MetaWeather: Few-Shot Weather-Degraded Image Restoration

Youngrae Kim(✉), Younggeol Cho, Thanh-Tung Nguyen, Seunghoon Hong, and Dongman Lee

School of Computing, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea
{youngrae.kim,rangewing,tungnt,seunghoon.hong,dlee}@kaist.ac.kr

Abstract. Real-world weather conditions are intricate and often occur concurrently. However, most existing restoration approaches are limited in their applicability to specific weather conditions in training data and struggle to generalize to unseen weather types, including real-world weather conditions. To address this issue, we introduce MetaWeather, a universal approach that can handle diverse and novel weather conditions with a single unified model. Extending a powerful meta-learning framework, MetaWeather formulates the task of weather-degraded image restoration as a few-shot adaptation problem that predicts the degradation pattern of a query image, and learns to adapt to unseen weather conditions through a novel spatial-channel matching algorithm. Experimental results on the BID Task II.A, SPA-Data, and RealSnow datasets demonstrate that the proposed method can adapt to unseen weather conditions, significantly outperforming the state-of-the-art multi-weather image restoration methods. Code is available at https://github.com/RangeWING/MetaWeather.

Keywords: Image restoration · Few-shot learning · Domain adaptation

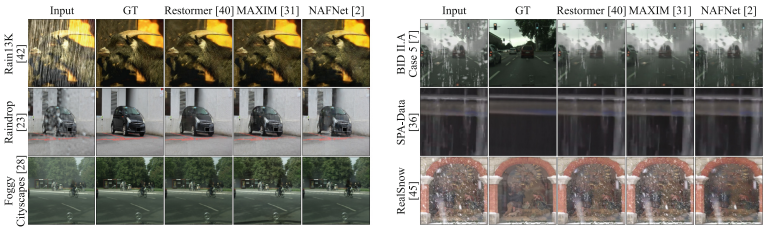
1 Introduction

Adverse weather conditions like rain, snow, and fog obfuscate visual clarity and degrade the quality of captured images and footage. Such degradation significantly and directly affects the performance of diverse vision tasks, including object detection, tracking, segmentation, and depth estimation [8,9,17,30]. The impact is especially critical for real-time outdoor vision applications, such as autonomous driving and surveillance, causing the systems to yield unreliable results. Therefore, weather-degraded image restoration has attracted much

Y. Kim and Y. Cho—These authors contributed equally.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-73464-9_13.

A. Leonardis et al. (Eds.): ECCV 2024, LNCS 15131, pp. 206–222, 2025.
https://doi.org/10.1007/978-3-031-73464-9_13



(a) Seen weather conditions **(b)** Unseen weather conditions

Fig. 1. (a). The state-of-the-art image restoration models perform well on the seen weather types in the training data, including rain [42], snow [18], raindrop [23], and fog [28]. (b). In contrast, these models show significant performance degradation on unseen weather types outside the scope of the training data, such as real rain (SPA-Data [36]) and real snow (RealSnow [45]), as well as co-occurrence of the weather types, rain, fog, and raindrops, in the training data (BID II.A Case 5 [7]). Best viewed with zoom and color.

attention from the research community in recent years. Recently, several works [15,33,39,45] have proposed methods for restoring images degraded by multiple types of weather conditions.

However, existing multi-weather degraded image restoration methods are unable to handle unseen weather conditions beyond the scope of their training data. Figure 1 shows that multi-weather degraded image restoration methods jointly trained on single-condition synthetic datasets, including rain [42], fog [28], snow [18], and raindrops [23], cannot deal with the unseen weather phenomena, which includes the co-occurrence of the trained conditions and the real weather-degraded footage. Our findings indicate that existing approaches do not adequately handle unseen weather conditions. A domain translation-based method [22] has been presented to cope with the real-world weather conditions. However, the method is unable to handle all cases of unseen weather conditions, as it only addresses the domain gap between synthetic and real data, rather than addressing more complex weather conditions. Restoring images degraded by unseen and arbitrary weather conditions has been under-explored, despite the imperative need.

To effectively restore such weather-degraded images, the following considerations should be met. First, the method should be data-efficient. In the real world, obtaining a sufficient number of clean images is often impractical. Several studies [1,36] have proposed filter-based methods for acquiring weather-degraded image pairs from real-world videos. However, due to the use of a stationary camera, they can only obtain one or a few pairs of degraded and clean images from a single location. Second, the method should be weather-agnostic. This is crucial because the deployed model may encounter various types of weather in real-world scenarios, including those outside the scope of the training data.

In this paper, we propose **MetaWeather**, a novel method for restoring images degraded by any unseen weather conditions with a single unified model. We treat the problem of image restoration for unseen weather as a few-shot adaptation problem, fully exploiting a small number of support pairs. MetaWeather

extends a flexible and universal meta-learning framework for dense prediction [11] to few-shot weather degraded image restoration. Specifically, given a few-shot support set of clean and degraded images, MetaWeather extracts the common degradation patterns of a query image, such as rain streaks and fog patterns, through spatial-channel matching. Then the identified degradation patterns are used to restore the query image. Our experiments demonstrate that matching the degradation patterns guides the architecture to learn more representation over a few-shot support set, thereby enabling more flexible adaptation. Additionally, we also observe that the spatial-channel matching better extracts the degradation patterns, where the spatial and channel dimensions primarily deal with weather particles and atmospheric effects, respectively.

Our contributions are summarized as follows:

- We propose MetaWeather, a novel few-shot weather-degraded image restoration method that enables restoration under any unseen and weather conditions. To the best of our knowledge, this is the first work introducing few-shot learning to the image restoration task for unseen weather conditions without any prior assumption.
- We extend the matching-based meta-learning framework to extract degradation patterns through a novel matching algorithm, namely spatial-channel matching, which effectively leverages weather-related representations from a few-shot support set.
- We evaluate MetaWeather on the BID Task II.A [7], SPA-Data [36], and RealSnow [45] datasets containing weather-degraded images under various combined weather conditions, real rain, and real snow, respectively. Extensive experimental results demonstrate that the proposed method effectively adapts to any unseen weather conditions, outperforming the state-of-the-art multi-weather image restoration methods.

2 Related Works

2.1 Generic Image Restoration

To handle multiple weather conditions in a single model, All-in-One [15] first introduces a generic model that has the ability to handle multiple weather degradations by using an individual encoder for each, which leads to high computational complexity. To eliminate the limitation of All-in-One, different methods [21,33,39,45] are proposed, which can handle several weather degradation patterns with a generic encoder and decoder. Similarly, generic models have been proposed to handle various restoration tasks, such as denoising, deblurring, and enhancement [2,3,13,31,40,43,44]. As they possess versatile architectures capable of handling a multitude of degradation types, these networks have demonstrated outstanding performances across diverse restoration tasks. However, despite their efficacy, these models exhibit notable performance deterioration when encountering new patterns of degradation as shown in Fig. 1.

2.2 Few-Shot Learning on Image Restoration

Recently, few-shot learning approaches are exploited in various vision tasks for mitigating the label scarcity problem [10,26,34]. Within these, the ones closest to image restoration are dense prediction tasks [11], such as semantic segmentation [29,35]. Although few-shot learning methods have been widely explored in such tasks, it is challenging to directly apply the proposed methods to image restoration tasks. This is because the image restoration model needs to learn how to preserve fine spatial details to restore the image [41], and it is a challenging to generalize these details with only a few images.

Prior efforts to incorporate few-shot learning into image restoration have been presented. MLDN [5] employ a meta-learning methodology to establish connections between rainy and clean images, enabling adaptation to new rain patterns. Similarly, FLUID [25] introduces a framework for few-shot adaptation in the deraining task. Although FLUID exhibits adaptability in rainy conditions, its focus is limited to rainy weather scenarios where a prior degradation distribution is presumed. Similarly, Liu *et al.* [16] propose a model agnostic meta-learning [4] based method for the dehazing task, enabling adaptable responses to new types of haze. These works demonstrate successful attempts to introduce few-shot learning in weather-degraded image restoration tasks. However, it is important to note that these methods presuppose knowledge of future degradation types, allowing them to establish predefined knowledge aimed at expected weather conditions. In contrast, our proposed method maximizes the utilization of degradation patterns within few-shot images, facilitating the flexible transfer of model representation to new degradation patterns, and thus avoiding task type limitations or requirements for prior degradation information.

3 MetaWeather

3.1 Problem Formulation

Following the prior works [14,18,24], we consider a degradation pattern that produces the degraded image $\mathbf{X} \in \mathbb{R}^{H\times W\times 3}$ from the clean image $\mathbf{Y} \in \mathbb{R}^{H\times W\times 3}$:

$$\mathbf{X} = \mathbf{T} \odot (\mathbf{Y} + \mathbf{P}) + (1 - \mathbf{T}) \odot \mathbf{A}, \tag{1}$$

where $\mathbf{T}$ is the transmission map, $\mathbf{P}$ is the sum of degradation particles, and $\mathbf{A}$ is the global atmospheric light of the scene.

Our objective is to learn a function $\mathcal{F} : \mathbb{R}^{H\times W\times 3} \rightarrow \mathbb{R}^{H\times W\times 3}$ that produces a clean image $\hat{\mathbf{Y}}$ of a query image $\mathbf{X}$ degraded by an arbitrary weather condition ω, given a few-shot support set of N labeled pairs sampled from the same degradation patterns $\mathbf{S}_\omega = \{(\mathbf{X}_i, \mathbf{Y}_i)\}_{i\leq N}$:

$$\hat{\mathbf{Y}} = \mathcal{F}(\mathbf{X}; \mathbf{S}_\omega). \tag{2}$$

3.2 Matching for Unseen Weather Conditions

In Eq. (2), a weather effect induced by ω is shared across the given few-shot support set $\mathbf{S}_\omega$ with the query image $\mathbf{X}$, rather than the background $\mathbf{Y}$. For instance, given two images of rain, the rain streaks are likely to be similar across the images, while the backgrounds may differ. Thus, exploiting the degradation patterns would be a more efficient approach than leveraging the backgrounds in a few-shot weather degraded image restoration.

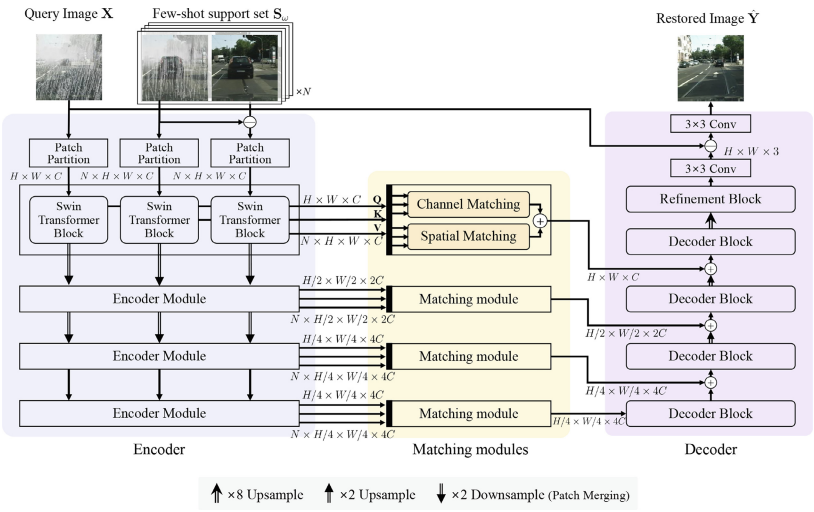


Fig. 2. Overall architecture of MetaWeather. MetaWeather consists of a hierarchical encoder-decoder design and matching module. Our matching module matches the degradation pattern between query and support set images, which enables MetaWeather to fully utilize a few-shot support set. The matching results are passed on to the decoder blocks at each level, and the extracted degradation pattern of the query image is then subtracted from the query image, resulting in the clean query image.

Inspired by this, we focus on the degradation pattern to enable effective utilization of a limited amount of data by leveraging the shared representations. We formulate the weather-degraded image restoration task as a task that produces a weather degradation pattern $\mathcal{G}(\mathbf{X})$ from an image $\mathbf{X}$:

$$\mathbf{Y} = \mathbf{X} - \mathcal{G}(\mathbf{X}), \quad \mathcal{G}(\mathbf{X}) = \mathbf{P} + (\mathbf{T}^{\circ -1} - 1) \odot (\mathbf{A} - \mathbf{X}), \tag{3}$$

where $\mathbf{T}^{\circ -1}$ is the element-wise inverse of $\mathbf{T}$.

To predict $\mathcal{G}(\mathbf{X})$ from $\mathbf{X}$, we employ a matching-based meta-learning framework [11] that predicts a query label via patch-wise matching. Given a support set $\{(\mathbf{X}_i, \mathbf{Y}_i)\}_{i \leq N} = \{(\mathbf{x}_k^i, \mathbf{y}_k^i)\}_{i \leq N, k \leq hw}$ and the patch size of $M = h \times w$, MetaWeather predicts the degradation pattern of a query image $\mathbf{X}^Q \in \mathbb{R}^{H \times W \times 3}$ by,

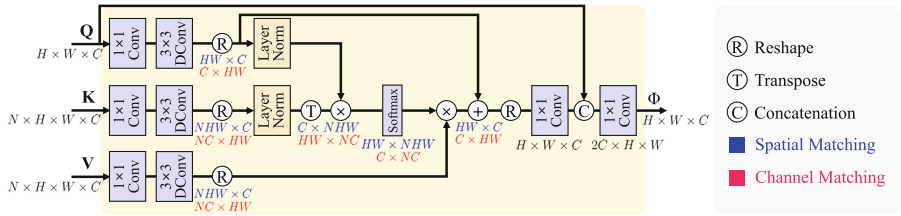


Fig. 3. Attention modules in our matching module.

$$f(\mathcal{G}(\mathbf{x}_j^Q)) = f(\mathbf{x}_j^Q - \mathbf{y}_j^Q) = \sum_{i=1}^{N}\sum_{k=1}^{M} \sigma(f(\mathbf{x}_j^Q), f(\mathbf{x}_k^i)) f(\mathbf{x}_k^i - \mathbf{y}_k^i), \tag{4}$$

where $\mathbf{X}^Q = \{\mathbf{x}_j^Q\}_{j \le M}$, f is the encoder, and σ is a similarity function. We extend the framework by replacing the query label $\mathbf{y}_j^Q$ with the degradation pattern $\mathcal{G}(\mathbf{x}_j^Q)$ and the label of the support set $\mathbf{y}_k^i$ with $\mathbf{x}_k^i - \mathbf{y}_k^i$. By introducing a decoder $h \approx f^{-1}$, the weather degradation pattern of the query image $\mathcal{G}(\mathbf{X}^Q)$ can be predicted from the encoded embeddings.

Spatial-Channel Matching. We present spatial-channel matching that takes into account the attribute of weather effects. Equation (3) decomposes the degradation pattern $\mathcal{G}(\mathbf{X})$ into degradation particles $\mathbf{P}$ and the atmospheric effect $(\mathbf{T}^{\circ -1} - 1) \odot (\mathbf{A} - \mathbf{X})$. These decomposed traits of weather effects suggest that the matching process should account for both the particles and the atmospheric effects. Spatial matching can address particles such as rain streaks, however, it is unsuitable for handling atmospheric effects since they are globally disseminated within an image. As atmospheric effects are often considered as a style and channel-wise features contain more style information compared to spatial features [6,12,32], we utilize channel information to extract atmospheric effects.

Motivated by these aspects, we devise the spatial-channel matching to extract the representation of atmospheric effects via channel-wise matching, while the degradation particles are extracted via spatial matching, in a decomposed manner. As channel matching would focus primarily on extracting atmospheric effects, spatial matching naturally prioritizes the extraction of degradation particles. By leveraging both dimensions in the matching module, the model can fully extract the weather representations from a few-shot support set and utilize them to predict the degradation patterns of incoming query images.

3.3 Implementation

The architecture of MetaWeather comprises a single-encoder, a single-decoder, and matching modules, as shown in Fig. 2. The encoder and decoder have hierarchical architectures. The encoder reduces spatial size while going deeper, while the decoder increases spatial size gradually following a U-Net-ike structure [27].

Encoder. The main objective of the encoder is to extract enriched and informative features of the query and the support set images to facilitate more effective matching between query and support images. These enriched features are passed on to the matching modules, where detailed comparisons and matching between the query and support images take place. The same encoder is shared to encode query and support set images. We adopt Swin Transformer [19] as our encoder. As the Swin Transformer hierarchically reduces spatial size and enlarges channel capacity, its intact structure can play the role of an encoder in a U-net-like structure, which has shown prominent performances on low-level vision tasks [2,37,40]. To build further generalized representation, we utilize SimMIM [38] pre-trained knowledge to initialize our encoder.

Matching Module. In order to make matching between the query image and support set images, the encoded features are bypassed from encoder blocks to matching modules. We employ channel and spatial attention in parallel in our matching module for more detailed matching results, as depicted in Fig. 3.

At each level of the encoder, given an encoded embedding $\mathbf{Q} \in \mathbb{R}^{H\times W\times C}$ of the query image $\mathbf{X^Q}$, and embeddings $\mathbf{K}, \mathbf{V} \in \mathbb{R}^{N\times H\times W\times C}$ of the support set $\{(\mathbf{X}_i, \mathbf{Y}_i)\}_{i\leq N}$, we project the embeddings into matrices $\mathbf{q} \in \mathbb{R}^{d\times l}, \mathbf{k}, \mathbf{v} \in \mathbb{R}^{Nd\times l}$ by $\mathbf{q} = \mathbf{W}_2^Q\mathbf{W}_1^Q\mathbf{Q}, \mathbf{k} = \mathbf{W}_2^K\mathbf{W}_1^K\mathbf{K}$, and $\mathbf{v} = \mathbf{W}_2^V\mathbf{W}_1^V\mathbf{V}$, where $\mathbf{W}_1^{(\cdot)}$ and $\mathbf{W}_2^{(\cdot)}$ are the 1×1 point-wise and the 3×3 depth-wise convolution, respectively. Note that $d = HW$ and $l = C$ for spatial matching and $d = C$ and $l = HW$ for channel matching. We apply layer normalization to the $\mathbf{q}$ and the $\mathbf{k}$ for effective matching, $\overline{\mathbf{q}} = \text{LayerNorm}(\mathbf{q}), \overline{\mathbf{k}} = \text{LayerNorm}(\mathbf{k})$. We predict the label of the query Φ by leveraging a modified Multi-Dconv Head Transposed Attention (MDTA) [40],

$$\Phi = \mathbf{W}_1^{\Phi}\text{Cat}\big(\text{MDTA}(\mathbf{Q}, \mathbf{K}, \mathbf{V}), \mathbf{Q}\big), \tag{5}$$

$$\text{where MDTA}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \mathbf{W}_1^O\big(\text{Softmax}(\overline{\mathbf{q}}\,\overline{\mathbf{k}}^\top/\alpha)\mathbf{v} + \mathbf{q}\big), \tag{6}$$

where α is a learnable temperature factor and $\text{Cat}(\cdot,\cdot)$ is the channel-wise concatenation operator. Consequently, the final output of the matching module is the sum of the outputs of channel and spatial matching modules.

Decoder. The decoder of our model receives results from the matching module for each hierarchical level. Subsequently, our decoder constructs the degradation pattern of the query image, utilizing the matched results obtained at each hierarchical level. After merging all the matched results, the output is further refined by a refinement block and a convolution layer. The refined output is subtracted from $\mathbf{X^Q}$, followed by a final convolution layer. We adopt the hierarchical decoder architecture presented in Chen *et al.* [2] as our decoder due to its seamless compatibility with our hierarchical architecture.

Training Strategy. We employ a conventional episodic meta-learning protocol to obtain generalized knowledge, following the training strategy of the universal

meta-learning framework [11]. We build general knowledge during the meta-train phase and then adapt the model to the given weather condition during the meta-test phase. Specifically, during the meta-train phase, we imitate the meta-test phase by partitioning batch data into query images and support set images. All parameters of our model are trained end-to-end during the meta-train phase.

Once the meta-training is complete, we adapt our architecture to unseen weather types during the meta-test phase. In the meta-test phase, we split the support set into two groups: half of the data serves as the query set, while the other half serves as the support set. With each iteration, we exchange the roles of the query and support set. We train only bias parameters of our model to preserve the knowledge acquired in the meta-train phase.

4 Experiments

4.1 Experimental Settings

Baselines. We compare our method with state-of-the-art generic image restoration methods. MPRNet [42], Restormer [40], MAXIM [31], NAFNet [2], and FocalNet [3] are initially proposed to be trained independently on individual specific degradation types. However, since they share the same architecture across tasks, they can seamlessly perform on our setting. AirNet [13], TransWeather [33], and WeatherDiffusion [21] are designed to be jointly trained with multiple types of degradation patterns, making their methods suitable to be our baseline models as well. Note that **all baselines are fine-tuned** with the same number of support set as ours during the meta-test phase.

Evaluation Metrics. We compare our method with the benchmark methods on PSNR (dB) and SSIM, following Zamir *et al.* [40] and Valanarasu *et al.* [33]. Higher values for both metrics indicate improved performance.

Implementation Details. Our method is implemented on the PyTorch framework, with training and testing performed on an NVIDIA RTX 3090 GPU. We use the Swin Transformer Base (Swin-B) model as the backbone for our encoder. We set the numbers of the attention heads in the matching modules as [4,8,16,16], from the top layer to the bottom layer. In the meta-train phase, we set the learning rate for the encoder as 10^{-5} and for the other parts of our model as 10^{-4}. During the meta-test phase, the learning rate for the bias parameters is set to 10^{-6}. We use the AdamW optimizer [20] with $\beta_1 = 0.9$ and $\beta_2 = 0.999$. An L1 loss function is used for both the meta-train and meta-test phases. In the meta-train phase, we set the batch size to 8, while the batch size is set to equal the number of few-shot images in the meta-test phase. In all experiments, the number of images in the support set is set to 1 by default. All images are resized to 224×224 pixels. The models are trained for 300k iterations during the meta-training phase and 20k iterations during the meta-test phase.

Table 1. Quantitative comparison on BID Task II.A dataset [7]. The **best** and second-best results are highlighted.

Model	GFLOPs	R+S		R+H_L		R+H_H		R+D+H_M		R+D+S+H_M		Average	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
MPRNet [42]	840.1	25.96	0.8397	22.75	0.8572	18.02	0.7291	18.70	0.7391	18.57	0.6682	20.80	0.7667
TransWeather [33]	**9.6**	24.27	0.7869	22.89	0.8394	18.20	0.7348	19.63	0.7365	19.45	0.6412	20.89	0.7478
AirNet [13]	230.7	25.34	0.7631	21.25	0.7704	17.00	0.6738	17.50	0.6404	18.42	0.5908	19.90	0.6877
MAXIM [31]	93.6	26.44	0.8086	21.11	0.8298	17.51	0.7054	15.49	0.6519	17.78	0.6270	19.67	0.7245
Restormer [40]	215.9	26.39	0.8501	22.81	0.8450	15.43	0.6733	19.92	0.7602	16.75	0.6713	20.26	0.7600
NAFNet [2]	24.5	26.82	0.8393	22.40	0.8508	17.78	0.7207	19.87	0.7602	16.95	0.6609	20.76	0.7664
WeatherDiffusion [21]	6415.5	25.54	0.8431	23.17	0.8077	21.17	0.7720	20.64	0.7623	20.80	**0.7735**	22.26	0.7917
FocalNet [3]	46.8	25.85	0.8164	23.47	0.8604	19.28	0.7566	20.46	0.7663	18.57	0.6707	21.53	0.7741
MetaWeather	16.1	**27.88**	**0.8574**	**25.85**	**0.8790**	**22.02**	**0.8137**	**23.24**	**0.8142**	**22.42**	0.7575	**24.28**	**0.8244**

4.2 Experiments on Co-occurring Weather Types

Datasets. In the meta-train phase, we train our proposed framework, MetaWeather, on a combination of basic weather-type datasets. And in the meta-test phase, we adapt our framework to datasets consisting of composited weather patterns. As a meta-train dataset, we use multiple datasets each of which includes one of the following degradation patterns: rain, snow, fog, and raindrop. The datasets are Rain13K dataset [42] for rain, Snow100K dataset [18] for snow, Foggy Cityscapes [28] for fog, Raindrop dataset [23] for raindrop.

We use the BID Task II.A dataset [7] for the meta-test datasets. The dataset has several different combinations of various synthetic weather degradation patterns. As shown in Fig. 1, the combination of basic weather types can be considered as a distinct weather type due to the noticeable decrease in the performance of existing methods. Each case contains a combination of weather degradation patterns as follows: rain streak + snow (R+S); rain streak + light haze (R+H_L); rain streak + heavy haze (R+H_H); rain streak + raindrop + moderate haze (R+D+H_M); rain streak + raindrop + snow + moderate haze (R+D+S+H_M).

Results. Table 1 compares the performance of our method with the baselines. MetaWeather achieves 2.02 dB performance increase in PSNR and 0.0327 in SSIM on average, compared to the second-best model, WeatherDiffusion, with the second best computational efficiency. Although TransWeather, AirNet, and WeatherDiffusion are proposed to deal with multiple weathers by training jointly with multiple datasets, they exhibit limited adaptability toward unseen weather types, whereas ours flexibly adapts to the weather conditions.

Figure 4 provides a qualitative comparison with the top five baselines in Table 1. Despite the mitigation of adverse weather artifacts by other state-of-the-art models, it is evident that degradation patterns are still present. In the case of R+S, the baseline models fail to remove the majority of adverse weather-related artifacts, particularly the rain streaks, whereas MetaWeather has removed most artifacts. Although WeatherDiffusion achieves the second-best performance in Table 1, it barely removes the rain streaks and snow particles compared to ours. Furthermore, as shown in cases of R+H_H, R+D+H_M,

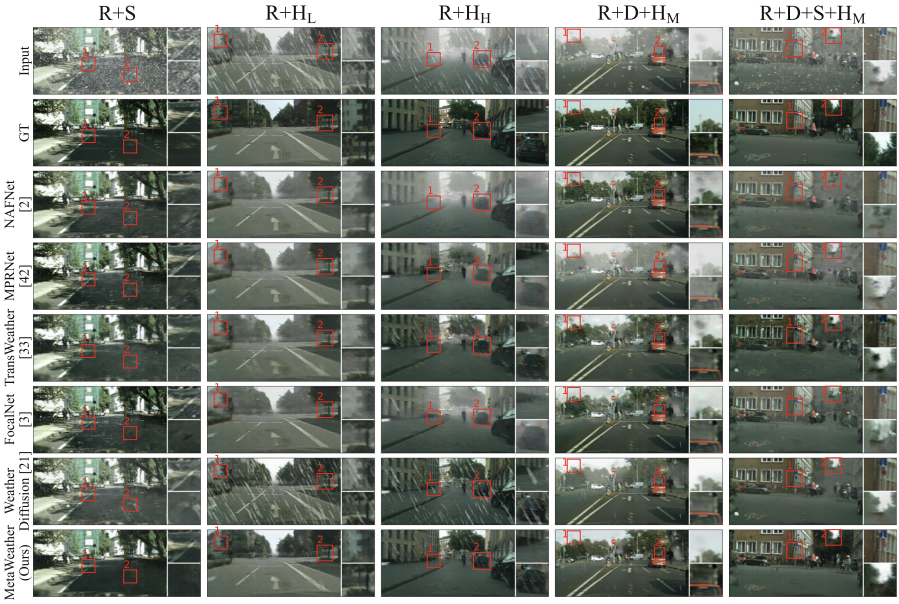


Fig. 4. Qualitative comparison on BID Task II.A dataset [7]. The results of top five baselines in Table 1 are sampled. Best viewed with zoom and color.

and R+D+S+H_M, existing models have limited ability to cope with the hazy effect, while MetaWeather removes almost all degradation patterns. In conclusion, these results indicate that MetaWeather enables flexible adaptation towards unseen new weather types.

4.3 Experiments on Real-World Scenario

Datasets. To further validate the effectiveness of our approach in real-world scenarios, we adapt and evaluate our method to the SPA-Data [36] (rain+veiling effect) and the RealSnow dataset [45] (snow+veiling effect), where the veiling effects naturally emerge due to the accumulation of weather patterns. The ground-truth images in these real-world datasets are obtained by applying a median filter over the frames to extract non-degraded parts of the images. More details are described in the supplementary material.

Baselines. In the real-world evaluation, we include the method proposed by Patil *et al.* [22] alongside the baselines used in Sect. 4.2. This method is intended for utilization in real-world weather-degraded image restoration only with pre-trained knowledge. Consequently, we assess this method in our real-world experiments without additional adaptation.

Table 2. Quantitative comparison on SPA-Data [36] and RealSnow [45] dataset. † denotes a method for real-world weather-degraded image restoration. The **best** results are highlighted.

Model	SPA-Data [35]		RealSnow [45]	
	PSNR	SSIM	PSNR	SSIM
MPRNet [42]	26.57	0.8355	26.84	0.8773
TransWeather [33]	26.62	0.8508	24.38	0.8354
AirNet [13]	27.72	0.8430	21.69	0.8244
MAXIM [31]	32.21	0.9002	27.30	0.8930
Restormer [40]	31.18	0.9089	29.65	0.9037
NAFNet [2]	32.16	0.9144	28.19	0.9014
WeatherDiffusion [21]	31.18	0.9014	26.32	0.8779
FocalNet [3]	29.03	0.8846	24.81	0.8527
Patil *et al.*† [22]	20.26	0.8291	20.22	0.8974
MetaWeather (Ours)	**34.81**	**0.9394**	**30.15**	**0.9222**

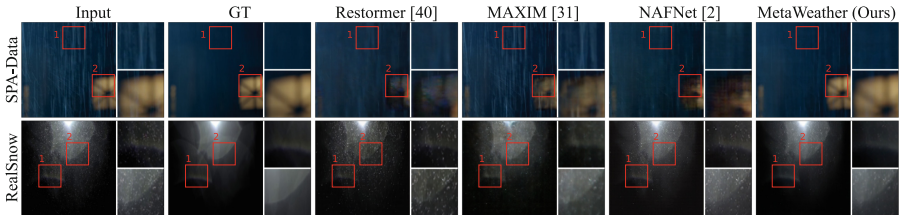


Fig. 5. Qualitative comparison on SPA-Data [36] and RealSnow [45] dataset. The results of top three baselines in Table 2 are sampled. Best viewed with zoom and color.

Results. As shown in Table 2, our method outperforms the baselines, similar to the previous results on unseen weather types, demonstrating its effectiveness in real-world scenarios as well. Surprisingly, although Patil *et al.* [22] is proposed for real-world weather degraded image restoration task, their method exhibits much lower performance on both datasets compared to the baselines and our approach, indicating the necessity of adaptation.

As depicted in Fig. 5, We qualitatively compare ours with the top three baselines in Table 2. While the baselines barely remove the rain streaks and snow particles, ours removes them almost perfectly. Additionally, while others distort the background parts, ours preserves the background, only removing the degradation patterns.

4.4 Ablation Study

Matched Features in Matching Modules. Table 3 shows the effect of the matched features within matching modules. We maintain the same architecture and training strategy for this experiment but change the matched features in the matching module to the background of the query image.

Table 3. Ablation study on matched features in matching modules. The **best** results are highlighted.

Matched features	R+S		R+H_L		R+H_H		R+D+H_M		R+D+S+H_M		Average	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Background	26.89	0.8220	23.41	0.8181	20.11	0.7436	21.23	0.7440	18.92	0.6567	22.11	0.7569
Degradation pattern	**27.88**	**0.8574**	**25.85**	**0.8790**	**22.02**	**0.8137**	**23.24**	**0.8142**	**22.42**	**0.7575**	**24.28**	**0.8244**

Table 4. Ablation study on spatial-channel matching. The **best** and second-best results are highlighted.

Spatial matching	Channel matching	R+S		R+H_L		R+H_H		R+D+H_M		R+D+S+H_M		Average	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
–	–	27.12	0.8369	24.79	0.8534	20.06	0.7611	21.41	0.7613	20.71	0.6997	22.82	0.7825
cre ✓	–	27.23	0.8282	23.75	0.8314	20.41	0.7675	22.33	0.7798	21.52	0.7088	23.05	0.7831
–	✓	27.85	0.8570	24.66	0.8696	**22.17**	0.8098	22.52	0.8012	22.31	0.7529	23.90	0.8181
✓	✓	**27.88**	**0.8574**	**25.85**	**0.8790**	22.02	**0.8137**	**23.24**	**0.8142**	**22.42**	**0.7575**	**24.28**	**0.8244**

Matching the degradation patterns significantly outperforms matching the background, with 2.17 dB in PSNR and 0.0675 in SSIM on average. This is because the predominant common factor across the few-shot sets is the degradation pattern, not the background, as mentioned in Sect. 3.2. In conclusion, learning to match the degradation pattern makes it much easier to generalize the representation of new weather effects compared to matching the background in few-shot weather degraded image restoration tasks.

Spatial-Channel Matching. Table 4 shows the impact of spatial and channel-wise matching in the matching modules. The use of matching mechanisms facilitates flexible adaptation in few-shot weather degraded image restoration tasks. The performance gap between the models increases as degradation patterns accumulate, suggesting that matching plays a key role in flexible adaptation.

However, spatial matching shows a limited performance increase, which is similar to the framework [11] we extend, showing a lack in restoring the veiling effects. Using channel matching outperforms using spatial matching solely, implying that leveraging channel-wise information is crucial in weather degraded image restoration tasks. Moreover, when both spatial and channel matching are used together, the performance improves further, indicating that matching in both dimensions is beneficial for learning the weather effect.

Figure 6 illustrates the effectiveness of spatial and channel-wise matching. For qualitative comparison on spatial-channel matching, we evaluate the variants of matching modules on the R+H_L weather type in BID Task II.A dataset. Figure 6(b) shows that using only spatial matching restores the rain streaks, while foggy degradation patterns still remain. Conversely, Fig. 6(c) illustrates that channel matching effectively removes the fog effect while the rain streaks still remain. Lastly, Fig. 6(d) depicts that utilizing spatial and channel matching together removes both degradation patterns. This result implies that matching in

(a) Input (b) Spatial Matching (c) Channel Matching (d) Spatial-Channel Matching

Fig. 6. Qualitative comparison on the effect of spatial-channel matching. Best viewed with zoom and color.

Table 5. Ablation study with respect to the VTM [11]. The **best** results are highlighted.

Method	R+S		R+H_L		R+H_H		R+D+H_M		R+D+S+H_M		Average	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
VTM [11]	26.99	0.8275	24.15	0.8310	19.15	0.7309	20.65	0.7388	20.03	0.6821	22.19	0.7621
+ Degradation pattern matching	27.23	0.8282	23.75	0.8314	20.41	0.7675	22.33	0.7798	21.52	0.7088	23.05	0.7831
+ Channel matching (**MetaWeather**)	**27.88**	**0.8574**	**25.85**	**0.8790**	**22.02**	**0.8137**	**23.24**	**0.8142**	**22.42**	**0.7575**	**24.28**	**0.8244**

each dimension effectively handles the weather effects in a decomposed manner, as described in Sect. 3.2.

Comparison with Visual Token Matching. We compare our proposed method to the universal meta-learning framework for dense prediction that we extend, named Visual Token Matching (VTM) [11]. For a fair comparison, we use the same encoder and decoder as ours to tailor the VTM to the task of weather-degraded image restoration. The matching modules only perform spatial matching of the background features to produce the background of the query image. In Table 5, our method significantly surpasses VTM in all cases due to the matching of degradation patterns instead of the background and the channel-wise matching. The consistent improvement in performance demonstrates that the proposed methods help the model to fully exploit the representation of new weather effects.

The Number of Shots in the Support Set. We vary the number of shots in the support set used in the adaptation phase, as presented in Table 6. We compare our method with the top three models in Table 1. Our method achieves the best average PSNR and SSIM regardless of the number of support set, indicating that our method is more effective at utilizing and generalizing knowledge extracted from a small number of images. Also, as the number of support set increases, the performance gap between our method and the others increases, while other baselines show limited performance improvements. This demonstrates the scalability and flexibility of our approach compared to baselines.

Robustness on Support Set Selection. Since our method utilizes a given support set during adaptation and inference, we verify whether our method is robust to the selection of the support set. To probe the robustness, we conduct the same experiments using five randomly selected support sets on all unseen

Table 6. Ablation study on the number of shots in the support set. The **best** and second-best results for each number of shots are highlighted.

# of shots	Model	R+S		$R+H_L$		$R+H_H$		$R+D+H_M$		$R+D+S+H_M$		Average	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
1	WeatherDiffusion [21]	25.54	0.8432	23.12	0.8066	21.17	0.7724	20.70	0.7646	20.81	**0.7734**	22.27	0.7920
	FocalNet [3]	25.85	0.8164	23.47	0.8604	19.28	0.7566	20.46	0.7663	18.57	0.6707	21.53	0.7741
	MetaWeather (Ours)	**27.88**	**0.8574**	**25.85**	**0.8790**	**22.02**	**0.8137**	**23.24**	**0.8142**	**22.42**	0.7575	**24.28**	**0.8244**
2	WeatherDiffusion [21]	25.54	0.8431	23.17	0.8077	21.17	0.7720	20.64	0.7623	20.80	0.7735	22.26	0.7917
	FocalNet [3]	26.58	0.8358	24.51	0.8636	20.16	0.7677	20.62	0.7599	18.79	0.6909	22.13	0.7836
	MetaWeather (Ours)	**28.54**	**0.8728**	**25.77**	**0.8876**	**22.58**	**0.8281**	**22.80**	**0.8188**	**22.16**	**0.7775**	**24.37**	**0.8370**
4	WeatherDiffusion [21]	25.49	0.8428	23.18	0.8077	21.19	0.7720	20.60	0.7616	20.80	0.7743	22.25	0.7917
	FocalNet [3]	26.46	0.8154	24.35	0.8719	19.76	0.7757	20.93	0.7815	20.76	0.7097	22.45	0.7908
	MetaWeather (Ours)	**28.62**	**0.8780**	**26.17**	**0.8928**	**23.40**	**0.8422**	**23.59**	**0.8349**	**22.99**	**0.7937**	**24.95**	**0.8483**

Table 7. Robustness on the selection of the support set. All performance numbers are averaged across cases in the BID Task II.A dataset for each selected support set.

Support set	# 1	# 2	# 3	# 4	# 5	Mean	Std.
PSNR	24.28	24.35	23.83	24.20	24.12	24.16	0.20
SSIM	0.8244	0.8265	0.8226	0.8256	0.8242	0.8247	0.0018

weather types in BID Task II.A dataset. Table 7 clearly shows that our method is robust to the selection of the support set, with tiny standard deviations in both PSNR and SSIM. This robustness stems from our method's ability to effectively extract and generalize representations obtained from a few-shot support set.

5 Conclusion

This paper presents MetaWeather, a few-shot learning method capable of restoring images degraded by unseen weather conditions. To fully exploit the given few-shot support set, we extend the matching-based meta-learning framework to predict weather degradation patterns with a novel matching algorithm. We guide our model to match the degradation patterns instead of the backgrounds, considering that the degradation patterns are more common across a weather than the backgrounds. We devise spatial-channel matching to effectively extract the representation from the support images, reflecting the decomposable attribute of the weather effects. Extensive experiments on co-occurring and real-world unseen weather types show that our method effectively handles unseen weather conditions with a few labeled samples, demonstrating its flexibility and applicability towards the real world.

Acknowledgment. This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.under grant RS-2019-II191126, Self-learning based Autonomic IoT Edge Computing).

References

1. Ba, Y., et al.: Not just streaks: towards ground truth for single image deraining. In: Avidan, S., Brostow, G., Cissé, M., Farinella, G.M., Hassner, T. (eds.) ECCV 2022. LNCS, vol. 13667, pp. 723–740. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-20071-7_42
2. Chen, L., Chu, X., Zhang, X., Sun, J.: Simple baselines for image restoration. In: Avidan, S., Brostow, G., Cissé, M., Farinella, G.M., Hassner, T. (eds.) ECCV 2022. LNCS, vol. 13667, pp. 17–33. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-20071-7_2
3. Cui, Y., Ren, W., Cao, X., Knoll, A.: Focal network for image restoration. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 13001–13011 (2023)
4. Finn, C., Abbeel, P., Levine, S.: Model-agnostic meta-learning for fast adaptation of deep networks. In: Precup, D., Teh, Y.W. (eds.) Proceedings of the 34th International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 70, pp. 1126–1135. PMLR (2017)
5. Gao, X., Wang, Y., Cheng, J., Xu, M., Wang, M.: Meta-learning based relation and representation learning networks for single-image deraining. Pattern Recogn. **120**, 108124 (2021)
6. Gatys, L.A., Ecker, A.S., Bethge, M.: Image style transfer using convolutional neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2414–2423 (2016)
7. Han, J., et al.: Blind image decomposition. In: Avidan, S., Brostow, G., Cissé, M., Farinella, G.M., Hassner, T. (eds.) ECCV 2022. LNCS, vol. 13678, pp. 218–237. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-19797-0_13
8. Hassaballah, M., Kenk, M.A., Muhammad, K., Minaee, S.: Vehicle detection and tracking in adverse weather using a deep learning framework. IEEE Trans. Intell. Transp. Syst. **22**(7), 4230–4242 (2021)
9. Huang, S.C., Le, T.H., Jaw, D.W.: DSNet: joint semantic learning for object detection in inclement weather conditions. IEEE Trans. Pattern Anal. Mach. Intell. **43**(8), 2623–2633 (2020)
10. Kang, B., Liu, Z., Wang, X., Yu, F., Feng, J., Darrell, T.: Few-shot object detection via feature reweighting. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 8420–8429 (2019)
11. Kim, D., Kim, J., Cho, S., Luo, C., Hong, S.: Universal few-shot learning of dense prediction tasks with visual token matching. In: The Eleventh International Conference on Learning Representations (2023)
12. Lee, S., Son, T., Kwak, S.: FIFO: learning fog-invariant features for foggy scene segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 18911–18921 (2022)
13. Li, B., Liu, X., Hu, P., Wu, Z., Lv, J., Peng, X.: All-in-one image restoration for unknown corruption. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 17452–17462 (2022)
14. Li, R., Cheong, L.F., Tan, R.T.: Heavy rain image restoration: integrating physics model and conditional adversarial learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1633–1642 (2019)
15. Li, R., Tan, R.T., Cheong, L.F.: All in one bad weather removal using architectural search. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3175–3185 (2020)

16. Liu, H., Wu, Z., Li, L., Salehkalaibar, S., Chen, J., Wang, K.: Towards multi-domain single image dehazing via test-time training. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 5821–5830. IEEE (2022)
17. Liu, W., Ren, G., Yu, R., Guo, S., Zhu, J., Zhang, L.: Image-adaptive YOLO for object detection in adverse weather conditions. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 36, no. 2, pp. 1792–1800 (2022)
18. Liu, Y.F., Jaw, D.W., Huang, S.C., Hwang, J.N.: DesnowNet: context-aware deep network for snow removal. IEEE Trans. Image Process. **27**(6), 3064–3073 (2018)
19. Liu, Z., et al.: Swin transformer: hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 10012–10022 (2021)
20. Loshchilov, I., Hutter, F.: Decoupled weight decay regularization. arXiv preprint arXiv:1711.05101 (2017)
21. Özdenizci, O., Legenstein, R.: Restoring vision in adverse weather conditions with patch-based denoising diffusion models. IEEE Trans. Pattern Anal. Mach. Intell. (2023)
22. Patil, P.W., Gupta, S., Rana, S., Venkatesh, S., Murala, S.: Multi-weather image restoration via domain translation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 21696–21705 (2023)
23. Qian, R., Tan, R.T., Yang, W., Su, J., Liu, J.: Attentive generative adversarial network for raindrop removal from a single image. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2482–2491 (2018)
24. Qian, R., Tan, R.T., Yang, W., Su, J., Liu, J.: Attentive generative adversarial network for raindrop removal from a single image. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2482–2491. IEEE (2018)
25. Rai, S.N., Saluja, R., Arora, C., Balasubramanian, V.N., Subramanian, A., Jawahar, C.: Fluid: few-shot self-supervised image deraining. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 3077–3086 (2022)
26. Ran, W., Yuan, W., Shibasaki, R.: Few-shot depth completion using denoising diffusion probabilistic model. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 6558–6566 (2023)
27. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
28. Sakaridis, C., Dai, D., Van Gool, L.: Semantic foggy scene understanding with synthetic data. Int. J. Comput. Vision **126**, 973–992 (2018)
29. Shaban, A., Bansal, S., Liu, Z., Essa, I., Boots, B.: One-shot learning for semantic segmentation. In: Proceedings of the British Machine Vision Conference, pp. 167.1–167.13. BMVA Press (2017)
30. Tremblay, M., Halder, S.S., de Charette, R., Lalonde, J.F.: Rain rendering for evaluating and improving robustness to bad weather. Int. J. Comput. Vision **129**(2), 341–360 (2021)
31. Tu, Z., et al.: Maxim: multi-axis MLP for image processing. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 5769–5780 (2022)
32. Ulyanov, D., Vedaldi, A., Lempitsky, V.: Instance normalization: the missing ingredient for fast stylization. arXiv preprint arXiv:1607.08022 (2016)

33. Valanarasu, J.M.J., Yasarla, R., Patel, V.M.: TransWeather: transformer-based restoration of images degraded by adverse weather conditions. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2353–2363 (2022)
34. Vinyals, O., Blundell, C., Lillicrap, T., Wierstra, D., et al.: Matching networks for one shot learning. In: Proceedings of Advances in Neural Information Processing Systems, vol. 29 (2016)
35. Wang, K., Liew, J.H., Zou, Y., Zhou, D., Feng, J.: PANet: few-shot image semantic segmentation with prototype alignment. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 9197–9206 (2019)
36. Wang, T., Yang, X., Xu, K., Chen, S., Zhang, Q., Lau, R.W.: 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, pp. 12262–12271. IEEE (2019)
37. Wang, Z., Cun, X., Bao, J., Zhou, W., Liu, J., Li, H.: UFormer: a general U-shaped transformer for image restoration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 17683–17693 (2022)
38. Xie, Z., et al.: SimMIM: a simple framework for masked image modeling. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 9653–9663 (2022)
39. Ye, T., et al.: Adverse weather removal with codebook priors. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 12653–12664 (2023)
40. Zamir, S.W., Arora, A., Khan, S., Hayat, M., Khan, F.S., Yang, M.H.: Restormer: efficient transformer for high-resolution image restoration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 5728–5739 (2022)
41. Zamir, S.W., et al.: Learning enriched features for real image restoration and enhancement. In: Vedaldi, A., Bischof, H., Brox, T., Frahm, J.-M. (eds.) ECCV 2020, Part XXV. LNCS, vol. 12370, pp. 492–511. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-58595-2_30
42. Zamir, S.W., et al.: Multi-stage progressive image restoration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 14821–14831 (2021)
43. Zhao, H., Gou, Y., Li, B., Peng, D., Lv, J., Peng, X.: Comprehensive and delicate: an efficient transformer for image restoration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 14122–14132 (2023)
44. Zhou, M., Huang, J., Guo, C.L., Li, C.: Fourmer: an efficient global modeling paradigm for image restoration. In: International Conference on Machine Learning, pp. 42589–42601. PMLR (2023)
45. Zhu, Y., et al.: 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, pp. 21747–21758 (2023)