Supplementary Material: Self-Aligned Video Deraining with Transmission-Depth Consistency

Wending Yan[1], Robby T. Tan[1,2], Wenhan Yang[3] and Dengxin Dai[4]
[1]National University of Singapore, [2]Yale-NUS College, [3]City University of Hong Kong ,[4]ETH Zurich
eleyanw@nus.edu.sg, robby.tan@{nus,yale-nus}.edu.sg, wyang34@cityu.edu.hk, dai@vision.ee.ethz.ch

1. More Experimental Results

In this supplementary material, we evaluate the performance of the proposed method in comparison with other methods [4, 5, 7, 3, 2, 6, 1] qualitatively, as shown in Figs. 1,2,3.

2. More Ablation Studies

To show the effectiveness of our feature-alignment encoder, we use the standard convolution layers to replace the deformable convolution layers. We extract the feature maps from the central frame and the adjacent frame, and calculate the error map between the central and adjacent feature maps. If the extracted feature maps are not aligned, the final outputs will have an overlapping issue. The feature error maps and derained outputs are shown in the second and third rows of Fig. 4. We observe that the feature error map is significantly reduced after using our feature-alignment encoder. The derained output without alignment is more blurry.

Our method is trained in a semi-supervised way. To show the necessity of our unsupervised losses, we train our method with only supervised losses in this ablation study. In the first row and second column of Fig.4, we show the rain-streak removal outputs with only supervised loss, and compare it with the outputs with the full module in the third column. We observe that most of the rain streaks are successfully removed (although there are few streaks left, see the left edge of the roof), but the output suffers from a strong color shift. This color shifting problem is solved after adding the self-learned consistency loss.

The third row and first column of Fig.4 shows the derained output with only the supervised losses. Like existing fully-supervised methods, this output suffers from the ambiguity between depth and water-droplet density. As a result, the nearby wall in this output is over-saturated, while the faraway trees are still foggy. Having included the depth information in our full module, the ambiguity is reduced and both nearby and faraway objects are better recovered.

3. Training Data

We provide examples of images used to train our networks. Fig. 5 shows the pairs of synthetic rainy images, and their groundtruths for the pairwise supervision. We render the synthetic images using the physical model. Paired synthetic rainstreak-free images and transmission maps for rendering are also shown. Rendered rain fall rates are randomly chosen within the range of 16 to 56 mm/hr. We use Gaussian distribution for the distribution of rain-streaks. Testing results on this synthetic rainy dataset is shown in Fig. 6

References

[1] Hang Dong, Jinshan Pan, Lei Xiang, Zhe Hu, Xinyi Zhang, Fei Wang, and Ming-Hsuan Yang. Multi-scale boosted dehazing network with dense feature fusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2157–2167, 2020. 1

[2] Kui Jiang, Zhongyuan Wang, Peng Yi, Baojin Huang, Yimin Luo, Jiayi Ma, and Junjun Jiang. Multi-scale progressive fusion network for single image deraining. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020. 1

[3] Tai-Xiang Jiang, Ting-Zhu Huang, Xi-Le Zhao, Liang-Jian Deng, and Yao Wang. Fastderain: A novel video rain streak removal method using directional gradient priors. *IEEE Transactions on Image Processing*, 28(4):2089–2102, 2018. 1

[4] Ruoteng Li, Loong-Fah Cheong, and Robby T Tan. Heavy rain image restoration: Integrating physics model and conditional adversarial learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1633–1642, 2019. 1

[5] Wenhan Yang, Jiaying Liu, and Jiashi Feng. Frame-consistent recurrent video deraining with dual-level flow. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1661–1670, 2019. 1

[6] Wenhan Yang, Robby T Tan, Shiqi Wang, and Jiaying Liu. Self-learning video rain streak removal: When cyclic consistency meets temporal correspondence. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1720–1729, 2020. 1

Figure 1: Qualitative comparisons with the state of the art methods on real rain-streak images. Zoom-in for better visualization.

[7] Rajeev Yasarla, Vishwanath A Sindagi, and Vishal M Patel. Syn2real transfer learning for image deraining using gaussian processes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2726–2736, 2020. 1

Figure 2: Qualitative comparisons with the state of the art methods on real rain-streak images. Zoom-in for better visualization.

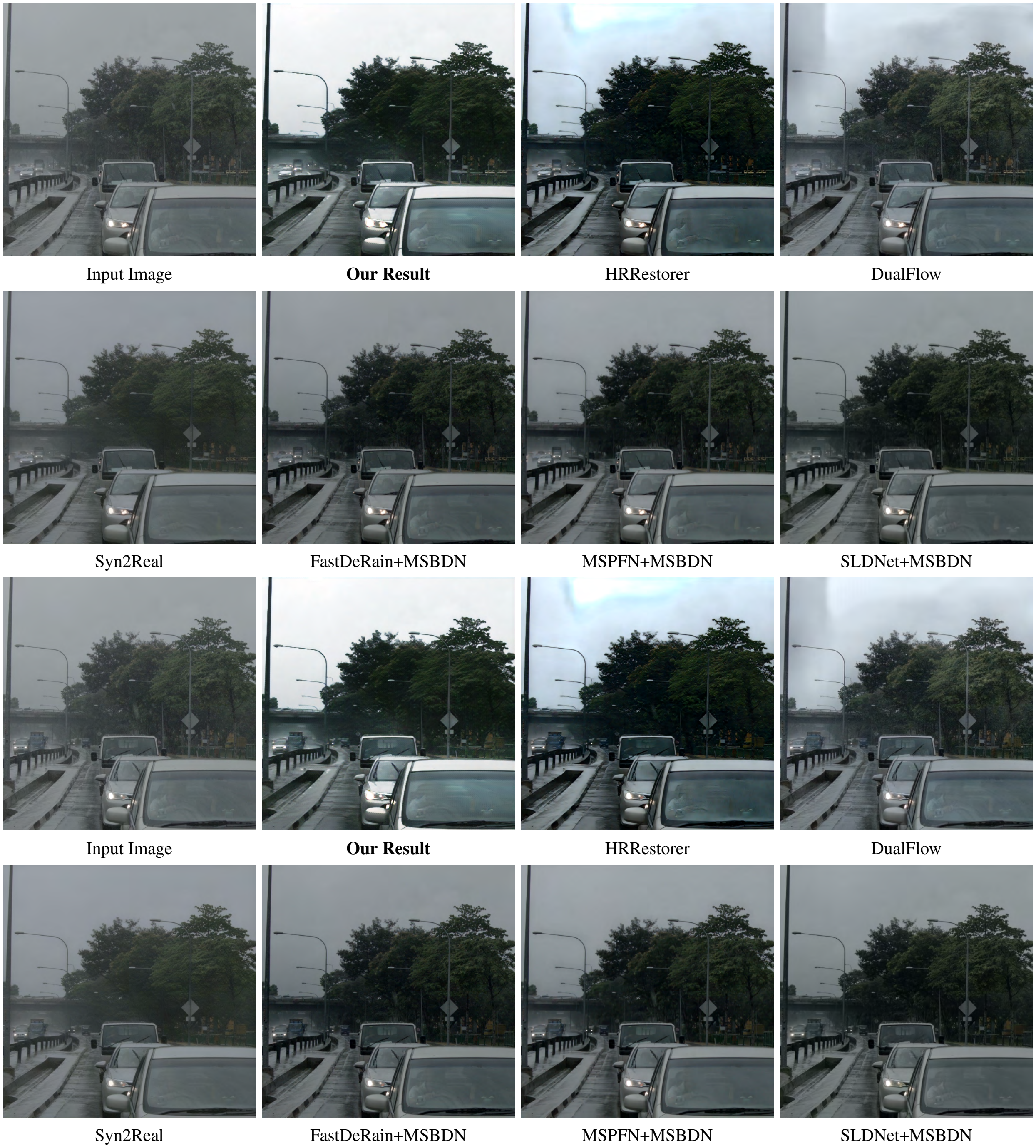


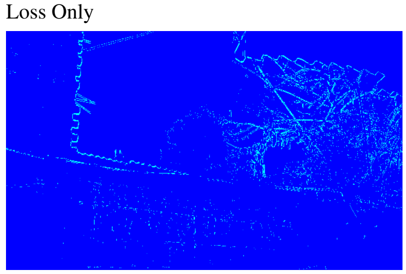
Figure 3: Qualitative comparisons with the state of the art methods on real rainy images. Zoom-in for better visualization.

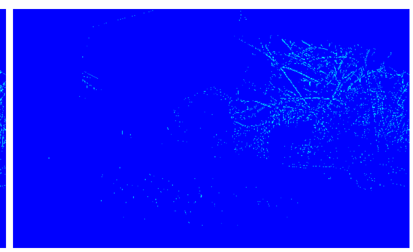

Input Central Frame


Rain-streak Removal Output With Supervised Loss Only


Rain-streak Removal Output With Full Module


Input Adjacent Frame


Feature Map Error Without Feature-Alignemnt


Feature Map Error With Full Module


Final Output With Supervised Loss Only


Final Output Without Feature-Alignemnt


Final Output With Full Module

Figure 4: Ablation studies on our feature-alignment encoder.

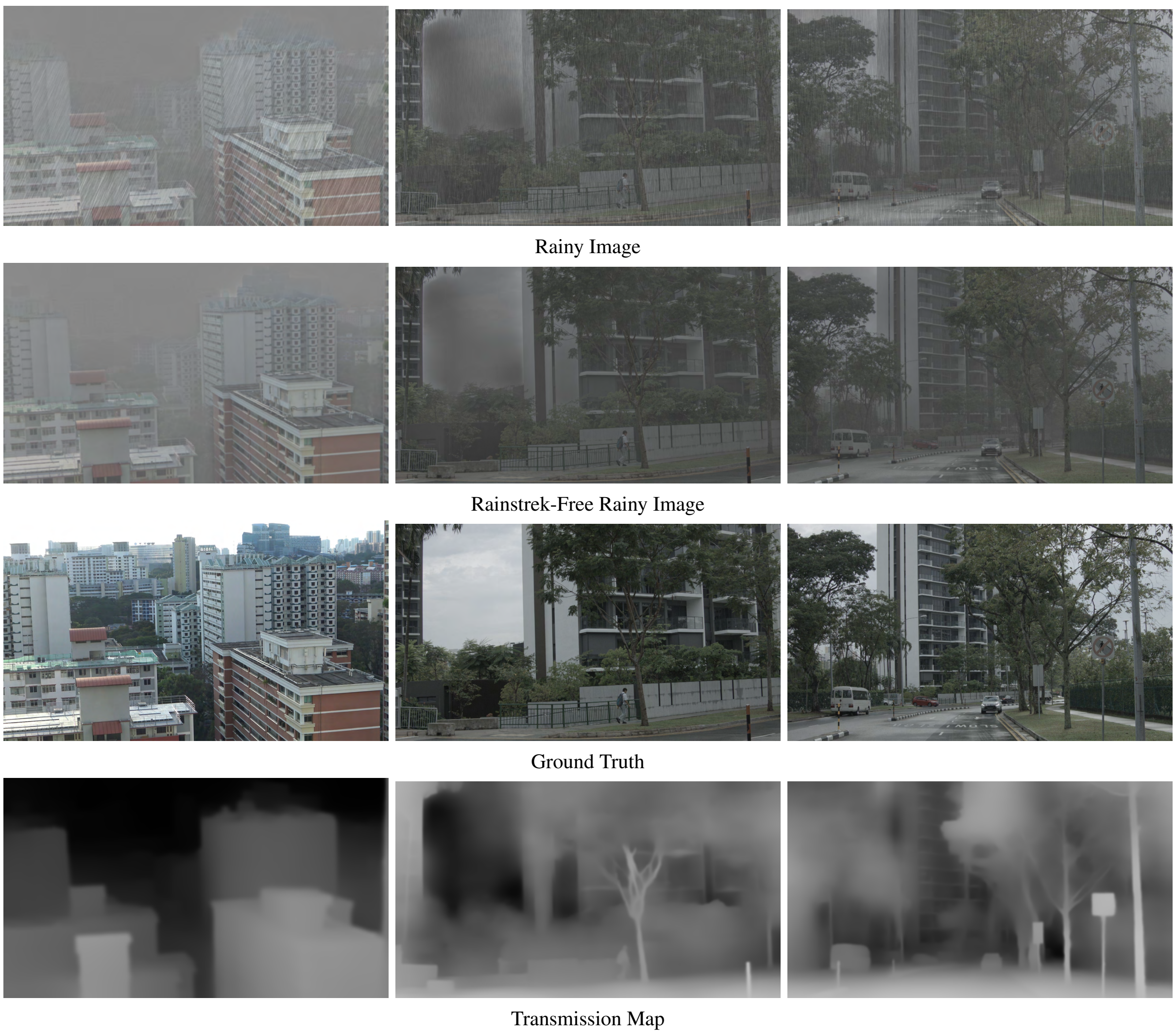


Figure 5: Examples of synthetic rainy image pairs used for supervised loss.

Rainy Image

Ours

Ground Truth

Figure 6: Testing reuslts on our synthetic rainy images.