

Direction-aware Feature-level Frequency Decomposition for Single Image Deraining

Sen Deng^{1,*}, Yidan Feng^{1,*}, Mingqiang Wei^{1,†}, Haoran Xie^{2,†}, Yiping Chen³, Jonathan Li³, Xiao-Ping Zhang⁴ and Jing Qin⁵

¹Nanjing University of Aeronautics and Astronautics, Nanjing, China

²Lingnan University, Hong Kong, China

³Xiamen Univeristy, Xiamen, China

⁴Ryerson University, Toronto, Canada

⁵Hong Kong Polytechnic University, Hong Kong, China

{sendeng, fengyidan}@nuaa.edu.cn, mingqiang.wei@gmail.com, hrxie@ln.edu.hk, chenyping@xmu.edu.cn, junli@uwaterloo.ca, xzhang@ryerson.ca, harry.qin@polyu.edu.hk

Abstract

We present a novel direction-aware feature-level frequency decomposition network for single image deraining. Compared with existing solutions, the proposed network has three compelling characteristics. First, unlike previous algorithms, we propose to perform frequency decomposition at feature-level instead of image-level, allowing both low-frequency maps containing structures and high-frequency maps containing details to be continuously refined during the training procedure. Second, we further establish communication channels between low-frequency maps and high-frequency maps to interactively capture structures from high-frequency maps and add them back to low-frequency maps and, simultaneously, extract details from low-frequency maps and send them back to high-frequency maps, thereby removing rain streaks while preserving more delicate features in the input image. Third, different from existing algorithms using convolutional filters consistent in all directions, we propose a direction-aware filter to capture the direction of rain streaks in order to more effectively and thoroughly purge the input images of rain streaks. We extensively evaluate the proposed approach in three representative datasets and experimental results corroborate our approach consistently outperforms state-of-the-art deraining algorithms.

1 Introduction

Images captured in rainy weather suffer from severe visibility degradation, which may impose great negative effects on many computer vision tasks, including object detection and tracking, autonomous driving, and semantic segmentation. In this regard, image deraining is an essential prerequisite for

many vision applications, seeking to recover the clean image from its complex entanglement with rain streaks. This problem is, however, very challenging and ill-posed, as the underlying background is totally unknown.

Many efforts have been dedicated to addressing the problem. Early investigations are mainly based on various image priors. One of the image priors closely related to deraining is that the main structures of an image are usually of low frequency while the details, such as rain streaks, are often of high frequency [Perona and Malik, 1990]. Naturally, the pioneering work of image deraining first adopts bilateral filtering to decompose the images into low frequency maps and high frequency maps, and then deal with rain streaks in the high frequency maps using dictionary learning [Fu *et al.*, 2011]. Later, more approaches have been proposed based on this prior [Wang *et al.*, 2017], as well as other image priors, such as Gaussian mixture model [Li *et al.*, 2016] and low-rank representations [Guo *et al.*, 2019]. However, these hand-crafted image priors are incapable of disentangling structures, particularly exquisite ones, from rain streaks, thereby reconstructing unsatisfactory clean images.

Recently, the performance of image deraining is boosted by deep convolutional neural networks (CNNs), which aim at capturing a variety of image characteristics by learning a complex model from massive data. The first deep learning-based deraining framework is proposed by [Fu *et al.*, 2017]. In this work, after a frequency-based decomposition operation, a three-layer CNN is directly adopted to extract rain streaks from the high-frequency maps. Later, more CNN models have been proposed by introducing either extra network modules [Ren *et al.*, 2019] or task-specific auxiliary information [Hu *et al.*, 2019a; Zhang and Patel, 2018] to guide the learning process, attempting to capture more powerful features to distinguish image structures and rain streaks. However, these models still have several shortcomings. First, these approaches perform frequency decomposition only at the image level, making the rain streaks mistakenly assigned to low-frequency maps difficult to be effectively removed and, simultaneously, the delicate structures assigned to high-frequency maps difficult to be recovered in the clean images.

*Joint first authors

†Corresponding authors: mingqiang.wei/hrxie2@gmail.com

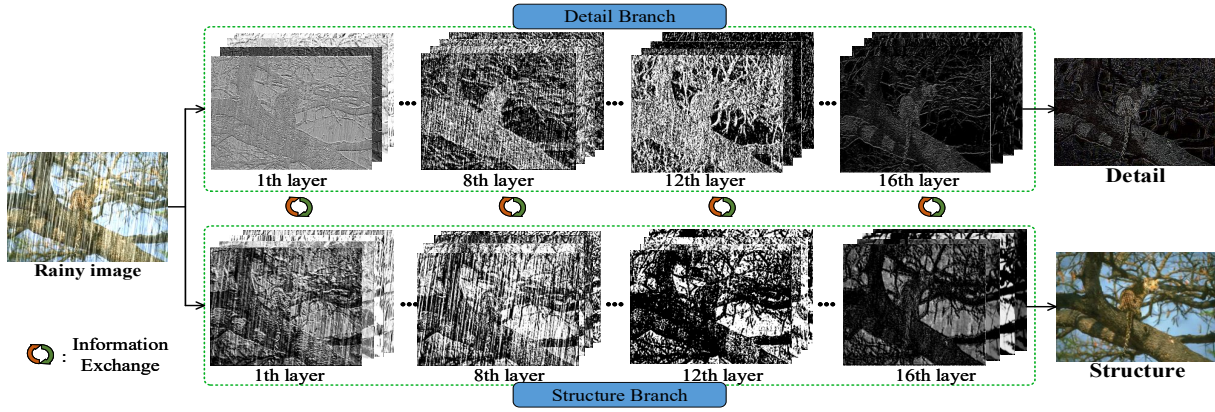


Figure 1: Visualization of feature maps in progressive frequency decomposition. The first row refers to the extraction of high-frequency details from the rainy image. As the layer goes deeper, the feature maps emphasize more and more on the foreground and are less entangled with rain streaks. While in the low-frequency structure branch, the network gradually obtains clean features highlighted on the background.

Second, there lack of interactive mechanisms between the low-frequency maps and high-frequency maps in the training procedure. Third, traditional convolutional filters are often consistent in all directions while the rain streaks in an image usually head in one direction, i.e. the wind direction; this property is ignored by most of current solutions.

In order to comprehensively address these shortcomings, in the paper, we propose a novel network with dual branches for single image deraining. Compared with existing solutions, the proposed network has three compelling characteristics. First, unlike previous algorithms, we propose to perform frequency decomposition at feature-level instead of image-level. By this way, the proposed network is able to generate low-frequency maps and high-frequency maps from feature maps at different layers, and hence allows these maps to be continuously refined during the training procedure (see Fig. 1). Second, we further establish communication channels between the dual branches, promoting the information propagation between low-frequency maps and high-frequency maps during the training procedure. Such a mechanism is not only helpful for separating more rain streaks from low-frequency maps to facilitate deraining but also useful for extracting more delicate features from high-frequency maps and add them back to low-frequency maps, enhancing clean image reconstruction. Third, most existing methods employ convolutional filters harmonious in all directions but ignore the fact that the rain streaks always head to the wind direction in an image, and hence are sub-optimal for image deraining. In order to take full advantage of this phenomenon, we propose a novel cross-median filter to capture the direction of rain streaks, aiming at producing more representative features to thoroughly purge the input image of rain streaks. We extensively evaluate the proposed network on three famous image deraining datasets. Experimental results demonstrate the effectiveness of the proposed network, consistently outperforming state-of-the-art approaches in most metrics. Our contributions can be summarized as:

- We propose a novel network with dual branches for single image deraining, conducting frequency decomposition

tion at feature level instead of image level so as to gradually and iteratively refine both low frequency maps and high frequency maps during training.

- We propose a new mechanism to promote interactions between low frequency maps and high frequency maps, facilitating both rain streak removal and fine features recovery; we further propose a novel direction-aware filter to more efficiently and effectively capture rain streaks in training.
- We set state-of-the-art performance of single image deraining on three famous datasets.

2 Related Work

2.1 Conventional Methods

Mainstream in conventional methods models image deraining as an image decomposition problem, where a rainy image is decomposed into a clean background layer and a rain streak layer [Fu *et al.*, 2011]. This strategy is followed by [Wang *et al.*, 2017], which introduce more prior knowledge, such as depth of field and color variance, for better extracting the rain streaks from the detail layer. In addition, many other image priors are exploited for image deraining. [Chen and Hsu, 2013] first adopts low rank representation to describe the non-local similarity in different rain patches, which is further explored by [Zhang and Patel, 2017]. [Luo *et al.*, 2015] considers the difference in rain streak layer and background layer, based on which they propose a novel discriminative sparse coding method. [Li *et al.*, 2016] exploits Gaussian mixture models for rain removal, which is learned on small patches that can accommodate a variety of background appearances and rain streak appearances. [Gu *et al.*, 2017] combines analysis sparse representation and synthesis sparse representation to better separate rain streaks and image textures.

2.2 Deep Learning-based Methods

Using frequency domain decomposition and residual connections, [Fu *et al.*, 2017] first employs a three-layer CNN to

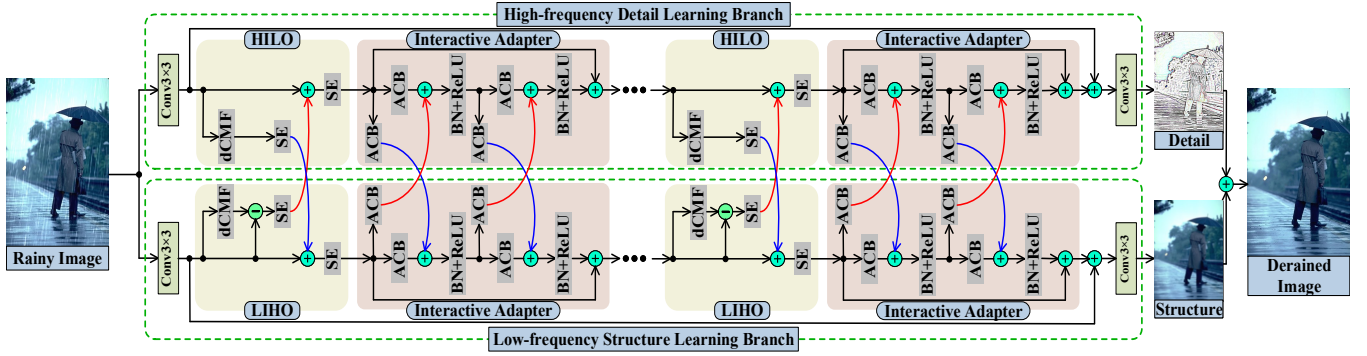


Figure 2: Our proposed network learns decomposed labels via two parallel yet interactive branches, where a detail learning branch keeps peeling off low-frequency components while reusing high-frequency features stripped from a structure learning branch, and vice versa. The two branches share similar structures. For detail learning branch, features are first fed into a High In Low Out (HILO) module, which extracts the low-frequency (blue arrow) using the proposed direction-aware Cross-Median Filter (dCMF) and accepts the high-frequency (red arrow) from a corresponding Low In High Out (LIHO) module, and then an Interactive Adapter based on Asymmetric Conv Block (ACB) is used for feature learning and adapting. This procedure is conducted iteratively for robust and effective learning on both clean details and structures.

extract rain streaks from the detail layer. Thereafter, advancing network modules are introduced, such as residual block [Deng *et al.*, 2020], dilated convolution [Deng *et al.*, 2020] and recursive block [Ren *et al.*, 2019]. Among them, [Yu *et al.*, 2019] and adopts a coarse-to-fine strategy by adding supervision on different learning stages. Due to the complexity of rain streaks and their composition with the background, several methods are proposed to separate the task using dual-path networks [Deng *et al.*, 2020; Pan *et al.*, 2018], or adopt a multi-stage strategy using recurrent neural networks to progressively recover the clean image [Li *et al.*, 2018; Ren *et al.*, 2019]. [Pan *et al.*, 2018] proposes to recover low frequency image structures and high frequency image details separately using two parallel network branches. [Deng *et al.*, 2020] takes advantage of another network branch to find back lost details. GAN is also exploited by [Zhang *et al.*, 2019] to refine the deraining results for more visual appealing effects. Besides, [Hu *et al.*, 2019a] builds a dataset to describe heavy rainy scenes using depth images to associate rain streaks and rainy haze. [Wang *et al.*, 2019] proposes a real rain dataset using video-based deraining results and adopt a directional IRNN to learn spatial attention for guiding the network. [Li *et al.*, 2019a] presents a comprehensive benchmark named MPID for evaluation of various deraining methods. [Zhu *et al.*, 2019] first utilizes CycleGAN for single image deraining. For removing different scales of rain streaks, [Yang *et al.*, 2020] designs a fractal band learning network trained with self-supervision for scale-robust rain streak removal.

3 Method

In this section, we introduce our proposed method built on frequency decomposition. Denote the rain-free label image as $\mathbf{I}_{gt}$, it is first decomposed into a low-frequency structure map $\tilde{\mathbf{S}}$ and a high-frequency detail map $\tilde{\mathbf{D}}$. Our goal is to accurately predict both $\tilde{\mathbf{S}}$ and $\tilde{\mathbf{D}}$ using a single rainy input $\mathbf{I}$, so as to recover a high-fidelity derained image with abundant details and minimum distortions. Our solution resorts to feature-level frequency decomposition along and across a

parallel network architecture. The extraction of a particular frequency is learned along each branch towards the decomposed label, while components of the other frequency are continuously delivered across the branch. We first interpret the multiple labels and their loss functions, and then go into details of the direction-aware Cross-Median Filter (dCMF) and the interactive adapter to explain how we isolate the different frequency to enhance communication between branches.

3.1 Label Decomposition and Loss Functions

Label decoupling in low-level vision is a strategy that models the final task as a composition of several easier sub-tasks [Wei *et al.*, 2020; Pan *et al.*, 2018], which we argue is particularly effective in image deraining due to the complex entanglement of rain streaks and image contents [Deng *et al.*, 2020]. In our case, the label image $\mathbf{I}_{gt}$ is decomposed into low- and high- frequencies using a low-pass image filter, where the high frequency part $\tilde{\mathbf{D}}$ contains abundant image details and the low frequency part $\tilde{\mathbf{S}}$ characterizes main image structures. We employ two network branches to deal with structures and details respectively. For the detail branch, we minimize the L_1 distance between the detail and the output of the high-frequency branch to preserve gradient discontinuities, while for the structure branch we enforce L_2 loss to encourage global smoothness

$$L_d = \|f(\mathbf{I}) - \tilde{\mathbf{D}}\|_1, \quad (1)$$

$$L_s = \|g(\mathbf{I}) - \tilde{\mathbf{S}}\|_2, \quad (2)$$

where $f(\cdot)$ and $g(\cdot)$ represent the detail branch and the structure branch, respectively. Furthermore, to ensure the fidelity and structure integrity in the composited derained image, we combine L_1 loss and SSIM loss to constrain the final result, which can be written as

$$L_r = \|\mathbf{I}_p - \mathbf{I}_{gt}\|_1 + 1 - SSIM(\mathbf{I}_p, \mathbf{I}_{gt}). \quad (3)$$

where $\mathbf{I}_p = f(\mathbf{I}) + g(\mathbf{I})$ is the output of the whole network as the prediction of the rain-free background. Given the aforementioned three kinds of loss, the overall loss can be formulated as

$$L_c = \lambda_1 L_d + \lambda_2 L_s + \lambda_3 L_r, \quad (4)$$

where λ_1 , λ_2 and λ_3 are the weighting parameters, which in our experiments are all fixed to be 1.

3.2 Feature-level Frequency Decomposition

As indicated in [Chen *et al.*, 2019], feature maps are also composed of different frequencies. While in low-level vision, as can be observed in Fig. 1, low- and high-frequency feature maps also reflect image structures and details, which are complementary and tightly correlated. To enhance communication across frequencies, we propose a direction-aware Cross-Median Filter (dCMF) to explicitly extract low-frequency components from an entanglement of background features and rain streak patterns with varying falling directions, and an interactive adapter to implicitly enhance feature decomposition through interactive connections.

Direction-aware Cross-Median Filter

Direction-aware Cross-Median Filter (dCMF) aims at separating different frequencies on rain-affected features in the communication paths across the branches. As shown in Fig. 2, in the HILO module, dCMF extracts low-frequency components, which are adapted by channel-wise attention using Squeeze-and-Excitation and then sent to the structure learning branch, while in the LIHO module, the residual of the dCMF filtering result is delivered to the high-frequency branch.

Our key observation in designing dCMF is that, rain patterns, not only in rainy images but also in their feature maps, have the properties of relatively high intensity and globally consistent orientation, which makes them distinguishable to background patterns in similar scales (see Fig. 1). To take full advantage of this prior, we make three additional modifications upon a naive low-pass image filter. First, we adopt median pooling rather than averaging to avoid extreme values brought by rain streaks in feature maps. Second, we replace 2D filtering kernels with 1D directional lines. It is intuitive that rain mainly falls in the vertical direction and leaves traces of globally vertical streaks. Suppose a 1D filtering line centered at a single rain streak, and the angle between them is denoted as α . It is clear that, when $\alpha = 90^\circ$ the filtering kernel is least affected by this rain streak and other possible rain directions in this rainy image. However, real scenes are much more complicated, the direction of rain streaks can be affected by many factors, such as wind and obstacles. To take this complexity into consideration, we enumerate different orientations of the 1D kernels as shown in the first column of Fig. 3 (a), extending CMF_1 to the set $\{CMF_1, CMF_2, CMF_3\}$ and leverage a self-attention mechanism to learn the importance of each direction. To enforce attention on both feature channels and different groups of filtering results, we adopt a similar strategy with [Li *et al.*, 2019b]: three groups of filtering results are aggregated through addition and global average pooling for computing the individual attentions for each group, and the weighted sum of the groups constitute

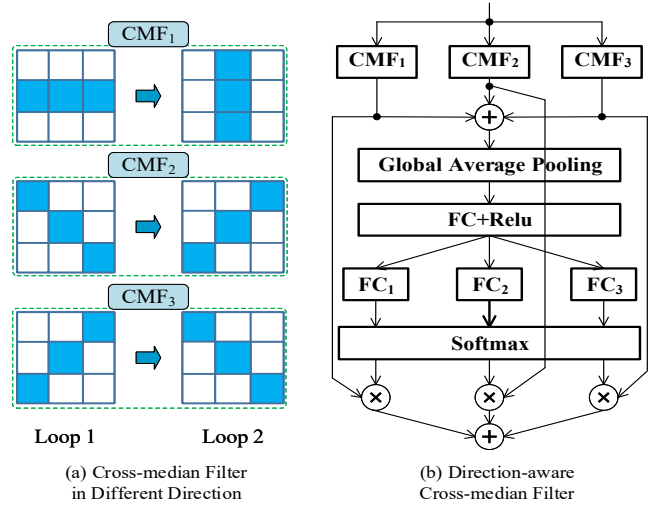


Figure 3: The detailed structure of direction-aware Cross-Median Filter (dCMF), where the self-attention module described in (b) determines the weights of each direction in (a).

the final output of dCMF. Third, each 1D kernel in CMF_i is followed by a crisscross counterpart to constitute a complete Cross-Median Filter (CMF). Each CMF has exactly the same receptive field with the corresponding 2D kernel, but is more robust against rain streaks since the second kernel operates on the feature maps with greatly reduced rain patterns.

Interactive Adapters

For frequency exchange across branches, dCMF leverages prior knowledge on rain streaks to explicitly compute low-frequency components, while the interactive adapter uses learnable convolutional kernels as the frequency filter. The behavior of the interactive adapter is guided by the decomposed labels as a complementary to dCMF-based frequency decomposition. As shown in Fig. 2, we adopt asymmetric convolution blocks (ACB) [Hu *et al.*, 2019b] to integrate features from different branches and automatically adjust the information exchange between branches. Suppose $z_d, z_s \in R^{F \times H \times W}$ are the input features of the adapter, Ψ_{d_i} and Ψ_{s_i} denotes the ACB unit, d and s refer to the detail branch and the structure branch, respectively. The basic function of the interactive adapter can be expressed as

$$z'_d = S(\Psi_{d_1}(z_d)) + \Psi_{s_2}(z_s), \quad (5)$$

where $S(\cdot)$ refers to batch normalization followed by rectified linear unit and $z'_d \in R^{F \times H \times W}$ is the output feature. For each interactive adapter, the above function is computed twice, and the second output feature can be obtained by simply replace z_d, z_s with z'_d, z'_s . Due to the symmetry in interactive adapters, the corresponding function in the structure branch can be easily inferred. Through these dual interaction functions, redundant information can be efficiently transferred to another path and thus encourages exploration on new features. In the adapter, computation occurs in parallel at both the dual branches and the interactive paths, which allows accurate decomposition in an information intensive but computation efficient way.

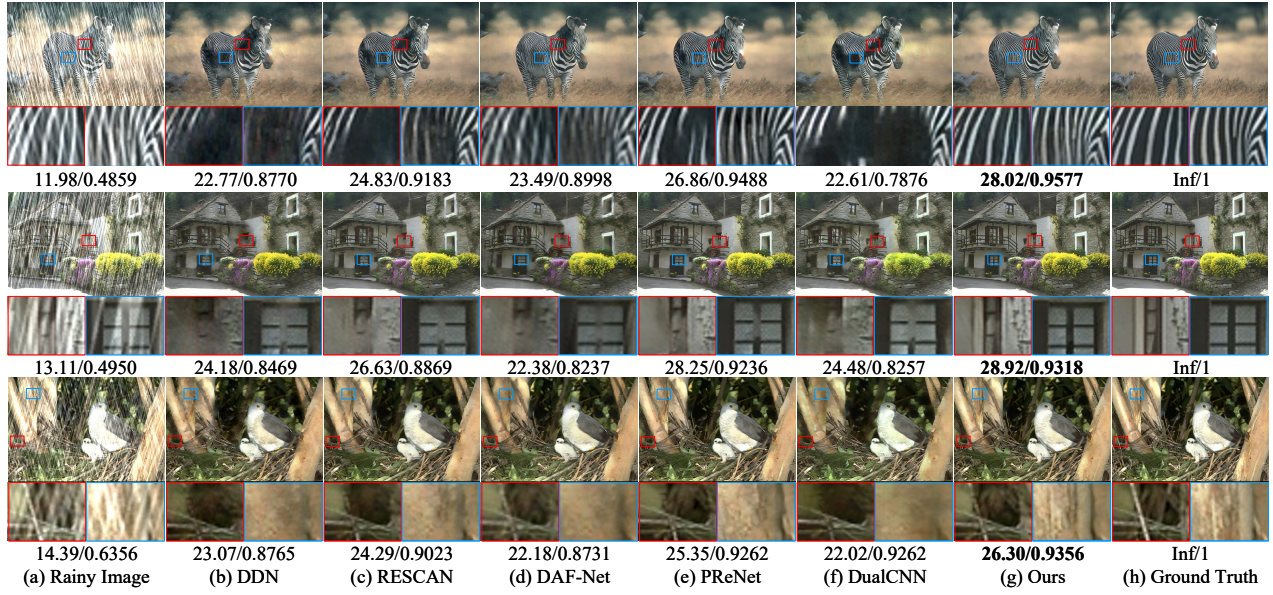


Figure 4: Image deraining results tested on the synthetic datasets. From (a)-(h): (a) the rainy images, and the deraining results of (b) DDN, (c) RESCAN, (d) DAF-Net, (e) PReNet, (f) DualCNN, (g) ours and (h) the ground truth, respectively.



Figure 5: Image deraining results tested on the real-world datasets. From (a)-(h): (a) the rainy images, and the deraining results of (b) DDN, (c) RESCAN, (d) SPA-Net, (e) PReNet, (f) DualCNN, and (g) ours, respectively.

4 Experiments and Results

In this section, we evaluate our method on three synthetic datasets: Rain200L, Rain200H [Yang *et al.*, 2017] and Rain800 [Zhang *et al.*, 2019]. Since no rain-free ground truths for real-world images are provided, we performed a user study on several real-world datasets. Please refer to the supplementary materials for more details.

4.1 Comparison with the State-of-the-Arts

We compare our method with two traditional methods: GMM [Li *et al.*, 2016] and DSC [Luo *et al.*, 2015], and six learning-

based methods: DDN [Fu *et al.*, 2017], RESCAN [Li *et al.*, 2018], DualCNN [Pan *et al.*, 2018], DAF-Net [Hu *et al.*, 2019a], SPA-Net [Wang *et al.*, 2019], PReNet [Ren *et al.*, 2019], and DRD-Net [Deng *et al.*, 2020].

The quantitative evaluation results of PSNR and SSIM are shown in Tab. 1. As can be observed, our proposed method mostly obtains the highest values of PSNR and SSIM than other methods on the synthetic datasets. The visual comparisons are shown in Fig. 4, from which one can observe that our method better remains the structure and preserves the detail of images.

Dataset	Rain200L		Rain200H		Rain800	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
GMM	27.16	0.8982	13.04	0.4673	24.04	0.8675
DSC	25.68	0.8751	13.17	0.4272	20.95	0.7530
DDN	33.01	0.9692	24.64	0.8489	24.04	0.8675
DualCNN	32.93	0.9575	24.09	0.7632	23.83	0.8395
RESCAN	37.07	0.9867	26.60	0.8974	24.09	0.8410
RWL	36.75	0.9632	26.89	0.8406	27.79	0.8795
DAF-Net	32.07	0.9641	24.65	0.8607	25.27	0.8895
SPA-Net	31.59	0.9652	23.04	0.8522	22.41	0.8382
PReNet	36.76	0.9796	28.08	0.8871	<u>26.61</u>	0.9015
DRD-Net	<u>37.15</u>	<u>0.9873</u>	28.16	<u>0.9201</u>	26.32	<u>0.9018</u>
Ours	37.74	0.9896	<u>28.09</u>	0.9316	26.86	0.9164

Table 1: Quantitative experiments evaluated on three recognized synthetic datasets. The 1st and 2nd best results boldfaced and underlined.

Furthermore, visual evaluation on a series of real-world rainy images is provided in Fig. 5, from which one can observe that our method can not only remove real rain streaks but also better preserve the image structures and details. As can be seen, challenging areas, such as the textures of the pillars and the border of the wall, are well preserved by our method.

4.2 Ablation Study

Ablation Study on Different Components: In Tab. 2, we show quantitative results in order to validate the effectiveness of: dual branch architecture, interactive adapter, direction-aware Cross-Median Filter in HILO and LIHO modules.

- **BL:** Baseline (BL) indicates that we use a single branch with the residual network to learn a rainy-to-derained function.
- **DBL:** Dual Baseline (DBL) indicates that we use two same branches without interaction for single image rain removal, which learns the detail image and the structure image respectively.
- **DBL+I:** Replacing the residual block with Interactive Adapter in DBL (We remove the HILO and LIHO from our proposed network).
- **DBL+I+O:** Adding HILO, LIHO and OC to DBL (Interactive adapter is replaced by the Octave Conv (OC) in our network).

Dataset	Metrics	BL	DBL	DBL+I	DBL+I+O	Ours
Rain200L	PSNR	35.57	36.33	36.96	<u>37.33</u>	37.74
	SSIM	0.9759	0.9864	0.9879	<u>0.9889</u>	0.9896
Rain200H	PSNR	26.20	27.03	<u>27.95</u>	27.58	28.09
	SSIM	0.8245	0.9212	<u>0.9310</u>	0.9260	0.9316
Rain800	PSNR	25.16	25.23	25.46	<u>26.19</u>	26.84
	SSIM	0.9008	0.9043	0.9034	<u>0.9086</u>	0.9164

Table 2: Quantitative comparison between our network and other network architectures on Rain200H.

Datasets	Metrics	MF	Gaussian	Our
Rain200H	PSNR	27.86	22.58	28.09
	SSIM	0.9290	0.8064	0.9316

Table 3: Quantitative evaluation. The result of our method by replacing the dCMF with the MF and the Gaussian filter on Rain200H, respectively.

Metrics	DSC	DDN	RES CAN	Dual CNN	DAF Net	PRe Net	SPA Net	Ours
PSNR	13.17	24.64	26.60	24.09	24.65	28.08	23.04	28.09
SSIM	0.4272	0.8489	0.8974	0.7632	0.8607	0.8871	0.8522	0.9316
Avg time	92.9s	0.03s	0.25s	0.06s	0.52s	0.20s	0.06s	0.31s

Table 4: Averaged time (in seconds) and performances of different methods on Rain200H.

Analysis on dCMF: To validate the effectiveness of our proposed dCMF in HILO and LIHO, we remove them from our method and the result can be found in the 'DBL+I' column of Tab. 2. In addition, we replace dCMF with ordinary 3×3 kernel median filter and Gaussian filter, as shown in Tab. 3. Generally speaking, both results show clear advantage of dCMF in the deraining task, and we also visually inspected that deraining results without dCMF suffer from heavier degradations.

Analysis on Interactive Adapter: In order to further analyze the necessity of the interactive adapter, we replace it with another frequency decomposition method: octave convolution (OC) [Chen *et al.*, 2019], in our network and the result is shown in Tab. 2, which demonstrates that not only can the interactive blocks between two branches improve the performance of the network but also can our interactive adapter outperforms the Octave Conv in the deraining task.

4.3 Running Time

We compare the running time of our method with different approaches on Rain200H. As shown in Tab. 4, our method is not the fastest one, but reaches a reasonable balance between performance and efficiency.

5 Conclusion

We propose an interactive dual-branch network where features of different frequencies are learned and exchanged to enhance the performance of single image deraining. The communication between high- and low- frequency branches relies on two key designs: (1) instead of using convolutional filters consistent in all directions, we propose direction-aware Cross-Median Filter to thoroughly purge rain patterns in frequency decomposition; (2) we present the interactive adapter to enhance feature learning and interaction towards decomposed labels.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Nos. 62032011, 61502137).

References

- [Chen and Hsu, 2013] Yi-Lei Chen and Chiou-Ting Hsu. A generalized low-rank appearance model for spatio-temporally correlated rain streaks. In *ICCV*, pages 1968–1975, 2013.
- [Chen *et al.*, 2019] Yunpeng Chen, Haoqi Fan, Bing Xu, Zhicheng Yan, Yannis Kalantidis, Marcus Rohrbach, Shuicheng Yan, and Jiashi Feng. Drop an octave: Reducing spatial redundancy in convolutional neural networks with octave convolution. In *ICCV*, pages 3435–3444, 2019.
- [Deng *et al.*, 2020] 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 *CVPR*, pages 14548–14557, 2020.
- [Fu *et al.*, 2011] Yu-Hsiang Fu, Li-Wei Kang, Chia-Wen Lin, and Chiou-Ting Hsu. Single-frame-based rain removal via image decomposition. In *ICASSP*, pages 1453–1456, 2011.
- [Fu *et al.*, 2017] Xueyang Fu, Jiabin Huang, Delu Zeng, Yue Huang, Xinghao Ding, and John Paisley. Removing rain from single images via a deep detail network. In *CVPR*, pages 3855–3863, 2017.
- [Gu *et al.*, 2017] Shuhang Gu, Deyu Meng, Wangmeng Zuo, and Lei Zhang. Joint convolutional analysis and synthesis sparse representation for single image layer separation. In *ICCV*, pages 1708–1716, 2017.
- [Guo *et al.*, 2019] Xianglin Guo, Xingyu Xie, Guangcan Liu, Mingqiang Wei, and Jun Wang. Robust low-rank subspace segmentation with finite mixture noise. *PR*, 93:55–67, 2019.
- [Hu *et al.*, 2019a] Xiaowei Hu, Chi-Wing Fu, Lei Zhu, and Pheng-Ann Heng. Depth-attentional features for single-image rain removal. In *CVPR*, pages 8022–8031, 2019.
- [Hu *et al.*, 2019b] Xinxin Hu, Kailun Yang, Lei Fei, and Kaiwei Wang. Acnet: Attention based network to exploit complementary features for rgbd semantic segmentation. In *ICIP*, pages 1440–1444. IEEE, 2019.
- [Li *et al.*, 2016] Yu Li, Robby T Tan, Xiaojie Guo, Jiangbo Lu, and Michael S Brown. Rain streak removal using layer priors. In *CVPR*, pages 2736–2744, 2016.
- [Li *et al.*, 2018] Xia Li, Jianlong Wu, Zhouchen Lin, Hong Liu, and Hongbin Zha. Recurrent squeeze-and-excitation context aggregation net for single image deraining. In *ECCV*, pages 254–269, 2018.
- [Li *et al.*, 2019a] Siyuan Li, Iago Breno Araujo, Wenqi Ren, Zhangyang Wang, Eric K Tokuda, Roberto Hirata Junior, Roberto Cesar-Junior, Jiawan Zhang, Xiaojie Guo, and Xiaochun Cao. Single image deraining: A comprehensive benchmark analysis. In *CVPR*, pages 3838–3847, 2019.
- [Li *et al.*, 2019b] Xiang Li, Wenhai Wang, Xiaolin Hu, and Jian Yang. Selective kernel networks. In *CVPR*, pages 510–519, 2019.
- [Luo *et al.*, 2015] Yu Luo, Yong Xu, and Hui Ji. Removing rain from a single image via discriminative sparse coding. In *ICCV*, pages 3397–3405, 2015.
- [Pan *et al.*, 2018] Jinshan Pan, Sifei Liu, Deqing Sun, Jiawei Zhang, Yang Liu, Jimmy Ren, Zechao Li, Jinhui Tang, Huchuan Lu, Yu-Wing Tai, et al. Learning dual convolutional neural networks for low-level vision. In *CVPR*, pages 3070–3079, 2018.
- [Perona and Malik, 1990] Pietro Perona and Jitendra Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE TPAMI*, 12(7):629–639, 1990.
- [Ren *et al.*, 2019] Dongwei Ren, Wangmeng Zuo, Qinghua Hu, Pengfei Zhu, and Deyu Meng. Progressive image deraining networks: a better and simpler baseline. In *CVPR*, pages 3937–3946, 2019.
- [Wang *et al.*, 2017] Yinglong Wang, Shuaicheng Liu, Chen Chen, and Bing Zeng. A hierarchical approach for rain or snow removing in a single color image. *IEEE TIP*, 26(8):3936–3950, 2017.
- [Wang *et al.*, 2019] 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 *CVPR*, pages 12270–12279, 2019.
- [Wei *et al.*, 2020] Jun Wei, Shuhui Wang, Zhe Wu, Chi Su, Qingming Huang, and Qi Tian. Label decoupling framework for salient object detection. In *CVPR*, pages 13025–13034, 2020.
- [Yang *et al.*, 2017] 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 *CVPR*, pages 1357–1366, 2017.
- [Yang *et al.*, 2020] Wenhan Yang, Shiqi Wang, Dejia Xu, Xiaodong Wang, and Jiaying Liu. Towards scale-free rain streak removal via self-supervised fractal band learning. In *AAAI*, pages 12629–12636, 2020.
- [Yu *et al.*, 2019] Weijiang Yu, Zhe Huang, Wayne Zhang, Litong Feng, and Nong Xiao. Gradual network for single image de-raining. In *ACM MM*, pages 1795–1804, 2019.
- [Zhang and Patel, 2017] He Zhang and Vishal M Patel. Convolutional sparse and low-rank coding-based rain streak removal. In *WACV*, pages 1259–1267, 2017.
- [Zhang and Patel, 2018] He Zhang and Vishal M Patel. Density-aware single image de-raining using a multi-stream dense network. In *CVPR*, pages 695–704, 2018.
- [Zhang *et al.*, 2019] He Zhang, Vishwanath Sindagi, and Vishal M Patel. Image de-raining using a conditional generative adversarial network. *IEEE TCSVT*, 2019.
- [Zhu *et al.*, 2019] Hongyuan Zhu, Xi Peng, Joey Tianyi Zhou, Songfan Yang, Vijay Chanderasekh, Liyuan Li, and Joo-Hwee Lim. Single image rain removal with unpaired information: A differentiable programming perspective. In *AAAI*, volume 33, pages 9332–9339, 2019.