Restore-R1: Efficient Image Restoration Agents via Reinforcement Learning with Multimodal LLM Perceptual Feedback

Supplementary Material

6. Experimental Details

6.1. Data

In this section, we show how to synthesize degraded images following existing work [74]. For dark images, the V channel value of the images in the HSV color space will be randomly decreased by one of the following strategies: linear mapping, Gamma correction, and subtracting a constant. For defocus blur, the images will be filtered with circular kernels with random radius as in [30]. For JPEG compression artifact, images are compressed with random quality factor, such as 5, 40, and 90. For noise, images are added with Poisson or Gaussian noise with random scale. For rain, images are first added with noise and filtered the noise with linear kernels with random directions as in [17]. For motion blur, images are filtered with linear kernels with random direction and radius as in [30]. For haze generation, we simulate degraded images using the atmospheric scattering model, where the global atmospheric light and scattering coefficient are randomly sampled following the settings in prior work [11, 21]. Following [74], we consider degradation combinations that are common in real-world scenarios, e.g., dark+noise, rain+haze, rain+dark, rain+haze+defocus blur + JPEG compression artifact. Table 4 shows the 15 degradation combinations used in our experiments.

6.2. Tool Set

For each degradation, we follow previous restoration agents [14, 74] and consider several open-sourced restoration models as the callable tools. These candidate tools are:

- *Brightening*: CLAHE [75], Gamma correction ($\gamma = 2/3$), constant shift (adding a constant 40).
- *Defocus deblurring*: DRBNet [40], IFAN [20], Restormer [69].
- *JPEG compression artifact removal*: SwinIR [25] (quality factor 40), FBCNN [13] (quality factor 90), FBCNN [13] (quality factor 5), FBCNN [13] (blind to quality factor).
- *Denoising*: SwinIR [25] (noise level 15), SwinIR [25] (noise level 50), MAXIM [49], MPRNet [68], Restormer [69], X-Restormer [5].
- *Deraining*: MAXIM [49], MPRNet [68], Restormer [69], X-Restormer [5].
- *Motion deblurring*: MAXIM [49], MPRNet [68], Restormer [69], X-Restormer [5].
- *Dehazing*: MAXIM [49], X-Restormer [5]; RIDCP [61], DehazeFormer [47].

6.3. Evaluation Metrics

We assess model performance using three full-reference metrics: PSNR, SSIM [56], LPIPS [72], and four no-reference metrics: MANIQA [62], CLIP-IQA [54], MUSIQ [16], and DeQA-Score [65]. We briefly introduce these metrics below:

- *PSNR*: A pixel-level metric that measures the mean squared error between the restored and reference images. Higher PSNR indicates better fidelity.
- *SSIM* [56]: Evaluates perceptual similarity by assessing luminance, contrast, and structural components between image pairs. It better aligns with human visual perception than PSNR.
- *LPIPS* [72]: A perceptual metric that uses deep neural network features to compare image similarity, capturing differences that are visually meaningful but not captured by PSNR or SSIM.
- *MANIQA* [62]: A no-reference image quality assessment model that leverages transformer-based architecture and multi-level semantic features to predict perceptual quality without needing a ground-truth image.
- *CLIP-IQA* [54]: A no-reference quality assessment model that uses CLIP embeddings to assess image quality based on its alignment with natural image statistics learned from large-scale vision-language pretraining.
- *MUSIQ* [16]: A transformer-based non-reference image quality assessment model that adapts to various resolutions and content types by processing image patches, offering strong generalization across diverse datasets.
- *DeQA-Score* [65]: The model DeQA-Score computes a continuous image-quality score by first having a MLLM process one or more input images, then outputting a soft-label distribution over discrete quality levels (rather than a simple one-hot label). It treats the human quality ratings as approximately Gaussian and trains the MLLM with a KL-divergence loss to match that soft label. At inference time, the predicted distribution over levels is combined (by a weighted sum over discrete rating tokens) to produce a final quality score that more accurately reflects continuous human judgments.

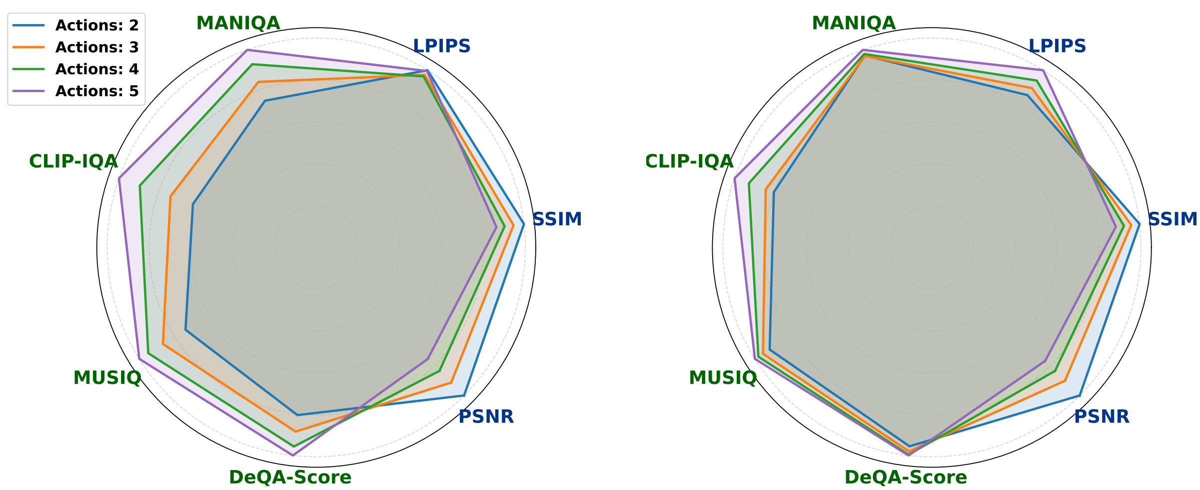
7. More Results

7.1. Qualitative Comparison

Figure 8 shows visual comparisons between our method and AirNet [22], PromptIR [37], InstructIR [6], and MiOIR(R) [17] under rain+haze and rain+dark+noise degradation

Table 4. Degradation data construction

Settings	# of Degradations	Case Number	Combinations
I	2	Case 1	dark+noise
		Case 2	defocus blur+JPEG compression artifact
		Case 3	motion blur + dark
		Case 4	noise+JPEG compression artifact
		Case 5	rain+haze
II	2	Case 6	haze+noise
		Case 7	motion blur+JPEG compression artifact
		Case 8	rain+dark
III	3	Case 9	dark+defocus blur+JPEG compression artifact
		Case 10	motion blur+defocus blur+noise
		Case 11	rain+dark+noise
		Case 12	rain+haze+noise
IV	>3	Case 13	haze+dark+motion blur+JPEG compression artifact
		Case 14	rain+haze+defocus blur+JPEG compression artifact
		Case 15	rain+motion blur+defocus blur+noise+JPEG compression artifact



(a) defocus blur+JPEG compression artifact

(b) noise+JPEG compression artifact

Figure 7. Illustration of distortion-perception tradeoff on (a) noise+jpeg compression artifact and (b) motion blur+defocus blur+noise.

cases. The results further demonstrate that our method effectively removes multiple co-occurring corruptions from degraded images and produces visual quality that is comparable to, or even exceeds, these supervised baselines.

7.2. Quantitative Comparison

Tabls 5, 6, 7 present the performance comparison between our method and the competing baselines across all degradation cases. As shown, even without access to ground-truth supervision, our method achieves competitive results on full-reference metrics and consistently outperforms all baselines on no-reference metrics, further demonstrating the effectiveness and robustness of our approach.

7.3. Perception-Distortion Tradeoff

Figures 7a and 7b illustrate the perception–distortion tradeoff under the defocus blur + JPEG compression and noise + JPEG compression degradation cases. These results lead to the same conclusion as discussed in Section 4.3.

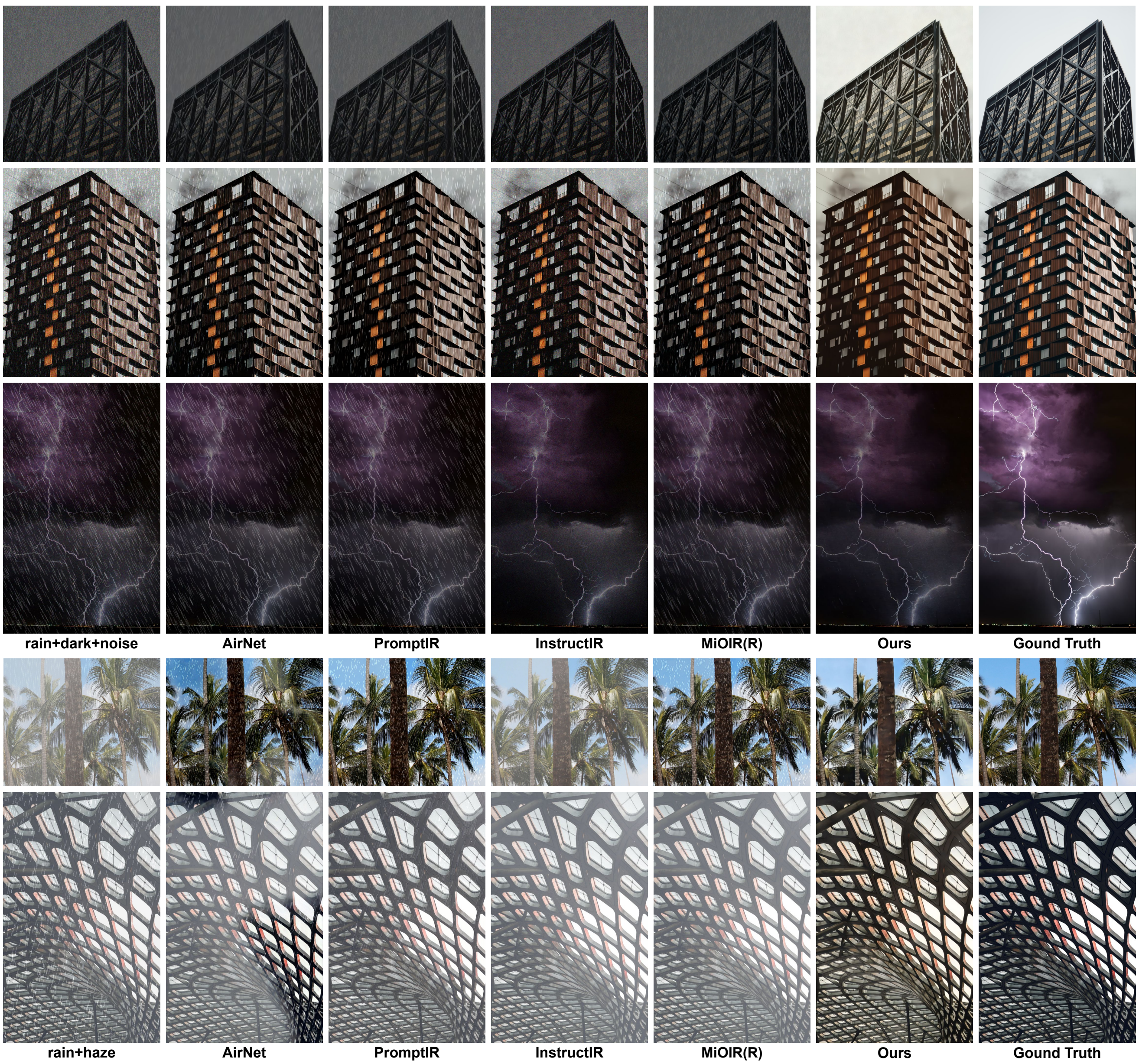


Figure 8. Qualitative comparison between our method and SOTA restoration baselines.

Table 5. Quantitative comparison across multiple mixed-degradation conditions. Arrows indicate the desired direction of improvement.

	Method	Full-Reference			No-Reference			
		PSNR ↑	SSIM ↑	LPIPS ↓	MANIQA ↑	CLIP-IQA ↑	MUSIQ ↑	DeQA-Score) ↑
Case 1	AirNet [22]	16.272	0.633	0.221	0.423	0.572	64.253	3.539
	PromptIR [37]	16.217	0.615	0.242	0.407	0.540	64.364	3.479
	InstructIR [6]	15.305	0.469	0.496	0.260	0.507	49.016	3.010
	MiOIR(R) [17]	16.563	0.632	0.233	0.312	0.612	59.093	3.561
	MiOIR(U) [17]	16.696	0.675	0.210	0.348	0.598	62.262	3.628
	DA-CLIP [29]	16.343	0.552	0.340	0.330	0.494	57.732	3.289
	AutoDIR [15]	16.771	0.658	0.257	0.362	0.590	63.672	3.625
	AgenticIR [74]	19.692	0.708	0.338	0.346	0.444	59.594	3.388
	Ours	17.995	0.664	0.362	0.399	0.631	68.240	3.871
Case 2	AirNet [22]	22.282	0.629	0.501	0.194	0.232	28.489	2.282
	PromptIR [37]	23.110	0.632	0.486	0.203	0.247	28.715	2.298
	InstructIR [6]	23.727	0.648	0.519	0.211	0.216	29.441	2.324
	MiOIR(R) [17]	23.960	0.653	0.503	0.209	0.256	30.628	2.353
	MiOIR(U) [17]	23.965	0.653	0.511	0.203	0.278	30.848	2.338
	DA-CLIP [29]	23.580	0.647	0.490	0.207	0.272	30.073	2.354
	AutoDIR [15]	23.915	0.654	0.439	0.215	0.269	36.215	2.571
	AgenticIR [74]	22.755	0.658	0.435	0.216	0.303	43.866	2.688
	Ours	22.338	0.634	0.456	0.225	0.302	44.950	2.832
Case 3	AirNet [22]	14.083	0.450	0.411	0.175	0.281	42.572	2.429
	PromptIR [37]	14.061	0.433	0.406	0.183	0.290	42.710	2.434
	InstructIR [6]	14.489	0.461	0.413	0.182	0.278	42.521	2.481
	MiOIR(R) [17]	17.294	0.561	0.371	0.176	0.281	43.473	2.497
	MiOIR(U) [17]	16.989	0.574	0.388	0.170	0.276	43.659	2.471
	DA-CLIP [29]	15.520	0.517	0.356	0.201	0.299	48.516	2.728
	AutoDIR [15]	15.984	0.567	0.319	0.254	0.364	56.248	3.340
	AgenticIR [74]	17.306	0.519	0.314	0.269	0.406	59.304	3.325
	Ours	15.691	0.482	0.393	0.345	0.582	66.178	3.797
Case 4	AirNet [22]	23.522	0.613	0.458	0.287	0.379	44.700	2.958
	PromptIR [37]	24.249	0.620	0.448	0.293	0.408	45.182	2.982
	InstructIR [6]	24.238	0.626	0.439	0.294	0.413	45.629	3.028
	MiOIR(R) [17]	25.453	0.671	0.381	0.320	0.464	52.704	3.137
	MiOIR(U) [17]	25.628	0.673	0.373	0.324	0.486	53.445	3.086
	DA-CLIP [29]	24.592	0.632	0.403	0.316	0.506	49.354	3.123
	AutoDIR [15]	24.221	0.622	0.436	0.295	0.437	45.982	3.030
	AgenticIR [74]	25.489	0.800	0.258	0.344	0.462	61.722	3.720
	Ours	24.834	0.770	0.281	0.374	0.487	63.852	3.774
Case 5	AirNet [22]	16.288	0.716	0.279	0.396	0.531	64.235	3.349
	PromptIR [37]	18.351	0.745	0.275	0.414	0.515	64.975	3.341
	InstructIR [6]	14.213	0.727	0.215	0.392	0.538	63.626	3.377
	MiOIR(R) [17]	16.578	0.776	0.197	0.395	0.539	65.445	3.444
	MiOIR(U) [17]	16.322	0.784	0.177	0.401	0.541	65.869	3.533
	DA-CLIP [29]	16.225	0.722	0.262	0.408	0.515	65.923	3.374
	AutoDIR [15]	16.136	0.721	0.280	0.402	0.497	65.782	3.365
	AgenticIR [74]	19.373	0.805	0.184	0.400	0.592	68.463	3.994
	Ours	16.705	0.687	0.333	0.403	0.624	71.796	4.010

Table 6. Quantitative comparison across multiple mixed-degradation conditions. Arrows indicate the desired direction of improvement.

	Method	Full-Reference			No-Reference			
		PSNR ↑	SSIM ↑	LPIPS ↓	MANIQA ↑	CLIP-IQA ↑	MUSIQ ↑	DeQA-Score) ↑
Case 6	AirNet [22]	14.781	0.657	0.313	0.373	0.492	63.081	3.293
	PromptIR [37]	14.783	0.661	0.310	0.355	0.467	62.443	3.320
	InstructIR [6]	14.456	0.446	0.646	0.248	0.469	43.749	2.770
	MiOIR(R) [17]	14.672	0.632	0.320	0.267	0.562	55.223	3.390
	MiOIR(U) [17]	14.830	0.665	0.285	0.302	0.487	60.079	3.361
	DA-CLIP [29]	14.503	0.602	0.397	0.313	0.399	56.516	3.134
	AutoDIR [15]	17.249	0.699	0.268	0.322	0.518	62.531	3.567
	AgenticIR [74]	18.725	0.699	0.349	0.315	0.433	59.596	3.580
	Ours	15.999	0.623	0.389	0.393	0.623	68.938	3.869
Case 7	AirNet [22]	20.800	0.623	0.401	0.198	0.249	34.574	2.591
	PromptIR [37]	21.494	0.632	0.385	0.207	0.271	34.707	2.630
	InstructIR [6]	21.738	0.640	0.400	0.210	0.255	36.812	2.665
	MiOIR(R) [17]	21.941	0.648	0.389	0.193	0.302	39.335	2.640
	MiOIR(U) [17]	21.949	0.648	0.393	0.186	0.326	39.704	2.619
	DA-CLIP [29]	21.664	0.640	0.384	0.193	0.295	38.169	2.641
	AutoDIR [15]	21.679	0.634	0.374	0.222	0.296	38.146	2.822
	AgenticIR [74]	20.397	0.625	0.383	0.212	0.349	48.192	2.922
	Ours	20.370	0.603	0.411	0.245	0.350	50.361	3.061
Case 8	AirNet [22]	16.006	0.647	0.263	0.399	0.576	63.726	3.409
	PromptIR [37]	16.178	0.687	0.207	0.411	0.585	64.320	3.512
	InstructIR [6]	16.329	0.692	0.190	0.400	0.574	63.881	3.548
	MiOIR(R) [17]	17.522	0.732	0.202	0.404	0.573	65.350	3.489
	MiOIR(U) [17]	16.580	0.715	0.156	0.409	0.585	65.647	3.592
	DA-CLIP [29]	16.917	0.669	0.266	0.419	0.548	66.502	3.438
	AutoDIR [15]	16.691	0.660	0.314	0.410	0.530	66.372	3.382
	AgenticIR [74]	22.132	0.832	0.169	0.402	0.561	68.141	3.800
	Ours	17.506	0.682	0.356	0.390	0.618	70.732	3.866
Case 9	AirNet [22]	14.814	0.458	0.544	0.189	0.192	28.382	2.131
	PromptIR [37]	15.014	0.446	0.534	0.196	0.202	28.532	2.128
	InstructIR [6]	15.434	0.483	0.550	0.203	0.182	29.241	2.165
	MiOIR(R) [17]	15.491	0.490	0.550	0.206	0.233	30.287	2.197
	MiOIR(U) [17]	15.483	0.489	0.555	0.201	0.260	30.538	2.184
	DA-CLIP [29]	15.778	0.496	0.543	0.206	0.244	30.166	2.224
	AutoDIR [15]	15.973	0.514	0.488	0.196	0.219	34.427	2.338
	AgenticIR [74]	18.798	0.584	0.495	0.196	0.262	41.299	2.504
	Ours	16.287	0.535	0.482	0.304	0.491	60.227	3.438
Case 10	AirNet [22]	21.250	0.559	0.536	0.196	0.277	29.854	2.096
	PromptIR [37]	21.269	0.562	0.562	0.196	0.254	28.904	2.086
	InstructIR [6]	19.104	0.347	0.780	0.103	0.305	22.685	1.882
	MiOIR(R) [17]	20.923	0.530	0.569	0.116	0.345	25.380	2.034
	MiOIR(U) [17]	21.225	0.566	0.526	0.153	0.278	26.593	2.090
	DA-CLIP [29]	20.474	0.486	0.624	0.163	0.295	25.290	2.019
	AutoDIR [15]	18.855	0.479	0.520	0.241	0.335	41.623	2.528
	AgenticIR [74]	20.034	0.546	0.525	0.195	0.292	38.196	2.401
	Ours	20.202	0.530	0.531	0.186	0.288	37.379	2.368

Table 7. Quantitative comparison across multiple mixed-degradation conditions. Arrows indicate the desired direction of improvement.

	Method	Full-Reference			No-Reference			
		PSNR ↑	SSIM ↑	LPIPS ↓	MANIQA ↑	CLIP-IQA ↑	MUSIQ ↑	DeQA-Score) ↑
Case 11	AirNet [22]	16.650	0.521	0.452	0.371	0.499	62.832	3.060
	PromptIR [37]	16.601	0.511	0.460	0.354	0.493	62.143	3.065
	InstructIR [6]	15.790	0.455	0.560	0.236	0.455	47.557	2.826
	MiOIR(R) [17]	16.734	0.529	0.438	0.285	0.525	56.743	3.200
	MiOIR(U) [17]	16.487	0.532	0.451	0.319	0.537	60.476	3.142
	DA-CLIP [29]	16.241	0.443	0.514	0.296	0.443	55.044	2.971
	AutoDIR [15]	16.776	0.575	0.387	0.324	0.554	62.473	3.373
	AgenticIR [74]	18.674	0.661	0.404	0.287	0.394	57.342	3.254
	Ours	17.464	0.596	0.440	0.343	0.569	66.904	3.582
Case 12	AirNet [22]	13.162	0.528	0.483	0.327	0.429	61.533	2.871
	PromptIR [37]	13.167	0.523	0.502	0.300	0.422	59.385	2.866
	InstructIR [6]	13.320	0.370	0.767	0.227	0.428	39.365	2.451
	MiOIR(R) [17]	13.151	0.508	0.490	0.252	0.474	52.839	3.000
	MiOIR(U) [17]	13.202	0.527	0.489	0.273	0.430	57.886	2.927
	DA-CLIP [29]	13.058	0.453	0.567	0.265	0.386	53.851	2.739
	AutoDIR [15]	15.875	0.592	0.410	0.292	0.474	60.303	3.228
	AgenticIR [74]	16.895	0.612	0.435	0.263	0.399	54.407	3.184
	Ours	16.648	0.591	0.447	0.345	0.542	66.515	3.600
Case 13	AirNet [22]	14.032	0.467	0.542	0.174	0.183	30.106	2.185
	PromptIR [37]	15.082	0.506	0.521	0.191	0.194	30.544	2.237
	InstructIR [6]	14.738	0.508	0.548	0.199	0.177	32.111	2.264
	MiOIR(R) [17]	14.787	0.513	0.544	0.187	0.245	34.559	2.285
	MiOIR(U) [17]	14.769	0.513	0.551	0.178	0.272	35.271	2.263
	DA-CLIP [29]	14.590	0.505	0.539	0.192	0.231	33.721	2.322
	AutoDIR [15]	14.821	0.509	0.525	0.207	0.216	33.228	2.372
	AgenticIR [74]	15.514	0.520	0.513	0.181	0.277	41.847	2.405
	Ours	14.671	0.476	0.517	0.276	0.448	56.298	3.168
Case 14	AirNet [22]	15.499	0.465	0.635	0.152	0.195	23.437	1.842
	PromptIR [37]	15.397	0.475	0.634	0.171	0.207	23.194	1.924
	InstructIR [6]	12.947	0.452	0.682	0.186	0.195	23.107	1.937
	MiOIR(R) [17]	12.791	0.449	0.693	0.188	0.213	23.952	1.967
	MiOIR(U) [17]	12.775	0.450	0.704	0.185	0.241	24.009	1.967
	DA-CLIP [29]	12.772	0.442	0.689	0.194	0.239	24.610	1.967
	AutoDIR [15]	13.324	0.428	0.645	0.205	0.229	30.584	2.069
	AgenticIR [74]	16.214	0.491	0.602	0.159	0.250	35.394	2.175
	Ours	15.120	0.447	0.570	0.266	0.472	57.365	3.217
Case 15	AirNet [22]	18.854	0.484	0.640	0.138	0.213	20.739	1.797
	PromptIR [37]	19.598	0.495	0.627	0.141	0.234	20.752	1.800
	InstructIR [6]	19.420	0.504	0.667	0.150	0.204	20.900	1.823
	MiOIR(R) [17]	19.245	0.500	0.662	0.159	0.218	21.385	1.838
	MiOIR(U) [17]	19.244	0.502	0.674	0.159	0.248	21.425	1.835
	DA-CLIP [29]	19.110	0.498	0.649	0.158	0.237	22.256	1.854
	AutoDIR [15]	18.808	0.456	0.576	0.196	0.248	32.703	2.155
	AgenticIR [74]	18.910	0.484	0.604	0.156	0.250	34.060	2.042
	Ours	18.750	0.497	0.605	0.156	0.248	35.768	2.207