DDPR: DOMINANT DEGRADATION PLANNING WITH LATENT RESIDUAL-GUIDED ADAPTIVE EDITING FOR ALL-IN-ONE IMAGE RESTORATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Generative models have shown strong potential for All-in-One image restoration. However, existing methods either process compound degradations in a single pass or lack dynamic reassessment of the dominant degradation, often leading to suboptimal restoration orders and residual degradations. They also apply uniform restoration across the image, without distinguishing regions requiring content preservation from those requiring stronger restoration, which may cause content hallucination. To address these limitations, we propose **D**ominant **D**egradation **P**lanning for image **R**estoration (DDPR), a unified framework with latent residual-guided adaptive editing. Specifically, the Dominant Degradation Planner identifies the current dominant degradation and reassesses the image after each restoration step, decomposing compound degradations into a dynamically planned sequence of restoration tasks. The Degradation-Aware Latent Residual Assessor predicts a region-level degradation gate from VAE latent residuals, guiding the Degradation-Gated Dual-Branch LoRA to adaptively balance the content-preservation and degradation-restoration branches. Experiments demonstrate that DDPR achieves SOTA performance among open-source methods across both synthetic and real-world degradation benchmarks. Code will be publicly released.

1 INTRODUCTION

All-in-One image restoration aims to address multiple degradations within a single model (Li et al., 2022; Potlapalli et al., 2023; Luo et al., 2024; Conde et al., 2024; Cui et al., 2025; Luo et al., 2025; Tian et al., 2025). More recently, generative models have shown strong potential for All-in-One image restoration by leveraging powerful image priors to reconstruct realistic details and generalize across diverse degradations. However, real-world images may suffer from multiple degradations simultaneously, whose visual effects can overlap, occlude, and interfere with one another (Zhang et al., 2017; Nah et al., 2017; Wang et al., 2019; Guo et al., 2022; Cai et al., 2023; Li et al., 2020; Dong et al., 2015), making compound-degradation restoration particularly challenging for existing restoration methods (Guo et al., 2024).

Most existing All-in-One methods directly produce the final restored image in a single forward pass, which may leave residual degradations or insufficiently recovered details when multiple degradations interfere, as shown in Figure 1(a) (Li et al., 2022; Potlapalli et al., 2023; Luo et al., 2024; Conde et al., 2024). Recent agentic approaches perform multi-step restoration by selecting and scheduling specialized restoration models (Chen et al., 2024; Zhu et al., 2025; Lin et al., 2025). However, the effectiveness of progressive restoration depends critically on the restoration order, since an inappropriate order may interfere with subsequent restoration and substantially degrade the final visual quality, as illustrated in Figure 1(b). Moreover, these approaches generally focus on model scheduling rather than explicitly re-identifying and prioritizing the current dominant degradation within a single unified restoration model after each step. In compound-degradation scenarios, dominant degradations may obscure weaker ones, which can become apparent only after the image has been partially restored (Guo et al., 2024; Mao et al., 2026; Zhu et al., 2025). Consequently, weaker degradations may be addressed at inappropriate stages or remain unresolved. Therefore, the restoration process should dynamically identify the current dominant degradation at each step and determine the subsequent restoration objective based on the updated image as restoration progresses.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

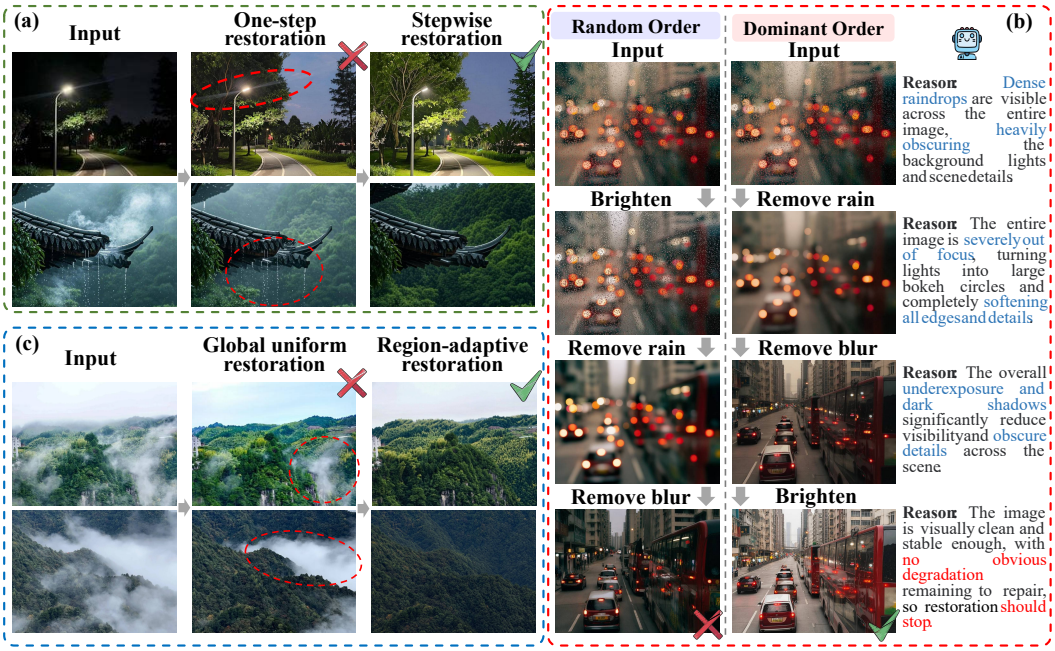

Figure 1: Motivation for DDPR. (a) Progressive restoration alleviates residual degradations. (b) Dominant degradation planning avoids suboptimal restoration sequences. (c) Region-adaptive restoration strengthens degraded regions while preserving relatively clean content.

Beyond task-level restoration order, degradation severity is often spatially non-uniform within an image (Liu et al., 2025). Severely degraded regions require stronger restoration, whereas mildly degraded or relatively clean regions should preserve their original structures and textures as much as possible. However, generative models typically rely on global textual instructions and thus lack an explicit mechanism to characterize region-level restoration requirements (Brooks et al., 2023; Wu et al., 2025). Conventional LoRA also applies shared low-rank adaptations across spatial locations, without explicitly accounting for local degradation severity (Hu et al., 2022; Ai et al., 2024). As a result, severely degraded regions may remain insufficiently restored, while relatively clean regions may undergo unnecessary modifications, as shown in Figure 1(c). Therefore, All-in-One image restoration further requires region-adaptive restoration capabilities (Liu et al., 2025).

To address these challenges, we propose DDPR, an All-in-One image restoration framework that combines dominant degradation planning with latent residual-guided adaptive editing, as illustrated in Figure 2. The Dominant Degradation Planner identifies the dominant degradation of the current image and reassesses the image after each restoration step, decomposing compound degradations into a dynamically planned sequence of single-step restoration tasks. For region-adaptive restoration, the Degradation-Aware Latent Residual Assessor (DLRA) learns region-level restoration requirements from VAE latent residuals and predicts a degradation gate at inference. Guided by this gate, the Degradation-Gated Dual-Branch LoRA (DGDB-LoRA) adaptively balances the content-preservation and degradation-restoration branches, strengthening restoration in severely degraded regions while reducing unnecessary modifications to relatively clean areas. Our contributions are summarized as follows:

- We introduce a Dominant Degradation Planner that dynamically identifies the current dominant degradation and decomposes compound restoration into sequential tasks.
- We propose a latent residual-guided adaptive editing framework with the Degradation-Aware Latent Residual Assessor and Degradation-Gated Dual-Branch LoRA for spatially adaptive restoration that balances content preservation and degradation removal.
- We construct RealComp-Bench with 221 carefully curated real-world compound-degraded images and validate DDPR across real-world single, synthetic compound, and real-world compound degradations to assess its robustness and generalization.

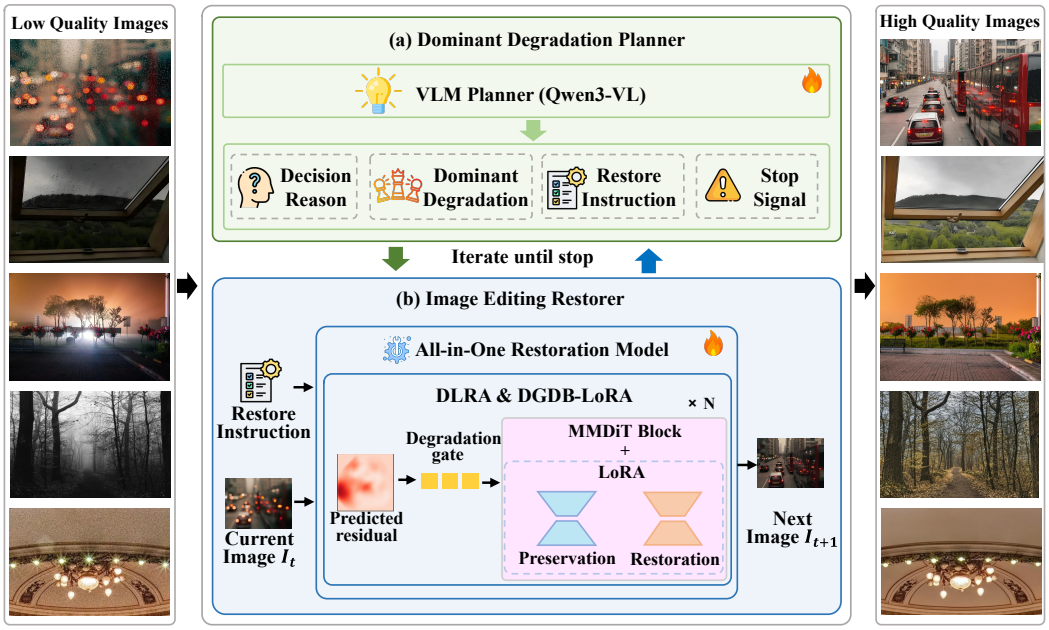

Figure 2: Overview of DDPR, which combines dominant degradation planning with latent residual-guided region-adaptive restoration in a unified progressive pipeline.

2 Related Work

2.1 All-in-One Image Restoration

Existing All-in-One image restoration methods extract degradation-aware conditions from inputs to modulate a shared restoration network (Kong et al., 2024; Gao et al., 2024; Zhang et al., 2026a; Cao et al., 2024; Wang et al., 2025). AirNet (Li et al., 2022) learns degradation representations through contrastive learning, PromptIR (Potlapalli et al., 2023) encodes input-dependent degradation information into dynamic prompts, and DA-CLIP (Luo et al., 2024) adapts a pretrained vision-language model to extract content and degradation features. AMIRNet (Zhang et al., 2023) models the commonality and differences among degradations using hierarchical degradation representations, whereas Neural Degradation Representation Learning (Yao et al., 2024) decomposes degradations into compositional neural representations to improve adaptation to unknown degradations.

With advances in generative and large-scale pretrained models (Cheng et al., 2026; Zhang et al., 2026b), AutoDIR (Jiang et al., 2024) combines degradation analysis with latent diffusion restoration, FoundIR (Li et al., 2025) leverages real-world degradation data to improve generalization in open-world scenarios, and RealRestorer (Yang et al., 2026) transfers image editing models to real-world image restoration. AllRestorer (Mao et al., 2026) and UniCoRN (Mandal et al., 2026) address coexisting degradations through compound-degradation conditioning or multi-branch control. However, these methods still largely generate results from the initial input in a single pass, making it difficult to handle complex and diverse compound degradations. In contrast, DDPR decomposes compound degradations into a sequence of single-step restoration tasks.

2.2 Planning-Based and Agentic Image Restoration

Recent studies have employed large multimodal models to organize complex image restoration processes (Zhou et al., 2025; Li et al., 2026). RestoreAgent (Chen et al., 2024) automatically analyzes the type and severity of image degradations and performs task identification, order planning, model selection, and execution. AgenticIR (Zhu et al., 2025) divides restoration into perception, scheduling, execution, reflection, and rescheduling, dynamically invoking restoration tools through a language model. JarvisIR (Lin et al., 2025) uses a vision-language model to coordinate multiple restoration experts, while MAIR (Jiang et al., 2026) reduces the search cost of complex restoration through collaboration between a scheduler and multiple degradation experts.

These methods demonstrate the effectiveness of progressive planning for compound-degradation restoration, but they typically select and combine multiple task-specific models from an external tool library. In contrast, DDPR does not perform multi-tool selection. Instead, the Dominant Degradation Planner identifies the degradation that most urgently requires restoration, while the same All-in-One image restoration model is used throughout all restoration steps.

3 METHOD

DDPR consists of a Dominant Degradation Planner and an All-in-One restoration model. As shown in Figure 2, the planner analyzes the current image I_t and jointly outputs the Decision Reason, Dominant Degradation, Restore Instruction, and Stop Signal. If the Stop Signal is True, the process terminates; otherwise, I_t and the restoration instruction are fed into the restoration model. DLRA predicts a region-level degradation gate to guide DGDB-LoRA in adaptively fusing the content-preservation and degradation-restoration branches, producing I_{t+1}. The planner then reassesses the updated image throughout the progressive restoration process, and the process repeats until the stopping condition is satisfied or the maximum number of iterations is reached.

3.1 DOMINANT DEGRADATION PLANNER

The Dominant Degradation Planner builds on Qwen3-VL-8B-Instruct (Bai et al., 2025) and uses LoRA fine-tuning. We construct trajectories using nine degradations: blur, compression artifacts, flare, haze, low-light, moiré patterns, noise, rain, and reflection. Given a clean image I_c, we sample n degradations from $\mathcal{D}$ and order them as $\Pi_0 = [d_1, d_2, \ldots, d_n]$. The compound-degradation function $\mathcal{A}(I_c, \Pi)$ sequentially applies the degradations in Π to I_c, yielding $I_0 = \mathcal{A}(I_c, \Pi_0)$. At step t, the remaining sequence is $\Pi_t = [d_{t,1}, d_{t,2}, \ldots, d_{t,K}]$, and the current image state is $I_t = \mathcal{A}(I_c, \Pi_t)$. For each candidate degradation $d \in \Pi_t$, we remove d from the current sequence while preserving the relative order of the remaining degradations, obtaining the candidate sequence $\Pi_t^{(-d)}$. The corresponding candidate image is then synthesized as $I_t^{(-d)} = \mathcal{A}(I_c, \Pi_t^{(-d)})$.

To identify the dominant degradation that should be prioritized, we use five no-reference image quality metrics, including Q-Align (Wu et al., 2024), MANIQA (Yang et al., 2022), MUSIQ (Ke et al., 2021), CLIP-IQA (Wang et al., 2023), and NIQE-quality (Mittal et al., 2013), to jointly measure the quality improvement of each candidate image over the current image. The candidate degradation with the highest score is selected as the current dominant degradation d_t. After determining d_t, we update the degradation sequence and image state as $\Pi_{t+1} = \Pi_t^{(-d_t)}$ and $I_{t+1} = I_t^{(-d_t)}$. Repeated candidate construction, quality evaluation, and sequence updates form a multi-step restoration trajectory $I_0 \to I_1 \to \cdots \to I_T$. When $\Pi_T = \varnothing$, all synthetic degradations are removed and $I_T = I_c$.

To improve real-world generalization, we collect a large number of real images from the Internet, each exhibiting a single degradation, and restore them using GPT-Image-2 and Nano Banana 2. Failed restoration samples are manually discarded, while high-quality and content-consistent results are retained as clean targets, thereby forming real-world single-step restoration pairs. The synthetic training data include samples containing one to four degradations. These samples are organized into progressive restoration trajectories according to their degradation-removal sequences. The resulting synthetic restoration trajectories and real-world single-step restoration pairs are used to train the Dominant Degradation Planner and the All-in-One restoration model separately.

3.2 LATENT RESIDUAL-GUIDED ADAPTIVE EDITING

Figure 3 illustrates the latent residual-guided adaptive editing framework for All-in-One image restoration. Built upon Qwen-Image-Edit-2511, the model incorporates a Degradation-Aware Latent Residual Assessor and a Degradation-Gated Dual-Branch LoRA.

Degradation-Aware Latent Residual Assessor. During training, we denote the current state image I_t in a progressive restoration trajectory as I_{deg}, and its corresponding single-step restoration target I_{t+1} as I_{tar}. The two images are first encoded into the VAE latent space:

$$z_{\mathrm{deg}} = \mathrm{VAE}(I_{\mathrm{deg}}), \quad z_{\mathrm{tar}} = \mathrm{VAE}(I_{\mathrm{tar}}). \tag{1}$$

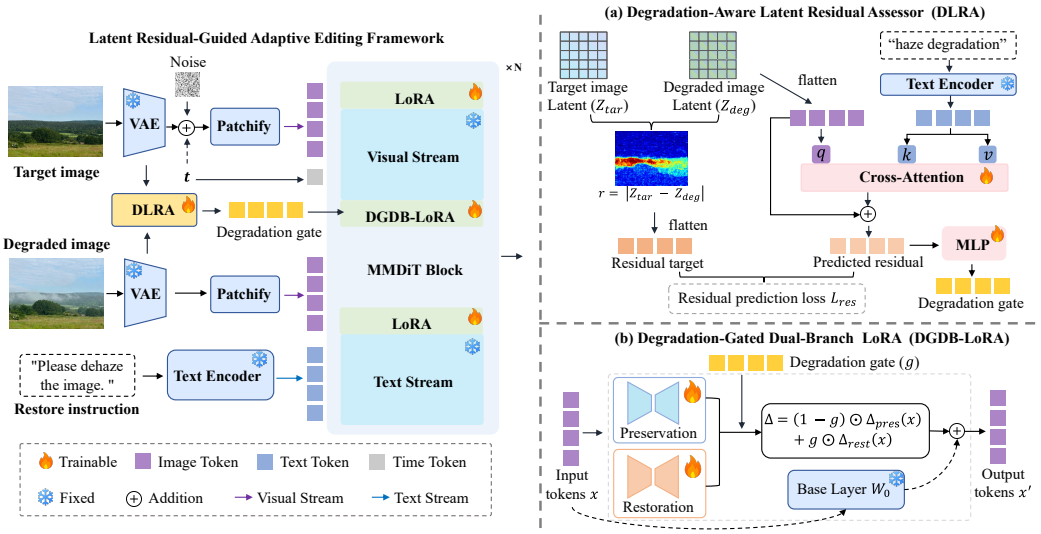

Figure 3: Architecture of the latent residual-guided adaptive editing framework.

The latent residual between the degraded image and its target image is defined as:

$$r = |z_{\text{tar}} - z_{\text{deg}}| \,. \tag{2}$$

Larger latent residuals indicate pronounced feature discrepancies between the current degradation state and the restoration target, suggesting stronger restoration requirements. Smaller residuals indicate that the corresponding regions should preserve their current structures and textures. To align the residual with image tokens, it is flattened into a token-level supervision target $\mathbf{r}_{\text{target}} \in \mathbb{R}^{B\times N\times C_r}$, where N denotes the number of tokens and C_r is the residual feature dimension.

In the gate prediction branch, the latent representation z_{deg} of the current degraded image is converted into visual tokens X_{deg}. Meanwhile, the current dominant degradation d_t is converted into a corresponding degradation prompt and encoded into the textual embedding $e(d_t)$. To condition the local residual assessment on the current restoration objective, we use the visual tokens as Query and the degradation-text tokens as Key and Value to perform cross-attention:

$$Q = f_q(X_{\text{deg}}),\ K = f_k(e(d_t)),\ V = f_v(e(d_t)), \tag{3}$$

$$A = \text{Softmax}\left(\frac{QK^\top}{\sqrt{d_k}}\right)V. \tag{4}$$

An MLP predicts the residual condition c from visual tokens fused with the cross-attention output:

$$c = \text{MLP}\left(\text{LN}(X_{\text{deg}} + A)\right), \tag{5}$$

where $\text{LN}(\cdot)$ denotes layer normalization.

During training, the residual condition c is supervised by the latent residual target r_{target}. A linear projection followed by a sigmoid function further predicts the degradation gate, which serves as the fusion coefficient for the subsequent dual-branch LoRA:

$$g = \sigma(W_g c + b_g), \quad g \in [0,1]^{B\times N\times 1}. \tag{6}$$

Figure 4 visualizes the predicted degradation gates under different degradation scenarios. Higher gate responses are mainly concentrated in regions strongly affected by the current degradation, whereas relatively stable content regions generally exhibit lower responses.

At inference time, the target image is unavailable, so the ground-truth latent residual r is not computed. The trained residual assessor predicts the residual condition c from the current image latent and the dominant-degradation prompt, and further maps c to the degradation gate g.

Degradation-Gated Dual-Branch LoRA. Standard LoRA shares the same low-rank mappings across spatial locations and lacks an explicit mechanism for modulating low-rank updates according to local degradation responses. To address this limitation, we divide the LoRA parameter updates into a content-preservation branch and a degradation-restoration branch, and use the degradation gate predicted by DLRA to adaptively fuse the two branches.

For a frozen linear layer with the original weight $W_0 \in \mathbb{R}^{d\times k}$ and an input token sequence x, the output of conventional LoRA is formulated as:

$$y = W_0 x + \Delta_{\text{LoRA}}(x) = W_0 x + BAx, \quad (7)$$

where $B \in \mathbb{R}^{d\times r}$ and $A \in \mathbb{R}^{r\times k}$ are low-rank matrices. Unlike conventional LoRA, DGDB-LoRA decomposes its update into the content-preservation and degradation-restoration branches:

$$\begin{aligned} \Delta_{\text{pres}}(x) &= B_{\text{pres}} A_{\text{pres}} x, \\ \Delta_{\text{rest}}(x) &= B_{\text{rest}} A_{\text{rest}} x. \end{aligned} \quad (8)$$

The two branches are adaptively fused according to the region-wise degradation gate g:

$$y = W_0 x + (1-g) \odot \Delta_{\text{pres}}(x) + g \odot \Delta_{\text{rest}}(x), \quad (9)$$

where $\odot$ denotes element-wise multiplication.

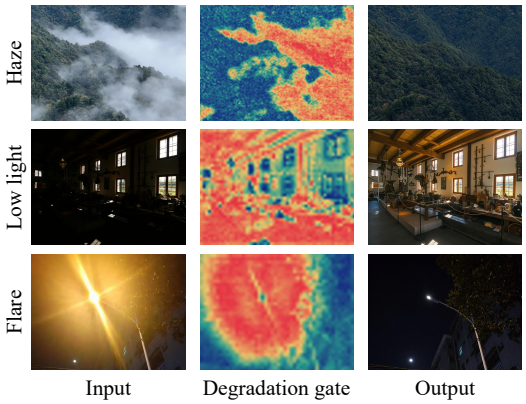

Figure 4: Visualization of predicted degradation gates, where higher responses indicate more severely degraded regions.

3.3 Optimization Objectives

The Dominant Degradation Planner is trained using standard supervised fine-tuning. This section mainly introduces the additional optimization objectives for the All-in-One restoration model. The model retains the original diffusion loss of Qwen-Image-Edit-2511, denoted as $\mathcal{L}_{\text{base}}$. On this basis, we introduce a residual prediction loss, a degradation gate loss, and a spatial smoothness loss to supervise DLRA in learning region-level restoration requirements.

Residual Prediction Loss. DLRA predicts the token-wise residual condition c from the current image features and dominant degradation condition, supervised by the latent residual target r_{target}:

$$\mathcal{L}_{\text{res}} = \text{MSE}(c, r_{\text{target}}). \quad (10)$$

This loss encourages c to capture region-level differences between the current and target images.

Degradation Gate Loss. To supervise the degradation gate, we average the latent residual target along the channel dimension and map it to $[0, 1]$:

$$g_{\text{target}} = \phi(\text{Mean}_c(r_{\text{target}})), \quad (11)$$

$$\mathcal{L}_{\text{gate}} = \|g - g_{\text{target}}\|_1, \quad (12)$$

where, $\text{Mean}_c(\cdot)$ denotes channel-wise averaging, and $\phi(\cdot)$ maps the result to the gate range. This loss aligns the predicted gate with local restoration requirements.

Spatial Smoothness Loss. To encourage spatial continuity of the degradation gate, we reshape it into a spatial gate map $G \in \mathbb{R}^{H_g \times W_g}$ and introduce a spatial smoothness loss:

$$\mathcal{L}_{\text{tv}} = \frac{\|\nabla_h G\|_1 + \|\nabla_w G\|_1}{H_g W_g}, \quad (13)$$

Table 1: Quantitative comparison on RealComp-Bench. CDFS denotes the Compound Degradation Fusion Score, while **CDRS denotes the Compound Degradation Restoration Score and serves as the primary metric for evaluating compound-degradation restoration**. The best and second-best results are marked in **bold** and underlined, respectively.

Method	Q-Align↑	MANIQA↑	MUSIQ↑	CLIP-IQA↑	**CDRS↑**	CDFS↑
Input	3.3634	0.3310	51.9654	0.4627	0.000	1.9860
GPT-Image-2	4.046	**0.448**	65.448	**0.636**	1.149	**2.777**
Nano Banana 2	3.985	0.414	62.802	0.572	0.985	2.608
Seedream-5.0-lite	4.058	0.432	64.658	0.618	0.878	2.684
Qwen-Image-Edit-2511	3.530	0.328	51.701	0.481	0.207	2.073
DDPR	**4.082**	0.410	**65.510**	0.589	**1.381**	2.747



Figure 5: Qualitative comparisons on RealComp-Bench. The first four columns show the progressive restoration trajectory and final result of DDPR, while the remaining columns present the results of competing methods.

where, ∇_h and ∇_w denote the horizontal and vertical finite-difference operators, respectively. This loss encourages neighboring regions to produce similar gate values, making the predicted degradation gate more consistent with the spatial distribution of degradations in the image.

Finally, the overall optimization objective of the All-in-One restoration model is defined as:

$$\mathcal{L} = \mathcal{L}_{\text{base}} + \lambda_{\text{res}}\mathcal{L}_{\text{res}} + \lambda_{\text{gate}}\mathcal{L}_{\text{gate}} + \lambda_{\text{tv}}\mathcal{L}_{\text{tv}}, \tag{14}$$

where λ_{res}, λ_{gate}, and λ_{tv} weight residual prediction, degradation gate, and spatial smoothness losses. Throughout experiments, we use 0.1, 0.05, and 0.001, respectively, for stable training.

4 Experiment

4.1 Experimental Setting

Datasets. We adopt three complementary benchmarks: RealComp-Bench for real-world compound degradations, AgenticIR Benchmark for synthetic compound degradations, and RealIR-Bench for real-world single degradations. RealComp-Bench is constructed in this work and contains 221 real-world compound-degraded images collected from the Internet, covering diverse combinations of blur, low-light, rain, haze, noise, flare, and reflection.

Metrics. We use PSNR, SSIM (Wang et al., 2004), and LPIPS (Zhang et al., 2018) as full-reference metrics, and Q-Align, MANIQA, CLIP-IQA, and MUSIQ as no-reference metrics. For RealIR-Bench, we follow RealRestorer's evaluation protocol (Yang et al., 2026) and use LPS, RS, and FS. Because RS evaluates only a single degradation, it cannot capture restoration performance under coexisting degradations. We therefore extend RS to the Compound Degradation Restoration Score (CDRS) and define the Compound Degradation Fusion Score (CDFS).

Let the degradation set involved in the input image be $D_{\text{in}} = \{d_1, d_2, \ldots, d_n\}$. For each degradation d, its input severity and restoration improvement are defined as:

$$\text{Severity}_d = 5 - V_d(I_{\text{in}}), \tag{15}$$

$$\mathrm{Diff}_d = \max(0,\, V_d(I_{\mathrm{out}}) - V_d(I_{\mathrm{in}}))\,, \tag{16}$$

where $V_d(\cdot)$ denotes the VLM score for degradation d, following the same computation as RS. CDRS weights the restoration improvement of each degradation according to its severity in the input:

$$\mathrm{CDRS} = \frac{\sum_{d\in D_{\mathrm{in}}} \mathrm{Severity}_d \cdot \mathrm{Diff}_d}{\sum_{d\in D_{\mathrm{in}}} \mathrm{Severity}_d}. \tag{17}$$

CDFS combines no-reference image quality with compound-degradation restoration performance:

$$\mathrm{CDFS} = \frac{\text{Q-Align}}{5} + \mathrm{MANIQA} + \frac{\mathrm{MUSIQ}}{100} \quad + \text{CLIP-IQA} + \frac{\mathrm{CDRS}}{5}. \tag{18}$$

Implementation Details. All experiments are implemented using PyTorch. For DDPR, we set the maximum number of iterations to $T_{\max} = 9$, matching the number of degradation types considered in this work. Additional experimental details are provided in the appendix.

4.2 Comparison with State-of-the-Art Methods

RealComp-Bench. Since most open-source methods cannot simultaneously handle all degradation types considered in this work, we mainly compare DDPR with GPT-Image-2, Nano Banana 2, Seedream-5.0-lite, and Qwen-Image-Edit-2511 (Wu et al., 2025), using a fixed prompt for one-step restoration. Table 1 reports the quantitative results. DDPR achieves the best CDRS and obtains the highest Q-Align and MUSIQ scores. Its CDFS is close to that of GPT-Image-2 and surpasses Nano Banana 2, Seedream-5.0-lite, and Qwen-Image-Edit-2511. Figure 5 presents qualitative comparisons among different methods on real-world examples.

Table 2: Quantitative results on the AgenticIR Benchmark. Higher is better for all metrics except LPIPS.

Degradations	Method	PSNR	SSIM	LPIPS↓	MANIQA	CLIP-IQA	MUSIQ
Group A	AirNet	19.13	0.6019	0.4283	0.2581	0.3930	42.46
	PromptIR	20.06	0.6088	0.4127	0.2633	0.4013	42.62
	MiOIR	20.84	0.6558	0.3715	0.2451	0.3933	47.82
	DA-CLIP	19.58	0.6032	0.4266	0.2418	0.4139	42.51
	InstructIR	18.03	0.5751	0.4429	0.2660	0.3528	45.77
	AutoDIR	19.64	0.6286	0.3967	0.2500	0.3767	47.01
	AgenticIR	21.04	0.6818	0.3148	0.3071	0.4474	56.88
	DDPR	**21.62**	**0.7016**	**0.1593**	**0.3851**	**0.6054**	**65.52**
Group B	AirNet	19.31	0.6567	0.3670	0.2882	0.4274	47.88
	PromptIR	20.47	0.6704	0.3370	0.2893	0.4289	48.10
	MiOIR	20.56	0.6905	0.3243	0.2638	0.4330	51.87
	DA-CLIP	18.56	0.5946	0.4405	0.2435	0.4154	43.70
	InstructIR	18.34	0.6235	0.4072	0.3022	0.3790	50.94
	AutoDIR	19.90	0.6643	0.3542	0.2534	0.3986	49.64
	AgenticIR	20.55	0.7009	0.3072	0.3204	0.4648	57.57
	DDPR	**21.63**	**0.7077**	**0.1620**	**0.3801**	**0.6119**	**65.16**
Group C	AirNet	17.95	0.5145	0.5782	0.1854	0.3113	30.12
	PromptIR	18.51	0.5166	0.5756	0.1906	0.3104	29.71
	MiOIR	15.63	0.4896	0.5376	0.1717	0.2891	37.95
	DA-CLIP	18.53	0.5320	0.5335	0.1916	0.3476	33.87
	InstructIR	17.09	0.5135	0.5582	0.1732	0.2537	33.69
	AutoDIR	18.61	0.5443	0.5019	0.2045	0.2939	37.86
	AgenticIR	18.82	0.5474	0.4493	0.2698	0.3948	48.68
	DDPR	**19.83**	**0.6448**	**0.2111**	**0.3441**	**0.5778**	**58.49**

AgenticIR Benchmark. The benchmark is divided into three groups. Groups A and B mainly contain two-degradation combinations, whereas Group C contains more challenging three-degradation combinations. We compare DDPR with AirNet (Li et al., 2022), PromptIR (Potlapalli et al., 2023), MiOIR (Kong et al., 2024), DA-CLIP (Luo et al., 2024), InstructIR (Conde et al., 2024), AutoDIR (Jiang et al., 2024), and AgenticIR (Zhu et al., 2025). Table 2 reports the quantitative results. DDPR achieves the best results across all metrics in all three groups, showing consistent gains in full-reference metrics (PSNR, SSIM, and LPIPS) and no-reference metrics (MANIQA, CLIP-IQA, and MUSIQ).

RealIR-Bench. Since each image's degradation type is known, we disable the planner and directly perform single-step restoration. We compare DDPR with GPT-Image-2, Nano Banana 2, Seedream-5.0-lite, Qwen-Image-Edit-2511 (Wu et al., 2025), DA-CLIP (Luo et al., 2024), FoundIR (Li et al., 2025), and RealRestorer (Yang et al., 2026). Table 3 reports FS on nine degradation types, while LPS, RS, and FS results are provided in the appendix. DDPR achieves the best average RS and FS among open-source methods, substantially outperforming Qwen-Image-Edit-2511. Its average FS approaches GPT-Image-2 and exceeds Nano Banana 2 and Seedream-5.0-lite. Because low LPS may result from conservative or negligible edits, it should not be interpreted alone. Figure 6 shows that some open-source methods obtain low LPS while performing little restoration.

4.3 Ablation Study

Table 3: FS results on nine RealIR-Bench tasks; higher is better. Complete LPS, RS, and FS results are provided in the appendix. The best and second-best open-source results are **bold** and underlined, respectively. "–" denotes unavailable results.

Method (Open-source)	Rain	Blur	Low-light	Haze	Reflection	Flare	Moiré	Noise	Compression	Avg.
GPT-Image-2 (No)	0.191	0.214	0.149	0.181	0.285	0.282	0.110	0.106	0.063	0.176
Nano Banana 2 (No)	0.208	0.207	0.150	0.174	0.201	0.183	0.111	0.094	0.061	0.154
Seedream-5.0-lite (No)	0.190	0.236	0.152	0.150	0.248	0.247	0.111	0.089	0.056	0.164
Qwen-Image-Edit-2511 (Yes)	0.151	0.003	0.016	0.028	0.126	0.125	0.096	0.060	0.027	0.070
DA-CLIP (Yes)	0.011	0.000	0.084	0.009	0.011	0.007	0.019	0.000	0.010	0.017
FoundIR (Yes)	0.043	0.000	0.017	0.004	–	–	–	0.008	0.007	–
RealRestorer (Yes)	0.135	0.159	0.110	0.090	0.230	0.247	**0.142**	0.090	**0.097**	0.144
DDPR (Yes)	**0.185**	**0.215**	**0.140**	**0.119**	**0.277**	**0.322**	0.113	**0.091**	0.064	**0.170**



Figure 6: Qualitative comparisons on RealIR-Bench. **Notably, some methods, such as DA-CLIP and FoundIR, produce outputs close to the inputs in several cases, resulting in limited restoration but lower LPS values.**

Table 4 reports the ablation results on RealComp-Bench. The variants include removing the Dominant Degradation Planner and performing one-step restoration with a fixed prompt (w/o Planner), replacing the dominant degradation order with a random order (w/o D-Order), and removing DGDB-LoRA (w/o DGDB) or DLRA individually.

Table 4: Ablation results on RealComp-Bench. The best results are highlighted in bold.

Setting	Q-Align	MANIQA	MUSIQ	CLIP-IQA	CDRS	CDFS
Input	3.3634	0.3310	51.9654	0.4627	0.000	1.9860
w/o Planner	3.6447	0.3668	57.075	0.5202	0.977	2.3821
w/o D-Order	3.9911	0.3888	61.5175	0.5537	1.277	2.6113
w/o DGDB	3.9441	0.3887	62.8685	0.5628	1.150	2.5990
w/o DLRA	3.9757	0.3942	64.3829	0.5633	1.287	2.6539
Ours (Full)	**4.0817**	**0.4096**	**65.5099**	**0.5887**	**1.381**	**2.7459**

Effectiveness of Restoration Planning. Removing the Planner consistently degrades all metrics, particularly CDRS and CDFS, demonstrating the importance of dynamic planning for real-world compound-degradation restoration. Using a random restoration order also reduces performance, indicating that the improvement arises not only from progressive restoration but also from selecting the dominant degradation according to the current image state.

Effectiveness of Adaptive Editing Components. Removing either DGDB-LoRA or DLRA decreases performance, validating the effectiveness of region-level residual assessment and gated dual-branch adaptation. The full DDPR achieves the best results across all metrics, demonstrating the complementary benefits of combining restoration planning and region-adaptive editing.

5 Conclusion

We proposed DDPR, a progressive All-in-One image restoration framework for compound degradations. Its Dominant Degradation Planner repeatedly identifies the dominant degradation and replans the restoration objective from the updated image. DLRA estimates region-level restoration requirements from latent residuals, while DGDB-LoRA adaptively balances content preservation and degradation restoration. Experiments across three benchmarks demonstrate strong performance under diverse degradations, approaching GPT-Image-2.

AI USE STATEMENT

Generative AI tools were used to assist with language polishing, LaTeX formatting, and limited code-related tasks during the preparation of this manuscript. All AI-assisted content was carefully reviewed, verified, and revised by the authors. The research ideas, methodology, experimental design, analysis, and conclusions were developed and validated by the authors. The authors take full responsibility for the final content of this work.

ETHICS STATEMENT

This work focuses on image restoration and does not involve human subjects, personally identifiable information, or sensitive personal data. We do not anticipate significant ethical concerns arising from the proposed method.

REPRODUCIBILITY STATEMENT

The method and optimization objectives are described in the main text. Appendix A describes training-data construction and planner supervision. Appendices C.1 and C.2 provide model configurations, training settings, and inference settings. Appendix C.3 provides the baseline prompt, and Appendix D reports additional results.

REFERENCES

Yuang Ai, Huaibo Huang, and Ran He. LoRA-IR: Taming low-rank experts for efficient all-in-one image restoration, 2024.

Shuai Bai, Yuxuan Cai, Ruizhe Chen, Keqin Chen, Xionghui Chen, Zesen Cheng, Lianghao Deng, Wei Ding, Chang Gao, Chunjiang Ge, Wenbin Ge, Zhifang Guo, Qidong Huang, Jie Huang, Fei Huang, Binyuan Hui, Shutong Jiang, Zhaohai Li, Mingsheng Li, Mei Li, Kaixin Li, Zicheng Lin, Junyang Lin, Xuejing Liu, Jiawei Liu, Chenglong Liu, Yang Liu, Dayiheng Liu, Shixuan Liu, Dunjie Lu, Ruilin Luo, Chenxu Lv, Rui Men, Lingchen Meng, Xuancheng Ren, Xingzhang Ren, Sibo Song, Yuchong Sun, Jun Tang, Jianhong Tu, Jianqiang Wan, Peng Wang, Pengfei Wang, Qiuyue Wang, Yuxuan Wang, Tianbao Xie, Yiheng Xu, Haiyang Xu, Jin Xu, Zhibo Yang, Mingkun Yang, Jianxin Yang, An Yang, Bowen Yu, Fei Zhang, Hang Zhang, Xi Zhang, Bo Zheng, Humen Zhong, Jingren Zhou, Fan Zhou, Jing Zhou, Yuanzhi Zhu, and Ke Zhu. Qwen3-VL technical report. *arXiv preprint arXiv:2511.21631*, 2025.

Tim Brooks, Aleksander Holynski, and Alexei A. Efros. InstructPix2Pix: Learning to follow image editing instructions. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 18392–18402, 2023.

Yuanhao Cai, Hao Bian, Jing Lin, Haoqian Wang, Radu Timofte, and Yulun Zhang. Retinexformer: One-stage retinex-based transformer for low-light image enhancement. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 12504–12513, 2023.

Jin Cao, Yi Cao, Li Pang, Deyu Meng, and Xiangyong Cao. HAIR: Hypernetworks-based all-in-one image restoration, 2024.

Haoyu Chen, Wenbo Li, Jinjin Gu, Jingjing Ren, Sixiang Chen, Tian Ye, Renjing Pei, Kaiwen Zhou, Fenglong Song, and Lei Zhu. RestoreAgent: Autonomous image restoration agent via multimodal large language models. In *Advances in Neural Information Processing Systems*, volume 37, pp. 110643–110666, 2024. doi: 10.52202/079017-3512.

Zihan Cheng, Liangtai Zhou, Dian Chen, Ni Tang, Xiaotong Luo, Yuan Xie, and Yanyun Qu. UniLDiff: Unlocking the power of diffusion priors for all-in-one image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 37465–37475, 2026.

Marcos V. Conde, Gregor Geigle, and Radu Timofte. InstructIR: High-quality image restoration following human instructions. In *European Conference on Computer Vision*, pp. 1–21. Springer, 2024. doi: 10.1007/978-3-031-72764-1_1.

Yuning Cui, Syed Waqas Zamir, Salman H. Khan, Alois Knoll, Mubarak Shah, and Fahad Shahbaz Khan. AdaIR: Adaptive all-in-one image restoration via frequency mining and modulation. In *International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=M5t0WvjfCg.

Chao Dong, Yubin Deng, Chen Change Loy, and Xiaoou Tang. Compression artifacts reduction by a deep convolutional network. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 576–584, 2015.

Hu Gao, Jing Yang, Ying Zhang, Ning Wang, Jingfan Yang, and Depeng Dang. Prompt-based ingredient-oriented all-in-one image restoration. *IEEE Transactions on Circuits and Systems for Video Technology*, 34(10):9458–9471, 2024. doi: 10.1109/TCSVT.2024.3398810.

Chun-Le Guo, Qixin Yan, Saeed Anwar, Runmin Cong, Wenqi Ren, and Chongyi Li. Image dehazing transformer with transmission-aware 3d position embedding. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5812–5820, 2022.

Yu Guo, Yuan Gao, Yuxu Lu, Huilin Zhu, Ryan Wen Liu, and Shengfeng He. OneRestore: A universal restoration framework for composite degradation. In *European Conference on Computer Vision*, pp. 255–272. Springer, 2024. doi: 10.1007/978-3-031-72655-2_15.

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. LoRA: Low-rank adaptation of large language models. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=nZeVKeeFYf9.

Xu Jiang, Gehui Li, Bin Chen, and Jian Zhang. Multi-agent image restoration. *International Journal of Computer Vision*, 134:205, 2026. doi: 10.1007/s11263-026-02792-5.

Yitong Jiang, Zhaoyang Zhang, Tianfan Xue, and Jinwei Gu. AutoDIR: Automatic all-in-one image restoration with latent diffusion. In *European Conference on Computer Vision*, pp. 340–359. Springer, 2024. doi: 10.1007/978-3-031-73661-2_19.

Junjie Ke, Qifei Wang, Yilin Wang, Peyman Milanfar, and Feng Yang. MUSIQ: Multi-scale image quality transformer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 5148–5157, 2021.

Xiangtao Kong, Chao Dong, and Lei Zhang. Towards effective multiple-in-one image restoration: A sequential and prompt learning strategy, 2024.

Bingchen Li, Xin Li, Yiting Lu, and Zhibo Chen. Hybrid agents for image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 22636–22647, 2026.

Boyun Li, Xiao Liu, Peng Hu, Zhongqin Wu, Jiancheng Lv, and Xi Peng. All-in-one image restoration for unknown corruption. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 17452–17462, 2022.

Chao Li, Yixiao Yang, Kun He, Stephen Lin, and John E. Hopcroft. Single image reflection removal through cascaded refinement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3565–3574, 2020.

Hao Li, Xiang Chen, Jiangxin Dong, Jinhui Tang, and Jinshan Pan. FoundIR: Unleashing million-scale training data to advance foundation models for image restoration. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 12626–12636, 2025.

Yunlong Lin, Zixu Lin, Haoyu Chen, Panwang Pan, Chenxin Li, Sixiang Chen, Kairun Wen, Yeying Jin, Wenbo Li, and Xinghao Ding. JarvisIR: Elevating autonomous driving perception with intelligent image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 22369–22380, 2025.

Shuaizheng Liu, Jianqi Ma, Lingchen Sun, Xiangtao Kong, and Lei Zhang. InstructRestore: Region-customized image restoration with human instructions. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

Wenyang Luo, Haina Qin, Zewen Chen, Libin Wang, Dandan Zheng, Yuming Li, Yufan Liu, Bing Li, and Weiming Hu. Visual-instructed degradation diffusion for all-in-one image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 12764–12777, 2025. doi: 10.1109/CVPR52734.2025.01191.

Ziwei Luo, Fredrik K. Gustafsson, Zheng Zhao, Jens Sjölund, and Thomas B. Schön. Controlling vision-language models for multi-task image restoration. In *International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=t3vnnLeajU.

Debabrata Mandal, Soumitri Chattopadhyay, Guansen Tong, and Praneeth Chakravarthula. UniCoRN: Latent diffusion-based unified controllable image restoration network across multiple degradations. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 1543–1553, 2026.

Jiawei Mao, Yu Yang, Xuesong Yin, Ling Shao, and Hao Tang. AllRestorer: All-in-one transformer for image restoration under adverse weather degradations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 48(6):6628–6641, 2026. doi: 10.1109/TPAMI.2026.3658598.

Anish Mittal, Rajiv Soundararajan, and Alan C. Bovik. Making a "completely blind" image quality analyzer. *IEEE Signal Processing Letters*, 20(3):209–212, 2013. doi: 10.1109/LSP.2012.2227726.

Seungjun Nah, Tae Hyun Kim, and Kyoung Mu Lee. Deep multi-scale convolutional neural network for dynamic scene deblurring. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3883–3891, 2017.

Vaishnav Potlapalli, Syed Waqas Zamir, Salman H. Khan, and Fahad Shahbaz Khan. PromptIR: Prompting for all-in-one image restoration. In *Advances in Neural Information Processing Systems*, volume 36, pp. 71275–71293, 2023. doi: 10.52202/075280-3121.

Xiangpeng Tian, Xiangyu Liao, Xiao Liu, Meng Li, and Chao Ren. Degradation-aware feature perturbation for all-in-one image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 28165–28175, 2025.

Huiqiang Wang, Mingchen Song, and Guoqiang Zhong. Dynamic degradation decomposition network for all-in-one image restoration, 2025.

Jianyi Wang, Kelvin C. K. Chan, and Chen Change Loy. Exploring CLIP for assessing the look and feel of images. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pp. 2555–2563, 2023. doi: 10.1609/aaai.v37i2.25353.

Tianyu Wang, Xin Yang, Ke Xu, Shaozhe Chen, Qiang Zhang, and Rynson W. H. Lau. Spatial attentive single-image deraining with a high quality real rain dataset. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 12270–12279, 2019.

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. doi: 10.1109/TIP.2003.819861.

Chenfei Wu, Jiahao Li, Jingren Zhou, Junyang Lin, Kaiyuan Gao, Kun Yan, Sheng-ming Yin, Shuai Bai, Xiao Xu, Yilei Chen, Yuxiang Chen, Zecheng Tang, Zekai Zhang, Zhengyi Wang, An Yang, Bowen Yu, Chen Cheng, Dayiheng Liu, Deqing Li, Hang Zhang, Hao Meng, Hu Wei, Jingyuan Ni, Kai Chen, Kuan Cao, Liang Peng, Lin Qu, Minggang Wu, Peng Wang, Shuting Yu, Tingkun Wen, Wensen Feng, Xiaoxiao Xu, Yi Wang, Yichang Zhang, Yongqiang Zhu, Yujia Wu, Yuxuan Cai, and Zenan Liu. Qwen-Image technical report, 2025.

Haoning Wu, Zicheng Zhang, Weixia Zhang, Chaofeng Chen, Liang Liao, Chunyi Li, Yixuan Gao, Annan Wang, Erli Zhang, Wenxiu Sun, Qiong Yan, Xiongkuo Min, Guangtao Zhai, and Weisi Lin. Q-Align: Teaching LMMs for visual scoring via discrete text-defined levels. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 54015–54029. PMLR, 2024.

Sidi Yang, Tianhe Wu, Shuwei Shi, Shanshan Lao, Yuan Gong, Mingdeng Cao, Jiahao Wang, and Yujiu Yang. MANIQA: Multi-dimension attention network for no-reference image quality assessment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1191–1200, 2022.

Yufeng Yang, Xianfang Zeng, Zhangqi Jiang, Fukun Yin, Jianzhuang Liu, Wei Cheng, Jinghong Lan, Shiyu Liu, Yuqi Peng, Gang Yu, and Shifeng Chen. RealRestorer: Towards generalizable real-world image restoration with large-scale image editing models, 2026.

Mingde Yao, Ruikang Xu, Yuanshen Guan, Jie Huang, and Zhiwei Xiong. Neural degradation representation learning for all-in-one image restoration. *IEEE Transactions on Image Processing*, 33:5408–5423, 2024. doi: 10.1109/TIP.2024.3456583.

Cheng Zhang, Yu Zhu, Qingsen Yan, Jinqiu Sun, and Yanning Zhang. All-in-one multi-degradation image restoration network via hierarchical degradation representation. In *Proceedings of the 31st ACM International Conference on Multimedia*, pp. 2285–2293, 2023. doi: 10.1145/3581783.3611825.

Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, 2017. doi: 10.1109/TIP.2017.2662206.

Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 586–595, 2018. doi: 10.1109/CVPR.2018.00068.

Xu Zhang, Jiaqi Ma, Guoli Wang, Qian Zhang, Huan Zhang, and Lefei Zhang. Perceive-IR: Learning to perceive degradation better for all-in-one image restoration. *IEEE Transactions on Image Processing*, 35:2018–2033, 2026a. doi: 10.1109/TIP.2025.3566300.

Xu Zhang, Huan Zhang, Guoli Wang, Qian Zhang, and Lefei Zhang. ClearAIR: A human-visual-perception-inspired all-in-one image restoration. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 12861–12869, 2026b. doi: 10.1609/aaai.v40i15.38284.

Yingjie Zhou, Jiezhang Cao, Zicheng Zhang, Farong Wen, Yanwei Jiang, Jun Jia, Xiaohong Liu, Xiongkuo Min, and Guangtao Zhai. Q-Agent: Quality-driven chain-of-thought image restoration agent through robust multimodal large language model, 2025.

Kaiwen Zhu, Jinjin Gu, Zhiyuan You, Yu Qiao, and Chao Dong. An intelligent agentic system for complex image restoration problems. In *International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=3RLxccFPHz.

APPENDIX

A DOMINANT DEGRADATION TRAINING DATA CONSTRUCTION

Figure S1 presents the complete training data construction pipeline of DDPR, including synthetic restoration trajectory construction, real-world restoration pair construction, training data construction for the Dominant Degradation Planner, and training data construction for the All-in-One restoration model. This section provides additional details on the data annotation and quality assessment procedures that are not fully described in the main paper.

A.1 CANDIDATE-STATE QUALITY ASSESSMENT

As illustrated in Figure S1(a), at step t, for each candidate degradation $d_{t,k}$ in the remaining degradation sequence, we construct the corresponding candidate state $I_t^{(-d_{t,k})}$. We then use five no-reference image quality assessment metrics, including Q-Align, MANIQA, MUSIQ, CLIP-IQA, and NIQE (Wu et al., 2024; Yang et al., 2022; Ke et al., 2021; Wang et al., 2023; Mittal et al., 2013), to measure the quality improvement of the candidate state over the current state I_t. The aggregated score of each candidate state is defined as:

$$S\left(I_t^{(-d_{t,k})}\right) = \frac{1}{5}\sum_{m=1}^{5} \frac{Q_m\left(I_t^{(-d_{t,k})}\right) - Q_m(I_t)}{|Q_m(I_t)|}, \tag{19}$$

where $Q_m(\cdot)$ denotes the m-th no-reference image quality assessment metric. For Q-Align, MANIQA, MUSIQ, and CLIP-IQA, we use their original quality scores. Since a lower NIQE score indicates better image quality, we convert it into a higher-is-better NIQE-quality score as:

$$Q_{\text{NIQE-quality}}(I) = \exp\left(-\frac{\text{NIQE}(I)}{10}\right). \tag{20}$$

After this conversion, all five metrics follow the same higher-is-better direction. Since the candidate score is computed using the relative quality improvement over the current state, the numerical scales of different metrics are normalized to a certain extent. We therefore do not introduce additional manually designed weights and instead assign equal weight to all five metrics, i.e., $1/5$ for each metric. The degradation corresponding to the candidate state with the highest aggregated score is selected as the current dominant degradation.

A.2 REAL-WORLD RESTORATION PAIR CONSTRUCTION

As shown in Figure S1(b), to improve the generalization capability of the restoration model to real-world degradations, we further collect real images from the Internet that exhibit a clear single degradation. For each image, GPT-Image-2 and Nano Banana 2 are independently used to generate two candidate restoration results.

The two candidates are then manually compared and inspected, and the result with better restoration quality and more faithful content preservation is selected as the final target image. The screening process mainly considers the following criteria:

- whether the target degradation is effectively removed;
- whether real objects and scene structures remain consistent;
- whether any object insertion, deletion, or local redrawing occurs;
- whether people, text, and fine-grained textures undergo unreasonable changes; and
- whether new artifacts, color shifts, or over-smoothing are introduced.

If neither candidate satisfies the above requirements, the sample is discarded. Finally, each real degraded image and its manually selected high-quality restoration result form a real-world single-step restoration pair. Unlike the synthetic trajectories, these data do not contain predefined multi-

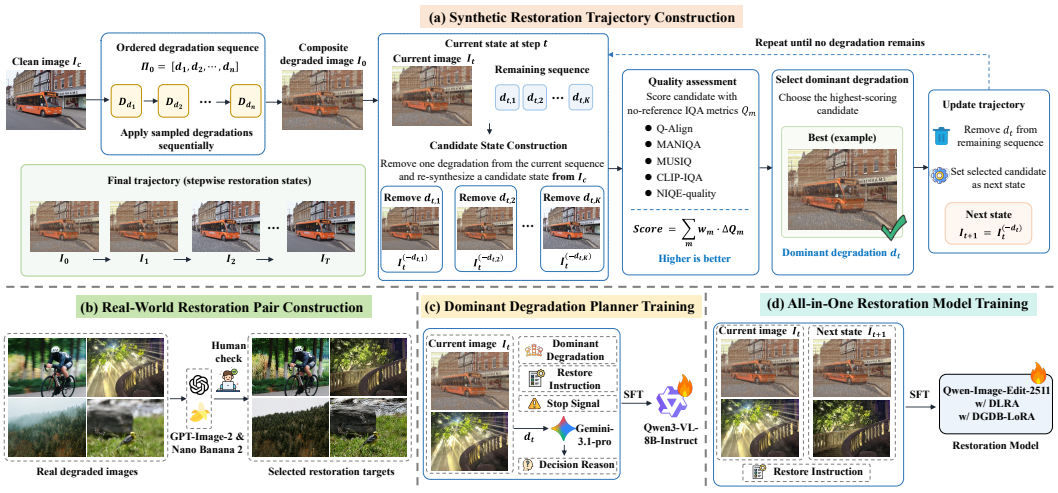

Figure S1: Training data construction pipeline of DDPR.

step degradation sequences and are mainly used to complement complex real-world degradation distributions that are difficult to cover through synthetic data alone.

Synthetic restoration trajectories and real-world restoration pairs play complementary roles during training. Synthetic trajectories provide controllable degradation combinations, explicit intermediate restoration states, and stepwise supervision, allowing a trajectory to be expanded into multiple training samples at restoration stages. This enables the Dominant Degradation Planner to learn how to adjust the next restoration decision according to changes in the current image state. Meanwhile, degradations in real images often exhibit more complex spatial distributions, severity variations, and imaging characteristics, which are difficult to fully simulate using predefined degradation models. Therefore, real-world restoration pairs are mainly used to complement the real degradation distributions not covered by synthetic data and improve the generalization capability of the All-in-One restoration model in real-world scenarios. Finally, adjacent states in synthetic trajectories and real-world restoration pairs jointly form single-step restoration samples, while each current state and its corresponding supervision fields are used to train the Dominant Degradation Planner. This data organization not only ensures the controllability of stepwise planning supervision, but also enhances the restoration model's ability to adapt to complex real-world degradation distributions.

A.3 Planner Supervision and Prompt Design

As shown in Figure S1(c), each training sample for the Dominant Degradation Planner consists of the current image and four supervision fields: *Decision Reason*, *Dominant Degradation*, *Restore Instruction*, and *Stop Signal*.

For each state in a synthetic restoration trajectory, the Dominant Degradation label is determined by the candidate-state quality assessment described in Section A.1. For a real-world single-degradation image, its known degradation category is directly used as the Dominant Degradation label.

For each Dominant Degradation, we use a fixed Restore Instruction mapping, as summarized in Table S1. To reduce prompt-formulation effects on restoration performance and ensure a fair comparison with RealRestorer (Yang et al., 2026), we adopt the same restoration instructions as RealRestorer. When the Dominant Degradation is `none`, no obvious degradation remains to be processed. In this case, the Restore Instruction is set to `none`, and the Stop Signal is set to `true`.

Decision Reason Generation. We use Gemini-3.1-pro to generate the *Decision Reason* field. Specifically, the current image and its predetermined Dominant Degradation label are provided to Gemini, which is instructed to generate a concise explanation based on the visible degradation cues in the current image. Gemini only generates the natural-language explanation and does not redetermine the dominant degradation type. This design ensures that the Decision Reason remains

Table S1: Mapping between degradation types and restoration instructions.

Task	Restore Instruction
None	none
Blur	Please deblur the image and make it sharper.
Moiré	Please remove the moiré patterns from the image.
Haze	Please dehaze the image.
Noise	Please remove noise from the image.
Reflection	Please remove the reflection from the image.
Compression	Please restore the image clarity and artifacts.
Flare	Please remove the lens flare and glare from the image.
Rain	Please remove the rain from the image and restore its clarity.
Low-light	Please restore this low-quality image, recovering its normal brightness and clarity.

consistent with the predetermined supervision label. When the Dominant Degradation is set to `none`, the Decision Reason explains that no obvious degradation remains to be processed.

Planner System Prompt. The system prompt used during planner training is provided below. The token `<image>` denotes the current image input, while `<restoration_instruction_mapping>` is a placeholder that is dynamically replaced by the predefined restoration-instruction dictionary in Table S1 before the prompt is fed to the planner.

```
<image>
You are a visual restoration planning model.

Given the current degraded image, decide the next restoration step.

Restoration instruction mapping:
<restoration_instruction_mapping>

Important rules:
- Select the most dominant remaining degradation in the current image.
- If no obvious remaining degradation should be repaired, output dominant_degradation as none
     and stop as true.
- dominant_degradation must be selected only from this fixed list: blur, haze, noise, rain,
    compression, lowlight, reflection, moire, flare, none.
- restore_instruction must exactly match the instruction associated with the selected
    dominant_degradation in the restoration instruction mapping above.

Output valid JSON only. The JSON must contain exactly these keys:
- dominant_degradation: one of blur, haze, noise, rain, compression, lowlight, reflection,
    moire, flare, none
- reason: a concise reason explaining why this is the current dominant degradation, or why
    restoration should stop
- restore_instruction: the restoration instruction for the next step
- stop: true or false
```

B CONSTRUCTION OF REALCOMP-BENCH

This section provides additional details on the combination distribution of RealComp-Bench. The benchmark contains 221 real-world compound-degraded images collected from the Internet, without corresponding clean reference images. It covers real-world degradation types including blur, low-light, haze, rain, noise, flare, and reflection. In total, RealComp-Bench contains 21 compound-degradation combinations, including 11 two-degradation combinations, nine three-degradation combinations, and one four-degradation combination, as summarized in Table S2.

These combinations cover different numbers of degradations and diverse degradation co-occurrence patterns, enabling the evaluation of restoration performance under varying levels of compound-degradation complexity. Since all images are collected from real-world scenes, their degradation severity, spatial distributions, and visual characteristics are more complex and closer to practical image quality problems. Therefore, RealComp-Bench complements synthetic compound-degradation benchmarks in terms of degradation realism, combination diversity, and scene complexity.

Table S2: Compound-degradation combinations included in RealComp-Bench.

Number of Degradations	Compound-Degradation Combinations
Two	lowlight+haze; lowlight+flare; lowlight+rain; blur+lowlight; haze+rain; blur+noise; noise+flare; blur+flare; haze+flare; blur+haze; blur+rain
Three	rain+flare+lowlight; haze+flare+lowlight; rain+haze+lowlight; blur+lowlight+flare; rain+blur+lowlight; lowlight+haze+noise; blur+noise+lowlight; reflection+rain+blur; reflection+flare+lowlight
Four	blur+rain+flare+lowlight

Table S3: Complete quantitative comparison on all nine RealIR-Bench tasks. The best and second-best results among open-source methods are marked in **bold** and underlined, respectively. **Since LPS measures the perceptual distance between the input and restored result, a low LPS may also arise when a model makes little change to the input; therefore, LPS should be interpreted jointly with RS and FS.** "–" denotes unavailable results.

Method (Open-source)	Rain			Blur			Low-light			Haze			Reflection		
	LPS↓	RS↑	FS↑	LPS↓	RS↑	FS↑	LPS↓	RS↑	FS↑	LPS↓	RS↑	FS↑	LPS↓	RS↑	FS↑
GPT-Image-2 (No)	0.498	1.906	0.191	0.575	2.520	0.214	0.516	1.543	0.149	0.585	2.178	0.181	0.417	2.440	0.285
Nano Banana 2 (No)	0.442	1.868	0.208	0.450	1.880	0.207	0.506	1.522	0.150	0.574	2.044	0.174	0.436	1.780	0.201
Seedream-5.0-lite (No)	0.433	1.679	0.190	0.500	2.360	0.236	0.428	1.326	0.152	0.495	1.489	0.150	0.381	2.000	0.248
Qwen-Image-Edit-2511 (Yes)	0.420	1.302	0.151	0.174	0.020	0.003	0.103	0.087	0.016	0.227	0.178	0.028	0.233	0.820	0.126
DA-CLIP (Yes)	0.045	0.057	0.011	0.025	0.000	0.000	0.224	0.543	0.084	**0.026**	0.044	0.009	**0.043**	0.060	0.011
FoundIR (Yes)	**0.039**	0.226	0.043	**0.022**	0.000	0.000	**0.025**	0.087	0.017	0.032	0.022	0.004	–	–	–
RealRestorer (Yes)	0.371	1.076	0.135	0.582	1.900	0.159	0.597	1.360	0.110	0.339	0.680	0.090	0.290	1.620	0.230
DDPR (Yes)	0.431	**1.623**	**0.185**	0.502	**2.160**	**0.215**	0.520	**1.457**	**0.140**	0.429	**1.044**	**0.119**	0.353	**2.140**	**0.277**

Method (Open-source)	Flare			Moiré			Noise			Compression			Avg.		
	LPS↓	RS↑	FS↑	LPS↓	RS↑	FS↑	LPS↓	RS↑	FS↑	LPS↓	RS↑	FS↑	LPS↓	RS↑	FS↑
GPT-Image-2 (No)	0.339	2.132	0.282	0.596	1.360	0.110	0.502	1.061	0.106	0.604	0.800	0.063	0.515	1.771	0.176
Nano Banana 2 (No)	0.255	1.226	0.183	0.580	1.320	0.111	0.449	0.857	0.094	0.494	0.600	0.061	0.465	1.455	0.154
Seedream-5.0-lite (No)	0.265	1.679	0.247	0.575	1.300	0.111	0.502	0.898	0.089	0.581	0.667	0.056	0.462	1.489	0.164
Qwen-Image-Edit-2511 (Yes)	0.227	0.811	0.125	0.598	1.200	0.096	0.473	0.571	0.060	0.180	0.167	0.027	0.293	0.573	0.070
DA-CLIP (Yes)	**0.027**	0.038	0.007	**0.197**	0.120	0.019	**0.037**	0.000	0.000	0.025	0.050	0.010	**0.072**	0.101	0.017
FoundIR (Yes)	–	–	–	–	–	–	0.074	0.041	0.008	**0.005**	0.033	0.007	–	–	–
RealRestorer (Yes)	0.239	1.623	0.247	0.563	**1.620**	**0.142**	0.478	0.863	0.090	0.547	**1.067**	**0.097**	0.445	1.312	0.144
DDPR (Yes)	0.294	**2.283**	**0.322**	0.577	1.340	0.113	0.496	**0.898**	**0.091**	0.464	0.600	0.064	0.452	**1.505**	**0.170**

C Additional Implementation Details

C.1 Planner Configuration

The Dominant Degradation Planner uses Qwen3-VL-8B-Instruct (Bai et al., 2025) and is fine-tuned with LoRA (Hu et al., 2022) using 21,331 supervised samples. During training, the vision encoder is frozen, while the multimodal projection and LoRA parameters on the language side are optimized. The LoRA rank, scaling factor, and dropout rate are set to 32, 64, and 0.05, respectively.

The planner is trained for two epochs with an initial learning rate of 1×10^{-4}. We adopt a cosine learning-rate schedule with a warm-up ratio of 0.03. The per-GPU batch size is 1, and gradients are accumulated over four steps, resulting in a batch size of 4. Training is conducted using BF16 precision, together with DeepSpeed ZeRO-2 and gradient checkpointing to reduce memory consumption.

During inference, the planner operates in BF16 precision, and the temperature is set to 0 to produce deterministic structured predictions.

C.2 Restoration Model Configuration

The All-in-One restoration model is built upon Qwen-Image-Edit-2511 (Wu et al., 2025) and incorporates the DLRA and DGDB-LoRA modules. We fine-tune the model using 42,268 image restoration pairs. The target resolution of the training images is set to 1024×1024, and the per-GPU batch size is set to 1. To improve the model's robustness to missing conditions, the text condition and input image condition are independently dropped with a probability of 0.1.

Standard LoRA modules are attached to the attention projection layers, conditional modulation layers, and text MLP layers of the DiT, with the LoRA rank set to 128. DGDB-LoRA is mainly deployed in the last linear layer of the image-stream MLP and the final output projection layer of the DiT. These positions are located at the end of the image-stream feature transformation and the final output stage of the model, respectively, allowing them to directly regulate image feature updates and output predictions based on the aggregated spatial, semantic, and degradation information. DLRA employs a four-head cross-attention module to fuse the textual degradation condition, with its hidden feature dimension set to 64. The model is optimized using AdamW with an initial learning rate of 1×10^{-4} and a weight decay of 0.01.

During inference, the trained LoRA weights are loaded with a LoRA scale of 1.0. The input condition image is resized while preserving its original aspect ratio, with its maximum edge limited to 1,024 pixels. The maximum edge of the generated image is also limited to 1,024 pixels. The restoration model uses 30 sampling steps, with the classifier-free guidance scale set to 1.0.

C.3 Baseline Prompt on RealComp-Bench

To ensure a fair comparison among different methods on RealComp-Bench, we use the same general restoration prompt for image editing models, including GPT-Image-2, Nano Banana 2, Seedream-5.0-lite, and Qwen-Image-Edit-2511:

> This image contains one or more of the following degradation effects: “blur”, “compression”, “flare”, “haze”, “lowlight”, “moire”, “noise”, “rain”, “reflection”. Please restore the degraded image and recover a clean natural image.

This prompt only provides the set of degradation types considered in this work and asks the model to restore the input into a clean and natural image. It does not reveal the actual degradation types present in the current image, the number of degradations, or the restoration order. Therefore, each baseline method must independently identify the compound degradations in the input image and complete the restoration, avoiding additional priors introduced by method-specific prompt information and thereby ensuring a fair comparison.

D Additional Experimental Results

This section reports the complete quantitative results of different methods on RealIR-Bench, including LPS, RS, and FS, as shown in Table S3. It should be noted that a lower LPS may also result from making only minor changes to the input image. Therefore, LPS should be interpreted jointly with RS and FS, rather than being used independently as an indicator of restoration quality.

Considering the overall results across the nine tasks, DDPR achieves the best average RS and FS among open-source methods and obtains the highest FS on most real-world degradation tasks, demonstrating stable degradation restoration capability. Although some methods achieve lower LPS values, their RS and FS remain relatively limited, indicating that they may only make minor modifications to the degraded inputs. In contrast, DDPR achieves a better balance between degradation removal and content preservation. Its overall performance is close to that of the closed-source image editing model GPT-Image-2 and outperforms the other open-source baselines.

E Additional Qualitative Results

Figures S2 and S3 present additional qualitative results.

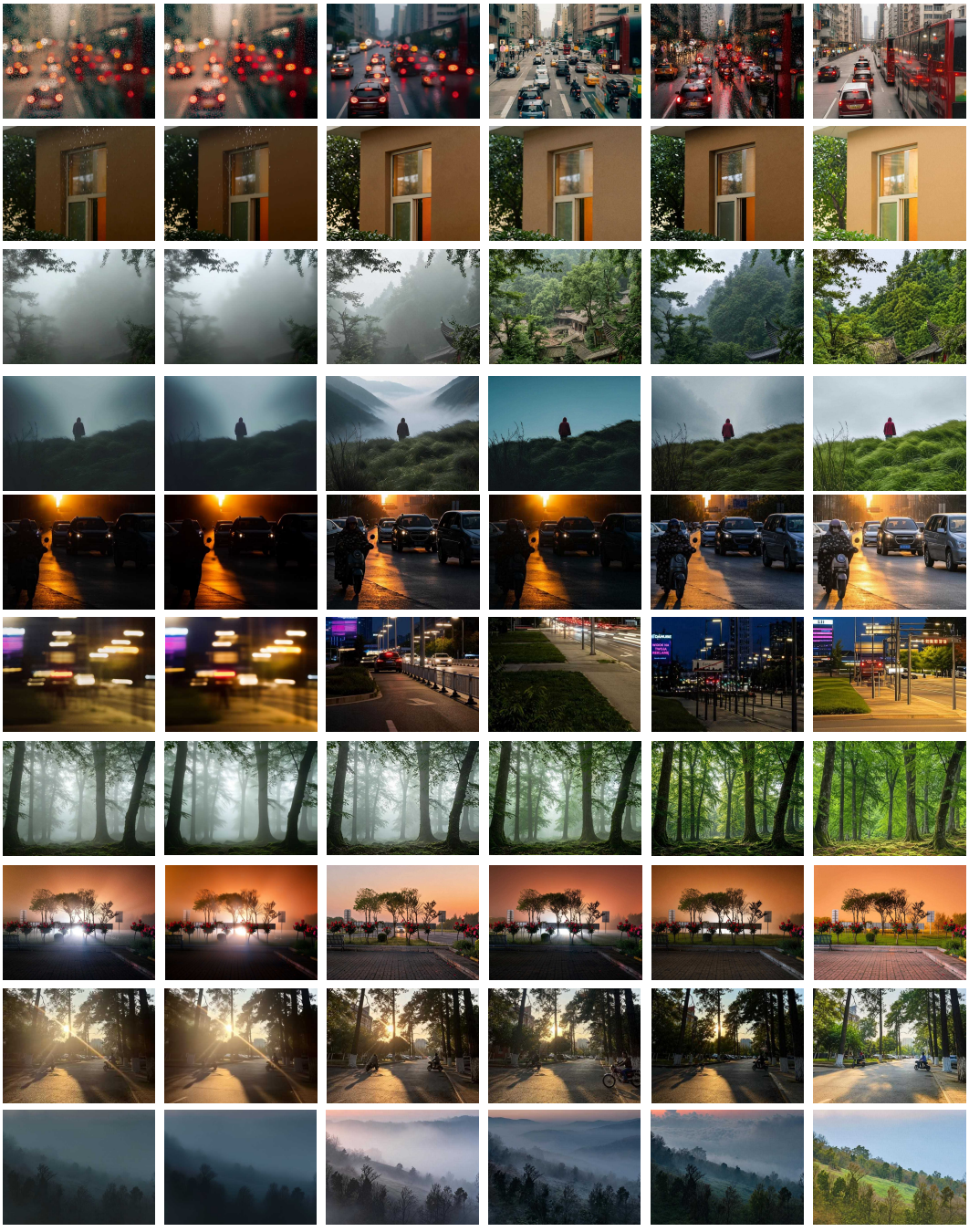

Figure S2: More Qualitative Results.

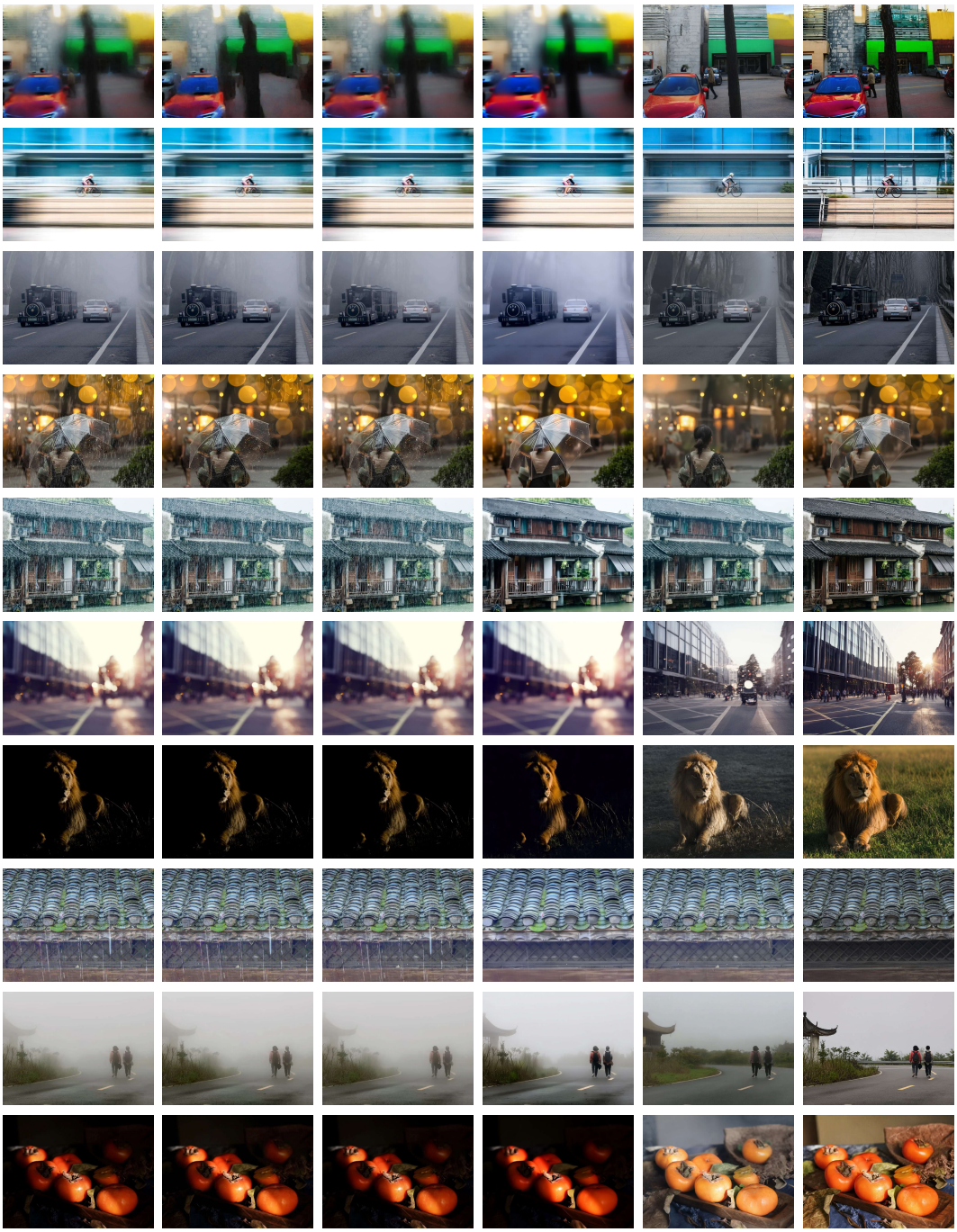

Figure S3: More Qualitative Results.