSI denoising are summarized and contrasted.
3) We comprehensively present an evaluation of the HSI denoising methods in simulated and real experiments. The classification results of denoised HSIs and execution efficiency are depicted for these methods.
4) Prospects are summarized for future works on HSI denoising. Both the challenges and improved directions are described through further combined model-driven with a data-driven strategy for the denoising task.

II. Problem Formulation

A. Noise Analysis

Before introducing existing HSI denoising methods, we need specifically give the noise analysis in different noisy HSIs. As stated in Section I, HSI suffers from both atmospheric imaging environment and sensor issues [16]. In consequence, most HSIs are polluted by mixed noise. The noise types mainly include Gaussian noise, Poisson noise, impulse noise, stripe noise, and deadline [17]. What is worse, the noise intensity and type usually vary from different bands in single-noisy HSIs [18].

As shown in Fig. 3, noise pollution in different HSIs (urban, EO-1, and GF-5) and various bands demonstrates the complexity and difficulty for HSI denoising. In urban and GF-5 HSIs, they suffer from stripe noise and Gaussian noise in different bands. EO-1 HSI is affected by stripe noise, random noise, and deadline.

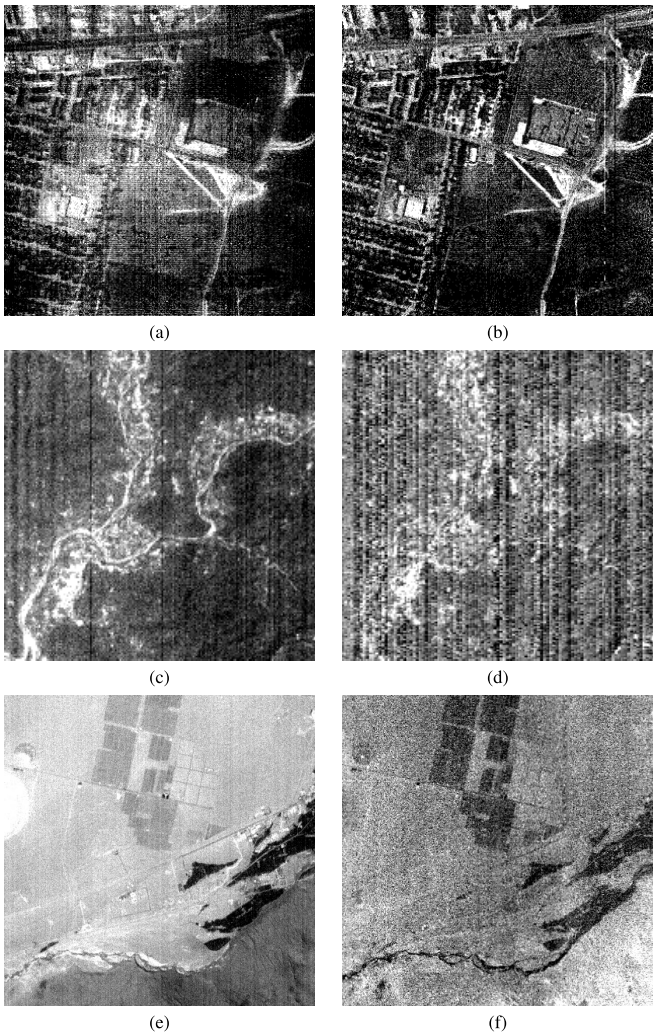


Fig. 3. Noise pollution in different HSIs and different bands. (a) Urban (103rd band). (b) Urban (188th band). (c) EO-1 (second band). (d) EO-1 (165th band). (e) GF-5 (first band). (f) GF-5 (328th band).

Therefore, five crucial points are significant when programming suitable HSI denoising algorithms, via the noise analysis in different noisy HSIs. These essential points are listed as follows.

1) HSI denoising should remove mixed noise in noisy bands, without existing noise residuals and artifacts.
2) HSI denoising should maintain the original feature in noise-free bands, without destroying clean spatial and spectral information.
3) HSI denoising should keep global spectral preservation in HSI, without generating noticeable spectral distortion.
4) HSI denoising should ensure generalization ability for different HSIs, without appearing overfitting problem.
5) HSI denoising should guarantee high efficiency and processing speed, without consuming a large amount of time.

B. General HSI Restoration Model

After the noise analysis in different noisy HSIs, we then formulate the general HSI restoration model before reviewing existing HSI denoising algorithms.

TABLE I
TAXONOMY OF EXISTING HSI DENOISING METHODS

Method Type	Subcategory	Representative methods	Pros and cons
Model-Driven	Non-Local Mean	BM4D [19]; 3D-NLM [20]	Pros: The structural similarity between different pixels and different non-localized plaques is considered. Cons: Spatial over-smoothness effect.
	Total Variation	SSAHTV [21]; SAMNLTV [22]; ASSTV [23]; SSTV [24]; HSSTV [25]; 3DCrTV [26]; E-3DTV [27]	Pros: The marginal features of HSI are effectively preserved. Cons: Model adaptability is poor.
	Sparse Representation	HSSNR [28]; NS-PCA [29]; WS-PCA [30]; WSRRR [31]; BDL [32]; 3D-NLS [33]; MTSNMF [34]; SRLR [35]; SSASR [36]; SSDSR [37]; FastHyDe [38]	Pros: Spectral redundancy are used. Cons: Spatial correlation of 3D data cannot be fully utilized.
	Low-Rank Matrix Approximation	LRMR [39]; NAILRMR [40]; GLRR [41]; NonLRMA [42]; RPCA-LRMF [43]; LRRSDS [44]; BLRMF [45]; SSLR [46]; NLRPnP [47]; NonRLRS [48]; SS-LRR [49]; FSSLRL [50]; F-LRNMF [51]; HyDRoS [52]; ISLRMA [53]; VLRMFmcl [54]; LRTV [55]; LLRSSTV [56]; LRCSSTV [57]; LRWTV [58]; BFTV [59]; LLRMTV [60]; NMoG [61]; BNLRMF [62]; GRLD [63]; DLR [64]; SSLR-SSTV [65]	Pros: Effectively extract the spatial-spectral redundancy of HSI, and retain internal spatial details and spectral information. Cons: Destroys the overall multidimensional information of HSI.
	Low-Rank Tensor Factorization	ITSReg [66]; NLR-CPTD [67]; R1TD [68]; LRTA [69]; GKTD [70]; TDL [71]; LLRT [72]; NTD-HALS [73]; LRTDTV [74]; LRTD [75]; TV-NLRTD [76]; LRTDGS [77]; WLRTR [78]; LTDL [79]; NGMeet [80]; LRTF-DFR [81]; FGSLR [82]; LRTR [83]; SSTV-LRTF [84]; MSt-SVD [85]; GTRPCA [86]; CTSD [87]; 3DTNN [88]; TenSRDe [89]; NonLRTA [90]; LTR-L_{1-2}SSTV [91]; FRCTR-PnP [92]; MLR-SSTV [93]; TLR-L_0TV [94]; NLTR [95]; LRTFL_0 [96]	Pros: Extract global spatial-spectral correlation, and make good use of low-rank and smoothness in cube data. Cons: Relies heavily on special priors and regularization, and takes a long time.
Data-Driven	2-D CNN	SSDRN [97]; HSID-CNN [98]; SSGN [99]; HSI-DeNet [100]; DeCS-Net [101]; HSI-SDeCNN [102]; DnRCNN [103]	Pros: Effectively remove the noise of the spatial dimension and spectral dimension. Cons: Cannot completely process the 3-D spectral-spatial structure in the HSI.
	3-D CNN	3D-ADCNN [104]; 3D-DCNN [105]; 3D-ADNet [106]; TRQ3DNet [107];	Pros: High-dimensional data is effectively processed to extract 3D spatial-spectral structures. Cons: 3-D network parameters are superabundant. And hard to train.
	Hybrid Networks	QRNN3D [108]; DSSGRN [109]; DELIL [110]; HyperMixNet [111]; FastHyMix [112]; Hider [113]	Pros: Effectively integrate different networks or models. Cons: Lack of training samples for noise-free HSI.
	Unsupervised Networks	DHP [114]; SIP [115]; SURE-CNN [116]; S2DIP [117]	Pros: No noise-free HSI training samples are required. Cons: The intrinsic prior structure of HSI (sparseness and low rankness) is ignored. And lacks interpretability.
Model-Data-Driven		NMF-DPR [118]; DSSBP [119]; SMDS-Net [120]; LR-Net [121]	Pros: Avoids the problems of time-consuming parameter adjustment and lack of training samples. Cons: The generalization ability is not stable for different noisy HSIs.

To distinguish data dimensions in this work, different values are, respectively, represented as scalar x (lower and italic), vector $\mathbf{x}$ (lower and overstriking), matrix $\mathbf{X}$ (upper and overstriking), and tensor $\mathcal{X}$ (upper and Euclid math font) formats. In terms of a third-order tensor $\mathcal{X}$ value, $\mathcal{X}(:, j, b)$, $\mathcal{X}(i, :, b)$, and $\mathcal{X}(i, j, :)$ severally stand for the (j, b)th column, (i, b)th row, and (i, j)th slice of tensor $\mathcal{X}$. Besides, $\mathbf{X}_{(m)}$ is denoted as the unfolding matrix format of the third-order tensor $\mathcal{X} \in \mathbb{R}^{k_1 \times k_2 \times k_3}$, with a size of $k_m \times \prod_{n \neq m} k_n$.

The HSI degraded procedure could be simply formulated as the following equation:

$$\mathcal{Y} = \mathcal{X} + \mathcal{N} + \mathcal{S} \tag{1}$$

where noisy HSI $\mathcal{Y} \in \mathbb{R}^{k_1 \times k_2 \times k_3}$ could be regarded as a natural tensor, with the spatial size $k_1 \times k_2$ and spectrum size k_3. $\mathcal{X} \in \mathbb{R}^{k_1 \times k_2 \times k_3}$ represents the clean HSI. $\mathcal{N} \in \mathbb{R}^{k_1 \times k_2 \times k_3}$ stands for the additive random noise like Gaussian noise and Poisson noise in HSI. $\mathcal{S} \in \mathbb{R}^{k_1 \times k_2 \times k_3}$ denotes the sparse noise, such as stripe noise and deadline in HSI.

To solve this general HSI restoration model, many scholars have proposed plenty of methods for HSI denoising in recent decades. Current HSI denoising methods can be generally split into three types: 1) model-driven; 2) data-driven; and 3) model-data-driven methods. As depicted in Table I, a taxonomy of exiting HSI denoising methods is classified with partial representative model-driven, data-driven, and model-data-driven methods.

Based on the existing literatures, Section III concentrates on model-driven methods. Section IV focuses on data-driven methods. And Section V provides several model-data-driven methods.

III. Model-Driven Methods

The model-driven strategy mainly utilizes the intrinsic characteristics of HSI, to remove mixed noise in HSI. These methods have excellent interpretability and embeddability. However, different HSIs have differences in data distribution, scene, and noise, resulting in strong parameter sensitivity of the model, which requires a lot of time for parameter adjustment. In addition, this kind of method has high computational complexity and time consuming. Up to now, nonlocal mean [19], [20], total variation (TV) [21], [22], [23], [24], [25], [26], [27], sparse representation [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], low-rank matrix approximation [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62], [63], [64], [65], and low-rank tensor factorization [66], [67], [68], [69], [70], [71], [72], [73], [74], [75], [76], [77], [78], [79], [80], [81], [82], [83], [84], [85], [86], [87], [88], [89], [90], [91], [92], [93], [94], [95], [96] have been employed for HSI denoising. The general HSI restoration framework could be concluded as

$$\arg\min_{\mathcal{X}} \frac{1}{2}\|\mathcal{Y} - \mathcal{X} - \mathcal{S}\|_F^2 + \alpha \cdot \mathcal{R}(\mathcal{X}) + \beta \cdot \mathcal{T}(\mathcal{S}) \tag{2}$$

where $\mathcal{R}(\mathcal{X})$ and $\mathcal{T}(\mathcal{S})$ represent two regularization terms, to utilize internal characteristics of HSI $\mathcal{X}$ and sparse noise $\mathcal{S}$. α and β are two balanced parameters of the two regularization terms. Specific descriptions of these five subcategories are summarized as follows.

A. Nonlocal Mean

The original nonlocal prior was proposed [122] in 1999. It is based on the high similarity between different nonlocal patches in natural images. Nonlocal mean simultaneously considers the numerical similarity between different pixels and the structural similarity between different nonlocal patches.

In terms of HSI denoising, nonlocal mean has also been introduced in recent years. Through imposing the nonlocal prior, (2) could be transformed as

$$\arg\min_{\mathcal{X}} \frac{1}{2}\|\mathcal{Y} - \mathcal{X} - \mathcal{S}\|_F^2 + \alpha \cdot \mathrm{NLM}(\mathcal{X}) \tag{3}$$

where $\mathrm{NLM}(\cdot)$ represents the nonlocal regularization term in HSI denoising. α is a balanced parameter of the regularization term. For instance, Maggioni et al. [19] developing a 4-D collaborative filtering paradigm (named BM4D) for volumetric data denoising. By means of the Fourier-domain transformation, BM4D calculates the weights of every cube group, which are then adaptively aggregated on their initial positions. Qian et al. [20] developed a novel 3-D nonlocal mean (3-D NLM) method for HSI denoising. The spatiospectral structure is jointly considered in 3-D NLM, which takes full advantage of spatiospectral correlation and nonlocal similarity. In addition, Zhao et al. [123] developed a 3-D orthogonal transformation (3-DOT) denoising method based on sparse spectral space and low-rank representation (SSSLRR). In this method, the SSSLRR of the 3-D discrete orthogonal wavelet transform is used.

Generally, the nonlocal mean strategy could produce smooth results on spatial dimension for HSI denoising. However, NLM-based methods easily tend to cause an over-smooth effect when removing the mixed noise in HSI. What is worse, the running-time of NLM-based HSI denoising methods is usually time consuming.

B. Total Variation

TV prior is first presented by Rudin et al. [124]. Specific to a gray image $\mathbf{M}$, TV prior is defined as follows:

$$\mathrm{TV}(\mathbf{M}) = \sum_{i,j} \sqrt{\left(\nabla_{i,j}^h \mathbf{M}\right)^2 + \left(\nabla_{i,j}^v \mathbf{M}\right)^2} \tag{4}$$

where $\nabla_{i,j}^h$ and $\nabla_{i,j}^v$, respectively, stand for the horizontal and vertical first-order difference operations at location (i, j). Due to the capacity of preserving edge features, the TV model has been successfully applied for hyperspectral image processing. In terms of HSI denoising, (2) could be rewritten as the following model by imposing the TV prior:

$$\arg\min_{\mathcal{X}} \frac{1}{2}\|\mathcal{Y} - \mathcal{X} - \mathcal{S}\|_F^2 + \alpha \cdot \mathrm{TV}(\mathcal{X}). \tag{5}$$

The most convenient strategy of introducing TV prior to HSI denoising is using the band-by-band way, where $\mathrm{TV}(\mathcal{X})$ can be transformed as $\sum_b^{k_3} \mathrm{TV}(\mathcal{X}(:, :, b))$. And α is a balanced

parameter of the TV($\mathcal{X}$). However, the noise location, type, and level vary with different bands in HSI. This will lead to insufficient or excessive denoising effects for each band. To solve this problem, several modified spatiospectral TV models are proposed for HSI denoising. Based on the typical 2-D TV prior, Yuan et al. [21] developed the spatiospectral adaptive hyperspectral TV (SSAHTV) model. SSAHTV simultaneously takes noise levels among different spectrums and spatial heterogeneity within a single band into consideration. On the basis of SSAHTV, [22] further proposed the spectral adaptive multidimensional nonlocal TV (SAMNLTV) method for HSI denoising. SAMNLTV exploits the redundancy between the high-relevant or neighboring bands, which considers the spectral continuity in HSIs. It introduces the nonlocal prior into a multidimensional TV model, to utilize the similarity of different patches in high-relevant or neighboring bands for HSI denoising.

Except for the multichannel TV model, the 3-D TV model has also been developed to jointly extract spatial and spectral features for HSI denoising. Aggarwal and Majumdar [24] presented the spatiospectral TV (SSTV) model for noise removal. SSTV joint exploits spatiospectral correlation, which leads to synthesis prior via the transform domain. It utilizes two L_1 norms by the horizontal and vertical finite differencing operator for the 3-D cube and employs a quadratic penalty function in the split Bregman optimization. Furthermore, Chang et al. [23] progressed on the anisotropic spectral–spatial TV (ASSTV) framework for HSI destriping. ASSTV treats multispectral images as 3-D cube data, which both considers the continuity in spectral dimension and the consistency in the spatial dimension. In addition, [25] designed a hybrid spatiospectral TV (HSSTV) model for mixed noise reduction. HSSTV model combines difference operators with hybrid L_{1-2} norm. The optimization scheme of HSSTV could be easily addressed via proximal splitting algorithms. Relied on 3-D TV, Sun et al. [26] improved the spatiospectral weights of TV regularization term, which proposed a 3-D weighted cross TV (3DCrTV) model for HSI noise removal. Peng et al. [27] progressed on an enhanced 3-D TV (E-3DTV), which estimates the global sparsity along all the bands for HSI denoising.

In summary, TV-based HSI denoising methods could achieve pleasing results. Compared with NLM-based methods, these methods perform predominantly on edge feature perseveration, due to the TV prior in the HSI restoration procedure. Nevertheless, although several TV-based methods bring in the weighted strategy, their adaptive ability is still weak for HSI denoising. Especially for noisy HSIs with heterogeneous and complex texture regions, TV-based methods usually perform poorly for mixed noise removal.

C. Sparse Representation

The target of sparse representation is to transform a current data vector $y \in \mathbb{R}^n$ into a linear weighting combination of a limited number of fundamental atoms. This representation is generally presented as [125]

$$y = \mathbf{D}x \tag{6}$$

where $\mathbf{D} \in \mathbb{R}^{n\times k}$ stands for the overcomplete dictionary matrix ($k > n$). $x \in \mathbb{R}^k$ is denoted as the sparse coefficients, which meets the condition that the item of nonzero ingredients in x is less than k [38]. Then, the sparse prior could be formulated as

$$\arg\min_{\mathbf{D},x} \frac{1}{2}\|y - \mathbf{D}x\|_F^2 + \alpha \cdot \|x\|_0 \tag{7}$$

where $\|x\|_0$ represents the sparse regularization term. α is a balanced parameter of the sparse regularization term.

In terms of HSI denoising, sparse representation has also achieved credible results. The sparse prior for HSI denoising could be denoted as

$$\arg\min_{\mathbf{D},\mathcal{X}} \frac{1}{2}\|\mathcal{Y} - \mathbf{D}\mathcal{X}\|_F^2 + \alpha \cdot \|\mathcal{X}\|_0 \tag{8}$$

where $\|\mathcal{X}\|_0$ represents the sparse regularization term. α is a balanced parameter of the sparse regularization term. Based on sparse representation theory, related HSI denoising methods can be classified into two types: fixed-dictionary-based methods and learned-dictionary-based methods. The earliest utilization of sparse representation for noise reduction in HSI could be retrospected to fixed-dictionary-based algorithms [28], [29], [30], [31]. They employ a wavelet dictionary in hybrid spatial–spectral noise reduction (HSSNR) [28], neighbor shrinking principal component analyzing (NS-PCA) [29], wavelet shrinking principal component analyzing (WS-PCA) [30], and wavelet-based sparse reduced-rank regression (WSRRR) [31]. These methods execute a shrinkage action for the threshold values to achieve the sparse condition by 2-D or 3-D wavelet filtering.

Apart from these fixed-dictionary-based methods, the learned-dictionary-based methods for HSI denoising have gradually developed in recent years. For instance, [32] proposed a Bayesian dictionary learning (BDL) model to solve the problem of HSI denoising and inpainting. BDL considers the realistic case in HSI and assumes the Bayesian prior noise distribution in HSI. Different from the BDL method, Qian and Ye [33] utilized the noisy HSI itself to learn the overcomplete dictionary, which presented a 3-D nonlocal sparse (3D-NLS) denoising model for the noisy HSI. 3D-NLS jointly combines the nonlocal spatiospectral structural sparse prior with the noise modeling, to effectively separate the clean information and noise in HSI. Later, [34] developed the multitask sparse nonnegative matrix factorization (MTSNMF) method, which integrates dictionary learning with sparse coding. MTSNMF also uses polluted HSI itself to train the dictionary. It shares a communal sparse coefficient matrix to jointly exploit the spatiospectral structure in HSI, and imports variance-stabilizing transformation to solve hybrid noise removal.

Relied on learned-dictionary-based methods, several modified methods have also been presented for HSI denoising. Zhao and Yang [35] progressed on a sparse coding framework with low-rank constraint (SRLR) to remove the polluted noise in HSI. Sparse prior is introduced to depict holistic redundancy and correlation on spatial dimension and local redundancy and correlation on spectral dimension. Low-rankness restraint is utilized to dispose of the redundancy on spectral dimension. SRLR is effective in decreasing spectral distortion for

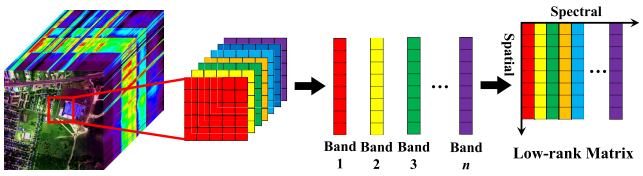


Fig. 4. Orderly transformed into low-rank matrix form for HSI.

HSI denoising. Lu et al. [36] developed a spectral–spatial adaptive sparse representation (SSASR) approach. First, adaptive spectrum partitioning is carried out to assemble the high-interrelated bands and isolate low-interrelated bands in SSASR. Second, SSASR employs spatial semblable element adaptively searching tactics to group semblable elements in local areas. Finally, SSASR introduces a sparse dictionary to adaptively manifest each subset of high-correlated bands and spatial elements. In addition, [37] raised the joint spectral–spatial distributed sparse representation (SSDSR) method. By imposing different weighted parameters of different noisy bands, SSDSR could adaptively denoise the noisy bands in HSI.

To sum up, sparse representation-based HSI denoising methods could effectively utilize spectral and spatial redundancy and correlation. However, sparse representation-based methods still exist the problem of how to adequately exploit the joint spectral–spatial redundancy in 3-D data. Therefore, the low-rank prior for HSI denoising is then increasingly developed to address this problem as follows.

D. Low-Rank Matrix Approximation

For a 2-D noisy image $\mathbf{Y}$, we assume that $\mathbf{Y}$ could be decomposed into two parts: a noise-free matrix $\mathbf{X}$ with low-rank structure and a noisy matrix $\mathbf{S}$ with sparse structure. To solve this inverse problem, the low-rank matrix approximation [126] could be introduced in the following:

$$\arg\min_{\mathbf{X},\mathbf{S}} \frac{1}{2}\|\mathbf{Y}-\mathbf{X}-\mathbf{S}\|_F^2 + \alpha\cdot\|\mathbf{X}\|_* + \beta\cdot\|\mathbf{S}\|_1 \tag{9}$$

where $\|\mathbf{X}\|_*$ stands for the matrix nuclear norm. α and β are two balanced parameters of two regularization terms. To import the low-rank matrix approximation, A 3-D noisy HSI $\mathcal{Y}$ could be orderly transformed as a 2-D matrix form through high-dimensional data unfolding operation. As shown in Fig. 4, the matrix can be formulated along the specified direction for HSI.

In recent years, low-rank matrix approximation methods have been successfully extended to HSI restoration tasks. Zhang et al. [39] first transferred the low-rank matrix recovery (LRMR) framework on HSI denoising. LRMR arranges a patch mode of HSI as a matrix format, to exploit intrinsic low rankness. It uses a go-decomposition strategy to settle this optimization issue. Later, He et al. [40] developed the noise-adjusted iterative low-rank matrix approximation (NAILRMA) model. NAILRMA takes noise intensity differences in different bands into account, which can adaptively protect the noise-free bands and denoise the noisy bands. Wang et al. [41] raised the group low-rank representative (GLRR) approach to address noise reduction issues in HSI. In consideration of the nonconvex optimization problem for noise removal, Chen et al. [42] presented the nonconvex low-rank matrix approximation (NonLRMA) method. NonLRMA depends on a nonconvex regularizer to improve the convergence performance. Xu et al. [43] employed a robust PCA low-rank matrix factorization (RPCA-LRMF) for HSI denoising. RPCA-LRMF adopts a fast matrix factorization algorithm to accelerate iterative optimization. Sun et al. [44] depended on a low-rank representation framework via spectral difference space (LRRSDS), which transforms the format of the polluted noise within HSI into the spectral difference strategy. In [45], the bilinear low-rank matrix factorization (BLRMF) model was raised for noise reduction. Ye et al. [127] proposed a novel smooth rank approximation (SRA) model, and it copes with mixture noises for HSIs.

Except for the above-mentioned methods, other priors have also been blended into low-rank matrix approximations like nonlocal prior, TV prior, and sparse prior. For example, Xue et al. [46] raised the spatial and spectral low-rankness regularization (SSLR) model. SSLR not only imports typical low-rankness prior to the spectral dimension but also leads to the nonlocal low-rank peculiarity along the spatial dimension. Zeng et al. [47] proposed a spatial nonlocal and local rank-constraint plug-and-play (NLRPnP) method, which splits HSI into nonlocal similar patches to effectively exploit local low-rankness in HSI. In [48], the nonconvex regularized low-rank and sparse (NonRLRS) model was presented. NonRLRS resolves HSI into a matrix, with low-rank terms and sparse regularization to promote the sparsity. Fan et al. [49] blended super-pixel segmentation with low-rank representation (SS-LRR). Later, Sun et al. [50] progressed on a subspace learning (FSSLRL) algorithm, which also embeds super-pixel segmentation into the HSI denoising framework. Chen et al. [51] invented a framelet-based low-rank (F-LRNMF) model for HSI denoising. F-LRNMF employs framelet regularization to efficiently guarantee the spatial textures and geometric features of denoised HSI. Gao et al. [52] realized the hyperspectral denoising task through robust principal component analysis and self-similarity (HyDRoS). Wei et al. [53] carried out the clusters with local similarity and nonlocal similarity structures from noisy HSI, which presents an intracluster structured low-rank matrix approximation (ISLRMA) method. In [54], the variational low-rank matrix factorization framework with multipatch collaborative learning (VLRMFmcl) was put forward for mixed noise removal in HSI.

In addition, TV prior has also been widely fused into a low-rank matrix approximation framework. He et al. [55] integrated TV constraint with low-rank matrix representation (LRTV) for HSI denoising. LRTV could simultaneously utilize the spectral low-rank characteristic via matrix nuclear norm and ensure spatial smoothness via TV regularization. Subsequently, He et al. [56] put forward a spatial–spectral TV regularization local low-rank (LLRSSTV) model. LLRSSTV adopts the global spatial–spectral TV terms, to further boost reconstructing smoothness in HSI. In [57], Wang et al.

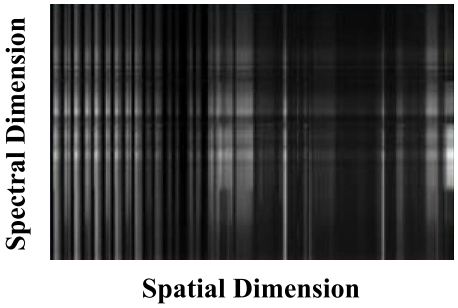


Fig. 5. Low-rank characteristic of three-order tensor in HSI.

united low-rank constraint into spectral–spatial TV (LRC-SSTV). Liu et al. [58] developed the low-rank weighted spectral–spatial TV (LRWTV) approach. LRWTV concentrates on the protection of spatial details via TV prior and spectral information via low-rank prior. Behind, Chen et al. [59] structured the bilinear factorization TV (BFTV) method, which combines bilinear matrix factorization with TV regularizer for HSI denoising. In [60], the local low-rank matrix TV (LLRMTV) method is presented. LLRMTV devotes the nonconvex penalty to HSI mixed noise removal.

In terms of noise modeling and stripe noise removal in HSI, several low-rank matrix approximation methods are also presented to better reveal the noise structure and characteristics in HSI. Chen et al. [61] devoted to the nonindependent mixture of Gaussian (NMoG) noise modeling for HSI denoising. NMoG fuses this noise modeling procedure with low-rank matrix decomposition, which achieves satisfactory results for real-noise scenarios. Du et al. [62] relied on the additive white Gaussian noise model, which proposed the bandwise noise low-rank matrix (BNLRMF) model. Lu et al. [63] raised the graph regularization low-rank destriping (GRLD) method. GRLD relies on the spectral low-rankness, to preserve the internal property of HSI. Later, Zhang et al. [64] progressed on the double low-rank (DLR) matrix factorization model for destriping. DLR collaboratively exploits low-rank prior in clean HSI, and low-rank structure of stripe noise. Subsequently, [65] employed stripe and spectral low-rankness, which integrated it into the spatiospectral TV (SSLR-SSTV) model. SSLR-SSTV works well for HSI destriping.

In summary, low-rank matrix approximation-based methods could effectively exploit the spectral and spatial redundancy of HSI. Nevertheless, low-rank matrix approximation-based methods ignore the original structure of 3-D cubes. To some degree, these methods destroy the integral multidimension information of HSI. To address this weakness, the low-rank tensor factorization framework has been quickly developed in the last five years. Detailed descriptions are elaborated in the following.

E. Low-Rank Tensor Factorization

HSI could be naturally treated as a three-order tensor [128]. Due to its redundancy and correlation on both spatial and spectral dimensions, low-rank characteristic in the three-order tensor is prominent and intrinsic, as shown in Fig. 5.

Therefore, a low-rank tensor factorization-based strategy is also appropriate for HSI restoration. General low-rank tensor factorization framework could be denoted as [129]

$$\arg\min_{\mathcal{X},\mathcal{S}} \frac{1}{2}\|\mathcal{Y}-\mathcal{X}-\mathcal{S}\|_F^2+\alpha\cdot\|\mathcal{X}\|_*+\beta\cdot\|\mathcal{S}\|_1 \tag{10}$$

where $\|\mathcal{X}\|_*$ represents the tensor nuclear norm. And the $\|\mathcal{S}\|_1$ represents a regularization term. α and β are two balanced parameters of two regularization terms. Generally, these low-rank tensor factorization methods can be distinguished through the tensor decomposition paths. Such as: 1) canonical polyadic (cp) decomposition; 2) Tucker decomposition; 3) tensor singular value decomposition (t-SVD); 4) tensor-ring decomposition; and 5) block term decomposition (BTD). Elaborate descriptions of each tensor decomposition strategy for HSI denoising are listed as follows.

1) CP Decomposition: In CP decomposition, an N-order tensor $\mathcal{X}$ could be transformed into a summation format with several rank-1 tensors with number r [130]

$$\mathcal{X}=\sum_{i=1}^{r}\lambda_i a_i^{(1)}\circ a_i^{(2)}\circ\cdots a_i^{(N)} \tag{11}$$

where $a_i^{(N)}$ stands for the unit vector. $\circ$ denotes the vector outer product operation. Based on CP decomposition, Xie et al. [66] proposed an intrinsic tensor sparsity regularization (ITSReg) model for multispectral image denoising. Xue et al. [67] developed a nonlocal low-rank regularized CP tensor decomposition (NLR-CPTD) method. NLR-CPTD adequately exploits the global correlation of spectrum and nonlocal similarity of spatial patches. In [68], the rank-1 tensor decomposition (R1TD) approach was presented for HSI noise reduction.

2) Tucker Decomposition: Different from CP decomposition, Tucker decomposition mainly concentrates on the principal component of the tensor. For a N-order tensor, its Tucker decomposition could be defined as [131]

$$\mathcal{X}=\mathcal{G}\times_1\mathbf{A}^{(1)}\times_2\mathbf{A}^{(2)}\cdots\times_{(N)}\mathbf{A}^{(N)} \tag{12}$$

where $\times_{(N)}$ represents mode-N tensor-matrix product operation. $\mathcal{G}$ stands for kernel element of $\mathcal{X}$. $\mathbf{A}^{(N)}$ refers to the matrix factor. For instance, Renard et al. [69] developed the low-rank tensor approximation (LRTA) approach for HSI denoising, which relies on joint spatial and spectral low-rank prior via Tucker decomposition. Karami et al. [70] proposed the genetic kernel tucker decomposition (GKTD) model to reduce the noise in HSI. GKTD employs the kernel element to transfer Tucker decomposition into the high-dimension characteristic field. In [71], the tensor dictionary learning (TDL) method was presented for multispectral image restoration. TDL embeds the group block-sparsity into the tucker decomposition. Chang et al. [72] progressed on a Laplacian low-rank tensor factorization (LLRT) approach, which alleviates spectral distortion problem via the hyper-Laplacian term for multispectral image denoising. Bai et al. [73] put forward a nonnegative tucker decomposition via the hierarchical alternative-least-square (NTD-HALS) method for HSI denoising. Wang et al. [74] introduced the TV prior to low-rank

tensor Tucker decomposition (LRTDTV). In LRTDTV, the L_1 norm is imposed for removing sparse noise in different bands. Furthermore, Chen et al. [75] depicts the high relevance of stripe noise, which developed a low-rank tensor decomposition (LRTD) for HSI destriping. Combining nonlocal Tucker decomposition with TV constraint (TV-NLRTD), [76] employs the overlapping 3-D cubes to jointly exploit the spectral–spatial correlation for HSI denoising. Chen et al. [77] raised a group sparse Tucker decomposition (LRTDGS) denoising approach. In [78], the weighting low-rank tensor recovery (WLRTR) was formulated for comprehensive HSI reconstruction. In consideration of the latent low-rank structure, [79] established the low-rank TDL (LTDL) approach for noise removal.

In addition, a subspace representation strategy is also developed for HSI denoising, which is assigned to a branch of Tucker decomposition [132]

$$\mathcal{X} = \mathcal{B} \times_3 \mathbf{A} \tag{13}$$

where $\mathcal{B}$ refers to the spatial reduced image. $\mathbf{A}$ stands for the spectral orthogonal matrix. For example, meeting the nonlocal prior with global spectral low-rankness (NGMeet), He et al. [80] relied on the mode-3 Tucker tensor decomposition for HSI restoration. In consideration of the latent property in the spatial reduced image and spectral orthogonal matrix, Zheng et al. [81] proposed the double factor regularized tensor factorization (LRTF-DFR) model for HSI denoising. In [82], the factor group sparsity-regularized low-rank factorization (FGSLR) approach is put forward. FGSLR employs group sparsity prior to better highlighting the low-rank trait.

3) t-SVD: For a Three-Order Tensor $\mathcal{X} \in \mathbb{R}^{k_1 \times k_2 \times k_3}$, It Can be Factorized as Three Components Through Tensor SVD:

$$\mathcal{X} = \mathcal{U} * \mathcal{S} * \mathcal{V}^{\top} \tag{14}$$

where the three tensors $\mathcal{U} \in \mathbb{R}^{k_2 \times k_2 \times k_3}$, $\mathcal{S} \in \mathbb{R}^{k_1 \times k_2 \times k_3}$, and $\mathcal{V} \in \mathbb{R}^{k_1 \times k_1 \times k_3}$ are the decomposed components of tensor $\mathcal{X}$. $*$ represents the tensor product operation. It should be noted that the kernel tensor $\mathcal{S}$ is sparse and diagonal. Based on this strategy, Fan et al. [83] formulated the low-rank tensor recovery (LRTR) approach, integrating tensor nuclear norm, and norm in a uniform convex relaxation schema. Later, Fan et al. [84] blended the spectral–spatial TV with LRTF (SSTV-LRTF) framework for HSI denoising. In [85], a multispectral t-SVD (MSt-SVD) method was developed for multispectral image denoising. Nie et al. [86] put forward a graph tensor robust PCA (GTRPCA) model. GTRPCA combines t-SVD with spectral graph regularization to accurately portray the spectral–spatial correlation. Zhao et al. [87] raised a constrained tubal rank and sparsity method (CTSD) to reduce hybrid noise in HSI.

In the t-SVD framework, Zheng et al. [88] presented the three-directional tensor nuclear norm (3DTNN) model. 3DTNN introduces a fibered-rank term into the t-SVD framework. In [89], a tensor subspace representative denoising (TenSRDe) approach was structured, which is inspired by the t-SVD strategy to utilize the intrinsic property of HSI. Furthermore, Lin et al. [90] formulated a nonconvex LRTA (NonLRTA) model, via estimating weighted pseudonorm in a three-order tensor. In combination with local low-rank tensor recovery with L_{1-2} spatiospectral TV (LTR-L_{1-2}SSTV), Zeng et al. [91] employed t-SVD to calculate the tensor nuclear norm of local patches, rather than the global HSI. Based on t-SVD, [92] presented the fibered rank-constrained tensor recovery model with an embedded plug-and-play (FRCTR-PnP) prior to HSI denoising. Wang et al. [93] combined multidirectional low-rankness with spatiospectral TV (MLR-SSTV). MLR-SSTV could effectively depict low-rank property within a three-order tensor. Furthermore, Wang et al. [94] cooperated tensor low-rank prior with L_0 TV (TLR-L_0TV), which uses weighting summation of tensor trace norm.

4) Tensor-Ring Decomposition: The objective of tensor-ring decomposition points to factorizing a high-order tensor into several circular tensor factors. The definition of tensor-ring decomposition is determined as follows:

$$\mathcal{X}(i_1, i_2, \ldots, i_N) = Tr\left(\prod_{k=1}^{N} \mathbf{G}^{(k)}(i_k)\right) \tag{15}$$

where $Tr(\cdot)$ represents the matrix trace function. $\mathbf{G}^{(k)}(i_k)$ refers to the i_kth lateral slice in kernel tensor $\mathcal{G}^{(k)}$. As an example, Chen et al. [95] developed the nonlocal tensor-ring decomposition (NLTR) model. In NLTR, tensor-ring decomposition factorizes HSI into several cyclic and contractive tensors.

5) Block Term Decomposition: BTD decomposes a tensor into several component tensors, and each component tensor is formed by the Tucker factorization

$$\mathcal{X} = \sum_{i=1}^{r} \mathcal{G}_i \times_1 \mathbf{A}_i{}^{(1)} \times_2 \mathbf{A}_i{}^{(2)} \cdots \times_{(N)} \mathbf{A}_i{}^{(N)} \tag{16}$$

where r represents the rank number of BTD. For example, Xiong et al. [96] first applied the BTD for HSI denoising, which proposed a gradient regularization low-rank tensor factorization (LRTFL_0) model. In LRTFL_0, spectral–spatial L_0 gradient regularization is embedded with BTD, to improve spectrum smoothness and preserve detailed textures.

Compared with low-rank matrix approximation, the low-rank tensor factorization strategy could better exploit the global spectral–spatial redundancy and correlation for HSI denoising. Nevertheless, the optimization procedure of low-rank tensor factorization is usually complicated and time consuming. And severely depend on special priors and regularizations. To address these issues, data-driven methods have been quickly developed in recent years. Detailed technical reviews are given in the following.

IV. Data-Driven Methods

Data-driven methods adaptively adjust the trainable parameters from external data and labels, via a deep learning framework [133]. With the quick development of deep learning, such as convolutional neural network (CNN) [134], [135], it has been extensively utilized in natural image denoising [136],

remote sensing image recovery [9], [137], [138], and reconstruction [10], [139], [140], [141]. In terms of HSI denoising, plenty of works have been proposed through a deep learning framework in recent years. These methods are classified into two types: Supervised and unsupervised networks for HSI denoising. Furthermore, the supervised networks can be specifically categorized into three types: 1) 2-D CNN [97], [98], [99], [100], [101], [102], [103]; 2) 3-D CNN [104], [105], [106], [107]; and 3) hybrid networks [108], [109], [110], [111], [112], [113]. Specific descriptions are summarized as follows.

A. 2-D CNN

For natural image denoising, 2-D CNN has been successfully applied, such as DnCNN [142]. In terms of HSI denoising, multichannel 2-D CNN has also been employed for spectral and spatial feature extraction as follows:

$$\mathbf{F}_{w,h,c} = \mathbf{X} \otimes \mathbf{C}_{w,h,c} \tag{17}$$

where $\otimes$ represents the convolution operation. $\mathbf{C}_{w,h,c}$ stands for the 2-D filter with the size of $w \times h$ and channel of c.

For instance, Xie et al. [97] proposed a spectral–spatial denoising residual network (SSDRN) for HSI noise removal. Based on the spectral difference, SSDRN relies on 2-D CNN and residual learning to denoise the polluted noise for the spectral difference in HSI. Yuan et al. [98] learns the nonlinear mapping between polluted and noise-free HSIs, which developed a spatial–spectral denoising CNN (HSID-CNN) model. In a spectral cube, the current band and its neighboring bands have high correlation and redundancy, which could be fully exploited for HSI denoising. Furthermore, Zhang et al. [99] utilized both spatial and spectral gradient as the input terms of cascaded 2-D CNN (SSGN) for mixed noise removal in HSI. Chang et al. [100] raised a multichannel 2-D CNN named HSI-DeNet for HSI denoising. HSI-DeNet directly learns the 2-D filters in each layer, via plenty of HSI clean-noisy samples. Based on the 2-D CNN framework, [101] presented the self-encoding denoising network (DeCS-Net). Maffei et al. [102] put forward a single-denoising CNN (HSI-SDeCNN), which uses 2-D CNN to exploit spectral–spatial correlation in HSI. Guan et al. [103] proposed an innovative deep recurrent CNN (DnRCNN) model. HSI striping is researched from the perspective of intraband and interband correlation exploration by a recursive CNN.

To sum up, 2-D CNN effectively removes HSI noise in both spatial and spectral dimensions via multichannel filters. The execution efficiency of 2-D CNN is usually high due to the end-to-end learning strategy and lightweight model [143]. However, 2-D CNN-based methods cannot integrally deal with 3-D spectral–spatial structure in HSI. Therefore, 3-D CNN-based methods for HSI denoising are developed as follows.

B. 3-D CNN

Compared with 2-D CNN, 3-D CNN can better deal with HSI [144]. It can simultaneously exploit the 3-D spectral–spatial structure in HSI, which benefits mixed noise removal in HSI. 3-D CNN-based methods use 3-D kernel filters

$$\mathcal{F}_{w,h,b,c} = \mathcal{X} \otimes \mathcal{C}_{w,h,b,c} \tag{18}$$

where $\mathbf{C}_{w,h,c}$ refers to the 3-D filter with the size of $w \times h \times b$ and channel of c. Without separating spatial and spectral information, 3-D CNN embodies the natural advantage for HSI denoising.

For instance, Liu and Lee [104] put forward the 3-D atrous denoising CNN (3D-ADCNN) model. 3DADCNN employs 3-D atrous filters to expand the receptive field, which captures both spatial and spectral features in HSI. It is flexible and robust. Based on the 3-D U-Net structure, Dong et al. [105] raised a 3-D HSI denoising CNN (3D-DCNN). It decomposes the 3-D filters as 2-D spatial filters and 1-D spectral filters, which effectively reduces the computing complexity. 3D-DCNN takes the transfer learning strategy to train the network parameter, using RGB images as the training samples, rather than HSIs. Later, Shi et al. [106] developed the 3-D attention network (3D-ADNet) for HSI restoration. 3D-ADNet introduces a self-attention unit into both the spatial and spectral directions. It can extract joint spectral–spatial features. Pang et al. [107] presented a new deep neural network, TRQ3DNet, which combines a CNN and converters for HSI denoising.

To sum up, 3D-CNN can simultaneously exploit the 3-D spectral–spatial structure in HSI, which benefits mixed noise removal in HSI. However, the 3-D network parameters are superabundant and hard to train.

C. Hybrid Networks

Except for 2-D CNN or 3-D CNN, several hybrid networks have also been proposed for HSI denoising through integrating different networks or modules. For instance, Zhong and Wang [145] developed a hybrid conditional random field (CRF) (for denoising) and multinomial logistic regression (MLR) (for classification) model. Wei et al. [108] structured a 3-D quasi-recurrent neural network (QRNN3D). QRNN3D simultaneously takes the global correlation along spectrum and spatiospectral correlation into account. It combines the 3-D CNN with a recurrent neural network, to conveniently solve different noisy HSI with arbitrary spectrum numbers. Cao et al. [109] raised a deep spatial–spectral global reasoning network (DSSGRN). DSSGRN embeds spectral module and spatial module into the U-Net architecture, to reason the relevant features in HSI. Fu et al. [110] combined deep external learning with internal learning (DELIL) for coded HSI reconstruction. Later, [111] developed a hyperspectral deep mixed network (HyperMixNet) for HSI restoration. HyperMixNet uses multiple mixed spatial and spectral convolutional modules, to excavate the spatial and spectral correlation and redundancy in HSI. Zhuang and Ng [112] introduced a fast and parameter-free hyperspectral image mixed noise removal method (termed FastHyMix). Chen et al. [113] proposed an efficient end-to-end transformer, named HSI denoising transformer (Hider), for mixed HSI noise removal.

D. Unsupervised Networks

The above-mentioned supervised HSI denoising methods need plenty of noise-free HSIs for network training and optimizing. However, noise-free HSIs are actually rare in practical environments. To address this inevitable issue, these unsupervised networks carry out the HSI denoising task without noise-free HSIs as training samples. In recent years, with the development of unsupervised methods for natural image denoising [146], [147], unsupervised methods for hyperspectral image denoising have also come into view [148].

For instance, Sidorov and Yngve Hardeberg [114] developed a deep hyperspectral prior (DHP) for HSI denoising and inpainting, to utilize intrinsic characteristics of 3-D CNN without any training. It uses the method of randomly starting parameter gradient descent to optimize the function and obtain the global optimal solution. Later, Imamura et al. [115] relied on a self-supervised training framework, which designed the zero-shot separable image prior (SIP). SIP uses the separable convolutional filter with depthwise and pointwise convolutional layers, to enclose the statistical property of HSI. It uses the Adam optimizer to train the network to obtain the global optimal solution. Based on unsupervised CNN, Nguyen et al. [116] proposed Stein's unbiased risk estimate CNN (SURE-CNN) for HSI denoising. SURE-CNN modifies the loss function with unknown noise variance, incorporating a dimensionality reduction step via SVD. It uses Adam optimization to train the network to optimize parameters. In addition, [117] introduced an unsupervised spatiospectral deep image prior (S2DIP). S2DIP learns the CNN model without any training data. It not only uses the ADMM algorithm to update CNN parameters but also employs one gradient descent step by using Adam in each update. Miao et al. [149] proposed an unsupervised DIP framework based on HSI classical spatiospectral decomposition. This method adopts two types of unsupervised DIPs: U-Net class networks and fully connected networks. They are used to model abundance maps and endmembers contained in HSI.

To sum up, the unsupervised network performs HSI denoising without noise-free HSI as a training sample to solve the inevitable problem of less noise-free HSI data, and the unsupervised network has lower complexity, so it is easy to obtain data. However, the intrinsic prior structure of HSI (sparseness and low-rankness) is ignored.

Data-driven HSI denoising methods usually perform more efficiently than model-driven methods, without carefully adjusting key parameters. Nevertheless, the supervised-based HSI denoising framework still needs a large number of clean HSIs as the label samples. While the clean HSIs are actually rare and hardly obtained in most imaging environments, except for this limitation, the simulated noise distribution for clean HSI samples is usually hard to agree with the actual noise distribution in noisy HSIs, because of the complicated imaging mechanism and degrading procedure. This issue greatly affects the generality of the data-driven methods, especially for HSI blind noise removal.

V. Model-Data-Driven Methods

As mentioned earlier, there are pros and cons, for both model-driven and data-driven methods. Could we conjointly combine the merits of model-driven with data-driven methods, and overcome the disadvantages of the two methods? Therefore, with the development of model-data-driven methods for natural image denoising [150], several model-data-driven methods have been developed [118], [119], [120], [121], to address both the advantages and disadvantages of model-driven and data-driven methods for HSI denoising.

For example, Lin et al. [118] coupled nonnegative matrix factorization with deep prior regularization (NMF-DPR) method for HSI denoising. NMF-DPR introduces the linear spectral mixing model. Then, the spatial prior is denoised by a single 2-D CNN with a proximal alternating linearized minimization method. Zhang et al. [119] developed the deep spatiospectral Bayesian posterior (DSSBP) model. DSSBP carries out noise estimation and distribution via the Bayesian framework. It combines three convolutional subnetworks with this Bayesian framework to learn the trainable parameters. Through estimating three posterior terms, DSSBP integrates model-driven into data-driven strategy. Xiong et al. [120] raised a subspace-based multidimensional sparse network (SMDS-Net) for noise removal in HSI. SMDS-Net simultaneously couples spectral low-rankness and spatial multidimensional sparsity into an unfolded convolutional network. In addition, Zhang et al. [121] proposed the unified low-rank spatiospectral network (LR-Net) for mixed noise reduction in HSI. LR-Net integrates the low-rankness of HSI into the convolutional network. It could effectively deal with random noise, impulse noise, and stripe noise within HSI.

To sum up, model-data-driven methods avoid the problems of time-consuming parameter adjustment and lack of training samples. However, the generalization ability is not stable for different noisy HSIs.

VI. Experiments and Analysis

In this part, the simulated and real-HSI denoising experiments are carried out to further manifest the performance of representative model-driven, data-driven, and model-data-driven methods. Later, HSI classification results via different HSI denoising methods are given. Finally, an efficiency analysis is described in real experiments, to further verify the practicability and reliability of different HSI denoising methods. Detailed descriptions are depicted as follows.

A. Simulated Experiments

In the simulated experiments, two noise-free HSIs are employed as the simulated data: Washington D.C. Mall HSI (outdoor data) and CAVE Toy HSI (indoor data). The W. DC Mall data and CAVE Toy data are separately cropped with the size of $200 \times 200 \times 191$ and the size of $512 \times 512 \times 31$ in the simulated experiments. The noise types of these HSIs include Gaussian noise, stripe noise, pepper noise, deadline noise, and mixed noise.

To validate the scene adaptation under multiple noisy environments, we simulate four different cases for both the W.

TABLE II
QUANTITATIVE EVALUATION INDEXES IN THE SIMULATED EXPERIMENTS (W. DC MALL DATASET)

Case	Index	Noisy	BM4D	LRMR	LRTV	NMoG	NGMeet	HSID	SSGN	DSSBP
Case 1	MPSNR	21.38	29.72	31.85	32.16	33.98	**34.27**	33.86	33.91	34.21
	MSSIM	0.752	0.948	0.969	0.972	0.983	0.986	0.981	0.982	**0.988**
	MSA	19.64	5.976	5.423	5.383	4.006	**3.842**	4.025	4.013	3.896
Case 2	MPSNR	20.85	27.69	29.42	29.66	30.45	31.68	26.74	28.86	**32.94**
	MSSIM	0.736	0.948	0.966	0.968	0.974	0.978	0.946	0.965	**0.982**
	MSA	18.82	5.969	4.913	4.875	4.786	4.652	6.890	5.381	**4.032**
Case 3	MPSNR	19.64	27.83	26.85	27.42	28.15	30.86	25.98	28.42	**31.16**
	MSSIM	0.718	0.951	0.946	0.949	0.953	0.971	0.934	0.958	**0.975**
	MSA	19.95	5.872	6.833	6.780	5.468	5.192	7.021	5.636	**4.289**
Case 4	MPSNR	17.84	25.94	25.37	26.48	27.03	**29.79**	24.16	27.68	29.45
	MSSIM	0.663	0.932	0.926	0.935	0.941	**0.958**	0.918	0.947	0.952
	MSA	21.08	6.997	7.048	6.924	6.187	5.614	7.653	6.059	**5.042**

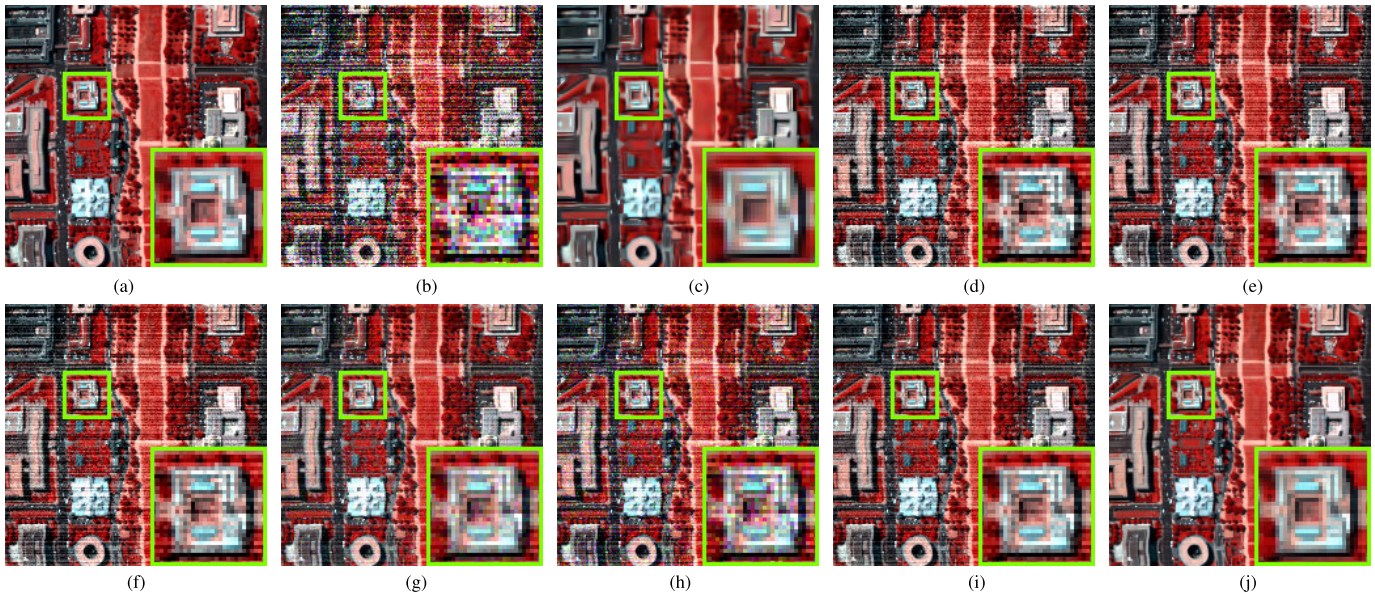


Fig. 6. HSI denoising results for bands (57, 27, and 17) of W. DC Mall HSI dataset in Case 3. (a) Clean (MPSNR/MSSIM). (b) Noisy (19.64/0.718). (c) BM4D (27.83/0.951). (d) LRMR (26.85/0.946). (e) LRTV (27.42/0.949). (f) NMoG (28.65/0.953). (g) NGMeet (30.86/0.971). (h) HSID-CNN (25.98/0.934). (i) SSGN (28.42/0.958). (j) DSSBP (31.16/0.975).

DC Mall and CAVE Toy HSI datasets. The added types of noise are Gaussian noise, stripe noise, and impulse noise. The specific simulated operations of these four cases are described as follows.

Case 1 (i.i.d. Gaussian Noise): Every band in these two HSIs is contaminated by i.i.d. Gaussian noise. For diverse bands, the variance σ_i of simulated Gaussian noise is equal for each other ($\sigma_i = 30$).

Case 2 (Non-i.i.d. Gaussian Noise): Every band in these two HSIs is contaminated by non-i.i.d. Gaussian noise. For diverse bands, the variance of simulated Gaussian noise is unequal for each other ($\sigma_i \in [0, 75]$). And it obeys the stochastic probability distribution.

Case 3 (Non-i.i.d. Gaussian + Stripe Noise): Based on Case 2, additive stripe noise is simulated in these two HSIs, which are simultaneously contaminated by non-i.i.d. noise in Case 2.

Case 4 (Non-i.i.d. Gaussian + Stripe + Impulse Noise): Based on Case 3, the additive impulse noise is simulated in these two HSIs, which are simultaneously contaminated by non-i.i.d. Gaussian and stripe noise in Case 3.

Besides, eight representative HSI denoising algorithms are regarded as the comparison methods. These algorithms include model-driven methods: BM4D [19], LRMR [39], LRTV [55], NMoG [61], and NGMeet [80]. Data-driven methods: HSID-CNN [98] and SSGN [99]. And model-data-driven methods: DSSBP [119].

With respect to the quantitative evaluation index for HSI denoising, we employ three indexes in the simulated experiments: mean peak-signal-to-noise ratio (MPSNR), mean structural similarity index measurement (MSSIM), and mean spectral angle measurement (MSA). MPSNR and MSSIM are utilized for evaluating the spatial recovery degree. MSAM is used for verifying the spectral perseveration degree. Generally,

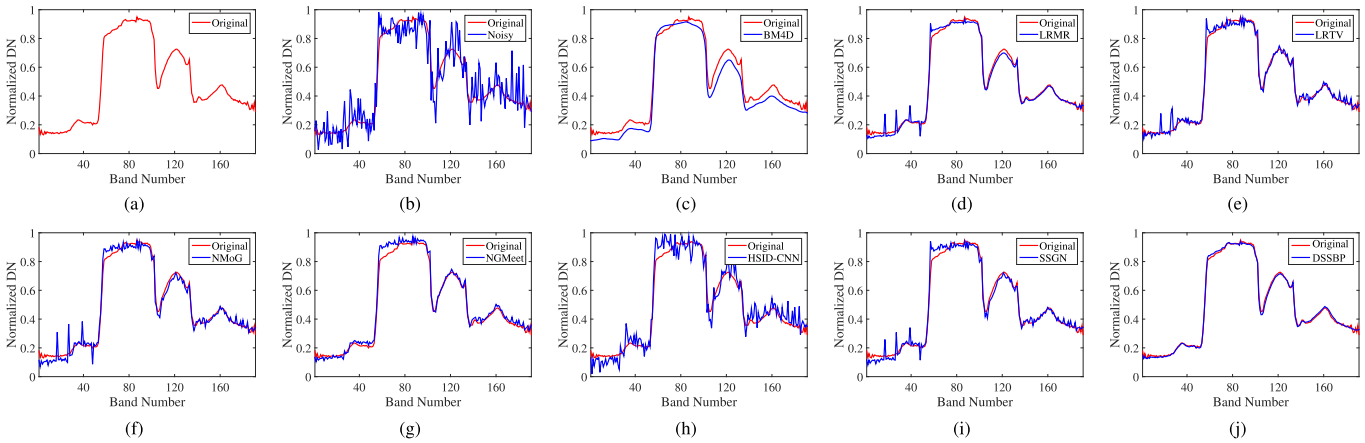


Fig. 7. Spectral curves for position (78 and 75) of W. DC Mall HSI dataset in Case 2. (a) Clean (MSA). (b) Noisy (18.82). (c) BM4D (5.969). (d) LRMR (4.913). (e) LRTV (4.875). (f) NMoG (4.786). (g) NGMeet (4.652). (h) HSID-CNN (6.890). (i) SSGN (5.381). (j) DSSBP (4.032).

TABLE III
QUANTITATIVE EVALUATION INDEXES IN THE SIMULATED EXPERIMENTS (CAVE TOY DATASET)

Case	Index	Noisy	BM4D	LRMR	LRTV	NMoG	NGMeet	HSID	SSGN	DSSBP
Case 1	MPSNR	25.66	32.76	32.08	33.69	34.17	**38.56**	33.81	34.79	38.32
	MSSIM	0.831	0.948	0.942	0.953	0.956	0.982	0.954	0.962	**0.983**
	MSA	16.29	7.059	7.637	6.901	6.878	**4.379**	7.122	6.058	4.659
Case 2	MPSNR	21.48	32.42	28.76	30.69	30.87	33.24	27.56	30.21	**34.67**
	MSSIM	0.796	0.946	0.914	0.935	0.939	0.951	0.863	0.938	**0.959**
	MSA	23.65	8.374	14.56	10.88	10.62	7.278	16.49	10.46	**5.831**
Case 3	MPSNR	19.94	29.25	26.83	29.94	30.06	**32.34**	25.71	30.66	31.87
	MSSIM	0.732	0.914	0.887	0.912	0.919	0.945	0.842	0.929	**0.954**
	MSA	25.14	12.75	15.06	11.68	11.27	6.428	17.36	9.147	**6.258**
Case 4	MPSNR	17.84	25.94	25.37	26.48	27.03	**29.79**	24.16	27.68	29.36
	MSSIM	0.663	0.932	0.926	0.935	0.941	**0.958**	0.918	0.947	0.943
	MSA	21.08	12.94	7.048	6.924	7.497	**6.614**	7.653	7.059	6.873

the higher the MPSNR/MSSIM index and the lower the MSA index is, the better the quality of HSI restoration is. As listed in Tables II and III, MPSNR, MSSIM, and MSA for each algorithm are given in the two simulated HSI datasets and four cases, respectively. The optimal index for each row is marked as a bold format in Tables II and III.

1) W. DC Mall Dataset: Table II displays the three objective evaluation indexes (MPSNR, MSSIM, and MSA) of eight contrast algorithms, under the four noisy cases for the W. DC Mall HSI dataset. Besides, the pseudocolor denoising results of bands (57, 27, and 17) in Case 3 are presented in Fig. 6.

In addition, the spectral curves of position (78 and 75) in the W. DC HSI dataset are also depicted in Fig. 7. For better comparisons, the original noise-free spectral curve is given in each denoising result through eight contrast methods.

As shown in Fig. 6(c), BM4D can generally eliminate the mixed noise in enlarged regions. LRMR, LRTV, NMoG, and NGMeet could effectively reduce random noise, as shown in Fig. 6(d)–(g). Nevertheless, residual stripe noise is obvious in the denoising results. In terms of the data-driven methods, HSID-CNN and SSGN perform well for HSI denoising, while the stripe noise and spectral distortion are problematic in Figs. 6(h)–(i) and 7(h)–(i). Model-data-driven method DSSBP outperforms on mixed noise removal and spectral preservation.

2) CAVE Toy Dataset: Table III lists the three quantitative evaluation indexes (MPSNR, MSSIM, and MSA) of eight contrast algorithms, under the four noisy cases for the CAVE Toy HSI dataset. Besides, the denoising results of the 26th-band in Case 2 are presented in Fig. 8.

As displayed in Fig. 8(c), BM4D could generally weaken the random noise in enlarged regions. Nevertheless, the over-smoothing phenomenon is still prominent in the denoised HSI. LRMR, LRTV, NMoG, and NGMeet could effectively wipe out random noise, as shown in Fig. 8(d)–(g), whereas residual stripe noise is evident in the denoising results. In terms of the data-driven methods, HSID-CNN and SSGN perform well in Fig. 8(h)–(i), while the spectral distortion is problematic in Table III. DSSBP can simultaneously remove the random noise without obvious residual noise in Fig. 8(j). And the spatial details are also clearly recovered especially for the tiny textures.

B. Real Experiments

To further validate the practicability and generalization of the above-mentioned HSI denoising methods, four real noisy

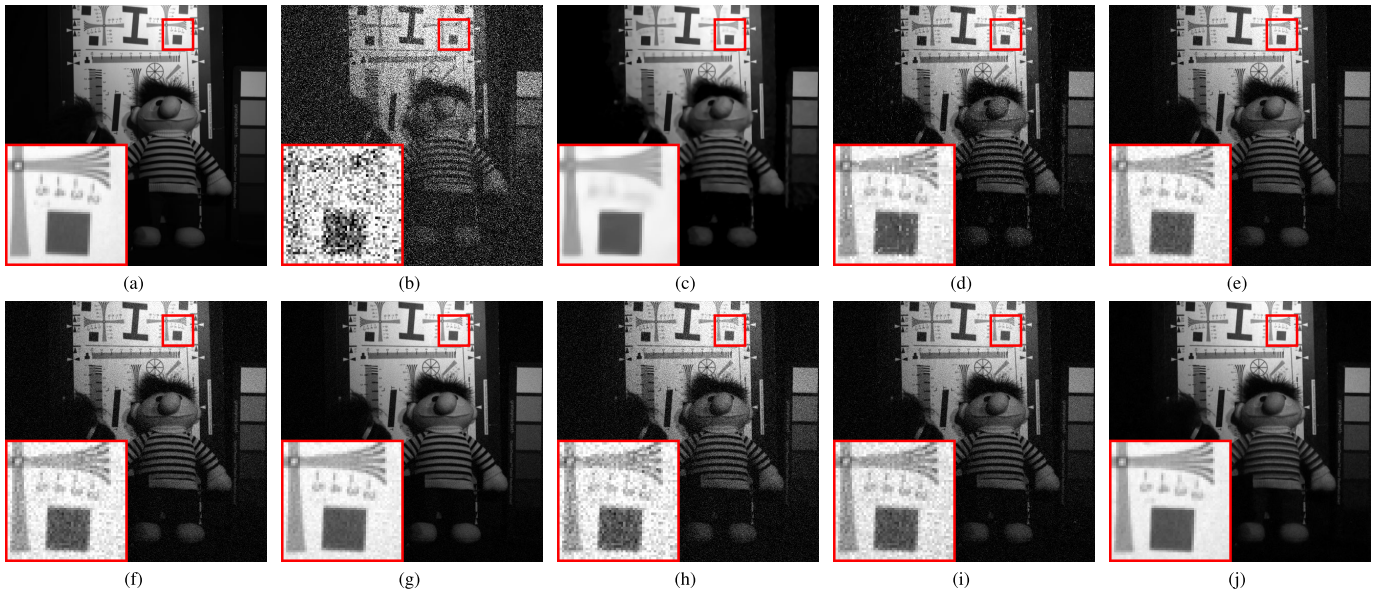


Fig. 8. HSI denoising results for band 26 of CAVE Toy HSI dataset in Case 2. (a) Clean (MPSNR/MSSIM). (b) Noisy (21.48/0.796). (c) BM4D (32.42/0.946). (d) LRMR (28.76/0.914). (e) LRTV (30.69/0.935). (f) NMoG (30.87/0.939). (g) NGMeet (33.24/0.951). (h) HSID-CNN (27.56/0.863). (i) SSGN (30.21/0.938). (j) DSSBP (34.67/0.959).

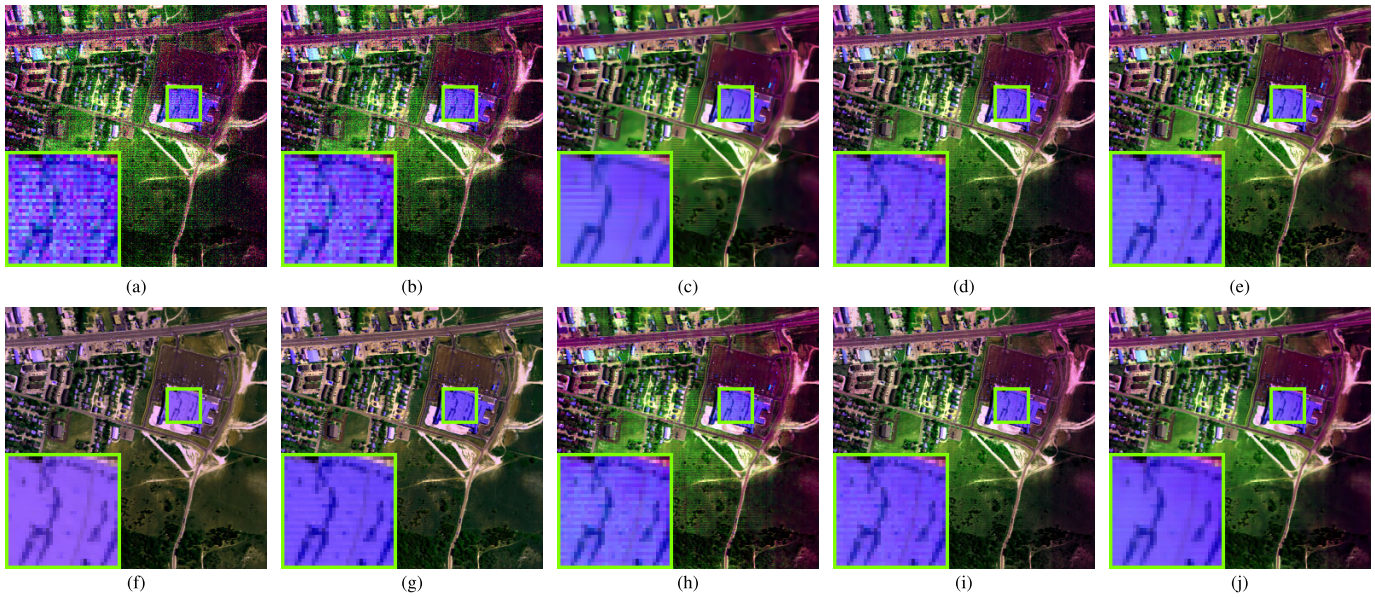


Fig. 9. HSI denoising results for bands (187, 104, and 24) of Urban HSI dataset in real experiments. (a) Noisy. (b) TDL. (c) BM4D. (d) LRMR. (e) LRTV. (f) NMoG. (g) NGMeet. (h) HSID-CNN. (i) SSGN. (j) DSSBP.

HSI datasets are utilized in our experiments: Urban ($307 \times 307 \times 188$), Indian Pines ($145 \times 145 \times 206$), GF-5 ($400 \times 400 \times 330$), and Zhuhai-1 ($400 \times 400 \times 32$). It should be highlighted that these noisy HSIs exist with a large diversity in spatial resolution, spectral resolution, spectral range, noise level, noise type, and scenario category. Generally speaking, the more bands the sensor has, the narrower the bandwidth is, and the higher the corresponding spectral resolution will be. Spatial resolution refers to the ability to display the smallest size of ground objects detected in remote sensing images. The larger the spatial resolution is, the better the spatial information can be captured. The spectral range of hyperspectral images generally includes a series of visible/near-infrared spectra, generally 400–1000 nm. For different HSIs, the spectral range is also different. The noise level of hyperspectral images is also different due to the influence of external factors, such as imager jitter and atmospheric interference. In addition, there are many kinds of noise, and the noise types of different hyperspectral images are also different. For example, HSIs contain Gaussian noise, strip noise, deadline noise, pulse noise, or mixed noise.

Nine HSI denoising algorithms: TDL, BM4D, LRMR, LRTV, NMoG, NGMeet, HSID-CNN, SSGN, and DSSBP are employed as contrast methods. For better distinguishing the restoration quality, enlarged figures for local details are also given in each result. Besides, the blind quality index Q-metrics of the four noisy HSIs are also recorded in Table IV. The optimal index is marked as the bold format in Table IV.

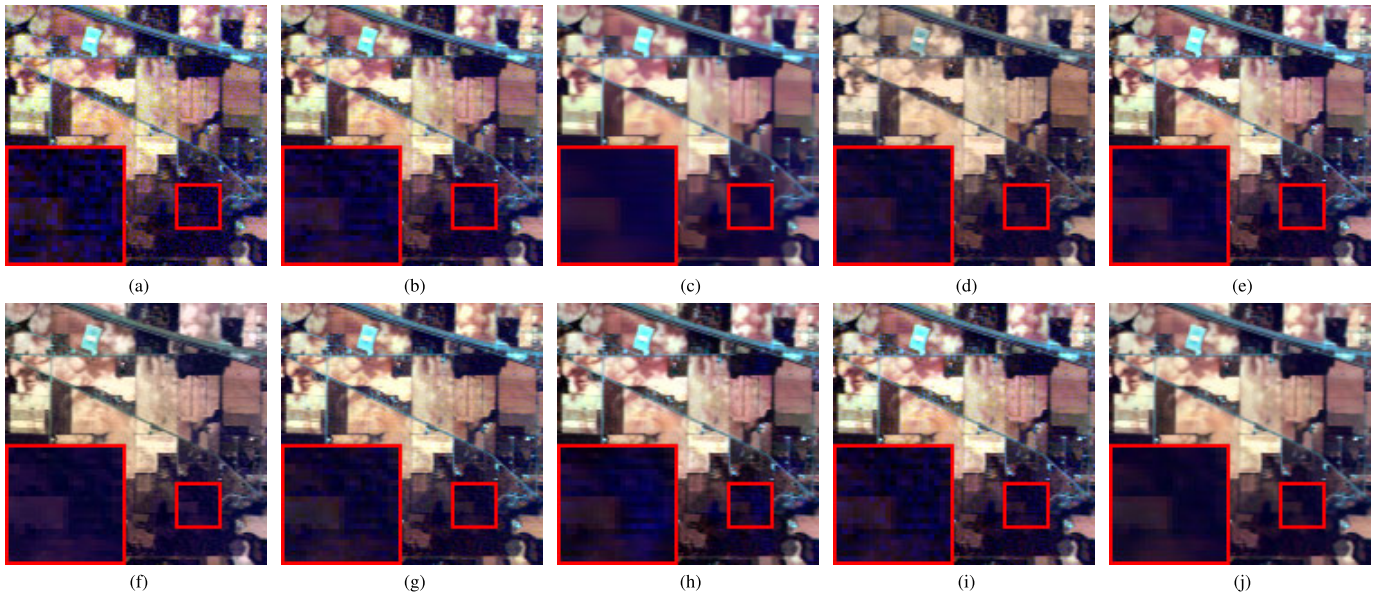


Fig. 10. HSI denoising results for bands (145, 24, and 2) of Indian Pines HSI dataset in real experiments. (a) Noisy. (b) TDL. (c) BM4D. (d) LRMR. (e) LRTV. (f) NMoG. (g) NGMeet. (h) HSID-CNN. (i) SSGN. (j) DSSBP.

The concrete HSI denoising results and analysis for each real dataset are described as follows.

1) Urban Dataset: As shown in Fig. 9, the HSI denoising results for bands (187, 104, and 24) of the Urban HSI dataset are listed through nine HSI denoising methods. This dataset is polluted by mixed noise, including random noise and stripe noise in Fig. 9(a).

TDL performs poorly for mixed noise removal in Fig. 9(b). BM4D, LRMR, LRTV, and NGMeet exist obvious stripe noise in Fig. 9(c)–(g). NMoG can effectively denoise the stripe noise and random noise in Fig. 9(f), while it leads to obvious spectral distortion issue. In HSID-CNN and SSGN, the residual noise is still salient in Fig. 9(h) and (i). The model-data-driven method DSSBP outperforms the Urban HSI dataset denoising, which could well deal with mixed noise in Urban HSI.

2) Indian Pines Dataset: As displayed in Fig. 10, HSI denoising results for bands (145, 24, and 2) of the Indian Pines HSI dataset are listed through nine contrast methods. This HSI is chiefly polluted with impulse noise.

BM4D could generally eliminate the impulse noise in enlarged regions in Fig. 10(c). Nevertheless, the over-smoothing phenomenon is still prominent in the denoised HSI. LRMR, LRTV, NMoG, NGMeet, HSID-CNN, and SSGN could effectively restrain polluted noise in Fig. 10(d)–(i). Model-data-driven method DSSBP can simultaneously remove the impulse noise without obvious residual noise in Fig. 10(j).

3) GF-5 Dataset: As depicted in Fig. 11, HSI denoising results for band 193 of GF-5 HSI are listed through nine contrast methods. This dataset is mainly polluted by random noise and stripe noise. Especially for the magnified regions, BM4D results in serious blurry effects and spatial detail missing issues. TDL, LRMR, LRTV, NMoG, NGMeet, HSID-CNN, SSGN, and DSSBP can effectively reduce mixed noise in HSI, as shown in Fig. 11(d)–(j). However, over-smoothing problems and residual stripe noise still exist in all these methods, to different degrees.

4) Zhuhai-1 Dataset: As shown in Fig. 12, HSI denoising results for band 32 of Zhuhai-1 HSI are listed through nine contrast methods: TDL, LRMR, LRTV, NMoG, NGMeet, HSID-CNN, SSGN, and DSSBP. This dataset is mainly polluted by random noise. Especially for the magnified regions, the model-data-driven method DSSBP outperforms on random noise removal and spatial details recovery. While BM4D results in serious blurry effects in Fig. 12(c). LRMR, LRTV, NMoG, NGMeet, HSID-CNN, and SSGN exist obvious residual noise or over-smoothing effects in Fig. 12(b)–(i). These HSI denoising results manifest that current methods are hard to guarantee the generalization ability for different HSIs.

C. Classification Results

Generally, HSI denoising is beneficial for classification, by removing the polluted noise in HSI. To ulteriorly validate the availability of different HSI restoration methods, classification results before and after HSI denoising are carried out in this section.

Based on the supervised SVM strategy, 16 ground-truth classes are utilized for verifying the classification precision, as portrayed in Fig. 13(a). Its training data include 10% of total data. The classification accuracy indexes (OA, AA, and kappa coefficient) are listed in Table V. These HSI classification results for Indian Pines data are displayed in Fig. 13(b)–(j). Compared with noisy HSI, different HSI denoising methods contribute to subsequent classification tasks. This also validates the effectiveness and meaning of HSI denoising. NGMeet achieves the highest classification results with 88.47% OA value, 89.95% AA value, and 0.8572 Kappa coefficient, as described in Table V.

D. Efficiency Analysis

To measure the reconstruction efficiency of the above-mentioned model-driven, data-driven, and model-data-driven

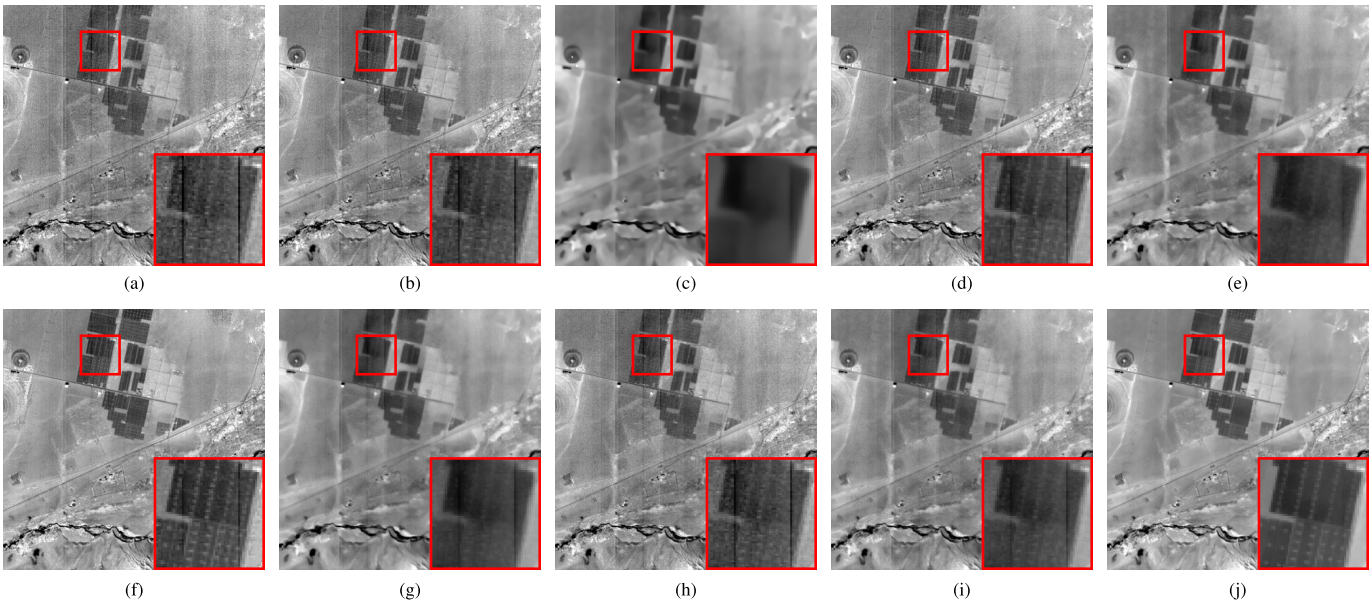


Fig. 11. HSI denoising results for band 193 of GF-5 HSI dataset in real experiments. (a) Noisy. (b) TDL. (c) BM4D. (d) LRMR. (e) LRTV. (f) NMoG. (g) NGMeet. (h) HSID-CNN. (i) SSGN. (j) DSSBP.

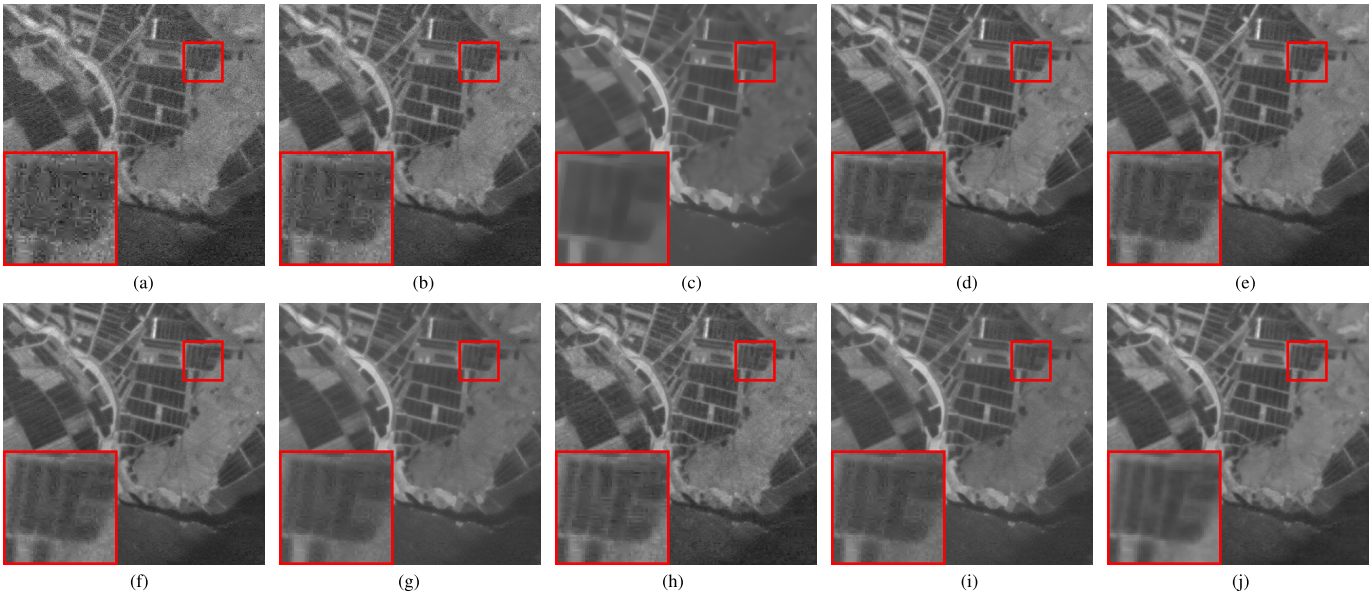


Fig. 12. HSI denoising results for band 32 of Zhuhai-1 HSI dataset in real experiments. (a) Noisy. (b) TDL. (c) BM4D. (d) LRMR. (e) LRTV. (f) NMoG. (g) NGMeet. (h) HSID-CNN. (i) SSGN. (j) DSSBP.

TABLE IV
QUANTITATIVE QUALITY EVALUATION OF Q-METRIC IN THE REAL EXPERIMENTS

HSI	TDL	BM4D	LRMR	LRTV	NMoG	NGMeet	HSID-CNN	SSGN	DSSBP
Urban	0.041	0.064	0.061	0.068	0.072	0.069	0.065	0.068	**0.074**
Indian	0.053	0.071	0.075	0.083	0.079	0.830	0.076	0.082	**0.085**
GF-5	0.038	0.042	0.052	0.045	0.060	0.048	0.046	0.053	**0.062**
Zhuhai-1	0.042	0.047	0.061	0.063	0.058	**0.065**	0.059	0.063	0.064

methods, an efficiency analysis is given in this section. These methods are executed with the same software and hardware environment (running software: MATLAB R2018a; language: Python 3.8; and RAM: 16 GB). Meanwhile, the running times of the above-mentioned methods in the real experiments are also recorded in Table VI. Compared with model-driven methods (BM4D, LRMR, LRTV, NMoG, and NGMeet), data-driven methods (HSID-CNN and SSGN) have

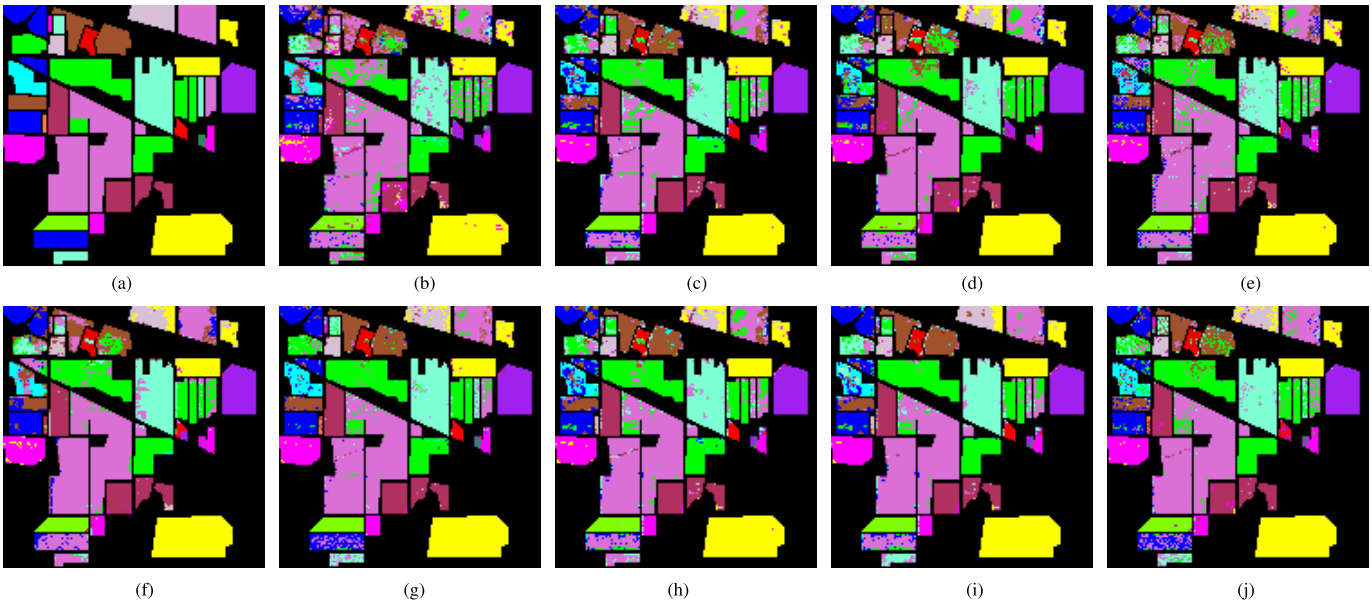


Fig. 13. HSI classification results for Indian Pines data before and after HSI denoising. (a) Ground truth. (b) Noisy. (c) BM4D. (d) LRMR. (e) LRTV. (f) NMoG. (g) NGMeet. (h) HSID-CNN. (i) SSGN. (j) DSSBP.

TABLE V
HSI CLASSIFICATION ACCURACY INDEXES FOR INDIAN PINES BEFORE AND AFTER HSI DENOISING

Index	Noisy	BM4D	LRMR	LRTV	NMoG	NGMeet	HSID-CNN	SSGN	DSSBP
OA	75.96%	83.97%	79.44%	80.59%	84.36%	**88.47%**	85.65%	86.01%	87.33%
AA	77.24%	86.19%	81.35%	83.42%	87.16%	**89.95%**	87.63%	87.82%	88.74%
Kappa	0.7220	0.8162	0.7641	0.7693	0.8124	**0.8572**	0.8338	0.8392	0.8449

TABLE VI
RUNNING-TIMES OF DIFFERENT HSI DENOISING METHODS IN THE REAL EXPERIMENTS (SECOND)

HSI	Noisy	BM4D	LRMR	LRTV	NMoG	NGMeet	HSID-CNN	SSGN	DSSBP
Urban	—	936.2	1155.6	789.4	1325.4	452.4	**8.7**	15.6	34.2
Indian	—	236.3	268.7	195.8	348.2	137.8	**2.6**	3.9	10.7
GF-5	—	1085.7	1723.5	1034.1	1297.6	893.9	**15.4**	21.4	48.6
Zhuhai-1	—	156.2	184.3	147.8	261.9	64.5	**2.2**	3.1	8.7

significant advantages in denoising the efficiency of HSI. Meanwhile, model-data-driven method DSSBP also performs well on execution efficiency for HSI denoising in Table VI. This profits from high-efficiency feature extraction via deep learning strategy [151].

VII. SUMMARY AND PROSPECTS

A. Summary

Mixed noise pollution in HSI severely disturbs subsequent interpretations and applications. In this review, we first describe the noise analysis in different noisy HSIs and conclude crucial points for programming HSI denoising algorithms. Then a general HSI restoration model is formulated for solving and optimization. Later, we technically review the existing HSI denoising methods, from model-driven strategy (nonlocal mean, TV, sparse representation, LRMA, and LRTF), data-driven strategy (2-D CNN, 3-D CNN, hybrid networks, and unsupervised networks), to model-data-driven strategy. The advantages and disadvantages of each strategy for HSI denoising are summarized and contrasted. Finally, we present a performance evaluation of the above-mentioned HSI denoising methods for various noisy HSIs in simulated and real experiments. The classification results of denoised HSIs and execution efficiency are depicted for these HSI denoising methods.

B. Prospects

Although existing HSI denoising methods achieve effective results, there are still problems, such as generalization ability, physical mechanism, and so on. We summarize possible prospects of HSI denoising methods as follows.

1) *Improving the Generalization Ability of HSI Denoising Methods:* Different noisy HSIs usually exist in large diversity. Therefore, it is meaningful and beneficial to consider the difference in self-information for universal HSI denoising. In addition, how to adaptively choose

the crucial parameters of HSI denoising methods is also significant.

2) *Integrating Model-Driven With Data-Driven Strategy More Effectively:* Besides, underlying features of HSI could be extracted via deep neural networks. Therefore, how to logically combine the model-driven and data-driven strategy is significant.
3) *Optimizing HSI Denoising Model via Self-Supervised or Unsupervised Learning:* Due to the number of noise-free HSIs is limited, therefore, exploiting a novel training strategy without plentiful labels is the necessary road for HSI denoising. The overfitting problem of supervised HSI denoising methods is also inevitable, which can be alleviated via self-supervised or unsupervised learning.
4) *Combining Low-Level Vision With High-Level Vision for HSI:* It should be highlighted that the purpose of HSI denoising is to improve subsequent interpretation. This strategy can further enhance the practicability and convenience of HSI applications, without multiple processing steps.

REFERENCES

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

[2] 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.

[3] R. N. Patro, S. Subudhi, P. K. Biswal, and F. Dell'acqua, "A review of unsupervised band selection techniques: Land cover classification for hyperspectral earth observation data," *IEEE Geosci. Remote Sens. Mag.*, vol. 9, no. 3, pp. 72–111, Sep. 2021.

[4] D. Hong et al., "Interpretable hyperspectral artificial intelligence: When nonconvex modeling meets hyperspectral remote sensing," *IEEE Geosci. Remote Sens. Mag.*, vol. 9, no. 2, pp. 52–87, Jun. 2021.

[5] C. Wu, J. Li, R. Song, Y. Li, and Q. Du, "HPRN: Holistic prior-embedded relation network for spectral super-resolution," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Apr. 3, 2023, doi: 10.1109/TNNLS.2023.3260828.

[6] B. Rasti, Y. Chang, E. Dalsasso, L. Denis, and P. Ghamisi, "Image restoration for remote sensing: Overview and toolbox," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 2, pp. 201–230, Jun. 2022.

[7] Q. Zhang, Q. Yuan, J. Li, Y. Wang, F. Sun, and L. Zhang, "Generating seamless global daily AMSR2 soil moisture (SGD-SM) long-term products for the years 2013–2019," *Earth Syst. Sci. Data*, vol. 13, no. 3, pp. 1385–1401, Mar. 2021.

[8] L. Bruzzone and C. Persello, "A novel approach to the selection of spatially invariant features for the classification of hyperspectral images with improved generalization capability," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 9, pp. 3180–3191, Sep. 2009.

[9] Q. Zhang, Q. Yuan, J. Li, Z. Yang, and X. Ma, "Learning a dilated residual network for SAR image despeckling," *Remote Sens.*, vol. 10, no. 2, p. 196, Jan. 2018.

[10] Q. Zhang, Q. Yuan, J. Li, Z. Li, H. Shen, and L. Zhang, "Thick cloud and cloud shadow removal in multitemporal imagery using progressively spatio-temporal patch group deep learning," *ISPRS J. Photogramm. Remote Sens.*, vol. 162, pp. 148–160, Apr. 2020.

[11] S. Han and J. P. Kerekes, "Overview of passive optical multispectral and hyperspectral image simulation techniques," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 11, pp. 4794–4804, Nov. 2017.

[12] J. P. Kerekes and J. E. Baum, "Hyperspectral imaging system modeling," *Lincoln Lab. J.*, vol. 14, no. 1, pp. 117–130, 2003.

[13] Y. Xiao, Q. Yuan, K. Jiang, J. He, Y. Wang, and L. Zhang, "From degrade to upgrade: Learning a self-supervised degradation guided adaptive network for blind remote sensing image super-resolution," *Inf. Fusion*, vol. 96, pp. 1566–2535, Aug. 2023.

[14] Y. Yuan, X. Zheng, and X. Lu, "Spectral–spatial kernel regularized for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3815–3832, Jul. 2015.

[15] Z. Xie, P. Duan, W. Liu, X. Kang, X. Wei, and S. Li, "Feature consistency-based prototype network for open-set hyperspectral image classification," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jan. 6, 2023, doi: 10.1109/TNNLS.2022.3232225.

[16] D. Letexier and S. Bourennane, "Noise removal from hyperspectral images by multidimensional filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 7, pp. 2061–2069, Jul. 2008.

[17] B. Rasti, P. Scheunders, P. Ghamisi, G. Licciardi, and J. Chanussot, "Noise reduction in hyperspectral imagery: Overview and application," *Remote Sens.*, vol. 10, no. 3, p. 482, Mar. 2018.

[18] L.-R. Gao, B. Zhang, X. Zhang, W.-J. Zhang, and Q.-X. Tong, "A new operational method for estimating noise in hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 1, pp. 83–87, Jan. 2008.

[19] 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.

[20] Y. Qian, Y. Shen, M. Ye, and Q. Wang, "3-D nonlocal means filter with noise estimation for hyperspectral imagery denoising," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2012, pp. 1345–1348.

[21] Q. Yuan, L. Zhang, and H. Shen, "Hyperspectral image denoising employing a spectral–spatial adaptive total variation model," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 3660–3677, Oct. 2012.

[22] J. Li, Q. Yuan, H. Shen, and L. Zhang, "Hyperspectral image recovery employing a multidimensional nonlocal total variation model," *Signal Process.*, vol. 111, pp. 230–248, Jun. 2015.

[23] Y. Chang, L. Yan, H. Fang, and C. Luo, "Anisotropic spectral–spatial total variation model for multispectral remote sensing image destriping," *IEEE Trans. Image Process.*, vol. 24, no. 6, pp. 1852–1866, Jun. 2015.

[24] 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.

[25] S. Takeyama, S. Ono, and I. Kumazawa, "Hyperspectral image restoration by hybrid spatio-spectral total variation," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2017, pp. 4586–4590.

[26] L. Sun, B. Jeon, Y. Zheng, and Z. Wu, "A novel weighted cross total variation method for hyperspectral image mixed denoising," *IEEE Access*, vol. 5, pp. 27172–27188, 2017.

[27] J. Peng, Q. Xie, Q. Zhao, Y. Wang, L. Yee, and D. Meng, "Enhanced 3DTV regularization and its applications on HSI denoising and compressed sensing," *IEEE Trans. Image Process.*, vol. 29, pp. 7889–7903, 2020.

[28] H. Othman and S.-E. Qian, "Noise reduction of hyperspectral imagery using hybrid spatial–spectral derivative-domain wavelet shrinkage," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp. 397–408, Feb. 2006.

[29] G. Chen and S.-E. Qian, "Denoising and dimensionality reduction of hyperspectral imagery using wavelet packets, neighbour shrinking and principal component analysis," *Int. J. Remote Sens.*, vol. 30, no. 18, pp. 4889–4895, Aug. 2009.

[30] G. Chen and S.-E. Qian, "Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 3, pp. 973–980, Mar. 2011.

[31] B. Rasti, J. R. Sveinsson, and M. O. Ulfarsson, "Wavelet-based sparse reduced-rank regression for hyperspectral image restoration," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6688–6698, Oct. 2014.

[32] Z. Xing, M. Zhou, A. Castrodad, G. Sapiro, and L. Carin, "Dictionary learning for noisy and incomplete hyperspectral images," *SIAM J. Imag. Sci.*, vol. 5, no. 1, pp. 33–56, Jan. 2012.

[33] Y. Qian and M. Ye, "Hyperspectral imagery restoration using nonlocal spectral–spatial structured sparse representation with noise estimation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 499–515, Apr. 2013.

[34] M. Ye, Y. Qian, and J. Zhou, "Multitask sparse nonnegative matrix factorization for joint spectral–spatial hyperspectral imagery denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2621–2639, May 2015.

[35] Y.-Q. Zhao and J. Yang, "Hyperspectral image denoising via sparse representation and low-rank constraint," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 296–308, Jan. 2015.

[36] T. Lu, S. Li, L. Fang, Y. Ma, and J. A. Benediktsson, "Spectral–spatial adaptive sparse representation for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 373–385, Jan. 2016.

[37] J. Li, Q. Yuan, H. Shen, and L. Zhang, "Noise removal from hyperspectral image with joint spectral–spatial distributed sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 9, pp. 5425–5439, Sep. 2016.

[38] L. Zhuang and J. M. Bioucas-Dias, "Fast hyperspectral image denoising and inpainting based on low-rank and sparse representations," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 730–742, Mar. 2018.

[39] 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.

[40] W. He, H. Zhang, L. Zhang, and H. Shen, "Hyperspectral image denoising via noise-adjusted iterative low-rank matrix approximation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 3050–3061, Jun. 2015.

[41] M. Wang, J. Yu, J. Xue, and W. Sun, "Denoising of hyperspectral images using group low-rank representation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 9, pp. 4420–4427, Sep. 2016.

[42] Y. Chen, Y. Guo, Y. Wang, D. Wang, C. Peng, and G. He, "Denoising of hyperspectral images using nonconvex low rank matrix approximation," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 5366–5380, Sep. 2017.

[43] F. Xu, Y. Chen, C. Peng, Y. Wang, X. Liu, and G. He, "Denoising of hyperspectral image using low-rank matrix factorization," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 7, pp. 1141–1145, Jul. 2017.

[44] L. Sun, B. Jeon, Y. Zheng, and Z. Wu, "Hyperspectral image restoration using low-rank representation on spectral difference image," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 7, pp. 1151–1155, Jul. 2017.

[45] H. Fan, J. Li, Q. Yuan, X. Liu, and M. Ng, "Hyperspectral image denoising with bilinear low rank matrix factorization," *Signal Process.*, vol. 163, pp. 132–152, Oct. 2019.

[46] J. Xue, Y. Zhao, W. Liao, and S. G. Kong, "Joint spatial and spectral low-rank regularization for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 1940–1958, Apr. 2018.

[47] H. Zeng, X. Xie, W. Kong, S. Cui, and J. Ning, "Hyperspectral image denoising via combined non-local self-similarity and local low-rank regularization," *IEEE Access*, vol. 8, pp. 50190–50208, 2020.

[48] T. Xie, S. Li, and B. Sun, "Hyperspectral images denoising via nonconvex regularized low-rank and sparse matrix decomposition," *IEEE Trans. Image Process.*, vol. 29, pp. 44–56, 2020.

[49] F. Fan, Y. Ma, C. Li, X. Mei, J. Huang, and J. Ma, "Hyperspectral image denoising with superpixel segmentation and low-rank representation," *Inf. Sci.*, vols. 397–398, pp. 48–68, Aug. 2017.

[50] L. Sun, B. Jeon, B. N. Soomro, Y. Zheng, Z. Wu, and L. Xiao, "Fast superpixel based subspace low rank learning method for hyperspectral denoising," *IEEE Access*, vol. 6, pp. 12031–12043, 2018.

[51] Y. Chen, T.-Z. Huang, X.-L. Zhao, and L.-J. Deng, "Hyperspectral image restoration using framelet-regularized low-rank nonnegative matrix factorization," *Appl. Math. Model.*, vol. 63, pp. 128–147, Nov. 2018.

[52] L. Gao, D. Yao, Q. Li, L. Zhuang, B. Zhang, and J. Bioucas-Dias, "A new low-rank representation based hyperspectral image denoising method for mineral mapping," *Remote Sens.*, vol. 9, no. 11, p. 1145, Nov. 2017.

[53] W. Wei, L. Zhang, Y. Jiao, C. Tian, C. Wang, and Y. Zhang, "Intracluster structured low-rank matrix analysis method for hyperspectral denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 866–880, Feb. 2019.

[54] S. Liu, J. Feng, and Z. Tian, "Variational low-rank matrix factorization with multi-patch collaborative learning for hyperspectral imagery mixed denoising," *Remote Sens.*, vol. 13, no. 6, p. 1101, Mar. 2021.

[55] 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.

[56] W. He, H. Zhang, H. Shen, and L. Zhang, "Hyperspectral image denoising using local low-rank matrix recovery and global spatial–spectral total variation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 713–729, Mar. 2018.

[57] Q. Wang, Z. Wu, J. Jin, T. Wang, and Y. Shen, "Low rank constraint and spatial spectral total variation for hyperspectral image mixed denoising," *Signal Process.*, vol. 142, pp. 11–26, Jan. 2018.

[58] H. Liu, P. Sun, Q. Du, Z. Wu, and Z. Wei, "Hyperspectral image restoration based on low-rank recovery with a local neighborhood weighted spectral–spatial total variation model," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 3, pp. 1409–1422, Mar. 2019.

[59] Y. Chen, J. Li, and Y. Zhou, "Hyperspectral image denoising by total variation-regularized bilinear factorization," *Signal Process.*, vol. 174, Sep. 2020, Art. no. 107645.

[60] Y. Yang, J. Zheng, S. Chen, and M. Zhang, "Hyperspectral image restoration via local low-rank matrix recovery and Moreau-enhanced total variation," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 6, pp. 1037–1041, Jun. 2020.

[61] Y. Chen, X. Cao, Q. Zhao, D. Meng, and Z. Xu, "Denoising hyperspectral image with non-i.I.d. noise structure," *IEEE Trans. Cybern.*, vol. 48, no. 3, pp. 1054–1066, Mar. 2018.

[62] B. Du, Z. Huang, and N. Wang, "A bandwise noise model combined with low-rank matrix factorization for hyperspectral image denoising," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1070–1081, Apr. 2018.

[63] X. Lu, Y. Wang, and Y. Yuan, "Graph-regularized low-rank representation for destriping of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 4009–4018, Jul. 2013.

[64] H. Zhang, J. Cai, W. He, H. Shen, and L. Zhang, "Double low-rank matrix decomposition for hyperspectral image denoising and destriping," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5502619.

[65] F. Yang, X. Chen, and L. Chai, "Hyperspectral image destriping and denoising using stripe and spectral low-rank matrix recovery and global spatial–spectral total variation," *Remote Sens.*, vol. 13, no. 4, p. 827, Feb. 2021.

[66] Q. Xie et al., "Multispectral images denoising by intrinsic tensor sparsity regularization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1692–1700.

[67] J. Xue, Y. Zhao, W. Liao, and J. C. Chan, "Nonlocal low-rank regularized tensor decomposition for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 5174–5189, Jul. 2019.

[68] X. Guo, X. Huang, L. Zhang, and L. Zhang, "Hyperspectral image noise reduction based on rank-1 tensor decomposition," *ISPRS J. Photogramm. Remote Sens.*, vol. 83, pp. 50–63, Sep. 2013.

[69] N. Renard, S. Bourennane, and J. Blanc-Talon, "Denoising and dimensionality reduction using multilinear tools for hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 138–142, Apr. 2008.

[70] A. Karami, M. Yazdi, and A. Zolghadre Asli, "Noise reduction of hyperspectral images using kernel non-negative tucker decomposition," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 487–493, Jun. 2011.

[71] 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.*, Jun. 2014, pp. 2949–2956.

[72] 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. (CVPR)*, Jul. 2017, pp. 5901–5909.

[73] X. Bai, F. Xu, L. Zhou, Y. Xing, L. Bai, and J. Zhou, "Nonlocal similarity based nonnegative tucker decomposition for hyperspectral image denoising," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 701–712, Mar. 2018.

[74] Y. Wang, J. Peng, Q. Zhao, Y. Leung, X. Zhao, and D. Meng, "Hyperspectral image restoration via total variation regularized low-rank tensor decomposition," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1227–1243, Apr. 2018.

[75] Y. Chen, T. Huang, and X. Zhao, "Destriping of multispectral remote sensing image using low-rank tensor decomposition," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 12, pp. 4950–4967, Dec. 2018.

[76] H. Zhang, L. Liu, W. He, and L. Zhang, "Hyperspectral image denoising with total variation regularization and nonlocal low-rank tensor decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 5, pp. 3071–3084, May 2020.

[77] Y. Chen, W. He, N. Yokoya, and T. Huang, "Hyperspectral image restoration using weighted group sparsity-regularized low-rank tensor decomposition," *IEEE Trans. Cybern.*, vol. 50, no. 8, pp. 3556–3570, Aug. 2020.

[78] Y. Chang, L. Yan, X. Zhao, H. Fang, Z. Zhang, and S. Zhong, "Weighted low-rank tensor recovery for hyperspectral image restoration," *IEEE Trans. Cybern.*, vol. 50, no. 11, pp. 4558–4572, Nov. 2020.

[79] X. Gong, W. Chen, and J. Chen, "A low-rank tensor dictionary learning method for hyperspectral image denoising," *IEEE Trans. Signal Process.*, vol. 68, pp. 1168–1180, 2020.

[80] W. He et al., "Non-local meets global: An integrated paradigm for hyperspectral image restoration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 4, pp. 2089–2107, Apr. 2022.

[81] Y. Zheng, T. Huang, X. Zhao, Y. Chen, and W. He, "Double-factor-regularized low-rank tensor factorization for mixed noise removal in hyperspectral image," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 12, pp. 8450–8464, Dec. 2020.

[82] Y. Chen, T. Huang, W. He, X. Zhao, H. Zhang, and J. Zeng, "Hyperspectral image denoising using factor group sparsity-regularized nonconvex low-rank approximation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–16, 2021.

[83] H. Fan, Y. Chen, Y. Guo, H. Zhang, and G. Kuang, "Hyperspectral image restoration using low-rank tensor recovery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 10, pp. 4589–4604, Oct. 2017.

[84] H. Fan, C. Li, Y. Guo, G. Kuang, and J. Ma, "Spatial–spectral total variation regularized low-rank tensor decomposition for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 6196–6213, Oct. 2018.

[85] Z. Kong and X. Yang, "Color image and multispectral image denoising using block diagonal representation," *IEEE Trans. Image Process.*, vol. 28, no. 9, pp. 4247–4259, Sep. 2019.

[86] Y. Nie, L. Chen, H. Zhu, S. Du, T. Yue, and X. Cao, "Graph-regularized tensor robust principal component analysis for hyperspectral image denoising," *Appl. Opt.*, vol. 56, no. 22, pp. 6094–6102, 2017.

[87] X.-L. Zhao, H. Zhang, T.-X. Jiang, M. K. Ng, and X.-J. Zhang, "Fast algorithm with theoretical guarantees for constrained low-tubal-rank tensor recovery in hyperspectral images denoising," *Neurocomputing*, vol. 413, pp. 397–409, Nov. 2020.

[88] Y. Zheng, T. Huang, X. Zhao, T. Jiang, T. Ma, and T. Ji, "Mixed noise removal in hyperspectral image via low-fibered-rank regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 734–749, Jan. 2020.

[89] J. Lin, T. Huang, X. Zhao, T. Jiang, and L. Zhuang, "A tensor subspace representation-based method for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 9, pp. 7739–7757, Sep. 2021.

[90] J. Lin, T.-Z. Huang, X.-L. Zhao, T.-H. Ma, T.-X. Jiang, and Y.-B. Zheng, "A novel non-convex low-rank tensor approximation model for hyperspectral image restoration," *Appl. Math. Comput.*, vol. 408, Nov. 2021, Art. no. 126342.

[91] H. Zeng, X. Xie, H. Cui, H. Yin, and J. Ning, "Hyperspectral image restoration via global L_{1-2} spatial–spectral total variation regularized local low-rank tensor recovery," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 4, pp. 3309–3325, Apr. 2021.

[92] Y. Liu, X. Zhao, Y. Zheng, T. Ma, and H. Zhang, "Hyperspectral image restoration by tensor fibered rank constrained optimization and plug-and-play regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–17, 2021.

[93] 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.

[94] M. Wang, Q. Wang, and J. Chanussot, "Tensor low-rank constraint and l_0 total variation for hyperspectral image mixed noise removal," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 3, pp. 718–733, Apr. 2021.

[95] Y. Chen, W. He, N. Yokoya, T. Huang, and X. Zhao, "Nonlocal tensor-ring decomposition for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 2, pp. 1348–1362, Feb. 2020.

[96] F. Xiong, J. Zhou, and Y. Qian, "Hyperspectral restoration via L_0 gradient regularized low-rank tensor factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 10410–10425, Dec. 2019.

[97] W. Xie, Y. Li, and X. Jia, "Deep convolutional networks with residual learning for accurate spectral–spatial denoising," *Neurocomputing*, vol. 312, pp. 372–381, Oct. 2018.

[98] Q. Yuan, Q. Zhang, J. Li, H. Shen, and L. Zhang, "Hyperspectral image denoising employing a spatial–spectral deep residual convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 1205–1218, Feb. 2019.

[99] Q. Zhang, Q. Yuan, J. Li, X. Liu, H. Shen, and L. Zhang, "Hybrid noise removal in hyperspectral imagery with a spatial–spectral gradient network," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 7317–7329, Oct. 2019.

[100] Y. Chang, L. Yan, H. Fang, S. Zhong, and W. Liao, "HSI-DeNet: Hyperspectral image restoration via convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 667–682, Feb. 2019.

[101] X. Liu, S. Mei, Z. Zhang, Y. Zhang, J. Ji, and Q. Du, "Decs-Net: Convolutional self-encoding network for hyperspectral image denoising," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 1951–1954.

[102] A. Maffei, J. M. Haut, M. E. Paoletti, J. Plaza, L. Bruzzone, and A. Plaza, "A single model CNN for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2516–2529, Apr. 2020.

[103] J. Guan, R. Lai, H. Li, Y. Yang, and L. Gu, "DnRCNN: Deep recurrent convolutional neural network for HSI destriping," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jan. 31, 2022, doi: 10.1109/TNNLS.2022.3142425.

[104] W. Liu and J. Lee, "A 3-D atrous convolution neural network for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5701–5715, Aug. 2019.

[105] W. Dong, H. Wang, F. Wu, G. Shi, and X. Li, "Deep spatial–spectral representation learning for hyperspectral image denoising," *IEEE Trans. Comput. Imag.*, vol. 5, no. 4, pp. 635–648, Dec. 2019.

[106] Q. Shi, X. Tang, T. Yang, R. Liu, and L. Zhang, "Hyperspectral image denoising using a 3-D attention denoising network," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 12, pp. 10348–10363, Dec. 2021.

[107] L. Pang, W. Gu, and X. Cao, "TRQ3DNet: A 3D quasi-recurrent and transformer based network for hyperspectral image denoising," *Remote Sens.*, vol. 14, no. 18, p. 4598, Sep. 2022.

[108] K. Wei, Y. Fu, and H. Huang, "3-D quasi-recurrent neural network for hyperspectral image denoising," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 363–375, Jan. 2021.

[109] X. Cao, X. Fu, C. Xu, and D. Meng, "Deep spatial–spectral global reasoning network for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–14, 2021.

[110] Y. Fu, T. Zhang, L. Wang, and H. Huang, "Coded hyperspectral image reconstruction using deep external and internal learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 7, pp. 3404–3420, Jul. 2022.

[111] K. Yorimoto and X. Han, "HyperMixNet: Hyperspectral image reconstruction with deep mixed network from a snapshot measurement," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2021, pp. 1184–1193.

[112] L. Zhuang and M. K. Ng, "FastHyMix: Fast and parameter-free hyperspectral image mixed noise removal," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Sep. 29, 2021, doi: 10.1109/TNNLS.2021.3112577.

[113] H. Chen, G. Yang, and H. Zhang, "Hider: A hyperspectral image denoising transformer with spatial–spectral constraints for hybrid noise removal," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Oct. 31, 2022, doi: 10.1109/TNNLS.2022.3215751.

[114] O. Sidorov and J. Y. Hardeberg, "Deep hyperspectral prior: Single-image denoising, inpainting, super-resolution," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019, pp. 3844–3851.

[115] R. Imamura, T. Itasaka, and M. Okuda, "Zero-shot hyperspectral image denoising with separable image prior," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019, pp. 1416–1420.

[116] H. V. Nguyen, M. O. Ulfarsson, and J. R. Sveinsson, "Hyperspectral image denoising using SURE-based unsupervised convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 4, pp. 3369–3382, Apr. 2021.

[117] Y. Luo, X. Zhao, T. Jiang, Y. Zheng, and Y. Chang, "Hyperspectral mixed noise removal via spatial–spectral constrained unsupervised deep image prior," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 9435–9449, 2021.

[118] B. Lin, X. Tao, and J. Lu, "Hyperspectral image denoising via matrix factorization and deep prior regularization," *IEEE Trans. Image Process.*, vol. 29, pp. 565–578, 2020.

[119] Q. Zhang, Q. Yuan, J. Li, F. Sun, and L. Zhang, "Deep spatio-spectral Bayesian posterior for hyperspectral image non-i.I.d. noise removal," *ISPRS J. Photogramm. Remote Sens.*, vol. 164, pp. 125–137, Jun. 2020.

[120] F. Xiong, J. Zhou, S. Tao, J. Lu, J. Zhou, and Y. Qian, "SMDS-Net: Model guided spectral–spatial network for hyperspectral image denoising," *IEEE Trans. Image Process.*, vol. 31, pp. 5469–5483, 2021.

[121] H. Zhang, H. Chen, G. Yang, and L. Zhang, "LR-Net: Low-rank spatial–spectral network for hyperspectral image denoising," *IEEE Trans. Image Process.*, vol. 30, pp. 8743–8758, 2021.

[122] A. A. Efros and T. K. Leung, "Texture synthesis by non-parametric sampling," in *Proc. 7th IEEE Int. Conf. Comput. Vis.*, Jan. 1999, pp. 1033–1038.

[123] B. Zhao, M. O. Ulfarsson, J. R. Sveinsson, and J. Chanussot, "Hyperspectral image denoising using spectral–spatial transform-based sparse and low-rank representations," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5522125.

[124] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, Nov. 1992.

[125] J. Peng et al., "Low-rank and sparse representation for hyperspectral image processing: A review," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 1, pp. 2–35, Jun. 2021.

[126] T. Xie, S. Li, and J. Lai, "Adaptive rank and structured sparsity corrections for hyperspectral image restoration," *IEEE Trans. Cybern.*, vol. 52, no. 9, pp. 8729–8740, Sep. 2022.

[127] H. Ye, H. Li, B. Yang, F. Cao, and Y. Tang, "A novel rank approximation method for mixture noise removal of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4457–4469, Jul. 2019.

[128] C. Li, Y. Ma, J. Huang, X. Mei, and J. Ma, "Hyperspectral image denoising using the robust low-rank tensor recovery," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 32, no. 9, pp. 1604–1612, Sep. 2015.

[129] X. Kong, Y. Zhao, J. Xue, and J. C.-W. Chan, "Hyperspectral image denoising using global weighted tensor norm minimum and nonlocal low-rank approximation," *Remote Sens.*, vol. 11, no. 19, p. 2281, Sep. 2019.

[130] X. Liu, S. Bourennane, and C. Fossati, "Denoising of hyperspectral images using the PARAFAC model and statistical performance analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 3717–3724, Oct. 2012.

[131] L. Sun and C. He, "Hyperspectral image mixed denoising using difference continuity-regularized nonlocal tensor subspace low-rank learning," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[132] Z. Huang, S. Li, L. Fang, H. Li, and J. A. Benediktsson, "Hyperspectral image denoising with group sparse and low-rank tensor decomposition," *IEEE Access*, vol. 6, pp. 1380–1390, 2018.

[133] X. X. Zhu et al., "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 8–36, Dec. 2017.

[134] R. Dian, S. Li, and X. Kang, "Regularizing hyperspectral and multispectral image fusion by CNN denoiser," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 3, pp. 1124–1135, Mar. 2021.

[135] J. Hu, T. Huang, L. Deng, T. Jiang, G. Vivone, and J. Chanussot, "Hyperspectral image super-resolution via deep spatiospectral attention convolutional neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 12, pp. 7251–7265, Dec. 2022.

[136] S. Anwar, N. Barnes, and L. Petersson, "Attention-based real image restoration," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Dec. 13, 2021, doi: 10.1109/TNNLS.2021.3131739.

[137] W. Xie, Y. Li, J. Hu, and D.-Y. Chen, "Trainable spectral difference learning with spatial starting for hyperspectral image denoising," *Neural Netw.*, vol. 108, pp. 272–286, Dec. 2018.

[138] Q. Zhang, Y. Dong, Q. Yuan, M. Song, and H. Yu, "Combined deep priors with low-rank tensor factorization for hyperspectral image restoration," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, pp. 1–5, Jan. 2023.

[139] Q. Zhang, Q. Yuan, C. Zeng, X. Li, and Y. Wei, "Missing data reconstruction in remote sensing image with a unified spatial–temporal–spectral deep convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4274–4288, Aug. 2018.

[140] Q. Zhang, Q. Yuan, Z. Li, F. Sun, and L. Zhang, "Combined deep prior with low-rank tensor SVD for thick cloud removal in multitemporal images," *ISPRS J. Photogramm. Remote Sens.*, vol. 177, pp. 161–173, Jul. 2021.

[141] C. Lin, Y. Liu, C. Chi, C. Hsu, H. Ren, and T. Q. S. Quek, "Hyperspectral tensor completion using low-rank modeling and convex functional analysis," *IEEE Trans. Neural Netw. Learn. Syst.*, pp. 1–15, 2023.

[142] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.

[143] Y. Xiao, X. Su, Q. Yuan, D. Liu, H. Shen, and L. Zhang, "Satellite video super-resolution via multiscale deformable convolution alignment and temporal grouping projection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5610819.

[144] S. Zhang, L. Wang, L. Zhang, and H. Huang, "Learning tensor low-rank prior for hyperspectral image reconstruction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 12001–12010.

[145] P. Zhong and R. Wang, "Jointly learning the hybrid CRF and MLR model for simultaneous denoising and classification of hyperspectral imagery," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 7, pp. 1319–1334, Jul. 2014.

[146] A. Majumdar, "Blind denoising autoencoder," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 1, pp. 312–317, Jan. 2019.

[147] A. Creswell and A. A. Bharath, "Denoising adversarial autoencoders," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 4, pp. 968–984, Apr. 2019.

[148] Q. Zhang, Q. Yuan, M. Song, H. Yu, and L. Zhang, "Cooperated spectral low-rankness prior and deep spatial prior for HSI unsupervised denoising," *IEEE Trans. Image Process.*, vol. 31, pp. 6356–6368, Oct. 2022.

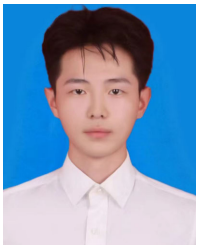
[149] Y. Miao, X. Zhao, X. Fu, J. Wang, and Y. Zheng, "Hyperspectral denoising using unsupervised disentangled spatiospectral deep priors," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5513916.

[150] F. Fang, J. Li, Y. Yuan, T. Zeng, and G. Zhang, "Multilevel edge features guided network for image denoising," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 9, pp. 3956–3970, Sep. 2021.

[151] Y. Xiao et al., "Space-time super-resolution for satellite video: A joint framework based on multi-scale spatial–temporal transformer," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 108, Apr. 2022, Art. no. 102731.

Qiang Zhang (Member, IEEE) received the B.E. degree in surveying and mapping engineering, and the M.E. and Ph.D. degrees in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2017, 2019, and 2022, respectively.

He is currently a Xinghai Associate Professor with the Center of Hyperspectral Imaging in Remote Sensing, College of Information Science and Technology, Dalian Maritime University, Dalian, China. He has authored or coauthored more than ten journal articles on Earth System Science Data (ESSD), ISPRS *Journal of Photogrammetry and Remote Sensing* (ISPRS P&RS), and IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS). His research interests include remote sensing information processing, computer vision, and machine learning. More details could be found at https://qzhang95.github.io.



Yaming Zheng received the B.S. degree from the School of Information Engineering (School of Software), Henan Animal Husbandry and Economics University, Zhengzhou, China, in 2022. He is currently pursuing the M.S. degree with the College of Information Science and Technology, Dalian Maritime University, Dalian, China.

His research interests include remote sensing information processing.

Qiangqiang Yuan (Member, IEEE) received the B.S. degree in surveying and mapping engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2006 and 2012, respectively.

In 2012, he joined the School of Geodesy and Geomatics, Wuhan University, where he is currently a Professor. He has authored or coauthored more than 90 research papers.

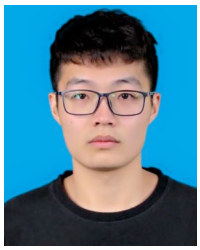
Prof. Yuan was a recipient of the Youth Talent Support Program of China in 2019. He was a recipient of the Best Reviewer Award of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS in 2019. He is an associate editor of five international journals. He has served as a referee for more than 40 journals.

Meiping Song (Member, IEEE) received the Ph.D. degree from the College of Computer Science and Technology, Harbin Engineering University, Harbin, China, in 2006.

She has been a Professor with the College of Information Science and Technology, Dalian Maritime University, Dalian, China, since 2020. Her research interests include remote sensing and hyperspectral image processing.

Haoyang Yu (Member, IEEE) received the B.S. degree in information and computing science from Northeastern University, Shenyang, China, in 2013, and the Ph.D. degree in cartography and geographic information system from the Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China, in 2019.

He is currently a Xing Hai Associate Professor with the Center of Hyperspectral Imaging in Remote Sensing, Information Science and Technology College, Dalian Maritime University, Dalian, China. His research interests include models and algorithms for hyperspectral image processing, analysis, and applications.



Yi Xiao received the B.S. degree from the School of Mathematics and Physics, China University of Geosciences, Wuhan, China, in 2020. He is currently pursuing the Ph.D. degree with the School of Geodesy and Geomatics, Wuhan University, Wuhan.

His research interests include remote sensing image super-resolution and computer vision.