

SINGLE IMAGE DE-RAINING WITH HIGH-LOW FREQUENCY GUIDANCE

Ying Zhang, Youjun Xiang, Lei Cai, Yuli Fu, Wanliang Huo, Junjun Xia*

School of Electronic and Information Engineering, South China University of Technology
Guangzhou 510641, China

ABSTRACT

Rain removal is a highly demanding task because a rainy image in computer lacks discriminative information to distinguish the image details from the rain streaks. In this paper, we present a new High-Low-Frequency Guided De-raining (HLFGD) method to remove the rain streaks clearly while reserve the image details. Specifically, the proposed HLFGD is built with three network branches, namely global-structure branch, de-raining branch, and edge-detail branch, which achieve the collaboration by concatenating intermediate features. Among them, the global-structure and edge-detail branches aim to explore the high-low frequency information, and the de-raining branch leverages the resulting spatial frequency information to restore the global structure of image and to retain fine edge details of objects during the de-raining process. Besides, a new architecture unit, called Residual Coordinate Attention Block (RCAB), is proposed to improve the effect of rain removal. Experimental results show the superiority of our method for image de-raining quantitatively and qualitatively.

Index Terms— Rain removal, high-low frequency, global structure, edge details, Coordinate Attention.

1. INTRODUCTION

Rain streaks in an image greatly degrade the visual quality, and could obstruct the performance of computer vision algorithms. Hence, image de-raining is expected to have broad application prospects, including video surveillance, autonomous driving, object tracking, and more.

To date, many image de-raining algorithms (*e.g.*, [1–13]) have been developed with encouraging performance. Despite the great improvements of image de-raining, they are usually hindered to both remove the rain streaks and reserve image details effectively, especially for those images with heavy rain. That is because the magnitude of image details is close to and even smaller than that of rain streaks, but a rainy image in computer lacks effective discriminative information to describe them separately [11]. As a result, the image details and

This work is partially supported by Natural Science Foundation of Guangdong Province (2019A1515010861), Guangzhou Technical Project (Grant: 201902020008), and NSFC (Grant: 61471174). *Corresponding author.

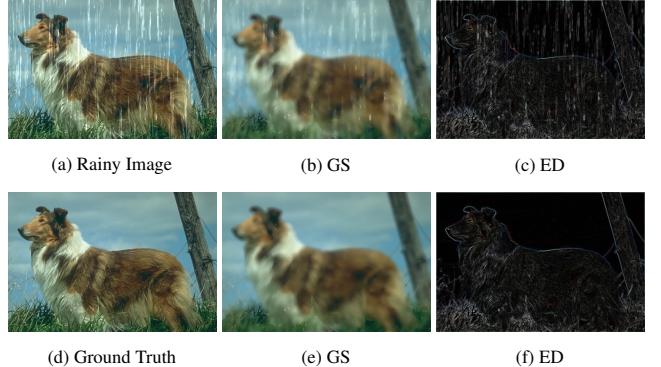


Fig. 1. High-low frequency components derived from a rainy image and its corresponding ground truth image. From (a)-(c): the rainy image and its corresponding global structure (GS) and edge details (ED). From (d)-(f): the ground truth image and its corresponding global structure (GS) and edge details (ED).

the rain streaks are likely to be removed simultaneously. To tackle this issue, [2,3] utilize image decomposition method to restore the high-frequency layer to reserve the image details, but their methods cannot deal with rain streaks with different directions and sizes. And [11] also aims to recover details, but the various and heavy rain streaks are still not well removed. On the other hand, [1,7,13] introduce the rain density detection, and [5,6,8–10,12,14] propose the recurrent methods to remove the rain streaks progressively. However, these methods suffer from over de-raining, losing the image details.

To address these limitations, we develop a new method, called High-Low-Frequency Guided De-raining (HLFGD), to remove the rain streaks while retain the background details effectively. The motivation behind our method lies in that the high frequency information is encoded with edge details (ED) of objects, while the low frequency information is encoded with the global structure (GS) [15]. As shown in Fig. 1, the global structure degraded by rain is commonly reflected in the large smoothed rain streaks. And the degradation of edge details is reflected in that the rain streaks and the edge of objects are overlapped intrinsically. Removing rain streaks from the low or high frequency domain is supposed to be easier than directly removing the rain streaks from a whole

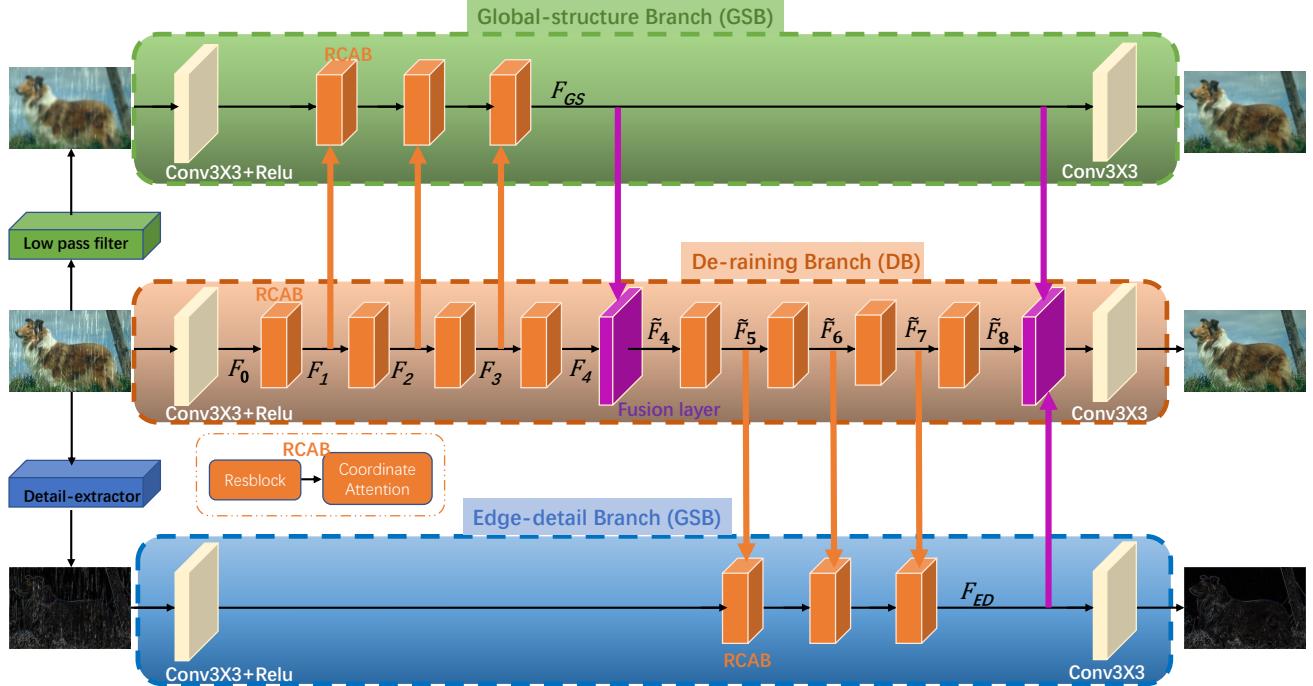


Fig. 2. The proposed HLFGD network containing three branches.

rainy image. With these considerations, we design a three-branch collaboration network architecture, which consists of global-structure branch (GSB), de-raining branch (DB), and edge-detail branch (EDB). GSB and EDB are used for auxiliary tasks for guiding rain removal processed by DB. GSB aims to resolve the global-structure map of rainy image to that of the rain-free counterpart. It aggregates multi-level representations from DB to output rain-free low frequency information to guide the first half de-raining branch to make the de-raining branch focus on the restoration of global structure. On the basis of well-retained global structure in de-raining branch, EDB resolves the edge details map of rainy image to that of the rain-free counterpart. It also aggregates multi-level representations from DB to output high frequency information, which will guide the latter half de-raining branch to make DB focus on the restoration of edge details. Furthermore, for better restoration of damaged frequency information caused by rain streaks, we propose Residual Coordinate Attention Block (RCAB) to help the network more accurately locate the area needed to be restored since Coordinate Attention [16] can capture long-range dependencies and preserve precise positional information [16]. Our main contributions are: 1) we design a new high-low frequency guided de-raining method for simultaneously de-raining and preserving image details; 2) we propose an effective Residual Coordinate Attention Block for obtaining more spatial information to improve the performance of rain removal; 3) we achieve clear image detail reservation and numerical improvements of our method over the state-of-the-arts.

2. PROPOSED METHOD

Fig. 2 displays the framework of the proposed HLFGD, from top to bottom: global-structure branch (GSB) for exploring low-frequency components, de-raining branch (DB) for image de-raining, and edge-detail branch (EDB) for exploring high-frequency components. In each branch, the first layer is used to extract the features of images, while the last layer is used to transform the 32-channel features to 3-channel RGB image. We preserve the original image resolution because there is no down sampling in the whole network.

We construct a basic block called Residual Coordinate Attention Block (RCAB) by connecting Coordinate Attention [16] after a Resblock [17]. All of the RCAB have 32 input channels and 32 output channels. The reason of utilizing Coordinate Attention [16] is that it can aggregate the features along the horizontal and vertical directions respectively, producing two direction-aware features for the object detection task or the semantic segmentation task to help the models locate the objects of interest more accurately [16]. We believe that, in de-raining task, when the attention along two spatial directions reweights the input tensor, it can help the network more accurately locate the position of objects needed to be restored, namely the areas covered by rain streaks. Thus, the performance of de-raining models will be improved, and this will be demonstrated in our ablation experiments.

2.1. De-raining Branch

DB containing 8 RCABs is used to successively focus on restoring low-frequency information and high-frequency in-

formation of the rainy image.

$$F_i = f_{RCAB,i}(F_{i-1}), 1 \leq i \leq 4 \quad (1)$$

where $f_{RCAB,i}(\cdot)$ and F_i respectively denote the i -th RCAB module and the features generated by i -th RCAB. F_0 denotes the original features extracted from the rainy image.

Then, F_4 and F_{GS} which represents the reconstructed global-structure features, are jointly processed by a 3×3 convolution layer as a fusion block. Then the fusion features are taken as input to the second half of DB.

$$\tilde{F}_i = f_{RCAB,i}(\tilde{F}_{i-1}), 5 \leq i \leq 8 \quad (2)$$

where \tilde{F}_4 represents the fusion features, carrying rich global structure information provided by GSB. Similarly, the second half of DB is guided by EDB, which aims at reconstructing the features of rain-free edge details F_{ED} . Finally, \tilde{F}_8 , F_{GS} and F_{ED} are fused to generate the final rain-free image.

2.2. Two guidance branches

Guidance branches contain GSB and EDB. As shown in Fig. 2, both two branches incorporate features from DB. This is motivated by the fact that DB with superior performance is supposed to generate well restored features of global structural and edge details, which can promote the recovery of low-frequency and high-frequency information respectively. Moreover, when we get features reconstructed by guidance branches, the obtained features representing restored global structure and edge details are integrated into DB to guide the rain-free image reconstruction in turn. In general, we use relatively simple recovery processes (i.e., GSB and EDB) to guide the recovery of the whole rainy image (i.e., DB).

Global-structure Branch: By low-pass filtering the original input, we can get its low-frequency component, that is, the global structure of the rainy image. Then, GSB receives the intermediate features from DB, and finally outputs the rain-free global structure. The restored global structure features will also be integrated into DB. Similarly, our low-frequency branch is also constructed by RCABs, because RCAB can help network sense the white smooth areas caused by the coverage of rain streaks, which is beneficial for GSB to better locate and restore the damaged low-frequency components.

Edge-detail Branch: The composition of EDB is the same to GSB. The high-frequency components extracted from the rainy image are restored by EDB. Similar to the guidance role of GSB, the high-frequency restoration processed by EDB guides the latter half of DB, so that the final rain-free image has good performance on keeping details. In this paper, we use gradients to represent high-frequency edge details of objects. The gradient of each pixel $G(x, y)$ in the image I can be calculated as below:

$$\begin{aligned} G_h(x, y) &= I(x+1, y) - I(x-1, y), \\ G_v(x, y) &= I(x, y+1) - I(x, y-1), \\ G(x, y) &= \sqrt{G_h(x, y)^2 + G_v(x, y)^2}, \end{aligned} \quad (3)$$

where $I(x, y)$ represents the pixel value at pixel position (x, y) . $G_h(x, y)$ and $G_v(x, y)$ respectively denote the horizontal and vertical gradient. We use a convolution layer with fixed parameters to realize the gradient extraction.

2.3. Loss function:

We adopt MSE and SSIM as our loss function, which involves the output results of three branches.

$$\begin{aligned} Loss &= MSE(I_{Rec}^{GS}, I_{gt}^{GS}) \\ &\quad + 1 - SSIM(I_{out}, I_{gt}) \\ &\quad + MSE(G_{Rec}, G_{gt}), \end{aligned} \quad (4)$$

where I_{out} , I_{gt} , I_{gt}^{GS} and G_{gt} respectively represent the output of HLFGID, ground truth, and its corresponding global-structure and edge-detail maps, respectively. I_{Rec}^{GS} and G_{Rec} respectively denote the global-structure and edge-detail maps reconstructed by two guidance branches.

3. EXPERIMENTS

In this section, we conduct experiments on three synthetic datasets, and compare the proposed method with multiple state-of-the-art methods, including PReNet [8], REHEN [9], MSPFN [12], CVID [13] and EffDeRain [18]. The de-raining performances on synthetic datasets are evaluated in terms of peak signal-to-noise (PSNR) [19] and structural similarity (SSIM) [20].

3.1. Implementation details

The proposed HLFGD method is implemented with Pytorch [21] and trained on an NVIDIA TITAN XP GPU. The patch size of training images is 160×160 . The ADAM [22] solver with an initial learning rate 1×10^{-3} is applied to train our models. When arriving at 30 and 50 epochs, the learning rate is decayed by multiplying 0.2. We totally train our models with 80 epochs.

3.2. Comparison on synthetic datasets

The synthetic datasets for comparisons include Rain1400 [2], Rain100L [1] and Rain100H [1]. We used the original image resolution during inference for all methods. The quantitative results of the proposed and other related methods are listed in Table 1. It can be observed that our method achieves superior PSNR and SSIM results and exceeds other competing methods on average performance. Besides, it can be seen from Fig. 4 that the de-rained image derived by our method retains more image details while accurately removing rain streaks.

3.3. Comparison on real data

Besides the synthesized images, we also demonstrate the de-raining effects on two real rainy images. As can be seen in



Fig. 3. Visual quality comparison on two real rainy images.

Table 1. Comparasion of PSNR and SSIM results by different methods on three synthetic datasets. The best results are marked in bold.

	PreNet [8]	REHEN [9]	CVID [13]	MSPFN [12]	EfDeRain [18]	Ours
CVPR19	ACMMM19	TIP20	CVPR20	AAAI21		
Dataset	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Rain100H	29.43	0.899	27.86	0.857	27.93	0.877
	28.70	0.882	30.37	0.901	30.44	0.912
Rain100L	37.51	0.975	37.45	0.981	37.83	0.988
	36.03	0.971	34.18	0.958	38.79	0.984
Rain1400	32.54	0.939	31.25	0.909	28.96	0.938
	30.39	0.881	32.33	0.937	33.18	0.949
Average	33.16	0.938	32.19	0.916	31.57	0.934
	31.71	0.878	32.29	0.932	34.14	0.948

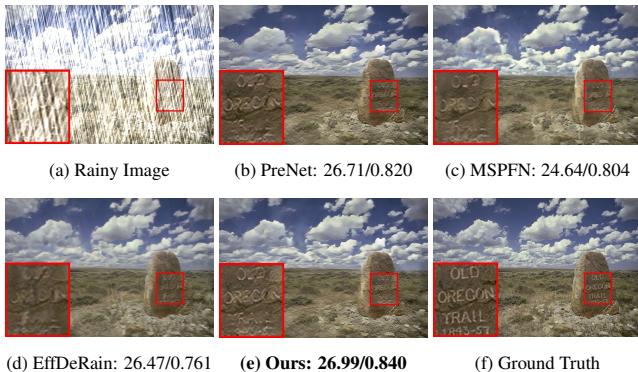


Fig. 4. De-rained image and PSNR(dB)/SSIM results on a synthesized rainy image from Rain100H [1].

Fig. 3, the competing methods either cannot completely remove the rain streaks or suffer from over de-raining issue, losing the image details, while the proposed method achieves visually more pleasing de-raining effects.

3.4. Ablation Study

Experimental results on different combinations of modules on Rain100H [1] can be shown in Table 2 and Fig 5. Note that DB w/o CA means that Coordinate Attention (CA) is removed from the RCAB. As can be seen from Table 2 that the performance of the de-raining branch (DB) decreases obviously after removing Coordinate Attention, which proves the effectiveness of RCAB in our de-raining network. By adding global-structure branch (GSB) to DB, the performance is effectively improved. In Fig 5 (c) and (d), the sharp area due

to the failure to remove the rain streaks is also smooth to a great extent, which illustrates that restoring low-frequency as an auxiliary task can promote the de-raining effect. Finally, the proposed HLFGD (DB+GSB+EDB) combining global-structure, edge-detail, and de-raining branches yields the best performance both in quantity and quality, realizing the effect of removing the rain streaks and restoring the same details as the ground truth as much as possible.

Table 2. Quantitative results of different operations on Rain100H [1]. The best results are marked in bold.

Modules	Different Combinations of Modules			
DB w/o CA	✓	✗	✗	✗
DB	✗	✓	✓	✓
GSB	✗	✗	✓	✓
EDB	✗	✗	✗	✓
PSNR	25.23	28.85	29.52	30.44
SSIM	0.841	0.878	0.893	0.912

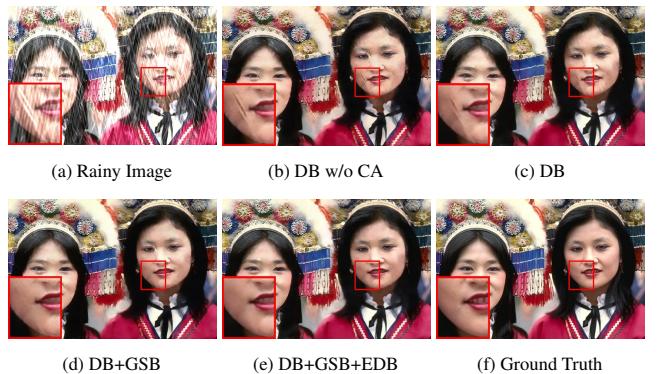


Fig. 5. Qualitative results of different operations on a synthesized rainy image from Rain100H [1].

4. CONCLUSION

In this paper, we propose a High-Low-Frequency Guided De-raining (HLFGD) method to remove the rain streaks and retain the image details simultaneously. The HLFGD consists of three network branches, one is designed to remove the rain streaks, the other two are developed to find back the global structure and edge details of an image.

5. REFERENCES

- [1] W. Yang, R. T. Tan, J. Feng, J. Liu, and S. Yan, “Deep joint rain detection and removal from a single image,” in *2017 IEEE CVPR*, 2017.
- [2] 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 CVPR*, 2017, pp. 3855–3863.
- [3] Xueyang Fu, Jiabin Huang, Xinghao Ding, Yinghao Liao, and John Paisley, “Clearing the skies: A deep network architecture for single-image rain removal,” *IEEE Transactions on Image Processing*, vol. 26, no. 6, pp. 2944–2956, 2017.
- [4] Zhiwen Fan, Huafeng Wu, Xueyang Fu, Yue Huang, and Xinghao Ding, “Residual-guide network for single image deraining,” in *Proceedings of the 26th ACM international conference on Multimedia*, 2018, pp. 1751–1759.
- [5] Ruoteng Li, Loong-Fah Cheong, and Robby T Tan, “Single image deraining using scale-aware multi-stage recurrent network,” *arXiv preprint arXiv:1712.06830*, 2017.
- [6] 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 ECCV*, 2018, pp. 254–269.
- [7] He Zhang and Vishal M Patel, “Density-aware single image de-raining using a multi-stream dense network,” in *Proceedings of the IEEE CVPR*, 2018, pp. 695–704.
- [8] 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 CVPR*, 2019, pp. 3937–3946.
- [9] Youzhao Yang and Hong Lu, “Single image deraining via recurrent hierarchy enhancement network,” in *Proceedings of the 27th ACM International Conference on Multimedia*, 2019, pp. 1814–1822.
- [10] Youzhao Yang and Hong Lu, “Single image deraining via recurrent hierarchy enhancement network,” in *Proceedings of the 27th ACM International Conference on Multimedia*, 2019, pp. 1814–1822.
- [11] 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*, 2020, pp. 14560–14569.
- [12] 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 CVPR*, 2020, pp. 8346–8355.
- [13] Yingjun Du, Jun Xu, Xiantong Zhen, Ming-Ming Cheng, and Ling Shao, “Conditional variational image deraining,” *IEEE Transactions on Image Processing*, vol. 29, pp. 6288–6301, 2020.
- [14] Youzhao Yang and Hong Lu, “A fast and efficient network for single image deraining,” in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021, pp. 2030–2034.
- [15] 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 *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 3435–3444.
- [16] Qibin Hou, Daquan Zhou, and Jiashi Feng, “Coordinate attention for efficient mobile network design,” in *Proceedings of the IEEE/CVF Conference on CVPR*, 2021, pp. 13713–13722.
- [17] 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*, 2016, pp. 770–778.
- [18] Qing Guo, Jingyang Sun, Felix Juefei-Xu, L. Ma, Xiaofei Xie, Wei Feng, and Yang Liu, “Efficientderain: Learning pixel-wise dilation filtering for high-efficiency single-image deraining,” in *AAAI*, 2021.
- [19] Quan Huynh-Thu and Mohammed Ghanbari, “Scope of validity of psnr in image/video quality assessment,” *Electronics letters*, vol. 44, no. 13, pp. 800–801, 2008.
- [20] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [21] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer, “Automatic differentiation in pytorch,” 2017.
- [22] Diederik P Kingma and Jimmy Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.