

Structure-Preserving Deraining with Residue Channel Prior Guidance

Qiaosi Yi¹, Juncheng Li¹, Qinyan Dai¹, Faming Fang^{1,2*}, Guixu Zhang¹, and Tiejong Zeng³

¹School of Computer Science and Technology, East China Normal University, Shanghai, China

²Key Laboratory of Advanced Theory and Application in Statistics
and Data Science, Ministry of Education

³Department of Mathematics, The Chinese University of Hong Kong, Hong Kong, China

qiaosiyijoyies@gmail.com, cvjunchengli@gmail.com, 649310204@qq.com

fmfang@cs.ecnu.edu.cn, gxzhang@cs.ecnu.edu.cn, zeng@math.cuhk.edu.hk

Abstract

Single image deraining is important for many high-level computer vision tasks since the rain streaks can severely degrade the visibility of images, thereby affecting the recognition and analysis of the image. Recently, many CNN-based methods have been proposed for rain removal. Although these methods can remove part of the rain streaks, it is difficult for them to adapt to real-world scenarios and restore high-quality rain-free images with clear and accurate structures. To solve this problem, we propose a Structure-Preserving Deraining Network (SPDNet) with RCP guidance. SPDNet directly generates high-quality rain-free images with clear and accurate structures under the guidance of RCP but does not rely on any rain-generating assumptions. Specifically, we found that the RCP of images contains more accurate structural information than rainy images. Therefore, we introduced it to our deraining network to protect structure information of the rain-free image. Meanwhile, a Wavelet-based Multi-Level Module (WMLM) is proposed as the backbone for learning the background information of rainy images and an Interactive Fusion Module (IFM) is designed to make full use of RCP information. In addition, an iterative guidance strategy is proposed to gradually improve the accuracy of RCP, refining the result in a progressive path. Extensive experimental results on both synthetic and real-world datasets demonstrate that the proposed model achieves new state-of-the-art results. Code: <https://github.com/Joyies/SPDNet>

1. Introduction

Single image deraining (SID) aims to reconstruct a visually pleasing image from its corresponding rain-streaks-degraded image. In the past years, various methods have been proposed for SID. For traditional methods, many researchers [11, 1, 3, 4, 41, 24, 27, 12] focused on exploring

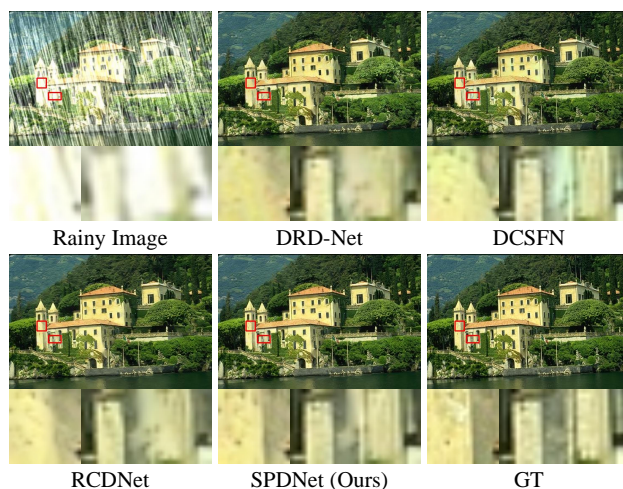


Figure 1: An example of image deraining results. Obviously, SPDNet can reconstruct high-quality image with clear and accurate structure.

the physical properties of the rain and background layers. Meanwhile, various priors have been proposed to regularize and separate them, such as layer priors with Gaussian mixture model (GMM) [24], discriminative sparse coding (DSC) [27], and joint convolutional analysis and synthesis sparse representation (JCAS) [12]. Although these methods can effectively remove the rain streaks, they require complex iterative optimization to find the best solution.

Recently, convolutional neural networks (CNN) have achieved significant success in many computer vision tasks [32, 13, 18, 29]. Moreover, many CNN-based methods [7, 8, 43, 23, 49, 15, 38, 50, 31, 45, 6, 37, 16] have been proposed for rain removal, such as DDN [8], RESCAN [23], PReNet [31], DRDNet [6], and RCDNet [37]. Albeit these methods have brought great performance improvements, they are difficult to remove all rain streaks and recover the structural information of images in complex scenarios (Fig. 1). This is because: (1). Most of them di-

*Corresponding author



A. Rainy Image

B. RCP of A

Figure 2: B is the residue channel prior (RCP) extracted from the rainy image. Obviously, even the RCP is extracted from the rainy image, it still contains clear structures.

rectly learn the mapping between the rainy image and the rain streaks layers. In other words, most methods predict rain streaks via the built CNN model and then subtract rain streaks from rainy images to get the final output. However, the density of rain streaks varies, which leads to excessive or insufficient removal of rain streaks, resulting in incomplete structural information of the reconstructed images. (2) These methods focus on learning the structure of rain streaks, but they pay less attention to learning the structure of objects and ignore the importance of image prior.

To address the aforementioned issues, we aim to explore an image prior that can protect the structure of the image, and introduce the prior into the model to guide high-quality image reconstruction. According to our investigation, we found that the Total Variation (TV) prior will smooth texture details in the restored images, the sparse prior is usually difficult to model because it requires other domain knowledge, and the edge prior is difficult to obtain from the rainy image since the off-the-shelf edge detectors are sensitive to the rain streaks. In contrast, residue channel prior (RCP [19, 21, 20]) show clear structures even extracted from the rainy image (Fig. 2). Moreover, compared with layer prior [27] and [15], RCP is the residual result of the maximum channel value and minimum channel value of the rainy image, calculated without any additional parameters. Hence, we adopt RCP to the image deraining task and propose a Structure-Preserving Deraining Network (SPDNet) with residue channel prior (RCP) guidance. SPDNet pays more attention to learning the background information and directly generates high-quality rain-free images with clear and accurate structures. To achieve this, an RCP guidance network and an iterative guidance strategy are proposed for structure-preserving deraining. Specifically, we design a wavelet-based multi-level module (WMLM) as the backbone of SPDNet to fully learn the background information in the rainy image. Meanwhile, an RCP extraction module and an interactive fusion module (IFM) are designed for RCP extraction and guidance, respectively. This is also the most important step in the SPDNet, which can highlight the structure of objects in the rainy images and promote generating high-quality rain-free images. In addition, we found that the RCP improves as image quality improves.

Therefore, we propose a progressive reconstruction method, which means that the model extracts more accurate RCPs in the intermediate reconstruction stage to further guide image reconstruction. Under the iterative guidance of RCP, SPDNet can reconstruct high-quality rain-free images.

The main contributions of this paper are as follows:

- We explore the importance of residue channel prior (RCP) for rain removal and propose a Structure-Preserving Deraining Network (SPDNet) with RCP guidance. Extensive experimental results show that SPDNet achieves new state-of-the-art results.
- We propose an RCP extraction module and an Interactive Fusion Module (IFM) for RCP extraction and guidance, respectively. Meanwhile, an iterative guidance strategy is designed for progressive image reconstruction.
- We design a Wavelet-based Multi-Level Module (WMLM) as the backbone of SPDNet to learn the background of the area covered by the rain streak.

2. Related Work

In the past few years, many excellent rain removal methods have been proposed, which greatly promoted the development of image deraining. This part mainly focuses on several classical single image rain removal methods. A more detailed introduction can be found in [44, 22, 36].

2.1. Traditional Methods

Traditional methods mainly use manually extracted features and priors to describe the features of rain streaks. For example, Kang *et al.* [17] proposed a method that can decompose an image into low- and high-frequency parts by using a bilateral filter. Luo *et al.* [27] presented a discriminative sparse coding for separating rain streaks from the rainy image. Li *et al.* [24] uses Gaussian mixture models (GMM) as the prior to separate the rain streaks. Wang *et al.* [39] combines image decomposition and dictionary learning to remove rain or snow from images. Zhu *et al.* [52] detects rain-dominant regions. The detected regions are utilized as a guidance image to help separate rain streaks from the background layer.

2.2. CNN-based Methods

Recently, many CNN-based image deraining networks have been proposed [46, 40, 35, 42, 44, 47, 16, 30, 33, 9], and they have greatly promoted the development of this field. For example, Fu *et al.* [7] proposed the firstly CNN-based model to remove rain streaks. Later, they [8] also proposed a deeper network based on a deep residual network and use the image domain knowledge to remove rain streaks. Yang *et al.* [43] proposed a recurrent deep network

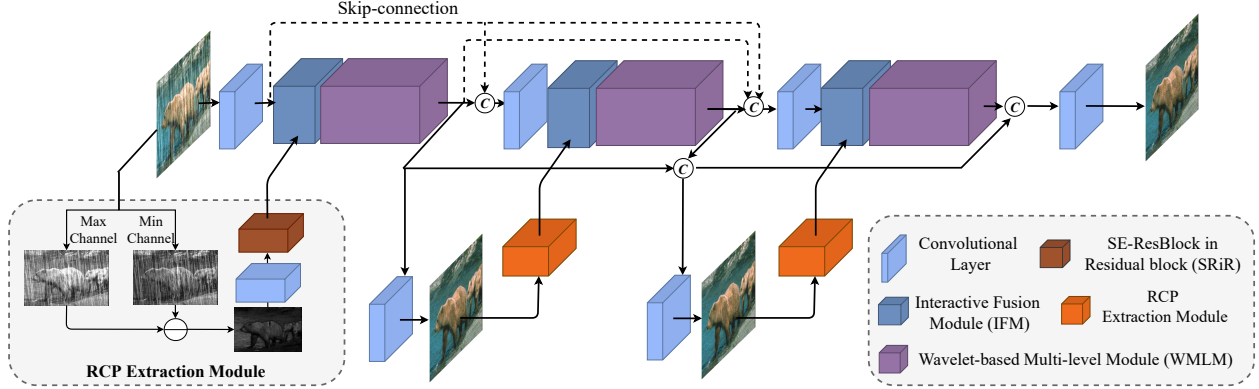


Figure 3: The overall architecture of the proposed Structure-Preserving Deraining Network (SPDNet).

for joint rain detection and removal to progressively remove rain streaks. Li *et al.* [23] proposed a recurrent squeeze-and-excitation context aggregation network to make full use of contextual information. Deng *et al.* [6] introduced a detail-recovery image deraining network, which is composed of a rain residual network and a detail repair network. Jiang *et al.* [16] designed a coarse-to-fine progressive fusion network and Wang *et al.* [34] proposed a cross-scale fusion network based on inner-scale connection block. Moreover, Ren *et al.* [31] proposed a progressive recurrent network (PReNet) by repeatedly unfolding a shallow ResNet. Wang *et al.* [37] introduced a rain convolutional dictionary network, which continuously optimizes the model to obtain better rain streaks and rain-free images.

Although these methods can remove part of the rain streaks, they still cannot effectively remove all the rain streaks in complex situations. Meanwhile, these methods can hardly protect the structural information of the image thus cannot reconstruct high-quality rain-free images. Therefore, a method that can effectively remove rain streaks and protect the structure of the object is essential.

3. Structure-Preserving Deraining Network

In this paper, we propose a Structure-Preserving Deraining Network (SPDNet). As shown in Fig. 3, SPDNet uses the wavelet-based feature extraction backbone as the main structure and introduces a residue channel prior (RCP) guided mechanism for structure-preserving deraining. In SPDNet, the wavelet-based feature extraction backbone is designed for background information learning, which consists of a series of wavelet-based multi-level modules (WMLM). However, it is difficult to maintain clear structures of the reconstructed image by only using the wavelet-based feature extraction backbone. Hence, RCP is introduced to the model to provide additional auxiliary information, so that the structural information of the reconstructed image is protected. More details will be introduced in the following sections.

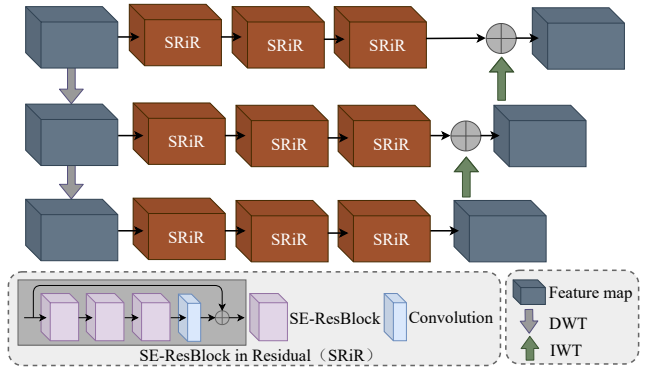


Figure 4: The architecture of the proposed Wavelet-based Multi-Level Module (WMLM).

3.1. Wavelet-based Feature Extraction Backbone

Due to the arbitrary size and density of the rain streaks, the occluded region and the degree of occlusion of the object in the rainy image are unknown. To solve this problem, many methods [16, 34, 48] adopt the multi-level strategy to fully learning the features under different scales. Following these methods, we also adopt the multi-level strategy in our model. However, directly using the downsampling operation or deconvolution operation will cause a lot of information to be lost. Therefore, we propose a wavelet-based multi-level module (WMLM), which adopts the discrete wavelet transform (DWT) and Inverse DWT (IWT) in place of the simple downsampling and deconvolution operations. Moreover, DWT can capture both frequency and location information of feature maps [26, 5, 25], which may be helpful in preserving detailed textures. The architecture of WMLM is shown in Fig. 4.

In WMLM, we firstly use DWT to obtain multiple rainy image features with different scales, and adopt a convolution to resize the feature channel ($\tilde{\mathcal{F}}_0, \tilde{\mathcal{F}}_1, \tilde{\mathcal{F}}_2$) as:

$$\begin{cases} \tilde{\mathcal{F}}_i = \tilde{\mathcal{F}}, & \text{if } i = 0, \\ \tilde{\mathcal{F}}_i = \text{Conv}(\text{DWT}(\tilde{\mathcal{F}}_{i-1})), & \text{if } i > 0, \end{cases} \quad (1)$$

where $\tilde{\mathcal{F}}$ denotes the input of WMLM. Next, we adopt three SE-ResBlock in Residual block (SRiR) [51] to learn the background information of the rainy image. As shown in Fig. 4, SRiR is similar to the ResBlock [13], which uses the SE-Resblock [14] to replace the convolution in ResBlock.

$$\tilde{\mathcal{F}}_i^s = SRiR(\tilde{\mathcal{F}}_i), \quad i = 0, 1, 2. \quad (2)$$

Lastly, we use a convolution to resize the smaller size features channel and up-sampled it by IWT. Meanwhile, they are added to the output of the previous level:

$$\tilde{\mathcal{F}}_{i-1}^s = IWT(Conv(\tilde{\mathcal{F}}_i^s)) + \tilde{\mathcal{F}}_{i-1}^s, \quad i = 2, 1, \quad (3)$$

With the help of WMLM, the model can effectively learn the background information of rainy images and reconstruct relatively clear images.

3.2. RCP Guided Structure-Preserving Deraining

Although the basic model can reconstruct relatively clear rain-free images, it is found that the object structure in the reconstructed image has also been damaged. In order to solve this problem, we recommend introducing additional image priors to protect the object structure. Therefore, we suggest introducing residue channel prior (RCP) to the model to achieve structure-preserving deraining. In addition, an Interactive Fusion Module (IFM) and an iterative guidance strategy are proposed to make full use of RCP information, so that high-quality rain-free images can be reconstructed.

3.2.1 Residue Channel Prior

As shown in Fig. 2, RCP is free of rain streaks and contains only a transformed version of the background details, which is the residual result of the maximum channel value and minimum channel value of the rainy image.

According to Li *et al.* [19], the colored-image intensity of a rain streak image can be expressed as:

$$\tilde{\mathcal{O}}(x) = t\beta_{rs}(x)\mathcal{B}\alpha + (T - t)\mathcal{R}\pi, \quad (4)$$

where $\tilde{\mathcal{O}}(x)$ is the color vector representing the colored intensity. β_{rs} is composed of refraction, specular reflection, and internal reflection coefficients of raindrops [10]. $\mathbf{B} = (B_r, B_g, B_b)^T$ represents light brightness and $\mathcal{B} = B_r + B_g + B_b$. $\mathbf{R} = (R_r, R_g, R_b)^T$ represents background reflection and $\mathcal{R} = R_r + R_g + R_b$. $\alpha = \mathbf{B}/\mathcal{B}$ and $\pi = \mathbf{R}/\mathcal{R}$ represents the chromaticities of \mathbf{B} and \mathbf{R} , respectively. T is the exposure time and t is the time for a raindrop to pass through pixel x . In the Eq.(4), the first term is the rain streak term and the second term is the background term. When employing any an existing color constancy algorithm to estimate α , we can get the following normalization step:

$$\mathcal{O}(x) = \frac{\tilde{\mathcal{O}}(x)}{\alpha} = O_{rs}(x)\mathbf{i} + O_{bg}(x), \quad (5)$$

where $\mathbf{i} = (1, 1, 1)^T$, $O_{rs} = t\beta_{rs}\mathcal{B}$, and $O_{bg} = (T - t)\mathcal{R}/\alpha$. The vector division is the element-wise division. When we normalize the image, the light chromaticity will be cancelled and the color effect of the spectral sensitivities will also be cancelled. Hence, based on Eq.(5), given a rain image \mathcal{O} , the residue channel prior \mathcal{P} of \mathcal{O} can be defined as:

$$\mathcal{P}(x) = \max_{c \in r, g, b} \mathcal{O}^c(x) - \min_{d \in r, g, b} \mathcal{O}^d(x). \quad (6)$$

Due to the rain streak term in Eq.(4) is achromatic, whose values are canceled when employing color constancy, residue channel prior can be free from rain streaks, as shown in Fig. 2. Moreover, according to Li *et al.* [19], without employing color constancy, the residue channel prior can still work. Since the dominant gray atmospheric light generated by a cloudy sky, the appearance of rain streaks is already achromatic in most cases. Based on this observation, RCP can extract a more complete and accurate object structure. Therefore, RCP has been introduced to the model for image deraining guidance.

3.2.2 RCP Extraction Module

In order to extract high-dimensional features of the RCP, we propose an RCP extraction module. As shown in Fig. 3, a convolution is firstly used to obtain the initial feature maps \mathcal{F}_p^{init} of \mathcal{P} . Furthermore, to reduce the noise in the initial features and enrich the semantic information of the feature, we use SE-ResBlock in Residual block (SRiR) to extract the deeper feature \mathcal{F}_p of RCP.

3.2.3 Interactive Fusion Module

Although RCP has clearer structural information than the rainy image, how to make full use of the RCP features to guide the model is still a challenging task. One simple solution is concatenating RCP features with image features together. However, this is not effective to guide the model deraining and may cause feature interference. To solve this problem, we propose an interactive fusion module (IFM) to progressively combine features together. As shown in Fig. 5, two convolution with 3×3 kernel size to map the rainy image feature \mathcal{F}_o and the RCP feature \mathcal{F}_p to $\hat{\mathcal{F}}_o$ and $\hat{\mathcal{F}}_p$.

$$\hat{\mathcal{F}}_o = Conv(\mathcal{F}_o), \quad (7)$$

$$\hat{\mathcal{F}}_p = Conv(\mathcal{F}_p), \quad (8)$$

Next, we calculate the similarity map \mathcal{S} between $\hat{\mathcal{F}}_o$ and $\hat{\mathcal{F}}_p$ by using element multiplication to enhance the background information of the rainy image destroyed by the rain streaks. Moreover, because the background of RCP is similar to the rainy image, the similarity map \mathcal{S} can also highlight the feature information in the prior features, thereby

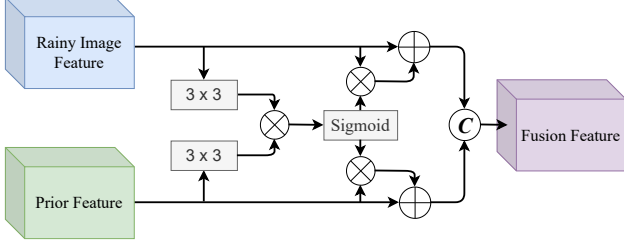


Figure 5: The architecture of the propose Interactive Fusion Module (IFM).

further strengthening the similar structure of the prior feature.

$$\mathcal{S} = \text{Sigmoid}(\widehat{\mathcal{F}}_o \otimes \widehat{\mathcal{F}}_p), \quad (9)$$

$$\mathcal{F}_o^s = \mathcal{S} \otimes \mathcal{F}_o, \quad (10)$$

$$\mathcal{F}_p^s = \mathcal{S} \otimes \mathcal{F}_p, \quad (11)$$

where \mathcal{F}_o^s denotes the activated feature of \mathcal{F}_o and \mathcal{F}_p^s denotes the activated feature of \mathcal{F}_p . Finally, the activated feature is add with the original feature and the output $\widetilde{\mathcal{F}}$ of IFM can be obtained by concatenating the features after addition:

$$\widetilde{\mathcal{F}}_o = \mathcal{F}_o^s + \mathcal{F}_o, \quad (12)$$

$$\widetilde{\mathcal{F}}_p = \mathcal{F}_p^s + \mathcal{F}_p, \quad (13)$$

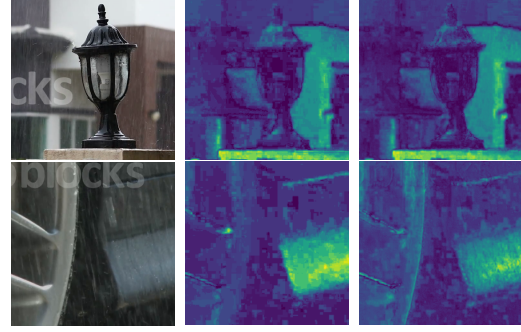
$$\widetilde{\mathcal{F}} = \text{Concat}(\widetilde{\mathcal{F}}_o, \widetilde{\mathcal{F}}_p), \quad (14)$$

3.2.4 Iterative Guidance Strategy

As shown in Fig. 6, with the improvement of image quality, the extracted RCP also shows better results. Based on this observation, we propose an iterative guidance strategy to obtain a clearer RCP and replace the RCP of rainy images, that is, a rain removal result \mathcal{B}_n can be obtained from the $WMLM_n$ output feature \mathcal{F}_n and the clearer structure of prior \mathcal{P}_{n+1} can be get from the \mathcal{B}_n .

$$\begin{cases} \mathcal{F}_n = WMLM_n(\widehat{\mathcal{F}}_n, \mathcal{P}_n; \theta_n), \\ \mathcal{B}_n = \text{Conv}(\mathcal{F}_n), \\ \mathcal{P}_{n+1} = \text{PEB}(\mathcal{B}_n), \quad \text{if } n = 1, 2, \end{cases} \quad (15)$$

where the θ_n represents the weights of $WMLM_n$, Conv stands of the output convolution, and n is the number of $WMLM$. In our proposed model, n is set to 3. The $\widehat{\mathcal{F}}_n$ is input feature map of the $WMLM_n$. Moverover, to enable the model to learn richer features and generate better results, we leverage the idea of *Ensemble Learning* to adopt the concat operation to fuse the output features of the previous WMLM with the current WMLM output feature.



(a) (b) (c)

Figure 6: Comparison between the RCP of rainy images and the RCP of output results. (a) is rainy images, (b) is the RCP of rainy images, and (c) is the RCP of output results. It is obvious observed that the structure of RCP of output results is more obvious than rainy images.

3.3. Loss Function

We employ \mathcal{L}_2 loss as our objective function. As mentioned above, under the iterative update strategy, the model will output three results. Thus, the comprehensive loss function of our proposed model can be formulated as:

$$\mathcal{L} = \sum_i \left\| \mathcal{B}_i - \widehat{\mathcal{B}} \right\|^2, \quad i = 1, 2, 3. \quad (16)$$

Where, $\widehat{\mathcal{B}}$ denotes the rain-free image(GT) and \mathcal{B}_i denotes output results in different stage.

4. Experiment and Discussions

4.1. Datasets

Synthetic Datasets: In our experiment, the *Rain200L/H*, *Rain800*, and *Rain1200* are used to validate our proposed method. The *Rain200L/H* are proposed by Yang *et al.* [43], which consists of 1800 training images and 200 test images. Zhang *et al.* [50] collect a synthetic dataset named *Rain800*, which contains 700 training images and 100 testing images. The *Rain1200* [49] consists of rainy images of different densities, including 12000 training sets and 1200 test sets.

Real-world Datasets: We use two real-word datasets to analyze the performance of all methods. The first one is the *SPA-Data* [38], which contains 638492 rainy/clear image pairs for training and the second one is supplied by Zhang *et al.* [50], which does not contain clear images.

Implementation Details: In our baseline, the number of WMLM is set to 3 and the number of RCP extraction module is set to 3. We use Adam optimizer with batch size of 16 and patch size of 128×128 for training on one NVIDIA Titan Xp GPU. The learning rate is set to 5×10^{-4} and we train the model for 300 epochs for *Rain200L/H*, *Rain800*,

Methods	Param	Time 128 × 128	Rain200L		Rain200H		Rain800		Rain1200		SPA-Data	
			PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
GMM[24]	—	27.961s	28.66	0.8652	14.50	0.4164	25.71	0.8020	25.81	0.8344	34.30	0.9428
DSC[27]	—	7.947s	27.16	0.8663	14.73	0.3815	22.61	0.7530	24.24	0.8279	34.95	0.9416
DDN[8]	0.06M	0.278s	34.68	0.9671	26.05	0.8056	25.87	0.8018	30.97	0.9116	36.16	0.9463
RESCAN[23]	0.15M	0.016s	36.09	0.9697	26.75	0.8353	26.58	0.8726	33.38	0.9417	38.11	0.9707
PreNet[31]	0.17M	0.012s	37.70	0.9842	29.04	0.8991	27.06	0.9026	33.17	0.9481	40.16	0.9816
DCSFN[34]	6.45M	0.253s	39.37	0.9854	29.25	0.9075	28.38	0.9072	<u>34.31</u>	<u>0.9545</u>	—	—
DRDNet[6]	2.72M	0.069s	39.05	0.9862	29.15	0.8921	28.21	0.9012	<u>34.02</u>	<u>0.9515</u>	40.89	0.9784
RCDNet[37]	3.17M	0.068s	<u>39.87</u>	<u>0.9875</u>	<u>30.24</u>	<u>0.9098</u>	<u>28.59</u>	<u>0.9137</u>	34.08	0.9532	<u>41.47</u>	<u>0.9834</u>
SPDNet(Ours)	3.04M	0.055s	40.59	0.9880	31.30	0.9217	30.21	0.9152	34.57	0.9561	43.55	0.9875

Table 1: Quantitative experiments evaluated on four recognized synthetic datasets. The best and the second best results have been boldfaced and underlined.

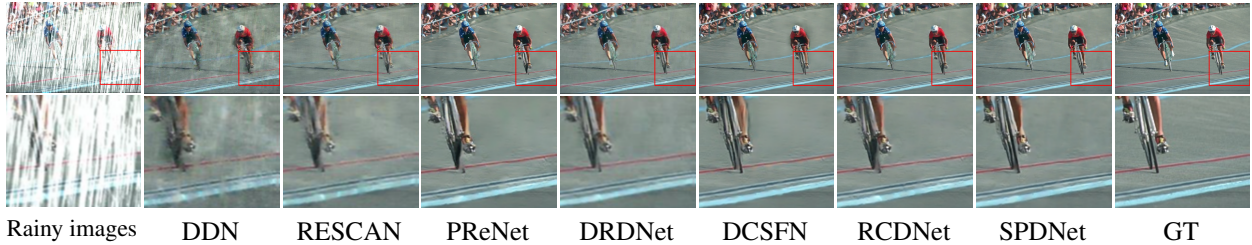


Figure 7: Image deraining results tested in the synthetic datasets. The first row is rainy image, the output of different methods, and GT. The second row is the zoom results of the red window. It is obvious that SPDNet can reconstruct rain-free image with clearer structure.

	RESCAN	PreNet	DRDNet	DCSFN	RCDNet	SPDNet
NIQE↓	3.7774	3.5891	3.8719	3.5326	3.5567	3.4603
PI↓	2.8069	2.7045	2.8980	2.6427	2.6946	2.6254

Table 2: Performance comparison on real-world dataset. ↓ means the better methods should achieve lower score.

and *Rain1200* datasets. For *SPA-Data*, we train the model for 6 epochs and the learning rate is set to 5×10^{-4} .

4.2. Comparison with the State-of-the-Arts

4.2.1 Synthesized Images

We compare our proposed method with the state-of-the-art single image deraining methods. Following RCDNet [37], we compute PSNR and SSIM in YCbCr space. Quantitative results are shown in Table 1 and the processing time of each method is an average time, which is calculated on 100 images of 128×128 size. It can be seen that SPDNet achieves remarkable improvements over these state-of-the-art methods. This substantiates the flexibility and generality of our proposed method in diverse rain types contained in these datasets. Furthermore, from Table 1, SPDNet is not only superior in performance but also superior to RCDNet [37] and DCSFN [34] in terms of parameter and processing time.

Fig. 7 illustrates the deraining performance of all competing methods on synthetic datasets. As shown, the deraining result of our proposed is better than that of other

methods in sufficiently removing the rain streaks and finely recovering the image textures. For other comparison methods, they tend to blur the image textures, or still leave some visible rain streaks.

4.2.2 Real-world Images

In order to demonstrate that our method is also applicable in real-world scenarios, we compare with other methods on SPA-Data and a real-world dataset. Table 1 compare the results on SPA-Data of all competing methods quantitatively. As shown, it is easy to see that SPDNet achieves an evident superior performance than other methods. Moreover, since the real-world dataset supplied by Zhang *et al.* [50] does not have the corresponding label, to validate the deraining performance, we use two non-reference indicators (NIQE [28] and PI [2]) to test the performance in this dataset. The results are shown in Table 2. It is obvious that SPDNet achieves the best results, which shows that SPDNet is not only suitable for synthetic datasets but also can repair rain in real-world scenarios. The visual results can be seen in Fig. 8. The deraining result of SPDNet is better than other methods, which finely recover the image textures and sufficiently removing the rain streaks.

4.3. Ablation Study

Effectiveness of the number of WMLM: To validate the influence of the number of WMLM on the performance,

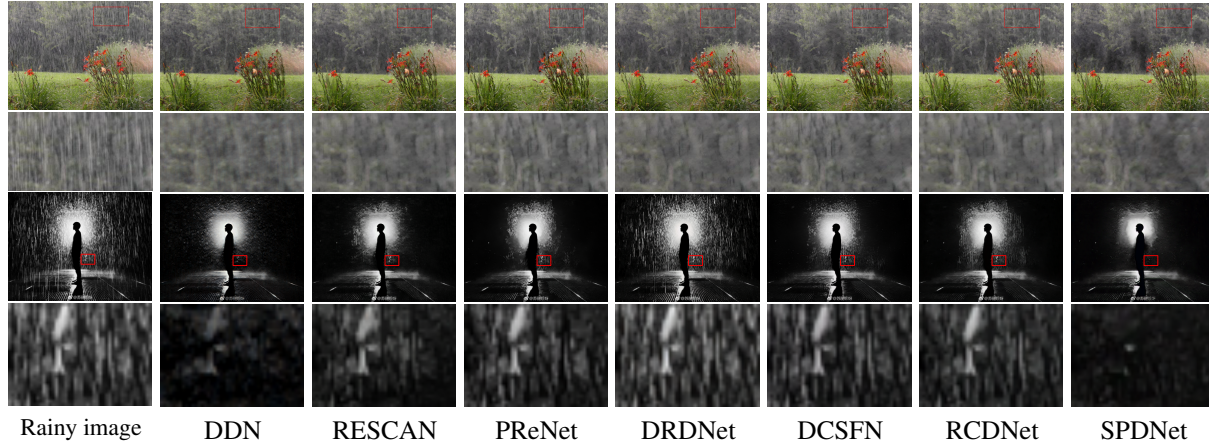


Figure 8: Image deraining results tested in the real-world dataset. The first row is rainy image and the output of different methods. The second row is the zoom results of the red window.

Method	Fusion method		Ensemble	Rain200H	
	IFM	Concat		PSNR	SSIM
w/o IFM		✓	✓	31.05	0.9167
w/o Ensemble	✓			30.98	0.9142
SPDNet	✓		✓	31.30	0.9217

Table 3: Ablation study on different settings of SPDNet on Rain200H.

Numbers	PSNR	SSIM	Param	Time
1	29.85	0.9006	0.98M	0.017s
2	30.90	0.9166	2.03M	0.039s
3	31.30	0.9217	3.04M	0.055s

Table 4: Explore the influence of different numbers of WLM on Rain200H. The time is the average run-time, calculated on 100 image of 128×128 size.

we conduct ablation experiments, which are shown in Table 4. When the number of WLM is equal to 3, the PSNR/SSIM is the best, but it will take a longer time to process the image. When the number of WLM is equal to 1, the performance is the lowest, but compared with some current methods, it is still better and the time consumed is shorter, which shows that WLM has a strong learning ability and may be able to serve real-time image rain removal task in future.

Effectiveness of Basic Components: In order to prove the performance of IFM and the impact of Ensemble(Sec. 3.2.4) on performance, we use Concat operation to instead the IFM and remove the Ensemble. The results are shown in Table 3. SPDNet with IFM achieves significant performance improvements over SPDNet without using IFM, which demonstrates that IFM is a more effective fusion method and can benefit to reconstruct clearer rain-free images. Furthermore, SPDNet with Ensemble outperforms SPDNet without Ensemble in PSNR and SSIM,

Iteration	RCP Update	PSNR	SSIM
0		30.56	0.9144
1		30.82	0.9161
2	✓	31.04	0.9197
3	✓	31.30	0.9217
3		31.12	0.9191

Table 5: Explore the influence of RCP on Rain200H. RCP Update means whether to use the iterative guidance strategy to update RCP and Iteration means the number of RCP guidance.

which shows that the Ensemble operation can enrich feature information so that the network can learn more useful information.

Effectiveness of the number of RCP guidance and Iterative Guidance Strategy: To demonstrate that RCP can effectively assist SPDNet to generate rain-free images with clearer structure, we conduct ablation experiments and the results are shown in Table 5. As shown, when more RCP guidance, the performance will be more superior. The visual comparison of different numbers of RCP guidance is shown in Fig. 10. It is observed that methods with RCP guidance can reconstruct clearer rain-free images with accurate structures, which indicates that the effectiveness of RCP in preserving structure. Moreover, to demonstrate the performance of iterative guidance strategy, we use RCP of rainy images to replace RCP of output results and the result is shown in the 4th row of Table 5. It can be found that the performance of the method that only uses RCP of rainy images is lower than SPDNet, which demonstrates that the effectiveness of iterative guidance strategy and clearer structure of RCP can reconstruct more high-quality images.

4.4. Application

To demonstrate that our proposed can benefit vision-based applications, we employ Google Vision API to eval-

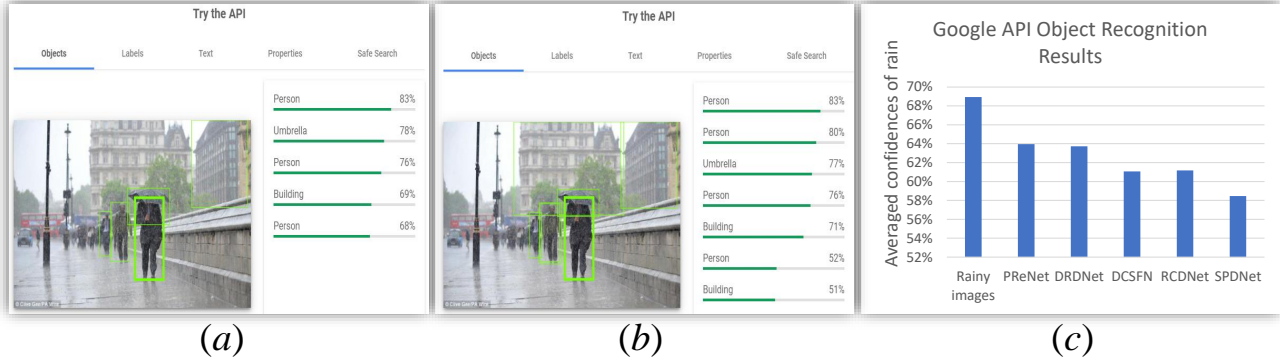


Figure 9: The deraining results are tested on the Google Vision API. *a*: object recognition result in the real-world rainy image, *b*: object recognition result after deraining by our proposed model, and *c*: the averaged confidences in recognizing rain from 30 sets of the real-world rainy images and output results of different methods. The lower averaged confidences, the better performance of deraining.

	Rainy	DRDNet	DCSFN	RCDNet	SPDNet
mAP-50 \uparrow	0.551	0.652	0.665	0.658	0.692

Table 6: The detection performance comparison on COCO350 dataset [16]. \uparrow means the better methods should achieve higher score.

uate the deraining results, as shown in Fig. 9. Fig. 9 (a) illustrates the object recognition result in the real-world rainy image and Fig. 9 (b) shows the object recognition result in the output result of SPDNet. Comparing Fig. 9 (a) and (b), the number of detections in our result is more, indicating that SPDNet can effectively improve the performance of detection and recognition. Furthermore, according to Deng *et al.* [6], we use Google Vision API to detect the confidences of rain in 30 sets of real-world rainy images and the results of five methods, as shown in Fig. 9 (c). The confidence of rain means that the probability of rainy weather. When confidence is lower, the rain is more light, which shows that the better deraining performance. As one can see, the confidence in recognizing rain from the output results of SPDNet is significantly reduced, which indicates that the deraining performance of SPDNet outperforms other methods in real-world datasets. Furthermore, we also use YOLOv3 to perform detection on COCO350 [16], as shown in Table 6 (mAP-50 means the mAP at .5 IOU metric). Compared with other models, the results generated by SPDNet effectively promote the detection performance, which further demonstrates the effectiveness of our proposed SPDNet.

5. Conclusion

In this paper, we have proposed a Structure-Preserving Deraining Network (SPDNet) with residue channel prior guidance. To achieve this, an RCP guidance network and an iterative guidance strategy are proposed for structure-preserving deraining. Specifically, an effective WMLM is proposed as the backbone to fully learn the background in-

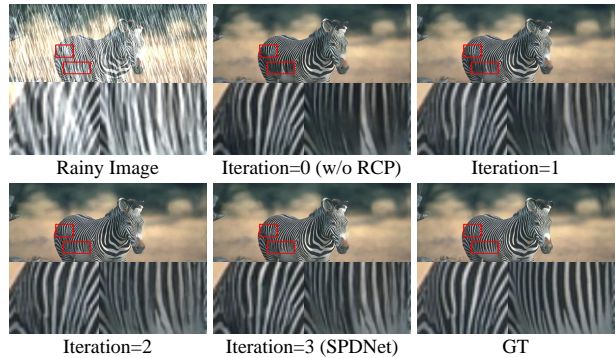


Figure 10: Comparison results on different numbers of RCP guidance. Obviously, compared with the method without RCP, the methods using RCR guidance can reconstruct the high-quality image with a clear and accurate structure.

formation. Meanwhile, the RCP is introduced as reference information to guide the learning of WMLM, and IFM is designed to make full use of the RCP information. In addition, based on an observation that the extracted RCP also shows better results with the improvement of deraining image quality, an iterative update strategy is proposed to improve the accuracy of RCP and then re-guide the learning of WMLM. The ablation study demonstrates the performance of RCP, IFM, and iterative update strategy. Experimental results on several synthetic deraining datasets and real-world scenarios have shown the great superiority of our proposed SPDNet over other top-performing methods.

Acknowledgment. This work was supported by the Key Project of the National Natural Science Foundation of China under Grant 61731009, the NSFC-RGC under Grant 61961160734, the National Natural Science Foundation of China under Grant 61871185, the Shanghai Rising-Star Program under Grant 21QA1402500, the Science Foundation of Shanghai under Grant 20ZR1416200, and the Open Research Fund of KLATASDS-MOE, ECNU.

References

- [1] Peter C Barnum, Srinivasa Narasimhan, and Takeo Kanade. Analysis of rain and snow in frequency space. *International journal of computer vision*, 86(2-3):256, 2010. [1](#)
- [2] Yochai Blau, Roey Mechrez, Radu Timofte, Tomer Michaeli, and Lihi Zelnik-Manor. The 2018 p1rm challenge on perceptual image super-resolution. In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, pages 0–0, 2018. [6](#)
- [3] Jérémie Bossu, Nicolas Hautiere, and Jean-Philippe Tarel. Rain or snow detection in image sequences through use of a histogram of orientation of streaks. *International journal of computer vision*, 93(3):348–367, 2011. [1](#)
- [4] Yi-Lei Chen and Chiou-Ting Hsu. A generalized low-rank appearance model for spatio-temporally correlated rain streaks. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1968–1975, 2013. [1](#)
- [5] Ingrid Daubechies. The wavelet transform, time-frequency localization and signal analysis. *IEEE transactions on information theory*, 36(5):961–1005, 1990. [3](#)
- [6] Sen Deng, Mingqiang Wei, Jun Wang, Yidan Feng, Luming Liang, Haoran Xie, Fu Lee Wang, and Meng Wang. Detail-recovery image deraining via context aggregation networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 14560–14569, 2020. [1](#), [3](#), [6](#), [8](#)
- [7] Xueyang Fu, Jiabin Huang, Xinghao Ding, Yinghao Liao, and John W. Paisley. Clearing the skies: A deep network architecture for single-image rain removal. *IEEE Trans. Image Process.*, 26(6):2944–2956, 2017. [1](#), [2](#)
- [8] Xueyang Fu, Jiabin Huang, Delu Zeng, Yue Huang, Xinghao Ding, and John Paisley. Removing rain from single images via a deep detail network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3855–3863, 2017. [1](#), [2](#), [6](#)
- [9] Xueyang Fu, Qi Qi, Zheng-Jun Zha, Yurui Zhu, and Xinghao Ding. Rain streak removal via dual graph convolutional network. In *Proc. AAAI Conf. Artif. Intell.*, pages 1–9, 2021. [2](#)
- [10] Kshitiz Garg and Shree K Nayar. Photometric model of a rain drop. In *CMU Technical Report*. Citeseer, 2003. [4](#)
- [11] Kshitiz Garg and Shree K Nayar. When does a camera see rain? In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, volume 2, pages 1067–1074. IEEE, 2005. [1](#)
- [12] Shuhang Gu, Deyu Meng, Wangmeng Zuo, and Lei Zhang. Joint convolutional analysis and synthesis sparse representation for single image layer separation. In *IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy, October 22-29, 2017*, pages 1717–1725, 2017. [1](#)
- [13] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. [1](#), [4](#)
- [14] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018. [4](#)
- [15] Xiaowei Hu, Chi-Wing Fu, Lei Zhu, and Pheng-Ann Heng. Depth-attentional features for single-image rain removal. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8022–8031, 2019. [1](#), [2](#)
- [16] Kui Jiang, Zhongyuan Wang, Peng Yi, Chen Chen, Baojin Huang, Yimin Luo, Jiayi Ma, and Junjun Jiang. Multi-scale progressive fusion network for single image deraining. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8346–8355, 2020. [1](#), [2](#), [3](#), [8](#)
- [17] Li-Wei Kang, Chia-Wen Lin, and Yu-Hsiang Fu. Automatic single-image-based rain streaks removal via image decomposition. *IEEE transactions on image processing*, 21(4):1742–1755, 2012. [2](#)
- [18] Juncheng Li, Yiting Yuan, Kangfu Mei, and Faming Fang. Lightweight and accurate recursive fractal network for image super-resolution. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, pages 0–0, 2019. [1](#)
- [19] Ruoteng Li, Robby T Tan, and Loong-Fah Cheong. Robust optical flow in rainy scenes. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 288–304, 2018. [2](#), [4](#)
- [20] Ruoteng Li, Robby T Tan, and Loong-Fah Cheong. All in one bad weather removal using architectural search. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3175–3185, 2020. [2](#)
- [21] Ruoteng Li, Robby T Tan, Loong-Fah Cheong, Angelica I Aviles-Rivero, Qingnan Fan, and Carola-Bibiane Schonlieb. Rainflow: Optical flow under rain streaks and rain veiling effect. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 7304–7313, 2019. [2](#)
- [22] Siyuan Li, Wenqi Ren, Feng Wang, Iago Breno Araujo, Eric K. Tokuda, Roberto Hirata Junior, Roberto M. Cesar-Jr., Zhangyang Wang, and Xiaochun Cao. A comprehensive benchmark analysis of single image deraining: Current challenges and future perspectives. *International Journal of Computer Vision*, 2021. [2](#)
- [23] Xia Li, Jianlong Wu, Zhouchen Lin, Hong Liu, and Hongbin Zha. Recurrent squeeze-and-excitation context aggregation net for single image deraining. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 254–269, 2018. [1](#), [3](#), [6](#)
- [24] Yu Li, Robby T Tan, Xiaojie Guo, Jiangbo Lu, and Michael S Brown. Rain streak removal using layer priors. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2736–2744, 2016. [1](#), [2](#), [6](#)
- [25] Lin Liu, Jianzhuang Liu, Shanxin Yuan, Gregory Slabaugh, Aleš Leonardis, Wengang Zhou, and Qi Tian. Wavelet-based dual-branch network for image demoiréing. In *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XIII 16*, pages 86–102. Springer, 2020. [3](#)
- [26] Pengju Liu, Hongzhi Zhang, Kai Zhang, Liang Lin, and Wangmeng Zuo. Multi-level wavelet-cnn for image restoration. In *Proceedings of the IEEE conference on computer*

- vision and pattern recognition workshops, pages 773–782, 2018. 3
- [27] Yu Luo, Yong Xu, and Hui Ji. Removing rain from a single image via discriminative sparse coding. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3397–3405, 2015. 1, 2, 6
- [28] Anish Mittal, Rajiv Soundararajan, and Alan C Bovik. Making a “completely blind” image quality analyzer. *IEEE Signal processing letters*, 20(3):209–212, 2012. 6
- [29] Long Peng, Aiwen Jiang, Qiaosi Yi, and Mingwen Wang. Cumulative rain density sensing network for single image derain. *IEEE Signal Processing Letters*, 27:406–410, 2020. 1
- [30] Rui Qian, Robby T Tan, Wenhan Yang, Jiajun Su, and Jiaying Liu. Attentive generative adversarial network for rain-drop removal from a single image. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2482–2491, 2018. 2
- [31] Dongwei Ren, Wangmeng Zuo, Qinghua Hu, Pengfei Zhu, and Deyu Meng. Progressive image deraining networks: A better and simpler baseline. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3937–3946, 2019. 1, 3, 6
- [32] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. 1
- [33] Cong Wang, Yutong Wu, Zhixun Su, and Junyang Chen. Joint self-attention and scale-aggregation for self-calibrated deraining network. In *Proceedings of the 28th ACM International Conference on Multimedia*, pages 2517–2525, 2020. 2
- [34] Cong Wang, Xiaoying Xing, Yutong Wu, Zhixun Su, and Junyang Chen. Dcsfn: Deep cross-scale fusion network for single image rain removal. In *Proceedings of the 28th ACM International Conference on Multimedia*, pages 1643–1651, 2020. 3, 6
- [35] Guoqing Wang, Changming Sun, and Arcot Sowmya. Erl-net: Entangled representation learning for single image deraining. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5644–5652, 2019. 2
- [36] Hong Wang, Qi Xie, Yichen Wu, Qian Zhao, and Deyu Meng. Single image rain streaks removal: a review and an exploration. *International Journal of Machine Learning and Cybernetics*, 11(4):853–872, 2020. 2
- [37] Hong Wang, Qi Xie, Qian Zhao, and Deyu Meng. A model-driven deep neural network for single image rain removal. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3103–3112, 2020. 1, 3, 6
- [38] Tianyu Wang, Xin Yang, Ke Xu, Shaozhe Chen, Qiang Zhang, and Rynson WH Lau. Spatial attentive single-image deraining with a high quality real rain dataset. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12270–12279, 2019. 1, 5
- [39] Yinglong Wang, Shuaicheng Liu, Chen Chen, and Bing Zeng. A hierarchical approach for rain or snow removing in a single color image. *IEEE Transactions on Image Processing*, 26(8):3936–3950, 2017. 2
- [40] Wei Wei, Deyu Meng, Qian Zhao, Zongben Xu, and Ying Wu. Semi-supervised transfer learning for image rain removal. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3877–3886, 2019. 2
- [41] Jing Xu, Wei Zhao, Peng Liu, and Xianglong Tang. An improved guidance image based method to remove rain and snow in a single image. *Computer and Information Science*, 5(3):49, 2012. 1
- [42] Wenhan Yang, Jiaying Liu, Shuai Yang, and Zongming Guo. Scale-free single image deraining via visibility-enhanced recurrent wavelet learning. *IEEE Transactions on Image Processing*, 28(6):2948–2961, 2019. 2
- [43] Wenhan Yang, Robby T Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, and Shuicheng Yan. Deep joint rain detection and removal from a single image. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1357–1366, 2017. 1, 2, 5
- [44] Wenhan Yang, Robby T Tan, Shiqi Wang, Yuming Fang, and Jiaying Liu. Single image deraining: From model-based to data-driven and beyond. *IEEE Transactions on pattern analysis and machine intelligence*, 2020. 2
- [45] Youzhao Yang and Hong Lu. Single image deraining via recurrent hierarchy enhancement network. In *Proceedings of the 27th ACM International Conference on Multimedia*, pages 1814–1822, 2019. 1
- [46] Rajeev Yasarla and Vishal M Patel. Uncertainty guided multi-scale residual learning-using a cycle spinning cnn for single image de-raining. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8405–8414, 2019. 2
- [47] Rajeev Yasarla, Vishwanath A Sindagi, and Vishal M Patel. Syn2real transfer learning for image deraining using gaussian processes. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2726–2736, 2020. 2
- [48] Qiaosi Yi, Juncheng Li, Faming Fang, Aiwen Jiang, and Guixu Zhang. Efficient and accurate multi-scale topological network for single image dehazing. *IEEE Transactions on Multimedia*, pages 1–1, 2021. 3
- [49] He Zhang and Vishal M Patel. Density-aware single image de-raining using a multi-stream dense network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 695–704, 2018. 1, 5
- [50] He Zhang, Vishwanath Sindagi, and Vishal M Patel. Image de-raining using a conditional generative adversarial network. *IEEE transactions on circuits and systems for video technology*, 30(11):3943–3956, 2019. 1, 5, 6
- [51] Yulun Zhang, Kungpeng Li, Kai Li, Lichen Wang, Bineng Zhong, and Yun Fu. Image super-resolution using very deep residual channel attention networks. In *Proceedings of the European conference on computer vision (ECCV)*, pages 286–301, 2018. 4
- [52] Lei Zhu, Chi-Wing Fu, Dani Lischinski, and Pheng-Ann Heng. Joint bi-layer optimization for single-image rain streak removal. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2545–2553, 2017. 2