Ingredient-oriented Multi-Degradation Learning for Image Restoration

Jinghao Zhang*, Jie Huang*, Mingde Yao, Zizheng Yang, Hu Yu, Man Zhou, Feng Zhao†
University of Science and Technology of China
{jhaozhang,hj0117,mdyao,yzz6000,yuhu520,manman}@mail.ustc.edu.cn, fzhao956@ustc.edu.cn

Abstract

Learning to leverage the relationship among diverse image restoration tasks is quite beneficial for unraveling the intrinsic ingredients behind the degradation. Recent years have witnessed the flourish of various All-in-one methods, which handle multiple image degradations within a single model. In practice, however, few attempts have been made to excavate task correlations in that exploring the underlying fundamental ingredients of various image degradations, resulting in poor scalability as more tasks are involved. In this paper, we propose a novel perspective to delve into the degradation via an ingredients-oriented rather than previous task-oriented manner for scalable learning. Specifically, our method, named Ingredients-oriented Degradation Reformulation framework (IDR), consists of two stages, namely task-oriented knowledge collection and ingredients-oriented knowledge integration. In the first stage, we conduct ad hoc operations on different degradations according to the underlying physics principles, and establish the corresponding prior hubs for each type of degradation. While the second stage progressively reformulates the preceding task-oriented hubs into single ingredients-oriented hub via learnable Principal Component Analysis (PCA), and employs a dynamic routing mechanism for probabilistic unknown degradation removal. Extensive experiments on various image restoration tasks demonstrate the effectiveness and scalability of our method. More importantly, our IDR exhibits the favorable generalization ability to unknown downstream tasks.

1. Introduction

Image restoration aims to recover the high-quality images from their degraded observations, which is a general term of a series of low-level vision tasks. In addition to achieving satisfactory visual effects in photography, image restoration is also widely used in many other real world scenarios, such as autopilot and surveillance. Complex environments put forward higher requirements for image restoration algorithms, when considering the variability and unknowability of the corruption types. Since most existing methods have been dedicated into single degradation removal, such as denoising [15, 24, 61], deraining [20, 52, 55], debluring [8, 40, 42], dehazing [26, 44, 45], low-light enhancement [14, 34, 50], *etc*., which do not satisfy the applications in real world scenarios.

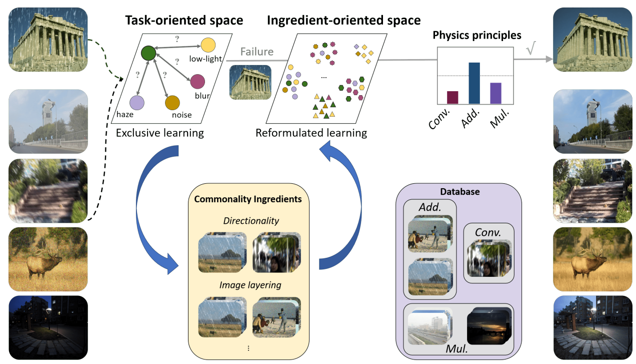

Figure 1. An illustration of our proposed ingredients-oriented degradation reformulation principle. Instead of previous task-oriented paradigm where each tasks are learned exclusively, we perform an ingredients-oriented paradigm to explore the correlation among diverse restoration tasks for scaleable degradation reformulated learning, where the *Conv.*, *Add.* and *Mul.* means the convolution, addition and multiplication.

Recently, all-in-one fashion methods have been coming to the fore, which handle multiple image degradations within a single model. These methods can be roughly categorized into two families, *i.e*., corruption-specific and corruption-agnostic. Representative studies of the former [2, 28] deal with different degradations via separate subnetworks, which demands pre-specification of corruption types, limiting the scope of further application. While the efforts in latter [25, 47] release the model from the prior of the corruption types, improving the flexibility in practice. However, both of them suffer from poor scalability as more tasks are involved, indicating that the diverse degradations are learned exclusively under the potential capability bottleneck, without touching the intrinsic correlation among them, which we referred as task-oriented paradigm.

*Both authors contributed equally to this research.
†Corresponding author.

To solve the above problem, we ask two questions: **i)** *'whether there are commonalities between different degradations?'* During past decades, few of works have been devoted to this field, [10] presented the interrelationship between image dehazing and low-light image enhancement. Going a step further, we envision that such association are widespread in various degradations, such as directionality in deblurring and deraining, unnatural image layering in deraining and denoising. Therefore, it is of great interest to consider the correlation among various restoration tasks for learning the intrinsic ingredients behind the degradation, which we referred as ingredient-oriented paradigm. **ii)** *'Whether an corrupted image definitely ascribed to only one type of degradation?'* In real world scenarios, it is hard to determine as multiple degradations may occur simultaneously, such as heavy rain typically accumulated with mist, or low-light combined with blur in night-time surveillance [63]. Therefore, it is inappropriate to learn each restoration task exclusively.

In this paper, we propose Ingredients-oriented Degradation Reformulation framework (IDR) for all-in-one image restoration, which provides a novel perspective via delving into the degradation and concentrating on the underlying fundamental ingredients. Specifically, the learning procedure of IDR consists of two stages, namely task-oriented knowledge collection and ingredients-oriented knowledge integration. We perform the above reformulation in the meta prior learning module (MPL) with the collaboration of both degradation representation and degradation operation, while the backbone network can be any transformer-based architecture. In the first stage, we conduct ad hoc operations for different degradations depending on the underlying physics principles, which pre-embedding the priors of disparate physics characteristics respectively. Meanwhile, separate task-oriented prior hubs are established for each type of degradation, responsible for excavating the specific degradation ingredients for compositional representation. While the second stage progressively reformulates the proceeding task-oriented hubs into single ingredients-oriented hub via learnable Principal Component Analysis (PCA), striving for commonalities among multiple degradations in terms of the ingredient-level, while preserving respective variance information as much as possible. Besides, a dynamic soft routing mechanism is employed in MPL for probabilistic unknown[1] degradation removal, according to the operation priors embedded in the first stage.

The contributions of this work are summarized as below:

- We rethink the current paradigm of all-in-one fashion methods, and propose to delve into the degradation for intrinsic ingredients excavation, in that improving the scalability of the model.
- We propose the Ingredients-oriented Degradation Reformulation framework (IDR) for image restoration, which consists of two stages, *i.e.*, task-oriented knowledge collection and ingredients-oriented knowledge integration, collaborating on both degradation representation and degradation operation.
- Extensive experiments are conducted to verify the effectiveness of our method. As far as we know, IDR is the first work to perform up to five image restoration tasks in an all-in-one fashion.

[1] Namely, the degradation types are not available in the second stage.

2. Related work

2.1. Image Restoration

Image restoration aims to restore degraded images to their clean counterparts, in that mitigating adverse circumstances of device or environment during imaging. Recent years have witnessed a great paradigm shift from conventional restoration methods to learning-based methods, due to their impressive performance on various image restoration tasks, such as denoising [15, 24, 61], deraining [20, 52, 55], debluring [8, 40, 42], dehazing [26, 44, 45], low-light enhancement [14, 34, 50], etc. Moreover, numerous general image restoration methods have also been proposed. [3] proposed a simple baseline for image restoration, which is a nonlinear activation free network. [4] deliberated the potential of normalization in low-level vision tasks. [43,58,59] proposed the collective goals of maintaining contextual information and spatial details. [13, 38] formulate the image restoration via unfolding strategy to deep into the rationality. With the flourish of vision transformers, their global modeling capability as well as the adaptability to input content have spawned a series of image restoration works, such as window-attention based [29, 49] , channel-interaction based [57] and latent-attention based [5].

Recently, all-in-one fashion methods have been prevalent, dealing with multiple degradations within a single model. [2] proposed a transformer based multi-head multi-tail framework for multiple degradation removal. [28] proposed a multi-encoder single-decoder network endowed with neural architecture search for several bad weather processing. [25] proposed a prior free network with contrastive learning that required none of task-specific heads or tails. [47] proposed to utilize weather type queries to deal with multiple degradation problems via single encoder decoder transformer. [30] proposed to learn task-agnostic prior for handling various image restoration tasks.

2.2. Multi-Task Learning

Multi-task learning [1] was introduced before the rise of deep learning, and has been applied to a series of fields, such as computer vision [31, 37], natural language processing [16], speech synthesis [51] and reinforcement learn-

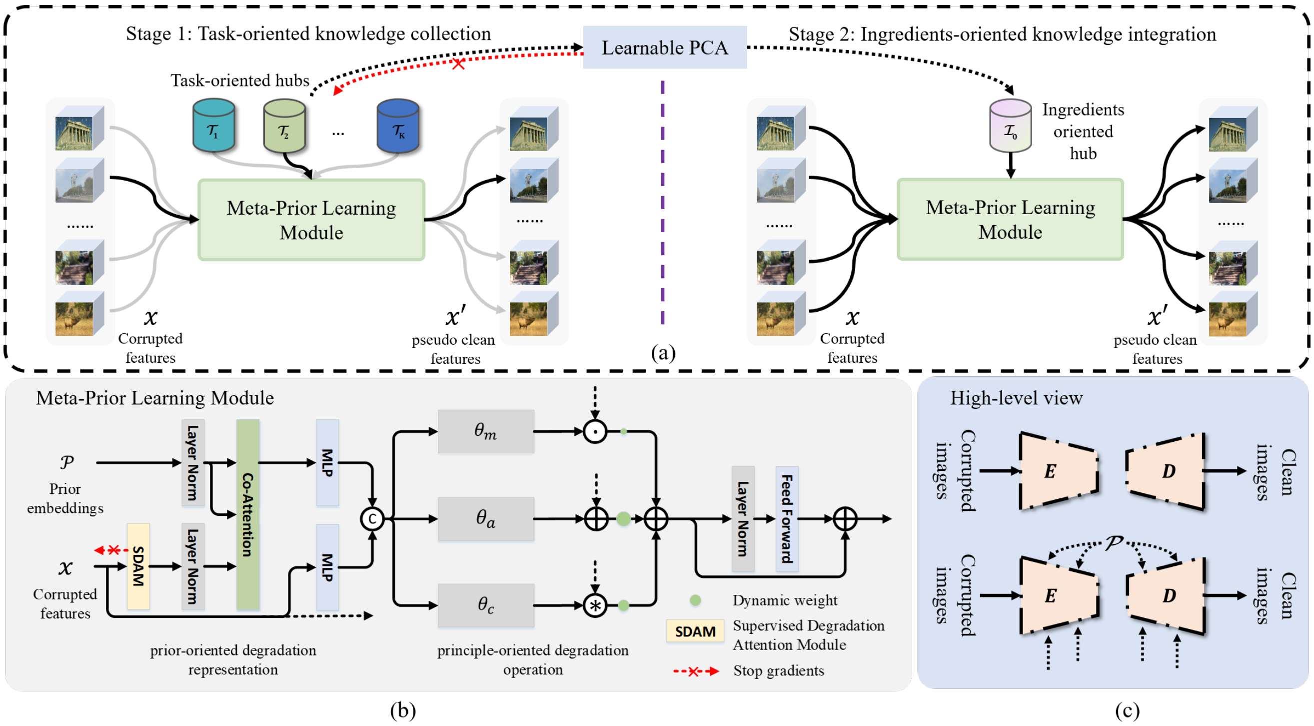

Figure 2. Overview of the proposed IDR. (a) The optimization procedure of the IDR which consists of two stages. (b) The architecture of the Meta-Prior Learning module. (c) High-level view of the IDR which ban be integrated into any transformer backbone networks via embedded Meta-Prior Learning modules (upward arrow) and collaboration with corresponding prior hub $\mathcal{P}$ (curved arrow).

ing [17]. As multi-task learning typically accompanied with optimization conflicts, a range of studies have been developed. [6] proposed to utilize gradient magnitudes to balance the loss functions for each task. [22] proposed a weighting mechanism by considering the homoscedastic uncertainty of each task. [54] formulated the meta learning with multiple objectives as multi-objective bi-level optimization problem, and devised a gradient-based optimization algorithm to find a common descent direction.

The particular bonus of multi-task learning lies in the relational exploitation of multiple tasks. [19, 21] proposed to cluster tasks into groups, resulting in a convex optimization formulation for multi-task learning. [32, 37] proposed to construct relationship networks and cross-stitch networks to discover the task relationships and learn optimal combination of shared and task-specific representations.

3. Method

In this section, we start with the formulation principles of various image degradations, and then introduce the core component of our IDR, *i.e.* meta-prior learning module, which can be explicitly embedded into any transformer architecture for practicability (Sec. 3.1). The optimization of IDR comprises a two-stage learning procedure: **(a)** Task-oriented knowledge collection (Sec. 3.2) and **(b)** Ingredients-oriented knowledge integration (Sec. 3.3). The optimization object is briefly presented in Sec. 3.4.

Degradation Formulation. In the literature, the image degradation process is generally defined as

$$\mathbf{y} = \phi(\mathbf{x}; \mathbf{A}) + \mathbf{N}, \tag{1}$$

where $\phi(\cdot)$ denotes the dergadation function and A represents its parameters, N represents the additive noise, y and x denote the degraded observation and latent clean image, respectively. When $\phi(\cdot)$ is the element-wise addition, the Eq. (1) can be reformulated as

$$\mathbf{y} = \mathbf{A} + \mathbf{x} + \mathbf{N} = \hat{\mathbf{A}} + \mathbf{x} \tag{2}$$

which is the general term of image deraining [48] and image denoising [9], and A represents the rain streak and i.i.d. zero-mean Gaussian noise, respectively. When $\phi(\cdot)$ is the element-wise multiplication,

$$\mathbf{y} = \mathbf{A} \cdot \mathbf{x} + \mathbf{N} = \mathbf{A} \cdot \mathbf{x} + \epsilon \cdot \mathbf{x} = \hat{\mathbf{A}} \cdot \mathbf{x}, \tag{3}$$

which is the general term of image dehazing [45] and low-light enhancement [50], according to the atmosphere scattering model [36, 41] and the Retinex theory [23]. And A represents the transmission map and illumination map, respectively. When $\phi(\cdot)$ represents the convolution,

$$\mathbf{y} = \mathbf{A} * \mathbf{x} + \mathbf{N} = \mathbf{A} * \mathbf{x} + \epsilon * \mathbf{x} = \hat{\mathbf{A}} * \mathbf{x}, \tag{4}$$

which is the general term of image deblurring [60], and A represents the blur kernel. Note that we re-param the noise term in Eqs. (2) to (4) to spotlight the peculiar physics principles, while more complex degradation can be formulated with the combination of the above functions, such as heavy rain with rain veiling effect. Basically, different degradations enjoy different compositional formulation principles, grounded on the above basic principles.

3.1. Meta-Prior Learning Module

The meta-prior learning module (MPL) aims at learning a set of prior embeddings to alleviate the awared corruption on the feature, and can be instantiated as task-prior learning and ingredient-prior learning in optimization. An illustration of MPL is depicted in Fig. 2, which consists of two parts, *i.e*. prior-oriented degradation representation and principle-oriented degradation operation.

Given the input of the degraded feature $x \in \mathbb{R}^{H \times W \times C}$ and a set of prior embeddings $\mathcal{P} \in \mathbb{R}^{N \times C_d}$, where N is the capacity of the prior hub and $N \ll C_d$, we first employ the supervised degradation attention module (SDAM) on x to dilute the background content while intensifying the latent subtle dergadation for content-agnostic prior learning, which is inspired by [59] with merely replaced degradation supervision. The process of the prior-oriented degradation representation can be formulated as

$$\mathrm{h_p} = \mathrm{MHCA}(\mathrm{LN}(\mathrm{SDAM}(x)), \mathrm{LN}(\mathcal{P}), \mathrm{LN}(\mathcal{P})) \tag{5}$$

where MHCA, LN denote the multi-head cross attention module and layer norm. Instead of taking additional priors as queries [30, 47], we borrow the idea from dictionary learning, which keep the few intrinsic prior embeddings as *keys* and release the tremendous degraded feature tokens as *queries* for potential knowledge compression. The degradation representation $h_p \in \mathbb{R}^{H \times W \times C}$ is the aggregation of the prior hub $\mathcal{P}$, according to the co-attention matrix modeled via the prior dictionary inquiry. We align the dimensions of $\mathcal{P}$ with x via the projection matrix.

Collaborated with the extracted h_p, the process of the principle-oriented degradation operation aims to correct the corrupted features depending on the underlying physics principles, which can be formulated as

$$\mathrm{x}' = \sum_{\mathrm{j} \in \{\mathrm{a,m,c}\}} \phi_\mathrm{j}(\mathrm{x}; \theta_\mathrm{j}([\mathrm{W_h h_p}, \mathrm{W_x x}])) \tag{6}$$

where $[\cdot]$ denotes the channel-wise concatenation, W_h, W_x are two individual MLPs for common space projection, ϕ_j represents the basic principle function including **a**ddition, **m**ultiplication and **c**onvolution, and θ_j denotes its parameter generation functions. Particularly, each θ_j is composed of a 3 × 3 convolution layer and two residual blocks, except for θ_c, which comprises an additional 1 × 1 convolution layer for kernel generation. The corrupted features x are corrected in virtue of the prior hub $\mathcal{P}$ and the underlying physics principles ϕ_j, producing the pseudo clean features $x' \in \mathbb{R}^{H \times W \times C}$. To further increase nonlinearity and to be consistent with the general transformer block, we retain the feed forward network (FFN) at the end of the MPL.

Algorithm 1 Learnable Principal Component Analysis

Input: Task-oriented prior embeddings $\{\mathcal{T}_k\}_{k=1}^K$
Output: Ingredient-oriented prior embedding $\mathcal{I}_0$

Stop the gradient for $\{\mathcal{T}_k\}_{k=1}^K$
for $\mathcal{T}_k$ in $\{\mathcal{T}_k\}_{k=1}^K$ **do**
    Calculate the singular value decomposition on $\mathcal{T}_k$, we have $\mathcal{T}_k = U_k S_k V_k^T$
Reformulate the$\{U_k\}_{k=1}^K$ and $\{V_k^T\}_{k=1}^K$ into $U_\mathcal{I}$ and $V_\mathcal{I}^T$ using Eq. (8).
Reformulate the $\{S_k\}_{k=1}^K$ into $S_\mathcal{I}$ using Eq. (9).
return $\mathcal{I}_0 = U_\mathcal{I} S_\mathcal{I} V_\mathcal{I}^T$

3.2. Task-oriented Knowledge Collection

The MPL introduced in Sec. 3.1 should extract the intrinsic degradation ingredients from the corrupted features, however, it is hard to directly learn the generalizable priors for diverse image degradations. Therefore, we first establish the independent task-oriented prior hubs for different restoration tasks to excavate the specific degradation ingredients, where the MPL is instantiated as tasks-prior learning in the first stage. The task-oriented prior hubs can be presented as $\{\mathcal{T}_k\}_{k=1}^K$, where K denotes the number of the tasks, $\mathcal{T}_k \in \mathbb{R}^{N \times C_d}$ is the set of k-th task-related prior embeddings, and the corresponding degraded features can be denoted as $\{x_k^{(i)}\}_{k=1}^K$. Consequently, the k-th task-related degradation representation $h_k^{(i)}$ can be expressed as

$$h_k^{(i)} = \sum_{j=k} MHCA(LN(SDAM(x_k^{(i)})), LN(\mathcal{T}_j), LN(\mathcal{T}_j)) \tag{7}$$

where only the specific $\mathcal{T}_k$ are interacted with the degraded features to exclude the leakage interference to other prior hubs for dedicated learning. We assign the principle function $\phi_j(\cdot\,; \theta_j)$ according to the specific degradation as discussed above, and only one type of principle is involved. We note that the collaboration between the independent prior learning and the corresponding hard principle allocation is quite favorable for the collection of the strong task-related priors, attributing to the principles-decoupled conformity.

3.3. Ingredient-oriented Knowledge Integration

Our main goal is to exploit the correlation among diverse image restoration tasks to realize the intrinsic generalizable degradation ingredients. Equipped with several task-oriented prior hubs $\{\mathcal{T}_k\}_{k=1}^K$, the second stage aims at progressively reformulates them into single ingredients-oriented hub $\mathcal{I}_0 \in \mathbb{R}^{N \times C_d}$, which is nontrivial to procure.

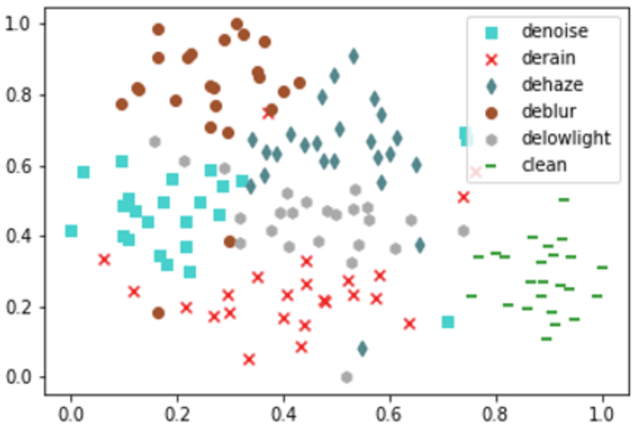

Figure 3. t-SNE visualization of the learned separate task-oriented prior hubs $\mathcal{T}_k$ in first optimization stage, where somewhat commonalities among them can be observed.

To this end, we propose the learnable principal component analysis that firstly decompose the respective $\mathcal{T}_k$ into multiple subspace via the singular value decomposition, and perform the integration at the singular vectors level. Meanwhile, the respective singular values matrices are freezed to dynamic adjust the propensity to specific degradation. The above design follow the interrelationship among the decomposed subspaces, which has been presented in [46]. The integration of the singular vectors can be formulated as

$$U_{\mathcal{I}} = \mathcal{G}_u(\{U_k\}_{k=1}^K),\ V_{\mathcal{I}}^T = \mathcal{G}_v(\{V_k^T\}_{k=1}^K) \tag{8}$$

where $\mathcal{T}_k = U_k S_k V_k^T$, and $U_{\mathcal{I}}, U_k \in \mathbb{R}^{N \times r}$, $S_k \in \mathbb{R}^{r \times r}$, $V_{\mathcal{I}}^T, V_k^T \in \mathbb{R}^{r \times C_d}$, $(r = min(N, C_d))$. Instead of taking the first N principle components of the $\{U_k\}_{k=1}^K$ and $\{V_k^T\}_{k=1}^K$ as the integration matrices, we employ the learnable $\mathcal{G}_u$ and $\mathcal{G}_v$ to progressively integrate them via neural networks. While the gradients of $\{U_k\}_{k=1}^K$ and $\{V_k^T\}_{k=1}^K$ are stopped to prevent undermining the originally learned task-oriented priors. The freezed $\{S_k\}_{k=1}^K$ are reformulated into $S_{\mathcal{I}} \in \mathbb{R}^{R \times R}$ adaptively, expressed as

$$S_{\mathcal{I}} = \sum_k \mathcal{M}(x_k) S_k \tag{9}$$

where $\mathcal{M}(\cdot)$ denotes the two-layer prediction head, providing dynamic weights to adjust the propensity of the reformulated $\mathcal{I}_0 = U_{\mathcal{I}} S_{\mathcal{I}} V_{\mathcal{I}}^T$ to specific degradation, as shown in Algorithm 1.

Consequently, the degradation representation $h_{\mathcal{I}}$ can be obtained via the interaction of the degraded features with the reformulated ingredients-oriented priors $\mathcal{I}_0$. Additionally, the dynamic soft routing mechanism is employed in place of the hard principle allocation, where the weights for each ϕ_j are derived from $\mathcal{M}(x_k)$ depending on the underlying physics principles. Thanks to the specialized learning in the first stage that endows the separate θ_j with specialized principle prior, the synergy of them can be more encyclopedic and are capable to handle more complex degradations.

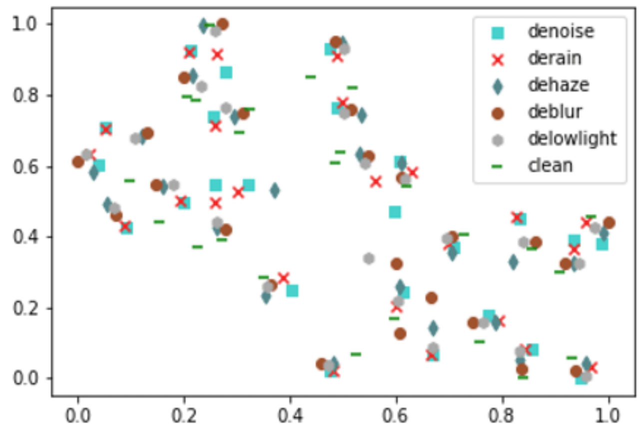

Figure 4. t-SNE visualization of the reformulated ingredients-oriented prior hub $\mathcal{I}_0$ with different degradation propensity $S_{\mathcal{I}}$ in the second optimization stage.

3.4. Optimization object

We optimize our IDR end-to-end with the combination of the reconstruction loss $\mathcal{L}_{rec}$ and classification loss $\mathcal{L}_{cls}$:

$$\mathcal{L}_{total} = \mathcal{L}_{rec} + \lambda_{cls}\mathcal{L}_{cls}, \tag{10}$$

The $\mathcal{L}_{rec}$ comprises the ℓ_1 loss between the restored image I and the ground-truth image Y, as well as the degraded supervision D introduced in each SDAM, formulated as

$$\mathcal{L}_{rec} = \mathcal{L}_1(I, Y) + \sum_{s \in \mathcal{C}} |\mathcal{L}_1(I_x^s, D^s)|, \tag{11}$$

where $\mathcal{C}$ denotes the set of the stages where MPL has been embedded into the backbone network, and I_x^s, D^s denote the restored image and the rescaled version of D in s stage. In addition, the cross-entropy loss is employed as $\mathcal{L}_{cls}$:

$$\mathcal{L}_{cls} = \sum_{s \in \mathcal{C}} |\mathcal{L}_{CE}(\mathcal{M}_s(x^s), y)|, \tag{12}$$

where x^s denotes the input corrupted feature of the s-th stage MPL, y denotes the task label of D, $\mathcal{M}_s(\cdot)$ denotes the s-th stage prediction head. The term of $\mathcal{L}_{cls}$ is included in the first stage, while the prediction results are served as the dynamic weights in the second optimization stage.

4. Experiments

In this section, we first clarify the experimental settings of our method, and then present the qualitative and quantitative comparison results with eleven state-of-the-art methods. Moreover, extensive experiments for ablation studies are conducted to verify the effectiveness of our method.

4.1. Implementation Details

Tasks and Metrics. We train our method on a combination of multiple image degradation datasets, following [25],

Table 1. **Quantitative results on five challenging image restoration datasets with state-of-the-arts general image restoration and all-in-one methods. The best and the second best results are marked in bold and underlined, respectively.**

Method	Rain100L [53]		SOTS [27]		BSD68 [35]		GoPro [39]		LOL [7]		**Average**		**Params**
	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑	
NAFNet [3]	35.56	0.967	25.23	0.939	31.02	0.883	26.53	0.808	20.49	0.809	27.76	0.881	17.11M
HINet [4]	35.67	0.969	24.74	0.937	31.00	0.881	26.12	0.788	19.47	0.800	27.40	0.875	88.67M
MPRNet [59]	**38.16**	**0.981**	24.27	0.937	31.35	**0.889**	26.87	0.823	20.84	0.824	28.27	0.890	15.74M
DGUNet [38]	36.62	0.971	24.78	0.940	31.10	0.883	27.25	0.837	**21.87**	0.823	28.32	0.891	17.33M
MIRNetV2 [56]	33.89	0.954	24.03	0.927	30.97	0.881	26.30	0.799	21.52	0.815	27.34	0.875	5.86M
SwinIR [29]	30.78	0.923	21.50	0.891	30.59	0.868	24.52	0.773	17.81	0.723	25.04	0.835	0.91M
Restormer [57]	34.81	0.962	24.09	0.927	31.49	0.884	27.22	0.829	20.41	0.806	27.60	0.881	26.13M
DL [11]	21.96	0.762	20.54	0.826	23.09	0.745	19.86	0.672	19.83	0.712	21.05	0.743	2.09M
Transweather [47]	29.43	0.905	21.32	0.885	29.00	0.841	25.12	0.757	21.21	0.792	25.22	0.836	37.93M
TAPE [30]	29.67	0.904	22.16	0.861	30.18	0.855	24.47	0.763	18.97	0.621	25.09	0.801	1.07M
AirNet [25]	32.98	0.951	21.04	0.884	30.91	0.882	24.35	0.781	18.18	0.735	25.49	0.846	8.93M
IDR (Ours)	35.63	0.965	**25.24**	**0.943**	**31.60**	0.887	**27.87**	**0.846**	21.34	**0.826**	**28.34**	**0.893**	15.34M

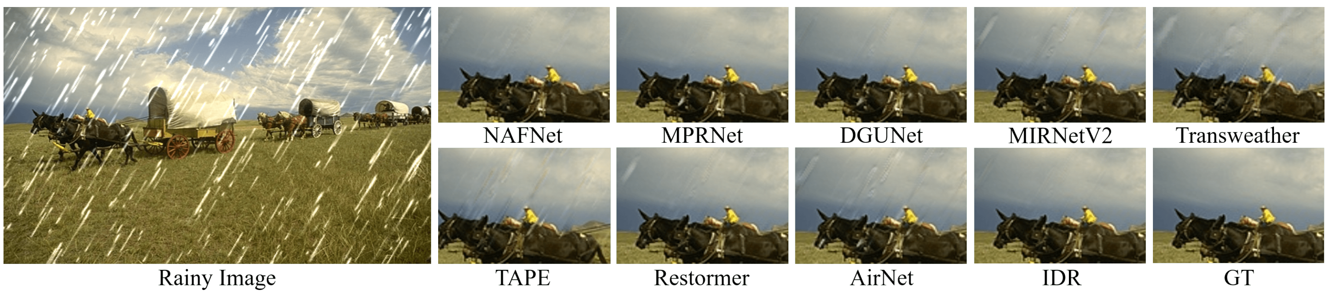

Figure 5. Visual comparison with state-of-the-art methods on Rain100L dataset. Please zoom in for details.

including Rain200L [53] for deraining, RESIDE [27] for dehazing, BSD400 [35] and WED [33] for denoising, GoPro [39] for deblurring and LOL [7] for low-light enhancement. For evaluation, Rain100L [53], SOTS-Outdoor [27], BSD68 [35], Urban100 [18], Kodak24 [12], GoPro [39], LOL [7] are utilized as the test sets. We report the Peak Signal to Noise Ratio (PSNR), Structural Similarity (SSIM) and Learned Perceptual Image Patch Similarity (LPIPS) [62] for numerical evaluation in our experiments.

Training. We implement our framework on single NVIDIA Geforce RTX 3090 GPU. The entire network is trained with Adam optimizer for 1200 epochs, and the initial learning rate is set to be 1×10^{-4}, gradually reduced to $1e^{-6}$ with the cosine annealing. We random crop 128x128 patch from original image as network input after data augumentation. The batch size is set as 8 with single degradation type in first stage while hybrid in the second stage. The label smoothing strategy is adopted in $\mathcal{L}_{cls}$ with $\epsilon = 0.1$, and the λ_{cls} is set to be 0.01. We adopt the trimmed restormer backbone with embedded MPLs at the end of multiple stages.

4.2. Comparison with state-of-the-art methods

We compare our IDR with seven general image restoration methods and four all-in-one fashion methods on five challenging image restoration tasks including deraining, dehazing, denoising, deblurring and low-light enhancement.

Table 1 reports the **quantitative comparison** results. Counterintuitively, the performance of the general image restoration methods are commonly exceeding the specialized all-in-one fashion methods as more tasks are involved. We attribute this to the large model size that endowed with more capability to handle complex mappings. However, our IDR reformulated this paradigm in a more efficient way. Consistent with existing methods [25, 57], Table 2 reports the denoising results at different noise ratio. Interestingly, it exhibits the comparable performance to previous individual task learning, compared with other degradations, suggesting the particular generality across diverse degradations.

Table 3 evaluates the performance of each method on **unknown tasks**, *i.e.* under-display camera (UDC) image restoration, without any fine-tuning. Typically, images captured under UDC system suffer from blurring due to the point spread function, and lower light transmission rate. The generalization ability of distinct methods is critically different. Comfortingly, our IDR demonstrates the favorable generalization ability via task-correlation excavation.

We present the **visual comparison** results of the foregoing image restoration tasks in Figs. 5 to 9. It can be observed that our IDR achieves steady performance in all tasks, compared with other methods. Due to the limited space, more bountiful results and the visual comparison on UDC dataset are provided in the supplementary material.

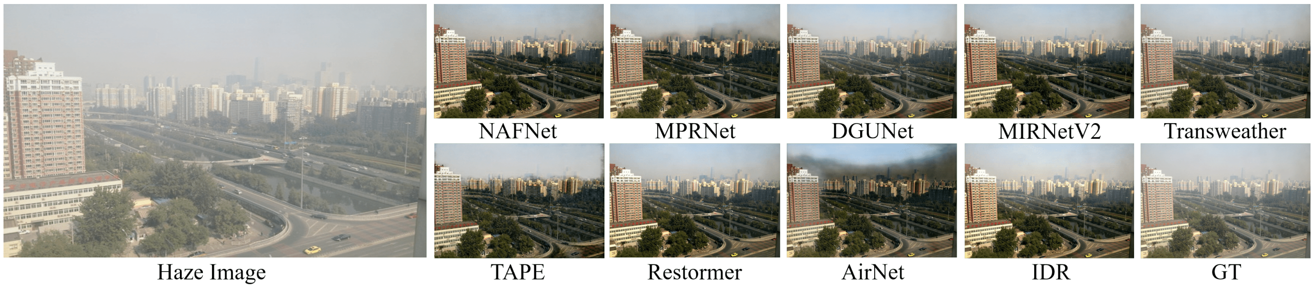

Figure 6. Visual comparison with state-of-the-art methods on SOTS dataset. Please zoom in for details.

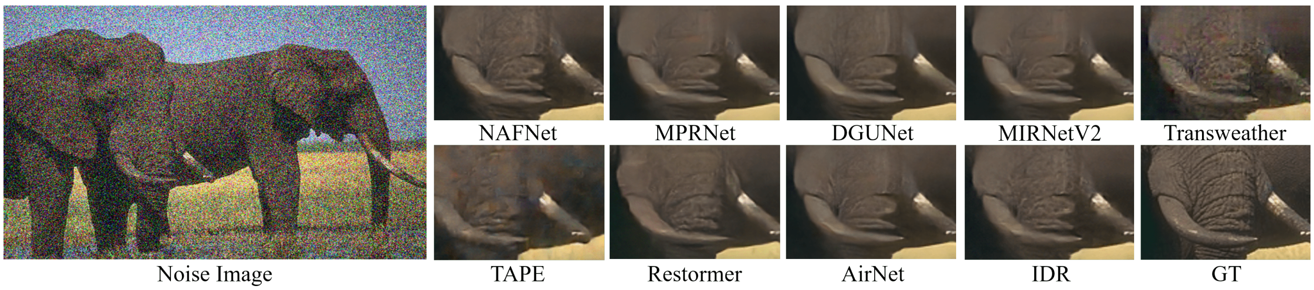

Figure 7. Visual comparison with state-of-the-art methods on BSD68 dataset. Please zoom in for details.

Table 2. **Quantitative results of image denoising on BSD68, Urban100 and Set12 datasets in terms of PSNR↑.**

Method	BSD68 [35]			Urban100 [18]			Kodak24 [12]		
	σ=15	σ=25	σ=50	σ=15	σ=25	σ=50	σ=15	σ=25	σ=50
NAFNet [3]	33.67	31.02	27.73	33.14	30.64	27.20	34.27	31.80	28.62
HINet [4]	33.72	31.00	27.63	33.49	30.94	27.32	34.38	31.84	28.52
MPRNet [59]	34.01	31.35	28.08	**34.13**	**31.75**	**28.41**	34.77	32.31	29.11
DGUNet [38]	33.85	31.10	27.92	33.67	31.27	27.94	34.56	32.10	28.91
MIRNetV2 [56]	33.66	30.97	27.66	33.30	30.75	27.22	34.29	31.81	28.55
SwinIR [29]	33.31	30.59	27.13	32.79	30.18	26.52	33.89	31.32	27.93
Restormer [57]	34.03	31.49	28.11	33.72	31.26	28.03	34.78	32.37	29.08
DL [11]	23.16	23.09	22.09	21.10	21.28	20.42	22.63	22.66	21.95
Transweather [47]	31.16	29.00	26.08	29.64	27.97	26.08	31.67	29.64	26.74
TAPE [30]	32.86	30.18	26.63	32.19	29.65	25.87	33.24	30.70	27.19
AirNet [25]	33.49	30.91	27.66	33.16	30.83	27.45	34.14	31.74	28.59
IDR (Ours)	**34.11**	**31.60**	**28.14**	33.82	31.29	28.07	**34.78**	**32.42**	**29.13**

Table 3. **Quantitative results of unknown tasks (under-display camera image restoration) on TOLED and POLED datasets.**

Method	TOLED [64]			POLED [64]		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
NAFNet [3]	26.89	0.774	0.346	10.83	0.416	0.794
HINet [4]	13.84	0.559	0.448	11.52	0.436	0.831
MPRNet [59]	24.69	0.707	0.347	8.34	0.365	0.798
DGUNet [38]	19.67	0.627	0.384	8.88	0.391	0.810
MIRNetV2 [56]	21.86	0.620	0.408	10.27	0.425	0.722
SwinIR [29]	17.72	0.661	0.419	6.89	0.301	0.852
Restormer [57]	20.98	0.632	0.360	9.04	0.399	0.742
DL [11]	21.23	0.656	0.434	13.92	0.449	0.756
Transweather [47]	25.02	0.718	0.356	10.46	0.422	0.760
TAPE [30]	17.61	0.583	0.520	7.90	0.219	0.799
AirNet [25]	14.58	0.609	0.445	7.53	0.350	0.820
IDR (Ours)	**27.91**	**0.795**	**0.312**	**16.71**	**0.497**	**0.716**

4.3. Abalation Studies

We perform the ablation experiments on the combined dataset to verify the effectiveness and scalability of our method. In Table 4, we quantitatively evaluate the **core components design and the two stage optimization procedure**. The metrics are reported on the average of all five datasets, from which we can make the following observations: **a)** It is difficult for the model to directly learn the ingredients-oriented priors without the assistance of the task specific optimization in stage 1. **b)** Compared with random integration (*i.e.* w/o *LPCA*), the proposed learnable principal component analysis is quite beneficial for the ingredients reformulation. **c)** The dynamic routing mechanism and the supervised degradation attention module are crucial for the overall performance improvement.

Table 5 evaluates the performance on all test sets with partially included train sets, where the R, H, N, B, L denotes the derain, dehaze, denoise, deblur and low-light enhancement, respectively. It can be observed that with more tasks involved, the performance ratains stable or even benefits, indicating the **scalability** of our method.

4.4. Discussion

We visualize the t-SNE statistics of the learned task-oriented prior embeddings $\{\mathcal{T}_k\}_{k=1}^{K}$ and the reformulated ingredients-oriented prior embeddings $\mathcal{I}_0$ in Figs. 3 and 4. In the first optimization stage, the learned respective task-oriented prior hubs are essentially separated, nevertheless, somewhat commonalities among them can be gleaned: **i)** The prior embeddings of the dehaze are crowding close to those of the low-light enhancement, suggesting their shared physics principles and the global degradation. **ii)** Despite the dissimilarity of the derain and deblur, few of embed-

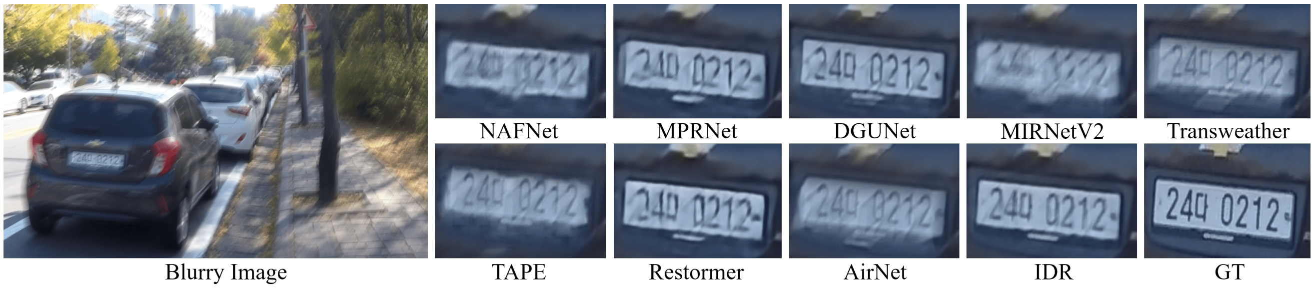

Figure 8. Visual comparison with state-of-the-art methods on GoPro dataset. Please zoom in for details.

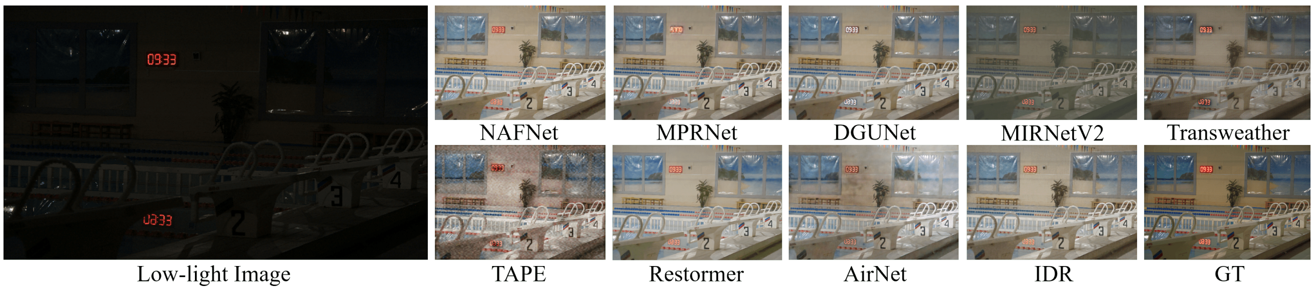

Figure 9. Visual comparison with state-of-the-art methods on LOL dataset. Please zoom in for details.

Table 4. **Ablation experiments on the components design.**

Method	*stage1*	*stage2*	*LPCA*	*Dyn.*	*SDAM*	PSNR↑	SSIM↑
a	✓		✓	✓	✓	28.06	0.880
b		✓	✓	✓	✓	27.92	0.878
c	✓	✓		✓	✓	27.98	0.883
d	✓	✓	✓		✓	28.07	0.881
e	✓	✓	✓	✓		28.23	0.889
f	✓	✓	✓	✓	✓	28.34	0.893

Table 5. **Ablation experiments for task scalability (PSNR↑).**

Tasks	Rain100L	SOTS	BSD68	GoPro	LOL
R+H+N	35.23	24.59	31.63	25.63	7.82
R+H+N+B	34.64	24.49	31.53	27.08	7.76
R+H+N+L	31.69	23.40	30.47	25.50	22.16
R+H+N+B+L	35.63	25.24	31.60	27.87	21.34

dings are interleaved, which may indicating that the potential directionality between them. **iii)** The noise are contiguous with all other types of embeddings, suggesting that the additive noise are widespread across diverse image degradations. **iv)** We further include an extra clean prior hub during training for reference, while the learned clean priors are apparently distinct with those corruption priors, indicating their natural oppositional essence.

Fig. 4 presents reformulated ingredients-oriented prior embeddings $\mathcal{I}_0$ with different degradation propensity $S_{\mathcal{I}}$ in the second stage. One can see that the learnable principal component analysis profoundly refreshes the distribution of various degradation priors, strengthening their correlations while preserving the respective variance information for diverse representation. In this way, the paradigm of multi-degradation learning is reformulated that dispose the constrains of the capability of the model for task-scaleable learning and improve the potential generalization ability.

4.5. Limitation and Future works

Despite the superior generalization ability and scalability that IDR have been made, it is of great interest to figure out the implication of the learned priors, and exploit their correlations for akin controllable degradation removal. Furthermore, how to leverage the clean image priors remains an open problem. Additionally, we hope IDR will be useful in promoting the further exploration of diverse degradation correlations for potential collaborative learning.

5. Conclusion

In this paper, we rethink the current paradigm of all-in-one fashion methods in image restoration, and propose to reformulate the degradation via a novel ingredients-oriented manner for task scalable learning. The proposed Ingredients-oriented Degradation Reformulation (IDR) framework consists of two stages, namely task-oriented knowledge collection and ingredients-oriented knowledge integration. Collaborating on both prior-oriented degradation representation and principle-oriented degradation operation, the learnable Principal Component Analysis (PCA) and the dynamic routing mechanism were proposed to realize the reformulation. Extensive experiments validate the effectiveness and scalability of the proposed method.

Acknowledgements

This work was supported by the JKW Research Funds under Grant 20-163-14-LZ-001-004-01, and the Anhui Provincial Natural Science Foundation under Grant 2108085UD12. We acknowledge the support of GPU cluster built by MCC Lab of Information Science and Technology Institution, USTC.

References

[1] Rich Caruana. Multitask learning. *Machine learning*, 28(1):41–75, 1997. 2

[2] Hanting Chen, Yunhe Wang, Tianyu Guo, Chang Xu, Yiping Deng, Zhenhua Liu, Siwei Ma, Chunjing Xu, Chao Xu, and Wen Gao. Pre-trained image processing transformer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12299–12310, 2021. 1, 2

[3] Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, and Jian Sun. Simple baselines for image restoration. *arXiv preprint arXiv:2204.04676*, 2022. 2, 6, 7

[4] Liangyu Chen, Xin Lu, Jie Zhang, Xiaojie Chu, and Chengpeng Chen. Hinet: Half instance normalization network for image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 182–192, 2021. 2, 6, 7

[5] Sixiang Chen, Tian Ye, Yun Liu, and Erkang Chen. Dualformer: Hybrid self-attention transformer for efficient image restoration. *arXiv preprint arXiv:2210.01069*, 2022. 2

[6] Zhao Chen, Vijay Badrinarayanan, Chen-Yu Lee, and Andrew Rabinovich. Gradnorm: Gradient normalization for adaptive loss balancing in deep multitask networks. In *International conference on machine learning*, pages 794–803. PMLR, 2018. 3

[7] Wenhan Yang Jiaying Liu Chen Wei, Wenjing Wang. Deep retinex decomposition for low-light enhancement. In *British Machine Vision Conference*. British Machine Vision Association, 2018. 6

[8] 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

[9] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on image processing*, 16(8):2080–2095, 2007. 3

[10] Xuan Dong, Yi Pang, and Jiangtao Wen. Fast efficient algorithm for enhancement of low lighting video. In *ACM SIGGRAPH 2010 Posters*, pages 1–1. 2010. 2

[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. 6, 7

[12] Rich Franzen. Kodak lossless true color image suite. *source: http://r0k. us/graphics/kodak*, 4(2), 1999. 6, 7

[13] Xueyang Fu, Zeyu Xiao, Gang Yang, Aiping Liu, Zhiwei Xiong, et al. Unfolding taylor's approximations for image restoration. *Advances in Neural Information Processing Systems*, 34:18997–19009, 2021. 2

[14] 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. 1, 2

[15] Shi Guo, Zifei Yan, Kai Zhang, Wangmeng Zuo, and Lei Zhang. Toward convolutional blind denoising of real photographs. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1712–1722, 2019. 1, 2

[16] Kazuma Hashimoto, Caiming Xiong, Yoshimasa Tsuruoka, and Richard Socher. A joint many-task model: Growing a neural network for multiple nlp tasks. *arXiv preprint arXiv:1611.01587*, 2016. 2

[17] Matteo Hessel, Hubert Soyer, Lasse Espeholt, Wojciech Czarnecki, Simon Schmitt, and Hado van Hasselt. Multi-task deep reinforcement learning with popart. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 3796–3803, 2019. 3

[18] 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, 7

[19] Laurent Jacob, Jean-philippe Vert, and Francis Bach. Clustered multi-task learning: A convex formulation. *Advances in neural information processing systems*, 21, 2008. 3

[20] 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, 2

[21] Zhuoliang Kang, Kristen Grauman, and Fei Sha. Learning with whom to share in multi-task feature learning. In *ICML*, 2011. 3

[22] Alex Kendall, Yarin Gal, and Roberto Cipolla. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7482–7491, 2018. 3

[23] Edwin H Land. The retinex theory of color vision. *Scientific american*, 237(6):108–129, 1977. 3

[24] Wooseok Lee, Sanghyun Son, and Kyoung Mu Lee. Apbsn: Self-supervised denoising for real-world images via asymmetric pd and blind-spot network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17725–17734, 2022. 1, 2

[25] 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, 2, 5, 6, 7

[26] 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. 1, 2

[27] 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

[28] 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. 1, 2

[29] 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, 7

[30] 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. *arXiv preprint arXiv:2203.06074*, 2022. 2, 4, 6, 7

[31] Shikun Liu, Edward Johns, and Andrew J Davison. End-to-end multi-task learning with attention. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1871–1880, 2019. 2

[32] Mingsheng Long, Zhangjie Cao, Jianmin Wang, and Philip S Yu. Learning multiple tasks with multilinear relationship networks. *Advances in neural information processing systems*, 30, 2017. 3

[33] 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

[34] 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. 1, 2

[35] 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*, volume 2, pages 416–423. IEEE, 2001. 6, 7

[36] Earl J McCartney. Optics of the atmosphere: scattering by molecules and particles. *New York*, 1976. 3

[37] Ishan Misra, Abhinav Shrivastava, Abhinav Gupta, and Martial Hebert. Cross-stitch networks for multi-task learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3994–4003, 2016. 2, 3

[38] 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. 2, 6, 7

[39] 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

[40] Seungjun Nah, Sanghyun Son, Suyoung Lee, Radu Timofte, and Kyoung Mu Lee. Ntire 2021 challenge on image deblurring. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 149–165, 2021. 1, 2

[41] Srinivasa G Narasimhan and Shree K Nayar. Vision and the atmosphere. *International journal of computer vision*, 48(3):233–254, 2002. 3

[42] Jinshan Pan, Haoran Bai, and Jinhui Tang. Cascaded deep video deblurring using temporal sharpness prior. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3043–3051, 2020. 1, 2

[43] Jinshan Pan, Sifei Liu, Deqing Sun, Jiawei Zhang, Yang Liu, Jimmy Ren, Zechao Li, Jinhui Tang, Huchuan Lu, Yu-Wing Tai, et al. Learning dual convolutional neural networks for low-level vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3070–3079, 2018. 2

[44] 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*, volume 34, pages 11908–11915, 2020. 1, 2

[45] Yuda Song, Zhuqing He, Hui Qian, and Xin Du. Vision transformers for single image dehazing. *arXiv preprint arXiv:2204.03883*, 2022. 1, 2, 3

[46] Yanpeng Sun, Qiang Chen, Xiangyu He, Jian Wang, Haocheng Feng, Junyu Han, Errui Ding, Jian Cheng, Zechao Li, and Jingdong Wang. Singular value fine-tuning: Few-shot segmentation requires few-parameters fine-tuning. *arXiv preprint arXiv:2206.06122*, 2022. 5

[47] 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. 1, 2, 4, 6, 7

[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. 3

[49] Zhendong Wang, Xiaodong Cun, Jianmin Bao, Wengang Zhou, Jianzhuang Liu, and Houqiang Li. Uformer: A general u-shaped transformer for image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17683–17693, 2022. 2

[50] 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. 1, 2, 3

[51] Zhizheng Wu, Cassia Valentini-Botinhao, Oliver Watts, and Simon King. Deep neural networks employing multi-task learning and stacked bottleneck features for speech synthesis. In *2015 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pages 4460–4464. IEEE, 2015. 2

[52] Jie Xiao, Xueyang Fu, Aiping Liu, Feng Wu, and Zheng-Jun Zha. Image de-raining transformer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022. 1, 2

[53] Wenhan Yang, Robby T Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, and Shuicheng Yan. Deep joint rain detection and removal from a single image. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1357–1366, 2017. 6

[54] Feiyang Ye, Baijiong Lin, Zhixiong Yue, Pengxin Guo, Qiao Xiao, and Yu Zhang. Multi-objective meta learn-

ing. *Advances in Neural Information Processing Systems*, 34:21338–21351, 2021. 3

[55] Yuntong Ye, Yi Chang, Hanyu Zhou, and Luxin Yan. Closing the loop: Joint rain generation and removal via disentangled image translation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2053–2062, 2021. 1, 2

[56] SW Zamir, A Arora, SH Khan, H Munawar, FS Khan, MH Yang, and L Shao. Learning enriched features for fast image restoration and enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022. 6, 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. 2, 6, 7

[58] 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 *European Conference on Computer Vision*, pages 492–511. Springer, 2020. 2

[59] 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. 2, 4, 6, 7

[60] Kaihao Zhang, Wenqi Ren, Wenhan Luo, Wei-Sheng Lai, Björn Stenger, Ming-Hsuan Yang, and Hongdong Li. Deep image deblurring: A survey. *International Journal of Computer Vision*, 130(9):2103–2130, 2022. 3

[61] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE transactions on image processing*, 26(7):3142–3155, 2017. 1, 2

[62] 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. 6

[63] Shangchen Zhou, Chongyi Li, and Chen Change Loy. Lednet: Joint low-light enhancement and deblurring in the dark. *arXiv preprint arXiv:2202.03373*, 2022. 2

[64] Yuqian Zhou, David Ren, Neil Emerton, Sehoon Lim, and Timothy Large. Image restoration for under-display camera. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9179–9188, 2021. 7