Test-Time Preference Optimization for Image Restoration

Bingchen Li[1]†, Xin Li[1]*, Jiaqi Xu[2], Jiaming Guo[2], Wenbo Li[3], Renjing Pei[2], Zhibo Chen[1]
[1]University of Science and Technology of China
[2]Huawei Noah's Ark Lab
[3]Joy Future Academy
lbc31415926@mail.ustc.edu.cn, xin.li@ustc.edu.cn, jiaqixuac@gmail.com

Abstract

Image restoration (IR) models are typically trained to recover high-quality images using L1 or LPIPS loss. To handle diverse unknown degradations, zero-shot IR methods have also been introduced. However, existing pre-trained and zero-shot IR approaches often fail to align with human preferences, resulting in restored images that may not be favored. This highlights the critical need to enhance restoration quality and adapt flexibly to various image restoration tasks or backbones without requiring model retraining and ideally without labor-intensive preference data collection. In this paper, we propose the first Test-Time Preference Optimization (TTPO) paradigm for image restoration, which enhances perceptual quality, generates preference data on-the-fly, and is compatible with any IR model backbone. Specifically, we design a training-free, three-stage pipeline: (i) generate candidate preference images online using diffusion inversion and denoising based on the initially restored image; (ii) select preferred and dispreferred images using automated preference-aligned metrics or human feedback; and (iii) use the selected preference images as reward signals to guide the diffusion denoising process, optimizing the restored image to better align with human preferences. Extensive experiments across various image restoration tasks and models demonstrate the effectiveness and flexibility of the proposed pipeline.

1 Introduction

Image restoration (IR) aims to remove unpleasant distortions from real-world low-quality (LQ) images and recover their high-quality (HQ) counterparts. Typical IR tasks include image denoising (Park, Yu, and Jeong 2019), image deblurring (Nah et al. 2021; Zhang et al. 2022a), image super-resolution (Li et al. 2024c,a), weather artifact removal (Yang et al. 2020; Li et al. 2018), and all-in-one restoration (Li et al. 2022). Despite the variety of IR models and advancements in zero-shot IR approaches (Kawar et al. 2022) to address unknown degradations, they often struggle to align with human preferences. For example, image restoration models optimized with standard L1 loss may achieve high PSNR scores but produce over-smoothed outputs that are less favored by users, highlighting the need for preference optimization.

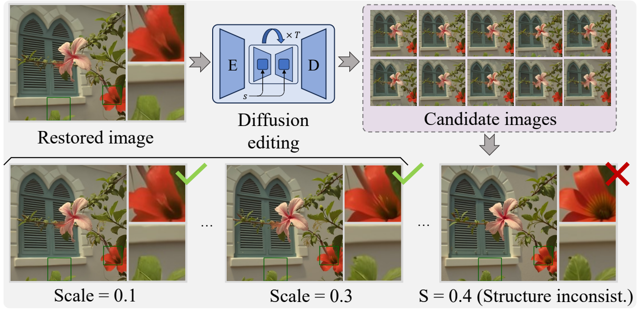

Figure 1: We use diffusion editing to generate candidate preference images. To preserve structural consistency, the noise scale is limited to a suitable range.

Increasing alignment between model outputs and human preferences has been actively explored in tasks like text-to-image (T2I) generation (Bie et al. 2024; Cao et al. 2024). Reinforcement learning (RL) is often used for this purpose by fine-tuning T2I models, leveraging human-annotated preference data, (*e.g.*, Pick-a-Pic (Kirstain et al. 2023)) to train reward models (Ouyang et al. 2022) or directly guide model fine-tuning (Wallace et al. 2024). While similar techniques could be applied to IR tasks, several challenges arise. First, existing image restoration models feature diverse architectural designs, with many being deterministic during inference, which makes them theoretically unsuitable as policy models under modern RL frameworks (*e.g.*, PPO (Schulman et al. 2017), DPO (Rafailov et al. 2023)). Furthermore, there are no publicly available preference datasets specifically tailored for IR tasks, as collecting such datasets with human annotations is both labor-intensive and challenging.

In this work, we propose Test-Time Preference Optimization (TTPO), a flexible approach to enhance restoration results for optimized perceptual quality, compatible with a wide range of IR tasks and backbones. The entire pipeline is broadly divided into three key stages: generation, selection, and optimization. In the first generation stage, multiple candidate preference images for restoration are generated on-the-fly for each image using pre-trained diffusion models. This is achieved by leveraging diffusion inversion (Meng et al. 2022) to modify the initially restored images while adjusting the scale of the added noise. Notably, us-

---

[*] Corresponding Author.

[†] Work done during an internship at Huawei Noah's Ark Lab.

ing a large noise scale can lead to structural inconsistencies, as shown in Fig. 1 (structural inconsistency), which is critical for IR tasks. To address this, we carefully limit the noise scale to an appropriate range when generating preference images. Furthermore, we ensemble multiple diffusion models in this process to enrich the diversity of candidate preference images. Note that this preference image generation process is model-agnostic and applicable to the initially restored images from any IR backbone.

In the second stage, selection, we identify preferred and dispreferred images from the generated restoration samples. While this process could ideally involve manual selection by humans to reflect true preferences, such involvement becomes impractical for large-scale datasets. To address this, we opt to use automated evaluation metrics to approximate human preferences. Specifically, no-reference image quality assessment (NR-IQA) metrics (Wu et al. 2024a; You et al. 2025; Wu et al. 2025) are well-suited for evaluating restored image quality, as ground truth images are typically unavailable during testing. To improve selection quality, we adopt a metric combination strategy that integrates top human-aligned metrics together to rank images, as individual metrics may deviate from human preferences.

In the final optimization stage, we refine the initially restored image using conditioned diffusion denoising, guided by the selected preference samples. This controlled optimization process enhances image details while incorporating preference alignment. Specifically, at each reverse denoising step, the clean image is firstly predicted by a pre-trained diffusion pipeline (Ho, Jain, and Abbeel 2020). To guide the actual subsequent denoising step, we modulate the process using preference samples by calculating the distance between the predicted clean image and the preferred/dispreferred samples. These distance measurements serve as reward signals, generating gradient descent directions that condition the denoising process, steering it toward the preferred direction while avoiding the dispreferred one. Additionally, the initially restored image is integrated as a constraint to preserve the overall image structure throughout the optimization. Various techniques, including frequency-based guidance decomposition and stage-wise guidance, are designed to enhance the optimization process.

Overall, the contributions of this paper can be summarized as follows:

- To the best of our knowledge, we introduce the first test-time preference optimization paradigm for image restoration, requiring neither model retraining nor extensive human preference data collection. Instead, the proposed pipeline can be seamlessly integrated with any existing IR model and directly optimizes restored images.
- We present an innovative approach that leverages diffusion-based editing to generate preference samples on-the-fly and employs conditioned diffusion denoising for preference optimization. Furthermore, our pipeline offers high flexibility by allowing human involvement in the selection of preference images.
- We conduct comprehensive experiments, including 12 quantitative comparisons and a user study, demonstrating the versatility of our proposed pipeline across a wide range of IR tasks and models, achieving improvements in both objective metrics and subjective human evaluations.

2 Related Works

2.1 Image Restoration Models

Image restoration (IR) has witnessed the development of diverse model architectures, ranging from traditional CNN-based (Dong et al. 2015) to recent Transformer-based (Chen et al. 2023; Quan et al. 2024), diffusion-based (Li et al. 2023; Ren et al. 2025; Qin et al. 2025), MLP-based (Tu et al. 2022; Li et al. 2024c), and state-space model-based (Guo et al. 2025) designs, reflecting the evolving needs for a better restoration fidelity and perceptual quality. Despite advances in network designs, the training objective for most IR models remains largely unchanged. Early works (Dong et al. 2015) primarily focus on optimizing pixel-wise fidelity using L1/MSE losses. However, such objectives often lead to over-smoothed outputs. To address this limitation, some subsequent approaches (Zhang et al. 2021; Wang et al. 2021) incorporate generative adversarial networks (GANs) into the training pipeline to produce more perceptually pleasing results. While numerous network designs have emerged for IR, little attention has been given to preference optimization, leading to restored images that often misalign with human perceptual preferences.

2.2 Zero-Shot Diffusion-Based Image Restoration

Among various IR models, there is a special type of diffusion-based IR model worth noting. Zero-shot diffusion-based IR models aim to leverage the strong generative prior in the pre-trained diffusion model (*e.g.*, DDPM (Ho, Jain, and Abbeel 2020)) to restore images. Song *et al.* (Song et al. 2022) propose to use gradient guidance to address inverse problems in medical imaging. ILVR (Choi et al. 2021) incorporates low-frequency condition information from the reference image as guidance. DDRM (Kawar et al. 2022) proposes an SVD-based decomposition solution to improve guidance information. DDNM (Wang, Yu, and Zhang 2023) further enhances the solution with range-null space decomposition. DPS (Chung et al. 2023) predicts noise-clean images at each denoising step, obtaining gradient guidance for the next step denoising through distance measurements. GDP (Fei et al. 2023) shares a similar idea, while TAO (Gou et al. 2024) further improves the distance measurements by incorporating GAN loss.

2.3 Diffusion Preference Optimization

Diffusion preference optimization typically aims to align pre-trained T2I generation models to generate more visually appealing results. Diffusion-DPO (Wallace et al. 2024) introduces DPO (Rafailov et al. 2023) into T2I generation and fine-tunes diffusion models with pairwise preference data to achieve better prompt alignment and aesthetic scores. MaPO (Hong et al. 2024) handles reference mismatch by replacing divergence regularization on the reference model with an amplification factor. Some recent works (Lee et al. 2025; Gu et al. 2024; Yang, Chen, and Zhou 2024) focus on reward design and further improve the alignments of T2I models. To improve the training process of diffusion-DPO, Curriculum DPO (Croitoru et al. 2025) proposes a curriculum learning approach, gradually introducing harder preference pairs to improve model alignment performance.

3 Preliminaries

3.1 Denoising Process

The diffusion process is the key operation in diffusion models, where the objective is to remove noise from the noisy data samples x_t to recover the clean data x_0, or conversely, to add noise in the forward process. Specifically, we focus on the Euler flow-matching scheduler employed in state-of-the-art diffusion models (Esser et al. 2024; Labs 2024):

$$x_{t'} = x_t + \Delta t \cdot v_\theta(x_t, t), \tag{1}$$

where each timestep $t \in [0,1], \Delta t = t - t'$ is the interval between two timesteps, $t', v_\theta(x_t, t)$ denotes the next sampled timestep and the learned velocity field, respectively. Notably, the clean data x_0 can be predicted by Eq. 1 by setting t' to 0: $x_0 = x_t + t \cdot v_\theta(x_t, t)$. Conversely, the intermediate noisy latent x_t cin the forward process can be sampled using Eq. 13 from (Esser et al. 2024):

$$x_t = (1-t)x_0 + t\epsilon, \tag{2}$$

where $\epsilon \sim \mathcal{N}(0, I)$ denotes standard Gaussian noise.

3.2 Preference Optimization

Preference optimization is typically achieved by training models on preference data triplets (c, x_w, x_l), where x_w and x_l represent the preferred (win) and dispreferred (lose) samples, and c denotes the condition used to generate these samples. The probability of such preferences is often formulated by Bradley-Terry model (Bradley and Terry 1952):

$$P(x_w \succ x_l \mid c) = \sigma(r(x_w, c) - r(x_l, c)), \tag{3}$$

where $\sigma(\cdot)$ is the sigmoid function and r denotes the reward model used to represent human preferences. Due to the challenges of learning such a reward model, Rafailov *et al.* (Rafailov et al. 2023) propose directly optimizing the model using preference data $(c, x_w, x_l) \sim \mathcal{D}$:

$$\mathcal{L}_{\text{DPO}} = -\mathbb{E}_{(c,x_w,x_l)\sim\mathcal{D}}\left[\log\sigma\left(\beta \cdot \log\frac{p_\theta(x_w \mid c)}{p_{\text{ref}}(x_w \mid c)} - \beta \cdot \log\frac{p_\theta(x_l \mid c)}{p_{\text{ref}}(x_l \mid c)}\right)\right], \tag{4}$$

where p_{ref}, p_θ denote the fixed pre-trained model and the fine-tuned model to align the preference, respectively, and β controls the regularization strength.

4 Methods

We propose the training-free test-time preference optimization (TTPO) approach for image restoration, which dynamically generates preference samples online and leverages them as reward signals to optimize the initially restored image from any image restoration backbone. This enhances visual preferences through a conditioned diffusion denoising process. TTPO consists of three stages: (i) generation of candidate preference images in Sec. 4.1; (ii) selection of preference images in Sec. 4.2; and (iii) optimization of restored images in Sec. 4.3.

4.1 Generation of Candidate Preference Images

In this stage, candidate preference images are generated for a given initially restored image from any image restoration backbone. The success of preference optimization often relies on human-annotated preference datasets (Kirstain et al. 2023). However, for image restoration (IR) tasks, such preference datasets are unavailable, and collecting preference data is highly labor-intensive. Therefore, we propose generating preference data $\mathcal{D}$ at test time in TTPO. Leveraging the powerful editing capabilities of pre-trained diffusion models, this stage employs diffusion inversion (Meng et al. 2022) to generate diverse preference samples.

Specifically, let y_0 represent the initially restored image from an IR model. First, we obtain its noisy latent $x_s = (1-s)x_0 + s\epsilon$ using Eq. 2, where $s \in [0,1]$ controls the scale of the added noise, and x_0 is the encoded latent via VAE encoder $\mathcal{E}$. Next, the edited version y_0^i is generated by progressively removing noise from x_s using diffusion denoising (Eq. 1) and decoding the latent back into an image in the end. Notably, the inference of pre-trained diffusion models implicitly incorporates rich generative priors, such as texture and detail generation, into the edited image. This ensures the diversity of the generated preference images. However, we observe that a large scale s, such as $s = 0.4$, leads to significant structural alternation (Fig. 1). To preserve structural consistency, we limit s to the range $[0.1, 0.3]$. Using this approach for a single sample, we construct the preference image set $\mathcal{D} = \{y_0, y_0^1, y_0^2, \ldots, y_0^N\}$ by leveraging multiple pre-trained diffusion models and varying s, where N is the number of generated preference samples.

4.2 Selection of Preference Images

After generating the preference samples, our goal is to select the preferred (win) sample y_w and the dispreferred (lose) sample y_l. While human involvement offers the most reliable indication of preference, it is labor-intensive and often impractical for large-scale applications. To overcome this challenge, we propose a simple yet effective alternative by utilizing no-reference image quality assessment (NR-IQA) metrics as a proxy for preference selection.

However, relying on a single NR-IQA metric may not always align well with human perception (Zhang et al. 2022b). To address this, we adopt a hybrid evaluation approach by combining multiple metrics, consistent with findings in (Chen et al. 2024; Lin et al. 2025). To implement this, we conduct a subjective experiment (details provided in the extended version) to identify the top three metrics from the `pyiqa` package (Chen and Mo 2022) that best align with human preferences. Using these selected metrics, we construct the hybrid metric by assigning a reward to each sample $y_0^i \sim \mathcal{D}$ based on the average of the Z-score normalized values from the three NR-IQA metrics:

$$r_{y_0^i} = \left(\text{IQA}_1 + \text{IQA}_2 + \text{IQA}_3\right) / 3. \tag{5}$$

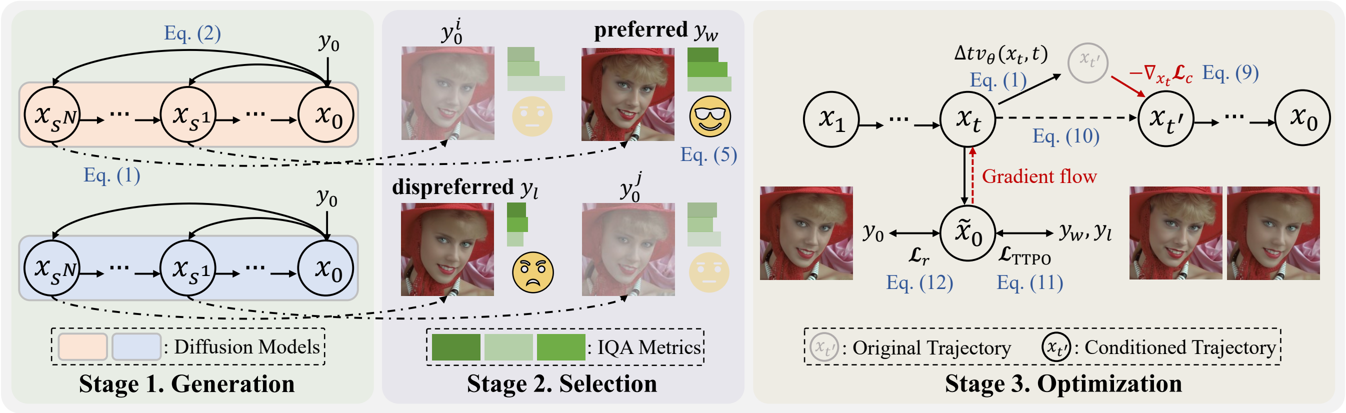

Figure 2: The pipeline of TTPO, which follows a generation-selection-optimization paradigm to perform test-time preference optimization for restored images. The pseudo codes are provided as Algorithm 1 in the extended version.

These rewards are then used to rank the candidate samples, with the top-ranked image selected as y_w and the bottom-ranked image as y_l.

4.3 Optimization of Restored Images

The core of TTPO lies in optimizing the restored image. In this stage, we leverage pre-trained diffusion models (Labs 2024) and integrate reward signals from the selected preference samples (y_w, y_l) via a conditioned diffusion process. This stage optimizes the initially restored image y_0, ultimately producing a perceptually enhanced version.

During the denoising process starting from noise, given the noisy data x_t at timestep t, the model follows the pre-trained denoising trajectory (*i.e.*, Eq. 1) to compute the update direction $\Delta t \cdot v_\theta(x_t, t)$ for the next timestep. Conditional information can be intuitively incorporated into this trajectory by introducing an additional correction term to this update direction. Note that the preference samples y_w, y_l are in the "clean" space at $t = 0$. Accordingly, we compute the conditional guidance $\mathcal{L}_c(x_0, y_w, y_l)$ at each timestep by first deriving the predicted clean data x_0, where $x_0 = x_t + t \cdot v_\theta(x_t, t)$, as defined in Eq. 1. We incorporate preference information into the conditioning process by modulating the reward signals through the calculation of the distance between x_0 and y_w, y_l as the relative rewards:

$$\mathcal{L}_{\text{TTPO}} = -\log \sigma(\beta d(x_0, \mathcal{E}(y_w)) - \beta d(x_0, \mathcal{E}(y_l))), \quad (6)$$

where d denotes a similarity-aware distance (*e.g.*, cosine similarity as the reward), β controls the strength of the reward signals, and $\mathcal{E}$ is VAE encoder. For distance metrics where lower values indicate greater similarity (*e.g.*, L1), a negative sign is applied to d to ensure a consistent reward direction. Notably, we also include y_0 itself as an additional condition to preserve structural consistency:

$$\mathcal{L}_r = \|x_0 - \mathcal{E}(y_0)\|_2^2, \quad (7)$$

$$\mathcal{L}_c(x_0, y_0, y_w, y_l) = \alpha \mathcal{L}_{\text{TTPO}} + \mathcal{L}_r, \quad (8)$$

where α is the weighting parameter. Next, we compute the correction term as:

$$-\nabla_{x_t}\mathcal{L}_c = \frac{\partial \mathcal{L}_c}{\partial x_t} = \frac{\partial \mathcal{L}_c}{\partial x_0} \cdot \frac{\partial x_0}{\partial x_t} = \frac{\partial \mathcal{L}_c}{\partial x_0}. \quad (9)$$

Here, we apply a stop gradient operation on $v_\theta(x_t, t)$ for two reasons: (i) to introduce a correction term that guides the update of x_t without interfering with the pre-trained denoising direction, and (ii) to avoid destabilizing the overall denoising process by altering the learned denoising trajectory. Detailed derivations are provided in the extended version. Finally, we integrate the correction term into Eq. 1 to compute the noisy data for the next timestep:

$$x_{t'} = x_t + \Delta t \cdot v_\theta(x_t, t) - g\nabla_{x_t}\mathcal{L}_c, \quad (10)$$

where g is the scaling factor. Intuitively, integrating conditional guidance into the denoising process enables the restored image to better align with preferences by maximizing the reward for preferred samples and minimizing it for dispreferred ones, all while preserving structural consistency.

A potential issue arises when y_l closely resembles y_0 and the structural consistency loss (Eq. 7) causes x_0 to be similar to y_l, resulting in an artificially inflated reward in Eq. 6. This undermines the effectiveness of preference guidance and may result in unstable optimization. To mitigate this, we apply frequency decomposition to x_0 and separately compute preference optimization and structural preservation in distinct frequency components. Specifically, we derive the preference optimization loss $\mathcal{L}_{\text{TTPO}}$ from the high-frequency components, where textures are most prominent. Simultaneously, we enforce structural consistency $\mathcal{L}_r$ using the low-frequency components. This process is formulated as:

$$\begin{aligned}\mathcal{L}_{\text{TTPO}} = -\log \sigma\Big(&\beta \mathcal{D}\big(\mathcal{F}^{\text{high}}(x_0), \mathcal{F}^{\text{high}}(\mathcal{E}(y_w))\big) \\ &- \beta \mathcal{D}\big(\mathcal{F}^{\text{high}}(x_0), \mathcal{F}^{\text{high}}(\mathcal{E}(y_l))\big)\Big),\end{aligned} \quad (11)$$

$$\mathcal{L}_r = \|\mathcal{F}^{\text{low}}(x_0) - \mathcal{F}^{\text{low}}(\mathcal{E}(y_0))\|_2^2, \quad (12)$$

where $\mathcal{F}^{\text{high}}(\cdot) = \texttt{FFT}(\cdot) \odot (1 - \mathcal{G}(D_0))$ and $\mathcal{F}^{\text{low}}(\cdot) = \texttt{FFT}(\cdot) \odot \mathcal{G}(D_0)$. $\mathcal{G}, D_0$ denote the Gaussian low-pass filter and the normalized cutoff frequency, respectively.

We divide the denoising process into three stages with distinct optimization objectives, inspired by findings that diffusion models generate structural information during early

	(a) DN (Kodak) - SwinIR (Liang et al. 2021)				(b) DN (Kodak) - Restormer (Zamir et al. 2022)				(c) DN (Kodak) - DMID (Li et al. 2024b)			
	y_0	y_{TTPO}	y_w	y_l	y_0	y_{TTPO}	y_w	y_l	y_0	y_{TTPO}	y_w	y_l
MUSIQ	73.790	**74.242**	73.816	70.210	74.059	$\underline{74.105}$	74.280	70.401	73.999	73.893	74.249	70.175
MANIQA	0.4565	**0.4912**	0.4828	0.4317	0.4655	**0.4928**	0.4858	0.4361	0.4664	**0.4832**	0.4765	0.4276
Q-Align	3.9830	$\underline{4.3201}$	4.3662	3.9122	4.5054	$\underline{4.6056}$	4.6706	4.3326	4.5736	$\underline{4.6276}$	4.6696	4.3391
NIQE ↓	4.1567	4.3001	4.1244	3.8924	4.2216	**3.9546**	3.9956	3.9148	4.1986	**4.0079**	4.0385	3.8991
CLIPIQA	0.5965	$\underline{0.6117}$	0.6333	0.6342	0.5937	$\underline{0.6015}$	0.6198	0.5938	0.6147	0.6097	0.6221	0.6030
LIQE	4.2576	**4.8690**	4.8051	4.6102	4.2599	$\underline{4.8121}$	4.8233	4.4915	4.4391	**4.8450**	4.8004	4.5659

	(d) CISR (BSD100) - SwinIR (Liang et al. 2021)				(e) CISR (BSD100) - HAT (Chen et al. 2023)				(f) RISR (RealSR) - StableSR (Wang et al. 2024)			
	y_0	y_{TTPO}	y_w	y_l	y_0	y_{TTPO}	y_w	y_l	y_0	y_{TTPO}	y_w	y_l
MUSIQ	57.720	$\underline{57.764}$	59.540	50.249	57.798	$\underline{58.066}$	59.588	50.785	65.883	$\underline{66.096}$	66.694	62.193
MANIQA	0.3499	**0.3745**	0.3691	0.2811	0.3510	**0.3778**	0.3702	0.2828	0.4275	$\underline{0.4474}$	0.4488	0.3852
Q-Align	2.9806	$\underline{3.0110}$	3.1874	2.6024	2.9831	$\underline{3.0291}$	3.1988	2.6006	3.2766	$\underline{3.3900}$	3.4917	3.1204
NIQE ↓	6.1102	6.2586	5.8336	5.9172	6.0029	6.1579	5.7606	5.8992	5.8812	$\underline{5.7778}$	5.4205	5.3647
CLIPIQA	0.5300	0.5284	0.5518	0.4397	0.5399	0.5328	0.5555	0.4456	0.6233	0.5518	0.5979	0.5382
LIQE	3.0100	$\underline{3.0819}$	3.2359	2.7823	3.0343	$\underline{3.1246}$	3.2639	2.8359	3.5831	$\underline{3.6191}$	3.7920	3.5152

	(g) RISR (RealSR) - OSEDiff (Wu et al. 2024b)				(h) DR (Rain100L) - Restormer (Zamir et al. 2022)				(i) DR (Rain100L) - PromptIR (Potlapalli et al. 2023)			
	y_0	y_{TTPO}	y_w	y_l	y_0	y_{TTPO}	y_w	y_l	y_0	y_{TTPO}	y_w	y_l
MUSIQ	69.089	$\underline{69.138}$	69.723	65.887	71.558	$\underline{72.181}$	72.802	69.099	70.555	$\underline{71.413}$	72.214	67.994
MANIQA	0.4717	**0.4837**	0.4816	0.4254	0.5200	**0.5424**	0.5422	0.4718	0.5063	**0.5478**	0.5437	0.4678
Q-Align	3.6962	$\underline{3.7458}$	3.8086	3.4310	4.0219	**4.2553**	4.2325	3.7953	3.8595	$\underline{4.0759}$	4.1588	3.6656
NIQE ↓	5.6474	5.8434	5.4493	5.4818	3.7239	$\underline{3.6696}$	3.6081	3.4614	3.1646	3.6794	3.4022	3.3319
CLIPIQA	0.6693	0.5875	0.6428	0.5724	0.7277	$\underline{0.7331}$	0.7440	0.7001	0.7691	0.7286	0.7623	0.7190
LIQE	4.1216	4.0495	4.1909	3.9410	4.7392	4.7255	4.7240	4.1047	4.7529	**4.7820**	4.7167	4.0718

	(j) LLIE (CDD11) - OKNet (Cui, Ren, and Knoll 2024)				(k) LLIE (CDD11) - WeatherDiff (Özdenizci and Legenstein 2023)				(l) LLIE (CDD11) - OneRestore (Guo et al. 2024)			
	y_0	y_{TTPO}	y_w	y_l	y_0	y_{TTPO}	y_w	y_l	y_0	y_{TTPO}	y_w	y_l
MUSIQ	67.832	$\underline{67.953}$	69.967	63.675	69.643	69.111	70.688	64.287	69.704	$\underline{70.050}$	71.607	65.578
MANIQA	0.3809	$\underline{0.3965}$	0.4261	0.3328	0.3575	$\underline{0.3671}$	0.3941	0.3126	0.4277	$\underline{0.4531}$	0.4799	0.3619
Q-Align	3.9201	$\underline{4.1880}$	4.3936	3.6550	4.1281	$\underline{4.2604}$	4.4086	3.7904	4.1983	$\underline{4.3741}$	4.5656	3.9126
NIQE ↓	2.7342	2.9758	2.9789	2.9956	2.7349	2.9613	2.9511	3.0477	3.3176	**3.3084**	3.3547	3.2066
CLIPIQA	0.5032	$\underline{0.5185}$	0.5786	0.4776	0.5184	$\underline{0.5322}$	0.5997	0.4816	0.6051	0.5862	0.6419	0.5409
LIQE	3.4979	$\underline{4.1540}$	4.3748	3.4762	3.9674	$\underline{4.3408}$	4.4619	3.7231	4.0552	$\underline{4.4837}$	4.5790	3.8407

Table 1: Quantitative comparisons across various IR models and IR tasks. We underline the results where the optimized image y_{TTPO} outperforms the restored image y_0 in evaluation metrics. If y_{TTPO} surpasses the preferred sample y_w, we further highlight the result in **bold**. Notably, we also provide numerical metric values for y_w and y_l in the Table as references.

timesteps, while fine details are refined in later steps (Rombach et al. 2022). In the first stage $(T_1, 1]$, only $\mathcal{L}_r$ is applied for structural guidance. In the middle stage $(T_2, T_1]$, both $\mathcal{L}_{\text{TTPO}}$ and $\mathcal{L}_r$ are employed. Finally, in the last stage $[0, T_2]$, only $\mathcal{L}_{\text{TTPO}}$ is used for fine-grained preference optimization.

Discussion about Pixel Space v.s. Latent Space. State-of-the-art diffusion models are commonly built upon Latent Diffusion Models (Rombach et al. 2022), which perform the denoising process in the latent space. While the pixel space can provide more precise spatial information, we opt to conduct conditioning in the latent space. This choice is motivated by two key factors: decoding to the pixel space at each timestep is computationally expensive, and, more importantly, we empirically observed that gradient backpropagation through the VAE decoder can be unstable.

Comparing $\mathcal{L}_{\text{TTPO}}$ to $\mathcal{L}_{\text{DPO}}$. While $\mathcal{L}_{\text{DPO}}$ (Rafailov et al. 2023) focuses on preference alignments during training, $\mathcal{L}_{\text{TTPO}}$ is designed for preference optimization at test time. Our method offers a more flexible and *model-agnostic* solution by directly optimizing the restored image (y_0), eliminating the need to retrain IR models.

5 Experiments

In this section, we present the main experimental results of TTPO and conduct ablation studies to analyze its key components. For detailed parameter configurations, please refer to our extended version.

5.1 Results on Various IR Models and IR Tasks

We consider several representative IR tasks to demonstrate the effectiveness of TTPO: (i) the most classical IR task, image denoising (DN); (ii) the classical image super-resolution (CISR) task; (iii) the practical real-world image super-resolution (RISR) task; (iv) the image deraining (DR) task; (v) low-light image enhancement (LLIE) task; and (vi) the zero-shot diffusion-based IR (ZSDIR) task. *Since our goal is to optimize the restored image, we primarily focus on methods whose visual results are publicly available.*

	y_0	DDRM	DDNM	TAO	y_{TTPO}	y_0	DDRM	DDNM	TAO	y_{TTPO}
	(a) DN - Restormer (Zamir et al. 2022)					(b) DN - DMID (Li et al. 2024b)				
MUSIQ	61.317	59.345	<u>60.509</u>	58.930	**61.054**	61.696	59.304	<u>60.908</u>	58.990	**61.098**
MANIQA	0.4690	0.4450	<u>0.4614</u>	0.4326	**0.4741**	0.4676	0.4437	<u>0.4611</u>	0.4338	**0.4767**
Q-Align	2.8213	2.7105	2.7722	<u>2.7829</u>	**2.9281**	2.8686	2.7248	2.8037	<u>2.8245</u>	**2.9687**
NIQE ↓	7.0202	7.7514	7.6399	**6.6802**	<u>7.0108</u>	6.7982	7.7266	7.5267	**6.6481**	<u>6.8480</u>
CLIPIQA	0.5600	0.5153	0.5225	**0.6064**	<u>0.5849</u>	0.5802	0.5365	0.5420	**0.6129**	<u>0.6011</u>
LIQE	3.8615	3.6078	3.6512	<u>4.1011</u>	**4.2702**	3.9228	3.6481	3.6975	<u>4.1588</u>	**4.3020**

Table 2: Quantitative comparisons between TTPO and previous ZSDIR methods, DDRM (Kawar et al. 2022), DDNM (Wang, Yu, and Zhang 2023), and TAO (Gou et al. 2024). Results are reported on Kodak24 (Franzen 1999) with the 256 × 256 resolution. For a fair comparison, we use the restored image y_0 (rather than the distorted image) as input to all ZSDIR methods. Each method is executed multiple times to identify the optimal hyperparameter settings. The best and second-best results are **bolded** and <u>underlined</u> (values of y_0 serve as reference).

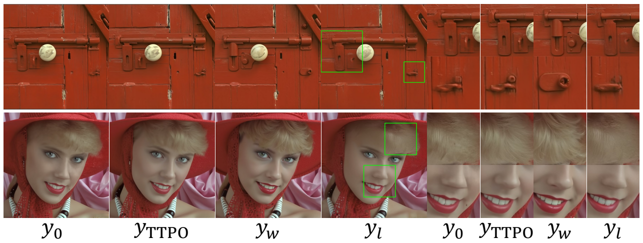

Figure 3: Qualitative comparisons between the initially restored image y_0, optimized image y_{TTPO}, preferred image y_w, and dispreferred image y_l.

Main Results. We present the quantitative results in Tab. 1. We report MUSIQ (Ke et al. 2021), MANIQA (Yang et al. 2022), and Q-Align (Wu et al. 2024a) in the Table, where these metrics are leveraged to select the preference images. Additionally, we include three more commonly adopted NR-IQA metrics, NIQE (Mittal, Soundararajan, and Bovik 2012), CLIPIQA (Wang, Chan, and Loy 2023), and LIQE (Zhang et al. 2023), to provide a comprehensive evaluation of the optimized images. Several conclusions can be drawn from the Table: (i) TTPO consistently improves over the restored image y_0 across almost all tasks and evaluation metrics. (ii) In many cases, TTPO achieves improvements not only on the metrics used during the preference selection stage but also on additional metrics that were not involved in the optimization. (iii) In some tasks, TTPO even surpasses the preferred image y_w. These three observations demonstrate the flexibility and effectiveness of TTPO against various IR models/tasks. We demonstrate the qualitative comparisons in Fig. 3. As observed, images optimized through preference guidance exhibit more realistic textures. For example, y_{TTPO} presents more natural representations of hair and teeth, without artifacts such as twisted lines. Notably, although y_w sometimes demonstrates finer texture details, it often introduces undesirable structural changes (*e.g.*, structure shifts in the door region). Therefore, y_w *cannot be directly used as the final result*, as structural consistency is a fundamental requirement in IR. In contrast, y_{TTPO} retains structural information with improved perceptual quality.

Comparing TTPO with ZSDIR Methods. Since ZSDIR methods are based on pre-trained diffusion models and do not require training, we compare our TTPO with them for *optimizing* restored images. We include three representative ZSDIR baselines: DDRM (Kawar et al. 2022), DDNM (Wang, Yu, and Zhang 2023), and TAO (Gou et al. 2024). Specifically, the restored image y_0 is used as input for all methods. For a fair comparison, all inputs are resized to 256×256, including those for TTPO. Each ZSDIR method is executed multiple times to determine the optimal hyperparameter settings. As shown in Tab. 2, TTPO achieves the best or second-best results across all metrics, which indicates the effectiveness of preference guidance.

Integrating TTPO with ZSDIR Methods. Notably, ZSDIR methods can be seamlessly integrated with TTPO to form a two-stage test-time restoration and optimization method for distorted images. We provide an example in Fig. 5, where the old photo is first restored by DDNM (Wang, Yu, and Zhang 2023) and optimized by TTPO. As shown, TTPO enhances the texture details and brings the old photo back to life.

User Study. To further demonstrate the preference optimization capability of TTPO, we conduct a user study on Kodak24 (Franzen 1999) dataset. We invite eight human experts to participate in this study, each with more than three years of research experience in low-level vision. Each expert is first asked to select preference images in stage two. We then optimize restored images based on their selections. Subsequently, these experts participate in stage three to compare the optimized images y_{human} with the following settings: (i) y_0; (ii) y_{metric} optimized by metric-selected preference images; (iii) y_{random} optimized by randomly selected preference images; and (iv) y_{reverse} optimized by switching the preference image pairs. We report the win rate of y_{human} in Fig. 6. As observed, y_{human} surpasses y_0 in most comparisons, which indicates the ability of TTPO to enhance the subjective quality of y_0. Additionally, a win rate near 50% between y_{human} and y_{metric} suggests that the NR-IQA metrics employed in TTPO align well with human preferences. More settings are detailed in the extended version.

Analysis of Conditions. We illustrate the curves of $\mathcal{L}_{\text{TTPO}}, \mathcal{L}_r, \mathcal{D}(x_0, \mathcal{E}(y_w))$, and $\mathcal{D}(x_0, \mathcal{E}(y_l))$ in Fig. 4. As demonstrated, both the preference guidance loss $\mathcal{L}_{\text{TTPO}}$

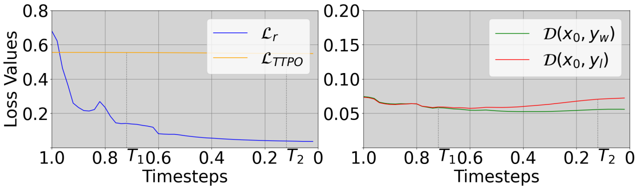

Figure 4: Loss curves for the upper image in Fig. 3 across all diffusion timesteps.

old photo y_0 (DDNM) y_{TTPO}

Figure 5: Integrating TTPO with the existing ZSDIR method DDNM (Wang, Yu, and Zhang 2023).

and the structural consistency loss $\mathcal{L}_r$ gradually decrease throughout the denoising process. Meanwhile, the distance between the clean data and the preferred image becomes smaller, while the distance to the dispreferred image grows larger. This demonstrates the effectiveness and correctness of the TTPO pipeline.

5.2 Ablation Studies

Unless stated otherwise, all experiments in this part are conducted on Kodak (Franzen 1999) denoised (DN) by SwinIR (Liang et al. 2021). We provide more ablation studies about parameter settings in the extended version.

Effectiveness of $\mathcal{L}_{\text{TTPO}}$, Frequency Decomposition, and Division of Denoising Steps. As demonstrated in Tab. 4, $\mathcal{L}_{\text{TTPO}}$ brings obvious improvements against (i). This indicates that our developed preference guidance effectively transfers the preference as a reward signal to guide the denoising process. TTPO achieves better performance with frequency decomposition in (iii), as it prevents the reward for y_l from being excessively high due to the similarity in low-frequency structural information. By dividing the denoising process into three stages and applying different guidance in (iv), the performance increases compared to (ii), as expected. This is because introducing preference guidance too early may interfere with the generation of structural details, leading to a suboptimal denoising process. By putting things together in (v), TTPO achieves the best performance, indicating the effectiveness of our optimization pipeline.

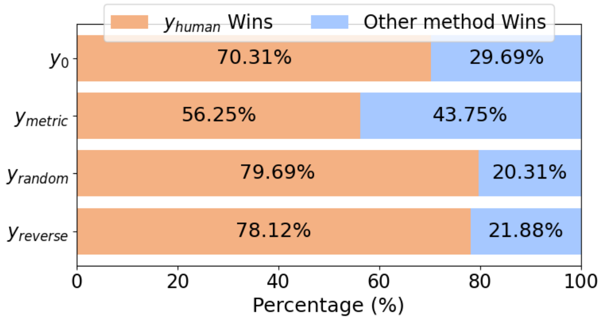

Figure 6: User study of y_{human} *vs.* other selection methods.

Model	MUSIQ	MANIQA	Q-Align	Inf. Time
SD3	74.709	0.4910	4.1135	26s
FLUX	74.242	0.4912	4.3201	45s

Table 3: Different diffusion models used in the third stage.

	$\mathcal{L}_r$	$\mathcal{L}_{\text{TTPO}}$	`FFT`	T_1, T_2	MUSIQ	MANIQA	Q-Align
(i)	✓				73.437	0.4701	4.1768
(ii)	✓	✓			73.895	0.4798	4.2397
(iii)	✓	✓	✓		74.103	0.4857	4.2422
(iv)	✓	✓		✓	74.051	0.4835	4.2585
(v)	✓	✓	✓	✓	**74.242**	**0.4912**	**4.3201**

Table 4: Effectiveness of each component in the third stage.

Flexibility of Diffusion Model in the Third Stage. Since the TTPO pipeline does not require any retraining, it provides flexibility in choosing the diffusion model used to align restored images with human preferences. We compare the performance between FLUX (Labs 2024) and SD3 (Esser et al. 2024) in Tab. 3. Since FLUX has more parameters than SD3, the latter can perform the TTPO process more efficiently, resulting in faster optimization. Therefore, users can choose the most appropriate diffusion model for preference optimization, depending on their trade-off between speed and image quality.

6 Conclusion

In this paper, we propose the first Test-Time Preference Optimization (TTPO) paradigm for image restoration. By generating preference data on-the-fly and leveraging a three-stage optimization pipeline, TTPO effectively aligns restored images with user preferences, eliminating the need for costly model retraining or labor-intensive data collection. Extensive experiments highlight the flexibility and effectiveness of TTPO across diverse image restoration tasks and backbones, demonstrating its capability to optimize images in various scenarios while enhancing perceptual quality and preserving structural consistency. Our method, however, has two main limitations: (i) the inference process of FLUX requires considerable computational resources, leading to relatively long runtime; and (ii) current NR-IQA metrics still fall short of fully capturing human preferences, making it difficult for metric-based selection to fully align the optimization results with human judgments. We leave improving the computational efficiency of TTPO and developing more accurate preference selection mechanisms as future works.

Acknowledgments

This work was supported in part by NSFC under Grant 623B2098, 62021001, and 62371434, the Postdoctoral Fellowship Program of CPSF under Grant Number GZC20252293, and the China Postdoctoral Science Foundation-Anhui Joint Support Program under Grant Number 2024T017AH. This work was also funded by Anhui Postdoctoral Scientific Research Program Foundation (No.2025A1015) and the Fundamental Research Funds for the Central Universities under Grant Number WK2100250064.

References

Bie, F.; Yang, Y.; Zhou, Z.; Ghanem, A.; Zhang, M.; Yao, Z.; Wu, X.; Holmes, C.; Golnari, P.; Clifton, D. A.; et al. 2024. Renaissance: A survey into ai text-to-image generation in the era of large model. *IEEE TPAMI*.

Bradley, R. A.; and Terry, M. E. 1952. Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*.

Cao, P.; Zhou, F.; Song, Q.; and Yang, L. 2024. Controllable generation with text-to-image diffusion models: A survey. *arXiv preprint arXiv:2403.04279*.

Chen, C.; and Mo, J. 2022. IQA-PyTorch: PyTorch Toolbox for Image Quality Assessment. [Online]. Available: https://github.com/chaofengc/IQA-PyTorch.

Chen, H.; Li, W.; Gu, J.; Ren, J.; Chen, S.; Ye, T.; Pei, R.; Zhou, K.; Song, F.; and Zhu, L. 2024. Restoreagent: Autonomous image restoration agent via multimodal large language models. *NeurIPS*.

Chen, X.; Wang, X.; Zhou, J.; Qiao, Y.; and Dong, C. 2023. Activating more pixels in image super-resolution transformer. In *CVPR*.

Choi, J.; Kim, S.; Jeong, Y.; Gwon, Y.; and Yoon, S. 2021. Ilvr: Conditioning method for denoising diffusion probabilistic models. *arXiv preprint arXiv:2108.02938*.

Chung, H.; Kim, J.; Mccann, M. T.; Klasky, M. L.; and Ye, J. C. 2023. Diffusion posterior sampling for general noisy inverse problems. In *ICLR*.

Croitoru, F.-A.; Hondru, V.; Ionescu, R. T.; Sebe, N.; and Shah, M. 2025. Curriculum direct preference optimization for diffusion and consistency models. In *CVPR*.

Cui, Y.; Ren, W.; and Knoll, A. 2024. Omni-kernel network for image restoration. In *Proceedings of the AAAI conference on artificial intelligence*, volume 38, 1426–1434.

Dong, C.; Loy, C. C.; He, K.; and Tang, X. 2015. Image super-resolution using deep convolutional networks. *IEEE TPAMI*.

Esser, P.; Kulal, S.; Blattmann, A.; Entezari, R.; Müller, J.; Saini, H.; Levi, Y.; Lorenz, D.; Sauer, A.; Boesel, F.; et al. 2024. Scaling rectified flow transformers for high-resolution image synthesis. In *ICML*.

Fei, B.; Lyu, Z.; Pan, L.; Zhang, J.; Yang, W.; Luo, T.; Zhang, B.; and Dai, B. 2023. Generative diffusion prior for unified image restoration and enhancement. In *CVPR*.

Franzen, R. 1999. Kodak Lossless True Color Image Suite. http://r0k.us/graphics/kodak/. Online accessed 24 Oct 2021.

Gou, Y.; Zhao, H.; Li, B.; Xiao, X.; and Peng, X. 2024. Test-time degradation adaptation for open-set image restoration. In *ICML*.

Gu, Y.; Wang, Z.; Yin, Y.; Xie, Y.; and Zhou, M. 2024. Diffusion-rpo: Aligning diffusion models through relative preference optimization. *arXiv preprint arXiv:2406.06382*.

Guo, H.; Guo, Y.; Zha, Y.; Zhang, Y.; Li, W.; Dai, T.; Xia, S.-T.; and Li, Y. 2025. Mambairv2: Attentive state space restoration. In *CVPR*.

Guo, Y.; Gao, Y.; Lu, Y.; Zhu, H.; Liu, R. W.; and He, S. 2024. Onerestore: A universal restoration framework for composite degradation. *arXiv preprint arXiv:2407.04621*.

Ho, J.; Jain, A.; and Abbeel, P. 2020. Denoising diffusion probabilistic models. *NeurIPS*.

Hong, J.; Paul, S.; Lee, N.; Rasul, K.; Thorne, J.; and Jeong, J. 2024. Margin-aware preference optimization for aligning diffusion models without reference. In *First Workshop on Scalable Optimization for Efficient and Adaptive Foundation Models*.

Kawar, B.; Elad, M.; Ermon, S.; and Song, J. 2022. Denoising diffusion restoration models. *NeurIPS*.

Ke, J.; Wang, Q.; Wang, Y.; Milanfar, P.; and Yang, F. 2021. Musiq: Multi-scale image quality transformer. In *ICCV*.

Kirstain, Y.; Polyak, A.; Singer, U.; Matiana, S.; Penna, J.; and Levy, O. 2023. Pick-a-pic: An open dataset of user preferences for text-to-image generation. *NeurIPS*.

Labs, B. F. 2024. FLUX. https://github.com/black-forest-labs/flux. Accessed: 2024-08-20.

Lee, K.; Li, X.; Wang, Q.; He, J.; Ke, J.; Yang, M.-H.; Essa, I.; Shin, J.; Yang, F.; and Li, Y. 2025. Calibrated multi-preference optimization for aligning diffusion models. In *CVPR*.

Li, B.; Li, X.; Zhu, H.; Jin, Y.; Feng, R.; Zhang, Z.; and Chen, Z. 2024a. Sed: Semantic-aware discriminator for image super-resolution. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 25784–25795.

Li, B.; Liu, X.; Hu, P.; Wu, Z.; Lv, J.; and Peng, X. 2022. All-in-one image restoration for unknown corruption. In *CVPR*.

Li, B.; Ren, W.; Fu, D.; Tao, D.; Feng, D.; Zeng, W.; and Wang, Z. 2018. Benchmarking single-image dehazing and beyond. *IEEE TIP*.

Li, T.; Feng, H.; Wang, L.; Zhu, L.; Xiong, Z.; and Huang, H. 2024b. Stimulating Diffusion Model for Image Denoising via Adaptive Embedding and Ensembling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Li, X.; Li, B.; Jin, Y.; Lan, C.; Zhu, H.; Ren, Y.; and Chen, Z. 2024c. UCIP: A universal framework for compressed image super-resolution using dynamic prompt. In *ECCV*.

Li, X.; Ren, Y.; Jin, X.; Lan, C.; Wang, X.; Zeng, W.; Wang, X.; and Chen, Z. 2023. Diffusion Models for Image Restoration and Enhancement–A Comprehensive Survey. *arXiv preprint arXiv:2308.09388*.

Liang, J.; Cao, J.; Sun, G.; Zhang, K.; Van Gool, L.; and Timofte, R. 2021. Swinir: Image restoration using swin transformer. In *ICCVW*.

Lin, Y.; Lin, Z.; Chen, H.; Pan, P.; Li, C.; Chen, S.; Wen, K.; Jin, Y.; Li, W.; and Ding, X. 2025. Jarvisir: Elevating autonomous driving perception with intelligent image restoration. In *CVPR*.

Meng, C.; He, Y.; Song, Y.; Song, J.; Wu, J.; Zhu, J.-Y.; and Ermon, S. 2022. Sdedit: Guided image synthesis and editing with stochastic differential equations. In *ICLR*.

Mittal, A.; Soundararajan, R.; and Bovik, A. C. 2012. Making a "completely blind" image quality analyzer. *IEEE Signal processing letters*.

Nah, S.; Son, S.; Lee, S.; Timofte, R.; Lee, K. M.; Chen, L.; Zhang, J.; Lu, X.; Chu, X.; Chen, C.; et al. 2021. NTIRE 2021 challenge on image deblurring. In *CVPRW*.

Ouyang, L.; Wu, J.; Jiang, X.; Almeida, D.; Wainwright, C.; Mishkin, P.; Zhang, C.; Agarwal, S.; Slama, K.; Ray, A.; et al. 2022. Training language models to follow instructions with human feedback. *NeurIPS*.

Özdenizci, O.; and Legenstein, R. 2023. Restoring vision in adverse weather conditions with patch-based denoising diffusion models. *IEEE TPAMI*.

Park, B.; Yu, S.; and Jeong, J. 2019. Densely connected hierarchical network for image denoising. In *CVPRW*.

Potlapalli, V.; Zamir, S. W.; Khan, S. H.; and Shahbaz Khan, F. 2023. Promptir: Prompting for all-in-one image restoration. *NeurIPS*.

Qin, X.; Wang, Z.; Li, F.; Chen, H.; Pei, R.; Li, W.; and Cao, X. 2025. CamEdit: Continuous Camera Parameter Control for Photorealistic Image Editing. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*.

Quan, Y.; Qin, X.; Pang, T.; and Ji, H. 2024. Siamese cooperative learning for unsupervised image reconstruction from incomplete measurements. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(7): 4866–4879.

Rafailov, R.; Sharma, A.; Mitchell, E.; Manning, C. D.; Ermon, S.; and Finn, C. 2023. Direct preference optimization: Your language model is secretly a reward model. *NeruIPS*.

Ren, Y.; Li, X.; Li, B.; Wang, X.; Guo, M.; Zhao, S.; Zhang, L.; and Chen, Z. 2025. Moe-diffir: Task-customized diffusion priors for universal compressed image restoration. In *European Conference on Computer Vision*, 116–134. Springer.

Rombach, R.; Blattmann, A.; Lorenz, D.; Esser, P.; and Ommer, B. 2022. High-resolution image synthesis with latent diffusion models. In *CVPR*.

Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; and Klimov, O. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.

Song, Y.; Shen, L.; Xing, L.; and Ermon, S. 2022. Solving inverse problems in medical imaging with score-based generative models. In *ICLR*.

Tu, Z.; Talebi, H.; Zhang, H.; Yang, F.; Milanfar, P.; Bovik, A.; and Li, Y. 2022. Maxim: Multi-axis mlp for image processing. In *CVPR*.

Wallace, B.; Dang, M.; Rafailov, R.; Zhou, L.; Lou, A.; Purushwalkam, S.; Ermon, S.; Xiong, C.; Joty, S.; and Naik, N. 2024. Diffusion model alignment using direct preference optimization. In *CVPR*.

Wang, J.; Chan, K. C.; and Loy, C. C. 2023. Exploring clip for assessing the look and feel of images. In *AAAI*.

Wang, J.; Yue, Z.; Zhou, S.; Chan, K. C.; and Loy, C. C. 2024. Exploiting diffusion prior for real-world image super-resolution. *International Journal of Computer Vision*, 132(12): 5929–5949.

Wang, X.; Xie, L.; Dong, C.; and Shan, Y. 2021. Real-esrgan: Training real-world blind super-resolution with pure synthetic data. In *ICCVW*.

Wang, Y.; Yu, J.; and Zhang, J. 2023. Zero-shot image restoration using denoising diffusion null-space model. In *ICLR*.

Wu, H.; Zhang, Z.; Zhang, W.; Chen, C.; Liao, L.; Li, C.; Gao, Y.; Wang, A.; Zhang, E.; Sun, W.; et al. 2024a. Q-align: Teaching lmms for visual scoring via discrete text-defined levels. In *ICML*.

Wu, R.; Sun, L.; Ma, Z.; and Zhang, L. 2024b. One-step effective diffusion network for real-world image super-resolution. *Advances in Neural Information Processing Systems*, 37: 92529–92553.

Wu, T.; Zou, J.; Liang, J.; Zhang, L.; and Ma, K. 2025. VisualQuality-R1: Reasoning-Induced Image Quality Assessment via Reinforcement Learning to Rank. *arXiv preprint arXiv:2505.14460*.

Yang, S.; Chen, T.; and Zhou, M. 2024. A dense reward view on aligning text-to-image diffusion with preference. In *ICML*.

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 *CVPR*.

Yang, W.; Tan, R. T.; Wang, S.; Fang, Y.; and Liu, J. 2020. Single image deraining: From model-based to data-driven and beyond. *IEEE TPAMI*.

You, Z.; Cai, X.; Gu, J.; Xue, T.; and Dong, C. 2025. Teaching large language models to regress accurate image quality scores using score distribution. In *CVPR*.

Zamir, S. W.; Arora, A.; Khan, S.; Hayat, M.; Khan, F. S.; and Yang, M.-H. 2022. Restormer: Efficient transformer for high-resolution image restoration. In *CVPR*.

Zhang, K.; Liang, J.; Van Gool, L.; and Timofte, R. 2021. Designing a practical degradation model for deep blind image super-resolution. In *ICCV*.

Zhang, K.; Ren, W.; Luo, W.; Lai, W.-S.; Stenger, B.; Yang, M.-H.; and Li, H. 2022a. Deep image deblurring: A survey. *IJCV*.

Zhang, W.; Li, D.; Min, X.; Zhai, G.; Guo, G.; Yang, X.; and Ma, K. 2022b. Perceptual attacks of no-reference image quality models with human-in-the-loop. *NeurIPS*.

Zhang, W.; Zhai, G.; Wei, Y.; Yang, X.; and Ma, K. 2023. Blind image quality assessment via vision-language correspondence: A multitask learning perspective. In *CVPR*.