Hyperspectral image destriping and denoising from a task decomposition view

Erting Pan [a], Yong Ma [a], Xiaoguang Mei [a], Jun Huang [a,*], Qihai Chen [b], Jiayi Ma [a]

[a] *Electronic Information School, Wuhan University, Wuhan 430072, China*
[b] *Tsinghua Innovation Center in Zhuhai, Zhuhai, Guangdong 519000, China*

ARTICLE INFO

ABSTRACT

The generalized mathematical model for HSI denoising or destriping lacks stability and uniqueness properties, failing to accurately portray the distribution and effects of stripes. Solutions following such a model would inevitably result in excessive destriping of strip-free areas, leading to the loss of texture detail. To remedy the above deficiencies, we reformulate the destriping task and introduce a novel solution from the task decomposition view. It is broken down into auxiliary sub-tasks involving stripe mask detection, stripe intensity estimation, and HSI restoration, which greatly reduces the difficulty of solving such an ill-posed problem. Based on this, we adopt a sequential multi-task learning framework and propose a stripes location-dependent restoration network, termed SLDR, which integrates the distribution and intensity features of stripes to achieve accurate destriping and high-fidelity restoration. Furthermore, we design a stripe attribute-aware estimator and a weighted total variation loss function to capture the unique properties of stripes and adaptively adjust the restoration weights of striped and non-striped regions. Extensive evaluation and comprehensive ablation studies on synthetic and practical scenes show the effectiveness and superiority of our model and architecture.

1. Introduction

Line-pattern stripe noise is a prevalent degradation in remote sensing images acquired from satellite-based and airborne imaging systems such as LANDSAT, and MODIS, especially more prone to arise in hyperspectral imaging systems, such as the Sentinel 2, EO-1 Hyperion imaging spectrometer, and the Chinese Gaofen-5 satellite [1–3]. As evident in Fig. 1, unlike random noise, stripe noise is majorly distributed along the scanning direction of the imaging system, has a certain width and extension length, and the overall grayscale value is brighter or darker than its normal neighboring lines. Typically, stripe noise results from the inconsistent response of the detection elements in the CCD, and accompanies the mainstream motion patterns of large-scale remote sensing imaging. It produces stripes or defect lines that share a fixed direction in the scene [4–6]. Particularly, imaging with whisk-broom or push-broom can evolve poor responses in the CCDs into horizontal or vertical stripes, respectively.

In general, stripes flow through the whole row or column of the remote sensing image, with a clear stripe edge different from the normal image, and the degradation intensity remains the same throughout. However, real-world stripes can be more complex, with random lengths or widths, drifting degradation intensity, and blurred edges. Observed

from Fig. 1, stripes in practical scenes can vary in distribution, intensity, and orientation. Beyond that, with its significant structural characteristics, widely or densely distributed stripes can both drown out the true radiance and natural texture features, or generate some artifacts related to the stripes. The resulting signal distortions and loss of details will lead to the failure of many subsequent interpretations and applications. In view of the fact that the striped pixels still contain radiance information, removing the stripes with high fidelity and restoring the true radiance information is, therefore, essential and of great significance to improve the potential value of remote sensing satellite data in various applications [7–10].

Since complex degradation not only greatly increases the difficulty of processing individual stripes, but also reduces the distribution regularity of stripes, it raises a higher demand for the destriping method to have universality. Fig. 2 shows three typical examples of stripe noise to mimic real-world complex cases. In concrete, the damage caused by stripe noise is apparently localized, and primarily the pixels around the stripes are suffering the most. Besides, the specific location of stripe noise may vary in bands of multi/hyperspectral remote sensing images due to small differences in imaging time. In addition, stripe noise may have various degrees of degradation across different land covers,

---

(a) Landsat 4 Multispectral Scanner   (b) Aqua MODIS (Band 27 and 36 from Pacific Ocean)

(c) Gaofen-5 (Band 150, 151, and 152 from Baoqing City)

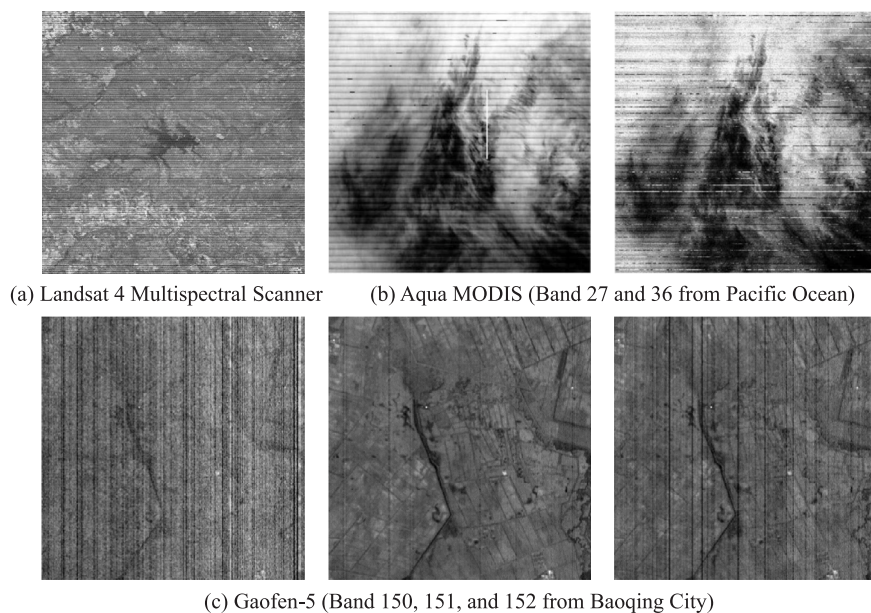**Fig. 1.** Examples of striping data collected from various satellite sensors.

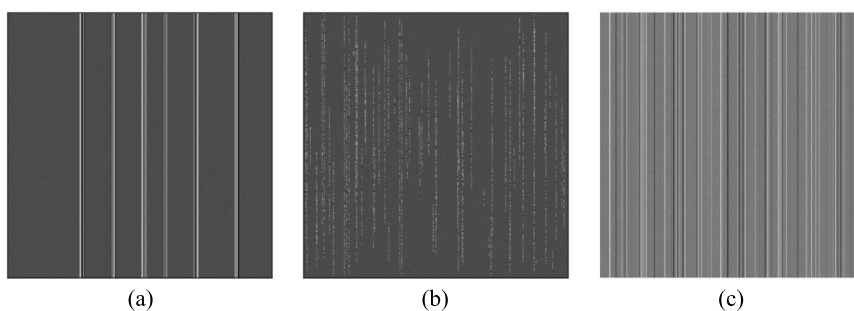(a)   (b)   (c)

Fig. 2. Examples of three typical complex stripes cases: (a) Stripes clusters; (b) Stripes in uncertain length; (c) Stripes mixed with random noise.

implying its nonlinear property. Hence, an effective destriping and denoising algorithm requires comprehensive analysis and consideration for complex stripes.

Over the past years, numerous scholars have investigated and developed a series of studies. Most existing solutions, based on traditional techniques involving statistical matching [11,12], digital filtering [4,6], and total variation [13,14], focus on modeling the image prior and are targeted to specific stripes, and these methods have limited applicability. Very recently, work on this topic has been driven increasingly by research on deep learning theory. Numerous studies have examined the data-driven mechanism and learning features in a supervised manner, achieving better results [15–17]. However, few studies have taken into account the potential roles of stripes distribution, and they ignore the relatively local corruption of stripes. Instead, they directly conduct average destriping [6,13], leading to excessive damage for stripe-free areas. These reasons result in their ill performance in recovering data with high fidelity.

In a nutshell, these methods achieve some degree of improvement, yet there still remains a legacy of unresolved issues. On the one hand, degradation caused by stripe noise can be complicated and diverse. The previously generalized mathematical model for additional noise $Y = X + N$ ignores some critical factors in real degraded images, such as differences in stripe distribution or orientation and image blur caused by stripes mixed with Gaussian noise. On the other hand, the corruption caused by stripes are apparently highly related to their location, only the stripe regions overlap with the background texture in the feature space. However, most methods destriping the whole

data directly, which leads to the loss of texture information in regions that are not contaminated by stripes, causing over-smoothing. Worse still, many existing methods are performed in a patchy manner over a limited receptive domain (limited spatial extent). Therefore, spatial background information is not available for larger regions, which has actually proven valuable for restoration.

To compensate for these limitations identified above in existing strategies, we endeavor to develop a novel image restoration model that can explicitly depict various stripe-based degradation in real scenes, including severe stripes and complex mixed noise, and then design an effective deep learning architecture based on it. Our ideas are as follows.

First, we consider pertinent to study inconsistent degradation of striped and non-striped regions. We begin by establishing a location-dependent mathematical destriping model. A stripe binary mask, where "1" indicates the presence of stripes and "0" otherwise, is integrated to model the location information of stripes. To deal with real noisy cases, our model also considers more complicated cases where stripes are mixed with random Gaussian noise. **Second**, following the reformulated location-dependent model, we decompose the destriping into several subtasks involving stripe mask detection, stripe intensity estimation, and clean data restoration. The detected stripe mask and the estimated stripe intensity can offer fresh insights to narrow the solution space for recovering degradation and enable the network to treat the striped and non-striped regions adaptively, achieving finer restoration. **Third**, with three subtasks of HSI destriping, a multi-task learning-based workflow, named stripes location-dependent restoration

network (SLDR), is organized. The highlight is that the three subtasks are sequentially chained up in the workflow, providing incremental support for the final HSI recovery. Although each of them has a specific objective function, the proposed method collaborates with them and trains the whole network with a joint loss function. The subsequent feature aggregator and decoder facilitate accurate stripe noise removal as well as high-quality HSI recovery. **Finally**, since the degradation caused by stripes are location-dependent, we design an adaptive weight term with a tailored bias for striped (more significant penalty) and non-striped (minor penalty) regions. It further lessens spatial details loss and alleviates the over-smooth phenomenon. Extensive experiments are conducted to demonstrate the superiority of our method on data with both synthesized and practical degradation. Particularly for some heavily degraded images, our method achieves considerably good results.

Our main contributions can be concluded in fourfold as follows:

- An effective solution based on task decomposition is proposed, which greatly reduces the difficulty of solving such an ill-posed problem.
- A stripes location-dependent restoration network for HSI restoration is presented, achieving accurate denoising and high-fidelity restoration.
- A stripe attribute-aware estimator is developed, capturing the unique properties of HSI stripes.
- An adaptive weight term with a tailored bias for the total variance penalty term is designed, weighting penalty to different degrees on striped and non-striped regions.

2. Related works

2.1. Hyperspectral image destriping

Early works solving such an ill-posed problem as HSI destriping mainly focus on finding appropriate priors. Some solutions heavily relied on statistical data properties, like histogram matching [11]. In contrast, some others have employed a variety of fundamental models, including frequency-domain transformation [4], low rank [18,19], sparse representation [20], and total variation [13], *etc*. These methods have alleviated this problem to some extent. However, it cannot be ignored that most of them concentrate on isolated research on specific stripes corruption of the given data and lack comprehensive and systematic analysis of different remote sensing images and various types of stripes noise.

Noteworthy, the recent literature is abundant in deep learning methods. For instance, Chang et al. [21] considered the destriping task of remote sensing images as an image decomposition problem and proposed a two-stream CNN. Based on this, HSI-DeNet [22] incorporates residual learning, dilation convolution and multi-channel filtering to improve the denoising performance on top of CNN. Zhong et al. [23] adopted a spaceborne-airborne data joint learning strategy to balance training samples and test samples and used deep CNN to fully describe the degradation of real HSIs in the spatial and spectral domains to restore the original radiance information better. Guan et al. [24] incorporated with the gated convolution strategy to capture the interband and inner band correlative features. Although methods based on deep learning have obvious advantages over traditional methods, such algorithms lack interpretability, rely on training samples, and have limited generalization capabilities.

On the basis of above limitations, we conclude that the existing methods are not accurate enough to deal with complex stripes, and their versatility is not ideal.

2.2. Hyperspectral image denoising

Scholars also have developed a series of approaches to denoise HSI, concentrating on modeling their typical properties such as nonlocal similarity [25,26], sparsity [27], and tensor structure [14,28]. For instance, Chang et al. [29] proposed a one-way low-rank tensor recovery method that additionally considers the non-local self-similarity. NMoG [30] defines a non-independent identically distributed (non-i.i.d.) mixture of the Gaussian to model the complicated HSI noise. However, these advanced model-driven methods are still limited by sacrificing computational efficiency to achieve high performance. Worse still, their performance also rely on manually optimized parameters to eliminate noise in fixed intensity or type, can hardly generalize to practical HSIs.

Very recently, more scholars have devoted to develop deep learning-based methods, which directly learn a nonlinear end-to-end mapping between noisy HSIs and clean HSIs. FastHyMix [31] exploits the low rankness and spatial correlation of HSIs by adding a deep image prior, which is extracted from a neural denoising network. Similarly, literature [32] proposes QRNN3D for deep denoising by introducing QRNN into the 3D U-net structure. T3SC [33] advocates a hybrid approach based on sparse coding principles to encode domain knowledge with handcrafted image priors in an unrolled optimization procedure. MAC-Net [34] takes the spectral low-rank model and spatial deep prior into account for HSI noise reduction. $\hat{N}$-Net [35] focus on intrinsic attributes of HSI noise distribution along the spectral dimension to achieve fine HSI restoration.

Although data-driven DL-based methods can recover the clean HSIs by training the networks on a large number of training data, they do not fully utilize the domain knowledge of HSIs such as the observation model and underlying characteristics.

2.3. Multi-task learning

Mainstream deep learning-based approaches aim to train a specific, reliable model for the objective function. However, features extracted by such single-task models tend to be valid only for that task and do not fully characterize the data. Once the task is closer to the complex real-world situation, it requires the model to learn more comprehensive and generalized features from the data. Multi-task learning (MTL) from machine learning theory supports learning multiple tasks simultaneously [36–38]. It introduces inductive bias by setting auxiliary tasks to improve the model generalization performance [39,40]. If tasks share complementary information, they act as regularizers for each other, which improves the prediction performance of each task [41]. In other words, MTL enhances the model generalization performance and prediction accuracy by leveraging the knowledge of all associated tasks and learning a shared representation across these tasks. It has been broadly used in many computer vision tasks, such as image classification [42], action recognition [43], anomalous event detection [44], *etc*.

Most MTL-based methods are designed based on existing CNN architectures, and the MTL is typically integrated with either hard or soft parameter sharing of hidden layers. For example, Cross-Stitch Networks [45] contain one standard feed-forward network per task, with cross-stitch units to allow features to be shared across tasks. The self-supervised approach of [46], based on the ResNet101 architecture, learns a regularized combination of features from different layers of a single shared network. The Multi-Task Attention Network [42] introduces a task-specific attention branch per task paired with the shared backbone. AdaShare [47] presents an adaptive sharing approach that decides what to share across which tasks to achieve the best recognition accuracy while considering resource efficiency.

In this work, to enhance the model generalization performance and achieve high-fidelity HSI restoration, we decompose the HSI destriping task into several subtasks and introduce the MTL-based scheme. Such a scheme not only encourages positive sharing among subtasks but also minimizes negative interference by using task-specific blocks.
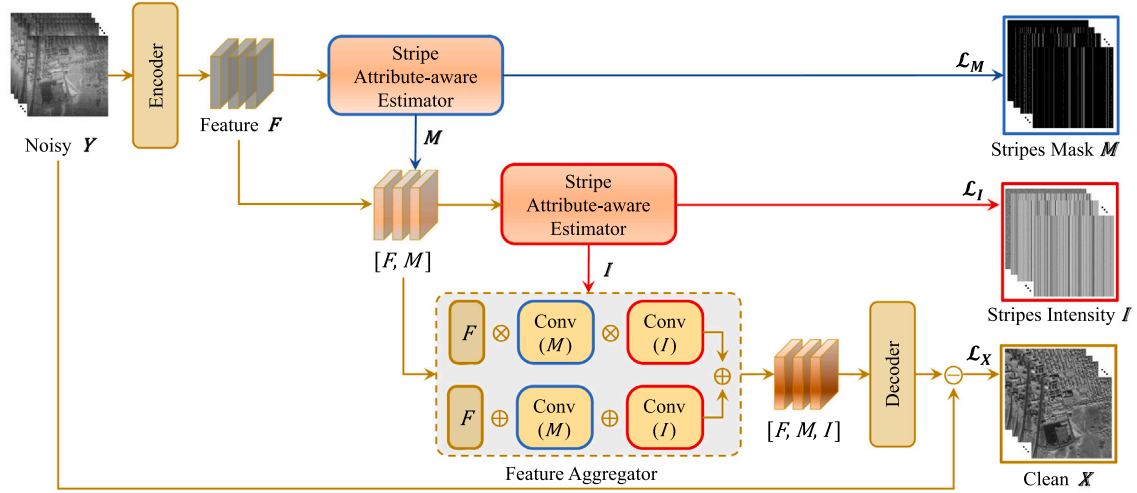
Fig. 3. Illustration of our proposed multi-task learning-based pipeline for HSI destriping. It decomposes the HSI destriping into three subtasks: stripe mask detection, stripe intensity estimation, and HSI restoration. Boxes with blue or red outlines are associated with M or I, respectively. Different colored outlines indicate that they employ blocks in the same structure without sharing parameters. Similarly, arrows in blue or red symbolize their corresponding information flow. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3. Methodology

3.1. Problem reformulation

Before elaborating on our method, we reformulate the task of HSI destriping, which aims to restore a clean HSI from a noisy observation. Mathematically, given a degraded HSI $Y \in \mathbb{R}^{b \times w \times h}$, the issue can be expressed as an inverse problem of the following widely used form:

$$X + S + D = Y, \tag{1}$$

where $S, D \in \mathbb{R}^{b \times w \times h}$ represents sparse stripes and random dense Gaussian noise which brings degradations to Y, b denotes the spectral dimension, w and h indicate the spatial size of HSI. Finding X from Eq. (1) is generally acknowledged to be an ill-posed problem. More unfortunately, the problem in the form Eq. (1) lacks stability and uniqueness properties, failing to portray the distribution and effects of stripes accurately. Solutions following such a model would result in an inevitable excessive denoising of strip-free areas, leading to the loss of texture detail.

To remedy the above deficiencies, we propose a novel location-dependent stripes model that enables to depict stripe location and stripe intensity separately as follows:

$$X + I \circ M + D = Y, \tag{2}$$

where $\circ$ symbolizes the pixel-wise multiplication, I denotes the stripe intensity, and M represents the binary stripe mask where "1" indicates the presence of stripes and "0" marks the stripe-free regions. With the introduced stripe intensity I and stripe mask M, the problem in the form Eq. (2) can be solved in a smaller solution space. It allows a different treatment between stripe and stripe-free regions, benefiting high-fidelity HSI restoration.

3.2. Task decomposition

According to HSI destriping task described as Eq. (2), given the observed noisy HSI Y, the desired solution is to estimate X, I, M. In other words, the HSI destriping task in Eq. (2) can be decomposed into stripes mask (M) detection, stripes intensity (I) estimation, and clean HSI (X) restoration. Such an inverse problem can be converted into an optimization problem based on a maximum-a-posteriori (MAP) estimation, which reads as:

$$arg \min_{X,I,M} \|X + I \circ M - Y\|_2^2 + P_m(M) + P_i(I) + P_x(X), \tag{3}$$

where the first term is regarded as a fidelity term, while $P_m(M)$, $P_i(I)$ and $P_x(X)$ act as regularization terms that encode prior information (*e.g.,* sparsity) on stripes mask, stripes intensity, and clean HSI, respectively.

In essence, the estimation of M, I, and X are highly associated. The high-fidelity restoration of X can derive benefits from accurately predicting M and I. Therefore, we incorporate the predictions of these three variables M, I, and X into a continuous and sequential process. As shown in Fig. 3, we decompose the HSI destriping task and construct a multi-task network that jointly and progressively performs HSI stripe noise detection, estimation and removal.

First, the noisy HSI observation Y is encoded by a structure encoder, where a residual dense block (termed RDB) is employed to initially explore structural feature F of the input. RDB [48] is proposed for image super-resolution with the merits of adaptively learning features from local and global via dense connection and hierarchical feature fusion. In this research, we utilize 3D convolution to customize the existing RDB into 3DRDB as structural feature encoder for HSI cube (refer to Fig. 4), which provides effective and reliable HSI structural feature F in the joint spectral and spatial domain. Hereby, on the basis of the extracted HSI structural feature F, we design an stripe attribute-aware estimator to explore intrinsic attributes of HSI stripes, and we utilize one for detecting predicting stripes mask M and another for estimating stripe intensities I. The concrete details will be introduced in Section 3.3. Next, we aggregate the extracted HSI structural feature F, the detected HSI stripe mask M, and the estimated stripe intensities I in the feature aggregation block according to Eq. (2). A decoder consisting of several 3D convolutional layers (see Fig. 4) is followed for further stripe noise estimation. Finally, the noisy input Y is integrated for the final noise removal and the clean HSI X reconstruction.

Additionally, we train and optimize the network by multiple loss functions, involving a specific weighted TV loss is designed for random dense Gaussian noise, *i.e.,* D in Eq. (2), which benefits in yielding a reliable and robust end-to-end HSI recovery model. The related description will be detailedly introduced in Section 3.4.

3.3. Stripe attribute-aware estimator

Intuitively, the degradation of HSI caused by stripe noise is region-dependent, demonstrating that accurate detection of damaged regions and estimation of their diverse severity facilitate identifying and removing stripe noises. Thus, the critical issue here turns to capture the
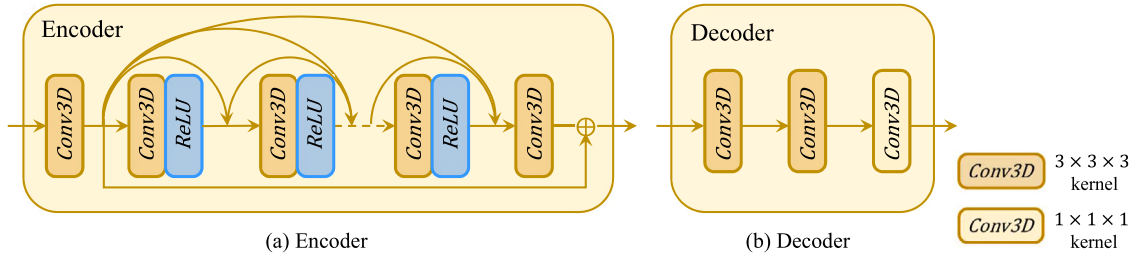
Fig. 4. Detailed architecture of encoder and decoder of the proposed method.
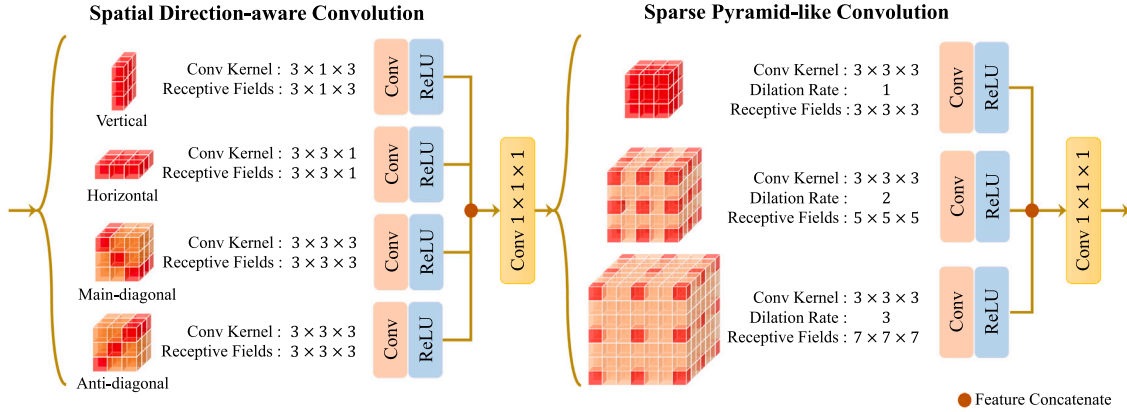

Fig. 5. The architecture of stripe attribute-aware estimator for estimating stripe mask M and stripe intensity I. It mainly consists of a spatial direction-aware convolution block and a sparse pyramid-like convolution block.

intrinsic attributes of HSI stripes, as mentioned above, unidirectionality and sparsity, which would directly contribute to predicting stripe noise mask M and stripe intensity I.

To this end, we propose a stripe attribute-aware estimator to learn the discriminative features of HSI stripes. It perceives and learns the special attributes of stripe noise through two main blocks: (1) a spatial direction-aware convolution block, which consisting of four specific direction-aware convolution kernels (including vertical, horizontal, main-diagonal, and anti-diagonal), empowers the model with the ability to robust handle stripe noise in different directions; (2) a sparse pyramid-like convolution block, comprising multiple 3D convolutional layers with different dilation rates and receptive fields, satisfies the attribute of stripe noise being sparsely distributed in both spatial and spectral domains.

Concretely, as shown in Fig. 5, the extracted HSI structural feature F is fed to the first block, *i.e.*, the spatial direction-aware convolution block, which contains four convolution paths with kernels in different directions. This block is tailored toward the unidirectional attribute of stripes. The vertical and horizontal convolution paths are simply with a kernel size of $3 \times 1 \times 3$ and $3 \times 3 \times 1$, which enables identifying HSI stripes caused by vertical push-broom and horizontal whisk-broom scanning. The main-diagonal and anti-diagonal convolution paths are implemented by initializing a convolution kernel of size $3 \times 3 \times 3$ and then filtering it with the corresponding identify matrix. Adding these two paths enables our model to deal with the oblique stripes caused by geometric correction. These four direction-aware convolution paths are concatenated to an additional $1 \times 1 \times 1$ convolution layer for further feature aggregation.

After that, considering the spatial–spectral sparsity of HSI stripes, we design a sparse pyramid-like convolution block, which consists of three 3D convolution paths with different dilated rates. The dilated convolution [49] updates the weight of pixels at dilation rate steps and thus enlarges their receptive fields without resolution sacrifice. At the same time, with larger dilation rates, the parameters of each convolution layer become more sparse, promoting the learning attribute of

sparsity in HSI stripes. As illustrated in Fig. 5, these three paths are initially setting with $3 \times 3 \times 3$ convolution and respectively dilated in rates $\{1, 2, 3\}$. And they have their receptive fields with the size of $3 \times 3 \times 3$, $5 \times 5 \times 5$, and $7 \times 7 \times 7$. Besides, a $1 \times 1 \times 1$ convolution layer is also followed to aggregate features from these three paths.

Particularly, it is worth mentioning that the designed stripe attribute-aware estimator is employed twice in our method. One is for detecting stripes mask M based on feature F, and another is for estimating stripes intensity I based on F and M. So, the two estimators do not share the parameters, and their output settings are different, too. The one for predicting binary mask M is set with $[2, b, w, h]$ and the other for estimating I with $[1, b, w, h]$.

3.4. Network training

Denote the inverse recovery functions modeled by the proposed network for stripes mask M, stripes intensity I, and stripes-free HSI X as $\mathcal{F}_M(\cdot), \mathcal{F}_I(\cdot), \mathcal{F}_X(\cdot)$, respectively. Symbolize all learned parameters of the network as Θ. We train the proposed multi-task-based network via end-to-end learning, optimizing by a combination of multiple loss functions involves mask loss $\mathcal{L}_M(\Theta)$, intensity loss $\mathcal{L}_I(\Theta)$, and recovery loss $\mathcal{L}_X(\Theta)$.

Since the value of stripe mask M is constrained as binary, we introduce cross-entropy loss function $\mathcal{L}_{CE}(\cdot)$ for $\mathcal{L}_M(\Theta)$ as below:

$$\mathcal{L}_M(\Theta) = \mathcal{L}_{CE}(\mathcal{F}_M(X; \Theta), (Y - X)). \tag{4}$$

Intensity loss $\mathcal{L}_I(\Theta)$ is formulated with $\mathcal{L}_{MAE}(\cdot)$ as:

$$\mathcal{L}_I(\Theta) = \mathcal{L}_{MAE}(\mathcal{F}_I(X; \Theta), Y - X). \tag{5}$$

In particular, considering the random Gaussian noise D that accompanies actual satellite imaging according to Eq. (3), we define the recovery loss $\mathcal{L}_X(\Theta)$ with two items as Eq. (6):

$$\mathcal{L}_X(\Theta) = \mathcal{L}_{MAE}(\mathcal{F}_X(X; \Theta), X) + \omega \cdot \mathcal{L}_{WTV}, \tag{6}$$

where one based on mean absolute error $\mathcal{L}_{\mathrm{MAE}}(\cdot)$ maintains fidelity while the other one $\mathcal{L}_{\mathrm{WTV}}$ designed on the basis of $\mathcal{L}_{\mathrm{TV}}(\cdot)$, which is specifically configured as a learnable term and capable of removing random Gaussian noise. ω is setting as $1e+2$ to adjust the weight of $\mathcal{L}_{\mathrm{WTV}}$. The $\mathcal{L}_{\mathrm{WTV}}$ is formulated as:

$$\mathcal{L}_{\mathrm{WTV}} = \arg\min \sum_i \sum_j W_{i,j} \sqrt{\sum_k [(\nabla_u \cdot X_{i,j,k})^2 + (\nabla_v \cdot X_{i,j,k})^2]}, \tag{7}$$

where i,j,k denotes the pixel-wise position in spatial and spectral dimensions. The highlight weight term $W_{i,j}$ can be expressed by the following equation:

$$W_{i,j} = \hat{M}_{i,j} + \epsilon, \tag{8}$$

where $W_{i,j}$ is updated according to the estimated $\hat{M}$. It indicates that stripe region can be further smoothing while ϵ in Eq. (8) ensures non-stripe region is processed with a slight smoothing, filtering the potential random Gaussian noise. $W_{i,j}$ empowers $\mathcal{L}_{\mathrm{WTV}}$ a learnable and versatile item.

We integrate $\mathcal{L}_{\mathrm{M}}(\Theta)$, $\mathcal{L}_{\mathrm{I}}(\Theta)$, and $\mathcal{L}_{\mathrm{X}}(\Theta)$ to train the network parameterized by Θ as Eq. (9), achieving jointly estimate M, I, and X based on noisy observation Y:

$$\mathcal{L}_{\mathrm{total}}(\Theta) = \lambda_1 \cdot \mathcal{L}_{\mathrm{M}}(\Theta) + \lambda_2 \cdot \mathcal{L}_{\mathrm{I}}(\Theta) + \lambda_3 \cdot \mathcal{L}_{\mathrm{X}}(\Theta), \tag{9}$$

where parameters λ_1, λ_2, and λ_3 are weighting factors for balanced training. To ensure balanced convergence, they are set to 1, 1, and $1e-2$, respectively, based on experiments. The network is trained to minimize the total loss $\mathcal{L}_{\mathrm{total}}(\Theta)$ by back-propagation.

4. Experiments

4.1. Experimental settings

4.1.1. Data

We conduct destriping experiments on data with synthetic stripe noises and real satellite HSIs. For training, datasets are from publicly available HSI dataset ICVL,[1] [50] where 120 out of 201 practically noise-free HSIs are sampled as the training set in our experiment, along with 5 for validation, the rest constitute the testing set. Each HSI cube in ICVL has a spatial resolution of 1392×1300 and a spectral dimension of 31. We also conduct synthetic mixture noises experiments on the publicly available HYDICE image of Washington D.C. Mall,[2] which has a resolution of $1280 \times 307 \times 191$. To further demonstrate the generalization ability of our model, we also test real HSIs collected from satellites, involving the EO1 Hyperion dataset[3] with 166 bands, the GF-5 Baoqing dataset and the GF-5 Wuhan dataset[4] with 155 bands acquired by the Advanced Hyperspectral Imager in GF-5 satellite [19].

4.1.2. Synthetic noise settings

We mimic a series of synthetic stripe noisy cases involving (1) random stripes; (2) random stripe clusters, which means that several adjacent stripes have similar intensities; (3) random stripes in uncertain length, where some broken stripes may possess arbitrary length; (4) a mixture of noises, where stripes coexists with non-i.i.d Gaussian noise. In addition, we configure some associated settings to maximize the imitation of practical noise contamination. Specifically, two-thirds of bands, 5% to 45% of columns or rows in the first three cases, and one-third of bands, 5% to 35% of columns or rows for the last case, are randomly selected to add stripe noise. Plus, the non-i.i.d. Gaussian noise in the last case is added to one-third of the bands with a random intensity ranging in the interval $[10, 70]$.

4.1.3. Training details

The training patches have a spatial size of 64×64 and the complete spectral dimension 31. To improve the robustness of training, we scale the patches with a rate in $\{1, 0.5, 0.25\}$ and randomly rotate them to further diversify the training set. Besides, 80 epochs are set up throughout the training process. The batch size (*i.e.* 16) is employed to stabilize the training. Parameters are initialized with Kaiming Initialization and updated by Adam Optimizer. The learning rate is set to 0.001 and decreases exponentially with epochs until the validation performance does not increase anymore. Noteworthy, all deep learning-based methods run on the same GPU RTX 3090Ti, while other methods run on the same CPU i9-10940H.

4.1.4. Evaluation metrics

We employ five mainstream quantitative metrics on the denoised results to measure the denoising performance, inclusive of mean peak signal-to-noise ratio (MPSNR), mean structure similarity (MSSIM), mean feature similarity (MFSIM), as well as mean spectral angle mapper (MSAM), mean relative dimensionless global error synthesis (MERGAS). The MPSNR, MSSIM, and MFSIM are used for spatial-based image quality evaluation, and a larger value represents better denoising performance. In contrast, MSAM and MERGAS are spectral-similarity-based metrics, and a smaller MSAM and MERGAS value implies better spectral fidelity for spectral dimension. Additionally, we also record the average running time of each comparative method under all noisy cases. Noteworthy, for deep learning-based methods, the running time refers to the average time consumed per HSI with the size of $512 \times 512 \times 31$ during the testing phase.

4.2. Results on HSIs with synthetic noises

To verify the effectiveness of our method in stripe noises removal, we compare it with several existing comparative HSI destriping methods and state-of-the-art methods for mixed HSI noises removal : LRTD [18], NMoG [30], LRTDGS [14], FastHyMix [31], SGIDN [23], QRNN3D [32], T3SC [33], MAC-Net [34], DnRCNN [24], $\hat{N}$-Net [35]. For traditional methods, we adopt the publicly available source code and the given parameters from the authors. For deep-learning based algorithms, we have reproduced them based on corresponding references and some public code. To ensure the fairness and credibility of the validation, we have retrained these algorithms using the experimental settings described in this paper.

4.2.1. Spatial quality comparisons

The visual results of some representative scenes are shown in Fig. 6. The first three cases, which cover different types of stripes, exhibit unstable destriping performance among the five conventional algorithms shown in the first-row. As evident in the results of LRTD in Fig. 6(a) and (c), LRTDGS in Fig. 6(b), and FastHyMix in Fig. 6(c), it is obvious that some stripes still remain, and these methods fail to completely remove all stripes. Moreover, LRTDGS in Figs. 6(a) and (d) presents over-smoothed and blurred results due to excessive destriping. In brief, these methods cannot cope with to different types or intensities of stripes corruption because they are initially trained for targeted noisy cases with specific parameter settings. Compared with advanced deep-learning-based methods such as SGIDN, QRNN3D, T3SC, and MAC-Net, DnRCNN, $\hat{N}$-Net, our proposed method is more delicate in retaining details, and the recovered image obtains optimal visual quality. Similar conclusions can also be drawn by observing the results of the last case with mixed noise. Notably, algorithms such as LRTD, SGIDN, and DnRCNN, designed only for stripe noise, fail to remove mixed noise and exhibit poor performance. In contrast, our proposed SLDR achieves the best denoising results with a superior model, achieves high-fidelity recovery, and exhibits fewer artifacts and sharper edges.

---

[1] http://icvl.cs.bgu.ac.il/hyperspectral/
[2] https://engineering.purdue.edu/biehl/MultiSpec/hyperspectral.html
[3] http://hipag.whu.edu.cn/resourcesdownload.html
[4] http://hipag.whu.edu.cn/dataset/

Fig. 6. Denoising results on ICVL dataset with synthetic complex noise. Examples for all the competing methods on band 5, 12 and 20 of the ICVL dataset under 5 complex noise cases are presented respectively. For each group of results, from top to bottom, left to right: Noisy Input, LRTD [18], NMoG [30], LRTDGS [14], FastHyMix [31], SGIDN [23], QRNN3D [32], T3SC [33], and MAC-Net [34], DnRCNN [24], $\hat{N}$-Net [35], Proposed SLDR.

4.2.2. Spectral fidelity comparisons

To verify the preeminent of our method in spectral fidelity, we select one typical pixel in different cases and draw the recovered spectral curves of comparing methods in Fig. 7. Apparently, compared to the other results, the spectra curve recovered by our proposed SLDR is much closer to the reference, and our proposed method yield optimal results in all synthetic cases on ICLV dataset. It reveals that our method accurately eliminates the negative effect of noise in the spectral domain and further confirms the advantage of our model in maintaining high spectral fidelity.

4.2.3. Quantitative comparisons

As the quantitative statistical results shown in Table 1 indicate, the compared methods, particularly the traditional methods LRTD and LRTDGS, exhibit unsatisfactory results, while NMoG and FastHyMix occasionally demonstrate competitive image restoration performance, with some metric values even ranking the top three. However, the performance of these methods is unstable because they can only handle specific noise degradation under customized parameter settings. On the other hand, SGIDN, a deep learning-based method specifically designed for stripe removal, shows a slight advantage in terms of time consumption and MFSIM index, with mediocre performance on other metrics. In contrast, our proposed SLDR demonstrates more solid performance, achieving an average gain of 21 dB in the MPSNR index and reaching 0.98 or 0.99 in MSSIM and MFSIM indices, even in cases severely affected by mixed noise, showing competitiveness. This demonstrates the robustness and effectiveness of our proposed SLDR, which decomposes HSI destriping into a multi-task learning-based framework, releasing the difficulty of restoration. However, it is worth noting that in the second and third cases, SLDR did not achieve the best performance in spectral MERGAS and in the fourth case, spectral MSAM indices, even dropping to lower values and ranked second.

In summary, our SLDR significantly improves spatial visual quality and spectral fidelity indices under all synthetic noise conditions, with a slight sacrifice in speed (0.27 s slower on average than SGIDN). This indicates that our proposed method is a promising approach for HSI destriping and restoration.
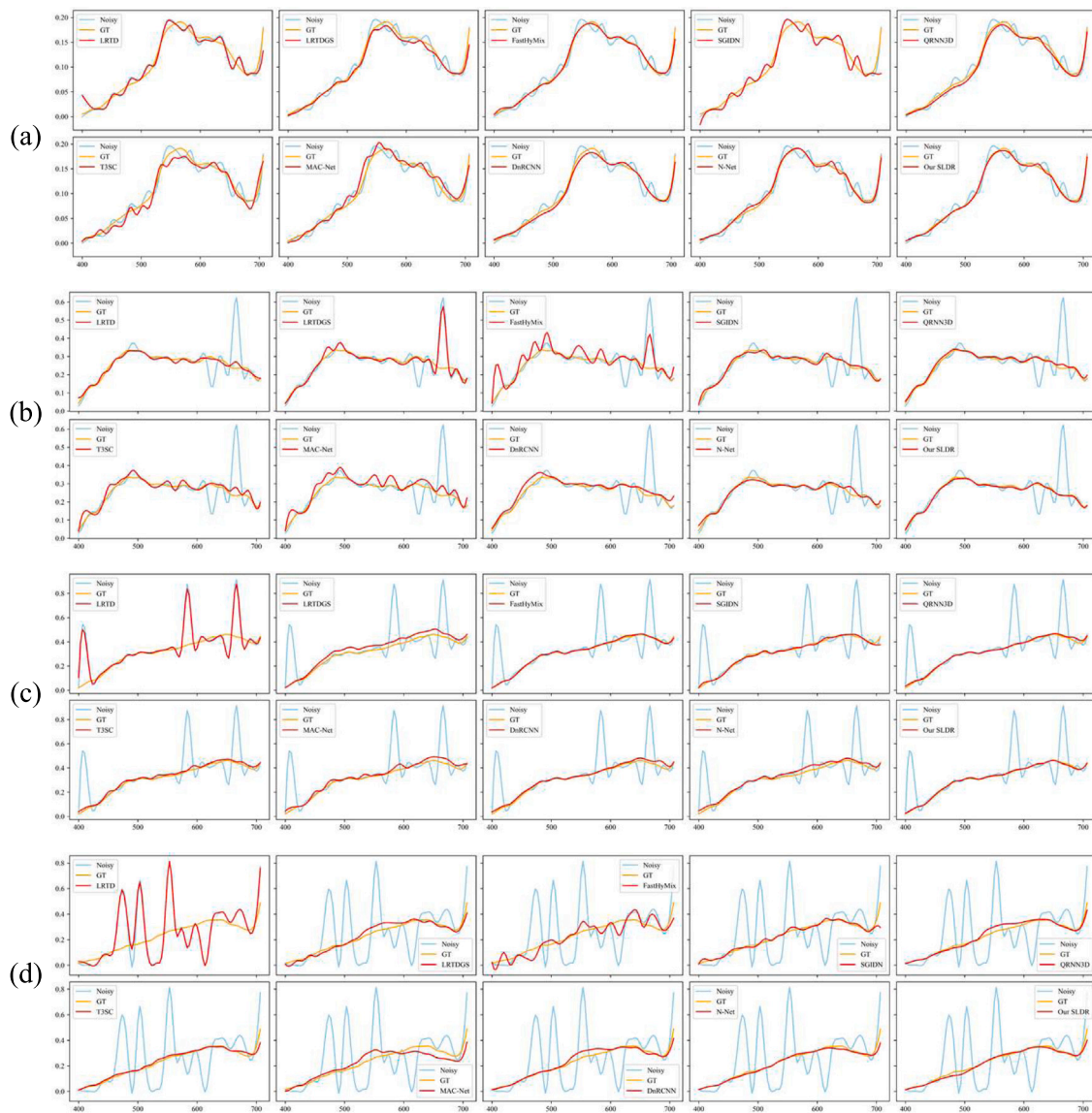
Fig. 7. Comparisons between the recovered spectra of the selected pixel by some advanced comparative methods under all noisy cases from the ICVL dataset are presented, respectively.

4.3. Denoising on remote sensing HSIs with synthetic noise

The experimental results shown in Fig. 8 and Table 2 demonstrate the effectiveness and robust performance of the proposed SLDR in handling complex noise degradation in remote sensing scenes.

Fig. 8 shows denoising results of ROIs on band 112 in the Washington D.C. Mall dataset. It can be seen that the proposed SLDR restores the cleanest denoised image and the most accurate spectral recovery compared to the other methods. Table 2 further quantitatively evaluates the performance of the proposed method and the other competing methods. FastHyMix achieves the best in MPSNR, MERGAS and Times, while NMoG gets the best on MSSIM and MFSIM. The proposed SLDR achieves the top three values of all metrics, indicating its robust denoising performance in remote sensing scenes. On the other hand, the performance of the other methods, neither traditional methods LRTD, LRTDGS, nor deep learning-based methods SGIDN, QRNN3D, and MAC-Net, is relatively poor, with lower and ordinary performance in all indicators. One possible reason is that with regard to all deep-learning-based methods, we directly adopted the trained model (all of these methods are trained on ICVL dataset) to test on this new dataset.

Overall, the experimental results demonstrate that the proposed method outperforms the other competing methods in handling complex noise degradations in remote sensing scenes. The generalization capability of the proposed SLDR to real-world HSIs and its lightweight scale are additional benefits. These results provide valuable insights for advancing the development of effective and robust HSI denoising methods.

4.4. Results on real satellite HSIs

We directly utilize the model trained on synthetic noise data from the ICVL dataset to test the real remote sensing datasets to verify its denoising performance and generalization ability. Figs. 10 and 11 illustrate denoising result on GF-5 Baoqing dataset and GF-5 Shanghai dataset, and Fig. 9 represents results on EO-1 Hyperion dataset. Due to environmental factors such as atmosphere, water, or hardware factors like photon shot within the sensor, these data sets are heavily contaminated with noise. We can easily observe complex noise in some bands involving random noises, stripes, and their mixtures, resulting in a big challenge to recover clean HSIs. Also, the degradations on these
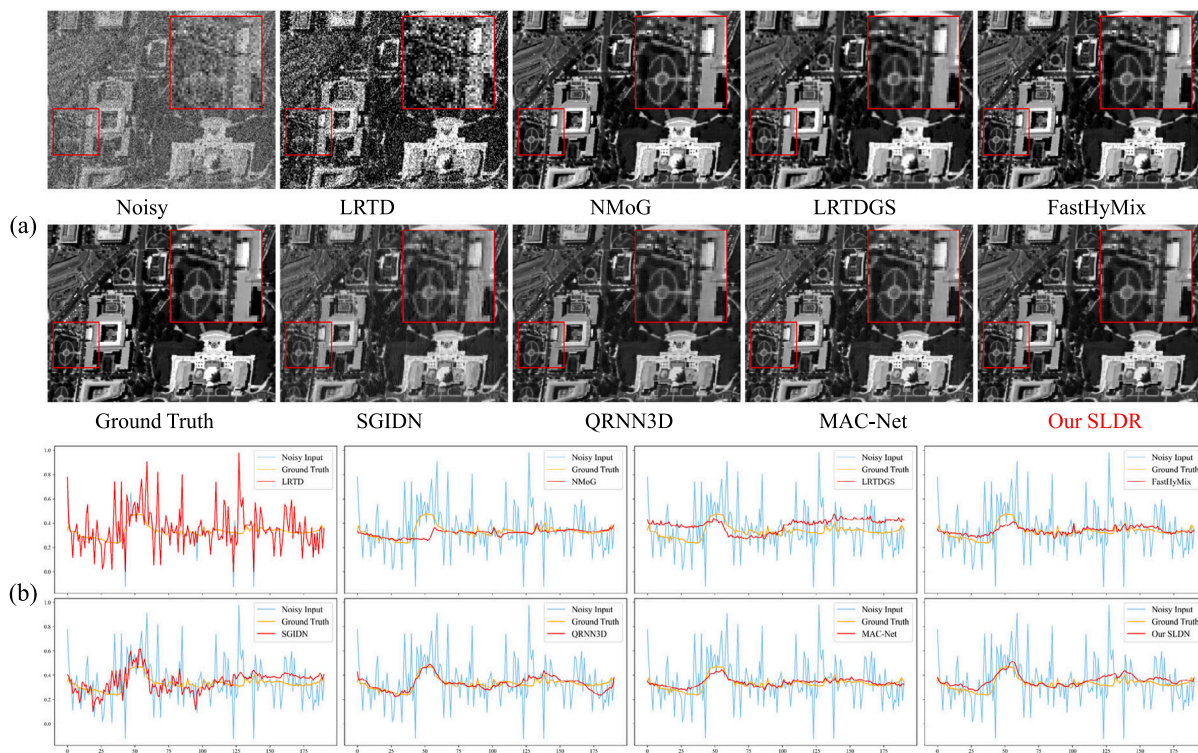
Fig. 8. Comparison in denoising results under complex synthetic mixture noise case on the Washington D.C. Mall dataset. (a) Denoising results of ROIs on band 112; (b) The spectra of the selected pixel recovered by comparative methods.
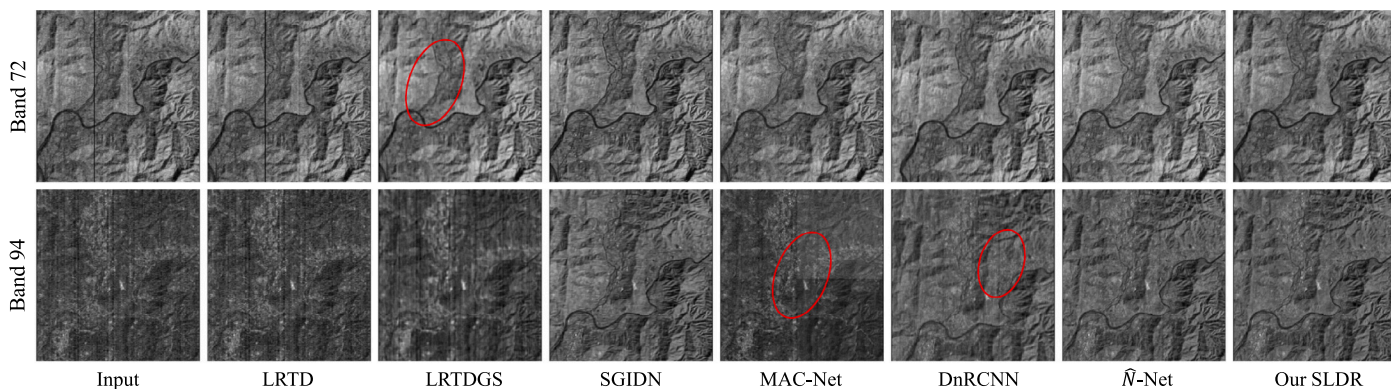

Fig. 9. Comparison in denoising results on EO-1 Hyperion dataset. For each group of results, from the left to right: Noisy Input, LRTD [18], NMoG [30], LRTDGS [14], SGIDN [23], QRNN3D [32], MAC-Net [34], and our proposed SLDR.

Fig. 10. Comparison in denoising results on GF-5 Wuhan dataset. For each group of results, from the left to right: Noisy Input, LRTD [18], NMoG [30], LRTDGS [14], SGIDN [23], QRNN3D [32], MAC-Net [34], and our proposed SLDR.
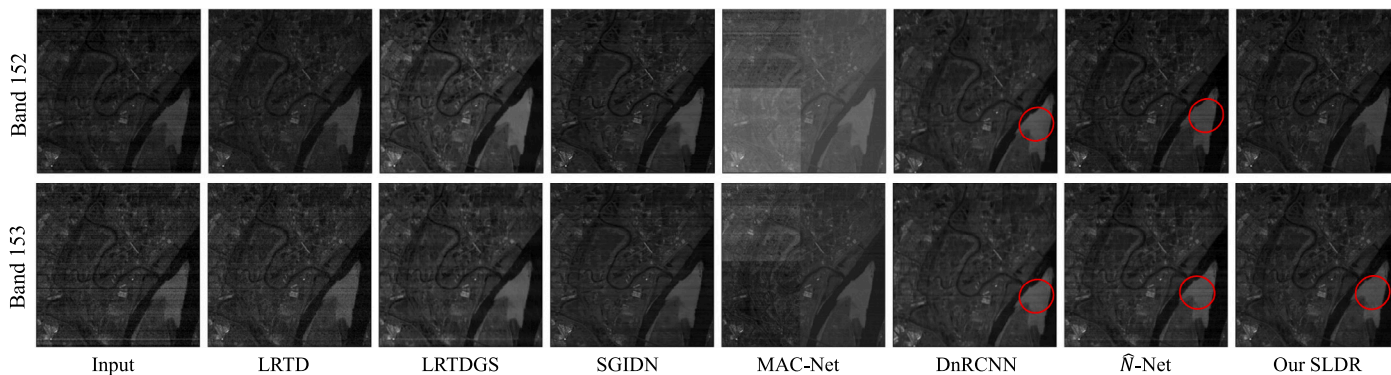
Table 1

Quantitative results of all competing methods under all synthetic stripe noise cases on the ICVL dataset. The best results are shown in red, the second best ones are shown in blue, and the third ones are shown in green.

| Metrics	Noisy	LRTD [18]	NMoG [30]	LRTDGS [14]	FastHyMix [31]	SGIDN [23]	QRNN3D [32]	T3SC [33]	MAC-Net [34]	DnRCNN [24]	$\hat{N}$-Net [35]	SLDR
Case 1: Random stripes												
MPSNR ↑	25.878	38.195	41.839	31.633	38.650	42.304	41.006	35.464	36.067	37.487	38.208	46.977
MSSIM ↑	0.5624	0.8963	0.9656	0.7392	0.7572	0.9705	0.9696	0.9054	0.9267	0.9442	0.9338	0.9885
MFSIM ↑	0.8166	0.9881	0.9945	0.8814	0.9010	0.9971	0.9910	0.9841	0.9867	0.9845	0.9761	0.9965
MERGAS ↓	24.2725	5.7626	3.4657	17.9397	13.9474	3.5950	3.2811	6.2184	7.8546	4.7506	4.7749	1.8716
MSAM ↓	0.5491	0.1747	0.0950	0.4584	0.3723	0.1407	0.1005	0.2082	0.1682	0.1254	0.1365	0.0541
Case 2: Random stripe clusters												
MPSNR↑	29.305	38.541	45.282	35.301	43.793	35.367	43.296	39.166	41.218	44.518	40.899	45.387
MSSIM	0.8475	0.9036	0.9846	0.9234	0.9746	0.8682	0.9838	0.9600	0.9762	0.9863	0.9738	0.9882
MFSIM ↑	0.9555	0.9899	0.9934	0.9601	0.9887	0.9727	0.9945	0.9897	0.9939	0.9960	0.9897	0.9949
MERGAS ↓	15.8683	5.2479	2.0872	10.9627	1.9896	5.9233	2.3289	3.6735	3.5689	2.0036	3.8231	2.0313
MSAM ↓	0.3476	0.1642	0.0562	0.2156	0.0468	0.1911	0.0689	0.1189	0.0811	0.0534	0.1027	0.0524
Case 3: Random stripes in uncertain length												
MPSNR ↑	27.751	28.144	43.938	39.826	45.155	38.093	45.167	40.590	39.616	41.200	34.713	49.227
MSSIM ↑	0.6353	0.6426	0.9745	0.9521	0.9851	0.9433	0.9865	0.9586	0.9667	0.9713	0.8368	0.9934
MFSIM ↑	0.8820	0.8905	0.9934	0.9770	0.9888	0.9850	0.9959	0.9921	0.9947	0.9947	0.9844	0.9984
MERGAS ↓	19.3117	18.1246	3.5657	3.3383	1.0883	4.8502	1.7300	3.4453	4.4488	3.0108	5.8024	1.1294
MSAM ↓	0.4225	0.4152	0.0834	0.0680	0.0462	0.1531	0.0473	0.1146	0.0947	0.0749	0.1541	0.0345
Case 4: Random stripes coexisting with Gaussian noise												
MPSNR ↑	19.978	24.082	38.628	35.209	34.021	39.416	43.012	41.259	37.151	41.202	44.404	44.981
MSSIM ↑	0.2773	0.3734	0.9376	0.8800	0.7984	0.9584	0.9787	0.9681	0.9223	0.9723	0.9839	0.9860
MFSIM ↑	0.6973	0.7381	0.9871	0.9423	0.9390	0.9855	0.9923	0.9904	0.9691	0.9895	0.9950	0.9954
MERGAS ↓	31.9248	30.7352	5.0105	6.1648	7.2914	4.3682	2.1064	2.8592	4.2165	2.5410	1.8708	1.8313
MSAM ↓	0.6842	0.6793	0.1215	0.1586	0.1995	0.1578	0.0523	0.0835	0.1191	0.0574	0.0542	0.0533
TIME(s) ↓	–	357.31	258.02	75.23	2.35	0.42	0.84	1.11	3.21	1.27	0.56	0.69

Table 2

Quantitative results of all competing methods under complex synthetic mixture noise case on Washington D.C. Mall dataset. The best results are shown in red, the second best ones are shown in blue, and the third ones are shown in green.

Metrics	Noisy	LRTD [18]	NMoG [30]	LRTDGS [14]	FastHyMix [31]	SGIDN [23]	QRNN3D [32]	MAC-Net [34]	SLDR
MPSNR ↑	18.710	22.778	34.516	27.271	34.793	26.105	28.288	34.111	34.581
MSSIM ↑	0.6138	0.6599	0.9818	0.9174	0.9783	0.8949	0.9499	0.9744	0.9751
MFSIM ↑	0.7810	0.8050	0.9855	0.9391	0.9825	0.9257	0.9511	0.9691	0.9780
MERGAS ↓	0.4880	0.4108	0.0842	0.0982	0.0702	0.1567	0.1132	0.0761	0.0790
MSAM ↓	21.7720	20.9618	4.0246	7.2009	3.3130	8.5653	6.5941	3.6026	4.2533
TIME(s) ↓	–	326.51	143.63	59.13	0.44	3.41	5.45	14.13	4.56

Table 3

Ablation study of the effectiveness of the task decomposition strategy in our method on the ICVL dataset. We adopt the results by MPSNR, MSSIM to evaluate the restoration performance, and the Params and FLOPs of these networks to measure the model complexity. The results of the proposed SLDR are shown in **bold**.

No.	Stripe mask detection	Stripe intensity estimation	HSI restoration	MPSNR	MSSIM	Params(#)	FLOPs
1	✗	✗	✓	45.788	0.9756	198.21K	25.17G
2	✗	✓	✓	46.213	0.9861	231.32K	29.37G
3	✓	✗	✓	47.175	0.9869	231.63K	29.41G
4	✓	✓	✓	**48.766**	**0.9972**	**264.74K**	**33.61G**
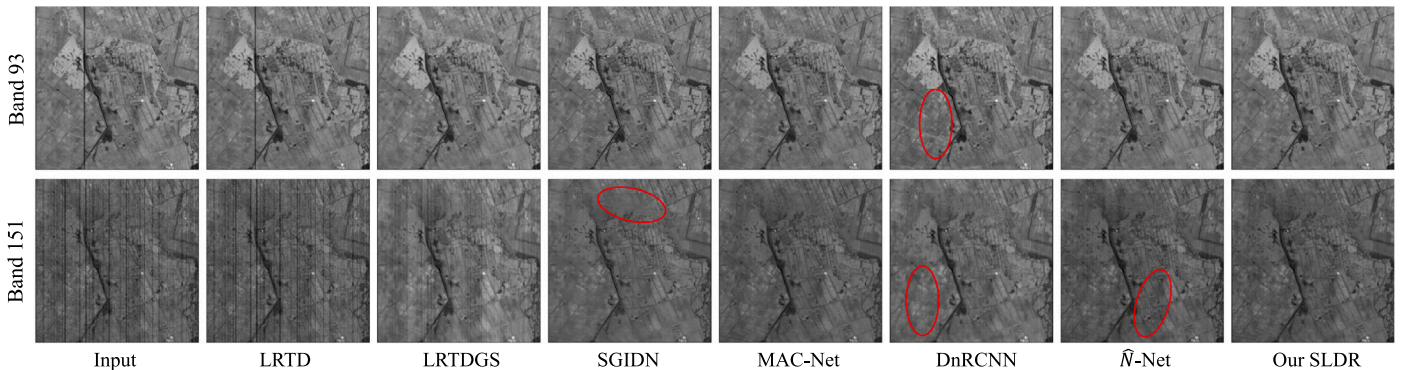
Fig. 11. Comparison in denoising results on GF-5 Baoqing dataset. For each group of results, from the left to right: Noisy Input, LRTD [18], NMoG [30], LRTDGS [14], SGIDN [23], QRNN3D [32], MAC-Net [34], and our proposed SLDR.
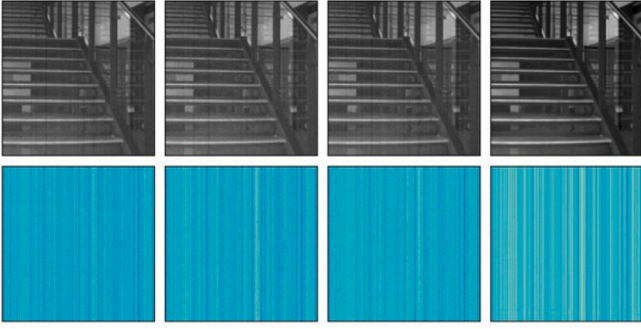
Fig. 12. The destriping results and the residual stripes maps in verifying the effectiveness of the task decomposition strategy in our method on the ICVL dataset. Figures from the left to right are corresponding to the ablation studies No.1 ~ 4 setting in Table 3, respectively.

datasets indicate that various types of stripe noise we simulated above are very close to the practical situation.

Aiming at a fine comparison of the generalization capabilities of the comparative algorithms, we select two bands in each dataset suffering from different intensities of noise contamination as examples. Some red marks circled in the figure above highlight areas, which are not very noticeable with residual noise or suffering from excessive denoising. Zooming in allows for a more detailed comparison. It is apparent that our method achieves better results in visualization due to its strong capability of stripes location-awareness, locating stripes more precisely and recovering texture and edge details more delicate. Moreover, compared to other advances destriping or denoising methods, which seems suffering from inadequate or excessive destriping, the proposed SLDR still performs stably in dealing with different contamination levels in various wavebands. Such results further confirm the robustness of our proposed method in a real-world scenario.

4.5. Ablation studies

To verify the effectiveness of our SLDR downright, we perform comprehensive ablation studies on the ICVL dataset. The ablation studies mainly concern the effectiveness of the task decomposition strategy, the advantages of the MTL-based framework, the effectiveness of the designed stripes-attribute aware block, as well as the incomes brought by the WTV loss function in the proposed SLDR. We adopt MPSNR, MSSIM, and the total number of parameters of the network as the evaluation metrics. The proposed SLDR is the benchmark of these ablation experiments.

4.5.1. Effectiveness of task decomposition strategy

A major contribution of this work is that we introduce multi-task learning based framework into HSI destriping. As aforementioned, we have decomposed HSI destriping into stripe mask detection, stripe intensity estimation, and HSI restoration. Table 3 provides results on whether or not adopting task decomposition strategy and whether or not taking all subtasks. The improvement of destriping performance in adding any of these subtasks is remarkable though there is a bit of additional parameters increasement. For instance, adding subtask of stripe mask detection brings about 0.5 dB incomes in MPSNR and 1% increases in MSSIM, adding subtask of stripe intensity estimation offers about 1.5 dB incomes in MPSNR and 1% increase in MSSIM, and adding both of them provides much higher increases in all metrics with nearly 33K sacrifices in model parameters. This is also demonstrated visually in Fig. 12, where the result for the fourth setting (*i.e.*, the proposed SLDR) is cleaner with almost no residual stripes due to the auxiliary learning of subtasks.
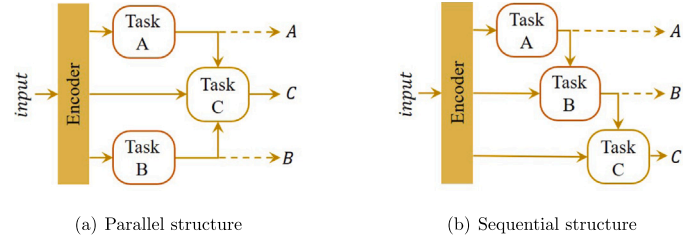

(a) Parallel structure (b) Sequential structure

Fig. 13. Illustrations of potential structures for MTL-based network.

Table 4
Ablation study of potential structures for MTL-based framework. We adopt the results by MPSNR, MSSIM to evaluate the restoration performance, and the Params and FLOPs of these networks to measure the model complexity. The results of the proposed SLDR are shown in **bold**.

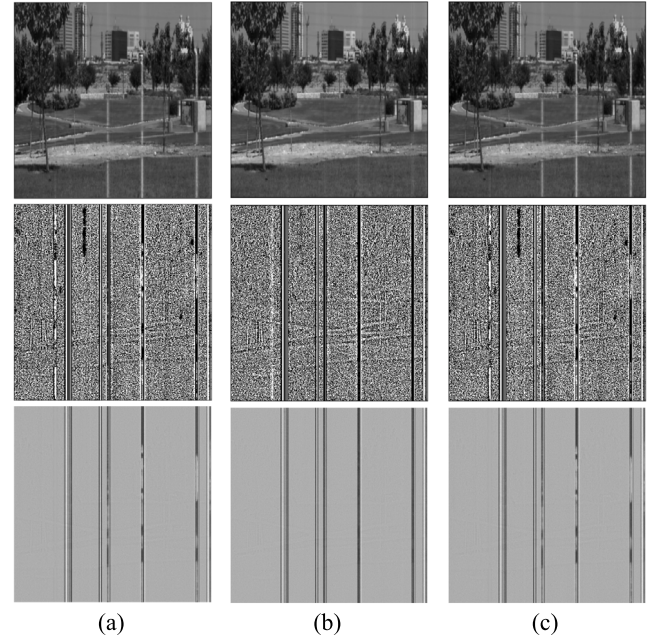| No.	Sequential	Parallel	MPSNR	MSSIM	Param(#)	FLOPs
1		PA-MI	46.567	0.9671	**264.74K**	**33.61G**
2	MIX		**48.766**	**0.9972**	**264.74K**	**33.61G**
3	IMX		47.067	0.9869	264.74K	33.61G

(a) (b) (c)

Fig. 14. Ablation study of potential structures for MTL-based framework. (a) PA-MI; (b) MIX (the proposed SLDR); (c) IMX. For each column of results, from top to bottom: restored X; detected stripe mask M; estimated stripe intensity I.

4.5.2. Advantage of different MTL-based network architectures

The selected architecture of the proposed SLDR is illustrated in Fig. 3. However, there are others potential MTL-based structures for cooperative estimating M, I and X in Eq. (3), as shown in Fig. 13. We have tested potential choice of network architectures, involving parallel structure and sequential ones in different orders.

- Parallel structure: As presented in Fig. 13(a), estimating stripe mask M and stripe intensity I in a parallel structure first, and then integrating them in restoring X, denoted as PA-MI.
- Sequential structure A: As presented in Fig. 13(b), restoring X followed by detecting stripe mask M and estimating stripe intensity I in order, which is marked as MIX and the final choice of out proposed solution.
- Sequential structure B: Restoring X followed by predicting I and M in order, marked as IMX.
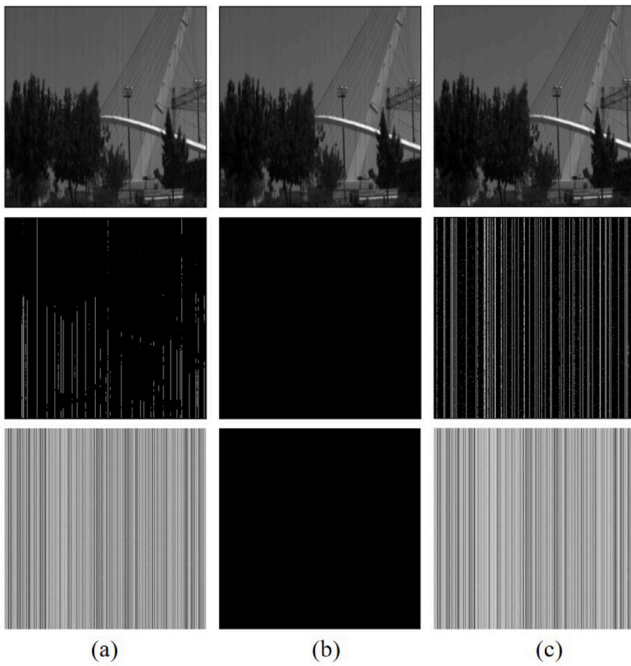
Fig. 15. Ablation study of the designed stripe-attribute aware estimator on the ICVL dataset. (a) Only with the direction-attribute aware block; (b) Only with the sparse-attribute aware block; (c) With the designed stripe-attribute aware estimator. For each column of results, from top to bottom: restored X; detected stripe mask M; estimated stripe intensity I.

Table 5
Ablation study of the designed stripe-attribute aware estimator on the ICVL dataset. We adopt the results by MPSNR, MSSIM to evaluate the restoration performance, and the Params and FLOPs of these networks to measure the model complexity. The results of the proposed SLDR are shown in **bold**.

| No.	Direction	Sparse	MPSNR	MSSIM	Param	FLOPs
1	✗	✓	43.773	0.9500	259.39K	32.92G
2	✓	✗	45.389	0.9837	220.91K	28.03G
3	✓	✓	**48.766**	**0.9972**	**264.74K**	**33.61G**

Table 6
Ablation study of the $\mathcal{L}_{\mathrm{WTV}}(\cdot)$ on the ICVL dataset. We adopt the results by MPSNR, MSSIM to evaluate the restoration performance, and the Params and FLOPs of these networks to measure the model complexity. The results of the proposed SLDR are shown in **bold**.

| No.	$\mathcal{L}_{\mathrm{TV}}(\cdot)$	$\mathcal{L}_{\mathrm{WTV}}(\cdot)$	MPSNR	MSSIM	Param(#)	FLOPs
1	✗	✗	48.446	0.9817	264.74K	33.61G
2	✓	✗	48.793	0.9872	264.74K	33.61G
3	✗	✓	**49.093**	**0.9974**	**264.74K**	**33.61G**

Table 4 offers quantitative results of these three structures, indicating that our final choice, named MIX, have better performance in MPSNR and MSSIM metrics without any extra computation cost. Besides, the predicted X, M and I of these three structures in Fig. 14 also reveals that MIX produced more precise results, indicating the superior of the proposed SLDR.

4.5.3. Effectiveness of the designed stripe attribute-aware estimator

This estimator consists of a direction attribute-aware block and a sparse attribute-aware block. To verify its effectiveness, we conduct ablation study on each block and present results in Table 5 and Fig. 15. Obviously, estimation of stripe noise with either attribute-aware block alone suffers from a certain deficiency, resulting in unsatisfactory recovery results. It also suggests that the designed estimator can precisely aware attributes of stripes, offers abundant information for subsequent destriping and high-fidelity restoration.
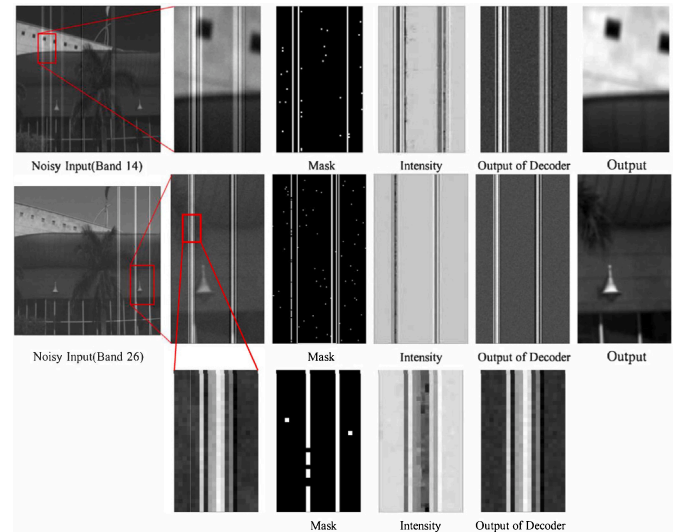
Fig. 16. Example of the stripe and deadline noise removal result and intermediate results (involving estimated mask, intensity, and the output of decoder) by the proposed method on ICVL dataset with synthetic degradation.
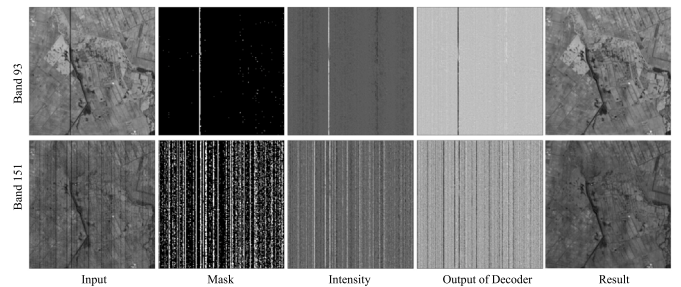
Fig. 17. Example of the stripe and deadline noise removal result and intermediate results (involving estimated mask, intensity, and the output of decoder) by the proposed method on GF5-Baoqing dataset.

4.5.4. Effectiveness of the WTV loss

Table 6 lists quantitative comparisons of ablating $\mathcal{L}_{\mathrm{TV}}(\cdot)$ and $\mathcal{L}_{\mathrm{WTV}}(\cdot)$. It is observed that, adding $\mathcal{L}_{\mathrm{TV}}(\cdot)$ benefits a bit on destriping results, and the updated $\mathcal{L}_{\mathrm{WTV}}(\cdot)$ boosts such benefits more without any increase of computation burden.

4.5.5. Effectiveness on handling deadlines or deadpixels

As shown in Fig. 16, it is apparent that the estimated mask M can locate the stripe degradation more accurately, which provides reliable prior information for subsequent stripe noise removal and the recovery of deadlines and dead pixels. With the guidance of the estimated M, the model can perceive the degradation and missing information in the image more precisely. In addition, the estimation of stripe noise with varying distributions and intensities in different bands, as shown in the figure, is also accurate, which demonstrates the robustness of our proposed method. Besides, we also add extend validation on real-world HSIs. As illustrated in Fig. 17, It is evident from the result of GF5 Baoqing dataset that the proposed method is also robust to real-world HSI restoration (either stripes or deadlines).

5. Conclusion

This work begins with reformulating the mathematical description of the HSI destriping and denoising task. We advocate a novel solution based on task decomposition, which comprises stripes mask detection, stripe intensity estimation, and HSI restoration. Our contribution can

be concluded below: (1) We raise a multi-task learning solution and introduce auxiliary sub-tasks, offering fresh insights and releasing the difficulty to narrow the solution space for recovering degradation. . (2) We propose a stripe location-dependent network, termed as SLDR, integrating the distribution and intensity of stripe noise to achieve more accurate restoration. (3) We design a stripe attribute-aware estimator that enables capturing directional and sparsity by combining different convolutional dictionary-like blocks. (4) We develop $\mathcal{L}_{\text{WTV}}(\cdot)$, which benefits in obtaining higher fidelity recovery results by adaptively adjusting the weights of restoring striped and non-striped regions. As demonstrated by abundant evaluation and comprehensive ablation studies on synthetic stripes and real-world HSIs, our proposed SLDR outperforms state-of-the-art methods designed for either destriping or denosing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgments

References

[1] M. Weinreb, R. Xie, J. Lienesch, D. Crosby, Destriping GOES images by matching empirical distribution functions, Remote Sens. Environ. 29 (2) (1989) 185–195.

[2] Y. Xu, J. Gong, X. Huang, X. Hu, J. Li, Q. Li, M. Peng, Luojia-HSSR: A high spatial-spectral resolution remote sensing dataset for land-cover classification with a new 3D-HRNet, Geo-Spatial Inf. Sci. (2022) 1–13.

[3] Y. Zhong, X. Wang, S. Wang, L. Zhang, Advances in spaceborne hyperspectral remote sensing in China, Geo-Spatial Inf. Sci. 24 (1) (2021) 95–120.

[4] R. Pande-Chhetri, A. Abd-Elrahman, De-striping hyperspectral imagery using wavelet transform and adaptive frequency domain filtering, ISPRS J. Photogramm. Remote Sens. 66 (5) (2011) 620–636.

[5] M. Wang, X. Zheng, J. Pan, B. Wang, Unidirectional total variation destriping using difference curvature in MODIS emissive bands, Infrared Phys. Technol. 75 (2016) 1–11.

[6] J. Jia, Y. Wang, X. Cheng, L. Yuan, D. Zhao, Q. Ye, X. Zhuang, R. Shu, J. Wang, Destriping algorithms based on statistics and spatial filtering for visible-to-thermal infrared pushbroom hyperspectral imagery, IEEE Trans. Geosci. Remote Sens. 57 (6) (2019) 4077–4091.

[7] D. Li, M. Wang, J. Jiang, China's high-resolution optical remote sensing satellites and their mapping applications, Geo-Spatial Inf. Sci. 24 (1) (2021) 85–94.

[8] H. Qin, W. Xie, Y. Li, K. Jiang, J. Lei, Q. Du, Weakly supervised adversarial learning via latent space for hyperspectral target detection, Pattern Recognit. 135 (2023) 109125.

[9] L. Fang, D. Zhu, J. Yue, B. Zhang, M. He, Geometric-spectral reconstruction learning for multi-source open-set classification with hyperspectral and LiDAR data, IEEE/CAA J. Autom. Sin. 9 (10) (2022) 1892–1895.

[10] Z. Shao, W. Wu, D. Li, Spatio-temporal-spectral observation model for urban remote sensing, Geo-Spatial Inf. Sci. 24 (3) (2021) 372–386.

[11] P. Rakwatin, W. Takeuchi, Y. Yasuoka, Stripe noise reduction in MODIS data by combining histogram matching with facet filter, IEEE Trans. Geosci. Remote Sens. 45 (6) (2007) 1844–1856.

[12] B. Nie, L. Yang, J. Jing, J. Zhou, Hyperspectral image destriping method based on time-frequency joint processing method, Optik 172 (2018) 317–327.

[13] Y. Chang, L. Yan, H. Fang, C. Luo, Anisotropic spectral-spatial total variation model for multispectral remote sensing image destriping, IEEE Trans. Image Process. 24 (6) (2015) 1852–1866.

[14] Y. Chen, W. He, N. Yokoya, T.-Z. Huang, Hyperspectral image restoration using weighted group sparsity-regularized low-rank tensor decomposition, IEEE Trans. Cybern. 50 (8) (2020) 3556–3570.

[15] E. Pan, Y. Ma, X. Mei, F. Fan, J. Huang, J. Ma, SQAD: Spatial-spectral quasi-attention recurrent network for hyperspectral image denoising, IEEE Trans. Geosci. Remote Sens. 60 (2022) 1–14.

[16] H. Sun, L. Zhang, L. Wang, H. Huang, Stochastic gate-based autoencoder for unsupervised hyperspectral band selection, Pattern Recognit. 132 (2022) 108969.

[17] E. Pan, Y. Ma, X. Mei, J. Huang, F. Fan, J. Ma, D2Net: Deep denoising network in frequency domain for hyperspectral image, IEEE/CAA J. Autom. Sin. 10 (3) (2022) 813–815.

[18] Y. Chen, T.-Z. Huang, X.-L. Zhao, Destriping of multispectral remote sensing image using low-rank tensor decomposition, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 11 (12) (2018) 4950–4967.

[19] H. Zhang, H. Chen, G. Yang, L. Zhang, LR-Net: Low-rank spatial-spectral network for hyperspectral image denoising, IEEE Trans. Image Process. 30 (2021) 8743–8758.

[20] Y. Chen, T.-Z. Huang, L.-J. Deng, X.-L. Zhao, M. Wang, Group sparsity based regularization model for remote sensing image stripe noise removal, Neurocomputing 267 (2017) 95–106.

[21] Y. Chang, L. Yan, T. Wu, S. Zhong, Remote sensing image stripe noise removal: From image decomposition perspective, IEEE Trans. Geosci. Remote Sens. 54 (12) (2016) 7018–7031.

[22] Y. Chang, L. Yan, H. Fang, S. Zhong, W. Liao, HSI-DeNet: Hyperspectral image restoration via convolutional neural network, IEEE Trans. Geosci. Remote Sens. 57 (2) (2018) 667–682.

[23] Y. Zhong, W. Li, X. Wang, S. Jin, L. Zhang, Satellite-ground integrated destriping network: A new perspective for EO-1 Hyperion and Chinese hyperspectral satellite datasets, Remote Sens. Environ. 237 (2020) 111416.

[24] J. Guan, R. Lai, H. Li, Y. Yang, L. Gu, DnRCNN: Deep recurrent convolutional neural network for HSI destriping, IEEE Trans. Neural Netw. Learn. Syst. (2022).

[25] J. Xue, Y. Zhao, W. Liao, J.C.-W. Chan, Nonlocal low-rank regularized tensor decomposition for hyperspectral image denoising, IEEE Trans. Geosci. Remote Sens. 57 (7) (2019) 5174–5189.

[26] R.C. Evangelista, D.H. Salvadeo, N.D. Mascarenhas, A new Bayesian Poisson denoising algorithm based on nonlocal means and stochastic distances, Pattern Recognit. 122 (2022) 108363.

[27] W. Sun, Q. Du, Hyperspectral band selection: A review, IEEE Geosci. Remote Sens. Mag. 7 (2) (2019) 118–139.

[28] W. He, Y. Chen, N. Yokoya, C. Li, Q. Zhao, Hyperspectral super-resolution via coupled tensor ring factorization, Pattern Recognit. 122 (2022) 108280.

[29] Y. Chang, L. Yan, S. Zhong, Hyper-laplacian regularized unidirectional low-rank tensor recovery for multispectral image denoising, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2017, pp. 4260–4268.

[30] Y. Chen, X. Cao, Q. Zhao, D. Meng, Z. Xu, Denoising hyperspectral image with non-iid noise structure, IEEE Trans. Cybern. 48 (3) (2018) 1054–1066.

[31] L. Zhuang, M.K. Ng, FastHyMix: Fast and parameter-free hyperspectral image mixed noise removal, IEEE Trans. Neural Netw. Learn. Syst. (2021) 1–15.

[32] K. Wei, Y. Fu, H. Huang, 3-D quasi-recurrent neural network for hyperspectral image denoising, IEEE Trans. Neural Netw. Learn. Syst. 32 (1) (2021) 363–375.

[33] T. Bodrito, A. Zouaoui, J. Chanussot, J. Mairal, A trainable spectral-spatial sparse coding model for hyperspectral image restoration, Adv. Neural Inf. Process. Syst. 34 (2021) 5430–5442.

[34] F. Xiong, J. Zhou, Q. Zhao, J. Lu, Y. Qian, MAC-Net: Model-aided nonlocal neural network for hyperspectral image denoising, IEEE Trans. Geosci. Remote Sens. 60 (2022) 1–14.

[35] E. Pan, Y. Ma, X. Mei, F. Fan, J. Ma, Hyperspectral image denoising via spectral noise distribution bootstrap, Pattern Recognit. 142 (2023) 109699.

[36] L. Torrey, J. Shavlik, Transfer learning, in: Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques, IGI global, 2010, pp. 242–264.

[37] S. Liu, Y. Liang, A. Gitter, Loss-balanced task weighting to reduce negative transfer in multi-task learning, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33, No. 01, 2019, pp. 9977–9978.

[38] L. Tang, Y. Deng, Y. Ma, J. Huang, J. Ma, SuperFusion: A versatile image registration and fusion network with semantic awareness, IEEE/CAA J. Autom. Sin. 9 (12) (2022) 2121–2137.

[39] S. Ruder, An overview of multi-task learning in deep neural networks, 2017, arXiv preprint arXiv:1706.05098.

[40] Y. Zhang, Q. Yang, A survey on multi-task learning, IEEE Trans. Knowl. Data Eng. (2021).

[41] S. Vandenhende, S. Georgoulis, M. Proesmans, D. Dai, L. Van Gool, Revisiting multi-task learning in the deep learning era, 2 (3), 2020, arXiv preprint arXiv:2004.13379.

[42] S. Liu, E. Johns, A.J. Davison, End-to-end multi-task learning with attention, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2019, pp. 1871–1880.

[43] O. Sener, V. Koltun, Multi-task learning as multi-objective optimization, Adv. Neural Inf. Process. Syst. 31 (2018).

[44] M.-I. Georgescu, A. Barbalau, R.T. Ionescu, F.S. Khan, M. Popescu, M. Shah, Anomaly detection in video via self-supervised and multi-task learning, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2021, pp. 12742–12752.

[45] I. Misra, A. Shrivastava, A. Gupta, M. Hebert, Cross-stitch networks for multi-task learning, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2016, pp. 3994–4003.

[46] C. Doersch, A. Zisserman, Multi-task self-supervised visual learning, in: Proceedings of the IEEE International Conference on Computer Vision, ICCV, 2017, pp. 2051–2060.

[47] X. Sun, R. Panda, R. Feris, K. Saenko, Adashare: Learning what to share for efficient deep multi-task learning, Adv. Neural Inf. Process. Syst. 33 (2020) 8728–8740.

[48] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, Y. Fu, Residual dense network for image super-resolution, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2018, pp. 2472–2481.

[49] F. Yu, V. Koltun, Multi-scale context aggregation by dilated convolutions, in: Y. Bengio, Y. LeCun (Eds.), 4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2–4, 2016, Conference Track Proceedings, 2016.

[50] B. Arad, O. Ben-Shahar, Sparse recovery of hyperspectral signal from natural RGB images, in: Proceedings of the European Conference on Computer Vision, ECCV, Springer, 2016, pp. 19–34.

Erting Pan received the B.S. degree in electrical engineering and its automation from the Northeast Normal University, Changchun, China, in 2018, the M.E. degree in electronic and communication engineering from Wuhan University, Wuhan, China, in 2020. She is currently studying for the Ph.D. degree in the Multi-spectral Vision Processing Lab, Electronic Information School, Wuhan University, Wuhan, China. Her current research interests include remote sensing image processing, computer vision and pattern recognition.

Yong Ma graduated from the Department of Automatic Control, Beijing Institute of Technology, Beijing, China, in 1997. He received the Ph.D. degree from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2003. His general field of research is in signal and systems. His current research projects include remote sensing of the Lidar and infrared, as well as Infrared image processing, pattern recognition, interface circuits to sensors and actuators. Between 2004 and 2006, he was a Lecturer at the University of the West of England, Bristol, U.K. Between 2006 and 2014, he was with the Wuhan National Laboratory for Optoelectronics, HUST, Wuhan, where he was a Professor of electronics. He is now a Professor with the Electronic Information School, Wuhan University.

Xiaoguang Mei received the B.S. degree in communication engineering from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2007, the M.S. degree in communications and information systems from Huazhong Normal University, Wuhan, in 2011, and the Ph.D. degree in circuits and systems from the HUST, in 2016. From 2010 to 2012, he was a Software Engineer with the 722 Research Institute, China Shipbuilding Industry Corporation, Wuhan. From 2016 to 2019, He was a PostDoctoral Fellow with the Electronic Information School, Wuhan University, Wuhan. He is currently an assistant professor with the Electronic Information School, Wuhan University. His research interests include hyperspectral imagery, machine learning, and pattern recognition.

Jun Huang received the B.S. and Ph.D. degrees from the Department of Electronic and Information Engineering, Huazhong University of Science and Technology, Wuhan, China, in 2008 and 2014, respectively. He is currently working as an Associate Professor at Electronic Information School, Wuhan University, China. His main research interest is infrared image processing and infrared spectrum processing.

Qihai Chen received his PhD. Degree from the School of Astronautics, Harbin Institute of Technology, Harbin, China, in 2009. He is currently the Director of the Space Institute at Tsinghua Innovation Center in Zhuhai, Zhuhai, China. His main research interest is Satellite image processing, navigation and control.

Jiayi Ma received the B.S. degree in information and computing science and the Ph.D. degree in control science and engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2008 and 2014, respectively. From 2012 to 2013, he was an Exchange Student with the Department of Statistics, University of California at Los Angeles, Los Angeles, CA, USA. He was a Post-Doctoral with the Electronic Information School, Wuhan University from August 2014 to November 2015, and received an accelerated promotion to Associate Professor and Full Professor in December 2015 and December 2018, respectively. He has authored or co-authored more than 120 refereed journal and conference papers, including IEEE TPAMI/TIP/TSP/TNNLS/TIE/TGRS /TCYB/TMM/TCSVT, IJCV, CVPR, ICCV, IJCAI, AAAI, ICRA, IROS, ACM MM, etc. His research interests include computer vision, machine learning, and pattern recognition.

Dr. Ma has been identified in the 2019 Highly Cited Researchers list from the Web of Science Group. He was a recipient of the Natural Science Award of Hubei Province (first class), the CAAI (Chinese Association for Artificial Intelligence) Excellent Doctoral Dissertation Award (a total of eight winners in China), and the CAA (Chinese Association of Automation) Excellent Doctoral Dissertation Award (a total of ten winners in China). He is an Editorial Board Member of Information Fusion and Neurocomputing, and a Guest Editor of Remote Sensing.