ODDR: One-Step Deshadow Diffusion via Reward Guidance

Junseong Shin[1] **Kijun Kim**[2] **Minseong Kim**[2] **Dongjin Kim**[3] **Tae Hyun Kim**[1,3†]

[1]Department of Artificial Intelligence, Hanyang University
[2]Department of Applied Artificial Intelligence, Hanyang University
[3]Department of Computer Science, Hanyang University
{junsung6140, kijunkim, idecotu, dongjinkim, taehyunkim}@hanyang.ac.kr
[†]Corresponding author

Abstract

Recent advances in deep learning for shadow removal have significantly enhanced image quality and realism. However, most approaches rely on real-world paired datasets, which are costly to collect and often limited in scene diversity, leading to limited generalization. To address these limitations, we propose One-step Deshadow Diffusion via Reward guidance (ODDR), a new framework that achieves efficient and high-fidelity shadow removal without relying on real-world paired supervision. Our method begins with One-step Deshadow Diffusion (ODD), a baseline model trained on synthetic shadow data for efficient one-step shadow-free reconstruction. We further adapt ODD into ODDR using ShadowReward. In contrast to traditional, annotation-heavy approaches, ShadowReward is the first reward model for shadow removal trained entirely without human annotation. It learns to mimic human perceptual judgments by ranking synthetically generated images with controlled degradations, such as texture distortion and boundary artifacts. This reward-guided fine-tuning enables ODDR to narrow the synthetic-to-real domain gap. Extensive experiments show that ODD achieves strong performance without relying on real-world paired supervision, and ODDR further improves the results, narrowing the gap to fully supervised methods trained on real-world paired data while maintaining higher computational efficiency as a single-step model. Our project page is available here.

1 Introduction

Shadows introduce strong illumination mismatches and visible boundary transitions that harm perceptual quality [77] and hinder downstream vision tasks, such as text recognition [37] and semantic segmentation [27]. The goal of shadow removal is to eliminate these artifacts and restore images that are both natural in appearance and suitable for subsequent applications like robotics [53, 2].

Recent learning-based approaches substantially improve shadow removal quality through CNN and Transformer-based architectures [10, 9, 14, 70]. However, many of these approaches require strict supervision using paired real-world data, necessitating precise alignment between shadow images and their shadow-free ground truths. Unfortunately, collecting such real-world pairs is labor-intensive and often unreliable: scene changes, exposure variation, and illumination fluctuations make alignment difficult, and the resulting datasets are typically limited in diversity. This limitation leads to a noticeable performance gap between controlled benchmarks and real-world scenarios, where shadows exhibit complex geometry and color interactions. While previous works have proposed alternatives to bypass the need for real-world paired datasets, these approaches often come with their own inherent drawbacks. GAN-based [21, 41, 29] approaches rely on adversarial objectives to translate shadowed images into shadow-free ones, but often suffer from unstable training and inconsistent restoration

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

quality. Synthetic dataset generation [23] provides paired supervision by rendering shadows onto clean images, yet models trained on synthetic pairs often struggle to remove real-world shadows due to domain gaps. Recent diffusion-based approaches [16] address perceptual quality using generative priors, but their iterative sampling procedures lead to high inference latency, which limits practical usage in time-sensitive applications. Therefore, we aim to elevate perceptual fidelity and inference efficiency without relying on real-world paired datasets, achieving this by preserving the strong perceptual priors of diffusion models.

In this study, we observe that although diffusion priors offer a robust initialization, achieving real-world generalization necessitates training on real shadow artifacts that synthetic datasets fail to capture. To address this, we introduce a two-stage framework that integrates a diffusion-prior backbone with a novel, shadow-specific perceptual reward model.

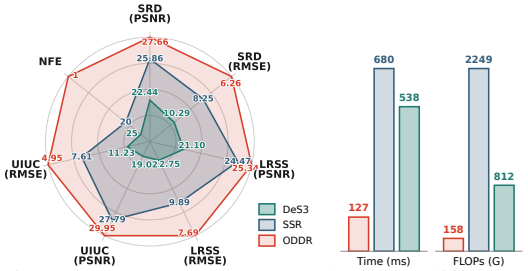

Figure 1: We compare our method with diffusion-based shadow removal models, DeS3 [26] and SSR [72] on SRD [50], LRSS [12], and UIUC [17] datasets, and analyze computational efficiency. Our ODDR achieves the highest restoration quality and simultaneously demands the least FLOPs, inference time, and NFE (Number of Function Evaluations).

In the first stage, we propose a baseline model, **O**ne-step **D**eshadow **D**iffusion (**ODD**), by adapting a pretrained text-to-image latent diffusion model to perform shadow removal in a single denoising step using only synthetic shadow and shadow-free pairs. This diffusion-prior-based training improves structural and color consistency while avoiding common instabilities of GAN-based training. In addition, compared to multi-step diffusion [15, 16], this one-step generation drastically reduces the number of function evaluations (NFE) at inference.

In the second stage, we refine ODD into **O**ne-step **D**eshadow **D**iffusion via **R**eward guidance (**ODDR**). Notably, we employ reward-guided optimization directly on images containing real shadows, without requiring corresponding shadow-free references. Specifically, we introduce **ShadowReward**, a reward model tailored to shadow removal that approximates human perceptual judgments without relying on manual annotation. ShadowReward is trained by ranking synthetically corrupted images with controlled degradations that mimic common shadow removal failures, such as texture distortion and boundary artifacts. During ODDR fine-tuning, the frozen ODD prediction acts as a pseudo target to preserve structural fidelity, while the reward term encourages outputs that receive higher scores from ShadowReward. By enabling real-world refinement without real-world paired datasets, this strategy narrows the synthetic-to-real domain gap and enhances perceptual quality without compromising the baseline content or color tone.

Fig. 1 summarizes the performance of ODDR compared to other diffusion-based models [26, 72], demonstrating its strong efficiency and robust generalization capabilities on unseen datasets [12, 17]. Through extensive experiments, we show that the proposed method outperforms existing methods that do not use real-world paired data, and substantially narrows the gap to fully supervised methods trained on real-world paired datasets.

2 Related Works

2.1 Single-Image Shadow Removal

Shadow removal methods learn to reconstruct shadow-free images from shadowed inputs by modeling illumination and appearance changes caused by shadows. Recent approaches utilizing CNNs enhance exposure balance and boundary smoothness via multi-exposure fusion [11], aggregation based on matting [10], style harmonization [61], and invertible mappings [78]. Transformer models are particularly effective in capturing global context and long-range dependencies, which facilitates consistent illumination restoration [14, 70, 71]. Additionally, diffusion-based approaches contribute to stabilizing texture recovery and preserving details by employing adaptive attention and latent diffusion with detail injection, resulting in more realistic appearances [15, 72]. However, diffusion variants, despite their perceptual improvements, come with significant inference costs. Moreover, these models are still prone to exposure bias in the training data [62, 50].

To overcome the reliance on paired datasets, various methods have been extensively explored for shadow removal. GAN-based approaches employ adversarial or cycle objectives, often with mask guidance, domain-classifier cues, lightness guidance, or non-cyclic unidirectional design [21, 25, 40, 32, 41, 29], enhancing color fidelity and penumbra smoothness without aligned pairs. BCDiff [16], a diffusion-based method, removes shadows by separating boundary and inner-shadow regions, using non-shadow context for boundary areas and intrinsic decomposition for inner-shadow recovery. SynShadow [23] introduces a physics-based synthetic shadow generation method that produces diverse shadow shapes and intensities without paired labels, improving the robustness of deshadowing networks. However, while such methods eliminate the need for paired data, they typically underperform in shadow removal compared to paired approaches.

2.2 One-Step Diffusion Models

Diffusion models [19, 52] generate high-quality images through iterative denoising steps that gradually refine random noise into realistic samples. While this iterative process enables strong generative capability, it also results in high inference latency and computational overhead in shadow removal [15, 72]. To address these issues, recent studies have explored one-step diffusion models [75, 65] that distill the entire multi-step denoising trajectory into a single forward pass. This architecture enables efficient task adaptation through lightweight fine-tuning mechanisms, such as LoRA [20]. These one-step models have demonstrated impressive efficiency and performance across various image restoration tasks such as super-resolution [58], JPEG artifact removal [13], and deblurring [39]. Inspired by these successes and by recent findings that pretrained diffusion models can be efficiently fine-tuned for single-step prediction [44], we adapt the one-step diffusion paradigm to shadow removal. Rather than the one-step formulation itself, our contribution lies in the architectural designs that make it effective for shadow removal and in the annotation-free reward guidance described below.

2.3 Reward Models and Preference Supervision

Reward modeling has become a central component in reinforcement learning with human feedback, where a learned reward function is optimized to align model outputs with human preferences [80, 6]. Such approaches have recently gained substantial attention in large language models [46, 68]. Beyond language modeling, preference-based learning has also been extended to various computer vision tasks, including face restoration [64], image generation [73, 36, 28, 67], video generation [34], and 3D generation [74]. These methods demonstrate that reward supervision can effectively improve perceptual quality by aligning model outputs with human judgments. However, existing reward models typically rely on preference datasets constructed from large-scale human annotations, which are costly and labor-intensive to collect. This dependency limits their scalability and practical applicability, particularly in low-level vision tasks where fine-grained perceptual differences are critical. To alleviate this limitation, recent image restoration studies have explored automated preference supervision using vision-language models (VLMs) [5] or perceptual Image Quality Assessment (IQA) metrics in super-resolution tasks [66, 49]. Nevertheless, such strategies have not yet been specifically tailored to the shadow removal problem. In this work, we introduce the first reward model designed for shadow removal. Notably, our proposed model is trained entirely without human annotations, yet provides human-aligned perceptual guidance to improve restoration quality.

3 Proposed Method

3.1 Overall Framework

Our proposed framework consists of two primary components: (i) One-step Deshadow Diffusion (ODD), a cost-effective one-step latent diffusion for deshadowing, and (ii) One-step Deshadow Diffusion via Reward guidance (ODDR), an enhanced version of ODD for real-world scenarios guided by a preference-based reward model (ShadowReward). As shown in Fig. 2, we first train ODD by adapting Stable Diffusion (SD) [52] using synthetic shadow and shadow-free pairs [23]. Subsequently, we refine ODD through ShadowReward, which provides perceptual feedback to assist in the fine-tuning process. This results in our final model, ODDR, which significantly enhances shadow removal performance in real scenarios. In the following, Sec. 3.2 describes the ODD baseline model, Sec. 3.3 introduces the reward-guided refinement for ODDR including the ShadowReward model and its human-annotation-free training procedure.

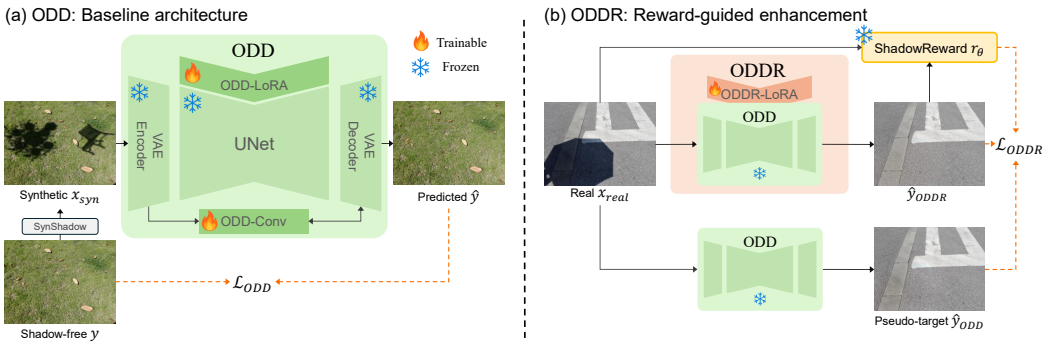

Figure 2: **Overall training pipeline.** (a) The baseline ODD is trained on synthetic shadow images (x_{syn}), optimizing the ODD-LoRA and ODD-Conv modules. (b) ODDR is then fine-tuned from ODD parameters using real-world shadow images (x_{real}) guided by ShadowReward (r_θ) model, optimizing ODDR-LoRA module.

3.2 One-Step Deshadow Diffusion (ODD)

3.2.1 Motivation

In general, deshadow models trained without real-world paired datasets struggle to match the performance of fully supervised approaches in real scenarios. To bridge this gap, we harness the image priors of Stable Diffusion (SD) [52] to construct our baseline model, ODD, which significantly enhances deshadowing quality without using the real-world paired dataset. However, SD relies on iterative denoising that requires many inference steps, leading to high computational cost and latency. Furthermore, because intermediate diffusion states remain distant from the final image, providing direct pixel-level feedback from the reward model is very challenging. To overcome these limitations, we adopt the one-step formulation from OSEDiff [65], which replaces stochastic multi-step denoising with a single deterministic mapping that directly predicts a shadow-free image.

Architecture The architecture of ODD is built upon SD, which consists of a latent variational autoencoder (VAE) and a UNet-based denoising network pre-trained on large-scale datasets [54], providing a robust foundation from large-scale pre-training. We introduce several key modifications to adapt this architecture for the one-step shadow removal task.

First, as shown in Fig. 2 (a), we adapt the SD UNet with LoRA [20], inserting rank-16 low-rank adapters into its attention and convolution blocks (ODD-LoRA). This partial adaptation enables the model to effectively adjust to the shadow removal task while keeping the original SD weights frozen. Despite this benefit, simply repurposing the SD architecture can lead to a loss of fine details, which are crucial for the shadow removal task. To address this limitation, we add lightweight two-layer convolutional skip connections (ODD-Conv) at each encoder-decoder level, which assist in maintaining fine details and recovering high-frequency textures. Furthermore, since our model does not take text prompts as input, we replace the original SD's text-conditioning modules with learnable self-attention layers.

Training Objective Given a shadow-free image y, we first synthesize a corresponding shadowed image x_{syn} through SynShadow [23] and train ODD to predict $\hat{y}$ that matches y by minimizing ℓ_1-based reconstruction and perceptual losses:

$$\mathcal{L}_{\text{ODD}} = \|\hat{y} - y\|_1 + \lambda_{\text{per}}\mathcal{L}_{\text{per}}(\hat{y}, y), \tag{1}$$

where $\mathcal{L}_{\text{per}}$ is the LPIPS loss [76] and λ_{per} is a hyperparameter balancing the two losses. This objective encourages structural fidelity and perceptual realism, serving as the initialization for subsequent reward-guided fine-tuning.

3.3 One-Step Deshadow Diffusion via Reward Guidance (ODDR)

Although ODD achieves effective reconstruction as a baseline, it still presents two significant drawbacks. First, because it is trained exclusively on synthetically shadowed images, a noticeable domain gap remains when applied to real-world shadows with complex lighting and color. Second, while LPIPS enforces feature-level alignment with the ground truth, it frequently overlooks visually salient artifacts inherent to shadow removal, including texture degradation, noise, color deviations, luminance fluctuations, and boundary seams. To overcome these limitations, we present ODDR, which adapts the baseline model (ODD) with a perceptual reward model (ShadowReward) to better align with human visual judgments.



Figure 3: An example list $\{x^{(j)}\}_{j=0}^{3}$ of the human-annotation-free dataset. Larger j means more degradation.

As illustrated in Fig. 2 (b), our process begins with two parallel instances of the pretrained ODD (top and bottom branches) with frozen weights. Given a real-world shadow input x_{real}, the bottom branch (ODD) generates a pseudo-target $\hat{y}_{\text{ODD}}$. Subsequently, the top branch (ODDR) undergoes partial fine-tuning through LoRA [20] (ODDR-LoRA), and is optimized by the objective function with a frozen reward function r_θ as follows:

$$\mathcal{L}_{\text{ODDR}} = \|\hat{y}_{\text{ODDR}} - \hat{y}_{\text{ODD}}\|_2 + \lambda_{\text{align}}\,\mathcal{L}_{\text{per}}(\hat{y}_{\text{ODDR}}, \hat{y}_{\text{ODD}}) - \lambda_r\, r_\theta(x_{\text{real}}, \hat{y}_{\text{ODDR}}), \tag{2}$$

where $\hat{y}_{\text{ODDR}}$ denotes the ODDR (top branch) output. To train ODDR without ground-truth shadow-free images, the loss function combines two main objectives. First, ℓ_2 loss and perceptual loss $\mathcal{L}_{\text{per}}$ (weighted by λ_{align}) maintain structural fidelity to the baseline prediction (ODD). Second, a frozen reward function r_θ (weighted by λ_r) subtracts a reward score that reflects human preference from the overall loss, successfully preserving data fidelity while guiding the model toward perceptually enhanced, boundary-consistent restorations.

Ranking Deshadow Preference (ShadowReward) In this work, we employ a reward model to fine-tune the baseline ODD model, which significantly enhances the perceptual quality of the final output. In the context of the reward model, human visual preference serves as a natural standard for assessing the quality of shadow removal results, and observers primarily evaluate the quality of restoration through boundary smoothness, color consistency, and texture realism [60]. However, collecting the human-annotated preference data required to train such reward models is a time-consuming and expensive process. To address this problem, we present **ShadowReward**, a reward model designed to mimic human preference without the need for human annotation to assess the quality of shadow removal.

Previous studies have shown that the level of image degradation influences perceptual preference, with more severely degraded images consistently judged as less favorable [47, 30]. Motivated by this observation, we design a training strategy that leverages controlled degradations to induce preference ordering, enabling the reward model to learn human-like judgments without explicit annotation. Specifically, to simulate scenarios where conventional shadow removal models generally do not perform well, we corrupt images with specific degrees of texture distortion, noise, chromatic shift, luminance change, and boundary artifacts. Subsequently, we train ShadowReward to rank these corrupted images according to their degree of perceptual degradation.

Human-Annotation-Free Dataset To train our ShadowReward model, we construct a Human-Annotation-Free dataset to enable preference learning without human labels. Specifically, following SynShadow [23], we first generate a synthetic shadow image x_{syn} from a corresponding shadow-free image y and a random shadow mask. We then create a list of degraded images $\{x^{(j)}\}_{j=0}^{J}$, where each $x^{(j)}$ exhibits a progressively increasing degree of degradation corresponding to its index j. As described in Fig. 3, a higher index indicates more degradation and lower perceptual quality, while $x^{(0)}$ is assigned as the clean image y. Through empirical evaluation, we set $J = 3$.

To manage the extent of each degradation, we initially establish a series of degradation types along with their respective degradation levels. These include texture distortion, additive noise, chromatic shift, luminance change, and boundary artifacts. Each is applied with a range of degradation levels to control degradation strength. This process allows us to construct diverse and perceptually ordered training samples without any human annotation. Detailed settings are provided in our Appendix (Sec. A).

ShadowReward Architecture In contrast to conventional reward models based on VLMs [51, 35], our shadow removal task requires nuanced spatial reasoning and sensitiv-

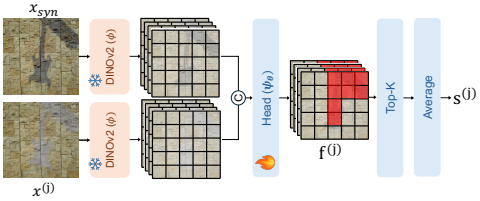

Figure 4: Overview of ShadowReward.

ity to local texture/illumination variations rather than high-level semantics. Thus, as shown in Fig. 4, we use DINOv2 [45] in our ShadowReward to capture spatial correspondence and discriminative local features for shadow removal. Specifically, we extract patch-level tokens from both the shadow input x_{syn} and a corrupted image $x^{(j)}$ using the DINOv2 encoder (ϕ). Next, these token sequences are concatenated and passed to a lightweight convolutional head ψ_θ, which produces a spatial map as follows:

$$f^{(j)} = \psi_\theta([\phi(x_{\text{syn}}), \phi(x^{(j)})]), \tag{3}$$

where $[\cdot, \cdot]$ denotes the concatenation operation. The concatenated features are processed by ψ_θ to capture the relative discrepancies between the shadowed and corrupted images, producing patch-wise perceptual features $f^{(j)}$. Then, $f^{(j)}$ is mapped into a final scalar reward $s^{(j)}$, which serves as a perceptual quality score for $x^{(j)}$. To compute this score, we apply Top-K pooling, averaging only the largest responses of $f^{(j)}$ so that the reward focuses on informative regions that best reflect perceptual differences. In practice, we set $K = 0.1$ (top 10% responses). The scalar reward is computed as

$$s^{(j)} = \frac{1}{|\mathcal{T}_K|} \sum_{i \in \mathcal{T}_K} f_i^{(j)}, \tag{4}$$

where $\mathcal{T}_K$ denotes the indices of the top-K responses in $f^{(j)}$.

Pre-training ShadowReward Assessing perceptual quality through ShadowReward entails comparing differently corrupted images and determining their relative quality. It is understood that determining relative rankings is easier and more dependable than scoring absolute quality. This is because measuring absolute quality is challenging due to the variety of images. To this end, we train ShadowReward r_θ to rank corrupted images $\{x^{(j)}\}_{j=0}^{J}$, assigning higher scores to those with better perceptual quality. To train ShadowReward concerning this ranking behavior, we utilize the ListMLE loss [69] denoted by $\mathcal{L}_{\text{list}}$. This loss function models the likelihood of achieving the correct sequence of a list of distorted images, and it is defined as follows:

$$\mathcal{L}_{\text{list}} = -\frac{1}{J+1} \sum_{i=0}^{J} \log \frac{\exp\big(s^{(i)}/\tau\big)}{\sum_{t=i}^{J} \exp\big(s^{(t)}/\tau\big)}, \tag{5}$$

where the temperature parameter τ regulates how decisively the model distinguishes between corrupted images in the ranking process. We set $\tau = 1$ by default. A higher score $s^{(j)}$ indicates better perceptual quality, meaning that images with fewer artifacts or distortions receive larger rewards. This loss enforces a higher reward for perceptually better images, which enables the reward model to assess the quality of shadow removal results. Note that our ShadowReward model can be trained separately from ODD and ODDR.

4 Experiments

4.1 Experimental Setup

Datasets We use the Adjusted ISTD (AISTD) dataset [62, 31] for training, which contains 1,870 triplets of shadow, shadow mask, and shadow-free images with a resolution of 480×640, where 1,330 triplets are used for training and 540 for testing. We utilize different components of the AISTD training set for each stage of our framework. ODD is trained solely on shadow-free images using synthetic shadow images generated from clean images in this dataset. For ODDR, we use only shadowed images from the training set. ShadowReward is trained on our Human-Annotation-Free dataset, which is constructed by augmenting AISTD shadow-free images with random shadow masks with SynShadow [23].

To evaluate the generalization ability of our model, we conduct tests on three unseen real-world datasets, SRD [50], LRSS [12], and UIUC [17]. The SRD and LRSS datasets contain 408 and 34 test pairs, respectively; LRSS consists of soft shadows, allowing us to assess robustness to faint shadow boundaries. For the UIUC dataset, we use 44 test pairs with consistent light conditions from the original 76 pairs. These datasets provide diverse illumination and shadow configurations that differ from AISTD, enabling a fair and rigorous evaluation of cross-dataset generalization.

Evaluation Metrics Following previous works [31, 11], we use three standard metrics: Root Mean Square Error (RMSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity (SSIM) [63].

Table 1: Quantitative comparison on AISTD [31] dataset. **Red** and blue denote the best and second-best results, evaluated among the methods trained without real-world pairs. ✓ denotes mask-free methods and ✗ indicates methods that require shadow masks. $\mathcal{D}$ indicates whether real-world paired data is used for training.

$\mathcal{D}$	Methods	Mask-free	Shadow Region			Non-Shadow Region			All Image		
			PSNR↑	SSIM↑	RMSE↓	PSNR↑	SSIM↑	RMSE↓	PSNR↑	SSIM↑	RMSE↓
	Real-world Shadow Image	–	20.83	0.927	39.01	37.46	0.985	2.40	20.46	0.894	8.40
w/ real-world pair	DHAN [10]	✗	32.92	0.988	9.6	27.15	0.971	7.4	25.66	0.956	7.8
	SP+M-Net [31]	✗	37.60	0.990	6.3	36.02	0.976	3.0	32.94	0.962	3.5
	Fu et al. [11]	✗	36.04	0.978	6.7	31.16	0.892	3.8	29.45	0.861	4.2
	BMNet [78]	✗	37.87	0.991	5.62	37.51	0.985	2.45	33.98	0.972	2.97
	SG-ShadowNet [61]	✗	36.80	0.990	5.93	35.57	0.978	2.92	32.46	0.962	3.41
	DeS3 [26]	✓	36.49	0.989	6.59	34.72	0.972	3.32	31.39	0.957	3.86
	OmniSR [71]	✓	37.01	0.990	6.57	37.82	0.984	2.40	33.35	0.970	3.08
	SSR [72]	✓	39.94	0.993	4.23	38.04	0.983	2.63	35.19	0.974	2.89
w/o real-world pair	MaskShadow-GAN [21]	✗	32.19	0.984	10.8	33.44	0.974	3.8	28.81	0.946	4.8
	LG-ShadowNet [40]	✗	32.45	0.982	9.65	33.73	0.975	3.25	29.22	0.947	4.30
	DCS [25]	✓	32.00	0.976	10.41	33.56	0.968	3.63	28.77	0.932	4.74
	Le et al. [32]	✗	33.09	0.983	9.67	35.35	0.978	2.82	30.15	0.951	3.94
	G2R [41]	✗	33.58	0.979	8.82	35.52	0.976	2.89	30.52	0.944	3.86
	BCDiff [16]	✗	35.91	0.986	7.6	37.27	**0.984**	2.4	32.73	0.962	3.3
	S3R-Net [29]	✓	32.08	0.983	12.03	33.61	0.978	3.55	29.05	0.952	9.94
	ODD (ours)	✓	36.67	0.990	6.27	**37.45**	**0.984**	**2.49**	33.32	0.971	3.11
	ODDR (ours)	✓	**38.37**	**0.991**	**5.36**	37.34	**0.984**	2.50	**34.20**	**0.972**	**2.97**

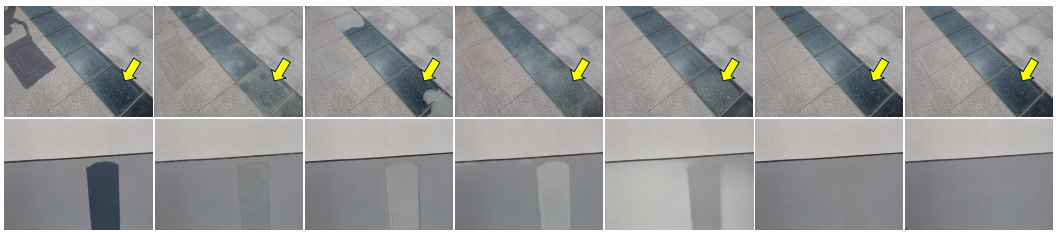

Figure 5: Visual comparison on AISTD [31]. Top: non-shadow preservation (yellow arrows mark competitor failures). Bottom: shadow removal. Ours performs well in both.

RMSE is computed in the LAB color space to measure color consistency, while PSNR and SSIM are calculated in the RGB color space. Lower RMSE indicates smaller reconstruction errors, while higher PSNR and SSIM denote better restoration quality. All metrics are computed on images of size 256×256 using the MATLAB evaluation scripts from previous work [79].

Implementation Details For ODD, we adopt the pretrained SD v2.1 model [52] as our backbone. The model is trained on 512×512 inputs, where smaller images are padded to this size using reflect padding. For optimization, we use the AdamW optimizer [43] with a batch size of 4 for 40K iterations, and a learning rate decaying from 1×10^{-4} to 1×10^{-6} via a cosine annealing scheduler [42]. For ShadowReward, we use 448×448 inputs, a batch size of 16, and the AdamW optimizer for 20K iterations with the same learning rate decay strategy as ODD. For ODDR, we use an identical setup to ODD, except that it is trained for 20K iterations with a fixed learning rate of 1×10^{-6}. All experiments are conducted on a single NVIDIA RTX A6000 GPU.

4.2 Quantitative and Qualitative Comparisons

We compare the proposed method with various shadow removal approaches across different training data settings. For methods trained with real-world pairs, we evaluate against DHAN [10], SP+M-Net [31], Fu et al. [11], BMNet [78], SG-ShadowNet [61], and OmniSR [71]. Moreover, we incorporate recent diffusion-based shadow removal models such as DeS3 [26] and SSR [72] to evaluate our single-step method over multi-step diffusion processes. For methods trained without real-world paired datasets, we compare with MaskShadow-GAN [21], DCS [25], LG-ShadowNet [40], G2R [41], BCDiff [16], and S3R-Net [29].

Quantitative Results Tab. 1 reports the main comparison results on the AISTD dataset, with metrics computed on the shadow, non-shadow, and whole-image regions. First, our ODD performs competitively among methods without using real-world pairs. Notably, our final reward-guided model (ODDR) further enhances shadow-region performance, increasing PSNR from 36.67 dB to 38.37 dB and reducing RMSE from 6.27 to 5.36. Since both alignment terms in Eq. (2) are minimized at

Table 2: Results on SRD, LRSS, and UIUC. The best results are highlighted in **red**.

Method	SRD [50]			LRSS [12]			UIUC [17]		
	PSNR↑	SSIM↑	RMSE↓	PSNR↑	SSIM↑	RMSE↓	PSNR↑	SSIM↑	RMSE↓
S3R-Net [29]	22.51	0.878	10.60	22.54	0.786	10.45	23.92	0.857	8.53
DCS [25]	21.99	0.868	11.47	22.10	0.774	11.63	23.94	0.852	9.69
G2R [41]	24.23	0.898	8.40	22.79	0.778	9.64	27.03	0.867	6.23
DeS3 [26]	22.44	0.900	10.29	21.10	0.767	12.75	19.02	0.822	11.23
SSR [72]	25.86	0.922	8.25	24.47	0.790	9.89	27.79	0.873	7.61
ODD	27.38	0.941	6.46	25.12	0.797	7.84	29.25	0.886	5.10
ODDR	**27.66**	**0.942**	**6.26**	**25.34**	**0.797**	**7.69**	**29.95**	**0.887**	**4.95**



Input DCS G2R SSR ODDR

Figure 6: Visual comparisons on SRD [50], LRSS [12], and UIUC [17] (top to bottom).

the ODD output, training with λ_r=0 reproduces ODD (Appendix Sec. E.1); this gain is therefore attributable to the reward.

While some methods trained on real-world paired datasets show higher PSNR in non-shadow regions than the input, this improvement often arises from illumination discrepancies in paired data, leading models to learn global brightness corrections rather than pure shadow removal [21]. Therefore, we measure cross-dataset generalization performance on three unseen real-world datasets, SRD, LRSS, and UIUC, without additional training. For a fair comparison, all results are obtained with the parameters trained on the AISTD dataset. As shown in Tab. 2, our ODD model achieves superior performance compared to conventional approaches, surpassing methods trained both with and without real-world paired data by a significant margin. Although ODD has strong cross-dataset generalization, ODDR consistently outperforms ODD and achieves higher PSNR and SSIM, and lower RMSE across all datasets, confirming its improved generalization performance. We note that the pretrained SD prior shared by ODD and ODDR contributes substantially to this generalization, while the reward adds a smaller but consistent refinement on top of it (+0.28 / +0.22 / +0.70 dB PSNR on SRD / LRSS / UIUC).

Qualitative Results Visual comparisons are presented in Fig. 5 for the AISTD dataset, and in Fig. 6 for the unseen SRD, LRSS, and UIUC datasets. On AISTD, methods trained without real-world paired data often reveal dark boundaries or exhibit color suppression around the boundary where a shadow fades into light, whereas our method produces smoother transitions and better tone harmonization across shadow and non-shadow regions. We observe a similar trend on the unseen SRD, LRSS, and UIUC datasets, where our model surpasses existing approaches in restoring natural illumination and suppressing color artifacts.

4.3 Ablation Study

Impact of ShadowReward Model To validate the effectiveness of our ShadowReward model, we conduct two complementary evaluations summarized in Tab. 3. For comparison, existing reward models such as Aesthetic Score Predictor [54], ImageReward [73], and PrefPaint [38] are used. Moreover, IQA metrics are also compared such as NR-IQAs proposed in [66] and PSNR. We perform an ablation study by substituting ShadowReward in the ODDR training pipeline (Fig. 2 (b)) with the aforementioned rewarding methods to compare their deshadowing performance on the AISTD dataset. Consequently, ODDR fine-tuned with ShadowReward outperforms the alternatives, achieving the best PSNR, SSIM, and RMSE values. The alignment with human perceptual preferences is further

Table 3: Comparison of different reward models and IQA metrics on the AISTD dataset. NR-IQAs denotes the combination of the non-reference IQA models proposed in [66]. ACC denotes the accuracy (%) of alignment with human preference ranking.

Method	ACC↑	PSNR↑	SSIM↑	RMSE↓
Aesthetic Score Predictor	12.50	33.29	0.971	3.11
ImageReward	18.75	33.29	0.971	3.11
PrefPaint	25.00	33.43	0.971	3.09
NR-IQAs	18.75	33.34	0.971	3.10
PSNR	56.25	–	–	–
ShadowReward (ours)	**68.75**	**34.20**	**0.972**	**2.97**

Table 4: Ablation on cumulative degradations for ShadowReward. Abbreviations: T (Texture distortion), N (Additive noise), C (Chromatic shift), L (Luminance change), B (Boundary artifact).

Degradation	Shadow		ALL	
	PSNR↑	RMSE↓	PSNR↑	RMSE↓
T	36.99	6.06	33.50	3.07
T, N	37.17	5.95	33.61	3.05
T, N, C	37.82	5.70	33.96	3.01
T, N, C, L	38.20	5.47	34.16	**2.97**
T, N, C, L, B	**38.37**	**5.36**	**34.20**	**2.97**

Table 5: Results of ODD under different conditioning strategies.

Condition	Shadow			ALL		
	PSNR↑	SSIM↑	RMSE↓	PSNR↑	SSIM↑	RMSE↓
Null	35.95	0.990	6.79	32.95	0.971	3.18
BLIP	36.29	0.990	6.54	33.10	0.971	3.16
ODD (ours)	**36.67**	**0.990**	**6.27**	**33.32**	**0.971**	**3.11**

Table 6: Comparison of efficiency among diffusion shadow removal models.

Method	NFE↓	Inference Time↓	FLOPs (G)↓
ShadowDiffusion [15]	25	581.96	937.15
DeS3 [26]	25	538.00	812.26
SSR [72]	20	680.64	2249.50
ODD (ours)	**1**	**113.18**	**147.22**
ODDR (ours)	**1**	127.41	158.21

analyzed in the **Human Preference Alignment** paragraph below. Note that PSNR-based results are shown as '–' because PSNR relies on ground-truth shadow-free images, which are not used during our training process.

Human Preference Alignment To quantitatively assess how accurately ShadowReward reflects human perceptual preferences for shadow removal, we perform an extensive user study involving 23 participants. Specifically, these individuals provided preference rankings for images generated by several shadow removal methods [40, 72, 41, 29], including our ODD and ODDR models, using the AISTD dataset. Subsequently, we rank the outcomes based on the predicted scores from each reward model and evaluate how closely these rankings align with human preferences.

In Tab. 3, the results for accuracy (ACC) confirm that ShadowReward aligns closely with human perceptual preferences. In contrast, existing reward models and rankings based on a non-reference image quality assessment metric (NR-IQA) exhibit low alignment accuracies (below 30%), as they rely on generic aesthetic or distortion cues rather than shadow-specific perceptual signals. Notably, a PSNR-based ranking yields an accuracy of 56.25% but necessitates ground-truth references for computation. ShadowReward reaches 68.75% alignment without any ground truth; while the 16-question study is too small to establish a statistically significant advantage over PSNR, it shows that ShadowReward is at least on par with pixel-level comparisons while requiring no reference image (see Appendix Sec. C for inter-annotator agreement). More details on human evaluation are provided in the Appendix Sec. C.

Effect of Degradations for Human-Annotation-Free Dataset Beyond comparing ShadowReward against alternative reward models, we further analyze how the diversity of synthetic corruptions used to train ShadowReward affects its effectiveness. The types are texture distortion, additive noise, chromatic shift, luminance change, and boundary artifacts, as detailed in Appendix Tab. 7. As shown in Tab. 4, incorporating multiple degradations progressively enhances deshadow performance. Employing all five types of degradation results in the highest PSNR and the lowest RMSE in both Shadow and ALL regions, indicating that a variety of jointly managed corruptions offer more robust supervision for learning.

Conditioning Strategy We next examine the architectural choices made in the ODD baseline. Since our ODD modifies the text-conditioning module of SD, we conduct an ablation study to validate this architectural change in Tab. 5. We first input a null prompt to ODD with the text-conditioning module to measure deshadow performance. We use BLIP [35] to extract text from the shadow image, input it to ODD with the text-conditioning module, and measure performance. We compare these results with our ODD, replacing the text-conditioning with self-attention. Conditioning on text generated by BLIP results in a slight enhancement compared to using a null prompt, whereas the suggested self-conditioning method (last row) delivers better performance across both Shadow and ALL regions.

Computational Cost We conduct an ablation on computational efficiency to compare the inference cost with other diffusion-based shadow removal models. All measurements are performed on a single NVIDIA RTX 4090 GPU with an input resolution of 256×256. As summarized in Tab. 6, multi-step diffusion models (ShadowDiffusion [15], DeS3 [26], and SSR [72]) require more than 20 denoising steps, resulting in large FLOPs and slow inference. Conversely, our single-step models (ODD and ODDR) decrease inference time by over fourfold while preserving similar performance levels.

5 Conclusion

We introduced One-step Deshadow Diffusion via Reward guidance (ODDR), an efficient framework for high-quality shadow removal that requires no real-world paired supervision. Our approach builds upon One-step Deshadow Diffusion (ODD), a baseline that performs single-step shadow-free reconstruction using only synthetic training data, and further refines it through a perceptual reward signal aligned with human preference. At the core of this process lies ShadowReward, a reward model trained on synthetically degraded images that mimic typical failure cases in shadow removal, guiding ODDR toward perceptually pleasing outputs without requiring ground-truth annotations. Extensive experiments demonstrate that our approach achieves state-of-the-art performance among methods without real-world paired supervision, shows strong generalization across diverse datasets, and remains highly efficient compared to multi-step diffusion-based models. Further analyses on reward-weight sensitivity, shadow-free inputs, and transfer beyond cast shadows are provided in Appendix Secs. E and G, and a discussion of limitations, representative failure cases, and broader impacts in Appendix Sec. H.

Acknowledgments and Disclosure of Funding

This research was supported by Culture, Sports and Tourism R&D Program through the Korea Creative Content Agency grant funded by the Ministry of Culture, Sports and Tourism in 2026 (Project Name: Development of AI-Based Animation Production Technology to Ensure Character and Scene Consistency and Continuity for Enhanced Efficiency and Quality, Project Number: RS-2026-25525207, Contribution Rate: 30%). This work was also supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2022-0-00156, Fundamental research on continual meta-learning for quality enhancement of casual videos and their 3D metaverse transformation), and IITP grant funded by the Korea government (MSIT) (No.RS-2020-II201373, Artificial Intelligence Graduate School Program (Hanyang University)).

References

[1] Abdelrahman Abdelhamed, Stephen Lin, and Michael S Brown. A high-quality denoising dataset for smartphone cameras. In *CVPR*, 2018.

[2] Luma Issa Abdul-Kreem et al. Shadow detection and elimination for robot and machine vision applications. *Scientific Visualization*, 2024.

[3] Codruta O Ancuti, Cosmin Ancuti, and Radu Timofte. NH-HAZE: An image dehazing benchmark with non-homogeneous hazy and haze-free images. In *CVPRW*, 2020.

[4] Yunhao Ba, Howard Zhang, Ethan Yang, Akira Suzuki, Arnold Pfahnl, Chethan Chinder Chandrappa, Celso M de Melo, Suya You, Stefano Soatto, Alex Wong, and Achuta Kadambi. Not just streaks: Towards ground truth for single image deraining. In *ECCV*, 2022.

[5] Shuai Bai, Keqin Chen, Xuejing Liu, Jialin Wang, Wenbin Ge, Sibo Song, Kai Dang, Peng Wang, Shijie Wang, Jun Tang, et al. Qwen2. 5-vl technical report. *arXiv preprint arXiv:2502.13923*, 2025.

[6] Yuntao Bai, Andy Jones, Kamal Ndousse, Amanda Askell, Anna Chen, Nova DasSarma, Dawn Drain, Stanislav Fort, Deep Ganguli, Tom Henighan, et al. Training a helpful and harmless assistant with reinforcement learning from human feedback. *arXiv preprint arXiv:2204.05862*, 2022.

[7] Tianrun Chen, Lanyun Zhu, Chaotao Deng, Runlong Cao, Yan Wang, Shangzhan Zhang, Zejian Li, Lingyun Sun, Ying Zang, and Papa Mao. Sam-adapter: Adapting segment anything in underperformed scenes. In *ICCVW*, 2023.

[8] Xinlei Chen, Saining Xie, and Kaiming He. An empirical study of training self-supervised vision transformers. In *ICCV*, 2021.

[9] Zipei Chen, Chengjiang Long, Ling Zhang, and Chunxia Xiao. Canet: A context-aware network for shadow removal. In *ICCV*, 2021.

[10] Xiaodong Cun, Chi-Man Pun, and Cheng Shi. Towards ghost-free shadow removal via dual hierarchical aggregation network and shadow matting gan. In *AAAI*, 2020.

[11] Lan Fu, Changqing Zhou, Qing Guo, Felix Juefei-Xu, Hongkai Yu, Wei Feng, Yang Liu, and Song Wang. Auto-exposure fusion for single-image shadow removal. In *CVPR*, 2021.

[12] Maciej Gryka, Michael Terry, and Gabriel J Brostow. Learning to remove soft shadows. *ACM TOG*, 34(5):1–15, 2015.

[13] Jinpei Guo, Zheng Chen, Wenbo Li, Yong Guo, and Yulun Zhang. Compression-aware one-step diffusion model for jpeg artifact removal. *arXiv preprint arXiv:2502.09873*, 2025.

[14] Lanqing Guo, Siyu Huang, Ding Liu, Hao Cheng, and Bihan Wen. Shadowformer: Global context helps shadow removal. In *AAAI*, 2023.

[15] Lanqing Guo, Chong Wang, Wenhan Yang, Siyu Huang, Yufei Wang, Hanspeter Pfister, and Bihan Wen. Shadowdiffusion: When degradation prior meets diffusion model for shadow removal. In *CVPR*, 2023.

[16] Lanqing Guo, Chong Wang, Wenhan Yang, Yufei Wang, and Bihan Wen. Boundary-aware divide and conquer: A diffusion-based solution for unsupervised shadow removal. In *ICCV*, 2023.

[17] Ruiqi Guo, Qieyun Dai, and Derek Hoiem. Paired regions for shadow detection and removal. *IEEE TPAMI*, 35(12):2956–2967, 2012.

[18] Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. *arXiv preprint arXiv:2207.12598*, 2022.

[19] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *NeurIPS*, 2020.

[20] Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, Weizhu Chen, et al. Lora: Low-rank adaptation of large language models. *ICLR*, 2022.

[21] Xiaowei Hu, Yitong Jiang, Chi-Wing Fu, and Pheng-Ann Heng. Mask-shadowgan: Learning to remove shadows from unpaired data. In *ICCV*, 2019.

[22] Eun Woo Im, Junseong Shin, Sungyong Baik, and Tae Hyun Kim. Deep variational Bayesian modeling of haze degradation process. In *CIKM*, 2023.

[23] Naoto Inoue and Toshihiko Yamasaki. Learning from synthetic shadows for shadow detection and removal. *IEEE TCSVT*, 31(11):4187–4197, 2020.

[24] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *CVPR*, 2017.

[25] Yeying Jin, Aashish Sharma, and Robby T Tan. Dc-shadownet: Single-image hard and soft shadow removal using unsupervised domain-classifier guided network. In *ICCV*, 2021.

[26] Yeying Jin, Wei Ye, Wenhan Yang, Yuan Yuan, and Robby T Tan. Des3: Adaptive attention-driven self and soft shadow removal using vit similarity. In *AAAI*, 2024.

[27] Cláudio Rosito Jung. Efficient background subtraction and shadow removal for monochromatic video sequences. *IEEE TMM*, 11(3):571–577, 2009.

[28] Yuval Kirstain, Adam Polyak, Uriel Singer, Shahbuland Matiana, Joe Penna, and Omer Levy. Pick-a-pic: An open dataset of user preferences for text-to-image generation. *NeurIPS*, 2023.

[29] Nikolina Kubiak, Armin Mustafa, Graeme Phillipson, Stephen Jolly, and Simon Hadfield. S3r-net: A single-stage approach to self-supervised shadow removal. In *CVPR*, 2024.

[30] Eric C Larson and Damon M Chandler. Most apparent distortion: full-reference image quality assessment and the role of strategy. *Journal of electronic imaging*, 2010.

[31] Hieu Le and Dimitris Samaras. Shadow removal via shadow image decomposition. In *ICCV*, 2019.

[32] Hieu Le and Dimitris Samaras. From shadow segmentation to shadow removal. In *ECCV*, 2020.

[33] Chia-Ming Lee, Yu-Fan Lin, Yu-Jou Hsiao, Jin-Hui Jiang, Yu-Lun Liu, and Chih-Chung Hsu. PhaSR: Generalized image shadow removal with physically aligned priors. In *CVPR*, 2026.

[34] Jiachen Li, Weixi Feng, Tsu-Jui Fu, Xinyi Wang, Sugato Basu, Wenhu Chen, and William Yang Wang. T2v-turbo: Breaking the quality bottleneck of video consistency model with mixed reward feedback. *NeurIPS*, 2024.

[35] Junnan Li, Dongxu Li, Caiming Xiong, and Steven Hoi. Blip: Bootstrapping language-image pre-training for unified vision-language understanding and generation. In *ICML*, 2022.

[36] Youwei Liang, Junfeng He, Gang Li, Peizhao Li, Arseniy Klimovskiy, Nicholas Carolan, Jiao Sun, Jordi Pont-Tuset, Sarah Young, Feng Yang, et al. Rich human feedback for text-to-image generation. In *CVPR*, 2024.

[37] Yun-Hsuan Lin, Wen-Chin Chen, and Yung-Yu Chuang. Bedsr-net: A deep shadow removal network from a single document image. In *CVPR*, 2020.

[38] Kendong Liu, Zhiyu Zhu, Chuanhao Li, Hui Liu, Huanqiang Zeng, and Junhui Hou. Prefpaint: Aligning image inpainting diffusion model with human preference. *NeurIPS*, 2024.

[39] Xiaoyang Liu, Yuquan Wang, Zheng Chen, Jiezhang Cao, He Zhang, Yulun Zhang, and Xiaokang Yang. One-step diffusion model for image motion-deblurring. *arXiv preprint arXiv:2503.06537*, 2025.

[40] Zhihao Liu, Hui Yin, Yang Mi, Mengyang Pu, and Song Wang. Shadow removal by a lightness-guided network with training on unpaired data. *IEEE TIP*, 30:1853–1865, 2021.

[41] Zhihao Liu, Hui Yin, Xinyi Wu, Zhenyao Wu, Yang Mi, and Song Wang. From shadow generation to shadow removal. In *CVPR*, 2021.

[42] Ilya Loshchilov and Frank Hutter. Sgdr: Stochastic gradient descent with warm restarts. *arXiv preprint arXiv:1608.03983*, 2016.

[43] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017.

[44] Gonzalo Martin Garcia, Karim Abou Zeid, Christian Schmidt, Daan de Geus, Alexander Hermans, and Bastian Leibe. Fine-tuning image-conditional diffusion models is easier than you think. In *WACV*, 2025.

[45] Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel Haziza, Francisco Massa, Alaaeldin El-Nouby, et al. Dinov2: Learning robust visual features without supervision. *arXiv preprint arXiv:2304.07193*, 2023.

[46] Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. Training language models to follow instructions with human feedback. *NeurIPS*, 2022.

[47] Nikolay Ponomarenko, Lina Jin, Oleg Ieremeiev, Vladimir Lukin, Karen Egiazarian, Jaakko Astola, Benoit Vozel, Kacem Chehdi, Marco Carli, Federica Battisti, et al. Image database tid2013: Peculiarities, results and perspectives. *Signal processing: Image communication*, 30: 57–77, 2015.

[48] Vaishnav Potlapalli, Syed Waqas Zamir, Salman Khan, and Fahad Shahbaz Khan. PromptIR: Prompting for all-in-one image restoration. *NeurIPS*, 2023.

[49] Junbo Qiao, Miaomiao Cai, Wei Li, Yutong Liu, Xudong Huang, Gaoqi He, Jiao Xie, Jie Hu, Xinghao Chen, and Shaohui Lin. Realsr-r1: Reinforcement learning for real-world image super-resolution with vision-language chain-of-thought. *arXiv preprint arXiv:2506.16796*, 2025.

[50] Liangqiong Qu, Jiandong Tian, Shengfeng He, Yandong Tang, and Rynson WH Lau. Deshadownet: A multi-context embedding deep network for shadow removal. In *CVPR*, 2017.

[51] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In *ICML*, 2021.

[52] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *CVPR*, 2022.

[53] Alzayat Saleh, Alex Olsen, Jake Wood, Bronson Philippa, and Mostafa Rahimi Azghadi. Fieldnet: Efficient real-time shadow removal for enhanced vision in field robotics. *ESWA*, 279: 127442, 2025.

[54] Christoph Schuhmann, Romain Beaumont, Richard Vencu, Cade Gordon, Ross Wightman, Mehdi Cherti, Theo Coombes, Aarush Katta, Clayton Mullis, Mitchell Wortsman, et al. Laion-5b: An open large-scale dataset for training next generation image-text models. *NeurIPS*, 2022.

[55] Junseong Shin and Tae Hyun Kim. PRISM: Latent composition consistency for single-image reflection removal. In *ECCV*, 2026.

[56] Junseong Shin, Seungwoo Chung, Yunjeong Yang, and Tae Hyun Kim. HazeFlow: Revisit haze physical model as ODE and non-homogeneous haze generation for real-world dehazing. In *ICCV*, 2025.

[57] Junseong Shin, Hyeonsu Jo, Daehyun Kim, and Tae Hyun Kim. PickMoment: Continuous-time single-image-to-video via learning deblurring and blur-to-video. *NeurIPS*, 2026.

[58] Lingchen Sun, Rongyuan Wu, Zhiyuan Ma, Shuaizheng Liu, Qiaosi Yi, and Lei Zhang. Pixel-level and semantic-level adjustable super-resolution: A dual-lora approach. In *CVPR*, 2025.

[59] Florin-Alexandru Vasluianu, Tim Seizinger, Zongwei Wu, Rakesh Ranjan, and Radu Timofte. Towards image ambient lighting normalization. In *ECCV*, 2024.

[60] Florin-Alexandru Vasluianu, Tim Seizinger, Zhuyun Zhou, Zongwei Wu, Cailian Chen, Radu Timofte, Wei Dong, Han Zhou, Yuqiong Tian, Jun Chen, et al. Ntire 2024 image shadow removal challenge report. In *CVPR*, 2024.

[61] Jin Wan, Hui Yin, Zhenyao Wu, Xinyi Wu, Yanting Liu, and Song Wang. Style-guided shadow removal. In *ECCV*, 2022.

[62] Jifeng Wang, Xiang Li, and Jian Yang. Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal. In *CVPR*, 2018.

[63] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE TIP*, 13(4):600–612, 2004.

[64] Bin Wu, Wei Wang, Yahui Liu, Zixiang Li, and Yao Zhao. Diffusionreward: Enhancing blind face restoration through reward feedback learning. *arXiv preprint arXiv:2505.17910*, 2025.

[65] Rongyuan Wu, Lingchen Sun, Zhiyuan Ma, and Lei Zhang. One-step effective diffusion network for real-world image super-resolution. *NeurIPS*, 2024.

[66] Rongyuan Wu, Lingchen Sun, Zhengqiang Zhang, Shihao Wang, Tianhe Wu, Qiaosi Yi, Shuai Li, and Lei Zhang. Dp^{2}o-sr: Direct perceptual preference optimization for real-world image super-resolution. *arXiv preprint arXiv:2510.18851*, 2025.

[67] Xiaoshi Wu, Yiming Hao, Keqiang Sun, Yixiong Chen, Feng Zhu, Rui Zhao, and Hongsheng Li. Human preference score v2: A solid benchmark for evaluating human preferences of text-to-image synthesis. *arXiv preprint arXiv:2306.09341*, 2023.

[68] Zeqiu Wu, Yushi Hu, Weijia Shi, Nouha Dziri, Alane Suhr, Prithviraj Ammanabrolu, Noah A Smith, Mari Ostendorf, and Hannaneh Hajishirzi. Fine-grained human feedback gives better rewards for language model training. *NeurIPS*, 2023.

[69] Fen Xia, Tie-Yan Liu, Jue Wang, Wensheng Zhang, and Hang Li. Listwise approach to learning to rank: theory and algorithm. In *ICML*, 2008.

[70] Jie Xiao, Xueyang Fu, Yurui Zhu, Dong Li, Jie Huang, Kai Zhu, and Zheng-Jun Zha. Homoformer: Homogenized transformer for image shadow removal. In *CVPR*, 2024.

[71] Jiamin Xu, Zelong Li, Yuxin Zheng, Chenyu Huang, Renshu Gu, Weiwei Xu, and Gang Xu. Omnisr: Shadow removal under direct and indirect lighting. In *AAAI*, 2025.

[72] Jiamin Xu, Yuxin Zheng, Zelong Li, Chi Wang, Renshu Gu, Weiwei Xu, and Gang Xu. Detail-preserving latent diffusion for stable shadow removal. In *CVPR*, pages 7592–7602, 2025.

[73] Jiazheng Xu, Xiao Liu, Yuchen Wu, Yuxuan Tong, Qinkai Li, Ming Ding, Jie Tang, and Yuxiao Dong. Imagereward: Learning and evaluating human preferences for text-to-image generation. *NeurIPS*, 2023.

[74] JunLiang Ye, Fangfu Liu, Qixiu Li, Zhengyi Wang, Yikai Wang, Xinzhou Wang, Yueqi Duan, and Jun Zhu. Dreamreward: Text-to-3d generation with human preference. In *ECCV*, 2024.

[75] Tianwei Yin, Michaël Gharbi, Richard Zhang, Eli Shechtman, Fredo Durand, William T Freeman, and Taesung Park. One-step diffusion with distribution matching distillation. In *CVPR*, 2024.

[76] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *CVPR*, 2018.

[77] Wuming Zhang, Xi Zhao, Jean-Marie Morvan, and Liming Chen. Improving shadow suppression for illumination robust face recognition. *IEEE TPAMI*, 41(3):611–624, 2018.

[78] Yurui Zhu, Jie Huang, Xueyang Fu, Feng Zhao, Qibin Sun, and Zheng-Jun Zha. Bijective mapping network for shadow removal. In *CVPR*, 2022.

[79] Yurui Zhu, Zeyu Xiao, Yanchi Fang, Xueyang Fu, Zhiwei Xiong, and Zheng-Jun Zha. Efficient model-driven network for shadow removal. In *AAAI*, 2022.

[80] Daniel M Ziegler, Nisan Stiennon, Jeffrey Wu, Tom B Brown, Alec Radford, Dario Amodei, Paul Christiano, and Geoffrey Irving. Fine-tuning language models from human preferences. *arXiv preprint arXiv:1909.08593*, 2019.

Appendix

This appendix provides additional details and analyses to complement the main paper.

- **Sec. A**: Construction of the Human-Annotation-Free Dataset, including degradation parameter settings and the effect of the number of degraded samples.
- **Sec. B**: Detailed hyperparameter configurations used for training ODD, ODDR, and ShadowReward.
- **Sec. C**: Full protocol, statistical reliability, and additional results of the human preference alignment experiment.
- **Sec. D**: Ablation studies on key design choices of ShadowReward, including the feature extractor backbone and spatial reduction strategy.
- **Sec. E**: Analysis of reward-guided fine-tuning, including reward attribution and loss-weight sensitivity, reward hacking, robustness to shadow-free inputs, alternative guidance strategies, training data source effects, and non-shadow region color consistency.
- **Sec. F**: Architecture and ablation of the ODD-Conv design for preserving fine-grained details.
- **Sec. G**: Transferability of ShadowReward to ambient lighting normalization and other low-level vision tasks.
- **Sec. H**: Limitations of the proposed framework and representative failure cases.
- **Sec. I**: Licenses of existing assets used in this work.
- **Sec. J**: Additional qualitative results and visual comparisons across multiple datasets.

A Details of Human-Annotation-Free Dataset

We construct the Human-Annotation-Free Dataset by synthesizing a set of perceptually ordered degraded images for each clean image y. Using a shadow matte m obtained from SynShadow [23], we apply degradations to the shadow region of y to simulate unsuccessful shadow removal artifacts. Specifically, we construct a list of progressively degraded images $\{x^{(j)}\}_{j=0}^{3}$, where $x^{(0)} = y$ represents the clean reference, and each $x^{(j)}$ for $j \geq 1$ applies increasingly severe distortions to the shadow region of y.

Table 7: Parameter ranges used to synthesize cumulative degradations (T, N, C, L, B) for the Human-Annotation-Free Dataset. Higher j corresponds to stronger degradation ($j = 0$ is the clean image). Here, r denotes the boundary dilation radius, Δh the hue shift, s the saturation scaling factor, b the brightness scaling factor, and γ the luminance gamma parameter. All parameters are uniformly sampled within the specified ranges. Larger values (or wider ranges) imply more severe distortions.

Type	**Operation**	$j = 1$	$j = 2$	$j = 3$
T	Gaussian blur	3×3	5×5	7×7
N	Gaussian noise	σ=0.004	σ=0.008	σ=0.012
C	Hue shift	$\Delta h = \pm 0.01$	$\Delta h = \pm 0.02$	$\Delta h = \pm 0.03$
	Saturation scale	$s \in [0.95, 1.05]$	$s \in [0.90, 1.10]$	$s \in [0.85, 1.15]$
	Brightness scale	$b \in [0.97, 1.03]$	$b \in [0.94, 1.06]$	$b \in [0.92, 1.08]$
L	Gamma correction	$\gamma \in [0.97, 1.03]$	$\gamma \in [0.94, 1.06]$	$\gamma \in [0.90, 1.10]$
B	Dilation + Gaussian blur	r=3, 3×3	r=7, 5×5	r=11, 7×7

A.1 Degradation Parameter Settings

To construct a list of degraded images, we define five degradation types that mimic the visual artifacts caused by unsuccessful deshadowing: texture distortion (T) applies Gaussian blur to simulate loss of detail, additive noise (N) injects Gaussian noise, chromatic shift (C) perturbs hue, saturation, and brightness, luminance change (L) applies gamma correction, and boundary artifacts (B) dilate the

Table 8: Impact of the degradation list length J on shadow removal performance.

List length (J)	Shadow		All	
	PSNR ↑	RMSE ↓	PSNR ↑	RMSE ↓
1	37.29	5.99	33.70	3.06
2	37.78	5.74	33.96	3.01
3	**38.37**	**5.36**	**34.20**	**2.97**
4	38.17	5.45	34.15	2.97

shadow boundary followed by Gaussian blur. All parameters are uniformly sampled within the ranges specified in Tab. 7, with level-dependent settings that produce progressively stronger degradations for $j = 1, 2, 3$.

In our experiments, we synthesize each training sample by applying multiple degradation types rather than relying on a single one. Depending on the degradation level j (1, 2, or 3), we randomly select 2, 3, or 4 degradation types from (T, N, C, L, B), respectively. This systematic increase in the number of applied types naturally yields more severe and complex degradations for the higher degradation levels. This design allows the model to encounter continuously varying and randomly mixed corruption patterns, enabling it to learn more stable and reliable perceptual rankings.

A.2 Length of the Degradation List

We also analyze how the length of the degradation list (J) used to train ShadowReward affects the final shadow removal performance of ODDR. As shown in Tab. 8, the shadow removal performance of ODDR improves as the list length grows from $J = 1$ to $J = 3$, achieving the best result (38.37 dB / 5.36 RMSE) at $J = 3$. However, using more than three degraded images leads to a slight drop in performance. This is primarily because extra distortions push the samples beyond the model's predictive bounds, rendering them irrelevant, while learning too many fine-grained rankings weakens the overall ranking signal.

B Additional Experimental Details

B.1 Synthetic Shadow Generation

For ODD training, we generate synthetic shadow images following SynShadow [23]. Given a clean image y and a randomly sampled shadow matte m, the synthetic shadow image is computed as:

$$x_{\text{syn}} = (1 - m) \odot y + m \odot (a \cdot y + b), \tag{6}$$

where $a < 1$ is a randomly sampled slope controlling shadow darkness, and b is a per-channel intercept. This affine shadow model produces diverse shadow intensities and color shifts by randomizing a and b for each training pair. The resulting (x_{syn}, y) pairs are used to train ODD in a supervised manner. Additionally, the same shadow matte m is reused during the construction of the Human-Annotation-Free Dataset (Sec. A) to apply degradations specifically to the shadow region when training ShadowReward.

B.2 Loss Hyperparameters

For ODD training (Eq. (1)), we use the LPIPS perceptual loss [76] with a weighting factor of $\lambda_{\text{per}} = 2$. When training ODDR (Eq. (2)), we additionally apply the alignment loss with $\lambda_{\text{align}} = 2$, and set the reward-guidance weight to $\lambda_r = 1 \times 10^{-1}$. For ShadowReward, we train the ListMLE-based ranking model [69] (Eq. (5)) using a temperature parameter of $\tau = 1$. All hyperparameters are empirically chosen and kept fixed across all experiments.

B.3 LoRA Configuration

We use LoRA [20] with rank 16 and Gaussian initialization to adapt the frozen Stable Diffusion UNet. LoRA modules are inserted into both the attention layers (`to_k`, `to_q`, `to_v`, `to_out`) and the convolutional layers (`conv1`, `conv2`, `conv_in`, `conv_shortcut`, `conv_out`), as well as the

feed-forward layers (`ff.net`) and projection layers (`proj_in`, `proj_out`). For ODD training, both ODD-LoRA and ODD-Conv are optimized while the original SD weights remain frozen. For ODDR fine-tuning, a separate ODDR-LoRA module is initialized and optimized while ODD-LoRA and ODD-Conv are frozen.

B.4 ShadowReward Architecture Details

We use DINOv2 ViT-B/14 [45] (patch size 14, embedding dimension 768) as the feature extractor, which is kept frozen during training. The convolutional head ψ_θ takes as input the concatenation of the shadow tokens, the corrupted image tokens, their element-wise difference, and its absolute value, resulting in an input dimension of $768 \times 4 = 3072$ channels. The head consists of: a 1×1 projection layer ($3072 \to 256$), a 3×3 depthwise separable convolution ($256 \to 128$, groups=128), a 1×1 pointwise convolution ($128 \to 64$), Dropout2d, and a final 1×1 convolution ($64 \to 1$) followed by a sigmoid activation. GELU activation is applied after each intermediate layer.

B.5 Data Augmentation

During training, we apply the following augmentations to the input images: (1) reflection padding to ensure a minimum resolution of 512×512; (2) random cropping to 512×512; (3) random horizontal flipping with probability 0.5; and (4) random rotation sampled from $\{90°, 180°, 270°\}$. These augmentations are applied consistently across all training stages (ODD, ODDR, and ShadowReward).

B.6 Inference Details

At inference time, the input image is first padded using reflection padding so that both spatial dimensions are multiples of 8, which is required by the VAE encoder. The model then performs a single-step denoising pass at timestep t=999 without classifier-free guidance (CFG) [18]. The output is cropped back to the original input resolution after decoding.

B.7 Training Cost

All experiments are conducted on a single NVIDIA RTX A6000 GPU. ODD training takes approximately 24 hours, while ODDR reward-guided fine-tuning converges in approximately 12 hours.

C Observations on Human Preference Alignment

C.1 User Study Details

To measure how well ShadowReward aligns with human preference, we conducted a user study collecting human rankings for shadow removal results. The survey interface, shown in Fig. 7, was designed to facilitate direct visual comparison.

- **Survey Structure / Interface**: As illustrated in Fig. 7, each question presented a 2×2 grid layout containing one input shadow image and three differently deshadowed results (labeled A, B, and C), where the three results were randomly selected from different shadow removal methods.
- **Ranking Task**: Participants were instructed to "rank images A, B, and C from the most effective shadow removal to the least effective." We utilized a matrix-style selection interface, ensuring that annotators explicitly assigned a unique rank (1st, 2nd, or 3rd) to each candidate without duplication.
- **Data Source**: The 16 input images were randomly selected from the AISTD test dataset [31]. The three comparison images for each question were sampled from a pool of six different models: four existing deshadowing models [40, 72, 41, 29] and our proposed ODD and ODDR models. The positions of the candidate images (A, B, C) were randomized for each question to prevent positional bias.
- **Participants**: A total of 23 annotators participated in the survey. The participant pool consisted of a diverse group, ranging from low-level vision experts (graduate students and

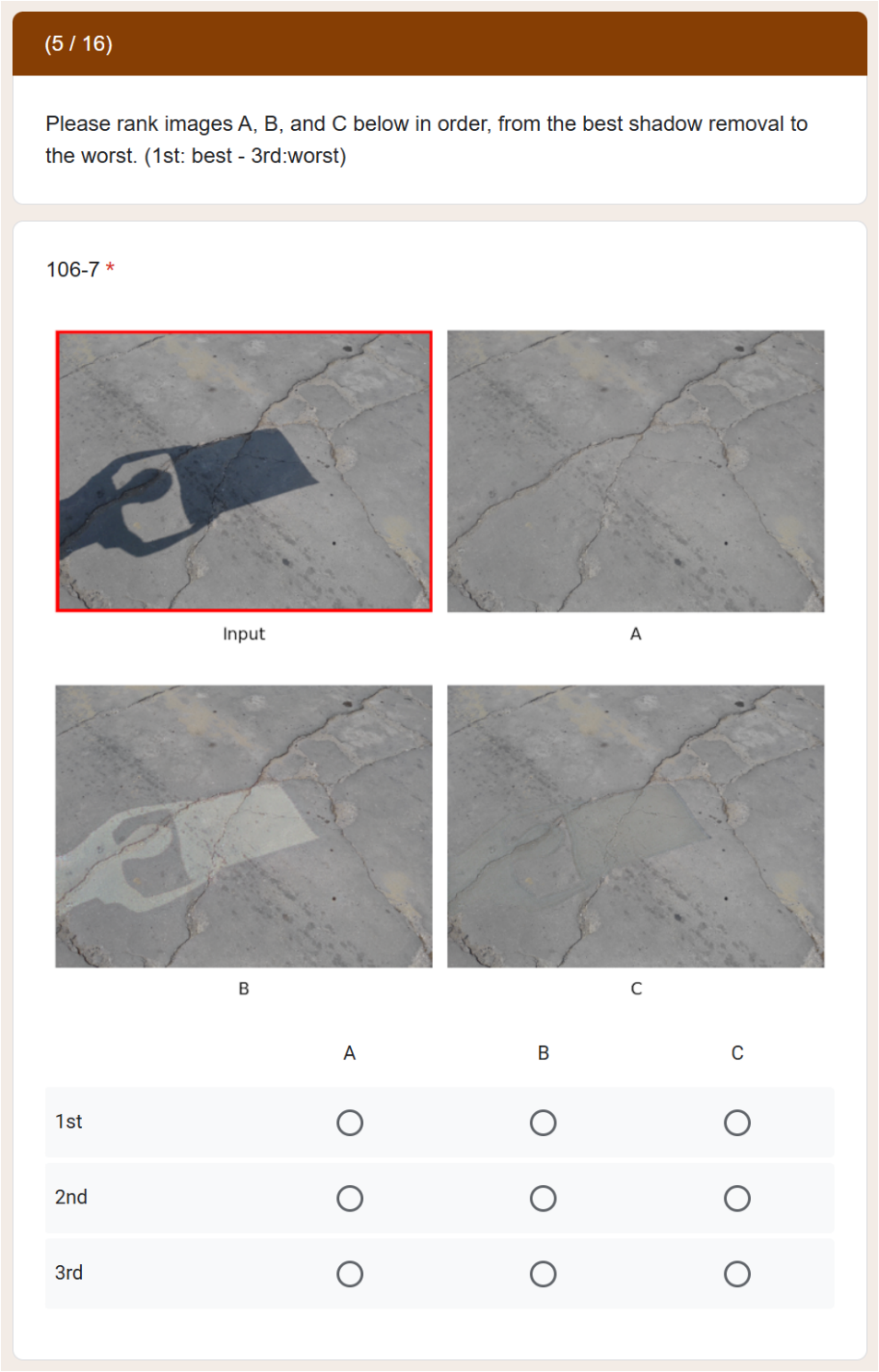

Figure 7: User study interface for human preference acquisition. Annotators rank three randomized results (A, B, C) against the input image (red border).



Reward Models	Image A	Image B	Image C	Ranking	
ImageReward	-0.8401 (#1)	-1.2754 (#3)	-1.2657 (#2)	A > C > B	✗
PickScore	19.9215 (#3)	19.9596 (#1)	19.9318 (#2)	B > C > A	✗
ShadowReward	**0.01313 (#1)**	**0.001294 (#2)**	**0.000998 (#3)**	**A > B > C**	✓

Figure 8: **Comparison of reward model ranking accuracy.** ShadowReward correctly ranks the shadow removal results in alignment with human preference ($A > B > C$), while existing reward models fail to match human perception.

Ph.D. researchers) to general users without technical background. This diversity ensures that the evaluation reflects both technical precision and general perceptual quality.

- **Ground Truth Establishment**: The ground-truth ranking for each question was determined by aggregating the responses via majority voting. This consensus ranking was subsequently used to calculate the reward model accuracy reported in Tab. 3.
- **Evaluation Setting**: Since every input is a real AISTD shadow image and every candidate is the output of a real shadow removal model, the reported accuracy measures ShadowReward on (real shadow, real model output) pairs, i.e., the setting in which it is used during ODDR fine-tuning, rather than on the synthetic degradations it is trained on.
- **Statistical Reliability**: To verify that the human consensus is a reliable reference, we measure the agreement of individual responses with the majority-vote ranking: annotators agree on 79.6% of full three-way rankings and on 86.7% of top choices, well above chance (1/6 for a full ranking). We note that the study contains only 16 questions; ShadowReward's accuracy (68.75%, 11/16) is therefore not statistically separable from PSNR-based ranking (56.25%, 9/16), and we interpret it as being on par with PSNR while requiring no ground truth. In contrast, the margin over existing reward models and NR-IQA metrics (all below 30%) is substantial.

C.2 Ranking Accuracy Comparison

To demonstrate that ShadowReward better captures human preference for shadow removal quality, we compare it against existing reward models commonly used in image generation: ImageReward [73] and PickScore [28]. As shown in Fig. 8, general-purpose reward models do not align with human rankings for shadow removal results. They often favor images with stronger global semantics, disregarding noticeable residual shadows or color distortions. In contrast, ShadowReward consistently aligns with human-preferred ordering, accurately prioritizing restorations that feature clean shadow removal and coherent illumination. This demonstrates that ShadowReward is specifically well-aligned with the requirements of shadow removal, providing task-aware perceptual judgments that general-purpose VLM-based reward models cannot achieve.

C.3 Perceptual Evaluation Beyond PSNR

Traditional metrics such as PSNR depend on strict pixel-level correspondence with paired ground truth and therefore become unreliable when misalignment occurs between two images. As illustrated in Fig. 9, mismatches stemming from transient objects (e.g., the shoe in the top row) or environmental movements (e.g., shifting grass in the bottom row) lead to unfairly low PSNR scores for high-quality restorations, even when the shadows are correctly removed. In contrast, ShadowReward provides a more perceptually reliable assessment by focusing on visual coherence rather than pixel-wise

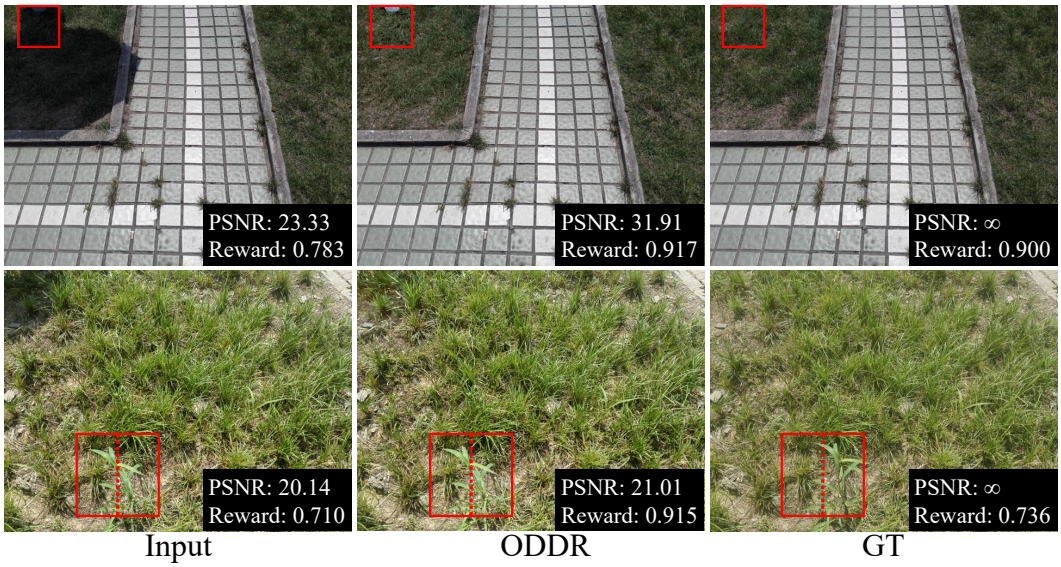

Figure 9: PSNR is unreliable under GT misalignment, while our ShadowReward (labeled 'Reward') provides perceptually consistent evaluation.

Table 9: Impact of different feature extractor backbones for ShadowReward.

Backbone	Shadow		All	
	PSNR ↑	RMSE ↓	PSNR ↑	RMSE ↓
CLIP	37.27	5.91	33.67	3.04
MoCov3	37.72	5.71	33.91	3.01
DINOv2	**38.37**	**5.36**	**34.20**	**2.97**

similarity. It assigns higher scores to restorations with natural tone and structure, and penalizes outputs with residual shadows or color inconsistency, regardless of ground-truth bias. These examples highlight that PSNR can fail to reflect human perceptual preference under misaligned conditions, whereas ShadowReward, which requires no ground truth, remains unaffected by such misalignment and thus serves as a useful complementary perceptual evaluator for shadow removal.

D Ablation Study of ShadowReward

We conduct ablation studies on key design choices of ShadowReward, including the feature extractor backbone and the spatial reduction strategy.

D.1 Feature Extractor

We investigate the impact of different feature extractors used in ShadowReward. As shown in Tab. 9, the reward model with DINOv2 [45] performs better than those with CLIP [51] and MoCov3 [8], achieving a shadow-region PSNR of 38.37 and an RMSE of 5.36, compared to 37.27/5.91 for CLIP and 37.72/5.71 for MoCov3. This indicates that DINOv2's superior capability in capturing local structural dependencies and illumination variations makes it more effective for evaluating shadow boundaries and color consistency.

D.2 Spatial Reduction Method

We further evaluate different spatial reduction strategies used for aggregating patch-level scores into the final perceptual reward. Tab. 10 shows that simple mean or max pooling methods underperform compared to Top-K pooling. Among the variants, Top-K pooling with K=0.1 (top 10% of spatial responses) yields the best balance between localization and robustness, achieving a Shadow PSNR of

Table 10: Impact of different spatial reduction methods for ShadowReward.

Method	Shadow		All	
	PSNR ↑	RMSE ↓	PSNR ↑	RMSE ↓
Mean	37.66	5.77	33.90	3.02
Max	37.62	5.56	33.89	3.02
Top-K(0.05)	38.28	5.48	34.19	**2.97**
Top-K(0.15)	38.08	5.51	34.12	2.98
Top-K(0.1)	**38.37**	**5.36**	**34.20**	**2.97**

Table 11: Reward attribution and reward-weight sensitivity on AISTD. λ_r=0 corresponds to anchor-only fine-tuning. λ_{align}=2 in all rows.

λ_r	Shadow			Non-Shadow			ALL		
	PSNR↑	SSIM↑	RMSE↓	PSNR↑	SSIM↑	RMSE↓	PSNR↑	SSIM↑	RMSE↓
ODD (base)	36.67	0.990	6.27	37.45	0.984	2.49	33.32	0.971	3.11
0 (anchor only)	36.63	0.990	6.29	37.45	0.984	2.50	33.30	0.971	3.11
0.01	36.91	0.991	6.13	37.45	0.984	2.48	33.47	0.971	3.08
0.1 (ours)	38.37	0.991	5.36	37.34	0.984	2.50	34.20	0.972	2.97
0.5	38.34	0.991	5.41	37.30	0.984	2.51	34.16	0.972	2.99
1.0	38.32	0.991	5.44	37.27	0.984	2.53	34.12	0.971	3.00
10.0	38.29	0.991	5.52	37.26	0.984	2.54	34.01	0.971	3.02
100.0	38.48	0.991	5.36	36.77	0.983	2.63	33.87	0.970	3.07

38.37 dB while maintaining an overall RMSE of 2.97. This demonstrates that selectively averaging the most informative spatial responses allows the reward model to focus on shadow boundaries while suppressing noise from irrelevant regions.

D.3 ListMLE Temperature

We tested $\tau \in \{0.5, 1.0, 1.5\}$ and observed pairwise preference accuracies of 0.939, 0.969, and 0.963 respectively on the AISTD test set (clean GT vs. ODD output), indicating that ShadowReward is robust to moderate variations in the ListMLE temperature; we therefore use τ=1.0 following standard practice [69].

E Analysis of Reward-Guided Fine-Tuning

In this section, we provide further analyses to validate the design choices and effectiveness of our reward-guided fine-tuning strategy.

E.1 Reward Attribution and Reward-Weight Sensitivity

A natural question is whether the ODD→ODDR improvement originates from the reward or from self-distillation toward the frozen ODD pseudo-target. Both alignment terms in Eq. (2), the ℓ_2 distance and the perceptual distance to $\hat{y}_{\text{ODD}}$, attain their minimum of zero exactly at the ODD output, i.e., at initialization. Consequently, with λ_r=0 the objective is already minimized and fine-tuning leaves the model essentially unchanged, while for λ_r>0 the reward is the only term that moves the model beyond ODD. Tab. 11 confirms this empirically: the anchor-only setting (λ_r=0) reproduces ODD (33.30 vs. 33.32 dB ALL PSNR), so the entire gain of ODDR (+1.70 dB shadow-region PSNR) is attributable to the reward. Note that the small value of λ_r relative to λ_{align} does not imply a small contribution, since the two weights multiply loss terms that live on different scales.

Performance is stable over a wide range of reward weights: ALL PSNR remains between 34.01 and 34.20 dB for $\lambda_r \in [0.1, 10]$, with λ_r=0.1 being the best setting.

Reward hacking. To locate where over-optimization begins, we further increase the weight to λ_r=100. Here, the shadow-region PSNR reaches its highest value (38.48 dB), but the non-shadow and ALL PSNR drop markedly (37.34→36.77 dB and 34.20→33.87 dB): the reward keeps improving the region it targets while overall reconstruction quality degrades, the typical signature of reward

Table 12: Alignment-weight sensitivity on AISTD (λ_r=0.1 in all rows).

λ_{align}	Shadow			Non-Shadow			ALL		
	PSNR↑	SSIM↑	RMSE↓	PSNR↑	SSIM↑	RMSE↓	PSNR↑	SSIM↑	RMSE↓
0	38.17	0.991	5.54	37.32	0.984	2.52	34.07	0.971	3.01
1.0	38.27	0.991	5.48	37.27	0.984	2.53	34.09	0.971	3.01
2.0 (ours)	38.37	0.991	5.36	37.34	0.984	2.50	34.20	0.972	2.97
5.0	37.36	0.991	5.87	37.45	0.984	2.48	33.74	0.972	3.04

Table 13: Robustness to shadow-free inputs. The 540 shadow-free images of the AISTD test set are fed to each model, and the output is compared with the clean input. Higher PSNR and lower RMSE indicate better preservation.

Shadow-free input	PSNR↑	RMSE↓
SSR [72]	40.08	2.08
ODDR (ours)	**45.10**	**0.83**

hacking. Our default λ_r=0.1 lies well within the stable regime, where the alignment terms protect non-shadow fidelity.

E.2 Alignment-Weight Sensitivity

Tab. 12 varies the alignment weight λ_{align}. ALL PSNR remains stable for $\lambda_{\text{align}} \in \{0, 1, 2\}$ (34.07 / 34.09 / 34.20 dB). At λ_{align}=5, the anchor dominates the reward and the shadow-region PSNR drops (38.37→37.36 dB), the expected trade-off of over-weighting the anchor. We therefore use λ_{align}=2.

E.3 Robustness to Shadow-Free and Soft Shadows

A shadow removal model should not alter content when there is little or nothing to remove. To test this, we pass the 540 shadow-free images of the AISTD test set through the model and compare each output with its clean input. As shown in Tab. 13, ODDR leaves clean inputs nearly unchanged (45.10 dB PSNR), well above the strongest baseline SSR (40.08 dB), indicating that it does not erase genuine detail in the absence of shadows. This is consistent with the non-shadow preservation observed in Tab. 1 and Fig. 5. For faint shadows, LRSS [12] consists of soft shadows, and ODDR achieves the best results on it (Tab. 2), showing that it handles soft shadow boundaries without over-editing.

E.4 Comparison with Alternative Guidance Strategies

To justify the use of ShadowReward over other possible guidance signals, we compare it against three alternative methods under the same unpaired fine-tuning protocol used for ODDR: (1) a SAM-based shadow detection model [7], which penalizes the shadow detection confidence in the restored output to encourage shadow-free results; (2) a conditional GAN discriminator [24], trained to distinguish unpaired shadow-free images from the restored images; and (3) a DINO-based discriminator adapted for one-step diffusion in low-level vision tasks.

As shown in Tab. 14, ShadowReward consistently outperforms all alternatives. The shadow detection approach only marginally improves over ODD, as it provides a binary shadow presence signal rather than fine-grained perceptual feedback. The conditional GAN discriminator suffers from unstable adversarial training dynamics in the unpaired setting and occasionally introduces new artifacts. The DINO-based discriminator performs reasonably well due to its strong structural and perceptual cue-capturing capabilities, yet it still falls short of ShadowReward by 0.58 dB in shadow-region PSNR. These results demonstrate that learning a preference ranking over restored results, as done by ShadowReward, provides a more effective supervision signal than discriminative or detection-based alternatives.

Table 14: Comparison with alternative guidance strategies on the AISTD dataset. All methods use the same ODD baseline and unpaired fine-tuning protocol.

Guidance Method	Shadow		ALL	
	PSNR↑	RMSE↓	PSNR↑	RMSE↓
Shadow Detection	36.69	6.26	33.33	3.11
Conditional GAN	36.79	6.40	33.36	3.12
DINO Discriminator	37.79	5.74	33.91	3.03
ShadowReward (ours)	**38.37**	**5.36**	**34.20**	**2.97**

Table 15: Effect of training data source overlap on the AISTD dataset. USR denotes a disjoint shadow-free dataset used to train ODD.

Method	ODD Train	Shadow		ALL	
		PSNR↑	RMSE↓	PSNR↑	RMSE↓
ODD	AISTD	36.67	6.27	33.32	3.11
ODDR	AISTD	**38.37**	**5.36**	**34.20**	**2.97**
ODD	USR	36.08	6.70	32.86	3.24
ODDR	USR	37.58	5.80	33.71	3.13



Figure 10: Visual comparisons of our ODD and ODDR on the AISTD [31] (top), LRSS [12] (middle), and UIUC [17] (bottom) datasets.

E.5 Effect of Training Data Source

A potential concern regarding data overlap is that ODD and ODDR both rely on the AISTD dataset: ODD uses its shadow-free images to synthesize training pairs, while ODDR uses its real shadow images for reward-guided fine-tuning. To verify that the improvement of ODDR is not simply due to this overlapping scene distribution, we train ODD using shadow-free images from a completely disjoint dataset, USR [21], while keeping ODDR fine-tuning on AISTD real shadow images as in the standard pipeline.

As shown in Tab. 15, even when ODD is trained on a different data source, ODDR still yields consistent improvements by raising the ALL PSNR from 32.86 dB to 33.71 dB, confirming that the gains stem from reward-guided optimization rather than reusing the same scene distribution. This result validates the generality of our reward-guided fine-tuning approach and demonstrates that ShadowReward provides an effective learning signal regardless of the baseline training data.

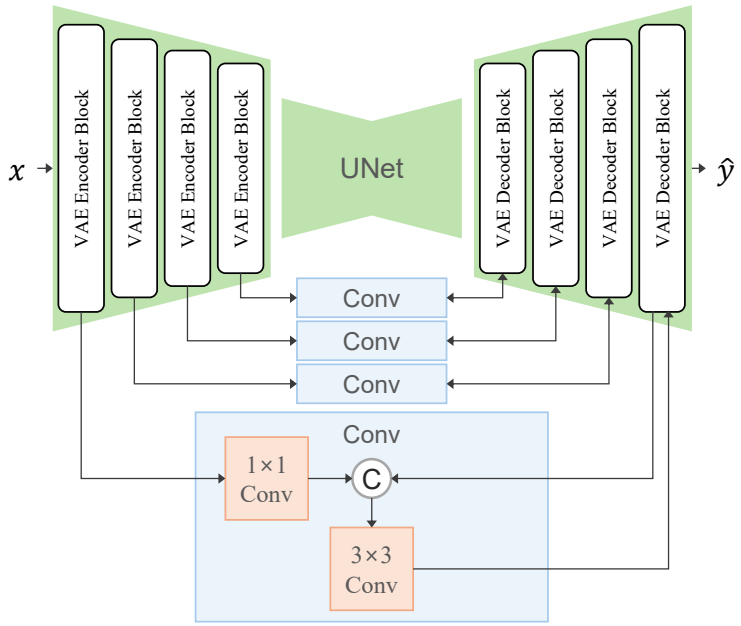

Figure 11: Detailed illustration of VAE skip-connection.

E.6 Why Does ODDR Preserve Non-Shadow Region Colors?

To visually evaluate the effectiveness of the ShadowReward model across different datasets, we provide visual comparisons of ODD and ODDR in Fig. 10. As shown in the first row, ODDR removes shadows more completely within the shadow regions, producing cleaner illumination and more natural color transitions compared to ODD. Moreover, ODDR restores boundary areas more cleanly in the second and last rows, showing fewer artifacts and better consistency between shadow and non-shadow regions. These findings show that our learned reward signal effectively guides the model to produce perceptually improved shadow-free reconstructions.

Beyond shadow-region quality, we also observe that ODD occasionally introduces undesirable color shifts in non-shadow regions, whereas ODDR effectively preserves the original color tone. This improvement can be attributed to the design of ShadowReward, which evaluates the overall perceptual quality of the restored image by comparing it against the shadow input. Since ShadowReward is trained to penalize perceptually unnatural artifacts, including chromatic distortions and luminance inconsistencies, it provides a gradient signal that discourages global color drift during ODDR fine-tuning. Combined with the alignment loss ($\mathcal{L}_{\text{per}}$ and ℓ_2) that anchors ODDR outputs to the ODD pseudo-target, the reward term selectively improves shadow regions while suppressing unintended modifications in non-shadow areas. This complementary supervision explains why ODDR achieves better color consistency across the entire image.

F ODD-Conv: Skip-Connection Design in the VAE

As described in Sec. 3.2, we introduce lightweight two-layer convolutional skip connections (ODD-Conv) at each encoder-decoder level of the VAE to mitigate the detail loss commonly observed in the original Stable Diffusion VAE. Specifically, as shown in Fig. 11, we extract intermediate encoder features and pass them through a 1×1 convolution to match the channel dimensions of the corresponding decoder features. The transformed encoder features are then concatenated with the decoder features, followed by a 3×3 convolution to fuse the information effectively. This design preserves fine-grained textures and structural details that would otherwise be lost during the VAE encoding-decoding process.

Ablation. As shown in Tab. 16, removing ODD-Conv sharply degrades quality: ALL PSNR drops from 33.32 to 26.16 dB, and the non-shadow region suffers the largest loss (-9.66 dB). Without the

Table 16: Ablation of ODD-Conv on AISTD. "w/o Conv" removes the VAE skip connections and is retrained from scratch.

Method	Shadow PSNR↑	Non-Shadow PSNR↑	ALL PSNR↑	ALL RMSE↓
ODD w/o Conv	33.14	27.79	26.16	8.46
ODD	**36.67**	**37.45**	**33.32**	**3.11**

Table 17: Transfer to ambient lighting normalization on Ambient6K [59]. ODDR refines ODD on Ambient6K images using only the unmodified ShadowReward (λ_r=0.1), without ground truth.

Method	PSNR↑	SSIM↑
DeS3 [26]	14.44	0.681
SSR [72]	15.64	0.704
ODD (zero-shot)	15.65	0.719
ODDR (reward-guided, no GT)	**17.16**	**0.747**
Δ (reward)	+1.51	+0.028

skip connections, details cannot bypass the lossy latent bottleneck, so even regions that should remain unchanged are corrupted.

G Transferability of ShadowReward

G.1 Ambient Lighting Normalization

Our main experiments focus on cast-shadow benchmarks. We further examine ambient lighting normalization, where the degradation is a global, smooth, color-dominated illumination field rather than a discrete masked shadow, using the Ambient6K benchmark [59]. Following the same recipe as for shadows, ODD serves as the base model and the ODDR stage refines it on the official training images of Ambient6K using only the unmodified ShadowReward (λ_r=0.1), with no ground truth; evaluation is performed on the official test split. As shown in Tab. 17, the reward stage improves the base model by +1.51 dB PSNR without any paired supervision, and ODDR also scores higher than the diffusion-based shadow removal methods DeS3 and SSR. We report this as an ODD→ODDR gain rather than a comparison with dedicated ambient lighting methods. The transfer is plausible because two of the five degradation types, chromatic shift and luminance change, model the same kind of color and brightness change that ambient lighting causes. Moreover, the shadow matte used by SynShadow only shapes how the training data of ShadowReward is constructed; at inference, both ODDR and ShadowReward operate without masks. The low absolute score of the AISTD-trained base, however, shows that complex, physically grounded lighting is not covered by our synthetic shadow pipeline (see Sec. H). Recent work on generalized shadow removal with physically aligned priors [33] is complementary to our annotation-free approach.

G.2 Other Low-Level Vision Tasks

The recipe behind ShadowReward is not specific to shadows: it requires only clean images and a taxonomy of degradations that reproduce the artifacts a restoration model tends to leave behind. To examine whether the learned reward itself transfers, we take PromptIR [48], an all-in-one restorer trained on synthetic data, and refine it with the unmodified ShadowReward on real-world benchmarks for three tasks: SIDD [1] (real sensor noise), GT-RAIN [4] (real rain), and NH-HAZE [3] (real non-homogeneous haze). We follow the official data splits of each benchmark, using only the degraded images for refinement and the ground truth solely for evaluation. Since the base model has never seen these real domains, the absolute scores are low; what matters is the gain from the reward. As shown in Tab. 18, the gain follows the coverage of ShadowReward's degradation types. Additive noise is one of the five types, so the reward improves denoising. Deraining, like deshadowing, requires recovering specific local content, and the reward likewise transfers. In contrast, ShadowReward never observes haze-type degradation and thus provides no signal for dehazing. Extending the degradation taxonomy with physically grounded haze modeling, as explored in recent dehazing work [22, 56], is a natural

Table 18: Transfer of ShadowReward to other low-level vision tasks. PromptIR [48], trained on synthetic data, is refined with the unmodified ShadowReward on real degraded images only (no ground truth). PSNR / SSIM.

Real-world benchmark	Base	+ShadowReward
Denoising (SIDD [1])	22.84 / 0.341	**23.35** / **0.344**
Deraining (GT-RAIN [4])	21.28 / 0.632	**22.06** / **0.636**
Dehazing (NH-HAZE [3])	12.33 / 0.444	12.32 / 0.444



Input ODDR GT

Figure 12: Example of failure case on highly textured/non-flat surfaces.

direction for future work. More broadly, since reward-guided refinement only requires a one-step restorer and a task-aware reward, it could similarly be applied to one-step models for other restoration tasks, such as dehazing [56], reflection removal [55], and deblurring [57].

H Limitations and Failure Cases

While ODDR achieves strong shadow removal performance without real-world paired supervision, several limitations remain. First, on extremely hard shadow boundaries with large intensity gaps, ODDR can leave faint halos or over-smooth the transition, since ShadowReward's boundary degradation (B) is trained with only mild dilation and Gaussian blur and thus does not fully cover such out-of-distribution edges. Second, on highly textured or non-flat surfaces (e.g., grass, foliage, or fabric), the recovered shadow region may exhibit inconsistent texture or subtle color drift, as synthetic degradations do not faithfully reproduce these surface-specific interactions (Fig. 12). Third, ShadowReward is trained on five synthetic degradation types (texture distortion, additive noise, chromatic shift, luminance change, and boundary artifact), and therefore provides limited supervision against artifacts that lie outside this taxonomy, such as semantic inconsistency or geometric misalignment introduced by the diffusion prior, or haze-type degradation, for which it provides no signal (Tab. 18). Fourth, since ODD is trained with synthetic shadows generated from AISTD, it handles complex scenes less well, including penumbra-dominated, multiple or overlapping, and self-shadows; on the physically grounded Ambient6K benchmark, the base model reaches only 15.65 dB PSNR (Tab. 17). Although ShadowReward still improves it, diversifying the synthetic shadow pipeline to cover such cases remains future work. Finally, our framework is developed under the standard single-image setting with a static scene; extensions to video shadow removal, multi-light scenes, and joint reasoning with scene geometry are promising directions for future work.

Failure cases of ShadowReward. Regarding the 5 of 16 user-study questions where ShadowReward disagrees with the human consensus, its ranking is least reliable on very faint shadows, where the residual shadow lies near the noise floor and the candidates differ only marginally; such cases are also difficult for human annotators to rank consistently. The remaining disagreements involve artifacts outside the five degradation types, such as diffusion-induced texture or structure changes, which we regard as the current coverage boundary of the taxonomy.

Broader impacts. ODDR is a low-level restoration method with limited misuse risk; like any image-editing tool, restored outputs should be treated as edited rather than original.

I Licenses of Existing Assets

The proposed framework builds on the following publicly available assets, which we use within the bounds of their respective licenses:

- **Stable Diffusion v2.1** [52] – CreativeML Open RAIL++-M License.
- **DINOv2** [45] – Apache License 2.0.
- **OSEDiff** [65] – Apache License 2.0.
- **LoRA** [20] – MIT License.
- **LPIPS** [76] – BSD 2-Clause License.
- **SynShadow** [23] – MIT License.
- **PromptIR** [48] – Academic Public License (non-commercial research use).
- **AISTD/ISTD+** [62, 31], **SRD** [50], **LRSS** [12], **UIUC** [17], **Ambient6K** [59], **SIDD** [1], **GT-RAIN** [4], and **NH-HAZE** [3] – released for academic / research use as specified by the original authors.

J Additional Visual Results

In this section, we provide additional visual results to further demonstrate the effectiveness and robustness of our proposed method across diverse scenes and conditions. As shown in Fig. 13, our method achieves significantly better visual quality on the AISTD [31] dataset than existing unsupervised shadow removal approaches, producing cleaner shadow-free regions and more consistent illumination. Moreover, even when compared with supervised methods, our results exhibit only marginal differences, highlighting the strong generalization capability of our framework despite the absence of paired supervision. Fig. 14 presents qualitative results on unseen datasets [12, 17], demonstrating the strong generalization ability of our ODDR framework. Despite never being trained on these datasets, ODDR consistently removes shadows while preserving natural color and structure, whereas existing methods often leave residual shadows or introduce color distortions. These results further confirm that ODDR is highly robust to domain shifts and remains effective even in challenging unseen scenarios.

To provide a more comprehensive analysis of our framework, we present additional visual comparisons between ODD and ODDR across multiple datasets. Fig. 15 shows qualitative results on the AISTD dataset, where ODDR produces cleaner shadow removal and more stable color consistency than ODD. Furthermore, as illustrated in Fig. 16 and Fig. 17, ODDR demonstrates superior performance on unseen datasets such as LRSS and UIUC, effectively preserving structural details while removing shadows that ODD occasionally fails to eliminate. These results collectively highlight the advantages of reward-guided optimization and confirm that ODDR delivers more robust and perceptually consistent restoration across both seen and unseen domains.

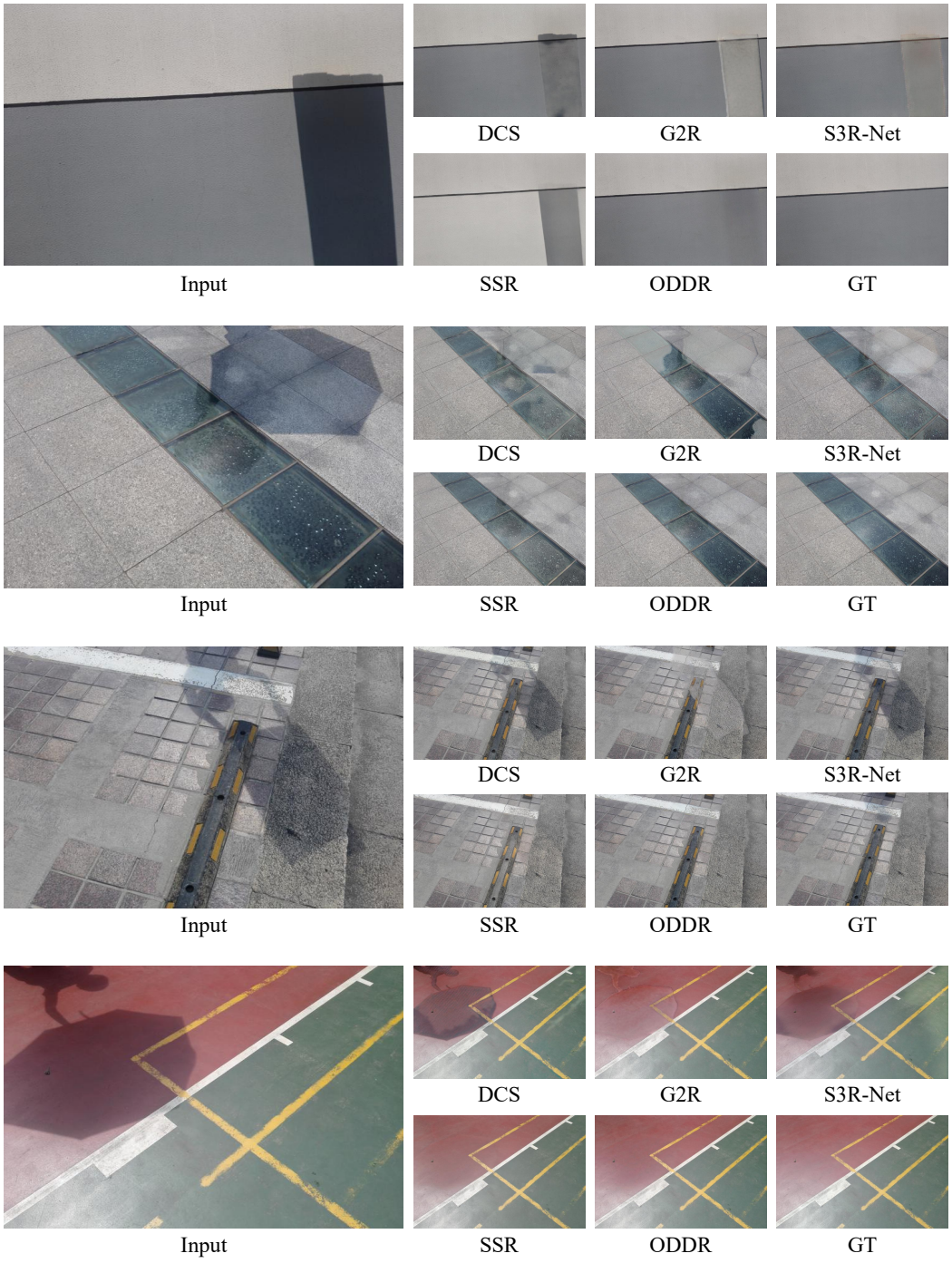

Figure 13: Additional visual comparison on the AISTD [31] dataset.

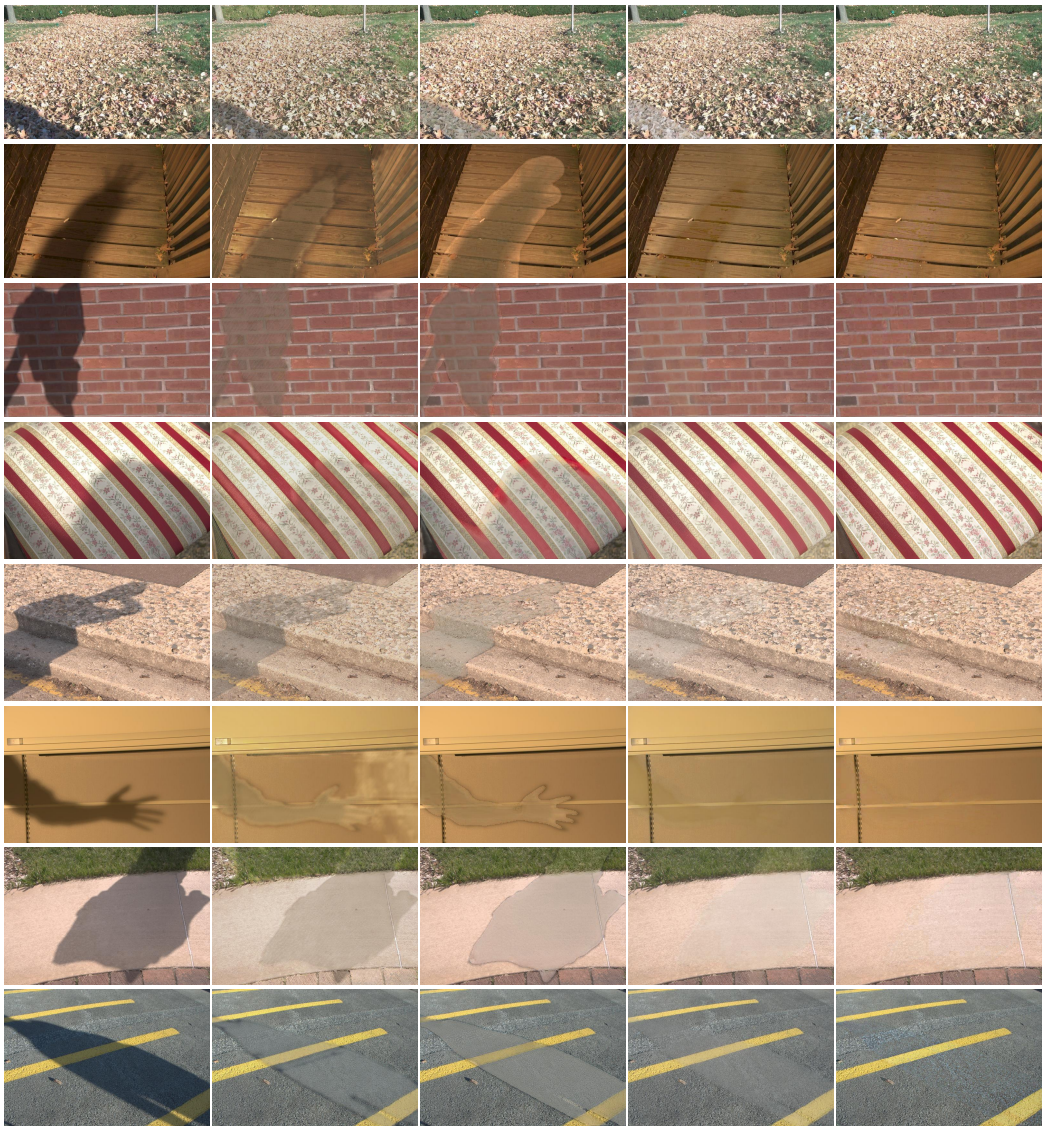

Figure 14: Additional visual comparisons on the LRSS [12] (top four rows) and UIUC [17] (bottom four rows) datasets.

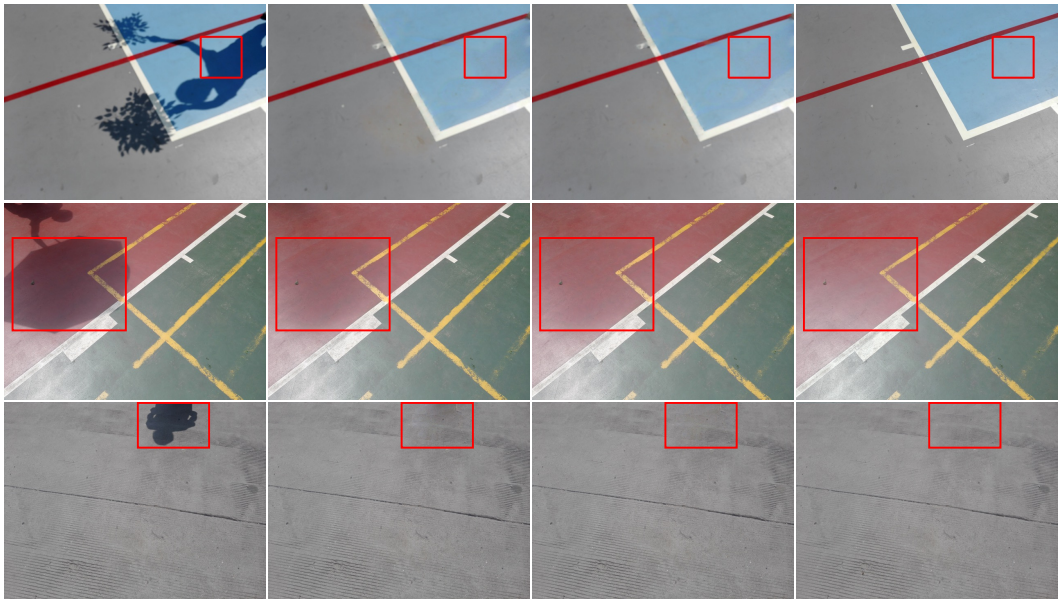

Figure 15: Additional visual comparison of ODD and ODDR on the AISTD [31] dataset.

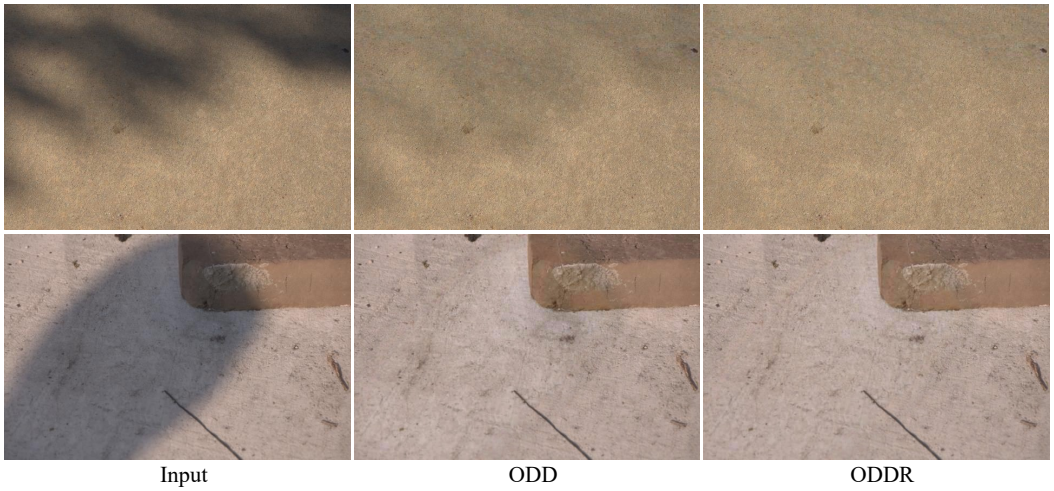

Figure 16: Additional visual comparison of ODD and ODDR on the unseen LRSS [12] dataset.

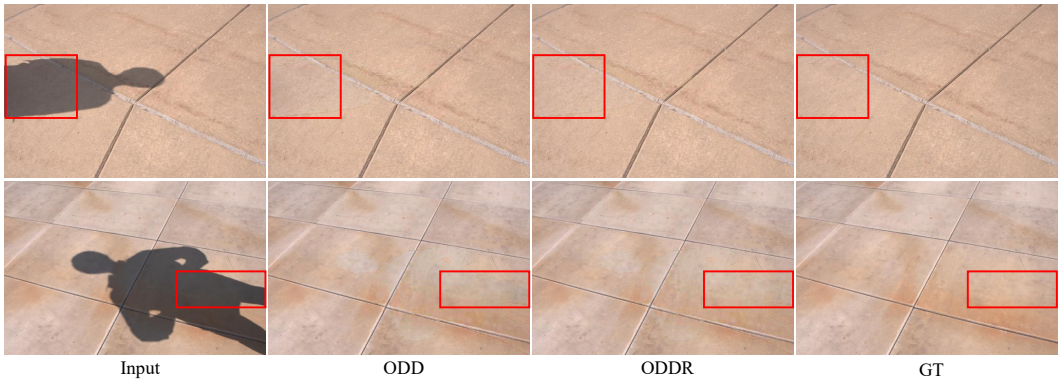

Figure 17: Additional visual comparison of ODD and ODDR on the unseen UIUC [17] dataset.

NeurIPS Paper Checklist

1. **Claims**

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

   Answer: [Yes]

   Justification: The abstract and introduction state our two main contributions, the One-step Deshadow Diffusion (ODD) baseline trained on synthetic pairs and its reward-guided refinement (ODDR) using ShadowReward, and these claims are directly supported by the quantitative results in Sec. 4 and the visual comparisons throughout the paper. The contribution of the reward is isolated by an anchor-only ablation (λ_r=0, Appendix Sec. E.1), and claims regarding the user study are stated with its limited sample size in mind (Appendix Sec. C).

   Guidelines:

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

2. **Limitations**

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

   Answer: [Yes]

   Justification: A dedicated Limitations and Failure Cases section is provided in the appendix (Sec. H), which discusses hard shadow boundaries, highly textured or non-flat surfaces, artifacts that lie outside the synthetic degradation taxonomy used to train ShadowReward (e.g., semantic inconsistency, geometric misalignment, and haze), complex scenes such as penumbra-dominated, overlapping, and self-shadows, failure cases of ShadowReward on very faint shadows, and the scope of the framework (single-image, static setting; video and multi-light scenarios are left as future work).

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.

- While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

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

   Answer: [N/A]

   Justification: The paper presents an empirical method for shadow removal and does not introduce formal theoretical results that require proofs.

   Guidelines:

   - The answer [N/A] means that the paper does not include theoretical results.
   - All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
   - All assumptions should be clearly stated or referenced in the statement of any theorems.
   - The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
   - Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
   - Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

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

   Answer: [Yes]

   Justification: The main paper describes the model architecture, training objectives, and evaluation protocol, and the appendix additionally provides hyperparameters, optimizer settings, training schedules, and dataset preprocessing details for ODD, ShadowReward, and ODDR.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.

(b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.

(c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).

(d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

   Justification: Code and checkpoints are released at `https://github.com/junsung6140/ODDR`. All datasets used in this work (AISTD, SRD, LRSS, UIUC, USR, Ambient6K, SIDD, GT-RAIN, and NH-HAZE) are publicly available, and our preprocessing follows their original protocols described in the appendix.

   Guidelines:

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

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: Training and test details, including data splits, input resolution, batch size, optimizer (AdamW), learning rate, training iterations, and loss weights for each of ODD, ShadowReward, and ODDR, are described in the experiments section and the appendix.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

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

   Answer: [No]

   Justification: Following the standard evaluation protocol in shadow removal literature, we report PSNR, SSIM, and RMSE on fixed benchmark test splits without error bars, since the test sets are deterministic and prior works use the same single-run evaluation. The reported gains are consistent across multiple datasets (AISTD, SRD, LRSS, UIUC), which supports the robustness of our results. For the user study, we additionally report inter-annotator agreement and explicitly note that its 16-question size does not allow a statistically significant comparison with PSNR-based ranking (Appendix Sec. C).

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.
   - It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
   - For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
   - If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

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

   Answer: [Yes]

   Justification: Training is conducted on a single NVIDIA RTX A6000 GPU, and the efficiency analysis (FLOPs, inference time, NFE) is measured on a single NVIDIA RTX 4090 GPU at 256×256 resolution, as stated in the experiments section.

   Guidelines:

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

9. **Code of ethics**

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

Answer: [Yes]

Justification: We have reviewed the NeurIPS Code of Ethics, and our research conforms to it. We use only publicly available datasets, the user study collects anonymous perceptual rankings without sensitive information, and the work is unlikely to cause direct harm.

Guidelines:

- The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

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

Answer: [Yes]

Justification: On the positive side, our method improves shadow removal quality and efficiency, which benefits downstream applications such as robotics, autonomous driving, OCR, and photo enhancement. On the negative side, like other image restoration models built on diffusion priors, it could in principle be misused as part of a pipeline for image manipulation, although the method itself only edits illumination artifacts and does not generate or alter scene content.

Guidelines:

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

11. **Safeguards**

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

Answer: [N/A]

Justification: Our model is a task-specific shadow removal network with low risk for misuse and does not release new scraped datasets or generative models capable of synthesizing arbitrary content.

Guidelines:

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

12. **Licenses for existing assets**

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

    Answer: [Yes]

    Justification: All third-party assets are cited in the paper and listed with their licenses in the appendix (Sec. I): Stable Diffusion v2.1 (CreativeML Open RAIL++-M), DINOv2 (Apache 2.0), OSEDiff (Apache 2.0), LoRA (MIT), LPIPS (BSD 2-Clause), SynShadow (MIT), PromptIR (Academic Public License), and the AISTD/ISTD+, SRD, LRSS, UIUC, Ambient6K, SIDD, GT-RAIN, and NH-HAZE datasets (academic / research use as specified by the original authors). We use each asset within the bounds of its license.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
    - If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [Yes]

    Justification: No new datasets are released with this submission. Code and weights are released at `https://github.com/junsung6140/ODDR` with a README and license (CC BY-NC 4.0).

    Guidelines:

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

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [Yes]

    Justification: We conducted a small-scale user study with 23 voluntary participants to evaluate alignment with human perceptual preferences. The appendix describes the study setup, the instructions given to participants, and the example interface used to collect rankings. Participation was voluntary and no monetary compensation was provided.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [No]

    Justification: The user study only asks participants to rank visual quality of natural-image shadow removal results and collects no personal or sensitive information, posing no greater than minimal risk; under our institution's guidelines this type of perceptual evaluation does not require formal IRB review.

    Guidelines:

    - The answer [N/A] 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.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: LLMs are not used as part of the core method. We use BLIP only for an auxiliary text-conditioning ablation in the analysis section, and any LLM usage is limited to writing assistance that does not affect the methodology or scientific contributions.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.