Deep Joint Rain Detection and Removal from a Single Image (Supplementary Material)

Wenhan Yang[1], Robby T. Tan[2,3], Jiashi Feng[2], Jiaying Liu[1], Zongming Guo[1], and Shuicheng Yan[4,2]
[1]Institute of Computer Science and Technology, Peking University, Beijing, P.R. China
[2] National University of Singapore, [3] Yale-NUS College [4]360 AI Institute

Contents

1. Comparisons Using Real Images

We compare the results of our method with some existing methods: ID [3], CNN [2], DSC [5] and LP [4]. These methods are not designed to handle rain streak accumulation, and thus suffer from it. Hence, in order to compare the results fairly, we use a defogging algorithm [1] to defog their results as post-processing. DSC follows a three-step "streak removal - accumuation removal - streak removal" route. Other methods only takes two steps. These choices are made to achieve best possible results of these methods.

(a) Rain Image

(b) ID

(c) LP

(d) DSC

(e) JORDER

Figure 1. Compared with other methods, JORDER removes most rain streaks.

(a) Rain Images

(b) CNN+DehazeNet

(c) LP+DehazeNet

(d) DSC+DehazeNet

(e) JORDER-R-DEVEIL

Figure 2. JORDER-R-DEVEIL removes most rain streaks.

(a) Rain Image

(b) ID+DehazeNet

(c) LP+DehazeNet

(d) DSC+DehazeNet

(e) JORDER-R-DEVEIL

Figure 3. JORDER-R-DEVEIL removes most rain streaks.

(a) Rain Image

(b) NN+DehazeNet

(c) LP+DehazeNet

(d) DSC+DehazeNet

(e) JORDER-R-DEVEIL

Figure 4. JORDER-R-DEVEIL removes more rain streaks than the existing methods do.

(a) Rain Image

(b) ID+DehazeNet

(c) LP+DehazeNet

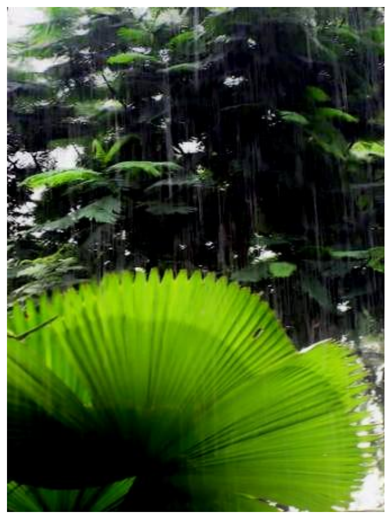
(d) DSC+DehazeNet

(e) JORDER-R-DEVEIL

Figure 5. JORDER-R-DEVEIL removes more rain streaks.

(a) Rain Image

(b) ID+DehazeNet



(c) LP+DehazeNet

(d) DSC+DehazeNet

(e) JORDER-R-DEVEIL

Figure 6. JORDER-R-DEVEIL successfully removes most rain streaks.

(a) Rain Image

(b) ID+DehazeNet

(c) LP+DehazeNet

(d) DSC+DehazeNet

(e) JORDER-R-DEVEIL

Figure 7. JORDER-R-DEVEIL removes most rain streaks.

(a) Rain Images

(b) ID

(c) LP

(d) DSC

(e) JORDER

Figure 8. Compared with other methods, JORDER removes most rain streaks.

(a) Rain Images

(b) ID+DehazeNet

(c) LP+DehazeNet

(d) DSC+DehazeNet

(e) JORDER-R-DEVEIL

Figure 9. JORDER-R-DEVEIL not only removes more rain streaks but also preserves texture details better.

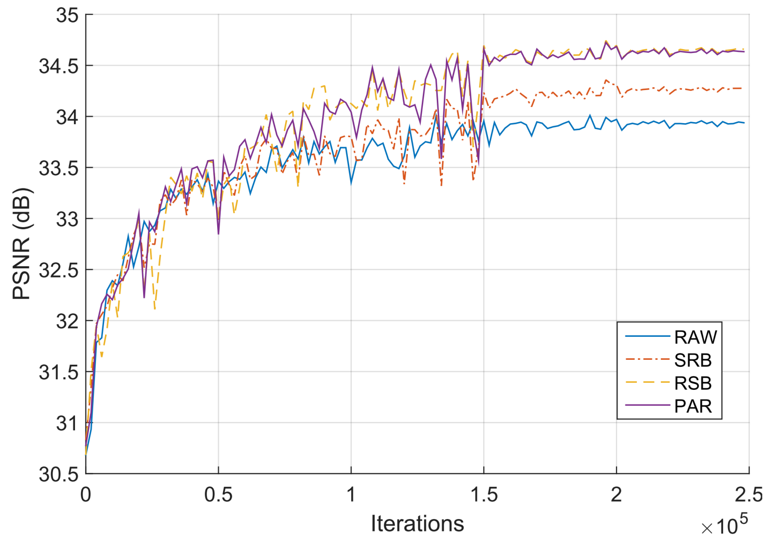


Figure 10. The training performance of the four networks.We drop the learning rate from 0.001 to 0.0001 when reaching 1.5×10^5 iterations and from 0.0001 to 0.00001 when reaching 2×10^5 iterations.

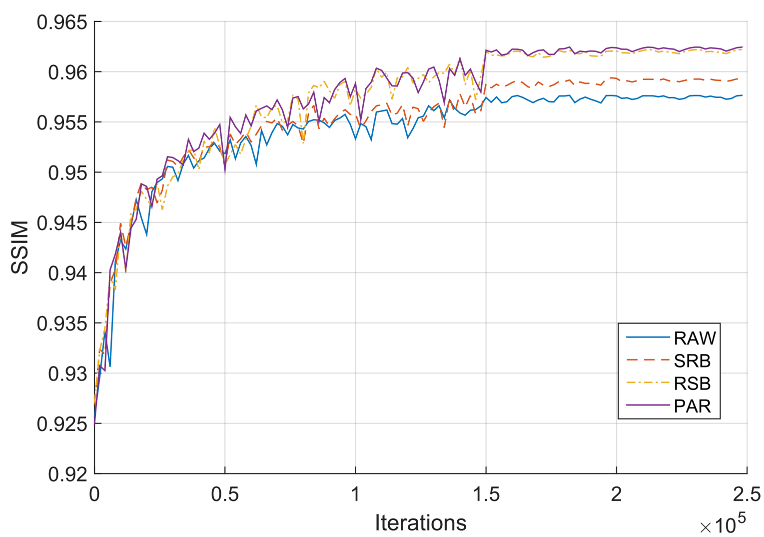


Figure 11. The training performance of the four networks. We drop the learning rate from 0.001 to 0.0001 when reaching 1.5×10^5 iterations.

2. Evaluations of Different Network Architectures

We compare the proposed framework in Figure 4 in the main body of the submission with other potential network structures that also jointly estimate $\mathbf{B}, \mathbf{S}, \mathbf{R}$. From the analysis in Section 4.1 based on the equation

$$\arg\min_{\mathbf{B},\mathbf{S},\mathbf{R}} ||\mathbf{O}-\mathbf{B}-\mathbf{SR}||_2^2 + P_b(\mathbf{B}) + P_s(\mathbf{S}) + P_r(\mathbf{R}), \quad (1)$$

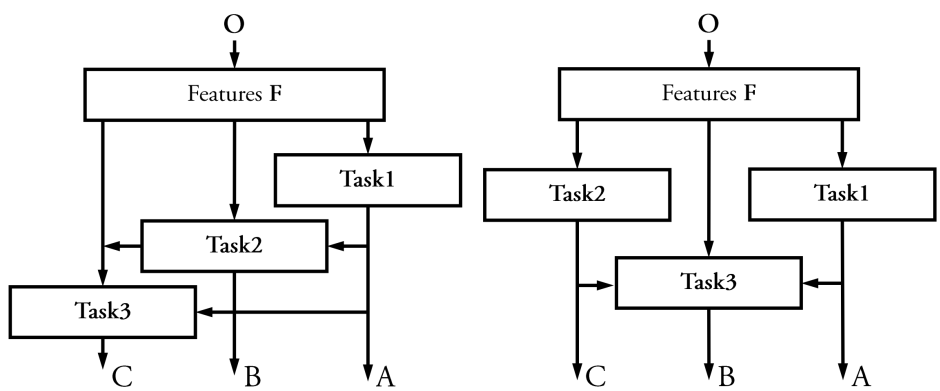
generally we have two choices for the network structures as shown in Figure 12: parallel and sequential. Specifically, considering the prediction order of the variables, there are three candidates:

1. The sequential structure (predicting $\mathbf{R}$, $\mathbf{S}$ and $\mathbf{B}$ in order), denoted as **RSB**, which is our final proposed architecture,

2. The parallel structure (predicting $\mathbf{S}$ and $\mathbf{R}$ based on $\mathbf{F}$), denoted as **PAR**.

3. The sequential structure (predicting $\mathbf{S}$, $\mathbf{R}$ and $\mathbf{B}$ in order), denoted as **SRB**.

Besides, we also compare with a vanilla version: 4) a raw four-layer ResNet with only one recurrence to directly predict the background image, denoted as **RAW**, and Note that, all the experiments here do not exploit the contextualized dilated convolution. The comparison of the potential structures is based on the basic model of a six-layer recurrent ResNet [6].



(a) Sequential Structure (b) Parrallel Structure

Figure 12. Potential choices for network structures.

We compare the training performance and objective quality of these five versions on Rain20L with PSNR and SSIM as the evaluation metrics as shown in Figure 10 and 11 as well as in Table 1. The experimental results clearly show the superiority of $\mathbf{PAR}$ and $\mathbf{RSB}$.

Table 1. PSNR and SSIM results of the five versions.

Version	**RSB**	**PAR**	**SRB**	**RAW**
PSNR	34.66	34.63	34.28	33.94
SSIM	0.9622	0.9624	0.9593	0.9576

3. Analysis for Contextualized Dilated Convolution

We look into the benefit of contextualized dilated convolution to the final performance. Three coupled versions are involved in the comparisons: 1) **PARD**, boosted **PAR** with contextualized dilated convolution; 2) **RSBD**, boosted **RSB** with contextualized dilated convolution. 3) **JORDER-** (10-layer **RSB**), the version of **JORDER** (10-layer **RSBD**) without contextualized dilated convolution. The training performance is also showed in Figs. 13 and 14. The comparison for objective quality is shown in Table 2. The experimental results clearly demonstrate the positive effect of the contextualized dilated convolution on the final objective performance. These results also reach an additional valuable conclusion, that **RSBD** structure is better at acquiring contextual information than **PARD**. This is the reason why our final network employs a **RSBD** structure.

Table 2. Objective evaluation for the effect of contextualized dilated convolution.

Metric	PAR	PARD	RSB	RSBD	JORDER-	JORDER	JORDER-	JORDER	JORDER-	JORDER
Dataset	Rain20L		Rain20L		Rain12		Rain100L		Rain100H	
PSNR	34.63	35.06	34.66	35.16	35.86	36.02	35.41	36.11	20.79	22.15
SSIM	0.9624	0.9655	0.9622	0.966	0.9412	0.9634	0.9632	0.9711	0.5978	0.6736

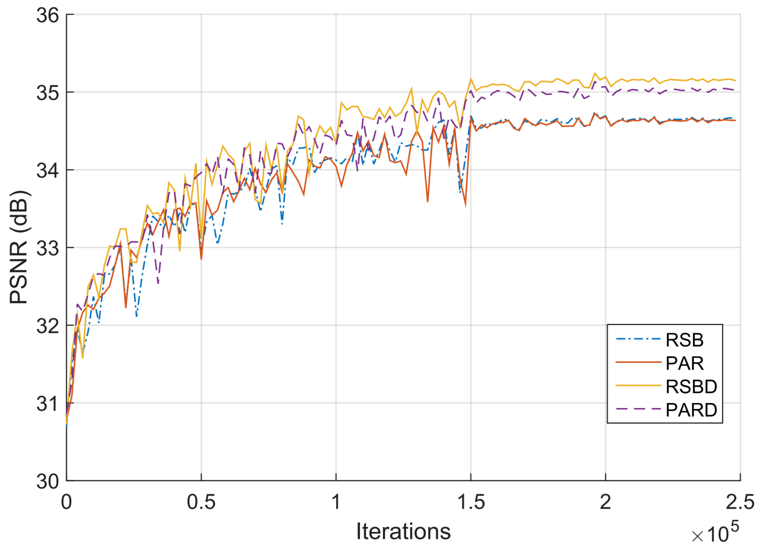


Figure 13. The training performance with and without contextualized convolutions. We drop the learning rate from 0.001 to 0.0001 when reaching 1.5×10^5 iterations and from 0.0001 to 0.00001 when reaching 2×10^5 iterations.

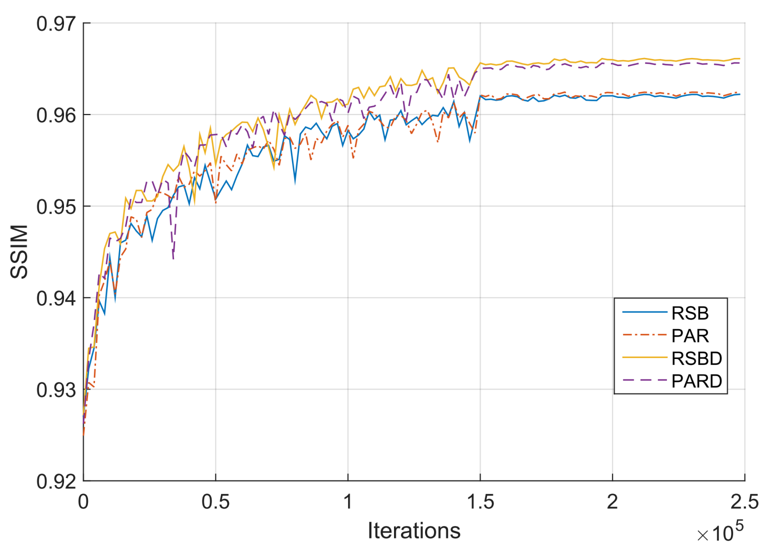


Figure 14. The training performance with and without contextualized convolutions. We drop the learning rate from 0.001 to 0.0001 when reaching 1.5×10^5 iterations.

4. Evaluations of Different Orders of Deraining (JOERER) and Deveiling

We also evaluate the order of joint accumulation removal and streak removal as shown in Figs 15-16. From the results, it is observed the order – streak removal - accumulation removal - streak removal is generally good. The reason is that some obvious rain streaks, noises and artifacts are removed in the first round deraining. Then, the accumulation removal cleans up the rain accumulation, enhances the contrast and visibility, and at the same time boosts weak rain streaks. The subsequent deraining removes these boosted rain streaks, as well as artifacts caused by the accumulation removal, making the results cleaner.

(a) Rain Images

(b) Deveil-Derain

(c) Derain

(d) Derain-Deveil

(e) Derain-Derain

(f) Derain-Deveil-Derain

Figure 15. Order evaluation for JORDER-R-DEVEIL.

(a) Rain Images

(b) Deveil-Derain

(c) Derain

(d) Derain-Deveil

(e) Derain-Derain

(f) Derain-Deveil-Derain

Figure 16. Order evaluation for JORDER-R-DEVEIL.

References

[1] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao. Dehazenet: An end-to-end system for single image haze removal. *IEEE Trans. on Image Processing*, PP(99):1–1, 2016.

[2] D. Eigen, D. Krishnan, and R. Fergus. Restoring an image taken through a window covered with dirt or rain. In *Proc. IEEE Int'l Conf. Computer Vision*, December 2013.

[3] L. W. Kang, C. W. Lin, and Y. H. Fu. Automatic single-image-based rain streaks removal via image decomposition. *IEEE Trans. on Image Processing*, 21(4):1742–1755, April 2012.

[4] Y. Li, R. T. Tan, X. Guo, J. Lu, and M. S. Brown. Rain streak removal using layer priors. In *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition*, pages 2736–2744, 2016.

[5] Y. Luo, Y. Xu, and H. Ji. Removing rain from a single image via discriminative sparse coding. In *Proc. IEEE Int'l Conf. Computer Vision*, pages 3397–3405, 2015.

[6] W. Yang, J. Feng, J. Yang, F. Zhao, J. Liu, Z. Guo, and S. Yan. Deep Edge Guided Recurrent Residual Learning for Image Super-Resolution. *ArXiv*, April 2016.