# Rebuttal Supplementary Evaluations and Results

Anonymous CVPR submission

Paper ID 3970

Dear AC and reviewers, We show additional evaluations and results in this document. Sec. 1 shows the rain removal output frames' PSNR/SSIM comparison with the two additional video based rain removal methods, i.e., *VST-ICCV17'* [4] and *TCL-TIP15'* [2]. Sec. 2 shows additional visual comparisons.

## 1. Derain Frame PSNR/SSIM Comparison

Table. 1 shows additional PSNR/SSIM evaluations for two more video based derain methods, i.e., *VST-ICCV17'* [4] and *TCL-TIP15'* [2]. The best performance for each video sequence has been highlighted in red, and the second best in blue. It can be seen from the table, the proposed *SPAC-CNN* algorithm's advantage holds among all additional competing methods. We also copy the results of *DDN-CVPR17'* [1] and *VMD-CVPR17'* from the manuscript for your reference.

Table 1: Rain removal performance comparison between different methods in terms of scene reconstruction PSNR/SSIM.

| *Camera Motion* | *Clip No.* | Rain | | *DDN-CVPR17* [1] | | *VMD-CVPR17* [3] | | *VST-ICCV17'* [4] | | *TCL-TIP15'* [2] | | *SPAC-CNN* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | – | | Image-Based | | Video-Based | | Video-Based | | Video-Based | | Video-Based | |
| | | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM |
| *panning unstable camera* | *a*1 | 28.46 | 0.94 | 28.02 | 0.95 | 26.96 | 0.92 | 26.14 | 0.94 | **29.87** | **0.96** | **29.78** | **0.97** |
| | *a*2 | 28.09 | 0.95 | 27.38 | **0.95** | 24.80 | 0.93 | 24.03 | 0.83 | **29.01** | **0.96** | **30.09** | **0.96** |
| | *a*3 | 27.84 | 0.93 | 27.41 | 0.94 | 26.45 | 0.90 | 20.50 | 0.70 | **28.82** | **0.95** | **29.75** | **0.96** |
| | *a*4 | 31.48 | 0.95 | 32.47 | **0.97** | 29.55 | 0.94 | 33.41 | 0.96 | **34.12** | **0.98** | **34.82** | **0.98** |
| | avg. *a* | 28.97 | 0.94 | 28.82 | 0.95 | 26.94 | 0.92 | 26.02 | 0.86 | **30.46** | **0.96** | **31.11** | **0.97** |
| *camera speed 20-30 km/h* | *b*1 | 28.72 | 0.92 | **29.48** | **0.96** | 24.09 | 0.84 | 22.25 | 0.76 | 28.07 | **0.94** | **31.19** | **0.96** |
| | *b*2 | 29.49 | 0.90 | 30.23 | 0.95 | 25.81 | 0.89 | 25.13 | 0.79 | **32.41** | **0.97** | **34.05** | **0.98** |
| | *b*3 | 31.04 | 0.95 | **31.39** | **0.97** | 26.12 | 0.90 | 22.08 | 0.84 | 28.29 | 0.94 | **33.73** | **0.98** |
| | *b*4 | 27.99 | 0.92 | 29.83 | **0.96** | 25.90 | 0.88 | 25.63 | 0.80 | **30.38** | 0.95 | **33.79** | **0.97** |
| | avg. *b* | 29.31 | 0.92 | **30.23** | **0.96** | 25.48 | 0.88 | 23.77 | 0.80 | 29.79 | 0.95 | **33.19** | **0.97** |

In the table, the results are calculated as average of all frames in each video sequence. Video-based methods generally perform better for Group *a* data with panning unstable camera. When camera motion becomes larger, video-based methods face greater challenge than image-based methods, which explains why *TCL-TIP15'* performs better for Group *a* whilst worse for Group *b* as compared with *DDN-CVPR17'*. However, our proposed *SPAC-CNN* proves to be able to produce best results among all competing methods for both Group *a* and Group *b* data.

## 2. Additional Visual Comparison

We show selected frames of each test video to highlight the performance of our method. Comparisons are carried out between *DDN-CVPR17'* [1] (single image-based), *TCL-TIP15'* [2] (video-based), *VST-ICCV17'* [4] (video-based), *VMD-CVPR17'* [3] (video-based), and the proposed *SPAC-CNN*.

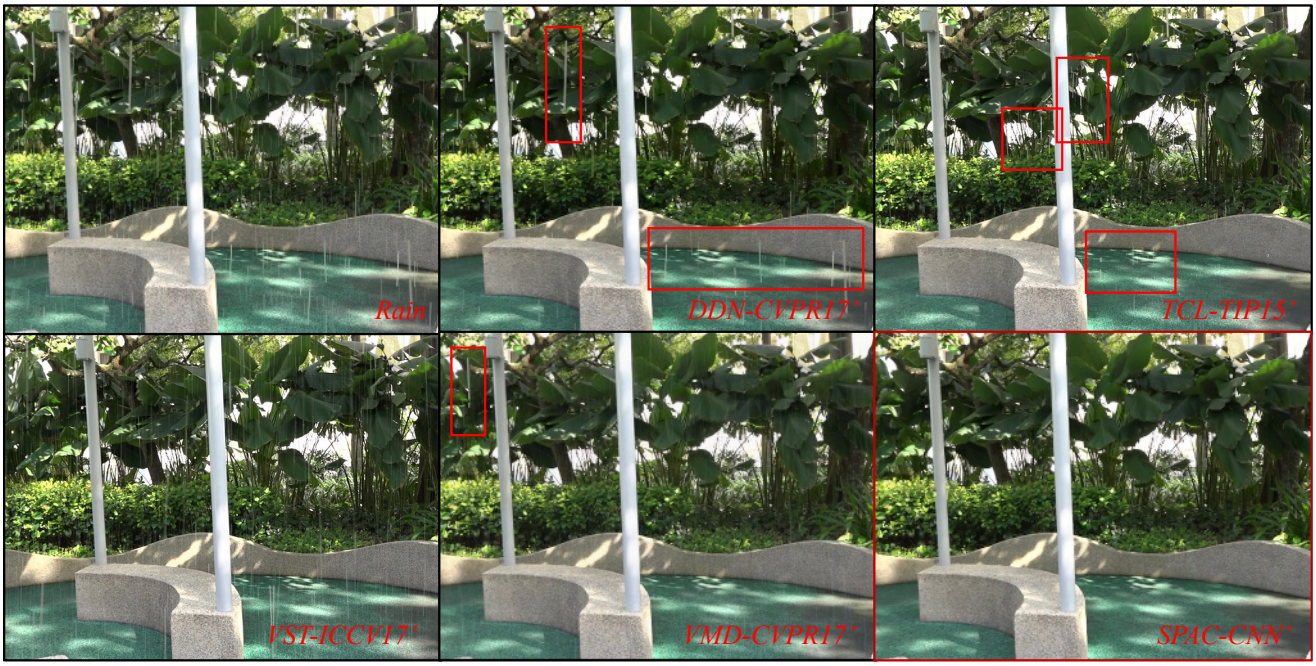


Figure 1: Derain output comparison for testing data *a*1: synthetic rain, panning unstable camera.

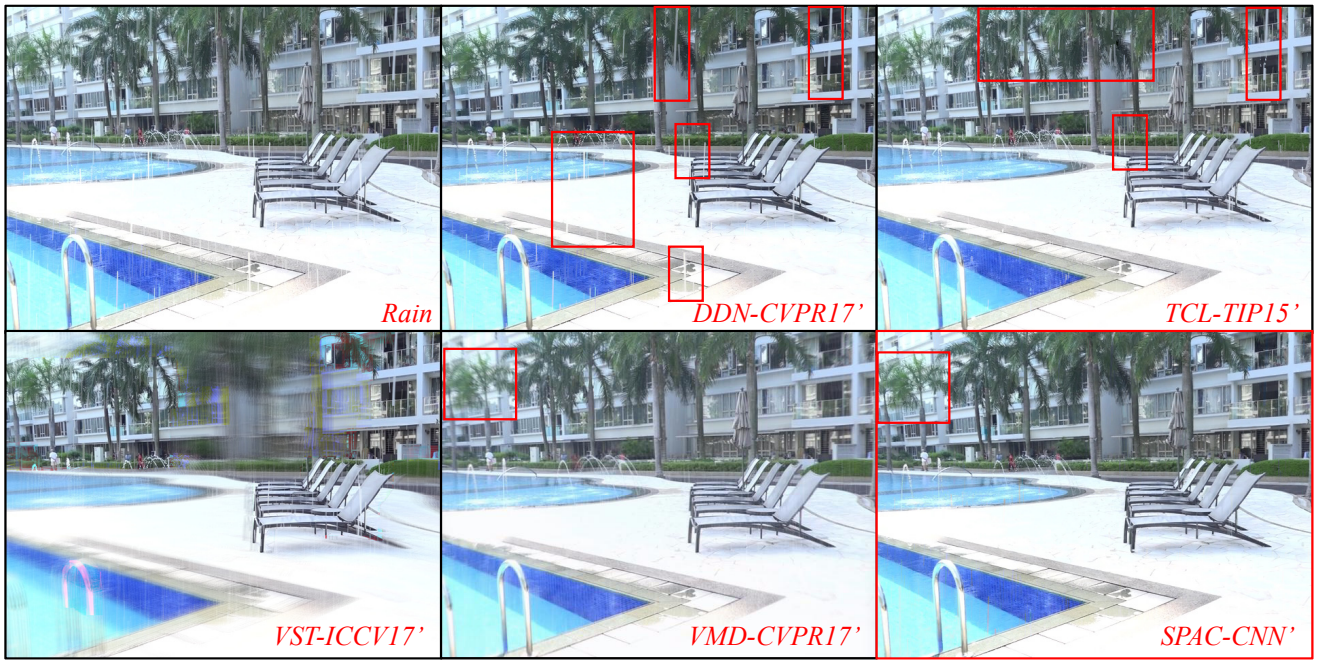


Figure 2: Derain output comparison for testing data *a*2: synthetic rain, panning unstable camera.

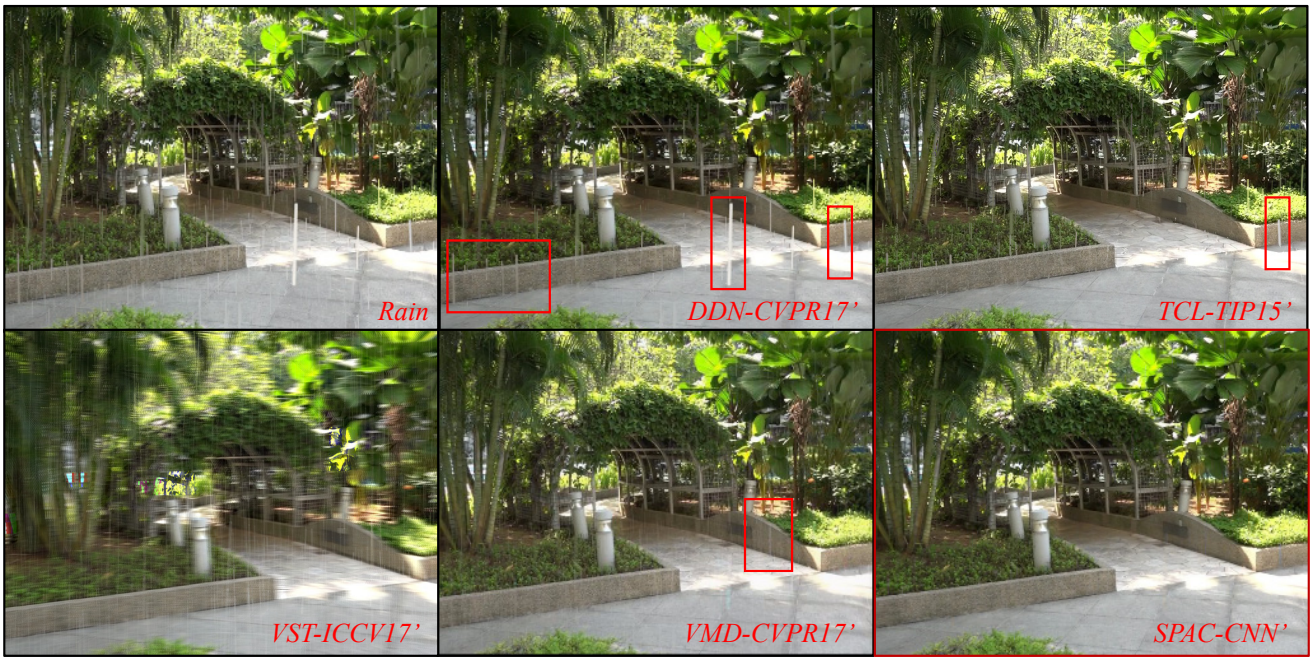


Figure 3: Derain output comparison for testing data *a*3: synthetic rain, panning unstable camera.

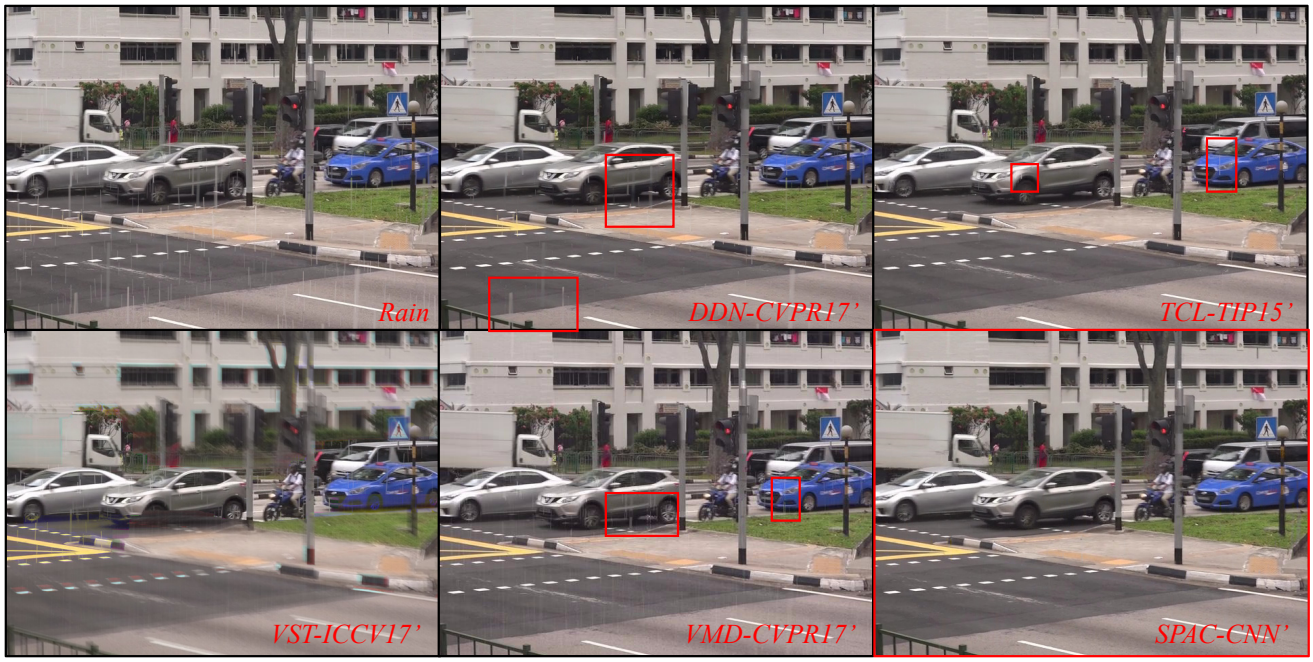


Figure 4: Derain output comparison for testing data *a*4: synthetic rain, stable camera, very dynamic scene.


Figure 5: Derain output comparison for testing data *b*1: synthetic rain, fast moving camera.

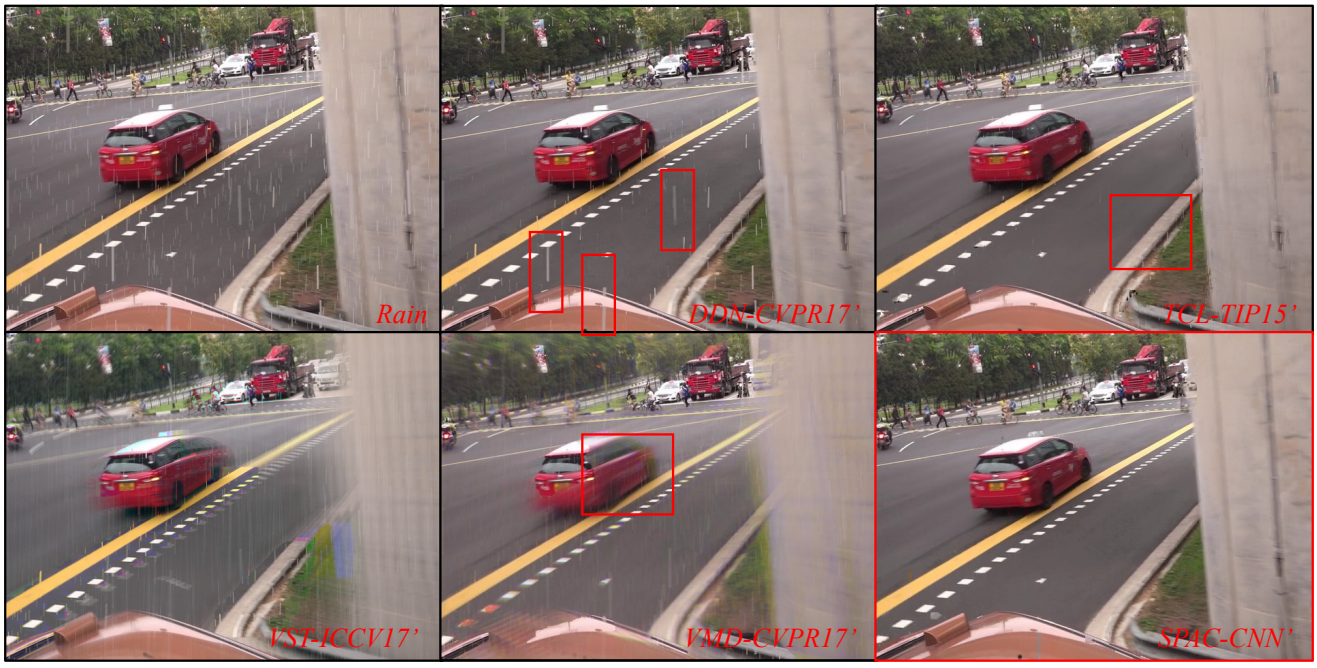


Figure 6: Derain output comparison for testing data *b*2: synthetic rain, fast moving camera.

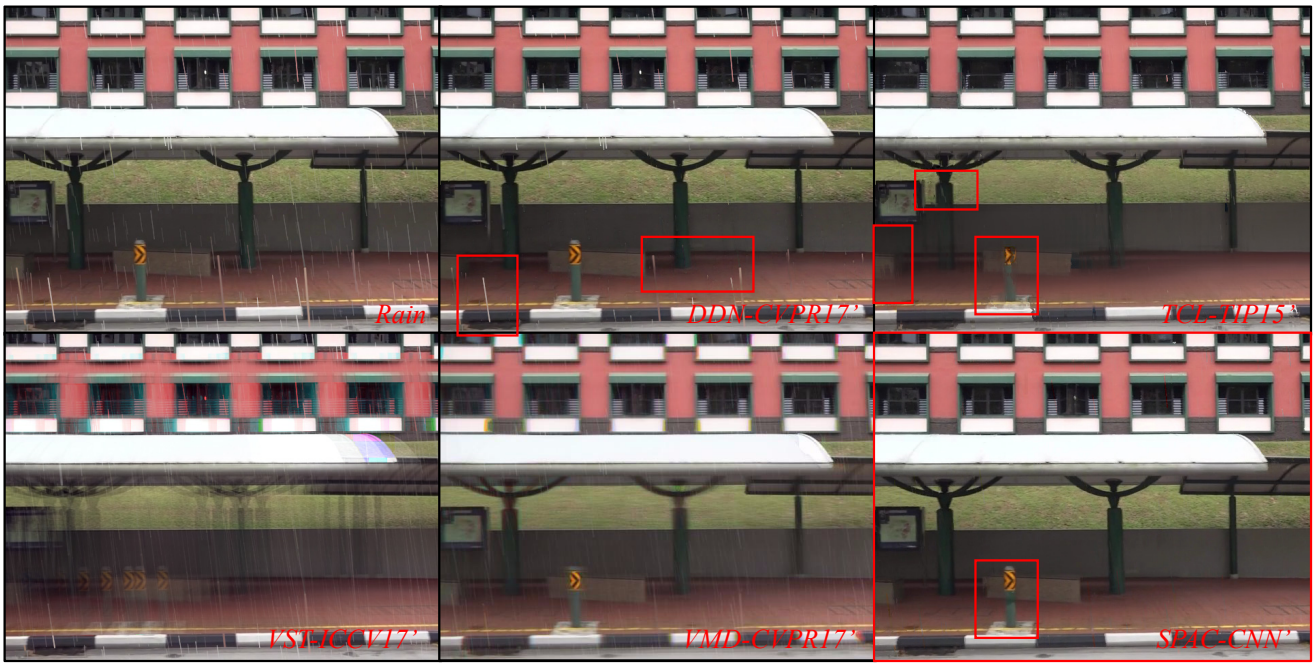


Figure 7: Derain output comparison for testing data *b*3: synthetic rain, fast moving camera.

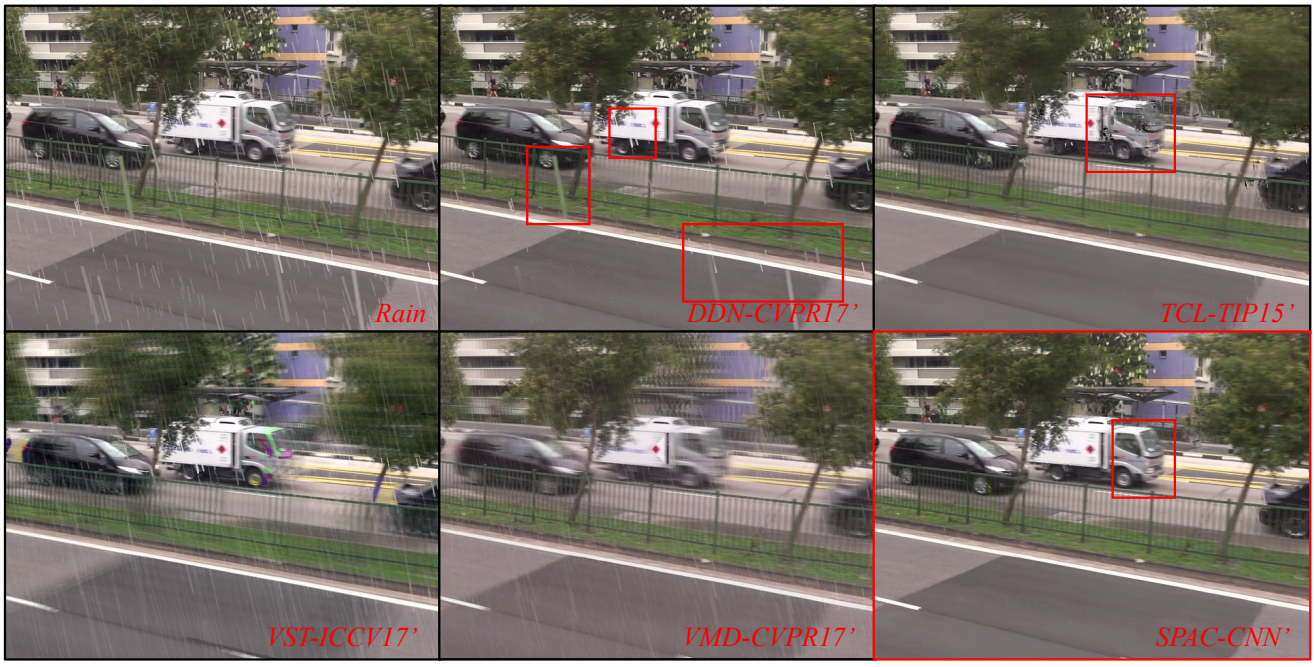


Figure 8: Derain output comparison for testing data *b*4: synthetic rain, fast moving camera.

## References

[1] X. Fu, J. Huang, D. Z. Y. Huang, X. Ding, and J. Paisley. Removing rain from single images via a deep detail network. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2017. 1, 2

[2] 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*, 24(9):2658–2670, Sept 2015. 1, 2

[3] W. Ren, J. Tian, Z. Han, A. Chan, and Y. Tang. Video desnowing and deraining based on matrix decomposition. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 4210–4219, 2017. 1, 2

[4] W. Wei, L. Yi, Q. Xie, Q. Zhao, D. Meng, and Z. Xu. Should we encode rain streaks in video as deterministic or stochastic? In *International Conference on Computer Vision*, pages 2535–2544, Oct 2017. 1, 2