HNDiff: Haze-Noise Diffusion for Image Dehazing

Jin-Ting He[1], Fu-Jen Tsai[2], Yan-Tsung Peng[3], Min-Hung Chen[4], Chia-Wen Lin[2], and Yen-Yu Lin[1]

[1] Department of Computer Science,
National Yang Ming Chiao Tung University, Taiwan
jinting.cs12@nycu.edu.tw, lin@cs.nycu.edu.tw
[2] National Tsing Hua University, Taiwan
fjtsai@gapp.nthu.edu.tw, cwlin@ee.nthu.edu.tw
[3] National Chengchi University, Taiwan
ytpeng@cs.nccu.edu.tw
[4] NVIDIA, Taiwan
minhungc@nvidia.com

Abstract. Existing diffusion-based methods have recently made significant progress in image dehazing. However, they typically neglect the physics of haze formation and reconstruct clean images from pure Gaussian noise, thereby limiting their restoration potential. To address this issue, we propose Haze-Noise Diffusion (HNDiff), a novel diffusion framework that embeds the atmospheric scattering model as an inductive bias. By grounding diffusion in physical principles, HNDiff ensures that the restoration aligns more closely with underlying mechanisms of haze formation. In its forward process, we introduce joint haze-noise diffusion with a haze-aware noise scheduler, which progressively adds both haze and noise to an image. Essentially, the scheduler adapts noise levels according to haze density, meaning that regions with heavier haze receive stronger noise injection to encourage content generation, while clearer regions receive lighter noise to better preserve details, which directly links the forward degradation process with the physics of haze. In the reverse process, we then derive a physically consistent dehazing-denoising process that simultaneously removes haze and noise to restore a clean image in a manner aligned with the forward degradation process. To further enhance practicality, we propose Latent HNDiff, which compiles clean latent priors that can be seamlessly integrated into existing dehazing networks to boost performance. Extensive experiments show that our work significantly improves leading dehazing backbones and achieves state-of-the-art results on benchmark datasets. The project page is available at https://jin-ting-he.github.io/HNDiff/.

Keywords: Image Dehazing · Haze-Noise Diffusion · Atmospheric Scattering Model

1 Introduction

Hazy weather conditions caused by atmospheric scattering frequently degrade image visibility by reducing contrast and obscuring scene details. Such degra-

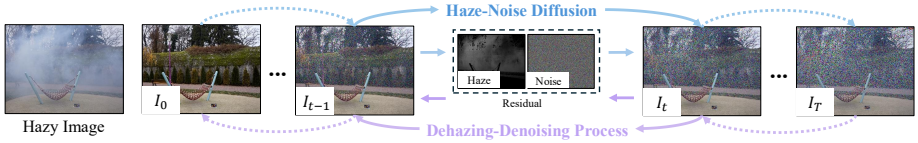

Fig. 1: HNDiff leverages the ASM inductive bias, progressively adding haze and noise in the forward process and removing them in the reverse process for image dehazing.

dation not only impairs human perception but also severely hinders the performance of many vision applications, such as object detection [20, 44], semantic segmentation [5, 47], and face recognition [21, 30]. To address the challenges, single image dehazing has emerged as a feasible solution to restore a clear image from a single hazy input. However, such a task remains highly ill-posed due to the complex interplay of scattering coefficients, atmospheric light, and scene depth.

Driven by advances in deep learning, CNN-based methods [4, 8, 10, 48] have achieved impressive results in image dehazing. Transformer-based approaches [9, 12, 14, 32, 33] further improved performance by exploring long-range dependencies and global context. Recently, Mamba-based methods [23, 58] have emerged as an efficient alternative with linear computational complexity. Despite the advancements, these methods still struggle in heavy haze scenarios, where most information is lost, leading to limited restoration quality.

In parallel, diffusion models [18, 36] have shown strong generative ability in image synthesis, producing results with rich details and sharp textures. Motivated by this progress, several studies [46, 53] have applied diffusion algorithms to image dehazing. Yet, conventional diffusion models are fundamentally misaligned with the nature of haze. That is, they reconstruct clean images from pure Gaussian noise, and their stochastic nature [54] often causes deviations from the original image, thereby reducing restoration fidelity. More importantly, they usually neglect the physical properties of haze formation, resulting in suboptimal restoration performance. According to the Atmospheric Scattering Model (ASM) [31], a hazy image results from attenuated scene radiance and global atmospheric light. Haze density may vary spatially across the image, as it depends on both the light scattering effect and scene depth, and intensifies with the increasing scattering coefficient. Thus, haze exhibits structured, spatially varying degradations, unlike the random Gaussian noise typically assumed in conventional diffusion models.

Based on this observation, we present *Haze-Noise Diffusion* (HNDiff), a new framework that redefines the forward process through a haze-noise diffusion mechanism. Instead of injecting only Gaussian noise, we integrate the ASM into the diffusion process, mimicking the physical formation of haze by highlighting its spatially varying characteristics.

In the forward process, HNDiff carries out the haze-noise diffusion mechanism, which gradually adds both Gaussian noise and haze to a clean image,

as illustrated in Fig. 1. To better control this process, we introduce the haze-aware noise scheduler, which dynamically adjusts the noise level according to haze density: hazier regions are assigned higher noise to boost generative capacity, while clearer regions receive less noise to preserve detail fidelity. Progressive haze diffusion and adaptive noise scheduling require transmission maps from ASM, which are generally unavailable. To overcome this limitation, we consider the haze residual, defined as the incremental haze accumulated as the scattering coefficient increases. We develop a continuous accumulation formulation to represent this residual implicitly in HNDiff and thus eliminate the need for explicit transmission maps. Through this design, the forward process remains ingeniously consistent with ASM, enabling progressive haze addition in tandem with adaptive noise injection.

In the reverse process, we derive the dehazing-denoising process, which is grounded by the physical principles of ASM and can implicitly approximate noise and haze residuals through dedicated estimators, thereby removing haze and noise to restore clean images. However, directly applying diffusion in the image space incurs substantial computational overhead and may suffer from fidelity issues in severely degraded regions due to the stochastic nature of the diffusion process. To address these problems, we propose Latent HNDiff, a prior generation network that integrates flexibly with dehazing backbones, allowing for more accurate and visually consistent restoration. This latent approach not only reduces computational cost but also enhances the applicability of the framework across diverse dehazing models.

The key contributions of our work are summarized as follows:

- We propose HNDiff, a novel diffusion-based framework that incorporates ASM as an inductive bias, specifically designed for image dehazing.
- HNDiff implements a haze-noise diffusion forward process that adds both haze and noise, and a corresponding dehazing-denoising reverse process with two dedicated estimators to respectively remove haze and noise.
- We design the haze-aware noise scheduler to adaptively adjust noise levels based on haze densities.
- Extensive experiments demonstrate that HNDiff consistently improves four representative dehazing models and achieves state-of-the-art performance on seven benchmark datasets.

2 Related work

2.1 Image Dehazing

CNN-based Dehazing. Deep learning has revolutionized image dehazing with CNN-based methods [4, 8, 10, 42, 48, 52] achieving impressive breakthroughs. For instance, Dong *et al.* [10] proposed a boosted decoder combined with a dense feature fusion module to progressively restore images. Wu *et al.* [48] introduced contrastive regularization within an autoencoder for efficient dehazing. More

recently, Cui *et al*. [8] presented a dual-domain selection mechanism and an efficient multi-scale network to further enhance restoration quality.

Transformer-based Dehazing. In addition to CNNs, Transformer-based methods have shown great promise in image dehazing by leveraging attention mechanisms to model long-range dependencies and global context [9,12,14,32,33,40,43]. For example, Qiu *et al*. [33] approximated softmax-attention with a Taylor expansion to achieve linear complexity for effective dehazing. Cui *et al*. [9] designed a multi-shape attention module with rectangle and dilated operations to enlarge receptive fields. Fang *et al*. [12] integrated phase and attention modules to leverage YCbCr textures for recovering clearer features.

Mamba-based Dehazing. Mamba-based methods have recently emerged as efficient alternatives for image dehazing, capturing global context with linear computational complexity. Zheng & Wu [58] combined convolution for local feature extraction with state space models to capture long-range dependencies in dehazing. Li *et al*. [23] designed an S-shaped stripe-based scanning strategy to better preserve locality and continuity for more effective restoration.

ASM-based Dehazing. Beyond architectural advances, several studies [6, 11, 19, 37, 38, 50, 52] explicitly exploit the Atmospheric Scattering Model (ASM) to improve dehazing. Wu *et al*. [50] designed an ASM-based data generation pipeline to synthesize hazy images for network training. Fang *et al*. [11] derived a cooperative unfolding network directly from ASM, jointly optimizing the transmission map and the clean image.

Despite these advancements, most dehazing methods are still trained end-to-end as direct regressors from hazy inputs to clean outputs. Although some ASM-based approaches incorporate the ASM as physical guidance to constrain this mapping, both regression-based and ASM-based methods often struggle under extremely dense haze, as the severe information loss makes it difficult to recover realistic high-frequency details. In contrast, our method couples the ASM with a diffusion process and leverages the generative capability of noise diffusion to compensate for missing content and restore plausible fine structures in heavily degraded regions.

2.2 Diffusion Models

Diffusion for Low-level Vision. Diffusion models [18,29,36,49,55] have shown strong generative capability in image synthesis and can produce results with rich details and sharp textures through forward noise diffusion and reverse denoising. This success has inspired much research exploring their potential in diffusion algorithms for various low-level vision tasks [13,15,17,24,26–28,34,35,51,56,57]. For example, He *et al*. [17] introduced a domain-adaptive blur condition to guide a diffusion-based blurring model, enabling the generation of domain-adaptive blurred videos. Xia *et al*. [51] employed diffusion models to extract compact priors used to guide a dynamic transformer for image recovery. Liu *et al*. [27] utilized a pre-trained diffusion model with task-specific priors for diverse image restoration tasks. Luo *et al*. [28] presented visual instruction-guided diffusion that models degradation patterns for all-in-one image restoration.

Diffusion for Image Dehazing. Within low-level vision, several studies have focused specifically on image dehazing using diffusion models [25, 27, 45, 46, 53]. For instance, Yang *et al*. [53] exploited the semantic latent space of a pre-trained diffusion model to guide dehazing without retraining. Wang *et al*. [46] combined diffusion-based hazy image generation with accelerated fidelity-preserving sampling for efficient dehazing. Liu *et al*. [25] leveraged diffusion models in the frequency domain with an amplitude residual encoder to enhance unpaired image dehazing. Despite these advances, these methods rely on conventional noise diffusion initialized with pure Gaussian noise, which disregards the physical properties of haze formation. As a result, the stochastic nature [54] of the process often leads to deviations from the target restoration fidelity, resulting in degraded performance and less consistent visual quality.

Degradation-aware Diffusion. Recently, a few studies [16, 26, 59] have explored degradation-aware diffusion for image restoration. For example, Liu *et al*. [26] proposed residual diffusion and operated on the difference between hazy and clean images. He *et al*. [16] designed blur diffusion that implicitly models the blur formation process through a dual-diffusion forward scheme for image deblurring. Zhou *et al*. [59] introduced a physics-guided dehazing diffusion by reformulating haze accumulation as a time-indexed process. However, these approaches either neglect the role of physical scene transmission in modeling haze degradation within the diffusion process or overlook haze density in the noise scheduling. To address these issues, our approach embeds ASM into the diffusion process as an inductive bias and adaptively adjusts noise addition according to haze density, achieving significantly improved dehazing performance.

3 Method

This section presents the proposed *Haze-Noise Diffusion* (HNDiff), a novel framework that integrates the Atmospheric Scattering Model (ASM) [31] into the diffusion process for image dehazing. As depicted in Fig. 1, HNDiff defines a physics-guided forward *Haze-Noise Diffusion* process equipped with a *Haze-Aware Noise Scheduler* (HANS) and a reverse *Dehazing–Denoising Process*. In the forward direction, an input image is progressively degraded by jointly introducing haze and Gaussian noise as the scattering coefficient increases, while HANS adaptively controls the noise level based on the local haze density, ensuring the corruption remains consistent with ASM-guided haze formation. In the reverse direction, HNDiff starts from a hazy input perturbed by Gaussian noise and iteratively removes both haze and noise in a manner consistent with the ASM. For high-fidelity restoration and better efficiency, as illustrated in Fig. 2, we further propose *Latent HNDiff*, where HNDiff serves as a prior generation network in the latent space, and the learned prior is injected into a dehazing backbone via a Feature Gating Module (FGM), enabling plug-and-play enhancement of existing dehazing architectures. In the following, Sec. 3.1 details the Haze-Noise Diffusion process and HANS, Sec. 3.2 presents the Dehazing–Denoising Process, and Sec. 3.3 describes Latent HNDiff together with the FGM integration.

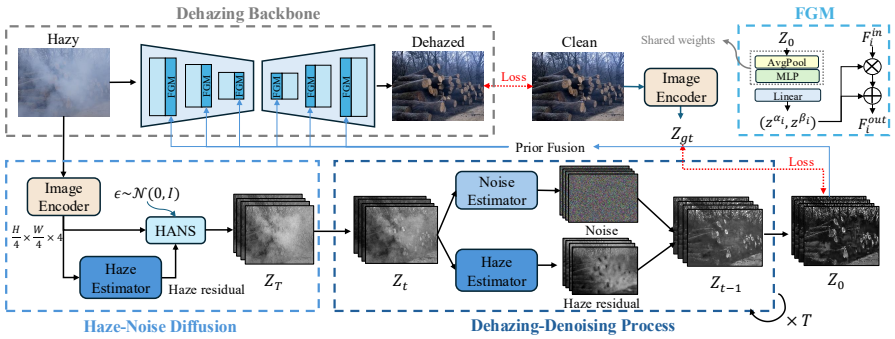

Fig. 2: Overview of Latent HNDiff. The framework starts by using an image encoder to extract latent priors from a hazy input. These priors undergo a haze-noise diffusion process to produce the diffused hazy and noisy representation Z_T. During the reverse process, dehazing and denoising are performed jointly by iteratively estimating both noise and haze residuals to recover clean priors Z_0. Lastly, the recovered priors are integrated into a dehazing backbone via the Feature Gating Module (FGM) to improve restoration quality.

3.1 Haze-Noise Diffusion

Haze in image formation arises from atmospheric scattering, where scene radiance is attenuated during transmission and blended with global atmospheric light, leading to reduced visibility and a loss of detail. This process can be mathematically modeled by the ASM:

$$I_H(x) = I_0(x)\,\tau(x) + A\,(1 - \tau(x)), \text{ where } \tau(x) = e^{-\sigma(x)d(x)} \tag{1}$$

with x denoting the pixel index. In Eq. (1), $I_H \in \mathbb{R}^{H\times W\times 3}$, $I_0 \in \mathbb{R}^{H\times W\times 3}$, $\tau \in \mathbb{R}^{H\times W\times 1}$, $A \in \mathbb{R}^3$, $\sigma \in \mathbb{R}^{H\times W\times 1}$, and $d \in \mathbb{R}^{H\times W\times 1}$ denote the hazy image, the clean scene radiance, the transmission map, the global atmospheric light, the scattering coefficient, and the scene depth, respectively. This formulation explicitly models haze formation, where larger scattering coefficients σ or greater depths d yield smaller transmission τ, increasing haze density and reducing scene visibility. However, existing diffusion-based dehazing methods typically adopt conventional diffusion models that add zero-mean Gaussian noise and drive the image toward pure noise, which does not reflect the structured, spatially varying haze described in Eq. (1). To bridge this gap, we introduce a *haze–noise diffusion* process in which ASM-based haze formation acts as a mean shift of the Gaussian corruption, so that the forward process follows a physically meaningful clean-to-hazy evolution while stochastic noise is injected around this trajectory.

Forward Haze-Noise Diffusion. In the forward process of HNDiff, we propose a haze-noise diffusion that embeds the physical haze formation process into noise diffusion. Specifically, a clean image is progressively degraded by both haze and Gaussian noise. The forward transition at time step t is defined as

$$I_t(x) = I_{t-1}(x)e^{-\alpha_t\sigma(x)d(x)} + A\left(1 - e^{-\alpha_t\sigma(x)d(x)}\right) + \beta_t(x)\epsilon_t(x), \tag{2}$$

where I_t denotes the intermediate hazy and noisy image at step t, α_t is the scaling factor, $\epsilon_t \sim \mathcal{N}(0, \mathbf{I})$ represents Gaussian noise, and $\beta_t(x)$ is the noise scaling coefficient at pixel x. As seen in Eq. (2), I_{t-1} is further degraded according to Eq. (1), with scene radiance attenuated by the α_t-scaled scattering coefficient, while noise is injected with another scaling coefficient $\beta_t(x)$.

Haze-Aware Noise Scheduler. Since haze density varies spatially, the generative capacity controlled by noise diffusion should also adapt across pixels. We therefore introduce a haze-aware noise scheduler, which defines the pixel-wise noise scaling coefficient as $\beta_t(x) = 1 - e^{-\alpha_t \sigma(x) d(x)}$. This design makes the injected noise explicitly dependent on haze density: pixels with heavier haze receive larger $\beta_t(x)$, thus introducing stronger noise that triggers diffusion to reconstruct severely degraded details; conversely, pixels with lighter haze yield smaller $\beta_t(x)$, injecting less noise to preserve content fidelity through regression. This adaptive scheduling enables the forward process to jointly model haze degradation and stochastic corruption.

Sampling Probability and Reparameterization. Inspired by [16, 26], we regard degradation in the forward process as a deterministic mean shift. From Eq. (2), each step from I_{t-1} to I_t can thus be expressed as a Gaussian transition, where the mean is shifted by haze and stochastic perturbations are introduced by Gaussian noise:

$$q(I_t(x) \mid I_{t-1}(x), \phi) \tag{3}$$

$$:= \mathcal{N}\Big(I_t(x) \;\Big|\; I_{t-1}(x) e^{-\alpha_t \sigma(x) d(x)} + A\left(1 - e^{-\alpha_t \sigma(x) d(x)}\right),\, \beta_t^2(x)\Big), \tag{4}$$

where $\phi = \{d(x), \sigma(x), A\}$. By iterating Eq. (4), we obtain a sequence of progressively hazy and noisy images $\{I_1, I_2, \ldots, I_T\}$ through a T-step diffusion process, with the complete forward sampling probability $q(I_{1:T}(x) \mid I_0(x), \phi) = \prod_{t=1}^{T} q(I_t(x) \mid I_{t-1}(x), \phi)$. However, existing dehazing datasets provide only hazy-clean image pairs and do not include ϕ (i.e., atmospheric light, scattering coefficients, and scene depth necessary to compute transmission).

To address this limitation, we apply the reparameterization trick [18] to Eq. (4) and obtain the conditional distribution after T steps as

$$q(I_T(x) \mid I_0(x), \phi) \tag{5}$$

$$= \mathcal{N}\Big(I_T(x) \;\Big|\; I_0(x) e^{-\sum_{t=1}^{T} \alpha_t \sigma(x) d(x)} + A\left(1 - e^{-\sum_{t=1}^{T} \alpha_t \sigma(x) d(x)}\right), \bar{\beta}_T^2(x)\Big), \tag{6}$$

where $\alpha_t = \frac{1}{T},\ \forall t \in \{1, 2, \ldots, T\}$, and $\bar{\beta}_T(x) = \sqrt{\frac{(1 - e^{-(1/T)\sigma(x)d(x)})(1 - e^{-2\sigma(x)d(x)})}{1 + e^{-(1/T)\sigma(x)d(x)}}}$. The complete derivation of Eq. (6) is provided in the supplementary materials (Sec. 1). It follows that the hazy and noisy image I_T can be sampled from $q(I_T \mid I_0)$ via

$$I_T(x) = I_0(x) e^{-\sigma(x)d(x)} + A\left(1 - e^{-\sigma(x)d(x)}\right) + \bar{\beta}_T(x)\epsilon(x) = I_H(x) + \bar{\beta}_T(x)\epsilon(x), \tag{7}$$

where I_T is generated in a single step by injecting noise into the hazy image I_H via the haze-aware noise scheduler. This formulation preserves the Gaussian

nature of the diffusion process while embedding ASM directly into the mean of the distribution through a physically grounded shift. As $\bar{\beta}_T(x)$ still relies on the transmission map, we introduce a learnable haze estimator to implicitly approximate it. The optimization details for the haze estimator are provided in Sec. 3.3.

3.2 Dehazing-Denoising Process

In the reverse generation procedure, we aim to progressively remove both haze and noise from the degraded observation I_T to recover the clean image I_0. Unlike conventional diffusion models that start from pure Gaussian noise, our method initializes from the hazy-noisy sample I_T drawn from the Gaussian distribution Eq. (6). Inspired by the deterministic sampling formulation in [39], we define the reverse transition distribution as

$$p_\theta(I_{t-1}(x) \mid I_t(x)) = q_\delta(I_{t-1}(x) \mid I_t(x), I_0(x), \phi). \tag{8}$$

The transition probability q_δ in Eq. (8) is defined as

$$q_\delta(I_{t-1}(x) \mid I_t(x), I_0(x), \phi) = \mathcal{N}\left(I_{t-1}(x) \mid \mu_t(x), \delta_t^2(x)\right), \quad \text{where} \tag{9}$$

$$\mu_t(x) = I_0(x)e^{-\sum_{s=1}^{t-1}\alpha_s\sigma(x)d(x)} + A\left(1 - e^{-\sum_{s=1}^{t-1}\alpha_s\sigma(x)d(x)}\right) \tag{10}$$

$$+ \sqrt{\bar{\beta}_{t-1}^2(x) - \delta_t^2(x)}\epsilon_{t-1}(x), \tag{11}$$

and $\delta_t^2 = \eta \cdot \frac{\beta_t^2\bar{\beta}_{t-1}^2}{\bar{\beta}_t^2}$ is a variance term that controls sampling stochasticity. When $\eta = 0$, this yields a deterministic sampling. From Eq. (6), we can derive

$$I_0(x) = (I_t - (1 - e^{-\sum_{s=1}^{t-1}\alpha_s\sigma(x)d(x)})A - \bar{\beta}_t\epsilon_t)e^{\sum_{s=1}^{t-1}\alpha_s\sigma(x)d(x)}. \tag{12}$$

By substituting Eq. (12) into Eq. (9) and simplifying, we obtain the sampling equation for $I_{t-1}(x)$ as

$$I_{t-1}(x) = \left(I_t(x) - N_t(x)\left(1 - e^{-\alpha_t\sigma(x)d(x)}\right)\right)e^{\alpha_t\sigma(x)d(x)}, \tag{13}$$

where $N_t(x) = A + \epsilon_t(x)$ denotes the atmospheric noise, which is composed of the atmospheric light term and a Gaussian noise term. To reconstruct I_0, we iterate Eq. (13) with two learnable estimators. One is the noise estimator $N_t^\theta(I_t, I_H, t)$, which approximates N_t. The other is the haze estimator $1 - e^{-\alpha_t o^\theta(I_t, I_H, t)}$, which approximates the residual transmission term $1 - e^{-\alpha_t\sigma d}$ (the complement of the transmission), where $o^\theta(I_t, I_H, t)$ is a learnable network estimating the scattering–depth product σd. Complete derivations of the variational lower bound and the sampling formulation are provided in the supplementary materials (Secs. 2 and 3). In the following, we detail the optimization of the haze estimator and noise estimators in the latent space.

3.3 Latent HNDiff

Performing diffusion-based restoration directly in image space, as noted in [7, 16, 36], can be computationally expensive and often result in slower, less stable optimization, with limited reconstruction fidelity. To address these challenges, and inspired by previous works [7, 16, 36, 51], we present Latent HNDiff. As illustrated in Fig. 2, our method runs the haze-noise diffusion procedure on an encoded representation extracted from the input image to generate informative priors, which are then injected into a dehazing backbone via the proposed Feature Gating Module (FGM). By incorporating physically grounded haze formation into the diffusion procedure, we encourage the learned priors to capture haze-aware cues that ultimately improve restoration quality.

Training strategy. Following the training scheme from [16], we first pretrain a dehazing backbone equipped with an Image Encoder (IE) and a Feature Gating Module (FGM) using paired data (I_H, I_0). The ground-truth prior is extracted as $Z_{gt} = \text{IE}(\text{Concat}(I_H, I_0))$, and fused into multi-scale features F_i^{in} through FGM:

$$z_{gt} = \text{MLP}(\text{AvgPool2D}(\text{Unshuffle}(Z_{gt}))) \in \mathbb{R}^{1\times C}; \tag{14}$$

$$F_i^{out} = F_i^{in} \times z^{\alpha_i} + z^{\beta_i}, \quad \text{where } (z^{\alpha_i}, z^{\beta_i}) = \text{Linear}(z_{gt}). \tag{15}$$

Next, we use the proposed HNDiff framework to learn a prior from the degraded representation Z_T (obtained using the haze-noise diffusion process) and then recover the clean priors Z_0, supervised by $\mathcal{L}_{\text{prior}} = \|Z_0 - Z_{gt}\|_1$. Lastly, we jointly fine-tune the IE, HNDiff, FGM, and the dehazing backbone, using the learned Z_0 to reconstruct the dehazed image I_{dehz}, supervised by I_0 under the standard backbone losses. More details are provided in supplementary materials (Sec. 4).

4 Experiments

4.1 Experimental Setup

Implementation Details. HNDiff is composed of four key components: the Image Encoder (IE), the Feature Gating Module (FGM), the Haze Estimator, and the Noise Estimator. The IE consists of six residual blocks and four CNN layers, while the FGM is implemented with a pooling operation and a lightweight MLP. Both the Haze Estimator and Noise Estimator share the same network architecture, which is a simplified U-Net [26]. In practice, we set the diffusion step to $T = 4$. The overall framework is optimized with the default hyperparameter and training protocol of each dehazing backbone (e.g., learning rate, number of epochs, batch size, and optimizer) to ensure fair comparisons.

Dehazing Models and Datasets. We adopt four state-of-the-art image dehazing models, including FocalNet [8], ConvIR [9], SGDN [12], and RIDCP [50] to validate the effectiveness of HNDiff. Following prior studies, we conduct

Table 1: Quantitative results on six benchmark datasets. Values in parentheses represent the improvements of HNDiff over the corresponding baselines.

Model	NH-HAZE		O-HAZE		Dense-HAZE		RW^{2}AH		SOTS-Indoor		SOTS-Outdoor	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
MSBDN [10]	17.97	0.659	24.36	0.749	15.13	0.555	21.51	0.595	33.67	0.985	33.48	0.982
FFA-Net [32]	18.13	0.647	22.12	0.770	15.70	0.549	18.73	0.556	36.39	0.989	33.57	0.984
Dehamer [14]	20.66	0.684	25.11	0.777	16.62	0.560	20.84	0.581	36.63	0.988	35.18	0.986
MB-Taylor [33]	20.43	0.688	25.05	0.788	16.66	0.560	21.37	0.608	40.71	0.992	37.42	0.989
FocalNet [8]	20.36	0.696	25.46	0.791	16.95	0.597	21.93	0.635	40.82	0.992	37.71	0.995
ConvIR [9]	20.65	0.692	25.25	0.784	16.86	0.600	21.99	0.640	41.53	0.994	37.95	0.994
SGDN [12]	20.13	0.680	24.59	0.778	16.60	0.571	22.24	0.631	41.01	0.992	36.22	0.986
FocalNet+HNDiff	**20.89** (+0.53)	**0.697** (+0.001)	**26.32** (+0.86)	**0.801** (+0.010)	**17.29** (+0.34)	**0.599** (+0.002)	**22.29** (+0.36)	**0.647** (+0.012)	**41.19** (+0.37)	**0.994** (+0.002)	**38.10** (+0.39)	**0.996** (+0.001)
ConvIR+HNDiff	**21.23** (+0.58)	**0.701** (+0.009)	**26.20** (+0.95)	**0.799** (+0.015)	**17.18** (+0.32)	**0.623** (+0.023)	**22.25** (+0.26)	**0.646** (+0.006)	**42.10** (+0.57)	**0.995** (+0.001)	**38.83** (+0.88)	**0.995** (+0.001)
SGDN+HNDiff	**20.64** (+0.51)	**0.686** (+0.006)	**25.40** (+0.81)	**0.782** (+0.004)	**17.17** (+0.57)	**0.611** (+0.040)	**22.81** (+0.57)	**0.653** (+0.022)	**41.47** (+0.46)	**0.995** (+0.003)	**37.10** (+0.88)	**0.991** (+0.005)
Avg Gains	**+0.54**	**+0.005**	**+0.87**	**+0.010**	**+0.41**	**+0.022**	**+0.40**	**+0.013**	**+0.47**	**+0.002**	**+0.72**	**+0.002**

Fig. 3: Qualitative results on the O-HAZE (left) and NH-HAZE (right) datasets. "Res" denotes residual maps between outputs and ground truth, where darker intensities indicate smaller errors.

experiments on two widely used synthetic dataset, SOTS-Indoor and SOTS-Outdoor [22], and five real-world benchmarks: NH-HAZE [3], O-HAZE [2], Dense-HAZE [1], RW^{2}AH [12], and RTTS [22]. The SOTS-Indoor dataset consists of 13,990 training pairs and 500 testing pairs. The SOTS-Outdoor dataset consists of 313,950 training pairs and 500 testing pairs. Both NH-HAZE and Dense-HAZE provide 50 training pairs and 5 testing pairs. O-HAZE offers 40 training pairs and 5 testing pairs. RW^{2}AH includes 1,406 training pairs and 352 testing pairs. RTTS contains 4,322 hazy images and is used exclusively for testing.

4.2 Experimental Results

Quantitative Results. Tab. 1 and Tab. 2 compare the dehazing performance of state-of-the-art methods and their HNDiff-enhanced versions, where the values in parentheses indicate the improvements made by HNDiff over the respective dehazing baselines. The results clearly demonstrate that HNDiff consistently boosts each baseline and outperforms existing state-of-the-art methods. Specifically, as shown in Tab. 1, HNDiff yields average PSNR improvements of +0.54, +0.87,

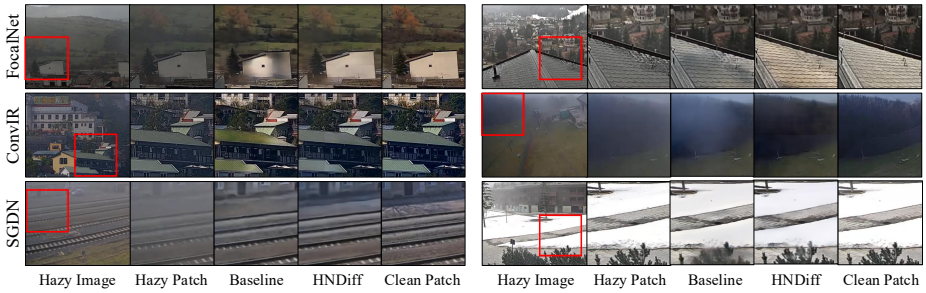

Fig. 4: Qualitative results on the RW^{2}AH dataset.

+0.41, +0.40, +0.47, and +0.72 dB on the NH-HAZE, O-HAZE, Dense-HAZE, RW^{2}AH, SOTS-Indoor, and SOTS-Outdoor test sets, respectively. Additionally, HNDiff achieves average PSNR improvements of +0.48, +0.59, and +0.63 on baselines FocalNet, ConvIR, and SGDN, respectively. Overall, HNDiff achieves an average gain of +0.57 dB in PSNR and +0.009 in SSIM across all datasets and baselines, highlighting its strong generalization ability and effectiveness as a prior generation network for image dehazing. In Tab. 2, the HNDiff-enhanced RIDCP achieves the best performance among the compared methods on the real-world RTTS benchmark, attaining 0.417 FADE, 5.08 NIMA, and 16.09 BRISQUE, which indicates strong generalization and improved perceptual quality under real-world haze.

Qualitative Results. We present qualitative comparisons between four baseline models and their HNDiff-enhanced counterparts. Fig. 3 shows the results on the NH-HAZE and O-HAZE test sets, including an additional “Res" column for better comparison. Fig. 4 presents the results on the RW^{2}AH test set. Fig. 5 provides results on the RTTS benchmark. The residual maps (“Res") are obtained by subtracting the ground truth from the model outputs, where lower intensities indicate smaller errors and thus higher reconstruction quality. As shown, HNDiff consistently produces cleaner and more visually compelling dehazed images compared to the baselines. By integrating HNDiff into the latent space of the dehazing networks, we exploit its capacity to model rich and realistic image priors while preserving fidelity to the underlying clean image structures. More qualitative results are provided in the supplementary materials (Sec. 8).

4.3 Ablation Studies

To assess the contributions of the proposed components in HNDiff, we conduct a series of ablation studies using FocalNet as the baseline dehazing model. Specifically, we evaluate the effectiveness of each component, compare the prior generator with different diffusion mechanisms, analyze computational overhead and performance gain, examine HNDiff against the baseline under comparable parameter counts, inspect the haze residual modeling in latent space, compare the dehazing results of applying HNDiff in image space and latent space, and

Table 2: Quantitative comparison on RTTS.

Method	FADE↓	NIMA↑	BRISQUE↓
Hazy image	2.484	4.33	37.01
MSBDN [10]	1.363	4.14	28.74
Dehamer [14]	1.895	3.87	33.87
DAD [37]	1.130	4.01	32.73
PSD [6]	0.920	4.35	25.24
RIDCP [50]	0.944	4.43	18.78
RIDCP+HNDiff	**0.417**	**5.08**	**16.09**

Table 3: Ablation study on the effectiveness of noise diffusion, haze diffusion, and HANS.

Model	Noise Diffusion	Haze Diffusion	HANS	PSNR
Net1				20.36
Net2	✓			20.46
Net3		✓		20.61
Net4	✓	✓		20.68
Net5	✓	✓	✓	**20.89**

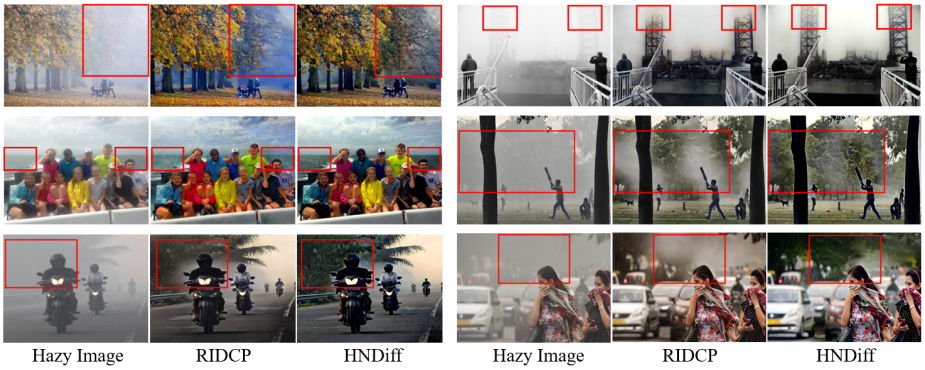

Fig. 5: Qualitative results on the RTTS dataset.

analyze the effect of varying the number of diffusion steps. All experiments are conducted using the default training configuration of FocalNet and evaluated on the NH-HAZE test set.

Effectiveness of Each Component. Our ablation study, detailed in Tab. 3, evaluates the contribution of each component in HNDiff. *Net1* denotes the baseline dehazing model. *Net2* represents a conventional DDPM-based variant that employs only noise diffusion, while *Net3* is a variant that incorporates only haze diffusion and omits noise diffusion. *Net4* adopts both noise and haze diffusion but excludes the haze-aware noise scheduler (HANS). Finally, *Net5* is the complete HNDiff design. The results show that both *Net2* and *Net3* surpass the baseline, demonstrating the individual benefits of noise and haze diffusion. Moreover, *Net5* achieves the best performance, indicating that the joint integration of both diffusion processes together with HANS provides complementary gains. These findings highlight the importance of incorporating haze-aware design in order to enhance dehazing effectiveness.

Comparison of Prior Generators with Different Diffusion Mechanisms. Tab. 4 evaluates the baseline dehazing model augmented with different prior generation methods, including U-Net, DDPM [18], RDDM [26], and our proposed HNDiff. *Net1* denotes the baseline model without a prior generator. *Net2* employs a U-Net to generate priors directly, without any diffusion process. *Net3*,

Table 4: Comparison of different prior generators.

Model	Prior Generator	PSNR	SSIM
Net1	N/A	20.36	0.696
Net2	U-Net	20.41	0.696
Net3	DDPM	20.46	0.695
Net4	RDDM	20.43	0.690
Net5	HNDiff	**20.89**	**0.697**

Table 5: Computational overhead, peak training VRAM, and PSNR gains of HNDiff.

Model	Params (M)	FLOPs (G)	VRAM (M)	Gains (dB)
FocalNet	3.74	30.53	2854	–
+ HNDiff	7.82	36.38	3932	**+0.48**
ConvIR	5.51	41.96	4506	–
+ HNDiff	9.59	45.39	5550	**+0.59**
SGDN	11.09	52.95	22980	–
+ HNDiff	15.30	56.38	23914	**+0.63**

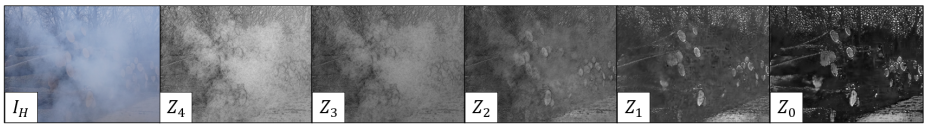

Fig. 6: Visualization of latent representations across reverse diffusion steps.

Fig. 7: Qualitative results of applying HNDiff in image space and latent space.

Net4, and *Net5* are the dehazing models enhanced with priors generated by DDPM, RDDM, and HNDiff, respectively. Although integrating the standard diffusion process (*Net3*) or the residual diffusion process (*Net4*) improves performance over the baseline, the gain is just comparable to that of *Net2*, which uses a U-Net without diffusion. In contrast, HNDiff explicitly embeds the atmospheric scattering model into the diffusion process, yielding consistent and superior improvements compared to both standard and residual diffusion mechanisms. We present the dehazed results of different diffusion mechanisms in the supplementary materials.

Analysis of Computational Overhead and Performance Gain. We evaluate the computational overhead using cropped 256×256 patches on a single NVIDIA RTX 4090 GPU with a batch size of 4. The reported VRAM denotes the memory required during training under this setting. As shown in Tab. 5, the PSNR gain is computed by averaging the improvements over the six datasets reported in Tab. 1. HNDiff consistently improves FocalNet, ConvIR, and SGDN by +0.48 dB, +0.59 dB, and +0.63 dB, respectively. Although our method introduces additional parameters and computational complexity, the achieved PSNR gains are substantial, demonstrating a favorable trade-off between computational cost and restoration performance.

Comparison Under Comparable Parameter Counts. To ensure a fair comparison under similar parameter budgets, we evaluate HNDiff against enlarged baseline variants, as reported in Tab. 7. Following the model-scaling

Table 6: Comparison of applying HNDiff in image space and latent space.

Metric	FocalNet	HNDiff (Image)	HNDiff (Latent)
PSNR (dB)	21.18	21.37	**21.52**
SSIM	0.5970	0.6166	**0.6254**
FLOPs (G)	30.53	65.59	36.38

Table 7: Comparison between HNDiff and baseline variants with comparable parameter counts.

	FocalNet	FocalNet$^+$	FocalNet*	HNDiff
Params (M)	3.74	8.40	8.28	7.82
FLOPs (G)	30.53	68.54	64.05	36.38
PSNR (dB)	20.36	20.37	20.51	**20.89**

study [41], we design these variants by scaling network width and depth. Specifically, *FocalNet*$^+$ increases the base channel size from 32 to 48, while *FocalNet** expands the number of residual blocks from 4 to 10. Although both variants substantially increase model complexity in terms of parameters and FLOPs, they yield only marginal PSNR gains over the baseline FocalNet. In contrast, HNDiff achieves the best performance of 20.89 dB in PSNR with a lower parameter count (7.82M) and significantly reduced FLOPs (36.38G). These results demonstrate that integrating the proposed diffusion prior is more effective than simply scaling the network capacity.

Analysis of Haze Residual Modeling in Latent Space. To verify that HNDiff models haze formation in latent space, we analyze diffusion prior outputs across reverse steps. Although the model is trained with $T = 4$ steps, we examine intermediate latent representations by performing $t \in [0, 1, 2, 3, 4]$ reverse steps starting from the fully hazy latent Z_4, resulting in the sequence $[Z_4, Z_3, \ldots, Z_0]$. For visualization, we compute the channel-wise mean of each latent and downsample I_H to the same spatial resolution only for visualization. As shown in Fig. 6, the representations progressively transition from hazy (Z_4) to clean (Z_0), confirming that HNDiff captures a progressive hazy-to-clean structure in the latent space and enables interpretable modeling.

Comparison of Applying HNDiff in Image Space and Latent Space. Tab. 6 compares an image-space variant, *HNDiff (Image)*, which applies our haze-noise diffusion directly to RGB images, and a latent-space variant, *HNDiff (Latent)*, which operates in the latent space of a FocalNet on the real-world RW^{2}AH dataset. The two variants use the same U-Net as the haze/noise estimator and both improve over the baseline, while *HNDiff (Latent)* achieves the best performance with only minimal additional FLOPs overhead, confirming its advantage in content fidelity. Fig. 7 further shows that *HNDiff (Latent)* produces results closer to the ground truth, supporting our choice of the latent formulation in the main experiments.

Analysis of Diffusion Step Setting. Fig. 8 analyzes the impact of varying diffusion steps T across three restoration backbones on the NH-HAZE and O-HAZE datasets. Without diffusion guidance ($T = 0$), performance is limited for all backbones. Increasing T improves the results, with FocalNet and SGDN peaking at $T = 4$, and ConvIR peaking at $T = 6$. Further increasing T brings no additional gains. These results show that our method achieves strong performance with only a few diffusion steps.

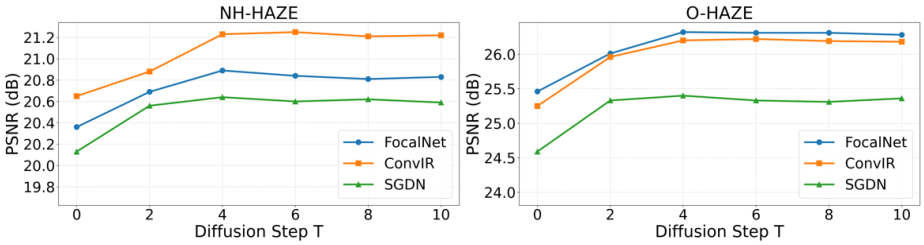

Fig. 8: Effect of the diffusion step T across three dehazing backbones on NH-HAZE and O-HAZE.

5 Limitations

HNDiff is tailored to the Atmospheric Scattering Model and thus cannot be directly applied to other degradations such as motion blur, raindrops, or low-light conditions. Extending it to these scenarios requires integrating degradation-specific priors (e.g., object motion, rain masks, exposure time), which we leave as future work.

6 Conclusion

We propose Haze-Noise Diffusion (HNDiff), a novel diffusion-based framework for image dehazing. HNDiff integrates the atmospheric scattering model into the diffusion framework, jointly performing haze diffusion and noise diffusion to account for the physical properties of haze formation. In the forward process, HNDiff progressively degrades a clean image by introducing both haze and noise through a haze-noise diffusion, with a haze-aware noise scheduler that adaptively adjusts noise levels according to haze density. In the reverse process, HNDiff restores the image by removing both haze and noise through its dehazing-denoising process. To enhance the existing dehazing methods, we incorporate HNDiff into their latent spaces as a prior generator, seamlessly integrating the learned prior into each encoder/decoder block via our proposed Feature Gating Module to generate higher-quality dehazed results. Extensive experimental results have demonstrated that our method effectively improves the performance of four state-of-the-art dehazing models across seven dehazing datasets.

Acknowledgments

This work was supported in part by the National Science and Technology Council (NSTC) under grants 114-2221-E-A49-038-MY3, 115-2634-F-A49-011, 113-2221-E-004-001-MY3, 114-2221-E-007-065-MY3 and by the NVIDIA Taiwan AI Research & Development Center (TRDC). This work was supported by H100 GPU computing resources donated by Wistron.

References

1. Ancuti, C.O., Ancuti, C., Sbert, M., Timofte, R.: Dense-haze: A benchmark for image dehazing with dense-haze and haze-free images. In: ICIP (2019)
2. Ancuti, C.O., Ancuti, C., Timofte, R., De Vleeschouwer, C.: O-haze: a dehazing benchmark with real hazy and haze-free outdoor images. In: CVPRW (2018)
3. Ancuti, C.O., Ancuti, C., Vasluianu, F.A., Timofte, R.: Ntire 2021 nonhomogeneous dehazing challenge report. In: CVPR (2021)
4. Bai, H., Pan, J., Xiang, X., Tang, J.: Self-guided image dehazing using progressive feature fusion. TIP (2022)
5. Benigmim, Y., Roy, S., Essid, S., Kalogeiton, V., Lathuilière, S.: Collaborating foundation models for domain generalized semantic segmentation. In: CVPR (2024)
6. Chen, Z., Wang, Y., Yang, Y., Liu, D.: Psd: Principled synthetic-to-real dehazing guided by physical priors. In: CVPR (2021)
7. Chen, Z., Zhang, Y., Liu, D., Gu, J., Kong, L., Yuan, X., et al.: Hierarchical integration diffusion model for realistic image deblurring. In: NeurIPS (2023)
8. Cui, Y., Ren, W., Cao, X., Knoll, A.: Focal network for image restoration. In: ICCV (2023)
9. Cui, Y., Ren, W., Cao, X., Knoll, A.: Revitalizing convolutional network for image restoration. TPAMI (2024)
10. Dong, H., Pan, J., Xiang, L., Hu, Z., Zhang, X., Wang, F., Yang, M.H.: Multi-scale boosted dehazing network with dense feature fusion. In: CVPR (2020)
11. Fang, C., He, C., Xiao, F., Zhang, Y., Tang, L., Zhang, Y., Li, K., Li, X.: Real-world image dehazing with coherence-based pseudo labeling and cooperative unfolding network. NeurIPS (2024)
12. Fang, W., Fan, J., Zheng, Y., Weng, J., Tai, Y., Li, J.: Guided real image dehazing using ycbcr color space. In: AAAI (2025)
13. Garber, T., Tirer, T.: Image restoration by denoising diffusion models with iteratively preconditioned guidance. In: CVPR (2024)
14. Guo, C.L., Yan, Q., Anwar, S., Cong, R., Ren, W., Li, C.: Image dehazing transformer with transmission-aware 3d position embedding. In: CVPR (2022)
15. He, C., Fang, C., Zhang, Y., Tang, L., Huang, J., Li, K., Guo, Z., Li, X., Farsiu, S.: Reti-diff: Illumination degradation image restoration with retinex-based latent diffusion model. In: ICLR (2025)
16. He, J.T., Tsai, F.J., Peng, Y.T., Chen, M.H., Lin, C.W., Lin, Y.Y.: Blurdm: A blur diffusion model for image deblurring. NeurIPS (2025)
17. He, J.T., Tsai, F.J., Wu, J.H., Peng, Y.T., Tsai, C.C., Lin, C.W., Lin, Y.Y.: Domain-adaptive video deblurring via test-time blurring. In: ECCV (2024)
18. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. NeurIPS (2020)
19. Huang, S.Y., Tsai, F.J., Lin, C.W., Lin, Y.Y.: Wtnet: A weather transfer network for domain-adaptive all-in-one adverse weather image restoration. In: BMVC (2025)
20. Kim, J., Cho, H., Kim, J., Tiruneh, Y.Y., Baek, S.: Sddgr: Stable diffusion-based deep generative replay for class incremental object detection. In: CVPR (2024)
21. Kim, M., Su, Y., Liu, F., Jain, A., Liu, X.: Keypoint relative position encoding for face recognition. In: CVPR (2024)
22. Li, B., Ren, W., Fu, D., Tao, D., Feng, D., Zeng, W., Wang, Z.: Benchmarking single-image dehazing and beyond. TIP (2018)

23. Li, B., Zhao, H., Wang, W., Hu, P., Gou, Y., Peng, X.: Mair: A locality-and continuity-preserving mamba for image restoration. In: CVPR (2025)
24. Li, G., Rao, C., Mo, J., Zhang, Z., Xing, W., Zhao, L.: Rethinking diffusion model for multi-contrast mri super-resolution. In: CVPR (2024)
25. Liu, C., Qi, L., Pan, J., Qian, X., Yang, M.H.: Frequency domain-based diffusion model for unpaired image dehazing. In: ICCV (2025)
26. Liu, J., Wang, Q., Fan, H., Wang, Y., Tang, Y., Qu, L.: Residual denoising diffusion models. In: CVPR (2024)
27. Liu, Y., Ke, Z., Liu, F., Zhao, N., Lau, R.W.: Diff-plugin: Revitalizing details for diffusion-based low-level tasks. In: CVPR (2024)
28. Luo, W., Qin, H., Chen, Z., Wang, L., Zheng, D., Li, Y., Liu, Y., Li, B., Hu, W.: Visual-instructed degradation diffusion for all-in-one image restoration. In: CVPR (2025)
29. Meng, C., He, Y., Song, Y., Song, J., Wu, J., Zhu, J.Y., Ermon, S.: Sdedit: Guided image synthesis and editing with stochastic differential equations. In: International Conference on Learning Representations (2022)
30. Mi, Y., Zhong, Z., Huang, Y., Ji, J., Xu, J., Wang, J., Wang, S., Ding, S., Zhou, S.: Privacy-preserving face recognition using trainable feature subtraction. In: CVPR (2024)
31. Narasimhan, S.G., Nayar, S.K.: Contrast restoration of weather degraded images. TPAMI (2003)
32. Qin, X., Wang, Z., Bai, Y., Xie, X., Jia, H.: Ffa-net: Feature fusion attention network for single image dehazing. In: AAAI (2020)
33. Qiu, Y., Zhang, K., Wang, C., Luo, W., Li, H., Jin, Z.: Mb-taylorformer: Multi-branch efficient transformer expanded by taylor formula for image dehazing. In: ICCV (2023)
34. Rajagopalan, S., Nair, N.G., Paranjape, J.N., Patel, V.M.: Gendeg: Diffusion-based degradation synthesis for generalizable all-in-one image restoration. In: CVPR (2025)
35. Rao, C., Li, G., Lan, Z., Sun, J., Luan, J., Xing, W., Zhao, L., Lin, H., Dong, J., Zhang, D.: Rethinking video deblurring with wavelet-aware dynamic transformer and diffusion model. In: ECCV (2024)
36. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B.: High-resolution image synthesis with latent diffusion models. In: CVPR (2022)
37. Shao, Y., Li, L., Ren, W., Gao, C., Sang, N.: Domain adaptation for image dehazing. In: CVPR (2020)
38. Shin, J., Chung, S., Yang, Y., Kim, T.H.: Hazeflow: Revisit haze physical model as ode and non-homogeneous haze generation for real-world dehazing. In: ICCV (2025)
39. Song, J., Meng, C., Ermon, S.: Denoising diffusion implicit models. In: ICLR (2021)
40. Song, Y., He, Z., Qian, H., Du, X.: Vision transformers for single image dehazing. TIP (2023)
41. Tan, M., Le, Q.: Efficientnet: Rethinking model scaling for convolutional neural networks. In: ICML (2019)
42. Tsai, F.J., Peng, Y.T., Lin, Y.Y., Lin, C.W.: Phatnet: A physics-guided haze transfer network for domain-adaptive real-world image dehazing. In: ICCV (2025)
43. Valanarasu, J.M.J., Yasarla, R., Patel, V.M.: Transweather: Transformer-based restoration of images degraded by adverse weather conditions. In: CVPR (2022)
44. Wang, J., Chen, Y., Zheng, Z., Li, X., Cheng, M.M., Hou, Q.: Crosskd: Cross-head knowledge distillation for object detection. In: CVPR (2024)

45. Wang, J., Wu, S., Yuan, Z., Tong, Q., Xu, K.: Frequency compensated diffusion model for real-scene dehazing. Neural Networks (2024)
46. Wang, R., Zheng, Y., Zhang, Z., Li, C., Liu, S., Zhai, G., Liu, X.: Learning hazing to dehazing: Towards realistic haze generation for real-world image dehazing. In: CVPR (2025)
47. Weber, S., Zöngür, B., Araslanov, N., Cremers, D.: Flattening the parent bias: Hierarchical semantic segmentation in the poincare ball. In: CVPR (2024)
48. Wu, H., Qu, Y., Lin, S., Zhou, J., Qiao, R., Zhang, Z., Xie, Y., Ma, L.: Contrastive learning for compact single image dehazing. In: CVPR (2021)
49. Wu, J.H., Tsai, F.J., Peng, Y.T., Tsai, C.C., Lin, C.W., Lin, Y.Y.: Id-blau: Image deblurring by implicit diffusion-based reblurring augmentation. In: CVPR (2024)
50. Wu, R.Q., Duan, Z.P., Guo, C.L., Chai, Z., Li, C.: Ridcp: Revitalizing real image dehazing via high-quality codebook priors. In: CVPR (2023)
51. Xia, B., Zhang, Y., Wang, S., Wang, Y., Wu, X., Tian, Y., Yang, W., Van Gool, L.: Diffir: Efficient diffusion model for image restoration. In: ICCV (2023)
52. Yang, Y., Wang, C., Liu, R., Zhang, L., Guo, X., Tao, D.: Self-augmented unpaired image dehazing via density and depth decomposition. In: CVPR (2022)
53. Yang, Z., Yu, H., Li, B., Zhang, J., Huang, J., Zhao, F.: Unleashing the potential of the semantic latent space in diffusion models for image dehazing. In: ECCV (2024)
54. Ye, T., Chen, S., Chai, W., Xing, Z., Qin, J., Lin, G., Zhu, L.: Learning diffusion texture priors for image restoration. In: CVPR (2024)
55. Zhang, L., Rao, A., Agrawala, M.: Adding conditional control to text-to-image diffusion models. In: ICCV (2023)
56. Zhang, Y., Zhang, J., Li, H., Wang, Z., Hou, L., Zou, D., Bian, L.: Diffusion-based blind text image super-resolution. In: CVPR (2024)
57. Zheng, D., Wu, X.M., Yang, S., Zhang, J., Hu, J.F., Zheng, W.S.: Selective hourglass mapping for universal image restoration based on diffusion model. In: CVPR (2024)
58. Zheng, Z., Wu, C.: U-shaped vision mamba for single image dehazing. arXiv preprint arXiv:2402.04139 (2024)
59. Zhou, S., Xie, X., Fan, B., Tian, J.: Physics-guided image dehazing diffusion. arXiv preprint arXiv:2504.21385 (2025)