Up-Restorer: When Unrolling Meets Prompts for Unified Image Restoration

Minghao Liu[1], **Wenhan Yang**[2], **Jinyi Luo**[1], **Jiaying Liu**[1*]

[1]Wangxuan Institute of Computer Technology, Peking University
[2]Peng Cheng Laboratory
lmhlmh@stu.pku.edu.cn, yangwh@pcl.ac.cn, luojinyi@stu.pku.edu.cn, liujiaying@pku.edu.cn

Abstract

All-in-one restoration needs to implicitly distinguish between different degradation conditions and apply specific prior constraints accordingly. To fulfill this goal, our work makes the first effort to create an all-in-one restoration via unrolling from the typical maximum a-posterior optimization function. This unrolling framework naturally leads to the construction of progressively solving models, which are equivalent to a diffusion enhancer taking as input dynamically generated prompts. Under a score-based diffusion model, the prompts are integrated for propagating and updating several context-related variables, *i.e.* transmission map, atmospheric light map, and noise or rain map progressively. Such a learned prompt generation process, which simulates the nonlinear operations in the unrolled solution, is combined with linear operations owning clear physics implications to make the diffusion models well-regularized and more effective in learning degradation-related visual priors. Experimental results demonstrate that our method achieves significant performance improvements across various image restoration tasks, realizing true all-in-one image restoration.

Introduction

Image restoration (IR) aims to reconstruct a high-quality (HQ) image from its degraded low-quality (LQ) version. It has a wide range of applications in various fields, including photography, surveillance analysis, medical imaging, forensics, *etc*. The goal of IR is to recover the original scene or object in its best possible quality, despite the presence of various forms of degradation. Traditional IR methods relied heavily on mathematical models and hand-crafted features, *e.g.* Wiener filtering, Total Variation (TV) regularization, and Non-Local Means (NLM) commonly used for IR tasks like denoising, deblurring, and inpainting. These methods are effective to a certain extent but often struggle with complex degradations, as they lack the capacity to model the intrinsic details of natural images.

With the rise of deep learning, image restoration witnessed a significant change. Recent deep learning-based IR approaches excel in addressing single kind of degradation, such as denoising (Zhang et al. 2017), deblurring (Ruan et al. 2022), adverse weather removal (Özdenizci and Legenstein 2023), and low-light enhancement (Guo et al. 2020). These task-specific methods have demonstrated impressive performance on individual degradation types.

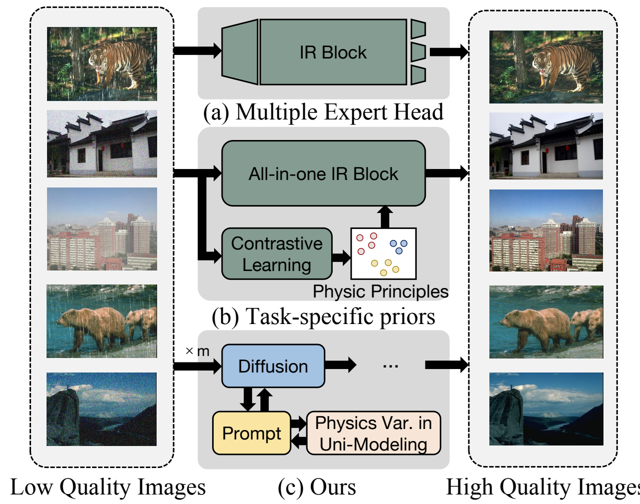

Figure 1: Comparison of existing all-in-one image restoration methods. (a) Multiple Expert Heads: Utilizes separate expert heads for different tasks, requiring task-specific training and storage. (b) Task-specific Priors: Employs a unified IR block with physical principles by contrastive learning to handle multiple tasks, but is limited by the need for multi-stage training. (c) Ours: Proposing a diffusion model guided by prompt blocks and physics variables derived from a unified degradation model to upgrade prompts, offering a more generalized and adaptable solution across various degradation scenarios.

However, real-world scenarios, such as autonomous driving and outdoor surveillance, often include complex and dynamic degradation that single degradation methods fail to handle effectively (Mao et al. 2017). Such complex visual signals obtained from real scenes have driven the need for all-in-one image restoration approaches that aim to handle multiple degradations within a unified model. Existing all-in-one methods leverage techniques such as (1) contrastive learning (Li et al. 2022), (2) task-specific sub-

*Corresponding Author.

networks (Park, Lee, and Chun 2023) that train a separate subnetwork for each type of degradation and then assess the proportion of each degradation presented in an image, (3) task-specific priors (Valanarasu, Yasarla, and Patel 2022) that learn priors that are strongly correlated with specific degradations, and then used for restoration, and (4) task-agnostic priors (Liu et al. 2022) that focus on learning general priors from natural images that are not tied to any specific type of degradation. While these approaches have shown promising results, they still face significant challenges in practical applications. Specifically, their heavy reliance on specific degradation modeling formulation and the complicated multi-stage training processes limit their practical adaptability and scalability to diverse real-world scenarios.

Diffusion models (Ho, Jain, and Abbeel 2020a), known for their powerful generative capabilities, have recently gained attention for various vision tasks, including image restoration, such as deraining (Cui et al. 2022), denoising (Choi et al. 2021), super-resolution (Saharia et al. 2023), deblurring (Whang et al. 2022), as well as adverse weather removal (Özdenizci and Legenstein 2023). These models progressively recover clean images from noisy ones through a forward diffusion process, unlike traditional CNNs and transformer-based methods that directly estimate clear images in a single pass. While diffusion models show great potential in restoration tasks, they lack awareness of degradation modeling and visual prior regularization, focusing more on transitioning between distributions than capturing intrinsic representations. This limitation prevents these models from fully adapting to handling different inputs, raising the requirement of additional constraints for that.

To address the issue, we trace back to the root of the problem, the Maximum A Posteriori (MAP) principle, for image restoration. The restoration solution is constrained to combining the fidelity term and the degradation prior regularization. It can be solved with an unrolling solution progressively, *e.g.* the proximal gradient (Combettes and Wajs 2005), augmented Lagrangian (Hestenes 1969), split Bregman (Goldstein and Osher 2009), and alternating direction method of multipliers (Bartz, Campoy, and Phan 2022), which have characteristics that are similar to the progressive nature of diffusion models. This inspires us to consider whether we can explore the possibility of integrating MAP-inspired techniques with diffusion models to enhance their performance. To this end, we make efforts to combine the strengths of diffusion models with a prompt-based network, unrolling the iterative process and breaking it into simpler sub-problems for faster and more stable convergence. The prompt network serves to estimate the degradation physics prior progressively and bypasses these factors for regularizing the diffusive process. Our method, inspired by MAP formulation, aims to make diffusion models more input-aware and effective in handling complex real-world degradation.

Our contributions are as follows:

- We propose a unified degradation model that combines multiple degradation models into one system and transforms it into an optimization problem using the score-based model. Inspired by ADMM, a prompt-based diffusion network, Up-Restorer is built to simulate and unroll an iterative optimization, breaking it into simpler sub-problems.
- We introduce a prompt-based optimization network that generates prompts and variables from the ADMM-derived equation to guide the diffusive restoration process. The prompts are applied to predict and bypass these variables during the process, dynamically adjusting each image's degradation characteristics for accurate and effective restorations.
- We integrate learnable query parameters into the Diffusion Transformer, allowing the model to adapt dynamically to varying degradations, which further improves flexibility and performance across diverse degradation scenarios.

Related Work

Multiple Degradations Image Restoration

Many image restoration methods focus on specific tasks using convolutional neural networks (Zhang and Patel 2018; Li et al. 2018b) or vision transformers (Zhang et al. 2020; Zamir et al. 2022), but they struggle to generalize across different degradation types. To overcome this, multi-task and all-in-one methods have been developed to handle diverse scenarios. Multi-task methods solve multiple tasks within a single model using separate modules, such as task-specific heads (Chen et al. 2021) or feature extractors (Li, Tan, and Cheong 2020). However, they may struggle with unknown degradations. All-in-one methods (Li et al. 2022; Potlapalli et al. 2023) offer a more flexible approach by handling various degradations within a single model without prior knowledge. These methods use techniques like contrastive learning (Li et al. 2022) or dynamic prompts (Potlapalli et al. 2023) to adapt to different conditions. Unlike previous methods, we integrate diffusion models with prompt-based mechanisms in a unified framework. We introduce a unified degradation model and simulate the ADMM optimization process with a prompt-based unrolling network, allowing dynamic adjustment to degradation characteristics. Learnable queries in the Diffusion Transformer further enhance flexibility and adaptability.

Prompt Learning for Image Restoration

Prompt learning, initially developed for incorporating additional text inputs into pre-trained large language models to influence output generation (Brown et al. 2020), has since evolved into a broader technique used in model training and fine-tuning (Wang et al. 2023b). In the field of image restoration (IR), approaches like ProRes (Ma et al. 2023) and PromptGIP (Liu et al. 2023) utilize additional images or image pairs as prompts to inform the model about the specific IR task. These methods represent explicit prompt learning.

However, real-world IR tasks often involve images for which the specific degradation type is unclear. This has led to the development of techniques that adaptively extract prompts from the input image itself (Li et al. 2022). PromptIR (Potlapalli et al. 2023) employs a classifier-based architecture to identify degradation details within images, though this approach still depends on additional degradation context, making it akin to explicit prompt learning.

In our work, we incorporate prompts into the ADMM-based iterative optimization framework, where prompts dynamically adjust parameters like the transmission map and noise map during each iteration. These prompts are generated adaptively based on the input features, guiding the restoration process in a unified manner across different degradation scenarios.

Unified Prompt-based Unrolling Restoration

Existing all-in-one image restoration methods have several neglected issues: 1) they fail to consider the commonalities and differences among various degradation types, resulting in limited performance; 2) existing methods are inefficient due to the usage of multiple expert heads, contrastive training, or complex multiple training stages. To avoid these shortcomings, we propose a unified image restoration modeling: 1) first to achieve all-in-one restoration by unrolling from the maximum a-posteriori optimization function, which excels at conditional modeling with prompts, handling various types of degradation better. 2) introducing learnable queries in the diffusion transformer with prompt-based modules to apply the physics prior constraint in the progressive optimization process.

Our method is rooted in MAP estimation and follows an ADMM-solver-based design, integrating the advantages of both data-driven and physics-based approaches. It holds even greater value in addressing universal restoration.

Preliminary: Score-based Diffusion Model

The score-based diffusion model (Ho, Jain, and Abbeel 2020b; Song et al. 2021b) is a generative model that adds noise to data through a diffusion process and learns to reverse this process to generate samples (Anderson 1982). Given a dataset with n-dimensional i.i.d. samples from an unknown distribution $p(\mathbf{x_0})$, the forward process is described by the stochastic differential equation (SDE):

$$\mathrm{d}\mathbf{x}_t = \mathbf{f}\left(\mathbf{x}_t, t\right)\mathrm{d}t + g_t \mathrm{d}\mathbf{w}_t, \tag{1}$$

where $\mathbf{f}$ is the drift coefficient, g_t the scalar diffusion coefficient, and $\mathbf{w}_t$ standard Brownian motion. As t progresses from 0 to T, $p(\mathbf{x}_0)$ evolves into $p(\mathbf{x}_T)$, approximating a standard Gaussian distribution $p_{\text{prior}}(\mathbf{x})$. The reverse-time SDE (Anderson 1982) that maps $p(\mathbf{x}_T)$ back to $p(\mathbf{x}_0)$ is:

$$\mathrm{d}\mathbf{x}_t = \left[\mathbf{f}\left(\mathbf{x}_t, t\right) - g_t^2 \nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t)\right]\mathrm{d}t + g_t \mathrm{d}\mathbf{w}_t. \tag{2}$$

To reverse the diffusion process, starting from $p(\mathbf{x}_T)$, the score function $\nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t)$ is parameterized as $\mathbf{s}_\theta(\mathbf{x}_t, t)$. The model is trained using conditional score matching (Vincent 2011), minimizing the loss:

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}\int_0^T E_{\mathbf{x}_t}\left[\lambda(t)\left\|\nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t) - \mathbf{s}_\theta(\mathbf{x}_t, t)\right\|^2\right]\mathrm{d}t \\ &\approx \frac{1}{2}\int_0^T E_{\mathbf{x}_0,\mathbf{x}_t}\left[\lambda(t)\left\|\nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t|\mathbf{x}_0) - \mathbf{s}_\theta(\mathbf{x}_t, t)\right\|^2\right]\mathrm{d}t,\end{aligned}$$

where $\lambda(t)$ is often set to g_t^2 for an optimal upper bound on the negative log-likelihood (Song et al. 2021a). The second line is used when the conditional probability $p(\mathbf{x}_t \mid \mathbf{x}_0)$ is available. To generate samples, we start with $\mathbf{x}_T$ from $p(\mathbf{x}_T) \approx p_{\text{prior}}(\mathbf{x})$ and solve Equation 2 iteratively to obtain $\mathbf{x}_0$.

The score-based model and diffusion model essentially perform similar tasks, with the score function $\mathbf{s}_\theta(\mathbf{x}_t, t)$ corresponding to the noise predicted at each time step in the diffusion model.

Unified Modeling Guided by Progressive Solution

To address the limitations of existing image restoration methods, we propose a unified degradation modeling approach. The degradation model is formulated as follows:

$$I = J \cdot t + A \cdot (1-t) + N, \tag{3}$$

where I is the degraded image, J is the clean image, t is the transmission map, A is the atmospheric light map, and N is the noise or rain map. This unified model can handle various types of degradation, including rain, haze, and noise. If N is zero, the model reduces to the dehazing problem:

$$I = J \cdot t + A \cdot (1-t), \tag{4}$$

and when $t = 1$, it reduces to a simple denoising and deraining model:

$$I = J + N. \tag{5}$$

Assume the clean image J and the degraded image I follow the model described above. The goal is to maximize the posterior probability $p(J|I)$, which can be achieved by minimizing the negative log-posterior:

$$\arg\min_J \left\{-\log p(I|J) - \log p(J)\right\}. \tag{6}$$

Here, $p(I|J)$ is the likelihood function and $p(J)$ is the prior distribution. Assuming Gaussian observation noise with σ as standard deviation, the likelihood function is:

$$p(I|J) = \exp\left(-\frac{1}{2\sigma^2}\|I - J \cdot t - A \cdot (1-t) - N\|^2\right).$$

In the score-based model, our diffusion model fits the score function $\mathbf{s}_\theta(J)$, which is the gradient of the log-data distribution $\mathbf{s}_\theta(J) \approx \nabla_J \log p(J)$. Thus, we can write:

$$\log p(J) = \int \mathbf{s}_\theta(J)\, dJ + C, \tag{7}$$

where C is an integration constant. Ignoring the constant term, the objective function becomes:

$$\arg\min_J \left\{\frac{\|I - J \cdot t - A \cdot (1-t) - N\|^2}{2\sigma^2} - \int \mathbf{s}_\theta(J)\, dJ\right\}.$$

We can use the Alternating Direction Method of Multipliers (ADMM) to solve this optimization problem because ADMM effectively handles complex optimization tasks by decomposing them into simpler subproblems, facilitating efficient convergence. Introducing an auxiliary variable U and

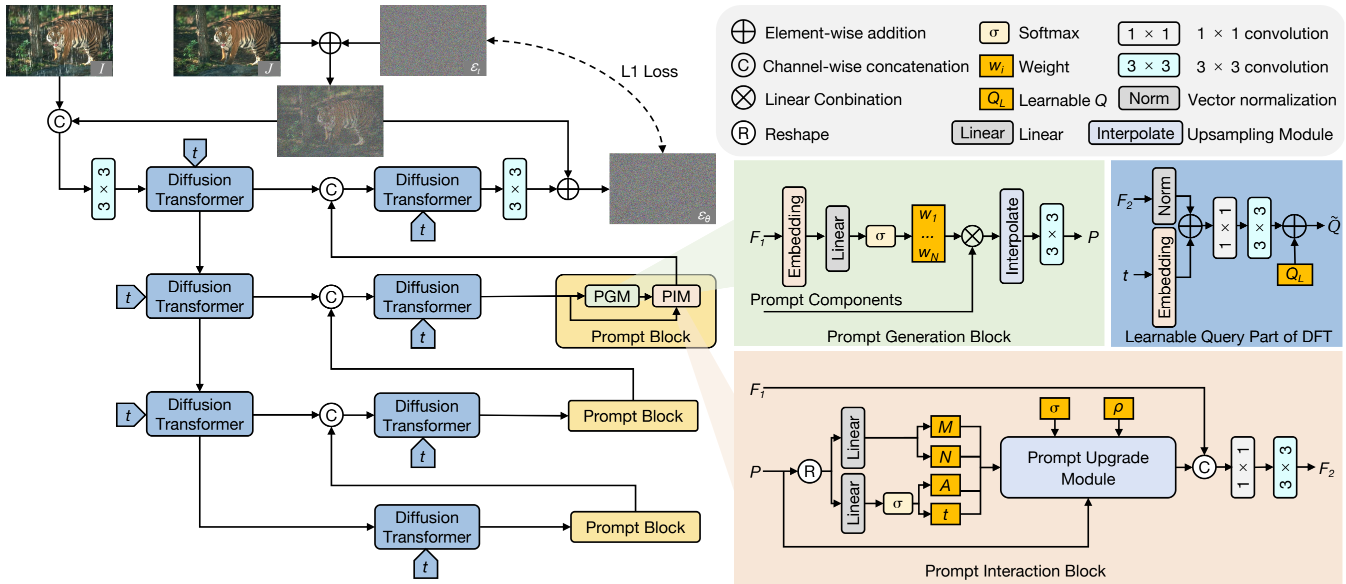

Figure 2: The framework employs a multi-level encoder-decoder structure in which Diffusion Transformers progressively process degraded images. At each level, Prompt Blocks generate and store prompts that estimate and bypass the physics degradation-related variables. These prompts, along with learnable queries in the Diffusion Transformer, are refined in the Prompt Upgrade Module like Eqn. (14) shows, which significantly boosts the model's capacity to adaptively restore images across different degradation scenarios.

Lagrange multipliers λ, the original problem is transformed into:

$$\begin{aligned} &\arg\min_{J,U}\ A(J)+B(U), \\ &A(J)=\frac{1}{2\sigma^2}\|I-J\cdot t-A\cdot(1-t)-N\|^2, \\ &B(U)=-\int \mathbf{s}_\theta(U)\,dU, \text{subject to} \quad J=U. \end{aligned} \tag{8}$$

The corresponding Lagrangian function is:

$$\begin{aligned} L(J,U,\lambda) &= \frac{1}{2\sigma^2}\|I-J\cdot t-A\cdot(1-t)-N\|^2 \\ &-\int \mathbf{s}_\theta(U)\,dU+\frac{\rho}{2}\|J-U\|^2+\lambda^T(J-U), \end{aligned} \tag{9}$$

where ρ is a penalty parameter. This optimization problem can be divided into three iterative steps, each considered as a simpler sub-optimization problem:

Step 1: Update J: Fixing U and λ, we update J by solving the following optimization problem:

$$\begin{aligned} J^{k+1} = \arg\min_J \Big\{ &\frac{1}{2\sigma^2}\|I-J\cdot t-A\cdot(1-t)-N\|^2 \\ &+\frac{\rho}{2}\|J-U^k\|^2+(\lambda^k)^T(J-U^k)\Big\}. \end{aligned}$$

This is a quadratic optimization problem in terms of J. Since it is a quadratic problem, it can be efficiently solved using standard methods for quadratic optimization. The solution is given by:

$$\begin{aligned} J^{k+1} = (\frac{t^2}{\sigma^2}+\rho)^{-1}(&\frac{t\cdot(I-A\cdot(1-t)-N)}{\sigma^2} \\ &+\rho U^k-\lambda^k) \quad \text{where} \quad \lambda^k=-\mathbf{s}_\theta(U^k). \end{aligned} \tag{10}$$

The optimization of J in this step is to refine the clean image based on the estimated physical variables A, N, t, and M. In the network, directly updating features in the latent and decoder feature space can disrupt the original diffusion model gradients. Therefore, we use prompts to perform these updates according to the iterative steps, implicitly storing key physics variables in the latent space for unified restoration, while guiding the model with these prompts.

Step 2: Update U: Fixing J and λ, we update U by solving the following optimization problem:

$$\begin{aligned} U^{k+1} = \arg\min_U \Big\{ &-\int \mathbf{s}_\theta(U)\,dU \\ &+\frac{\rho}{2}\|J^{k+1}-U\|^2+(\lambda^k)^T(J^{k+1}-U)\Big\}. \end{aligned} \tag{11}$$

We have U^{k+1} to satisfy:

$$\begin{aligned} U^{k+1} &= J^{k+1}+\frac{\mathbf{s}_\theta(U)+\lambda^k}{\rho} \\ &= J^{k+1}+\frac{\mathbf{s}_\theta(U^{k+1})-\mathbf{s}_\theta(U^k)}{\rho}. \end{aligned} \tag{12}$$

The second iteration step is to compute U^{k+1} that satisfies Eqn.(12), corresponding to the step of denoising process.

Step 3: Update Lagrange Multipliers λ: Finally, we update the Lagrange multipliers λ to enforce the constraint $J=U$. This is done by solving:

$$\lambda^{k+1}=\lambda^k+\rho(J^{k+1}-U^{k+1})=-\mathbf{s}_\theta(U^{k+1}). \tag{13}$$

The third iteration step is equal to the noise estimation network.

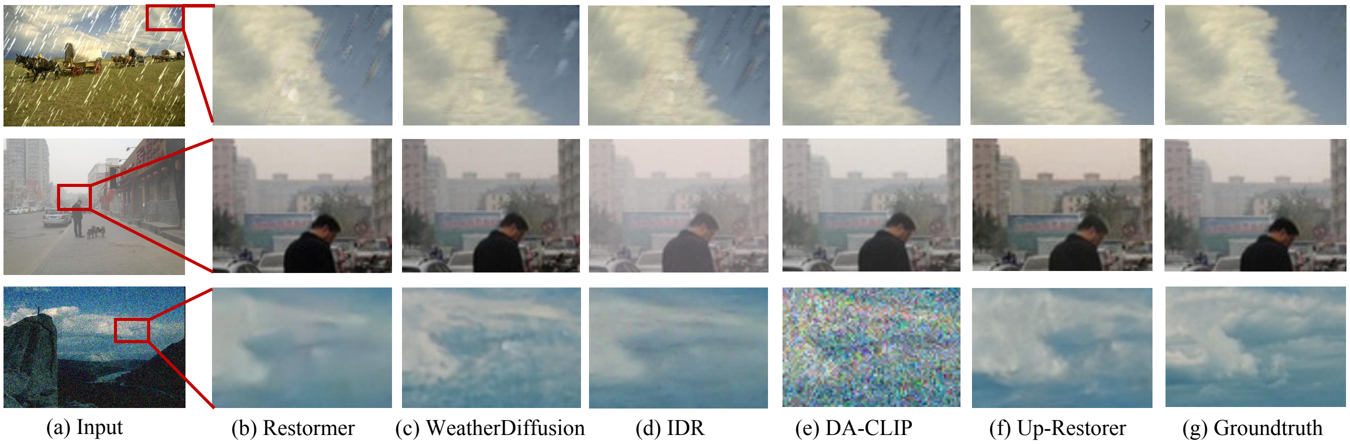

Figure 3: Visual results comparison with other all-in-one image restoration methods.

In the whole framework, each module corresponds to an ADMM step: 1. Update J is the Prompt Upgrade Module that simulates ADMM's Update J by incorporating degradation characteristics. 2. Update U is the denoising process. 3. Update λ is the noise estimation network.

It aligns with three iterative equations, with all steps being data-driven. The last two steps use the diffusion model's powerful generative capability, while the first step automatically learns degradation characteristics to optimize J, enabling Up-Restorer to learn various degradation types. The Prompt Upgrade Module is designed as follows:

Prompt-Based Unrolling Network

To simulate the iterative process derived in the previous subsection, we design a prompt-based unrolling network. This network generates prompts that dynamically adjust the parameters in our unified model during the restoration process.

The prompt generation block creates prompts based on the input features. The input features are first averaged to form an embedding vector, which is then processed through a linear layer followed by a softmax activation to generate prompt weights. These weights scale the learnable prompt parameters, creating a tailored prompt P for each input image. The prompt is then refined using a convolutional layer to match the dimensions of the input features.

In our unified model, the iterative equation Eqn. (10) is simulated using this prompt-based approach. Here, t represents the transmission map, which ranges between 0 and 1. It is obtained by processing the input through a linear layer followed by a sigmoid activation. N represents the noise map and is derived directly from another linear layer. The term $\left(\frac{t^2}{\sigma^2}+\rho\right)^{-1}$ is treated as M, which is learnable parameter map, and A is also modeled as a learnable map. At each stage of the network, the input features are used to generate the prompt, which then produces the parameters t, N, M, and A. These parameters are utilized to update the prompt P according to the iterative equation Eqn. (10) and Eqn. (13).

$$P = M\left(\frac{t\cdot(\mathbf{x}_t - A\cdot(1-t) - N)}{\sigma^2} + \rho P - \mathbf{s}_\theta(\mathbf{x}_t, t)\right). \tag{14}$$

Specifically, the generated prompt is used to adjust the feature maps by incorporating the degradation-specific information, ensuring that the model dynamically adapts to the characteristics of each input image. This adaptive mechanism allows the network to effectively handle diverse types of image degradation.

The adjusted feature maps are then concatenated with the original input features and processed through a series of convolutional layers to produce the final restored image. This prompt-based unrolling network thus effectively simulates the ADMM optimization process, providing a robust and efficient solution for unified image restoration.

Two-Stage Training with Combined Loss Functions

To optimize our diffusion model for image restoration, we follow the C2F-DFT (Wang et al. 2023a) to implement a two-stage training process based on ADMM iterative steps 2 and 3, ensuring the auxiliary variable U closely aligns with the main variable J.

Stage 1: Diffusion Model Training

In the first stage, the diffusion model is trained to minimize the L1 loss between the predicted and actual noise:

$$L_{\text{diff}} = E_{t\sim[1,T],\mathbf{x}_0,\epsilon_t} \left\|\epsilon_t - \epsilon_\theta(\mathbf{x}_t, \mathbf{y}, t)\right\|_1, \tag{15}$$

where ϵ_θ denotes the noise estimation network.

Stage 2: Image Restoration Optimization

The second stage optimizes restored images by minimizing a combination of SSIM and L1 losses:

$$L_{\text{res}} = \alpha(1-\text{SSIM}(J, \mathbf{x}_{\text{gt}})) + (1-\alpha)\left\|J - \mathbf{x}_{\text{gt}}\right\|_1, \tag{16}$$

where J is the restored image, and $\mathbf{x}_{\text{gt}}$ is the ground truth.

By iteratively updating U and J through these stages, our method effectively incorporates degradation-specific information, ensuring robust and high-quality restoration across various degradation types.

Method	Dehazing on SOTS (Li et al. 2018a)	Deraining on Rain100L (Fan et al. 2019)	Denoising on BSD68 dataset (Martin et al. 2001)) $\sigma = 15$	$\sigma = 25$	$\sigma = 50$	Average
BRDNet	23.23/0.895	27.42/0.895	32.26/0.898	29.76/0.836	26.34/0.836	27.80/0.843
LPNet	20.84/0.828	24.88/0.784	26.47/0.778	24.77/0.748	21.26/0.552	23.64/0.738
FDGAN	24.71/0.924	29.89/0.933	30.25/0.910	28.81/0.868	26.43/0.776	28.02/0.883
MPRNet	25.28/0.954	33.57/0.954	33.54/0.927	30.89/0.880	27.56/0.779	30.17/0.899
Restormer	29.30/0.976	33.23/0.948	32.95/0.917	30.91/0.885	27.55/0.785	30.79/0.902
DL	26.92/0.391	32.62/0.931	33.05/0.914	30.41/0.861	26.90/0.740	29.98/0.875
AirNet	27.94/0.962	34.90/0.967	33.92/0.933	31.26/0.888	28.00/0.797	31.20/0.910
PromptIR	<u>30.58</u>/<u>0.974</u>	<u>36.37</u>/<u>0.972</u>	<u>33.98</u>/<u>0.933</u>	<u>31.31</u>/<u>0.888</u>	<u>28.06</u>/<u>0.799</u>	<u>32.06</u>/<u>0.913</u>
WeatherDiffusion	25.54/0.971	31.53/0.961	33.26/0.926	30.51/0.871	26.75/0.745	29.32/0.895
DA-CLIP	26.83/0.962	35.85/0.972	31.43/0.889	25.05/0.605	18.33/0.308	27.50/0.747
IDR	21.05/0.856	33.53/0.947	32.90/0.911	30.38/0.866	27.03/0.765	28.98/0.869
Up-Restorer	**30.68/0.977**	**36.74/0.978**	**33.99/0.933**	**31.33/0.888**	**28.07/0.799**	**32.16/0.915**

Table 1: Comparisons under All-in-one restoration setting: single model trained on a combined set of images originating from different degradation types. When averaged across different tasks, our method provides a significant gain of 0.10 dB over the previous all-in-one method PromptIR (Potlapalli et al. 2023). The best and second-best methods are **highlighted** and <u>underlined</u>.

Dehazing Results on SOTS									
Method	DehazeNet	MSCNN	AODNet	EPDN	FDGAN	AirNet	Restormer	PromptIR	Up-Restorer
PSNR	22.46	22.06	20.29	22.57	23.15	23.18	30.87	**31.31**	<u>31.12</u>
SSIM	0.851	0.908	0.877	0.863	0.921	0.900	0.969	**0.973**	<u>0.972</u>
Deraining Results on Rain100L									
Method	DIDMDN	UMR	SIRR	MSPFN	LPNet	AirNet	Restormer	PromptIR	Up-Restorer
PSNR	23.79	32.39	32.37	33.50	33.61	34.90	36.74	<u>37.04</u>	**39.52**
SSIM	0.773	0.921	0.926	0.948	0.958	0.977	0.978	<u>0.979</u>	**0.986**

Table 2: Dehazing and Deraining results in the single-task setting on the SOTS benchmark dataset and Rain100L.

Learnable Queries in Diffusion Transformer

To enhance the flexibility and adaptability of our Diffusion Transformer (DFT), we integrate additional learnable query parameters into the self-attention mechanism. This integration allows the model to dynamically adapt to various image degradations, improving its performance in handling diverse and complex scenarios.

Overall Pipeline: Given paired clean and degraded images $\{x, y\} \in R^{H \times W \times 3}$, where $H \times W$ denotes the spatial dimensions, we first obtain the noise sample $x_t \in R^{H \times W \times 3}$ by adding Gaussian noise $\epsilon_t \sim \mathcal{N}(0, I)$ at time step t on the clean image x. The noisy image x_t is concatenated with the degraded image y to form the input $H \in R^{H \times W \times 6}$. This input is then encoded using a 3×3 convolution to obtain the embedding feature $F_0 \in R^{H \times W \times C}$. The feature F_0 is processed through the DFT blocks (DFTBs) in a hierarchical manner, with the time step t encoded into a feature T which is embedded into the DFTBs. Skip connections are used to link features at the same level in the encoder and decoder. Finally, a 3×3 convolution is used to produce the residual image, which is added to x_t to obtain the estimated noise $\hat{H} \in R^{H \times W \times 3}$.

Learnable Queries: We enhance the self-attention mechanism by adding learnable query parameters Q_L. These learnable queries are added to the original queries, allowing the model to better adapt to specific degradations:

$$\tilde{Q} = Q + Q_L, \tag{17}$$

where Q_L are the learnable queries randomly initialized. The modified attention mechanism becomes:

$$A(\tilde{Q}, K, V) = V \cdot \text{Softmax}\left(\frac{K \cdot \tilde{Q}}{\beta}\right). \tag{18}$$

where β is a learnable scaling parameter.

Integrating learnable queries into the Diffusion Transformer enables our model to dynamically encode degradation-specific information, optimizing the restoration process for effective recovery across various degradation types. This approach ensures robust generalization, leading to higher-quality restorations.

Experiments

To demonstrate the effectiveness of our proposed method, we evaluate it on three primary image restoration tasks: image dehazing, image deraining, and image denoising. Following the methodology in (Li et al. 2022) and (Potlapalli et al. 2023), we perform experiments under two distinct settings: **(a)** All-in-One, and **(b)** Single-Task. In the All-in-One setting, a unified model is trained to handle all three types of degradations. Conversely, in the Single-task setting, separate models are trained for each specific restoration task.

Designs			Denoising on BSD68 dataset (Martin et al. 2001)			Deraining on	Dehazing on
Learnable Q	Prompt Block	Physics Variables	$\sigma = 15$	$\sigma = 25$	$\sigma = 50$	Rain100L (Fan et al. 2019)	SOTS (Li et al. 2018a)
✓	✗	✗	33.57/0.923	31.08/0.882	27.68/0.779	34.72/0.959	29.23/0.970
✓	✓	✗	33.66/0.926	31.09/0.885	27.72/0.785	34.64/0.957	29.51/0.972
✗	✓	✓	33.80/0.929	31.14/0.879	27.74/0.790	35.68/0.966	29.74/0.973
✓	✓	✓	**33.99/0.933**	**31.33/0.888**	**28.07/0.799**	**36.74/0.978**	**30.68/0.977**

Table 3: Performance of the Up-Restorer network, when remove different modules.

Prompt Block Position	PSNR	SSIM
levels 1+3	35.93	0.971
levels 1+2	36.26	0.975
levels 2+3	36.55	0.976
levels 1+2+3	**36.74**	**0.978**

Table 4: Prompt blocks position results on Rain100L.

Datasets and Evaluation Metrics

Building upon the work of (Li et al. 2022), we conduct comprehensive experiments across five image restoration tasks: image denoising at noise levels $\sigma = 15, 25, 50$, image deraining, and image dehazing.

For these tasks, we use datasets: BSD400, BSD68 (Martin et al. 2001), and WED (Ma et al. 2016) for image denoising, Rain100L (Yang et al. 2020) for image deraining, and RESIDE (Li et al. 2018a) for image dehazing. Following the division in (Li et al. 2022), BSD400 and WED are used for training while BSD68 is used for testing with 68 ground truth images. For image deraining, we use the 1800 rain-clean paired images and 100 testing pairs provided in the Rain100L dataset. For image dehazing, we use the Outdoor Training Set (OTS) for training and the Synthetic Objective Testing Set (SOTS) for testing from the RESIDE dataset.

We evaluate performance using Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM). In the results tables, the best and second-best methods are **highlighted** and <u>underlined</u>, respectively.

Implementation Details

Our framework consists of a four-level encoder-decoder structure, with each level's Diffusion Transformer containing different numbers of Transformer blocks: [4, 6, 6, 8] from the first to the fourth level. We use a Prompt Block between every two consecutive decoder levels, totaling three Prompt Blocks in the network. Each Prompt Block contains five Prompt Components with a channel dimension of 1. The model is trained in a multi-degradation all-in-one setting with batch sizes of 120 and 20 for the two training phases, and a batch size of 120 in the single degradation setting. The network is optimized using the Adam optimizer (β_1=0.9, β_2=0.999), with the learning rate reduced to 0.01 after 50,000 epochs. Training the model to convergence requires only one day on a single 4090 GPU, using 128×128 cropped patches as input.

Method or Module	Parameters
DA-CLIP	48,975,747 + 246,213,634 (CLIP)
IDR	12,262,448
Restormer	26,126,644
WeatherDiffusion	85,606,147
C2F-DFT (Backbone)	25,363,344
Prompt Upgrade Module	2,148,576
Learnable Query	500,352
UpRestorer (Total)	28,012,272

Table 5: Comparison of model parameters across various methods and modules.

Model Parameter Comparison

We also compare the parameter counts of various methods and components of our model. The Prompt and Prompt Upgrade Module adds only 8% to the total number of parameters, while the learnable queries contribute an additional 2%. As a result, our final Up-Restorer model remains lightweight, with the majority of the parameters coming from the backbone C2F-DFT. Table 5 summarizes the parameter counts of our method and several recent models.

Multiple Degradation All-in-One Results

We compared the proposed Up-Restorer with several general image restoration methods (Tian, Xu, and Zuo 2020; Gao et al. 2019; Dong et al. 2020; Zamir et al. 2021, 2022) and specialized all-in-one methods (Fan et al. 2019; Li et al. 2022; Potlapalli et al. 2023; Özdenizci and Legenstein 2023; Luo et al. 2023; Zhang et al. 2023). The results are presented in Table 1. Compared to previous methods, our algorithm achieves the best overall performance when averaged across different restoration tasks. The visual examples in Figure 1 show that Up-Restorer effectively removes rain, haze, and noise from various degraded input images and benefits from the powerful generative capabilities of the diffusion model, producing clearer and more visually pleasing details than other methods. For other tasks, we also extend to deblurring, desnowing, and super-resolution tasks. For deblurring, we used the GoPro dataset; for desnowing, we selected 611 images from the Snow100K dataset; and for super-resolution, we used the RealSR dataset. The performance was evaluated using PSNR and SSIM metrics. Table 7 presents the results of our Up-Restorer model compared with several state-of-the-art methods across these tasks. Figure 4 shows the transmission map t after training, it closely matches the actual input image, indicating that the module has indeed learned

Method	Rain100L			SOTS			BSD68 ($\sigma = 50$)		
	LPIPS ↓	HFD-IQA ↑	Hyper-IQA ↑	LPIPS ↓	HFD-IQA ↑	Hyper-IQA ↑	LPIPS ↓	HFD-IQA ↑	Hyper-IQA ↑
Ours	0.0145	22.9767	56.73	0.0153	24.9012	61.36	0.1822	21.8093	50.78
DA-CLIP	0.0173	22.8600	56.33	0.0213	24.7841	60.54	0.7820	19.4388	44.23
IDR	0.0663	22.7916	55.66	0.2786	24.6320	56.38	0.2314	21.3319	46.41
Restormer	0.0635	22.6849	55.32	**0.0152**	**24.9466**	60.64	0.2393	21.0835	47.83
WeatherDiffusion	0.0293	22.7988	56.12	0.1304	24.9056	60.57	0.3208	21.5863	50.40

Table 6: Perceptual and blind image quality metrics for Deraining (Rain100L), Dehazing (SOTS), and Denoising (BSD68).

Dataset	Snow100K_S	RealSR	GoPro
Up-Restorer	32.78/0.9515	32.39/0.9047	29.00/0.8808
TransWeather	32.51/0.9341	-	-
Restormer	-	-	32.92/0.961
RCAN	-	33.87/0.922	-

Table 7: PSNR/SSIM for deblurring (GoPro), desnowing (Snow100K), and super-resolution (RealSR).

essential degradation attributes.

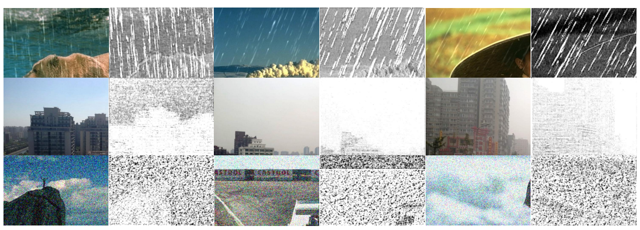

Figure 4: The transmission map t for different tasks after training. The three lines represent rain, haze, and noise degradation. The odd-numbered columns are input images, and even-numbered columns are their corresponding transmission maps in the model.

Single Degradation One-by-One Results

In this section, we evaluate the performance of Up-Restorer in a single-task setting, where separate models are trained for different restoration tasks within the original framework. This evaluation demonstrates the adaptability of the content and the network architecture's learning capability on individual tasks. Table 2 shows the results for the dehazing task compared with single image dehazing methods (Cai et al. 2016; Ren et al. 2016; Li et al. 2017; Qu et al. 2019), and the deraining task compared with some single image deraining methods (Zhang and Patel 2018; Yasarla and Patel 2019; Wei et al. 2019; Jiang et al. 2020). Our Up-Restorer achieves strong results across all tasks, particularly excelling in the deraining task, where it outperforms the PromptIR (Potlapalli et al. 2023) method by 2.48 dB, showcasing the model's robust learning capability in rain removal.

Additional Perceptual and No-Reference Metrics

To further evaluate the performance of our method, we incorporated three additional metrics: LPIPS, HFD-IQA (Wu et al. 2017), and Hyper-IQA (Su et al. 2020). LPIPS is a widely used metric that evaluates perceptual quality based on a ground truth reference, while HFD-IQA and Hyper-IQA are no-reference image quality assessment metrics. For HFD-IQA, we trained the required SVR model using degraded images from the TID2013 dataset, with their NQM metrics as features extracted by a pre-trained ResNet50 model. For Hyper-IQA, we utilize the pre-trained model on the Koniq-10k dataset, as provided by the authors.

Table 6 compares the performance of our method with recent models, including DA-CLIP, IDR, Restormer, and WeatherDiffusion, across the three perceptual metrics for tasks such as deraining (Rain100L), dehazing (SOTS), and denoising (BSD68, $\sigma = 50$). As shown in the table, our method consistently achieves superior results in terms of perceptual quality, outperforming existing models across all metrics.

Ablations Studies

In the ablation study, we assess the impact of incorporating Prompt Blocks at different levels of our model, specifically during the transition from the latent space to the final decoder. These Prompt Blocks, inserted at levels 1, 2, and 3, show the best performance when applied at all levels, as seen in Table 4, demonstrating their cumulative effect in enhancing the model's overall image restoration capability. We also examine the influence of three key components: learnable queries, Prompt Blocks, and physics variables for unified modeling. Testing on the Rain100L dataset, as shown in Table 3, indicates that all three components positively impact performance.

Conclusion

In this paper, we presented Up-Restorer, a unified image restoration framework that effectively addresses multiple degradation types. Our approach leverages the innovative use of prompt-based mechanisms to simulate iterative optimization processes within a diffusion model framework. By integrating the Alternating Direction Method of Multipliers with diffusion models, we developed a novel iterative solving process that dynamically adjusts model parameters in response to varying degradation conditions. These combined innovations result in a powerful, adaptable solution for all-in-one image restoration tasks, as demonstrated by our extensive experimental evaluations.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China under Grant 62332010, in part by the Program of Beijing Municipal Science and Technology Commission Foundation under Grant Z241100003524010, in part by Guangdong Basic and Applied Basic Research Foundation under Grant 2024A1515010454, and in part by the AI Joint Lab of Future Urban Infrastructure sponsored by Fuzhou Chengtou New Infrastructure Group.

References

Anderson, B. D. 1982. Reverse-time diffusion equation models. *Stochastic Processes and their Applications*.

Bartz, S.; Campoy, R.; and Phan, H. M. 2022. An Adaptive Alternating Direction Method of Multipliers. *Journal of Optimization Theory and Applications*.

Brown, T.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J. D.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; et al. 2020. Language models are few-shot learners. *Proceedings of Advances in Neural Information Processing Systems (NeurIPS)*.

Cai, B.; Xu, X.; Jia, K.; Qing, C.; and Tao, D. 2016. Dehazenet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing (TIP)*.

Chen, H.; Wang, Y.; Guo, T.; Xu, C.; Deng, Y.; Liu, Z.; Ma, S.; Xu, C.; Xu, C.; and Gao, W. 2021. Pre-trained image processing transformer. In *CVPR*.

Choi, J.; Kim, S.; Jeong, Y.; Gwon, Y.; and Yoon, S. 2021. ILVR: Conditioning Method for Denoising Diffusion Probabilistic Models. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*.

Combettes, P. L.; and Wajs, V. R. 2005. Signal Recovery by Proximal Forward-Backward Splitting. *Multiscale Modeling & Simulation*.

Cui, X.; Wang, C.; Ren, D.; Chen, Y.; and Zhu, P. 2022. Semi-supervised image deraining using knowledge distillation. *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*.

Dong, Y.; Liu, Y.; Zhang, H.; Chen, S.; and Qiao, Y. 2020. FD-GAN: Generative adversarial networks with fusion-discriminator for single image dehazing. In *Proceedings of the AAAI Conference on Artificial Intelligence*.

Fan, Q.; Chen, D.; Yuan, L.; Hua, G.; Yu, N.; and Chen, B. 2019. A general decoupled learning framework for parameterized image operators. *IEEE transactions on pattern analysis and machine intelligence (TPAMI)*.

Gao, H.; Tao, X.; Shen, X.; and Jia, J. 2019. Dynamic scene deblurring with parameter selective sharing and nested skip connections. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Goldstein, T.; and Osher, S. 2009. The Split Bregman Method for L1-Regularized Problems. *SIAM Journal on Imaging Sciences*.

Guo, C.; Li, C.; Guo, J.; Loy, C. C.; Hou, J.; Kwong, S.; and Cong, R. 2020. Zero-reference deep curve estimation for low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Hestenes, M. R. 1969. Multiplier and gradient methods. *Journal of Optimization Theory and Applications*.

Ho, J.; Jain, A.; and Abbeel, P. 2020a. Denoising Diffusion Probabilistic Models. In *Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS)*.

Ho, J.; Jain, A.; and Abbeel, P. 2020b. Denoising diffusion probabilistic models. *Proceedings of Advances in Neural Information Processing Systems (NeurIPS)*.

Jiang, K.; Wang, Z.; Yi, P.; Chen, C.; Huang, B.; Luo, Y.; Ma, J.; and Jiang, J. 2020. Multi-scale progressive fusion network for single image deraining. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Li, B.; Liu, X.; Hu, P.; Wu, Z.; Lv, J.; and Peng, X. 2022. All-in-one image restoration for unknown corruption. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Li, B.; Peng, X.; Wang, Z.; Xu, J.; and Feng, D. 2017. Aod-net: All-in-one dehazing network. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 4770–4778.

Li, B.; Ren, W.; Fu, D.; Tao, D.; Feng, D.; Zeng, W.; and Wang, Z. 2018a. Benchmarking single-image dehazing and beyond. *IEEE Transactions on Image Processing (TIP)*.

Li, R.; Pan, J.; Li, Z.; and Tang, J. 2018b. Single image dehazing via conditional generative adversarial network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Li, R.; Tan, R. T.; and Cheong, L.-F. 2020. All in One Bad Weather Removal Using Architectural Search. In *CVPR*.

Liu, L.; Xie, L.; Zhang, X.; Yuan, S.; Chen, X.; Zhou, W.; Li, H.; and Tian, Q. 2022. Tape: Task-agnostic prior embedding for image restoration. In *ECCV*, 447–464.

Liu, Y.; Chen, X.; Ma, X.; Wang, X.; Zhou, J.; Qiao, Y.; and Dong, C. 2023. Unifying Image Processing as Visual Prompting Question Answering. *arXiv preprint arXiv:2310.10513*.

Luo, Z.; Gustafsson, F. K.; Zhao, Z.; Sjölund, J.; and Schön, T. B. 2023. Controlling vision-language models for universal image restoration. *arXiv preprint arXiv:2310.01018*.

Ma, J.; Cheng, T.; Wang, G.; Zhang, Q.; Wang, X.; and Zhang, L. 2023. Prores: Exploring degradation-aware visual prompt for universal image restoration. *arXiv preprint arXiv:2306.13653*.

Ma, K.; Duanmu, Z.; Wu, Q.; Wang, Z.; Yong, H.; Li, H.; and Zhang, L. 2016. Waterloo exploration database: New challenges for image quality assessment models. *TIP*.

Mao, J.; Xiao, T.; Jiang, Y.; and Cao, Z. 2017. What can help pedestrian detection? In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (CVPR)*.

Martin, D.; Fowlkes, C.; Tal, D.; and Malik, J. 2001. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*.

Özdenizci, O.; and Legenstein, R. 2023. Restoring vision in adverse weather conditions with patch-based denoising diffusion models. *IEEE transactions on pattern analysis and machine intelligence (TPAMI)*.

Özdenizci, O.; and Legenstein, R. 2023. Restoring Vision in Adverse Weather Conditions With Patch-Based Denoising Diffusion Models. *IEEE transactions on pattern analysis and machine intelligence (TPAMI)*.

Park, D.; Lee, B. H.; and Chun, S. Y. 2023. All-in-One Image Restoration for Unknown Degradations Using Adaptive Discriminative Filters for Specific Degradations. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (CVPR)*.

Potlapalli, V.; Zamir, S. W.; Khan, S.; and Khan, F. S. 2023. PromptIR: Prompting for All-in-One Blind Image Restoration. *Proceedings of Advances in Neural Information Processing Systems (NeurIPS)*.

Qu, Y.; Chen, Y.; Huang, J.; and Xie, Y. 2019. Enhanced pix2pix dehazing network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 8160–8168.

Ren, W.; Liu, S.; Zhang, H.; Pan, J.; Cao, X.; and Yang, M.-H. 2016. Single image dehazing via multi-scale convolutional neural networks. In *Proceedings of the European Conference on Computer Vision (ECCV)*.

Ruan, L.; Chen, B.; Li, J.; and Lam, M. 2022. Learning to deblur using light field generated and real defocus images. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Saharia, C.; Ho, J.; Chan, W.; Salimans, T.; Fleet, D. J.; and Norouzi, M. 2023. Image Super-Resolution via Iterative Refinement. *IEEE transactions on pattern analysis and machine intelligence (TPAMI)*.

Song, Y.; Durkan, C.; Murray, I.; and Ermon, S. 2021a. Maximum likelihood training of score-based diffusion models. *Proceedings of Advances in Neural Information Processing Systems (NeurIPS)*.

Song, Y.; Sohl-Dickstein, J.; Kingma, D. P.; Kumar, A.; Ermon, S.; and Poole, B. 2021b. Score-based generative modeling through stochastic differential equations. *Proceedings of International Conference on Learning Representations (ICLR)*.

Su, S.; Yan, Q.; Zhu, Y.; Zhang, C.; Ge, X.; Sun, J.; and Zhang, Y. 2020. Blindly assess image quality in the wild guided by a self-adaptive hyper network. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR)*.

Tian, C.; Xu, Y.; and Zuo, W. 2020. Image denoising using deep CNN with batch renormalization. *Neural Networks*.

Valanarasu, J. M. J.; Yasarla, R.; and Patel, V. M. 2022. Transweather: Transformer-based restoration of images degraded by adverse weather conditions. In *CVPR*, 2353–2363.

Vincent, P. 2011. A connection between score matching and denoising autoencoders. *Neural computation*.

Wang, L.; Yang, Q.; Wang, C.; Wang, W.; Pan, J.; and Su, Z. 2023a. Learning A Coarse-to-Fine Diffusion Transformer for Image Restoration. *arXiv preprint arXiv:2308.08730*.

Wang, X.; Wang, W.; Cao, Y.; Shen, C.; and Huang, T. 2023b. Images speak in images: A generalist painter for in-context visual learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Wei, W.; Meng, D.; Zhao, Q.; Xu, Z.; and Wu, Y. 2019. Semi-supervised transfer learning for image rain removal. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Whang, J.; Delbracio, M.; Talebi, H.; Saharia, C.; Dimakis, A. G.; and Milanfar, P. 2022. Deblurring via Stochastic Refinement. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (CVPR)*.

Wu, J.; Zeng, J.; Liu, Y.; Shi, G.; and Lin, W. 2017. Hierarchical feature degradation based blind image quality assessment. In *Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCVW)*.

Yang, F.; Yang, H.; Fu, J.; Lu, H.; and Guo, B. 2020. Learning texture transformer network for image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*.

Yasarla, R.; and Patel, V. M. 2019. Uncertainty guided multi-scale residual learning-using a cycle spinning cnn for single image de-raining. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Zamir, S. W.; Arora, A.; Khan, S.; Hayat, M.; Khan, F. S.; and Yang, M.-H. 2022. Restormer: Efficient Transformer for High-Resolution Image Restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Zamir, S. W.; Arora, A.; Khan, S.; Hayat, M.; Khan, F. S.; Yang, M.-H.; and Shao, L. 2021. Multi-Stage Progressive Image Restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Zhang, D.; Zhang, H.; Tang, J.; Wang, M.; Hua, X.; and Sun, Q. 2020. Feature pyramid transformer. In *Proceedings of the European conference on computer vision (ECCV)*.

Zhang, H.; and Patel, V. M. 2018. Density-aware single image de-raining using a multi-stream dense network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Zhang, J.; Huang, J.; Yao, M.; Yang, Z.; Yu, H.; Zhou, M.; and Zhao, F. 2023. Ingredient-Oriented Multi-Degradation Learning for Image Restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Zhang, K.; Zuo, W.; Chen, Y.; Meng, D.; and Zhang, L. 2017. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *TIP*.