Boosting Image De-Raining via Central-Surrounding Synergistic Convolution

Long Peng, Yang Wang*, Xin Di, PeizheXia, Xueyang Fu, Yang Cao, Zheng-Jun Zha

University of Science and Technology of China
longp2001@mail.ustc.edu.cn, ywang120@ustc.edu.cn

Abstract

Rainy images suffer from quality degradation due to the synergistic effect of rain streaks and accumulation. The rain streaks are anisotropic and show a specific directional arrangement, while the rain accumulation is isotropic and shows a consistent concentration distribution in local regions. This distribution difference makes unified representation learning for rain streaks and accumulation challenging, which may lead to structure distortion and contrast degradation in the deraining results. To address this problem, a central-surrounding mechanism inspired Synergistic Convolution (SC) is proposed to extract rain streaks and accumulation features simultaneously. Specifically, the SC consists of two parallel novel convolutions: Central-Surrounding Difference Convolution (CSD) and Central-Surrounding Addition Convolution (CSA). In CSD, the difference operation between central and surrounding pixels is injected into the feature extraction process of convolution to perceive the direction distribution of rain streaks. In CSA, the addition operation between central and surrounding pixels is injected into the feature extraction process of convolution to facilitate the modeling of rain accumulation properties. The SC can be used as a general unit to substitute Vanilla Convolution (VC) in current de-raining networks to boost performance. To reduce computational costs, CSA and CSD in SC are merged into a single VC kernel by our parameter equivalent transformation before inferencing. Evaluations of twelve de-raining methods on nine public datasets demonstrate that our proposed SC can **comprehensively improve the performance of twelve de-raining networks under various rainy conditions without changing the original network structure or introducing extra computational costs**. Even for the current SOTA methods, SC can further achieve SOTA++ performance. The source codes will be publicly available.

Introduction

Images captured under rainy conditions often suffer from quality degradation due to the effect of rain streaks and rain accumulation, which will cause unpleasant visual perception and hurt the performance of outdoor computer vision systems, such as video surveillance (Yang et al. 2022b) and object detection (He et al. 2017). Thus, restoring images from rain is an essential pre-processing step for both human vision and computer vision systems and has drawn much research attention in recent years (Guo et al. 2021; Chen et al. 2023b; Wang, Ma, and Liu 2023; Chen et al. 2023a; Guo et al. 2023; Zhang et al. 2023; Zou et al. 2022).

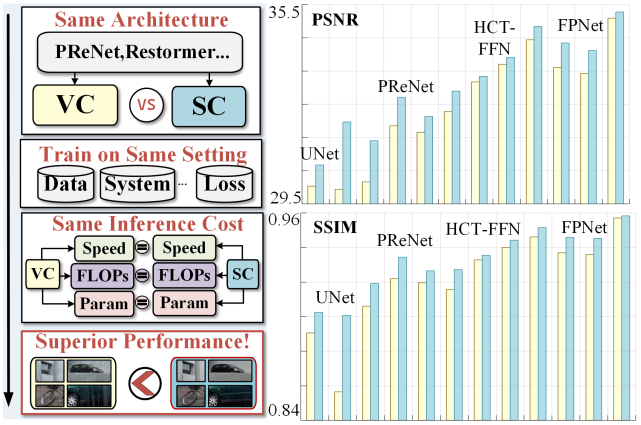

Figure 1: (a) Models equipped with SC trained in the same environment as VC exhibit the same inference cost but superior performance. (b) Under the boost of SC, twelve existing deraining methods have achieved comprehensive performance improvements without increasing computational costs during inferencing.

The effects of rain streaks and accumulation on the image can be formulated as follows (Yang et al. 2017):

$$O = T \odot (B + \sum_{i}^{n} S_i) + (1 - T) \odot A. \quad (1)$$

where O denotes the captured rainy images. B denotes the background. S_i represents the rain-streak layer that has the same direction distribution. A is the global atmospheric light, and T is the atmospheric transmission. i indexes the rain streak layer, and n is the maximum number of the rain streak layers. The rain streaks are anisotropic and have large variations in orientation, which cause structure distortion in the background. In contrast, the rain accumulation is isotropic and has smooth variations across regions, which narrows down the dynamic range of the image, especially

*Yang Wang is the corresponding author

in heavy rainy conditions. The goal of image deraining is to remove rain streaks and enhance the dynamic range of the image simultaneously (Yang et al. 2020; Li, Cheong, and Tan 2019; Wen et al. 2024; Zhang et al. 2024).

To achieve this goal, researchers try to make assumptions about the statistics of rain streaks. Among traditional methods, researchers devise specialized regularities for minimization and propose diverse priors by exploring the physical properties of rain direction, such as Gaussian mixture model (Li et al. 2016) and image decomposition (Kang, Lin, and Fu 2011). Benefiting from the significant success of deep learning in de-raining, various networks are proposed to learn the statistical regularities of rain streaks and accumulations from datasets (Li, Cheong, and Tan 2019; Chen et al. 2023b). For better deraining results, researchers propose to inject some assumptions or priors related to rain streaks into network design and propose various modules and architectures, such as SPANet (Wang et al. 2019a) and SPDNet (Yi et al. 2021). However, these methods mainly focus on rain streak removal, which has difficulty in learning unified representations for anisotropic rain streaks and isotropic rain accumulation, leading to structure distortion and contrast degradation in the derained results.

To alleviate the above issues, this paper proposes a novel Synergistic Convolution (SC) to **simultaneously extract the feature of rain streaks and accumulation**. It is inspired by the central-surrounding mechanism of human vision (Chao-Yi et al. 1991), which can help humans perceive contrast and direction variation more easily. Specifically, our proposed SC contains two parallel convolutions, namely central-surrounding difference convolution (CSD) and central-surrounding addition convolution (CSA). The CSD calculates gradient information in eight directions and adaptively perceives the distribution of rain streaks in all directions using learnable weights. The CSA adds the value of the central pixel to enhance the response of the smooth signal when extracting the contrast relevant properties over the smoothed area. Furthermore, the proposed SC can be used as a general unit to replace the Vanilla Convolution (VC) in various de-raining networks to simultaneously extract the feature of rain streaks/accumulation when training. To reduce the computational costs, we merge the CSA and CSD in SC into a VC kernel during inferencing. To demonstrate the effectiveness of SC, we evaluate it on twelve de-raining methods in nine publicly available datasets. Under the boosting of the proposed SC, the existing deraining methods can **achieve SOTA++ performance without introducing extra computational costs**, as shown in Fig. 1.

The contributions can be summarized as follows:

(1) Inspired by the central-surrounding mechanism in human vision, a novel central-surrounding addition convolution (CSA) and central-surrounding difference convolution (CSD) are proposed to extract isotropic rain streaks and anisotropic rain accumulation simultaneously.

(2) With the synergy of CSA and CSD in SC, the properties of rain streaks and accumulation can be learned more comprehensively. Further, the proposed SC can be used as a basic unit and generalized to various networks to boost performance without introducing extra computational costs.

(3) Extensive experiments on nine datasets and twelve deraining methods demonstrate that SC can comprehensively boost performance. Additionally, we achieve SOTA++ based on existing methods under the boosting of SC.

Related Work

Single image de-raining has been studied for a long time (Wang, Ma, and Liu 2023; Peng et al. 2024c,b; Chen et al. 2023c; Peng et al. 2024a; Lin et al. 2024; Zheng, Lu, and Narasimhan 2024; Chen et al. 2024; Gu, Wang, and Li 2023; Wang et al. 2024, 2023; Zhao et al. 2024), which can be divided into traditional and deep learning-based methods. Traditional methods introduce some image priors and hand-extracted features. However, it can only deal with specific rainy artifacts and is hard to deal with complex real scenes, leading to poor generalization ability. Benefiting from the rapid development of deep learning, many learning-based image rain removal methods have been proposed. A typical method is to introduce the physical properties of rain streaks into the network and module design. For example, according to the directional characteristics of rain, Wang *et al.* (Wang et al. 2019a) propose a spatial attentive network to remove rain streaks in a local-to-global manner. Yi *et al.* (Yi et al. 2021) propose a structure-preserving de-raining network by using residue channel prior. However, these networks cannot simultaneously extract the features of rain streaks and accumulation. A feasible solution is the structural re-parameterization parallel framework (Ding et al. 2021, 2019) and dynamic convolution (Yang et al. 2022a), which can extract different features through different parallel branches simultaneously during training and then reduce the computational cost by equivalently converting parameters in the parallel framework during inferencing. However, they can't explicitly guide the modeling of rain streaks and accumulation related features by utilizing the isotropic and arrangement properties of rain degradation. And, they may damage de-raining performance since multiple traditional convolutions in parallel may conflict with each other.

The central surrounding mechanism (Chao-Yi et al. 1991; Jiang et al. 2023; Yu et al. 2020; Chen, He, and Lu 2024) in the human visual system is a mechanism formed after a long period of evolution, which can help human eyes perceive high-frequency detail and reserve low-frequency information under various challenging conditions. Inspired by this mechanism, we propose a novel central-surrounding synergistic convolution to learn a unified representation of rain streaks and rain accumulation for single-image de-raining.

Proposed Method

To explicitly guide the modeling of anisotropic rain streaks and isotropic rain accumulation simultaneously, we propose a novel Synergistic Convolution (SC). It is inspired by the central-surrounding mechanism of human vision, which can automatically magnify small differences by comparing the central and the surroundings and help humans perceive direction variation, as shown in Fig. 2 (a). Taking the image of Victor Vasarely Optical Illusion in Fig. 2 (b) as an example, the human eye will observe a bright "X" in the di-

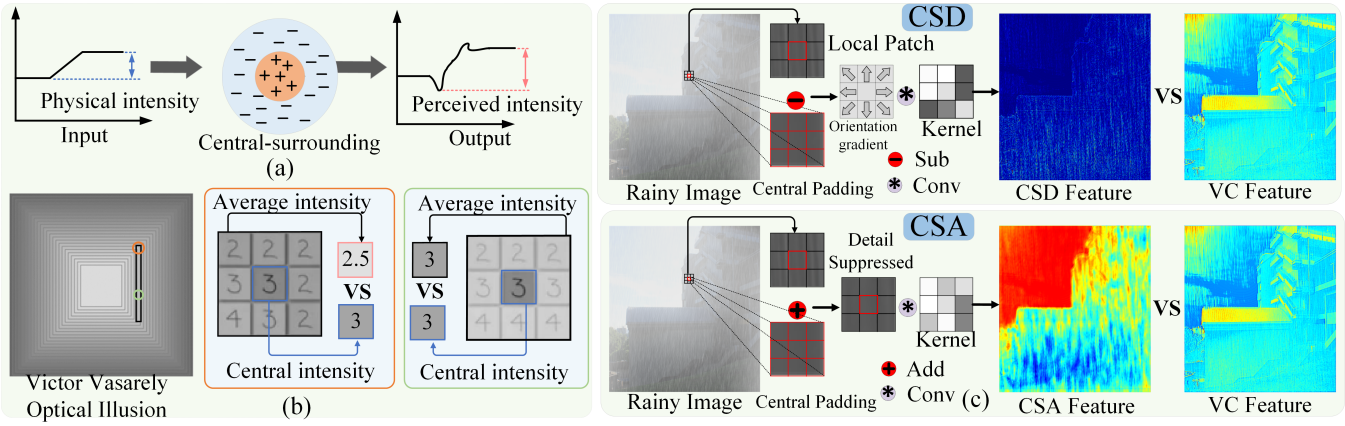

Figure 2: (a) The central-surrounding mechanism of human vision. (b) The example of Victor Vasarely's Optical Illusion. (c) The scheme illustration of our proposed CSD and CSA.

agonal direction. Actually, the pixel intensities within the black box are all the same. However, the average intensity of the surrounding pixels of the orange circle is 2.5, which is much lower than that of the green circle, as shown in Fig. 2 (b). The difference between the center and surroundings allows the human eyes to perceive a greater stimulus intensity than physical changes. In this paper, we inject the above comparison operation into the convolution kernel and devise a Central-Surrounding Difference (CSD) kernel and Central-Surrounding Addition (CSA) kernel to perceive the rain streak distribution and rain accumulation properties, respectively, as shown in Fig. 2 (c). With the synergy of CSD and CSA, the network can perceive the rain-relevant properties more comprehensively.

Central-Surrounding Difference Convolution

Rain streaks are anisotropic and have a strong directional property. This physical property is widely used to design image rain removal algorithms, such as GMM(Li et al. 2016) and SPANet(Wang et al. 2019a). However, the network and module are difficult to guide the internal feature extraction process within the network. Different from previous works, we propose to inject the above physical property into the feature extraction process at each convolution kernel to explicitly guide rain-relevant feature extraction.

To extract various directional gradients of rain streaks, we introduce the difference operation to calculate the directional gradient between the central and surrounding pixels in the receptive field. Taking $k \times k$ convolution as an example, the gradients in each direction are calculated as follows:

$$G(x_{p_i}) = x_{p_i} - x_{p_c} \qquad i = 1, 2, 3...k^2. \tag{2}$$

where x_{p_c} represents the intensity of the central pixel, x_{p_i} represents the intensity of the pixels in the surrounding area, and $G(x_{p_i})$ represents the central gradient function. Through the above difference operation, we can obtain directional gradient relationships between the central pixel and surrounding pixels. For example, with a 3×3 convolution kernel, we can get eight different directional gradient results. These directional gradients provide candidates for rain streak direction representation. Then, all the candidate gradients are weighted, as follows:

$$F(p_c) = \sum_{p_i \in \mathbb{R}} W(p_i) \cdot G(x_{p_i}). \tag{3}$$

where $W(p_i)$ is the weight for each direction, which can be updated through back-propagation. $F(p_c)$ represents the convolution output of p_c. With continuous training, the weights are constantly updated, and the weights on the most representative directional gradients will be improved, while the weights on the rain streak's irrelevant positions will be suppressed, as shown in Fig. 2 (c).

Central-Surrounding Addition Convolution

Rain accumulation is a common phenomenon that forms a strong veiling and reduces the visibility of the captured images. It often occurs in 1) heavy rain or rainstorm conditions and other scenes with high densities of rain streaks and 2) the distant area of rainy scenes. Rain accumulation removal is an urgent problem for the single-image de-raining task (Yang et al. 2020). A representative method is to use the rain model in Eq. 1 to learn the imaging parameters through the CNN network to reconstruct the de-raining image (Li, Cheong, and Tan 2019). However, the rain accumulation is isotropic and smooth, and the statistical modeling process for rain accumulation is susceptible to tiny textures (Wang et al. 2019b). To alleviate this problem, we propose to add the central value as the isotropic smoothing component to suppress the anisotropic properties of tiny textures in the process of convolution, as follows:

$$F(p_c) = \sum_{p_i \in \mathbb{R}} W(p_i) \cdot (x_{p_i} + x_l). \tag{4}$$

where x_l represents the isotropic smoothing component of local regions. An intuitive option is to use the average intensity of local regions covered by the receptive field as x_l. However, mean filtering needs to be performed for each local region, which will increase computational costs and

memory requirements. To solve this, based on the local smoothness assumption, this paper proposes to use the central pixel intensity of local regions to replace the average intensity. To verify that the central pixel intensity has a similar property to the smoothing component, we conduct the following statistical experiments. Firstly, we randomly select 20,000 rainy images and feature images obtained by (Ren et al. 2019) from the public datasets. Secondly, we divide these images/features into 1,500,000 images/features patches with 3×3 resolution. Finally, we calculate the difference between the central pixel intensity p_c and the mean intensity of all pixels p_i within each patch, as follows:

$$\delta = \left| p_c - \frac{1}{n} \cdot \sum_i p_i \right|. \tag{5}$$

where n represents the number of pixels within each patch. Through the statistical experiments, we find that over 90% of δ is less than 0.071/0.048 (the maximum value is 1) in the image/feature space, which indicates that in the local receptive field, the intensity of the central pixel is very close to the mean in the image and feature space. Therefore, we can use the central intensity to replace the mean intensity and then suppress tiny structures within the local regions to extract isotropic statistical properties of rain accumulation, and the formula Eq. 4 can be rewritten as:

$$F(p_c) = \sum_{p_i \in \mathbb{R}} W(p_i) \cdot (x_{p_i} + x_{p_c}). \tag{6}$$

Based on the above methods, we can capture rain streaks' relevant features through CSD and perceive rain accumulation distribution through CSA. To fully exploit the synergistic properties of CSA and CSD, we connect them in parallel to extract rain streaks and accumulation features simultaneously, denoted as SC, which can perceive the properties of rain streaks and accumulation more comprehensively, as shown in Fig. 2 (c). Further, SC can be used as a general unit to replace the VC in the existing methods to improve image de-raining performance without changing the network structure. Specifically, given a deraining network, we use SC to replace the VC kernels within the network during training. With the assistance of SC, the network designed for rain streak removal can improve the ability to remove rain accumulation. However, this parallel structure will increase the model complexity and parameters, resulting in low efficiency. To reduce computational costs during inferencing, CSD and CSA in SC are merged into a single VC kernel by our proposed equivalent transformation:

$$\begin{aligned} F(p_c) &= \underbrace{\sum_{p_i \in \mathbb{R}} W_{csd}(p_i) \cdot (x_{p_i} - x_{p_c})}_{CSD} + \underbrace{\sum_{p_i \in \mathbb{R}} W_{csa}(p_i) \cdot (x_{p_i} + x_{p_c})}_{CSA} \\ &= \underbrace{\sum_{p_i \in \mathbb{R}} W_{csd}(p_i) \cdot x_{p_i} + \sum_{p_i \in \mathbb{R}} W_{csa}(p_i) \cdot x_{p_i}}_{k \times k\ convolution} \\ &\quad + \underbrace{x_{p_c} \cdot \left(\sum_{p_i \in \mathbb{R}} W_{csa}(p_i) - \sum_{p_i \in \mathbb{R}} W_{csd}(p_i) \right)}_{term}. \end{aligned} \tag{7}$$

where W_{csd} and W_{csa} represent the convolution kernels of CSD and CSA. $F(p_c)$ represents feature responses. Since x_{p_c} is the central pixel, $term$ in Eq. 7 can be seen as a 1×1 convolution. Then, we extend the 1×1 convolution to the $k \times k$ convolution W_c (Ding et al. 2019, 2021):

$$\mathrm{W}_c(p_i) = \begin{cases} \mathrm{sum}(W_{csa}) - \mathrm{sum}(W_{csd}) & if \quad p_i = p_c \\ 0 & if \quad p_i \neq p_c. \end{cases} \tag{8}$$

Further, Eq. 7 can be expressed as:

$$\begin{aligned} F(p_c) = &\underbrace{\sum_{p_i \in \mathbb{R}} W_{csd}(p_i) \cdot x_{p_i}}_{k \times k\ convolution} + \underbrace{\sum_{p_i \in \mathbb{R}} W_{csa}(p_i) \cdot x_{p_i}}_{k \times k\ convolution} + \\ &\underbrace{\left(\sum_{p_i \in \mathbb{R}} W_c(p_i) \cdot x_{p_i} \right)}_{k \times k\ convolution}. \end{aligned} \tag{9}$$

Finally, we fuse all $k \times k$ kernels into a single $k \times k$ kernel W_{all} by the linearity of convolution (Ding et al. 2021):

$$\begin{aligned} F(p_c) &= \sum_{p_i \in \mathbb{R}} (W_{csd}(p_i) + W_{csa}(p_i) + W_c(p_i)) \cdot x_{p_i} \\ &= \sum_{p_i \in \mathbb{R}} W_{all}(p_i) \cdot x_{p_i}. \end{aligned} \tag{10}$$

where W_{all} is the sum of W_{csd}, W_{csa} and W_c. In the inference phase, we can use W_{all} to replace SC equivalently, which can significantly reduce computational costs while maintaining the same performance as SC.

Experiments and Analysis

Experimental Settings

We evaluate the effectiveness of our proposed method on nine public single-image de-raining datasets, including both synthetic and real datasets: Rain12(Li et al. 2016), Rain200H(Yang et al. 2017), Rain200L(Yang et al. 2017), Rain1200(Zhang and Patel 2018), Rain12600(Fu et al. 2017), Outdoor-Rain(Li, Cheong, and Tan 2019), JORDER-R(Yang et al. 2017), ID-CGAN-R(Zhang, Sindagi, and Patel 2019) and SIRR-R(Wei et al. 2019). Following previous works (Li, Cheong, and Tan 2019; Yi et al. 2021; Chen et al. 2023b), we use reference metrics of PSNR and SSIM to evaluate the performance with ground truth. For real datasets without ground truth, we use non-reference metrics to evaluate. Referring to previous works (Yi et al. 2021; Chen et al. 2024), we use four kinds of non-reference metrics, including the NIQE, BRISQUE, PIQE, and PI. In all experiments, we keep the training settings (*e.g.,* model framework, loss function, and active function) the same as the original official public code, except that the VC is replaced by the SC on eight NVIDIA RTX3090 GPUs at Pytorch.

Comparison methods. To fully verify the effectiveness and generality of our proposed synergistic convolution, we select twelve kinds of de-raining methods, including both classic and SOTA: NLEDN (Li et al. 2018a), RESCAN (Li et al. 2018b), PReNet (Ren et al. 2019), UNet (Ronneberger, Fischer, and Brox 2015), Syn2Real (Yasarla, Sindagi, and Patel 2020), SPANet (Wang et al. 2019a), MPRNet (Zamir et al. 2021), HCT-FFN (Chen et al. 2023c), FPNet (Guo et al. 2022), SPDNet (Yi et al. 2021), Restormer (Zamir et al. 2022), DRSformer (Chen et al. 2023b).

	Rain12	Rain200H	Rain200L	Rain12600	Rain1200	Outdoor-Rain
NLEDN	36.706/0.950	28.640/0.871	37.960/0.978	32.050/0.918	34.020/0.928	24.303/0.873
NLEDN*	**37.233/0.956**	**29.314/0.883**	**38.675/0.981**	**32.801/0.927**	**34.216/0.931**	**25.151/0.886**
UNet	34.614/0.942	25.350/0.822	33.420/0.932	30.520/0.892	31.760/0.891	24.561/0.864
UNet*	**34.867/0.949**	**25.960/0.828**	**33.950/0.943**	**30.950/0.899**	**32.120/0.900**	**26.182/0.896**
RESCAN	36.540/0.957	27.450/0.821	35.080/0.959	30.940/0.882	32.000/0.892	17.634/0.628
RESCAN*	**36.945/0.957**	**28.440/0.873**	**38.678/0.982**	**32.838/0.930**	**33.928/0.929**	**20.977/0.733**
PReNet	36.610/0.960	29.040/0.890	37.120/0.976	32.750/0.927	33.370/0.919	22.222/0.860
PReNet*	**37.410/0.964**	**29.746/0.910**	**38.534/0.983**	**33.297/0.935**	**33.750/0.936**	**23.530/0.878**
SPANet	35.920/0.958	26.270/0.867	35.790/0.965	30.580/0.907	32.120/0.912	20.290/0.828
SPANet*	**37.130/0.966**	**27.935/0.887**	**38.410/0.983**	**30.846/0.910**	**32.762/0.919**	**21.338/0.850**
Syn2Real	35.811/0.948	28.540/0.874	35.300/0.968	32.450/0.923	33.310/0.916	24.500/0.889
Syn2Real*	**36.162/0.959**	**28.912/0.880**	**36.325/0.975**	**32.647/0.926**	**33.686/0.926**	**25.082/0.894**
HCT-FFN	37.175/0.956	29.046/0.885	38.720/0.983	33.078/0.932	34.394/0.935	26.607/0.906
HCT-FFN*	**37.462/0.961**	**29.330/0.891**	**38.813/0.983**	**33.188/0.932**	**34.543/0.937**	**26.712/0.909**
MPRNet	36.557/0.954	30.760/0.908	39.890/0.985	33.460/0.928	34.240/0.933	25.663/0.909
MPRNet*	**36.854/0.954**	**31.630/0.925**	**40.340/0.987**	**33.850/0.941**	**35.390/0.945**	**26.626/0.920**
SPDNet	37.063/0.951	31.300/0.922	40.590/0.988	33.270/0.919	34.570/0.956	25.433/0.904
SPDNet*	**37.431/0.965**	**31.410/0.926**	**40.620/0.988**	**33.460/0.921**	**34.700/0.957**	**25.821/0.908**
FPNet	37.794/0.962	30.165/0.914	39.946/0.987	33.105/0.923	34.600/0.933	26.005/0.904
FPNet*	**38.014/0.963**	**31.033/0.923**	**40.671/0.988**	**33.373/0.934**	**35.157/0.943**	**27.804/0.923**
Restormer	37.851/0.967	31.392/0.916	40.581/0.987	34.040/0.934	35.201/0.936	28.817/0.929
Restormer*	**38.195/0.968**	**31.567/0.924**	**40.831/0.989**	**34.359/0.944**	**35.482/0.943**	**29.374/0.934**
DRSformer	38.015/0.968	32.173/0.932	41.232/0.989	35.354/0.964	34.354/0.959	29.433/0.931
DRSformer*	**38.245/0.969**	**32.313/0.935**	**41.366/0.989**	**35.419/0.965**	**34.627/0.960**	**29.679/0.932**

Table 1: The PSNR↑ / SSIM↑ performance for twelve de-raining methods with VC-based (bold) and SC-based (**bold***).

Quantitative and Qualitative Results

In Table 1, we report the PSNR/SSIM of ten de-raining baselines on six benchmarks. The performance of all CNN-based and CNN-VIT hybrid methods is improved after utilizing the SC, demonstrating that our proposed SC is not sensitive to the architecture and can be plugged into various networks to boost performance. Note that, CNN-based MPRNet and SPDNet, and CNN-VIT hybrid methods HCT-FFN and Restormer are state-of-the-art de-raining methods. Utilizing our SC, the performance of these networks is further improved, and a new SOTA is reached. For example, the PSNR/SSIM score of Restormer on the Rain12600 dataset improved from 34.04/0.934 to 34.359/0.944. The performance on different de-raining datasets is comprehensively improved, demonstrating that our proposed SC can effectively perceive the properties of rain streaks and accumulation and can be used in diverse rain conditions. Furthermore, all these performance gains are cost-free without introducing extra computational costs. Furthermore, to demonstrate the effectiveness of SC for real rainy images, we also evaluate our SC on three real de-raining datasets, as shown in Table 2. Following (Yi et al. 2021), we use the best pre-trained model to evaluate real rainy images and measure image quality from different perspectives using four no-reference evaluation metrics. The network based on our proposed SC can outperform the original network, demonstrating that the features of rain streaks and accumulation extracted by our SC can effectively be generalized to the real scene. We prove that our method aligns with the original model's structure, parameters, and inferencing time. By comparing inference time, parameters, and FLOPs of VC and SC versions on four classic deraining baselines in Tab. 3, we show that the SC-based models have identical computational costs to VC-based models, demonstrating no additional computational overhead while maintaining superior performance.

Fig. 3 shows the qualitative improvement after using our SC. We observe that the network incorporated with SC enhances the contrast more effectively without rain accumulation residual, as shown in Fig. 3 (a). This demonstrates that our SC accurately models the properties of rain accumulation. In addition, the VC-based PReNet, SPANet, and MPRNet will damage the detail of the background and result in dark regions, and the Restormer tends to over-smooth the tiny details, as shown in the red box of Fig. 3 (b). In contrast, after using our SC, these methods can successfully remove the rain streaks without damaging the structural and textural image details. This is because our SC can effectively identify the rain streak regions and remove them accurately. Fig. 4 shows the qualitative comparison of real rainy images. We observe that the rain streaks can be removed more thoroughly and contrast can be effectively enhanced without introducing any artifacts. This demonstrates the generality of our methods for real scenes.

Ablation Study

To demonstrate the effectiveness of SC in simultaneously extracting rain streaks and accumulation properties, we compare SC with several different settings: RepVGG block used in (Ding et al. 2021), asymmetric convolution used in (Ding et al. 2019), CDC (Yu et al. 2020), two VCs in parallel, VC in parallel with CSD, and VC in parallel with

Figure 3: Qualitative comparison on a rainy image from Outdoor-Rain and Rain200H.

	SIRR-R				JORDER-R				ID-CGAN-R			
	PI↓	NIQE↓	PIQE↓	BRISQUE↓	PI↓	NIQE↓	PIQE↓	BRISQUE↓	PI↓	NIQE↓	PIQE↓	BRISQUE↓
PReNet	3.126	4.209	32.766	31.154	2.458	3.625	31.125	29.015	2.801	3.961	31.860	28.498
PReNet*	**3.003**	**3.787**	**30.204**	**29.872**	**2.221**	**3.312**	**29.039**	**27.450**	**2.714**	**3.551**	**29.387**	**26.127**
MPRNet	3.076	4.120	31.283	30.049	2.418	3.475	29.746	28.383	2.761	3.763	29.718	27.255
MPRNet*	**2.926**	**3.583**	**30.562**	**27.203**	**2.378**	**3.149**	**29.306**	**22.764**	**2.655**	**3.409**	**29.392**	**24.000**
NLEDN	2.835	3.421	26.794	26.501	2.221	3.111	25.664	22.901	2.571	3.306	26.218	22.577
NLEDN*	**2.822**	**3.398**	**24.393**	**24.533**	**2.081**	**3.102**	**23.206**	**21.167**	**2.551**	**3.212**	**24.740**	**22.012**
FPNet	2.697	3.504	28.626	25.706	2.216	3.232	25.132	20.342	2.443	3.368	27.649	20.435
FPNet*	**2.320**	**3.188**	**28.595**	**25.531**	**2.213**	**3.139**	**24.942**	**22.332**	**2.235**	**3.119**	**26.541**	**20.407**
Restormer	2.859	3.344	27.809	24.950	2.364	3.072	25.441	23.243	2.496	2.994	26.202	20.469
Restormer*	**2.836**	**3.326**	**27.588**	**24.741**	**2.353**	**3.069**	**25.138**	**22.545**	**2.488**	**2.813**	**26.163**	**20.285**
RESCAN	3.171	3.841	32.924	27.341	2.606	3.535	31.590	26.325	2.742	3.324	30.462	24.012
RESCAN*	**3.139**	**3.814**	**32.653**	**27.227**	**2.568**	**3.473**	**30.778**	**26.311**	**2.633**	**3.150**	**30.413**	**23.965**

Table 2: Quantitative results on non-reference metrics (including the NIQE, BRISQUE, PIQE, and PI) on three real datasets.

Methods	SPDNet	**Ours**	Restormer	**Ours**
Time(s)	0.042	**0.042(+0)**	0.100	**0.100(+0)**
Param(M)	2.982	**2.982(+0)**	26.09	**26.09(+0)**
FLOPs(G)	6.059	**6.059(+0)**	8.812	**8.812(+0)**
PSNR(dB)	33.7	**33.9(+0.2)**	34.44	**34.8(+0.4)**

Table 3: Compared to VC (bold), using SC (**Ours**) improves performance without extra computational cost.

	VC	VV	VD	VA	ACB	RepB	CDC	**SC**
O	22.2	22.5	22.9	23.0	17.4	15.9	21.9	**23.5**
H	29.0	29.0	29.3	29.4	25.9	24.2	25.4	**29.7**
L	37.1	37.1	37.8	37.3	30.4	26.3	32.4	**38.5**

Table 4: Ablation experiments for our proposed SC.

CSA, denoted as RepB, ACB, VV, VD, and VA, respectively. We choose PReNet as the baseline and replace each VC kernel with the above settings within the network. The experiments are conducted on the Outdoor-Rain (O), Rain200H (H), and Rain200L (L) datasets, and the PSNR results are reported in Table 4. The model's performance is extraordinarily degraded after using ACB, RepB, and CDC. This is probably because they are designed for high-level vision tasks and cannot boost the modeling ability of rain streaks and accumulation simultaneously. We can observe that directly replacing each convolution in the original network with VV may slightly improve the performance. This is because the network parameters have been doubled, and the learning capacity has been improved. However, VV cannot always guarantee that two VCs are optimized towards the same goal, and it will drop the performance, as shown in the results on the Rain200H dataset in Table 4. The performance of VD is higher than that of VV. This is because the CSD can perceive the isotropic directional gradient distribution of rain streaks, providing better rain streak characteristics. Similarly, the CSA can suppress the influence of tiny structures and perceive the isotropic distribution of rain

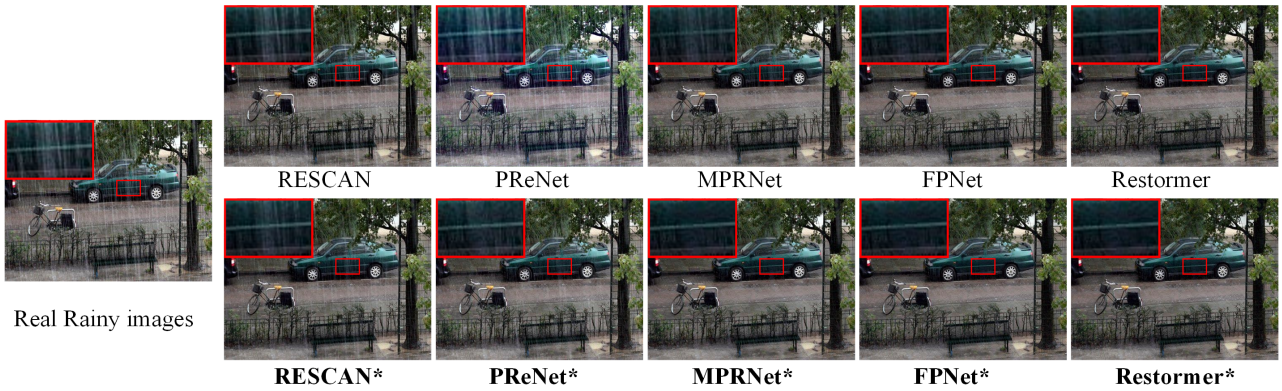

Figure 4: Qualitative comparison on a real rainy image from SIRR-Real (Wei et al. 2019). Zoom for a better view.

	L_1	H_1	H_2	H_3	Image
VV	26.5	44.1	44.5	43.3	28.2
VA	27.2(+0.7)	44.4	44.6	43.4	28.8
VD	26.7	45.3(+1.2)	45.6(+1.1)	44.6(+1.2)	28.9

Table 5: Performance analysis of CSA and CSD. The **bold** represents performance improvement compared to VV.

Tasks	PSNR↑	Time(s)↓	Param(M)↓	FLOPs(G)↓
SR	33.57	0.04	7.24	13.052
	33.99(+0.42)	**0.04**(+0)	**7.24**(+0)	**13.052**(+0)
LIIE	22.31	0.01	0.58	0.002
	22.98 (+0.67)	**0.01**(+0)	**0.58**(+0)	**0.002**(+0)
MD	32.68	0.03	16.14	9.620
	33.27(+0.59)	**0.03**(+0)	**16.14**(+0)	**9.620**(+0)

Table 6: Extend our SC to other low-level vision tasks.

accumulation, which delivers higher performance than VV. Finally, our SC with CSD and CSA achieved the best improvement, which validates the synergistic relationship between CSA and CSD.

Analysis and Discussion

Performance analysis. To verify the effectiveness of CSD and CSA for modeling rain streaks and accumulation-related features, we conduct the following statistical experiment on the deraining results of Outdoor-Rain of Restormer on the settings of VV, VA, and VD in Tab. 4. Based on the fact that rain streaks/accumulation are mainly distributed in the image structure/flat region, we use the Laplacian pyramid to decompose the deraining results into three high-frequency structure components (H_1, H_2, and H_3) and one low-frequency flat component(L_1). Then, we calculate the PSNR between each component and the corresponding GT as shown in Tab. 5. The components of structure and flat are consistently improved by CSA and CSD, which demonstrates the effectiveness of the proposed CSA and CSD.

Extend to other low-level vision tasks. Considering the superiority of SC in simultaneously modeling low-frequency contrast-related representations and high-frequency detail representations, we further explore the potential of SC in other low-level vision tasks, specifically including three representative tasks: **blind image super-resolution (SR)** DASR (Wang et al. 2021) on the Set5, **low-light image enhancement (LIIE)** ENC (Huang et al. 2022) the LOL, and **motion deblurring (MD)** MIMO-UNet (Cho et al. 2021) on the GoPro. As shown in Tab. 6, SC also demonstrates significant performance enhancement without degradation in other low-level vision tasks.

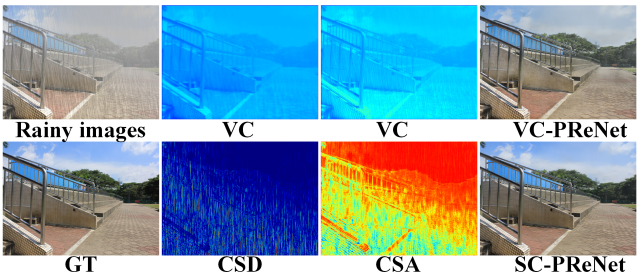

Figure 5: Feature visualization of CSD and CSA.

Conclusion

In this paper, we introduce a novel central-surrounding synergistic convolution (SC) for single image de-raining, which can be used to learn a unified representation for rain streaks and rain accumulation removal. Without introducing extra computational costs, our SC can be plugged into various networks to improve the modeling capability of rain streaks and accumulation. Extensive experiments on various popular de-raining benchmarks, including both synthetic and real, demonstrate that SC can comprehensively improve the performance of twelve existing methods under various rainy conditions. Even for the current SOTA deraining networks, SC can further achieve SOTA++ performance without introducing extra computational costs.

In future work, considering that the integration of SC with current VC-based architectures might not be optimal, we plan to develop a novel architecture tailored for our SC to fully leverage its potential in image de-raining, and we will further enhance SC to expand its impact in low-level vision.

Acknowledgments

This work was supported by the Natural Science Foundation of China under Grants 62225207,62436008 and 62206262.

References

Chao-Yi, L.; Xing, P.; Yi-Xiong, Z.; et al. 1991. Role of the extensive area outside the X-cell receptive field in brightness information transmission. *Vision Research*, 31(9): 1529–1540.

Chen, H.; Chen, X.; Wu, C.; Zheng, Z.; Pan, J.; and Fu, X. 2024. Towards Ultra-High-Definition Image Deraining: A Benchmark and An Efficient Method. *arXiv preprint arXiv:2405.17074*.

Chen, S.; Ye, T.; Bai, J.; Chen, E.; Shi, J.; and Zhu, L. 2023a. 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 (ICCV)*, 13106–13117.

Chen, X.; Li, H.; Li, M.; and Pan, J. 2023b. Learning A Sparse Transformer Network for Effective Image Deraining. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5896–5905.

Chen, X.; Pan, J.; Lu, J.; Fan, Z.; and Li, H. 2023c. Hybrid cnn-transformer feature fusion for single image deraining. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, 378–386.

Chen, Z.; He, Z.; and Lu, Z.-M. 2024. DEA-Net: Single image dehazing based on detail-enhanced convolution and content-guided attention. *IEEE Transactions on Image Processing*.

Cho, S.-J.; Ji, S.-W.; Hong, J.-P.; Jung, S.-W.; and Ko, S.-J. 2021. Rethinking coarse-to-fine approach in single image deblurring. In *Proceedings of the IEEE/CVF international conference on computer vision*, 4641–4650.

Ding, X.; Guo, Y.; Ding, G.; and Han, J. 2019. Acnet: Strengthening the kernel skeletons for powerful cnn via asymmetric convolution blocks. In *Proceedings of the IEEE/CVF international conference on computer vision*, 1911–1920.

Ding, X.; Zhang, X.; Ma, N.; Han, J.; Ding, G.; and Sun, J. 2021. Repvgg: Making vgg-style convnets great again. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 13733–13742.

Fu, X.; Huang, J.; Zeng, D.; Huang, Y.; Ding, X.; and Paisley, J. 2017. Removing rain from single images via a deep detail network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3855–3863.

Gu, Y.; Wang, C.; and Li, J. 2023. Incremental image deraining via associative memory. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, 685–693.

Guo, Q.; Sun, J.; Juefei-Xu, F.; Ma, L.; Xie, X.; Feng, W.; Liu, Y.; and Zhao, J. 2021. Efficientderain: Learning pixel-wise dilation filtering for high-efficiency single-image deraining. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, 1487–1495.

Guo, X.; Fu, X.; Zhou, M.; Huang, Z.; Peng, J.; and Zha, Z.-J. 2022. Exploring fourier prior for single image rain removal. In *Proceedings of the 30th International Joint Conferences on Artificial Intelligence*, 935–941.

Guo, Y.; Xiao, X.; Chang, Y.; Deng, S.; and Yan, L. 2023. From Sky to the Ground: A Large-scale Benchmark and Simple Baseline Towards Real Rain Removal. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 12097–12107.

He, K.; Gkioxari, G.; Dollár, P.; and Girshick, R. 2017. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, 2961–2969.

Huang, J.; Liu, Y.; Fu, X.; Zhou, M.; Wang, Y.; Zhao, F.; and Xiong, Z. 2022. Exposure normalization and compensation for multiple-exposure correction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 6043–6052.

Jiang, K.; Liu, W.; Wang, Z.; Zhong, X.; Jiang, J.; and Lin, C.-W. 2023. Dawn: Direction-aware attention wavelet network for image deraining. In *Proceedings of the 31st ACM international conference on multimedia*, 7065–7074.

Kang, L.-W.; Lin, C.-W.; and Fu, Y.-H. 2011. Automatic single-image-based rain streaks removal via image decomposition. *IEEE transactions on image processing*, 21(4): 1742–1755.

Li, G.; He, X.; Zhang, W.; Chang, H.; Dong, L.; and Lin, L. 2018a. Non-locally enhanced encoder-decoder network for single image de-raining. In *Proceedings of the 26th ACM international conference on Multimedia*, 1056–1064.

Li, R.; Cheong, L.-F.; and Tan, R. T. 2019. Heavy rain image restoration: Integrating physics model and conditional adversarial learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1633–1642.

Li, X.; Wu, J.; Lin, Z.; Liu, H.; and Zha, H. 2018b. Recurrent squeeze-and-excitation context aggregation net for single image deraining. In *Proceedings of the European conference on computer vision (ECCV)*, 254–269.

Li, Y.; Tan, R. T.; Guo, X.; Lu, J.; and Brown, M. S. 2016. Rain streak removal using layer priors. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2736–2744.

Lin, B.; Jin, Y.; Yan, W.; Ye, W.; Yuan, Y.; Zhang, S.; and Tan, R. T. 2024. NightRain: Nighttime Video Deraining via Adaptive-Rain-Removal and Adaptive-Correction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, 3378–3385.

Peng, L.; Cao, Y.; Pei, R.; Li, W.; Guo, J.; Fu, X.; Wang, Y.; and Zha, Z.-J. 2024a. Efficient Real-world Image Super-Resolution Via Adaptive Directional Gradient Convolution. *arXiv preprint arXiv:2405.07023*.

Peng, L.; Cao, Y.; Sun, Y.; and Wang, Y. 2024b. Lightweight Adaptive Feature De-drifting for Compressed Image Classification. *IEEE Transactions on Multimedia*.

Peng, L.; Li, W.; Pei, R.; Ren, J.; Wang, Y.; Cao, Y.; and Zha, Z.-J. 2024c. Towards Realistic Data Generation for Real-World Super-Resolution. *arXiv preprint arXiv:2406.07255*.

Ren, D.; Zuo, W.; Hu, Q.; Zhu, P.; and Meng, D. 2019. Progressive image deraining networks: A better and simpler baseline. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 3937–3946.

Ronneberger, O.; Fischer, P.; and Brox, T. 2015. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, 234–241. Springer.

Wang, C.; Pan, J.; Wang, W.; Dong, J.; Wang, M.; Ju, Y.; and Chen, J. 2023. PromptRestorer: A Prompting Image Restoration Method with Degradation Perception. In *NeurIPS*.

Wang, C.; Pan, J.; Wang, W.; Fu, G.; Liang, S.; Wang, M.; Wu, X.-M.; and Liu, J. 2024. Correlation Matching Transformation Transformers for UHD Image Restoration. In *AAAI*, volume 38, 5336–5344.

Wang, L.; Wang, Y.; Dong, X.; Xu, Q.; Yang, J.; An, W.; and Guo, Y. 2021. Unsupervised degradation representation learning for blind super-resolution. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 10581–10590.

Wang, T.; Yang, X.; Xu, K.; Chen, S.; Zhang, Q.; and Lau, R. W. 2019a. 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*, 12270–12279.

Wang, Y.; Cao, Y.; Zha, Z.-J.; Zhang, J.; Xiong, Z.; Zhang, W.; and Wu, F. 2019b. Progressive retinex: Mutually reinforced illumination-noise perception network for low-light image enhancement. In *Proceedings of the 27th ACM international conference on multimedia*, 2015–2023.

Wang, Y.; Ma, C.; and Liu, J. 2023. SmartAssign: Learning a Smart Knowledge Assignment Strategy for Deraining and Desnowing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 3677–3686.

Wei, W.; Meng, D.; Zhao, Q.; Xu, Z.; and Wu, Y. 2019. Semi-supervised transfer learning for image rain removal. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 3877–3886.

Wen, Y.; Gao, T.; Zhang, J.; Zhang, K.; and Chen, T. 2024. From heavy rain removal to detail restoration: A faster and better network. *Pattern Recognition*, 148: 110205.

Yang, H.; Chen, Q.; Fu, K.; Zhu, L.; Jin, L.; Qiu, B.; Ren, Q.; Du, H.; and Lu, Y. 2022a. Boosting medical image segmentation via conditional-synergistic convolution and lesion decoupling. *Computerized Medical Imaging and Graphics*, 101: 102110.

Yang, W.; Tan, R. T.; Feng, J.; Liu, J.; Guo, Z.; and Yan, S. 2017. Deep joint rain detection and removal from a single image. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1357–1366.

Yang, W.; Tan, R. T.; Wang, S.; Fang, Y.; and Liu, J. 2020. Single image deraining: From model-based to data-driven and beyond. *IEEE Transactions on pattern analysis and machine intelligence*, 43(11): 4059–4077.

Yang, W.; Tan, R. T.; Wang, S.; Kot, A. C.; and Liu, J. 2022b. Learning to Remove Rain in Video With Self-Supervision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Yasarla, R.; Sindagi, V. A.; and Patel, V. M. 2020. Syn2real transfer learning for image deraining using gaussian processes. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2726–2736.

Yi, Q.; Li, J.; Dai, Q.; Fang, F.; Zhang, G.; and Zeng, T. 2021. Structure-preserving deraining with residue channel prior guidance. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 4238–4247.

Yu, Z.; Zhao, C.; Wang, Z.; Qin, Y.; Su, Z.; Li, X.; Zhou, F.; and Zhao, G. 2020. Searching central difference convolutional networks for face anti-spoofing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5295–5305.

Zamir, S. W.; Arora, A.; Khan, S.; Hayat, M.; Khan, F. S.; and Yang, M.-H. 2022. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5728–5739.

Zamir, S. W.; Arora, A.; Khan, S.; Hayat, M.; Khan, F. S.; Yang, M.-H.; and Shao, L. 2021. Multi-stage progressive image restoration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 14821–14831.

Zhang, F.; You, S.; Li, Y.; and Fu, Y. 2023. Learning Rain Location Prior for Nighttime Deraining. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 13148–13157.

Zhang, H.; and Patel, V. M. 2018. Density-aware single image de-raining using a multi-stream dense network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 695–704.

Zhang, H.; Sindagi, V.; and Patel, V. M. 2019. Image deraining using a conditional generative adversarial network. *IEEE transactions on circuits and systems for video technology*, 30(11): 3943–3956.

Zhang, R.; Yu, J.; Chen, J.; Li, G.; Lin, L.; and Wang, D. 2024. A Prior Guided Wavelet-Spatial Dual Attention Transformer Framework for Heavy Rain Image Restoration. *IEEE Transactions on Multimedia*.

Zhao, H.; Zhang, J.; Chen, Z.; Zhao, S.; and Tao, D. 2024. Unimix: Towards domain adaptive and generalizable lidar semantic segmentation in adverse weather. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 14781–14791.

Zheng, S.; Lu, C.; and Narasimhan, S. G. 2024. TPSeNCE: Towards Artifact-Free Realistic Rain Generation for Deraining and Object Detection in Rain. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 5394–5403.

Zou, W.; Wang, Y.; Fu, X.; and Cao, Y. 2022. Dreaming To Prune Image Deraining Networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 6023–6032.