

Import Vertical Characteristic of Rain Streak for Single Image Deraining

Zhexin Zhang

Hangzhou Dianzi University

Jiajun Ding

Hangzhou Dianzi University

Jun Yu (✉ yujun@hdu.edu.cn)

Hangzhou Dianzi University

Yiming Yuan

Hangzhou Dianzi University

Jianping Fan

Hangzhou Dianzi University

Research Article

Keywords: Single Image Deraining, Vertical Module, Contrastive Learning

Posted Date: April 20th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1501451/v2>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Import Vertical Characteristic of Rain Streak for Single Image Deraining

Zhexin Zhang · Jiajun Ding · Jun Yu · Yiming Yuan · Jianping Fan*

Received: date / Accepted: date

Abstract Recently, deep convolutional neural networks show good effect for single image deraining. These networks always adopt the conventional convolution method to extract features, which may neglect the characteristic of rain streak. A novelty vertical module is proposed to focus on the vertical characteristic of rain streak. Such module uses $1 \times X$ convolution kernel to extract the vertical information of rain streaks and a $X \times X$ convolution kernel to keep relative location information. Use this module in the front of deraining network can better detach rain streaks from background. In addition, the contrastive learning is employed to improve the performance of the model. Extensive experimental results demonstrated the superiority of the deraining methods with the proposed methods in comparison with the base ones.

Keywords Single Image Deraining · Vertical Module · Contrastive Learning ·

Zhexin Zhang
Zhuoyue Honors College, Hangzhou Dianzi University
E-mail: zzx@hdu.edu.cn

Jiajun Ding
Computer & Software School, Hangzhou Dianzi University
E-mail: djj@hdu.edu.cn

Jun Yu
Computer & Software School, Hangzhou Dianzi University
E-mail: yujun@hdu.edu.cn

Yiming Yuan
Zhuoyue Honors College, Hangzhou Dianzi University
E-mail: yimingyuan@hdu.edu.cn

Jianping Fan* (Corresponding Author)
Computer & Software School, Hangzhou Dianzi University
E-mail: fanjianping@hdu.edu.cn

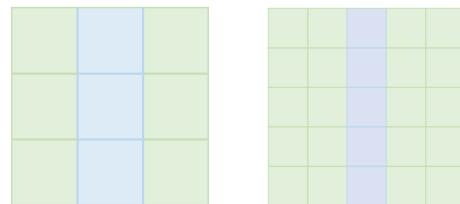


Fig. 1 3×3 and 5×5 Vertical Module

1 Introduction

Rain is a common natural phenomenon. while shooting with this weather, the images captured outdoors under the influence of this weather will have many disadvantages, such as rain streaks in the image will have a certain obscuring and distorting effect on the background scene or the heavy rain obscures the important things.

In image processing, the rain streaks and other noise in the image will produce certain obscuration and distortion of the background scene, making the key information in the image more disturbed, which is not conducive to the post-use of the image.

However, many computer vision applications for outdoor scenes generally require high quality and clear images. Occluded and distorted images due to rain can make these computer vision applications less effective or lose their original functionality. Based on these demands, use computer programs such as deraining[1] and dehazing[2,3] to make images clearer have developed rapidly in recent years. With the development of deep learning, deep learning-based derain network has been widely used. [1] proposed a progressively optimized residual network progressive ResNet (PRN) and progressively optimized recurrent network (PReNet) to complete the image rain removal. [4] proposed bilateral re-

current network further improves the accuracy of deraining network. To generate higher quality derained images, [5] Ding et al. proposed a recurrent distributed feedback network (DFN). The new feedback block is proposed to use high-level information to correct low-level input.

Lots of deraining networks before have not fully exploited the vertically oriented features of rain marks. In order to make reasonable use of this feature to make the network achieve better results, we propose vertical module in this paper in the figure 1. The convolution kernel of vertical module is in the shape of a vertical strip, this allows the convolution to better extract feature information in the vertical direction. Adding vertical convolution before the deraining network can better extract the low-level information of rain streaks.

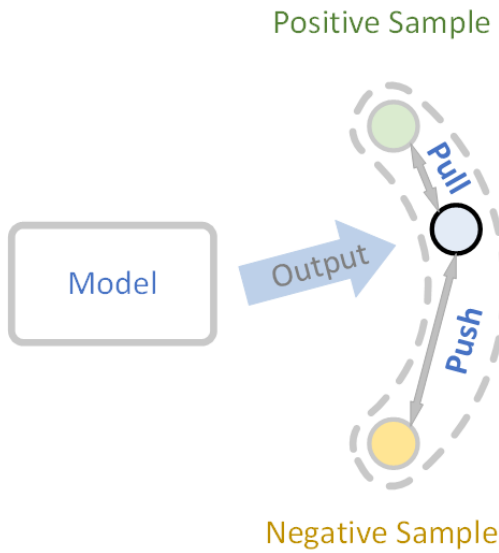


Fig. 2 Schematic diagram of contrastive learning

To gain better performance, proposed frame uses a contrastive learning method in the figure 2. Contrastive learning has been widely used in various deep learning networks and achieved good results [6,7,8]. The proposed method use a norain picture and a raining picture (same as the input of the model) as samples, compare the output of model and the sample picture with the pre-trained VGG to calculate a new loss. Make the training model close to positive sample(norain picture) and far away from Negative sample(raining picture) to gain better training model.

2 Related Work

In this section, we present a briefly over for the approaches and results of deraining algorithms.

These studies can be divided into three parts: deep network-based deraining methods, special convolution and contrastive learning.

2.1 Developing of Deraining Models

A rainy image is consists of a clean image and a rain layer. Linear summation deraining models have been used in video deraing [9, 10, 11, 12, 13]. Inspired by the success of its success in video editing, linear summation model have been developed in combination with proper regularizers on both background image and rain layer. With the help of Gaussian mixture model and other models, linear summation image deraining models have shown significant promise [14]. A further improved model, the screen blend [15,16] model, can remove rain streaks with the use of the discriminative dictionary learning to separate rain streaks.

Deep learning-based models for removing rain marks have been developed after 2017. From Derainnet to DNN [17, 18], a residual neural network based on processing the relationship between the rain and the background layers was proposed by Fu.

Based on previous studies, Ren et al. [1] provided a better and simpler baseline deraining network by improving the network architecture, the input and output, and the loss functions, and repeatedly using ResNet proposed progressive ResNet to take advantage of the recursive computation. Based on bilateral LSTMs, BRN [4] is also proposed later to improve the interplay between rain layers and background layers.

2.2 Special Convolution

Asymmetric convolutions is a type of special convolutions which are compression and acceleration-optimized for computational complexity on the basis of square-kernel convolutional.

Previous studies [19,20] have shown that a typical convolution with a kernel $d \times d$ can be replaced with two convolution with $d \times 1$ and $1 \times d$ kernel to reduce the parameters. This result, however, is based on the fact that the learned kernels' eigenvalues' intrinsic rank is not equal to one, therefore, applying the new kernel instead of the typical one would result in significant information loss [21]. Denton et al.[19] developed an optimization algorithm with a low-rank approximation by using SVD algorithm and successfully reduced losses of this method. Jin et al. [21]Jin et al. used structural constraints to separate 2D kernels and obtained comparable performance as conventional CNN with 2 times speed-up. Asymmetric convolutions also have a good effect on reducing

the amount of parameters.

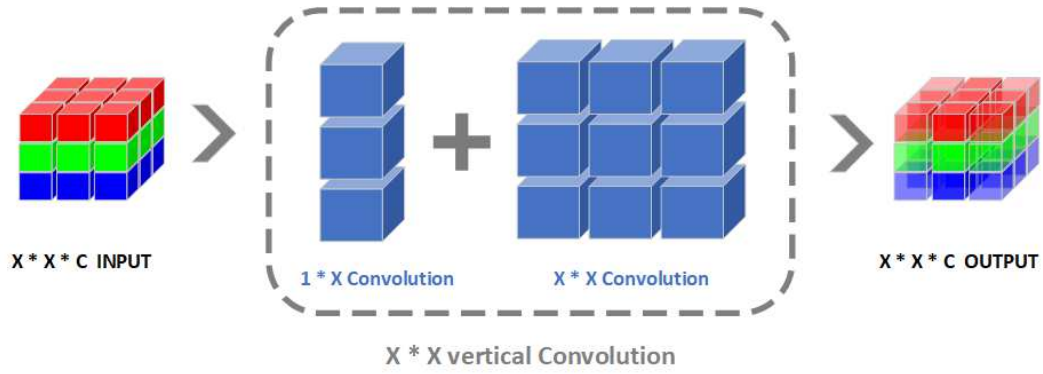


Fig. 3 vertical module

EDANet [22] replaced 3×3 convolution kernel with a similar method and reduced the relevant part by one-third of the parameters computations with minor performance degradation. Ding et al. propose Asymmetric Convolution Blocks(ACNet[23] the method replaced the traditional 3×3 convolution in the existing network with the mix of 1×3 convolution, 3×1 convolution and 3×3 convolution as the output of the convolutional layer.

2.3 Contrastive Learning

Contrastive learning is widely used in self-supervised learning [24,25,26,27,28] and deep learning-based image processing models [6,28,27,26,29]. Contrastive learning uses a positive point and a negative point to judge the output loss of the training model. Contrastive learning makes the output of the model close to the positive result and away from the negative result faster and more accurate by changing the method of calculating the loss value. It has been shown that contrastive learning leads to a better effect in image-to-image translating. Recently, Wu et al. [6] proposed a new sampling method and a novel pixel-wise contrastive loss with contrastive learning and used the method to image dehazing with success. We propose that the similar model would be applicable to deraining network.

3 Deraining Frame with vertical module and contrastive learning

This section describes the composition of the proposed deraining framework in the present study. The proposed method first uses vertical module to make the network better extract features, then use contrastive learning to optimize the training of the model.

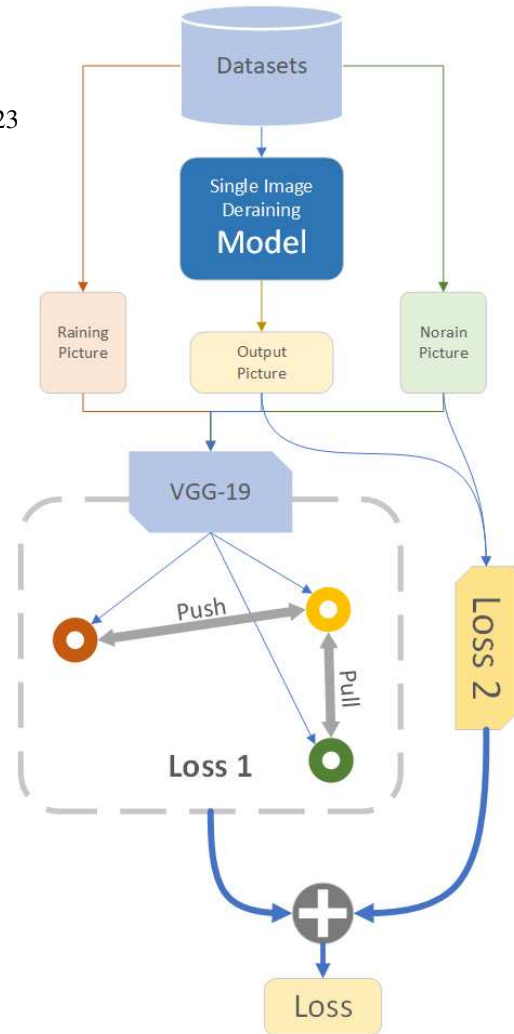


Fig. 4 Contrastive learning for Single Image Deraining Method

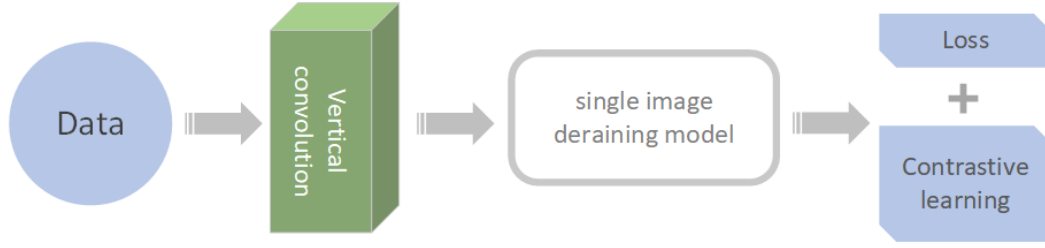


Fig. 5 training model schematic

3.1 Vertical Module for Single Image Derain

3.1.1 Structure of vertical module

Special convolutions provide an efficient way to process and extract information. According to the information to be processed, special convolution can improve the accuracy or efficiency of the convolutional extraction of information or reduce the amount of operations and parameters by changing the way the convolution operates. To improve the effect of single image deraining network, we propose vertical module 3 to extract low-level information of the rain image. Due to gravity, rain streaks are usually in the vertical direction, the proposed method uses a $1 \times X$ strip convolution kernel to extract the vertical features of rain streaks. Simultaneously, in order to preserve the positional relationship between rain streaks and surrounding images, an $X \times X$ convolution kernel is also used to extract relevant spatial position information. The proposed method superimpose the output of each convolution to form the final vertical module to achieve better extraction of the vertical direction features of rain streaks without losing position information.

3.1.2 vertical module with Single image derain Networks

In order to improve the effect of the single image deraining networks, vertical module is proposed to add to the network to improve the effect. Vertical module is more effective in extracting low-level information. Due to the operation of the neural network, the low-level information including rain streaks has been destroyed in the middle and rear parts of the neural network, which is not conducive to the proposed module to detach the rain streaks from background. But in the front section of the neural network, the data contains vertical features of rain marks obviously can make our modules work better. After considering many factors, we found adding vertical module in front of the deraining neural networks to process rain streaks directly can gain better effect in detaching rain streaks from background and remove them. Based on this, the proposed method add the vertical module to the single image deraining neural network to better improving the effect of original single image deraining method.



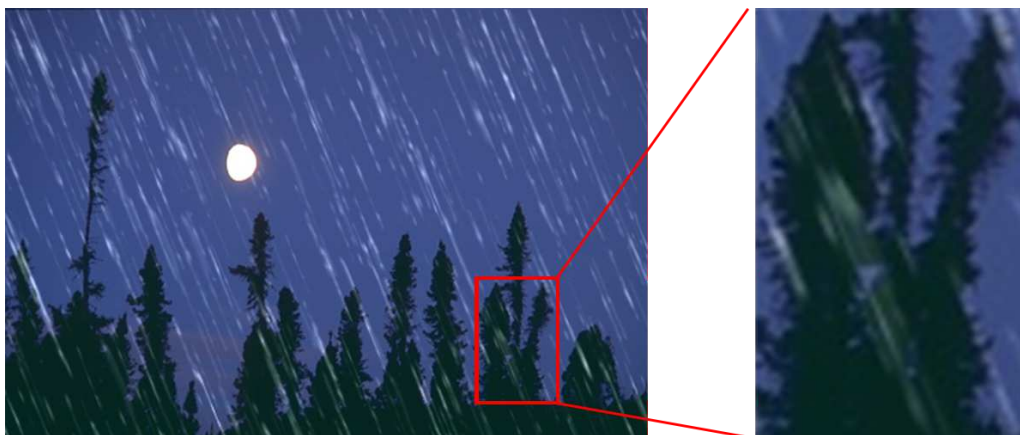
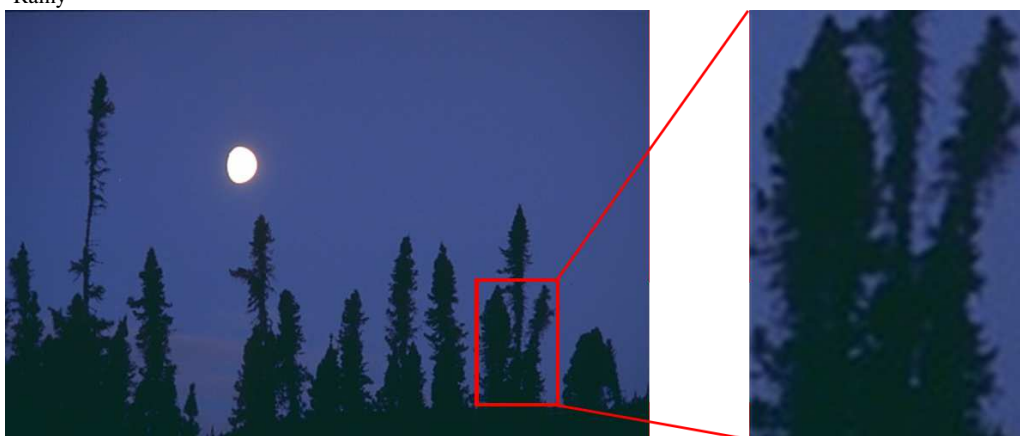
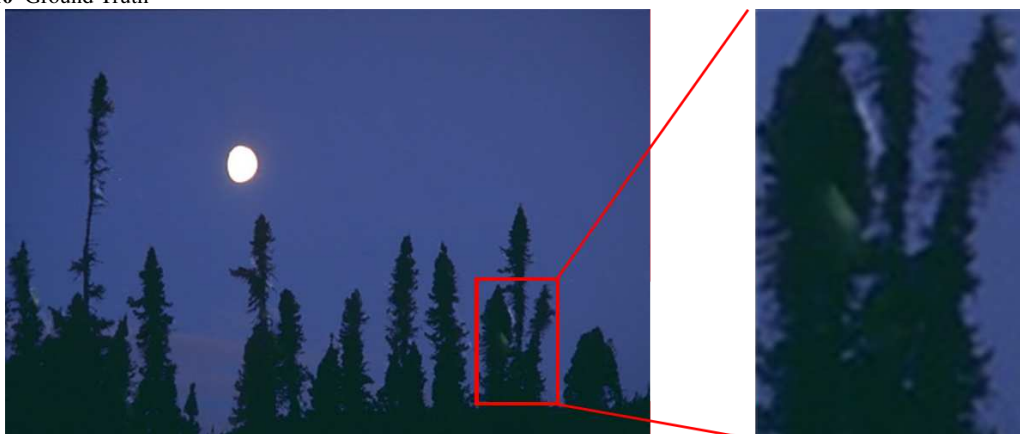
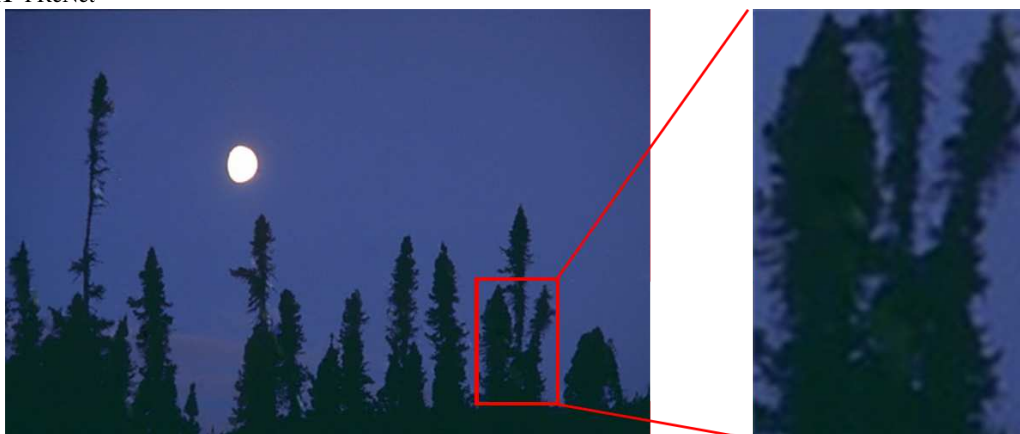
Fig. 6 A Clear Image in Rain100L



Fig. 7 Rainy image 004 in Rain100L



Fig. 8 Rainy image 003 in Rain100L

**Fig. 9** Rainy**Fig. 10** Ground Truth**Fig. 11** PReNet**Fig. 12** PReNet With Proposed Frame

PSNR/SSIM	GMM	DDN	SSIR	PRN	PReNet	PReNet with Proposed frame
Rain100L	28.66/0.865	32.16/0.936	32.37/0.926	36.99/0.977	37.48/0.979	37.88/0.980
Rainlight	-	21.92/0.764	32.20/0.929	-	37.93/0.983	38.33/0.982
Rain12	32.02/0.855	31.78/0.900	34.02/0.935	36.62/0.952	36.66/0.961	37.17/0.962

Table 1 Experiments on Other Datasets

3.2 Contrastive Learning for Deraining Method

Contrastive learning aims to learn a representation to pull the result with positive sample in some metric space and separate apart the representation between negative sample. The application of this method to rain removal involves two parts: to find suitable positive and negative samples, and to find a proper method to determine the difference between the network output and the sample. In this paper, we use common deraining datasets such as Rain100L, Derain datasets usually contains a certain number of rain image such as Fig.7,8 and their corresponding clear image such as Fig.6. In the contrastive learning of the single image deraining neural network, we define clear images as positive sample and raining images as negative samples. Contrastive learning can make the image output by the model closer to the norain image and away from the raining image to achieve the purpose of single image deraining. In order to better extract the characteristics of the image to determine the difference between the network output and the sample, we use the pre-trained VGG-19 [30] to extract the features of the picture to more accurately and efficiently judge the difference between the output picture and sample pictures; See Fig. 5 for a schematic illustration of the proposed training model. In addition, we compared that L1 loss work better than L2 loss[31] and choose L1 loss to cooperate with the proposed module.

4 Experimental Results

In this section, we compare the method of the proposed frame with other methods, including PReNet[1], BRN, et al. In addition, we also compare the proposed framework with non-contrastive learning and other methods with special convolutions, including ACNet[23], dynamic convolution [32], et al.

4.1 Datasets and Training Settings

We conduct experiments on multiple datasets , including Rain100L[33], Rain12[14], Rainlight[34], et al. Datasets such as Rain100L are widely used in rain removal tasks. Rain100L contain 100 pictures of light or heavy rain respectively, simulating various rainy scenes. Fig. 6 shows a norain picture in the dataset, and Figs. 7 and 8 show the rainy versions, respectively.

4.2 Environment and Implementation

We used the pytorch framework to train the proposed model. We refer to the method in Ref. [23] for the optimized implementation of the vertical module. We refer to [6] for the optimized implementation of our proposed contrastive learning and the adjustmend of the scale coefficients of the contrastive learning loss and SSIM loss according to the training situation.

The model was trained with 2 NVIDIA titan Xp graphics card and intel i7-6950X CPU with 128GB RAM. Our dataset processor uses h5py to pack the dataset into groups, and then feed each group of data into the network separately for training, similar to PReNet [1]. The training batch size is set to 18 to run under a suitable computing power, and run each experiment for 200 epochs to make the model achieve a more stable effect. We use 0.001 as the initial learning rate for model training and automatically adjust it multiple times during training.

4.3 Experiments on Rain100L Datasets

Rain100L is a widely used single image deraining datasets with a training sets of 200 images and a test sets of 100 images. The proposed method uses the single image derain network PReNet and the PReNet with our proposed framework to train on the Rain100L training set and test on Rain100L to reflect the role of the proposed single image derain framework.

	PRN	PReNet	PReNet with Proposed Frame
PSNR	36.99	37.48	37.88
SSIM	0.977	0.979	0.98

Table 2 PSNR and SSIM of three method

Table 2 shows that the proposed frame plays a great role in the task of single image deraining. The vertical module in the proposed framework can better extract the features of the vertical direction of the rain streak in the light rain image, resultinh in an improved accuracy in the rain removal. Adding this module to the single-image deraining network can improve the peak signal-to-noise ratio of the output image and improve the image quality.



Fig. 13 Rainy

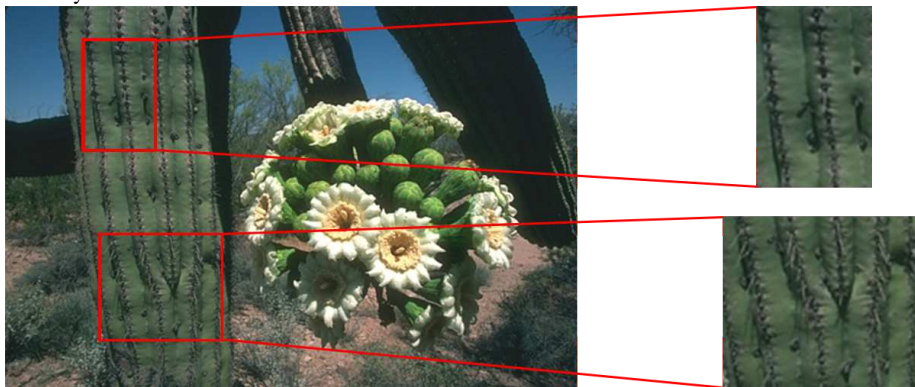


Fig. 14 Ground Truth

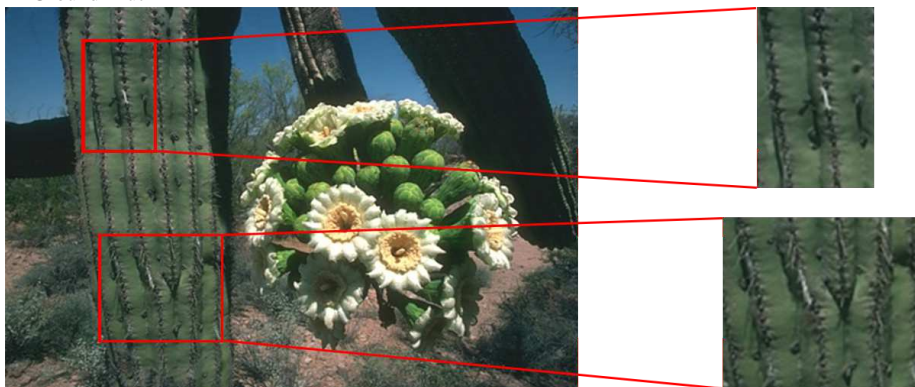


Fig. 15 PReNet

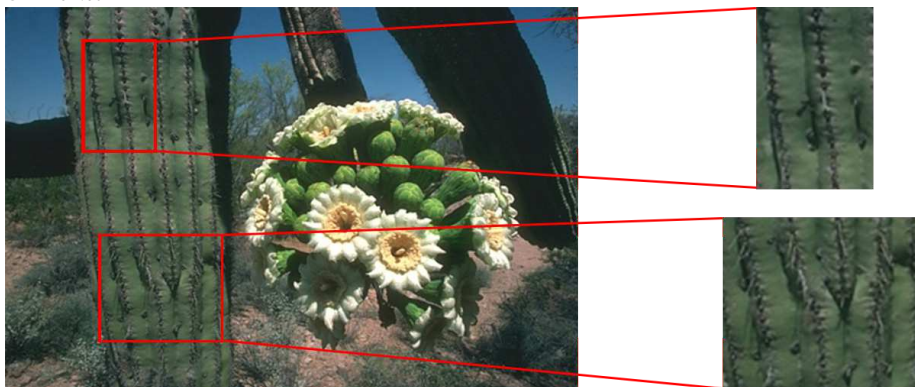


Fig. 16 PReNet With Proposed Frame

We select one of the output pictures from the models trained with dataset Rain100L. In Figs. 11 and 12, comparisons between the outputs of the original model and of the model with our proposed framework suggest that the model with our proposed framework can better recognize and remove rain streaks through vertical module.

4.4 Experiments on Other Datasets

To further test the proposed method, we have also tested the proposed model with (test) datasets RainLight and Rain12. Specifically, we trained the proposed model on PReNet and tested the model on these test datasets and compare our results with PReNet and other deraining networks.

Experimental data in table 1 shows that our proposed framework performs well for the task of removing rain. The frame results in an enhanced Peak Signal to Noise Ratio from 37.48 to 37.87 in dataset Rain100L and from 36.66 to 37.1657 in dataset Rain12.

Experimental data in table 4.3 shows that with the help of the vertical module, the experimental model preserves many vertical and bar-shaped image details. These phenomena reflect the enhancement of rain streaks recognition in the vertical direction by our proposed vertical module, which can better distinguish rain streaks from background.

4.5 Ablation Study

To better understand the contribution of vertical module and contrastive learning to image deraining in network model, we changed the network structure and trained two models to compare the results with the network structure proposed in this study. We trained one of the models with only vertical module and the other with only contrastive learning.

	PSNR	SSIM
PReNet	37.48	0.979
Only vertical module	37.7093	0.9793
Only Contrastive Learning	37.8046	0.9793
proposed frame	37.8763	0.9795

Table 3 Ablation study

From the experiment results in table 3, we find that the vertical module and contrastive learning can indeed lead to improvements in single image deraining, and the proposed framework, which integrates the advantages of both methods, results in further improved results.

5 Conclusion

We have developed a novel framework for single image deraining. A novelty vertical module was proposed in order to detach rain streak from background. The proposed vertical module includes an $1 \times X$ and a $X \times X$ convolution kernel to gain better derain effect when used in front of derain networks. In the meantime, we introduce contrastive learning to make the direction more accurate and the training more effective. Extensive tests on a widely used deraining datasets demonstrate the validity and effectiveness of our approach.

References

1. D. Ren, W. Zuo, Q. Hu, P. Zhu, and D. Meng, "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.
2. H. Khan, B. Xiao, W. Li, and N. Muhammad, "Recent advancement in haze removal approaches," *Multimedia Systems*, pp. 1–24, 2021.
3. W. Ren, L. Zhou, and J. Chen, "Unsupervised single image dehazing with generative adversarial network," *Multimedia Systems*, pp. 1–11, 2022.
4. D. Ren, W. Shang, P. Zhu, Q. Hu, D. Meng, and W. Zuo, "Single image deraining using bilateral recurrent network," *IEEE Transactions on Image Processing*, vol. 29, pp. 6852–6863, 2020.
5. J. Ding, H. Guo, H. Zhou, J. Yu, and B. Jiang, "Distributed feedback network for single-image deraining," *Information Sciences*, 2021.
6. H. Wu, Y. Qu, S. Lin, J. Zhou, R. Qiao, Z. Zhang, Y. Xie, and L. Ma, "Contrastive learning for compact single image dehazing," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10551–10560, 2021.
7. P. Khosla, P. Teterwak, C. Wang, A. Sarna, Y. Tian, P. Isola, A. Maschinot, C. Liu, and D. Krishnan, "Supervised contrastive learning," *Advances in Neural Information Processing Systems*, vol. 33, pp. 18661–18673, 2020.
8. B. Dai and D. Lin, "Contrastive learning for image captioning," *Advances in Neural Information Processing Systems*, vol. 30, 2017.
9. K. Garg and S. K. Nayar, "Detection and removal of rain from videos," in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, vol. 1, pp. I–I, IEEE, 2004.
10. T.-X. Jiang, T.-Z. Huang, X.-L. Zhao, L.-J. Deng, and Y. Wang, "A novel tensor-based video rain streaks removal approach via utilizing discriminatively intrinsic priors," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4057–4066, 2017.
11. T.-X. Jiang, T.-Z. Huang, X.-L. Zhao, L.-J. Deng, and Y. Wang, "Fastderain: A novel video rain streak removal method using directional gradient priors," *IEEE Transactions on Image Processing*, vol. 28, no. 4, pp. 2089–2102, 2018.
12. J.-H. Kim, J.-Y. Sim, and C.-S. Kim, "Video deraining and desnowing using temporal correlation and low-rank matrix completion," *IEEE Transactions on Image Processing*, vol. 24, no. 9, pp. 2658–2670, 2015.
13. M. Li, Q. Xie, Q. Zhao, W. Wei, S. Gu, J. Tao, and D. Meng, "Video rain streak removal by multiscale convolutional sparse coding," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6644–6653, 2018.

14. Y. Li, R. T. Tan, X. Guo, J. Lu, and M. S. Brown, "Rain streak removal using layer priors," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2736–2744, 2016.
15. Y. Luo, Y. Xu, and H. Ji, "Removing rain from a single image via discriminative sparse coding," in *Proceedings of the IEEE international conference on computer vision*, pp. 3397–3405, 2015.
16. D. Ren, W. Zuo, D. Zhang, L. Zhang, and M.-H. Yang, "Simultaneous fidelity and regularization learning for image restoration," *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 1, pp. 284–299, 2019.
17. X. Fu, J. Huang, X. Ding, Y. Liao, and J. 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.
18. X. Fu, J. Huang, D. Zeng, Y. Huang, X. Ding, and J. Paisley, "Removing rain from single images via a deep detail network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3855–3863, 2017.
19. E. L. Denton, W. Zaremba, J. Bruna, Y. LeCun, and R. Fergus, "Exploiting linear structure within convolutional networks for efficient evaluation," *Advances in neural information processing systems*, vol. 27, 2014.
20. M. Jaderberg, A. Vedaldi, and A. Zisserman, "Speeding up convolutional neural networks with low rank expansions," *arXiv preprint arXiv:1405.3866*, 2014.
21. J. Jin, A. Dundar, and E. Culurciello, "Flattened convolutional neural networks for feedforward acceleration," *arXiv preprint arXiv:1412.5474*, 2014.
22. S.-Y. Lo, H.-M. Hang, S.-W. Chan, and J.-J. Lin, "Efficient dense modules of asymmetric convolution for real-time semantic segmentation," in *Proceedings of the ACM Multimedia Asia*, pp. 1–6, 2019.
23. X. Ding, Y. Guo, G. Ding, and J. Han, "Acnet: Strengthening the kernel skeletons for powerful cnn via asymmetric convolution blocks," in *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 1911–1920, 2019.
24. P. Sermanet, C. Lynch, Y. Chebotar, J. Hsu, E. Jang, S. Schaal, S. Levine, and G. Brain, "Time-contrastive networks: Self-supervised learning from video," in *2018 IEEE international conference on robotics and automation (ICRA)*, pp. 1134–1141, IEEE, 2018.
25. Y. Tian, D. Krishnan, and P. Isola, "Contrastive multiview coding," in *European conference on computer vision*, pp. 776–794, Springer, 2020.
26. O. Henaff, "Data-efficient image recognition with contrastive predictive coding," in *International Conference on Machine Learning*, pp. 4182–4192, PMLR, 2020.
27. K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, "Momentum contrast for unsupervised visual representation learning," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 9729–9738, 2020.
28. T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in *International conference on machine learning*, pp. 1597–1607, PMLR, 2020.
29. J.-B. Grill, F. Strub, F. Altché, C. Tallec, P. Richemond, E. Buchatskaya, C. Doersch, B. Avila Pires, Z. Guo, M. Gheshlaghi Azar, *et al.*, "Bootstrap your own latent-a new approach to self-supervised learning," *Advances in Neural Information Processing Systems*, vol. 33, pp. 21271–21284, 2020.
30. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
31. H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss functions for image restoration with neural networks," *IEEE Transactions on computational imaging*, vol. 3, no. 1, pp. 47–57, 2016.
32. Y. Chen, X. Dai, M. Liu, D. Chen, L. Yuan, and Z. Liu, "Dynamic convolution: Attention over convolution kernels," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11030–11039, 2020.
33. W. Yang, R. T. Tan, J. Feng, J. Liu, Z. Guo, and S. Yan, "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.
34. W. Yang, R. T. Tan, J. Feng, Z. Guo, S. Yan, and J. Liu, "Joint rain detection and removal from a single image with contextualized deep networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 42, no. 6, pp. 1377–1393, 2019.