DEGRADATION-ADAPTIVE INVERSE OPERATOR LEARNING FOR ALL-IN-ONE IMAGE RESTORATION

Anonymous authors
Paper under double-blind review

ABSTRACT

All-in-one image restoration aims to recover images corrupted by heterogeneous degradations using a single unified model. Most existing approaches directly predict clean images or residuals, often relying on degradation-specific prompts or additional conditioning modules to distinguish different restoration tasks. In this work, we explore a different perspective and formulate all-in-one restoration as input-conditioned inverse operator learning. Instead of directly regressing restored pixels, we predict spatially adaptive inverse operators that transform degraded observations into clean estimates. Based on this formulation, we propose OPIR, an all-in-one image restoration framework that infers degradation representations directly from the input and uses them to modulate a shared restoration operator. This input-conditioned modulation allows the model to adapt to heterogeneous degradations without requiring degradation labels or manually specified task prompts at inference. To progressively improve difficult regions, we further introduce an operator-uncertainty-guided refinement strategy, where the distribution and statistical responses of the predicted local operators are used to identify pixels that require stronger restoration. A second restoration stage then selectively refines these challenging regions. In addition, an efficient multi-scale operator implementation reuses compact predicted kernels across different receptive fields, substantially reducing the computational overhead of spatially adaptive filtering. Extensive experiments demonstrate that OPIR consistently achieves superior performance across multiple all-in-one restoration benchmarks, while also remaining highly competitive on task-aligned restoration settings.

1 INTRODUCTION

Image restoration (IR) aims to recover a latent clean image J from an observed degraded image I. The degradation process can be generally formulated as

$$I = \mathcal{D}_t(J), \tag{1}$$

where $\mathcal{D}_t(\cdot)$ denotes a task-dependent degradation process and t represents the degradation type. IR is inherently ill-posed, since a degraded observation may correspond to multiple plausible clean solutions. Traditional methods He et al. (2011); Rong et al. (2024) address this problem by introducing task-specific priors, modeling degradation processes, and solving the corresponding inverse problems. However, their performance often relies on prior knowledge of degradation mechanisms, which limits adaptability to spatially varying or unknown degradations.

Deep learning methods Su et al. (2025); Cui et al. (2026); Zeng et al. (2025) have significantly advanced image restoration by learning image priors directly from data. According to the number of degradation types handled by a single model, existing methods can be broadly divided into task-specific, task-aligned, and all-in-one approaches. Task-specific methods Su et al. (2025); Guo et al. (2025b) optimize specialized models for individual degradations. Task-aligned methods Cui et al. (2023a); Gu et al. (2025) employ a shared architecture across different restoration tasks. In contrast, all-in-one image restoration (AIOIR) methods Potlapalli et al. (2023); Zeng et al. (2025) jointly learn from multiple degradation types and aim to restore diverse degraded inputs using a unified model.

Despite substantial progress, AIOIR remains challenging because heterogeneous degradations require different restoration behaviors. For example, denoising suppresses stochastic perturbations,

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.

deraining removes structured interference, deblurring recovers spatially distributed structures, while dehazing corrects large-scale visibility degradation. Jointly optimizing these heterogeneous objectives may therefore introduce cross-task interference. Existing AIOIR methods mainly alleviate this issue through degradation-aware feature adaptation, including learned prompts, degradation representations, feature disentanglement, and vision-language priors Potlapalli et al. (2023); Zeng et al. (2025); Luo et al. (2025). Although effective, these approaches primarily adapt intermediate features, while the final restoration is still commonly generated through a shared image- or residual-prediction mapping.

This motivates us to consider AIOIR from a different perspective: rather than only adapting features, can the restoration operation itself be adapted to the observed degradation? Instead of explicitly estimating $\mathcal{D}_t$ and analytically computing its inverse, we learn an input-conditioned spatial restoration operator from paired training data. Specifically, we propose OPIR, an input-conditioned inverse operator learning framework for all-in-one image restoration. OPIR extracts degradation characteristics directly from the input and uses them to modulate a shared spatial restoration operator. Since the conditioning information is inferred from the observation, the framework can adapt to heterogeneous degradation conditions without requiring explicit degradation labels during inference.

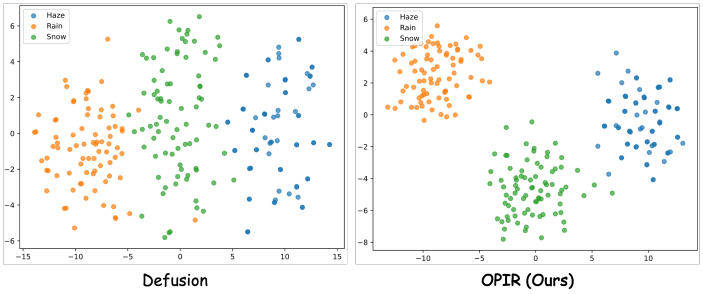

Figure 1: t-SNE visualization of degradation representations learned by OPIR and Defusion.

The predicted operator further provides information about the amount of transformation required at different spatial locations. Weakly degraded regions are expected to require operators close to an identity transformation, whereas severely corrupted regions generally require stronger spatial redistribution. We therefore introduce an operator uncertainty map derived from the statistical responses of the predicted local operators to characterize local restoration difficulty. Based on this measure, OPIR adopts a two-stage strategy: the first stage produces an initial restoration and its uncertainty map, while the second stage further refines challenging regions under uncertainty guidance. To improve efficiency, we additionally reuse a compact base operator across multiple dilated neighborhoods and adaptively fuse the multi-scale responses, enlarging the effective receptive field without predicting large spatial kernels. As shown in Figure 1, OPIR learns more distinguishable degradation representations. Extensive experiments under three-task and five-task all-in-one restoration settings further demonstrate favorable restoration performance.

The main contributions of this work are summarized as follows:

1. We formulate all-in-one image restoration as input-conditioned inverse operator learning. Instead of directly predicting restored pixels or residuals, OPIR learns spatially adaptive restoration operators conditioned on degraded observations, enabling heterogeneous inputs to induce different restoration behaviors within a shared model.

2. We introduce a degradation-adaptive operator modeling mechanism that automatically infers degradation representations from the input to modulate the predicted restoration operator, enabling adaptation to different degradation conditions without explicit degradation labels during inference.

3. We propose an operator-uncertainty-guided progressive restoration strategy, where statistical responses of the predicted operator characterize local restoration difficulty and guide second-stage refinement.

4. We develop an efficient multi-scale operator implementation that reuses compact predicted kernels across multiple receptive fields, improving computational efficiency while preserving effective spatial context.

2 Related Work

2.1 Image Restoration

Image restoration (IR) aims to recover high-quality images from degraded inputs, a challenging and inherently ill-posed problem. Traditional approaches He et al. (2011); Rong et al. (2024) tackle this by introducing handcrafted priors to constrain the solution space. However, these priors rely heavily on expert knowledge and often exhibit limited adaptability, resulting in poor generalization across diverse degradation scenarios.

With the rapid development of deep learning, numerous data-driven IR methods Chen et al. (2026); Gao et al. (2024c; 2025a); Peng et al. (2026) have been proposed to learn image priors from data and outperform traditional approaches. According to the number of degradation types handled by a single model, these methods can be broadly divided into task-specific, task-aligned, and all-in-one approaches. Task-specific methods Su et al. (2025); Guo et al. (2025b) design dedicated models for individual degradations. ALGNet Gao et al. (2024a) introduces local feature extraction for deblurring, while XYScanNet Liu et al. (2025) models spatial dependencies through alternating scanning. EfDeRain+ Guo et al. (2025b) formulates deraining as predictive filtering. Diffusion-based methods have also been explored, including UPID-EDM Wen et al. (2024) with vision-language priors, PGH^{2}Net Su et al. (2025) with image priors for dehazing, and Diff-Unmix Zeng et al. (2024) for self-supervised denoising. Task-aligned methods Cui et al. (2023a); Li et al. (2026) employ general restoration architectures across different degradation tasks. MambaIR Guo et al. (2024b) and MambaIRV2 Guo et al. (2025a) exploit state-space modeling, while SFNet Cui et al. (2023d) and FSNet Cui et al. (2023a) perform adaptive frequency selection. ACL Gu et al. (2025) introduces efficient linear-attention-based restoration, and MHNet Gao et al. (2025b) adopts a mixed hierarchical architecture for richer structural representation. All-in-one methods Mao et al. (2026); Zeng et al. (2025) jointly train a single model on multiple degradation types. PromptIR Potlapalli et al. (2023) and CAPTNet Gao et al. (2024b) employ prompt learning, while NDR Yao et al. (2024) learns shared degradation representations. Recent methods further exploit degradation disentanglement and large-scale priors: VLU-Net Zeng et al. (2025) uses vision-language features, AutoDIR Jiang et al. (2024) combines VLM guidance with latent diffusion, AdaIR Cui et al. (2025) separates degradation and content information, Defusion Luo et al. (2025) introduces visual-instruction-guided degradation diffusion, and Perceive-IR Zhang et al. (2025) models both degradation type and severity. Despite their different designs, most existing AIOIR methods perform degradation adaptation primarily in the feature space while retaining a shared image- or residual-prediction paradigm. In contrast, we investigate degradation adaptation directly in the restoration operator space, where input-dependent degradation representations modulate spatial restoration operators for heterogeneous degradations.

2.2 Adaptive Filtering and Operator-Based Restoration

Adaptive filtering techniques have been widely used in image restoration due to the local coherence, continuity, and structural properties of natural images. Traditional methods rely on handcrafted kernels and priors to process images locally, including adaptive Awate & Whitaker (2006), recursive Kundur & Hatzinakos (1998), bilateral Zhang & Gunturk (2008), and collaborative 3-D transform-domain approaches Dabov et al. (2007). While effective in certain cases, these methods often struggle with complex degradations such as rain streaks, where bilateral kernels may blur the image and guided kernels can introduce artifacts. With the rise of deep learning, predictive kernel methods have emerged as a powerful alternative. Kernel Prediction Networks (KPN) Mildenhall et al. (2018) learn per-pixel, image-adaptive kernels through deep convolutional networks, enabling flexible, data-driven restoration. These approaches have been successfully applied to various image restoration tasks Guo et al. (2025c; 2021); Fu et al. (2021). However, most existing methods overlook multi-scale modeling, and the effectiveness of the predicted kernels is seldom explicitly evaluated, which can result in distorted details. To alleviate this issue, EfDeRain+ Guo et al. (2025b) introduces a weight-sharing multi-scale dilated filtering strategy. However, such predictive-filtering methods are predominantly designed for individual restoration tasks, where the required restoration behavior remains relatively consistent across inputs. Our work extends operator-based restoration to the all-in-one setting by studying input-conditioned operator adaptation and operator-guided progressive refinement across heterogeneous degradations.

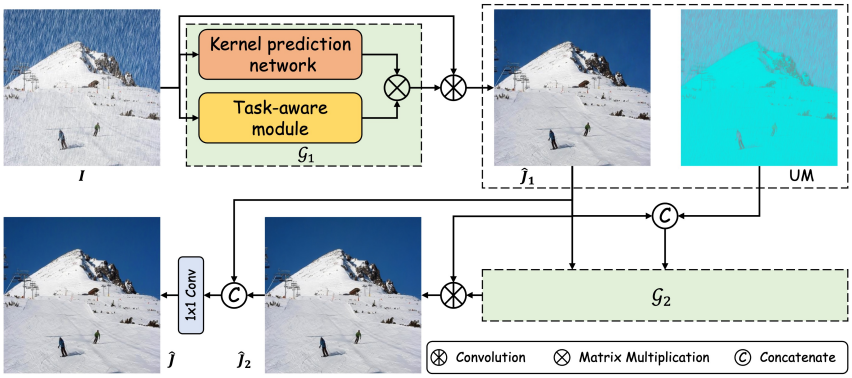

Figure 2: The overall architecture of the proposed OPIR is composed of two stages.

3 METHOD

3.1 PROBLEM DEFINITION

As shown in Eq. 1, the restoration objective is to estimate the latent clean image J from a degraded observation I under unknown and task-dependent degradation mechanisms. Although different degradation processes have different physical interpretations, they can be viewed from a common inverse-problem perspective. For notational convenience, we use the following generic abstraction $I \approx H_t J + b_t$, where H_t denotes a task-dependent degradation operator and b_t summarizes additional degradation components. Under this unified view, deraining, desnowing, dehazing, deblurring and denoising can all be interpreted as inverse problems with unknown degradation operators. Instead of explicitly estimating H_t and b_t, we directly learn a task-aware inverse mapping:

$$\hat{J} = \mathcal{G}(I; \theta_t) \tag{2}$$

where $\mathcal{G}(\cdot)$ denotes a shared restoration backbone modulated by task-aware parameters θ_t. From an operator-learning perspective, the model learns a family of inverse operators $\mathcal{G}_t \approx \mathcal{D}_t^{-1}$. Accordingly, image restoration can be formulated as $\min ||\mathcal{G}(I; \theta_t) - J||$. To further enhance restoration performance, as shown in Figure 2, we design a two-stage strategy. In the first stage, the model predicts not only the inverse mapping but also an uncertainty estimation map. In the second stage, this uncertainty map is leveraged to focus on severely degraded regions, enabling an all-in-one restoration framework across diverse degradation types. Accordingly, Eq. 2 can be rewritten as:

$$\hat{J} = \mathcal{G}_2\big(\mathcal{G}_1(I; \theta_{t1}); \theta_{t2}\big) \tag{3}$$

In this paper, our objective is to learn $\mathcal{G}_1(\cdot)$ and $\mathcal{G}_2(\cdot)$. The functions $\mathcal{G}_1(\cdot)$ and $\mathcal{G}_2(\cdot)$ share the same network architecture and are defined as:

$$\mathcal{G}_1(I; \theta_t) = \mathrm{TAIOM}(I) = \big(\mathrm{KPN}(I) \otimes \mathrm{TAM}(I)\big) \circledast I \tag{4}$$

where TAIOM denotes the task-aware inverse operator modeling. Here, KPN refers to the kernel prediction network, which is implemented using a U-Net Chen et al. (2022) architecture and adaptively predicts pixel-wise restoration kernels to process different image regions according to their respective degradation characteristics. TAM denotes the task-aware module.

3.2 TASK-AWARE INVERSE OPERATOR MODELING

Different restoration tasks require different restoration behaviors. Instead of learning an independent operator for each degradation type, we employ a shared kernel prediction network and adapt its predicted operator according to degradation information inferred from the current input.

Let $K = \mathrm{KPN}(I)$ denote the spatially varying base operator. The task-aware module generates an input-dependent modulation term, resulting in:

$$\tilde{K} = \mathcal{T}\big(K; \theta_t\big) = \mathrm{KPN}(I) \odot \mathrm{TAM}(I), \tag{5}$$

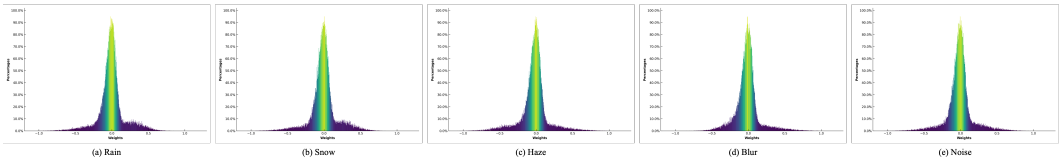

Figure 3: Distribution of predicted operator weights over the training samples. The empirical expectation is approximately zero, supporting the zero-centered assumption used in Eq. 6.

where $\theta(I)$ denotes task-aware information extracted from the observation rather than an externally provided task label. This formulation preserves a shared operator predictor while allowing its restoration behavior to vary with the degradation condition.

Restoration uncertainty estimation. Besides generating the restoration result, the predicted operator also provides information about the difficulty of reconstructing each spatial location. Specifically, as shown in Figure 3, we observe that the predicted kernel weights approximately follow an underlying distribution P with a zero-centered expectation μ.

For a target pixel p, the predicted local kernel assigns k^2 weights to pixels within its neighborhood $\mathcal{N}(p)$. Following the statistical characteristics of predictive kernels, the average response of these weights relative to the global expectation reflects the difficulty of reconstructing the target pixel. We therefore define the uncertainty response as:

$$UM(p) = -\frac{1}{k^2} \sum_{q \in \mathcal{N}(p)} \left(K_p(q) - \mu\right), \tag{6}$$

where μ denotes the expectation of the predicted kernel weights and is approximately zero in practice. The negative sign reverses the kernel response so that a larger $UM(p)$ consistently indicates higher reconstruction difficulty.

For relatively clean regions, the predicted kernel generally assigns a dominant positive weight to the center pixel while keeping neighboring weights close to zero, resulting in a lower uncertainty response. In more severely degraded regions, stronger compensatory, including negative, weights are required for neighboring pixels, leading to a higher uncertainty response. By computing $UM(p)$ over all pixels, we obtain an uncertainty map UM that highlights challenging regions and provides guidance for subsequent refinement. Based on this formulation, we adopt a two-stage inverse operator modeling strategy. In the first stage, the task-aware inverse operator predicts an initial restoration result $\hat{J}_1$ together with the corresponding uncertainty map $\hat{J}_1 = \tilde{K}_1 \circledast I$.

In the second stage, the same framework further refines the intermediate restoration conditioned on both the restored image and the estimated uncertainty map:

$$\begin{aligned} \hat{J}_2 &= \tilde{K}_2 \circledast \hat{J}_1, \\ \tilde{K}_2 &= \mathcal{T}(K_2; \theta_{t2}) = KPN(\hat{J}_1 \copyright UM) \otimes TAM(\hat{J}_1) \end{aligned} \tag{7}$$

where $\copyright$ denotes channel-wise concatenation. By explicitly incorporating task-aware adaptation and uncertainty-guided refinement, TAIOM can more effectively focus on ambiguous or severely degraded regions, enabling progressively improved restoration results while capturing degradation-specific characteristics and maintaining strong generalization across diverse restoration tasks.

Multi-scale inverse operator. For degradations extending over relatively large spatial regions, a small local neighborhood may not provide sufficient context. Directly predicting large dense kernels, however, considerably increases both kernel dimensionality and computational cost. Following Guo et al. (2025b), we adopt a weight-shared multi-scale filtering strategy that reuses a compact base kernel over different dilated neighborhoods. To recover a clean image $\hat{J}$ from an observed degraded image $I \in \mathbb{R}^{H \times W \times C}$, instead of predicting separate kernels $K^{(s)}$ for each scale s, we define a single shared base kernel $K \in \mathbb{R}^{H \times W \times 3 \times 3}$ and reuse it across multiple scales via dilation:

$$K_p^{(s)}(\Delta) = K_p(\delta), \quad \Delta = s \cdot \delta, \quad \delta \in \{-1, 0, 1\}^2 \tag{8}$$

where positions not covered by this mapping are set to zero. This design avoids explicit materialization of $K^{(s)}$ and keeps the predicted kernel dimensionality at $\mathcal{O}(9HW)$ while avoiding explicit construction of large scale-specific kernels. The scale responses are evaluated in parallel through batched tensor operations, resulting in low practical overhead. The output at scale s for pixel p is:

$$\hat{J}_p^{(s)} = \sum_{\delta \in \{-1,0,1\}^2} K_p(\delta) \circledast I(p + s\delta) \tag{9}$$

To further reduce computation, we note that kernel weights are shared across scales. Let $\Delta_s = \{s \cdot \delta \mid \delta \in \{-1, 0, 1\}^2\}$ denote sampled offsets, and $I_{\text{sampled}} \in \mathbb{R}^{H \times W \times 9 \times S}$ the tensor of sampled image values for S scales. Then the multi-scale convolution can be written as:

$$\hat{J}_p^{(s)} = \sum_{i=1}^{9} K_p(\delta_i) \circledast I_{\text{sampled}}(p, i, s) \tag{10}$$

Finally, we adopt a learnable fusion $\alpha_p^{(s)}$ across scales:

$$\hat{J}_p = \sum_{s=1}^{S} \alpha_p^{(s)} \otimes \hat{J}_p^{(s)}, \quad \sum_{s=1}^{S} \alpha_p^{(s)} = 1, \quad \alpha_p^{(s)} \geq 0 \tag{11}$$

In implementation, sampling and filtering at different scales are performed in parallel, avoiding explicit construction of large dense kernels and maintaining a low computational overhead.

3.3 Task-Aware Module

Different image restoration tasks correspond to distinct degradation processes, each characterized by different physical formation mechanisms and statistical properties. As a result, a single fixed inverse operator is often insufficient to handle multiple degradations effectively. To address this, we introduce a task-aware module (TAM), which adapts a shared base kernel K through lightweight, task-specific parameters θ_t, enabling flexible modeling of diverse degradation types within a unified framework. TAM further introduces spatially-varying modulation to account for the fact that degradation is often non-uniform across an image. For instance, rain streaks may appear in localized regions, and blur may differ across motion patterns. Let p index a pixel and $\mathcal{N}(p)$ denote its local neighborhood. The modulated kernel at pixel p is defined as:

$$\tilde{K}_p(q) = K_p(q) \otimes m_p(q; \theta_t), \quad q \in \mathcal{N}(p) \tag{12}$$

where $m_p(q; \theta_t)$ represents the task-aware modulation weight predicted by the TAM. This modulation function is implemented as a lightweight MLP, which takes task embeddings and optionally local features as input to produce adaptive spatial weights. The design enables fine-grained control over how the shared kernel is reshaped under different degradation contexts. Given the modulated kernel, the restored pixel is obtained through a local adaptive convolution:

$$\hat{J}_p = \sum_{q \in \mathcal{N}(p)} \tilde{K}_p(q) \circledast I(q) \tag{13}$$

This formulation allows the restoration operator to dynamically adjust both globally and locally, while maintaining a shared backbone representation. Compared with learning separate models or independent kernels for each task, the proposed TAM significantly improves parameter efficiency and computational scalability, while preserving strong adaptability across heterogeneous image restoration scenarios.

4 Experiments

In this section, we first present the comparison results, followed by the ablation studies. The experimental setup and more experiments are provided in Appendix A.2 and A.3.

Table 1: Quantitative results in the all-in-one (R+S+H) setting.

Type	Methods	Deraining		Desnowing		Dehazing		Average	
		PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑
Specific	EfDeRain+ Guo et al. (2025b)	31.18	0.912	29.96	0.938	33.25	0.971	31.46	0.940
	PGH^{2}Net Su et al. (2025)	29.89	0.907	32.18	0.943	35.55	0.983	32.21	0.944
	ALGNet Gao et al. (2024a)	29.32	0.905	31.56	0.944	33.11	0.977	31.33	0.942
Aligned	MHNet Gao et al. (2025b)	30.69	0.906	32.03	0.943	34.59	0.980	32.44	0.943
	PPTformer Wang et al. (2025)	30.88	0.910	32.80	0.945	34.66	0.980	32.78	0.945
	ACL Gu et al. (2025)	31.11	0.912	32.91	0.948	34.81	0.981	32.94	0.947
All-in-One	VLU-Net Zeng et al. (2025)	31.07	0.910	33.02	0.952	35.43	0.984	33.17	0.949
	Perceive-IR Zhang et al. (2025)	31.13	0.911	33.11	0.952	35.35	0.983	33.20	0.949
	Defusion Luo et al. (2025)	30.98	0.910	33.29	0.953	35.52	**0.985**	33.26	0.949
	OPIR(Ours)	**31.96**	**0.914**	**34.14**	**0.959**	**35.96**	0.984	**34.02**	**0.952**

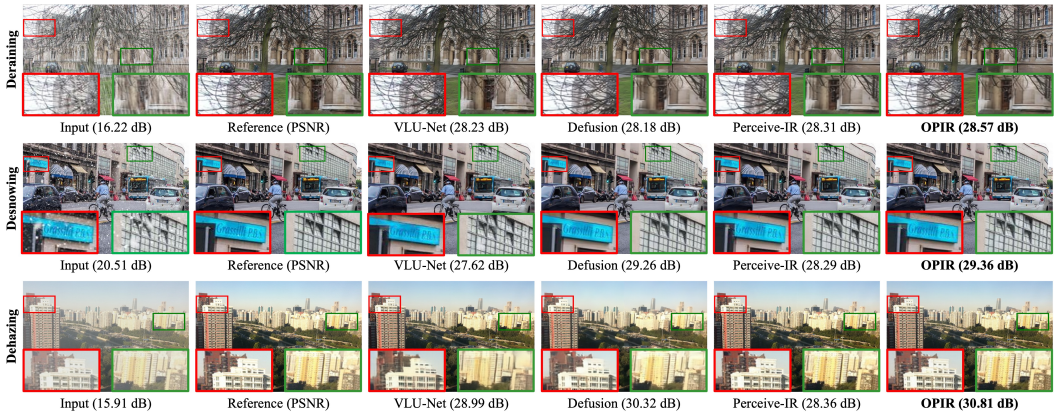

Figure 4: Qualitative results under the all-in-one (R+S+H) experimental setup.

4.1 All-in-one Setting Results

Table 1 compares task-specific, task-aligned, and all-in-one image restoration methods under the all-in-one (R+S+H) setting. The results demonstrate that OPIR achieves strong and balanced performance across different restoration tasks. Compared with task-specific methods, OPIR shows better overall generalization. Although methods such as PGH^{2}Net Su et al. (2025) and EfDeRain+ Guo et al. (2025b) perform well on their target degradations, their performance drops on other tasks. In contrast, OPIR improves the best task-specific method by +1.81 dB in average PSNR. Compared with task-aligned methods, OPIR also achieves consistently higher accuracy. In particular, it outperforms ACL Gu et al. (2025), the strongest task-aligned baseline, by +1.08 dB in average PSNR, indicating that task-aware operator modeling better captures differences among heterogeneous degradations. Among all-in-one methods, OPIR surpasses Defusion Luo et al. (2025) by +0.76 dB in average PSNR and achieves the best PSNR on all three tasks, with gains of +0.74 dB on deraining, +0.85 dB on desnowing, and +0.44 dB on dehazing. These results highlight the effectiveness of task-aware inverse operator modeling over existing prompt- or diffusion-based designs. As shown in Figure 4, OPIR also produces sharper restorations that are visually closer to the ground truth. When extended to the more challenging (R+S+H+B+N) setting, Table 2 shows that OPIR maintains its performance advantage.

Table 2: Quantitative results in the all-in-one (R+S+H+B+N) setting.

Methods	Deraining		Desnowing		Dehazing		Deblurring		Denoising		Average	
	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑
EfDeRain+ Guo et al. (2025b)	30.89	0.909	29.42	0.925	33.01	0.952	26.48	0.804	30.62	0.880	30.08	0.894
PGH^{2}Net Su et al. (2025)	29.52	0.896	31.88	0.932	34.34	0.976	25.96	0.789	30.97	0.882	30.53	0.895
ECFNet Gao et al. (2026)	30.44	0.890	32.54	0.932	33.99	0.978	27.58	0.836	31.10	0.882	31.13	0.904
ACL Gu et al. (2025)	30.62	0.893	32.52	0.926	33.98	0.978	28.59	0.847	31.21	0.882	31.38	0.905
Defusion Luo et al. (2025)	30.66	0.905	33.33	0.952	34.65	**0.981**	29.98	0.889	31.50	0.886	32.02	0.922
AllRestorer Mao et al. (2026)	30.71	0.903	33.17	0.951	34.57	0.978	29.63	0.882	31.39	0.885	31.89	0.920
BaryIR Tang et al. (2026)	30.82	0.907	33.05	0.948	34.60	0.979	29.52	0.881	31.32	0.884	31.86	0.920
OPIR(Ours)	**31.82**	**0.912**	**33.91**	**0.953**	**34.68**	0.980	**30.02**	**0.891**	31.51	**0.889**	**32.39**	**0.925**

Table 3: Real-world results in the all-in-one (R+S+H+B+N) setting.

Methods	RealRain-1k-L			RTTS			SIDD	
	PSNR ↑	SSIM ↑	LPIPS ↓	FADE ↓	BRISQUE ↓	NIMA ↑	PSNR ↑	SSIM ↑
EfDeRain+ Guo et al. (2025b)	24.05	0.762	0.398	1.393	28.865	4.457	24.07	0.475
PGH^{2}Net Su et al. (2025)	21.65	0.747	0.419	1.352	28.501	<u>4.625</u>	24.00	0.466
ACL Gu et al. (2025)	23.96	0.757	0.408	1.360	28.962	4.319	<u>24.66</u>	0.473
VLU-Net Zeng et al. (2025)	27.22	0.899	0.375	<u>1.242</u>	**23.490**	4.564	24.31	0.471
Defusion Luo et al. (2025)	<u>27.37</u>	**0.903**	<u>0.371</u>	1.277	<u>23.520</u>	4.615	24.52	<u>0.495</u>
OPIR(Ours)	**27.59**	<u>0.902</u>	**0.369**	**1.208**	23.523	**4.703**	**24.72**	**0.499**

Table 4: Ablation study on individual components of OPIR.

Method	Stage	UM	TAM	Multi-scale kernel	PSNR	△ PSNR
(a)	1				32.31	-
(b)	1		✔		33.26	+ 0.95 dB
(c)	1			✔	32.95	+ 0.64 dB
(d)	1		✔	✔	33.52	+ 1.21 dB
(e)	2		✔	✔	33.77	+ 1.46 dB
(f)	2	✔	✔	✔	34.02	+ 1.71 dB

4.2 Zero-shot generalization in real-world scenes

We evaluate methods trained under the (R+S+H+B+N) setting and directly test their zero-shot generalization to real-world degradations. As shown in Table 3, OPIR achieves the best overall deraining performance on RealRain-1k-L Li et al. (2022), with 27.59 dB PSNR and 0.369 LPIPS while maintaining competitive SSIM. On SIDD Abdelhamed et al. (2018), OPIR obtains the best denoising performance with 24.72 dB PSNR. On RTTS Li et al. (2018b), it achieves the lowest FADE score of 1.208 and the highest NIMA score of 4.703, together with competitive BRISQUE performance. These results demonstrate the strong zero-shot generalization of OPIR to real-world degradations.

4.3 Ablation Studies

4.3.1 Effect of each component

We perform ablation studies under the all-in-one setting and conduct a step-by-step analysis by progressively incorporating the proposed modules to validate the effectiveness of each component, as summarized in Table 4. Starting from the single-stage baseline without any proposed modules (a), the model achieves 32.31 dB PSNR. Adding the task-aware module (b) improves PSNR by 0.95 dB, showing the benefit of conditioning restoration on task-specific information. Introducing multi-scale kernel prediction (c) yields a 0.64 dB gain, confirming the effectiveness of multi-scale modeling. Combining the task-aware module and multi-scale kernel prediction in a single stage (d) further improves PSNR to 33.52 dB, a total gain of 1.21 dB over the baseline, indicating that the two components are complementary. Extending the model to a two-stage architecture (e) brings an additional 0.25 dB improvement, demonstrating the benefit of progressive refinement. Finally, incorporating the uncertainty map (f) achieves the best performance of 34.02 dB, corresponding to a total gain of 1.71 dB over the baseline. This result verifies that uncertainty-aware modeling better captures spatially varying degradation difficulty and improves restoration robustness.

4.3.2 Effect of different combinations of restoration tasks

To further investigate the influence of task composition, we evaluate restoration performance under different task combinations. As shown in Table 5, the model performs best when trained on a single degradation, while performance generally decreases as more tasks are introduced, reflecting the difficulty of jointly modeling heterogeneous degradation distributions. Task combinations also affect individual restoration performance. For example, (R+S+H+N) achieves better deraining and desnowing results than (R+S+B+N), suggesting that some degradations provide complementary supervision during multi-task learning. Haze and noise are both globally distributed degradations and may encourage more transferable representations, whereas motion blur primarily affects structural continuity and may introduce stronger conflicts with other restoration objectives.

Table 5: Effect of different combinations of restoration task.

Task					Deraining	Desnowing	Dehazing	Deblurring	Denoising
R	S	H	B	N					
✔					33.15	-	-	-	-
	✔				-	34.58	-	-	-
		✔			-	-	40.54	-	-
			✔		-	-	-	33.42	-
				✔	-	-	-	-	31.82
✔	✔				32.72	34.43	-	-	-
✔		✔			32.45	-	37.96	-	-
	✔	✔			-	34.50	38.32	-	-
			✔	✔	-	-	-	32.81	31.75
✔	✔	✔			32.31	34.22	36.17	-	-
✔	✔		✔		32.09	34.13	-	31.45	-
✔	✔			✔	32.16	34.19	-	-	31.68
✔	✔	✔	✔		31.93	33.97	35.02	30.49	-
✔	✔	✔		✔	31.95	33.99	35.13	-	31.57
✔	✔		✔	✔	31.86	33.93	-	30.52	31.54
✔	✔	✔	✔	✔	31.82	33.91	34.68	30.02	31.51

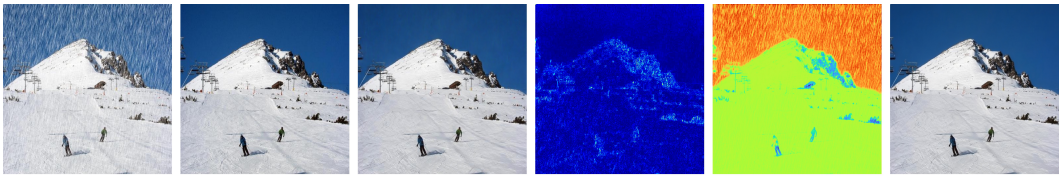

Figure 5: Effect of the uncertainty map.

4.3.3 Effect of the uncertainty map

To further evaluate the effectiveness of the uncertainty map (UM), we provide a visualization in Figure 5. In severely degraded regions, the kernel tends to assign stronger compensatory, including negative, weights to neighboring pixels, which reduces the average kernel response and consequently produces a higher uncertainty value after the sign reversal in Eq. 6. Regions with large reconstruction errors largely coincide with high-uncertainty regions, indicating that UM can effectively identify challenging degraded areas and reflect pixel-level reconstruction difficulty.

The distribution of $UM(p)$ is closely related to the estimated restoration kernel. In relatively clean regions, the predicted kernel typically assigns a dominant positive weight to the center pixel, while neighboring weights remain close to zero, corresponding to lower uncertainty. In severely degraded regions, the kernel tends to assign stronger compensatory weights to neighboring pixels, resulting in higher uncertainty responses. These observations show that UM can effectively locate difficult degradation regions and provide useful guidance for Stage 2 to further refine the restoration results.

5 Conclusion

We propose OPIR, an efficient all-in-one image restoration framework based on input-conditioned inverse operator learning. OPIR predicts spatially adaptive restoration operators from degraded observations, enabling the model to adapt to heterogeneous degradation conditions within a unified framework. OPIR further adopts a two-stage restoration strategy, where the first stage produces an initial restoration together with an operator uncertainty map, and the second stage refines challenging regions under difficulty guidance. A shared operator prediction network with task-aware modulation improves adaptation across different degradation types, while the efficient multi-scale operator implementation maintains a favorable computational cost. Extensive experiments demonstrate that OPIR achieves strong restoration performance.

AI USE STATEMENT

In this work, we used generative AI tools to assist with manuscript editing and scientific figure preparation. GPT-based tools were used to improve the language, grammar, and readability of the manuscript without changing the underlying scientific content. The underlying data and numerical results in the figure were not generated or modified by AI.

All AI-assisted outputs were manually reviewed and verified by the authors. The authors take full responsibility for the final manuscript, including all text, figures, experimental results, scientific claims, and conclusions.

REFERENCES

Abdelrahman Abdelhamed, Stephen Lin, and Michael S Brown. A high-quality denoising dataset for smartphone cameras. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1692–1700, 2018.

Eirikur Agustsson and Radu Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1122–1131, 2017. doi: 10.1109/CVPRW.2017.150.

Pablo Arbeláez, Michael Maire, Charless Fowlkes, and Jitendra Malik. Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(5):898–916, 2011. doi: 10.1109/TPAMI.2010.161.

S.P. Awate and R.T. Whitaker. Unsupervised, information-theoretic, adaptive image filtering for image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(3):364–376, 2006.

Jun Chen, Wei Yu, Xin Tian, Jun Huang, and Jiayi Ma. Mdbfusion++: A visible and infrared image fusion framework capable for motion deblurring. *IEEE Transactions on Image Processing*, 35: 3647–3661, 2026.

Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, and Jian Sun. Simple baselines for image restoration. *ECCV*, 2022.

Sixiang Chen, Tian Ye, Yun Liu, Taodong Liao, Jingxia Jiang, Erkang Chen, and Peng Chen. Msp-former: Multi-scale projection transformer for single image desnowing. In *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1–5, 2023. doi: 10.1109/ICASSP49357.2023.10095605.

Wei-Ting Chen, Hao-Yu Fang, Jian-Jiun Ding, Cheng-Che Tsai, and Sy-Yen Kuo. Jstasr: Joint size and transparency-aware snow removal algorithm based on modified partial convolution and veiling effect removal. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXI 16*, pp. 754–770. Springer, 2020.

Wei-Ting Chen, Hao-Yu Fang, Cheng-Lin Hsieh, Cheng-Che Tsai, I Chen, Jian-Jiun Ding, Sy-Yen Kuo, et al. All snow removed: Single image desnowing algorithm using hierarchical dual-tree complex wavelet representation and contradict channel loss. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 4196–4205, 2021.

Zixuan Chen, Zewei He, and Zhe-Ming Lu. Dea-net: Single image dehazing based on detail-enhanced convolution and content-guided attention. *IEEE Transactions on Image Processing*, 2024.

Lark Kwon Choi, Jaehee You, and Alan Conrad Bovik. Referenceless prediction of perceptual fog density and perceptual image defogging. *IEEE Transactions on Image Processing*, 24(11): 3888–3901, 2015.

Yuning Cui, Wenqi Ren, Xiaochun Cao, and Alois Knoll. Image restoration via frequency selection. *TPAMI*, pp. 1–16, 2023a.

Yuning Cui, Wenqi Ren, Xiaochun Cao, and Alois Knoll. Focal network for image restoration. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 13001–13011, 2023b.

Yuning Cui, Wenqi Ren, Sining Yang, Xiaochun Cao, and Alois Knoll. Irnext: Rethinking convolutional network design for image restoration. In *Proceedings of the 40th International Conference on Machine Learning*, 2023c.

Yuning Cui, Yi Tao, Zhenshan Bing, Wenqi Ren, Xinwei Gao, Xiaochun Cao, Kai Huang, and Alois Knoll. Selective frequency network for image restoration. In *ICLR*, 2023d.

Yuning Cui, Syed Waqas Zamir, Salman Khan, Alois Knoll, Mubarak Shah, and Fahad Shahbaz Khan. AdaIR: Adaptive all-in-one image restoration via frequency mining and modulation. In *The Thirteenth International Conference on Learning Representations*, 2025.

Yuning Cui, Wenqi Ren, Boxin Shi, Jianhou Gan, and Alois Knoll. Cdir: Lora-inspired attention for efficient composite degradation image restoration. *IEEE Transactions on Image Processing*, 35: 3998–4011, 2026. doi: 10.1109/TIP.2026.3682023.

Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16 (8):2080–2095, 2007.

Rich Franzen. Kodak lossless true color image suite. *source: http://r0k. us/graphics/kodak*, 4(2):9, 1999.

Lan Fu, Changqing Zhou, Qing Guo, Felix Juefei-Xu, Hongkai Yu, Wei Feng, Yang Liu, and Song Wang. Auto-exposure fusion for single-image shadow removal. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10566–10575, 2021.

Xueyang Fu, Jiabin Huang, Delu Zeng, Yue Huang, Xinghao Ding, and John Paisley. Removing rain from single images via a deep detail network. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1715–1723, 2017. doi: 10.1109/CVPR.2017.186.

Hu Gao, Bowen Ma, Ying Zhang, Jingfan Yang, Jing Yang, and Depeng Dang. Learning enriched features via selective state spaces model for efficient image deblurring. In *Proceedings of the 32nd ACM International Conference on Multimedia*, pp. 710–718, 2024a.

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, 2024b.

Hu Gao, Jing Yang, Ying Zhang, Jingfan Yang, Bowen Ma, and Depeng Dang. Learning optimal combination patterns for lightweight stereo image super-resolution. In *Proceedings of the 32nd ACM International Conference on Multimedia*, pp. 5566–5574, 2024c.

Hu Gao, Bowen Ma, Ying Zhang, Jingfan Yang, Jing Yang, and Depeng Dang. Frequency domain task-adaptive network for restoring images with combined degradations. *Pattern Recognition*, 158:111057, 2025a.

Hu Gao, Ying Zhang, Jing Yang, and Depeng Dang. Mixed hierarchy network for image restoration. *Pattern Recognition*, 161:111313, 2025b.

Hu Gao, Bowen Ma, Ying Zhang, Jingfan Yang, Jing Yang, Xingjian Wang, and Depeng Dang. Emphasizing crucial features for efficient image restoration. *Pattern Recognition*, pp. 113575, 2026.

Yubin Gu, Yuan Meng, Jiayi Ji, and Xiaoshuai Sun. Acl: Activating capability of linear attention for image restoration. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 17913–17923, 2025.

Hang Guo, Jinmin Li, Tao Dai, Zhihao Ouyang, Xudong Ren, and Shu-Tao Xia. Mambair: A simple baseline for image restoration with state-space model. *arXiv preprint arXiv:2402.15648*, 2024a.

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

Hang Guo, Yong Guo, Yaohua Zha, Yulun Zhang, Wenbo Li, Tao Dai, Shu-Tao Xia, and Yawei Li. Mambairv2: Attentive state space restoration. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 28124–28133, 2025a.

Qing Guo, Xiaoguang Li, Felix Juefei-Xu, Hongkai Yu, Yang Liu, and Song Wang. Jpgnet: Joint predictive filtering and generative network for image inpainting. In *Proceedings of the 29th ACM International conference on multimedia*, pp. 386–394, 2021.

Qing Guo, Hua Qi, Jingyang Sun, Felix Juefei-Xu, Lei Ma, Di Lin, Wei Feng, and Song Wang. Efficientderain+: Learning uncertainty-aware filtering via rainmix augmentation for high-efficiency deraining. *International Journal of Computer Vision*, 133(4):2111–2135, 2025b.

Qing Guo, Zhijie Wang, Lubo Wang, Haotian Dong, Felix Juefei-Xu, Di Lin, Lei Ma, Wei Feng, and Yang Liu. Carvenet: Carving point-block for complex 3d shape completion. *IEEE Transactions on Multimedia*, 27:1047–1058, 2025c.

Xin Guo, Xi Wang, Xueyang Fu, and Zheng-Jun Zha. Deep unfolding network for image desnowing with snow shape prior. *IEEE Transactions on Circuits and Systems for Video Technology*, 35(5): 4740–4752, 2025d.

Kaiming He, Jian Sun, and Xiaoou Tang. Single image haze removal using dark channel prior. *TPAMI*, 2011.

Jia-Bin Huang, Abhishek Singh, and Narendra Ahuja. Single image super-resolution from transformed self-exemplars. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5197–5206, 2015.

Xingyu Jiang, Ning Gao, Hongkun Dou, Xiuhui Zhang, Xiaoqing Zhong, Yue Deng, and Hongjue Li. Global modeling matters: A fast, lightweight, and effective baseline for efficient image restoration. *IEEE Transactions on Image Processing*, 35:2740–2754, 2026.

Yitong Jiang, Zhaoyang Zhang, Tianfan Xue, and Jinwei Gu. Autodir: Automatic all-in-one image restoration with latent diffusion. In *Computer Vision – ECCV 2024: 18th European Conference, Milan, Italy, September 29–October 4, 2024, Proceedings, Part XL*, pp. 340–359, 2024.

D. Kingma and J. Ba. Adam: A method for stochastic optimization. *Computer Science*, 2014.

D. Kundur and D. Hatzinakos. A novel blind deconvolution scheme for image restoration using recursive filtering. *IEEE Transactions on Signal Processing*, 46(2):375–390, 1998.

Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang. Benchmarking single-image dehazing and beyond. *TIP*, 28(1):492–505, 2018a.

Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang. Benchmarking single-image dehazing and beyond. *IEEE transactions on image processing*, 28 (1):492–505, 2018b.

Chongyi Li, Chunle Guo, Wenqi Ren, Runmin Cong, Junhui Hou, Sam Kwong, and Dacheng Tao. An underwater image enhancement benchmark dataset and beyond. *IEEE transactions on image processing*, 29:4376–4389, 2019.

Hongyu Li, Tianyi Xu, Zengyou Wang, Xiantong Zhen, Ran Gu, David Zhang, and Jun Xu. Diagnosing and improving vector-quantization-based blind image restoration. *IEEE Transactions on Image Processing*, 35:844–857, 2026. doi: 10.1109/TIP.2026.3651985.

Wei Li, Qiming Zhang, Jing Zhang, Zhen Huang, Xinmei Tian, and Dacheng Tao. Toward real-world single image deraining: A new benchmark and beyond. *arXiv preprint arXiv:2206.05514*, 2022.

Yaowei Li, Hang An, Tong Zhang, Xiaoxuan Chen, Bo Jiang, and Jinshan Pan. Omni-deblurring: Capturing omni-range context for image deblurring. *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 1–1, 2025.

Yu Li, Robby T. Tan, Xiaojie Guo, Jiangbo Lu, and Michael S. Brown. Rain streak removal using layer priors. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2736–2744, 2016. doi: 10.1109/CVPR.2016.299.

Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 136–144, 2017.

Hanzhou Liu, Chengkai Liu, Jiacong Xu, Peng Jiang, and Mi Lu. Xyscannet: An interpretable state space model for perceptual image deblurring. In *Proceedings of the Computer Vision and Pattern Recognition Conference (CVPR)*, pp. 779–789, 2025.

Yun-Fu Liu, Da-Wei Jaw, Shih-Chia Huang, and Jenq-Neng Hwang. Desnownet: Context-aware deep network for snow removal. *IEEE Transactions on Image Processing*, 27(6):3064–3073, 2018. doi: 10.1109/TIP.2018.2806202.

I. Loshchilov and F. Hutter. Sgdr: Stochastic gradient descent with warm restarts. 2016.

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 (CVPR)*, pp. 12764–12777, June 2025.

Kede Ma, Zhengfang Duanmu, Qingbo Wu, Zhou Wang, Hongwei Yong, Hongliang Li, and Lei Zhang. Waterloo exploration database: New challenges for image quality assessment models. *IEEE Transactions on Image Processing*, 26(2):1004–1016, 2016.

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

D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, volume 2, pp. 416–423 vol.2, 2001.

Ben Mildenhall, Jonathan T Barron, Jiawen Chen, Dillon Sharlet, Ren Ng, and Robert Carroll. Burst denoising with kernel prediction networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2502–2510, 2018.

Anish Mittal, Anush K Moorthy, and Alan C Bovik. Blind/referenceless image spatial quality evaluator. In *2011 conference record of the forty fifth asilomar conference on signals, systems and computers (ASILOMAR)*, pp. 723–727. IEEE, 2011.

Seungjun Nah, Tae Hyun Kim, and Kyoung Mu Lee. Deep multi-scale convolutional neural network for dynamic scene deblurring. *CVPR*, pp. 257–265, 2016.

Karen Panetta, Chen Gao, and Sos Agaian. Human-visual-system-inspired underwater image quality measures. *IEEE journal of oceanic engineering*, 41(3):541–551, 2015.

Xiting Peng, Yandi Zhang, Xiaoyu Zhang, Junyi Wang, Shi Bai, Yubo Cao, Hongye Chen, and Tianshu Li. Fcs-ednet: Exploring magnetic particle imaging deblurring with neural network. *IEEE Transactions on Image Processing*, 35:480–494, 2026. doi: 10.1109/TIP.2025.3648497.

Vaishnav Potlapalli, Syed Waqas Zamir, Salman Khan, and Fahad Shahbaz Khan. Promptir: Prompting for all-in-one blind image restoration. *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.

Jianxiang Rong, Hua Huang, and Jia Li. Imu-assisted accurate blur kernel re-estimation in non-uniform camera shake deblurring. *IEEE Transactions on Image Processing*, 33:3823–3838, 2024.

Ziyi Shen, Wenguan Wang, Xiankai Lu, Jianbing Shen, Haibin Ling, Tingfa Xu, and Ling Shao. Human-aware motion deblurring. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 5571–5580, 2019.

Xiongfei Su, Siyuan Li, Yuning Cui, Miao Cao, Yulun Zhang, Zheng Chen, Zongliang Wu, Zedong Wang, Yuanlong Zhang, and Xin Yuan. Prior-guided hierarchical harmonization network for efficient image dehazing. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 7042–7050, 2025.

Hossein Talebi and Peyman Milanfar. Nima: Neural image assessment. *IEEE transactions on image processing*, 27(8):3998–4011, 2018.

Xiaole Tang, Xiaoyi He, Jiayi Xu, Xiang Gu, and Jian Sun. Learning continuous wasserstein barycenter space for generalized all-in-one image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–16, 2026. doi: 10.1109/TPAMI.2026.3669121.

Cong Wang, Jinshan Pan, Liyan Wang, and Wei Wang. Intra and inter parser-prompted transformers for effective image restoration. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 7609–7618, 2025.

Yuanbo Wen, Tao Gao, and Ting Chen. Unpaired photo-realistic image deraining with energy-informed diffusion model. In *Proceedings of the 32nd ACM International Conference on Multimedia*, pp. 360–369, 2024.

Miao Yang and Arcot Sowmya. An underwater color image quality evaluation metric. *IEEE transactions on image processing*, 24(12):6062–6071, 2015.

Wenhan Yang, Robby T. Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, and Shuicheng Yan. Deep joint rain detection and removal from a single image. *CVPR*, pp. 1685–1694, 2016.

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.

Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Multi-stage progressive image restoration. In *CVPR*, 2021.

Haijin Zeng, Jiezhang Cao, Kai Zhang, Yongyong Chen, Hiep Luong, and Wilfried Philips. Unmixing diffusion for self-supervised hyperspectral image denoising. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 27820–27830, 2024.

Haijin Zeng, Xiangming Wang, Yongyong Chen, Jingyong Su, and Jie Liu. Vision-language gradient descent-driven all-in-one deep unfolding networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7524–7533, June 2025.

He Zhang and Vishal M. Patel. Density-aware single image de-raining using a multi-stream dense network. *CVPR*, pp. 695–704, 2018.

He Zhang, Vishwanath A. Sindagi, and Vishal M. Patel. Image de-raining using a conditional generative adversarial network. *TCSVT*, 30:3943–3956, 2017.

Ming Zhang and Bahadir K. Gunturk. Multiresolution bilateral filtering for image denoising. *IEEE Transactions on Image Processing*, 17(12):2324–2333, 2008.

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 conference on computer vision and pattern recognition*, pp. 586–595, 2018.

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*, pp. 1–1, 2025.

A Appendix

A.1 Loss Function

Consistent with prior work, we optimize OPIR jointly in the spatial and frequency domains. The overall objective function is formulated as:

$$
\begin{aligned}
L &= \sum_{i=1}^{3} L_c(\hat{J}_i, \overline{I}) + \delta L_e(\hat{J}_i, \overline{I}) + \lambda L_f(\hat{J}_i, \overline{I}) \\
L_c &= \sqrt{||\hat{J}_i - \overline{I}||^2 + \epsilon^2} \\
L_e &= \sqrt{||\triangle \hat{J}_i - \triangle \overline{I}||^2 + \epsilon^2} \\
L_f &= ||\mathcal{F}(\hat{J}_i) - \mathcal{F}(\overline{I})||_1
\end{aligned}
\tag{14}
$$

where i indexes the input–output image pairs, and $\overline{I}$ denotes the corresponding ground-truth images. $\mathcal{L}_c$ is the Charbonnier loss, with the constant ϵ empirically set to 0.001 in all experiments. $\mathcal{L}_e$ is an edge-aware loss, where $\triangle$ denotes the Laplacian operator, encouraging accurate reconstruction of high-frequency structures. $\mathcal{L}_f$ represents the frequency-domain loss, in which $\mathcal{F}(\cdot)$ denotes the fast Fourier transform. The weighting factors λ and δ are set to 0.1 and 0.05, respectively, following Zamir et al. (2021); Cui et al. (2023a).

A.2 Experimental Setup

We conduct experiments under both the all-in-one and task-aligned settings.

A.2.1 Datasets

For the task-aligned setting, **i)** Image deraining: we train the deraining model using 13,712 paired clean–rain images collected from multiple datasets Yang et al. (2016); Zhang et al. (2017); Fu et al. (2017); Li et al. (2016). The trained OPIR is evaluated on several benchmark test sets, including Rain100H Yang et al. (2016), Rain100L Yang et al. (2016), Test100 Zhang et al. (2017), and Test1200 Zhang & Patel (2018). **ii)** Image desnowing: we adopt the Snow100K Liu et al. (2018), SRRS Chen et al. (2020), and CSD Chen et al. (2021) datasets. To ensure consistency with the training strategy of previous work Cui et al. (2023a), we randomly sample 2,500 image pairs for training and 2,000 images for evaluation. **iii)** Image dehazing: we assess the performance of our OPIR on daytime synthetic datasets from RESIDE Li et al. (2018a), which include the indoor training set (ITS), the outdoor training set (OTS), and the synthetic objective testing set (SOTS). The model is trained separately on ITS and OTS, and then evaluated on the corresponding SOTS-Indoor and SOTS-Outdoor test sets, each containing 500 paired images. **iv)** Image deblurring: we train OPIR on the GoPro dataset Nah et al. (2016), which consists of 2,103 training image pairs and 1,111 evaluation pairs. **v)** Image denoising: we train a single OPIR model capable of handling various noise levels on a composite dataset. This dataset comprises 800 images from DIV2K Agustsson & Timofte (2017), 2,650 images from Flickr2K Lim et al. (2017), 400 images from BSD500 Arbeláez et al. (2011), and 4,744 images from WED Ma et al. (2016). The noisy images are generated by adding additive white Gaussian noise with a random noise level σ chosen from the set [15, 25, 50] to the clean images. Testing is conducted on the CBSD68 Martin et al. (2001), Urban100 Huang et al. (2015), and Kodak24 Franzen (1999) benchmark datasets.

For the all-in-one setting, we summarize two common all-in-one training sets, namely "Rain+Snow+Haze(R+S+H)" and "Rain+Snow+Haze+Blur+Noise(R+S+H+B+N)" by combining the above image restoration datasets. During evaluation, we use Test100 Zhang et al. (2017) for deraining, Snow100K Liu et al. (2018) for desnowing, SOTS-Outdoor Li et al. (2018a) for dehazing, GoPro Nah et al. (2016) for deblurring, and CBSD68 Martin et al. (2001) for denoising.

A.2.2 Evaluation Metrics

We evaluate performance using both reference-based and no-reference metrics. The reference-based metrics include Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and

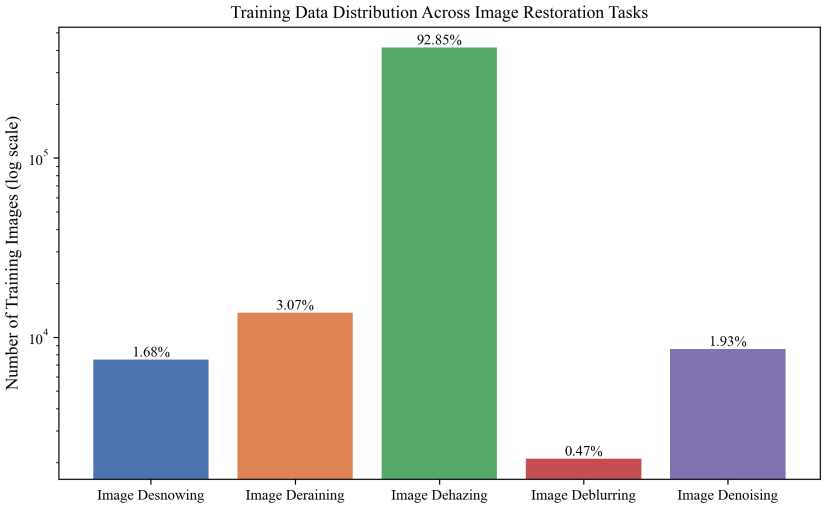

Figure 6: Sample sizes of datasets of different degradation types.

Learned Perceptual Image Patch Similarity (LPIPS) Zhang et al. (2018). The no-reference metrics include the Underwater Colour Image Quality Evaluation Metric (UCIQE) Yang & Sowmya (2015), Underwater Image Quality Measure (UIQM) Panetta et al. (2015), Fog Aware Density Evaluator (FADE) Choi et al. (2015), Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) Mittal et al. (2011), and Neural Image Assessment (NIMA) Talebi & Milanfar (2018). Among them, UCIQE and UIQM are specifically designed for underwater image restoration evaluation, while FADE, BRISQUE, and NIMA are commonly used to assess dehazing performance in real-world scenarios. For PSNR, SSIM, UCIQE, UIQM, and NIMA, higher values indicate better performance, whereas lower values are preferred for LPIPS, FADE, and BRISQUE. In the tables, the best and second-best results are highlighted in bold and underlined, respectively.

A.2.3 Training details

All models are optimized using the Adam optimizer Kingma & Ba (2014) with $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The learning rate is initialized at 2×10^{-4} and progressively reduced to 1×10^{-7} according to a cosine annealing strategy Loshchilov & Hutter (2016). During training, we randomly crop 256×256 patches with a batch size of 32 and train the models for 4×10^5 iterations. Horizontal and vertical flipping are applied for data augmentation. To ensure fair comparisons, all deep learning-based baselines are fine-tuned or retrained following the hyperparameter configurations reported in their original works.

A.2.4 Training Data Distribution

As illustrated in Figure 6, the training data distribution across different image restoration tasks is highly imbalanced. Image dehazing dominates the dataset, accounting for approximately 92.85% of the total training samples, while image deraining, image desnowing, image deblurring and image denoising contribute only 3.07%, 1.68%, 0.47% and 1.93%, respectively. This severe imbalance poses a significant challenge for learning an all-in-one restoration model, as the model tends to be biased toward tasks with a larger proportion of data, potentially leading to suboptimal performance on underrepresented tasks.

To mitigate this issue, we propose a task-aware data augmentation strategy (TDAS) that aims to achieve a more balanced effective data distribution across all tasks. Specifically, we adopt a multi-level feature augmentation approach, which operates not only in the input image space but also at intermediate feature representations extracted by the backbone network. At the image level, standard augmentation techniques such as random cropping, horizontal and vertical flipping, rotation, and color jittering are applied to increase the diversity of underrepresented task samples. At the feature level, we introduce perturbations and transformations on intermediate feature maps. These

Table 6: Image deraining results in the task-aligned setting.

Methods	Test100		Test1200		Rain100H		Rain100L		Average	
	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑
FSNet Cui et al. (2023a)	31.05	0.919	33.08	0.916	31.77	0.906	38.00	0.972	33.48	0.928
MHNet Gao et al. (2025b)	31.25	0.901	33.45	0.925	31.08	0.899	40.04	0.985	33.96	0.928
PPTformer Wang et al. (2025)	31.48	0.922	33.39	0.911	31.77	0.907	39.33	0.983	33.99	0.931
ACL Gu et al. (2025)	31.51	0.914	33.27	0.928	32.22	0.920	39.18	0.983	34.05	0.936
EfDeRain+ Guo et al. (2025b)	31.10	0.911	33.12	0.925	**34.57**	**0.957**	39.03	0.972	34.46	**0.941**
OPIR(Ours)	**33.15**	**0.923**	**35.19**	**0.941**	32.38	0.913	**41.01**	**0.986**	**35.43**	**0.941**

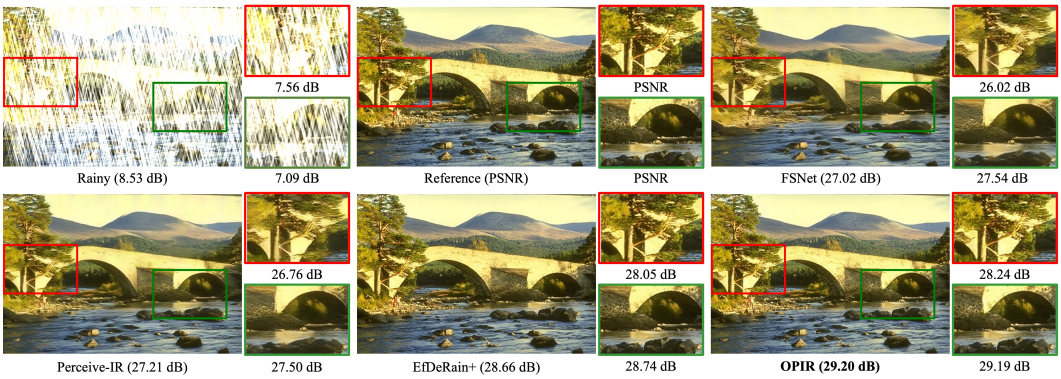

Figure 7: Qualitative results under the task-aligned experimental setup. Our OPIR is able to reconstruct finer and sharper details.

augmentations encourage the model to learn more robust and task-invariant representations, thereby improving the generalization on tasks with fewer samples. Furthermore, to prevent overfitting on the minority tasks while maintaining the performance on the dominant task, we employ a task-balanced sampling scheme, which adjusts the probability of sampling images from each task during training based on the inverse of their original proportions. Combined with the multi-level feature augmentation, this strategy ensures that each task contributes more equitably to the training process, effectively alleviating the bias caused by the original dataset imbalance.

A.3 More Experiments

A.3.1 Task-aligned Setting Results

To demonstrate that the proposed OPIR is effective not only for all-in-one restoration but also for task-specific scenarios, we conduct experiments on five representative image restoration tasks: image deraining, image desnowing, image dehazing, image deblurring and image denoising.

Image Deraining Following the protocol in prior work Gao et al. (2025b), we evaluate PSNR and SSIM metrics on the Y channel of the YCbCr color space for the image deraining task. Table 6 reports the quantitative results of image deraining under the task-aligned setting. Our proposed OPIR consistently outperforms all compared methods across most datasets. Specifically, on Test100 Zhang et al. (2017) and Test1200 Zhang & Patel (2018), OPIR achieves 33.15 dB and 35.19 dB in PSNR, respectively, surpassing previous state-of-the-art methods by a clear margin. On Rain100L Yang et al. (2016), OPIR reaches 41.01 dB PSNR, demonstrating superior capability in handling light rain conditions. Although EfDeRain+ Guo et al. (2025b) attains the highest scores on Rain100H Yang et al. (2016), our method still maintains competitive performance while achieving the best average PSNR (35.43 dB) and SSIM (0.941) across all datasets. These results indicate that OPIR not only effectively restores diverse rain streak patterns but also provides stable and robust performance in various deraining scenarios. Figure 7 illustrates that the images restored by our OPIR effectively reduce color distortion compared to other state-of-the-art methods. Moreover, our approach is able to reconstruct finer and sharper details in the degraded regions.

Table 7: Image desnowing results in the task-aligned setting.

Methods	CSD		SRRS		Snow100K	
	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑
NAFNet Chen et al. (2022)	33.13	0.96	29.72	0.94	32.41	0.95
FocalNet Cui et al. (2023b)	37.18	**0.99**	31.34	**0.98**	33.53	0.95
MSP-Former Chen et al. (2023)	33.75	0.96	30.76	0.95	33.43	**0.96**
IRNeXt Cui et al. (2023c)	37.29	**0.99**	31.91	**0.98**	33.61	0.95
PEUNet Guo et al. (2025d)	37.28	0.97	31.89	**0.98**	34.11	**0.96**
PW-FNet Jiang et al. (2026)	-	-	32.74	**0.98**	34.50	0.95
OPIR(Ours)	**37.74**	0.98	**33.10**	**0.98**	**34.58**	**0.96**

Table 8: Image dehazing results in the task-aligned setting.

Methods	SOTS-Indoor		SOTS-Outdoor	
	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑
SFNet Cui et al. (2023d)	41.24	0.996	40.05	**0.996**
IRXNext Cui et al. (2023c)	41.21	0.996	39.18	**0.996**
DEA-Net-CR Chen et al. (2024)	41.31	0.995	36.59	0.990
Defusion Luo et al. (2025)	41.65	0.995	37.41	0.993
PGH^{2}Net Su et al. (2025)	41.70	0.996	37.52	0.989
OPIR(Ours)	**42.21**	**0.997**	**40.54**	0.990

Image Desnowing Table 7 presents quantitative comparisons of image desnowing performance on three benchmark datasets: CSD, SRRS, and Snow100K. Our proposed OPIR consistently achieves the highest PSNR across all datasets, with 37.74 dB on CSD, 33.10 dB on SRRS, and 34.58 dB on Snow100K. In particular, OPIR achieves a 0.45 dB improvement on the CSD dataset Chen et al. (2021), an impressive 2.07 dB gain on the SRRS dataset Chen et al. (2020) and a 0.97 dB improvement on the Snow100K dataset Liu et al. (2018) when compared with IRNeXt Cui et al. (2023c). These results highlight the model's superior capability in handling diverse snow patterns and background complexities. Furthermore, relative to MSP-Former Chen et al. (2023), a task-specific architecture tailored for image desnowing, OPIR exhibits substantial quantitative gains of 1.15 dB, 2.34 dB, and 3.99 dB on the Snow100K, SRRS, and CSD datasets, respectively. This demonstrates that our adaptive multi-degradation design not only generalizes effectively to snow removal tasks but also achieves state-of-the-art restoration quality across different data distributions. These results demonstrate that OPIR not only excels in desnowing across multiple datasets but also provides robust and reliable performance.

Image Dehazing Table 8 reports quantitative comparisons on the SOTS Li et al. (2018a) benchmark for image dehazing. OPIR consistently outperforms all competing methods on both the SOTS-Indoor and SOTS-Outdoor subsets. On SOTS-Indoor, OPIR achieves a PSNR of 42.21 dB, surpassing the previous best method, PGH^{2}Net Su et al. (2025), by 0.51 dB. On the more challenging SOTS-Outdoor dataset, OPIR attains 40.54 dB PSNR, exceeding the strongest baseline SFNet Cui et al. (2023d) by 0.49 dB, while maintaining comparable SSIM performance. These consistent improvements demonstrate that OPIR more effectively removes haze while preserving fine structures

Table 9: Image deblurring results in the task-aligned setting.

Methods	GoPro		HIDE	
	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑
MPRNet Zamir et al. (2021)	32.66	0.959	30.96	0.939
FSNet Cui et al. (2023a)	33.29	0.963	31.05	0.941
MambaIR Guo et al. (2024a)	33.21	0.962	31.01	0.939
ACL Gu et al. (2025)	33.25	**0.964**	-	-
Omni-Deblurring Li et al. (2025)	33.29	0.963	31.65	**0.947**
OPIR(Ours)	**33.42**	**0.964**	**31.69**	**0.947**

Table 10: Image denoising results in the task-aligned setting.

Methods	CBSD68			Kodak24			Urban100		
	15	25	50	15	25	50	15	25	50
SFNet Cui et al. (2023d)	34.09	31.49	28.02	34.93	32.42	29.25	34.19	32.01	29.03
FSNet Cui et al. (2023a)	34.11	31.51	28.01	34.95	32.42	29.27	34.15	32.04	29.15
Perceive-IR Zhang et al. (2025)	34.38	31.74	28.53	34.84	32.50	29.16	34.86	32.55	29.42
VLU-Net Zeng et al. (2025)	34.35	31.72	28.46	-	-	-	**34.92**	**32.71**	**29.61**
OPIR(Ours)	**34.51**	**31.82**	**28.54**	**34.99**	**32.96**	**29.67**	34.89	32.52	29.42

Table 11: Results on the training-unseen degradation type in the all-in-one (R+S+H+B+N) setting .

Methods	UIEB			C60	
	PSNR ↑	SSIM ↑	LPIPS ↓	UCIQE↑	UIQW↑
EfDeRain+ Guo et al. (2025b)	17.66	0.741	0.377	0.480	1.866
ECFNet Gao et al. (2026)	20.49	0.856	0.223	0.539	2.517
ACL Gu et al. (2025)	20.94	0.867	0.201	0.552	2.485
VLU-Net Zeng et al. (2025)	21.58	0.889	0.172	0.556	2.493
Perceive-IR Zhang et al. (2025)	21.77	0.891	0.169	**0.557**	2.555
OPIR(Ours)	**21.98**	**0.895**	**0.162**	**0.557**	**2.561**

and visual fidelity, highlighting its robustness and strong generalization capability across both indoor and outdoor dehazing scenarios.

Image Deblurring Tables 9 present the performance of various image deblurring methods on the synthetic GoPro Nah et al. (2016) and HIDE Shen et al. (2019) datasets. Compared to the previous state-of-the-art Omni-Deblurring Li et al. (2025), our OPIR achieves a 0.13 dB gain on the GoPro dataset. Although trained exclusively on the GoPro dataset, our model achieves a 0.64 dB PSNR improvement over FSNet Cui et al. (2023a) on the HIDE dataset, demonstrating strong generalization capability.

Image Denoising Table 10 presents the performance of OPIR model under different noise levels (15, 25, and 50). Our approach consistently delivers strong results across multiple datasets and noise intensities. In particular, at the challenging noise level of 50 on the Kodak24 dataset Franzen (1999), OPIR surpasses FSNet Cui et al. (2023a) by 0.40 dB.

A.3.2 Generalization to unknown degradation type

We utilize the all-in-one methods trained under the "R+S+H+B+N" setting and directly evaluate them on unseen underwater image restoration tasks Li et al. (2019). As shown in Table 11, our OPIR achieves the best overall performance. These results demonstrate the superior generalization capability of OPIR, particularly in handling previously unseen degradation scenarios.

A.3.3 Generalization to unknown degradation severity

As mentioned in the dataset section, the model is trained using noise levels within the range of [15, 25, 50]. To further evaluate the robustness and generalization ability of the proposed method under unseen degradation severities, we additionally conduct experiments on substantially higher noise

Table 12: Results on the training-unseen degradation severity in the all-in-one (R+S+H+B+N) setting .

Methods	CBSD68		Urban100	
	60	100	60	100
VLU-Net Zeng et al. (2025)	27.11	20.59	27.64	21.53
Perceive-IR Zhang et al. (2025)	27.13	20.65	27.65	21.55
Defusion Luo et al. (2025)	27.11	**20.72**	**27.67**	21.49
OPIR(Ours)	**27.25**	**20.72**	27.66	**21.62**

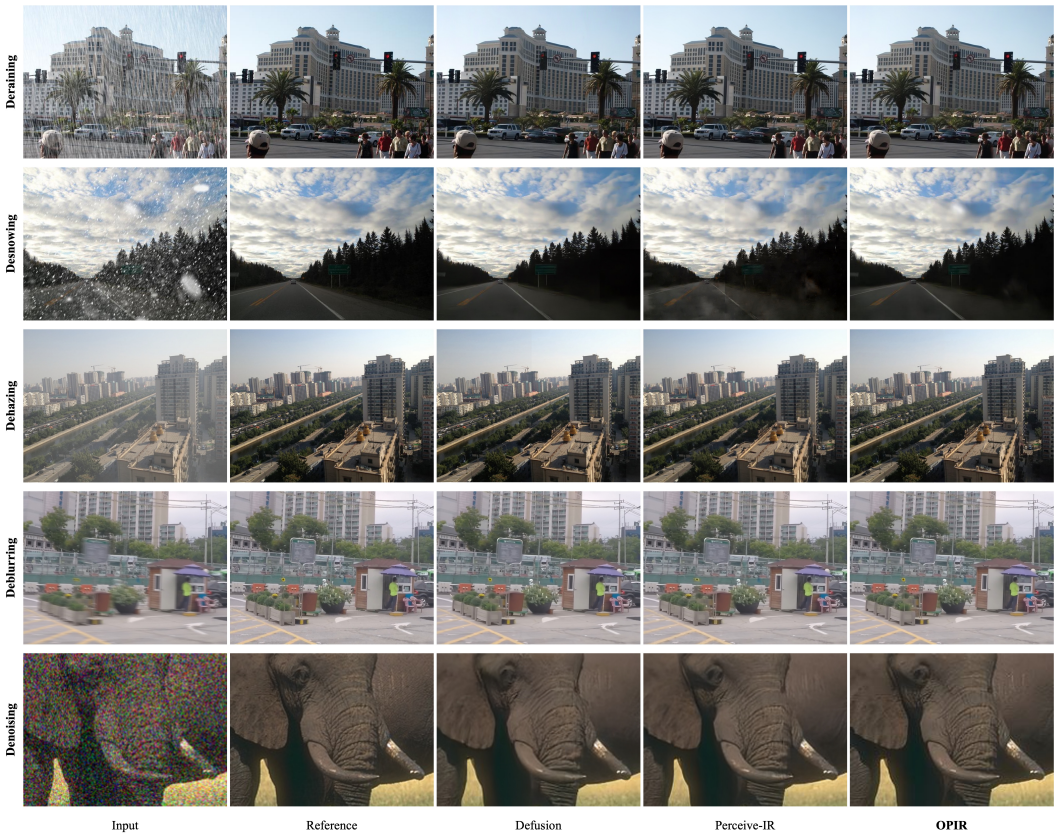

Figure 8: Qualitative results under the all-in-one experimental setup (R+S+H+B+N).

Table 13: The impact of task-aware data augmentation strategy.

Methods	Deraining PSNR ↑	Desnowing PSNR ↑	Dehazing PSNR ↑	Average PSNR ↑	Average △ PSNR
w/ TDAS	31.96	34.14	35.96	34.02	-
w/o TDAS	30.02	33.05	37.58	33.55	- 0.47 dB

levels of 60 and 100, which are not included during training. The corresponding results are reported in Table 12. It can be observed that OPIR consistently achieves superior or highly competitive performance across different datasets and unseen noise intensities. These results demonstrate that OPIR can effectively generalize to degradation severities beyond the training distribution, exhibiting strong robustness under unknown and highly challenging noise conditions.

A.3.4 Additional Visual Results

Figure 8 presents qualitative comparisons of our method under the all-in-one experimental setting (R+S+H+B+N), where a single unified model is used to handle multiple degradation types. Figure 9 shows the corresponding visual results under the task-aligned setting, in which the model is explicitly adapted to each specific degradation. As can be seen, our approach effectively removes diverse degradations across both settings, while consistently reconstructing images with sharp structural details, natural textures, and visually pleasing overall appearance. These results demonstrate the robustness and strong generalization capability of the proposed method under different restoration paradigms.



Figure 9: Qualitative results under the task-aligned experimental setup.

Table 14: Ablation study on EfDeRain+ Guo et al. (2025b).

Method	PSNR	△ PSNR
EfDeRain+ Guo et al. (2025b)	31.46	-
add Task-aware module	33.11	+ 1.65 dB
replace with Multi-scale kernel	31.75	+ 0.29 dB
replace with Multi-scale kernel + Task-aware module	33.69	+ 2.23 dB

A.3.5 Effect of the data augmentation strategy

Table 13 presents the effect of the proposed TDAS on the overall restoration performance. As shown, incorporating TDAS improves the PSNR from 33.55 dB to 34.02 dB, yielding a gain of 0.47 dB. This result demonstrates that the task-aware augmentation effectively alleviates the imbalance among different restoration tasks. In contrast, without TDAS, the model tends to be biased toward the dominant task, leading to inferior performance.

A.4 Discussion

EfDeRain+ Guo et al. (2025b) tackles the limitations of naive cascaded predictive filtering by introducing an uncertainty map derived from the statistics of predicted kernels, which enables a second-stage refinement that focuses on regions that are difficult to reconstruct. While this strategy is effective for single-task deraining, it remains inherently task-specific. Both predictive networks are trained exclusively for rain removal, and the uncertainty formulation is based on assumptions about kernel weight distributions that are closely coupled with rain streak characteristics. Consequently, extending EfDeRain+ Guo et al. (2025b) to other restoration tasks requires retraining separate models, which restricts its general applicability.

Both EfDeRain+ and OPIR exploit statistics of predicted kernels to estimate pixel-wise reconstruction difficulty. EfDeRain+ introduces this mechanism for single-task deraining, where the kernel statistics are analyzed under rain-specific predictive filtering. In OPIR, we retain this efficient kernel-statistics-based uncertainty estimation but apply it within a shared task-aware inverse operator framework for heterogeneous restoration tasks. Thus, our main distinction lies not in introducing a new uncertainty statistic itself, but in integrating uncertainty-guided refinement with degradation-adaptive operator modeling in the all-in-one setting.

A further distinction lies in uncertainty modeling. EfDeRain+ Guo et al. (2025b) estimates uncertainty by averaging predicted kernel weights, motivated by the observation that rain-corrupted pixels often exhibit characteristic positive–negative weight patterns. Although effective for deraining, this formulation implicitly depends on task-specific kernel distributions. In contrast, OPIR derives uncertainty from the statistical response of the predicted inverse operator, reflecting the strength of local dependencies required to reconstruct each pixel. This operator-centric uncertainty formulation is task-agnostic and characterizes reconstruction difficulty from a physical perspective, making it applicable to a broader range of degradations beyond rain.

The main distinction between the two methods lies in the scope and conditioning of predictive filtering. EfDeRain+ uses uncertainty-guided cascaded filtering for single-task deraining, whereas OPIR integrates the same efficient refinement principle into a task-aware inverse operator framework that jointly handles heterogeneous degradation types.

Finally, although both approaches employ weight-sharing multi-scale dilated filtering for computational efficiency, EfDeRain+ Guo et al. (2025b) applies this strategy within a single-task deraining pipeline. OPIR integrates multi-scale filtering into a task-aware inverse operator framework, enabling adaptive receptive field selection at the pixel level while preserving low computational complexity. Moreover, OPIR generates offsets for all scales within a single loop, performs multi-scale convolution in parallel, and further reduces redundant computation through learnable fusion weights.

To further demonstrate the advantages of our approach, we adopt EfDeRain+ Guo et al. (2025b) as the baseline and progressively replace or incorporate the modules proposed in this paper. This controlled modification strategy allows us to directly assess the effectiveness of each component.

Table 15: The evaluation of model computational complexity.

Method	Time(s)	Flops(G)	PSNR
VLU-Net Zeng et al. (2025)	0.743	143	33.17
PromptIR Potlapalli et al. (2023)	1.012	134	33.04
Perceive-IR Zhang et al. (2025)	0.682	144	33.20
OPIR(Ours)	**0.174**	**47**	**34.02**

Table 14 reports an ablation study based on EfDeRain+ Guo et al. (2025b), analyzing the contributions of task-aware inverse operator modeling and multi-scale kernel design. Introducing the task-aware module alone leads to a substantial performance gain of +1.65 dB, which clearly surpasses the improvement obtained by replacing the single-scale kernel with a multi-scale kernel (+0.29 dB). This observation suggests that explicitly modulating the predicted inverse operator according to task characteristics plays a more crucial role than simply enlarging the receptive field. When both components are jointly applied, the proposed method achieves an overall improvement of +2.23 dB compared to the baseline.

A.5 Resource Efficient

We further analyze the computational efficiency of OPIR by comparing its runtime and FLOPs with recent state-of-the-art methods. As shown in Table 15, OPIR not only achieves leading restoration performance but also maintains a highly efficient computational profile, demonstrating its suitability for practical and real-time applications. In particular, it surpasses the previous best-performing method, Perceive-IR Zhang et al. (2025), by 0.82 dB in restoration quality, while requiring only 25.5% of its inference time. This significant reduction in computational cost, together with improved performance, highlights the strong efficiency advantage of OPIR and its potential for deployment in latency-sensitive scenarios.

A.6 Limitations and Future work

Although OPIR achieves strong performance across various all-in-one and task-aligned image restoration benchmarks, while maintaining a favorable balance between efficiency and generalization, several limitations still remain. First, the current framework is primarily trained on predefined combinations of degradation types, which may limit its robustness when encountering more complex and unseen real-world mixed degradations. Second, although the introduced uncertainty modeling improves the second-stage refinement, it is largely based on explicit statistical discrepancies, and may not fully capture higher-level semantic degradation awareness. In addition, while the proposed operator computation is optimized for efficiency, its performance on extremely high-resolution images or resource-constrained devices still leaves room for further improvement in practical deployment scenarios.

In future work, we aim to explore more general and flexible degradation modeling strategies, such as integrating data-driven learning with physical priors to enhance robustness under unknown degradations. We also plan to investigate more powerful semantic-level uncertainty modeling mechanisms, enabling joint pixel-wise and semantic guidance for restoration. Furthermore, developing more lightweight and hardware-friendly operator implementations would be an important direction to improve real-time performance on edge devices. Finally, extending the current framework to broader low-level vision tasks, such as video restoration and multimodal degradation scenarios, could further enhance its generality and practical applicability.