Both Diverse and Realism Matter:
Physical Attribute and Style Alignment for Rainy Image Generation
Supplementary Material

Appendix

In this supplementary material, we present the details of our network design in section A, and demonstrate our generation results based on clear images as well as the manipulation based on real rain in section B. Then in section C, we provide visualization results on both paired SPA data and unpaired FCRealRain data, showing the improvement of our generation for real rain removal. In the end, we further discuss the generalization of deraining methods under different augmentation models in section D.

A. Implementation of Network Architecture

The decomposition module is structured similarly to [11] and consists of 32 Resblocks. For the attribute encoder, we adopt the UNet [6] architecture to regress the location of the rain layer with the same dimensions as the rainy image. Another UNet architecture with global pooling is used to regress the remaining attributes, such as angle, length, and brightness. The style encoder employs four downsample blocks to regress the mean and variance to represent the distribution. We use the PatchGAN architecture [3] for the discriminator networks.

B. More Visualization on Rain Generation

Diverse Generation on Clean Image: We show the visualization results of generated rainy images in Figure. 1 to demonstrate the effectiveness of physical alignment and controllable generation network (PCGNet). With a clean image, we can generate a variety of realistic and diverse results as expected properties by manually specifying certain attribute parameters, such as angle, length, brightness, and others. Meanwhile, since the introduced style alignment, the PCGNet could learn the rain accumulation and rain veil from the real rain reconstruction, which is hard to precisely describe and generate by the hand-crafted synthetic model.

Manipulative Generation on Rainy Image: SPA [8] and FCRealRain [2] are high-quality datasets of real rain images without corresponding attribute labels. Thus, we take the SPA and FCRealRain datasets as the target of real rain respectively and conduct the attribute manipulation on the corresponding rainy images. Figure. 2, 3, 4, 5 show the manipulation results as the length, angle, and brightness of the rain attribute varies. It is noteworthy that PCGNet not only learns to manipulate the generated results with target attributes, but also comprehends the distinctive rain styles of different datasets, thereby reducing the domain gap between the generated results and target rain.

Table 1. The cross dataset validation for generalization of different augmentation model.

Index	Rain1400 $\longrightarrow$ Rain800					SPA $\longrightarrow$ GT-Rain				
	-O	-T	-V	-J	P	-O	-T	-V	-J	P
PSNR	22.4	21.84	21.12	22.45	23.78	20.18	20.09	20.89	20.54	21.57
SSIM	0.7845	0.7854	0.7801	0.8012	0.8325	0.5045	0.5732	0.5819	0.5662	0.6017

C. Promotion of Generated Image for Rain Remval

We provide more qualitative comparisons of deraining performance on SPA [8] in Figure. 6, FCRealRain [2] in Figure. 7 and Figure. 8 with the latest deraining methods: DSC [5], DDN [4], JORDER-E [9], and MPRNet [12].

Regarding the paired rain/clean dataset of SPA, the existing datasets are augmented by PCGNet with modulating attributes, effectively mitigating the bias of rain streaks in the original datasets. Consequently, the deraining methods are able to acquire a more robust feature representation for rain removal, ultimately yielding a more compelling performance on test data. With respect to the unpaired datasets of FCRealRain, the integration of real rain into the training process of PCGNet significantly reduces the domain gap between the generation results and real rain. Empirical results show that the model trained on the generation results of PCGNet exhibits a more compelling performance.

D. Additional Discussing

Generalization of different augmentation models. To better demonstrate the impact of different augmentation methods on the generalization of rain removal models, we trained the JORDER-E method on the Rain1400 and SPA datasets using various augmentation methods such as tradition augmentation (-T), VRGNet [7] (-V), JRGR [10] (-J), PCGNet (-P), and validated on the Rain800 [13] and GT-Rain [1] test sets as shown in Tabel 1. The results indicate that the JRGR method based on realism outperforms the VRGNet based on diversity in terms of generalization. Furthermore, the proposed method which simutalenoues takes realism and diversity into consideration demonstrates superior generalization performance on both synthetic and real rain data, further confirming the effectiveness of the PCGNet.

References

[1] Yunhao Ba, Howard Zhang, Ethan Yang, Akira Suzuki, Arnold Pfahnl, Chethan Chinder Chandrappa, Celso de Melo, Suya You, Stefano Soatto, Alex Wong, and Achuta Kadambi. Not just streaks: Towards ground truth for single image deraining. In *Eur. Conf. Comput. Vis.*, 2022.

[2] Yi Chang, Yun Guo, Yuntong Ye, Changfeng Yu, Lin Zhu, Xile Zhao, Luxin Yan, and Yonghong Tian. Unsupervised deraining: Where asymmetric contrastive learning meets self-similarity. *arXiv preprint arXiv:2211.00837*, 2022.

[3] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 1125–1134, 2017.

[4] Xia Li, Jianlong Wu, Zhouchen Lin, Hong Liu, and Hongbin Zha. Recurrent squeeze-and-excitation context aggregation net for single image deraining. In *Eur. Conf. Comput. Vis.*, pages 254–269, 2018.

[5] Yu Luo, Yong Xu, and Hui Ji. Removing rain from a single image via discriminative sparse coding. In *Int. Conf. Comput. Vis.*, pages 3397–3405, 2015.

[6] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[7] Hong Wang, Zongsheng Yue, Qi Xie, Qian Zhao, Yefeng Zheng, and Deyu Meng. From rain generation to rain removal. *IEEE Conf. Comput. Vis. Pattern Recog.*, 2021.

[8] 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 *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 12270–12279, 2019.

[9] Wenhan Yang, Robby T Tan, Jiashi Feng, Zongming Guo, Shuicheng Yan, and Jiaying Liu. Joint rain detection and removal from a single image with contextualized deep networks. *IEEE Trans. Pattern Anal. Mach. Intell.*, 42(6):1377–1393, 2019.

[10] Yuntong Ye, Yi Chang, Hanyu Zhou, and Luxin Yan. Closing the loop: Joint rain generation and removal via disentangled image translation. *IEEE Conf. Comput. Vis. Pattern Recog.*, 2021.

[11] Changfeng Yu, Yi Chang, Yi Li, Xile Zhao, and Luxin Yan. Unsupervised image deraining: Optimization model driven deep cnn. In *ACM Int. Conf. Multimedia*, pages 2634–2642, 2021.

[12] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Multi-stage progressive image restoration. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 14821–14831, 2021.

[13] He Zhang, Vishwanath Sindagi, and Vishal M Patel. Image de-raining using a conditional generative adversarial network. *IEEE transactions on circuits and systems for video technology*, 30(11):3943–3956, 2019.

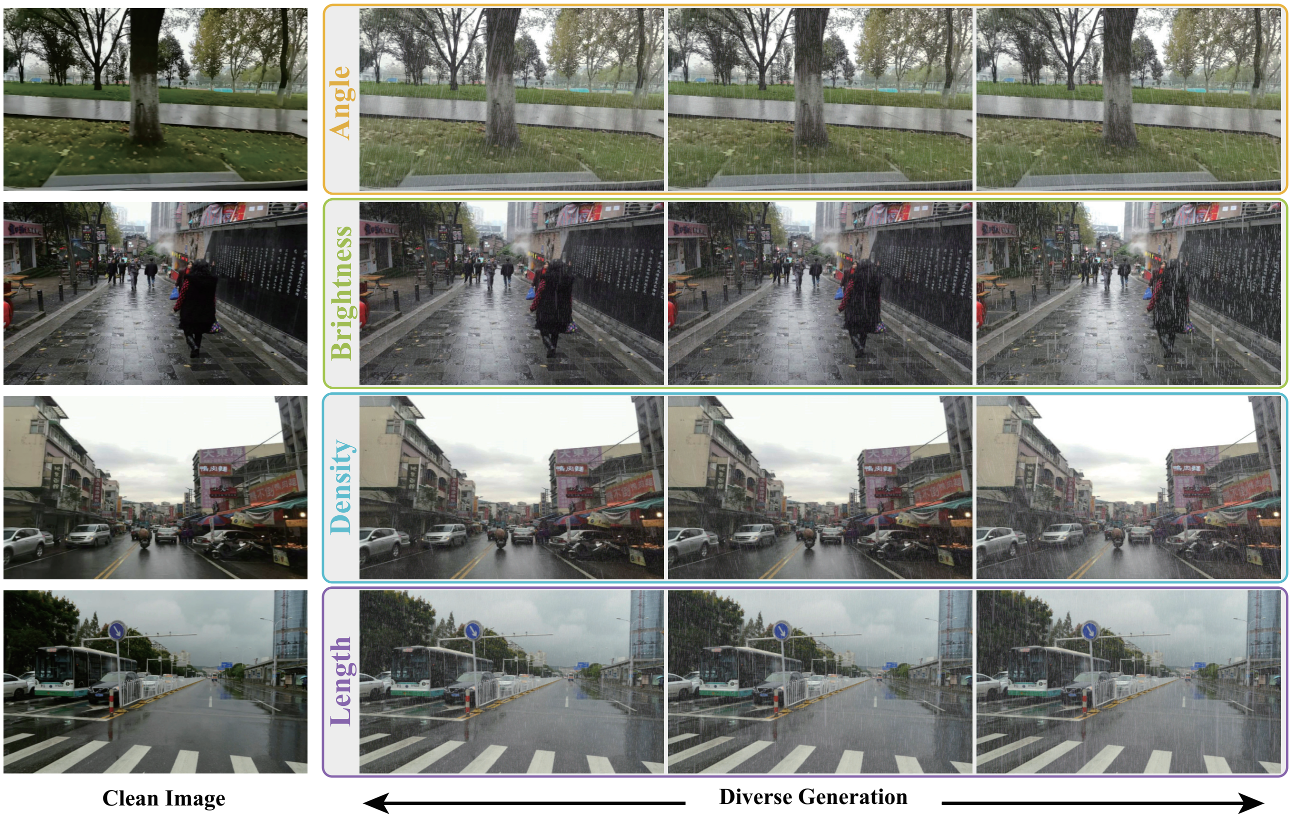

Figure 1. Given a clean image (the first column), we can generate diverse and realistic rainy images (the second to the fourth column) by setting the attributes empirically.

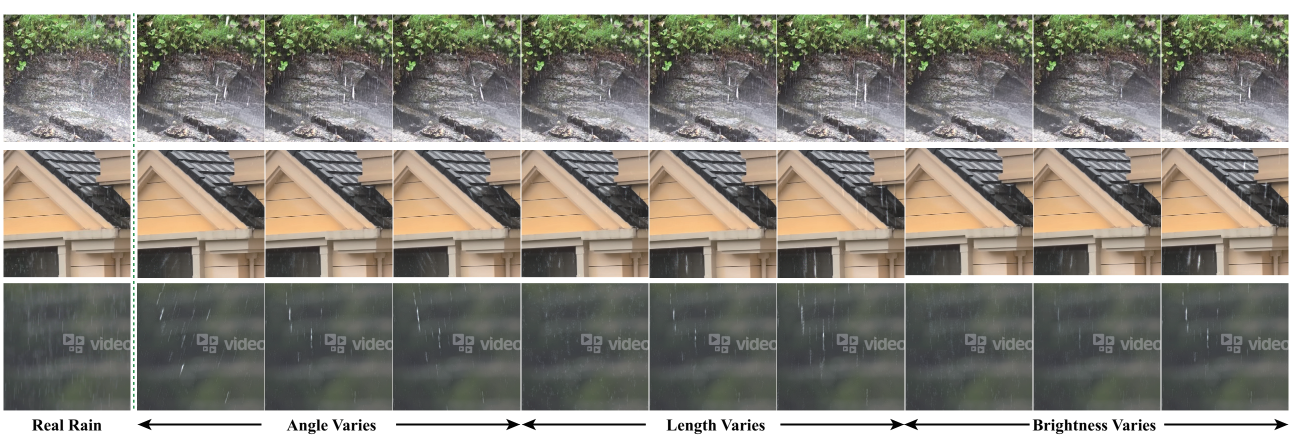

Figure 2. SPA dataset manipulation. Given the real rainy image from SPA (the 1^{st} column), we can precisely estimate the attributes of the rain and manipulate similar generation results following required properties as the angle (the 2^{nd} - 4^{th} column), length (the 5^{th} - 7^{th} column), brightness (the 8^{th} - 10^{th} column) attributes of rain streak vary.

Figure 3. FCRealRain dataset manipulation on angle attribute. Given the real rain image from FCRealRain (the 1^{st} column), we can precisely estimate the attribute of the rain and manipulate the similar generation results following required properties as the angle (the 2^{nd} - 5^{th} column) attribute of rain streak varies.

Figure 4. FCRealRain dataset manipulation on length attribute. Given the real rain image from FCRealRain (the 1^{st} column), we can precisely estimate the attribute of the rain and manipulate the similar generation results following required properties as the length (the 2^{nd} - 5^{th} column) attribute of rain streak varies.

Figure 5. FCRealRain dataset manipulation on brightness attribute. Given the real rain image from FCRealRain (the 1^{st} column), we can precise to estimate the attribute of the rain and manipulate the similar generation results following required properties as the brightness (the 2^{nd} - 5^{th} column) attribute of rain streak varies.

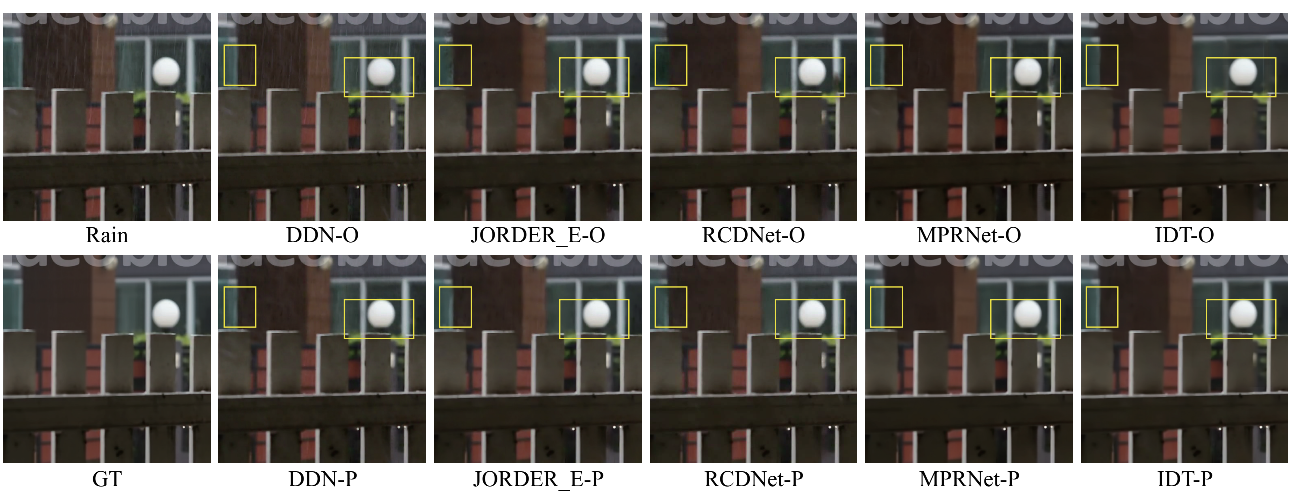

Figure 6. Comparison of deraining results between original and augmented SPA datasets. The first row is the results of deraining methods trained on the original SPA dataset, and the second row is the corresponding results of the model trained on the augmented SPA dataset by PCGNet.

Figure 7. Comparison for real rain removal. The '-O' in the first column are the results of deraining methods with the released model for real rain removal in the paper, and '-P' in the second column are the results of the corresponding model trained on the generation results by PCGNet, which conducts the FCRealRain as the imitative objects in the training.

Figure 8. Comparison for real rain removal. The '-O' in the first column are the results of deraining methods with the released model for real rain removal in the paper, and '-P' in the second column are the results of the corresponding model trained on the generation results by PCGNet, which conducts the FCRealRain as the imitative objects in the training.