Supplementary Materials of "AligNeRF: High-Fidelity Neural Radiance Fields via Alignment-Aware Training"

1. Effectiveness of Alignment-aware Training

To further understand how much the alignment-aware training improves NeRF's performance, we quantitatively analyze the performance of NeRF models trained with different training strategy. Concretely speaking, we set 9 models and each model are optimized with the same training iterations. However, instead of adopting alignment-aware training in the entire fine-tuning stage, we only include it in a sub-stage and keep the standard training for the rest time. The percentage of alignment-aware training iterations in the total training iterations range from 0.1 to 0.9, as shown in Fig. 1. By comparing three metrics, the experiments demonstrate that longer alignment-aware training strategy can consistently improve NeRF's performance.

2. Detailed Experimental Results

To present detailed scores on each scene, we include the expanded version of the main results on comparing the proposed methods with the previous version, as shown in Table 1 and 2. Meanwhile, we include depth map visualization in the project pages: https://yifanjiang19.github.io/alignerf.

References

[1] Jonathan T Barron, Ben Mildenhall, Matthew Tancik, Peter Hedman, Ricardo Martin-Brualla, and Pratul P Srinivasan. Mip-nerf: A multiscale representation for anti-aliasing neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5855–5864, 2021.

[2] Jonathan T Barron, Ben Mildenhall, Dor Verbin, Pratul P Srinivasan, and Peter Hedman. Mip-nerf 360: Unbounded anti-aliased neural radiance fields. *arXiv preprint arXiv:2111.12077*, 2021.

[3] Peter Hedman, Julien Philip, True Price, Jan-Michael Frahm, George Drettakis, and Gabriel Brostow. Deep blending for free-viewpoint image-based rendering. *ACM Transactions on Graphics (TOG)*, 37(6):1–15, 2018.

[4] Georgios Kopanas, Julien Philip, Thomas Leimkühler, and George Drettakis. Point-based neural rendering with per-view optimization. In *Computer Graphics Forum*, volume 40, pages 29–43. Wiley Online Library, 2021.

[5] 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 *European conference on computer vision*, pages 405–421. Springer, 2020.

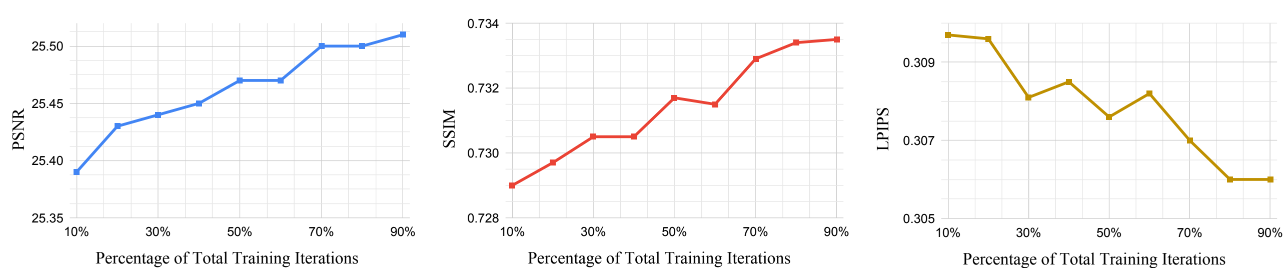

Figure 1. Analysis of the effectiveness made by alignment-aware training. We evaluate 9 models trained with different strategy. The percentage of total training iterations shows how much the alignment-aware training strategy takes comparing to the whole fine-tuning stage. Three metrics are reported and the results demonstrate that adopting more alignment-aware training can produce better NeRF models.

Method	PSNR				
	bicycle	*flowers*	*garden*	*stump*	*treehill*
NeRF [5]	21.76	19.40	23.11	21.73	21.28
mip-NeRF [1]	21.69	19.31	23.16	23.10	21.21
Deep Blending [3]	21.09	18.13	23.61	24.08	20.80
Point-Based Neural Rendering [4]	21.64	19.28	22.50	23.90	20.98
Instant-NGP [6]	22.78	19.18	25.25	24.79	22.45
Stable View Synthesis [8]	22.79	20.15	25.99	24.39	21.72
mip-NeRF [1] w/bigger MLP	22.90	20.79	25.85	23.64	21.71
NeRF++ [9] w/bigger MLPs	23.75	21.11	25.91	25.48	22.77
mip-NeRF-360 [2]	24.46	21.45	26.94	26.40	22.53
Ours	24.75	21.61	27.07	26.69	22.63

Method	SSIM				
	bicycle	*flowers*	*garden*	*stump*	*treehill*
NeRF [5]	0.455	0.376	0.546	0.453	0.459
NeRF w/ DONeRF [7] param.	0.454	0.379	0.542	0.522	0.461
mip-NeRF [1]	0.454	0.373	0.543	0.517	0.46
NeRF++ [9]	0.526	0.453	0.635	0.594	0.530
Deep Blending [3]	0.466	0.320	0.675	0.634	0.523
Point-Based Neural Rendering [4]	0.608	0.487	0.735	0.651	0.579
Instant-NGP [6]	0.540	0.378	0.709	0.654	0.546
Stable View Synthesis [8]	0.663	0.541	0.818	0.683	0.606
mip-NeRF [1] w/bigger MLP	0.612	0.514	0.777	0.643	0.577
NeRF++ [9] w/bigger MLPs	0.630	0.533	0.761	0.687	0.597
mip-NeRF-360 [2]	0.690	0.572	0.815	0.747	0.621
Ours	0.7052	0.588	0.825	0.765	0.632

Method	LPIPS				
	bicycle	*flowers*	*garden*	*stump*	*treehill*
NeRF [5]	0.536	0.529	0.415	0.551	0.546
NeRF w/ DONeRF [7] param.	0.542	0.539	0.436	0.492	0.545
mip-NeRF [1]	0.541	0.535	0.422	0.490	0.538
NeRF++ [9]	0.455	0.466	0.331	0.416	0.466
Deep Blending [3]	0.377	0.476	0.231	0.351	0.383
Point-Based Neural Rendering [4]	0.313	0.372	0.197	0.303	0.325
Stable View Synthesis [8]	0.243	0.317	0.137	0.281	0.286
Instant-NGP [6]	0.397	0.441	0.255	0.339	0.420
mip-NeRF [1] w/bigger MLP	0.372	0.407	0.205	0.357	0.401
NeRF++ [9] w/bigger MLPs	0.356	0.395	0.223	0.328	0.386
mip-NeRF-360 [2]	0.293	0.348	0.165	0.254	0.337
Ours	0.285	0.323	0.152	0.236	0.320

Table 1. We present an expanded version of Table 3 in our main manuscript. We report the detailed scores on each scene separately, on the low-resolution dataset (1280 $\times$ 840).

[6] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Trans. Graph.*, 41(4):102:1–102:15, July 2022.

[7] Thomas Neff, Pascal Stadlbauer, Mathias Parger, Andreas Kurz, Joerg H Mueller, Chakravarty R Alla Chaitanya, Anton Kaplanyan, and Markus Steinberger. Donerf: Towards real-time rendering of compact neural radiance fields using depth oracle networks. In *Computer Graphics Forum*, volume 40, pages 45–59. Wiley Online Library, 2021.

[8] Gernot Riegler and Vladlen Koltun. Stable view synthesis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages

PSNR											
Method	Iterations	Standard					Warped				
		bicycle	*flowers*	*garden*	*stump*	*treehill*	*bicycle*	*flowers*	*garden*	*stump*	*treehill*
NeRF [5]	1x	21.30	18.89	22.92	23.02	21.50	-	-	-	-	-
mip-NeRF [1]	1x	21.33	18.96	22.72	22.96	20.82	-	-	-	-	-
mip-NeRF [1] bigger	1x	21.66	19.27	24.45	23.64	20.76	22.10	19.78	25.49	23.62	21.22
mip-NeRF-360 [2]	1x	23.68	20.83	25.83	26.25	21.96	24.38	21.63	27.31	26.81	22.78
Ours	1x	23.82	20.89	25.95	26.36	22.18	24.57	21.86	27.5	26.96	23.00
NeRF [5]	4x	21.49	19.02	23.19	23.19	21.09	-	-	-	-	-
mip-NeRF [1]	4x	21.56	19.23	23.12	23.18	21.10	-	-	-	-	-
mip-NeRF 360 [2]	4x	24.17	20.71	26.26	26.19	22.09	24.96	21.44	27.90	26.78	23.05
Ours	4x	24.47	20.94	26.43	26.53	22.42	25.44	21.92	28.25	27.18	23.32

SSIM											
Method	Iterations	Standard					Warped				
		bicycle	*flowers*	*garden*	*stump*	*treehill*	*bicycle*	*flowers*	*garden*	*stump*	*treehill*
NeRF [5]	1x	0.464	0.363	0.503	0.561	0.478	-	-	-	-	-
mip-NeRF [1]	1x	0.491	0.386	0.509	0.530	0.504	-	-	-	-	-
mip-NeRF [1] bigger	1x	0.534	0.431	0.678	0.634	0.549	0.568	0.471	0.733	0.656	0.597
mip-NeRF-360 [2]	1x	0.637	0.515	0.734	0.733	0.600	0.679	0.565	0.794	0.760	0.661
Ours	1x	0.640	0.521	0.738	0.739	0.605	0.684	0.579	0.801	0.767	0.667
NeRF [5]	4x	0.471	0.370	0.519	0.568	0.483	-	-	-	-	-
mip-NeRF [1]	4x	0.503	0.401	0.541	0.587	0.518	-	-	-	-	-
mip-NeRF 360 [2]	4x	0.669	0.530	0.764	0.744	0.617	0.714	0.579	0.827	0.773	0.693
Ours	4x	0.684	0.548	0.769	0.762	0.624	0.735	0.610	0.837	0.792	0.691

LPIPS											
Method	Iterations	Standard					Warped				
		bicycle	*flowers*	*garden*	*stump*	*treehill*	*bicycle*	*flowers*	*garden*	*stump*	*treehill*
NeRF [5]	1x	0.717	0.700	0.554	0.583	0.768	-	-	-	-	-
mip-NeRF [1]	1x	0.562	0.567	0.513	0.574	0.548	-	-	-	-	-
mip-NeRF [1] bigger	1x	0.493	0.506	0.324	0.434	0.475	0.480	0.494	0.301	0.429	0.459
mip-NeRF-360 [2]	1x	0.385	0.435	0.268	0.330	0.417	0.368	0.419	0.244	0.318	0.394
Ours	1x	0.381	0.424	0.262	0.341	0.416	0.357	0.399	0.232	0.318	0.389
NeRF [5]	4x	0.669	0.668	0.513	0.559	0.741	-	-	-	-	-
mip-NeRF [1]	4x	0.547	0.551	0.473	0.511	0.530	-	-	-	-	-
mip-NeRF 360 [2]	4x	0.348	0.421	0.232	0.310	0.384	0.332	0.406	0.210	0.300	0.351
Ours	4x	0.328	0.392	0.225	0.307	0.381	0.303	0.366	0.194	0.282	0.349

Table 2. We present an expanded version of Table 1 in our main manuscript. We report the detailed scores on each scene separately, on the high-resolution dataset (2560×1680).

12216–12225, 2021.

[9] Kai Zhang, Gernot Riegler, Noah Snavely, and Vladlen Koltun. Nerf++: Analyzing and improving neural radiance fields. *arXiv preprint arXiv:2010.07492*, 2020.