Global Structure-Aware Diffusion Process for Low-Light Image Enhancement (*Supplementary Material*)

1 Inference Details

Generation. Following [1, 2], we apply a more flexible control over the number of generation steps and noise schedule during the reverse process after embedding $\bar{\alpha}_t$ into the diffusion model. To be specific, we utilized 20 and 10 generation steps when testing on LOLv1 and LOLv2 faster inference, respectively, and empirically selected $1-\alpha_0 = 6\times10^{-4}$ and $1-\alpha_T = 0.88$ for LOLv1, $1-\alpha_0 = 9\times10^{-4}$ and $1-\alpha_T = 0.85$ for LOLv2-real, and $1-\alpha_0 = 2\times10^{-3}$ and $1-\alpha_T = 0.84$ for LOLv2-synthetic in the reverse process.

Inference time. Table S1 provides the inference time of different methods. Although our method consumes a little more time owing to the inherent computational burden associated with diffusion-based methods, our work shows promising performance compared to other methods, as evidenced by the quantitative results in Table 1 and visual results in Figs. 3, and 4 of our manuscript, and more visual results in the following Figs. S1, S2, and S3. We will also try to apply other techniques [3, 4] to speed-up the reverse process.

Table S1: Comparisons of inference time of the recent SOTA methods on the LOLv1 dataset. Note that the inference time per image was conducted using an RTX3080 GPU on the same server.

Methods	LLFlow [6]	SNR-Aware [8]	LLFormer [5]	Ours
Inference time (s)	0.40	0.09	0.31	0.55

2 More Visual Results

Figs. S1, S2, and S3 present more enhanced results on the LOLv1 [7], LOLv2 [9] datasets, obtained by our proposed method and other three most recent SOTA approaches, i.e., LLFlow [6], SNR-Aware [8], and LLFormer [5]. It can be observed that our method consistently outperforms compared approaches by effectively suppressing artifacts and revealing image details, leading to visually appealing results that are more faithful to the original scene.

3 Ablation Studies of Matrix Rank across the Reverse Process

Fig. S4 presents the matrix rank across clusters from the intermediate-generated results of the reverse process. The empirical findings reveal that in the absence of the proposed global structure-aware regularization, the rank of a cluster of patches tends to manifest either lower (the first and second rows of Fig. S4) or higher singular values (the last two rows of Fig. S4) relative to the ground truth (GT). Lower singular values denote the loss of reconstructing distinct components within similar patches, whereas higher values indicate the emergence of dissimilar contents or noise. Nonetheless, the incorporation of the proposed regularization technique consistently captures an appropriate rank in matrix, thereby facilitating the reconstruction of the global structure in images.

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

Figure S1: Visual comparisons of the enhanced results by different methods on LOLv1.

Figure S2: Visual comparisons of the enhanced results by different methods on LOLv2-real.

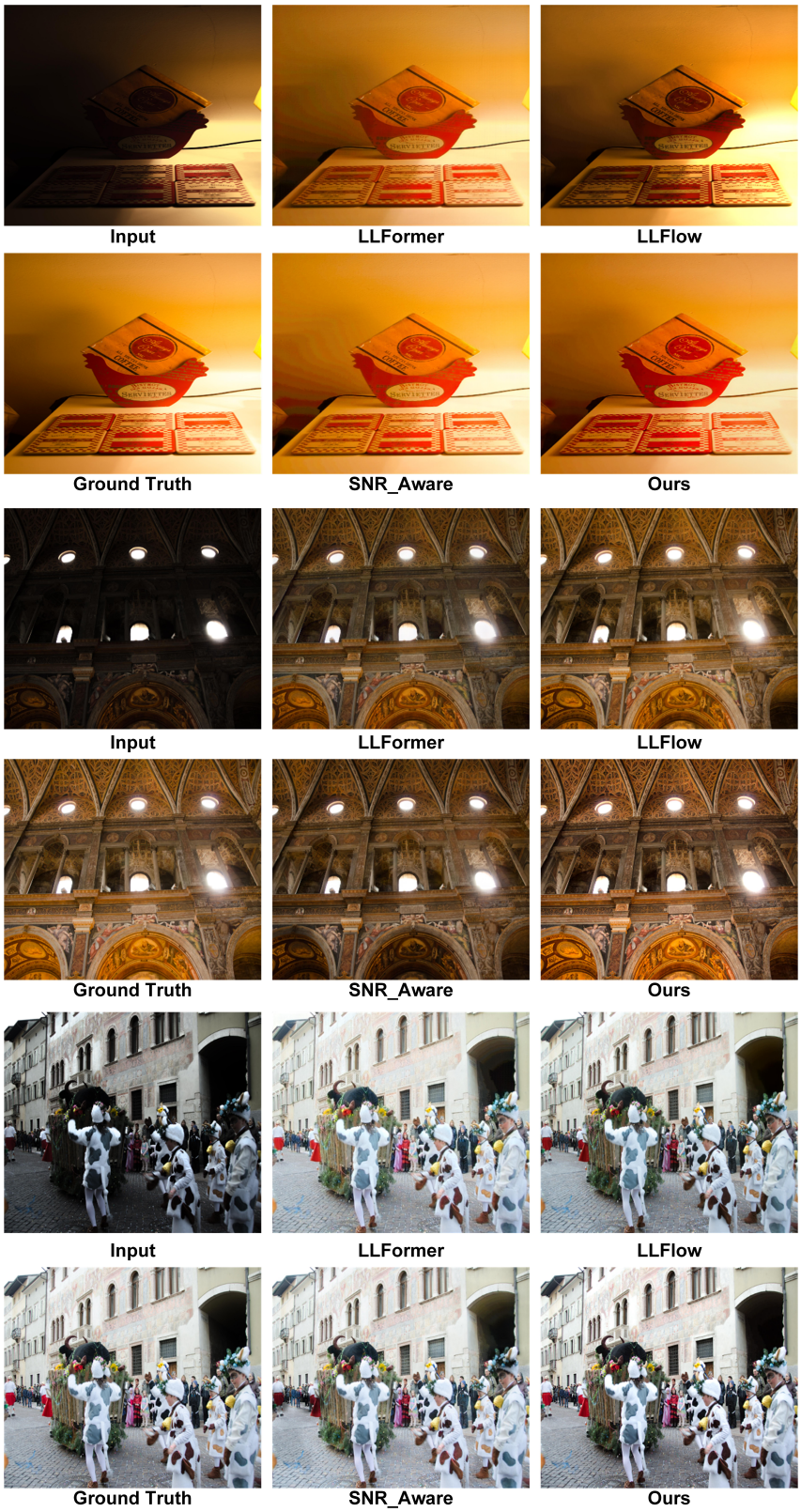

Figure S3: Visual comparisons of the enhanced results by different methods on LOLv2-synthetic.

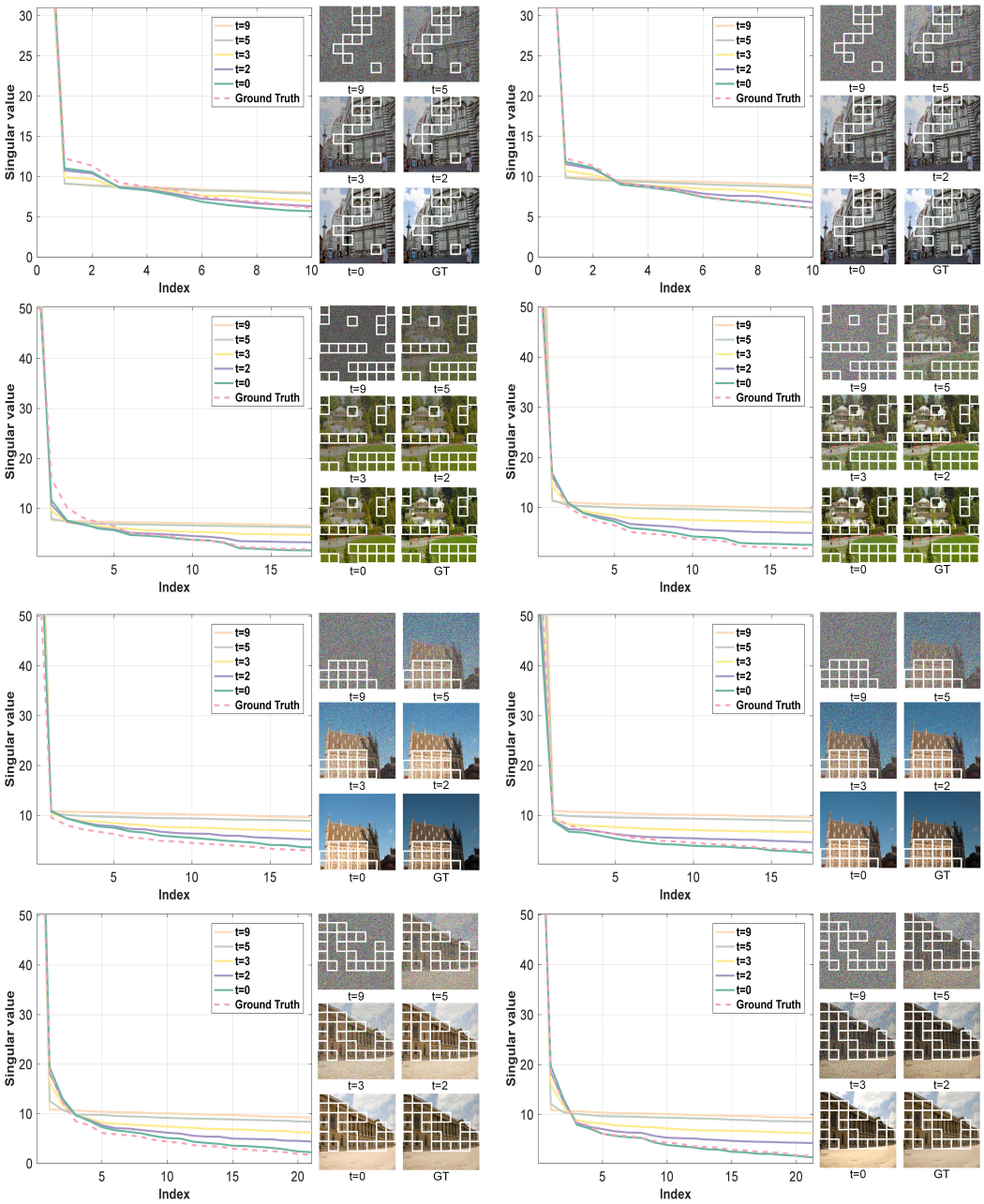

Figure S4: The distribution of the singular values of a cluster from the intermediate enhanced results over different timesteps (t) of the reverse process. Note that we illustrate the rank of a cluster of patches, which are highlighted with a white box.

References

[1] Nanxin Chen, Yu Zhang, Heiga Zen, Ron J Weiss, Mohammad Norouzi, and William Chan. Wavegrad: Estimating gradients for waveform generation. In *Proceedings of International Conference on Learning Representations*, 2021.

[2] Chitwan Saharia, Jonathan Ho, William Chan, Tim Salimans, David J Fleet, and Mohammad Norouzi. Image super-resolution via iterative refinement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(4):4713–4726, 2022.

[3] Tim Salimans and Jonathan Ho. Progressive distillation for fast sampling of diffusion models. *arXiv preprint arXiv:2202.00512*, 2022.

[4] Yang Song, Prafulla Dhariwal, Mark Chen, and Ilya Sutskever. Consistency models. In *Proceedings of International Conference on Machine Learning*, 2023.

[5] Tao Wang, Kaihao Zhang, Tianrun Shen, Wenhan Luo, Bjorn Stenger, and Tong Lu. Ultra-high-definition low-light image enhancement: A benchmark and transformer-based method. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 2654–2662, 2023.

[6] Yufei Wang, Renjie Wan, Wenhan Yang, Haoliang Li, Lap-Pui Chau, and Alex Kot. Low-light image enhancement with normalizing flow. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 2604–2612, 2022.

[7] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. *arXiv preprint arXiv:1808.04560*, 2018.

[8] Xiaogang Xu, Ruixing Wang, Chi-Wing Fu, and Jiaya Jia. Snr-aware low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17714–17724, 2022.

[9] Wenhan Yang, Wenjing Wang, Haofeng Huang, Shiqi Wang, and Jiaying Liu. Sparse gradient regularized deep retinex network for robust low-light image enhancement. *IEEE Transactions on Image Processing*, 30:2072–2086, 2021.