Appendix

A Additional Comparisons

Sparse-view reconstruction. To further evaluate our proposed Deceptive-NeRF and Deceptive-3DGS in few-view reconstruction tasks, we compared our method against baseline approaches FreeNeRF [6] and DiffusioNeRF [5] in novel view synthesis on mip-NeRF 360 [1] and Hypersim [2] datasets, with 10 input views. As illustrated in Figure A and Figure B, our method's synthesized novel

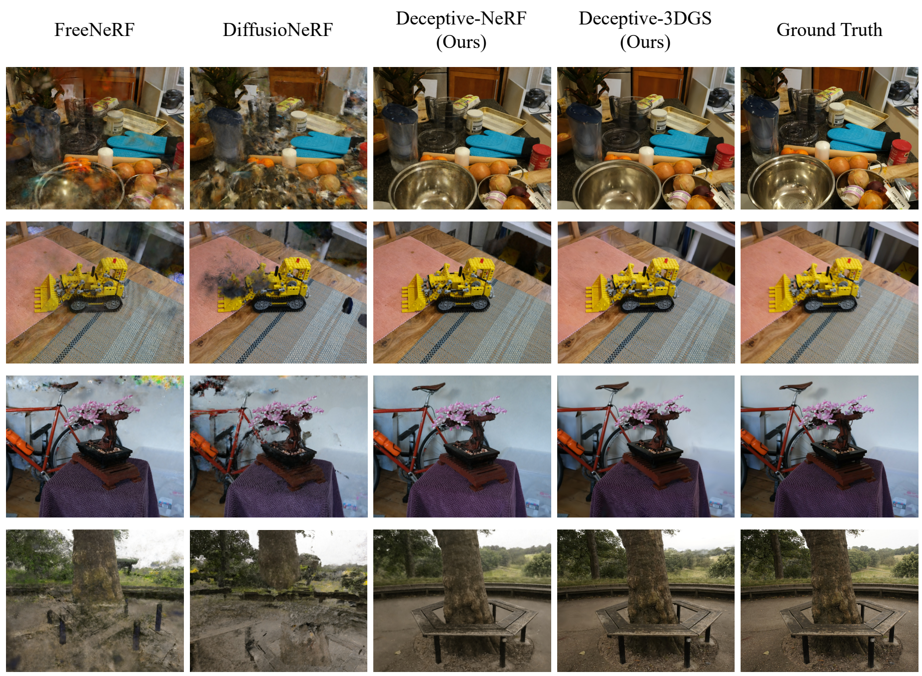

Fig. A: Qualitative comparisons of few-view reconstruction on mip-NeRF 360 dataset [1]. For each scene, we reconstruct with 10 input views. Our methods can provide high-quality reconstructions, whereas baseline methods may yield completely unreasonable and incorrect reconstructions.

views do not produce floating artifacts and achieve better restoration of distant scenes. Even in the challenging 360-degree scenes, our method consistently delivers high-quality reconstructions, outperforming baseline methods that may result in wholly unreasonable and incorrect reconstructions. This showcases our method's advanced few-view reconstruction capabilities over the baselines.

Novel view super-resolution. We further validate the capability of our method to perform novel-view super-resolution in 360-degree scenes. Utilizing 20 input views and downsampling the input images by a factor of 4 for each scene on the mip-NeRF 360 dataset [1], our method, Deceptive-3DGS, achieves high-quality super-resolution at novel viewpoints. It recovers details of objects, such as Lego

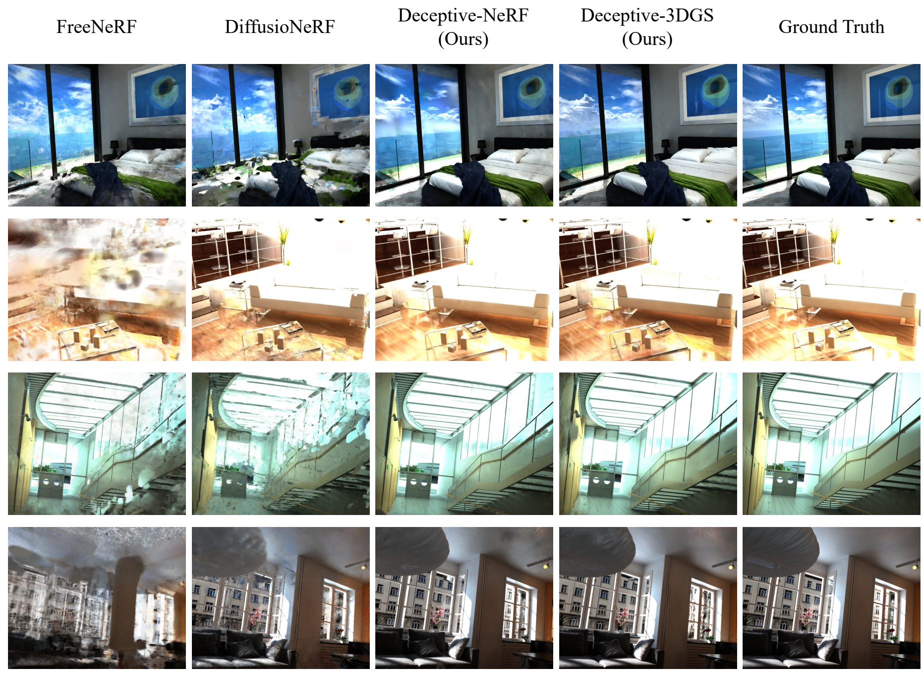

Fig. B: Qualitative comparisons of few-view reconstruction on Hypersim dataset [2]. For each scene, we reconstruct with 10 input views. Novel views synthesized by our proposed Deceptive-NeRF and Deceptive-3DGS do not exhibit floating artifacts. They offer better restoration of distant scenes (such as the sky and buildings viewed seen through the windows).

toys and flower petals, more effectively than competing approaches like FreeNeRF [6] and NeRF-SR [4].

B Additional Evaluations

"Regularizer" v.s. "view densifier" Different from the straightforward utilization of the 2D diffusion model as a "scorer" for synthesized novel views to regularize NeRF/3DGS training, our approach uses it to generate pseudo-observations to densify observations. To better validate the advantages of this choice, we evaluate the reconstruction quality and efficiency between the two methods of using diffusion models to enhance 3D reconstruction. We experiment on the Hypersim dataset [2] with 10 input views. In Table A, the "view densifier" refers to our proposed Deceptive-NeRF, while the "regularizer" denotes a variant of our method that uses a diffusion prior to regularize NeRF training. Our proposed "view densifier" approach outperforms the "regularizer" in all metrics of rendering quality. Moreover, our approach achieves nearly ten times faster training speed and increased rendering speed. This improvement is primarily because our method does not require inferring the diffusion model at every training step.

Number of input views. We run Deceptive-NeRF on Scene 027_003 from Hypersim across a range of 2 to 10 input views and examine the testing PSNRs

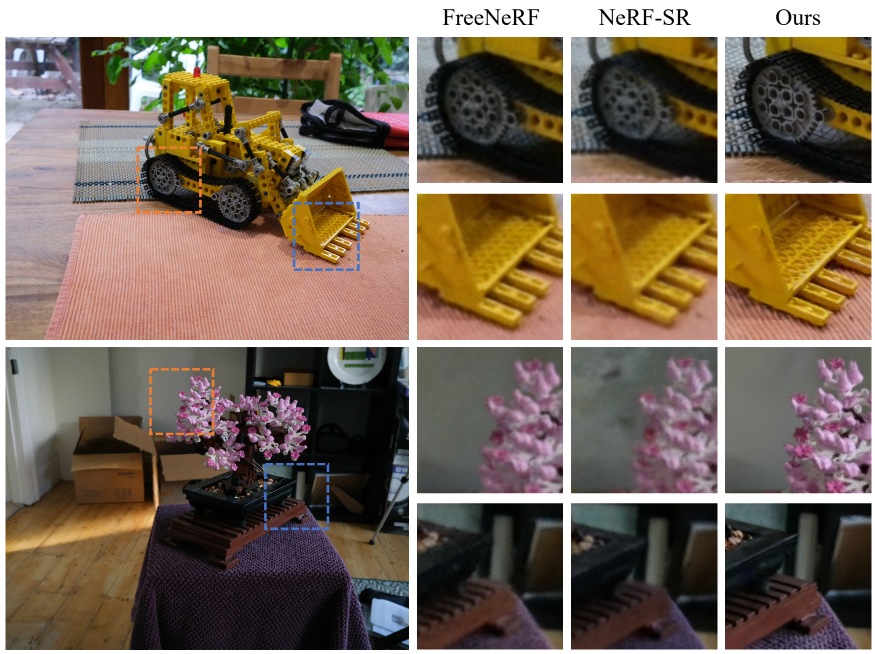

Fig. C: Qualitative comparisons of novel view super-resolution on mip-NeRF 360 dataset [1]. For each scene, we utilize 20 input views and downsample the input images by a factor of 4. Our method (Deceptive-3DGS) manages to achieve high-quality super-resolution at novel viewpoints, recovering details of objects such as Lego toys and flower petals more effectively than baseline approaches.

Table A: Quantitative comparison of two methods to utilize 2D diffusion models for 3D reconstruction.

Method	PSNR (↑)	SSIM (↑)	LPIPS (↓)	Training Time (↓)	Rendering FPS (↑)
"Regularizer"	19.31	0.710	0.253	2h40min	1.1
"View Densifier" (ours)	**20.44**	**0.748**	**0.173**	**17min**	**5.3**

of initial NeRFs and final NeRFs (10x densification). As depicted in Figure D, our pseudo-observations consistently enhance the reconstruction quality, demonstrating that our method can work effectively with as few as 3 input images.

Geometry recovery. We quantitatively evaluate geometry recovery on scan 65 of the DTU dataset with 9 input views, applying TSDF (truncated signed distance function) Fusion to extract meshes from trained Deceptive-NeRF and baseline methods. In Table B, we report their Chamfer distances to the ground truth, and our approach achieves higher reconstruction accuracy than baselines.

Uncertainty guidance. We conducted an ablation study on our uncertainty guidance on the Hypersim dataset [2]. As shown in Table C, without the guidance of our proposed uncertainty measure, there is a significant decline in reconstruction quality. Without uncertainty guidance, the deceptive diffusion model cannot effectively remove artifacts in coarse renderings, leading to inconsisten-

Table B: Quantitative comparison on geometry recovery.

	NeRF	FreeNeRF	DiffusioNeRF	Deceptive-NeRF (Ours)
Chamfer Distance (↓)	3.75	1.93	3.63	**1.86**

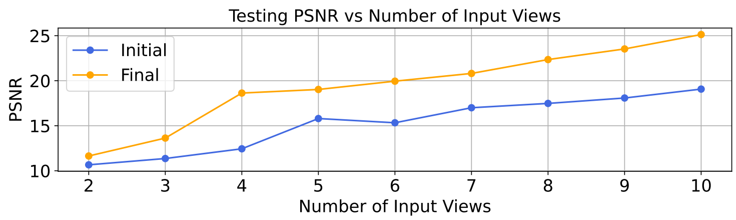

Fig. D: Initial and final testing PSNRs across different numbers of input views.

cies between pseudo-observations and input images. In Figure E, we visualize the uncertainty map and compare the coarse and final renderings under its guidance. Areas of high uncertainty (highlighted in brighter colors) correspond to artifacts, which are removed in the final rendering.

Table C: Quantitative ablation study on the uncertainty guidance.

Method	PSNR (↑)	SSIM (↑)	LPIPS (↓)
w/o uncertainty	18.61	0.703	0.28
w/ uncertainty (ours)	**20.44**	**0.748**	**0.173**

C Experimental Details

We adopt Nerfacto and Splatfacto from NerfStudio [3] as the backbones for Deceptive-NeRF and Deceptive-3DGS, respectively, utilizing the default proposal sampling, scene contraction, and appearance embeddings. We randomly initialize the Gaussians for Deceptive-3DGS, without utilizing Colmap point clouds. We alternately generate pseudo-observations and train scene representations five times, ultimately densifying observations to ten times their original amount. We randomly sample novel views $\{\phi^i_{\text{pseudo}}\}$ within the bounding box defined by the outermost input cameras.

References

1. Barron, J.T., Mildenhall, B., Verbin, D., Srinivasan, P.P., Hedman, P.: Mip-nerf 360: Unbounded anti-aliased neural radiance fields. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)(2022)

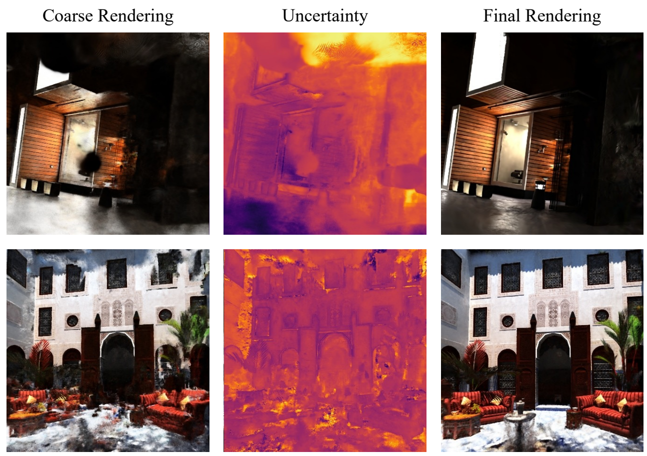

Fig. E: Visualizations of uncertainty guidance. Under the uncertainty measure's guidance, coarse rendering artifacts are eliminated in the pseudo-observations. In uncertainty maps, high uncertainty is indicated by brighter colors, while low uncertainty is shown in darker colors.

2. Roberts, M., Ramapuram, J., Ranjan, A., Kumar, A., Bautista, M.A., Paczan, N., Webb, R., Susskind, J.M.: Hypersim: A photorealistic synthetic dataset for holistic indoor scene understanding. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 10912–10922 (2021)
3. Tancik, M., Weber, E., Ng, E., Li, R., Yi, B., Kerr, J., Wang, T., Kristoffersen, A., Austin, J., Salahi, K., et al.: Nerfstudio: A modular framework for neural radiance field development. arXiv preprint arXiv:2302.04264 (2023)
4. Wang, C., Wu, X., Guo, Y.C., Zhang, S.H., Tai, Y.W., Hu, S.M.: Nerf-sr: High quality neural radiance fields using supersampling. In: Proceedings of the 30th ACM International Conference on Multimedia. pp. 6445–6454 (2022)
5. Wynn, J., Turmukhambetov, D.: DiffusioNeRF: Regularizing neural radiance fields with denoising diffusion models. In: arxiv (2023)
6. Yang, J., Pavone, M., Wang, Y.: FreeNeRF: Improving few-shot neural rendering with free frequency regularization. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)(2023)