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*]

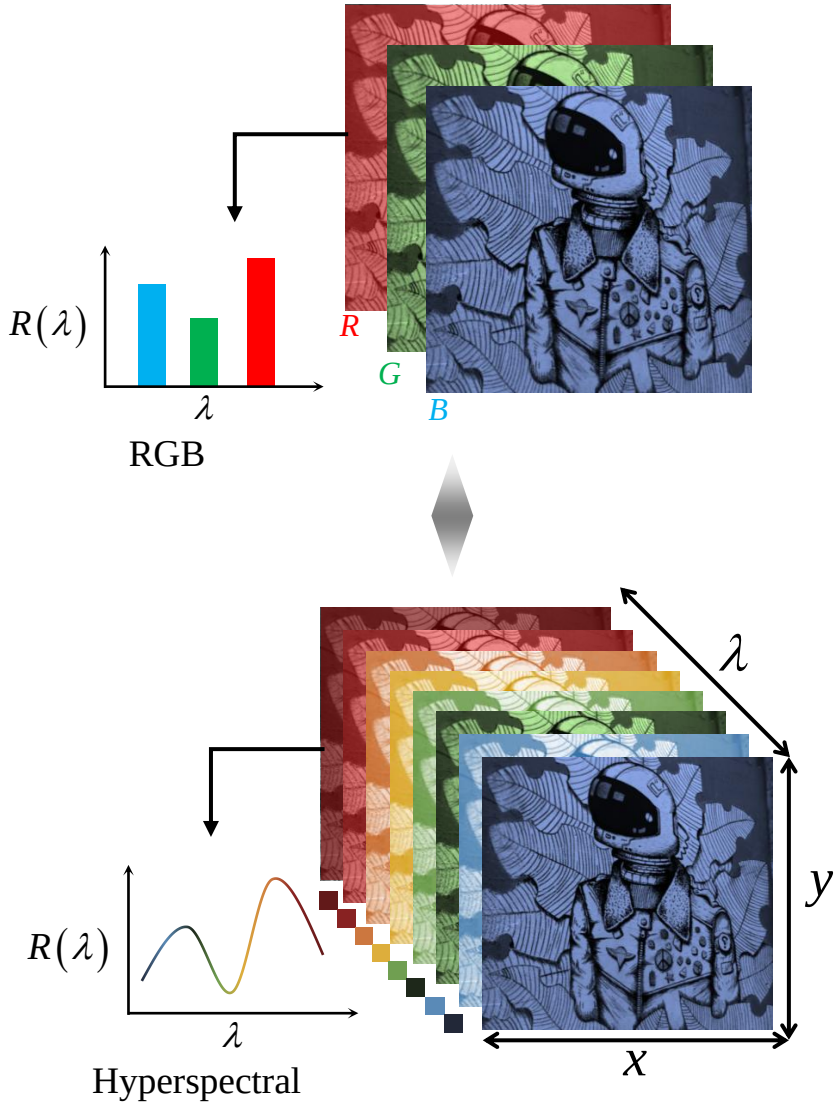
[1]Wuhan University, [2]Jiangxi Normal University, [3]The University of Tokyo,

[4]RIKEN Center for Advanced Intelligence Project

{wuzhehui, weihe1990}@whu.edu.cn, chen yong@jxnu.edu.cn, yokoya@k.u-tokyo.ac.jp

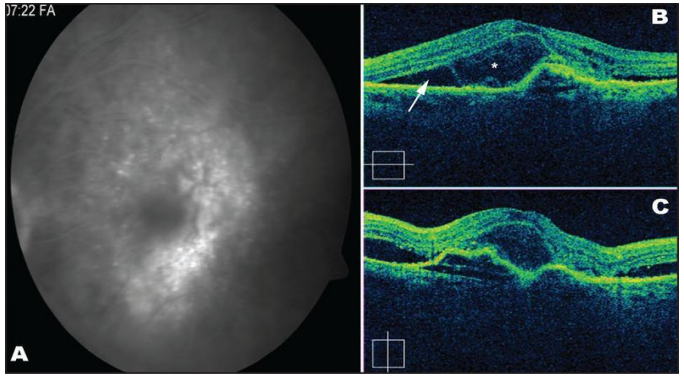
Contents

Background



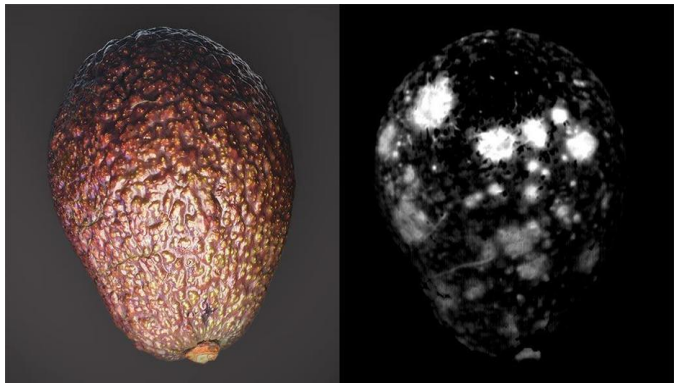
➢ Differences Between HSIs and RGB Images

Agricultural Monitoring



Medical Diagnosis

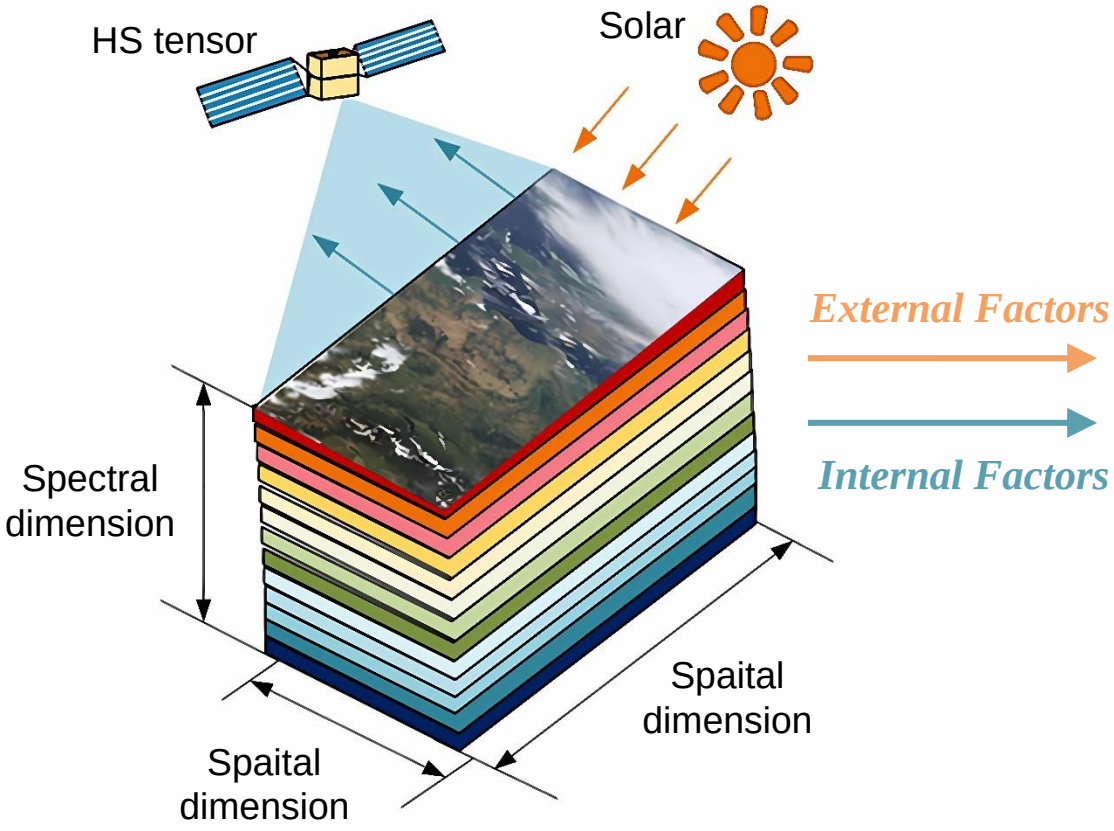
Classification of Scenes

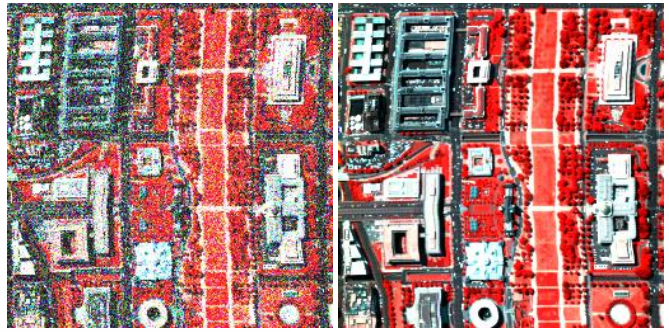


Food Safety Testing

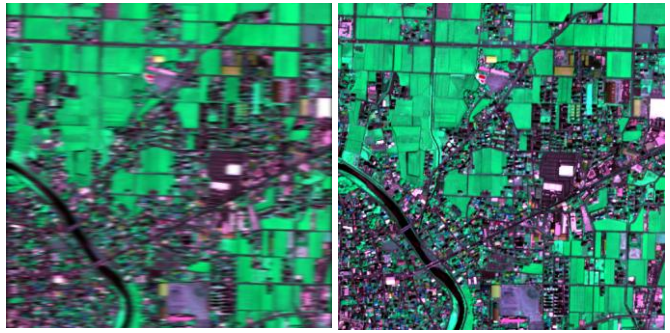
➢ Applications of Hyperspectral Images (HSIs)

HSI Degradation



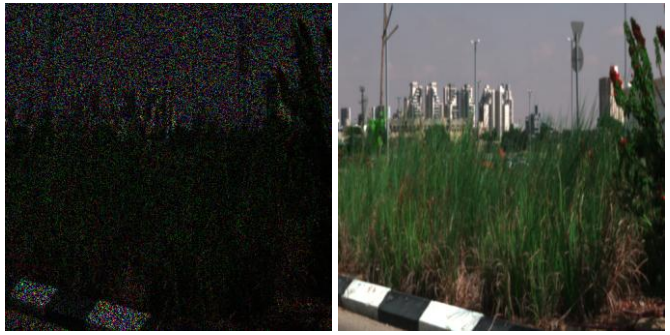


Denoising



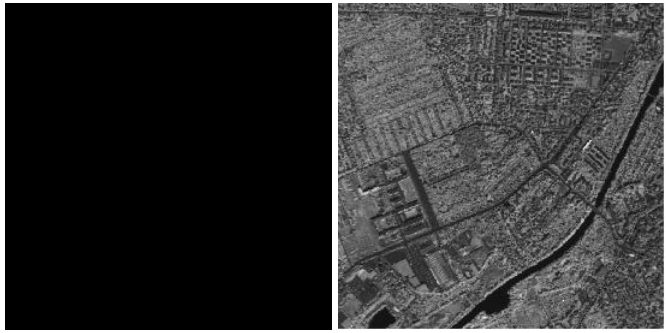
Deblurring

Super Resolution



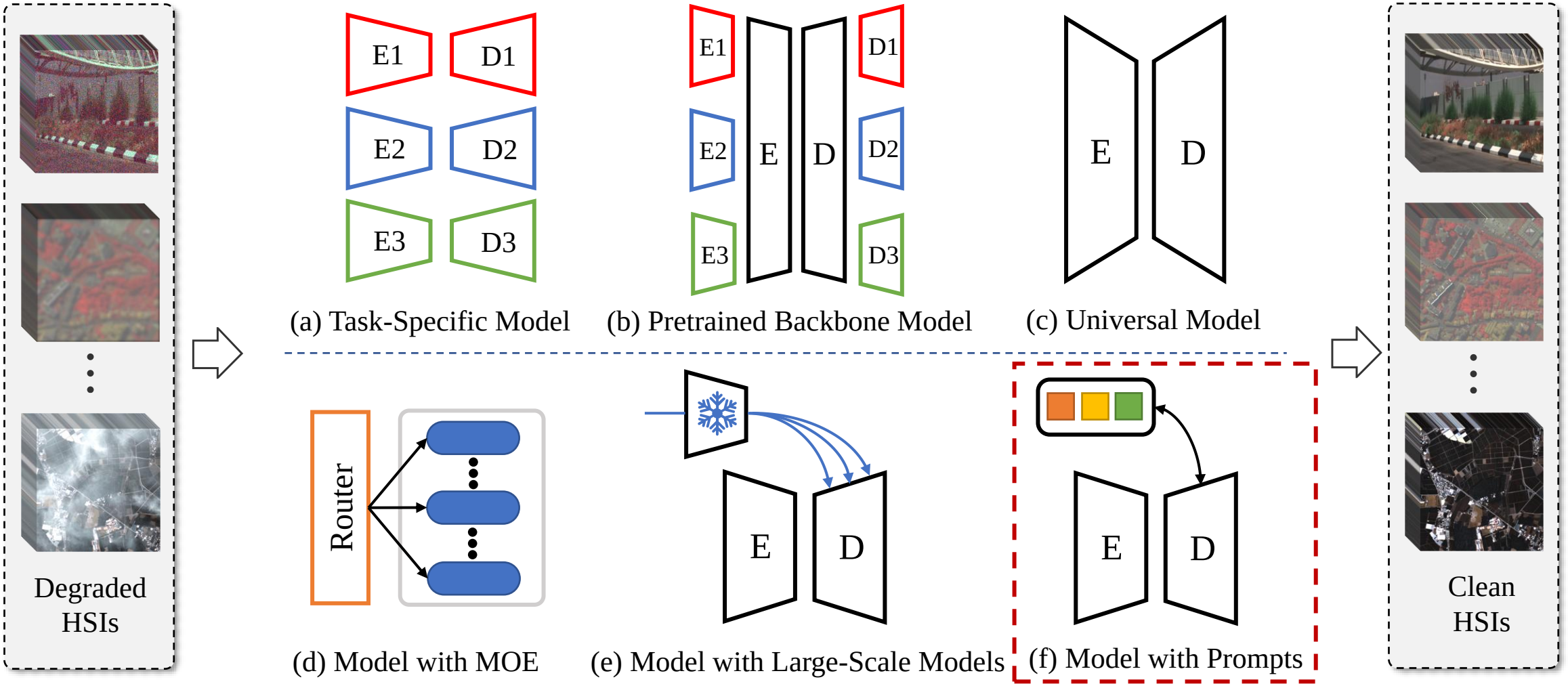
Inpainting

Dehazing

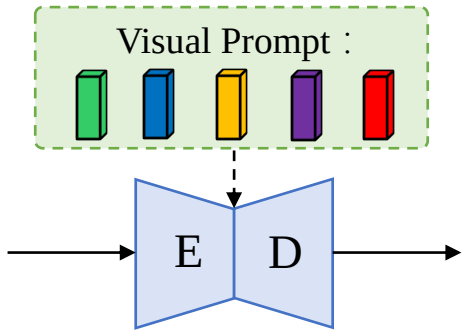


Band Completion

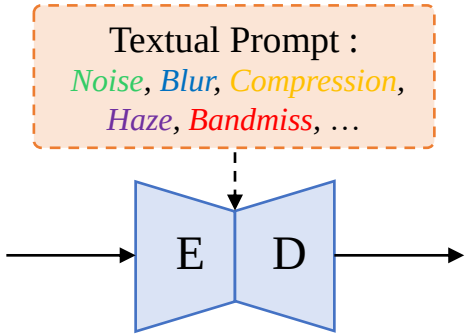
Related Works



Motivation



(a) Model with Visual Prompt

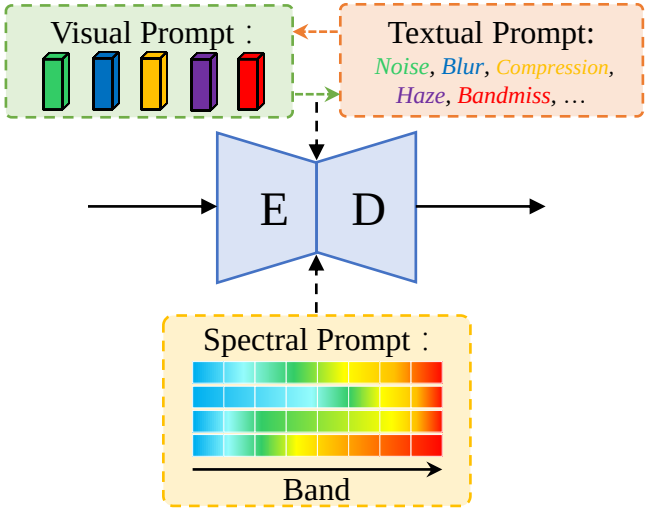


(b) Model with Textual Prompt

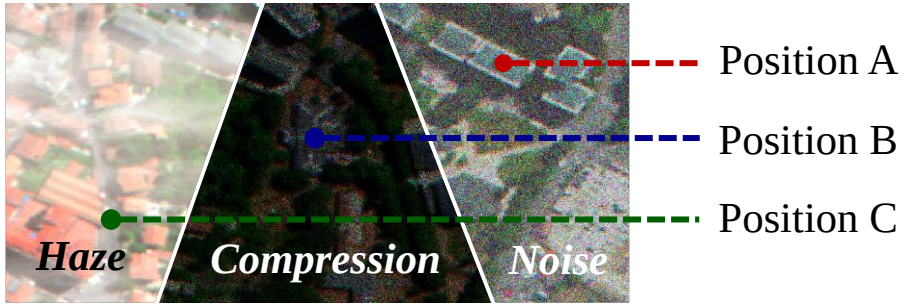
Single-degradation type prediction accuracy of VLMS

- Qwen2.5-VL: 77.6%
- GPT-o4-mini-high: 81.9%
- EarthDial: 83.4%

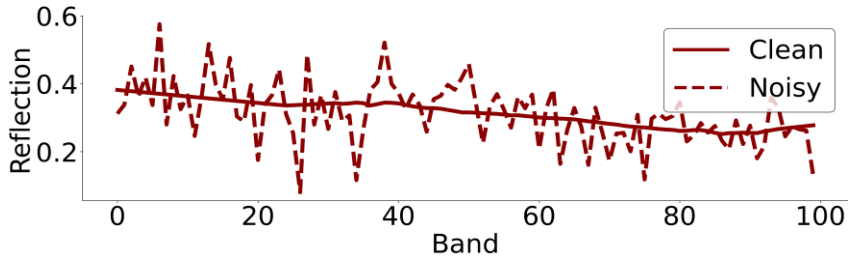
(c) Model with Dual Prompts

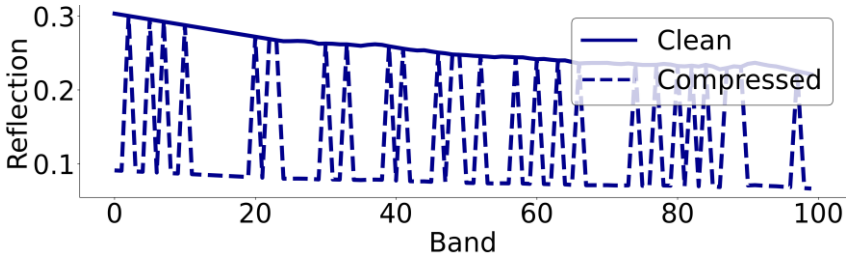


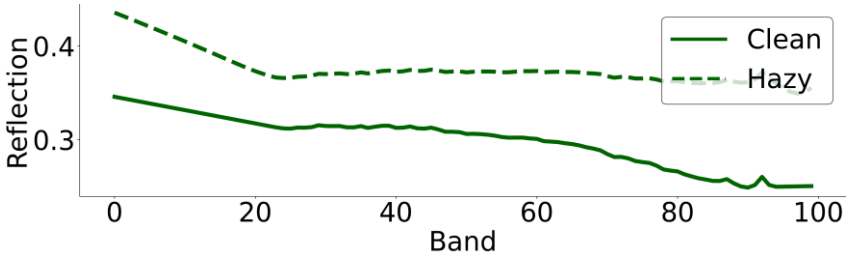
(d) Model with Tri-Modal Prompts



Pavia Center



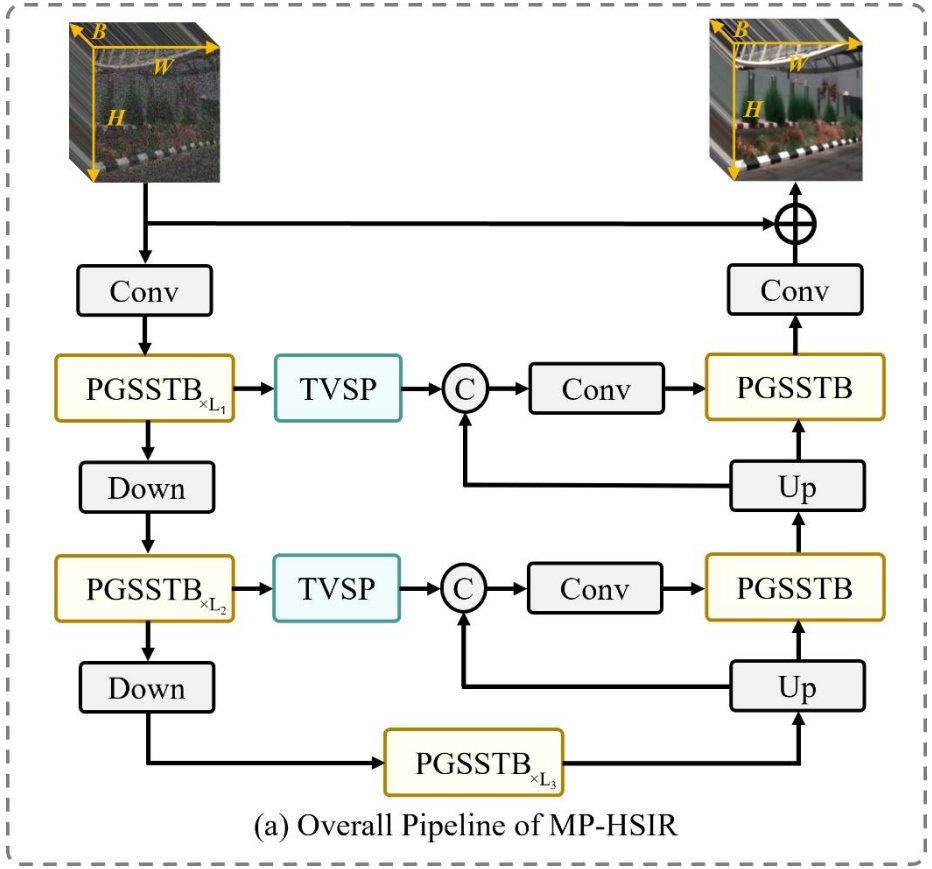




Contents

Methodology

➢ Overall Architecture



PGSSTB	Prompt-Guided Spatial-Spectral Transformer Block
TVSP	* Text-Visual Synergistic Prompt Module
SSA	Spatial Self-Attention
PGSSA	* Prompt-Guided Spectral Self-Attention
GMLP	Gated Multi-Layer Perceptron

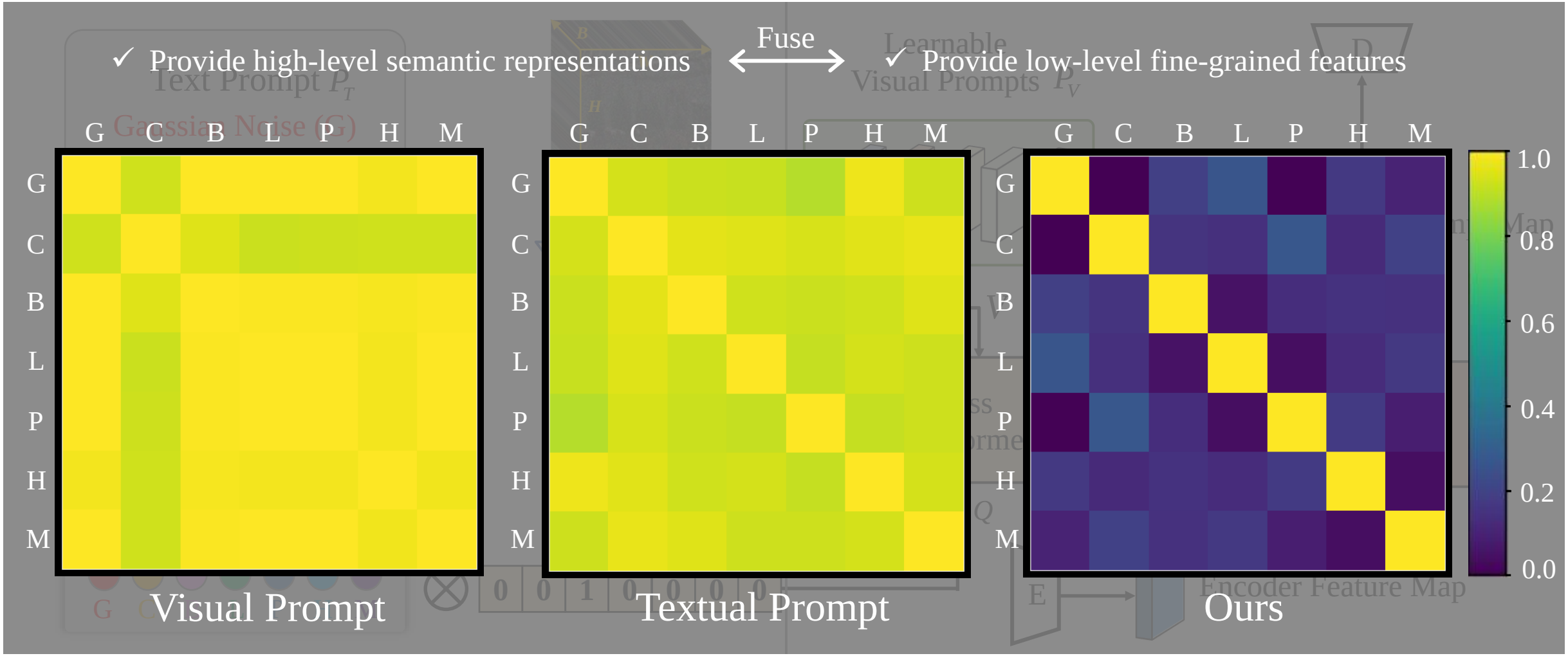
The key part is indicated by *

➢ How to incorporate degradation information?

➢ How to reconstruct spectral information?

Methodology

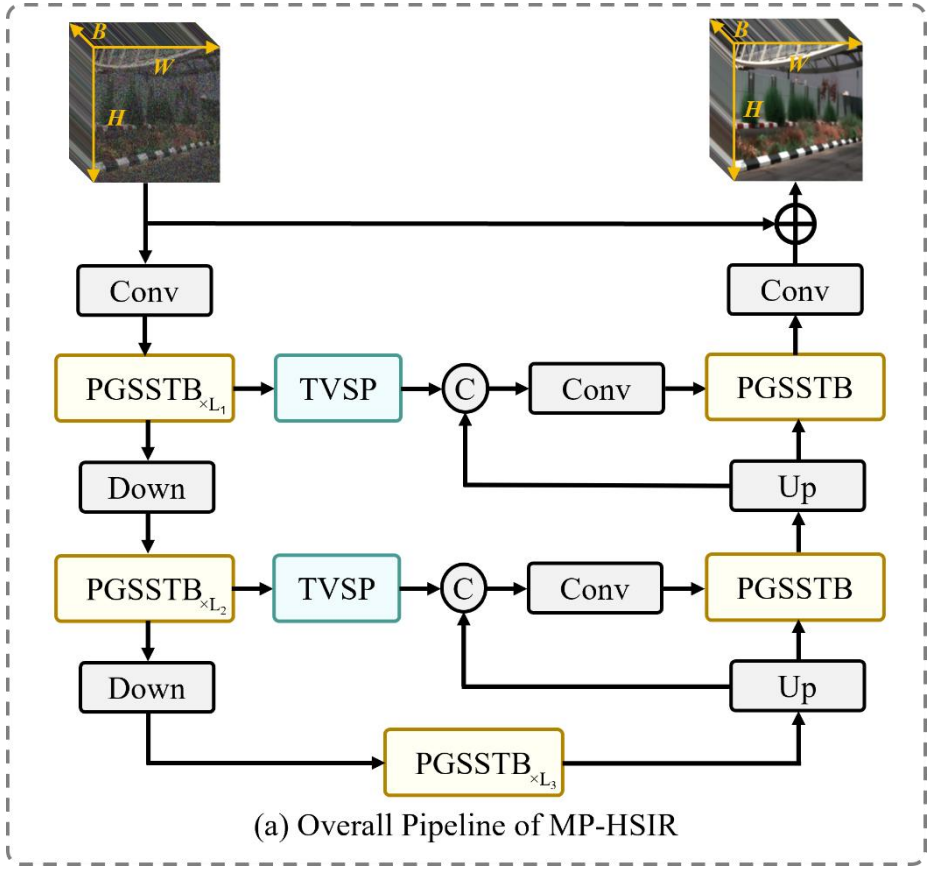
➢ How to incorporate degradation information?



Text-Visual Synergistic Prompt Module (TVSP)

Methodology

➢ Overall Architecture



PGSSTB	Prompt-Guided Spatial-Spectral Transformer Block
TVSP	* Text-Visual Synergistic Prompt Module
SSA	Spatial Self-Attention
PGSSA	* Prompt-Guided Spectral Self-Attention
GMLP	Gated Multi-Layer Perceptron

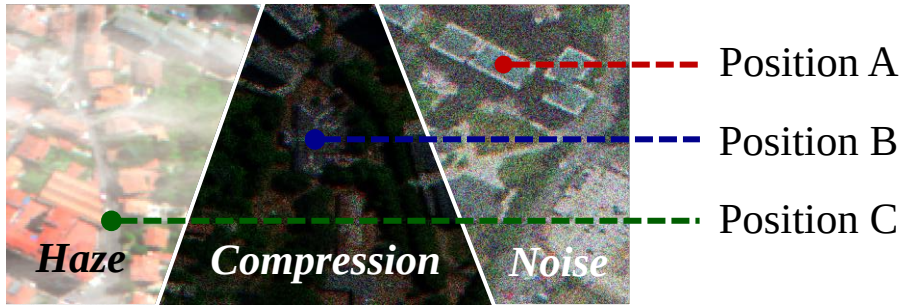
The key part is indicated by *

➢ How to incorporate degradation information?

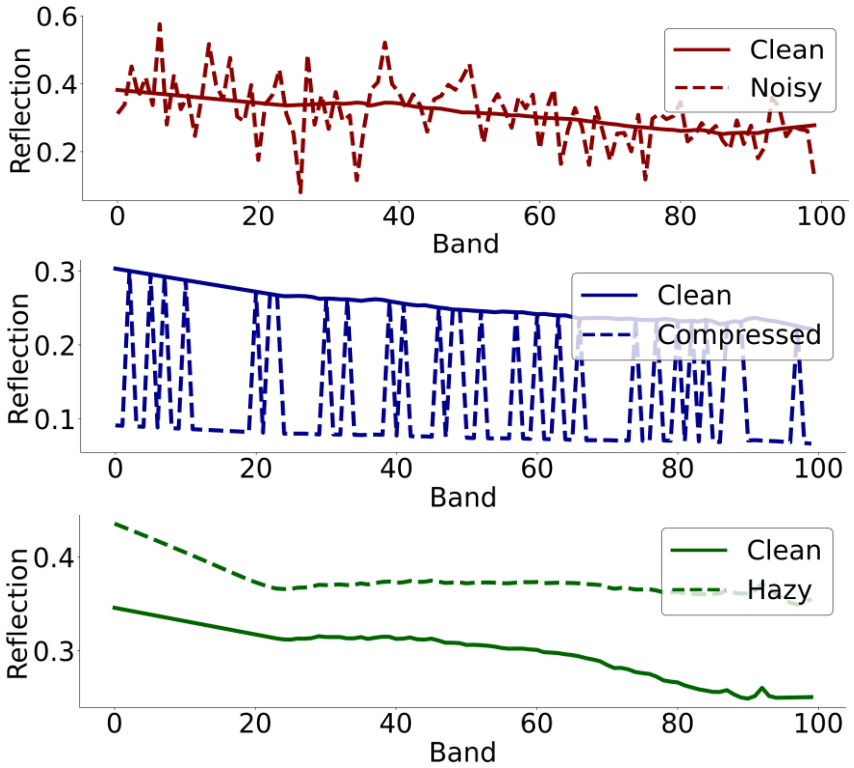
➢ How to reconstruct spectral information?

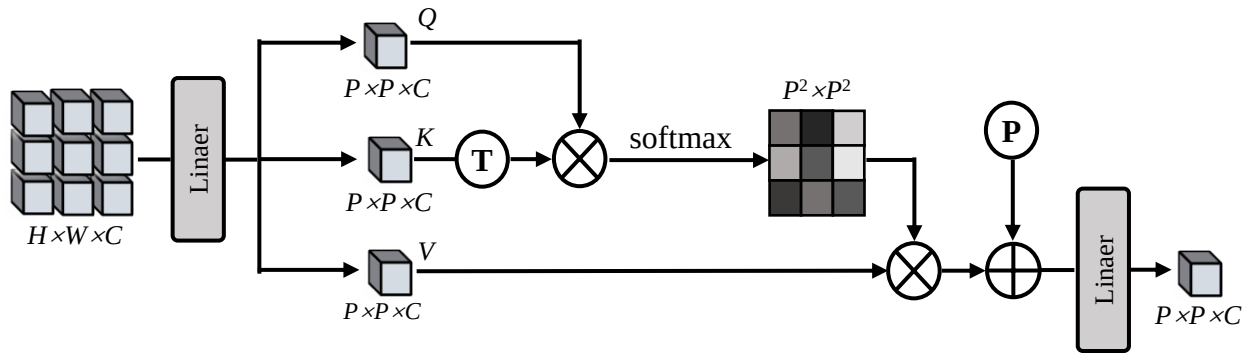
Methodology

➢ How to reconstruct spectral information?

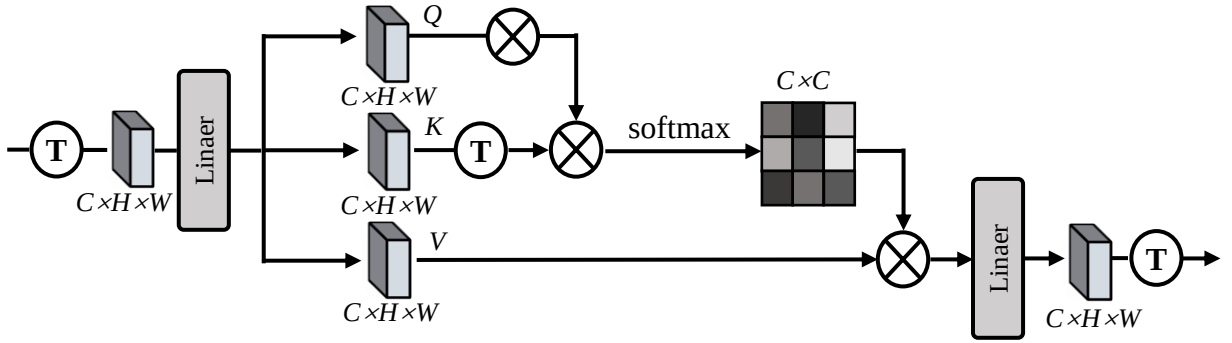


Pavia Center





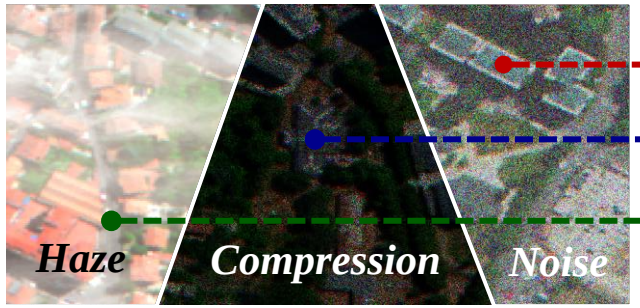
Spatial Self-Attention



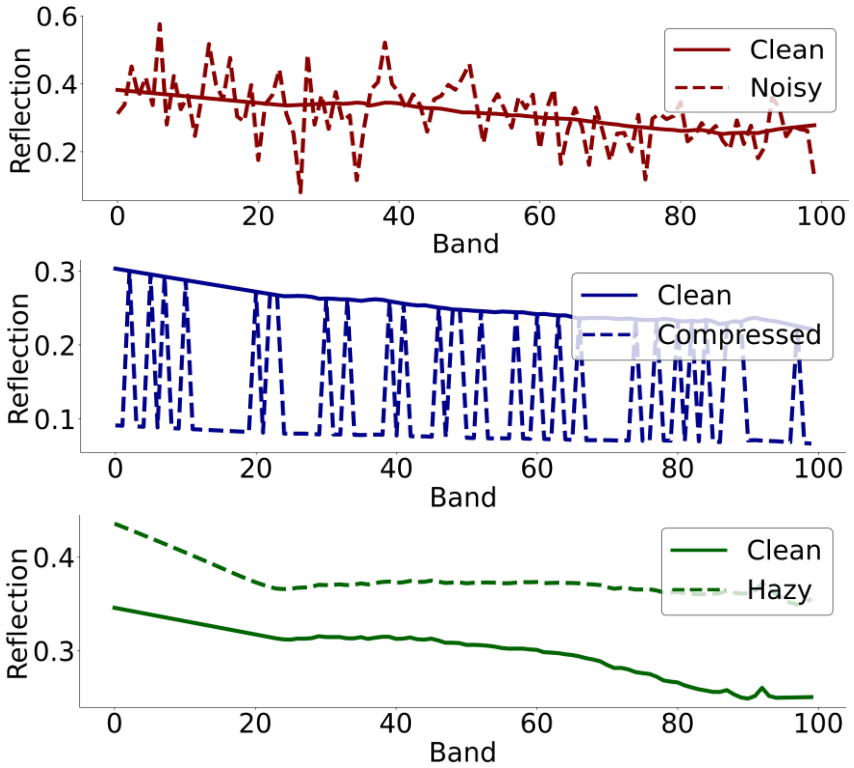
Spectral Self-Attention

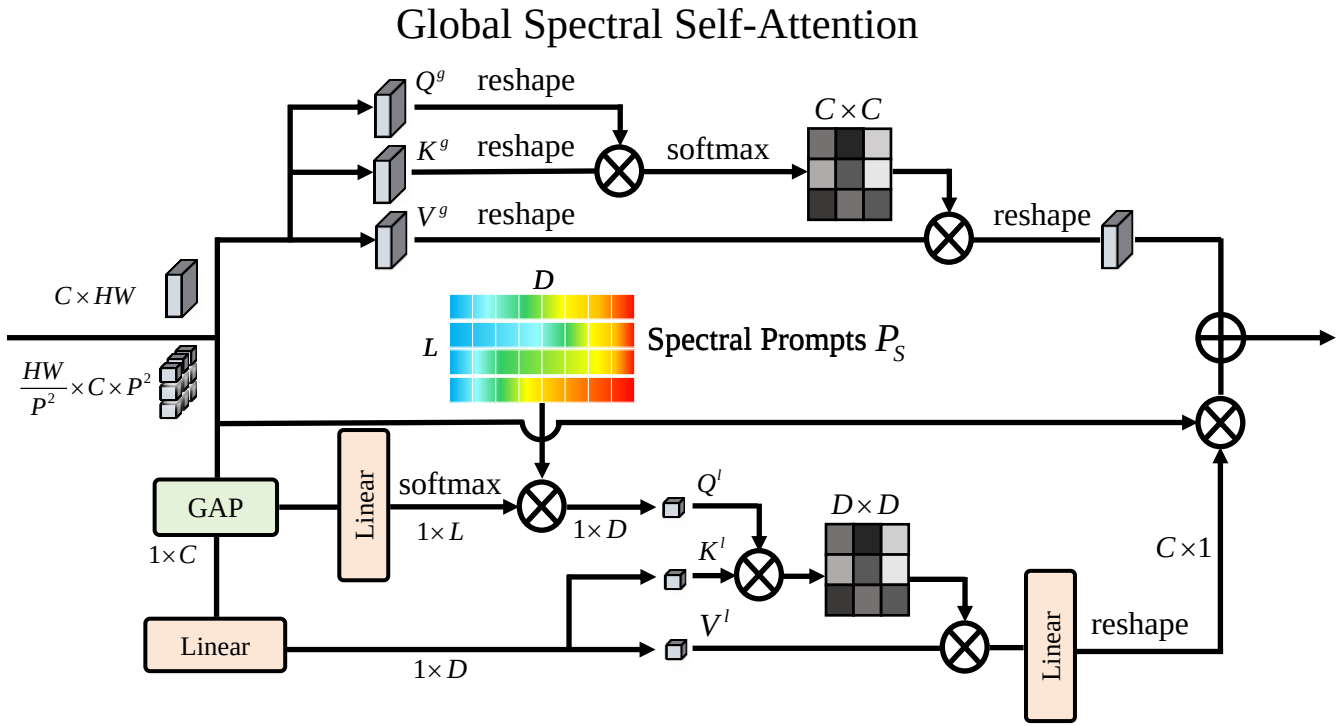
Methodology

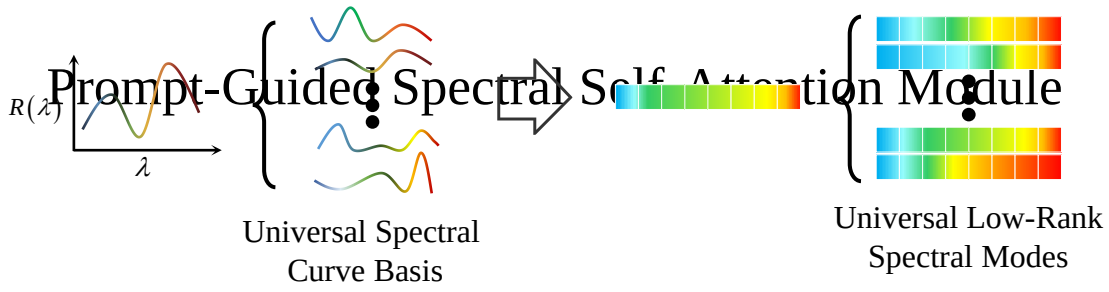
➢ How to reconstruct spectral information?



Pavia Center

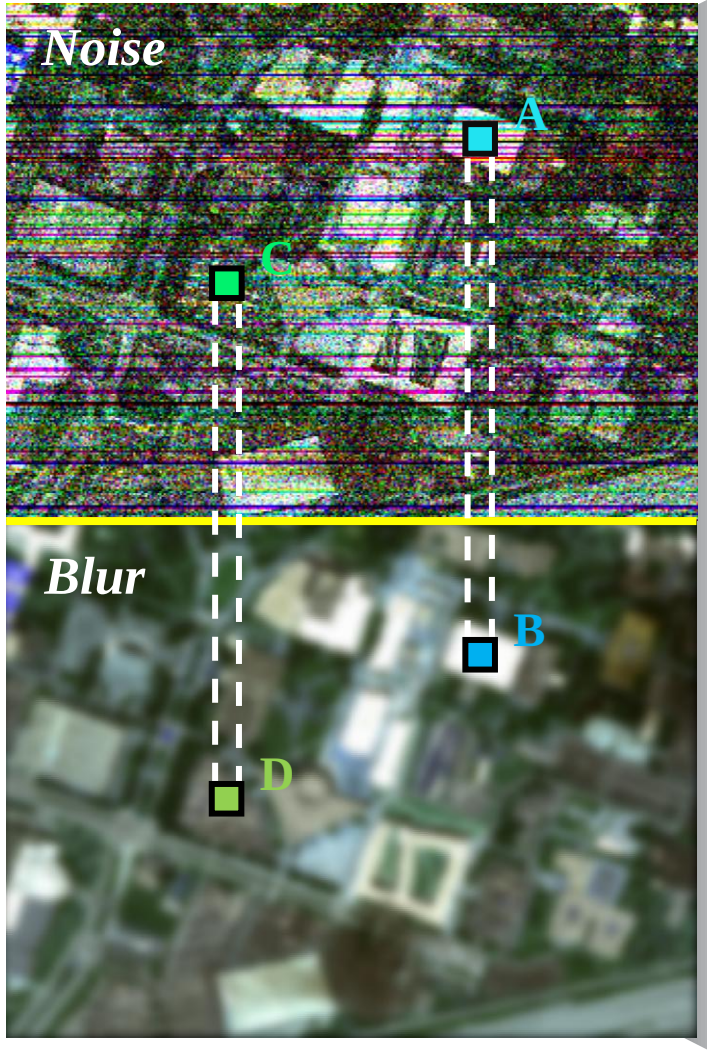


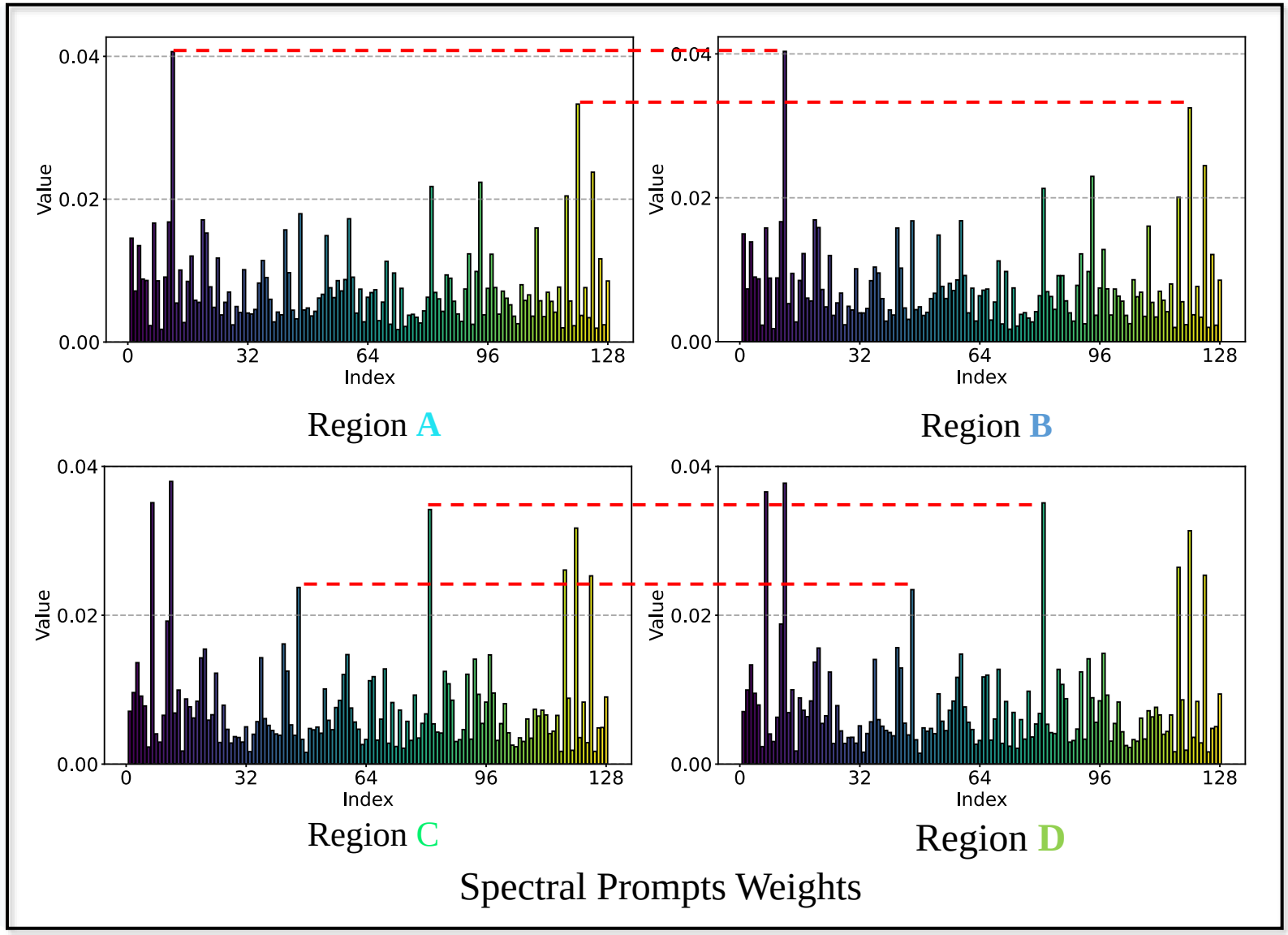




Methodology

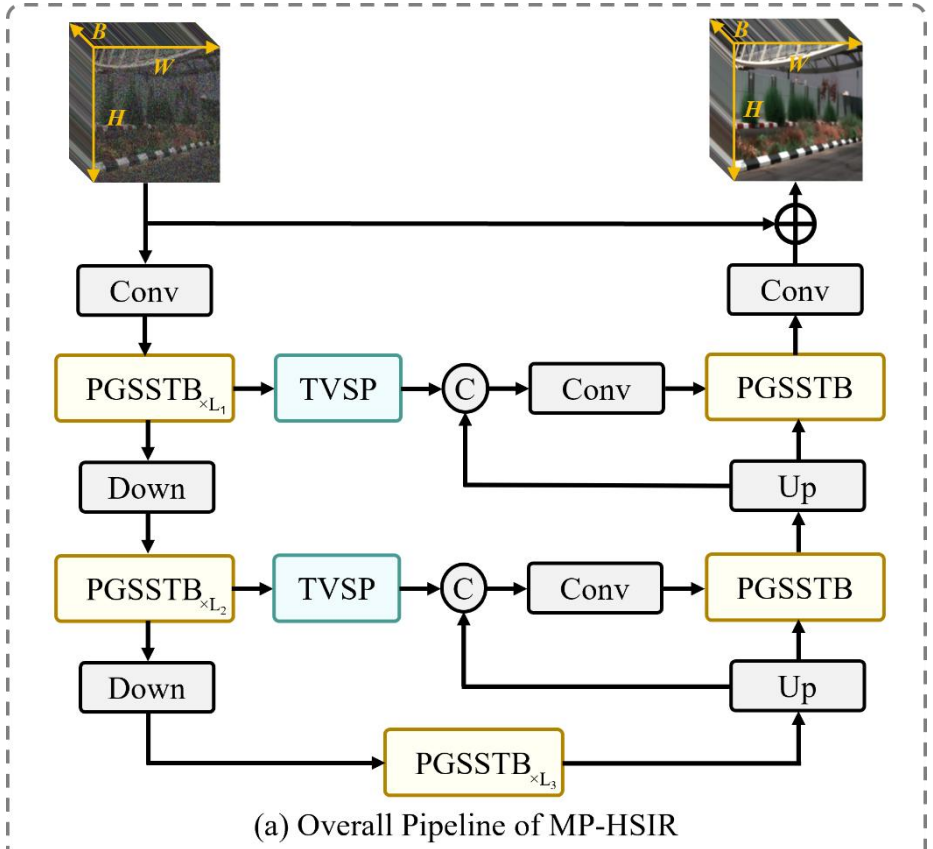
➢ How to reconstruct spectral information?



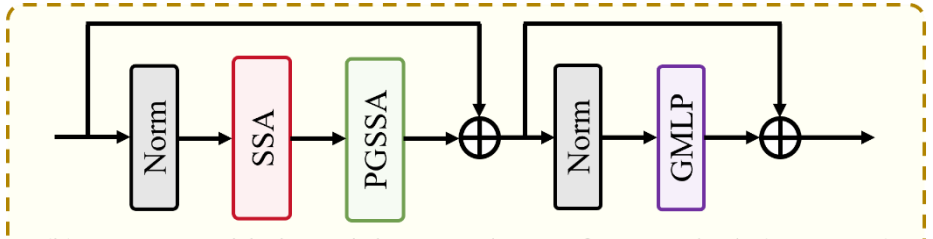


Methodology

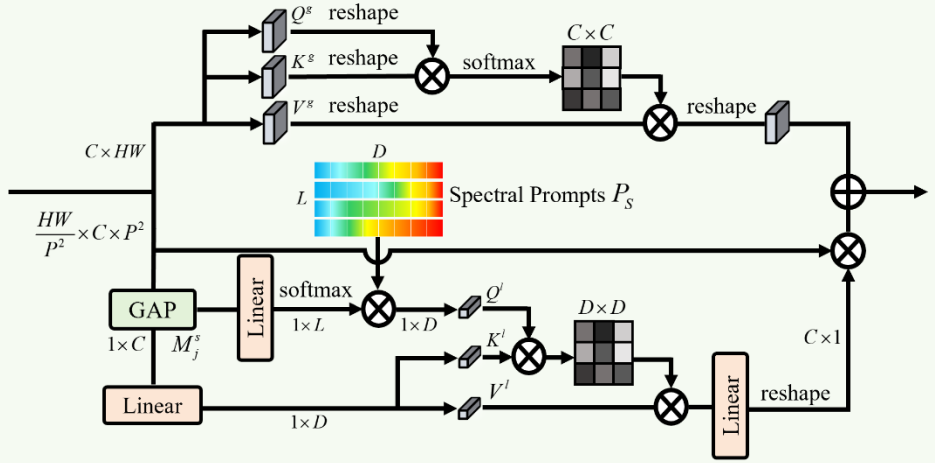
➢ Overall Architecture



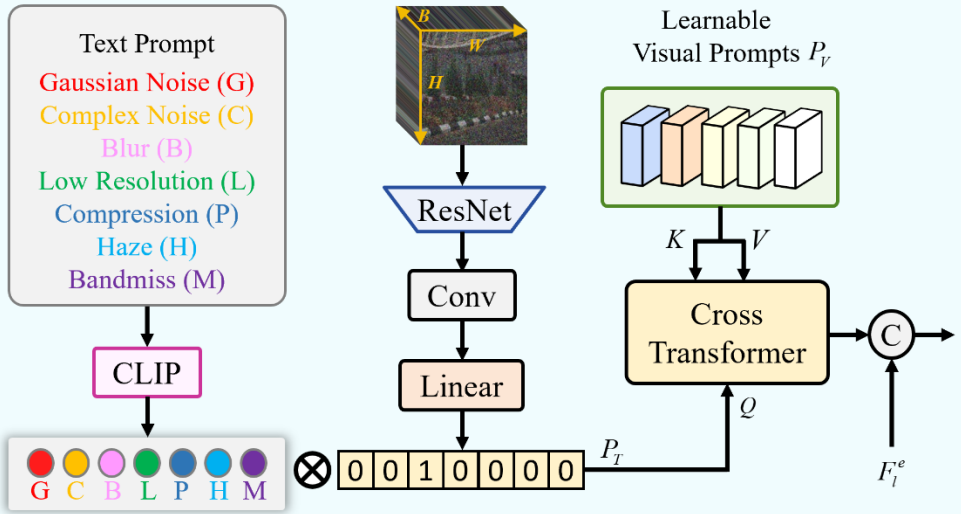
(a) Overall Pipeline of MP-HSIR



(b) Prompt-Guided Spatial-Spectral Transformer Block (PGSSTB)



(d) Prompt-Guided Spectral Self-Attention Module (PGSSA)



(c) Text-Visual Synergistic Prompt Module (TVSP)

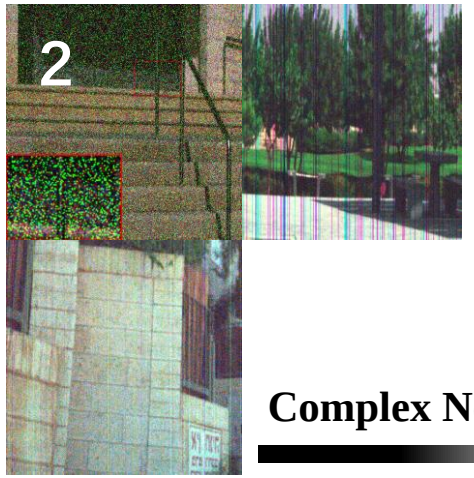
Contents

Datasets

Table 1. Properties of 13 Natural Scene and Remote Sensing Hyperspectral Datasets.

	Dataset	Sensor	Wavelength (nm)	Channels	Size	Amount	GSD (m)
Natural Scene	ARAD_1K	Specim IQ	400–700	31	482×512	1000	/
	ICVL	Specim PS Kappa DX4	400–700	31	1392×1300	201	/
Remote Sensing	Xiong'an	Unknown	400–1000	256	3750×1580	1	0.5
	WDC	Hydice	400–2400	191	1208×307	1	5
	PaviaC	ROSIS	430–860	102	1096×715	1	1.3
	PaviaU	ROSIS	430–860	103	610×340	1	1.3
	Houston	ITRES CASI-1500	364–1046	144	349×1905	1	2.5
	Chikusei	HH-VNIR-C	343–1018	128	2517×2335	1	2.5
	Eagle	AsiaEAGLE II	401–999	128	2082×1606	1	1
	Berlin	Unknown	455–2447	111	6805×1830	1	3.6
	Urban	Hydice	400-2500	210	307×307	1	2
	APEX	Unknown	350-2500	285	1000×1500	1	2
	EO-1	Hyperion	357-2567	242	3471×991	1	30

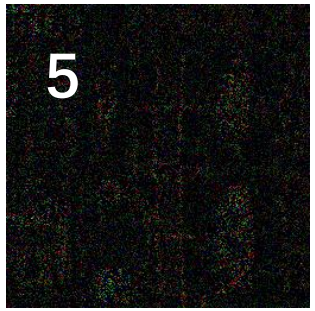
All-in-One Experiments



Complex Noise

Blur

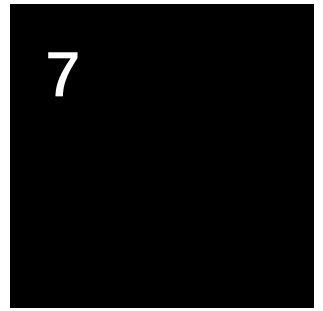
Low Resolution



Inpaint

Haze

Gaussian Noise



Bandmiss

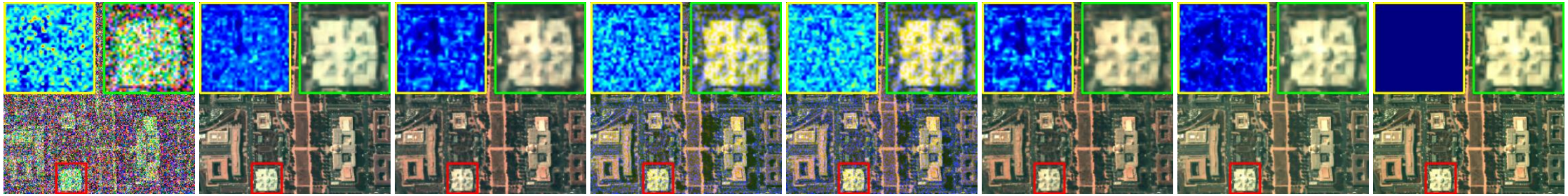
Experimental Settings

Loss：L1
optimizer：AdamW ($\beta1 = 0.9, \beta2 = 0.999$)
Batch_size：32
Learning rate：2×10^{-4} / 1×10^{-4} → 1×10^{-6}
Epochs：100 / 300
Patch_size：64×64

All-in-One Results

Table 2. **[All-in-one]** Quantitative comparison of all-in-one and state-of-the-art task-specific methods on 7 HSI restoration tasks. ***Gaussian Denoising*** & ***Complex Denoising***

Type	Method	*Gaussian Denoising (Sigma = 30, 50, 70)*			Method	*Complex Denoising (Case = 1, 2, 3, 4)*		
		ICVL [3]	ARAD [4]	Xiong'an [68]		ICVL [3]	ARAD [4]	WDC [77]
		PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑		PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑
Task Specific	QRNN3D [60]	39.99 / 0.947	39.18 / 0.932	36.06 / 0.829	QRNN3D [60]	41.60 / 0.966	41.11 / 0.960	30.82 / 0.868
	SST [31]	41.31 / 0.959	40.86 / 0.955	37.53 / 0.850	SST [31]	42.43 / 0.971	41.95 / 0.967	32.71 / 0.889
	SERT [32]	41.55 / 0.967	41.09 / 0.959	37.83 / 0.859	SERT [32]	43.31 / 0.976	42.87 / 0.973	33.31 / 0.903
	LDERT [33]	41.92 / 0.969	41.47 / 0.965	38.14 / 0.865	LDERT [33]	43.42 / 0.977	43.02 / 0.974	33.49 / 0.904
All in One	AirNet [30]	39.76 / 0.943	39.19 / 0.920	31.94 / 0.668	AirNet [30]	40.68 / 0.959	40.09 / 0.945	28.07 / 0.726
	PromptIR [50]	40.25 / 0.953	39.69 / 0.945	32.99 / 0.684	PromptIR [50]	41.29 / 0.965	40.71 / 0.954	28.83 / 0.735
	PIP [34]	40.87 / 0.958	40.25 / 0.952	32.64 / 0.674	PIP [34]	41.67 / 0.968	41.17 / 0.962	28.50 / 0.724
	HAIR [7]	40.51 / 0.956	39.91 / 0.949	32.54 / 0.679	HAIR [7]	40.63 / 0.958	40.15 / 0.947	28.18 / 0.729
	InstructIR [12]	41.02 / 0.960	40.32 / 0.954	31.74 / 0.666	InstructIR [12]	40.12 / 0.957	39.81 / 0.952	27.49 / 0.707
	PromptHSI [28]	40.65 / 0.960	40.04 / 0.956	38.07 / 0.881	PromptHSI [28]	39.14 / 0.955	38.75 / 0.936	33.77 / 0.912
	MP-HSIR (Ours)	41.62 / 0.964	41.09 / 0.960	38.81 / 0.897	MP-HSIR (Ours)	42.29 / 0.971	41.99 / 0.969	34.07 / 0.918



Input / 14.07 | SST / 30.53 | LDERT / 30.74 | PromptIR / 26.38 | InstructIR / 24.63 | PromptHSI / 30.78 | Ours / 31.36 | GT / PSNR (dB)

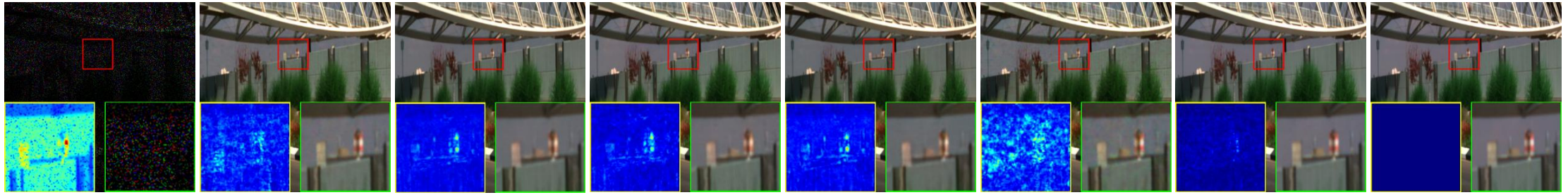
Figure 2. **[All-in-one]** Visual comparison results on ***Complex Denoising***.

All-in-One Results

Type	Method	*Gaussian Deblurring (Radius = 9, 15, 21 & 7, 11, 15)*			Method	*Super-Resolution (Scale = 2, 4, 8)*		
		ICVL [3]	PaviaC [22]	Eagle [49]		ARAD [4]	PaviaU [22]	Houston [63]
		PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑		PSNR / SSIM ↑	PSNR / SSIM ↑	PSNR / SSIM ↑
Task Specific	Stripformer [55]	46.03 / 0.988	37.13 / 0.913	41.96 / 0.958	SNLSR [21]	36.05 / 0.898	30.55 / 0.730	31.70 / 0.787
	FFTformer [25]	46.65 / 0.988	37.96 / 0.921	42.76 / 0.962	MAN [57]	36.88 / 0.911	30.92 / 0.733	32.03 / 0.791
	LoFormer [45]	47.15 / 0.989	37.72 / 0.917	42.59 / 0.962	ESSAformer [74]	37.40 / 0.918	31.34 / 0.738	32.37 / 0.793
	MLWNet [13]	47.66 / 0.990	39.01 / 0.928	44.04 / 0.969	SRFormer [76]	38.02 / 0.922	31.80 / 0.751	32.80 / 0.803
All in One	AirNet [30]	47.21 / 0.989	37.76 / 0.918	42.42 / 0.963	AirNet [30]	36.73 / 0.910	30.93 / 0.735	32.00 / 0.791
	PromptIR [50]	47.67 / 0.990	38.72 / 0.927	43.74 / 0.968	PromptIR [50]	37.37 / 0.918	31.53 / 0.746	32.73 / 0.799
	PIP [34]	47.52 / 0.990	38.52 / 0.924	43.17 / 0.967	PIP [34]	38.36 / 0.926	31.77 / 0.749	32.89 / 0.805
	HAIR [7]	46.45 / 0.988	37.93 / 0.919	42.77 / 0.964	HAIR [7]	36.84 / 0.915	31.47 / 0.745	32.54 / 0.796
	InstructIR [12]	34.05 / 0.798	27.03 / 0.602	32.68 / 0.766	InstructIR [12]	36.51 / 0.913	31.34 / 0.741	32.52 / 0.793
	PromptHSI [28]	32.52 / 0.817	36.73 / 0.912	39.75 / 0.954	PromptHSI [28]	35.03 / 0.904	30.75 / 0.732	31.85 / 0.775
	MP-HSIR (Ours)	48.07 / 0.990	39.00 / 0.927	43.72 / 0.968	MP-HSIR (Ours)	38.25 / 0.924	31.97 / 0.760	33.06 / 0.810

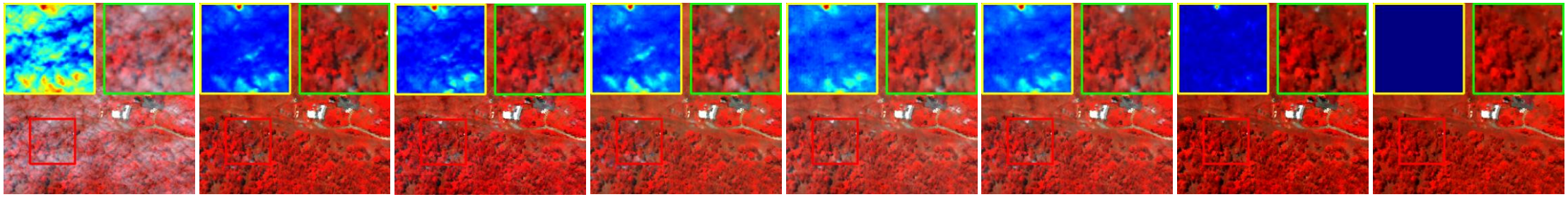
Method	*Inpainting (Mask Rate = 0.7, 0.8, 0.9)*		Method	*Dehazing (Omega = 0.5, 0.75, 1)*		Method	*Completion (Rate = 0.1, 0.2 0.3)*	
	ICVL [3]	Chikusei [71]		PaviaU [24]	Eagle [51]		ARAD [4]	Berlin [49]
	PSNR / SSIM ↑	PSNR / SSIM ↑		PSNR / SSIM ↑	PSNR / SSIM ↑		PSNR / SSIM ↑	PSNR / SSIM ↑
NAFNet [8]	44.39 / 0.987	39.82 / 0.953	SGNet [45]	34.28 / 0.963	37.22 / 0.976	NAFNet [8]	47.04 / 0.995	38.35 / 0.907
Restormer [74]	45.79 / 0.990	36.33 / 0.899	SCANet [18]	36.59 / 0.978	39.64 / 0.985	Restormer [74]	48.34 / 0.995	35.07 / 0.606
DDS2M [48]	42.18 / 0.969	34.94 / 0.887	MB-Taylor [53]	37.99 / 0.983	41.03 / 0.991	SwinIR [37]	49.75 / 0.995	35.45 / 0.886
HIR-Diff [50]	38.91 / 0.949	37.65 / 0.916	DCMPNet [77]	37.20 / 0.985	40.24 / 0.990	MambaIR [16]	50.26 / 0.995	36.12 / 0.888
AirNet [32]	42.60 / 0.981	37.46 / 0.919	AirNet [32]	35.59 / 0.965	38.83 / 0.981	AirNet [32]	45.27 / 0.992	35.91 / 0.624
PromptIR [52]	46.38 / 0.990	38.07 / 0.930	PromptIR [52]	37.41 / 0.982	40.73 / 0.992	PromptIR [52]	46.60 / 0.994	40.45 / 0.652
PIP [36]	43.37 / 0.982	38.43 / 0.930	PIP [36]	37.62 / 0.982	40.74 / 0.990	PIP [36]	47.36 / 0.993	37.20 / 0.668
HAIR [7]	44.02 / 0.982	38.05 / 0.927	HAIR [7]	36.76 / 0.978	40.86 / 0.992	HAIR [7]	45.08 / 0.992	38.01 / 0.650
InstructIR [12]	44.07 / 0.986	36.11 / 0.907	InstructIR [12]	34.72 / 0.971	37.65 / 0.982	InstructIR [12]	51.31 / 0.997	35.97 / 0.580
PromptHSI [30]	41.48 / 0.972	37.33 / 0.946	PromptHSI [30]	36.77 / 0.974	39.78 / 0.984	PromptHSI [30]	47.34 / 0.994	43.36 / 0.972
MP-HSIR (Ours)	51.53 / 0.996	43.63 / 0.979	MP-HSIR (Ours)	39.59 / 0.986	42.41 / 0.995	MP-HSIR (Ours)	56.48 / 0.999	49.50 / 0.987

All-in-One Results



Input / 14.20 | NAFNet / 40.82 | Restormer / 42.16 | PromptIR / 42.65 | InstructIR / 41.43 | PromptHSI / 38.52 | Ours / 47.67 | GT / PSNR (dB)

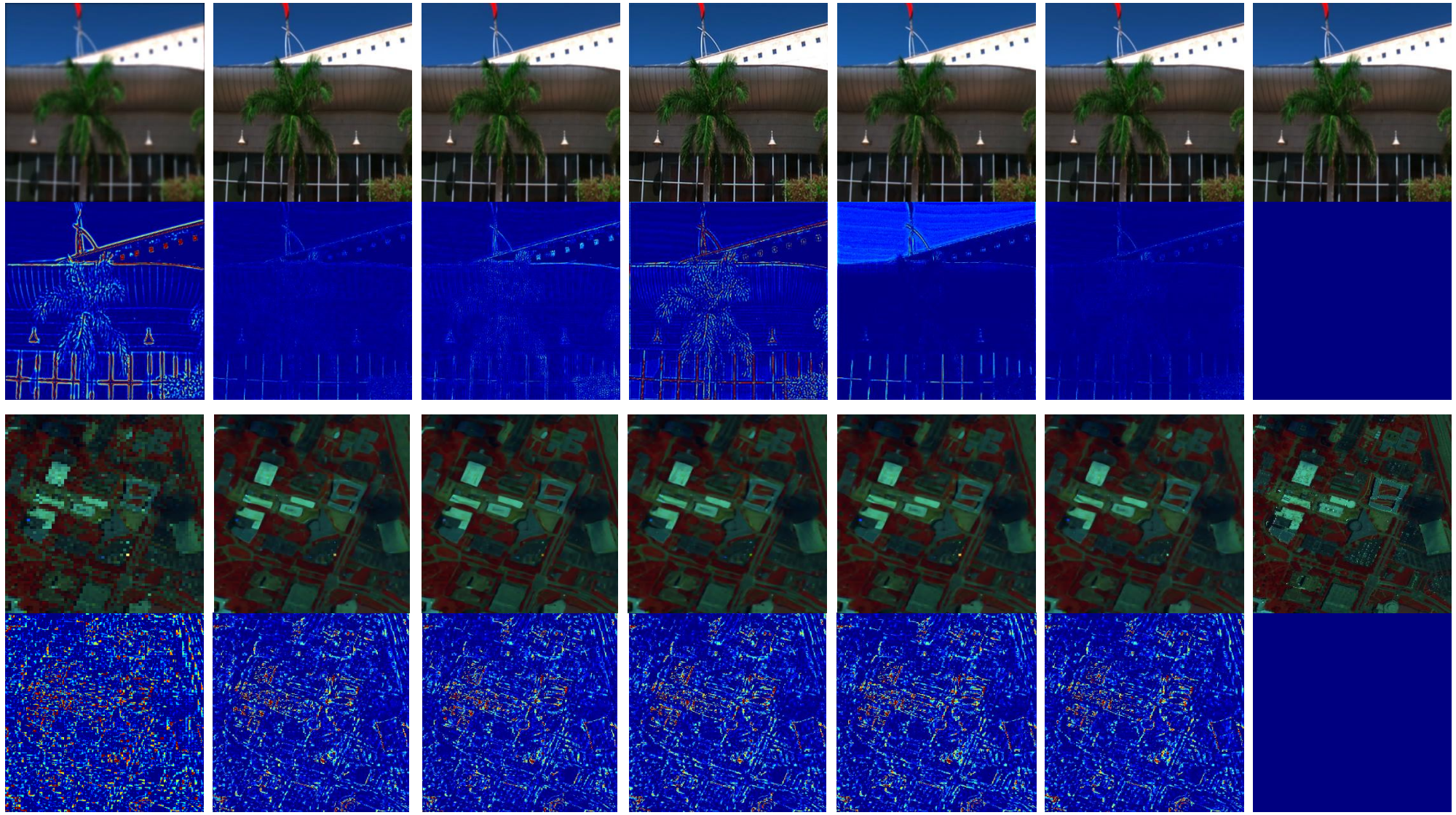
Figure 3. **[All-in-one]** Visual comparison results on ***Inpainting***.



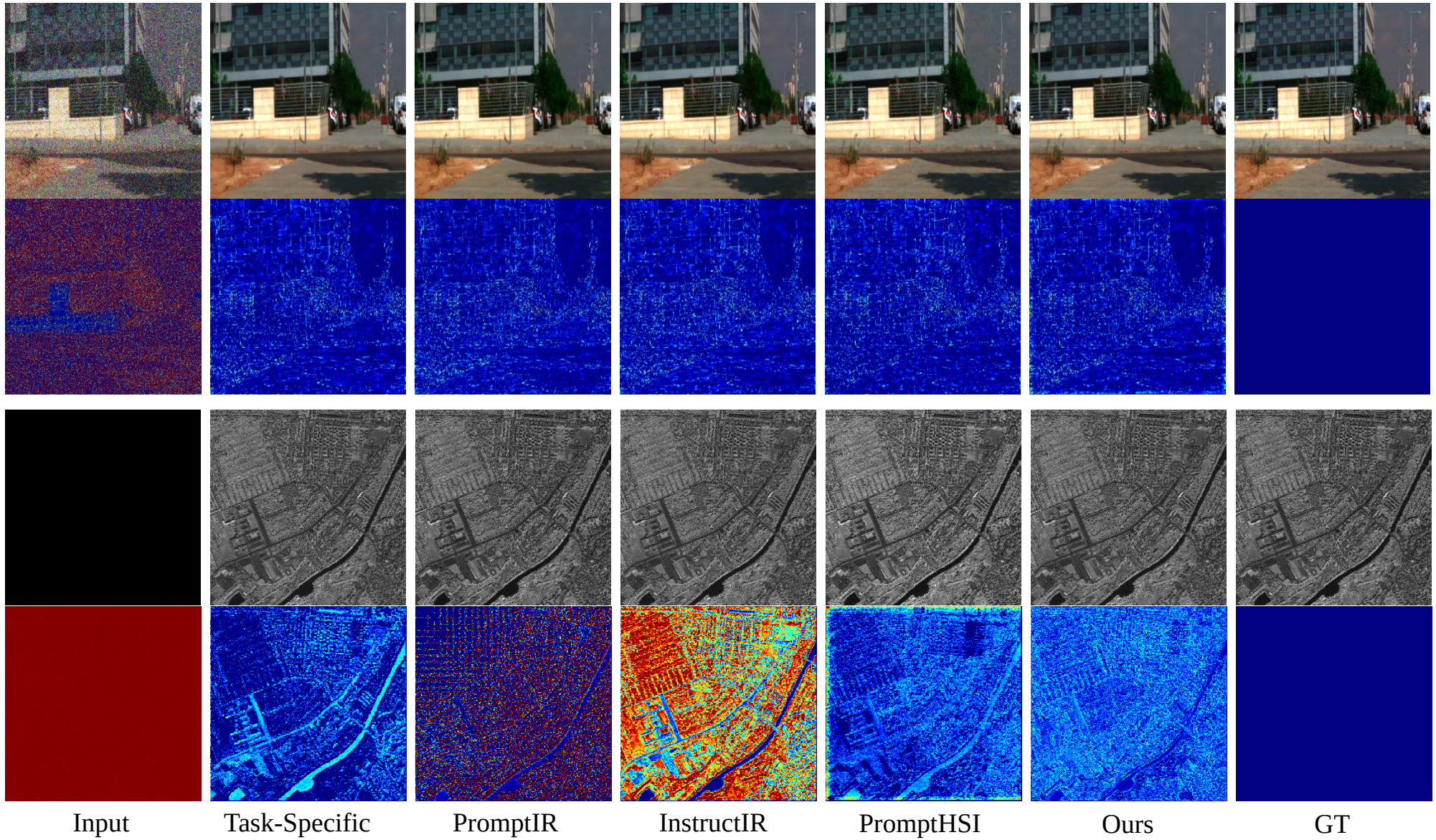
Input / 18.24 | MB-Taylor / 38.34 | DCMPNet / 37.68 | PromptIR / 37.35 | InstructIR / 33.76 | PromptHSI / 37.82 | Ours / 39.49 | GT / PSNR (dB)

Figure 4. **[All-in-one]** Visual comparison results on ***Dehazing***.

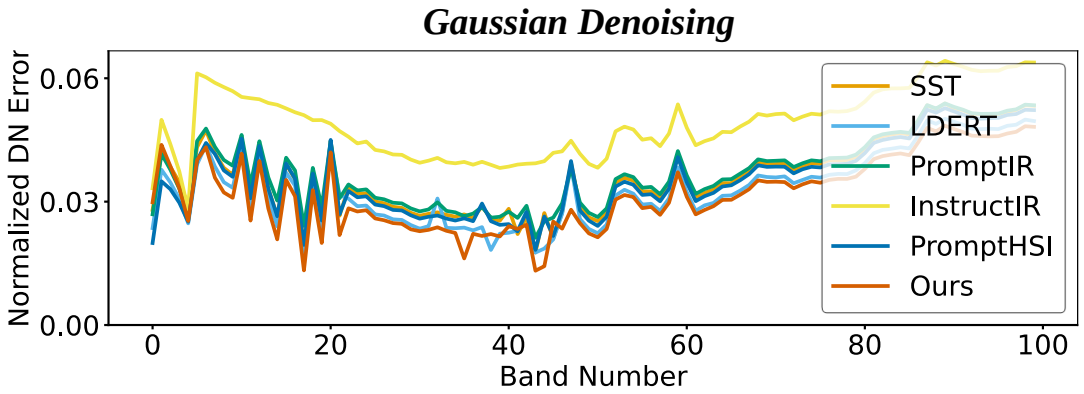
All-in-One Results

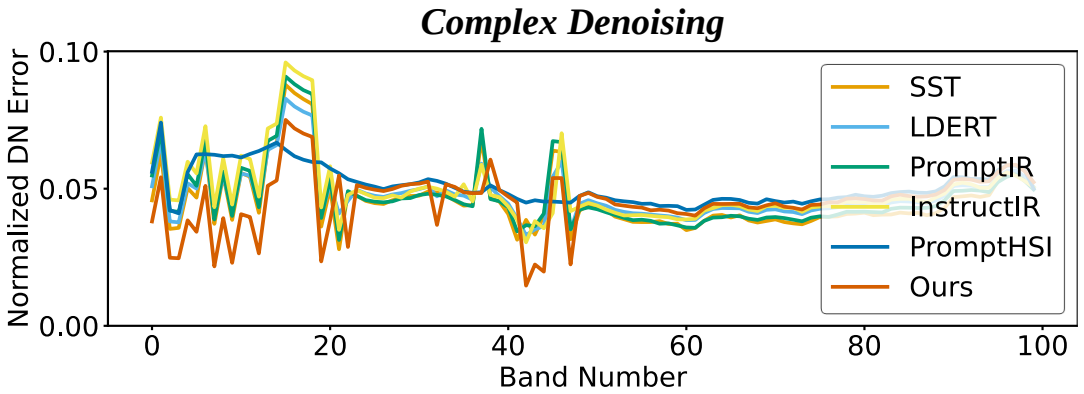


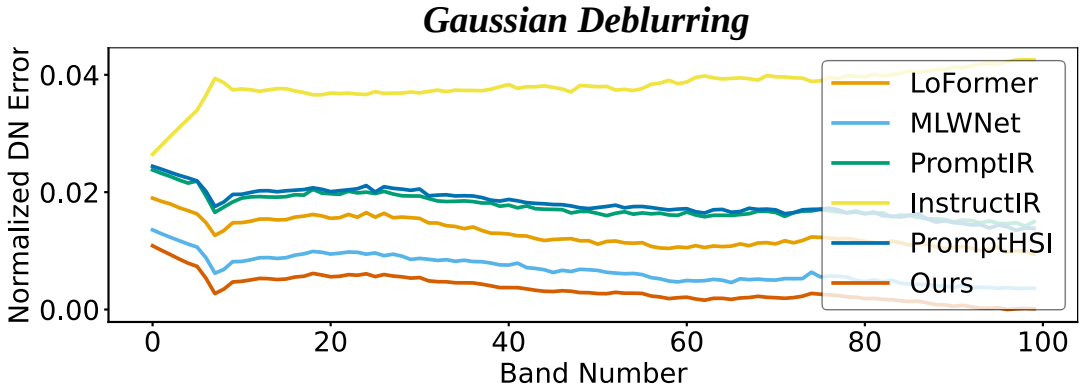
All-in-One Results

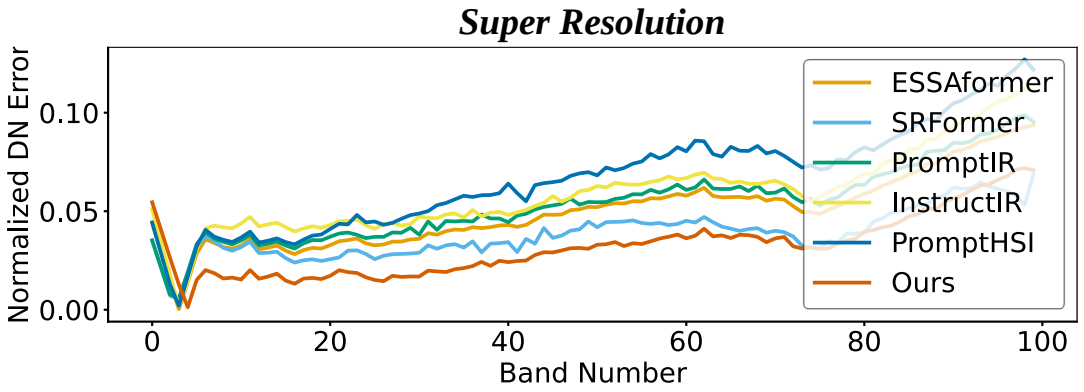


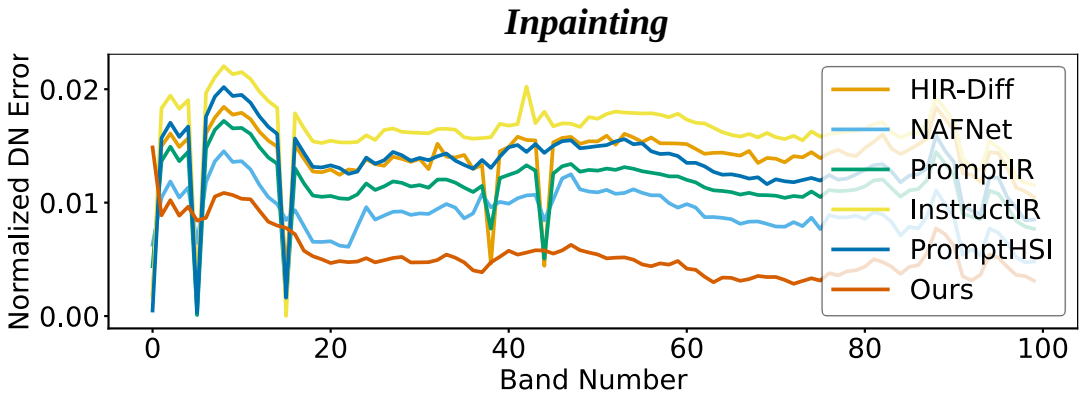
All-in-One Results

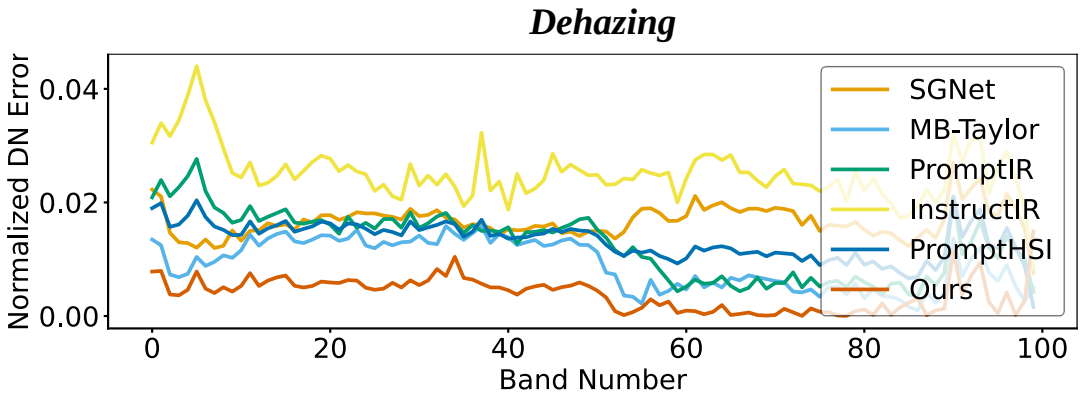












Generalization Results

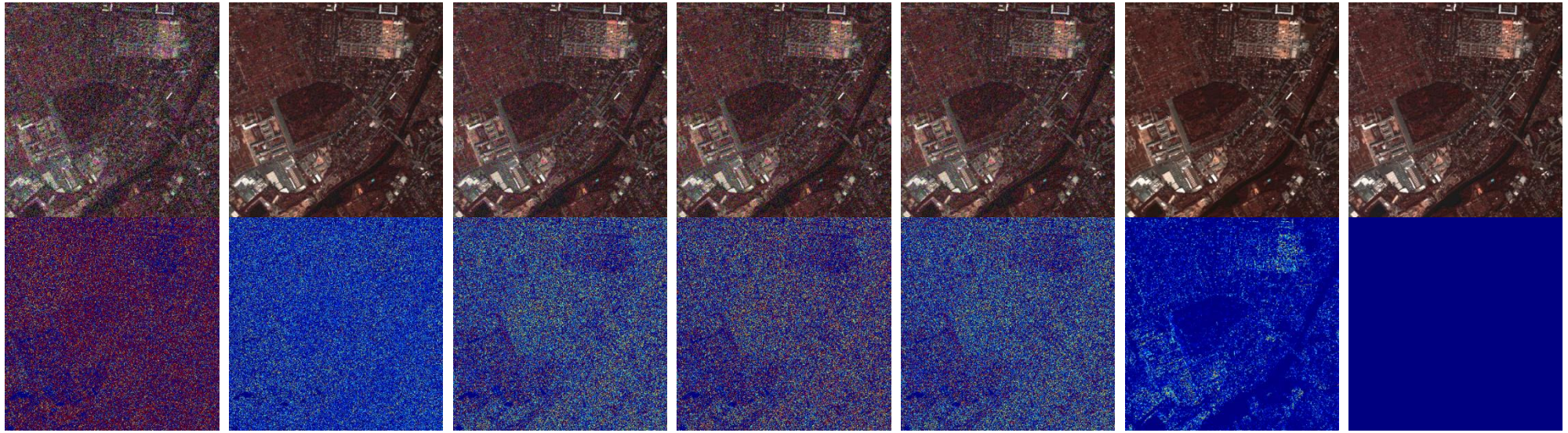
Type	Method	*Motion Deblurring*		Methods	*Poisson Denoising*	
		PSNR ↑	SSIM ↑		PSNR ↑	SSIM ↑
Task Specific	Stripformer [55]	37.91	0.948	QRNN3D [60]	35.98	0.910
	FFTformer [25]	38.42	0.950	SST [31]	37.51	0.914
	LoFormer [45]	38.94	0.953	SERT [32]	37.46	0.913
	MLWNet [13]	39.28	0.959	LDERT [33]	37.92	0.915
All in One	AirNet [30]	37.43	0.947	AirNet [30]	33.91	0.896
	PromptIR [50]	38.77	0.951	PromptIR [50]	34.83	0.902
	PIP [34]	38.55	0.950	PIP [34]	34.52	0.901
	HAIR [7]	37.58	0.948	HAIR [7]	34.38	0.901
	InstructIR [12]	36.58	0.942	InstructIR [12]	33.96	0.895
	PromptHSI [28]	36.79	0.945	PromptHSI [28]	37.25	0.912
	MP-HSIR (Ours)	40.06	0.965	MP-HSIR (Ours)	38.56	0.922

Table 3. **[Generalization]** The results of ***Motion Deblurring*** on the ICVL dataset and ***Poisson Denoising*** on the Houston dataset. Task-specific methods are trained on the entire dataset, while all-in-one methods are fine-tuned using only 5% of the data.

Motion Deblurring ⟶ Few-shot (5%)

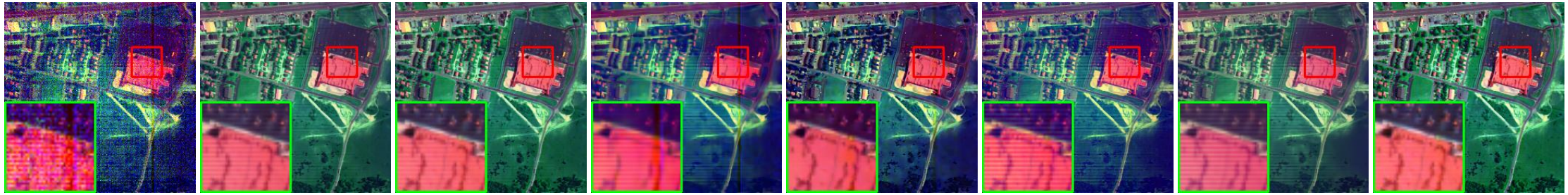
Poisson Denoising ⟶ Zero-shot

Poisson Denoising



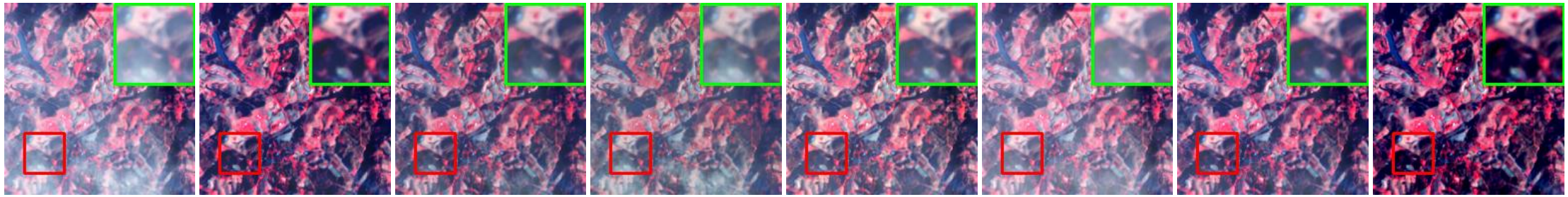
Input | Task-Specific | PromptIR | InstructIR | PromptHSI | Ours | GT

Real-World Results



Input SST LDERT AirNet PromptIR InstructIR PromptHSI Ours

Figure 5. **[Real-world]** Visual comparison results on ***Real Denoising***.



Input MB-Taylor DCMPNet AirNet PromptIR InstructIR PromptHSI Ours

Figure 6. **[Real-world]** Visual comparison results on ***Real Dehazing***.

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**

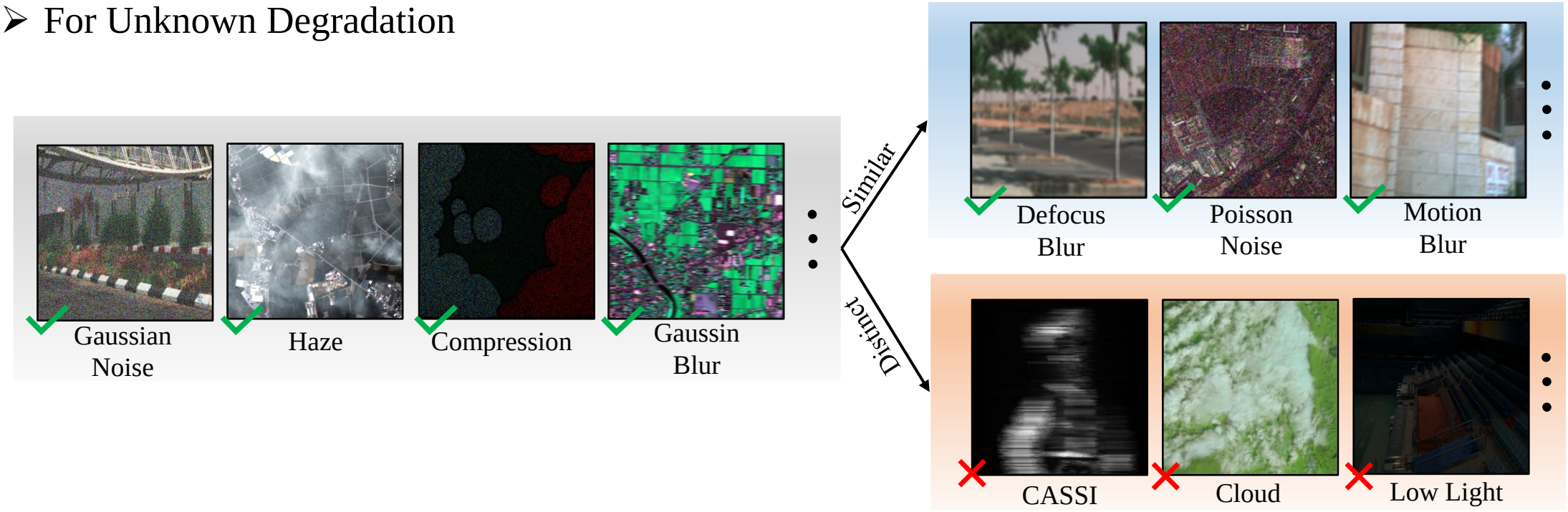
Table 4. **No-reference assessment** (QSFL) on real datasets.

Metric	Origin	PromptIR	InstructIR	PromptHSI	**MP-HSIR**
Accuracy (%)	80.2	82.7	81.9	83.2	**84.3**

Table 5. **Application evaluation**: Urban dataset classification.

Model Scalability

➢ For Unknown Degradation



➢ For Composite Degradation

Method	B+N	H+N	H+B	H+B+N
PromptIR	21.46 / 28.49	20.84 / 26.88	21.05 / 25.23	15.61 / 23.30
InstructIR	19.18 / 25.77	19.43 / 26.32	18.72 / 23.36	14.12 / 21.13
PromptHSI	21.14 / 31.14	20.75 / 29.61	20.55 / 26.89	15.65 / 24.52
MP-HSIR	**23.08 / 34.67**	**22.72 / 32.86**	**22.64 / 31.42**	**18.79 / 28.99**

Table 6. PSNR of composite degradations (**before / after fine-tunig**). Combinations of noise(N), blur(B), and haze(H).

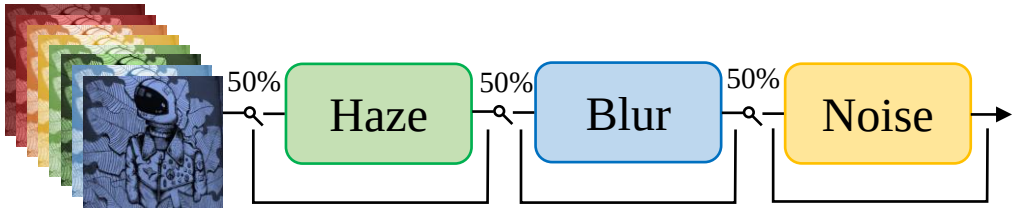
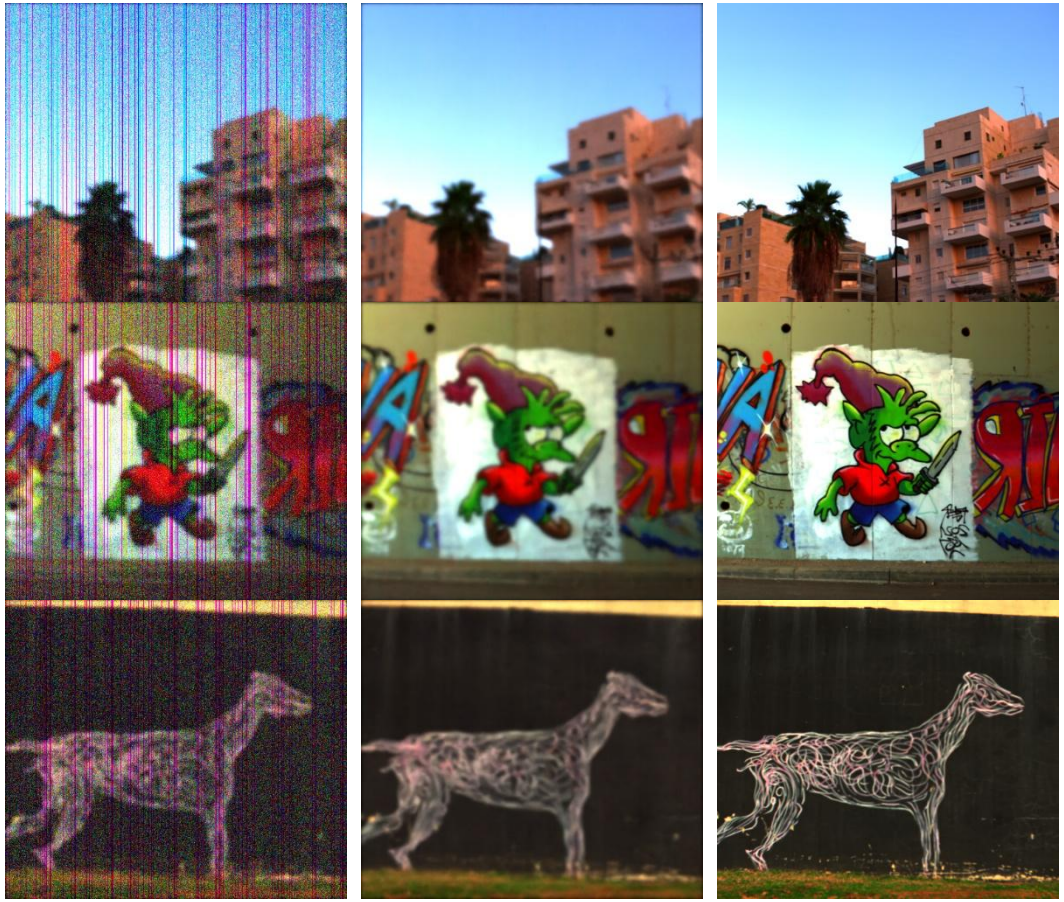


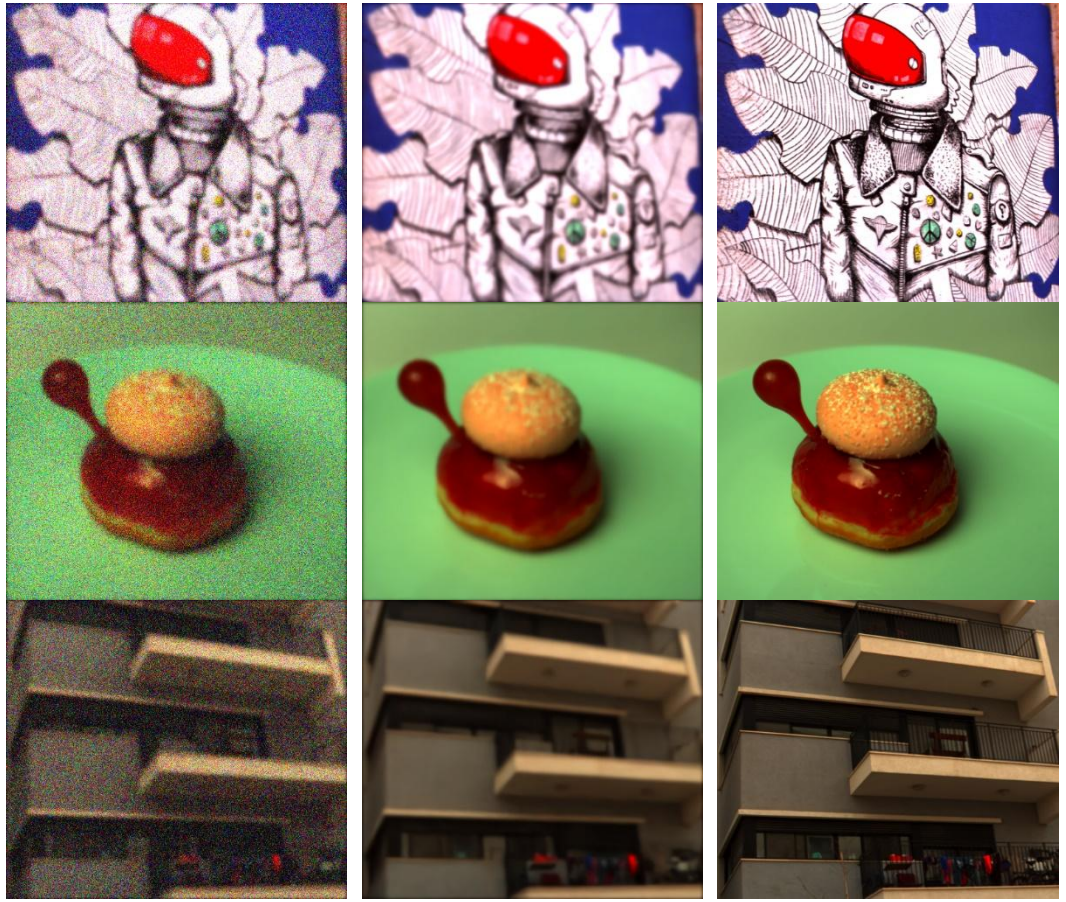
Figure 7. Illustration of the composite degradation synthesis pipeline.

Controllable Results



Degraded Denoised GT

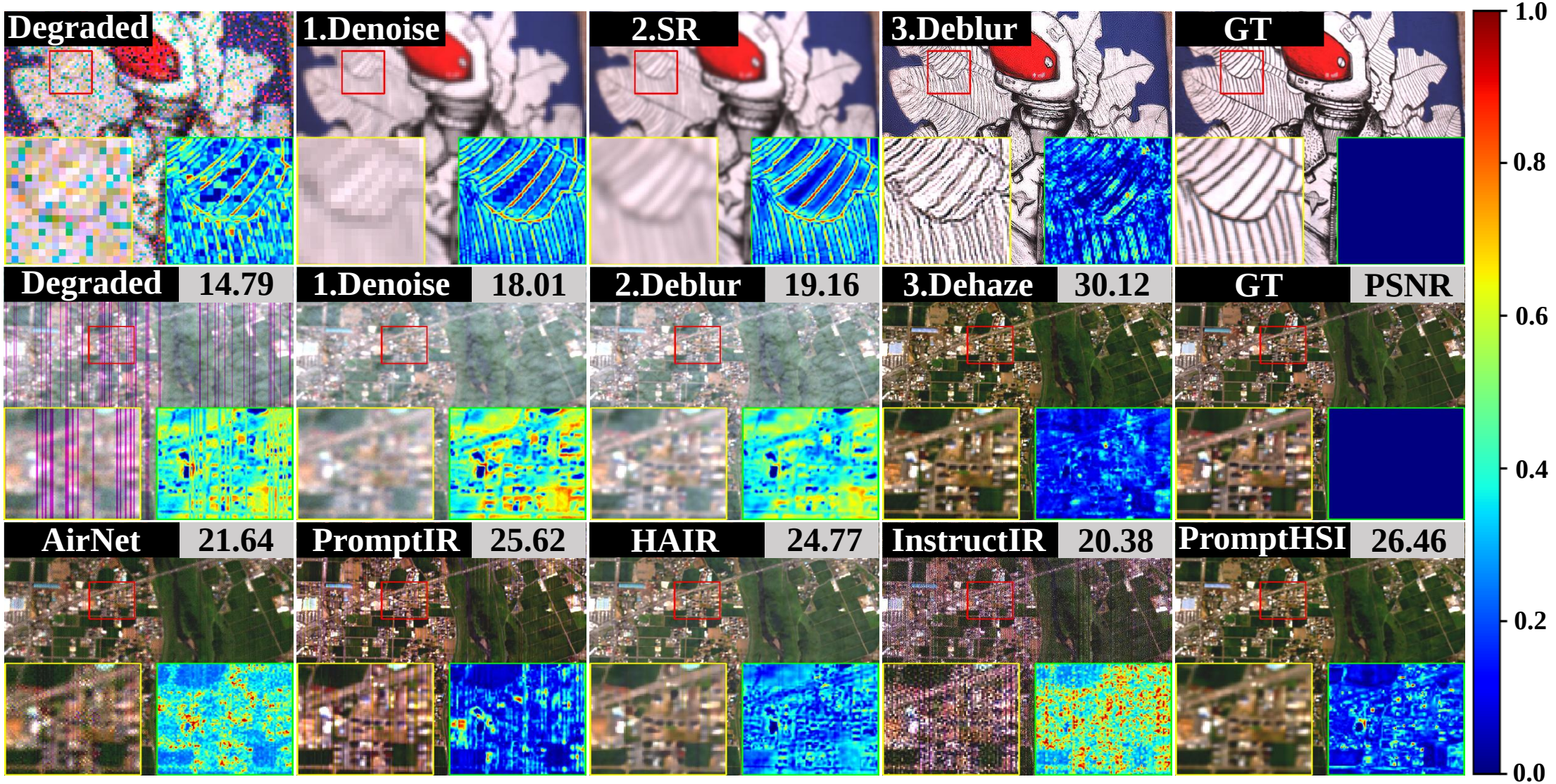
Figure 8. **[Controllable Reconstruction]** Removing ***Complex Noise*** from ***Complex Noise*** + ***Gaussian Blur*** Degradation.



Degraded Denoised GT

Figure 9. **[Controllable Reconstruction]** Removing ***Gaussian Noise*** from ***Gaussian Noise*** + ***Gaussian Blur*** Degradation.

Composite Degradations



Ablation Study & Complexity Analysis

Table 7. 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)	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 8. Ablation study to verify the effectiveness of modules on Chikusei dataset in ***Inpainting*** task with mask ratio = 0.9.

Method	PSNR ↑	SSIM ↑	Params (M)
Baseline (Only Spatial SA)	39.24	0.963	20.93
+ Textual Prompt P_T	39.62	0.964	21.51
+ Visual Prompt P_V	39.57	0.964	23.68
+ Textual Prompt P_T + Visual Prompt P_V	39.90	0.964	24.26
+ Global Spectral SA + P_T + P_T	40.63	0.969	30.07
+ Local Spectral SA + P_T + P_V	40.46	0.968	24.43
+ Local Spectral SA + P_T + P_V + Spectral Prompt P_S	41.05	0.971	25.10
Full Model	**41.98**	**0.974**	30.91

Table 9. Ablation study to verify the effectiveness of modules on PaviaU dataset in ***Dehazing*** task with Omega = 1.0.

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

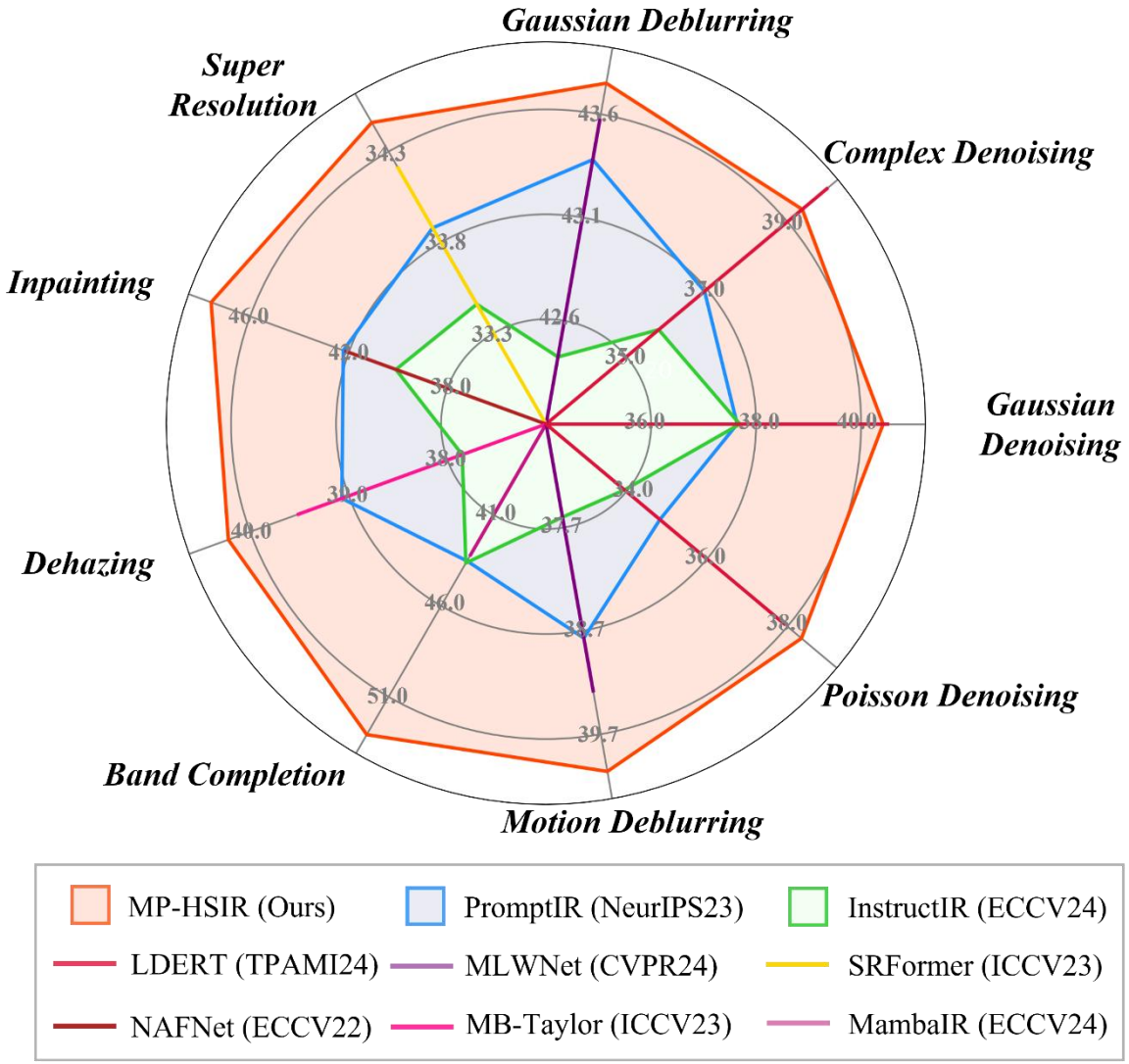
Table 10. Model complexity comparisons

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

Contents

Conclusion

- We propose a novel multi-prompt framework for all-in-one HSI restoration, integrating spectral, textual, and visual prompts. The spectral prompt provides universal low-rank spectral patterns to enhance spectral reconstruction.
- We introduce a text-visual synergistic prompt that combines semantic representations from textual prompts with fine-grained features from visual prompts, enhancing the controllability, interpretability, and degradation adaptability of the HSI restoration process.
- Extensive experiments on 9 HSI restoration tasks and real-world scenarios demonstrate that MP-HSIR significantly outperforms compared all-in-one methods and surpasses state-of-the-art task-specific approaches in multiple tasks.



Many thanks!

High-resolution information intelligent Processing and Application Group (HiPAG)

EarthVision and Application (EVA)