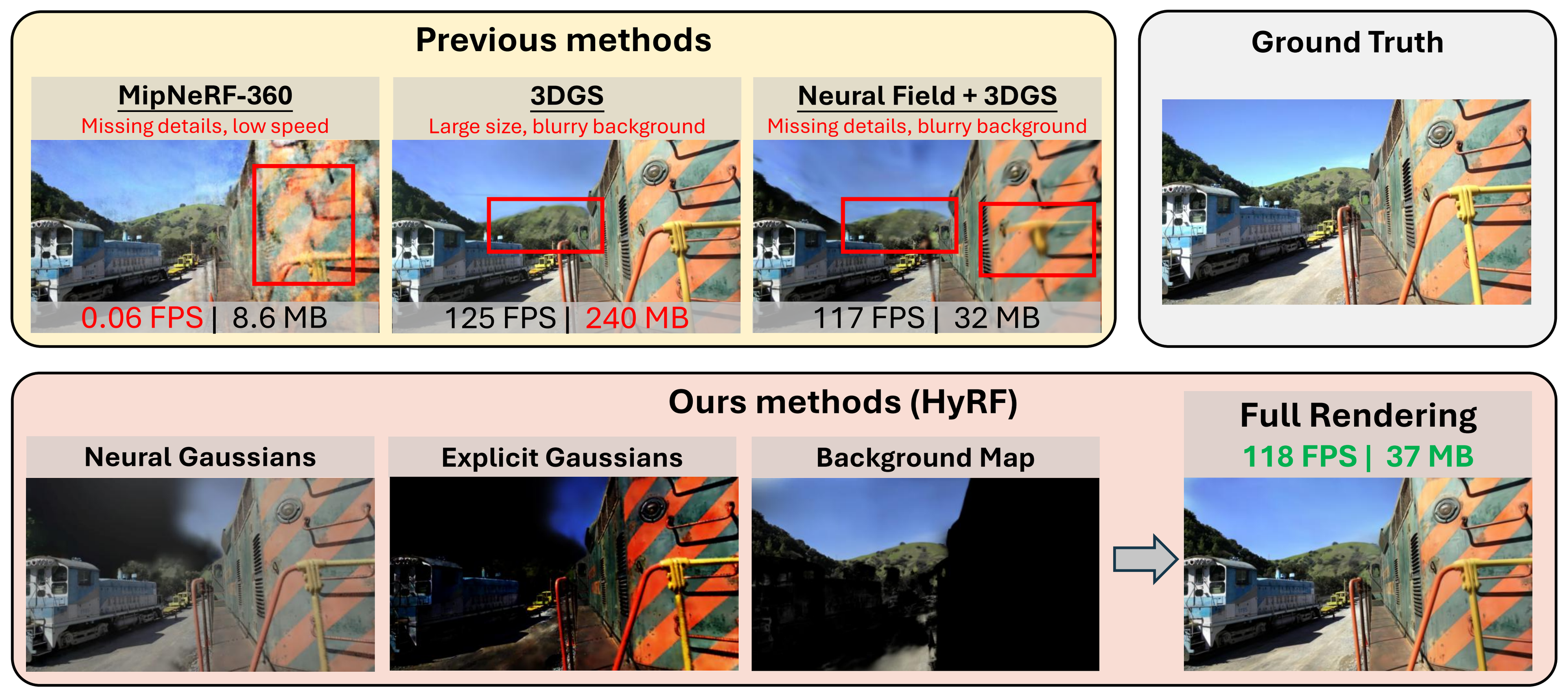
HyRF: Hybrid Radiance Fields for Memory-efficient and High-quality Novel View Synthesis

Zipeng Wang, Dan Xu
The Hong Kong University of Science and Technology

Website ArXiv

香港科技大學 THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

NEURAL INFORMATION PROCESSING SYSTEMS



The Trilemma of Scene Representations

Previous methods cannot simultaneously achieve **high rendering quality**, **real-time rendering** and **compact model size**.

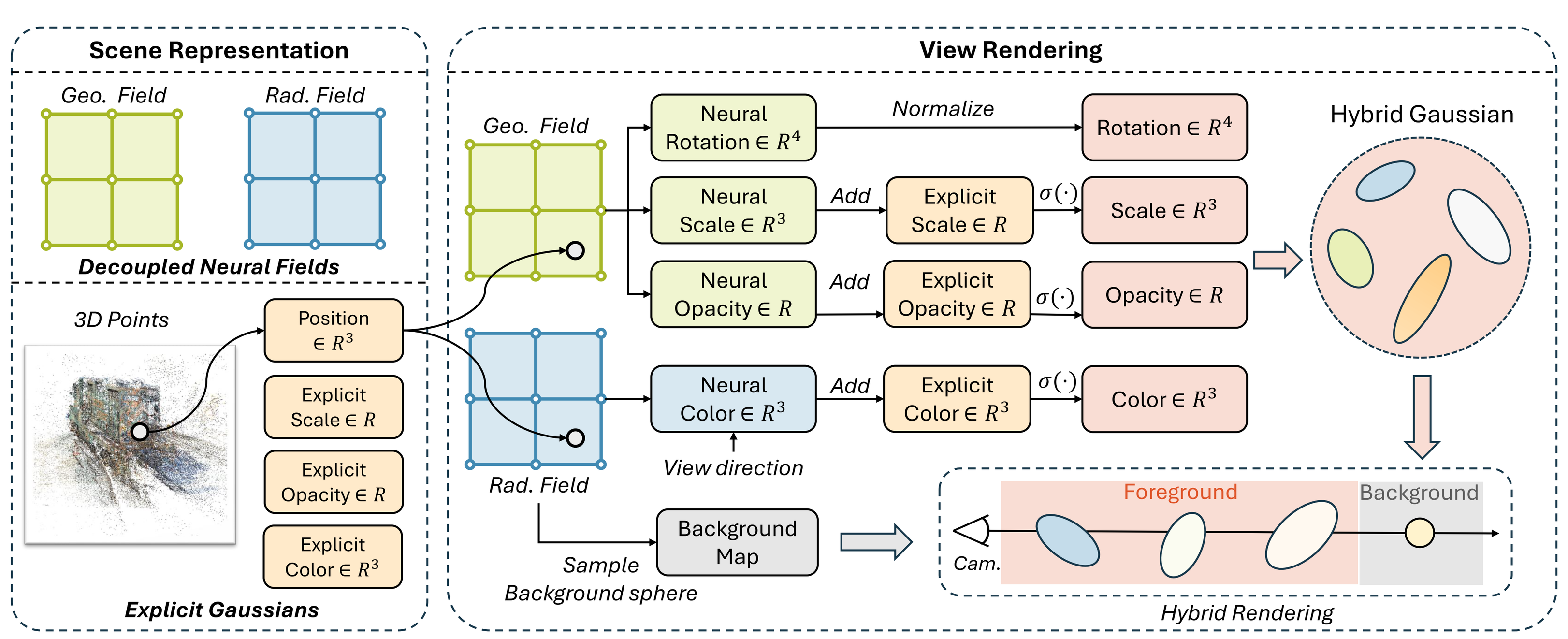
- NeRF-based methods (e.g. MipNeRF-360) struggle with fine details and **slow rendering speeds**.
- 3DGS face challenges of **large model sizes**.
- A naive combination of neural fields and 3DGS leads to **loss of high-frequency information.**

Hybrid Radiance Fields (HyRF)

TLDR: Radiance fields with **SOTA quality**, **NeRF size** and **3DGS speed.**

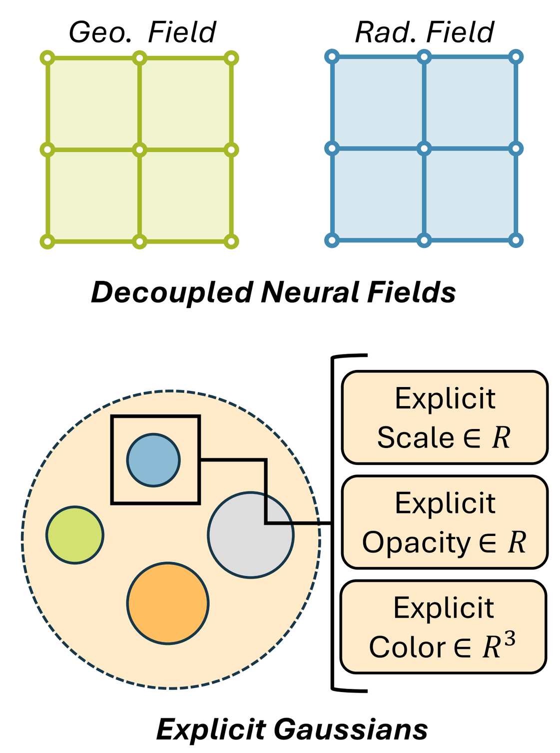
- A novel integration of **neural fields** with **explicit compact Gaussians**, preserving high-frequency details while minimizing memory overhead.
- A **dual-field architecture** that improves the modeling of Gaussian properties by disentangling geometry and view-dependent effects.
- A **hybrid rendering strategy** that reduces computational overhead and improves rendering quality for backgrounds.

Method Overview



By synergistically combining neural fields, explicit Gaussians, and neural background map, we achieve competitive or superior performance in both visual quality and model compactness, while maintaining real-time rendering.

Scene Representation

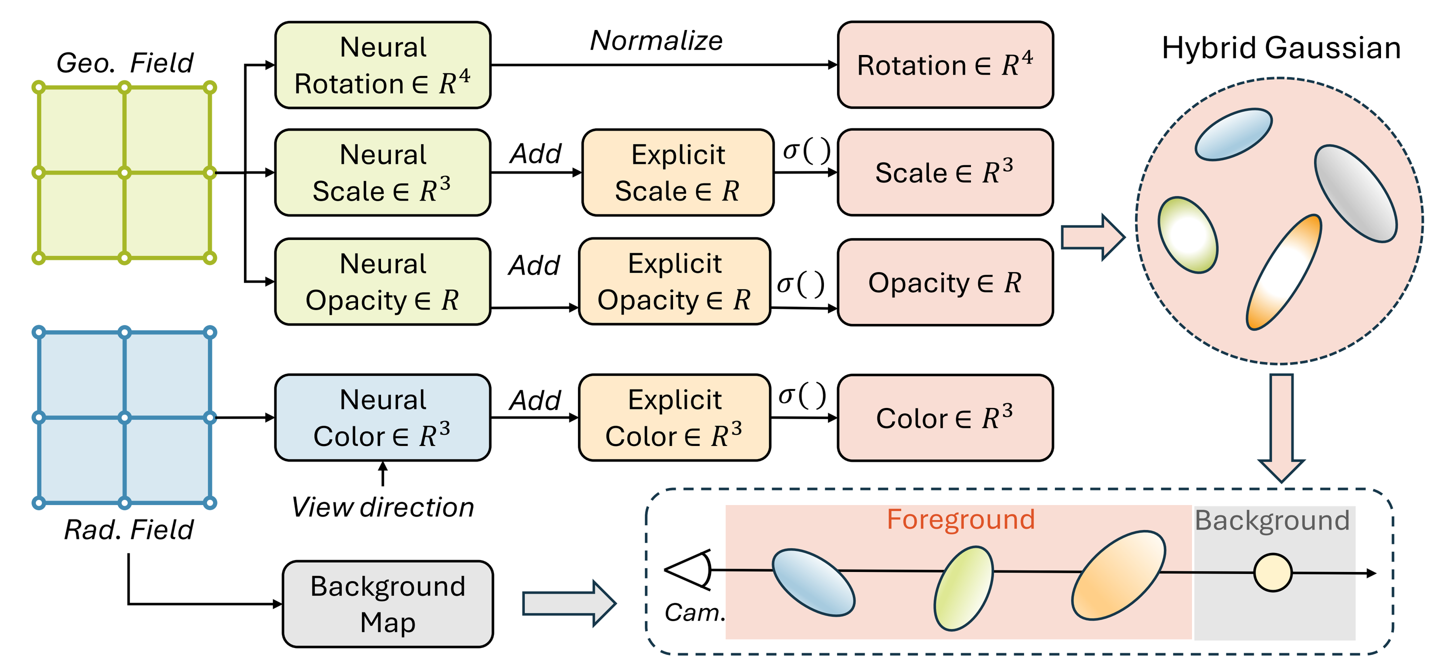


Our method represents a scene using:

(a) Decoupled Neural Fields. Predicting all Gaussian properties through a single neural field fails to achieve satisfactory performance. To address this issue, we predicts geometry properties (scale, opacity and rotation) and appearance property (view-dependent color) with two separate neural fields.

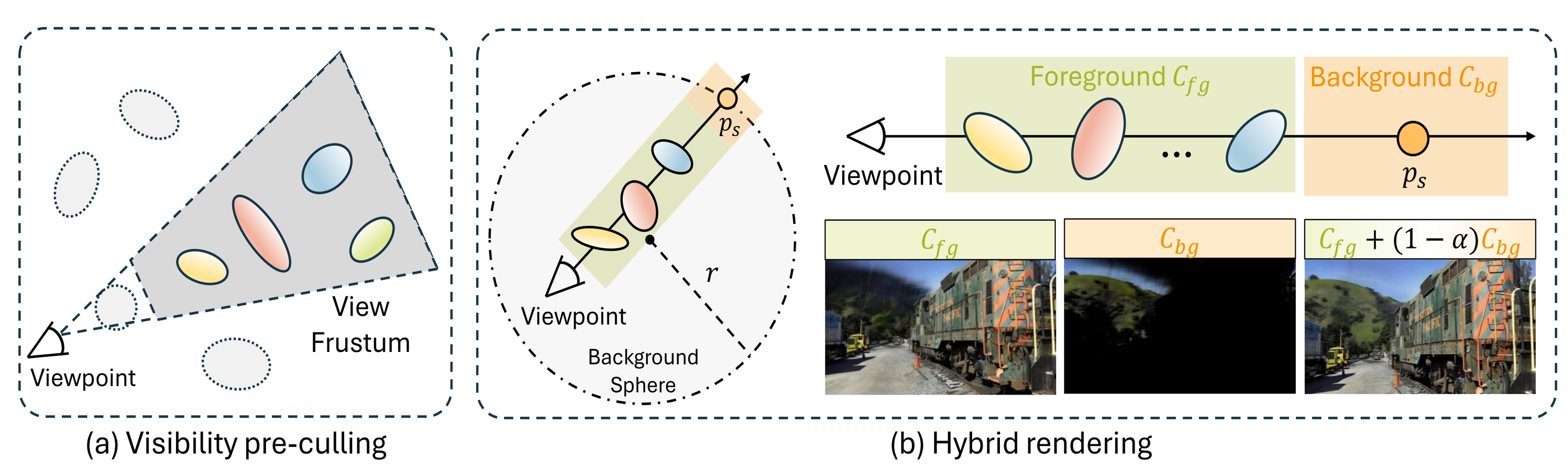
(b) Explicit Gaussians. Each Gaussian also holds an explicit set of 3D Gaussians each holds only 8 parameters, including 3D positions, 3D diffuse color, 1D isotropic scale and 1D opacity.

Hybrid Gaussians



Grid-based neural fields often overlook high-frequency scene components such as intrinsic structures. We address this problem by aggregating the predicted properties from neural fields with explicit properties stored in each Gaussian. These Gaussian properties are then integrated into the efficient 3DGS rasterizer.

Rendering Pipeline



(a) Visibility Pre-Culling. We first determine whether each Gaussian lies within the current view frustum before applying neural field decoding.

(b) Background Map. For each camera ray, we:

1. Compute its intersection point with a background sphere,
2. Sample the radiance field at the intersection point.
3. Composite the foreground and background colors using alpha blending.

Comparison Results

Dataset	Mip-NeRF360 [3]					Tanks&Temples [16]					Deep Blending [13]				
	PSNR↑	SSIM↑	LPIPS↓	FPS↑	Size(MB)↓	PSNR↑	SSIM↑	LPIPS↓	FPS↑	Size(MB)↓	PSNR↑	SSIM↑	LPIPS↓	FPS↑	Size(MB)↓
Plenoxels [8]	23.08	0.626	0.463	6.79	2150	21.08	0.719	0.379	13.0	2355	23.06	0.795	0.510	11.2	2764
Instant-NGP [27]	25.59	0.699	0.331	9.43	48	21.92	0.745	0.305	14.4	48	24.96	0.817	0.390	2.79	48
M-NeRF360 [3]	27.69	0.792	0.237	0.06	**8.6**	22.22	0.759	0.257	0.14	**8.6**	29.40	0.901	0.245	0.09	**8.6**
3DGS [15]	27.21	0.815	0.214	**117**	734	23.14	0.841	0.183	**130**	411	29.41	0.903	0.243	112	676
Scaffold-GS [23]	27.39	0.806	0.252	86	244	23.96	**0.853**	0.177	94	86.5	30.21	0.906	0.254	**120**	66
Ours	**27.78**	**0.816**	**0.211**	102	49	**24.02**	0.844	**0.176**	106	39	**30.37**	**0.910**	**0.241**	114	34

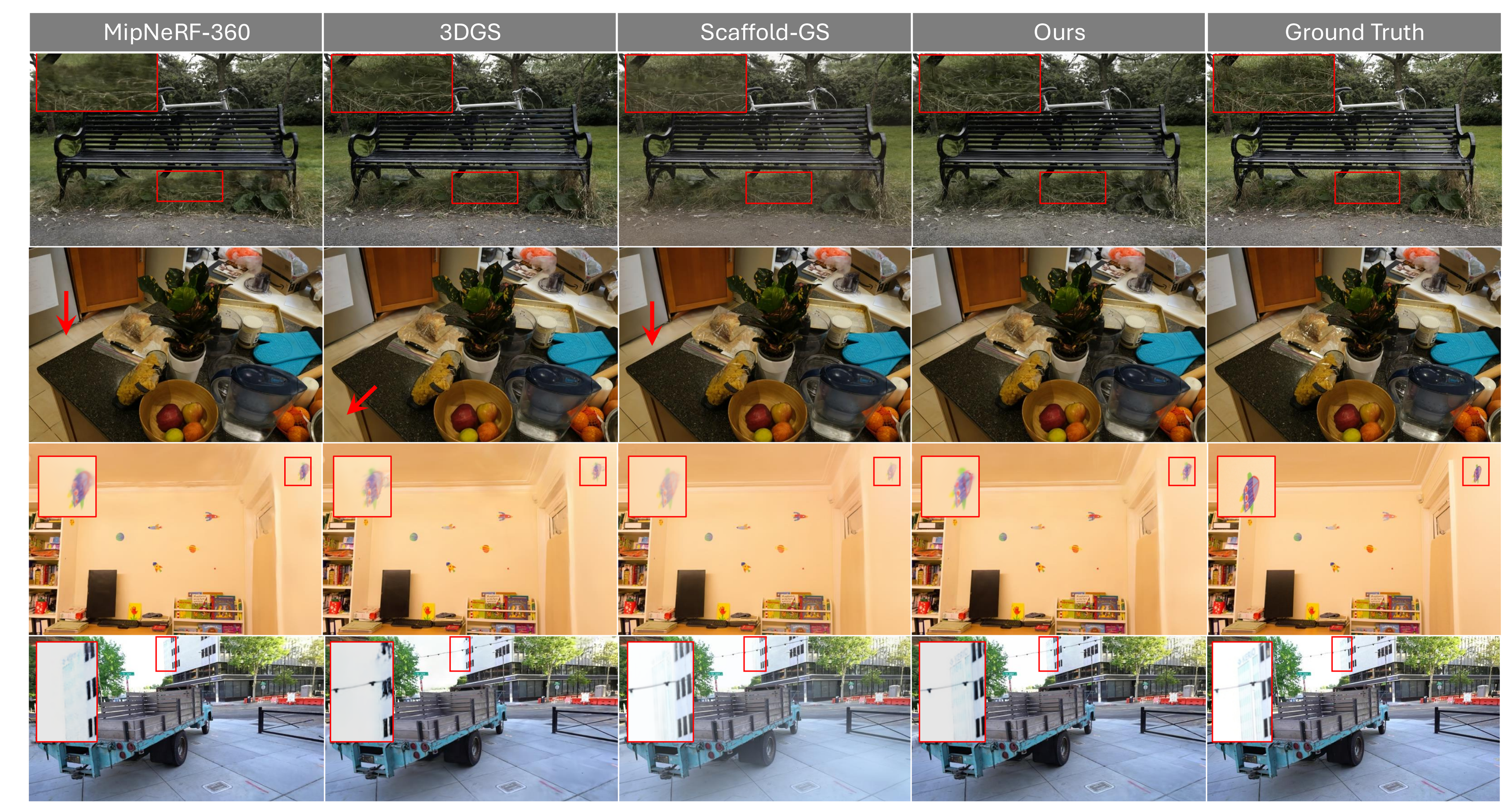
Quantitative evaluation on standard dataset

Dataset	Mill19 [43]					Urbanscene3D [20]				
	PSNR↑	SSIM↑	LPIPS↓	FPS↑	Size(MB)↓	PSNR↑	SSIM↑	LPIPS↓	FPS↑	Size(MB)↓
MegaNeRF [43]	22.50	0.55	0.510	<0.01	32	23.84	0.699	0.440	<0.01	32
SwitchNeRF [56]	22.93	0.571	0.485	<0.01	**17**	24.54	0.725	0.418	<0.01	**17**
3DGS [15]	22.41	0.695	0.348	**81**	1566	21.41	0.763	0.287	**84**	935
Scaffold-GS [23]	22.33	0.658	0.339	36	560	20.25	0.729	0.295	34	435
Ours	**23.52**	**0.709**	**0.319**	75	215	**24.68**	**0.791**	**0.272**	77	202

Quantitative evaluation on large-scale dataset

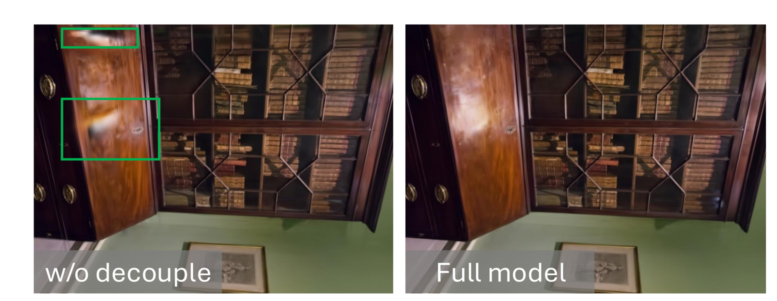
	PSNR↑	Size(MB)↓
MipNeRF [2]	32.63	**2.4**
Instant-NGP [49]	33.18	12
3DGS [15]	33.32	53
Scaffold-GS [23]	33.68	23
Ours	**33.72**	13

Synthetic dataset

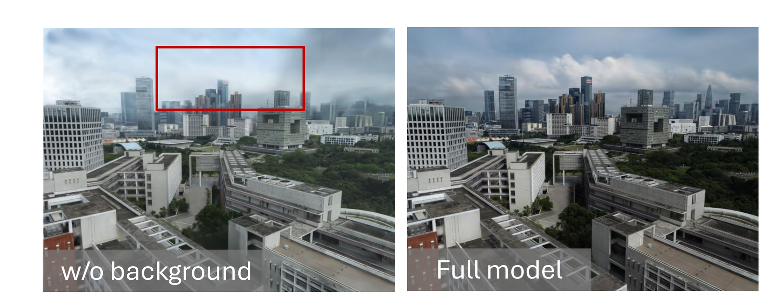


Qualitative comparisons of our method against previous approaches

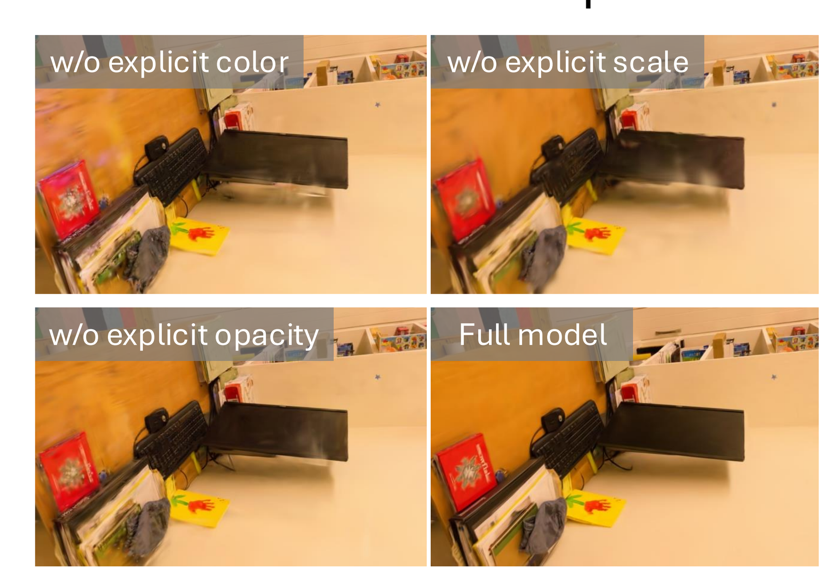
Model Ablation & Analysis



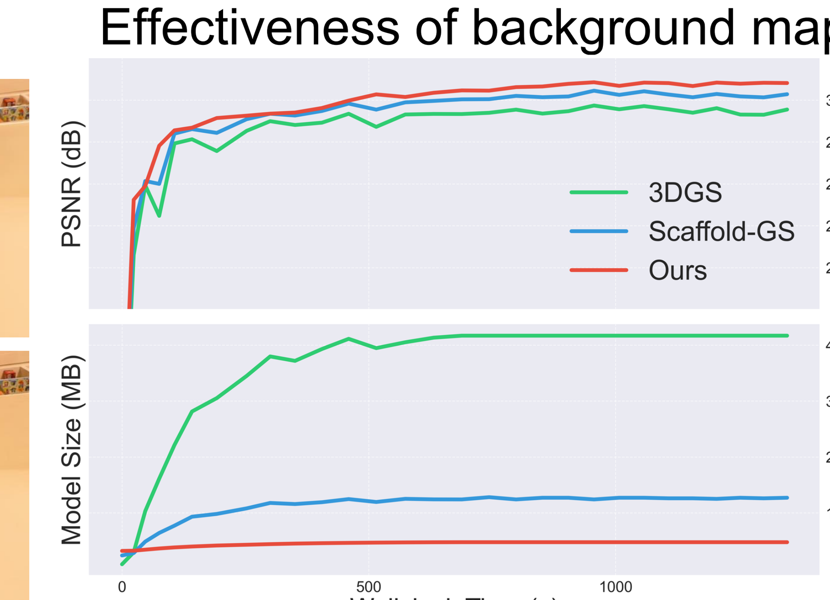
Effectiveness of decouple fields



Effectiveness of background map



Effectiveness of explicit Gaussians



Comparison of PSNR and model size during training