Multi-scale and Multi-stage Deraining Network with Fourier Space Loss

Zhaoyong Yan, Liyan Ma(✉), Xiangfeng Luo(✉), and Yan Sun

School of Computer Engineering and Science, Shanghai University, Shanghai, China
{liyanma,luoxf}@shu.edu.cn

Abstract. The goal of rain streak removal is to recover the rain-free background scenes of an image degraded by rain streaks. Most current deep convolutional neural networks methods have achieved dramatic performance. However, these methods still cannot capture the discriminative features to well distinguish the rain streaks and the important image content. To solve this problem, we propose a Multi-scale and Multi-stage deraining network in the end-to-end manner. Specifically, we design a multi-scale rain streak extraction module to capture complex rain streak features across different scales through the multi-scale selection kernel attention mechanism. In addition, multi-stage learning is used to extract deeper feature representations of rain streak and fuse different stages of background information. Furthermore, we introduce a Fourier space loss function to reduce the loss of high-frequency information in the background image and improve the quality of deraining results. Extensive experiments demonstrate that our network performs favorably against the state-of-the-art deraining methods.

Keywords: Image deraining · Multi-scale · Multi-stage · Fourier space loss

1 Introduction

Image quality is degraded when images are captured in rainy conditions, which not only affects human visual perception but also drastically drop the performance of downstream vision tasks, such as optical flow estimation [21], object recognition [1] and object tracking [19]. Therefore, it is important to study how to effectively remove rain streaks from the given rain image.

The de-rain problem is usually represented by the following linear superimposition model [17]:

$$O = B + R, \tag{1}$$

where O denotes the rain image, R is the rain layer, and B is the background layer, which is generally called the rain-free image.

In recent decades, many methods have been proposed for single image deraining (SID). The tradition methods are layer priors with Gaussian mixture model (GMM) [16], discriminative sparse coding (DSC) [17], and joint convolutional analysis and synthesis sparse representation (JCAS) [6]. Although the above SID

D.-T. Dang-Nguyen et al. (Eds.): MMM 2023, LNCS 13834, pp. 575–586, 2023.
https://doi.org/10.1007/978-3-031-27818-1_47

methods can achieve good performance, they still cannot separate the rain layer from rainy images completely. Besides, the handcraft low-level representation produced by strong prior assumptions, leads to bad generalization performance for rain streaks removal.

Recently, deep learning has been widely used in computer vision and has shown competitive results. Various SID methods based on convolutional neural network(CNN) have been proposed [13,15,23]. In order to adequately separate the rain streaks from the rain images, Li et al. [13] adopted Squeeze-and-Excitation network to explore the relationship between channels and rain streaks of the rain images. Wang et al. [23] designed an iterative algorithm using proximal gradient descent technique to extract rain streaks. In addition, Li et al. [15] applied rain embedding loss and rectified local contrast normalization to SID. However, the above methods often tend to excessive or insufficient removal of rain streaks, resulting in the loss of structural information of the reconstructed images. The main reason maybe that these methods do not consider the characteristics of rain streaks semantically related to the image content [7], such as the different shapes, directions, transparency across the image, which leads to inadequate learning ability of their basic modules to rain streaks.

In the paper, we propose a novel MS2DNet by learning rain streaks distribution to solve the above mentioned difficulties. We note that the similarity of patch-level rain streaks exists not only in the same scale but also in different scales. Therefore, we construct a multi-scale feature extraction module to effectively capture the local and global features of rain streaks. To be specific, for the multi-scale feature extraction module, we first use convolution with SE residuals in residual blocks [30] to extract rain streaks information from different scales. Secondly, we use a selective feature fusion mechanism [14] to align and adaptively extract features from different scales. Finally, the extracted rain streak features are fused again and the rain streaks map is sent as the background module to predict the rain-free image. In addition, we also introduce frequency features by Fourier transform to further improve the derain quality. In this paper, we perform a series of experimental validation on MS2DNet. The quantitative and qualitative results show the superior performance of the MS2DNet. The main contributions of this work are as follows:

- We propose a novel MS2DNet to incorporate multi-scale information with multi-stage learning strategy for SID. Our MS2DNet can extract the rain streak features from rainy images by using a multi-scale rain streaks extraction module (MREM) that can adaptively select and fuse the rain streak features extracted from different scales to improve the performance. To further improve the performance, we propose a multi-stage learning strategy to extract deeper information about rain streaks and to fuse background information from different stages.
- To preserve the frequency domain information as well as preserving content and structure information of the image, we introduce the frequency space loss by calculating the frequency components with the fast Fourier transform (FFT).
- Extensive experiments demonstrate that our approach achieves outstanding performance on synthetic and real-world rain datasets.

2 Related Work

2.1 Single Image Deraining

Most single image derain methods acquire the background layer by separating the rain layer from the rain image. These approaches can be divided into two categories: model-driven and data-driven.

The model-driven approaches use a priori knowledge derived from statistical images to extract rain streaks. Kang et al. [11] decomposed the rain streaks into the high-frequency part, and then extracted the rain streaks by dictionary learning. This method successfully extracts the rain streaks. However, the separated boundaries can not be obtained precisely, resulting in blurred background information. To solve the above problem, Zhang et al. [27] combined sparsity code with low-rank representation to obtain clear background layer and rain layer.

The data-driven approachs use the powerful fitting capabilities of CNN to distinguish the difference between background and rain streaks. Yang et al. [25] proposed a SID model based on recurrent network to remove rain streaks. RESCAN [13] designs a recurrent network to obtain rain streaks, where the dilated convolution is used to learn the contextual information. To full use multi-scale collaborative representation for rain streaks from the perspective of input image scales, Jing et al. [8] designed a progressive network to learn the intermediate information of the image, which consists of multiple residual blocks, while introducing the LSTM. To better capture the characteristics of rain streaks, Wang et al. [23] proposed a convolutional dictionary to extract rain streaks and replace complex optimization processes with proximal gradient algorithms.

2.2 Multi-scale Learning

There is a high similarity of rain streaks between same scale and different scales in natural environments. How to use this property to achieve effective extraction of rain streaks has attracted the attention of many researchers. For example, Li et al. [12] proposed a scale-aware multi-stage convolutional neural network to learn rain streaks at different scales. Fu et al. [4] proposed multi-branch pyramidal networks to learn rain streaks features in a specific pyramidal spatial. However, this method ignored the correlated information among these pyramid layers. Zheng et al. [31] proposed a cascaded pyramidal network instead of the optimization method of coarse to fine. Unfortunately, some details and scale features are lost due to the excessive focus on higher-level features. Different from the above methods, we design a new multi-scale feature extraction and fusion module, in which rich information of different scale obtained by residual attention module extraction is fused by adaptive selection network. As a result, we are able to obtain more accurate rain streaks distribution.

2.3 Multil-stage Learning

It is proved that extracting features through multiple stages is more beneficial for feature representation [13,22,29]. For example, Nah et al. [18] divided the image recovery task into three stages: coarse, medium and fine, and used the information from the first two stages to retain detailed features. Fu et al. [4] attempted to capture complex rain streaks in different stages based on pyramidal division into networks. To tackle the optimization problem from a coarse to fine, a cascaded pyramidal network is proposed by Zheng et al. [31]. However, it is difficult to make full use of feature information by simply cascading multiple stages. Different from the above methods, we propose a multi-stage learning strategy to adaptively fuse the useful information of the previous stage rain-free background image to guide the deraining in later stages. Such a simple implementation preserves the desired fine texture in the final output image.

3 Proposed Method

3.1 The Overall Structure of MS2DNet

The proposed Multi-scale and Multi-Stage deraining Network (MS2DNet) consists of N stages. Each stage is composed of Multi-scale rain streaks extraction module (MREM) and Background recover module (BRM). The architecture of MS2DNet is shown in Fig. 1. MREM is designed for extracting rain streaks information, which is mainly composed of SE-residual in residual block [30] (SRiR) and selective kernel feature fusion mechanism [14] (SKFF). In order to further improve the quality of deraining images, BRM is introduced to recover the background image through different stages to avoid low-quality background images caused by excessive or incomplete deraining.

3.2 Multi-scale Rain Streaks Extract Module

Rain streaks extraction is a critical step for the rain removal. The multi-scale residual block [26] was used for image enhancement which extracts features and fuses information from different streams through multiple attention mechanisms. Inspired by this, we design a novel multi-scale residual extract feature module to effectively capture rain streaks information. Our MREM consists of three convolution streams connected in parallel and SRiR is used for feature extraction on different scales. In order to fully exploit the representation power of the network, we use SKFF to adaptively fuse and select features from different scales.

In MREM module, as shown in Fig. 1, we firstly apply downsampling operation (DS) with different times to obtain the multi-scale rain feature maps, and utilize a convolution operation to adjust the channels of feature according to different scales. It can be mathematically described as:

$$\begin{cases} \tilde{R} = Conv(R^{s-1}), & 1 \leq s, \\ \tilde{R}_i = \tilde{R}, & \text{if } i = 0, \\ \tilde{R}_i = Conv(DS(\tilde{R}_{i-1})), & \text{if } i > 0, \end{cases} \tag{2}$$

where R^{s-1} denotes rain layer from $s-1$ stage, $\tilde{R}$ is the input of MREM, $\tilde{R}_i$ represents the rain feature maps of different scales, DS denotes downsampling

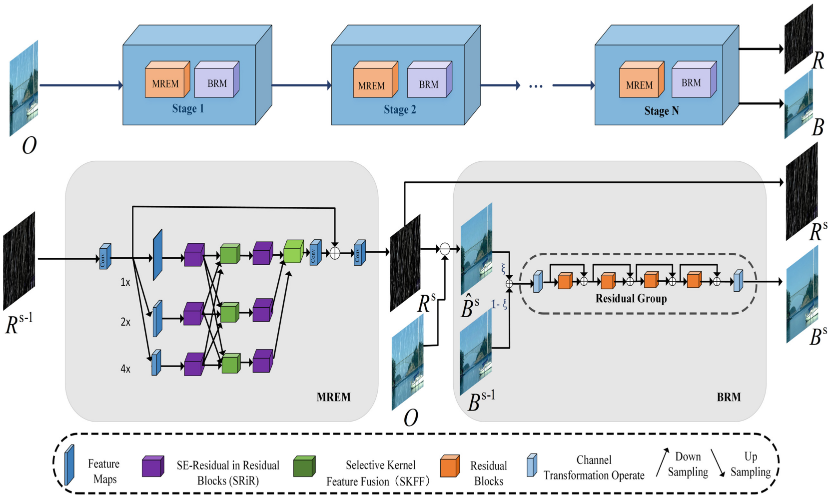


Fig. 1. The overall framework of the proposed Multi-scale and Multi-stage Deraining Network (MS2DNet).

operation. Then, we employ SRiR to extract effective rain feature maps. In order to fully exchange information and obtain contextual information from different scales, a feature fuse and select module is also applied as following:

$$\begin{cases} \tilde{R}_i^r = SRiR(\tilde{R}_i), & \text{if } i = 0, 1, 2, \\ \tilde{R}_i^k = SKFF(UDS(Conv(\tilde{R}_i^r))), & \text{if } i = 0, 1, 2, \\ \hat{R}_i^r = SRiR(\tilde{R}_i^k), & \text{if } i = 0, 1, 2, \end{cases} \tag{3}$$

where UDS denotes upsampling or downsampling operation which is used to adjust the scale of rain features. After obtaining the effective rain feature maps $\hat{R}_i^k$, and again, we adopt feature fusion module to fuse and select informative features from different scales as following:

$$\hat{R}^k = SKFF(US(Conv(\hat{R}_i^r))), \quad i = 0, 1, 2, \tag{4}$$

$$R^s = Conv(Conv(\hat{R}^k) + \tilde{R}). \tag{5}$$

3.3 Background Recover Module

After obtaining the rain layer R^s from MREM, we need to recover the background B^s. However, obtaining background information directly can lead to contain some residual rain streaks or lose some image content. Inspired by [23], we develop a fusion module that recovers the background $\hat{B}^s$ with rain maps R^s, and fuses $\hat{B}^s$ with the previously stage B^{s-1} by the parameters of the learnable ξ to get the derained results. The background fusion module (shown in Fig. 1) can be defined as:

$$\begin{cases} \hat{B}^s = O - R^s, \\ B^s = G((1-\xi)B^{s-1} + \xi \hat{B}^s), \end{cases} \tag{6}$$

where O represents rain image and G is a residual group consisting of several residual blocks.

3.4 Training Loss for Deraining Network

In order to obtain derained images with high quality perception and texture from the deraining model, the loss of the model is formulated as:

$$L_S = L_r + \alpha L_p + \beta L_f, \tag{7}$$

where L_r represents the reconstruction loss, L_p represents the perceptual loss, L_f represents the Fourier space loss that aims to preserve the image details as much as possible under the supervision of groundtruth image, α and β are the balancing weights.

Reconstruction Loss. To obtain the reconstruction result, we use mean squares error [2] ($L2$) to measure the distance between ground truth image and deraining result as:

$$L_r = ||P(O) - B||_2, \tag{8}$$

where $P(O)$ is the predicted image, and B represents the ground truth image.

Perceptual Loss. The perceptual loss [10] (PL) is used as a feature-level constraint to obtain the more semantic-level features of the acquired image. Perceptual loss can be defined as:

$$L_p = \frac{1}{C_i H_i W_i} \left\| \varphi_i(P(O)) - \varphi_i(B) \right\|_2^2, \tag{9}$$

where φ_i denotes the convolution operation used to extract features at layer i, and $C_i \times H_i \times W_i$ denotes the feature maps shape of the output at layer i. We use the $L2$ to measure the inter-feature distance, and in this loss we use the pre-trained VGG19 model.

Fourier Space Loss. Although having good performance in the image reconstruction, the above two loss functions do not fully recover the details of the groundtruth image. To solve the above problem, some methods [5,9] adopt frequency information to compensate for missing texture details. Following these methods, we introduce Fourier space loss (FSL) to the deraining network to preserve texture information. The image $I(x, y)$ are transformed into Fourier space by applying the Fast Fourier transform (FFT). We calculate the amplitude difference $L_f, |.|$ and phase difference $L_f, \angle$ of all frequency components between output image and ground truth image. The averaged differences is computed as the total frequency loss L_f as following:

$$L_f, |.| = \frac{2}{UV} \sum_{u=0}^{U/2-1} \sum_{v=0}^{V-1} \left| \left| \hat{Y} \right|_{u,v} - |Y|_{u,v} \right|, \tag{10}$$

$$L_f, \angle = \frac{2}{UV} \sum_{u=0}^{U/2-1} \sum_{v=0}^{V-1} \left| \angle \hat{Y}_{u,v} - \angle Y_{u,v} \right|, \tag{11}$$

$$L_f = \frac{1}{2} L_f, |.| + \frac{1}{2} L_f, \angle, \tag{12}$$

where $\hat{Y}_{u,v}$ represents the spectrum of the recovered image, and $Y_{u,v}$ represents the spectrum of the ground truth image.

4 Experiment and Result

4.1 DataSet

To evaluate the effectiveness of our proposed method, the experiments on the synthetic dataset and the real dataset are conducted seperately. Rain200H [25] and Rain200L [25] contains 1800 training images and 200 testing images. Rain100L [25] contains 200 training images and 100 testing images. Rain12 [16] has only 12 images for testing. Real-world dataset [28] includes 167 rain images for testing. The model trained on Rain200H is used to test on Rain12 and real-world datasets.

4.2 Implementation Details

In this paper, we adopt the Adam optimizer with the Hyperparameters $\beta_1 = 0.9$, $\beta_2 = 0.999$. The learning rate is 5×10^{-4}, and the batch size is set to be 16. The weight parameters α and β are 0.1 and 10 in the loss function, respectively. We randomly crop the patch of which size is 128×128 as the input of the deraining network and the total epoch is 200. Our implementation is based on the Pytorch platform and we train our model on an NVIDIA V100.

4.3 Comparisons with the State-of-the-arts

The proposed MS2DNet is compared to several state-of-the-art image deraining methods, including RESCAN [13], SpaNet [24], PreNet [20], MSPFN [8], RCD-Net [23], RLNet [3] and ECNetLL [15]. We use some common evaluation metrics including PSNR and SSIM for quantitative deraining.

Result on Synthetic Datasets. The quantitative results on the synthetic datasets are presented in Table 1. We can notice that the proposed MS2DNet achieves remarkable improvement over the existing state-of-the-art methods. Meanwhile, it can be observed that MS2DNet has a better generalization ability than other models via comparing the results on Rain12. We also provided the visual results on synthetic datasets in Fig. 2. We can observe that our method can obtain better recovered images and obtain more accurate structure and cleaner background, especially in the cropped areas.

Result on Real-World Datasets. We also verify the robustness on real-world datasets. Since there are no corresponding groundtruth images to these real-world datasets, the de-rained results are only evaluated by human visual perception. For the two illustrated examples in Fig. 3, we can find that MS2DNet has effectively remove rain streaks and restore better rain-free images. These results show the strength of our methods to distinguish the rain streaks and the important image content.

Table 1. PSNR and SSIM comparisons on four benchmark datasets. Bold indicates 1st rank. * indicates the results directly copied from [3] since the authors do not provide the full codes.

Methods	Rain200H		Rain200L		Rain100L		Rain12	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RESCAN [13]	26.66	0.841	36.99	0.978	36.58	0.970	32.96	0.954
SpaNet [24]	25.48	0.858	36.07	0.977	27.85	0.881	33.21	0.954
PreNet [20]	27.52	0.866	34.26	0.966	36.28	0.979	35.09	0.940
MSPFN [8]	25.55	0.803	30.36	0.921	33.50	0.948	34.25	0.946
RCDNet [23]	28.69	0.890	38.40	0.984	**38.60**	**0.983**	31.03	0.906
RLNet* [3]	28.87	0.895	–	–	37.38	0.980	–	–
ECNetLL [15]	28.85	0.898	38.35	0.984	38.26	0.982	32.35	0.962
MS2DNet (Ours)	**29.91**	**0.908**	**39.28**	**0.986**	38.39	**0.983**	**36.70**	**0.966**

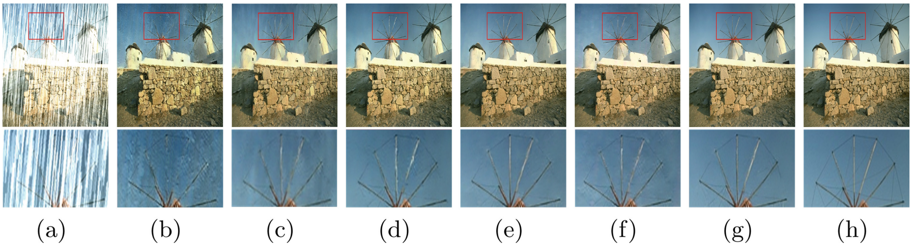


Fig. 2. Visual comparison of the synthetic example on Rain200H. (a) Rainy image. Draining results by (b) SpaNet, (c) PreNet, (d) MSPFN, (e) RCDNet, (f) ECNetLL, (g) MS2DNet, (h) GT.

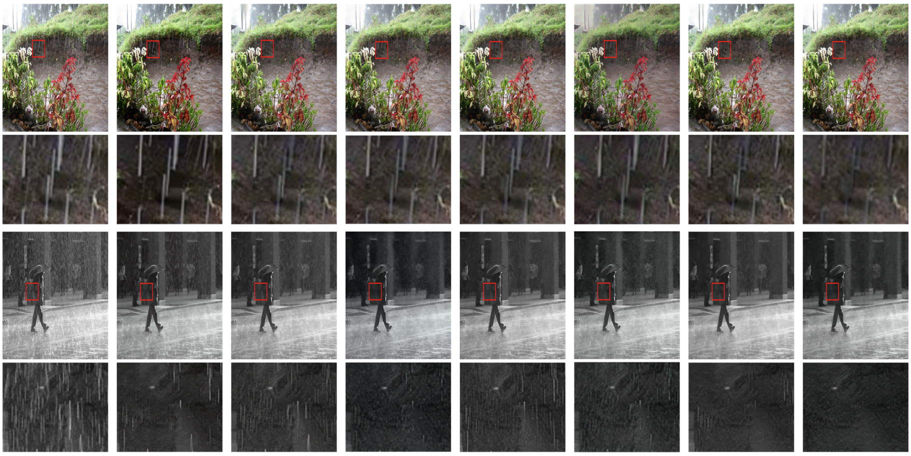
4.4 Ablation Study

We conduct the ablation studies to discuss the effectiveness of different components in MS2DNet.

Effectiveness of the Different Modules

To validate the superiority of the MS2DNet, we conduct ablation study and analyze different components of the module. MREM+BRM denotes the MS2DNet.

The results of the different modules are shown in Table 2. We can observe that only the MERM module, we get (PSNR:29.37, SSIM:0.896), which can outperform the compared SOTA RLNet [3] (PSNR:28.87, SSIM:0.895). The addition of the BRM module can further improve the rain removal performance.



(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 3. Visual comparison of the competing methods on real-world dataset. (a) Rainy image. Draining results by (b) RESCAN, (c) SpaNet, (d) PreNet, (e) MSPFN, (f) RCDNet, (g) ECNetLL, (h) MS2DNet.

Table 2. Ablation study on the effectiveness of the different modules on Rain200H.

Modules	BRM	MREM	MREM+BRM	PSNR	SSIM
BRM	✓			28.07	0.881
MREM		✓		29.37	0.896
MREM+BRM			✓	**29.91**	**0.908**

Effectiveness of the Fourier Space Loss. To validate the effectiveness of the Fourier space loss function, we conduct ablation experiments and analyze the effectiveness of the mentioned loss functions for rain removal. Table 3 presents the results of different loss combinations. We can observe that the Fourier space loss improves the results compared to the perceptual loss, and approximates the results which obtained by a combination of all loss functions.

Number of Stages in the MS2DNet. To explore the influence of different numbers of stages, we conduct ablation experiments and the results are shown in Table 4, where N indicates MS2DNet with different numbers of stages. We can observe that our method with 3 stages achieves the best performance. We set the stages to be 3 for the balance of the computation cost and performance of our algorithm.

Table 3. Ablation study on the effectiveness of loss functions on Rain200H.

Loss	MSE	PL	FSL	PSNR	SSIM
L_r	✓			29.24	0.897
$L_r + L_p$	✓	✓		29.70	0.905
$L_r + L_f$	✓		✓	29.86	**0.908**
$L_r + L_p + L_f$	✓	✓	✓	**29.91**	**0.908**

Table 4. Ablation study on stage (MS2DNet) numbers on Rain200H.

Stage	N = 2	N = 3	N = 4
PSNR	29.52	**29.91**	29.89
SSIM	0.901	**0.908**	0.906

5 Conclusion

In this paper, we have proposed a multi-scale and multi-stage deraining network with Fourier space loss. In order to obtain high-quality derained images, a new multi-scale and multi-stage deraining network is designed to extract complex rain streaks and fuse background information from different stages. Especially, the selection kernel attention mechanism is beneficial to fuse and select useful rain streaks information from different scales. In addition, in order to eliminate the blurring of the reconstruction content brought by the popular pixel-level loss function, we introduce a Fourier space loss to reduce the loss of high-frequency information and improve the deraining quality. The experiments in this paper fully demonstrate the effectiveness of our proposed method on synthetic and real-world datasets.

Acknowledgements. This work was supported in part by the Shanghai Municipal Natural Science Foundation (No. 21ZR1423300), in part by "Shuguang Program" supported by Shanghai Education Development Foundation and Shanghai Municipal Education Commission (No. 20SG40), in part by program of Shanghai Academic Research Leader (No. 20XD1421700).

References

1. Barbu, A., et al.: ObjectNet: a large-scale bias-controlled dataset for pushing the limits of object recognition models. In: Advances in Neural Information Processing Systems, vol. 32 (2019)

2. Burger, H.C., Schuler, C.J., Harmeling, S.: Image denoising: can plain neural networks compete with BM3D? In: 2012 IEEE Conference on Computer Vision and Pattern Recognition, pp. 2392–2399. IEEE (2012)
3. Chen, C., Li, H.: Robust representation learning with feedback for single image deraining. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 7742–7751 (2021)
4. Fu, X., Liang, B., Huang, Y., Ding, X., Paisley, J.: Lightweight pyramid networks for image deraining. IEEE Trans. Neural Netw. Learn. Syst. **31**(6), 1794–1807 (2019)
5. Fuoli, D., Van Gool, L., Timofte, R.: Fourier space losses for efficient perceptual image super-resolution. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 2360–2369 (2021)
6. Gu, S., Meng, D., Zuo, W., Zhang, L.: Joint convolutional analysis and synthesis sparse representation for single image layer separation. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1708–1716 (2017)
7. Guo, Q., et al.: EfficientDeRain: learning pixel-wise dilation filtering for high-efficiency single-image deraining. arXiv preprint arXiv:2009.09238 (2020)
8. Jiang, K., et al.: Multi-scale progressive fusion network for single image deraining. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8346–8355 (2020)
9. Jiang, L., Dai, B., Wu, W., Change Loy, C.: Focal frequency loss for generative models. arXiv e-prints, p. arXiv-2012 (2020)
10. Johnson, J., Alahi, A., Fei-Fei, L.: Perceptual losses for real-time style transfer and super-resolution. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9906, pp. 694–711. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46475-6_43
11. Kang, L.W., Lin, C.W., Fu, Y.H.: Automatic single-image-based rain streaks removal via image decomposition. IEEE Trans. Image Process. **21**(4), 1742–1755 (2011)
12. Li, R., Cheong, L.F., Tan, R.T.: Single image deraining using scale-aware multi-stage recurrent network. arXiv preprint arXiv:1712.06830 (2017)
13. Li, X., Wu, J., Lin, Z., Liu, H., Zha, H.: Recurrent squeeze-and-excitation context aggregation net for single image deraining. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11211, pp. 262–277. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01234-2_16
14. Li, X., Wang, W., Hu, X., Yang, J.: Selective kernel networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 510–519 (2019)
15. Li, Y., Monno, Y., Okutomi, M.: Single image deraining network with rain embedding consistency and layered LSTM. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 4060–4069 (2022)
16. Li, Y., Tan, R.T., Guo, X., Lu, J., Brown, M.S.: Rain streak removal using layer priors. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2736–2744 (2016)
17. Luo, Y., Xu, Y., Ji, H.: Removing rain from a single image via discriminative sparse coding. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 3397–3405 (2015)
18. Nah, S., Hyun Kim, T., Mu Lee, K.: Deep multi-scale convolutional neural network for dynamic scene deblurring. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3883–3891 (2017)

19. Ning, G., et al.: Spatially supervised recurrent convolutional neural networks for visual object tracking. In: 2017 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1–4. IEEE (2017)
20. Ren, D., Zuo, W., Hu, Q., Zhu, P., Meng, D.: Progressive image deraining networks: a better and simpler baseline. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3937–3946 (2019)
21. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. In: Advances in Neural Information Processing Systems, vol. 28 (2015)
22. Tao, X., Gao, H., Shen, X., Wang, J., Jia, J.: Scale-recurrent network for deep image deblurring. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 8174–8182 (2018)
23. Wang, H., Xie, Q., Zhao, Q., Meng, D.: A model-driven deep neural network for single image rain removal. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3103–3112 (2020)
24. Wang, T., Yang, X., Xu, K., Chen, S., Zhang, Q., Lau, R.W.: Spatial attentive single-image deraining with a high quality real rain dataset. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 12270–12279 (2019)
25. Yang, W., Tan, R.T., Feng, J., Liu, J., Guo, Z., Yan, S.: Deep joint rain detection and removal from a single image. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1357–1366 (2017)
26. Zamir, S.W., et al.: Learning enriched features for real image restoration and enhancement. In: Vedaldi, A., Bischof, H., Brox, T., Frahm, J.-M. (eds.) ECCV 2020. LNCS, vol. 12370, pp. 492–511. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-58595-2_30
27. Zhang, H., Patel, V.M.: Convolutional sparse and low-rank coding-based rain streak removal. In: 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1259–1267. IEEE (2017)
28. Zhang, H., Sindagi, V., Patel, V.M.: Image de-raining using a conditional generative adversarial network. IEEE Trans. Circuits Syst. Video Technol. **30**(11), 3943–3956 (2019)
29. Zhang, H., Dai, Y., Li, H., Koniusz, P.: Deep stacked hierarchical multi-patch network for image deblurring. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 5978–5986 (2019)
30. Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., Fu, Y.: Image super-resolution using very deep residual channel attention networks. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11211, pp. 294–310. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01234-2_18
31. Zheng, Y., Yu, X., Liu, M., Zhang, S.: Residual multiscale based single image deraining. In: BMVC, p. 147 (2019)