Reversing Flow for Image Restoration

Haina Qin[1,2*♡], Wenyang Luo[1,2*], Libin Wang[3†], Dandan Zheng[3],
Jingdong Chen[3], Ming Yang[3], Bing Li[1,2], Weiming Hu[1,2]
[1]State Key Laboratory of Multimodal Artificial Intelligence Systems (MAIS), CASIA;
[2]School of Artificial Intelligence, University of Chinese Academy of Sciences; [3]Ant Group;
{qinhaina2020@,luowenyang2020@,bli@nlpr.,wmhu@nlpr.}ia.ac.cn lbin.wlb@antgroup.com

Abstract

Image restoration aims to recover high-quality (HQ) images from degraded low-quality (LQ) ones by reversing the effects of degradation. Existing generative models for image restoration, including diffusion and score-based models, often treat the degradation process as a stochastic transformation, which introduces inefficiency and complexity. In this work, we propose ResFlow, a novel image restoration framework that models the degradation process as a deterministic path using continuous normalizing flows. ResFlow augments the degradation process with an auxiliary process that disambiguates the uncertainty in HQ prediction to enable reversible modeling of the degradation process. ResFlow adopts entropy-preserving flow paths and learns the augmented degradation flow by matching the velocity field. ResFlow significantly improves the performance and speed of image restoration, completing the task in fewer than four sampling steps. Extensive experiments demonstrate that ResFlow achieves state-of-the-art results across various image restoration benchmarks, offering a practical and efficient solution for real-world applications.

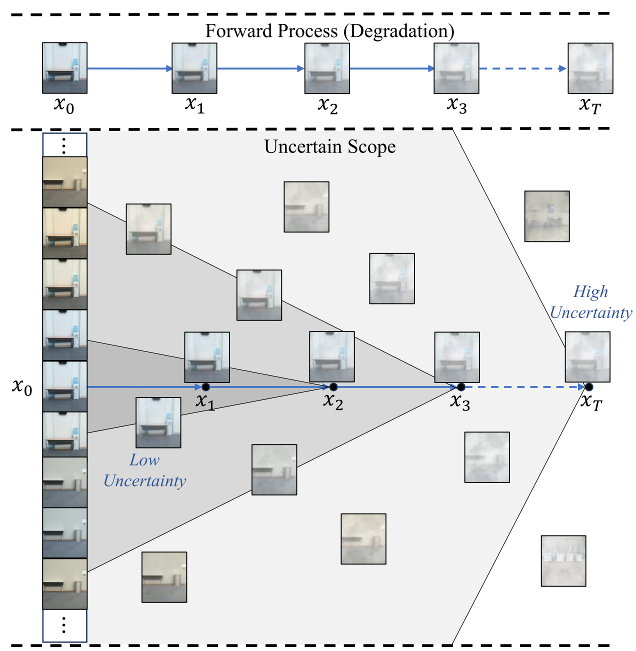

Figure 1. Image restoration is an ill-posed problem. The degradation process incurs decreasing mutual information between HQ and intermediate images, and multiple HQ images can degrade to similar or the same LQ images when their variations diminish in the process. Gray regions represent the uncertainty range from an intermediate state. As the image moves from LQ to HQ, its mutual information with HQ increases, and the uncertainty scope shrinks.

1. Introduction

Image restoration [24, 39, 51, 77, 96, 100] refers to recovering high-quality (HQ) images from degraded, low-quality (LQ) ones by reversing the effects associated with image degradation to reconstruct the original object. Many problems in image restoration, such as dehaze [17, 27, 124], weather removal [21, 60, 76, 101, 126], denoise [53, 71, 117], and artifact removal [23, 35].

Image restoration is ill-posed because it erases information in the HQ images. Given an LQ image, its corresponding HQ image is not necessarily unique, as depicted in Fig. 1. Image degradation gradually removes details from the HQ images. The corrupted image can result from multiple HQ images with different degradation processes. In fact, degradation is a Markov chain that transits from HQ to LQ images, which is subject to the data processing inequality (DPI) [11]: the mutual information between HQ and intermediate images decreases as more distortion is applied to the image. The ambiguous multi-correspondence between HQ and LQ images creates *uncertainty scopes* for LQ images defined as the collection of possible HQ images given the LQ images. The uncertainty scope should be disambiguated to make the degradation process reversible by

[*] Equal contribution. † Project lead. Project Page.
[♡] Work done during internship at Ant Group.

selecting exactly one HQ from the uncertainty scope given the intermediate image.

Generally, restoration methods navigate the uncertainty scope with prior knowledge besides the LQ images to reverse the degradation effects. The natural image distribution from diffusion and score-based generative models [31, 91, 93] serves as a strong prior for many approaches [48, 86, 104]. These methods define degradation as a conditional stochastic process that stochastically diffuses the HQ image into noise and learns a score function conditioned on the LQ image to reverse this process. The reverse process generates the HQ images under the guidance of the LQ images, where the LQ images provide structural and semantic hints. However, starting the reverse process from Gaussian noise is unnecessary and inefficient be cause it regenerates the structures already known in the LQ images.

To address this inefficiency, some studies incorporate prior knowledge of the degraded image directly into the forward stochastic process to enhance the efficiency of the reverse process [38, 55, 57, 64, 65, 89, 112]. For example, DDRM [38] progressively denoises samples stochastically to achieve the desired output. IR-SDE [64] models the degradation process using mean-reverting stochastic differential equations (SDEs), while I2SB [55] constructs a Schrödinger bridge between the clean and degraded data distributions. ResShift [112] transfers residuals from degraded low-resolution images to high-resolution ones for restoration in the latent space. RDDM [57] and Resfusion [89] introduce residual terms in the forward process. However, these approaches still treat the degradation process as a progressively diffusing stochastic forward process, which seems unnecessary and introduces additional complexity and inefficiency. Given that the degraded image is already known, the degradation process could be redefined as a deterministic forward process.

In this paper, we propose ResFlow, a novel general framework that reverses the deterministic paths between HQ and LQ images for image restoration. ResFlow models the forward process as a deterministic continuous normalizing flow [15, 54], directly simulating the path from high-quality images to low-quality ones. The image restoration process begins directly from the degraded image in the reverse process. The construction of the reverse process is non-trivial due to the existence of uncertainty scope: multiple plausible clear images may correspond to the same degraded image. We take into consideration the decrease of mutual information during degradation and augment the degradation process with an auxiliary process that couples with the uncertainty scope to disambiguate the velocity of the backward process. We also derive a flow path based on the intuition of entropy conservation in reversible processes. The deterministic flow path allows ResFlow to achieve better performance and faster generation speeds, completing image restoration in fewer than four sampling steps.

Our contributions are summarized as follows:

- We present ResFlow, a novel image restoration framework that reverses the deterministic degradation path between HQ and LQ images for image restoration, which achieves better performance and faster inference.
- ResFlow reverses image degradation by augmenting the degradation process with an auxiliary process that couples with the uncertainty scope, and adopts an entropy-preserving flow path in the reverse process.
- We conduct experiments on various image restoration tasks and datasets, and results demonstrate the effectiveness of ResFlow that sets new state-of-the-art performance for image restoration.

2. Related Work

Image restoration can be formulated as an inverse problem [9, 100] where a high quality (HQ) image is reconstructed given its degraded version (rainy [21, 101, 126], snowy [60, 76], hazy [17, 27, 124], noisy [53, 71, 117], compressed [23, 35], etc.) dubbed low quality (LQ) image. Image degradations usually erase information from HQ images; thus, the inverse problem (i.e., image restoration) is ill-posed in the sense that one LQ image may come from multiple HQ images by different degradations [9, 77, 100]. To tackle this problem, generative models are widely adopted to provide prior knowledge [24, 39, 51, 77, 100] about the distribution of HQ images.

Early works follow generative adversarial nets (GANs) [25, 70] and train a discriminator to guide the LQ images towards the distribution of HQ images [44, 96–98]. Variational autoencoders (VAEs) [40] are another choice that optimizes the evidence lower bound (ELBO) of the restored images [52, 90, 123]. These approaches predict the HQ images in a single step and suffer from the quality-diversity dilemma [105]: the adversarial training of GANs causes unstable training and mode collapse problem [7, 26, 67, 85]. In contrast, the mean-squared error (MSE) in VAE training leads to blurry and low-quality predictions [12, 120, 122].

Recent methods perform multi-step or iterative prediction [20, 57, 59, 65]: splitting the degradation process into small steps and investing each step is more tractable and easier to learn. Many works build upon diffusion models [31, 91, 93] or Schrödinger bridges [55] and model degradation processes as stochastic paths that transit between HQ and LQ images [38, 56, 57, 59, 64, 65, 89, 104, 110, 112, 121]. The uncertainty introduced by random transitions slows down training and inference. Other works [30, 63, 99] approximate the degradations as a sequence of reversible deterministic steps, dubbed normalizing flows [42, 78]. Taking the limit of the number of steps and making each step infinitesimally small, one obtains an ordinary differential equation (ODE), also known as a continuous

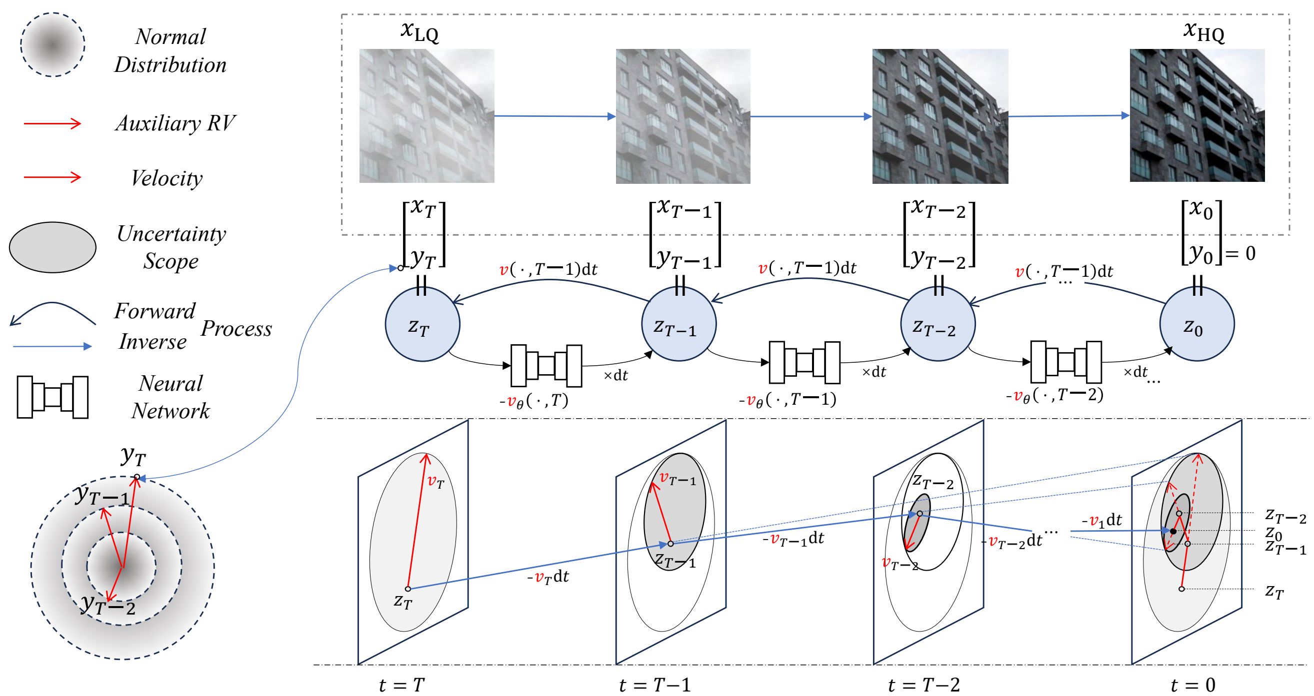

Figure 2. Framework of ResFlow. *RV* stands for *random variable*. The state $\boldsymbol{z}_t$ consists of a data component $\boldsymbol{x}_t$ that transits between HQ and LQ images and an auxiliary component $\boldsymbol{y}_t$ that disambiguates the velocity to ensure invertibility. The forward process is defined by interpolation, while the reverse process is learned by matching the velocity field. The lower part depicts the transition of ResFlow. Image degradation is usually non-reversible due to decreasing mutual information; thus, the velocity is uncertain for $\boldsymbol{x}_t$. The range reachable by possible velocity is dubbed uncertainty scope. As $\boldsymbol{x}_t$ approaches $\boldsymbol{x}_0$, the uncertainty in estimating velocity decreases, and so does the typical set of $\boldsymbol{y}_t$.

normalizing flow [15, 54, 75]. InDI [20] incrementally estimates the HQ images, which is equivalent to solving a "residual flow" that becomes sensitive to error in prediction near HQ, leading to sub-optimal results. These flow-based methods contradict the ill-posed and irreversible nature of the image degradation, leading to inferior performance.

Our work defines the degradations from HQ to LQ images by deterministic paths. Acknowledging the decrease of mutual information in image degradation that results in uncertain correspondence between LQ and HQ images, we augment the degradation process with an *auxiliary process* that couples with the uncertainty scope and guides the restoration directions. Our augmented degradation process is fully reversible and can be modeled by a deterministic ODE dubbed *degradation flow*. We learn the velocity field of the degradation flow by matching it with the ground-truth flow path [54], which leads to fast training and inference and achieves better restoration.

3. Method

3.1. Reversing Flow for Image Restoration

We tackle the image restoration problem by *inversion*: learning the degradation process that corrupts a high-quality (HQ) image $\boldsymbol{x}_{\text{HQ}}$ into a low-quality (LQ) image $\boldsymbol{x}_{\text{LQ}}$, and "reversing" the learned process to restore the HQ image from the LQ image. The learned degradation process should be 1) *reversible* to allow inverting LQ images to HQ images and 2) *tractable* to enable efficient training and inference.

A natural choice is to model the degradation process by an ordinary differential equation (ODE) on random process $\{\boldsymbol{z}_t \,|\, 0 \le t \le 1\}$:

$$\frac{\partial \boldsymbol{z}_t}{\partial t} = \boldsymbol{v}(\boldsymbol{z}_t, t); \quad 0 \le t \le 1, \tag{1}$$

where $\boldsymbol{v}$ is the velocity field, $\boldsymbol{z}_0$ corresponds to the HQ image, and $\boldsymbol{z}_1$ to the LQ image. Eq. (1) is also known as a continuous normalizing flow in the literature [15, 54].

However, Eq. (1) cannot directly apply to image degradation because it describes an reversible process. In contrast, image degradation is generally irreversible. This can be illustrated from the perspective of mutual information: a random process described by Eq. (1) preserves mutual information while image degradation does not. First, we have the following proposition.

Proposition 1. *Given random process $\boldsymbol{z}_t$ defined by Eq. (1), denote the mutual information as* $\mathrm{MI}(\cdot,\cdot)$*, then for any ref-*

erence random variable $\boldsymbol{r}$ and any $0 \leq t_1, t_2 \leq 1$, we have

$$\mathrm{MI}(\boldsymbol{z}_{t_1}, \boldsymbol{r}) = \mathrm{MI}(\boldsymbol{z}_{t_2}, \boldsymbol{r}). \tag{2}$$

Proof. See the supplementary material. □

The reference random variable $\boldsymbol{r}$ can be arbitrary state $\boldsymbol{z}_t \neq \boldsymbol{z}_0$, indicating that simulating Eq. (1) does not lose any information: everything we known about $\boldsymbol{z}_0$ remains in $\boldsymbol{z}_t$. However, this property does not hold for the ill-posed image degradation. Generally, the mutual information between the intermediate states and the HQ images decreases as more distortion is applied to the images as a consequence of the data processing inequality (DPI) [11]: as $\boldsymbol{x}_t$ approaches $\boldsymbol{x}_{\mathrm{LQ}}$, it shares less mutual information with $\boldsymbol{x}_{\mathrm{HQ}}$, so generally less is known about HQ images. As a result, an LQ image can correspond to multiple HQ images through different degradation processes. Fig. 1 shows an example where adding haze to different HQ images blends their content and leads to similar or the same LQ images. Noise is another example that eventually converts HQ images into indistinguishable noise.

The collection of possible HQ images given the LQ images is dubbed the uncertainty scope. The uncertainty scope should be disambiguated to make the degradation process reversible by selecting exactly one HQ from the uncertainty scope given the intermediate image. We achieve this by augmenting the degradation process with an auxiliary process $\{\boldsymbol{y}_t \,|\, 0 \leq t \leq 1\}$ that couples with the uncertainty scope and evolves with the degraded image. The ODE states $\{\boldsymbol{z}_t\}$ now become

$$\boldsymbol{z}_t^\top = [\boldsymbol{x}_t^\top; \boldsymbol{y}_t^\top], \boldsymbol{z}_0^\top = [\boldsymbol{x}_{\mathrm{HQ}}^\top; \boldsymbol{y}_0^\top], \boldsymbol{z}_1^\top = [\boldsymbol{x}_{\mathrm{LQ}}^\top; \boldsymbol{y}_1^\top]. \tag{3}$$

Conceptually, $\boldsymbol{y}_t$ encodes the "information loss" caused by image degradation. According to DPI, the mutual information from the coupling between it and the HQ images increases as $\boldsymbol{z}_t$ approaches $\boldsymbol{z}_1$ to keep $\mathrm{MI}(\boldsymbol{z}_t, \boldsymbol{z}_0)$ constant. This makes the choice of $\{\boldsymbol{y}_t\}$ non-trivial because $\boldsymbol{y}_1$ should have the maximal mutual information with $\boldsymbol{x}_0$, which is equivalent to knowing the ground-truth $\boldsymbol{x}_0$ in prior and thus infeasible. The next section details how we parameterize the augmented flow and tackle this problem.

3.2. Parameterization

Learning to invert the degradation process by the augmented flow (Eqs. (1) and (3)) requires obtaining the coupling between the auxiliary $\boldsymbol{y}_t$ and $\boldsymbol{x}_0$ that conceptually maps to the degraded mutual information. While one can manually determine such coupling, it is not necessarily optimal. Instead, inspired by recent advances in transport-based generative modeling techniques [54, 58], we *learn the deterministic coupling starting from an arbitrary coupling between $\boldsymbol{y}_t$ and $\boldsymbol{x}_0$.*

Given a pair of HQ and LQ images, the only bound of the degradation process (Eq. (1)) is that it begins with the HQ image and ends with the LQ image as formally given by Eq. (3). Usually, the ground-truth "natural" degradation paths between HQ and LQ images are infeasible to acquire, so we instead define the flow paths of $\boldsymbol{z}_t$ as geodesics in the Euclidean space:

$$\boldsymbol{z}_t = \alpha_t \boldsymbol{z}_0 + \sigma_t \boldsymbol{z}_1, \quad \alpha_t, \sigma_t \geq 0, \tag{4}$$
$$\text{such that } \alpha_0 = \sigma_1 = 1, \;\; \alpha_1 = \sigma_0 = 0. \tag{5}$$

We name $\{\alpha_t, \sigma_t\}$ as the degradation schedule defining the flow paths' dynamics. Given the degradation schedules, the augmented degradation process Eqs. (1) and (3) can be parameterized by a neural network $\boldsymbol{v}_\theta$ that estimates the velocity field $\boldsymbol{v}$:

$$\frac{\partial [\boldsymbol{x}_t^\top; \boldsymbol{y}_t^\top]^\top}{\partial t} = \boldsymbol{v}_\theta(\boldsymbol{x}_t, \boldsymbol{y}_t, t). \tag{6}$$

The auxiliary $\boldsymbol{y}_t$ is chosen to be Gaussian at $t = 1$ as it maximizes entropy and zero at $t = 0$ where the restoration ends. During training, $\{\boldsymbol{y}_t\}$ are independently coupled with both $\boldsymbol{x_0}$ and $\boldsymbol{x}_1$. However, the trained velocity network $\boldsymbol{v}_\theta$ induces a deterministic coupling between $\boldsymbol{x}_0$, $\boldsymbol{x}_1$ and $\boldsymbol{y}_t$ as defined by Eq. (6). This coupling disambiguates the velocity when multiple possible HQ images exist given an $\boldsymbol{x}_t$ as in Figs. 1 and 2. In such cases, the velocity network $\boldsymbol{v}_\theta$ uniquely maps $\boldsymbol{y}_t$ to one of the possible velocities.

We also note that the image $\boldsymbol{x}_t$ and auxiliary $\boldsymbol{y}_t$ in $\boldsymbol{z}_t$ do not necessarily have to use the same degradation schedule. Denoting the respective degradation schedules of $\boldsymbol{x}_t$ and $\boldsymbol{y}_t$ as $\{\alpha_t^{\boldsymbol{x}}, \sigma_t^{\boldsymbol{x}}\}$ and $\{\alpha_t^{\boldsymbol{y}}, \sigma_t^{\boldsymbol{y}}\}$, we propose the following entropy-preserving degradation schedule:

$$\alpha_t^{\boldsymbol{x}} = 1 - \sigma_t^{\boldsymbol{x}}, \quad \sigma_t^{\boldsymbol{x}} = t \tag{7}$$
$$\alpha_t^{\boldsymbol{y}} = 1 - \sigma_t^{\boldsymbol{y}}, \quad \sigma_t^{\boldsymbol{y}} = \beta \cdot (1 - t + \beta)^{-1}. \tag{8}$$

Here, $\beta = 10$ is a hyperparameter. This schedule moves $\boldsymbol{x}_t$ in straight lines (geodesics in Euclidean space). It keeps the entropy of $\boldsymbol{z}_t$ constant throughout the flow paths based on the intuition that the entropy remains the same for the reversible process (see the supplementary material for the derivation).

An alternative parameterization is to estimate the expectation $\mathbb{E}\left[\boldsymbol{z}_0 \,|\, \boldsymbol{z}_t\right]$ with a neural network [20]. However, this approach is equivalent to a time-weighted version of Eq. (6) that leads to high discretization errors near $t = 0$ when solved numerically.

Table 1. Synthetic datasets. Desnowing, Deraining, and Dehazing results on Snow100K [60], Outdoor-Rain [49], and Dense-Haze [4].

Method	Desnowing			Method	Deraining			Method	Dehazing	
	PSNR↑	SSIM↑	LPIPS↓		PSNR↑	SSIM↑	LPIPS↓		PSNR↑	SSIM↑
SPANet[95]	23.70	0.793	0.104	CycleGAN[125]	17.62	0.656	-	DehazeNet[13]	13.84	0.43
JSTASR[16]	25.32	0.807	0.059	pix2pix[34]	19.09	0.710	-	AOD-Net[47]	13.14	0.41
RESCAN[50]	26.08	0.810	0.054	HRGAN[49]	21.56	0.85	0.154	SGID[8]	12.49	0.51
DesnowNet[60]	27.17	0.898	0.070	PCNet[36]	26.19	0.901	0.132	MSBDN[22]	15.37	0.49
MPRNet[114]	29.76	0.894	0.049	MPRNet[114]	28.03	0.919	0.089	FFA-Net[79]	14.39	0.45
NAFNet[14]	30.06	0.901	0.051	NAFNet[14]	29.59	0.902	0.085	AECR-Net[103]	15.80	0.47
Restormer[115]	30.52	0.909	0.047	Restormer[115]	29.97	0.921	0.074	DeHamer[27]	16.62	0.56
SnowDiff64[76]	30.43	0.914	0.035	RainHazeDiff64[76]	28.38	0.932	0.067	PMNet[108]	16.79	0.51
SnowrDiff128[76]	30.28	0.900	0.038	RainHazeDiff128[76]	26.84	0.915	0.071	FocalNet[19]	17.07	**0.63**
DTPM-4[109]	30.92	**0.917**	0.034	DTPM-4[109]	30.99	0.934	0.0635	MB-Taylor[80]	16.44	0.56
ResFlow(Ours)	**31.86**	**0.917**	**0.030**	**ResFlow(Ours)**	**32.82**	**0.936**	**0.0514**	**ResFlow(Ours)**	**17.12**	0.59

3.3. Optimization and Inference

The parameterized augmented degradation ODE Eq. (6) is learned by matching the velocity field [54, 58]:

$$\min_{\theta} \mathbb{E}_{\boldsymbol{x}_0,\boldsymbol{x}_1,\boldsymbol{y}_0,\boldsymbol{y}_1} \left[\int_0^1 \lambda(t) \left\| \boldsymbol{v}_\theta(\boldsymbol{x}_t, \boldsymbol{y}_t, t) - \dot{\boldsymbol{z}}_t \right\|^2 \mathrm{d}t \right], \tag{9}$$

$$\text{where} \quad \dot{\boldsymbol{z}}_t = \begin{bmatrix} \dot{\boldsymbol{x}}_t \\ \dot{\boldsymbol{y}}_t \end{bmatrix} = \begin{bmatrix} \dot{\alpha}_t^{\boldsymbol{x}} \boldsymbol{x}_0 + \dot{\sigma}_t^{\boldsymbol{x}} \boldsymbol{x}_1 \\ \dot{\alpha}_t^{\boldsymbol{y}} \boldsymbol{y}_0 + \dot{\sigma}_t^{\boldsymbol{y}} \boldsymbol{y}_1 \end{bmatrix}. \tag{10}$$

Here, $\lambda(t)$ is a time-dependent loss weighting function, $\|\cdot\|$ is the L2 norm. Optimizing Eq. (10) is efficient as it does not require simulating Eq. (6) as in traditional simulation-based methods [15, 30, 63]. Moreover, Eq. (10) provides additional benefit that any convex transport cost induced by the coupling between $\boldsymbol{z}_0$ and $\boldsymbol{z}_1$ is guaranteed to be non-increasing [58] combined with Eq. (8).

We also propose a loss weighting scheme for image degradation that emphasizes time t close to 1, which empirically improves the quality of estimated HQ images:

$$\lambda(t) = \left(\cos\left(\frac{\pi}{2}(t-2) \right) + 1 \right)^{\gamma}, \tag{11}$$

where $\gamma = 1.75$. The rationale is that when $\boldsymbol{x}_t$ is close to the LQ image $\boldsymbol{x}_1$, its mutual information with the HQ image $\boldsymbol{x}_0$ decreases, making the velocity prediction more difficult near $t = 1$. Thus, we increase the loss weighting as t increases to balance the gradients towards those from the harder tasks.

After training by Eq. (10), we restore the LQ image $\boldsymbol{x}_1$ by first sampling $\boldsymbol{y}_1 \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$, then numerically solving Eq. (6) from $t = 1$ to $t = 0$ and obtain $\hat{\boldsymbol{x}}_0$ as the predicted HQ image. The intermediate $\hat{\boldsymbol{y}}_t$ is discarded and replaced with the ground-truth $\boldsymbol{y}_t$ given by Eq. (5). Conceptually, $\boldsymbol{y}_t$ is asymptotically mapped to the corrupted information during inference.

4. Experiments

4.1. Settings

To evaluate ResFlow's performance, we conducted experiments on five major image restoration tasks, including desnowing, draining, dehazing, denoising, and JPEG compression artifact removal, using synthetic and real-world datasets. For image desnowing, we conducted experiments on the Snow100K [60] dataset and the RealSnow [126] dataset. We used the Outdoor-Rain [49] dataset and the LHP [28] dataset to evaluate deraining performance. The Dense-Haze [4] and NH-HAZE [5] datasets were employed to assess image dehazing. We utilized the SIDD [1] dataset for real-world denoising. The DPDD [2] dataset was used to test single-image defocus deblurring. The LVE1 [87] and BSD500 [69] datasets were used to verify the removal of JPEG artifacts through entropy validation. We computed several distortion and perception-based metrics, including PSNR, SSIM [102], MAE, and LPIPS [120].

We adopted the same U-Net architecture [83] as DDPM [31] to predict the velocity in Eq. (1) for all tasks. Timestep t is embedded and injected into U-Net blocks via adaptive layer normalization [106]. The model was trained using the Adam optimizer [41] on 256-resolution image crops and tested on full-resolution images. The learning rate and other hyperparameters are detailed in Appendix C. We employed a uniform time schedule and performed only four sampling steps for all datasets.

4.2. Main Results

Desnowing Results. We report our method's quantitative performance on synthetic and real-world datasets in Tabs. 1 and 2. Overall, our method outperforms state-of-the-art algorithms across all datasets. Specifically, on the synthetic Snow100K-L [60] dataset, our method surpasses DTPM [109] by 0.86dB in PSNR. Additionally, our ap-

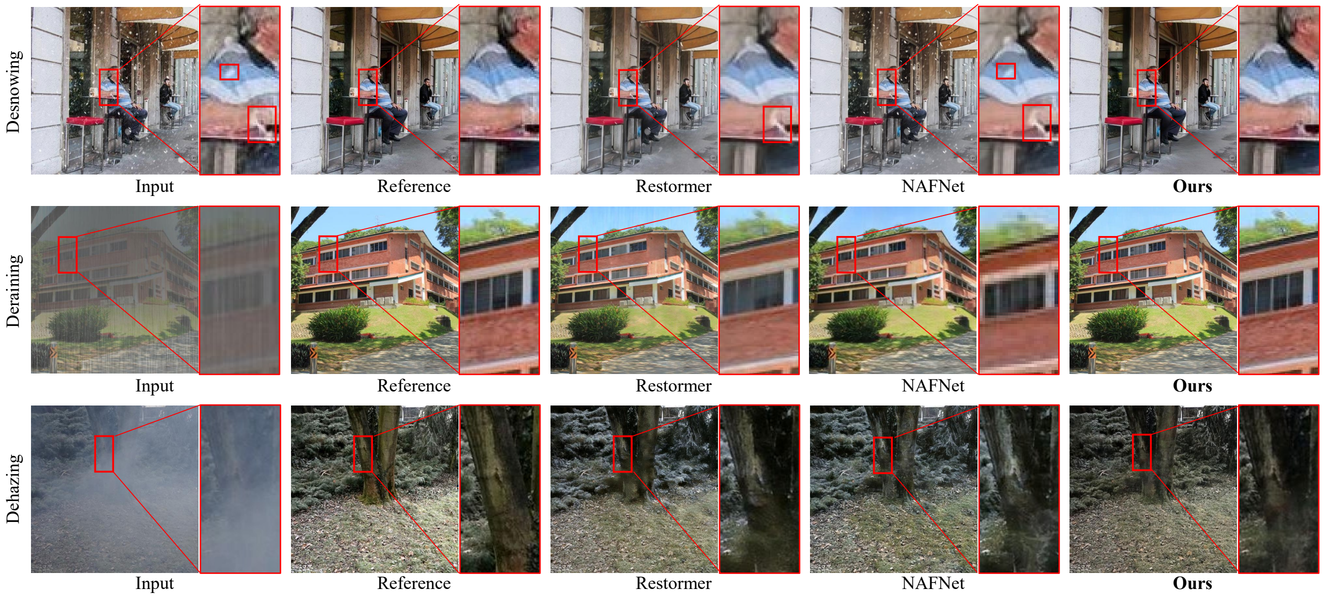

Figure 3. Dehazing, Deraining, and Desnowing results. The part of the image is methodized to observe the local details clearly. From left to right: input blurry images, reference images, and the predicted images obtained by Restormer [115], NAFNet [14], and our ResFlow, respectively. Note that both Restormer and NAFNet retain some artifacts in the output images.

Table 2. Real-world datasets. Dehazing results on NH-HAZE [5], Denoising results on SIDD [1], Deraining results on LHP [28], and Desnowing results on RealSnow [126]

Method	Denoising		Method	Dehazing		Method	Deraining		Method	Desnowing	
	PSNR	SSIM		PSNR	SSIM		PSNR	SSIM		PSNR	SSIM
DAGL[72]	38.94	0.953	DehazeNet[13]	16.62	0.52	SPANet[95]	31.19	0.934	MIRNetv2[116]	31.39	0.916
DeamNet[81]	39.47	0.957	AOD-Net[47]	15.40	0.57	PReNet[82]	32.13	0.917	ART[118]	31.05	0.913
MIRNet[113]	39.72	0.959	AECR-Net[103]	19.88	0.72	RCDNet[32]	32.34	0.915	Restormer[115]	31.38	**0.923**
DANet[111]	39.47	0.957	DeHamer[27]	20.66	0.68	MPRNet[114]	33.34	0.930	NAFNet[14]	31.44	0.919
Restormer[115]	40.02	0.960	PMNet[108]	20.42	0.73	SCD-Former[28]	34.33	**0.946**	WGWS-Net[126]	31.37	0.919
Xformer[119]	39.98	0.960	FocalNet[19]	20.43	**0.79**	IDT[35]	33.02	0.931	TransWeather[94]	31.13	0.922
ResFlow(Ours)	**42.26**	**0.962**	**ResFlow(Ours)**	**21.44**	**0.79**	**ResFlow(Ours)**	**34.54**	0.939	**ResFlow(Ours)**	**31.63**	0.919

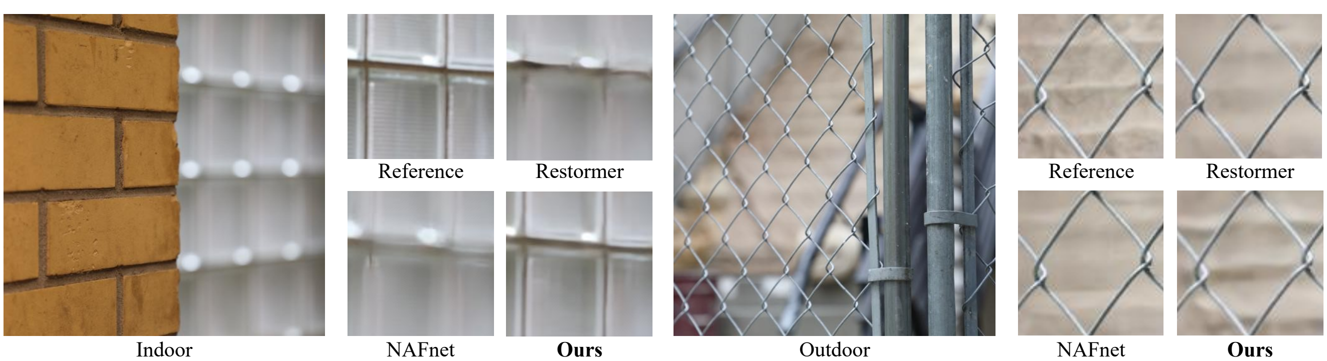

Figure 4. Single-image defocus deblurring results on the DPDD [2] dataset. The part of the image is methodized to observe the local details clearly. From left-top to right-bottom: input blurry images, reference images, and the predicted images obtained by Restormer [115], NAFNet [14], and our ResFlow, respectively. See the supplementary material for extra visualizations.

Table 3. Single-image Defocus Deblurring comparisons on the DPDD [2] datasets.

Method	**Indoor Scenes**				**Outdoor Scenes**				**Combined**			
	PSNR↑	SSIM↑	MAE↓	LPIPS↓	PSNR↑	SSIM↑	MAE↓	LPIPS↓	PSNR↑	SSIM↑	MAE↓	LPIPS↓
DPDNet[2]	26.54	0.816	0.031	0.239	22.25	0.682	0.056	0.313	24.34	0.747	0.044	0.277
KPAC[92]	27.97	0.852	0.026	0.182	22.62	0.701	0.053	0.269	25.22	0.774	0.040	0.227
DeepRFT[68]		-				-			25.71	0.801	0.039	0.218
IFAN[46]	28.11	0.861	0.026	0.179	22.76	0.72	0.052	0.254	25.37	0.789	0.039	0.217
DRBNet[84]		-				-			25.73	0.791	-	0.183
Restormer[115]	28.87	0.882	0.025	0.145	23.24	0.743	0.050	0.209	25.98	0.811	0.038	0.178
EBDB[37]		-				-			23.45	0.683	0.049	0.336
DMENet[45]		-				-			23.41	0.714	0.051	0.349
JNB[88]		-				-			23.84	0.715	0.048	0.315
FocalNet[19]	29.10	0.876	0.024	0.173	23.41	0.743	0.049	0.246	26.18	0.808	0.037	0.210
DTPM-4[109]		-				-			25.98	0.823	0.038	0.153
ResFlow(Ours)	**29.81**	**0.907**	**0.022**	**0.096**	**24.25**	**0.782**	**0.046**	**0.166**	**26.96**	**0.842**	**0.034**	**0.131**

proach performs well in more challenging real-world scenarios, achieving the best performance across most metrics. On the RealSnow [126] dataset, our method exceeds NAFNet [14] by 0.21dB in PSNR. Visual examples of several methods are provided in Fig. 3, where our method demonstrates superior removal of snowflakes and enhanced detail quality compared to other approaches. Crucially, our method successfully removes all the snowflakes dispersed in the whole image, while the compared methods leave out some snowflakes in their predicted images. We attribute this to our method's ability to perform multi-step restoration: single-step prediction can be imprecise when the degradation is strong, while our multi-step method can gradually remove the degradation, each step generating better results than in the previous steps.

Deraining Results. The numerical results for the synthetic Outdoor-Rain [49] dataset and the real-world LHP [28] dataset are presented in Tables 1 and 2. Our model exhibits strong deraining capabilities, achieving better or comparable results across all metrics. Our method demonstrates a significant 1.83dB improvement in PSNR over NAFNet [14] on the Outdoor-Rain dataset. Moreover, our method outperforms SCD-Former [28] on the more challenging real-world dataset by 0.20dB in PSNR. Visual results in Fig. 3 show that our method produces high-quality images resembling ground-truth images with no artifacts. In particular, our restored images faithfully retain more details, such as the bricks of the walls, than the compared methods, while the compared methods tend to produce smoothed appearances. When the correspondence between HQ and LQ images is ambiguous, one-step estimation converges to the mean of the HQ images conditioned on the LQ images. In contrast, our method introduces the auxiliary variable to disambiguate the HQ images and is able to preserve the sharp details.

Dehazing Results. As shown in Table 1, our method surpasses FocalNet [19] by 0.05dB in PSNR on the synthetic Dense-Haze [4] dataset. On the more challenging real-world NH-HAZE [5] dataset (Table 2), our approach achieves a notable 1.01dB improvement in PSNR over FocalNet [19], highlighting our method's superior performance in real-world scenarios. Visual results in Fig. 3 illustrate that our method effectively removes haze while preserving original details. Similar to the hazy example in Fig. 4, our model recovers more details than other methods, such as the tree's textures. This again demonstrates the promising performance of reversible flows.

Real Denoising Results. As shown in Tab. 2, our method achieves the best PSNR/SSIM results on the real-world SIDD citeabdelhamed2018high denoising dataset. In particular, our method improves PSNR by 2.24dB over Restormer [115], demonstrating its superior performance in real denoising scenarios compared to other state-of-the-art methods.

Defocus Deblurring Results. Tab. 3 presents the quantitative comparison on the DPDD [2] dataset. Our method achieves state-of-the-art results across all metrics compared to existing algorithms. Specifically, our method surpasses FocalNet [19] by 0.78dB in the overall category and significantly outperforms DTPM [109] by 0.98dB. Moreover, in both indoor and outdoor scenes, our method exceeds the second-best results by 0.71dB and 0.84dB, respectively. These results demonstrate the superior performance of our method in all scenarios. Fig. 4 provides visual comparisons.

4.3. Ablation Study

In this section, we perform an ablation study on the components of our methods, including experiments and analyses

Figure 5. Reserve process performance curves, averaged on 32 samples from the desnowing dataset. The right shows the intermediate results of two example images.

Table 4. Color JPEG compression artifact (QF=10) removal on BSD500 [69] and LIVE1 [87] datasets.

Method	BSD500 [69]		LIVE1[87]	
	PSNR	SSIM	PSNR	SSIM
QGAC[23]	27.74	0.802	27.62	0.804
FBCNN[18]	27.85	0.799	27.77	0.803
IPT[76]	27.57	0.792	27.37	0.799
SwinIR[107]	27.62	0.789	27.45	0.796
DAGN[66]	28.07	0.799	27.95	0.807
ResFlow(Ours)	**28.21**	**0.835**	**27.95**	**0.830**

Table 5. Ablation Study. The average performance on four real-world and synthetic datasets are reported.

Method		Average	
Degradation Schedule	**Auxiliary Variable**	**PSNR**	**SSIM**
Entropy-preserving	Gaussian	25.51	0.804
Constant	Guassian	23.81	0.786
Entropy-preserving	Constant	23.77	0.780
Injection of Auxiliary Variable		**PSNR**	**SSIM**
Adapter		25.51	0.804
Add		24.04	0.794
Channel-wise concatenation		24.94	0.798

about the reverse restoration process, which is the center of image restoration, and the formulation and effects of the auxiliary variable we introduce to enable flow-based degradation modeling.

Reverse Restoration Process. ResFlow restores images by reversing the degradation process from HQ to LQ images. Specifically, ResFlow progressively removes degradation and noise over several timesteps. We provide specific restoration examples in Fig. 5, where ResFlow tends to remove degradation along the shortest transport path. Additionally, Fig. 5 shows the performance curve for ResFlow in image desnowing. The results show that deblurring performance improves gradually with the number of steps, converging in the final few steps.

Reverse Auxiliary Variable. The auxiliary variable $\boldsymbol{y}_t$ is chosen as a Gaussian distribution at $t = 1$ and zero at $t = 0$ at the end of the restoration. We learn the coupling between the auxiliary variable and the uncertainty scope using transport-based generative modeling techniques. By preserving entropy for the reversible process, we provide a schedule for sampling the auxiliary variable at different time steps. Ablation studies on the sampling distribution, schedule, and coupling scheme are presented in Tab. 5. The results show that a Gaussian distribution, which maximizes entropy, helps the model learn the coupling between the auxiliary variable and the uncertainty scope. Additionally, entropy-preserving sampling schedules improve model performance compared to fixed distributions. Finally, we validate different augmentation forms for coupling the auxiliary variable with input, with the generalized augmentation form output by the Adaptor showing the best results.

5. Conclusion

In this paper, we introduced ResFlow, a new framework for image restoration that models the degradation process as a deterministic and invertible flow. By departing from the common practice of stochastic degradation in generative models, ResFlow eliminates unnecessary complexity and allows for a more direct inversion process. By augmenting the degradation process with an auxiliary mechanism, ResFlow disambiguates the uncertainty scope inherent to ill-posed restoration tasks, allowing for fast and efficient inversion of the degradation process. Through extensive experiments, we demonstrated that ResFlow outperforms existing approaches and sets new state-of-the-art, achieving superior restoration quality across several tasks with a few sampling steps. The experimental results across various datasets validate the effectiveness of ResFlow, making it a promising approach for practical image restoration tasks, from weather removal to denoising and beyond. Future work will explore extending ResFlow to more complex degradation models and applying it to video restoration tasks.

References

[1] Abdelrahman Abdelhamed, Stephen Lin, and Michael S Brown. A high-quality denoising dataset for smartphone cameras. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1692–1700, 2018. 5, 6, 3, 4

[2] Abdullah Abuolaim and Michael S Brown. Defocus deblurring using dual-pixel data. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part X 16*, pages 111–126. Springer, 2020. 5, 6, 7, 3, 4

[3] Eirikur Agustsson and Radu Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2017. 3

[4] Codruta O Ancuti, Cosmin Ancuti, Mateu Sbert, and Radu Timofte. Dense-haze: A benchmark for image dehazing with dense-haze and haze-free images. In *2019 IEEE international conference on image processing (ICIP)*, pages 1014–1018. IEEE, 2019. 5, 7, 2, 4

[5] Codruta O Ancuti, Cosmin Ancuti, and Radu Timofte. Nh-haze: An image dehazing benchmark with non-homogeneous hazy and haze-free images. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pages 444–445, 2020. 5, 6, 7, 2, 3, 4

[6] 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. 3

[7] Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks. In *International conference on machine learning*, pages 214–223. PMLR, 2017. 2

[8] Haoran Bai, Jinshan Pan, Xinguang Xiang, and Jinhui Tang. Self-guided image dehazing using progressive feature fusion. *IEEE Transactions on Image Processing*, 31:1217–1229, 2022. 5

[9] Mark R Banham and Aggelos K Katsaggelos. Digital image restoration. *IEEE signal processing magazine*, 14(2):24–41, 1997. 2

[10] Arpit Bansal, Eitan Borgnia, Hong-Min Chu, Jie Li, Hamid Kazemi, Furong Huang, Micah Goldblum, Jonas Geiping, and Tom Goldstein. Cold diffusion: Inverting arbitrary image transforms without noise. *Advances in Neural Information Processing Systems*, 36:41259–41282, 2023. 3

[11] Normand J Beaudry and Renato Renner. An intuitive proof of the data processing inequality. *arXiv preprint arXiv:1107.0740*, 2011. 1, 4

[12] Yochai Blau and Tomer Michaeli. The perception-distortion tradeoff. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 6228–6237, 2018. 2

[13] Bolun Cai, Xiangmin Xu, Kui Jia, Chunmei Qing, and Dacheng Tao. Dehazenet: An end-to-end system for single image haze removal. *IEEE transactions on image processing*, 25(11):5187–5198, 2016. 5, 6

[14] 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. 5, 6, 7

[15] Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018. 2, 3, 5

[16] Wei-Ting Chen, Hao-Yu Fang, Jian-Jiun Ding, Cheng-Che Tsai, and Sy-Yen Kuo. Jstasr: Joint size and transparency-aware snow removal algorithm based on modified partial convolution and veiling effect removal. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXI 16*, pages 754–770. Springer, 2020. 5

[17] Xiang Chen, Hao Li, Mingqiang Li, and Jinshan Pan. Learning a sparse transformer network for effective image deraining. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5896–5905, 2023. 1, 2

[18] 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. 8

[19] Yuning Cui, Wenqi Ren, Xiaochun Cao, and Alois Knoll. Focal network for image restoration. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 13001–13011, 2023. 5, 6, 7

[20] Mauricio Delbracio and Peyman Milanfar. Inversion by direct iteration: An alternative to denoising diffusion for image restoration. *arXiv preprint arXiv:2303.11435*, 2023. 2, 3, 4

[21] Sen Deng, Mingqiang Wei, Jun Wang, Yidan Feng, Luming Liang, Haoran Xie, Fu Lee Wang, and Meng Wang. Detail-recovery image deraining via context aggregation networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 14560–14569, 2020. 1, 2

[22] Hang Dong, Jinshan Pan, Lei Xiang, Zhe Hu, Xinyi Zhang, Fei Wang, and Ming-Hsuan Yang. Multi-scale boosted dehazing network with dense feature fusion. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2157–2167, 2020. 5

[23] Max Ehrlich, Larry Davis, Ser-Nam Lim, and Abhinav Shrivastava. Quantization guided jpeg artifact correction. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VIII 16*, pages 293–309. Springer, 2020. 1, 2, 8

[24] Ben Fei, Zhaoyang Lyu, Liang Pan, Junzhe Zhang, Weidong Yang, Tianyue Luo, Bo Zhang, and Bo Dai. Generative diffusion prior for unified image restoration and enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9935–9946, 2023. 1, 2

[25] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. *Communications of the ACM*, 63(11):139–144, 2020. 2

[26] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron C Courville. Improved training of wasserstein gans. *Advances in neural information processing systems*, 30, 2017. 2

[27] Chun-Le Guo, Qixin Yan, Saeed Anwar, Runmin Cong, Wenqi Ren, and Chongyi Li. Image dehazing transformer with transmission-aware 3d position embedding. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5812–5820, 2022. 1, 2, 5, 6

[28] Yun Guo, Xueyao Xiao, Yi Chang, Shumin Deng, and Luxin Yan. From sky to the ground: A large-scale benchmark and simple baseline towards real rain removal. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 12097–12107, 2023. 5, 6, 7, 2, 4

[29] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. 3

[30] Leonhard Helminger, Michael Bernasconi, Abdelaziz Djelouah, Markus Gross, and Christopher Schroers. Generic image restoration with flow based priors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 334–343, 2021. 2, 5

[31] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020. 2, 5, 3

[32] Huaibo Huang, Aijing Yu, Zhenhua Chai, Ran He, and Tieniu Tan. Selective wavelet attention learning for single image deraining. *International Journal of Computer Vision*, 129(4):1282–1300, 2021. 6

[33] Xun Huang and Serge Belongie. Arbitrary style transfer in real-time with adaptive instance normalization. In *Proceedings of the IEEE international conference on computer vision*, pages 1501–1510, 2017. 3

[34] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017. 5

[35] Jiaxi Jiang, Kai Zhang, and Radu Timofte. Towards flexible blind jpeg artifacts removal. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4997–5006, 2021. 1, 2, 6

[36] Kui Jiang, Zhongyuan Wang, Peng Yi, Chen Chen, Zheng Wang, Xiao Wang, Junjun Jiang, and Chia-Wen Lin. Rain-free and residue hand-in-hand: A progressive coupled network for real-time image deraining. *IEEE Transactions on Image Processing*, 30:7404–7418, 2021. 5

[37] Ali Karaali and Claudio Rosito Jung. Edge-based defocus blur estimation with adaptive scale selection. *IEEE Transactions on Image Processing*, 27(3):1126–1137, 2017. 7

[38] Bahjat Kawar, Michael Elad, Stefano Ermon, and Jiaming Song. Denoising diffusion restoration models. *Advances in Neural Information Processing Systems*, 35:23593–23606, 2022. 2

[39] Geonung Kim, Kyoungkook Kang, Seongtae Kim, Hwayoon Lee, Sehoon Kim, Jonghyun Kim, Seung-Hwan Baek, and Sunghyun Cho. Bigcolor: Colorization using a generative color prior for natural images. In *European Conference on Computer Vision*, pages 350–366. Springer, 2022. 1, 2

[40] Diederik P Kingma. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013. 2

[41] Diederik P Kingma. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 5

[42] Durk P Kingma and Prafulla Dhariwal. Glow: Generative flow with invertible 1x1 convolutions. *Advances in neural information processing systems*, 31, 2018. 2

[43] Alexander Kraskov, Harald Stögbauer, and Peter Grassberger. Estimating mutual information. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 69(6): 066138, 2004. 1, 2

[44] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690, 2017. 2

[45] Junyong Lee, Sungkil Lee, Sunghyun Cho, and Seungyong Lee. Deep defocus map estimation using domain adaptation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 12222–12230, 2019. 7

[46] Junyong Lee, Hyeongseok Son, Jaesung Rim, Sunghyun Cho, and Seungyong Lee. Iterative filter adaptive network for single image defocus deblurring. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2034–2042, 2021. 7

[47] Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng. Aod-net: All-in-one dehazing network. In *Proceedings of the IEEE international conference on computer vision*, pages 4770–4778, 2017. 5, 6

[48] Haoying Li, Yifan Yang, Meng Chang, Shiqi Chen, Huajun Feng, Zhihai Xu, Qi Li, and Yueting Chen. Srdiff: Single image super-resolution with diffusion probabilistic models. *Neurocomputing*, 479:47–59, 2022. 2

[49] 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. 5, 7, 2, 4

[50] 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. 5

[51] Xinqi Lin, Jingwen He, Ziyan Chen, Zhaoyang Lyu, Bo Dai, Fanghua Yu, Wanli Ouyang, Yu Qiao, and Chao Dong. Diffbir: Towards blind image restoration with generative diffusion prior. *arXiv preprint arXiv:2308.15070*, 2023. 1, 2

[52] Xinmiao Lin, Yikang Li, Jenhao Hsiao, Chiuman Ho, and Yu Kong. Catch missing details: Image reconstruction with frequency augmented variational autoencoder. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1736–1745, 2023. 2

[53] 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

[54] Yaron Lipman, Ricky TQ Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. *arXiv preprint arXiv:2210.02747*, 2022. 2, 3, 4, 5

[55] Guan-Horng Liu, Arash Vahdat, De-An Huang, Evangelos A Theodorou, Weili Nie, and Anima Anandkumar. I^2sb: Image-to-image schrödinger bridge. *arXiv preprint arXiv:2302.05872*, 2023. 2

[56] Haipeng Liu, Yang Wang, Biao Qian, Meng Wang, and Yong Rui. Structure matters: Tackling the semantic discrepancy in diffusion models for image inpainting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8038–8047, 2024. 2

[57] Jiawei Liu, Qiang Wang, Huijie Fan, Yinong Wang, Yandong Tang, and Liangqiong Qu. Residual denoising diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2773–2783, 2024. 2, 3

[58] Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. *arXiv preprint arXiv:2209.03003*, 2022. 4, 5

[59] Yuhao Liu, Zhanghan Ke, Fang Liu, Nanxuan Zhao, and Rynson WH Lau. Diff-plugin: Revitalizing details for diffusion-based low-level tasks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4197–4208, 2024. 2

[60] Yun-Fu Liu, Da-Wei Jaw, Shih-Chia Huang, and Jenq-Neng Hwang. Desnownet: Context-aware deep network for snow removal. *IEEE Transactions on Image Processing*, 27(6): 3064–3073, 2018. 1, 2, 5, 4

[61] I Loshchilov. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017. 3

[62] Ilya Loshchilov and Frank Hutter. Sgdr: Stochastic gradient descent with warm restarts. *arXiv preprint arXiv:1608.03983*, 2016. 3

[63] Andreas Lugmayr, Martin Danelljan, Fisher Yu, Luc Van Gool, and Radu Timofte. Normalizing flow as a flexible fidelity objective for photo-realistic super-resolution. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pages 1756–1765, 2022. 2, 5

[64] Ziwei Luo, Fredrik K Gustafsson, Zheng Zhao, Jens Sjölund, and Thomas B Schön. Image restoration with mean-reverting stochastic differential equations. *arXiv preprint arXiv:2301.11699*, 2023. 2

[65] Ziwei Luo, Fredrik K Gustafsson, Zheng Zhao, Jens Sjölund, and Thomas B Schön. Refusion: Enabling large-size realistic image restoration with latent-space diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1680–1691, 2023. 2

[66] Li Ma, Yifan Zhao, Peixi Peng, and Yonghong Tian. Sensitivity decouple learning for image compression artifacts reduction. *IEEE Transactions on Image Processing*, 2024. 8

[67] Xudong Mao, Qing Li, Haoran Xie, Raymond YK Lau, Zhen Wang, and Stephen Paul Smolley. Least squares generative adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2794–2802, 2017. 2

[68] Xintian Mao, Yiming Liu, Wei Shen, Qingli Li, and Yan Wang. Deep residual fourier transformation for single image deblurring. *arXiv preprint arXiv:2111.11745*, 2(3):5, 2021. 7

[69] 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. 5, 8

[70] Mehdi Mirza. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*, 2014. 2

[71] Nick Moran, Dan Schmidt, Yu Zhong, and Patrick Coady. Noisier2noise: Learning to denoise from unpaired noisy data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12064–12072, 2020. 1, 2

[72] Chong Mou, Jian Zhang, and Zhuoyuan Wu. Dynamic attentive graph learning for image restoration. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 4328–4337, 2021. 6

[73] Chong Mou, Xintao Wang, Liangbin Xie, Yanze Wu, Jian Zhang, Zhongang Qi, and Ying Shan. T2i-adapter: Learning adapters to dig out more controllable ability for text-to-image diffusion models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 4296–4304, 2024. 3

[74] Axi Niu, Kang Zhang, Trung X Pham, Jinqiu Sun, Yu Zhu, In So Kweon, and Yanning Zhang. Cdpmsr: Conditional diffusion probabilistic models for single image superresolution. *arXiv preprint arXiv:2302.12831*, 2023. 3

[75] Derek Onken, Samy Wu Fung, Xingjian Li, and Lars Ruthotto. Ot-flow: Fast and accurate continuous normalizing flows via optimal transport. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 9223–9232, 2021. 3

[76] Ozan Özdenizci and Robert Legenstein. Restoring vision in adverse weather conditions with patch-based denoising diffusion models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(8):10346–10357, 2023. 1, 2, 5, 8, 4

[77] Xingang Pan, Xiaohang Zhan, Bo Dai, Dahua Lin, Chen Change Loy, and Ping Luo. Exploiting deep generative prior for versatile image restoration and manipulation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(11):7474–7489, 2021. 1, 2

[78] George Papamakarios, Eric Nalisnick, Danilo Jimenez Rezende, Shakir Mohamed, and Balaji Lakshminarayanan. Normalizing flows for probabilistic modeling and inference. *Journal of Machine Learning Research*, 22(57):1–64, 2021. 2

[79] 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. 5

[80] Yuwei Qiu, Kaihao Zhang, Chenxi Wang, Wenhan Luo, Hongdong Li, and Zhi Jin. Mb-taylorformer: Multi-branch efficient transformer expanded by taylor formula for image dehazing. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 12802–12813, 2023. 5

[81] 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. 6

[82] Dongwei Ren, Wangmeng Zuo, Qinghua Hu, Pengfei Zhu, and Deyu Meng. Progressive image deraining networks: A better and simpler baseline. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 3937–3946, 2019. 6

[83] 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. 5

[84] Lingyan Ruan, Bin Chen, Jizhou Li, and Miuling Lam. Learning to deblur using light field generated and real defocus images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 16304–16313, 2022. 7

[85] Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. Improved techniques for training gans. *Advances in neural information processing systems*, 29, 2016. 2

[86] Shuyao Shang, Zhengyang Shan, Guangxing Liu, LunQian Wang, XingHua Wang, Zekai Zhang, and Jinglin Zhang. Resdiff: Combining cnn and diffusion model for image super-resolution. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 8975–8983, 2024. 2

[87] H Sheikh. Live image quality assessment database release 2. *http://live. ece. utexas. edu/research/quality*, 2005. 5, 8, 3

[88] Jianping Shi, Li Xu, and Jiaya Jia. Just noticeable defocus blur detection and estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 657–665, 2015. 7

[89] Zhenning Shi, Haoshuai Zheng, Chen Xu, Changsheng Dong, Bin Pan, Xueshuo Xie, Along He, Tao Li, and Huazhu Fu. Resfusion: Denoising diffusion probabilistic models for image restoration based on prior residual noise. *arXiv e-prints*, pages arXiv–2311, 2023. 2

[90] Jae Woong Soh and Nam Ik Cho. Variational deep image restoration. *IEEE Transactions on Image Processing*, 31: 4363–4376, 2022. 2

[91] Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International conference on machine learning*, pages 2256–2265. PMLR, 2015. 2

[92] Hyeongseok Son, Junyong Lee, Sunghyun Cho, and Seungyong Lee. Single image defocus deblurring using kernel-sharing parallel atrous convolutions. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2642–2650, 2021. 7

[93] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020. 2

[94] 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

[95] Tianyu Wang, Xin Yang, Ke Xu, Shaozhe Chen, Qiang Zhang, and Rynson WH Lau. Spatial attentive single-image deraining with a high quality real rain dataset. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 12270–12279, 2019. 5, 6

[96] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. Esrgan: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European conference on computer vision (ECCV) workshops*, pages 0–0, 2018. 1, 2

[97] Xintao Wang, Yu Li, Honglun Zhang, and Ying Shan. Towards real-world blind face restoration with generative facial prior. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9168–9178, 2021.

[98] Xintao Wang, Liangbin Xie, Chao Dong, and Ying Shan. Real-esrgan: Training real-world blind super-resolution with pure synthetic data. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 1905–1914, 2021. 2

[99] Yufei Wang, Renjie Wan, Wenhan Yang, Haoliang Li, Lap-Pui Chau, and Alex Kot. Low-light image enhancement with normalizing flow. In *Proceedings of the AAAI conference on artificial intelligence*, pages 2604–2612, 2022. 2

[100] Yinhuai Wang, Jiwen Yu, and Jian Zhang. Zero-shot image restoration using denoising diffusion null-space model. *arXiv preprint arXiv:2212.00490*, 2022. 1, 2

[101] 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, 2

[102] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility

to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004. 5

[103] 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. 5, 6

[104] Bin Xia, Yulun Zhang, Shiyin Wang, Yitong Wang, Xinglong Wu, Yapeng Tian, Wenming Yang, and Luc Van Gool. Diffir: Efficient diffusion model for image restoration. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 13095–13105, 2023. 2

[105] Zhisheng Xiao, Karsten Kreis, and Arash Vahdat. Tackling the generative learning trilemma with denoising diffusion gans. *arXiv preprint arXiv:2112.07804*, 2021. 2

[106] Jingjing Xu, Xu Sun, Zhiyuan Zhang, Guangxiang Zhao, and Junyang Lin. Understanding and improving layer normalization. *Advances in neural information processing systems*, 32, 2019. 5

[107] Wenhan Yang, Jiaying Liu, Shuai Yang, and Zongming Guo. Scale-free single image deraining via visibility-enhanced recurrent wavelet learning. *IEEE Transactions on Image Processing*, 28(6):2948–2961, 2019. 8

[108] Tian Ye, Yunchen Zhang, Mingchao Jiang, Liang Chen, Yun Liu, Sixiang Chen, and Erkang Chen. Perceiving and modeling density for image dehazing. In *European conference on computer vision*, pages 130–145. Springer, 2022. 5, 6

[109] Tian Ye, Sixiang Chen, Wenhao Chai, Zhaohu Xing, Jing Qin, Ge Lin, and Lei Zhu. Learning diffusion texture priors for image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2524–2534, 2024. 5, 7

[110] Fanghua Yu, Jinjin Gu, Zheyuan Li, Jinfan Hu, Xiangtao Kong, Xintao Wang, Jingwen He, Yu Qiao, and Chao Dong. Scaling up to excellence: Practicing model scaling for photo-realistic image restoration in the wild. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 25669–25680, 2024. 2

[111] Zongsheng Yue, Qian Zhao, Lei Zhang, and Deyu Meng. Dual adversarial network: Toward real-world noise removal and noise generation. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part X 16*, pages 41–58. Springer, 2020. 6

[112] Zongsheng Yue, Jianyi Wang, and Chen Change Loy. Resshift: Efficient diffusion model for image super-resolution by residual shifting. *Advances in Neural Information Processing Systems*, 36, 2024. 2

[113] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Learning enriched features for real image restoration and enhancement. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXV 16*, pages 492–511. Springer, 2020. 6

[114] 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. 5, 6

[115] 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. 5, 6, 7, 4

[116] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Learning enriched features for fast image restoration and enhancement. *IEEE transactions on pattern analysis and machine intelligence*, 45(2):1934–1948, 2022. 6

[117] Dan Zhang, Fangfang Zhou, Yuwen Jiang, and Zhengming Fu. Mm-bsn: Self-supervised image denoising for real-world with multi-mask based on blind-spot network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4189–4198, 2023. 1, 2

[118] Jiale Zhang, Yulun Zhang, Jinjin Gu, Yongbing Zhang, Linghe Kong, and Xin Yuan. Accurate image restoration with attention retractable transformer. *arXiv preprint arXiv:2210.01427*, 2022. 6

[119] Jiale Zhang, Yulun Zhang, Jinjin Gu, Jiahua Dong, Linghe Kong, and Xiaokang Yang. Xformer: Hybrid x-shaped transformer for image denoising. *arXiv preprint arXiv:2303.06440*, 2023. 6

[120] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 586–595, 2018. 2, 5

[121] Yuzhe Zhang, Jiawei Zhang, Hao Li, Zhouxia Wang, Luwei Hou, Dongqing Zou, and Liheng Bian. Diffusion-based blind text image super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 25827–25836, 2024. 2

[122] Hang Zhao, Orazio Gallo, Iuri Frosio, and Jan Kautz. Loss functions for image restoration with neural networks. *IEEE Transactions on computational imaging*, 3(1):47–57, 2016. 2

[123] Dihan Zheng, Xiaowen Zhang, Kaisheng Ma, and Chenglong Bao. Learn from unpaired data for image restoration: A variational bayes approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(5):5889–5903, 2022. 2

[124] Chu Zhou, Minggui Teng, Yufei Han, Chao Xu, and Boxin Shi. Learning to dehaze with polarization. *Advances in neural information processing systems*, 34:11487–11500, 2021. 1, 2

[125] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017. 5

[126] Yurui Zhu, Tianyu Wang, Xueyang Fu, Xuanyu Yang, Xin Guo, Jifeng Dai, Yu Qiao, and Xiaowei Hu. Learning

weather-general and weather-specific features for image restoration under multiple adverse weather conditions. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 21747–21758, 2023. 1, 2, 5, 6, 7, 4