

Supplementary Material:

MFNeRF: Memory Efficient NeRF with Mixed-Feature Hash Table

Yongjae Lee
Arizona State University
Tempe, AZ 85281
ylee298@asu.edu

Li Yang
University of North Carolina at Charlotte
Charlotte, NC 28223
lyang50@uncc.edu

Deliang Fan
Arizona State University
Tempe, AZ 85281
dfan@asu.edu

A. Experiment detail

We present detailed experimental results for each scene. Our experiments are conducted on the Synthetic NeRF [3], MipNeRF360 [1], Tanks&Temples [2], and Blended-MVS [4] datasets. In general, we use the hyperparameter settings listed in Tab. 5 as the default across these datasets and their scenes. Scene-specific settings are detailed in the subsequent paragraphs. For our experiment, we use the python implementation of InstantNGP, named ‘ngp-pl’¹.

Hyper-parameter	Description	Value
F	Number of features per hash table entry	2
L	Number of feature grids	16
M	Number of mixed-feature hash tables	8
N_{\min}	Coarsest resolution of feature grid	16
N_{\max}	Finest resolution of feature grid	1024
lr	Learning rate	$2 \cdot 10^{-2}$
decay	Learning rate decay factor	$1 \cdot 10^{-2}$
iters	Number of optimization steps	20k
scale	Scene scaling factor	1
downsample	Image downsampling ratio	1
rgb_{ch}	The number of hidden layer channels of color MLP	128
seed	Random seed	1337
batch	Batch size of training rays	16384

Table 5. Common hyperparameter settings across the datasets

The Synthetic NeRF dataset comprises eight scenes: Chair, Drums, Ficus, Hotdog, Lego, Materials, Mic, and Ship. Each scene provides 400 images with a resolution of 800×800 . These images are divided as follows: 100 for train, 100 for evaluation, and 200 for test. As the evalu-

ation dataset is generally not utilized, we have omitted it, focusing solely on the training and testing sets. We employ a consistent set of hyperparameters across all eight scenes, as listed in Tab. 5. The results are presented in Tab. 7 and Fig. 6.

The MipNeRF360 dataset comprises seven scenes with backgrounds. These scenes are Bicycle (194 images), Bonsai (292 images), Counter (240 images), Garden (185 images), Kitchen (279 images), Room (311 images), and Stump (125 images). Since the scenes have different resolutions, we use a different downsampling ratio for each scene to fit within a resolution range between 1k and 1.6k. We detail the scene-specific settings in Tab. 6. Settings not listed in this table are consistent with those in Tab. 5. The results can be found in Tab. 8 and Fig. 7.

Scene	Original resolution	Scale	Downsample	Batch
Bicycle	4946×3286	16	0.25	8192
Bonsai	3118×2078	16	0.5	8192
Counter	3115×2076	16	0.5	8192
Garden	5187×3361	16	0.25	8192
Kitchen	3115×2078	4	0.5	8192
Room	3114×2075	4	0.5	8192
Stump	4978×3300	64	0.25	8192

Table 6. Scene-specific hyperparameter settings for MipNeRF360 [1]

The Tanks&Temples dataset comprises five scenes with backgrounds removed (whited out). These scenes are Barn (336/48 images for train/test), Caterpillar (322/46 images), Family (133/19 images), Ignatius (230/33 images), and Truck (218/32 images). All images has the same resolution of 1920×1080 . We used the same setting to Synthetic NeRF dataset as listed in Tab. 5, except for $N_{\max} = 512$. The results can be found in Tab. 9 and Fig. 8.

The BlendedMVS dataset comprises four scenes with

¹https://github.com/kweal23/ngp_pl

removed backgrounds. These scenes are Character (96/14 images for train/test), Fountain (79/12 images), Jade (50/8 images), and Statues (51/8 images). All images has the same resolution of 768×576 . We used the same setting to Synthetic NeRF dataset as listed in Tab. 5, except Character scene. For Character scene, we use learning rate of $8e-3$ and decay of $2e-2$. The results can be found in Tab. 10 and Fig. 9.



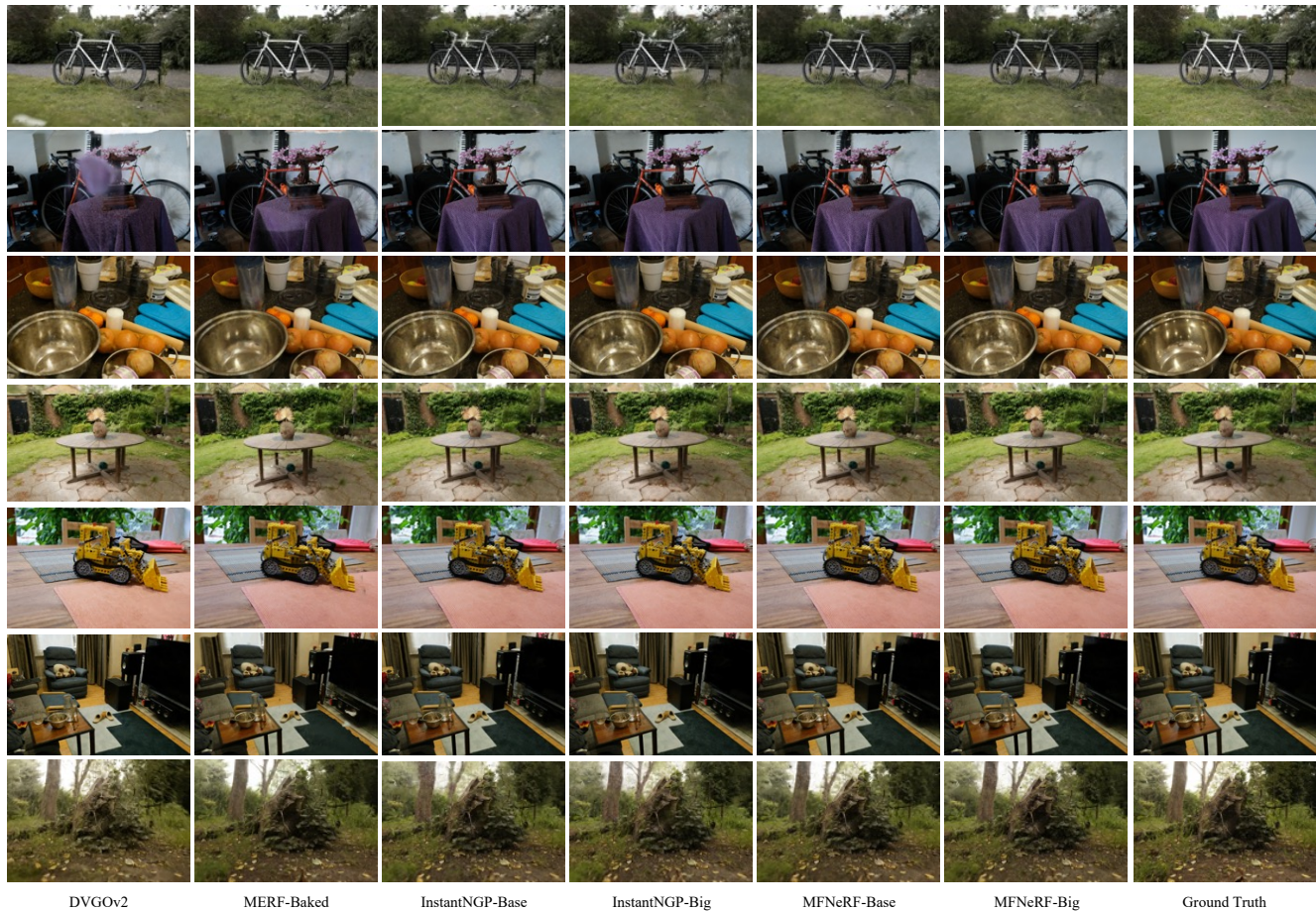
Figure 6. Qualitative result on Synthetic NeRF [3]

Metrics	Methods	Chair	Drums	Ficus	Hotdog	Lego	Materials	Mic	Ship	Avg.
PSNR	TensoRF	35.77	25.99	34.10	37.50	36.54	30.11	34.89	30.70	33.20
	DVGOv1	34.11	25.50	32.79	36.71	34.45	29.57	33.16	29.15	31.93
	DVGOv2	34.06	25.39	32.57	36.71	34.67	29.58	33.16	29.05	31.90
	CCNeRF	34.35	24.74	30.04	35.95	33.68	29.00	33.56	28.46	31.22
	FactorFields	35.59	25.72	34.33	37.05	35.69	29.40	36.69	30.29	33.09
	FastLearning	33.39	25.38	31.95	34.66	33.74	28.17	34.22	23.74	30.66
	InstantNGP-Base	35.48	25.64	33.86	37.37	35.88	29.47	35.69	30.44	32.98
	InstantNGP-Big	35.55	25.76	34.06	37.32	35.97	29.55	35.49	30.40	33.01
	MFNeRF-Base	35.37	25.88	34.06	37.40	35.86	29.56	35.72	30.36	33.03
MFNeRF-Big	35.56	25.79	34.16	37.41	36.04	29.47	35.82	30.58	33.11	
SSIM	TensoRF	0.985	0.937	0.983	0.983	0.983	0.952	0.988	0.894	0.963
	DVGOv1	0.977	0.931	0.979	0.980	0.975	0.951	0.983	0.879	0.957
	DVGOv2	0.976	0.929	0.977	0.980	0.976	0.950	0.983	0.878	0.956
	CCNeRF	0.977	0.918	0.962	0.975	0.969	0.935	0.984	0.855	0.947
	FactorFields	0.985	0.938	0.983	0.982	0.979	0.946	0.991	0.885	0.961
	FastLearning	0.977	0.931	0.977	0.973	0.973	0.935	0.985	0.842	0.949
	InstantNGP-Base	0.984	0.931	0.981	0.980	0.980	0.945	0.989	0.890	0.960
	InstantNGP-Big	0.982	0.932	0.982	0.981	0.981	0.946	0.989	0.889	0.960
	MFNeRF-Base	0.980	0.932	0.982	0.980	0.979	0.945	0.989	0.889	0.960
MFNeRF-Big	0.985	0.932	0.982	0.981	0.980	0.946	0.990	0.889	0.961	
LPIPS (VGG)	TensoRF	0.022	0.071	0.022	0.030	0.018	0.058	0.015	0.139	0.047
	DVGOv1	0.027	0.077	0.024	0.034	0.028	0.058	0.017	0.161	0.053
	DVGOv2	0.028	0.080	0.025	0.034	0.027	0.059	0.018	0.161	0.054
	CCNeRF	0.036	0.110	0.054	0.057	0.036	0.080	0.030	0.190	0.074
	FactorFields	0.019	0.068	0.020	0.032	0.024	0.067	0.012	0.153	0.049
	FastLearning	0.033	0.082	0.031	0.048	0.033	0.082	0.021	0.204	0.067
	InstantNGP-Base	0.022	0.078	0.026	0.036	0.021	0.067	0.017	0.131	0.050
	InstantNGP-Big	0.021	0.077	0.025	0.035	0.020	0.065	0.017	0.124	0.048
	MFNeRF-Base	0.024	0.078	0.025	0.035	0.023	0.069	0.017	0.135	0.051
MFNeRF-Big	0.021	0.077	0.025	0.034	0.020	0.067	0.016	0.128	0.048	
Training Time	TensoRF	654.13	651.47	822.97	834.85	746.50	1174.84	563.15	1009.09	807.13
	DVGOv1	537.86	524.40	565.63	659.03	564.34	654.65	482.72	721.90	588.82
	DVGOv2	197.04	195.08	209.19	257.88	221.87	263.90	171.81	333.70	231.31
	CCNeRF	866.86	946.49	1011.66	946.94	944.16	1317.62	751.28	1120.82	988.23
	FactorFields	617.27	544.23	571.26	649.41	630.36	798.91	513.02	705.99	628.81
	FastLearning	171.02	168.18	166.96	193.36	188.97	179.53	169.27	251.99	186.16
	InstantNGP-Base	204.45	194.03	164.00	255.86	226.90	221.35	163.97	380.94	226.44
	InstantNGP-Big	298.34	287.26	240.69	375.75	335.96	323.64	242.78	540.10	330.57
	MFNeRF-Base	194.47	185.03	149.60	249.70	222.51	217.78	150.26	392.20	220.19
MFNeRF-Big	255.46	240.37	194.30	330.68	291.79	282.67	195.18	521.75	289.03	
FPS	TensoRF	1.20	1.07	1.13	1.09	1.09	0.88	1.20	0.77	1.05
	DVGOv1	2.41	2.15	2.31	1.73	2.10	1.70	2.44	1.40	2.03
	DVGOv2	6.68	5.57	6.92	5.22	5.74	4.95	6.33	3.17	5.57
	CCNeRF	1.02	0.76	0.93	0.99	0.94	0.87	1.02	0.67	0.90
	FactorFields	0.79	0.79	0.84	0.82	0.77	0.74	0.81	0.68	0.78
	FastLearning	32.17	28.97	29.28	26.38	27.23	16.27	33.00	9.36	25.33
	InstantNGP-Base	59.03	32.73	46.85	33.54	38.93	19.90	51.55	10.44	36.62
	InstantNGP-Big	43.61	26.64	39.37	26.51	31.23	16.78	43.18	8.02	29.42
	MFNeRF-Base	61.82	34.81	47.91	35.68	39.37	19.57	52.55	10.54	37.78
MFNeRF-Big	50.84	29.24	42.90	28.63	34.07	16.88	45.70	8.54	32.10	
Memory	TensoRF	67.09	67.14	70.00	82.19	67.99	83.77	66.12	70.06	71.79
	DVGOv1	208.92	209.55	209.33	209.03	208.11	210.31	209.52	208.68	209.18
	DVGOv2	205.06	205.68	205.48	205.18	205.25	206.43	204.47	204.82	205.30
	CCNeRF	874.90	957.24	1020.22	956.21	953.01	1326.66	760.03	1133.90	997.77
	FactorFields	20.44	20.43	20.41	20.45	20.46	20.49	20.47	20.41	20.44
	FastLearning	54.98	54.98	54.98	54.98	54.98	54.98	54.98	54.98	54.98
	InstantNGP-Base	112.63	112.63	112.63	112.63	112.63	112.63	112.63	112.63	112.63
	InstantNGP-Big	300.09	300.09	300.09	300.09	300.09	300.09	300.09	300.09	300.09
	MFNeRF-Base	74.91	74.91	74.91	74.91	74.91	74.91	74.91	74.91	74.91
MFNeRF-Big	173.63	173.63	173.63	173.63	173.63	173.63	173.63	173.63	173.63	

Table 7. Scene by scene results on the synthetic dataset (Synthetic NeRF [3])

Metrics	Methods	Bicycle	Bonsai	Counter	Garden	Kitchen	Room	Stump	Avg.
PSNR	DVGOv2	22.13	27.77	25.72	24.33	26.00	28.40	23.57	25.42
	InstantNGP-Base	23.38	29.65	26.62	25.35	29.13	29.47	23.31	26.70
	InstantNGP-Big	23.28	30.11	26.72	25.85	29.41	29.69	23.26	26.90
	MFNeRF-Base	23.59	29.70	26.61	25.14	28.94	29.08	24.07	26.73
	MFNeRF-Big	23.61	30.20	26.77	25.51	29.44	29.44	23.80	26.97
SSIM	DVGOv2	0.481	0.829	0.783	0.628	0.709	0.852	0.580	0.694
	InstantNGP-Base	0.546	0.883	0.804	0.682	0.846	0.868	0.563	0.742
	InstantNGP-Big	0.570	0.898	0.820	0.734	0.866	0.880	0.589	0.765
	MFNeRF-Base	0.545	0.877	0.790	0.653	0.836	0.863	0.578	0.735
	MFNeRF-Big	0.569	0.892	0.812	0.714	0.862	0.875	0.598	0.760
LPIPS (VGG)	DVGOv2	0.510	0.402	0.435	0.376	0.400	0.425	0.460	0.430
	InstantNGP-Base	0.490	0.304	0.376	0.312	0.251	0.360	0.466	0.366
	InstantNGP-Big	0.459	0.274	0.345	0.268	0.224	0.329	0.441	0.334
	MFNeRF-Base	0.488	0.327	0.399	0.341	0.264	0.375	0.465	0.380
	MFNeRF-Big	0.464	0.288	0.361	0.288	0.228	0.342	0.438	0.344
Training Time	DVGOv2	772.04	707.37	749.62	716.16	681.92	722.30	811.65	737.29
	InstantNGP-Base	422.65	394.50	522.34	488.76	391.78	284.82	814.09	474.13
	InstantNGP-Big	764.00	694.17	743.79	747.52	606.19	453.29	1259.91	752.69
	MFNeRF-Base	388.18	414.87	535.11	558.65	408.27	290.03	673.01	466.87
	MFNeRF-Big	620.49	634.85	698.61	736.27	569.26	456.45	988.32	672.04
FPS	DVGOv2	0.72	0.50	0.45	0.72	0.51	0.49	0.66	0.58
	InstantNGP-Base	4.55	5.08	3.72	5.00	6.29	5.06	2.58	4.61
	InstantNGP-Big	3.32	3.43	3.29	3.80	5.12	4.74	1.76	3.64
	MFNeRF-Base	4.62	4.55	3.62	5.13	6.50	4.74	3.11	4.61
	MFNeRF-Big	3.47	3.74	3.18	4.34	5.54	4.14	2.21	3.80
Memory	DVGOv2	1640.53	1640.53	1640.53	1640.53	1640.53	1640.53	1640.53	1640.53
	InstantNGP-Base	173.51	173.51	173.51	173.51	151.60	151.60	194.66	170.27
	InstantNGP-Big	443.70	443.70	443.70	443.70	399.60	399.60	482.66	436.67
	MFNeRF-Base	124.84	124.84	124.84	124.84	106.56	106.56	145.48	122.57
	MFNeRF-Big	268.84	268.84	268.84	268.84	234.56	234.56	289.48	261.99

Table 8. Scene by scene results on the unbounded dataset (MipNeRF360 [1])



DVGOv2 MERF-Baked InstantNGP-Base InstantNGP-Big MFNeRF-Base MFNeRF-Big Ground Truth

Figure 7. Qualitative result on MipNeRF360 [1].

Metrics	Methods	Barn	Caterpillar	Family	Ignatius	Truck	Avg
PSNR	InstantNGP-Base	27.78	25.98	34.36	28.11	27.54	28.75
	InstantNGP-Big	27.85	26.17	34.50	28.29	27.63	28.89
	MFNeRF-Base	27.75	26.03	34.25	28.14	27.61	28.76
	MFNeRF-Big	27.87	25.81	34.33	28.11	27.62	28.75
SSIM	InstantNGP-Base	0.866	0.913	0.964	0.946	0.912	0.920
	InstantNGP-Big	0.876	0.915	0.966	0.948	0.917	0.925
	MFNeRF-Base	0.865	0.909	0.963	0.945	0.909	0.918
	MFNeRF-Big	0.881	0.915	0.965	0.947	0.921	0.926
LPIPS (VGG)	InstantNGP-Base	0.239	0.153	0.053	0.077	0.136	0.132
	InstantNGP-Big	0.223	0.149	0.050	0.075	0.130	0.125
	MFNeRF-Base	0.250	0.163	0.057	0.078	0.142	0.138
	MFNeRF-Big	0.224	0.156	0.053	0.077	0.134	0.129
Memory	InstantNGP-Base	112.63	112.63	112.63	112.63	112.63	112.63
	InstantNGP-Big	300.09	300.09	300.09	300.09	300.09	300.09
	MFNeRF-Base	77.10	77.10	77.10	77.10	77.10	77.10
	MFNeRF-Big	192.97	192.97	192.97	192.97	192.97	192.97

Table 9. Scene by scene results on Tanks&Temples [2]

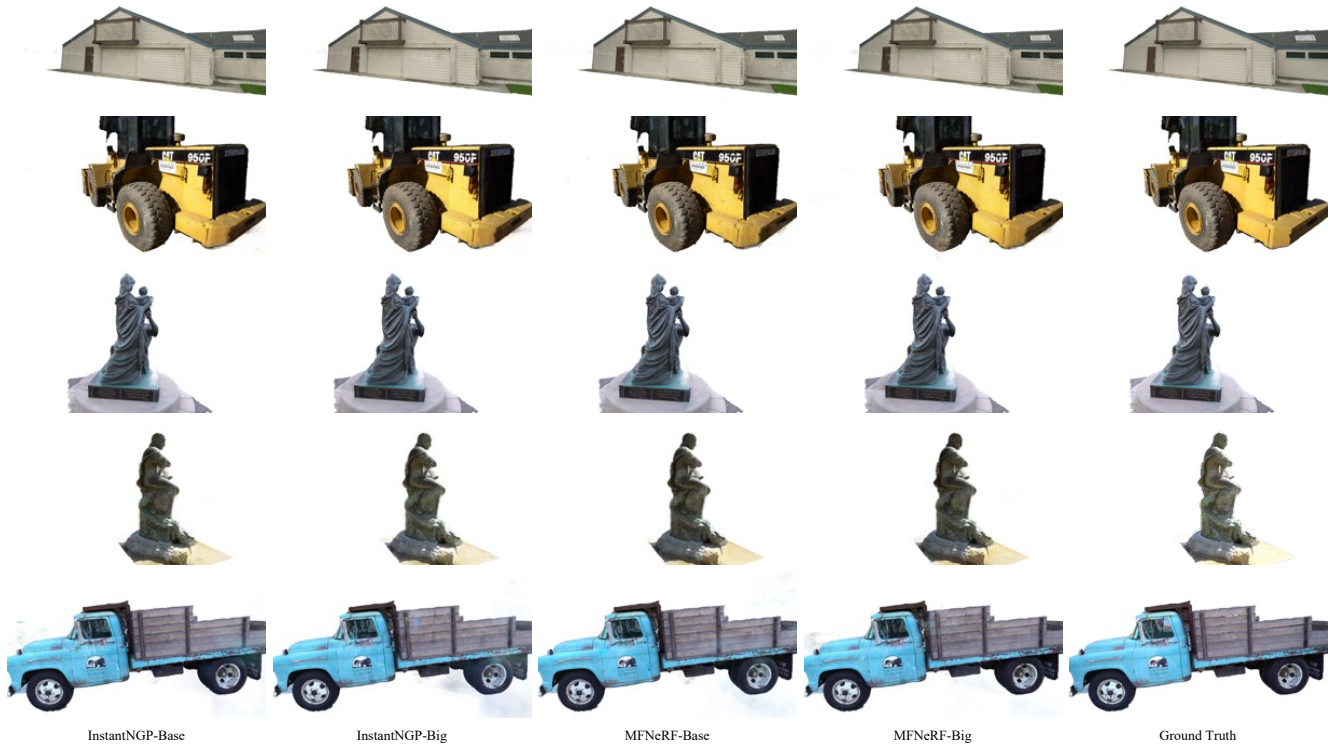


Figure 8. Qualitative result on Tanks&Temples [2].

Metrics	Methods	Character	Fountain	Jade	Statues	Avg.
PSNR	InstantNGP-Base	30.75	26.92	25.68	26.03	27.35
	InstantNGP-Big	30.78	26.67	25.69	26.66	27.45
	MFNeRF-Base	30.72	26.81	25.76	26.23	27.38
	MFNeRF-Big	30.75	26.78	25.89	25.91	27.33
SSIM	InstantNGP-Base	0.968	0.927	0.890	0.903	0.922
	InstantNGP-Big	0.968	0.924	0.884	0.910	0.922
	MFNeRF-Base	0.967	0.924	0.888	0.902	0.920
	MFNeRF-Big	0.968	0.925	0.887	0.901	0.920
LPIPS (VGG)	InstantNGP-Base	0.037	0.097	0.123	0.097	0.088
	InstantNGP-Big	0.038	0.100	0.130	0.092	0.090
	MFNeRF-Base	0.039	0.105	0.121	0.097	0.090
	MFNeRF-Big	0.038	0.101	0.130	0.098	0.092
Memory	InstantNGP-Base	112.63	112.63	112.63	112.63	112.63
	InstantNGP-Big	300.09	300.09	300.09	300.09	300.09
	MFNeRF-Base	74.91	74.91	74.91	74.91	74.91
	MFNeRF-Big	173.63	173.63	173.63	173.63	173.63

Table 10. Scene by scene results on BlendedMVS [4]

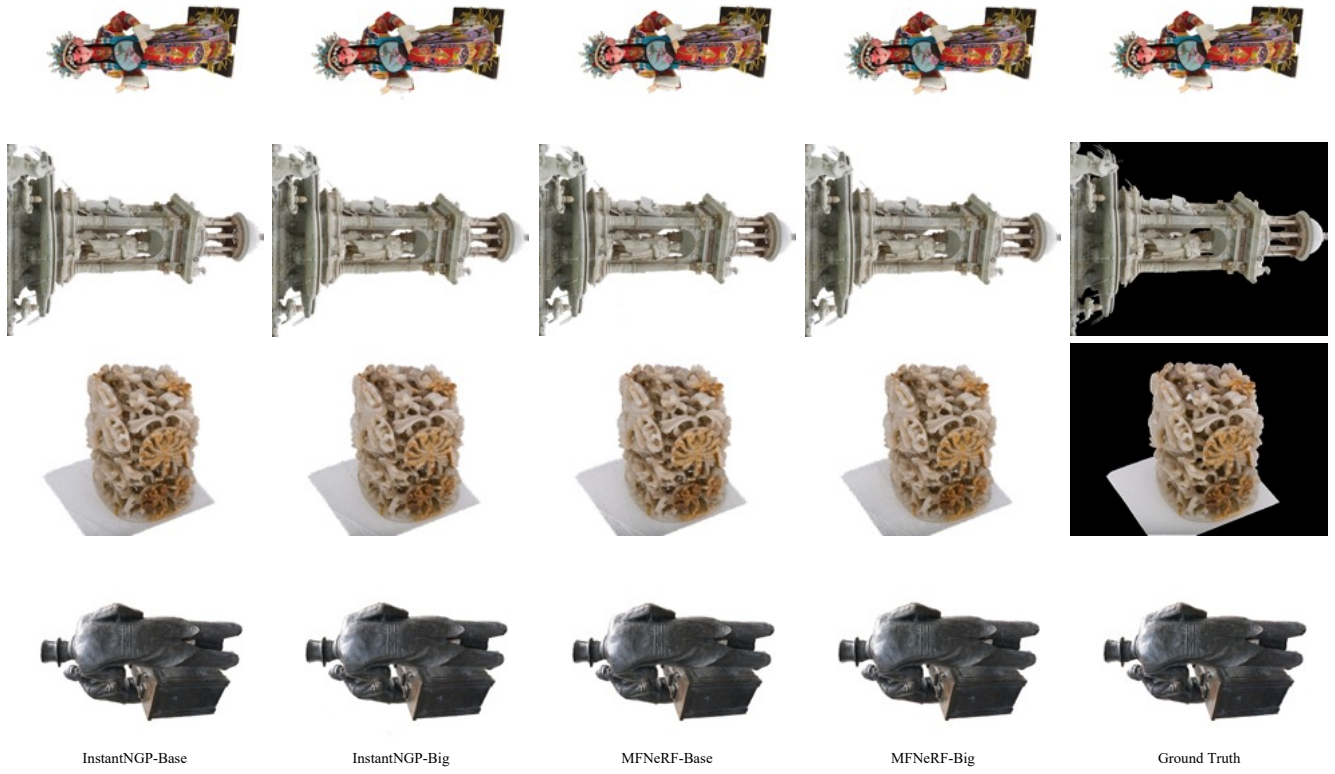


Figure 9. Qualitative result on BlendedMVS [4].

References

- [1] Jonathan T Barron, Ben Mildenhall, Dor Verbin, Pratul P Srinivasan, and Peter Hedman. Mip-NeRF 360: Unbounded Anti-Aliased Neural Radiance Fields. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5460–5469, New Orleans, LA, USA, 6 2022. IEEE. [1](#), [5](#), [6](#)
- [2] Arno Knapitsch, Jaesik Park, Qian-Yi Zhou, and Vladlen Koltun. Tanks and temples: benchmarking large-scale scene reconstruction. *ACM Trans. Graph.*, 36(4):1–13, 7 2017. [1](#), [6](#), [7](#)
- [3] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm, editors, *Computer Vision – ECCV 2020*, pages 405–421, Cham, 2020. Springer International Publishing. [1](#), [3](#), [4](#)
- [4] Y Yao, Z Luo, S Li, J Zhang, Y Ren, L Zhou, T Fang, and L Quan. BlendedMVS: A Large-Scale Dataset for Generalized Multi-View Stereo Networks. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1787–1796, Los Alamitos, CA, USA, 6 2020. IEEE Computer Society. [1](#), [7](#), [8](#)