Degradation-Aware Feature Perturbation for All-in-One Image Restoration

Xiangpeng Tian, Xiangyu Liao, Xiao Liu, Meng Li, Chao Ren*
College of Electronics and Information Engineering, Sichuan University, China
{tianxp, liaoxiangyu1, liux, limeng_scu}@stu.scu.edu.cn, chaoren@scu.edu.cn

Abstract

All-in-one image restoration aims to recover clear images from various degradation types and levels with a unified model. Nonetheless, the significant variations among degradation types present challenges for training a universal model, often resulting in task interference, where the gradient update directions of different tasks may diverge due to shared parameters. To address this issue, motivated by the routing strategy, we propose DFPIR, a novel all-in-one image restorer that introduces Degradation-aware Feature Perturbations(DFP) to adjust the feature space to align with the unified parameter space. In this paper, the feature perturbations primarily include channel-wise perturbations and attention-wise perturbations. Specifically, channel-wise perturbations are implemented by shuffling the channels in high-dimensional space guided by degradation types, while attention-wise perturbations are achieved through selective masking in the attention space. To achieve these goals, we propose a Degradation-Guided Perturbation Block (DGPB) to implement these two functions, positioned between the encoding and decoding stages of the encoder-decoder architecture. Extensive experimental results demonstrate that DFPIR achieves state-of-the-art performance on several all-in-one image restoration tasks including image denoising, image dehazing, image deraining, motion deblurring, and low-light image enhancement. Our codes are available at https://github.com/TxpHome/DFPIR.

1. Introduction

Image restoration, a fundamental task in computer vision, has been widely studied, with a primary focus on addressing individual degradation types, such as noise, haze, rain, or blur. Recently, deep learning-based approaches have achieved remarkable progress in single-degradation restoration tasks, including denoising [15, 26, 35, 36, 42, 61], deraining [2, 17, 22, 48, 54], dehazing [5, 9, 40, 45, 51], and deblurring [7, 19, 37, 37, 60]. While these single-degradation solutions perform well under specific conditions, they often struggle to generalize effectively to multiple degradation scenarios. Consequently, recent research has shifted toward multi-degradation restoration techniques, delivering state-of-the-art performance for known combinations of degradations [3, 33, 56, 57]. However, these approaches typically require separate networks for each degradation type, leading to large model sizes and high computational demands.

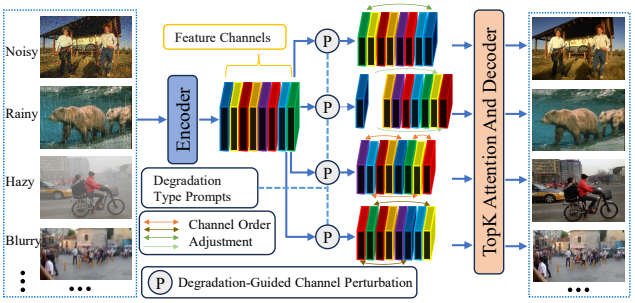

Figure 1. This figure demonstrates the channel-wise perturbation method of DFPIR, where channel shuffle assigns unique channel orders for different degradation types.

Recently, all-in-one approaches(also known as multi-degradation or multi-task image restoration) have gained prominence by addressing multiple image degradations within a unified model [6, 21, 38, 39, 49, 55, 59]. While these methods have achieved state-of-the-art results, they often overlook the relationships and distinctive characteristics among different degradation types due to shared parameters. For example, MedIR [55] demonstrates that the gradient update directions between different tasks are inconsistent or even opposite. According to the characteristics of existing methods, current all-in-one image restoration methods can be roughly divided into two categories: (1) one solution is to modify the parameter space to fit the model for different degradations [21, 38, 49, 59]; (2) and the other solution is to modify the feature space to align with the shared parameter space [8, 39, 55]. Both approaches require incorporating degradation information into the network to mitigate the interference between different degra-

*Corresponding author

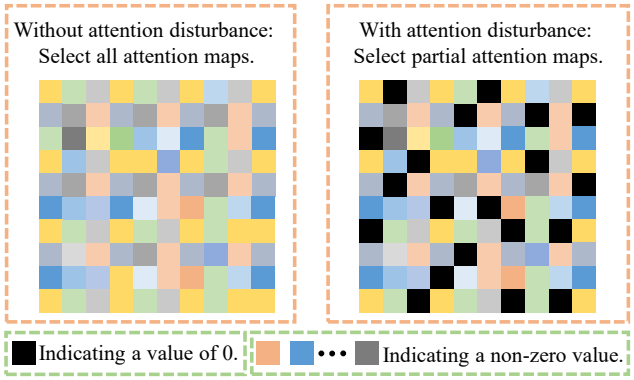

Figure 2. This figure showcases our attention-wise perturbation, where DFPIR applies attention selection to perturb image features, discarding a portion of attention for each degradation type.

dations. Although the approach (1) can effectively enhance network performance, it usually requires a large number of additional degradation parameters or a more complex network structure, increasing computational complexity.

Compared to the approach (1), the method (2) is more prevalent in the "All-in-One" image restoration framework. The approach (2) typically involves introducing a degradation type prompt, which modulates features to align with the shared network parameters. Specifically, PromptIR [39] conducts multi-degradation processing by introducing additional implicit prompts. However, although this implicit cue in the feature domain utilizes the inherent features of the image, it neglects the influence of degradation types, making it challenging to reduce the mutual influence among multiple degradation types, which ultimately leads to suboptimal results. MedIR [55] introduces multiple experts into the feature space modulation to implement a task-adaptive routing strategy. Although this "hard" routing strategy effectively reduces the influence between tasks, it may overlook the inherent features of the image among multiple degradations. InstructIR [8] introduces text commands for multi-task restoration, showcasing the potential of using text prompts to guide image restoration. However, InstructIR [8] modulates features through channel attention using text prompts, which may make it challenging to mitigate the mutual influence among different degradations.

To overcome these challenges, this paper analyzes the design philosophy and rationale behind the success of prompts and presents a degradation-aware feature perturbation approach to adjust the feature space to align with the unified parameter space for all-in-one image restoration. Our method modulates features through perturbations which include channel-wise and attention-wise perturbations to align with shared network parameters or structures, guided by degradation type prompts. Specifically, channel-wise perturbations involve shuffling the channels in high-dimensional space rather than using the conventional channel attention mechanism (see Fig.1). This shuffling approach preserves the inherent features of the image while reducing the mutual influence of degradation features. While attention-wise perturbations involve selectively masking the features after channel shuffling to achieve channel-adapted attention perturbation (see Fig.2). This approach not only preserves the inherent feature information of the image but also adaptively reduces the mutual influence between different degradation types, achieving a good balance between influence and inherent features. The main contributions of our method are as follows:

- We propose a novel degradation-aware feature perturbation all-in-one restoration framework, DFPIR, which adaptively adjusts the feature space in high-dimensional space to align with the unified parameter space, guided by degradation type prompts.
- We design a Degradation-Guided Perturbation Block (DGPB), consisting of a Degradation-Guided Channel Perturbation Module (DGCPM) and a Channel-Adapted Attention Perturbation Module (CAAPM), to apply perturbation modulation to features along the channel and attention dimensions, aligning with the encoder-decoder architecture with unified parameters.
- Extensive experiments demonstrate that our network achieves state-of-the-art performance in all-in-one image restoration. Especially for all-in-one restoration, 0.45dB PSNR improvement is obtained in comparison with InstructIR [8].

2. Related Works

Single Task Image Restoration. Single task image restoration seeks to recover a high-quality image from its low-quality version. Due to the ill-posed nature of the task, early methods primarily focused on designing effective handcrafted priors to constrain the solution space. In contrast, recent deep learning-based approaches have significantly advanced the performance of various image restoration tasks, such as denoising [13, 26, 35, 36, 42, 43], deraining [2, 22, 24, 27, 47], dehazing [5, 9, 29, 40, 45, 51], deblurring [4, 7, 18, 19, 37, 37, 60], and low-light enhancement [12, 31, 52], by learning generalizable priors from large-scale datasets. We focus on general purpose restoration models [3, 25, 57], as these architectures can be independently trained for a variety of tasks. NAFNet [3] simplifies the network structure with lightweight channel attention and gated mechanisms, offering an alternative to non-linear activations. Meanwhile, Restormer [57] leverages transformer architectures to enhance low-level restoration tasks while minimizing computational overhead. In this study, we adopt Restormer [57] as the backbone for our DFPIR model due to its efficient design and high performance across multiple restoration tasks. However, these models are primarily designed for single-degradation scenarios, limiting their ef-

fectiveness when applied directly to all-in-one restoration tasks.

All-In-One Image Restoration. Multi-task image restoration aims to address multiple tasks using the same network design [8, 21, 38, 39, 49, 55, 59]. Compared to single-task image restoration, the key challenge in multi-task image restoration is how to reduce the mutual influence of different degradation features while preserving the inherent features of the image. One approach is to adjust the network's parameter space to accommodate different degradation types. Several works have explored these strategies to handle diverse degradations. Li et al. [23] introduce a single-encoder, multi-decoder framework targeting weather-based degradations using the Rain-Haze-Snow dataset. Chen et al. [6] propose a two-stage knowledge transfer mechanism, employing a multi-teacher, single-student approach to handle various degradation types. Li et al. [21] propose an all-in-one framework capable of restoring multiple degraded images without requiring prior knowledge of degradation types or levels. Zhang et al. [59] propose an ingredient-oriented strategy that supports up to five restoration tasks within a single model, significantly enhancing scalability. Similarly, Zhang et al. [58] introduce a representation learning network guided by degradation classification, using its strong classification capabilities to effectively steer the restoration process. Another approach is to introduce image or degradation prompts to modulate the features, adapting to a unified parameter space [8, 39, 55]. PromptIR [39] encodes degradation-specific information through prompts, using them to dynamically guide the restoration network. In contrast, InstructIR [8] allows for model-driven image editing based on instructions that specify the desired actions, rather than relying on text labels, captions, or descriptions of input and output images.

3. Proposed Method

In "All-In-One" image restoration, the objective is to develop a single model $\mathbf{M}$ capable of restoring a clean image $\widehat{\mathbf{I}}$ from a degraded input image $\mathbf{I}$ affected by a degradation $\mathbf{D}$. Within this framework, under the same network parameters and architecture, different types of degradation mutually influence each other. For example, degraded images contain both inherent image features and degradation features. While inherent image features help the network learn latent parameters, thereby enhancing restoration performance, degradation features may negatively impact each other, reducing restoration effectiveness. In order to fully utilize the inherent features of the image while reducing the influence of different degradation features, motivated by the "hard" routing strategy [55] and PromptIR [39], we propose

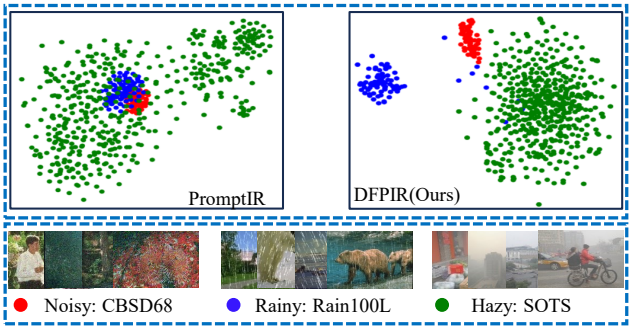

Figure 3. The figure provides t-SNE plots of the intermediate features from DFPIR (our method) and PromptIR on the test datasets (CBSD68, Rain100L, and SOTS) under the three-task setting. In our model, the features for each task exhibit tighter clustering, highlighting the effectiveness of our degradation-aware feature perturbation strategy in enhancing restoration performance.

a degradation-aware feature perturbation modulation network, DFPIR, that adjusts the feature space through perturbation modulation to align with the shared-parameter network structure, thereby enhancing multi-degradation image restoration performance. The perturbation is divided into two parts: channel-wise perturbation and attention-wise perturbation. Channel-wise perturbation is implemented through channel shuffle, while attention-wise perturbation is achieved by selecting a portion of the attention maps. The details of these two perturbation mechanisms are provided in Section 3.3. The pipeline of the proposed DFPIR framework is illustrated in Fig.4. In the following sections, we describe the overall structure of DFPIR in detail.

3.1. Overall Pipeline

Given a degraded input image $\mathbf{I} \in R^{H\times W\times 3}$, the DFPIR initially extracts shallow features $\mathbf{F_0} \in R^{H\times W\times C}$ using a 3×3 convolution layer; where $H \times W$ is the spatial size and C denotes the channels. Next, these features $\mathbf{F_0}$ undergo a 4-level encoder-decoder network, transforming into deep features $\mathbf{F_r} \in R^{H\times W\times 2C}$. Each level of the encoder-decoder employs multiple Transformer blocks [57], where the number of blocks gradually increases from the top level to the bottom level, facilitating a computationally efficient design. Beginning with the high-resolution input, the encoder aims to gradually decrease spatial resolution while increasing channel capacity, resulting in a low-resolution latent representation $\mathbf{F_e} \in R^{\frac{H}{8}\times\frac{W}{8}\times 8C}$. Given the low-resolution latent features $\mathbf{F_e}$, the objective of the decoder is to progressively restore the high-resolution clean output. To assist the recovery process, the encoder features are concatenated with the decoder features via skip connections [44]. We insert our Degradation-Guided Perturbation Block (DGPB) between the encoder and decoder, specifically within the skip-connection stage, to perturb the encoded feature space, which is guided by degradation type

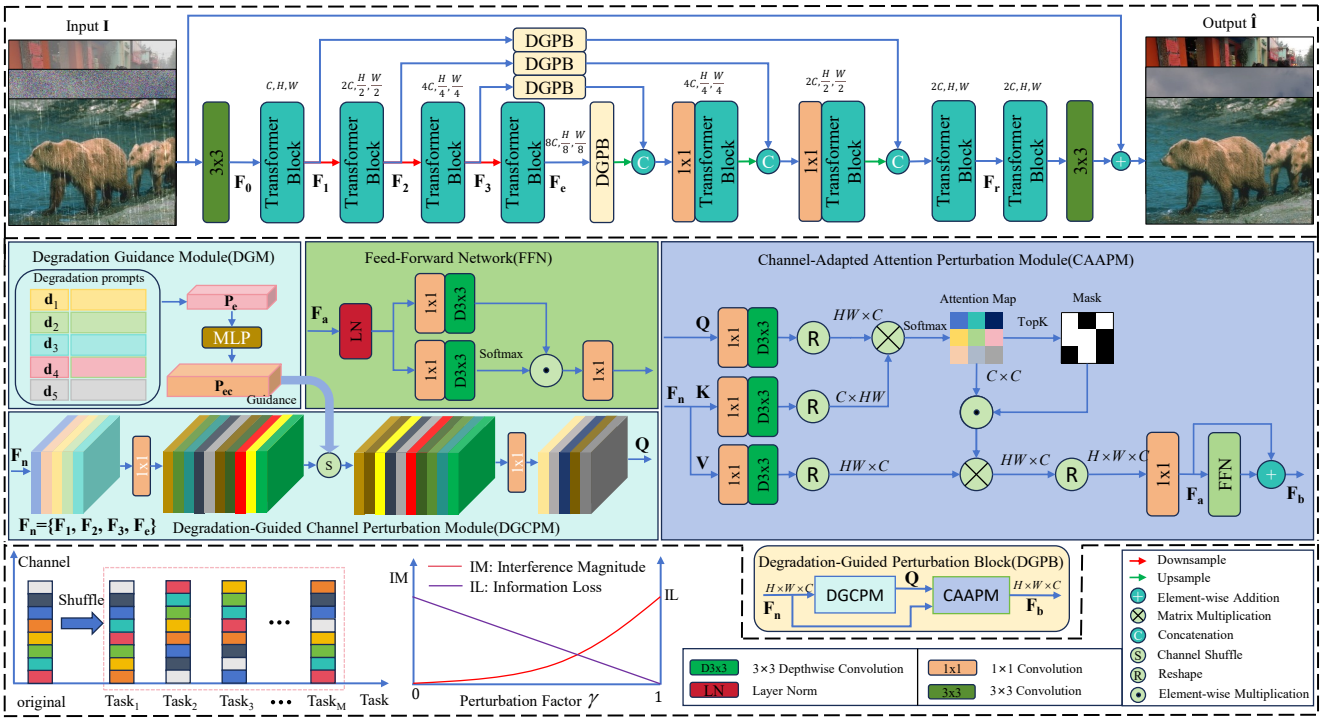

Figure 4. Overview of the DFPIR framework. We employ Restormer [57], an encoder-decoder network with transformer blocks in the encoding and decoding stages, as our backbone. The primary component of the framework, the Degradation-Guided Perturbation Block (DGPB), consists of two submodules, i.e., Degradation-Guided Channel Perturbation Module (DGCPM) and Channel-Adapted Attention Perturbation Module (CAAPM). The DGCPM module introduces conventional dimensional perturbations to image features in the form of channel shuffling, guided by degradation-type prompts. The CAAPM module applies attention perturbation to the channel-shuffled features through a top-K masking strategy.

prompts, and align it with the shared-parameter decoder. To obtain degradation type prompts, we utilize a pre-trained CLIP [41] model to encode the textual degradation type descriptions. In the following sections, we describe the proposed DGPB and its core building modules in detail.

3.2. Degradation-Guided Perturbation Block (DGPB)

In all-in-one setting, shared parameters are challenging to handle all-in-one image restoration effectively. We specifically design a Degradation-Guided Perturbation Block (DGPB) to apply perturbations to the encoded features, better aligning them with the shared decoder in multi-task scenarios, guided by the degradation type prompts (as shown in Fig.4). Given as inputs both the image feature $\mathbf{F_n} \in R^{\widehat{H}\times\widehat{W}\times\widehat{C}}$ and degradation type prompts $\mathbf{P_e} \in T^{\widehat{L}\times 1}$, the overall process of DGPB is defined as:

$$\mathbf{F_b} = \mathrm{CAAPM}(\mathrm{DGCPM}(\mathbf{F_n}, \mathrm{DGM}(\mathbf{P_e})), \mathbf{F_n}) \quad (1)$$

where $\mathbf{F_n}$ represents the encoder output. The DGPB comprises two core components: the Degradation-Guided Channel Perturbation Module (DGCPM) and the Channel-Adapted Attention Perturbation Module (CAAPM).

3.2.1 Degradation-Guided Channel Perturbation Module (DGCPM)

In DGCPM, the objective is to add perturbations to the feature channels through channel shuffling, guided by degradation type prompts. However, directly shuffling the channels on the feature $\mathbf{F_n}$ may make it difficult to converge due to excessive perturbation and may also affect the quality of the reconstruction. To address this issue, we initially expand the number of channels in the image feature $\mathbf{F_n}$ by 2x to introduce channel perturbations in a higher-dimensional channel space, guided by the Degradation Guidance Module(DGM). After channel expansion, the size of feature $\mathbf{F_n}$ becomes $\mathbf{F_{2n}} \in R^{\widehat{H}\times\widehat{W}\times 2\widehat{C}}$. In DGM, the degradation type prompts ($\mathbf{P_e}$) are used to adaptively apply channel shuffle. To maintain the dimensionality of $\mathbf{P_e}$ consistent with the channel dimensionality of $\mathbf{F_{2n}}$, we employ a Multi-Layer Perceptron (MLP) which consisting of two linear layers to achieve dimension matching. Thus, MLP transforms the input feature $\mathbf{P_e}$ into $\mathbf{P_{ec}} \in T^{2\widehat{C}\times 1}$. **For degradation type prompts, see the supplementary document.** In channel shuffle stage, we get the index value corresponding to the top-K ($\mathrm{K} = 2\widehat{\mathrm{C}}$) value of $\mathbf{P_{ec}}$, and use these index values to reorder the channels. Finally, halve the number of channels

Table 1. Comparison to state-of-the-art on three tasks. PSNR (dB, ↑) and SSIM (↑) metrics are reported on the full RGB images. On average PSNR, our DFPIR provides a significant gain of 0.45 dB over the previous all-in-one method InstructIR [8].

Method	Dehazing on SOTS	Deraining on Rain100L	Denoising on CBSD68 dataset			Average
			$\sigma = 15$	$\sigma = 25$	$\sigma = 50$	
DL [11]	26.92 / 0.391	32.62 / 0.931	33.05 / 0.914	30.41 / 0.861	26.90 / 0.740	29.98 / 0.875
FDGAN [10]	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 [56]	25.28 / 0.954	33.57 / 0.954	33.54 / 0.927	30.89 / 0.880	27.56 / 0.779	30.17 / 0.899
AirNet [21]	27.94 / 0.962	34.90 / 0.967	33.92 / 0.933	31.26 / 0.888	28.00 / 0.797	31.20 / 0.910
Restormer [57]	30.43 / 0.975	36.55 / 0.975	33.84 / 0.931	31.18 / 0.885	27.90 / 0.790	31.98 / 0.911
PromptIR [39]	30.58 / 0.974	36.37 / 0.972	33.98 / 0.933	31.31 / 0.888	28.06 / 0.799	32.06 / 0.913
InstructIR [8]	30.22 / 0.959	37.98 / 0.978	34.15 / 0.933	31.52 / 0.890	28.30 / 0.804	32.43 / 0.913
DFPIR(Ours)	31.87 / 0.980	38.65 / 0.982	34.14 / 0.935	31.47 / 0.893	28.25 / 0.806	32.88 / 0.919

after the shuffle to keep the channel count consistent before and after the transformation. Overall, the DGCPM process is summarized as:

$$\mathbf{Q} = Convh(\mathbf{CS}_{topK}(Convk(\mathbf{F_n})|\mathrm{DGM}(\mathbf{P_e}))) \quad (2)$$

where the operation $\mathbf{CS}topK\,(\bullet|\mathrm{DGM}(\mathbf{Pe}))$ denotes top-K channel shuffling guided by $\mathrm{DGM}(\mathbf{P_e})$, while $Convk\,(\bullet)$ and $Convh\,(\bullet)$ represent the operations of doubling and halving the feature channels, respectively. After DGCPM, we obtain feature $\mathbf{Q} \in R^{\widehat{H}\times\widehat{W}\times\widehat{C}}$.

3.2.2 Channel-Adapted Attention Perturbation Module(CAAPM)

Although the shuffled features can adapt well to specific degradations, directly using them for reconstruction may not yield optimal results. This is because the shuffled features carry degradation-type information but lack interaction with the original feature information. To address this, we designed the Channel-Adapted Attention Perturbation Module (CAAPM). CAAPM has two main functions: facilitating information interaction between the shuffled and original features, and adding perturbations in the attention dimension. Motivated by Restormer [57], we design a cross-attention mechanism in the channel dimension to aggregate the shuffled and original features. To add perturbations to the attention map, we introduce a mask matrix $\mathbf{M} \in R^{\widehat{C}\times\widehat{C}}$ to select part of the attention map from each row using a top-K approach with a perturbation factor parameter γ. The process of obtaining the perturbed attention map $\mathbf{PAM} \in R^{\widehat{H}\widehat{W}\times\widehat{C}}$ can be described as:

$$\mathbf{PAM(Q,K,V)} = \mathrm{softmax}(\mathbf{M} \odot (\mathbf{Q}_*\mathbf{K}_*^T \Big/ \sqrt{d_k}))\mathbf{V}_* \quad (3)$$

where the query $\mathbf{Q}_*$ is derived from shuffled feature $\mathbf{Q}$, and the key $\mathbf{K}_*$ and value $\mathbf{V}_*$ are derived from original features $\mathbf{F_n}$. $\mathbf{M}$ represents the selection mask matrix. $\odot$ represents the element-wise multiplication. Then, we use a 1×1 convolution to obtain the feature $\mathbf{F_a}$ with attention perturbations. Finally, we obtain the final output through an FFN network. This can be expressed as:

$$\mathbf{F_b} = \mathbf{F}_\mathrm{a} + \mathrm{FFN}(\mathbf{F_a}) \quad (4)$$

The selection of the perturbation factor γ needs to comprehensively consider the extent of interference magnitude between tasks and information loss. Fig.4 illustrates the relationship between γ, interference size, and information loss. In our paper, the parameter is γ fixed at 0.9, and the ablation study presents the experimental results for different values of γ.

3.3. Analysis of Our Feature Perturbation Strategy

The task-aware feature perturbation strategy we propose involves both channel and attention dimensions. By perturbing these two dimensions, it not only preserves the inherent features of the image but also reduces the mutual influence of degradation features. Fig.3 illustrates that our model excels in learning discriminative degradation contexts.

Channel-wise Perturbation Strategy. Directly performing channel shuffle in a low-dimensional space causes excessive perturbation, making network training difficult to converge. Therefore, we propose a degradation-aware adaptive channel shuffle strategy (as shown in Fig.1), which adaptively reorders the feature channels for each task in a high-dimensional feature space. Given a image feature which contains N channels $\mathbf{F_{nc}} = \{C_1, C_2, \cdots, C_r, \cdots, C_N\}\,(\mathrm{r} \leq N)$ and M tasks, our channel shuffle strategy adopts different channel orders for each degradation restoration task. For any given task $\mathrm{T_m}(\mathrm{m} \leq \mathrm{M})$, the channel index vector after channel shuffle is defined as $\mathbf{S}^N = \{S_1^N, S_2^N, \cdots, S_N^N\}$. Finally, the feature obtained after channel shuffle is $\mathbf{F_m} = \{C_{S_1^N}, C_{S_2^N}, \cdots, C_{S_N^N}\}$. The adjustment of this channel order is adaptively random under the guidance of degradation prompts. This random adaptive shuffle strategy reduces interference between feature channels while preserving the

Table 2. Comparison to state-of-the-art on five tasks. PSNR (dB, ↑) and SSIM (↑) metrics are reported on the full RGB images with (*) denoting general image restorers, others are specialized all-in-one approaches. Denoising results are reported for the noise level $\sigma = 25$.

Method	Dehazing on SOTS	Deraining on Rain100L	Denoising on CBSD68	Deblurring on Gopro	Low-light Enh. on LOL	Average
DGUNet* [33]	24.78 / 0.940	36.62 / 0.971	31.10 / 0.883	27.25 / 0.837	21.87 / 0.823	28.32 / 0.891
SwinIR* [25]	21.50 / 0.891	30.78 / 0.923	30.59 / 0.868	24.52 / 0.773	17.81 / 0.723	25.04 / 0.835
Restormer* [57]	24.09 / 0.927	34.81 / 0.962	31.49 / 0.884	27.22 / 0.829	20.41 / 0.806	27.60 / 0.881
NAFNet* [3]	25.23 / 0.939	35.56 / 0.967	31.02 / 0.883	26.53 / 0.808	20.49 / 0.809	27.76 / 0.881
DL [11]	20.54 / 0.826	21.96 / 0.762	23.09 / 0.745	19.86 / 0.672	19.83 / 0.712	21.05 / 0.743
Transweather [46]	21.32 / 0.885	29.43 / 0.905	29.00 / 0.841	25.12 / 0.757	21.21 / 0.792	25.22 / 0.836
TAPE [28]	22.16 / 0.861	29.67 / 0.904	30.18 / 0.855	24.47 / 0.763	18.97 / 0.621	25.09 / 0.801
AirNet [21]	21.04 / 0.884	32.98 / 0.951	30.91 / 0.882	24.35 / 0.781	18.18 / 0.735	25.49 / 0.846
IDR [59]	25.24 / 0.943	35.63 / 0.965	31.60 / 0.887	27.87 / 0.846	21.34 / 0.826	28.34 / 0.893
InstructIR [8]	27.10 / 0.956	36.84 / 0.973	31.40 / 0.887	29.40 / 0.886	23.00 / 0.836	29.55 / 0.907
DFPIR(Ours)	31.64 / 0.979	37.62 / 0.978	31.29 / 0.889	28.82 / 0.873	23.82 / 0.843	30.64 / 0.913

inherent characteristics of the image (as channel shuffling does not result in channel feature loss), thereby improving reconstruction quality.

Attention-wise Perturbation Strategy. The main purpose of channel shuffle is to minimize the influence of degradation features while preserving the inherent characteristics of the image. However, this reduction is not thorough enough, meaning its effectiveness is limited. To address this, we directly discard a portion of the attention in the attention maps to further mitigate the impact of multiple degradation features (as shown in Fig.2). Specifically, we compute the transposed cross attention map $\mathbf{A}_{tt} \in R^{\widehat{C}\times\widehat{C}}$. We select a portion of the attention map values from each column of $\mathbf{A}_{tt}$ using the top-K method to generate a disturbance mask matrix $\mathbf{M} \in R^{\widehat{C}\times\widehat{C}}$, where the values that are not selected are set to 0, and the remaining values are set to 1. Then, the matrix $\mathbf{M}$ is multiplied element-wise with $\mathbf{A}_{tt}$ to obtain a new attention map $\mathbf{A}_{ttn}$, achieving perturbation in the attention dimension. In other words, $\mathbf{A}_{ttn} = \mathbf{A}_{tt} \odot \mathbf{M}$, where $\mathbf{A}_{tt}$ is $\mathbf{Q}_*\mathbf{K}_*^T/\sqrt{d_k}$ in equation 3.

4. Experiments

In this section, we follow the protocols of prior state-of-the-art works [21, 39] to conduct experiments under two settings: (a) all-in-one and (b) single-task. For the all-in-one setting, a unified model is trained to handle multiple degradation types, with experiments conducted across three and five distinct degradations. In contrast, the single-task setting involves training separate models, each specialized for a specific restoration task. The image quality metrics—PSNR and SSIM [16]—for the top-performing methods are highlighted in red and the second-best results are highlighted in blue in the result tables. **Single-task results and more visual experiments are in the supplementary material.**

4.1. Experimental Settings

Datasets. In line with previous work [21, 39], we prepare datasets tailored for various restoration tasks. For single-task image denoising, we merge images from the BSD400 [1] and WED [30] datasets to train the model. Testing is conducted on the CBSD68 [32] and Urban100 [14] datasets. For image dehazing, we use the SOTS [20] dataset, while Rain100L [53] is employed for image deraining. Deblurring and low-light enhancement tasks utilize the GoPro [34] and LOL-v1 [34] datasets, respectively. In the all-in-one setting, a unified model is trained on the combined training datasets mentioned above and tested directly across multiple restoration tasks. **Additional dataset details can be found in the supplementary materials.**

Implementation Details. Our DFPIR provides an end-to-end trainable solution, removing the necessity for pre-training any individual components. Following the configuration of PromptIR [39], our DFPIR architecture features a 4-level encoder-decoder structure, comprising varying numbers of Transformer blocks at each level, specifically [4, 6, 6, 8] from level-1 to level-4. We integrate our Degradation-Guided Perturbation Block (DGPB) between the encoder and decoder, with a total of four DGPBs distributed throughout the network. We conduct our experiments using PyTorch on a single NVIDIA GeForce RTX 3090 GPU. For training, we run for 80 epochs with an initial learning rate of $1e^{-4}$, followed by fine-tuning for 5 epochs at a learning rate of $1e^{-5}$. Initially, we set the patch size to 128^2 with the batch size of 5, and for fine-tuning, the patch size is adjusted to 192^2 with a batch size of 3. The network is optimized using the L1 loss function in conjunction with the Adam optimizer (with parameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$). We train on cropped patches and augment the dataset with random horizontal and vertical flips.

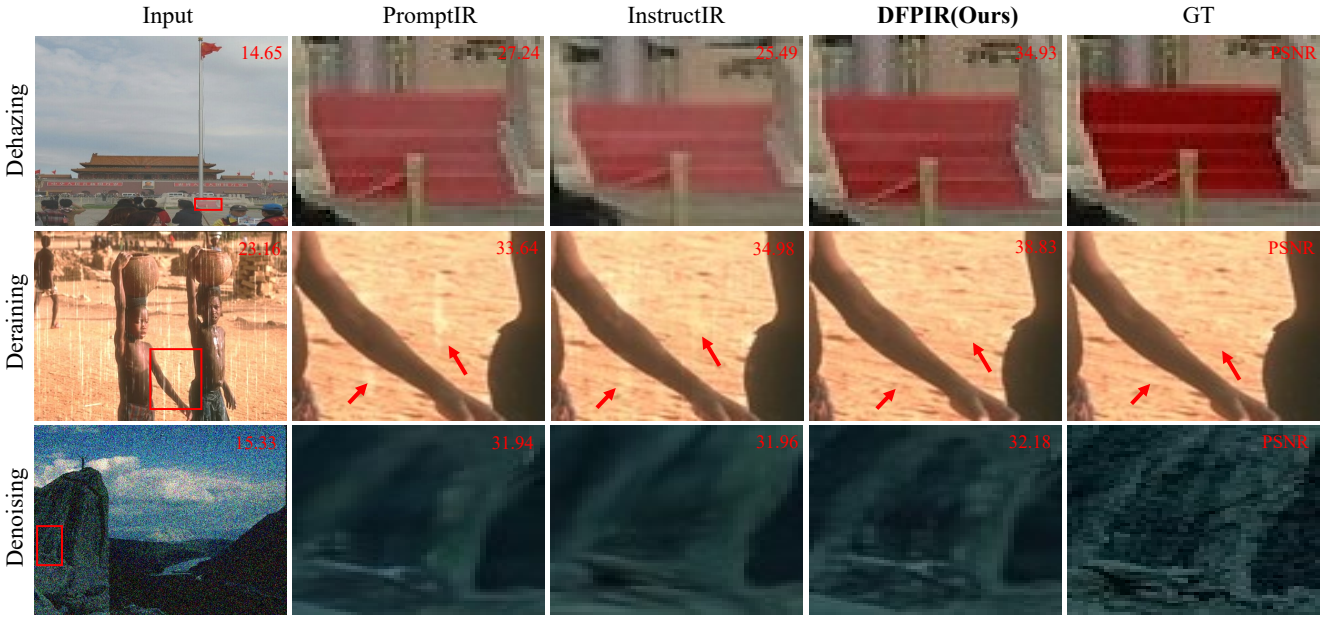

Figure 5. Visual comparison of DFPIR with state-of-the-art methods on challenging cases for the All-in-One setting considering three degradations. Zoom in for better view.

4.2. Results Comparisons on Three Tasks

We assess the performance of our all-in-one DFPIR across three distinct restoration tasks: dehazing, deraining, and denoising. Our DFPIR is compared with various general image restoration methods, including Restormer [57], FDGAN [10], and MPRNet [56], as well as specialized all-in-one approaches such as DL [11], AirNet [21], PromptIR [39], and InstructIR [8]. As shown in Tab. 1, the proposed DFPIR consistently outperforms the other competing methods. On average across the different restoration tasks, our algorithm achieves a performance improvement of 0.45 dB over the previous best method, InstructIR [8], and 0.82 dB over the second-best approach, PromptIR [39]. Specifically, DFPIR demonstrates a notable increase of 0.67 dB on the deraining task and 1.65 dB on the dehazing task compared to InstructIR [8].

4.3. Results Comparisons on Five Tasks

To further validate the method's effectiveness in addressing a broader range of tasks, building upon the recent research conducted by IDR [59] and InstructIR [8], we extend our investigation into the efficacy of DFPIR by conducting experiments across five restoration tasks: dehazing, deraining, denoising, deblurring, and low-light image enhancement. To achieve this, we train a comprehensive DFPIR on combined datasets compiled for five distinct tasks. These encompass datasets from the previously mentioned three-task scenario, alongside additional datasets: GoPro [34] for motion deblurring and LOL [50] for low-light image enhancement. Tab. 2 demonstrates that DFPIR achieves a 1.09 dB improvement compared to the recent leading method InstructIR [8], when averaged across five restoration tasks. Additionally, we compare our method to general image restoration models that were trained in the same All-in-One setting. Notably, our method surpasses Restormer[57] and NAFNet [3] on average PSNR by 3.04 dB and 2.88 dB, respectively, validating the effectiveness of our approach in handling multiple degradations.

4.4. Visual Results

Visual examples illustrating our results for dehazing, deraining, and denoising are provided in Fig.5. Compared to PromptIR [39] and InstructIR [8], our method is more effective in challenging dehazing scenarios. Additionally, our model's deraining results are closer to the ground-truth images. Furthermore, in image denoising, our model recovers more details from heavily degraded noisy inputs. We visualized the results of channel shuffling, as shown in Fig.7. After channel shuffling, the channel order for each task changed accordingly, which validates the effectiveness of our proposed channel shuffling strategy. Additionally, we also visualized the feature perturbations before and after for the $\mathbf{F}_1$ layer, as shown in Fig.6. As can be seen, after channel dimension perturbation (DGCPM), the network extracts the image's inherent features from multi-degradation scenarios, reducing the impact of degradation-specific features. With the addition of perturbation in the attention dimension (DGCPM+CAAPM), the intrinsic detail features of the image are further enhanced, while the degradation features are more effectively suppressed. This demonstrates the efficacy

Figure 6. Feature visualization. DGCPM extracts the intrinsic features of the image while suppressing degradation characteristics. DGCPM + CAAPM enhances intrinsic features while further reducing the impact of degradation. Zoom in for better view.

of our proposed channel and attention perturbation strategy.

4.5. Ablations Studies

We conduct several ablation experiments to demonstrate the effectiveness of our proposed degradation-guided perurbation block. We report the results of training an all-in-one model on combined datasets from three restoration tasks. **More detailed ablation experiments can be found in the supplementary materials.**

Impact of key components. As illustrated in the Tab. 3(a), using channel attention (Method (a)) directly improves by 0.36 dB compared to the baseline [57], but it is 0.15 dB lower than channel shuffle (Method (c)). This also validates the effectiveness of the channel shuffle strategy we proposed. Channel shuffle preserves inherent image features with degradation info but offers limited reduction in cross-degradation interference. By applying attention-wise perturbation, restoration quality is significantly enhanced (DGCPM+CAAPM). The average PSNR increased from 32.49 to 32.88, reflecting an improvement of 0.39 dB. However, Method (b) (CA+CAAPM) results in a lower performance than DFPIR, indicating that the perturbations in channel and attention dimensions produce a synergistic enhancement effect.

Impact of parameter γ. As illustrated in the Tab. 3(b), if the perturbation in the attention dimension is too high ($\gamma = 0.5$) or absent ($\gamma = 1.0$), the performance is not optimal. This is because excessive perturbation, while reducing the interference between images with different degradations, increases information loss, leading to suboptimal performance. Similarly, if the perturbation is too small, the interference between tasks becomes more significant, resulting in suboptimal performance as well.

5. Conclusion

In this paper, we present a novel all-in-one image restoration framework, DFPIR, which introduces Degradation-aware Feature Perturbations (DFP) to adjust the feature space in alignment with a unified parameter space. Our approach incorporates both channel-wise and attention-wise perturbations, dynamically guided by degradation type prompts. Specifically, channel-wise perturbations are achieved by shuffling channels in a high-dimensional space, while attention-wise perturbations are implemented through selective masking in the attention space. To effectively realize these operations, we design a Degradation-Guided Perturbation Block (DGPB), which integrates channel shuffling and attention masking, strategically placed between the encoding and decoding stages of the encoder-decoder architecture. The proposed DGPB demonstrates its efficacy in enhancing comprehensive image restoration when integrated into a state-of-the-art model, yielding notable improvements in the all-in-one restoration setting.

Table 3. Ablation study results on three tasks. Average PSNR (dB, ↑) and SSIM (↑) metrics are reported on the full RGB images.The CA stands for Channel Attention.

(a) Impact of key components.

Method	CA	DGCPM	CAAPM	PSNR/SSIM
baseline	×	×	×	31.98 / 0.911
(a)	✓	×	×	32.34 / 0.914
(b)	✓	×	✓	32.65 / 0.917
(c)	×	✓	×	32.49 / 0.910
DFPIR	×	✓	✓	32.88 / 0.919

(b) Impact of γ.

γ	PSNR / SSIM
0.5	32.67 / 0.913
0.7	32.81 / 0.919
0.8	32.84 / 0.919
0.9	32.88 / 0.919
1.0	32.83 / 0.919

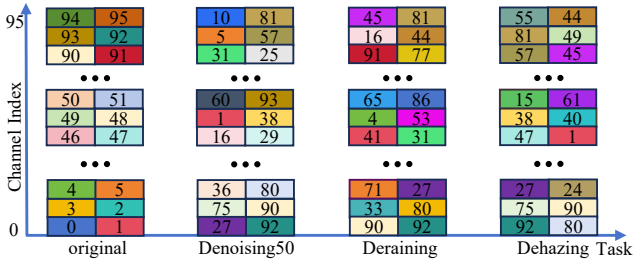

Figure 7. Channel shuffle visualization. After channel shuffling, the feature channel order for each task has changed compared to the original order.

6. Acknowledgment

This work was supported in part by the National Natural Science Foundation of China under Grant 62171304 and partly by the Natural Science Foundation of Sichuan Province under Grant 2024NSFSC1423, Cooperation Science and Technology Project of Sichuan University and Dazhou City under Grant 2022CDDZ-09, the TCL Science and Technology Innovation Fund under grant 25JZH008, and the Young Faculty Technology Innovation Capacity Enhancement Program of Sichuan University under Grant 2024SCUQJTX025.

References

[1] Pablo Arbelaez, Michael Maire, Charless Fowlkes, and Jitendra Malik. Contour detection and hierarchical image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 33(5):898–916, 2010. 6

[2] Chenghao Chen and Hao Li. Robust representation learning with feedback for single image deraining. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 7742–7751, 2021. 1, 2

[3] Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, and Jian Sun. Simple baselines for image restoration. In *European conference on computer vision*, pages 17–33. Springer, 2022. 1, 2, 6, 7

[4] Lufei Chen, Xiangpeng Tian, Shuhua Xiong, Yinjie Lei, and Chao Ren. Unsupervised blind image deblurring based on self-enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 25691–25700, 2024. 2

[5] Wei-Ting Chen, Hao-Yu Fang, Jian-Jiun Ding, and Sy-Yen Kuo. Pmhld: Patch map-based hybrid learning dehazenet for single image haze removal. *IEEE Transactions on Image Processing*, 29:6773–6788, 2020. 1, 2

[6] Wei-Ting Chen, Zhi-Kai Huang, Cheng-Che Tsai, Hao-Hsiang Yang, Jian-Jiun Ding, and Sy-Yen Kuo. Learning multiple adverse weather removal via two-stage knowledge learning and multi-contrastive regularization: Toward a unified model. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17653–17662, 2022. 1, 3

[7] Sung-Jin Cho, Seo-Won Ji, Jun-Pyo Hong, Seung-Won Jung, and Sung-Jea Ko. Rethinking coarse-to-fine approach in single image deblurring. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 4641–4650, 2021. 1, 2

[8] Marcos V Conde, Gregor Geigle, and Radu Timofte. Instructir: High-quality image restoration following human instructions. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2024. 1, 2, 3, 5, 6, 7

[9] Sourya Dipta Das and Saikat Dutta. Fast deep multi-patch hierarchical network for nonhomogeneous image dehazing. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pages 482–483, 2020. 1, 2

[10] Yu Dong, Yihao Liu, He Zhang, Shifeng Chen, and Yu Qiao. Fd-gan: Generative adversarial networks with fusion-discriminator for single image dehazing. In *Proceedings of the AAAI conference on artificial intelligence*, pages 10729–10736, 2020. 5, 7

[11] Qingnan Fan, Dongdong Chen, Lu Yuan, Gang Hua, Nenghai Yu, and Baoquan Chen. A general decoupled learning framework for parameterized image operators. *IEEE transactions on pattern analysis and machine intelligence*, 43(1): 33–47, 2019. 5, 6, 7

[12] Chunle Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, and Runmin Cong. Zero-reference deep curve estimation for low-light image enhancement. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1780–1789, 2020. 2

[13] Jie Huang, Xiao Liu, Yizhong Pan, Xiaohai He, and Chao Ren. Casapunet: Channel affine self-attention based progressively updated network for real image denoising. *IEEE Transactions on Industrial Informatics*, 2022. 2

[14] Jia-Bin Huang, Abhishek Singh, and Narendra Ahuja. Single image super-resolution from transformed self-exemplars. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5197–5206, 2015. 6

[15] Tao Huang, Songjiang Li, Xu Jia, Huchuan Lu, and Jianzhuang Liu. Neighbor2neighbor: Self-supervised denoising from single noisy images. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 14781–14790, 2021. 1

[16] Quan Huynh-Thu and Mohammed Ghanbari. Scope of validity of psnr in image/video quality assessment. *Electronics letters*, 44(13):800–801, 2008. 6

[17] Kui Jiang, Zhongyuan Wang, Peng Yi, Chen Chen, Baojin Huang, Yimin Luo, Jiayi Ma, and Junjun Jiang. Multi-scale progressive fusion network for single image deraining. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 8346–8355, 2020. 1

[18] Orest Kupyn, Volodymyr Budzan, Mykola Mykhailych, Dmytro Mishkin, and Jiří Matas. Deblurgan: Blind motion deblurring using conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 8183–8192, 2018. 2

[19] Orest Kupyn, Tetiana Martyniuk, Junru Wu, and Zhangyang Wang. Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 8878–8887, 2019. 1, 2

[20] Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang. Benchmarking single-image dehazing and beyond. *IEEE Transactions on Image Processing*, 28(1):492–505, 2018. 6

[21] Boyun Li, Xiao Liu, Peng Hu, Zhongqin Wu, Jiancheng Lv, and Xi Peng. All-in-one image restoration for unknown corruption. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17452–17462, 2022. 1, 3, 5, 6, 7

[22] Ruoteng Li, Loong-Fah Cheong, and Robby T Tan. Heavy rain image restoration: Integrating physics model and conditional adversarial learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1633–1642, 2019. 1, 2

[23] Ruoteng Li, Robby T Tan, and Loong-Fah Cheong. All in one bad weather removal using architectural search. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 3175–3185, 2020. 3

[24] Xia Li, Jianlong Wu, Zhouchen Lin, Hong Liu, and Hongbin Zha. Recurrent squeeze-and-excitation context aggregation net for single image deraining. In *Proceedings of the European conference on computer vision (ECCV)*, pages 254–269, 2018. 2

[25] Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. Swinir: Image restoration using swin transformer. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 1833–1844, 2021. 2, 6

[26] Xin Lin, Chao Ren, Xiao Liu, Jie Huang, and Yinjie Lei. Unsupervised image denoising in real-world scenarios via self-collaboration parallel generative adversarial branches. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 12642–12652, 2023. 1, 2

[27] Xin Lin, Jingtong Yue, Sixian Ding, Chao Ren, Lu Qi, and Ming-Hsuan Yang. Dual degradation representation for joint deraining and low-light enhancement in the dark. *IEEE transactions on circuits and systems for video technology*, 35(3):2461–2473, 2024. 2

[28] Lin Liu, Lingxi Xie, Xiaopeng Zhang, Shanxin Yuan, Xiangyu Chen, Wengang Zhou, Houqiang Li, and Qi Tian. Tape: Task-agnostic prior embedding for image restoration. In *European Conference on Computer Vision*, pages 447–464. Springer, 2022. 6

[29] Xiaohong Liu, Yongrui Ma, Zhihao Shi, and Jun Chen. Griddehazenet: Attention-based multi-scale network for image dehazing. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 7314–7323, 2019. 2

[30] Kede Ma, Zhengfang Duanmu, Qingbo Wu, Zhou Wang, Hongwei Yong, Hongliang Li, and Lei Zhang. Waterloo exploration database: New challenges for image quality assessment models. *IEEE Transactions on Image Processing*, 26 (2):1004–1016, 2016. 6

[31] Long Ma, Tengyu Ma, Risheng Liu, Xin Fan, and Zhongxuan Luo. Toward fast, flexible, and robust low-light image enhancement. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5637–5646, 2022. 2

[32] David Martin, Charless Fowlkes, Doron Tal, and Jitendra Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, pages 416–423. IEEE, 2001. 6

[33] Chong Mou, Qian Wang, and Jian Zhang. Deep generalized unfolding networks for image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17399–17410, 2022. 1, 6

[34] Seungjun Nah, Tae Hyun Kim, and Kyoung Mu Lee. Deep multi-scale convolutional neural network for dynamic scene deblurring. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3883–3891, 2017. 6, 7

[35] Yizhong Pan, Chao Ren, Xiaohong Wu, Jie Huang, and Xiaohai He. Real image denoising via guided residual estimation and noise correction. *IEEE Transactions on Circuits and Systems for Video Technology*, 33(4):1994–2000, 2022. 1, 2

[36] Yizhong Pan, Xiao Liu, Xiangyu Liao, Yuanzhouhan Cao, and Chao Ren. Random sub-samples generation for self-supervised real image denoising. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 12150–12159, 2023. 1, 2

[37] Dongwon Park, Dong Un Kang, Jisoo Kim, and Se Young Chun. Multi-temporal recurrent neural networks for progressive non-uniform single image deblurring with incremental temporal training. In *European Conference on Computer Vision*, pages 327–343. Springer, 2020. 1, 2

[38] Dongwon Park, Byung Hyun Lee, and Se Young Chun. All-in-one image restoration for unknown degradations using adaptive discriminative filters for specific degradations. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5815–5824. IEEE, 2023. 1, 3

[39] Vaishnav Potlapalli, Syed Waqas Zamir, Salman Khan, and Fahad Khan. Promptir: Prompting for all-in-one image restoration. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023. 1, 2, 3, 5, 6, 7

[40] Xu Qin, Zhilin Wang, Yuanchao Bai, Xiaodong Xie, and Huizhu Jia. Ffa-net: Feature fusion attention network for single image dehazing. In *Proceedings of the AAAI conference on artificial intelligence*, pages 11908–11915, 2020. 1, 2

[41] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In *International conference on machine learning*, pages 8748–8763. PMLR, 2021. 4

[42] Chao Ren, Xiaohai He, Chuncheng Wang, and Zhibo Zhao. Adaptive consistency prior based deep network for image denoising. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 8596–8606, 2021. 1, 2

[43] Chao Ren, Yizhong Pan, and Jie Huang. Enhanced latent space blind model for real image denoising via alternative optimization. *Advances in Neural Information Processing Systems*, 35:38386–38399, 2022. 2

[44] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pages 234–241. Springer, 2015. 3

[45] Yuanjie Shao, Lerenhan Li, Wenqi Ren, Changxin Gao, and Nong Sang. Domain adaptation for image dehazing. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2808–2817, 2020. 1, 2

[46] Jeya Maria Jose Valanarasu, Rajeev Yasarla, and Vishal M Patel. Transweather: Transformer-based restoration of images degraded by adverse weather conditions. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2353–2363, 2022. 6

[47] Guoqing Wang, Changming Sun, and Arcot Sowmya. Erl-net: Entangled representation learning for single image deraining. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5644–5652, 2019. 2

[48] Hong Wang, Qi Xie, Qian Zhao, and Deyu Meng. A model-driven deep neural network for single image rain removal. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 3103–3112, 2020. 1

[49] Yinglong Wang, Chao Ma, and Jianzhuang Liu. Smartassign: Learning a smart knowledge assignment strategy for deraining and desnowing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3677–3686, 2023. 1, 3

[50] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. *arXiv preprint arXiv:1808.04560*, 2018. 7

[51] Haiyan Wu, Yanyun Qu, Shaohui Lin, Jian Zhou, Ruizhi Qiao, Zhizhong Zhang, Yuan Xie, and Lizhuang Ma. Contrastive learning for compact single image dehazing. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10551–10560, 2021. 1, 2

[52] Wenhui Wu, Jian Weng, Pingping Zhang, Xu Wang, Wenhan Yang, and Jianmin Jiang. Uretinex-net: Retinex-based deep unfolding network for low-light image enhancement. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5901–5910, 2022. 2

[53] Fuzhi Yang, Huan Yang, Jianlong Fu, Hongtao Lu, and Baining Guo. Learning texture transformer network for image super-resolution. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5791–5800, 2020. 6

[54] Wenhan Yang, Robby T Tan, Shiqi Wang, Yuming Fang, and Jiaying Liu. Single image deraining: From model-based to data-driven and beyond. *IEEE Transactions on pattern analysis and machine intelligence*, 43(11):4059–4077, 2020. 1

[55] Zhiwen Yang, Haowei Chen, Ziniu Qian, Yang Yi, Hui Zhang, Dan Zhao, Bingzheng Wei, and Yan Xu. All-in-one medical image restoration via task-adaptive routing. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 67–77. Springer, 2024. 1, 2, 3

[56] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Multi-stage progressive image restoration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 14821–14831, 2021. 1, 5, 7

[57] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5728–5739, 2022. 1, 2, 3, 4, 5, 6, 7, 8

[58] Cheng Zhang, Yu Zhu, Qingsen Yan, Jinqiu Sun, and Yanning Zhang. All-in-one multi-degradation image restoration network via hierarchical degradation representation. In *Proceedings of the 31st ACM International Conference on Multimedia*, pages 2285–2293, 2023. 3

[59] Jinghao Zhang, Jie Huang, Mingde Yao, Zizheng Yang, Hu Yu, Man Zhou, and Feng Zhao. Ingredient-oriented multi-degradation learning for image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5825–5835, 2023. 1, 3, 6, 7

[60] Kaihao Zhang, Wenhan Luo, Yiran Zhong, Lin Ma, Bjorn Stenger, Wei Liu, and Hongdong Li. Deblurring by realistic blurring. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2737–2746, 2020. 1, 2

[61] Magauiya Zhussip, Shakarim Soltanayev, and Se Young Chun. Training deep learning based image denoisers from undersampled measurements without ground truth and without image prior. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10255–10264, 2019. 1