Supplementary Materials for MP-HSIR: A Multi-Prompt Framework for Universal Hyperspectral Image Restoration

Zhehui Wu[1] Yong Chen[2] Naoto Yokoya[3,4] Wei He[1,4*]
[1] Wuhan University [2] Jiangxi Normal University [3] The University of Tokyo
[4] RIKEN Center for Advanced Intelligence Project
{wuzhehui, weihe1990}@whu.edu.cn, chen_yong@jxnu.edu.cn, yokoya@k.u-tokyo.ac.jp

A. Degradation Preictor

The proposed universal multi-prompt framework, MP-HSIR, can perform specific restoration tasks either based on human instructions or autonomously via the degradation predictor. In this work, we employ a ResNet-34 network with fast Fourier convolution [6] to train classification models separately on natural scene and remote sensing hyperspectral datasets. The binary cross-entropy loss used for training is defined as follows:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log p_i + (1-y_i)\log\left(1-p_i\right)\right], \quad (1)$$

where N denotes the number of samples, p_i represents the output after sigmoid activation, and y_i is the corresponding ground-truth label. The training was conducted with a batch size of 64, using the same optimizer as in the restoration experiments. The initial learning rate was set to 1×10^{-4} and progressively decreased to 1×10^{-6} via cosine annealing [24]. The model was trained for 1000 epochs on the natural scene hyperspectral dataset and 4000 epochs on the remote sensing hyperspectral dataset.

Table 1 presents the accuracy and precision of the degradation predictor for both hyperspectral datasets. The results show that the predictor achieves 100% accuracy and precision across all tasks, demonstrating MPIR-HSI's effectiveness in supporting all trained blind restoration tasks.

Task	Natural Scene		Remote Sensing	
	Accuracy ↑	Precision ↑	Accuracy ↑	Precision ↑
Gaussian Denoising	100.00	100.00	100.00	100.00
Complex Denoising	100.00	100.00	100.00	100.00
Gaussian Deblurring	100.00	100.00	100.00	100.00
Super-Resolution	100.00	100.00	100.00	100.00
Inpainting	100.00	100.00	100.00	100.00
Dehazing	100.00	100.00	100.00	100.00
Band Completion	100.00	100.00	100.00	100.00

Table 1. Accuracy and precision results of the degradation predictor in degradation classification

B. Dataset Details

This section provides a comprehensive overview of the 13 datasets used across 9 hyperspectral image (HSI) restoration tasks and real-world scenarios, as summarized in Table 2.

ARAD [2]. The ARAD dataset, derived from the NTIRE 2022 Spectral Recovery Challenge, was collected using a Specim IQ hyperspectral camera. It consists of 1,000 images, with 900 allocated for training and 50 for testing.

ICVL [1]. The ICVL dataset was obtained using a Specim PS Kappa DX4 hyperspectral camera combined with a rotating stage for spatial scanning. It contains 201 images, with 100 for training and 50 for testing, ensuring no scene overlap.

Xiong'an [38]. The Xiong'an dataset was captured using an imaging spectrometer developed by the Chinese Academy of Sciences. Three central regions of size 512 × 512 were randomly cropped for testing, while the remaining areas were used for training.

WDC [44]. The Washington DC (WDC) dataset was captured by a Hydice sensor. A central region of size 256 × 256 was selected for testing, with the remainder used for training.

PaviaC [14]. The Pavia Center (PaviaC) dataset was acquired using a ROSIS sensor, following the same partitioning strategy as the WDC dataset.

PaviaU [14]. The Pavia University (PaviaU) dataset was also collected using a ROSIS sensor, with the same partitioning strategy as WDC.

Houston [36]. The Houston dataset was obtained using an ITRES CASI-1500 sensor, employing the same partitioning strategy as WDC.

Chikusei [39]. The Chikusei dataset was captured using

*Corresponding author

Type	Dataset	Sensor	Wavelength (nm)	Channels	Size	GSD (m)
Natural HSIs	ARAD_1K	Specim IQ	400–700	31	482×512	/
	ICVL	Specim PS Kappa DX4	400–700	31	1392×1300	/
Remote Sensing HSIs	Xiong'an	Unknown	400–1000	256	3750×1580	0.5
	WDC	Hydice	400–2400	191	1208×307	5
	PaviaC	ROSIS	430–860	102	1096×715	1.3
	PaviaU	ROSIS	430–860	103	610×340	1.3
	Houston	ITRES CASI-1500	364–1046	144	349×1905	2.5
	Chikusei	HH-VNIR-C	343–1018	128	2517×2335	2.5
	Eagle	AsiaEAGLE II	401–999	128	2082×1606	1
	Berlin	Unknown	455–2447	111	6805×1830	3.6
	Urban	Hydice	400-2500	210	307×307	2
	APEX	Unknown	350-2500	285	1000×1500	2
	EO-1	Hyperion	357-2567	242	3471×991	30

Table 2. Properties of 13 Natural Scene and Remote Sensing Hyperspectral Datasets.

an HH-VNIR-C sensor. Four 512 × 512 regions were randomly cropped for testing, with the remaining areas used for training.

Eagle [30]. The Eagle dataset was collected using an AsiaEAGLE II sensor, following the same partitioning strategy as WDC.

Berlin [28]. The Berlin dataset utilizes only the HyMap image from the BerlinUrbGrad dataset. A 512 × 512 central region was randomly cropped for testing, while the remaining data were used for training.

Urban [3]. The Urban dataset, collected using a Hydice sensor, is specifically used for real-world denoising experiments.

APEX [15]. The APEX dataset exhibits characteristics similar to the Urban dataset and is primarily used for fine-tuning pre-trained models.

EO-1 [8]. The EO-1 dataset was captured by the Hyperion sensor. Ten scenes were collected for testing, with 67 invalid bands removed, retaining 175 valid bands for real-world dehazing experiments.

All datasets underwent min-max normalization, and training samples were uniformly cropped to 64 × 64.

C. Detailed Experimental Setup

In this section, we provide a detailed description of the experimental settings for the 9 HSI restoration tasks.

Gaussian Denoising. Each image was corrupted by zero-mean independent and identically distributed (i.i.d.) Gaussian noise with sigma ranging from 30 to 70. For testing, sigma = 30, 50, and 70 were selected for evaluation.

Complex Denoising. Each image was corrupted with one of the following four noise scenarios:

1) Case 1 (*Non-i.i.d. Gaussian Noise*): All bands were corrupted by non-i.i.d. Gaussian noise with standard deviations randomly selected from 10 to 70.

2) Case 2 (*Gaussian Noise + Stripe Noise*): All bands were corrupted by non-i.i.d. Gaussian noise, and one-third of the bands were randomly selected to add column stripe noise with intensities ranging from 5% to 15%.

3) Case 3 (*Gaussian Noise + Deadline Noise*): The noise generation process was similar to Case 2, but stripe noise was replaced by deadline noise.

4) Case 4 (*Gaussian Noise + Impulse Noise*): All bands were corrupted by non-i.i.d. Gaussian noise, and one-third of the bands were randomly selected to add impulse noise with intensities ranging from 10% to 70%.

Gaussian Deblurring. An empirical formula was used to calculate the standard deviation σ based on the Gaussian kernel size K_S, formulated as:

$$\sigma = 0.3 \times \left(\frac{K_S - 1}{2} - 1 \right) + 0.8. \tag{2}$$

For natural hyperspectral datasets, K_S was set to 9, 15, and 21, while for remote sensing hyperspectral datasets, K_S was set to 7, 11, and 15.

Super-Resolution. Bicubic interpolation was used to downsample the images, with downscaling factors of 2, 4, and 8. To ensure that the input and output image sizes of the all-in-one model remained consistent, an unpooling operation was applied to resize the downsampled HSIs to their original dimensions.

Inpainting. Random masks with rates of 0.7, 0.8, and 0.9 were applied to each image for the inpainting task.

Dehazing. To realistically simulate haze contamination, the haze synthesis method from [11] was adopted. Specifi-

Type	Methods	*Gaussian Denoising* (ICVL [1])			*Gaussian Denoising* (ARAD [2])			*Gaussian Denoising* (Xiong'an [38])		
		Sigma = 30	Sigma = 50	Sigma = 70	Sigma = 30	Sigma = 50	Sigma = 70	Sigma = 30	Sigma = 50	Sigma = 70
		PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑
Task Specific	QRNN3D [35]	42.18 / 0.967	39.70 / 0.942	38.09 / 0.933	41.67 / 0.967	39.15 / 0.935	36.71 / 0.894	37.86 / 0.870	36.03 / 0.825	34.29 / 0.792
	SST [19]	43.32 / 0.976	41.09 / 0.952	39.51 / 0.949	43.02 / 0.972	40.58 / 0.951	38.99 / 0.941	39.26 / 0.878	37.34 / 0.848	35.99 / 0.824
	SERT [20]	43.53 / 0.978	41.32 / 0.966	39.82 / 0.956	43.21 / 0.975	40.84 / 0.959	39.21 / 0.945	39.54 / 0.885	37.58 / 0.859	36.37 / 0.833
	LDERT [21]	44.12 / 0.982	41.68 / 0.968	39.95 / 0.957	43.74 / 0.979	41.35 / 0.966	39.32 / 0.950	39.92 / 0.889	37.96 / 0.868	36.54 / 0.838
All in One	AirNet [18]	42.02 / 0.966	39.68 / 0.942	37.59 / 0.923	41.39 / 0.963	39.08 / 0.933	37.09 / 0.903	34.04 / 0.700	31.61 / 0.665	30.17 / 0.639
	PromptIR [31]	42.40 / 0.971	40.14 / 0.954	38.20 / 0.934	41.84 / 0.967	39.55 / 0.947	37.67 / 0.921	34.90 / 0.715	32.76 / 0.680	31.31 / 0.657
	PIP [22]	43.00 / 0.974	40.69 / 0.958	38.94 / 0.941	42.33 / 0.970	40.07 / 0.953	38.36 / 0.933	34.51 / 0.704	32.43 / 0.671	30.98 / 0.647
	HAIR [4]	42.53 / 0.972	40.23 / 0.957	38.78 / 0.939	42.03 / 0.968	39.76 / 0.950	37.95 / 0.928	34.51 / 0.712	32.22 / 0.675	30.89 / 0.650
	InstructIR [7]	42.99 / 0.974	40.84 / 0.960	39.23 / 0.946	42.21 / 0.970	40.16 / 0.955	38.60 / 0.938	33.79 / 0.703	31.47 / 0.662	29.96 / 0.633
	PromptHSI [17]	42.61 / 0.976	40.27 / 0.960	39.08 / 0.945	41.90 / 0.971	39.84 / 0.959	38.37 / 0.938	39.54 / 0.902	37.80 / 0.877	36.87 / 0.864
	MP-HSIR (Ours)	43.62 / 0.977	41.41 / 0.963	39.82 / 0.951	43.12 / 0.975	40.88 / 0.960	39.28 / 0.946	40.55 / 0.922	38.70 / 0.896	37.17 / 0.874

Table 3. **[All-in-one]** Quantitative comparison of all-in-one and state-of-the-art task-specific methods under different Gaussian noise levels on ***Gaussian denoising*** tasks. The best and second-best performances are highlighted in red and blue, respectively.

cally, 100 haze masks were extracted from the cirrus band of Landsat-8 OLI and superimposed onto the original image according to the wavelength ratio to generate haze-affected HSIs, modeled as:

$$I_i = J_i e^{\left(\frac{\lambda_1}{\lambda_i}\right)^{\gamma} \ln t_1} + A\left(1 - e^{\left(\frac{\lambda_1}{\lambda_i}\right)^{\gamma} \ln t_1}\right), \quad (3)$$

where I is the hazy HSI, J is the clear HSI, A is the global atmospheric light, λ is the wavelength, and γ is the spatial function, which is set to 1. The reference transmission map t_1 is calculated from the cirrus band reflectance:

$$t_1 = 1 - \omega B_9, \quad (4)$$

where ω is a weighting factor controlling the haze intensity, and B_9 is the cirrus band reflectance. In the experiments, ω was set to [0.5, 0.75, 1], corresponding to different levels of haze contamination.

Band Completion. A certain proportion of bands were discarded for each image, with discard rates of 0.1, 0.2, and 0.3. The experimental results were evaluated only on the missing bands.

Motion Deblurring. The pre-trained model was fine-tuned and tested on this task with a blur kernel radius of 15 and a blur angle of 45 degrees.

Poisson Denoising. The pre-trained model was directly tested on this task with a Poisson noise intensity scaling factor of 10.

D. Visualization Results of Prompts

In Table 1, we present the similarity matrices of prompt vectors for different degradation types, comparing PromptIR [31], InstructIR [7], and the proposed method. PromptIR relies exclusively on visual prompts, InstructIR utilizes textual prompts for guidance, and the proposed method integrates

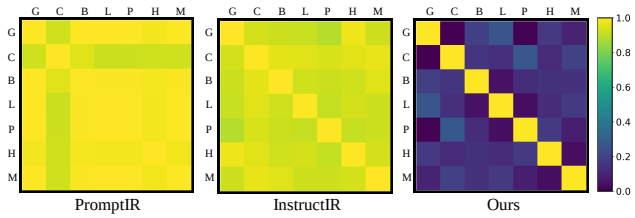


Figure 1. Similarity matrices of prompt vector for seven HSI restoration tasks, with task names as in manuscript.

text-visual synergistic prompts to enhance degradation modeling. As illustrated in the figure, both PromptIR and InstructIR struggle to effectively differentiate between various HSI degradation types, whereas our method demonstrates distinct separation for each individual task, highlighting its superior capability in handling diverse degradation scenarios.

E. More Experimental Results

In this section, we present additional experimental results, including the results of more quantitative results, model efficiency, more results of ablation study, controllability analysis, and more visual comparisons.

E.1. More Quantitative Results

In this section, we provide a detailed quantitative comparison for 7 all-in-one HSI restoration tasks. The experimental results across different degradation levels are systematically presented for each task, including Gaussian denoising in Table 3, complex denoising in Table 4, Gaussian deblurring in Table 5, super-resolution in Table 6, inpainting in Table 8, dehazing in Table 9, and band completion in Table 10.

In addition, we further evaluate the performance on real-world hyperspectral data using the no-reference metric QSFL [37]. The results on the Urban and EO-1 datasets are reported in Table 7, providing a more comprehensive assessment of generalization ability under practical conditions.

Type	Methods	*Complex Denoising* (ICVL [1])				*Complex Denoising* (ARAD [2])				*Complex Denoising* (WDC [44])			
		Case = 1	Case = 2	Case = 3	Case = 4	Case = 1	Case = 2	Case = 3	Case = 4	Case = 1	Case = 2	Case = 3	Case = 4
		PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑
Task Specific	QRNN3D [35]	42.24 / 0.969	41.98 / 0.968	41.62 / 0.968	40.55 / 0.960	41.74 / 0.966	41.55 / 0.965	41.38 / 0.964	39.76 / 0.946	31.98 / 0.885	31.77 / 0.882	31.48 / 0.878	28.05 / 0.822
	SST [19]	43.38 / 0.976	42.69 / 0.973	42.51 / 0.972	41.16 / 0.964	42.84 / 0.973	42.38 / 0.971	42.01 / 0.970	40.56 / 0.956	33.85 / 0.907	33.69 / 0.905	33.37 / 0.901	29.92 / 0.841
	SERT [20]	43.96 / 0.980	43.48 / 0.978	43.45 / 0.977	42.37 / 0.969	43.56 / 0.978	43.19 / 0.976	42.88 / 0.974	41.85 / 0.963	34.48 / 0.922	34.26 / 0.920	33.98 / 0.915	30.53 / 0.854
	LDERT [21]	44.04 / 0.981	43.57 / 0.979	43.55 / 0.979	42.51 / 0.971	43.67 / 0.979	43.33 / 0.977	43.06 / 0.974	42.02 / 0.965	34.65 / 0.923	34.42 / 0.920	34.13 / 0.917	30.74 / 0.856
All in One	AirNet [18]	42.11 / 0.968	41.24 / 0.964	40.89 / 0.961	38.49 / 0.942	41.62 / 0.965	40.83 / 0.959	40.31 / 0.957	37.59 / 0.905	29.02 / 0.752	28.93 / 0.744	28.67 / 0.740	25.67 / 0.667
	PromptIR [31]	42.76 / 0.973	41.93 / 0.969	41.43 / 0.969	39.04 / 0.947	42.26 / 0.969	41.54 / 0.964	40.90 / 0.964	38.12 / 0.919	29.84 / 0.761	29.71 / 0.754	29.39 / 0.747	26.38 / 0.676
	PIP [22]	42.96 / 0.974	42.13 / 0.970	41.38 / 0.969	40.19 / 0.959	42.39 / 0.971	41.71 / 0.966	41.11 / 0.966	39.45 / 0.943	29.57 / 0.751	29.31 / 0.745	29.17 / 0.740	25.95 / 0.658
	HAIR [4]	41.78 / 0.965	41.47 / 0.965	40.68 / 0.958	38.58 / 0.943	41.19 / 0.959	40.95 / 0.961	40.53 / 0.959	37.92 / 0.909	29.38 / 0.756	29.22 / 0.750	28.73 / 0.744	25.40 / 0.664
	InstructIR [7]	41.29 / 0.963	40.89 / 0.961	39.94 / 0.958	38.46 / 0.945	40.72 / 0.960	40.38 / 0.957	39.94 / 0.956	38.21 / 0.934	28.66 / 0.736	28.47 / 0.730	28.19 / 0.725	24.63 / 0.637
	PromptHSI [17]	40.61 / 0.967	40.36 / 0.965	39.30 / 0.960	36.27 / 0.927	40.22 / 0.951	39.98 / 0.955	39.42 / 0.953	35.39 / 0.884	34.93 / 0.931	34.87 / 0.930	34.49 / 0.928	30.78 / 0.857
	MP-HSIR (Ours)	43.07 / 0.975	42.46 / 0.972	42.20 / 0.972	41.42 / 0.966	42.74 / 0.973	42.32 / 0.971	41.93 / 0.970	40.98 / 0.960	35.21 / 0.933	34.99 / 0.931	34.72 / 0.928	31.36 / 0.880

Table 4. **[All-in-one]** Quantitative comparison of all-in-one and state-of-the-art task-specific methods under different cases on ***Complex denoising*** tasks. The best and second-best performances are highlighted in red and blue, respectively.

Type	Methods	*Gaussian Deblurring* (ICVL [1])			*Gaussian Deblurring* (PaviaC [14])			*Gaussian Deblurring* (Eagle [30])		
		Radius = 9	Radius = 15	Radius = 21	Radius = 7	Radius = 11	Radius = 15	Radius = 7	Radius = 11	Radius = 15
		PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑
Task Specific	Stripformer [33]	49.21 / 0.994	45.85 / 0.989	43.04 / 0.980	38.73 / 0.934	36.95 / 0.914	35.72 / 0.891	44.68 / 0.974	41.37 / 0.958	39.83 / 0.942
	FFTformer [16]	49.83 / 0.994	46.43 / 0.989	43.69 / 0.981	39.82 / 0.942	37.63 / 0.919	36.44 / 0.901	45.51 / 0.980	42.18 / 0.962	40.59 / 0.944
	LoFormer [26]	50.35 / 0.995	46.94 / 0.991	44.15 / 0.983	39.54 / 0.941	37.41 / 0.918	36.21 / 0.892	45.32 / 0.979	42.03 / 0.961	40.42 / 0.946
	MLWNet [9]	50.68 / 0.996	47.54 / 0.991	44.76 / 0.983	40.85 / 0.949	38.92 / 0.929	37.28 / 0.901	46.61 / 0.983	43.68 / 0.969	41.83 / 0.955
All in One	AirNet [18]	50.64 / 0.995	47.10 / 0.991	43.89 / 0.982	39.45 / 0.940	37.76 / 0.921	36.08 / 0.893	44.83 / 0.977	42.32 / 0.966	40.11 / 0.946
	PromptIR [31]	51.00 / 0.996	47.70 / 0.992	44.32 / 0.983	40.41 / 0.948	38.62 / 0.929	37.13 / 0.904	46.17 / 0.981	43.45 / 0.968	41.60 / 0.955
	PIP [22]	51.05 / 0.996	47.29 / 0.991	44.21 / 0.983	40.31 / 0.944	38.51 / 0.926	36.73 / 0.902	45.90 / 0.982	42.75 / 0.968	40.86 / 0.951
	HAIR [4]	49.43 / 0.994	46.06 / 0.989	43.87 / 0.981	39.97 / 0.943	38.18 / 0.923	35.64 / 0.891	45.14 / 0.979	42.69 / 0.966	40.48 / 0.947
	InstructIR [7]	24.19 / 0.533	33.44 / 0.878	44.51 / 0.983	19.66 / 0.329	25.12 / 0.583	36.32 / 0.894	23.35 / 0.477	34.05 / 0.872	40.64 / 0.949
	PromptHSI [17]	25.16 / 0.619	30.89 / 0.852	41.52 / 0.980	38.73 / 0.936	36.63 / 0.913	34.84 / 0.888	42.45 / 0.972	39.34 / 0.952	37.46 / 0.938
	MP-HSIR (Ours)	51.53 / 0.996	47.60 / 0.992	45.07 / 0.982	40.85 / 0.949	38.95 / 0.928	37.19 / 0.905	46.26 / 0.982	43.54 / 0.969	41.36 / 0.952

Table 5. **[All-in-one]** Quantitative comparison of all-in-one and state-of-the-art task-specific methods under different blur kernel radius on ***Gaussian deblurring*** tasks. The best and second-best performances are highlighted in red and blue, respectively.

Type	Methods	*Super-Resolution* (ARAD [2])			*Super-Resolution* (PaviaU [14])			*Super-Resolution* (Houston [36])		
		Scale = 2	Scale = 4	Scale = 8	Scale = 2	Scale = 4	Scale = 8	Scale = 2	Scale = 4	Scale = 8
		PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑
Task Specific	SNLSR [13]	43.93 / 0.980	34.56 / 0.902	29.67 / 0.813	34.58 / 0.869	29.85 / 0.719	27.23 / 0.601	34.91 / 0.908	31.33 / 0.782	28.86 / 0.671
	MAN [34]	44.81 / 0.985	35.35 / 0.912	30.49 / 0.830	34.92 / 0.872	30.26 / 0.723	27.59 / 0.604	35.26 / 0.911	31.68 / 0.785	29.15 / 0.677
	ESSAformer [41]	45.32 / 0.988	36.02 / 0.927	30.85 / 0.838	35.47 / 0.879	30.60 / 0.728	27.96 / 0.606	35.60 / 0.913	31.94 / 0.787	29.57 / 0.679
	SRFormer [43]	45.84 / 0.989	36.73 / 0.931	31.48 / 0.845	35.92 / 0.887	31.08 / 0.745	28.41 / 0.620	36.15 / 0.920	32.44 / 0.805	29.81 / 0.684
All in One	AirNet [18]	44.82 / 0.985	35.26 / 0.919	30.12 / 0.828	34.85 / 0.871	30.19 / 0.724	27.76 / 0.609	35.22 / 0.911	31.65 / 0.786	29.13 / 0.676
	PromptIR [31]	45.33 / 0.988	36.00 / 0.927	30.77 / 0.838	35.57 / 0.883	30.81 / 0.735	28.20 / 0.619	35.96 / 0.915	32.38 / 0.798	29.85 / 0.684
	PIP [22]	46.01 / 0.989	37.34 / 0.939	31.73 / 0.853	35.71 / 0.885	31.02 / 0.744	28.20 / 0.618	36.10 / 0.917	32.55 / 0.806	30.02 / 0.692
	HAIR [4]	43.77 / 0.984	35.89 / 0.924	30.87 / 0.836	35.49 / 0.882	30.79 / 0.736	28.13 / 0.616	35.81 / 0.913	32.17 / 0.795	29.64 / 0.680
	InstructIR [7]	43.47 / 0.984	35.46 / 0.921	30.61 / 0.834	35.25 / 0.879	30.71 / 0.732	28.05 / 0.613	35.68 / 0.909	32.15 / 0.789	29.73 / 0.681
	PromptHSI [17]	40.25 / 0.975	35.41 / 0.931	29.35 / 0.806	34.84 / 0.871	30.13 / 0.722	27.27 / 0.602	35.34 / 0.912	31.62 / 0.778	28.59 / 0.635
	MP-HSIR (Ours)	46.72 / 0.991	36.88 / 0.939	31.14 / 0.843	36.27 / 0.894	31.26 / 0.757	28.38 / 0.630	36.57 / 0.926	32.68 / 0.813	29.92 / 0.690

Table 6. **[All-in-one]** Quantitative comparison of all-in-one and state-of-the-art task-specific methods under different downsampling scales on ***Super-Resolution*** tasks. The best and second-best performances are highlighted in red and blue, respectively.

Dataset	PromptIR	InstructIR	PromptHSI	**MP-HSIR**
Urban	14.95	15.56	<u>12.34</u>	**11.42**
EO-1	<u>17.99</u>	19.71	18.13	**16.54**

Table 7. No-reference quality assessment on real datasets.

E.2. Model Efficiency

In this section, we present the parameter counts and computational costs of the all-in-one models for both natural scene and remote sensing hyperspectral datasets. Notably, the network width for remote sensing datasets is 1.5 times greater than that for natural scene datasets across all models. As demonstrated in Table 11, our method achieves a lower parameter count while maintaining competitive computational efficiency.

Furthermore, to evaluate the practical inference performance, Table 12 reports the average inference time of each method on remote sensing datasets. Our method maintains a favorable trade-off between efficiency and accuracy, demonstrating its suitability for large-scale deployment.

Type	Methods	*Inpainting* (ICVL [1])			*Inpainting* (Chikusei [39])		
		Rate = 0.7	Rate = 0.8	Rate = 0.9	Rate = 0.7	Rate = 0.8	Rate = 0.9
		PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑
Task Specific	NAFNet [5]	45.03 / 0.989	44.65 / 0.988	43.50 / 0.985	40.34 / 0.952	40.15 / 0.955	38.97 / 0.952
	Restormer [40]	46.51 / 0.991	46.00 / 0.991	44.85 / 0.988	36.52 / 0.902	36.34 / 0.892	36.13 / 0.903
	DDS2M [27]	45.41 / 0.989	43.34 / 0.983	37.80 / 0.935	36.77 / 0.906	35.23 / 0.901	32.83 / 0.854
	HIR-Diff [29]	41.82 / 0.973	38.87 / 0.949	36.04 / 0.924	38.59 / 0.923	37.96 / 0.920	36.41 / 0.904
All in One	AirNet [18]	43.10 / 0.983	43.06 / 0.983	41.65 / 0.977	38.12 / 0.919	37.86 / 0.921	36.39 / 0.918
	PromptIR [31]	46.96 / 0.992	46.93 / 0.992	45.24 / 0.988	38.86 / 0.925	38.30 / 0.931	37.05 / 0.934
	PIP [22]	44.37 / 0.985	43.47 / 0.983	42.26 / 0.978	38.74 / 0.922	38.58 / 0.930	37.98 / 0.938
	HAIR [4]	44.83 / 0.983	44.30 / 0.983	42.92 / 0.981	38.43 / 0.921	38.28 / 0.928	37.43 / 0.932
	InstructIR [7]	44.85 / 0.989	44.29 / 0.987	43.08 / 0.983	36.30 / 0.904	36.18 / 0.908	35.84 / 0.909
	PromptHSI [17]	42.83 / 0.983	41.72 / 0.976	39.89 / 0.956	38.99 / 0.966	37.64 / 0.952	35.35 / 0.920
	MP-HSIR (Ours)	53.06 / 0.997	51.94 / 0.996	49.60 / 0.994	44.75 / 0.981	44.06 / 0.981	42.08 / 0.975

Table 8. **[All-in-one]** Quantitative comparison of all-in-one and state-of-the-art task-specific methods under different mask rates on ***Inpainting*** tasks. The best and second-best performances are highlighted in red and blue, respectively.

Type	Methods	*Dehazing* (PaviaU [14])			*Dehazing* (Eagle [30])		
		Omega = 0.5	Omega = 0.75	Omega = 1.0	Omega = 0.5	Omega = 0.75	Omega = 1.0
		PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑
Task Specific	SGNet [25]	36.52 / 0.974	34.10 / 0.964	32.22 / 0.949	39.43 / 0.989	37.33 / 0.976	34.90 / 0.962
	SCANet [12]	39.01 / 0.986	36.54 / 0.978	34.21 / 0.969	41.92 / 0.991	39.68 / 0.987	37.31 / 0.978
	MB-Taylor [32]	40.51 / 0.991	38.03 / 0.984	35.44 / 0.975	43.76 / 0.995	40.97 / 0.992	38.36 / 0.986
	DCMPNet [42]	39.63 / 0.993	37.14 / 0.985	34.82 / 0.976	42.93 / 0.995	40.15 / 0.991	37.64 / 0.985
All in One	AirNet [18]	38.61 / 0.982	35.65 / 0.967	32.51 / 0.947	41.88 / 0.991	38.68 / 0.982	35.92 / 0.969
	PromptIR [31]	40.34 / 0.991	37.43 / 0.983	34.47 / 0.971	43.55 / 0.995	40.69 / 0.992	37.94 / 0.988
	PIP [22]	40.30 / 0.991	37.64 / 0.983	34.93 / 0.971	43.21 / 0.994	40.95 / 0.991	38.07 / 0.985
	HAIR [4]	39.47 / 0.989	36.79 / 0.980	34.02 / 0.965	43.38 / 0.995	40.54 / 0.992	38.67 / 0.988
	InstructIR [7]	38.24 / 0.986	34.57 / 0.974	31.36 / 0.954	40.90 / 0.991	38.07 / 0.985	33.99 / 0.971
	PromptHSI [17]	38.62 / 0.982	36.48 / 0.975	35.22 / 0.964	40.88 / 0.986	40.49 / 0.984	37.98 / 0.981
	MP-HSIR (Ours)	42.64 / 0.993	39.46 / 0.988	36.68 / 0.978	45.66 / 0.997	42.24 / 0.995	39.34 / 0.992

Table 9. **[All-in-one]** Quantitative comparison of all-in-one and state-of-the-art task-specific methods under different haze levels on ***Dehazing*** tasks. The best and second-best performances are highlighted in red and blue, respectively.

Type	Methods	*Band Completion* (ARAD [2])			*Band Completion* (Berlin [28])		
		Rate = 0.1	Rate = 0.2	Rate = 0.3	Rate = 0.1	Rate = 0.2	Rate = 0.3
		PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑
Task Specific	NAFNet [5]	47.82 / 0.996	47.02 / 0.995	46.29 / 0.994	39.71 / 0.972	37.51 / 0.874	37.84 / 0.875
	Restormer [40]	49.24 / 0.997	48.29 / 0.995	47.49 / 0.993	34.98 / 0.605	35.24 / 0.606	35.00 / 0.607
	SwinIR [23]	50.95 / 0.997	49.80 / 0.995	48.49 / 0.993	36.88 / 0.950	34.69 / 0.855	34.78 / 0.854
	MambaIR [10]	51.46 / 0.998	50.32 / 0.995	49.01 / 0.993	37.54 / 0.953	35.38 / 0.857	35.43 / 0.855
All in One	AirNet [18]	46.14 / 0.994	45.21 / 0.992	44.46 / 0.990	37.62 / 0.691	35.26 / 0.586	34.86 / 0.595
	PromptIR [31]	47.68 / 0.996	46.71 / 0.994	45.41 / 0.992	42.81 / 0.707	39.55 / 0.910	39.00 / 0.640
	PIP [22]	48.35 / 0.995	47.37 / 0.994	46.37 / 0.991	38.60 / 0.706	36.58 / 0.657	36.43 / 0.641
	HAIR [4]	46.27 / 0.994	44.92 / 0.992	44.04 / 0.990	40.04 / 0.705	36.54 / 0.607	37.45 / 0.639
	InstructIR [7]	52.66 / 0.998	51.37 / 0.997	49.90 / 0.996	36.17 / 0.606	35.33 / 0.559	36.40 / 0.576
	PromptHSI [17]	49.05 / 0.996	47.09 / 0.993	45.89 / 0.992	47.11 / 0.997	43.14 / 0.973	39.82 / 0.956
	MP-HSIR (Ours)	57.83 / 0.999	56.61 / 0.999	54.99 / 0.998	52.14 / 0.999	49.20 / 0.997	47.26 / 0.965

Table 10. **[All-in-one]** Quantitative comparison of all-in-one and state-of-the-art task-specific methods under different mask rates on ***Band Completion*** tasks. The best and second-best performances are highlighted in red and blue, respectively.

Methods	Natural Scene		Remote Sensing	
	Params (M)	FLOPS (G)	Params (M)	FLOPS (G)
AirNet [18]	5.82	19.04	12.23	43.79
PromptIR [31]	33.00	10.03	72.60	22.21
PIP [22]	27.80	10.66	58.26	22.08
HAIR [4]	7.68	2.72	17.28	6.46
InstructIR [7]	68.82	2.81	154.03	6.57
PromptHSI [17]	25.90	10.10	50.89	21.91
MP-HSIR (Ours)	13.88	14.40	30.91	32.74

Table 11. Model complexity comparisons

Metric	PromptIR	InstructIR	PromptHSI	**MP-HSIR**
Inference Time (s)	0.087	**0.065**	0.282	0.083

Table 12. Inference time for remote sensing scenes (64×64×100).

Method	PSNR ↑	SSIM ↑	Params (M)
Baseline (Only Spatial SA)	33.78	0.782	20.93
+ Textual Prompt P_T	34.53	0.807	21.51
+ Visual Prompt P_V	34.47	0.805	23.68
+ Textual Prompt P_T + Visual Prompt P_V	34.92	0.822	24.26
+ Global Spectral SA + P_T + P_T	35.20	0.835	30.07
+ Local Spectral SA + P_T + P_V	35.82	0.846	24.43
+ Local Spectral SA + P_T + P_V + Spectral Prompt P_S	36.67	0.863	25.10
Full Model	**37.17**	**0.874**	30.91

Table 13. Ablation study to verify the effectiveness of modules on Xiong'an dataset in ***Gaussian denoising*** task with sigma = 70.

Method	PSNR ↑	SSIM ↑	Params (M)
Baseline (Only Spatial SA)	30.10	0.788	20.93
+ Textual Prompt P_T	37.64	0.928	21.51
+ Visual Prompt P_V	37.42	0.925	23.68
+ Textual Prompt P_T + Visual Prompt P_V	38.91	0.942	24.26
+ Global Spectral SA + P_T + P_T	39.20	0.944	30.07
+ Local Spectral SA + P_T + P_V	39.79	0.948	24.43
+ Local Spectral SA + P_T + P_V + Spectral Prompt P_S	40.51	0.950	25.10
Full Model	**41.36**	**0.952**	30.91

Table 14. Ablation study to verify the effectiveness of modules on Eagle dataset in ***Gaussian deblurring*** task with radius = 15.

E.3. More results of Ablation Study

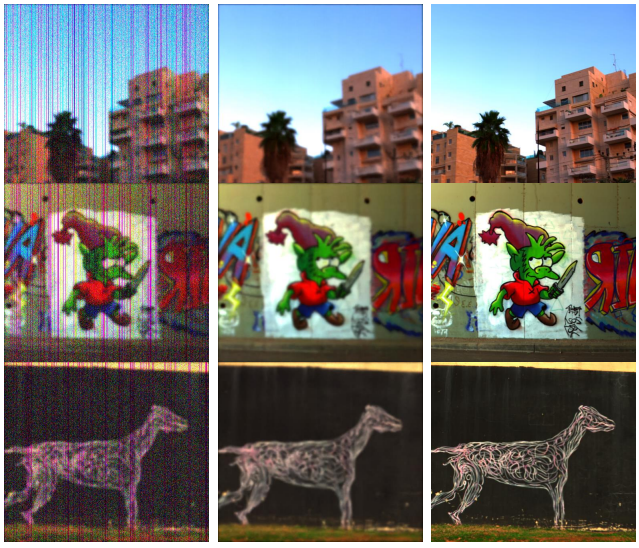
In this section, we present ablation studies on the textual prompts P_T, learnable visual prompts P_V, global spectral self-attention, local spectral self-attention, and spectral prompts P_S across multiple tasks, as shown in Tables 13, 14, and 15. Overall, the addition of each module progressively improves the two accuracy metrics across all tasks.

E.4. Controllable Results

In this section, we demonstrate the controllable restoration capability of the proposed method. Specifically, we conduct controlled restoration tasks under two composite degradation scenarios: Gaussian noise with Gaussian blur and complex noise with Gaussian blur, aiming to remove noise while

Method	PSNR ↑	SSIM ↑	Params (M)
Baseline (Only Spatial SA)	32.52	0.967	20.93
+ Textual Prompt P_T	34.24	0.965	21.51
+ Visual Prompt P_V	34.13	0.964	23.68
+ Textual Prompt P_T + Visual Prompt P_V	34.92	0.969	24.26
+ Global Spectral SA + P_T + P_T	35.53	0.973	30.07
+ Local Spectral SA + P_T + P_V	35.64	0.974	24.43
+ Local Spectral SA + P_T + P_V + Spectral Prompt P_S	36.13	0.976	25.10
Full Model	**36.68**	**0.978**	30.91

Table 15. Ablation study to verify the effectiveness of modules on PaviaU dataset in ***Dehazing*** task with Omega = 1.0.



Mixed Degradation | Controllable Recovery | GT

Figure 2. Controllable Reconstruction: Removing ***Gaussian*** from ***Gaussian Noise*** + ***Gaussian Blur*** Degradation.

preserving blur. As illustrated in Figures 2 and 3, our method can precisely remove specific degradation types through accurate guidance from textual prompts, highlighting its controllability and interpretability.

Building on this controllable restoration paradigm, we further explore sequential degradation removal, where multiple degradations are addressed step by step under prompt guidance. As shown in Figure 4, our method achieves superior results in this more challenging setting, outperforming other approaches in both visual quality and flexibility.

E.5. More Visual Results

In this section, we present further visual results for each task, including all-in-one experiments, generalization testing, and real-world scenarios. As shown in Figures 5, 6, 7, and 8, the visualization results indicate that our method achieves the best performance in restoring texture details and structural features.

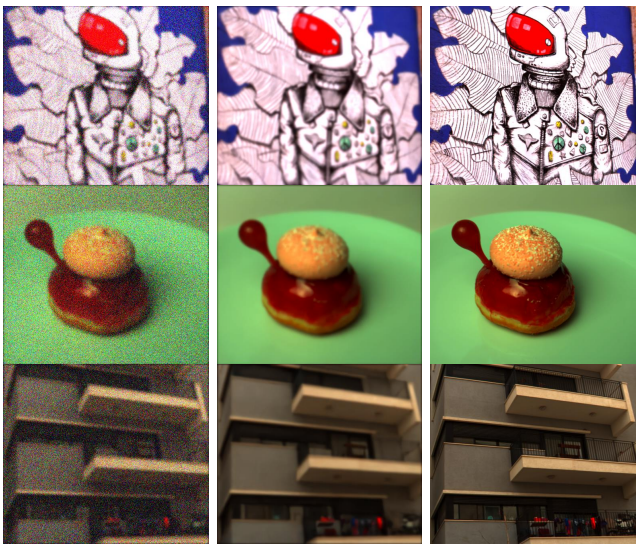


Figure 3. Controllable Reconstruction: Removing ***Complex Noise*** from ***Complex Noise + Gaussian Blur*** Degradation.

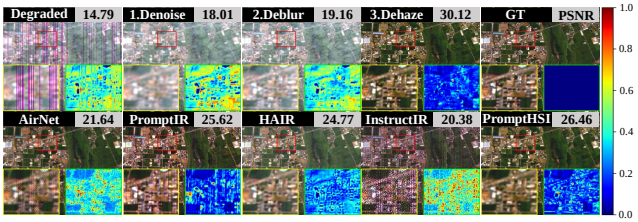


Figure 4. Progressive results (ours) and final results (others).

References

[1] Boaz Arad and Ohad Ben-Shahar. Sparse recovery of hyperspectral signal from natural rgb images. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, the Netherlands, October 11–14, 2016, Proceedings, Part VII 14*, pages 19–34. Springer, 2016. 1, 3, 4, 5

[2] Boaz Arad, Radu Timofte, Rony Yahel, Nimrod Morag, Amir Bernat, Yuanhao Cai, Jing Lin, Zudi Lin, Haoqian Wang, Yulun Zhang, et al. Ntire 2022 spectral recovery challenge and data set. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 863–881, 2022. 1, 3, 4, 5

[3] M. F. Baumgardner, L. L. Biehl, and D. A. Landgrebe. 220-band aviris hyperspectral image data set: June 12, 1992 indian pine test site 3, 2015. Available: https://purr.purdue.edu/publications/1947/1. 2

[4] Jin Cao, Yi Cao, Li Pang, Deyu Meng, and Xiangyong Cao. Hair: Hypernetworks-based all-in-one image restoration. *arXiv preprint arXiv:2408.08091*, 2024. 3, 4, 5, 6

[5] Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, and Jian Sun. Simple baselines for image restoration. In *European Conference on Computer Vision*, pages 17–33. Springer, 2022. 5

[6] Lu Chi, Borui Jiang, and Yadong Mu. Fast fourier convolution. *Advances in Neural Information Processing Systems*, 33:4479–4488, 2020. 1

[7] Marcos V Conde, Gregor Geigle, and Radu Timofte. Instructir: High-quality image restoration following human instructions. In *European Conference on Computer Vision*, pages 1–21. Springer, 2024. 3, 4, 5, 6

[8] Mark A Folkman, Jay Pearlman, Lushalan B Liao, and Peter J Jarecke. Eo-1/hyperion hyperspectral imager design, development, characterization, and calibration. *Hyperspectral Remote Sensing of the Land and Atmosphere*, 4151:40–51, 2001. 2

[9] Xin Gao, Tianheng Qiu, Xinyu Zhang, Hanlin Bai, Kang Liu, Xuan Huang, Hu Wei, Guoying Zhang, and Huaping Liu. Efficient multi-scale network with learnable discrete wavelet transform for blind motion deblurring. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2733–2742, 2024. 4

[10] Hang Guo, Jinmin Li, Tao Dai, Zhihao Ouyang, Xudong Ren, and Shu-Tao Xia. Mambair: A simple baseline for image restoration with state-space model. In *European Conference on Computer Vision*, pages 222–241. Springer, 2024. 5

[11] Jianhua Guo, Jingyu Yang, Huanjing Yue, Hai Tan, Chunping Hou, and Kun Li. Rsdehazenet: Dehazing network with channel refinement for multispectral remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 59 (3):2535–2549, 2020. 2

[12] Yu Guo, Yuan Gao, Wen Liu, Yuxu Lu, Jingxiang Qu, Shengfeng He, and Wenqi Ren. Scanet: Self-paced semi-curricular attention network for non-homogeneous image dehazing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1885–1894, 2023. 5

[13] Qian Hu, Xinya Wang, Junjun Jiang, Xiao-Ping Zhang, and Jiayi Ma. Exploring the spectral prior for hyperspectral image super-resolution. *IEEE Transactions on Image Processing*, 2024. 4

[14] Xin Huang and Liangpei Zhang. A comparative study of spatial approaches for urban mapping using hyperspectral rosis images over pavia city, northern italy. *International Journal of Remote Sensing*, 30(12):3205–3221, 2009. 1, 4, 5

[15] Klaus I Itten, Francesco Dell'Endice, Andreas Hueni, Mathias Kneubühler, Daniel Schläpfer, Daniel Odermatt, Felix Seidel, Silvia Huber, Jürg Schopfer, Tobias Kellenberger, et al. Apex-the hyperspectral esa airborne prism experiment. *Sensors*, 8 (10):6235–6259, 2008. 2

[16] Lingshun Kong, Jiangxin Dong, Jianjun Ge, Mingqiang Li, and Jinshan Pan. Efficient frequency domain-based transformers for high-quality image deblurring. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5886–5895, 2023. 4

[17] Chia-Ming Lee, Ching-Heng Cheng, Yu-Fan Lin, Yi-Ching Cheng, Wo-Ting Liao, Chih-Chung Hsu, Fu-En Yang, and Yu-Chiang Frank Wang. Prompthsi: Universal hyperspectral image restoration framework for composite degradation. *arXiv preprint arXiv:2411.15922*, 2024. 3, 4, 5, 6

[18] Boyun Li, Xiao Liu, Peng Hu, Zhongqin Wu, Jiancheng Lv, and Xi Peng. All-in-one image restoration for unknown

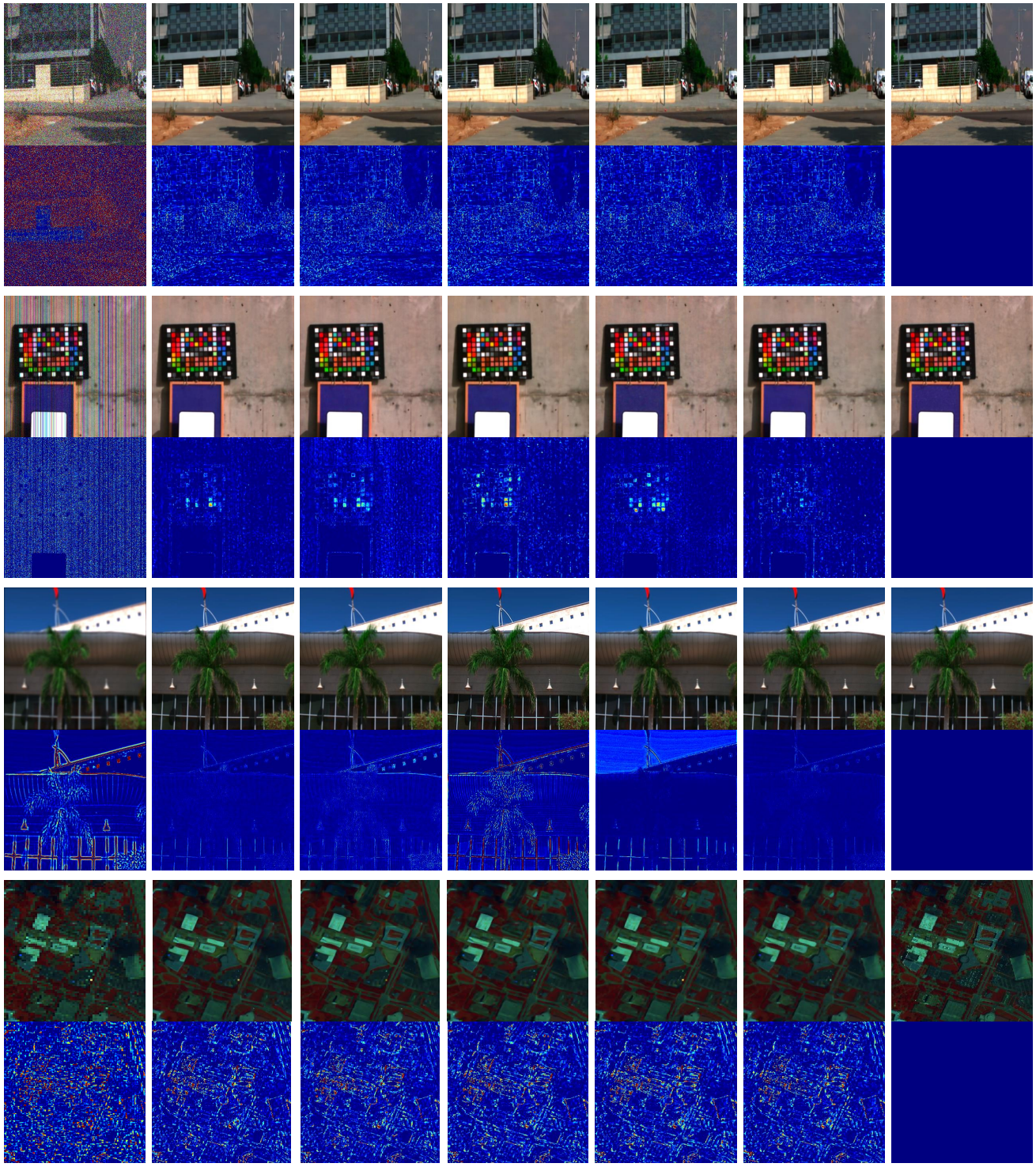


Figure 5. Visual comparison results of ***Gaussian denoising***, ***complex denoising***, ***Gaussian deblurring***, and ***super-resolution***, including the corresponding residual maps. Task-Specific represents the optimal task-specific method.

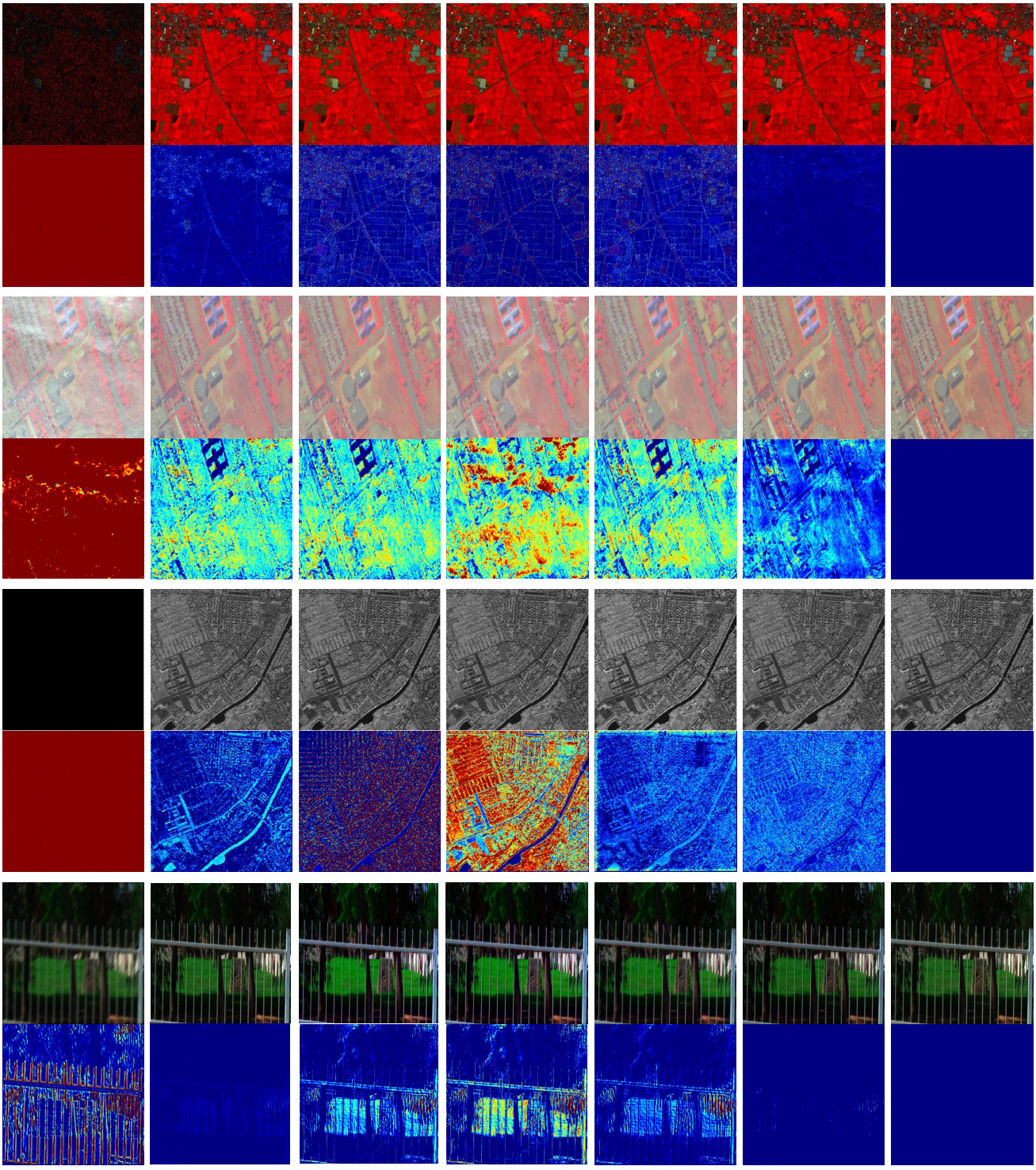


Figure 6. Visual comparison results of ***Inpainting***, ***Dehazing***, ***Band Completion***, and ***Motion Deblurring***, including the corresponding residual maps. Task-Specific represents the optimal task-specific method.

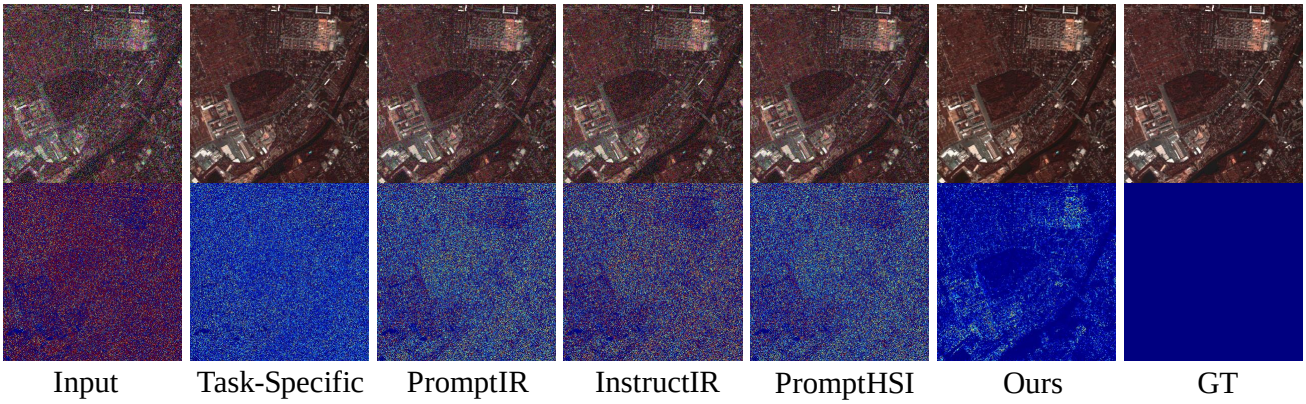


Figure 7. Visual comparison results of ***Poisson Denoising***, including the corresponding residual maps. Task-Specific represents the optimal task-specific method.

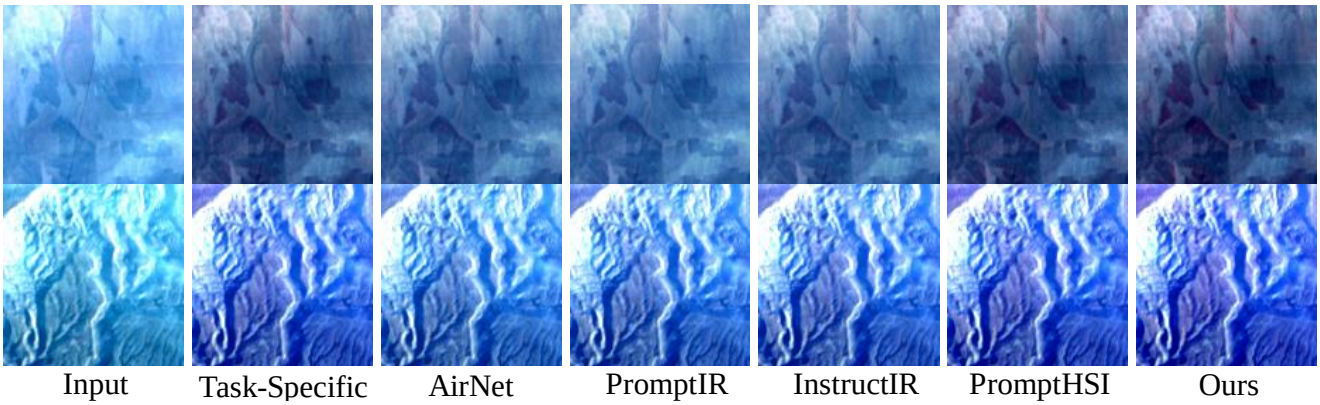


Figure 8. Visual comparison results of ***Real Dehazing***. Task-Specific represents the optimal task-specific method.

corruption. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17452–17462, 2022. 3, 4, 5, 6

[19] Miaoyu Li, Ying Fu, and Yulun Zhang. Spatial-spectral transformer for hyperspectral image denoising. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 1368–1376, 2023. 3, 4

[20] Miaoyu Li, Ji Liu, Ying Fu, Yulun Zhang, and Dejing Dou. Spectral enhanced rectangle transformer for hyperspectral image denoising. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5805–5814, 2023. 3, 4

[21] Miaoyu Li, Ying Fu, Tao Zhang, Ji Liu, Dejing Dou, Chenggang Yan, and Yulun Zhang. Latent diffusion enhanced rectangle transformer for hyperspectral image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024. 3, 4

[22] Zilong Li, Yiming Lei, Chenglong Ma, Junping Zhang, and Hongming Shan. Prompt-in-prompt learning for universal image restoration. *arXiv preprint arXiv:2312.05038*, 2023. 3, 4, 5, 6

[23] Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. Swinir: Image restoration using swin transformer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 1833–1844, 2021. 5

[24] Ilya Loshchilov and Frank Hutter. Sgdr: Stochastic gradient descent with warm restarts. *arXiv preprint arXiv:1608.03983*, 2016. 1

[25] Xiaofeng Ma, Qunming Wang, and Xiaohua Tong. A spectral grouping-based deep learning model for haze removal of hyperspectral images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 188:177–189, 2022. 5

[26] Xintian Mao, Jiansheng Wang, Xingran Xie, Qingli Li, and Yan Wang. Loformer: Local frequency transformer for image deblurring. In *Proceedings of the 32nd ACM International Conference on Multimedia*, pages 10382–10391, 2024. 4

[27] Yuchun Miao, Lefei Zhang, Liangpei Zhang, and Dacheng Tao. Dds2m: Self-supervised denoising diffusion spatio-spectral model for hyperspectral image restoration. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 12086–12096, 2023. 5

[28] Akpona Okujeni, Sebastian van der Linden, and Patrick Hostert. Berlin-urban-gradient dataset 2009: An enmap preparatory flight campaign, 2016. Available: https://doi.org/10.2312/enmap.2016.0021. 2, 5

[29] Li Pang, Xiangyu Rui, Long Cui, Hongzhong Wang, Deyu Meng, and Xiangyong Cao. Hir-diff: Unsupervised hyperspectral image restoration via improved diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3005–3014, 2024. 5

[30] Kabir Yunus Peerbhay, Onisimo Mutanga, and Riyad Ismail. Commercial tree species discrimination using airborne aisa eagle hyperspectral imagery and partial least squares discriminant analysis (pls-da) in kwazulu–natal, south africa. *ISPRS Journal of Photogrammetry and Remote Sensing*, 79:19–28, 2013. 2, 4, 5

[31] Vaishnav Potlapalli, Syed Waqas Zamir, Salman H Khan, and Fahad Shahbaz Khan. Promptir: Prompting for all-in-one image restoration. *Advances in Neural Information Processing Systems*, 36:71275–71293, 2023. 3, 4, 5, 6

[32] Yuwei Qiu, Kaihao Zhang, Chenxi Wang, Wenhan Luo, Hongdong Li, and Zhi Jin. Mb-taylorformer: Multi-branch efficient transformer expanded by taylor formula for image dehazing. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 12802–12813, 2023. 5

[33] Fu-Jen Tsai, Yan-Tsung Peng, Yen-Yu Lin, Chung-Chi Tsai, and Chia-Wen Lin. Stripformer: Strip transformer for fast image deblurring. In *European Conference on Computer Vision*, pages 146–162. Springer, 2022. 4

[34] Yan Wang, Yusen Li, Gang Wang, and Xiaoguang Liu. Multi-scale attention network for single image super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5950–5960, 2024. 4

[35] 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. 3, 4

[36] Hao Wu and Saurabh Prasad. Convolutional recurrent neural networks for hyperspectral data classification. *Remote Sensing*, 9(3):298, 2017. 1, 4

[37] Jingxiang Yang, Yong-Qiang Zhao, Chen Yi, and Jonathan Cheung-Wai Chan. No-reference hyperspectral image quality assessment via quality-sensitive features learning. *Remote Sensing*, 9(4):305, 2017. 3

[38] CEN Yi, Lifu Zhang, Xia Zhang, WANG Yueming, QI Wenchao, TANG Senlin, and Peng Zhang. Aerial hyperspectral remote sensing classification dataset of xiongan new area (matiwan village). *National Remote Sensing Bulletin*, 24(11): 1299–1306, 2020. 1, 3

[39] Naoto Yokoya and Akira Iwasaki. Airborne hyperspectral data over chikusei. *Space Appl. Lab., Univ. Tokyo, Tokyo, Japan, Tech. Rep. SAL-2016-05-27*, 5(5):5, 2016. 1, 5

[40] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5728–5739, 2022. 5

[41] Mingjin Zhang, Chi Zhang, Qiming Zhang, Jie Guo, Xinbo Gao, and Jing Zhang. Essaformer: Efficient transformer for hyperspectral image super-resolution. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 23073–23084, 2023. 4

[42] Yafei Zhang, Shen Zhou, and Huafeng Li. Depth information assisted collaborative mutual promotion network for single image dehazing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2846–2855, 2024. 5

[43] Yupeng Zhou, Zhen Li, Chun-Le Guo, Song Bai, Ming-Ming Cheng, and Qibin Hou. Srformer: Permuted self-attention for single image super-resolution. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 12780–12791, 2023. 4

[44] Feiyun Zhu. Hyperspectral unmixing: ground truth labeling, datasets, benchmark performances and survey. *arXiv preprint arXiv:1708.05125*, 2017. 1, 4