Revitalizing Real Image Deraining via a Generic Paradigm towards Multiple Rainy Patterns

Xin Li[1] , **Yuxin Feng**[1] , **Fan Zhou**[1,2] , **Yun Liang**[3] and **Zhuo Su**[1,2,*]
[1]School of Computer Science and Engineering, National Engineering Research Center of Digital Life, Sun Yat-sen University
[2]Research Institute of Sun Yat-sen University in Shenzhen
[3]College of Mathematics and Informatics, South China Agricultural University
suzhuo3@mail.sysu.edu.cn

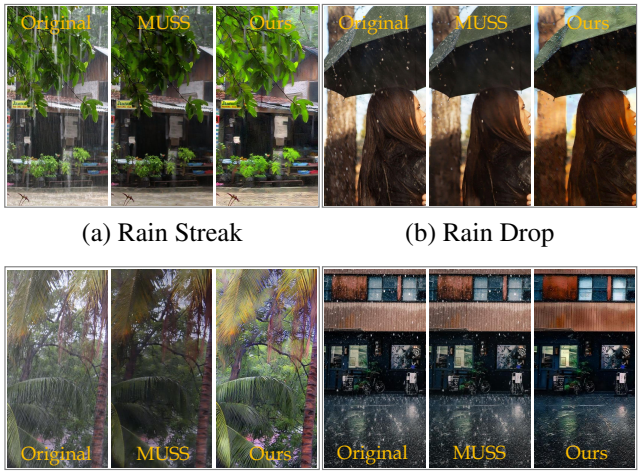
Abstract

Synthetic data-driven methods perform well on image rain removal task, but they still face many challenges in real rainfall scenarios due to the complexity and diversity of rainy patterns. In this paper, we propose a new generic paradigm for real image deraining from the perspective of synthesizing data covering more rainy patterns and constructing image rain removal networks with strong generalization performance. Firstly, instead of simply superimposing rain layers, we integrate various rainy patterns and design a phenomenal pipeline that incorporates multiple degradation types. Secondly, we construct a Patterns-aware Rain Removal Network (PRRN), which learns from both synthetic and real data simultaneously. In addition, to eliminate the inevitable distribution differences between synthetic and real data, we design a new Multi-representation Inter-domain Alignment Module (MIAM) in PRRN. By using multiple parallel submodules, MIAM achieves alignment of data domains in multiple feature subspaces. Based on several authoritative objective evaluation metrics, we successfully validate the effectiveness and robustness of the proposed method in real scenarios through extensive experiments carried out on five challenging real datasets.

1 Introduction

Due to the variety and intricacy of real rainfall scenarios and the ill-posed nature of rain removal problem, existing image rain removal methods tend to focus on specific rainy patterns only, and conduct research on a single or several synthetic rain datasets. However, as shown in Figure 1, real rainfall scenarios usually contain a variety of different rainy patterns, such as rain streak, rain drop, rain haze and rain block, and there is still a large research gap in the current work addressing multiple rainy patterns. Next, we comprehensively introduce some representative work in existing rain removal methods according to different types of supervision, as well as the main contributions made by this paper.

*Corresponding author: Zhuo Su



(a) Rain Streak (b) Rain Drop

(c) Rain Haze (d) Rain Block

Figure 1: Multiple rainy patterns in real scenarios. (a)-(d) show four common rainy patterns, namely, rain streak, rain drop, rain haze and rain block, and deraining results of both PRRN and MUSS [Huang *et al.*, 2023] methods on different rainy patterns.

1.1 Related Work

Synthetic Data-driven Method. Based on supervised learning of synthetic data pairs, image rain removal methods have achieved good deraining performance by designing elaborate network structures [Fu *et al.*, 2017b; Li *et al.*, 2018] and combining relevant priors [Chen and Hsu, 2013; Sun *et al.*, 2014] of rain images.

Depending on whether deraining results are generated directly or not, researchers subdivide synthetic data-driven rain removal methods into end-to-end methods [Guo *et al.*, 2022; Jiang *et al.*, 2022] and rain image model-based methods [Hu *et al.*, 2019; Gao *et al.*, 2023]. End-to-end methods utilise deep neural networks to extract and separate key features of the rain and background layers separately, enabling the generation of clean derained images directly from the background features. Model-based methods obtain the final derained image by designing a model close to the composition of real rain images and generating components that make up the model one by one.

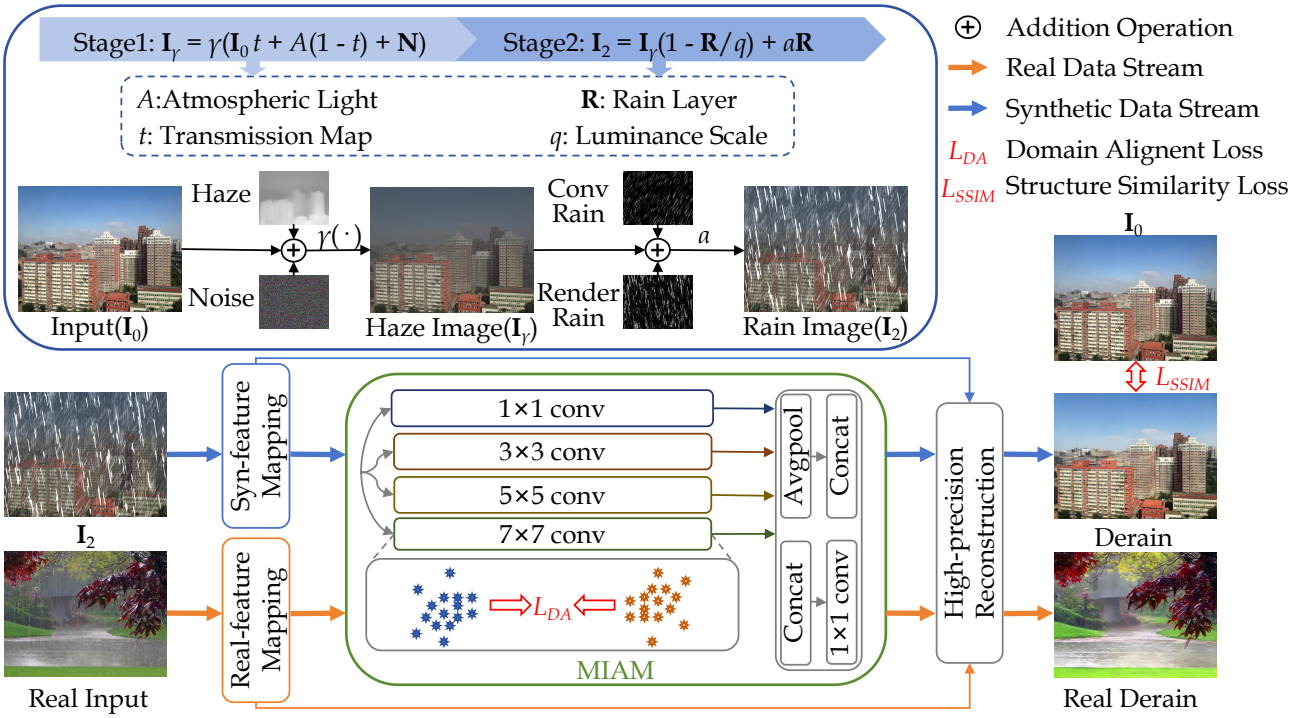


Figure 2: The new generic real image rain removal paradigm comprises two parts: the multi-degradation type data synthesis pipeline and the Patterns-aware Rain Removal Network (PRRN). The blue box depicts the overall framework of the data generation pipeline, consisting of both a haze generation stage and a rain layer superposition stage. The green box represents the Multi-representation Inter-domain Alignment Module (MIAM) in the PRRN, which is comprised of four parallel submodules with different receptive field sizes.

The deraining performance of these fully supervised methods depends largely on the synthesized dataset used. Although they have achieved excellent performance on synthesized data, their effectiveness is greatly compromised when there is a large data distribution difference between the used data and the real data.

Real Data-driven Method. In response to the inability of synthetic rain datasets to cover a wide range of complex rainy patterns, researchers have begun to introduce real data to enhance the generalization ability of image rain removal methods. According to whether synthetic data is used in the training process, real data-driven rain removal methods can be divided into semi-supervised methods [Cui *et al.*, 2022; Huang *et al.*, 2023] and unsupervised methods [Li *et al.*, 2019a; Wei *et al.*, 2021].

Semi-supervised methods often first learn an initial network based on a large amount of synthetic data, and then improve the generalization performance of the network by introducing real data and reducing the difference in feature distribution between different types of data. Unsupervised methods mostly employ the idea of generative adversarial networks [Radford *et al.*, 2015; Jin *et al.*, 2018]. The generator can not only obtain the derained image from the rain image, but also generate the corresponding rain image from the clean background to enrich the network's training data. Subsequently, the discriminator leverages the evaluation result to guide the generator's optimization direction.

Existing semi-supervised and unsupervised methods facilitate the task of image rain removal by effectively capturing rain information from real data. However, the insufficient amount of data mining has resulted in mediocre performance of these methods in rain removal.

1.2 Our Contributions

Faced with the problems of insufficient generalization ability under real scenarios and the inability to process diverse rainy patterns in existing image rain removal methods, we propose to explore them from both data and network levels. Our proposed method not only achieves effective response to multiple complex rainy patterns, but also achieves good generalization performance under multiple types of real scenarios. Our main contributions can be summarized as follows.

- We reconstruct a degradation model for real rain images, and propose a phenomenological degradation pipeline to simulate rain-world images in real scenes. The newly constructed synthetic dataset not only contains multiple rainy patterns, but also accounts for a variety of influences such as illumination and noise.
- A patterns-aware rain removal network is designed to fully leverage the global feature extraction capability of the self-attention mechanism and the local spatial exploration capability of the convolutional operation. Our network achieves effective alignment between the synthetic data domain and the real data domain at the feature level.

- A new generic paradigm for rain removal in real images is proposed. Extensive experiments on five challenging real-world datasets fully confirm the effectiveness and stability of our proposed approach in dealing with multiple rain patterns in real scenarios.

2 Data Preparation

Redesigning a more comprehensive data generation pipeline has proven to be an effective approach in multiple low-level vision tasks [Wang *et al.*, 2021; Meng *et al.*, 2023]. Therefore, we propose to construct a new real rain image degradation pipeline to solve the image rain removal problem at data level. The pipeline takes full account of multiple rainy patterns such as rain streak and rain drop produced by rainfall, as well as indirect effects of rainy weather on the global light and noise levels in the surrounding environment. Based on this pipeline, we synthesize a complex rain dataset MPRain which contains multiple rainy patterns.

As shown in the blue box in Figure 2, the data synthesis pipeline consists of a haze generation stage based on the atmospheric light scattering model and a rain layer superposition stage based on the rain image model. We separately describe each of these two stages below.

2.1 Haze Generation

In rainy weather, the increase of surface moisture and the decrease of temperature often form haze. Therefore, we first take into account the effect of haze when synthesizing rain images. Formula (1) shows the traditional atmospheric light scattering model for haze synthesis.

$$\mathbf{H} = \mathbf{I}_0 t + A(1-t), \tag{1}$$

where $\mathbf{I}_0$, $\mathbf{H}$, t, and A denote the input image, the generated haze image, the transmission map and the global atmospheric light value, respectively. On this basis, as shown in formula (2), we synthesize the haze image on the classical RESIDE dataset [Li *et al.*, 2017] by introducing tiny variables and considering environmental factors.

$$\mathbf{I}_\gamma = \gamma(\mathbf{I}_0 t + A(1-t) + \mathbf{N}), \tag{2}$$

here $\mathbf{I}_\gamma$, $\mathbf{N}$, and $\gamma(\cdot)$ indicate the generated haze image, ambient noise and Gamma correction operation respectively. Transmission map $t = e^{-\beta d}$ reflects the proportion of light that reaches the imaging device after particle attenuation, in which β is the haze density control parameter, d is the depth map obtained by the depth estimation method proposed in [Zhang *et al.*, 2023]. The specific details of haze image synthesis are shown below.

Transmission Map. During the construction of haze images, to control the general concentration of the haze, we initialize β to a constant 2 and add a disturbance factor β_1 with a variation range of [-0.5, 0.5) to change the intensity of the haze between different images. The adjusted expression for t is shown below:

$$t = e^{-(2+\beta_1)d}, \beta_1 \in [-0.5, 0.5). \tag{3}$$

Multi-interference Environment. Taking into account that rainy weather brings about cloud accumulation in the atmosphere, humidity rise, and water splashes, causing dimmer brightness and higher noise levels, we incorporate random noise $\mathbf{N}$ and Gamma correction into our data construction process to mimic these factors' effects on the imaging of hazy images. Since most of the natural noise is Gaussian noise, we first introduce Gaussian noise $\mathbf{N}$ with mean 0 and standard deviation 0.2 into the haze image, and then adjust the global color level of the image by Gamma correction.

2.2 Rain Layer Superimposition

In addition to inducing haze, rainfall also distorts images by introducing diverse rainy patterns with varying shapes, including rain streak, raindrop, rain block and so on. In the rain layer superposition stage, we adopt two rainy patterns generation approaches, convolution and rendering. The rain image generation process satisfies the rain image model [Hu *et al.*, 2019] in formula (4).

$$\mathbf{I}_2 = \mathbf{I}_\gamma(1 - \mathbf{R}/q) + \alpha\mathbf{R}, \tag{4}$$

in which $\mathbf{I}_2$, $\mathbf{I}_\gamma$, and $\mathbf{R}$ indicate the output rain image, the input haze image, and the rain layer with values in the range [0, 255], respectively. Both q and α are luminance control factors for the rain layer. Rain layer $\mathbf{R} = \mathbf{R}_p \cdot t_r$, here $\mathbf{R}_p$ denotes rain pattern layer containing multiple rainy patterns, t_r represents rain layer transmission map. The value of t_r is affected by the depth of the scene, as shown in formula (5), when the depth of the scene is greater than d_0, rainy patterns in the image are gradually blurred with the increase of depth.

$$t_r = e^{-cmax\{d_0, d\}}, \tag{5}$$

here c and d_0 are both constants that control the transparency of the $\mathbf{R}$ and the minimum depth at which blurring occurs, severally. The details of the synthesis of $\mathbf{R}_p$ are shown below.

Convolution Generation. Convolution operation combined with Gaussian filtering can synthesize sharp rain streaks with largely controllable factors such as position, length, width and direction. Combined with the observation that rain streaks are also denser in regions with greater depth of scene, we generate the rain streak layer $\mathbf{R}_s$ by stacking the rain streak layer by layer based on the depth map.

$$\mathbf{R}_s = \sum_k^K Streak(x) \cdot mask(k, x), \tag{6}$$

where $Streak(x)$ is the sub-rain-streak layer randomly generated based on the convolutional approach, K represents the number of superimposed sub-rain-streak layers and the value is a random integer of [1, 3], and $mask(k, x)$ is the mask of the kth sub-rain-streak layer. As shown in formula (7), $mask(\cdot, x)$ achieves a tight correlation between the density of rain streaks and the depth of the scene by gradually increasing the minimum depth at which rain streaks appear in the sub-rain-streak layer.

$$mask(k, x) = \begin{cases} 0, & d(x) < d_k, \\ 1, & d(x) \geq d_k, \end{cases} \tag{7}$$

in which $d_k = max\{d(x)\}/K * k, k \in [1, K]$ denotes the minimum depth at which rain streaks appear in the kth sub-rain-streak layer.

Dataset	Streak	Haze	Noise	LR	DV
Rain200L	✓	×	×	1.05	1,800/200
Rain200H	✓	×	×	1.52	1,800/200
Rain14000	✓	×	×	1.11	12,600/1,400
Rain-render	✓	✓	×	1.48	9,432/1,188
MPRain	✓	✓	✓	0.90	10,305/492

Table 1: Comparison of synthetic datasets. The table provides a detailed comparison of rainy patterns, environmental factors and data volumes across multiple synthetic datasets.

Render Generation. To ensure the similarity between synthesized rainy patterns and real rainy patterns, we also synthesize rainy patterns using the realism rendering method proposed in [Yang *et al.*, 2016]. The final rain pattern layer $\mathbf{R}_p$ is obtained by superimposing many different rendered rainy patterns at multiple scales and combining them with the convolutional synthetic rain streak layer $\mathbf{R}_s$.

2.3 Dataset Comparison

Table 1 provides a detailed comparison of differences in rainy patterns, environmental factors, and data volumes between MPRain and other four synthetic datasets: Rain200L, Rain200H [Yang *et al.*, 2016], Rain14000 [Fu *et al.*, 2017a], and Rain-render [Tremblay *et al.*, 2020]. The Luminance Ratio (LR) represents the brightness comparison between the rain image and the Ground Truth (GT) in the test set, while the Data Volume (DV) signifies the ratio between the number of training images and test images in the dataset.

As can be seen in Table 1, MPRain not only contains many rainy patterns, but also integrates a variety of factors such as noise and light in the environment. Compared with the GT, the synthetic rain image in the MPRain has a lower brightness, which is not only more in line with people's visual feeling on rainy days, but also prompts methods using MPRain to generate clearer and brighter deraining results.

3 Patterns-aware Rain Removal Network

To cope with distributional differences between synthetic and real data, we innovatively design a Patterns-aware Rain Removal Network, which employs a semi-supervised learning strategy and exhibits strong deraining performance on a wide range of types of real data. The overall framework of PRRN is shown in Figure 2, which primarily consists of three components: the self-attentive feature mapping module, the multi-representation inter-domain alignment module and the high-precision image reconstruction module, which are used for the extraction of image features, the feature alignment between different data domains and the reconstruction of derained images, respectively. Below, we introduce these three modules and the training process of PRRN in detail.

3.1 Self-attention Feature Mapping Module

The self-attentive feature mapping module consists of two parts: the shallow feature mapping part only uses a convolutional operation with kernel size of 3 × 3, and the deep feature extraction part contains four basic blocks connected in series. Considering the excellent performance of the self-attention mechanism [Feng *et al.*, 2023; Chen *et al.*, 2023b] in feature extraction, the multi-receptive field feature extraction module based on self-attention proposed in [Liu *et al.*, 2021; Liang *et al.*, 2022] is selected as the basic block. In addition, skip connections are used between basic blocks to ensure the correct transfer of shallow features in the network.

3.2 Multi-representation Inter-domain Alignment Module

The specific structure MIAM is shown as Figure 3, which comprises four parallel submodules. Four submodules are composed of 1, 1, 2 and 3 convolutional layers with kernel sizes of 1×1, 3×3, 5×5 and 7×7, respectively. The outputs of parallel submodules are simultaneously subjected to (1) average pooling and concatenation operations to obtain the hidden-layer code for feature alignment, and (2) concatenation and convolution operations to obtain the multi-representation output feature for image reconstruction. By employing a meticulous design of multiple submodules, MIAM effectively extracts multi-scale information from input features across various receptive fields. This not only enables the network to capture local information of variable sizes in diverse rainy patterns, but also ensures precise feature alignment in multiple feature subspaces.

3.3 High-precision Image Reconstruction Module

The high-precision image reconstruction module first processes the multi-representation output features using a 3×3 convolution layer. It then adds the resulting outputs to shallow features, and finally employs another 3×3 convolution layer to restore the added features to 3 channels, thus obtaining the final derained image. The addition of shallow features further ensures that the initially extracted important information is not lost during the propagation of the network. What is more, throughout the processing flow of PRRN, feature maps remain the same size as the original image, which effectively avoids information loss due to up and down sampling.

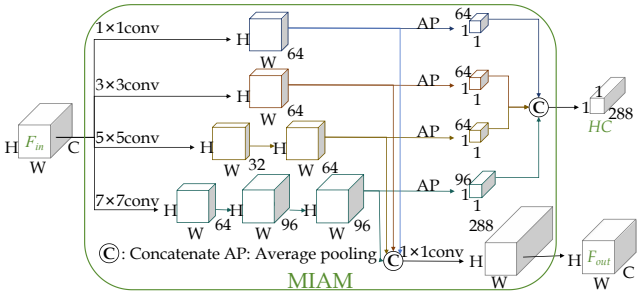


Figure 3: MIAM structure diagram. MIAM contains four parallel submodules, which carry out convolution processing on input feature (F_{in}) respectively to obtain four intermediate results. By performing different operations on these results, we obtain the Hidden-layer Code (HC) and the multi-presentation output feature (F_{out}), separately.

Method	DDN-SIRR		IVIPC_DQA		MPID		SPA		NTU	
-	Brisque↓	Entropy↑	Brisque↓	Entropy↑	Brisque↓	Entropy↑	Brisque↓	Entropy↑	Brisque↓	Entropy↑
Original	31.53	7.434	29.25	7.444	31.78	7.547	35.36	7.189	28.66	7.291
PReNet	24.63	7.418	26.09	7.411	29.26	7.547	31.29	7.156	24.72	7.322
MRFAN	28.61	7.422	25.91	7.426	30.45	7.545	30.49	7.157	25.18	7.288
IDT	27.80	7.413	26.89	7.426	30.98	7.544	31.81	7.154	24.97	7.312
DRS	27.80	7.401	26.00	7.307	31.73	7.538	33.47	7.053	27.34	7.288
AMRR	31.66	7.462	28.16	7.373	27.93	7.532	32.32	7.216	26.11	7.410
TRNR	26.15	7.413	25.48	7.413	29.99	7.544	30.92	7.157	24.08	7.344
UDR	32.38	7.338	35.79	7.357	35.33	7.512	38.60	7.093	28.31	7.302
SIRR	25.39	7.409	24.42	7.353	26.14	7.484	29.08	7.088	23.75	7.351
Syn2real	27.51	7.319	26.00	7.307	30.74	7.464	31.75	7.084	24.40	7.271
MOSS	27.17	7.378	28.89	7.392	30.57	7.523	32.64	7.127	25.57	7.340
SSID-KD	25.93	7.401	26.62	7.403	30.12	7.533	32.96	7.145	25.00	7.324
MUSS	26.26	7.229	26.39	7.302	29.94	7.465	30.66	7.088	26.63	7.355
Ours-teacher	27.00	7.558	24.94	7.544	25.40	7.682	29.47	7.359	25.98	7.593
Ours-student	24.57	7.596	23.71	7.536	24.86	7.707	29.24	7.404	19.70	7.627

Table 2: The performance of the proposed method is compared with several existing state-of-the-art rain removal methods on five real rain datasets. The evaluation metrics use non-reference metrics Brisque and Entropy. By comparison, red markers indicate the best performance, while blue markers indicate the second best performance.

3.4 Training Strategy and Loss Function

The teacher-student network [Su *et al.*, 2021; Cui *et al.*, 2022] is widely used in the field of image restoration due to its powerful ability to integrate synthetic and real data information. Based on the idea of the teacher-student network, PRRN is trained in two stages to obtain a teacher model and a student model, respectively, both of which share the same network architecture.

Stage 1. In the first stage, PRRN uses pairs of synthetic data to pretrain a teacher network. The network is trained on our proposed MPRain dataset for 520 epochs to achieve stable image restoration performance. The loss function $L_{teacher}$ for the teacher network during training is defined as the structure similarity loss L_{SSIM} between the derained image and the GT, as indicated in formula (8).

$$\begin{aligned} L_{SSIM}(x,y) &= 1 - SSIM(x,y) \\ &= 1 - \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \end{aligned} \tag{8}$$

where x, and y denote the GT and the derained image, μ_x, μ_y and σ_x, σ_y denote means and standard deviations of the GT and the derained image, severally. Both of C_1 and C_2 are constants taken as 0.0001 and 0.0009.

Stage 2. In the second stage, PRRN first initializes a student network using the trained teacher network. In the following training steps, the student network not only utilizes the synthetic data in the MPRain dataset, but also introduces 516 unpaired real rain images to improve the network's adaptability to real rainfall scenarios.

For the learning of synthetic data, the student network employs the same structural similarity loss L_{SSIM} as the teacher network. For the learning of real data, we design a new inter-domain alignment loss L_{DA} based on MIAM. As indicated in formula (9), the computation of L_{DA} is performed in each of four submodules of MIAM. Firstly, the KL divergence [Hou *et al.*, 2020] between the synthetic feature and the real feature in each submodule is independently calculated. Subsequently, L_{DA} is obtained by combining these individual calculations in equal proportion. The introduction of L_{DA} enables multi-scale alignment of synthetic and real data in multiple feature subspaces.

$$L_{DA} = \sum_{i=1}^{4} KL(HC(syn,i), HC(real,i)), \tag{9}$$

here $HC(\cdot,i), i \in [1,4]$ denotes the hidden-layer code extracted from the ith submodule of MIAM for the synthetic or real data. As in formula (10), the overall loss $L_{student}$ of the student network consists of structural similarity loss L_{SSIM} of the synthesized data and inter-domain alignment loss L_{DA} of the real data. The network is trained for 20 epochs under the dual guidance of these losses.

$$L_{student} = \lambda_1 L_{SSIM} + \lambda_2 L_{DA}, \tag{10}$$

in which λ_1 and λ_2 are constants, taken as 20 and 0.001, severally. Eventually, the total loss L_{total} of PRRN is represented in formula (11).

$$L_{total} = L_{teacher} + L_{student}. \tag{11}$$

4 Experimental Results

The experiment is structured into two parts. The first part compares the deraining performance of the proposed method with existing state-of-the-art rain removal methods across various real-world rainfall scenarios. It comprehensively verifies the effectiveness and generalization of the proposed method in authentic conditions. In the second part, ablation experiments are conducted on the individual modules and loss function of the proposed method. This analysis aims to validate the efficacy of each module as well as the design of the loss function.

4.1 Implementation Details

The network is trained on RTX 4090 with initial learning rates of 0.0005 for the teacher network and 0.0001 for the

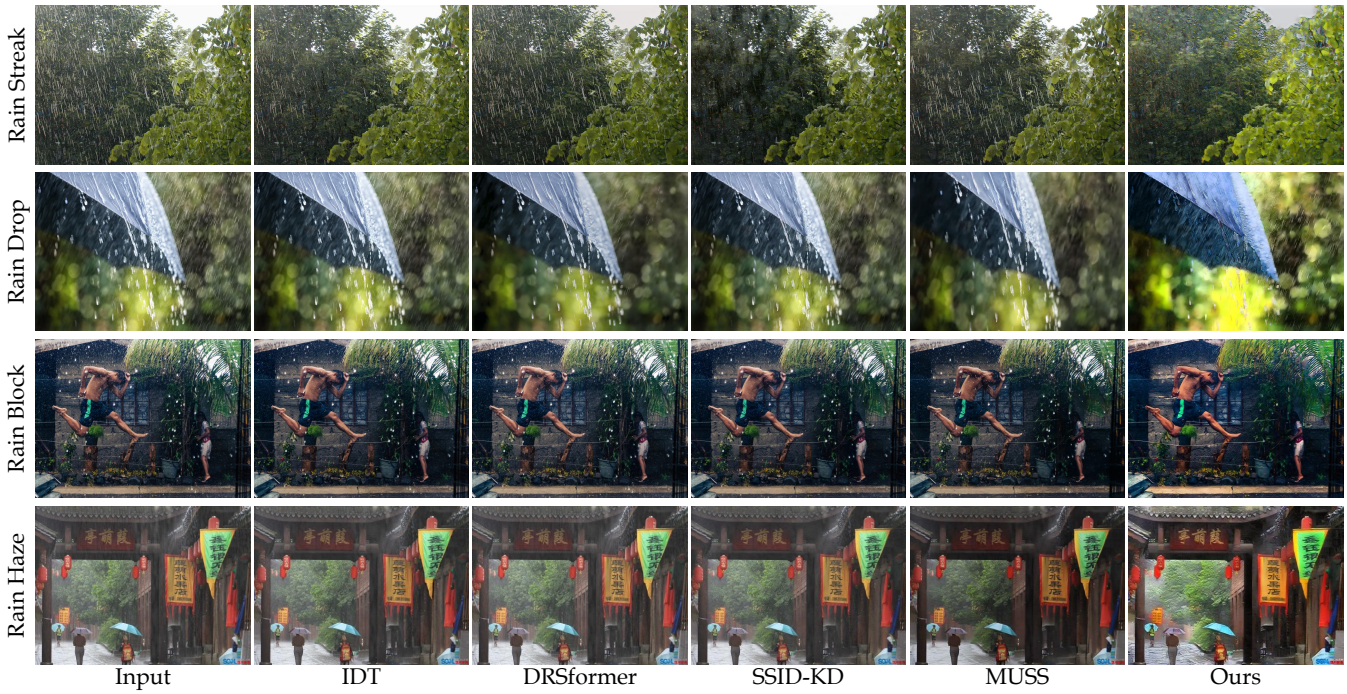


Figure 4: Comparison of deraining performance with existing state-of-the-art rain removal methods under multiple real rainy patterns.

student network. The learning rates are adjusted using MultiStepLR and the Adam optimizer with default settings is used to optimize network parameters. During training, images are cropped to sizes of 128×128, and data augmentation is achieved through random rotation and flipping.

Benchmarks Descriptions and Metrics. Experiments are conducted on five publicly available real datasets, DDN-SIRR [Wei *et al.*, 2018a], IVIPC_DQA [Wu *et al.*, 2019], MPID [Li *et al.*, 2019b], SPA [Wang *et al.*, 2019] and NTU [Chen *et al.*, 2018]. These datasets contain 147, 206, 185, 146, and 103 images, respectively. Among them, only the rain streak part of images are selected in the MPID dataset, and only 103 frames from the real rain video are randomly selected in the NTU dataset. Experimental performance is evaluated using two non-reference metrics Brisque [Mittal *et al.*, 2012] and Entropy [González *et al.*, 2006].

Comparison Methods. Comparison methods in the experiment are as follows. Fully supervised methods include universal rain removal methods PReNet [Ren *et al.*, 2019], MR-FAN [Liang *et al.*, 2022], IDT [Xiao *et al.*, 2022] and DRSformer [Chen *et al.*, 2023b], and the combined removal methods AMRR [He *et al.*, 2020], TRNR [Ran *et al.*, 2021] and UDR [Chen *et al.*, 2023a] for multiple disturbance factors. Semi-supervised methods SIRR [Wei *et al.*, 2018b], Syn2real [Yasarla *et al.*, 2020], MOSS [Huang *et al.*, 2021], SIDD-KD [Cui *et al.*, 2022] and MUSS [Huang *et al.*, 2023].

4.2 Comparison with State-of-the-art Methods

Quantitative Comparisons. Table 2 compares the deraining performance of PRRN with multiple existing state-of-the-art methods on five real rain datasets. The deraining effect is reflected by two non-reference metrics Brisque and Entropy, where a smaller value of Brisque indicates a clearer image and a higher value of Entropy indicates a more detailed image. As can be seen from Table 2, whether comparing fully supervised methods or semi-supervised methods, PRRN's student network can demonstrate optimal or suboptimal performance on most datasets, which strongly proves the effectiveness and robustness of our proposed method in a variety of different real scenarios. In addition, compared with the teacher network, the student network trained on real data shows a significant improvement in deraining performance, indicating that our method can effectively capture key information from real rain data.

Qualitative Comparisons. Figure 4 provides a comprehensive visual comparison of deraining capabilities of PRRN with other advanced rain removal methods under various real rainy patterns. The figure highlights the proficiency of PRRN in accurately identifying and eliminating various rain patterns in real-life scenarios, yielding final images that are not only clear but also visually appealing.

Figure 5 compares PRRN's performance against other rain removal methods for handling complex rainy patterns involving multiple degradation factors. From details of displayed images, it can be seen that PRRN can not only deal with multiple composite rainy patterns, but also restore better texture details of the image compared to other multi-factor joint rain removal methods.

4.3 Ablation Study

A series of ablation experiments are conducted to verify the effectiveness of our proposed real rain image degradation pipeline and patterns-aware rain removal network.

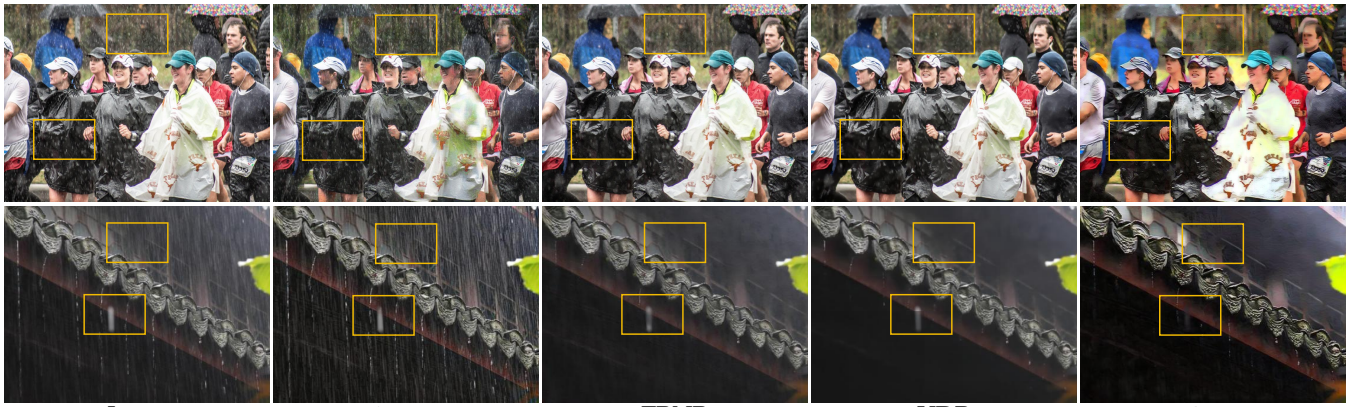


Figure 5: Performance comparison of rain removal methods considering multiple degradation factors in complex rain scenarios.

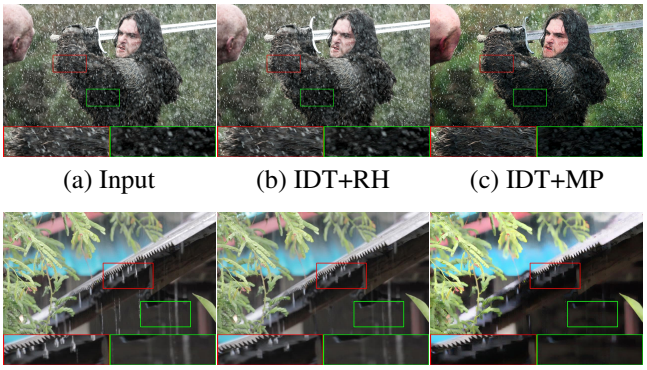


Figure 6: Comparison of deraining performance of IDT [Xiao *et al.*, 2022] and MRAFN [Liang *et al.*, 2022] in real scenarios, after training with Rain200H (RH) and MPRain (MP) datasets, respectively.

Effectiveness of the Degradation Pipeline. We validate the image degradation pipeline's effectiveness by assessing the impact of MPRain on the deraining performance of several advanced rain removal methods. Figures 6 (a)-(c) and (d)-(f) show deraining effects of IDT and MRFAN tested on real rain images after training using Rain200H and MPRain datasets respectively. It can be seen that both methods of training on the MPRain achieve better deraining performance on real-world rainy image, which fully demonstrates the practicality of the MPRain and the effectiveness of the degradation pipeline of the real rain image.

Effectiveness of MIAM and L_{DA}. Table 3 verifies the effectiveness of sequentially adding MIAM and L_{DA} into PRRN through the change in values before and after the DDN-SIRR dataset. The table shows that MIAM and L_{DA} both enhance PRRN's training performance to some degree.

Method	DDN-SIRR	
-	Brisque↓	Entropy↑
w/o MIAM	27.39	7.491
MIAM+w/o L_{DA}	$27.00_{\downarrow 0.390}$	$7.558_{\uparrow 0.067}$
MIAM+w/ L_{DA}	$24.57_{\downarrow 2.820}$	$7.596_{\uparrow 0.105}$

Table 3: Ablation experiments of MIAM and L_{DA}. The table records the performance improvement of the backbone on the DDN-SIRR dataset after adding MIAM and L_{DA} sequentially.

5 Concluding Remarks

In this paper, we propose a novel image deraining paradigm that aims to tackle challenging real-world rain removal problems more effectively. Through our innovative "phenomenological degradation pipeline", we have synthesized a broad and diverse rain dataset MPRain, which covers a wide variety of rainy patterns and provides a strong guarantee for practical applications. To better learn from synthetic and real data, we propose a patterns-aware rain removal network. By designing a multi-representation learning module and an inter-domain alignment loss, the network realizes the alignment of synthetic and real data in multiple feature subspaces. Numerous experimental results strongly demonstrate the paradigm's exceptional performance and its potential for addressing complex rain removal in real settings.

Limitations. Given the current method's propensity to overexpose and lose sky detail, we plan to incorporate two techniques. The first is to reasonably control the luminance difference between the rain image and the ground truth in the synthesized dataset, to avoid overly obvious luminance enhancement in the result. The second is to separate high and low frequency parts of the image to eradicate rain textures while safeguarding low-frequency details of the sky.

Acknowledgments

This work was supported by the Key-Area Research and Development Program of Guangdong Province (No. 2021B0101410003), the Guangdong Basic and Applied Basic Research Foundation (No. 2024A1515011563), and the Key R&D Project of Guangzhou (No. 202206010091).

References

[Chen and Hsu, 2013] Yi-Lei Chen and Chiou-Ting Hsu. A generalized low-rank appearance model for spatio-temporally correlated rain streaks. *2013 IEEE International Conference on Computer Vision*, pages 1968–1975, 2013.

[Chen *et al.*, 2018] Jie Chen, Cheen-Hau Tan, Junhui Hou, Lap-Pui Chau, and He Li. Robust video content alignment and compensation for rain removal in a cnn framework. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6286–6295, 2018.

[Chen *et al.*, 2023a] Sixiang Chen, Tian Ye, Jinbin Bai, Erkang Chen, Jun Shi, and Lei Zhu. Sparse sampling transformer with uncertainty-driven ranking for unified removal of raindrops and rain streaks. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 13106–13117, 2023.

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

[Cui *et al.*, 2022] Xin Cui, Cong Wang, Dongwei Ren, Yunjin Chen, and Pengfei Zhu. Semi-supervised image deraining using knowledge distillation. *IEEE Transactions on Circuits and Systems for Video Technology*, 32:8327–8341, 2022.

[Feng *et al.*, 2023] Yuxin Feng, Xiaozhe Meng, Fan Zhou, Weisi Lin, and Zhuo Su. Real-world non-homogeneous haze removal by sliding self-attention wavelet network. *IEEE Trans. Circuits Syst. Video Technol.*, 33(10):5470–5485, 2023.

[Fu *et al.*, 2017a] Xueyang Fu, Jiabin Huang, Delu Zeng, Yue Huang, Xinghao Ding, and John Paisley. Removing rain from single images via a deep detail network. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1715–1723, 2017.

[Fu *et al.*, 2017b] Xueyang Fu, Jiabin Huang, Delu Zeng, Yue Huang, Xinghao Ding, and John William Paisley. Removing rain from single images via a deep detail network. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1715–1723, 2017.

[Gao *et al.*, 2023] Xinjian Gao, Yang Wang, and Meng Wang. Macroscopic-and-microscopic rain streaks disentanglement network for single-image deraining. *IEEE Transactions on Image Processing*, 32:2663–2677, 2023.

[González *et al.*, 2006] Rafael Corsino González, Richard E. Woods, and Steven L. Eddins. Digital image processing using matlab. 2006.

[Guo *et al.*, 2022] Xin Guo, Xueyang Fu, Man Zhou, Zhen Huang, Jialun Peng, and Zhengjun Zha. Exploring fourier prior for single image rain removal. In *International Joint Conference on Artificial Intelligence*, 2022.

[He *et al.*, 2020] Da He, Xiaoyu Shang, and Jiajia Luo. Adherent mist and raindrop removal from a single image using attentive convolutional network. *Neurocomputing*, 505:178–187, 2020.

[Hou *et al.*, 2020] Yifan Hou, Jian Zhang, James Cheng, Kaili Ma, Richard T. B. Ma, Hongzhi Chen, and Ming Yang. Measuring and improving the use of graph information in graph neural networks. In *International Conference on Learning Representations*, 2020.

[Hu *et al.*, 2019] Xiaowei Hu, Chi-Wing Fu, Lei Zhu, and Pheng-Ann Heng. Depth-attentional features for single-image rain removal. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8014–8023, 2019.

[Huang *et al.*, 2021] Huaibo Huang, Aijing Yu, and Ran He. Memory oriented transfer learning for semi-supervised image deraining. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 7728–7737, 2021.

[Huang *et al.*, 2023] Huaibo Huang, Mandi Luo, and Ran He. Memory uncertainty learning for real-world single image deraining. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(3):3446–3460, 2023.

[Jiang *et al.*, 2022] Kui Jiang, Zhongyuan Wang, Zheng Wang, Peng Yi, Junjun Jiang, Jinsheng Xiao, and Chia-Wen Lin. Danet: Image deraining via dynamic association learning. In *International Joint Conference on Artificial Intelligence*, 2022.

[Jin *et al.*, 2018] Xin Jin, Zhibo Chen, Jianxin Lin, Zhikai Chen, and Wei Zhou. Unsupervised single image deraining with self-supervised constraints. *2019 IEEE International Conference on Image Processing (ICIP)*, pages 2761–2765, 2018.

[Li *et al.*, 2017] 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*, 28:492–505, 2017.

[Li *et al.*, 2018] Xia Li, Jianlong Wu, Zhouchen Lin, Hong Liu, and Hongbin Zha. Recurrent squeeze-and-excitation context aggregation net for single image deraining. In Vittorio Ferrari, Martial Hebert, Cristian Sminchisescu, and Yair Weiss, editors, *Computer Vision – ECCV 2018*, Cham, 2018. Springer International Publishing.

[Li *et al.*, 2019a] Ruoteng Li, Loong Fah Cheong, and Robby T. Tan. Heavy rain image restoration: Integrating physics model and conditional adversarial learning. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1633–1642, 2019.

[Li *et al.*, 2019b] Siyuan Li, Iago Breno Araujo, Wenqi Ren, Zhangyang Wang, Eric K. Tokuda, Roberto Hirata Junior, Roberto Marcondes Cesar Junior, Jiawan Zhang, Xiaojie Guo, and Xiaochun Cao. Single image deraining: A comprehensive benchmark analysis. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3833–3842, 2019.

[Liang *et al.*, 2022] Songliang Liang, Xiaozhe Meng, Zhuo Su, and Fan Zhou. Multi-receptive field aggregation net-

work for single image deraining. *J. Vis. Commun. Image Represent.*, 84:103469, 2022.

[Liu *et al.*, 2021] Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. Swin transformer: Hierarchical vision transformer using shifted windows. *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 9992–10002, 2021.

[Meng *et al.*, 2023] Xiaozhe Meng, Yuxin Feng, Fan Zhou, Yun Liang, and Zhuohan Su. Towards real-world haze removal with uncorrelated graph model. *J. Vis. Commun. Image Represent.*, 96:103927, 2023.

[Mittal *et al.*, 2012] Anish Mittal, Anush K. Moorthy, and Alan Conrad Bovik. No-reference image quality assessment in the spatial domain. *IEEE Transactions on Image Processing*, 21:4695–4708, 2012.

[Radford *et al.*, 2015] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *CoRR*, abs/1511.06434, 2015.

[Ran *et al.*, 2021] Wu Ran, Bohong Yang, Peirong Ma, and Hong Lu. Trnr: Task-driven image rain and noise removal with a few images based on patch analysis. *IEEE Transactions on Image Processing*, 32:721–736, 2021.

[Ren *et al.*, 2019] Dongwei Ren, Wangmeng Zuo, Qinghua Hu, Peng Fei Zhu, and Deyu Meng. Progressive image deraining networks: A better and simpler baseline. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3932–3941, 2019.

[Su *et al.*, 2021] Zhipeng Su, Yixiong Zhang, Jianghong Shi, and Xiao-Ping Zhang. Recurrent network knowledge distillation for image rain removal. *IEEE Transactions on Cognitive and Developmental Systems*, PP:1–1, 2021.

[Sun *et al.*, 2014] Shao-Hua Sun, Shang-Pu Fan, and Y. Wang. Exploiting image structural similarity for single image rain removal. *2014 IEEE International Conference on Image Processing (ICIP)*, pages 4482–4486, 2014.

[Tremblay *et al.*, 2020] Maxime Tremblay, Shirsendu Sukanta Halder, Raoul de Charette, and Jean-François Lalonde. Rain rendering for evaluating and improving robustness to bad weather. *International Journal of Computer Vision*, 129:341 – 360, 2020.

[Wang *et al.*, 2019] Tianyu Wang, Xin Yang, Ke Xu, Shaozhe Chen, Qiang Zhang, and Rynson W. H. Lau. Spatial attentive single-image deraining with a high quality real rain dataset. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 12262–12271, 2019.

[Wang *et al.*, 2021] Xintao Wang, Liangbin Xie, Chao Dong, and Ying Shan. Real-esrgan: Training real-world blind super-resolution with pure synthetic data. *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, pages 1905–1914, 2021.

[Wei *et al.*, 2018a] Wei Wei, Deyu Meng, Qian Zhao, Zongben Xu, and Ying Wu. Semi-supervised transfer learning for image rain removal. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3872–3881, 2018.

[Wei *et al.*, 2018b] Wei Wei, Deyu Meng, Qian Zhao, Zongben Xu, and Ying Wu. Semi-supervised transfer learning for image rain removal. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3872–3881, 2018.

[Wei *et al.*, 2021] Yanyan Wei, Zhao Zhang, Yang Wang, Mingliang Xu, Yi Yang, Shuicheng Yan, and Meng Wang. Deraincyclegan: Rain attentive cyclegan for single image deraining and rainmaking. *IEEE Transactions on Image Processing*, 30:4788–4801, 2021.

[Wu *et al.*, 2019] Qingbo Wu, Lei Wang, King Ngi Ngan, Hongliang Li, Fanman Meng, and Linfeng Xu. Subjective and objective de-raining quality assessment towards authentic rain image. *IEEE Transactions on Circuits and Systems for Video Technology*, 30:3883–3897, 2019.

[Xiao *et al.*, 2022] Jie Xiao, Xueyang Fu, Aiping Liu, Feng Wu, and Zhengjun Zha. Image de-raining transformer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45:12978–12995, 2022.

[Yang *et al.*, 2016] Wenhan Yang, Robby T. Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, and Shuicheng Yan. Deep joint rain detection and removal from a single image. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1685–1694, 2016.

[Yasarla *et al.*, 2020] Rajeev Yasarla, Vishwanath A. Sindagi, and Vishal M. Patel. Syn2real transfer learning for image deraining using gaussian processes. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2723–2733, 2020.

[Zhang *et al.*, 2023] Ning Zhang, Francesco Nex, George Vosselman, and Norman Kerle. Lite-mono: A lightweight cnn and transformer architecture for self-supervised monocular depth estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 18537–18546, June 2023.