

RESPLAT: DEGRADATION-AGNOSTIC FEED-FORWARD GAUSSIAN SPLATTING VIA SELF-GUIDED RESIDUAL DIFFUSION: *Supplementary Material*

Youngho Yoon & Kuk-Jin Yoon

Visual Intelligence Lab., KAIST, South Korea
 {dudgh1732, kjyoon}@kaist.ac.kr

1 OVERVIEW

Our supplementary material delves deeper into the proposed method, offering additional results and discussions that were not thoroughly addressed in the main paper:

1. Inference time for novel view synthesis.
2. Adaptation to other models.
3. Additional novel view synthesis results.
4. Additional Qualitative Comparisons.
5. Video Results of novel view synthesis.

2 INFERENCE TIME

In this section, we analyze the inference time required by various methods to perform novel view synthesis (NVS) in the presence of corruption. Table 1 reports the inference time for each method, measured in seconds. Gaussian Splatting-based methods, including ours, achieve inference times **within 1 second**, whereas GAURA Gupta et al. (2024) exhibits a significantly higher inference time of 33.03 seconds per frame. Other methods, such as AiRnet Li et al. (2022), PromptIR Potlapalli et al. (2024), and DiffUIR Zheng et al. (2024), are evaluated based on the total time required to perform NVS through image restoration and MVSGaussian Liu et al. (2025). Notably, despite being diffusion-based models, both DiffUIR Zheng et al. (2024) and ours drastically reduce inference time by completing the process within three steps using DDIM Song et al. (2020) sampling.

Table 1: Inference time for novel view synthesis with three source-view inputs (512×512 images).

Model	AiRnet	PromptIR	DiffUIR	GAURA	Ours
Time (sec.)	0.558	0.475	0.814	33.03	0.816

3 ADAPTATION TO OTHER MODELS

We use DiffUIR Zheng et al. (2024) as a widely used structure for degradation-agnostic restoration in the main paper. Our modular framework also allows easy replacement with other recent models. As shown in Table 2, applying our method to FoundIR Li et al. (2024) and DGSolver Wang et al. (2025) yields consistent gains, showing its robustness. The table presents averaged results over five degradation scenarios in the LLFF degradation dataset, further demonstrating that our method can be robustly adapted to diffusion-based universal image restoration approaches.

4 ADDITIONAL NOVEL VIEW SYNTHESIS RESULTS

In this section, we analyze the results of NVS under varying numbers of source-view inputs. We present the NVS results for **three** source-view inputs in Table 1 of the main paper. Additionally, Table 3 and Table 4 present the NVS outcomes for different types of corruption when provided with **two** and **four** source-view inputs, respectively.

Table 2: Mean results over five degradation scenarios.

Method	Novel View Synthesis			Image Restoration		
	PSNR(\uparrow)	SSIM(\uparrow)	LPIPS(\downarrow)	PSNR(\uparrow)	SSIM(\uparrow)	LPIPS(\downarrow)
FoundIR	21.52	0.8133	0.2537	25.44	0.9153	0.1815
Ours(w/ FoundIR)	22.12	0.8246	0.2332	26.88	0.9287	0.1595
DGSolver	21.94	0.8274	0.2345	26.78	0.9349	0.1560
Ours(w/ DGSolver)	22.15	0.8332	0.2240	27.32	0.9413	0.1461

The results in Table 3 and Table 4 demonstrate the performance of various methods under five corruption types: motion blur, snow, haze, low-light, and rain. A consistent trend is observed where increasing the number of source-view inputs improves the quality of NVS across all metrics (PSNR, SSIM, and LPIPS), highlighting the importance of input redundancy in mitigating the effects of corruption.

Specifically, ReSplat consistently achieves the best or near-best performance across all corruption types, benefiting significantly from additional source views. For example, under motion blur, ReSplat improves its PSNR from 22.29 to 23.31 and its SSIM from 0.7809 to 0.8126 when increasing the source views from two to four.

In more challenging conditions, such as haze and low-light, methods like AiRnet and GAURA struggle to maintain performance, with relatively lower PSNR and SSIM values. In contrast, ReSplat demonstrate robustness, leading in metrics such as SSIM (e.g., 0.8308 in low-light with four inputs). Consequently, the results underline the efficacy of methods like ReSplat in leveraging multiple source views to enhance the robustness and accuracy of NVS under various corruptions.

5 ADDITIONAL QUALITATIVE COMPARISONS

To further evaluate the effectiveness of our proposed method, we compare its image restoration performance against state-of-the-art methods across five types of degradations: haze, low-light, snow, rain, and motion blur, using the LLFF degradation dataset. Figure 1 illustrates the visual results of these methods.

For haze removal, our method demonstrates superior clarity and color restoration, closely matching the ground truth (GT). Competing methods, while effective to varying degrees, fail to recover finer details and exhibit either oversaturation or blurriness.

Under low-light conditions, our approach effectively restores brightness and color fidelity, ensuring clear visibility of textures and structures. In contrast, AiRnet Li et al. (2022) introduces significant noise, while PromptIR Potlapalli et al. (2024) and DiffUIR Zheng et al. (2024) struggle to adequately enhance contrast and color balance.

In snow degradation, ours achieves an impressive reduction in occlusions caused by snow particles while maintaining sharpness and natural tones in the background. Other methods either fail to completely remove snow artifacts or result in excessive smoothing that distorts scene details.

For rain removal, ours excels in accurately reconstructing background details, surpassing other approaches in preserving structural consistency. While AiRNet Li et al. (2022) and PromptIR Potlapalli et al. (2024) mitigate rain streaks, they introduce artifacts or distortions, which are absent in our outputs.

Lastly, in the case of motion blur, ours exhibits remarkable capability in recovering image sharpness and object edges, significantly outperforming other approaches. AiRNet Li et al. (2022) and PromptIR Potlapalli et al. (2024) produce partially blurred outputs, and DiffUIR Zheng et al. (2024) shows limitations in handling complex motion patterns. These qualitative results confirm the robustness and generalization capability of ours across diverse degradation types, positioning it as a state-of-the-art solution for universal image restoration tasks.

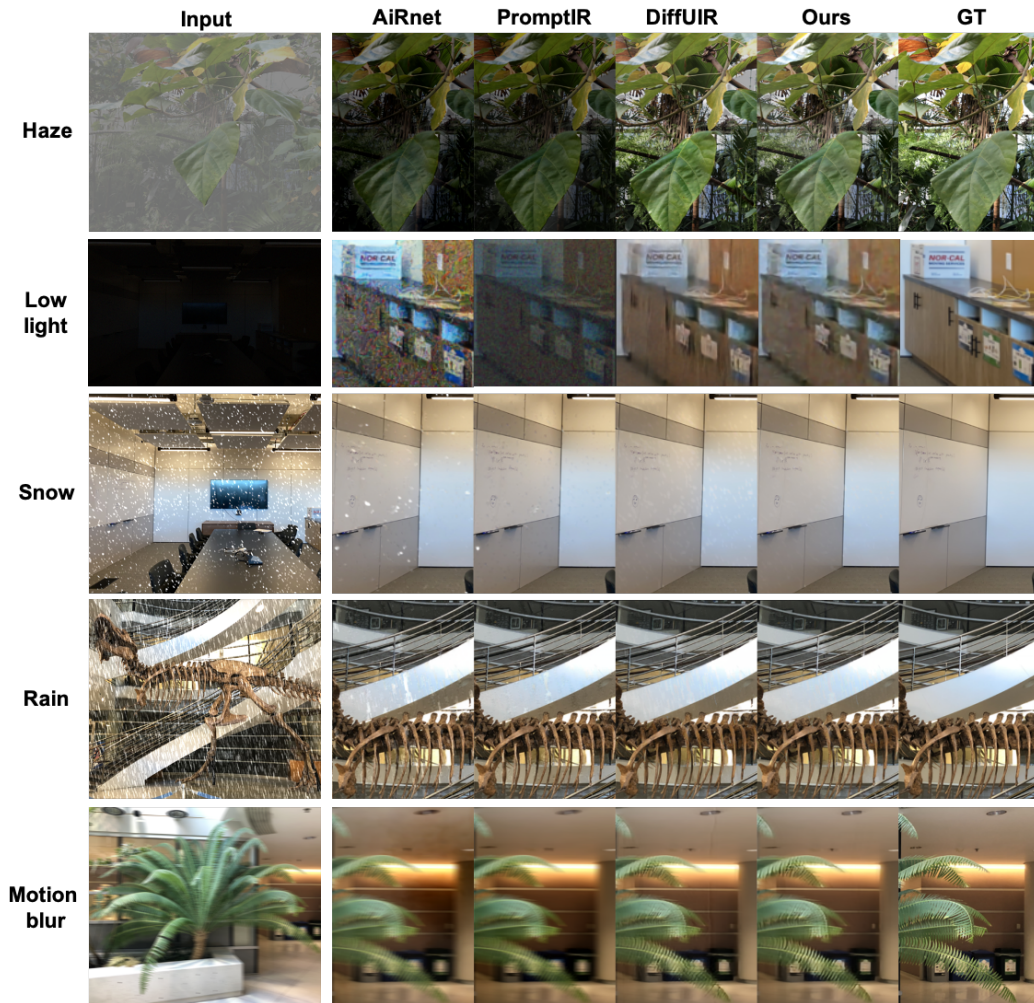


Figure 1: Visual Comparisons of universal image restoration results of 5 types (motion blur, snow, haze, low-light, rain) on LLFF degradation dataset.

Table 3: Novel View Synthesis results with **two source-view** inputs of 5 types (motion blur, snow, haze, low-light, rain) on LLFF degradation dataset. The **best** scores and **second best** scores are highlighted with their respective colors.

Method	Type	Novel View Synthesis		
		PSNR(\uparrow)	SSIM(\uparrow)	LPIPS(\downarrow)
AiRnet	Motion Blur	19.33	0.6659	0.4534
PromptIR		19.19	0.6614	0.4491
GAURA		19.35	0.6589	0.5311
DiffUIR		22.20	0.7709	0.3466
ReSplat		22.29	0.7803	0.3280
AiRnet	Snow	19.58	0.6646	0.3210
PromptIR		19.87	0.6796	0.2844
GAURA		19.61	0.6876	0.4788
DiffUIR		23.30	0.8290	0.2039
ReSplat		23.28	0.8319	0.1923
AiRnet	Haze	9.049	0.3707	0.3951
PromptIR		9.675	0.4547	0.3505
GAURA		16.71	0.6815	0.4972
DiffUIR		21.07	0.8198	0.1947
ReSplat		21.33	0.8230	0.1904
AiRnet	Low-light	6.296	0.0680	0.6207
PromptIR		6.366	0.0807	0.6276
GAURA		14.97	0.6254	0.5537
DiffUIR		18.25	0.7985	0.2749
ReSplat		19.02	0.7997	0.2595
AiRnet	Rain	19.86	0.6753	0.3678
PromptIR		20.16	0.6998	0.3220
GAURA		18.88	0.6377	0.5328
DiffUIR		22.74	0.8098	0.2721
ReSplat		23.02	0.8208	0.2419

Table 4: Novel View Synthesis results with **four source-view** inputs of 5 types (motion blur, snow, haze, low-light, rain) on LLFF degradation dataset. The **best** scores and **second best** scores are highlighted with their respective colors.

Method	Type	Novel View Synthesis		
		PSNR(\uparrow)	SSIM(\uparrow)	LPIPS(\downarrow)
AiRnet	Motion Blur	20.11	0.6895	0.4249
PromptIR		20.04	0.6871	0.4208
GAURA		21.78	0.7443	0.4105
DiffUIR		22.39	0.7741	0.3505
ReSplat		23.31	0.8126	0.3033
AiRnet	Snow	20.22	0.6852	0.3026
PromptIR		20.42	0.6975	0.2729
GAURA		23.34	0.8247	0.2970
DiffUIR		24.17	0.8555	0.1925
ReSplat		24.71	0.8710	0.1675
AiRnet	Haze	9.159	0.3840	0.3891
PromptIR		9.784	0.4651	0.3507
GAURA		17.44	0.7339	0.4408
DiffUIR		21.64	0.8402	0.1960
ReSplat		22.20	0.8549	0.1760
AiRnet	Low-light	6.301	0.0680	0.6020
PromptIR		6.367	0.0805	0.6240
GAURA		15.39	0.6810	0.4979
DiffUIR		18.86	0.8264	0.2644
ReSplat		19.56	0.8308	0.2429
AiRnet	Rain	20.49	0.6988	0.3416
PromptIR		20.71	0.7174	0.2991
GAURA		22.30	0.7843	0.3862
DiffUIR		23.02	0.8211	0.2756
ReSplat		24.22	0.8562	0.2153

6 VIDEO RESULTS

We present video results for five degradation types: motion blur, snow, haze, low light, and rain. The video results are included as part of the supplementary materials for further reference. Each video includes outputs from AiRnet Li et al. (2022), PromptIR Potlapalli et al. (2024), DiffUIR Zheng et al. (2024), and ours. Our approach consistently outperforms others, ensuring sharp, artifact-free restoration and temporal coherence across frames. For motion blur and haze, our method provides clear and stable results, while competing methods show residual artifacts or flickering. In snow and rain scenarios, our method effectively removes occlusions while preserving detail. For low-light, we achieve superior brightness enhancement and stability, avoiding noise seen in other methods. These results demonstrate the robustness of our method in dynamic scenarios. Furthermore, this consistent performance across varying degradation types reinforce the adaptability and effectiveness of our approach, making it well-suited for practical applications in real-world novel view synthesis tasks.

REFERENCES

- Vinayak Gupta, Rongali Simhachala Venkata Girish, Ayush Tewari, Kaushik Mitra, et al. Gaura: Generalizable approach for unified restoration and rendering of arbitrary views. *arXiv preprint arXiv:2407.08221*, 2024.
- Boyun Li, Xiao Liu, Peng Hu, Zhongqin Wu, Jiancheng Lv, and Xi Peng. All-in-one image restoration for unknown corruption. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 17452–17462, 2022.
- Hao Li, Xiang Chen, Jiangxin Dong, Jinhui Tang, and Jinshan Pan. Foundir: Unleashing million-scale training data to advance foundation models for image restoration. *arXiv preprint arXiv:2412.01427*, 2024.
- Tianqi Liu, Guangcong Wang, Shoukang Hu, Liao Shen, Xinyi Ye, Yuhang Zang, Zhiguo Cao, Wei Li, and Ziwei Liu. Mvsgaussian: Fast generalizable gaussian splatting reconstruction from multi-view stereo. In *European Conference on Computer Vision*, pp. 37–53. Springer, 2025.

- Vaishnav Potlapalli, Syed Waqas Zamir, Salman H Khan, and Fahad Shahbaz Khan. Promptir: Prompting for all-in-one image restoration. *Advances in Neural Information Processing Systems*, 36, 2024.
- Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. *arXiv preprint arXiv:2010.02502*, 2020.
- Hebaixu Wang, Jing Zhang, Haonan Guo, Di Wang, Jiayi Ma, and Bo Du. Dgsolver: Diffusion generalist solver with universal posterior sampling for image restoration. *arXiv preprint arXiv:2504.21487*, 2025.
- Dian Zheng, Xiao-Ming Wu, Shuzhou Yang, Jian Zhang, Jian-Fang Hu, and Wei-Shi Zheng. Selective hourglass mapping for universal image restoration based on diffusion model. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 25445–25455, 2024.