
UIR-LoRA: Achieving Universal Image Restoration through Multiple Low-Rank Adaptation

Anonymous Author(s)

Affiliation

Address

email

Abstract

1 Existing unified methods typically treat multi-degradation image restoration as a
2 multi-task learning problem. Despite performing effectively compared to single
3 degradation restoration methods, they overlook the utilization of commonalities
4 and specificities within multi-task restoration, thereby impeding the model’s per-
5 formance. Inspired by the success of deep generative models and fine-tuning tech-
6 niques, we proposed a universal image restoration framework based on multiple
7 low-rank adapters (LoRA) from multi-domain transfer learning. Our framework
8 leverages the pre-trained generative model as the shared component for multi-
9 degradation restoration and transfers it to specific degradation image restoration
10 tasks using low-rank adaptation. Additionally, we introduce a LoRA composing
11 strategy based on the degradation similarity, which adaptively combines trained
12 LoRAs and enables our model to be applicable for mixed degradation restoration.
13 Extensive experiments on multiple and mixed degradations demonstrate that the
14 proposed universal image restoration method not only achieves higher fidelity and
15 perceptual image quality but also has better generalization ability than other unified
16 image restoration models.

17 1 Introduction

18 In the wild, a range of distortions commonly appear in captured images, including noise[56], blur[14,
19 47, 6], low light[58, 22, 8], and various weather degradations[15, 51, 54, 45]. As a fundamental task
20 in low-level vision, image restoration aims to eliminate these distortions and recover sharp details and
21 original scene information from corrupted images. With the assistance of deep learning, an abundance
22 of restoration approaches [56, 3, 54, 2, 16, 14, 53] have made significant progress in eliminating
23 single degradation from images. However, these approaches typically require additional training from
24 scratch on specific image pairs in multi-degraded scenarios, which leads to inconvenience in usage
25 and limited generalization ability.

26 For simplicity and practicality, some existing works [15, 31, 55] consider training a unified model
27 (also called all-in-one model) to handle multiple degradations as multi-task learning. These studies
28 primarily explore how to discern degradation from the image and integrate it into the restoration
29 network. Nevertheless, these methods share all parameters across different degradations, resulting in
30 gradient conflicts [40, 52] that hinder further improvement of unified models’ performance.

31 Digging deeper, the underlying issue lies in that the similarities among different image restoration
32 tasks and the inherent specificity of each degradation are not well considered and utilized in the
33 training. This limitation drives us to seek solutions for multi-degradation restoration by leveraging
34 both commonalities and specificities.

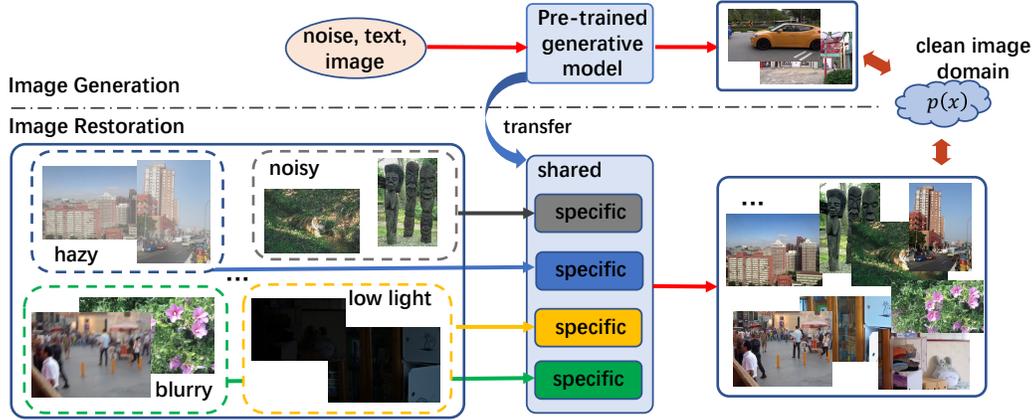


Figure 1: Motivation of our work. A pre-trained generative model serves as the shared component and minimal parameters are added to model the specificity of each degradation restoration task.

35 Inspired by the successes of deep generative models[37, 36, 35] and fine-tuning techniques[11, 10, 4],
 36 we propose addressing the aforementioned issue from the perspective of multi-domain transfer
 37 learning, as presented in Figure 1. The pre-trained generative model exhibits powerful capabilities,
 38 implying rich prior knowledge of clear image distribution $p(x)$, which is exactly what is needed
 39 for image restoration. Since image prior $p(x)$ is degradation-agnostic and applicable to all types
 40 of degraded images, the pre-trained generative model is an excellent candidate for serving as the
 41 shared component for multiple degradation restoration. To model the transition from the clean image
 42 domain to different degraded image domains, minimal specific parameters are required to fine-tune
 43 the pre-trained model for each degradation restoration task. This approach not only isolates conflicts
 44 between each degradation task but also ensures efficiency and performance during training.
 45 Following the idea of multi-domain transfer learning, we proposed a universal image restoration
 46 framework based on multiple low-rank adaptations, named UIR-LoRA. In our framework, the pre-
 47 trained SD-turbo [39] serves as the shared fundamental model for multiple degradation restoration
 48 tasks due to its powerful one-step generation capability and extensive image priors. Subsequently,
 49 we incorporate the low-rank adaptation (LoRA) technique [11] to fine-tune the base model for each
 50 specific image restoration task. This involves augmenting low-dimensional parameter matrices on
 51 selected layers within the base model, ensuring efficient fine-tuning while maintaining independence
 52 between LoRAs for each specific degradation. Additionally, we propose a LoRA composition strategy
 53 based on degradation similarity. We calculate the similarity between degradation features extracted
 54 from degraded images and existing degradation types, utilizing it as weights for combining different
 55 LoRA experts. This strategy enables our method to be applicable for restoring mixed degradation
 56 images. Moreover, we conducted extensive experiments and compared our approach with several
 57 existing unified image restoration methods. The experimental results demonstrate that our method
 58 achieves superior performance in the restoration of various degradations and mixed degradations. Not
 59 only does our approach outperform existing methods in terms of distortion and perceptual metrics,
 60 but it also exhibits significant improvements in visual quality.

61 Our contributions can be summarized as follows:

- 62 • From the perspective of multi-domain transfer learning, we propose a novel universal image
 63 restoration framework based on multiple low-rank adaptations. It leverages the pre-trained
 64 generative model as the shared component for multi-degradation restoration and employs
 65 distinct LoRAs for multiple degradations to efficiently transfer to specific degradation
 66 restoration tasks.
- 67 • We introduce a LoRAs composition strategy based on the degradation similarity, which
 68 adaptively combines trained LoRAs and enables our model to be applicable for mixed
 69 degradation restoration.
- 70 • Through extensive experiments on multiple and mixed degradations, we demonstrate that the
 71 proposed universal image restoration method not only achieves higher fidelity and perceptual
 72 image quality but also has better generalization ability than other unified models.

73 2 Related Work

74 2.1 Image Restoration

75 **Specific Degradation Restoration.** According to degradation type, image restoration tasks are
76 categorized into different groups, including denoising, deblurring, inpainting, draining *etc.* Most
77 existing image restoration methods [2, 53, 16, 56, 5, 14] mainly address the issue with a single
78 degradation. Traditional approaches [27, 28, 7] have proposed image priors. While these priors can
79 be applied to different degraded images, their capability is limited. Due to the remarkable capability
80 of the deep neural network (DNN), numerous DNN-based methods [2, 53, 16] have been proposed
81 to tackle image restoration tasks. While DNN-based methods have made significant progress, they
82 struggle with multiple degradations and mixed degradations, since they typically require retraining
83 from scratch on data with the same degradation.

84 **Universal degradation restoration.** Increasing attention is currently focused on developing a
85 unified model to process multiple degradations. For example, AirNet[15] explores the degradation
86 representation in latent space for separating them in the restoration network. PromptIR[31] utilizes a
87 prompt block to extract the degradation-related features to improve the performance. Daclip-IR[20]
88 introduces the clip-based encoder to distinguish the type of degradation and extract the semantics
89 information from distorted images and embed them into a diffusion model to generate high-quality
90 images. Despite the advancements, these unified models still have limitations. They also require
91 retraining all parameters when unseen degradations arrive and have limited performance due to the
92 gradient conflict.

93 2.2 Low-Rank Adaptation

94 LoRA [11] is proposed to fine-tune large models by freezing the pre-trained weights and introducing
95 trainable low-rank matrices. This fine-tuning method leverages the property of "intrinsic dimension"
96 within neural networks, lowering the rank of additional matrices and making the re-training process
97 efficient. Concretely, given a weight matrices $W \in \mathbb{R}^{n \times m}$ in pre-trained model θ_p , two trainable
98 matrices $B \in \mathbb{R}^{n \times r}$ and $A \in \mathbb{R}^{r \times m}$ are inserted into the layer to represent the LoRA $\Delta W = BA$,
99 where r is the rank and satisfy $r \ll \min(n, m)$, the updated weights W' are calculated by

$$W' = W + \Delta W. \quad (1)$$

100 By applying LoRA in pre-trained models, numerous image generation methods [29, 13], show
101 superior performance in the field of image style and semantics concept transferring. Additionally,
102 fine-tuning methods like ControlNet [57], T2i-adapter [24] are also commonly employed in large-
103 scale pre-trained generative models such as Stable Diffusion [37], SDXL [30], and Imagen [38].

104 2.3 Mixture of Experts

105 Mixture of Experts (MoE) [41, 49, 48] is an effective approach to scale up neural network capacity to
106 improve performance. Specifically, MoE integrates multiple feed-forward networks into a transformer
107 block, where each feed-forward network is regarded as an expert. A gating function is introduced to
108 model the probability distribution across all experts in the MoE layer. The gating function is trainable
109 and determines the activation of specific experts within the MoE layer based on top-k values. Broadly
110 speaking, our framework aligns with the concept of MoE. However, unlike traditional MoE layers, we
111 employ the more efficient LoRA as experts in selected frozen layers and utilize a degradation-aware
112 router across all selected layers to uniformly activate experts, reducing learning complexity and
113 avoiding conflicts among different image restoration tasks on experts.

114 3 Methodology

115 3.1 Problem Definition

116 This paper seeks to develop a novel universal image restoration framework capable of handling
117 diverse forms of image degradation in the wild by fine-tuning the pre-trained generative model.
118 Consider a set of T image restoration tasks $D = \{D^k\}_{k=1}^T$, where $D^k = \{(x_i, y_i)\}_{i=1}^{n_k}$ is the training
119 dataset containing n_k images pairs of the k -th image degradation task. Within the set of tasks D ,

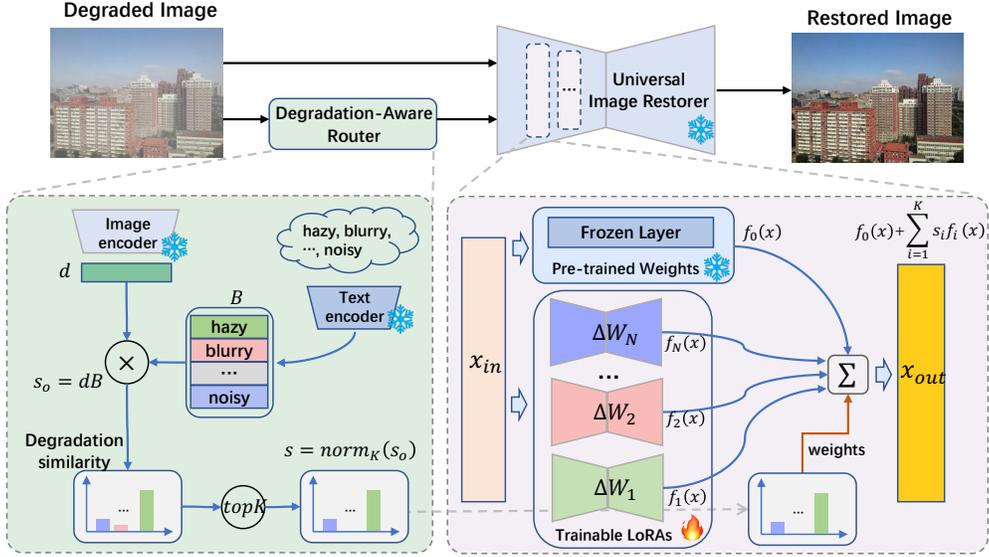


Figure 2: Overview of UIR-LoRA. UIR-LoRA consists of two components: a degradation-aware router and a universal image restorer. The router calculates degradation similarity in the latent space of CLIP, while the restorer utilizes the similarity provided by the router to combine LoRAs and frozen base model and restore images with multiple or mixed degradations.

120 each task D^k only has a specific type of image degradation, with no intersection between any two
 121 tasks. Given a pre-trained generative model θ_p with frozen parameters, our objective is to learn a
 122 set of composite $\{\theta_k\}_{k=1}^T$ to construct a unified model f_θ that performs well on multi-degradation
 123 restoration and mixed degradation restoration by transferring learning, where $\theta = \theta_p + \sum_{k=1}^T s_k \theta_k$
 124 and s_k represents the composite weight for θ_k . The trainable $\{\theta_k\}_{k=1}^T$ can be optimized through
 125 minimizing the overall image reconstruction loss:

$$L = E_{(x,y) \in D} l(f_\theta(x), y). \quad (2)$$

126 We will present how to design and optimize the trainable $\{\theta_k\}_{k=1}^T$ and construct the composite
 127 weights s in the next sections.

128 3.2 Overview of Universal Framework

129 Inspired by transferring learning, we introduce a novel universal image restoration framework based
 130 on multiple low-rank adaptations, named UIR-LoRA. Referring to Figure 2, our framework consists
 131 of two main components, namely degradation-aware router and universal image restorer, respectively.
 132 The degradation-aware router first extracts the degradation feature from input degraded images and
 133 then calculates the similarity probabilities s with existing degradations in the latent space of CLIP
 134 model [35, 20]. For the universal image restorer, it comprises a pre-trained generative model θ_p and
 135 T trainable LoRAs $\{\theta_k\}_{k=1}^T$. This design is primarily motivated by two considerations: firstly, the
 136 pre-trained generative model contains extensive image priors that are degradation-agnostic and can
 137 be shared across all types of degraded images. Secondly, each LoRA can independently capture
 138 specific characteristics of each degradation without gradient conflicts. In practice, the pre-trained
 139 SD-turbo [39] is employed as the frozen base model in our framework and each LoRA θ_k serves
 140 as an expert responsible for transferring the frozen base model to a specific degradation restoration
 141 task D^k . By adjusting the value of Top-K parameter within the degradation-aware router, different
 142 combinations of LoRAs in the universal image restorer can be activated, enabling the removal of a
 143 specific degradation and mixed degradation in multi-degraded scenarios.

144 3.3 Degradation-Aware Router

145 The Degradation-Aware Router is designed to provide the restorer with weights for LoRA combination
 146 based on degradation confidence. Following Daclip-ir [20], we utilized the pre-trained image encoder
 147 in CLIP [35] to obtain the degradation vector $d \in \mathbb{R}^{1 \times z}$ from the input degraded image x , where z is
 148 degradation length in latent space. Differing from Daclip-ir [20], we use the degradation vector and
 149 existing degradations to calculate the similarity, instead of directly embedding the degradation vector
 150 into the restoration network in Daclip-ir [20]. The existing degradations refer to the vocabulary bank
 151 of diverse degradation types that we introduce in the router, such as "noisy", "blurry" and "shadowed".
 152 This vocabulary bank is highly compact and flexible when adding new degradation types. Similarly,
 153 by applying the text encoder of CLIP [35], the vocabulary bank can be encoded into the degradation
 154 bank $B \in \mathbb{R}^{z \times T}$ in the latent space. As presented in Figure 2, the original degradation similarity
 155 $s_o \in \mathbb{R}^{1 \times T}$ is calculated by:

$$s_o = dB. \quad (3)$$

156 Building upon the original similarity, we adopt a more flexible and controllable Top-K strategy
 157 to modify s_o . Specifically, we select the Top-K largest values from the original similarity s_o , and
 158 normalize them to reallocate the weights for LoRAs. The reallocation process can be formulated as :

$$s = \frac{s_o \cdot M_K}{\sum s_o \cdot M_K}, \quad (4)$$

159 where M_K represents a binary mask with the same length as s_o , where it is 1 when the corresponding
 160 value in s_o is among the Top-K, otherwise it is 0. With a smaller value of K , the restorer activates
 161 fewer LoRAs, reducing its computational load. For instance, with $K = 1$, only the most similar
 162 LoRA is activated and it yields effective results when s is accurate, but performance noticeably
 163 declines with inaccurate s . Conversely, as K increases, the restorer exhibits higher tolerance to s and
 164 the combination of LoRAs allows it to handle mixed degradation.

165 3.4 Universal Image Restorer

166 Our universal image restorer consists of a pre-trained generative model θ_p and a set of LoRAs
 167 $\{\theta_k\}_{k=1}^T$. As illustrated in Figure 2, our universal image restorer takes the degraded image x and
 168 similarity s predicted by the degradation-aware router as inputs. It then activates relevant LoRAs
 169 based on s to recover the degraded image along with the frozen base model. Since one of our
 170 objectives is to ensure that each LoRA serves as an expert in processing a specific degradation, the
 171 number of LoRAs in the restorer aligns with the number of degradation types, T . In practice, we
 172 select multiple layers from the base model, For a selected layer W of the pre-trained base model, a
 173 sequence of trainable matrices $\{\Delta W_k\}_{k=1}^T$ are added into this layer, and the parameters of all chosen
 174 layers L form a complete LoRA $\theta_k = \{\Delta W_k^j\}_{j \in L}$. As previously explained, each LoRA is a unique
 175 expert responsible for a specific degradation. Drawing inspiration from Mixture of Expert (MoE), we
 176 aggregate the outputs of each expert rather than directly merging parameters in [11]. Therefore, given
 177 the input feature x_{in} of the current layer and the similarity s , the total output x_{out} of this modified
 178 layer can be expressed as

$$x_{out} = f_o(x_{in}) + \sum_{i=1}^K s_i f_i(x_{in}), \quad (5)$$

179 where $f_i(x_{in})$ denotes the result of i -th trainable matrix W_i , particularly $f_o(x_{in})$ is output of the
 180 frozen base layer. From the equation 5, it can be observed that the introduced LoRAs interact with the
 181 frozen base model at intermediate feature layers in our restorer. This interaction forces the restorer
 182 to leverage the image priors of the pre-trained generative model and eliminate degradation with the
 183 assistance of LoRAs. In contrast to employing stable diffusion [37] directly as a post-processing
 184 technique, our restorer yields results closer to the true scene without introducing inaccurate structural
 185 details. Since each W is implemented using two low-rank matrices like the formula 1, the total
 186 trainable parameters of our framework are much smaller than that of the pre-train generative model.

187 3.5 Training and Inference Procedure

188 During the training phase, for the efficient training of the universal image restorer, we ensure that
 189 each batch is sampled from the same degradation type D^k , and activate the corresponding LoRA θ^k

Table 1: Comparison of the restoration results over ten different datasets. The best results are marked in boldface.

Model	Distortion		Perceptual		Complexity	
	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	FID \downarrow	Param /M	Runtime /s
SwinIR [16]	23.37	0.731	0.354	104.37	15.8	0.66
NAFNet [2]	26.34	0.847	0.159	55.68	67.9	0.54
Restormer [53]	26.43	0.850	0.157	54.03	26.1	0.14
AirNet [15]	25.62	0.844	0.182	64.86	7.6	1.50
PromptIR [31]	27.14	0.859	0.147	48.26	35.6	1.19
IR-SDE [21]	23.64	0.754	0.167	49.18	36.2	5.07
DiffBIR [17]	21.01	0.618	0.263	91.03	363.2	5.95
Daclip-IR [20]	27.01	0.794	0.127	34.89	295.2	4.09
UIR-LoRA (Ours)	28.08	0.864	0.104	30.58	95.2	0.44

190 for training. Since the dataset D is organized by degradation type without overlap and each LoRA
 191 is assigned to handle each type of degradation correspondingly, the overall optimization process in
 192 equation 2 can be decomposed into independent optimization processes for each degradation. This
 193 design and training process circumvent task conflicts among multiple degradations and makes it
 194 possible to use suitable loss functions for the specific degradation. Due to the availability of accurate
 195 s during training and the use of pre-trained encoders from CLIP [35] and Daclip-ir [20] in our router,
 196 the router was not utilized during training.

197 In the inference phase, the similarity s is unknown and needs to be estimated from the degraded
 198 image. The estimated similarity s serves as a reference in our framework and can also be manually
 199 specified by users. Subsequently, our universal image restorer composite LoRAs and recovers the
 200 input image with the guidance of s .

201 4 Experiments

202 4.1 Experimental Setting

203 **Datasets.** We validate the effectiveness of our framework in multiple and mixed degradation scenarios.
 204 In the case of multiple degradations, we follow Daclip-IR [20] and construct a dataset using 10
 205 different single degradation datasets. Briefly, the composite dataset comprises a total of 52800 image
 206 pairs for training and 2490 image pairs for testing. The degradation types included are commonly
 207 encountered in image restoration, such as blur, noise, shadow, JPEG compression, and weather
 208 degradations. For mixed degradations, we utilize two degradation datasets, REDS [25] and LOLBlur
 209 [58]. In REDS, the images are distorted by JPEG compression and blur, and those images in LOLBlur
 210 have blur and low light. For more details about datasets in our experiments, please refer to **Appendix**.

211 **Metrics.** The objective of the image restoration task is to output images with enhanced visual quality
 212 while maintaining high fidelity to the original scene information. This differs from image generation
 213 tasks, which prioritize visual quality. Therefore, to thoroughly evaluate the effectiveness of our
 214 method, we utilize reference-based image quality assessment techniques from both distortion and
 215 perceptual perspectives, including PSNR, SSIM, and LPIPS, as well as FID.

216 **Comparison Methods.** In the experiments, we primarily compare with several state-of-the-art
 217 methods in image restoration, which fall into two categories: regression model and generative model.
 218 Regression models include NAFNet [2], Restormer [53], as well as AirNet [15] and PromptIR [31]
 219 proposed for multiple degradation restoration. DiffBIR [17], IR-SDE [21] and Daclip-IR [20] are
 220 generative models built upon the diffusion model [9].

221 4.2 Implementation Details

222 During the training, we adapt an AdamW optimizer to update the weights of trainable parameters in
 223 our model. Before training LoRA for specific degradation, we add skip-connections in the VAE of
 224 SD-turbo[39] like [29, 44] and train them with multiple degraded images. We set the initialization

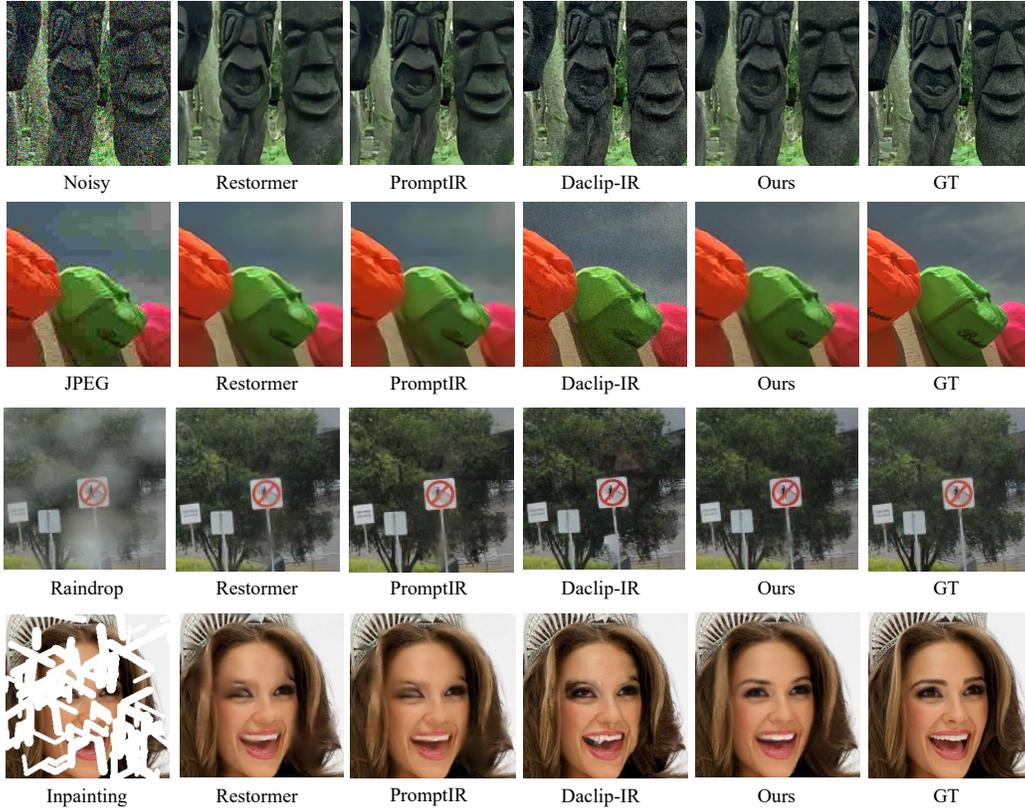


Figure 3: Qualitative comparison on multiple degraded images.

225 learning rate to $2e-4$ and decrease it with CosineAnnealingLR . We trained every LoRA for 80K
 226 iterations with batch size 8 and we keep the same hyper-parameters when training different LoRAs.
 227 The default rank of LoRAs in VAE and Unet is 4 and 8, respectively.

228 4.3 Multiple Image Restoration

229 For fair comparisons, all methods are trained and tested on the multiple degradation dataset. The
 230 results are presented in Table 1. We can find that our model, UIR-LoRA, considerably surpasses all
 231 compared image restoration approaches across four metrics. This indicates that our approach can
 232 balance generating clear structures and details while ensuring the restored images closely resemble the
 233 original information of the scene. The visual comparison results depicted in Figure 7 also confirm this
 234 assertion. Regression models such as NAFNet [2] and Restormer [53], lacking extensive image priors,
 235 tend to produce blurred and over-smoothed images, leading to inferior visual outcomes. Conversely,
 236 generative models Daclip-IR [20] excessively prioritize perceptual quality, yielding artifacts and
 237 noise that diverge from the actual scene information. Our approach integrates the strengths of both
 238 categories of methods, enabling strong performance in both distortion and perceptual aspects

239 4.4 Mixed Image Restoration

240 To evaluate the transferability of UIR-LoRA, we conduct some experiments on mixed degradation
 241 datasets from REDS[25] and LOLBlur [58]. Each image in these two datasets contains more than one
 242 type of degradation, like blur, jpeg compression, noise, and low light. We test the mixed degraded
 243 images using models trained on multiple degradations and set K to 2 in the router. As shown in
 244 Table 2, our method achieves superior results in both distortion and perceptual quality, particularly
 245 on the LOLBlur dataset. We also provide visual comparison results, as illustrated in Figure 4, our
 246 approach effectively enhances the low-light image compared to SOTA methods, highlighting its
 247 stronger transferability in the wild. More visual results can be found in **Appendix**.

Table 2: Comparison of the restoration results on mixed degradation datasets. The best results are marked in boldface.

Model	REDS				LOLBlur			
	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	FID \downarrow	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	FID \downarrow
SwinIR	21.53	0.676	0.449	116.80	10.06	0.320	0.619	124.52
NAFNet	25.06	0.721	0.412	122.12	10.57	0.397	0.477	85.77
Restormer	23.15	0.713	0.413	118.61	12.77	0.479	0.478	87.23
PromptIR	24.98	0.712	0.424	128.11	9.09	0.275	0.560	91.68
DiffBIR	20.70	0.598	0.377	122.76	9.86	0.288	0.611	125.41
Daclip-IR	24.30	0.699	0.337	95.29	14.52	0.599	0.358	68.10
UIR-LoRA	25.11	0.718	0.315	89.79	18.16	0.690	0.318	61.55

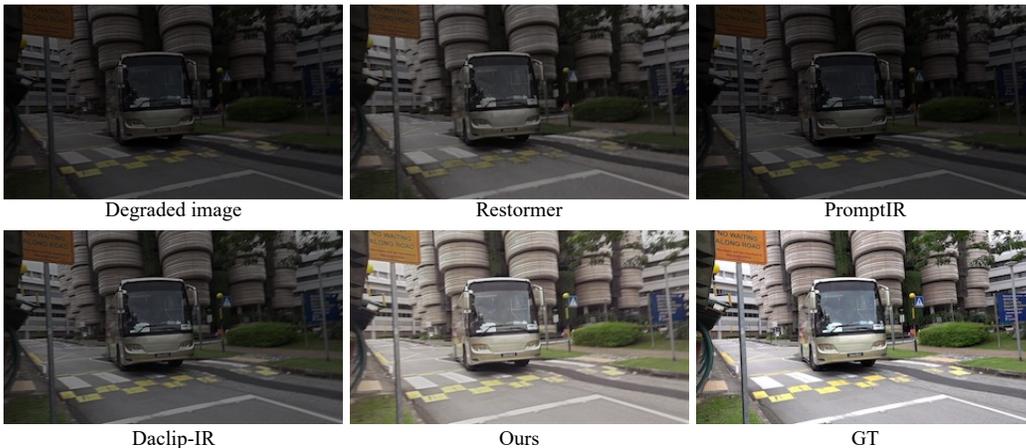


Figure 4: Qualitative comparison on multiple degraded images.

248 4.5 Ablation Study

249 **Complexity Analysis.** We compare model complexity with SOTA models. The comparison results
 250 are shown in Table 1, where we report the number of trainable parameters and the runtime for a
 251 256×256 image on an A100 GPU. The complexity of UIR-LoRA is comparable to regression models
 252 like NAFNet [2] and significantly more efficient than generative models like Daclip-IR [20].

253 **Effectiveness of Degradation-Aware Router.** The degradation-aware router plays a crucial role in
 254 determining which LoRAs are activated in the inference. To comprehensively demonstrate the impact
 255 of the router, we conduct experiments with different selection strategies. As illustrated in Table 3,
 256 we have five strategies: "random" indicates activating a LoRA at random, "average" denotes using
 257 average weights to activate all LoRAs, and "Top-1", "Top-2" and "All" correspond to setting K in the
 258 router to 1, 2, and 10, respectively. From the comparison of these results, we can see that the random
 259 and average strategies result in poorer performance while using the strategy based on degradation

Table 3: Impact of strategies in router

Strategy	Multiple Degradation				Mixed Degradation			
	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	FID \downarrow	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	FID \downarrow
Random	17.52	0.617	0.388	126.48	10.35	0.323	0.577	104.84
Average	17.62	0.617	0.370	129.06	9.28	0.277	0.549	106.05
Top-1	28.06	0.864	0.105	30.62	18.04	0.683	0.321	61.65
Top-2	28.05	0.864	0.105	30.60	18.16	0.690	0.318	61.55
All	28.05	0.864	0.105	30.61	18.16	0.691	0.318	61.58

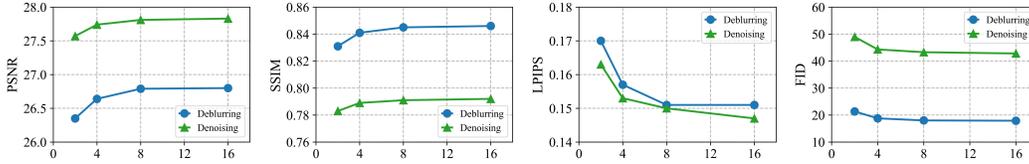


Figure 5: The impact of LoRA’s rank on deblurring and denoising tasks.

260 similarity achieves better outcomes. This suggests that the transferability between different types
 261 of degradation is limited and that specific parameters are needed to address their particularities.
 262 Furthermore, the selection of the K value also affects the model’s performance. When an image has
 263 only one type of degradation, a smaller K value can result in comparable performance with lower
 264 inference costs. However, for mixed degradations, a larger K value is required to handle the more
 265 complex situation.

266 **Impact of LoRA’s Rank.** Within our framework, LoRA is utilized to facilitate the transfer from
 267 the pre-trained generative model to the image restoration task. In order to investigate the impact of
 268 LoRA’s rank on the performance of image restoration, we conduct experiments using deblurring
 269 and denoising tasks chosen from ten distinct degradation categories. We set the initial rank to 2 and
 270 incrementally increase the value by a factor of 2. The performance changes are depicted in Figure 5.
 271 It is evident that as the rank grows, the restoration results improve in distortion and perceptual quality,
 272 and at the same time, the number of trainable parameters also increases. Once the rank value exceeds
 273 4, the performance improvement becomes progressively marginal. Therefore, we set the default rank to
 274 4 in our restorer to balance between performance and complexity.

Table 4: The accuracy of predicted degradation type

	PSNR↑	SSIM ↑	LPIPS ↓	FID ↓	Accuracy ↑
Original	26.66	0.839	0.159	18.72	91.6
Modified	26.87	0.842	0.155	18.42	99.2

275 **Impact of Predicted Degradation.**

276 The resizing operation on input images in CLIP models [20, 35] may lead to inaccurate predictions
 277 of degradation types, especially for blurry images. To reduce its negative impact on performance, we
 278 introduce a simple way that uses the degradation vector of the image crop without resizing to correct
 279 the potential error in the resized image. Table 4 is the comparison conducted on blurry images from
 280 GoPro dataset. It can be observed that our model with modified operation has higher accuracy and
 281 better performance for deblurring.

282 **5 Conclusion**

283 In this paper, we propose a universal image restoration framework based on multiple low-rank
 284 adaptation, named UIR-LoRA, from the perspective of multi-domain transfer learning. UIR-LoRA
 285 utilizes a pre-trained generative model as the frozen base model and transfers its abundant image
 286 priors to different image restoration tasks using the LoRA technique. Moreover, we introduce a
 287 LoRAs’ composition strategy based on the degradation similarity that allows UIR-LoRA applicable
 288 for multiple and mixed degradations in the wild. Extensive experiments on universal image restoration
 289 tasks demonstrate the effectiveness and better generalization capability of our proposed UIR-LoRA.

290 **6 Limitation and Discussion**

291 Although our UIR-LoRA has achieved remarkable performance in image restoration tasks under both
 292 multiple and mixed degradations, it still has limitations and problems for further exploration. For
 293 instance, adding new trainable parameters into the network for unseen degradations is unavoidable in
 294 image restoration tasks, although UIR-LoRA is already more efficient and flexible compared to other
 295 approaches.

References

- 296
- 297 [1] Eirikur Agustsson and Radu Timofte. Ntire 2017 challenge on single image super-resolution:
298 Dataset and study. In *Proceedings of the IEEE conference on computer vision and pattern
299 recognition workshops*, pages 126–135, 2017.
- 300 [2] Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, and Jian Sun. Simple baselines for image
301 restoration. In *European conference on computer vision*, pages 17–33. Springer, 2022.
- 302 [3] Liangyu Chen, Xin Lu, Jie Zhang, Xiaojie Chu, and Chengpeng Chen. Hinet: Half instance
303 normalization network for image restoration. In *Proceedings of the IEEE/CVF Conference on
304 Computer Vision and Pattern Recognition*, pages 182–192, 2021.
- 305 [4] Shoufa Chen, Chongjian Ge, Zhan Tong, Jiangliu Wang, Yibing Song, Jue Wang, and Ping Luo.
306 Adaptformer: Adapting vision transformers for scalable visual recognition. *Advances in Neural
307 Information Processing Systems*, 35:16664–16678, 2022.
- 308 [5] Xueyang Fu, Jiabin Huang, Xinghao Ding, Yinghao Liao, and John Paisley. Clearing the
309 skies: A deep network architecture for single-image rain removal. *IEEE Transactions on Image
310 Processing*, 26(6):2944–2956, 2017.
- 311 [6] Dong Gong, Jie Yang, Lingqiao Liu, Yanning Zhang, Ian Reid, Chunhua Shen, Anton Van
312 Den Hengel, and Qinfeng Shi. From motion blur to motion flow: A deep learning solution
313 for removing heterogeneous motion blur. In *Proceedings of the IEEE conference on computer
314 vision and pattern recognition*, pages 2319–2328, 2017.
- 315 [7] Shuhang Gu, Lei Zhang, Wangmeng Zuo, and Xiangchu Feng. Weighted nuclear norm min-
316 imization with application to image denoising. In *Proceedings of the IEEE conference on
317 computer vision and pattern recognition*, pages 2862–2869, 2014.
- 318 [8] Chunle Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, and
319 Runmin Cong. Zero-reference deep curve estimation for low-light image enhancement. In
320 *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages
321 1780–1789, 2020.
- 322 [9] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances
323 in neural information processing systems*, 33:6840–6851, 2020.
- 324 [10] Neil Houlsby, Andrei Giurgiu, Stanislaw Jastrzebski, Bruna Morrone, Quentin De Laroussilhe,
325 Andrea Gesmundo, Mona Attariyan, and Sylvain Gelly. Parameter-efficient transfer learning
326 for nlp. In *International conference on machine learning*, pages 2790–2799. PMLR, 2019.
- 327 [11] Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang,
328 Lu Wang, and Weizhu Chen. LoRA: Low-rank adaptation of large language models. In
329 *International Conference on Learning Representations*, 2022.
- 330 [12] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of gans for
331 improved quality, stability, and variation. *arXiv preprint arXiv:1710.10196*, 2017.
- 332 [13] Nupur Kumari, Bingliang Zhang, Richard Zhang, Eli Shechtman, and Jun-Yan Zhu. Multi-
333 concept customization of text-to-image diffusion. In *Proceedings of the IEEE/CVF Conference
334 on Computer Vision and Pattern Recognition*, pages 1931–1941, 2023.
- 335 [14] Orest Kupyn, Volodymyr Budzan, Mykola Mykhailych, Dmytro Mishkin, and Jiří Matas.
336 Deblurgan: Blind motion deblurring using conditional adversarial networks. In *Proceedings of
337 the IEEE conference on computer vision and pattern recognition*, pages 8183–8192, 2018.
- 338 [15] Boyun Li, Xiao Liu, Peng Hu, Zhongqin Wu, Jiancheng Lv, and Xi Peng. All-in-one image
339 restoration for unknown corruption. In *Proceedings of the IEEE/CVF Conference on Computer
340 Vision and Pattern Recognition*, pages 17452–17462, 2022.
- 341 [16] Jingyun Liang, Jiezhong Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. Swinir:
342 Image restoration using swin transformer. In *Proceedings of the IEEE/CVF international
343 conference on computer vision*, pages 1833–1844, 2021.

- 344 [17] Xinqi Lin, Jingwen He, Ziyang Chen, Zhaoyang Lyu, Ben Fei, Bo Dai, Wanli Ouyang, Yu Qiao,
345 and Chao Dong. Diffbir: Towards blind image restoration with generative diffusion prior. *arXiv*
346 *preprint arXiv:2308.15070*, 2023.
- 347 [18] Yun-Fu Liu, Da-Wei Jaw, Shih-Chia Huang, and Jenq-Neng Hwang. Desnownet: Context-aware
348 deep network for snow removal. *IEEE Transactions on Image Processing*, 27(6):3064–3073,
349 2018.
- 350 [19] Andreas Lugmayr, Martin Danelljan, Andres Romero, Fisher Yu, Radu Timofte, and Luc
351 Van Gool. Repaint: Inpainting using denoising diffusion probabilistic models. In *Proceedings*
352 *of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11461–11471,
353 2022.
- 354 [20] Ziwei Luo, Fredrik K Gustafsson, Zheng Zhao, Jens Sjölund, and Thomas B Schön. Controlling
355 vision-language models for universal image restoration. *arXiv preprint arXiv:2310.01018*, 2023.
- 356 [21] Ziwei Luo, Fredrik K Gustafsson, Zheng Zhao, Jens Sjölund, and Thomas B Schön. Image
357 restoration with mean-reverting stochastic differential equations. *International Conference on*
358 *Machine Learning*, 2023.
- 359 [22] Long Ma, Tengyu Ma, Risheng Liu, Xin Fan, and Zhongxuan Luo. Toward fast, flexible, and
360 robust low-light image enhancement. In *Proceedings of the IEEE/CVF conference on computer*
361 *vision and pattern recognition*, pages 5637–5646, 2022.
- 362 [23] David Martin, Charless Fowlkes, Doron Tal, and Jitendra Malik. A database of human seg-
363 mented natural images and its application to evaluating segmentation algorithms and measuring
364 ecological statistics. In *Proceedings of the IEEE/CVF International Conference on Computer*
365 *Vision*, pages 416–423, 2001.
- 366 [24] Chong Mou, Xintao Wang, Liangbin Xie, Yanze Wu, Jian Zhang, Zhongang Qi, and Ying Shan.
367 T2i-adapter: Learning adapters to dig out more controllable ability for text-to-image diffusion
368 models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages
369 4296–4304, 2024.
- 370 [25] Seungjun Nah, Sungyong Baik, Seokil Hong, Gyeongsik Moon, Sanghyun Son, Radu Timofte,
371 and Kyoung Mu Lee. Ntire 2019 challenge on video deblurring and super-resolution: Dataset
372 and study. In *Proceedings of the IEEE/CVF conference on computer vision and pattern*
373 *recognition workshops*, pages 0–0, 2019.
- 374 [26] Seungjun Nah, Tae Hyun Kim, and Kyoung Mu Lee. Deep multi-scale convolutional neural
375 network for dynamic scene deblurring. In *Proceedings of the IEEE Conference on Computer*
376 *Vision and Pattern Recognition*, pages 3883–3891, 2017.
- 377 [27] Jinshan Pan, Zhe Hu, Zhixun Su, and Ming-Hsuan Yang. Deblurring text images via l0-
378 regularized intensity and gradient prior. In *Proceedings of the IEEE Conference on Computer*
379 *Vision and Pattern Recognition*, pages 2901–2908, 2014.
- 380 [28] Jinshan Pan, Deqing Sun, Hanspeter Pfister, and Ming-Hsuan Yang. Blind image deblurring
381 using dark channel prior. In *Proceedings of the IEEE conference on computer vision and pattern*
382 *recognition*, pages 1628–1636, 2016.
- 383 [29] Gaurav Parmar, Taesung Park, Srinivasa Narasimhan, and Jun-Yan Zhu. One-step image
384 translation with text-to-image models. *arXiv preprint arXiv:2403.12036*, 2024.
- 385 [30] Dustin Podell, Zion English, Kyle Lacey, Andreas Blattmann, Tim Dockhorn, Jonas Müller, Joe
386 Penna, and Robin Rombach. Sdxl: Improving latent diffusion models for high-resolution image
387 synthesis. *arXiv preprint arXiv:2307.01952*, 2023.
- 388 [31] Vaishnav Potlapalli, Syed Waqas Zamir, Salman Khan, and Fahad Khan. Promptir: Prompting
389 for all-in-one image restoration. In *Thirty-seventh Conference on Neural Information Processing*
390 *Systems*, 2023.

- 391 [32] Rui Qian, Robby T Tan, Wenhan Yang, Jiajun Su, and Jiaying Liu. Attentive generative
392 adversarial network for raindrop removal from a single image. In *Proceedings of the IEEE*
393 *conference on computer vision and pattern recognition*, pages 2482–2491, 2018.
- 394 [33] Xu Qin, Zhilin Wang, Yuanchao Bai, Xiaodong Xie, and Huizhu Jia. Ffa-net: Feature fusion
395 attention network for single image dehazing. In *Proceedings of the AAAI conference on artificial*
396 *intelligence*, pages 11908–11915, 2020.
- 397 [34] Liangqiong Qu, Jiandong Tian, Shengfeng He, Yandong Tang, and Rynson WH Lau. Deshad-
398 ownet: A multi-context embedding deep network for shadow removal. In *Proceedings of the*
399 *IEEE conference on computer vision and pattern recognition*, pages 4067–4075, 2017.
- 400 [35] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal,
401 Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual
402 models from natural language supervision. In *International conference on machine learning*,
403 pages 8748–8763. PMLR, 2021.
- 404 [36] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark
405 Chen, and Ilya Sutskever. Zero-shot text-to-image generation. In *International conference on*
406 *machine learning*, pages 8821–8831. Pmlr, 2021.
- 407 [37] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-
408 resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF*
409 *conference on computer vision and pattern recognition*, pages 10684–10695, 2022.
- 410 [38] Chitwan Saharia, William Chan, Saurabh Saxena, Lala Li, Jay Whang, Emily L Denton,
411 Kamyar Ghasemipour, Raphael Gontijo Lopes, Burcu Karagol Ayan, Tim Salimans, et al.
412 Photorealistic text-to-image diffusion models with deep language understanding. *Advances in*
413 *neural information processing systems*, 35:36479–36494, 2022.
- 414 [39] Axel Sauer, Dominik Lorenz, Andreas Blattmann, and Robin Rombach. Adversarial diffusion
415 distillation. *arXiv preprint arXiv:2311.17042*, 2023.
- 416 [40] Ozan Sener and Vladlen Koltun. Multi-task learning as multi-objective optimization. *Advances*
417 *in neural information processing systems*, 31, 2018.
- 418 [41] Noam Shazeer, Azalia Mirhoseini, Krzysztof Maziarz, Andy Davis, Quoc Le, Geoffrey Hinton,
419 and Jeff Dean. Outrageously large neural networks: The sparsely-gated mixture-of-experts
420 layer. *arXiv preprint arXiv:1701.06538*, 2017.
- 421 [42] H Sheikh. Live image quality assessment database release 2. <http://live.ece.utexas.edu/research/quality>, 2005.
- 423 [43] Radu Timofte, Eirikur Agustsson, Luc Van Gool, Ming-Hsuan Yang, and Lei Zhang. Ntire 2017
424 challenge on single image super-resolution: Methods and results. In *Proceedings of the IEEE*
425 *conference on computer vision and pattern recognition workshops*, pages 114–125, 2017.
- 426 [44] Jianyi Wang, Zongsheng Yue, Shangchen Zhou, Kelvin CK Chan, and Chen Change Loy. Ex-
427 ploiting diffusion prior for real-world image super-resolution. *arXiv preprint arXiv:2305.07015*,
428 2023.
- 429 [45] Yinglong Wang, Chao Ma, and Jianzhuang Liu. Smartassign: Learning a smart knowledge
430 assignment strategy for deraining and desnowing. In *Proceedings of the IEEE/CVF Conference*
431 *on Computer Vision and Pattern Recognition*, pages 3677–3686, 2023.
- 432 [46] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for
433 low-light enhancement. *arXiv preprint arXiv:1808.04560*, 2018.
- 434 [47] Jay Whang, Mauricio Delbracio, Hossein Talebi, Chitwan Saharia, Alexandros G Dimakis,
435 and Peyman Milanfar. Deblurring via stochastic refinement. In *Proceedings of the IEEE/CVF*
436 *Conference on Computer Vision and Pattern Recognition*, pages 16293–16303, 2022.
- 437 [48] Xun Wu, Shaohan Huang, and Furu Wei. Mole: Mixture of lora experts. In *The Twelfth*
438 *International Conference on Learning Representations*, 2023.

- 439 [49] Yuan Xie, Shaohan Huang, Tianyu Chen, and Furu Wei. Moec: Mixture of expert clusters. In
440 *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 13807–13815,
441 2023.
- 442 [50] Wenhan Yang, Robby T Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, and Shuicheng Yan. Deep
443 joint rain detection and removal from a single image. In *Proceedings of the IEEE conference on*
444 *computer vision and pattern recognition*, pages 1357–1366, 2017.
- 445 [51] Wenhan Yang, Robby T Tan, Shiqi Wang, Yuming Fang, and Jiaying Liu. Single image deraining:
446 From model-based to data-driven and beyond. *IEEE Transactions on pattern analysis and*
447 *machine intelligence*, 43(11):4059–4077, 2020.
- 448 [52] Tianhe Yu, Saurabh Kumar, Abhishek Gupta, Sergey Levine, Karol Hausman, and Chelsea Finn.
449 Gradient surgery for multi-task learning. *Advances in Neural Information Processing Systems*,
450 33:5824–5836, 2020.
- 451 [53] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and
452 Ming-Hsuan Yang. Restormer: Efficient transformer for high-resolution image restoration. In
453 *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages
454 5728–5739, 2022.
- 455 [54] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-
456 Hsuan Yang, and Ling Shao. Multi-stage progressive image restoration. In *Proceedings of the*
457 *IEEE/CVF conference on computer vision and pattern recognition*, pages 14821–14831, 2021.
- 458 [55] Cheng Zhang, Yu Zhu, Qingsen Yan, Jinqiu Sun, and Yanning Zhang. All-in-one multi-
459 degradation image restoration network via hierarchical degradation representation. In *Proceed-*
460 *ings of the 31st ACM International Conference on Multimedia*, pages 2285–2293, 2023.
- 461 [56] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian
462 denoiser: Residual learning of deep cnn for image denoising. *IEEE transactions on image*
463 *processing*, 26(7):3142–3155, 2017.
- 464 [57] Lvmin Zhang, Anyi Rao, and Maneesh Agrawala. Adding conditional control to text-to-image
465 diffusion models. In *Proceedings of the IEEE/CVF International Conference on Computer*
466 *Vision*, pages 3836–3847, 2023.
- 467 [58] Shangchen Zhou, Chongyi Li, and Chen Change Loy. Lednet: Joint low-light enhancement and
468 deblurring in the dark. In *European conference on computer vision*, pages 573–589. Springer,
469 2022.

470 **A Appendix**

471 **A.1 More Details about Datasets**

472 For multiple degradations, we follow Daclip-IR [20] to construct the dataset, which includes a total
 473 of ten distinct degradation types: blurry, hazy, JPEG-compression, low-light, noisy, raindrop, rainy,
 474 shadowed, snowy, and inpainting. The data sources and data splits for each degradation type are
 475 illustrated in Table 5.

Table 5: Details of the datasets with ten different image degradation types

Dataset	Train		Test	
	Sources	Num	Sources	Num
Blurry	GoPro[26]	2 103	GoPro	1 111
Hazy	RESIDE-6k[33]	6 000	RESIDE-6k	1 000
JPEG	DIV2K[1] and Flickr2K[43]	3 550	LIVE1[42]	29
Low-light	LOL[46]	485	LOL	15
Noisy	DIV2K and Flickr2K	3 550	CBSD68[23]	68
Raindrop	RainDrop[32]	861	RainDrop	58
Rainy	Rain100H[50]	1 800	Rain100H	100
Shadowed	SRD[34]	2 680	SRD	408
Snowy	Snow100K-L[18]	1 872	Snow100K-L	601
Inpainting	CelebaHQ[12]	29 900	CelebaHQ and RePaint[19]	100

476 For mixed degradations, we utilize images from REDS[25] and LOLBlur[58] to evaluate the trans-
 477 ferability of models. We sample 60 images from REDS and 200 images from LOLBlur dataset for
 478 testing. The degraded images from REDS dataset feature a variety of realistic scenes and objects,
 479 which suffer from both motion blurs and compression. And the images from LOLBlur dataset cover
 480 a range of real-world dynamic dark scenarios with mixed degradation of low light and blurs.

481 **A.2 More Visual Results**

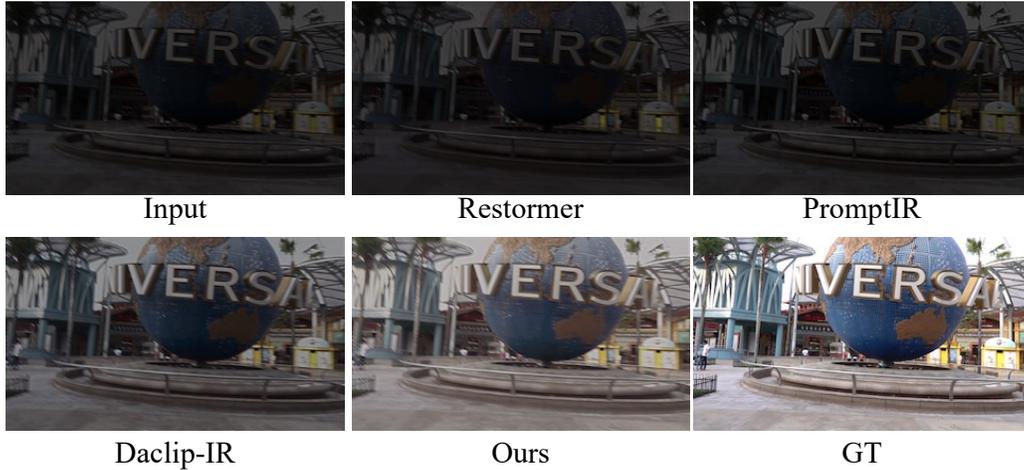


Figure 6: Qualitative comparison on mixed degraded images from LOLBlur dataset.

482 **A.3 Details about Metrics on Multiple Dagradaation**

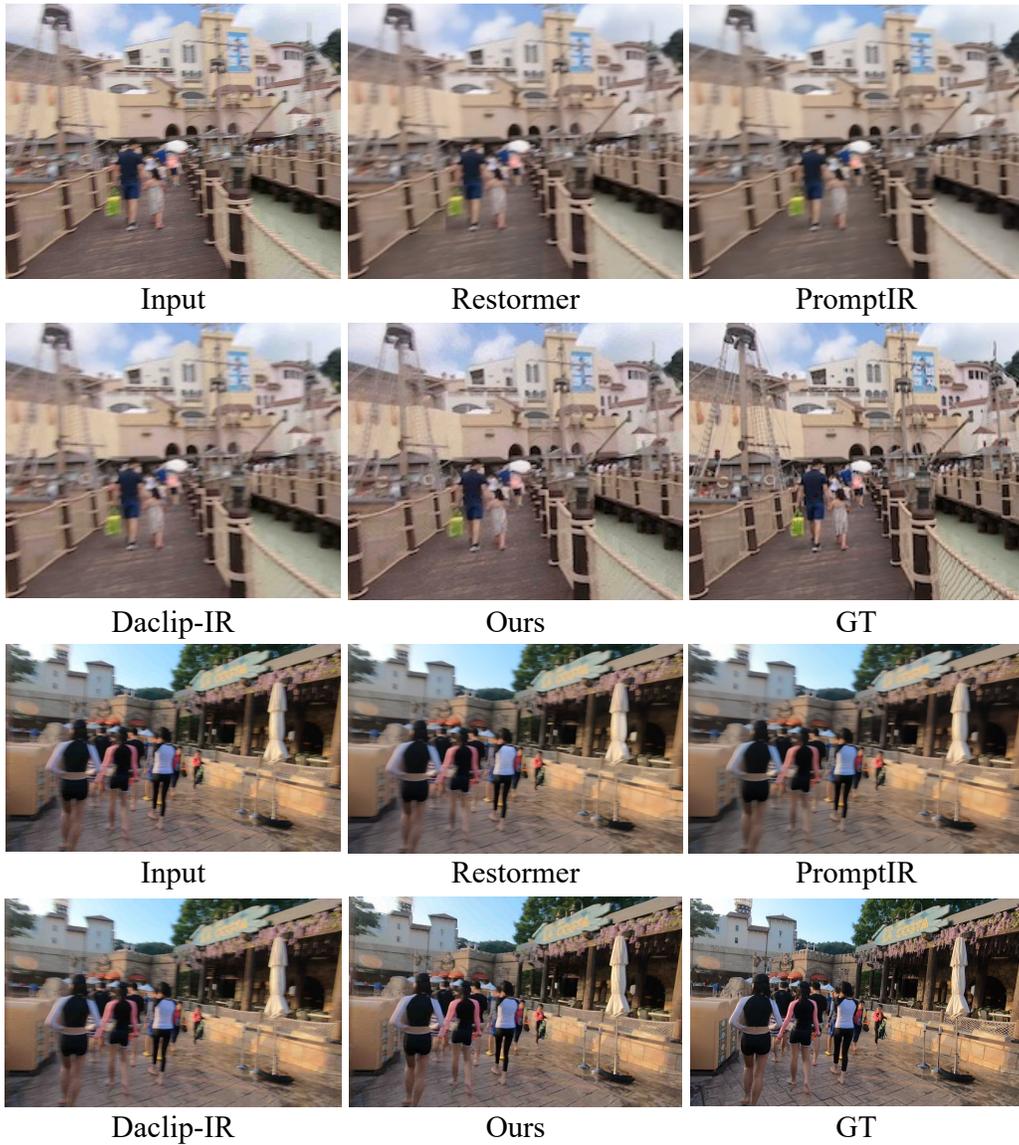


Figure 7: Qualitative comparison on mixed degraded images from REDS dataset.

Table 6: Comparison of the restoration results over ten different datasets on *PSNR*

	Blurry	Hazy	JPEG	Low-light	Noisy	Raindrop	Rainy	Shadowed	Snowy	Inpainting	Average
SwinIR	24.49	23.49	24.44	19.59	25.13	24.64	22.07	23.97	21.86	24.05	23.37
NAFNet	26.12	24.05	26.81	22.16	27.16	30.67	27.32	24.16	25.94	29.03	26.34
Restormer	26.34	23.75	26.90	22.17	27.25	30.85	27.91	23.33	25.98	29.88	26.43
AirNet	26.25	23.56	26.98	14.24	27.51	30.68	28.45	23.48	24.87	30.15	25.62
PromptIR	26.50	25.19	26.95	23.14	27.56	31.35	29.24	24.06	27.23	30.22	27.14
IR-SDE	24.13	17.44	24.21	16.07	24.82	28.49	26.64	22.18	24.70	27.56	23.64
DiffBIR	22.79	20.52	22.39	16.96	21.60	23.22	21.04	22.27	20.63	18.77	21.01
Daclip-IR	27.03	29.53	23.70	22.09	24.36	30.81	29.41	27.27	26.83	28.94	27.01
Ours	26.66	30.28	27.15	22.45	27.74	30.51	28.26	28.63	28.09	30.88	28.06

Table 7: Comparison of the restoration results over ten different datasets on *SSIM*

	Blurry	Hazy	JPEG	Low-light	Noisy	Raindrop	Rainy	Shadowed	Snowy	Inpainting	Average
SwinIR	0.758	0.848	0.734	0.735	0.690	0.758	0.623	0.757	0.665	0.743	0.731
NAFNet	0.804	0.926	0.780	0.809	0.768	0.924	0.848	0.839	0.869	0.901	0.847
Restormer	0.811	0.915	0.781	0.815	0.762	0.928	0.862	0.836	0.877	0.912	0.850
AirNet	0.805	0.916	0.783	0.781	0.769	0.926	0.867	0.832	0.846	0.911	0.844
PromptIR	0.815	0.933	0.784	0.829	0.774	0.931	0.876	0.842	0.887	0.918	0.859
IR-SDE	0.730	0.832	0.615	0.719	0.640	0.822	0.808	0.667	0.828	0.876	0.754
DiffBIR	0.695	0.761	0.607	0.665	0.395	0.682	0.573	0.568	0.566	0.678	0.618
Daclip-IR	0.810	0.931	0.532	0.796	0.579	0.882	0.854	0.811	0.854	0.894	0.794
Ours	0.839	0.962	0.782	0.826	0.789	0.908	0.857	0.862	0.893	0.916	0.864

Table 8: Comparison of the restoration results over ten different datasets on *LPIPS*

	Blurry	Hazy	JPEG	Low-light	Noisy	Raindrop	Rainy	Shadowed	Snowy	Inpainting	Average
SwinIR	0.347	0.180	0.392	0.362	0.439	0.353	0.481	0.335	0.388	0.265	0.354
NAFNet	0.284	0.043	0.303	0.158	0.216	0.082	0.180	0.138	0.096	0.085	0.159
Restormer	0.282	0.054	0.300	0.156	0.215	0.083	0.170	0.145	0.095	0.072	0.157
AirNet	0.279	0.063	0.302	0.321	0.264	0.095	0.163	0.145	0.112	0.071	0.182
PromptIR	0.267	0.051	0.269	0.140	0.230	0.078	0.147	0.143	0.082	0.068	0.147
IR-SDE	0.198	0.168	0.246	0.185	0.232	0.113	0.142	0.223	0.107	0.065	0.167
DiffBIR	0.269	0.158	0.244	0.273	0.442	0.187	0.309	0.261	0.236	0.246	0.263
Daclip-IR	0.140	0.037	0.317	0.114	0.272	0.068	0.085	0.118	0.072	0.047	0.127
Ours	0.159	0.021	0.204	0.126	0.153	0.048	0.112	0.103	0.070	0.056	0.105

Table 9: Comparison of the restoration results over ten different datasets on *FID*

	Blurry	Hazy	JPEG	Low-light	Noisy	Raindrop	Rainy	Shadowed	Snowy	Inpainting	Average
SwinIR	53.84	35.43	83.33	156.55	126.87	111.64	186.60	70.22	79.51	139.71	104.37
NAFNet	42.99	15.73	71.88	73.94	82.08	56.43	86.35	47.32	35.76	44.32	55.68
Restormer	39.08	15.34	72.68	78.22	87.14	50.97	78.16	48.33	33.45	36.96	54.03
AirNet	41.23	21.91	78.56	154.2	93.89	52.71	72.07	64.13	64.13	32.93	64.86
PromptIR	36.5	10.85	73.02	67.15	84.51	44.48	61.88	43.24	28.29	32.69	48.26
IR-SDE	29.79	23.16	61.85	66.42	79.38	50.22	63.07	50.71	34.63	32.61	49.18
DiffBIR	37.84	31.83	66.07	150.96	127.27	81.27	133.60	74.09	53.62	154.02	91.03
Daclip-IR	14.13	5.66	42.05	52.23	64.71	38.91	52.78	25.48	27.26	25.73	34.89
Ours	18.72	5.92	37.23	62.21	44.36	23.77	44.30	23.39	22.77	23.50	30.62

Table 10: Impact of rank in LoRAs

Rank	Deblurring				Denoising			
	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	FID \downarrow	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	FID \downarrow
2	26.35	0.831	0.170	21.35	27.57	0.783	0.163	48.98
4	26.64	0.841	0.157	18.79	27.74	0.789	0.153	44.32
8	26.79	0.845	0.151	18.01	27.81	0.791	0.150	43.29
16	26.80	0.846	0.151	17.90	27.83	0.792	0.147	42.82

483 **NeurIPS Paper Checklist**

484 **1. Claims**

485 Question: Do the main claims made in the abstract and introduction accurately reflect the
486 paper's contributions and scope?

487 Answer: [\[Yes\]](#)

488 Justification: The main claims made in the abstract and introduction accurately reflect the
489 paper's contributions and scope.

490 Guidelines:

- 491 • The answer NA means that the abstract and introduction do not include the claims
492 made in the paper.
- 493 • The abstract and/or introduction should clearly state the claims made, including the
494 contributions made in the paper and important assumptions and limitations. A No or
495 NA answer to this question will not be perceived well by the reviewers.
- 496 • The claims made should match theoretical and experimental results, and reflect how
497 much the results can be expected to generalize to other settings.
- 498 • It is fine to include aspirational goals as motivation as long as it is clear that these goals
499 are not attained by the paper.

500 **2. Limitations**

501 Question: Does the paper discuss the limitations of the work performed by the authors?

502 Answer: [\[Yes\]](#)

503 Justification: The paper does discuss the limitations of the work in Sections 6.

504 Guidelines:

- 505 • The answer NA means that the paper has no limitation while the answer No means that
506 the paper has limitations, but those are not discussed in the paper.
- 507 • The authors are encouraged to create a separate "Limitations" section in their paper.
- 508 • The paper should point out any strong assumptions and how robust the results are to
509 violations of these assumptions (e.g., independence assumptions, noiseless settings,
510 model well-specification, asymptotic approximations only holding locally). The authors
511 should reflect on how these assumptions might be violated in practice and what the
512 implications would be.
- 513 • The authors should reflect on the scope of the claims made, e.g., if the approach was
514 only tested on a few datasets or with a few runs. In general, empirical results often
515 depend on implicit assumptions, which should be articulated.
- 516 • The authors should reflect on the factors that influence the performance of the approach.
517 For example, a facial recognition algorithm may perform poorly when image resolution
518 is low or images are taken in low lighting. Or a speech-to-text system might not be
519 used reliably to provide closed captions for online lectures because it fails to handle
520 technical jargon.
- 521 • The authors should discuss the computational efficiency of the proposed algorithms
522 and how they scale with dataset size.
- 523 • If applicable, the authors should discuss possible limitations of their approach to
524 address problems of privacy and fairness.
- 525 • While the authors might fear that complete honesty about limitations might be used by
526 reviewers as grounds for rejection, a worse outcome might be that reviewers discover
527 limitations that aren't acknowledged in the paper. The authors should use their best
528 judgment and recognize that individual actions in favor of transparency play an impor-
529 tant role in developing norms that preserve the integrity of the community. Reviewers
530 will be specifically instructed to not penalize honesty concerning limitations.

531 **3. Theory Assumptions and Proofs**

532 Question: For each theoretical result, does the paper provide the full set of assumptions and
533 a complete (and correct) proof?

534 Answer: [\[Yes\]](#)

535 Justification: The paper provides the full set of assumptions and a complete (and correct)
536 proof in Sections 3.

537 Guidelines:

- 538 • The answer NA means that the paper does not include theoretical results.
- 539 • All the theorems, formulas, and proofs in the paper should be numbered and cross-
540 referenced.
- 541 • All assumptions should be clearly stated or referenced in the statement of any theorems.
- 542 • The proofs can either appear in the main paper or the supplemental material, but if
543 they appear in the supplemental material, the authors are encouraged to provide a short
544 proof sketch to provide intuition.
- 545 • Inversely, any informal proof provided in the core of the paper should be complemented
546 by formal proofs provided in appendix or supplemental material.
- 547 • Theorems and Lemmas that the proof relies upon should be properly referenced.

548 4. Experimental Result Reproducibility

549 Question: Does the paper fully disclose all the information needed to reproduce the main ex-
550 perimental results of the paper to the extent that it affects the main claims and/or conclusions
551 of the paper (regardless of whether the code and data are provided or not)?

552 Answer: [Yes]

553 Justification: The paper provides a comprehensive description of the experimental setting in
554 Sections 4.1 and implementation details in Sections 4.2, which are crucial for reproducing
555 the main results.

556 Guidelines:

- 557 • The answer NA means that the paper does not include experiments.
- 558 • If the paper includes experiments, a No answer to this question will not be perceived
559 well by the reviewers: Making the paper reproducible is important, regardless of
560 whether the code and data are provided or not.
- 561 • If the contribution is a dataset and/or model, the authors should describe the steps taken
562 to make their results reproducible or verifiable.
- 563 • Depending on the contribution, reproducibility can be accomplished in various ways.
564 For example, if the contribution is a novel architecture, describing the architecture fully
565 might suffice, or if the contribution is a specific model and empirical evaluation, it may
566 be necessary to either make it possible for others to replicate the model with the same
567 dataset, or provide access to the model. In general, releasing code and data is often
568 one good way to accomplish this, but reproducibility can also be provided via detailed
569 instructions for how to replicate the results, access to a hosted model (e.g., in the case
570 of a large language model), releasing of a model checkpoint, or other means that are
571 appropriate to the research performed.
- 572 • While NeurIPS does not require releasing code, the conference does require all submis-
573 sions to provide some reasonable avenue for reproducibility, which may depend on the
574 nature of the contribution. For example
 - 575 (a) If the contribution is primarily a new algorithm, the paper should make it clear how
576 to reproduce that algorithm.
 - 577 (b) If the contribution is primarily a new model architecture, the paper should describe
578 the architecture clearly and fully.
 - 579 (c) If the contribution is a new model (e.g., a large language model), then there should
580 either be a way to access this model for reproducing the results or a way to reproduce
581 the model (e.g., with an open-source dataset or instructions for how to construct
582 the dataset).
 - 583 (d) We recognize that reproducibility may be tricky in some cases, in which case
584 authors are welcome to describe the particular way they provide for reproducibility.
585 In the case of closed-source models, it may be that access to the model is limited in
586 some way (e.g., to registered users), but it should be possible for other researchers
587 to have some path to reproducing or verifying the results.

588 5. Open access to data and code

589 Question: Does the paper provide open access to the data and code, with sufficient instruc-
590 tions to faithfully reproduce the main experimental results, as described in supplemental
591 material?

592 Answer: [No]

593 Justification: Upon acceptance of the paper, we will release the code under an open-source
594 license, which will allow the community to access and verify the experimental results.

595 Guidelines:

- 596 • The answer NA means that paper does not include experiments requiring code.
- 597 • Please see the NeurIPS code and data submission guidelines ([https://nips.cc/
598 public/guides/CodeSubmissionPolicy](https://nips.cc/public/guides/CodeSubmissionPolicy)) for more details.
- 599 • While we encourage the release of code and data, we understand that this might not be
600 possible, so “No” is an acceptable answer. Papers cannot be rejected simply for not
601 including code, unless this is central to the contribution (e.g., for a new open-source
602 benchmark).
- 603 • The instructions should contain the exact command and environment needed to run to
604 reproduce the results. See the NeurIPS code and data submission guidelines ([https:
605 //nips.cc/public/guides/CodeSubmissionPolicy](https://nips.cc/public/guides/CodeSubmissionPolicy)) for more details.
- 606 • The authors should provide instructions on data access and preparation, including how
607 to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- 608 • The authors should provide scripts to reproduce all experimental results for the new
609 proposed method and baselines. If only a subset of experiments are reproducible, they
610 should state which ones are omitted from the script and why.
- 611 • At submission time, to preserve anonymity, the authors should release anonymized
612 versions (if applicable).
- 613 • Providing as much information as possible in supplemental material (appended to the
614 paper) is recommended, but including URLs to data and code is permitted.

615 6. Experimental Setting/Details

616 Question: Does the paper specify all the training and test details (e.g., data splits, hyper-
617 parameters, how they were chosen, type of optimizer, etc.) necessary to understand the
618 results?

619 Answer: [Yes]

620 Justification: The paper provides detailed information on all training and test aspects,
621 including datasets, metrics, comparison methods, hyperparameters, the type of optimizer
622 used, and other relevant details necessary to understand the results in Sections 4.1 and 4.2.

623 Guidelines:

- 624 • The answer NA means that the paper does not include experiments.
- 625 • The experimental setting should be presented in the core of the paper to a level of detail
626 that is necessary to appreciate the results and make sense of them.
- 627 • The full details can be provided either with the code, in appendix, or as supplemental
628 material.

629 7. Experiment Statistical Significance

630 Question: Does the paper report error bars suitably and correctly defined or other appropriate
631 information about the statistical significance of the experiments?

632 Answer: [Yes]

633 Justification: The paper reports error bars appropriately and includes correctly defined infor-
634 mation regarding the statistical significance of the experiments, ensuring the transparency
635 and reliability of the results.

636 Guidelines:

- 637 • The answer NA means that the paper does not include experiments.
- 638 • The authors should answer "Yes" if the results are accompanied by error bars, confi-
639 dence intervals, or statistical significance tests, at least for the experiments that support
640 the main claims of the paper.

- 641 • The factors of variability that the error bars are capturing should be clearly stated (for
- 642 example, train/test split, initialization, random drawing of some parameter, or overall
- 643 run with given experimental conditions).
- 644 • The method for calculating the error bars should be explained (closed form formula,
- 645 call to a library function, bootstrap, etc.)
- 646 • The assumptions made should be given (e.g., Normally distributed errors).
- 647 • It should be clear whether the error bar is the standard deviation or the standard error
- 648 of the mean.
- 649 • It is OK to report 1-sigma error bars, but one should state it. The authors should
- 650 preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis
- 651 of Normality of errors is not verified.
- 652 • For asymmetric distributions, the authors should be careful not to show in tables or
- 653 figures symmetric error bars that would yield results that are out of range (e.g. negative
- 654 error rates).
- 655 • If error bars are reported in tables or plots, The authors should explain in the text how
- 656 they were calculated and reference the corresponding figures or tables in the text.

657 8. Experiments Compute Resources

658 Question: For each experiment, does the paper provide sufficient information on the com-
 659 puter resources (type of compute workers, memory, time of execution) needed to reproduce
 660 the experiments?

661 Answer: [Yes]

662 Justification: The paper provides detailed information regarding the computer resources
 663 used in Sections 4.1, and time of execution in Table 1.

664 Guidelines:

- 665 • The answer NA means that the paper does not include experiments.
- 666 • The paper should indicate the type of compute workers CPU or GPU, internal cluster,
- 667 or cloud provider, including relevant memory and storage.
- 668 • The paper should provide the amount of compute required for each of the individual
- 669 experimental runs as well as estimate the total compute.
- 670 • The paper should disclose whether the full research project required more compute
- 671 than the experiments reported in the paper (e.g., preliminary or failed experiments that
- 672 didn't make it into the paper).

673 9. Code Of Ethics

674 Question: Does the research conducted in the paper conform, in every respect, with the
 675 NeurIPS Code of Ethics <https://neurips.cc/public/EthicsGuidelines?>

676 Answer: [Yes]

677 Justification: The research conducted in the paper is in full compliance with the NeurIPS
 678 Code of Ethics.

679 Guidelines:

- 680 • The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
- 681 • If the authors answer No, they should explain the special circumstances that require a
- 682 deviation from the Code of Ethics.
- 683 • The authors should make sure to preserve anonymity (e.g., if there is a special consid-
- 684 eration due to laws or regulations in their jurisdiction).

685 10. Broader Impacts

686 Question: Does the paper discuss both potential positive societal impacts and negative
 687 societal impacts of the work performed?

688 Answer: [Yes]

689 Justification: The paper provides both the potential benefits and the risks associated with the
 690 research, ensuring a comprehensive assessment of its societal implications.

691 Guidelines:

- 692 • The answer NA means that there is no societal impact of the work performed.
- 693 • If the authors answer NA or No, they should explain why their work has no societal
- 694 impact or why the paper does not address societal impact.
- 695 • Examples of negative societal impacts include potential malicious or unintended uses
- 696 (e.g., disinformation, generating fake profiles, surveillance), fairness considerations
- 697 (e.g., deployment of technologies that could make decisions that unfairly impact specific
- 698 groups), privacy considerations, and security considerations.
- 699 • The conference expects that many papers will be foundational research and not tied
- 700 to particular applications, let alone deployments. However, if there is a direct path to
- 701 any negative applications, the authors should point it out. For example, it is legitimate
- 702 to point out that an improvement in the quality of generative models could be used to
- 703 generate deepfakes for disinformation. On the other hand, it is not needed to point out
- 704 that a generic algorithm for optimizing neural networks could enable people to train
- 705 models that generate Deepfakes faster.
- 706 • The authors should consider possible harms that could arise when the technology is
- 707 being used as intended and functioning correctly, harms that could arise when the
- 708 technology is being used as intended but gives incorrect results, and harms following
- 709 from (intentional or unintentional) misuse of the technology.
- 710 • If there are negative societal impacts, the authors could also discuss possible mitigation
- 711 strategies (e.g., gated release of models, providing defenses in addition to attacks,
- 712 mechanisms for monitoring misuse, mechanisms to monitor how a system learns from
- 713 feedback over time, improving the efficiency and accessibility of ML).

714 11. Safeguards

715 Question: Does the paper describe safeguards that have been put in place for responsible
 716 release of data or models that have a high risk for misuse (e.g., pretrained language models,
 717 image generators, or scraped datasets)?

718 Answer: [NA]

719 Justification: The paper does not introduce assets that carry a high risk for misuse, therefore,
 720 no specific safeguards for data or model release are required.

721 Guidelines:

- 722 • The answer NA means that the paper poses no such risks.
- 723 • Released models that have a high risk for misuse or dual-use should be released with
- 724 necessary safeguards to allow for controlled use of the model, for example by requiring
- 725 that users adhere to usage guidelines or restrictions to access the model or implementing
- 726 safety filters.
- 727 • Datasets that have been scraped from the Internet could pose safety risks. The authors
- 728 should describe how they avoided releasing unsafe images.
- 729 • We recognize that providing effective safeguards is challenging, and many papers do
- 730 not require this, but we encourage authors to take this into account and make a best
- 731 faith effort.

732 12. Licenses for existing assets

733 Question: Are the creators or original owners of assets (e.g., code, data, models), used in
 734 the paper, properly credited and are the license and terms of use explicitly mentioned and
 735 properly respected?

736 Answer: [Yes]

737 Justification: The paper meticulously cites all external assets in references, including code,
 738 dataset, and models, acknowledging the contributions of their creators and respecting the
 739 associated licenses and terms of use.

740 Guidelines:

- 741 • The answer NA means that the paper does not use existing assets.
- 742 • The authors should cite the original paper that produced the code package or dataset.
- 743 • The authors should state which version of the asset is used and, if possible, include a
- 744 URL.

- 745 • The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- 746 • For scraped data from a particular source (e.g., website), the copyright and terms of
- 747 service of that source should be provided.
- 748 • If assets are released, the license, copyright information, and terms of use in the
- 749 package should be provided. For popular datasets, paperswithcode.com/datasets
- 750 has curated licenses for some datasets. Their licensing guide can help determine the
- 751 license of a dataset.
- 752 • For existing datasets that are re-packaged, both the original license and the license of
- 753 the derived asset (if it has changed) should be provided.
- 754 • If this information is not available online, the authors are encouraged to reach out to
- 755 the asset’s creators.

756 13. New Assets

757 Question: Are new assets introduced in the paper well documented and is the documentation

758 provided alongside the assets?

759 Answer: [Yes]

760 Justification: The novel universal image restoration framework introduced in the paper is

761 well documented, and the documentation is provided alongside the model in Sections 3,

762 offering comprehensive details for replication and application.

763 Guidelines:

- 764 • The answer NA means that the paper does not release new assets.
- 765 • Researchers should communicate the details of the dataset/code/model as part of their
- 766 submissions via structured templates. This includes details about training, license,
- 767 limitations, etc.
- 768 • The paper should discuss whether and how consent was obtained from people whose
- 769 asset is used.
- 770 • At submission time, remember to anonymize your assets (if applicable). You can either
- 771 create an anonymized URL or include an anonymized zip file.

772 14. Crowdsourcing and Research with Human Subjects

773 Question: For crowdsourcing experiments and research with human subjects, does the paper

774 include the full text of instructions given to participants and screenshots, if applicable, as

775 well as details about compensation (if any)?

776 Answer: [NA]

777 Justification: The paper does not engage in crowdsourcing experiments or research with

778 human subjects therefore, it does not include participant instructions, screenshots, or details

779 about compensation.

780 Guidelines:

- 781 • The answer NA means that the paper does not involve crowdsourcing nor research with
- 782 human subjects.
- 783 • Including this information in the supplemental material is fine, but if the main contribu-
- 784 tion of the paper involves human subjects, then as much detail as possible should be
- 785 included in the main paper.
- 786 • According to the NeurIPS Code of Ethics, workers involved in data collection, curation,
- 787 or other labor should be paid at least the minimum wage in the country of the data
- 788 collector.

789 15. Institutional Review Board (IRB) Approvals or Equivalent for Research with Human

790 Subjects

791 Question: Does the paper describe potential risks incurred by study participants, whether

792 such risks were disclosed to the subjects, and whether Institutional Review Board (IRB)

793 approvals (or an equivalent approval/review based on the requirements of your country or

794 institution) were obtained?

795 Answer: [NA]

796
797
798
799
800
801
802
803
804
805
806
807
808
809

Justification: The paper does not involve research with human subjects, so there are no participant risks to disclose, and no Institutional Review Board (IRB) approvals or equivalent reviews were required.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.