Genuine Knowledge from Practice:
Diffusion Test-Time Adaptation for Video Adverse Weather Removal

Supplementary Material

This is a supplementary material for Genuine Knowledge from Practice: Diffusion Test-Time Adaptation for Video Adverse Weather Removal.

We provide the following materials in this manuscript:

- Sec. 1 more details of our designed method.
- Sec. 2 computational costs comparison.
- Sec. 3 visualization results
- Sec. 4 future work

1. More Details

In our proposed method, we adopt NAFNet [1] as the backbone of the denoising network. NAFNet is a very simple but efficient baseline for image restoration task. We adapt it to video restoration tasks based on its original setting. Specifically, it is enlarged by increasing the width to 64, the number of blocks of NAFNet to 44. The additional MLP layer comprised of SimpleGate and a linear layer is incorporated into each block for time embedding as shown in Figure 1.

2. Computational Costs Comparison

In this section, we showcase our speed advantages against diffusion-based restoration methods. We conduct all inference stages on an RTX4090 GPU to ensure a fair comparison. The results of FLOPs and run-time on a video clip of 5 frames with 256×256 patch size are displayed in Table 1. This demonstrates that our approach achieves superior model complexity and run-time performance after incorporating the adaptation mechanism while obtaining promising restoration performance.

3. More Visualization Results

We present more visual comparisons against state-of-the-art methods on synthetic and real datasets to demonstrate the excellent visual performance and generalization ability of Diff-TTA in removing arbitrary adverse weather conditions. The restored results of consecutive frames with diverse weather conditions are also displayed in the supplementary videos.

Visualization comparison on seen datasets. We provide more qualitative comparisons between our Diff-TTA and SOTA methods. The results are shown in Figure 2. Our Diff-TTA achieves the best visual quality by removing adverse components and recovering more background details.

Visualization comparison on unseen datasets. We provide qualitative comparisons between our Diff-TTA and SOTA methods. The results are shown in Figure 3, 4.

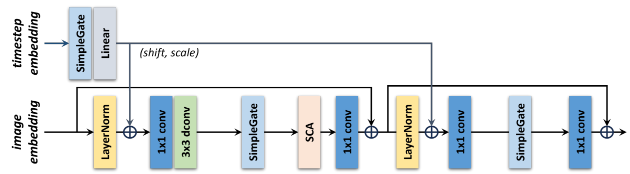

Figure 1. Detailed structure of the Denoising NAFNet's block.

Table 1. Comparison of FLOPs and Run-time on a video clip of 5 frames with 256×256 patch size.

Method	IR-SDE [2]	Refusion [3]	WeatherDiffusion [4]	Diff-TTA(ours)
FLOPs/G	1896.67×100 steps	372.38×100 steps	1242.01×1000 steps	385.14×25 steps
Run-time/s	23.42	12.55	542.76	6.01

Our Diff-TTA achieves the best visual quality by removing unknown adverse components and recovering more background details.

Visualization comparison on real-world data. We provide more qualitative comparisons between our Diff-TTA and SOTA methods. The results are shown in Figure 5. Our Diff-TTA achieves the best visual quality by removing unknown adverse components and recovering more background details, further validating its generalization.

4. Future Work

In the future, we plan further to accelerate the inference speed of Diff-TTA for real-time applications. Additionally, we will apply test-time adaptation to other restoration tasks.

References

[1] Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, and Jian Sun. Simple baselines for image restoration. In *European Conference on Computer Vision*, pages 17–33. Springer, 2022. 1

[2] Ziwei Luo, Fredrik K Gustafsson, Zheng Zhao, Jens Sjölund, and Thomas B Schön. Image restoration with mean-reverting stochastic differential equations. *arXiv preprint arXiv:2301.11699*, 2023. 1

[3] Ziwei Luo, Fredrik K Gustafsson, Zheng Zhao, Jens Sjölund, and Thomas B Schön. Refusion: Enabling large-size realistic image restoration with latent-space diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1680–1691, 2023. 1

[4] Ozan Özdenizci and Robert Legenstein. Restoring vision in adverse weather conditions with patch-based denoising diffusion models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023. 1

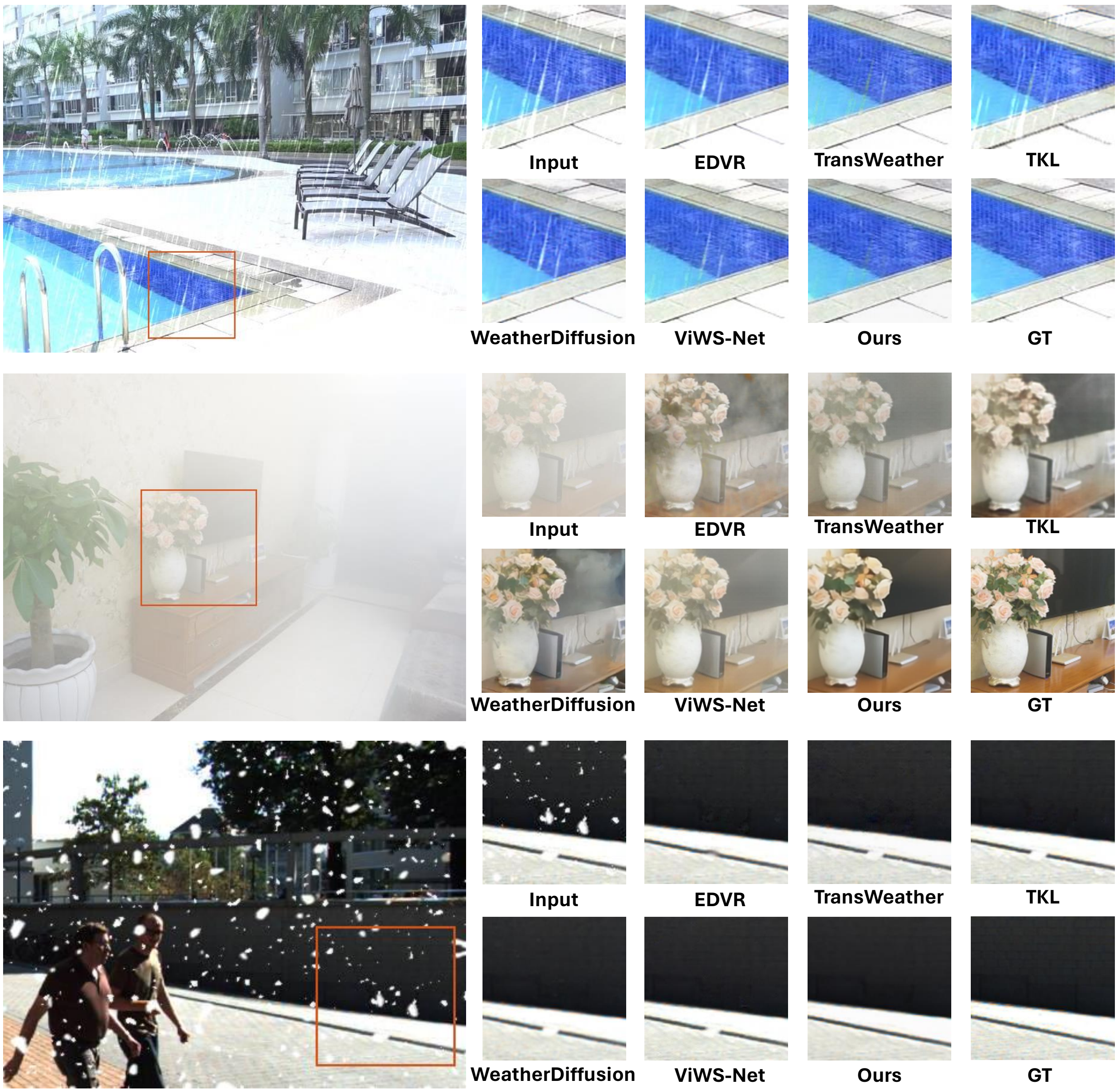

Figure 2. More visual comparisons on seen weather conditions.

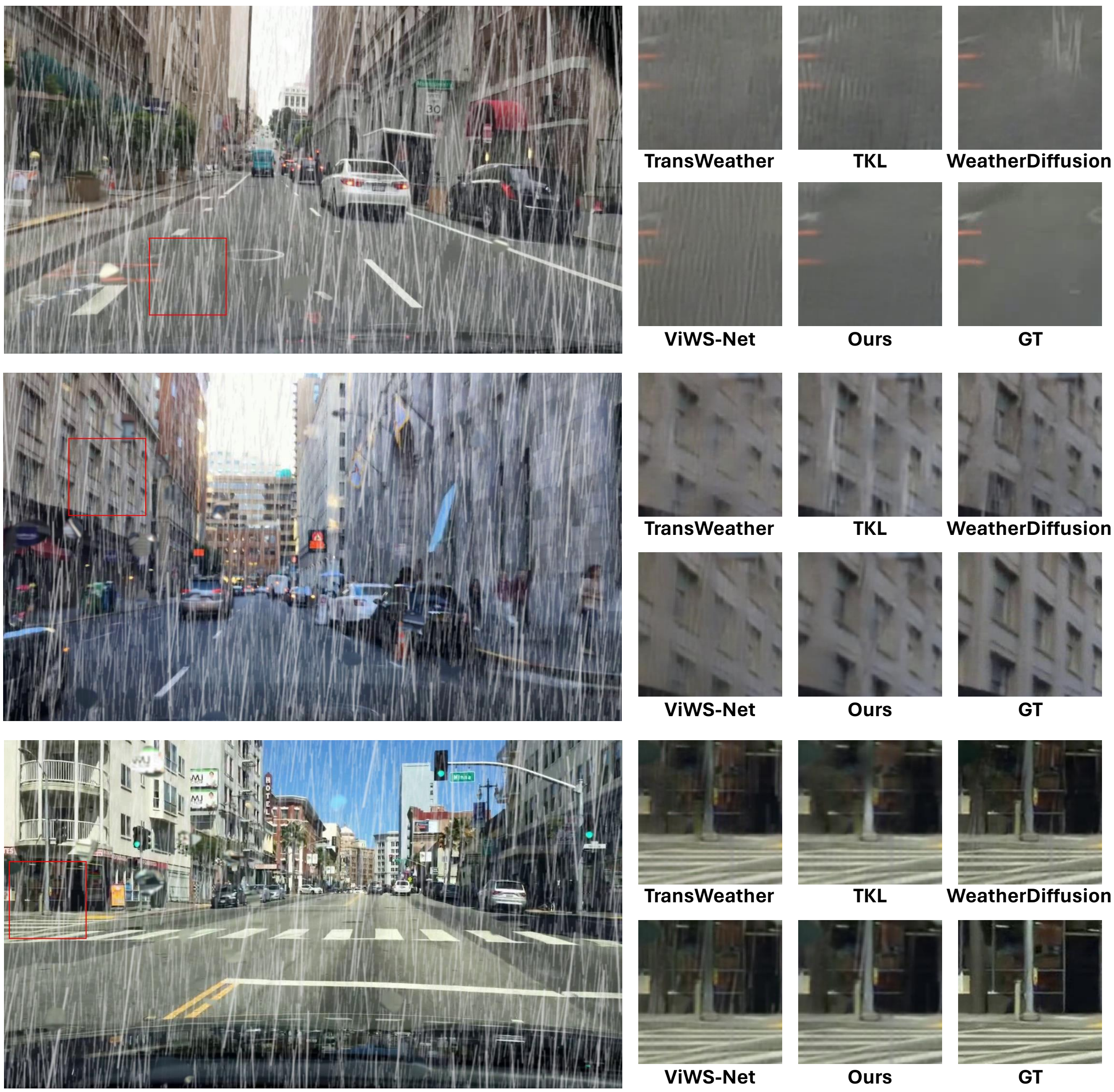

Figure 3. More visual comparisons on unseen weather conditions.

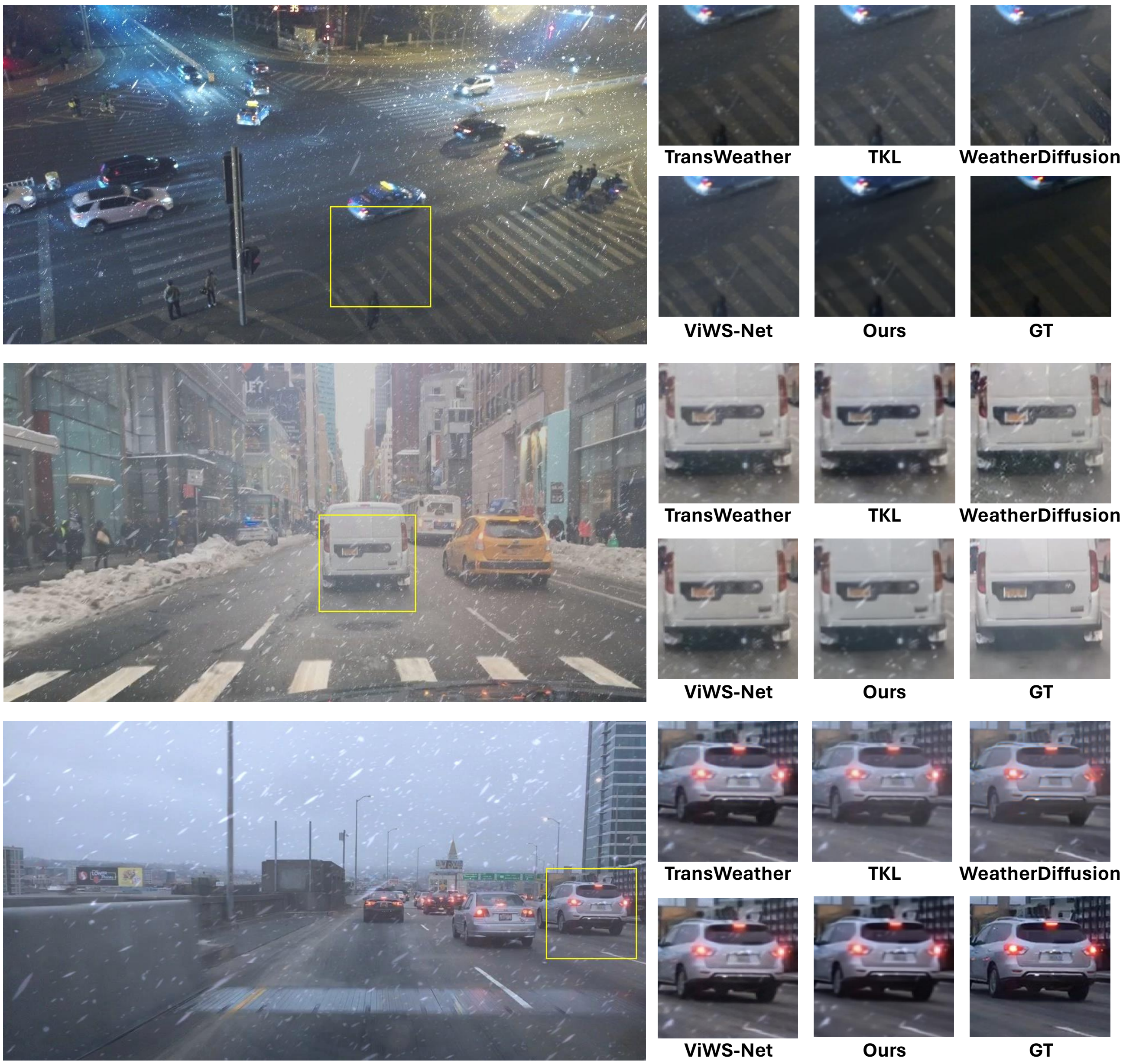

Figure 4. More visual comparisons on unseen weather conditions.

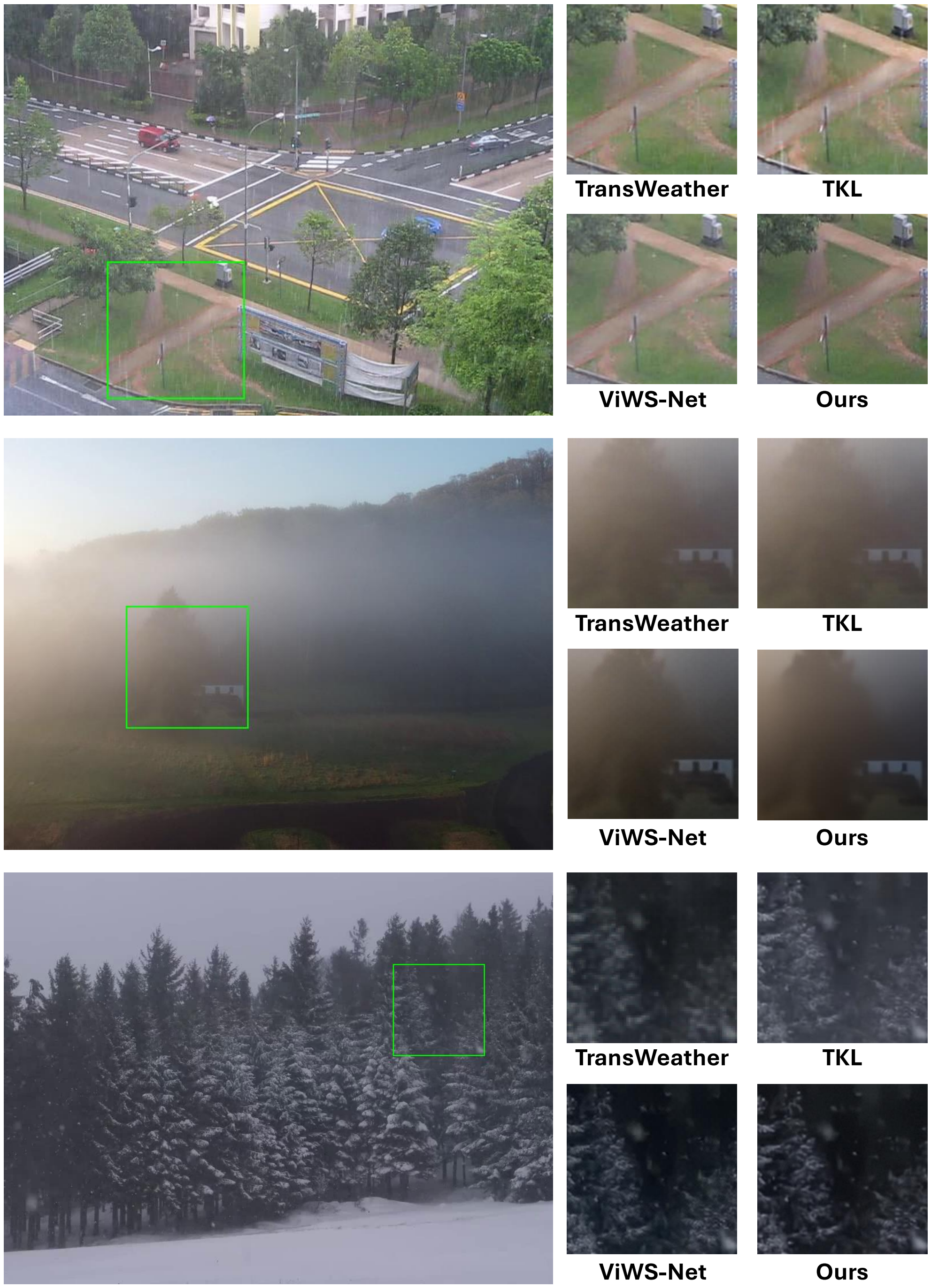

Figure 5. More visual comparisons on real-world scenarios.