

Diffusion Once and Done: Degradation-Aware LoRA for Efficient All-in-One Image Restoration

Ni Tang^{*1}, Xiaotong Luo^{*2}, Zihan Cheng¹, Liangtai Zhou¹, Dongxiao Zhang³, Yanyun Qu^{1,4†}

¹School of Informatics, Xiamen University

²The Hong Kong Polytechnic University

³Jimei University

⁴Key Laboratory of Multimedia Trusted Perception and Efficient Computing, Ministry of Education of China
nitang@stu.xmu.edu.cn, yyqu@xmu.edu.cn

Abstract

Diffusion models have revealed powerful potential in all-in-one image restoration (AiOIR), which is talented in generating abundant texture details. The existing AiOIR methods either retrain a diffusion model or fine-tune the pretrained diffusion model with extra conditional guidance. However, they often suffer from high inference costs and limited adaptability to diverse degradation types. In this paper, we propose an efficient AiOIR method, Diffusion Once and Done (DOD), which aims to achieve superior restoration performance with only one-step sampling of Stable Diffusion (SD) models. Specifically, multi-degradation feature modulation is first introduced to capture different degradation prompts with a pretrained diffusion model. Then, parameter-efficient conditional low-rank adaptation integrates the prompts to enable the fine-tuning of the SD model for adapting to different degradation types. Besides, a high-fidelity detail enhancement module is integrated into the decoder of SD to improve structural and textural details. Experiments demonstrate that our method outperforms existing diffusion-based restoration approaches in both visual quality and inference efficiency.

Code — <https://github.com/tonia86/DOD>

Introduction

All-in-One Image Restoration (AiOIR) aims to handle multiple degradation tasks within a unified framework, thereby enhancing the robustness and deployment efficiency of image restoration models. The existing methods typically improve model adaptability through scalable network architectures (Wang et al. 2024; Zeng et al. 2025; Tian et al. 2025), multi-task learning strategies (Wu et al. 2024), and integrating task-specific and task-agnostic priors (Yan et al. 2025; Luo et al. 2023). However, most of these methods primarily focus on optimizing distortion-based metrics (e.g., PSNR and SSIM), while ignoring the equally important perceptual quality (e.g., LPIPS and MANIQA).

Recently, diffusion-based AiOIR methods (Xiong et al. 2025; Zheng et al. 2024; Jiang et al. 2024; Ai et al. 2024)

^{*}These authors contributed equally.

[†]Correspondence author

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

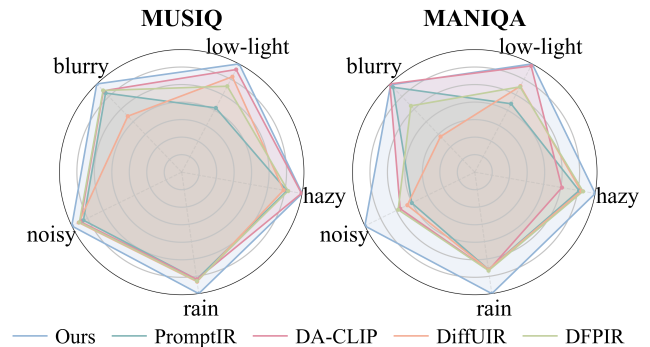


Figure 1: Comparison on five restoration tasks using MANIQA and MUSIQ shows that our method consistently outperforms others.

have gained increasing attention, owing to the strong potential in modeling complex natural image distributions. It favors more realistic outputs with superior perceptual quality. Existing diffusion-based approaches primarily fall into two categories: 1) Train diffusion models from scratch with the specific dataset (Xiong et al. 2025; Zheng et al. 2024). While offering the generative ability for customized tasks, they demand significant training expense and large-scale datasets to obtain a robust model. 2) Fine-tune the pre-trained diffusion models (e.g., Stable Diffusion (Rombach et al. 2022)) (Jiang et al. 2024; Ai et al. 2024). By means of the powerful diffusion prior in structure and texture generation, these approaches typically promote image restoration effects via adding conditional guidance or extra enhancement modules, while they still struggle to adapt to different degradation types. Note that these two types of methods rely on multi-step sampling, resulting in low inference efficiency that hinders practical deployment. Hence, it is desirable to propose an efficient diffusion-based method for robust AiOIR, which can reduce burdensome inference costs.

In this paper, we aim to design an efficient one-step diffusion method based on the pretrained Stable Diffusion (SD) for AiOIR, which can not only accelerate the inference but also maintain a superior restoration effect. However, it would face several challenges: 1) how to effectively char-

acterize diverse image degradation patterns to optimize and guide diffusion models for robust image restoration; 2) how to perform model fine-tuning while preserving the original parameter priors of the SD; 3) how to compensate for the high-quality detail loss inherent to one-step sampling.

To address the above problems, we propose a novel framework, Diffusion Once and Done (DOD), which is designed as a degradation-aware one-step diffusion solution for efficient and high-quality image restoration. Inspired by the efficient fine-tuning capability of Low-Rank Adaptation (LoRA) (Hu et al. 2022), we design a degradation-aware LoRA strategy that dynamically modulates the model’s feature space under varying degradation conditions, enabling effective modeling of diverse degradation patterns.

Specifically, our DOD includes Multi-degradation Feature Modulation (MFM), Parameter-efficient conditional LoRA (PLA), and High-fidelity Detail Enhancement (HDE). MFM aims to extract separate degradation information from intermediate features of the pretrained DDPM (Ho, Jain, and Abbeel 2020), which can effectively distinguish different degraded types. PLA integrates the degradation prompts with LoRA for feature space refinement, enabling a single set of parameters to adapt to different degradation types. In HDE, we design the detail enhancement module to promote the texture and detail reconstruction.

The whole method is optimized in a two-stage way to achieve fast and high-quality restoration. The two-stage training strategy ensures both computational efficiency through parameter-efficient tuning and performance superiority via degradation-aware adaptation and detail refinement. Extensive experiments demonstrate that our DOD achieves superior performance across multiple AiOIR benchmarks (as illustrated in Fig. 1). Notably, it reduces the sampling steps from multiple iterations to just one step, significantly improving inference efficiency. Meanwhile, with minimal parameter fine-tuning, our approach attains exceptional restoration quality, showcasing its strong practical value and deployment potential for real-world image restoration applications.

In summary, our contributions are as follows:

- We propose a one-step diffusion model for AiOIR that leverages abundant priors from SD, achieving high-quality restoration with significantly reduced inference time and trainable parameters.
- We introduce a degradation-aware LoRA strategy, which adaptively modulates the feature space based on degradation cues extracted from a pre-trained diffusion model, enabling effective removal of diverse degradations.
- We further incorporate a detail enhancement module in the decoder of SD to improve the restoration of textures and structural details. Besides, we adopt a two-stage optimization strategy to enable stable training.
- Extensive experiments show that our DOD achieves the best perceptual metrics with higher reconstruction accuracy on multiple image restoration tasks.

Related Works

All-in-one Image Restoration

Existing AiOIR methods typically improve generalization by architectural design and prompt mechanisms. For example, AirNet (Li et al. 2022) introduces a degradation encoder, TransWeather (Valanarasu, Yasarla, and Patel 2022) uses a query-based Transformer, PromptIR (Potlapalli et al. 2023) and DA-CLIP (Luo et al. 2023) enhance restoration via lightweight prompts and semantic guidance. Recent works such as VLU-Net (Zeng et al. 2025), DFPIR (Tian et al. 2025), and AdaIR (Cui et al. 2024) have pushed forward the progress of AiOIR by incorporating visual-language models, frequency-aware perception, and adaptive strategies.

With the success of diffusion models in image generation, some studies have explored their application in restoration. One group retrains diffusion models for multi-degradation tasks (e.g., WeatherDiffusion (Özdenizci and Legenstein 2023), JCDM (Yue et al. 2025), DA2Diff (Xiong et al. 2025), DiffUIR (Zheng et al. 2024)), but these approaches are costly and underutilize pre-trained priors. Others fine-tune large models like SD (e.g., AutoDIR (Jiang et al. 2024), MPerceiver (Ai et al. 2024), Diff-Restorer (Zhang et al. 2024b)), which leverage strong priors but often involve heavy parameters and slow inference. To address these issues, we propose a novel AiOIR method based on SD, combining LoRA for efficient fine-tuning with a one-step diffusion mechanism to significantly accelerate inference while maintaining generation quality.

Fine-tuning Strategies for AiOIR

Rapid advances in large generative models like diffusion models have made leveraging their priors vital for image restoration. Existing methods fall into two categories: structure-guided approaches using ControlNet (Zhang, Rao, and Agrawala 2023) and lightweight fine-tuning with LoRA (Hu et al. 2022). ControlNet injects structural priors (e.g., edges, depth) into diffusion models for conditional generation and is widely used in restoration. Key examples include DiffBIR (Lin et al. 2024), AdaptBIR (Liu et al. 2024), and Diff-Restorer (Zhang et al. 2024b), which improve structural awareness and adaptability. However, these methods often have complex designs and high inference costs.

In contrast, LoRA offers a more efficient and flexible alternative by applying low-rank decomposition to the weight matrices of pre-trained models, requiring only a small number of parameters to be fine-tuned. LoRA-IR (Ai, Huang, and He 2024) and UIR-LoRA (Zhang et al. 2024a) explore LoRA for multi-degradation modeling, but most are built on CNN or Transformer backbones, with limited integration into diffusion-based frameworks. Moreover, they often rely on stacked modules, leading to structural redundancy and slower response in complex degradation scenarios. To address these challenges, this paper presents an AiOIR method that, for the first time, deeply integrates LoRA’s lightweight adaptation mechanism with the SD model.

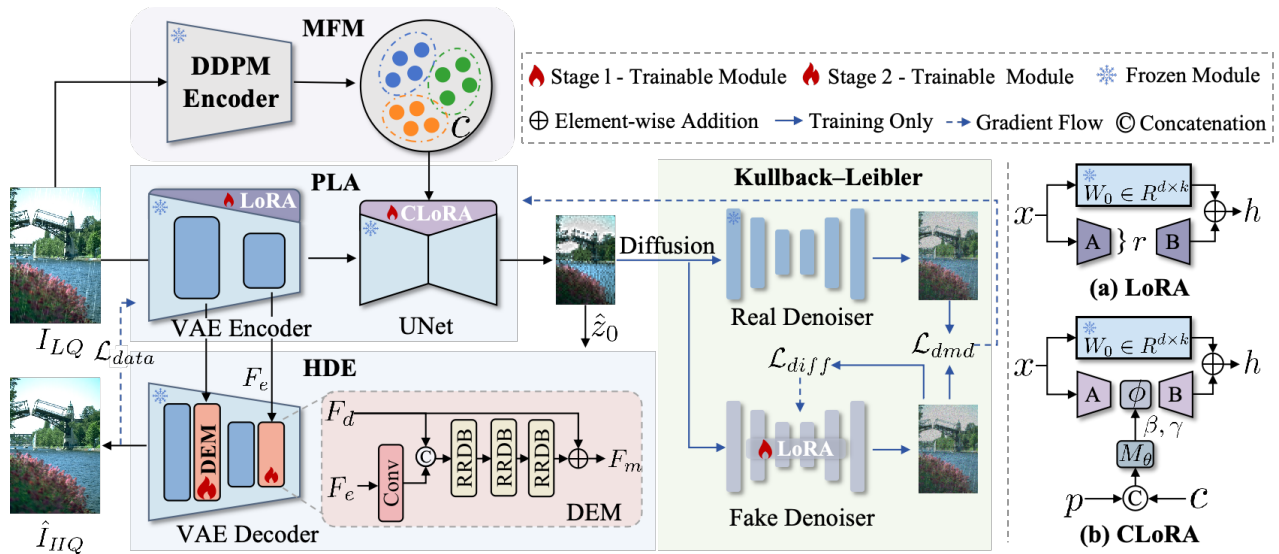


Figure 2: Overview of DOD. DOD enhances SD for AiOIR by injecting LoRA layers into the VAE encoder and conditional LoRA into the UNet. A detail enhancement module is also inserted into the VAE decoder to refine visual details. On the right, (a) shows standard LoRA, while (b) illustrates our proposed conditional LoRA.

Proposed Method

The DOD model is built upon the SD model, leveraging its exceptional image generation capabilities. To restore images degraded by various types, our model integrates three core modules: Multi-degradation Feature Modulation (MFM), Parameter-efficient conditional LoRA (PLA), and High-fidelity Detail Enhancement (HDE). These modules are designed to flexibly adjust the diffusion priors to accommodate different forms of image degradation.

As shown in Fig. 2, the overall architecture of the model is designed as follows: The low-quality degraded image I_{LQ} is first input into the MFM module to extract degradation feature c . Then, the extracted degradation features, along with the degraded image, are fed into the PLA module. The goal is to effectively remove the degradation with the aid of degradation information, thereby obtaining the potentially restored image features \hat{z}_0 . Subsequently, these restored latent features are passed into the HDE module to enhance the structural fidelity and texture quality further, ultimately producing a clear and high-quality restored image \hat{I}_{HQ} .

Multi-degradation Feature Modulation

Recent studies suggest that recognizing degradation types is crucial for improving both the generalization and performance of image restoration models (Hu et al. 2025). To enable rapid adaptation to various tasks, we introduce a degradation-aware module that enhances the SD model’s perception of input degradations. We focus on the h-space, which refers to the intermediate feature space of the UNet within pre-trained DDPM (Ho, Jain, and Abbeel 2020). This space has been shown to contain rich semantic information and offer strong controllability in high-level vision tasks (Kwon, Jeong, and Uh 2022; Jeong, Kwon, and Uh 2024). More recent work, such as DiffL2ID (Yang et al. 2024), fur-

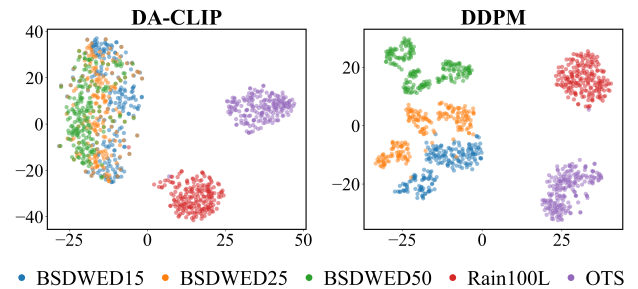


Figure 3: t-SNE visualization of degradation features extracted by DA-CLIP and DDPM. DA-CLIP learns degradation-aware prompts by fine-tuning CLIP, while DDPM features are taken from intermediate layers of a pretrained DDPM. Compared to DA-CLIP, DDPM features provide clearer separation across degradation types and noise levels, indicating stronger discriminative capacity.

ther reveals that h-space can also encode low-level degradation features in the context of image dehazing.

Building on these findings, we conduct further experiments and observe that at the diffusion timestep $t = 0$, the features in the h-space are clearly separable across different degradation types, as illustrated by the t-SNE visualization in Fig. 3. Motivated by this observation, we directly leverage the intermediate features from the pre-trained DDPM as the degradation-aware representation c , avoiding additional module training and enabling efficient and generalizable degradation modeling.

Parameter-efficient Conditional LoRA

To enable degradation-aware restoration with minimal parameter overhead, we introduce PLA that leverages LoRA to adapt both the VAE encoder and the UNet in SD, ensuring the generation of high-quality, restored posterior features. In contrast to existing methods (Zhang et al. 2024a) that typically train multiple LoRA modules independently and combine them based on degradation similarity, we propose a conditional LoRA (CLoRA) mechanism to fine-tune the UNet in SD. By embedding degradation information extracted from MFM into the LoRA structure, CLoRA enables feature space refinement, allowing a single set of parameters to flexibly handle diverse degradation types.

To prepare for our conditional design, we first revisit the standard LoRA formulation as illustrated in Fig. 2(a). Standard LoRA fine-tuning freezes the original weight matrix $W_0 \in \mathbb{R}^{d \times k}$ and introduces a pair of low-rank matrices $A \in \mathbb{R}^{r \times k}$ and $B \in \mathbb{R}^{d \times r}$ to form the adapted weight:

$$W_0 + \Delta W = W_0 + BA, \quad (1)$$

where the rank r is chosen such that $r \ll \min(d, k)$. The forward pass of this layer becomes:

$$h = W_0x + \Delta Wx = W_0x + BAx, \quad (2)$$

typically, B is initialized to zero so that LoRA does not affect the model behavior at the start of fine-tuning.

Based on this standard formulation, we introduce a conditional control mechanism that modulates the low-rank embeddings via a transformation function $\phi(Ax | c)$, where c is the degradation information encoded by the MFM module. The overall structure of this CLoRA is illustrated in Fig. 2(b). Specifically, the output of the CLoRA layer is formulated as:

$$h = W_0x + \Delta Wx = W_0x + B\phi(Ax | c), \quad (3)$$

where $\phi(Ax | c)$ adopts a simple affine form:

$$\phi(Ax | c) = \gamma_c \odot Ax + \beta_c, \quad (4)$$

where \odot denotes the element-wise multiplication, and the affine parameters γ_c and β_c are learned from the degradation condition c . This CLoRA module offers two key advantages: it enables dynamic, data-dependent parameter adjustment based on the specific degradation characteristics of the low-quality input, and it operates without requiring any manual configuration or user-provided degradation labels.

However, sharing the same affine parameters across all LoRA layers limits the flexibility of fine-tuning. Conversely, if assigning a different affine parameter for each LoRA layer, it causes a significant increase in the total number of trainable parameters due to the construction of a dedicated MLP per layer. To enable flexible parameterization of CLoRA in SD comprising N layers, we introduce a set of learnable prompts $\mathcal{P} = \{p_i\}_{i=1}^N$. The affine parameters for the i -th layer, γ_i and β_i , are generated as:

$$\gamma_i, \beta_i = M_\theta([c, p_i]), \quad (5)$$

where $[\cdot, \cdot]$ denotes a concatenation operation, M_θ is a lightweight network composed of several MLP and normalization layers. Compared to methods such as ControlNet, which duplicate the entire UNet encoder, our approach is superior in both parameter efficiency and computational cost.

High-fidelity Detail Enhancement

To balance inference efficiency and restoration quality, we replace the conventional multi-step sampling in diffusion models with a one-step strategy, which substantially reduces computational overhead but may underperform in recovering fine details. To compensate for this limitation, we introduce a detail enhancement module (DEM) in the VAE decoder to further refine textures and enhance image fidelity.

Specifically, features extracted by the encoder are first passed through a convolutional layer initialized with zeros to align their dimensions with the decoder features. These aligned features are then concatenated with the corresponding decoder features and fed into a series of residual-in-residual dense blocks (RRDB) (Wang et al. 2018) to extract detail-aware representations. The output is subsequently fused with the original decoder features to compensate for the detail loss caused by one-step diffusion. Additionally, a learnable dynamic weight w is introduced to control the refinement strength. The overall process can be formulated as:

$$F_m = F_d + \text{RRDB}([\text{Conv}(F_e), F_d]; \theta) \times w, \quad (6)$$

where F_e and F_d denote the features from the VAE encoder and decoder, respectively, and F_m is the final fused feature. $\text{Conv}(\cdot)$ represents a convolutional layer initialized with zeros and $\text{RRDB}(\cdot; \theta)$ denotes the RRDB module with learnable parameters θ .

Optimization

To stabilize the training process, we adopt a two-stage strategy to optimize the whole model.

Stage 1 aims to efficiently adapt the SD model for image restoration tasks. In this stage, we train only the PLA module while keeping the VAE decoder frozen, decoding the generated posterior feature \hat{z}_0 to obtain the restored image \hat{I}_{HQ} . During training, standard data consistency losses are employed to measure the error of alignment between the reconstructed and clean images at both pixel and semantic levels. To enable one-step diffusion, degradation-aware image restoration while preserving the generative capacity of the pretrained model, we further incorporate the distribution matching distillation (DMD) loss (Yin et al. 2024), which enhances both generative quality and generalization ability.

1) *Data consistency loss.* The data consistency loss $\mathcal{L}_{\text{data}}$ consists of a pixel-wise reconstruction term and a perceptual similarity term, formulated as:

$$\mathcal{L}_{\text{data}} = \lambda_{\text{rec}} \cdot \mathcal{L}_{\text{rec}} + \lambda_{\text{lpips}} \cdot \mathcal{L}_{\text{lpips}}, \quad (7)$$

$$\mathcal{L}_{\text{rec}} = \|\hat{I}_{HQ} - I_{HQ}\|_2, \quad (8)$$

$$\mathcal{L}_{\text{lpips}} = \|\Phi(\hat{I}_{HQ}) - \Phi(I_{HQ})\|_2, \quad (9)$$

where \hat{I}_{HQ} denotes the generated image and I_{HQ} is the ground-truth clean image. $\Phi(\cdot)$ represents perceptual features extracted by VGG (Simonyan and Zisserman 2014) network. λ_{rec} and λ_{lpips} balance the two loss terms.

2) *Distribution matching distillation loss.* To enable the network to rapidly generate samples that are indistinguishable from real images, the loss function (Yin et al. 2024)

minimizes the KL divergence between the real image feature distribution p_{real} and the generated image feature distribution p_{fake} .

$$\mathcal{L}_{\text{dmd}} = \text{KL}(p_{\text{real}} \parallel p_{\text{fake}}). \quad (10)$$

In practice, we directly backpropagate through the KL divergence, whose gradient is computed by estimating the scores of the real and generated distributions. To this end, we employ a pair of diffusion denoisers to model the score functions of real and fake distributions after Gaussian diffusion. As illustrated in Fig. 2, both denoisers adopt the UNet architecture from SD and are initialized with the same pretrained weights. The real denoiser keeps its parameters frozen throughout training, while the fake denoiser is dynamically updated according to the evolving distribution of generated samples. It is jointly optimized with the image restoration network using the standard noise prediction loss $\mathcal{L}_{\text{diff}}$ from SD. Please refer to the supplementary material for the detailed loss.

The overall loss for stage 1 of DOD formulated as:

$$\mathcal{L}_{\text{stage1}} = \mathcal{L}_{\text{data}} + \mathcal{L}_{\text{dmd}}. \quad (11)$$

Stage 2 focuses on recovering the missing fine details from the latent features reconstructed in the previous stage. To achieve this, the PLA module fine-tuned in Stage 1 is frozen, and only the DEMs, which are inserted into each layer of the VAE decoder as part of the HDE, are trained. A combination of reconstruction loss and structural similarity loss is employed to enhance perceptual quality and detail fidelity.

Structural similarity loss. To promote structural integrity and contrast preservation, we use SSIM-based loss (Wang et al. 2004):

$$\mathcal{L}_{\text{ssim}} = 1 - \text{SSIM}(\hat{I}_{HQ}, I_{HQ}), \quad (12)$$

which penalizes discrepancies in luminance, contrast, and structural patterns between \hat{I}_{HQ} and I_{HQ} .

The total loss at stage 2 is:

$$\mathcal{L}_{\text{stage2}} = \lambda_{\text{rec}} \cdot \mathcal{L}_{\text{rec}} + \lambda_{\text{ssim}} \cdot \mathcal{L}_{\text{ssim}}, \quad (13)$$

where λ_{rec} and λ_{ssim} are hyperparameters that balance each component’s contribution.

Experiments

Experimental Settings

Datasets. Following (Tian et al. 2025; Li et al. 2022), we construct task-specific datasets under two standard AiOIR settings. (1) *Rain-haze-noise*: BSD400 (Arbelaez et al. 2010) and WED (Ma et al. 2016) for denoising (tested on BSD68 (Martin et al. 2001)), Rain100L (Yang et al. 2017) for deraining, and RESIDE (Li et al. 2018) (OTS/SOTS) for dehazing. (2) *Rain-haze-noise-blur-dark*: additional GoPro (Nah, Hyun Kim, and Mu Lee 2017) for deblurring and LOL (Wei et al. 2018) for low-light enhancement.

Metrics. We report both full-reference metrics (PSNR, SSIM (Wang et al. 2004)) and perceptual metrics (LPIPS (Zhang et al. 2018), DISTS (Ding et al. 2020), FID (Heusel et al. 2017)). We also include common no-reference metrics: NIQE (Zhang, Zhang, and Bovik 2015), MANIQA (Yang

et al. 2022), MUSIQ (Ke et al. 2021), and CLIPQA (Wang, Chan, and Loy 2023).

Implementation Details. All experiments were conducted on an NVIDIA RTX 3090 using PyTorch. The LoRA rank in the VAE encoder and U-Net is 8, and SD 2.1-base is used as the T2I backbone. We set $\lambda_{\text{rec}} = 1$, $\lambda_{\text{lpiips}} = 2$, and $\lambda_{\text{ssim}} = 0.2$. Additional settings are provided in the supplementary material.

Comparison with State-of-the-Arts

Compared methods. We benchmark our method against nine state-of-the-art AiOIR baselines: AirNet (Li et al. 2022), PromptIR (Potlapalli et al. 2023), LORA-IR (Ai, Huang, and He 2024), DA-CLIP (Luo et al. 2023), DiffUIR (Zheng et al. 2024), VLU-Net (Zeng et al. 2025), DFPIR (Tian et al. 2025), AdaIR (Cui et al. 2024), and AutoDIR (Jiang et al. 2024). Among them, DA-CLIP and DiffUIR train diffusion models from scratch in the pixel domain, whereas AutoDIR is built upon the pre-trained SD. The remaining approaches follow non-generative paradigms.

Quantitative comparisons. Table 1 presents the average quantitative results of our method across three image restoration tasks. It is observed that non-diffusion methods perform strongly on traditional fidelity metrics such as PSNR and SSIM, with AdaIR and DFPIR achieving the highest scores of 32.98 dB (PSNR) and 0.9162 (SSIM), respectively. However, these methods fall short on perceptual quality metrics, making it challenging to satisfy the demands of real-world visual experience.

In contrast, diffusion-based methods, exemplified by DA-CLIP and our approach, show clear advantages in perceptual metrics. Their restoration results better align with human visual preferences in terms of detail preservation and structural integrity. Notably, our method achieves the best scores across five perceptual metrics, demonstrating superior perceptual quality. Moreover, while maintaining perceptual performance comparable to DA-CLIP, our method significantly improves fidelity, with PSNR and SSIM increasing by 1.60 dB and 0.034 respectively, further validating its strengths in restoring image details and preserving structure.

Furthermore, we extend the evaluation to a more challenging five-task setting by incorporating deblurring and low-light enhancement. Table 2 summarizes the performance of each task under two perceptual quality metrics, MUSIQ and MANIQA. Our method consistently achieves the best or second-best results across all tasks, with average improvements of approximately 0.16% to 2.75% on MUSIQ and 0.57% to 8.24% on MANIQA, highlighting strong generalization capability. Particularly in perceptually demanding tasks such as deblurring and low-light enhancement, our method leads in perceptual scores, further confirming its effectiveness in restoring details across diverse degradation types. These results demonstrate that our unified framework not only excels in the three-task setting but also exhibits high adaptability and robustness in complex multi-task scenarios.

Qualitative comparisons. Fig. 4 presents a visual comparison between our method and representative approaches across three tasks. Although existing methods can reduce

	Method	Venue	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	DISTS \downarrow	CLIQQA \uparrow	NIQE \downarrow	MUSIQ \uparrow	MANIQA \uparrow	FID \downarrow
Non-Diff	AirNet	CVPR2022	31.11	0.9068	0.0925	0.0950	0.6405	3.3911	66.98	0.6470	41.80
	PromptIR	NeurIPS23	32.18	0.9124	0.0859	0.0864	0.6409	4.0061	67.21	0.6569	38.64
	LORA-IR	arXiv2024	30.98	0.9031	0.0889	0.0809	0.6352	3.4133	67.22	0.6568	30.30
	VLU-Net	CVPR2025	32.55	0.9157	0.0796	0.0785	0.6433	3.6559	67.09	0.6698	34.78
	DFPIR	CVPR2025	32.75	0.9162	0.0758	0.0758	0.6626	3.5938	67.34	0.6679	34.52
	AdaIR	ICLR2025	32.98	0.9155	0.0825	0.0820	0.6435	3.6464	67.50	0.6685	37.41
Diff	DA-CLIP	ICLR2024	30.27	0.8780	0.0664	0.0650	0.6580	3.2783	67.31	0.6663	23.09
	DiffUIR	CVPR2024	31.89	0.9010	0.0959	0.0964	0.6163	3.7371	67.51	0.6570	37.51
	Ours	-	31.87	0.9122	0.0620	0.0626	0.6660	3.5212	68.85	0.7080	19.54

Table 1: Quantitative comparison on average performance across three restoration tasks: dehazing, deraining, and Gaussian denoising ($\sigma = 15, 25, 50$). The best results are marked in boldface.

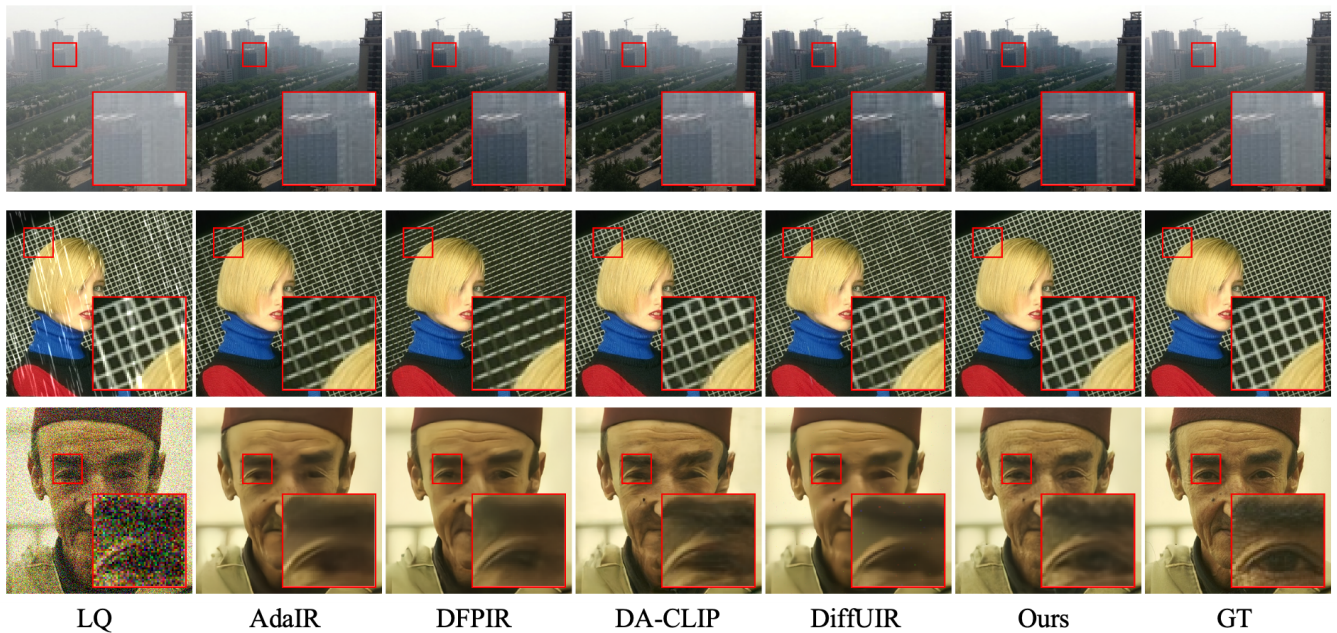


Figure 4: Visual comparisons with state-of-the-art AiOIR methods on three tasks. Please zoom in for a better view.

image degradation to some extent, the enlarged regions highlight their deficiencies in detail restoration. For instance, in the denoising task, while various methods successfully remove noise, the resulting images are overly smooth and lack detail, negatively impacting visual quality. In contrast, our method not only effectively removes noise but also better preserves fine details such as facial textures. This demonstrates the superior stability and performance of our method across multiple image restoration tasks. Additional visual comparisons can be found in the supplementary material.

Inference speed comparison. We compare four representative diffusion-based restoration methods in terms of inference steps and time. In Table 3, our method achieves the fastest inference with just one step. Although WeatherDiff uses only 25 steps, its patch-based overlapping strategy makes it much slower. Compared to DA-CLIP, which

offers similar visual quality, our method is about 84 \times faster, demonstrating clear efficiency advantages.

Ablation Study

Effectiveness of CLoRA and DEM. The results in Table 4 demonstrate that CLoRA and DEM play complementary roles in image restoration. Introducing conditional LoRA improves MUSIQ by 2.19 and reduces DISTS by 22%, indicating a positive impact on structural consistency and perceptual quality. DEM, on the other hand, focuses on enhancing fine details. When applied independently, it increases PSNR by 8.78 dB and MUSIQ by 3.49, highlighting its effectiveness in reconstructing low-level features. This is further illustrated in the visual comparisons shown in Fig. 5, where DEM successfully restores fine-grained textures of the original image. When combined, the two modules achieve the best performance across PSNR, DISTS, and

Method	Dehazing		Deraining		Denoising		Deblurring		Low-light		
	MUSIQ \uparrow	MANIQA \uparrow	MUSIQ \uparrow	MANIQA \uparrow	MUSIQ \uparrow	MANIQA \uparrow	MUSIQ \uparrow	MANIQA \uparrow	MUSIQ \uparrow	MANIQA \uparrow	
Non-Diff	PromptIR	65.43	0.6829	68.00	0.6679	67.29	0.6295	36.60	0.4250	64.76	0.5922
	AdaIR	65.83	0.6923	68.31	0.6720	67.65	0.6593	33.41	0.3763	69.43	0.6309
	VLU-Net	65.81	0.6920	68.45	0.6740	68.38	0.6669	32.48	0.3568	65.21	0.5948
	DFPIR	65.61	0.6898	68.29	0.6700	68.06	0.6563	36.93	0.3978	69.19	0.6234
Diff	DA-CLIP	67.88	0.6524	68.60	0.6695	67.88	0.6524	36.89	0.4298	72.59	0.6619
	DiffUIR	64.84	0.6864	68.40	0.6682	68.19	0.6383	33.63	0.3529	71.13	0.6246
	Ours	67.99	0.7115	70.49	0.7104	69.50	0.7219	37.77	0.4286	73.75	0.6657

Table 2: Quantitative comparison of no-reference metrics MUSIQ and MANIQA across five restoration tasks: dehazing, deraining, denoising ($\sigma = 25$), deblurring, and low-light enhancement. The best results are marked in boldface.

Method	WeatherDiff	AutoDIR	DA-CLIP	DiffUIR	Ours
Step \downarrow	25	200	100	3	1
Time(s) \downarrow	212.534	26.602	17.664	0.303	0.211

Table 3: Comparison of inference steps and time. All methods are tested with an input image of size 512×512 .

MUSIQ, suggesting a strong synergistic effect. These findings confirm that the proposed design effectively balances fidelity and perceptual quality, and validate its applicability to multi-task image restoration.

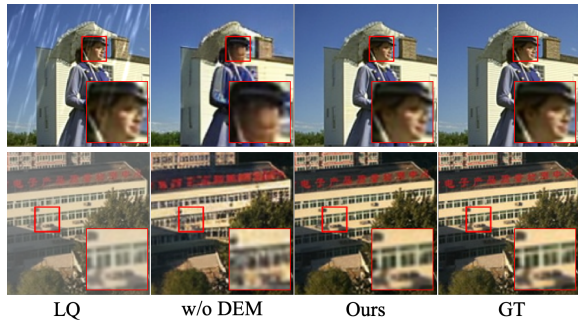


Figure 5: With DEM, fine structures are well preserved; without it, degradation is removed but key details are lost.

CLoRA	DEM	PSNR \uparrow	DISTS \downarrow	NIQE \downarrow	MUSIQ \uparrow
\times	\times	23.05	0.1121	4.4220	64.81
\checkmark	\times	24.60	0.0873	3.7072	67.00
\times	\checkmark	31.83	0.0666	3.5118	68.30
\checkmark	\checkmark	31.87	0.0626	3.5212	68.85

Table 4: Ablation study of CLoRA and DEM.

Effect of CLoRA placement. As shown in Table 5, applying it solely to the UNet leads to the best performance, achieving a PSNR of 31.87 dB with corresponding improvements in DISTS and MUSIQ. This suggests that CLoRA is more effective when used to modulate the generation process closely related to degradation. In contrast, introducing conditional modulation into the VAE encoder may disrupt its

Method	PSNR \uparrow	DISTS \downarrow	NIQE \downarrow	MUSIQ \uparrow
VAE Encoder + UNet	30.68	0.0666	3.4106	68.58
UNet only	31.87	0.0626	3.5212	68.85

Table 5: Performance comparison of CLoRA placement.

representation ability and negatively impact overall restoration quality.

The number of LoRA rank. As shown in Table 6, we evaluate the impact of different LoRA ranks across three tasks when jointly fine-tuning the VAE encoder and UNet. A small rank $r = 4$ leads to the model collapse directly, while a large rank $r = 16$ brings a certain degree of performance drop. In contrast, a rank of 8 achieves the best balance between stability and performance. Thus, we set the LoRA rank to 8 for both the fine-tuning of VAE encoder and UNet.

Rank	PSNR \uparrow	DISTS \downarrow	NIQE \downarrow	MUSIQ \uparrow
4	—	—	—	—
8	31.87	0.0620	3.5212	68.85
16	31.68	0.0626	3.5409	68.13

Table 6: Comparison of different LoRA ranks.

Conclusion

This paper presents an efficient all-in-one image restoration method, dubbed Diffusion Once and Done (DOD). By incorporating a degradation-aware LoRA fine-tuning strategy and a two-stage structural optimization mechanism, the method enables image restoration with various degradation types. Combined with distribution matching distillation, it produces high-quality restored images using one-step sampling while significantly reducing fine-tuning costs. Experimental results demonstrate that the proposed method achieves strong performance in subjective quality and objective fidelity across multiple restoration tasks. Despite the promising performance of one-step diffusion in image restoration tasks, the computational cost remains a limiting factor for deployment on resource-constrained devices. In future work, we plan to explore model compression techniques such as quantization to further reduce inference overhead.

Acknowledgments

This work is supported in part by the National Natural Science Foundation of China under Grant 62176224; in part by the Science and Technology on Sonar Laboratory under Grant 2024-JCJQ-LB-32/07; in part by the Doctoral Research Initiation Fund of Jimei University under Grant ZQ2023022.

References

- Ai, Y.; Huang, H.; and He, R. 2024. LoRA-IR: Taming Low-Rank Experts for Efficient All-in-One Image Restoration. *arXiv preprint arXiv:2410.15385*.
- Ai, Y.; Huang, H.; Zhou, X.; Wang, J.; and He, R. 2024. Multimodal prompt perceiver: Empower adaptiveness generalizability and fidelity for all-in-one image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 25432–25444.
- Arbelaez, P.; Maire, M.; Fowlkes, C.; and Malik, J. 2010. Contour detection and hierarchical image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 33(5): 898–916.
- Cui, Y.; Zamir, S. W.; Khan, S.; Knoll, A.; Shah, M.; and Khan, F. S. 2024. Adair: Adaptive all-in-one image restoration via frequency mining and modulation. *arXiv preprint arXiv:2403.14614*.
- Ding, K.; Ma, K.; Wang, S.; and Simoncelli, E. P. 2020. Image quality assessment: Unifying structure and texture similarity. *IEEE transactions on pattern analysis and machine intelligence*, 44(5): 2567–2581.
- Heusel, M.; Ramsauer, H.; Unterthiner, T.; Nessler, B.; and Hochreiter, S. 2017. GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium. In *Advances in Neural Information Processing Systems*.
- Ho, J.; Jain, A.; and Abbeel, P. 2020. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33: 6840–6851.
- Hu, E. J.; Shen, Y.; Wallis, P.; Allen-Zhu, Z.; Li, Y.; Wang, S.; Wang, L.; Chen, W.; et al. 2022. Lora: Low-rank adaptation of large language models. *ICLR*, 1(2): 3.
- Hu, J.; Jin, L.; Yao, Z.; and Lu, Y. 2025. Universal image restoration pre-training via degradation classification. *arXiv preprint arXiv:2501.15510*.
- Jeong, J.; Kwon, M.; and Uh, Y. 2024. Training-free content injection using h-space in diffusion models. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 5151–5161.
- Jiang, Y.; Zhang, Z.; Xue, T.; and Gu, J. 2024. Autodir: Automatic all-in-one image restoration with latent diffusion. In *European Conference on Computer Vision*, 340–359. Springer.
- Ke, J.; Wang, Q.; Wang, Y.; Milanfar, P.; and Yang, F. 2021. Musiq: Multi-scale image quality transformer. In *Proceedings of the IEEE/CVF international conference on computer vision*, 5148–5157.
- Kwon, M.; Jeong, J.; and Uh, Y. 2022. Diffusion models already have a semantic latent space. *arXiv preprint arXiv:2210.10960*.
- Li, B.; Liu, X.; Hu, P.; Wu, Z.; Lv, J.; and Peng, X. 2022. All-in-one image restoration for unknown corruption. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 17452–17462.
- Li, B.; Ren, W.; Fu, D.; Tao, D.; Feng, D.; Zeng, W.; and Wang, Z. 2018. Benchmarking single-image dehazing and beyond. *IEEE Transactions on Image Processing*, 28(1): 492–505.
- Lin, X.; He, J.; Chen, Z.; Lyu, Z.; Dai, B.; Yu, F.; Qiao, Y.; Ouyang, W.; and Dong, C. 2024. Diffbir: Toward blind image restoration with generative diffusion prior. In *European Conference on Computer Vision*, 430–448. Springer.
- Liu, Y.; He, J.; Liu, Y.; Lin, X.; Yu, F.; Hu, J.; Qiao, Y.; and Dong, C. 2024. AdaptBIR: Adaptive Blind Image Restoration with latent diffusion prior for higher fidelity. *Pattern Recognition*, 155: 110659.
- Luo, Z.; Gustafsson, F. K.; Zhao, Z.; Sjölund, J.; and Schön, T. B. 2023. Controlling vision-language models for universal image restoration. *arXiv preprint arXiv:2310.01018*, 3(8).
- Ma, K.; Duanmu, Z.; Wu, Q.; Wang, Z.; Yong, H.; Li, H.; and Zhang, L. 2016. Waterloo exploration database: New challenges for image quality assessment models. *IEEE Transactions on Image Processing*, 26(2): 1004–1016.
- Martin, D.; Fowlkes, C.; Tal, D.; and Malik, J. 2001. 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, 416–423. IEEE.
- Nah, S.; Hyun Kim, T.; and Mu Lee, K. 2017. Deep multi-scale convolutional neural network for dynamic scene deblurring. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3883–3891.
- Özdenizci, O.; and Legenstein, R. 2023. Restoring vision in adverse weather conditions with patch-based denoising diffusion models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(8): 10346–10357.
- Potlapalli, V.; Zamir, S. W.; Khan, S. H.; and Shahbaz Khan, F. 2023. Promptir: Prompting for all-in-one image restoration. *Advances in Neural Information Processing Systems*, 36: 71275–71293.
- Rombach, R.; Blattmann, A.; Lorenz, D.; Esser, P.; and Ommer, B. 2022. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 10684–10695.
- Simonyan, K.; and Zisserman, A. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Tian, X.; Liao, X.; Liu, X.; Li, M.; and Ren, C. 2025. Degradation-Aware Feature Perturbation for All-in-One Image Restoration. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, 28165–28175.

- Valanarasu, J. M. J.; Yasarla, R.; and Patel, V. M. 2022. Transweather: Transformer-based restoration of images degraded by adverse weather conditions. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2353–2363.
- Wang, H.-H.; Tsai, F.-J.; Lin, Y.-Y.; and Lin, C.-W. 2024. TANet: Triplet Attention Network for All-In-One Adverse Weather Image Restoration. In *Proceedings of the Asian Conference on Computer Vision*, 835–851.
- Wang, J.; Chan, K. C.; and Loy, C. C. 2023. Exploring clip for assessing the look and feel of images. In *Proceedings of the AAAI conference on artificial intelligence*, 2, 2555–2563.
- Wang, X.; Yu, K.; Wu, S.; Gu, J.; Liu, Y.; Dong, C.; Qiao, Y.; and Change Loy, C. 2018. Esrgan: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European conference on computer vision (ECCV) workshops*, 0–0.
- Wang, Z.; Bovik, A. C.; Sheikh, H. R.; and Simoncelli, E. P. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4): 600–612.
- Wei, C.; Wang, W.; Yang, W.; and Liu, J. 2018. Deep retinex decomposition for low-light enhancement. *arXiv preprint arXiv:1808.04560*.
- Wu, G.; Jiang, J.; Jiang, K.; and Liu, X. 2024. Harmony in diversity: Improving all-in-one image restoration via multi-task collaboration. In *Proceedings of the 32nd ACM International Conference on Multimedia*, 6015–6023.
- Xiong, J.; Yan, X.; Wang, Y.; Zhao, W.; Zhang, X.-P.; and Wei, M. 2025. DA2Diff: Exploring Degradation-aware Adaptive Diffusion Priors for All-in-One Weather Restoration. *arXiv preprint arXiv:2504.05135*.
- Yan, Q.; Jiang, A.; Chen, K.; Peng, L.; Yi, Q.; and Zhang, C. 2025. Textual prompt guided image restoration. *Engineering Applications of Artificial Intelligence*, 155: 110981.
- Yang, S.; Wu, T.; Shi, S.; Lao, S.; Gong, Y.; Cao, M.; Wang, J.; and Yang, Y. 2022. Maniqa: Multi-dimension attention network for no-reference image quality assessment. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 1191–1200.
- Yang, W.; Tan, R. T.; Feng, J.; Liu, J.; Guo, Z.; and Yan, S. 2017. Deep joint rain detection and removal from a single image. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1357–1366.
- Yang, Z.; Yu, H.; Li, B.; Zhang, J.; Huang, J.; and Zhao, F. 2024. Unleashing the Potential of the Semantic Latent Space in Diffusion Models for Image Dehazing. In *European Conference on Computer Vision*, 371–389. Springer.
- Yin, T.; Gharbi, M.; Zhang, R.; Shechtman, E.; Durand, F.; Freeman, W. T.; and Park, T. 2024. One-step diffusion with distribution matching distillation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 6613–6623.
- Yue, Y.; Yu, M.; Yang, L.; and Liu, T. 2025. Joint Conditional Diffusion Model for image restoration with mixed degradations. *Neurocomputing*, 129512.
- Zeng, H.; Wang, X.; Chen, Y.; Su, J.; and Liu, J. 2025. Vision-Language Gradient Descent-driven All-in-One Deep Unfolding Networks. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, 7524–7533.
- Zhang, C.; Gong, D.; He, J.; Zhu, Y.; Sun, J.; and Zhang, Y. 2024a. UIR-LoRA: Achieving Universal Image Restoration through Multiple Low-Rank Adaptation. *arXiv preprint arXiv:2409.20197*.
- Zhang, L.; Rao, A.; and Agrawala, M. 2023. Adding conditional control to text-to-image diffusion models. In *Proceedings of the IEEE/CVF international conference on computer vision*, 3836–3847.
- Zhang, L.; Zhang, L.; and Bovik, A. C. 2015. A feature-enriched completely blind image quality evaluator. *IEEE Transactions on Image Processing*, 24(8): 2579–2591.
- Zhang, R.; Isola, P.; Efros, A. A.; Shechtman, E.; and Wang, O. 2018. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 586–595.
- Zhang, Y.; Zhang, H.; Chai, X.; Cheng, Z.; Xie, R.; Song, L.; and Zhang, W. 2024b. Diff-restorer: Unleashing visual prompts for diffusion-based universal image restoration. *arXiv preprint arXiv:2407.03636*.
- Zheng, D.; Wu, X.-M.; Yang, S.; Zhang, J.; Hu, J.-F.; and Zheng, W.-S. 2024. Selective hourglass mapping for universal image restoration based on diffusion model. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 25445–25455.