Hyperspectral Image Denoising Using Uncertainty-Aware Adjustor

Jiahua Xiao , Xing Wei*

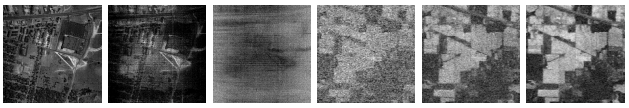
School of Software Engineering, Xi'an Jiaotong University
xjh847286495@stu.xjtu.edu.cn, weixing@mail.xjtu.edu.cn

Abstract

Hyperspectral image (HSI) denoising has achieved promising results with the development of deep learning. A mainstream class of methods exploits the spatial-spectral correlations and recovers each band with the aids of neighboring bands, collectively referred to as spectral auxiliary networks. However, these methods treat entire adjacent spectral bands equally. In theory, clearer and nearer bands tend to contain more reliable spectral information than noisier and farther ones with higher uncertainties. How to achieve spectral enhancement and adaptation of each adjacent band has become an urgent problem in HSI denoising. This work presents the UA-Adjustor, a comprehensive adjustor that enhances denoising performance by considering both the band-to-pixel and enhancement-to-adjustment aspects. Specifically, UA-Adjustor consists of three stages that evaluate the importance of neighboring bands, enhance neighboring bands based on uncertainty perception, and adjust the weight of spatial pixels in adjacent bands through estimated uncertainty. For its simplicity, UA-Adjustor can be flexibly plugged into existing spectral auxiliary networks to improve denoising behavior at low cost. Extensive experimental results validate that the proposed solution can improve over recent state-of-the-art (SOTA) methods on both simulated and real-world benchmarks by a large margin.

1 Introduction

Hyperspectral image (HSI) captures plentiful spectral information of the same scene with contiguous wavebands and provides an internal perspective to analyze the observed targets, unlike gray-scale or RGB images. Thanks to the advantages of hyperspectral imaging [Cai *et al.*, 2022b], a series of tasks based on HSI have been proposed, such as object recognition [Cai *et al.*, 2021a], classification [Hong *et al.*, 2021; Zhong *et al.*, 2021], detection [Liu *et al.*, 2021] and so on. However, due to photon effects, atmosphere, and other inevitable degradation factors, the practical HSIs are often suffered from various complex noises such as stripe, deadline, impulse, and more. These degradations largely hinder the analysis and interpretation of HSI. Therefore, HSI denoising is a critical pre-processing step before the downstream HSI applications.



(a) Urban 137-th band (b) Urban 138-th band (c) Urban 139-th band (d) IPs 103-th band (e) IPs 104-th band (f) IPs 105-th band

Figure 1: Visual presentation of the adjacent bands contaminated by inconsistent and unknown noise on the real-world datasets Urban and Indian Pines (IPs).

As the booming of the deep learning era, more researches [Wei *et al.*, 2020; Rui *et al.*, 2021; Cao *et al.*, 2021] focus on extending the Deep Convolutional Neural Network (DCNN) to the HSI denoising task. Most of these methods attempt to extract and utilize rich spectral information to restore noisy HSI. Among them, the framework of most mainstream networks [Yuan *et al.*, 2018; Yuan *et al.*, 2021; Kan *et al.*, 2022] is to recover one band at a time with the spectral complementary information from a fixed number K of adjacent bands and performs the HSI denoising task band by band. These spectral auxiliary networks consider spatial-spectral correlation simultaneously and achieve impressing performance compared with directly applying the grayscale or RGB image restoration methods in the HSI denoising task.

However, achieving such improvements often requires complex network architectures and training strategies. Therefore, we aim to enhance denoising performance without modifying their original network structure. In existing methods, adjacent bands are typically treated equally, assuming that spectral information is uniformly distributed. However, it is widely known that the noise intensity and distribution vary across bands, as shown in Fig. 1, making it challenging to optimize denoising performance. Additionally, pixels in different spatial areas exhibit varying levels of degradation, indicating the varying importance of visual information. Therefore, for the current denoising band, the uncertainties of neighboring bands and pixels are not consistent. Based on these obser-

*Corresponding author

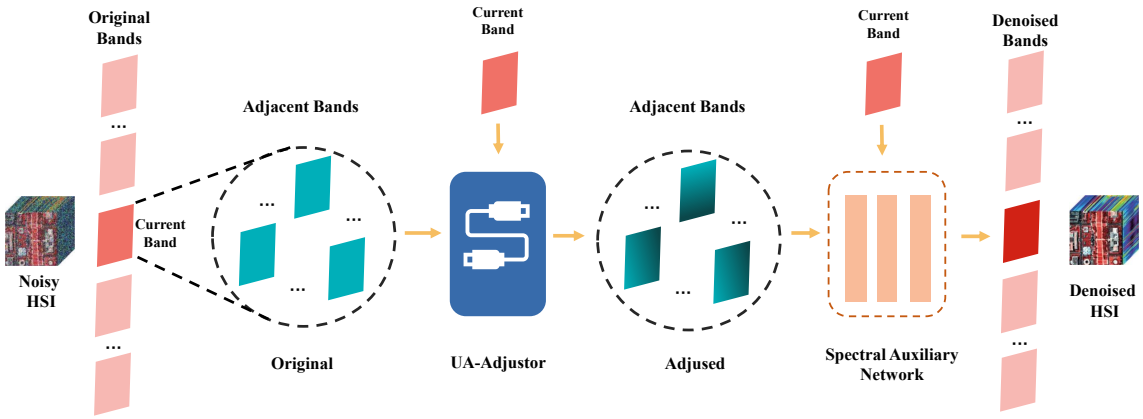


Figure 2: Overview of HSI denoising process with the proposed UA-Adjustor. UA-Adjustor is a plug-and-play module that can be effortlessly inserted into existing spectral auxiliary networks.

vations, we propose to enhance and adjust the adjacent spectral bands via uncertainty perception, leveraging their valuable information for network denoising.

In this paper, we take a novel perspective on HSI denoising and propose a general uncertainty-aware adjustor (UA-Adjustor), which can be flexibly crafted on top of existing spectral auxiliary networks effortlessly to improve their performance significantly. An overview of the HSI denoising process with the proposed UA-Adjustor is illustrated in Fig. 2. It can be seen that UA-Adjustor generates adjusted adjacent bands preceding noise removal, which is decoupled from the spectral auxiliary network. The flowchart of UA-Adjustor is detailed in Fig. 3, including three designed modules, i.e., *Adjacent Derivative Function* (ADF), *Spectral Significance Module* (SSM) and *Uncertainty Estimator* (UE). ADF is designed to make the interaction between two bands by performing a novel bidirectional derivative, SSM is proposed to generate a list of significant weights of the adjacent bands, and UE is devised to estimate the uncertainty of spatial pixels adaptively. The end-to-end learning steps of UA-Adjustor can be described as follows. In the first stage, we evaluate the importance of each neighboring band. The more informative spectral bands tend to impose higher weights to benefit subsequent feature learning. In the second stage, we enhance each band by aggregating the spatial content of a short local range neighboring bands based on uncertainty awareness. In the last stage, we aim to adjust the weight of each spatial pixel in the enhanced adjacent bands by its uncertainty and produce the final spectral auxiliary bands. Unlike channel-wise attention modules like SE-Net [Hu *et al.*, 2018] or S-MSA [Cai *et al.*, 2022a], UA-Adjustor is a more comprehensive adjustor that enhances denoising performance by considering both the band-to-pixel and enhancement-to-adjustment aspects.

The main contributions are summarized as follows:

1. A general UA-Adjustor is proposed, which can be flexibly crafted on existing spectral auxiliary networks and improve denoising performance without any network structural modification.
2. We propose three progressive stages in UA-Adjustor, i.e., evaluating the significance of neighboring bands, enhancing neighboring bands based on uncertainty perception, and adjusting the weight of spatial pixels in adjacent bands using estimated uncertainty.
3. Experimental results on several simulated and real-world benchmarks demonstrate that the proposed UA-Adjustor can improve the denoising performance of recent SOTA networks with a huge boost which achieves a great trade-off between efficiency and performance.

2 Related Work

Different from grayscale or RGB image denoising methods [Cai *et al.*, 2021b], hyperspectral image (HSI) denoising techniques face the challenge of inconsistent and unknown noise across abundant spectral bands. Consequently, most HSI denoising methods leverage the rich spectral information to restore noisy HSIs.

A prevalent category of HSI denoising approaches involves restoring each band in the noisy HSI using adjacent bands as auxiliary information, commonly referred to as spectral auxiliary networks. Among these methods, [Yuan *et al.*, 2018] proposes a joint spatial-spectral learning strategy that captures multiscale spatial-spectral features to recover noisy HSIs. [Maffei *et al.*, 2019] devises a downsampling kernel to achieve fast denoising without losing performance and uses the noise-level map to make the network more robust. [Yuan *et al.*, 2021] attempts to estimate the noise intensity of each band and extracts multiscale spatial–spectral features to retain the texture details in HSI. [Kan *et al.*, 2022] aims to decompose high-frequency features from the original features to better suppress noise. While these methods have shown promising results, they treat all adjacent bands equally, limiting the network's ability to attain optimal performance. To overcome this limitation, we propose a UA-Adjustor and elaborate on its workflow in the subsequent sections.

3 Proposed Method

In this section, we first introduce the preliminary and overall pipeline of UA-Adjustor. Afterwards, we elaborate on each

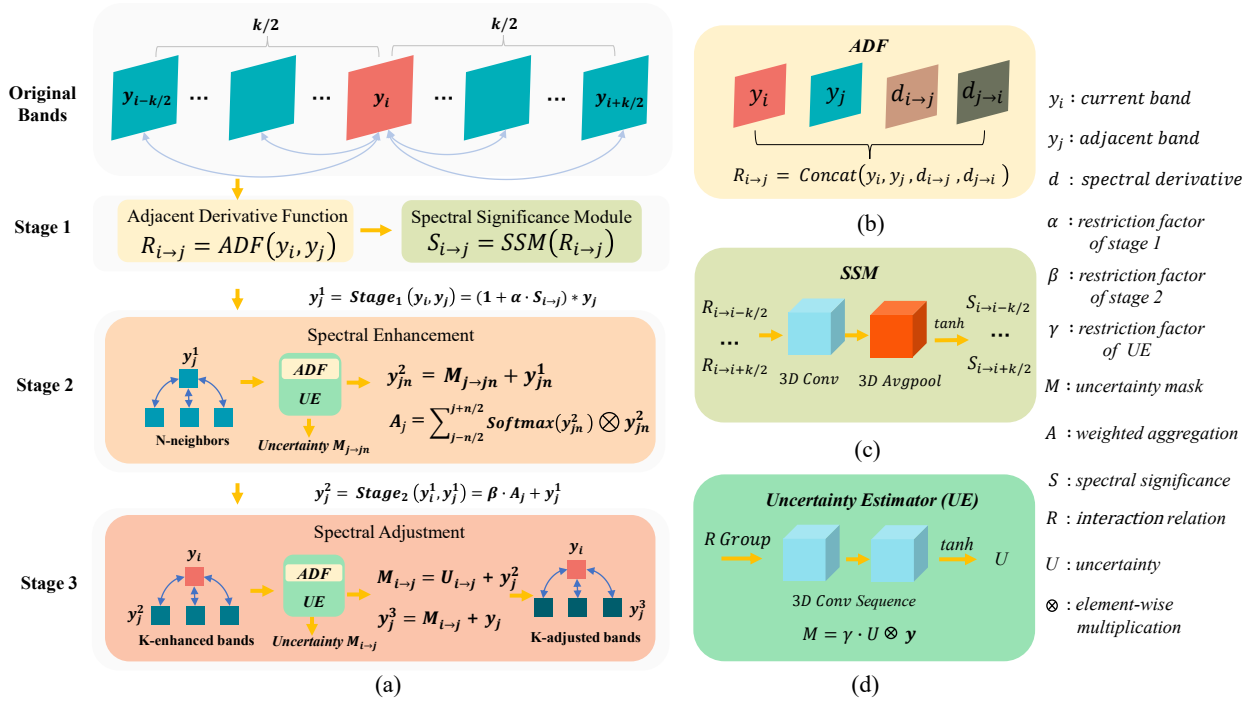


Figure 3: Flowchart of the proposed UA-Adjustor and the details of each component.

function and module in UA-Adjustor. Finally, we describe the workflow of the three progressive stages in detail.

3.1 Preliminary

The HSI can be treated as a 3-D image data of size $H \times W \times B$, where H, W, and B denote the spatial height, spatial width, and number of spectral bands of the HSI. The goal of HSI denoising is to give clean results $\{x_i\}_{i=1}^{B}$ for all bands of noisy HSI $\{y_i\}_{i=1}^{B}$. Suppose the current band to be denoised is y_i and $\{y_j\}_{j=i-k/2}^{i+k/2}$ indicates the fixed number K of adjacent bands. The workflow of the previous spectral auxiliary networks can be represented as:

$$\{x_i\}_{i=1}^{B} = \left\{ denoiser(y_i, \{y_j\}_{j=i-k/2}^{i+k/2}) \right\}_{i=1}^{B}. \quad (1)$$

In UA-Adjustor, we aim to adjust the neighboring bands $\{y_j\}_{j=i-k/2}^{i+k/2}$ based on uncertainty perception. The workflow of UA-Adjustor can be denoted as:

$$\{y_j^3\}_{j=i-k/2}^{i+k/2} = adjustor(y_i, \{y_j\}_{j=i-k/2}^{i+k/2}), \quad (2)$$

where the y_j^3 is the j-th adjusted neighboring band after the third stage. In the following, the functions and modules of each stage will be explained.

3.2 Adjacent Derivative Function

As noted in previous studies [Zhang *et al.*, 2019; Guan *et al.*, 2022], it has been discovered that the spectral derivative can effectively attenuate irrelevant scene details while emphasizing noise components. In these works, the spectral derivative is defined as the forward difference along the spectral dimension between two adjacent bands, as given below:

$$d_{i \to i-1} = y_i - y_{i-1}. \quad (3)$$

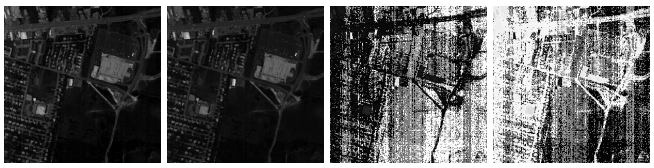
However, we contend that the constraints on the direction of spectral difference and the specific pairing of spectral bands, as defined in Eq. 3, are not applicable for the extraction of long-range spectral information. Here, y_i represents the current band, while y_j represents an adjacent band that can be located either in front of or behind y_i. In order to capture diverse spectral correspondences and ensure consistency, we introduce the concept of bidirectional derivatives, which extend the spectral derivative to incorporate both forward and backward differences between any pair of bands. This bidirectional derivative, illustrated in Fig. 4, effectively suppresses irrelevant and redundant spectral information, while highlighting the spectral disparities. To better leverage the spectral correlation, we propose the integration of the bidirectional derivative into an adjacent derivative function (ADF). The ADF explicitly binds the current denoising band with each of its neighbors as a spectrum pair and reasons the relations between every pair to exploit only valuable information for stronger denoising. Mathematically, the ADF is expressed as follows:

$$R_{i \to j} = ADF(y_i, y_j) = concat(y_i, y_j, d_{i \to j}, d_{j \to i}), \quad (4)$$

where $i \in [0, B]$, $j \in [i - k/2, i + k/2]$, and $R_{i \to j}$ is the interaction relation between y_i and y_j.

3.3 Spectral Significance Module

To efficiently evaluate the significance of different spectral bands, we introduce a lightweight spectral significance mod-



(a) y_i (b) y_j (c) $d_{i\to j}$ (d) $d_{j\to i}$

Figure 4: Visual presentation of adjacent derivative function on the real-world Urban dataset, where $i = 0$ and $j = 3$.

ule (SSM) comprising a 3-D convolution operation, a 3-D adaptive average pooling operation, and a hyperbolic tangent activation function (Tanh). The SSM takes the interaction relation group $\{R_{i\to j}\}_{j=i-k/2}^{i+k/2}$, generated by the ADF, as input and produces a significance list $\{S_{i\to j}\}_{j=i-k/2}^{i+k/2}$ for the adjacent bands:

$$\{S_{i\to j}\}_{j=i-k/2}^{i+k/2} = \{SSM(R_{i\to j})\}_{j=i-k/2}^{i+k/2}. \quad (5)$$

3.4 Uncertainty Estimator

To capture the uncertainty of each spatial pixel in the adjacent bands, we introduce a lightweight uncertainty estimator (UE) comprising two 3-D convolutions and a hyperbolic tangent activation function (Tanh). The UE takes the relation group $\{R_{i\to j}\}_{j=i-k/2}^{i+k/2}$ as input and dynamically learns the uncertainties $\{U_{i\to j}\}_{j=i-k/2}^{i+k/2}$ of different pixels in the adjacent bands during the training process. The workflow of the UE can be described as follows:

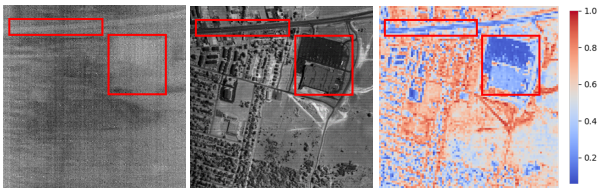
$$\{U_{i\to j}\}_{j=i-k/2}^{i+k/2} = \gamma \cdot \{UE(R_{i\to j})\}_{j=i-k/2}^{i+k/2}, \quad (6)$$

where γ represents the restriction factor, which is a learnable parameter. As illustrated in Fig. 5a, the current denoising band y_i is heavily affected by high-intensity noise. To enhance the noise removal process, we leverage the relatively clean adjacent band y_j (see Fig. 5b) as a spectral auxiliary band. Notably, we observe that the outline within the red rectangles in Fig. 5c is clearly visible in Fig. 5a. Consequently, these areas are assigned lower uncertainty values, while higher uncertainties are assigned to other pixels. In contrast, Fig. 5d shows the current denoising band y_i contaminated by low-intensity noise. Here, the relatively clean adjacent band y_j (see Fig. 5e) exhibits lower uncertainties in most areas (see Fig. 5f). As a result, pixels with higher certainty are given greater attention to extract valuable feature.

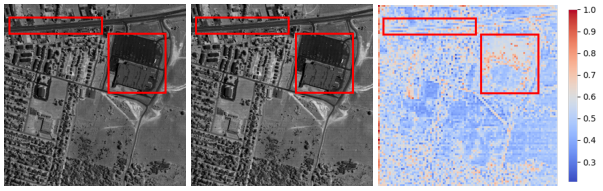
3.5 Spectral Significance Evaluation

Considering that the spectral information from different neighboring bands may contribute differently to the denoising process of the current band, we propose to evaluate the importance of each neighboring band. Consequently, we introduce the concept of modified adjacent bands $\{y_j^1\}_{j=i-k/2}^{i+k/2}$, which can be calculated as follows:

$$\{y_j^1\}_{j=i-k/2}^{i+k/2} = \{(1 + \alpha \cdot S_{i\to j}) * y_j\}_{j=i-k/2}^{i+k/2}, \quad (7)$$



(a) y_i - 103th (b) y_j - 104th (c) Uncertainty 104th



(d) y_i - 100th (e) y_j - 101th (f) Uncertainty 101th

Figure 5: Visual presentation of current denoising band y_i, adjacent band y_j and its uncertainty estimated by UE on the Urban dataset.

where α represents the restriction factor, which is a learnable parameter. $S_{i\to j}$ is the significance weight learned by SSM and y_j^1 denotes the modified y_j of stage 1.

3.6 Spectral Enhancement

The second stage aims to enhance the adjacent spectral bands by leveraging the correlation and constraint between bands. However, enhancing the entire spectral range is computationally impractical due to the large number of bands in HSI. To address this, a novel spectral enhancement strategy based on uncertainty perception is proposed. This strategy involves the introduction of UE to estimate the uncertainty of each spatial pixel in a short local band range. Based on the estimated uncertainty, the adjacent bands are modified accordingly, generating uncertainty masks that represent spectral gain information. To normalize the uncertainty masks within a short local range, the softmax function is applied to these bands. This ensures that the enhancement is balanced and takes into account the relative importance of different pixels within the local range. Finally, the weighted uncertainty masks are aggregated to obtain the final enhancement information, which is then added to the corresponding band. This stage follows the workflow described bellow, allowing for efficient and effective enhancement of the adjacent spectral bands in the HSI:

$$\{M_{j\to j_n}\}_{j_n=j-n/2}^{j+n/2} = \{U_{j\to j_n} \otimes y_{j_n}^1\}_{j_n=j-n/2}^{j+n/2}, \quad (8)$$

$$A_j = \sum_{j_n=j-n/2}^{j+n/2} softmax(M_{j\to j_n}) \otimes M_{j\to j_n}, \quad (9)$$

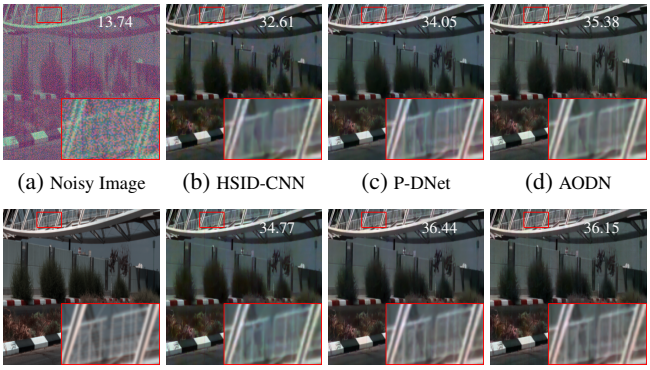
where n is the number of short local range, y_{j_n} is the neighboring bands of y_j, $U_{j\to j_n}$ is the uncertainty of y_{j_n} learned by UE, $M_{j\to j_n}$ is the uncertainty mask of y_{j_n}, $\otimes$ is element-wise multiplication and A_j is the aggregated enhancement information. Thus the enhanced spectral bands of this stage can be calculated as follows:

$$\{y_j^2\}_{j=i-k/2}^{i+k/2} = \{\beta \cdot A_j + y_j^1\}_{j=i-k/2}^{i+k/2}, \quad (10)$$

where β denotes the restriction factor and is a learnable parameter, the y_j^2 is the enhanced y_j of stage 2.

Methods	Case 1			Case 2			Case 3			Case 4		
	PSNR	SSIM	SAM	PSNR	SSIM	SAM	PSNR	SSIM	SAM	PSNR	SSIM	SAM
Noisy (ICVL)	17.65	0.1386	0.8684	17.56	0.1385	0.8806	14.88	0.1024	0.9110	13.82	0.0792	0.9271
HSID-CNN [Yuan *et al.*, 2018]	38.11	0.9450	0.1061	37.83	0.9439	0.1038	36.60	0.9263	0.1457	35.10	0.9123	0.1681
HSID-CNN (Ours)	39.37	0.9491	0.1047	39.43	0.9506	0.1019	38.27	0.9377	0.1315	37.05	0.9305	0.1382
P-DNet [Yuan *et al.*, 2021]	39.40	0.9526	0.0810	39.15	0.9511	0.0820	37.72	0.9218	0.1440	36.53	0.9084	0.1631
P-DNet (Ours)	40.72	0.9613	0.0826	40.87	0.9623	0.0806	39.71	0.9546	0.1005	38.48	0.9457	0.1110
AODN [Kan *et al.*, 2022]	39.94	0.9503	0.1052	39.85	0.9492	0.1031	39.07	0.9424	0.1206	38.05	0.9356	0.1287
AODN (Ours)	40.89	0.9607	0.0827	40.97	0.9617	0.0824	39.78	0.9534	0.0995	38.36	0.9428	0.1052

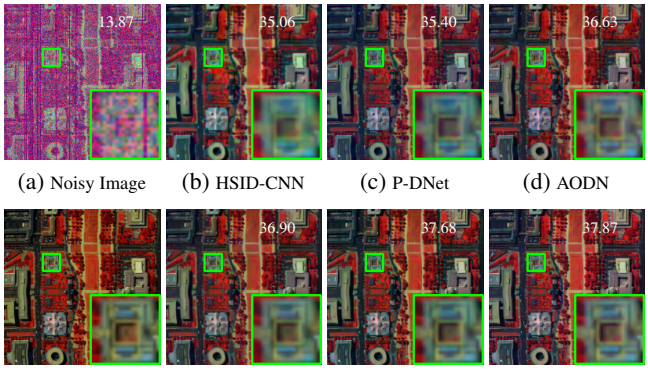
Table 1: Quantitative evaluation the recent SOTA spectral auxiliary methods under different complex noise cases on the ICVL dataset.



(a) Noisy Image (b) HSID-CNN (c) P-DNet (d) AODN

(e) GT Image (f) HSID-CNN (Ours) (g) P-DNet (Ours) (h) AODN (Ours)

Figure 6: Denoising results (PSNR, dB) and visual comparison of the simulated HSI from the ICVL dataset under complex noise case 4. The visual image is synthesized by HSI bands 23, 12, 9.



(a) Noisy Image (b) HSID-CNN (c) P-DNet (d) AODN

(e) GT Image (f) HSID-CNN (Ours) (g) P-DNet (Ours) (h) AODN (Ours)

Figure 7: Denoising results (PSNR, dB) and visual comparison of the simulated remotely sensed dataset WDCM under complex noise case 4. The visual image is synthesized by HSI bands 78, 45, 12.

3.7 Spectral Adjustment

Up till now, the enhanced adjacent spectral bands with more robustness are generated through the preceding two stages. In this stage, we aim to estimate the uncertainty of each spatial pixel in enhanced neighboring bands for current denoising and produce final spectral auxiliary bands. Thus the UE is also used in this stage. The workflow of this stage can be formulated as follows:

$$\{M_{i\to j}\}_{j=i-k/2}^{i+k/2} = \{U_{i\to j} \otimes y_j^2\}_{j=i-k/2}^{i+k/2}, \tag{11}$$

where $U_{i\to j}$ is the uncertainty of y_j learned by UE and the $M_{i\to j}$ is the uncertainty mask of y_j. Ultimately, the final adjacent spectral bands can be adjusted as follows:

$$\{y_j^3\}_{j=i-k/2}^{i+k/2} = \{M_{i\to j} + y_j\}_{j=i-k/2}^{i+k/2}, \tag{12}$$

where the y_j^3 is the adjusted y_j of stage 3.

4 Experiments and Discussions

To illustrate the effectiveness of the proposed UA-Adjustor, in this section, we conduct comprehensive experiments on both simulated and real-world benchmarks. More results and future works are explained in the supplementary materials.

4.1 Experiment Setups

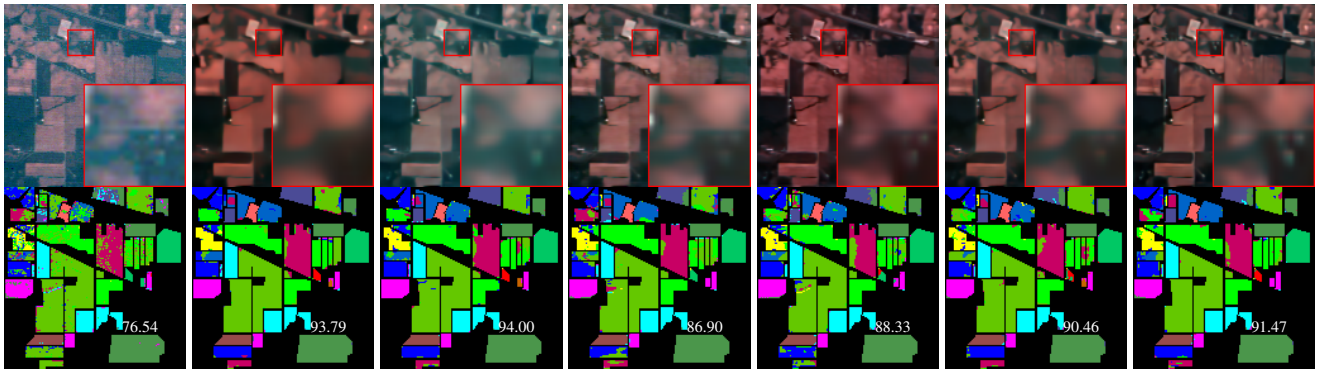
Dataset and Evaluation Setup

We integrate UA-Adjustor into several state-of-the-art systems, namely HSID-CNN [Yuan *et al.*, 2018], P-DNet [Yuan *et al.*, 2021], and AODN [Kan *et al.*, 2022], to evaluate its effectiveness. The training process is performed on the CAVE dataset [Park *et al.*, 2007], and we evaluate these models on the ICVL dataset [Arad and Ben-Shahar, 2016] and the remotely sensed dataset WDCM[1] for simulation experiments. For real-data experiments, we use the real-world datasets Urban [Mnih and Hinton, 2010] and Indian Pines [Landgrebe, 2003]. The details of dataset processing can be found in the supplementary materials. To assess the performance, we follow the commonly used protocols in HSI denoising methods and employ metrics such as peak signal-to-noise ratio (PSNR), structure similarity (SSIM) [Wang *et al.*, 2004], and spectral angle mapper (SAM) [Yuhas *et al.*, 1993]. For real-world datasets, we also use overall accuracy (OA) to evaluate the classification accuracy, which indirectly reflects the denoising performance. Additionally, we consider the parameters (Params) and runtime to evaluate the efficiency of UA-Adjustor in the ablation study.

Simulated Noise Setup

To simulate the various noisy situations in corrupted real-world HSIs, we generate four types of complex noise following the settings in [Wei *et al.*, 2023] (i.e., Case 1: Non-i.i.d. Gaussian + Stripe noise; Case 2: Non-i.i.d. Gaussian + Deadline noise; Case 3: Non-i.i.d. Gaussian + Impulse noise and

[1]https://engineering.purdue.edu/~biehl/MultiSpec/hyperspectral.html



(a) Noisy Image (b) HSID-CNN (c) HSID-CNN(Ours) (d) P-DNet (e) P-DNet (Ours) (f) AODN (g) AODN (Ours)

Figure 8: Visual comparison and classification results (OA) of the real-world HSI from the Indian Pines dataset. The visual image is synthesized by HSI bands 27, 3, 1.

Case 4: Mixture noise) to contaminate clean HSIs.

Implementation Details

Following common protocols in existing spectral auxiliary networks, the number K of adjacent bands is set to 24. The number N of the short local range is set to 4. The working mechanism of UA-Adjustor is to be crafted on spectral auxiliary networks and these two parts are trained as a whole architecture. In practice, the complex noises randomly selected from Case 1 to Case 3 are added to the training set. For head-to-head comparisons, we fairly train baseline networks and their counterparts with UA-Adjustor integrated in the same experimental setup. All experiments are conducted on the NVIDIA GeForce RTX 3090 GPU.

4.2 Experimental Results and Analysis

In this subsection, we fairly compare the recent SOTA spectral auxiliary networks and their original networks with the UA-Adjustor integrated.

Simulation Experiments

Table 1 presents the performance comparison on the ICVL dataset across four types of complex noise cases. The integration of UA-Adjustor significantly improves the performance of each compared method. Notably, HSID-CNN and P-DNet achieve an absolute improvement of **1.95 dB** in Case 4, which is the most challenging scenario and not learned in training process. AODN leverages the UA-Adjustor to enhance denoising performance by **0.31** to **1.12 dB**. Furthermore, Fig. 6 provides visual comparisons under complex noise Case 4. It is evident that methods integrated with UA-Adjustor can better preserve the texture of the original images. To further assess the effectiveness of UA-Adjustor, we also evaluate the entire architecture on the WDCM remotely sensed dataset. The PSNR results and visual comparisons under complex noise Case 4 are shown in Fig. 7. Our method demonstrates satisfactory performance, with a significant improvement of **2.28 dB** observed in P-DNet. HSID-CNN and AODN also achieve improvements of **1.84 dB** and **1.24 dB**, respectively.

Real-Data Experiments

We further evaluate the impact of UA-Adjustor on real-world noisy HSI datasets, namely Urban and Indian Pines. The assessments on the Urban dataset can be found in the supplementary materials. Given that image quality directly influences the performance of downstream HSI applications, we employ classification algorithms [Pedregosa *et al.*, 2011] to objectively evaluate the denoising results of real-world HSIs.

The quantitative and qualitative evaluation results on the Indian Pines dataset are depicted in Fig. 8, the methods integrated with UA-Adjustor effectively remove noise and recover lost details without introducing noticeable oversmoothing artifacts. Consequently, they generate superior classification results compared to the original denoising network.

Ablation Study and Discussions

We conduct a series of ablation studies to examine the impact of different settings of the UA-Adjustor and reveal their contribution to our final performance. All ablations are validated on WDCM benchmark.

Analysis of progressive stages. We conducted experiments to thoroughly validate the functionality of each stage in UA-Adjustor. The denoising results on the WDCM dataset are presented in Table 2. The original P-DNet and AODN serve as the backbones. From the table, it can be observed that the introduction of stage 1, which involves spectral significance evaluation, leads to a performance improvement of **1.60 dB** for P-DNet and **0.68 dB** for AODN. Subsequently, the addition of the modified adjacent bands from the first stage into the second enhancement stage results in further improvements of **1.74 dB** and **0.55 dB** for P-DNet and AODN, respectively, compared to their original versions. Finally, after the adjustment stage, P-DNet achieves an absolute gain of **2.28 dB**, while AODN achieves a gain of **1.24 dB**, compared to their respective backbones. Importantly, each stage introduces only a few parameters, making the lightweight UA-Adjustor require a total of only **0.0089 M** of additional parameters. As the stages progress, P-DNet and AODN integrated with UA-Adjustor demonstrate a significant improvement in denoising performance, while the increase in time

Methods	PSNR (dB)	Params (M)	Time (ms)
P-DNet	35.40	1.2791	25.88
+ stage 1	37.00	1.2878	27.57
+ stage 2	37.14	1.2880	29.08
+ stage 3	37.68	1.2880	32.99
AODN	36.63	2.0753	43.56
+ stage 1	36.97	2.0840	45.27
+ stage 2	37.18	2.0842	46.79
+ stage 3	37.87	2.0842	50.68

Table 2: The efficiency comparisons of the SOTA methods P-DNet and AODN with the integration of UA-Adjustor at different stages.

ADF	PSNR	SSIM	SAM
–	35.06	0.8284	0.1419
$d_{i\to j}$	36.39	0.8749	0.1229
$d_{j\to i}$	36.54	0.8828	0.1210
y_i, y_j	36.53	0.8769	0.1191
$y_i, y_j, d_{i\to j}$	36.78	0.8890	0.1181
$y_i, y_j, d_{j\to i}$	36.71	0.8822	0.1152
$y_i, y_j, d_{i\to j}, d_{j\to i}$	36.90	0.8887	0.1130

Table 3: Different combination results of adjacent derivative function. The '–' denotes denoising result of backbone HSID-CNN.

costs remains within an acceptable range, especially for larger models.

Experiments on ADF. To evaluate the effectiveness of the proposed ADF, including the bidirectional derivative, we conducted ablation studies on different combinations of spectral bands and their derivatives in Eq. 4. The denoising results and efficiency evaluation on the WDCM dataset are presented in Table 3, using the HSID-CNN integrated with UA-Adjustor. As observed, one basic idea is to employ either $d_{i\to j}$ or $d_{j\to i}$ as input, resulting in an improvement of 1.30 dB and 1.48 dB on HSID-CNN. However, this approach has limitations as it only uses spectral derivatives and loses image content. Conversely, combining (y_i, y_j) without spectral derivative information leads to a limited improvement of 1.47 dB. Therefore, we attempt to combine (y_i, y_j) with either $d_{i\to j}$ or $d_{j\to i}$, resulting in improvements of 1.72 dB and 1.65 dB, respectively. Finally, the most comprehensive combination of y_i, y_j, $d_{i\to j}$, and $d_{j\to i}$ is adopted as our adjacent derivative function, achieving improvement **1.85 dB** and optimal performance.

Analysis of short local range N. We conduct an ablation study to see how long the short local range N is optimal in the spectral enhancement stage. As shown in Table 4a, N = 2 costs the lowest runtime but achieves unsatisfactory denoising results. As the value of N increases, the denoising performance shows an initial improvement, but then starts to decline due to a larger range may introduce interference with high uncertainty. Ultimately, we observe that using a short local range of N = 4 offers the best trade-off between efficiency and performance.

N	PSNR	Time (ms)
2	36.48	17.91
6	36.71	18.59
8	36.62	18.92
4	36.90	18.24

(a) Short local range N

K	PSNR	Time (ms)
12	37.03	45.68
18	37.71	47.17
30	37.68	52.29
24	37.87	50.68

(b) Adjacent bands K

Table 4: Ablations of different N on HSID-CNN and K on AODN.

Methods	PSNR	Params (M)	Time (ms)
Ori.	35.40	1.2791	25.88
SE	35.92	1.2794	29.73
S-MSA	36.53	1.2818	28.01
Ours	37.68	1.2880	32.99

Table 5: Ablations of different attention modules on P-DNet.

Analysis of adjacent bands K. We also conduct an ablation study to determine the optimal length of the adjacent bands, denoted as K. Table 4b presents the results of ablation. Similar to the analysis of short local range, when K is set to 12, we observe the lowest runtime, but the denoising results are unsatisfactory. Ultimately, the K is set to 24 provides the best trade-off between efficiency and performance.

Attention scheme comparison. We compare UA-Adjustor with other attention modules and present the results in Table 5. The baseline model achieves a denoising performance of 35.40 dB with 1.2791 M Params and 25.88 ms time cost. We apply the SE-Net, S-MSA, and our UA-Adjustor, respectively. By using these methods, the model's denoising performance improves by 0.52, 1.13, and **2.28 dB**, while adding 0.0003, 0.0027, and 0.0089 M Params and 3.85, 2.13, and 7.11 ms time cost, respectively. UA-Adjustor achieves the most significant improvement in denoising performance with only three learnable 3D-conv layers, which maintains a great balance between efficiency and performance. However, it is anticipated that the denoising performance improvement may be limited when HSI is affected by low levels of simple noise, which can disrupt this balance.

5 Conclusion

In this work, we discuss the problems of existing HSI denoising methods and present UA-Adjustor, an approach that derives an effective strategy of adjacent spectral bands adjustment based on uncertainty perception so as to generate more suitable and powerful adjacent spectral bands on a per-band basis. The UA-Adjustor is general and can be plugged into existing spectral auxiliary networks effortlessly, which contains three progressive stages to achieve improved performance, i.e., spectral significance evaluation, spectral enhancement, and spectral adjustment. Extensive results indicate that the UA-Adjustor can improve over recent state-of-the-art methods on both simulated and real-world datasets.

Acknowledgments

This research is supported by National Key Research and Development Project of China grant number 2020AAA0105600, National Natural Science Foundation of China grant number 62006183, and Fundamental Research Funds for the Central Universities under grant Numbers xhj032021017-04. We also thank Yantao Ji and Jiehua Zhang for discussions.

References

[Arad and Ben-Shahar, 2016] Boaz Arad and Ohad Ben-Shahar. Sparse recovery of hyperspectral signal from natural rgb images. In *European Conference on Computer Vision*, pages 19–34. Springer, 2016.

[Cai *et al.*, 2021a] Weiwei Cai, Zhanguo Wei, Runmin Liu, Yuan Zhuang, Yan Wang, and Xin Ning. Remote sensing image recognition based on multi-attention residual fusion networks. *ASP Transactions on Pattern Recognition and Intelligent Systems*, 1(1):1–8, 2021.

[Cai *et al.*, 2021b] Yuanhao Cai, Xiaowan Hu, Haoqian Wang, Yulun Zhang, Hanspeter Pfister, and Donglai Wei. Learning to generate realistic noisy images via pixel-level noise-aware adversarial training. *Advances in Neural Information Processing Systems*, 34:3259–3270, 2021.

[Cai *et al.*, 2022a] Yuanhao Cai, Jing Lin, Xiaowan Hu, Haoqian Wang, Xin Yuan, Yulun Zhang, Radu Timofte, and Luc Van Gool. Mask-guided spectral-wise transformer for efficient hyperspectral image reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17502–17511, 2022.

[Cai *et al.*, 2022b] Yuanhao Cai, Jing Lin, Haoqian Wang, Xin Yuan, Henghui Ding, Yulun Zhang, Radu Timofte, and Luc V Gool. Degradation-aware unfolding half-shuffle transformer for spectral compressive imaging. *Advances in Neural Information Processing Systems*, 35:37749–37761, 2022.

[Cao *et al.*, 2021] Xiangyong Cao, Xueyang Fu, Chen Xu, and Deyu Meng. Deep spatial-spectral global reasoning network for hyperspectral image denoising. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–14, 2021.

[Guan *et al.*, 2022] Juntao Guan, Rui Lai, Huanan Li, Yintang Yang, and Lin Gu. Dnrcnn: Deep recurrent convolutional neural network for hsi destriping. *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[Hong *et al.*, 2021] Danfeng Hong, Zhu Han, Jing Yao, Lianru Gao, Bing Zhang, Antonio Plaza, and Jocelyn Chanussot. Spectralformer: Rethinking hyperspectral image classification with transformers. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–15, 2021.

[Hu *et al.*, 2018] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018.

[Kan *et al.*, 2022] Ziwen Kan, Suhang Li, Mingzheng Hou, Leyuan Fang, and Yi Zhang. Attention-based octave network for hyperspectral image denoising. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15:1089–1102, 2022.

[Landgrebe, 2003] David A Landgrebe. *Signal theory methods in multispectral remote sensing*, volume 24. John Wiley & Sons, 2003.

[Liu *et al.*, 2021] Jun Liu, Zengfu Hou, Wei Li, Ran Tao, Danilo Orlando, and Hongbin Li. Multipixel anomaly detection with unknown patterns for hyperspectral imagery. *IEEE Transactions on Neural Networks and Learning Systems*, 2021.

[Maffei *et al.*, 2019] Alessandro Maffei, Juan M Haut, Mercedes Eugenia Paoletti, Javier Plaza, Lorenzo Bruzzone, and Antonio Plaza. A single model cnn for hyperspectral image denoising. *IEEE Transactions on Geoscience and Remote Sensing*, 58(4):2516–2529, 2019.

[Mnih and Hinton, 2010] Volodymyr Mnih and Geoffrey E Hinton. Learning to detect roads in high-resolution aerial images. In *European conference on computer vision*, pages 210–223. Springer, 2010.

[Park *et al.*, 2007] Jong-Il Park, Moon-Hyun Lee, Michael D Grossberg, and Shree K Nayar. Multispectral imaging using multiplexed illumination. In *2007 IEEE 11th International Conference on Computer Vision*, pages 1–8. IEEE, 2007.

[Pedregosa *et al.*, 2011] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830, 2011.

[Rui *et al.*, 2021] Xiangyu Rui, Xiangyong Cao, Qi Xie, Zongsheng Yue, Qian Zhao, and Deyu Meng. Learning an explicit weighting scheme for adapting complex hsi noise. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6739–6748, 2021.

[Wang *et al.*, 2004] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.

[Wei *et al.*, 2020] Kaixuan Wei, Ying Fu, and Hua Huang. 3-d quasi-recurrent neural network for hyperspectral image denoising. *IEEE transactions on neural networks and learning systems*, 32(1):363–375, 2020.

[Wei *et al.*, 2023] Xing Wei, Jiahua Xiao, and Yihong Gong. Blind hyperspectral image denoising with degradation information learning. *Remote Sensing*, 15(2):490, 2023.

[Yuan *et al.*, 2018] Qiangqiang Yuan, Qiang Zhang, Jie Li, Huanfeng Shen, and Liangpei Zhang. Hyperspectral image denoising employing a spatial–spectral deep residual convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2):1205–1218, 2018.

[Yuan *et al.*, 2021] Yuan Yuan, Hanwen Ma, and Ganchao Liu. Partial-dnet: A novel blind denoising model with noise intensity estimation for hsi. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–13, 2021.

[Yuhas *et al.*, 1993] Roberta H Yuhas, Joseph W Boardman, and Alexander FH Goetz. Determination of semi-arid landscape endmembers and seasonal trends using convex geometry spectral unmixing techniques. In *JPL, Summaries of the 4th Annual JPL Airborne Geoscience Workshop. Volume 1: AVIRIS Workshop*, 1993.

[Zhang *et al.*, 2019] Qiang Zhang, Qiangqiang Yuan, Jie Li, Xinxin Liu, Huanfeng Shen, and Liangpei Zhang. Hybrid noise removal in hyperspectral imagery with a spatial–spectral gradient network. *IEEE Transactions on Geoscience and Remote Sensing*, 57(10):7317–7329, 2019.

[Zhong *et al.*, 2021] Zilong Zhong, Ying Li, Lingfei Ma, Jonathan Li, and Wei-Shi Zheng. Spectral–spatial transformer network for hyperspectral image classification: A factorized architecture search framework. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–15, 2021.