Arbitrary-Scale 3D Gaussian Super-Resolution

Huimin Zeng[1*], Yue Bai[1], Yun Fu[1,2]

[1]Department of Electrical and Computer Engineering, Northeastern University
[2]Khoury College of Computer Science, Northeastern University

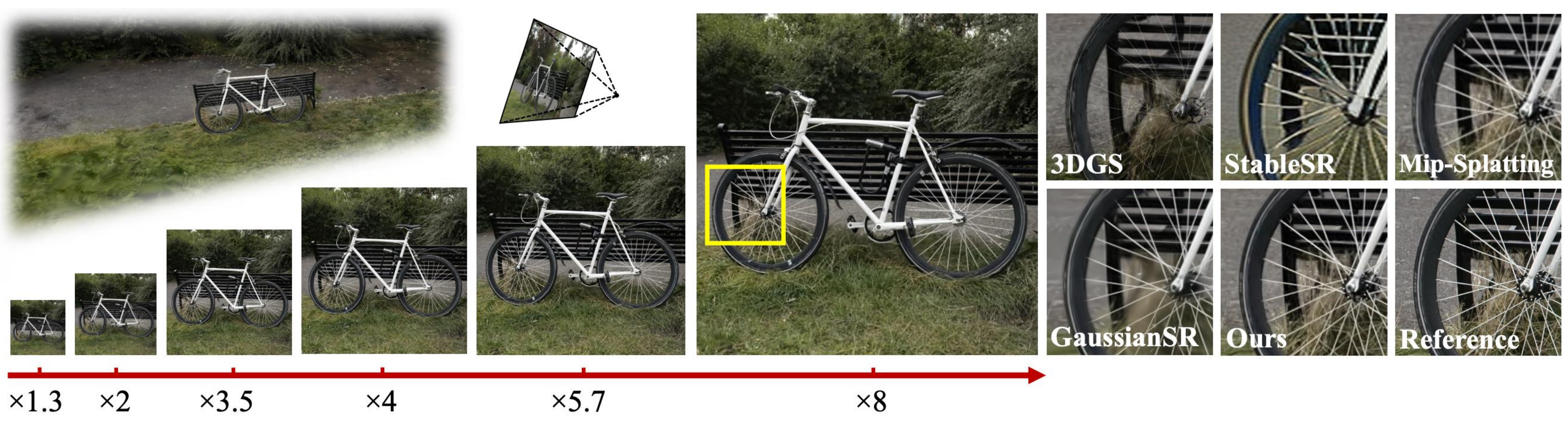
Northeastern University

Smile lab — Synergetic Media Learning Lab

Introduction

Background

- HRNVS needs flexibility to adjust accuracy based on available resource.
- Existing 3DGS SR method handle HRNVS a fixed integer scale factors.
- Cascaded solutions complicates the framework and slows rendering.

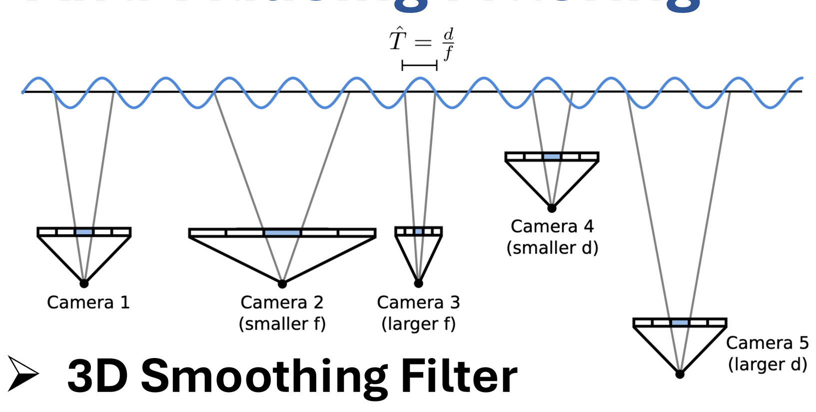


- **Favorable solution**
  - Conduct 3D super-resolution of arbitrary scales using a single unified model.

Challenges

- Anti-aliasing NVS at various scale factors.
- Constraining fine details of HR results without ground truth.
- Maintaining structural consistency across multiple scale factors.

Anti-Aliasing Filtering



Nyquist-Shannon sampling theorem:

$$\hat{\nu} \geq 2\nu$$

$$\hat{r}_i = \max\left(\left\{\mathbb{I}_k\left(G_i^{3D}\right)\cdot\frac{f_k}{d_k}\right\}_{k=1}^{K}\right)$$

- **3D Smoothing Filter**

$$G_i^{3D}(\boldsymbol{x})_{mip} = \sqrt{\frac{|\boldsymbol{\Sigma}_i|}{|\boldsymbol{\Sigma}_i+\frac{\gamma}{\hat{r}_i}\cdot\mathbf{I}|}}\, e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_i)^\top\left(\boldsymbol{\Sigma}_i+\frac{\gamma}{\hat{r}_i}\cdot\mathbf{I}\right)^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_i)}$$

- **2D Mip Filter**

$$G_i^{2D}(\hat{\boldsymbol{x}})_{mip} = \sqrt{\frac{|\hat{\boldsymbol{\Sigma}}_i|}{|\hat{\boldsymbol{\Sigma}}_i+\varepsilon\cdot\mathbf{I}|}}\, e^{-\frac{1}{2}(\hat{\boldsymbol{x}}-\hat{\boldsymbol{\mu}}_i)^\top\left(\hat{\boldsymbol{\Sigma}}_i+\varepsilon\cdot\mathbf{I}\right)^{-1}(\hat{\boldsymbol{x}}-\hat{\boldsymbol{\mu}}_i)}$$

Method

Scale-Aware Rendering for anti-aliasing rendering

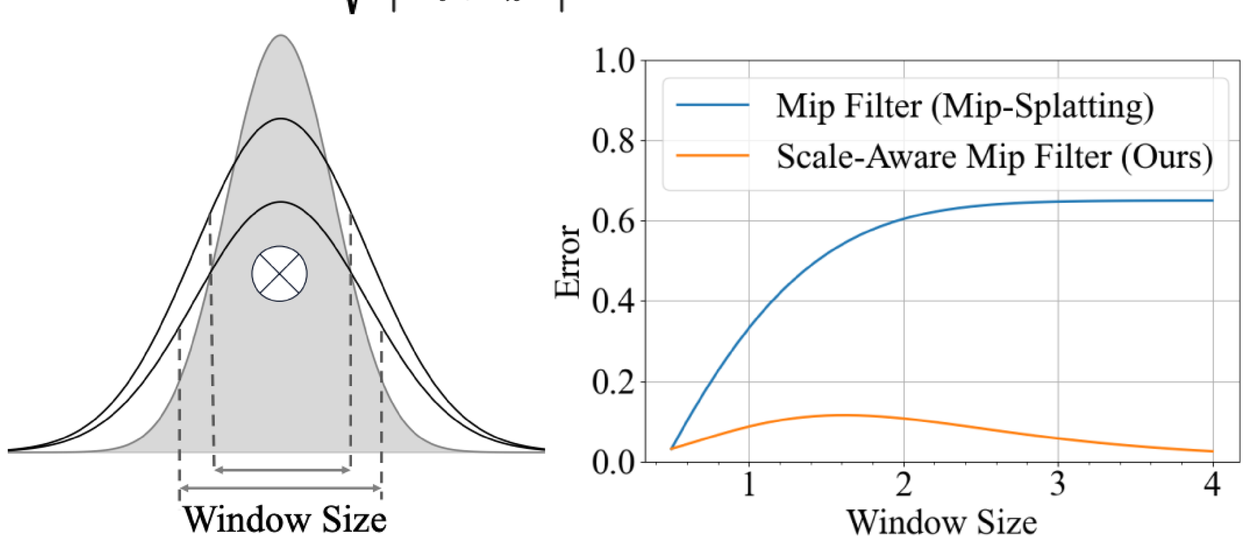
- 3D Scale-Aware Smoothing Filter

$$\hat{r}_i(s) = \max\left(\left\{\mathbb{I}_k\left(G_i^{3D}\right)\cdot\frac{f_k\cdot s_k}{d_k}\right\}_{k=1}^{K}\right)$$

- 2D Scale-Aware Mip Filter

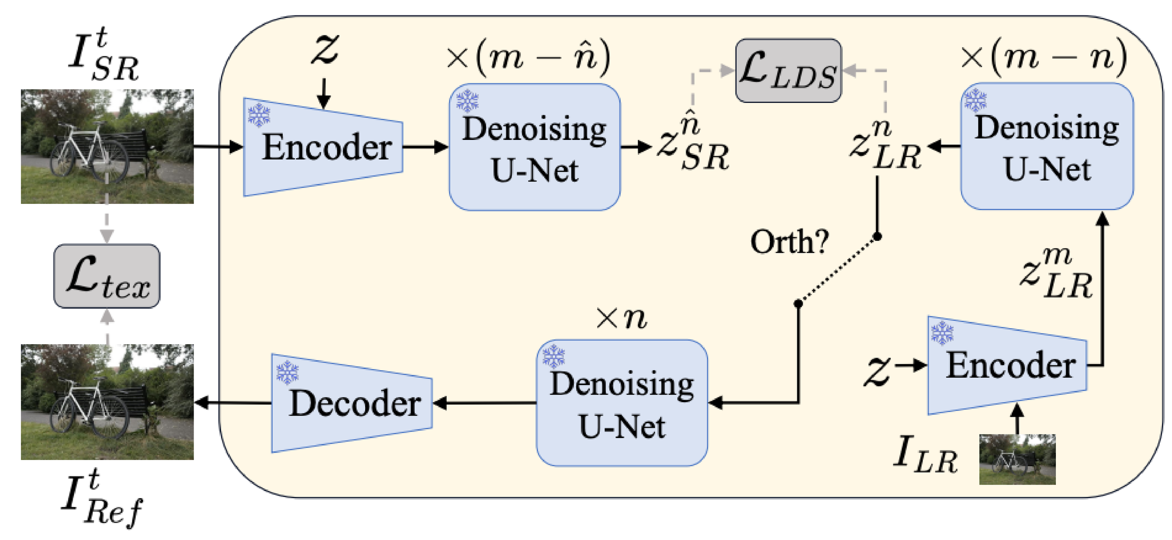
$$\varepsilon_k = \frac{\varepsilon}{s_k}$$

$$G_i^{2D}(\hat{\boldsymbol{x}})_{mip} = \sqrt{\frac{|\hat{\boldsymbol{\Sigma}}_i|}{|\hat{\boldsymbol{\Sigma}}_i+\varepsilon_k\cdot\mathbf{I}|}}\, e^{-\frac{1}{2}(\hat{\boldsymbol{x}}-\hat{\boldsymbol{\mu}}_i)^\top\left(\hat{\boldsymbol{\Sigma}}_i+\varepsilon_k\cdot\mathbf{I}\right)^{-1}(\hat{\boldsymbol{x}}-\hat{\boldsymbol{\mu}}_i)}$$



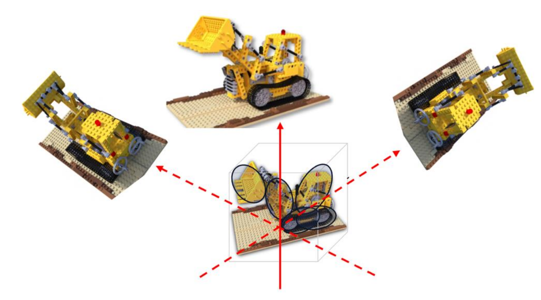
Generative Prior-guided Optimization for constraining details without GT

- Generative Latent Distillation



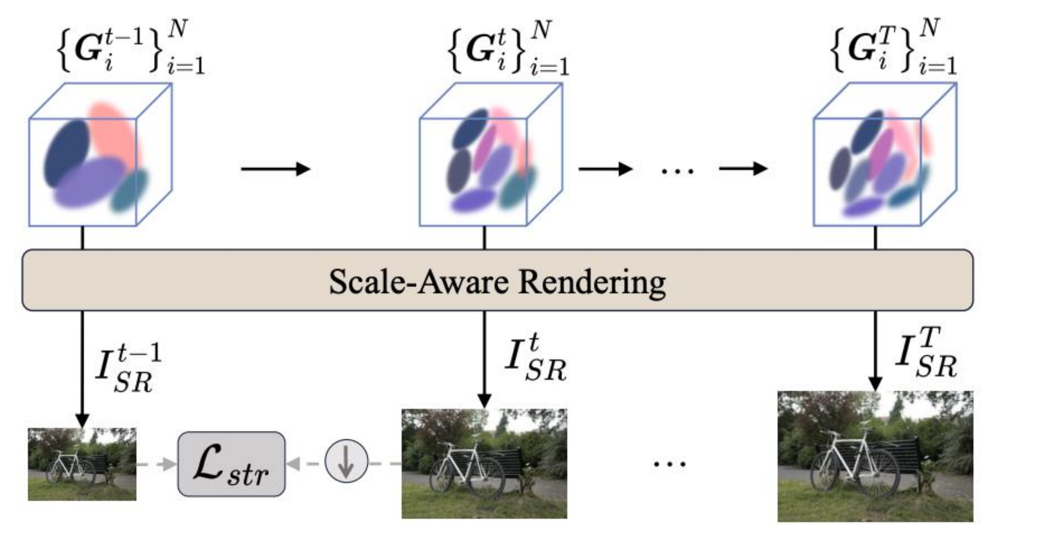
$$\nabla_\theta\mathcal{L}_{\mathrm{LDS}}(\theta) = \mathbb{E}_{\hat{n}}\left[w(\hat{n})\cdot\left(\epsilon_\phi\left(z_{SR}^{\hat{n}}; I_{SR}^t, \hat{n}\right) - \epsilon_\phi\left(z_{LR}^{n}; I_{LR}, n\right)\right)\frac{\partial I_{SR}^t}{\partial\theta}\right]$$

- Orthogonal Reference Refinement



$$\mathcal{L}_{tex} = \mathbb{I}_{ortho}\cdot\|I_{SR}^t - I_{Ref}^t\|^2$$

Progressive Super-Resolving for preserving structural consistency



$$\mathcal{L}_{str} = (1-\lambda)\,\mathcal{L}_{\mathrm{MSE}}\left(\mathcal{D}\left(I_{SR}^t, \frac{s^i}{s^{i-1}}\right), I_{SR}^{t-1}\right) + \lambda\,\mathcal{L}_{\mathrm{D\text{-}SSIM}}\left(\mathcal{D}\left(I_{SR}^t, \frac{s^i}{s^{i-1}}\right), I_{SR}^{t-1}\right)$$

Loss function

$$\mathcal{L} = \lambda_1\mathcal{L}_{LDS} + \lambda_2\mathcal{L}_{tex} + \lambda_3\mathcal{L}_{str}$$

Results

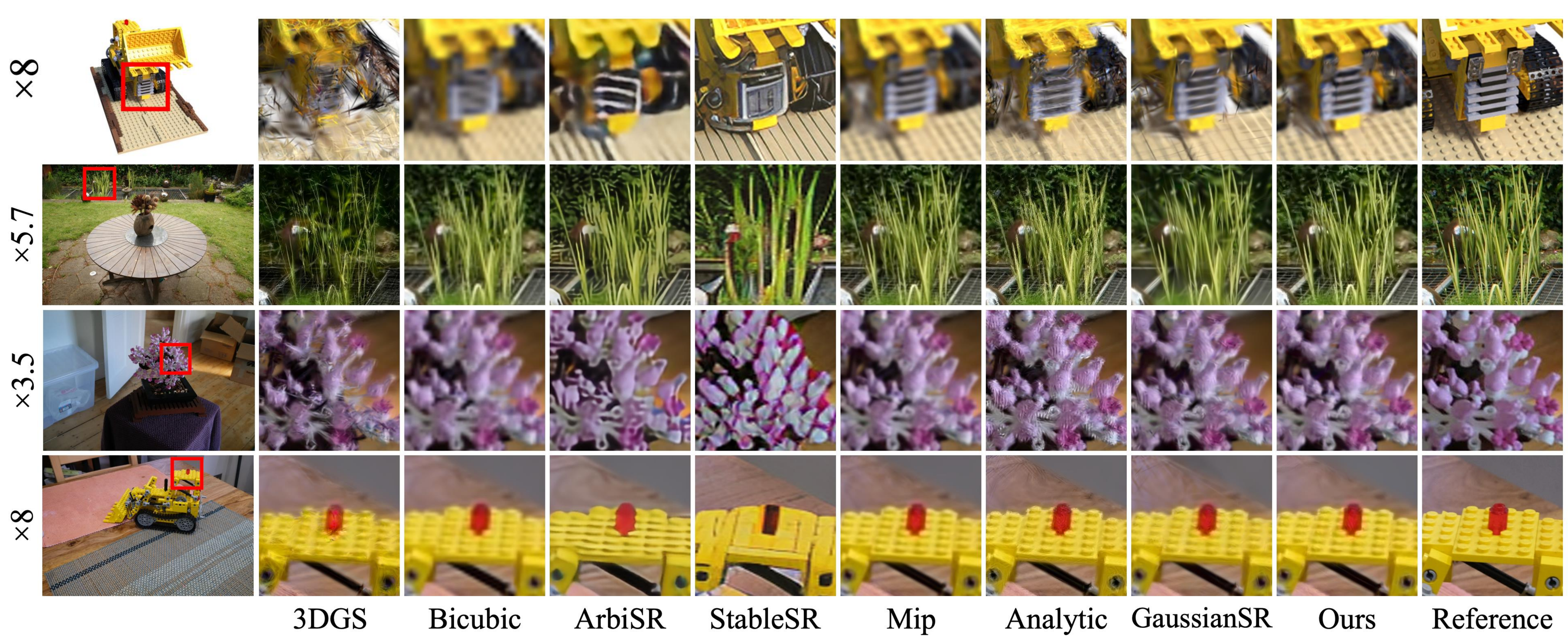
Quantitative Results (partial)

	Method	×2			×4			×8			×3.5			×5.7		
		PSNR↑	SSIM↑	FID↓	PSNR↑	SSIM↑	FID↓	PSNR↑	SSIM↑	FID↓	PSNR↑	SSIM↑	FID↓	PSNR↑	SSIM↑	FID↓
Blender	3DGS	19.66	0.834	129.567	17.84	0.789	208.166	16.86	0.797	244.212	18.10	0.792	198.147	17.28	0.790	229.619
	Bicubic	20.36	0.853	132.513	19.67	0.831	178.170	19.24	0.831	182.420	19.76	0.833	176.546	19.46	0.829	181.065
	ArbiSR	19.34	0.827	152.266	18.23	0.803	185.134	17.98	0.817	174.873	18.33	0.803	185.430	18.08	0.808	179.436
	StableSR	18.63	0.788	120.385	18.25	0.785	95.086	17.98	0.797	94.957	18.31	0.785	95.720	18.12	0.790	94.853
	Mip-Splatting	23.33	0.896	78.907	22.25	0.870	109.435	21.57	0.858	113.124	22.40	0.874	108.101	21.92	0.862	111.875
	Analytic-Splatting	25.23	0.915	103.452	23.57	0.873	141.302	22.51	0.847	146.244	23.82	0.881	139.486	23.02	0.857	144.521
	GaussianSR	23.25	0.885	94.017	23.03	0.868	118.021	22.37	0.856	117.618	23.13	0.871	117.458	22.73	0.861	118.465
	Ours	25.60	0.925	66.515	24.32	0.899	86.270	23.40	0.879	87.984	24.53	0.904	85.788	23.87	0.888	87.429

Higher Scale Factors (FID)

Method	Blender		Mip-NeRF360	
	×12	×16	×12	×16
3DGS	256.287	262.810	92.282	95.412
Bicubic	182.535	182.538	47.473	47.483
ArbiSR	171.130	169.713	54.522	54.416
StableSR	94.952	94.936	106.997	106.996
Mip-Splatting	113.081	113.060	53.736	53.732
Analytic-Splatting	147.154	147.488	35.363	32.368
GaussianSR	113.725	112.556	110.681	107.030
Ours	87.842	87.797	38.828	38.824

Qualitative Results



Efficiency Analysis

Method	Rendering (ms)	Throughput (FPS)	Storage Size (GB)	Training (min)	Memory (MB)
3DGS	13	74	0.99	10	718
Bicubic	34	29	-	-	-
ArbiSR	3225	0.31	-	-	-
StableSR	10890	0.13	-	-	-
Mip	19	52	0.99	12	858
Analytic	33	30	1.20	29	822
GaussianSR	8	126	0.56	256	8274
Ours	12	85	0.79	57	7160

Ablation Studies