

KiloNeRF

Juhyun Lee

2021. 10.

OEQE Lab

Seoul National University

*Optical Engineering and Quantum Electronics Lab.,
Electrical and Computer Engineering,
Seoul National University, Republic of Korea*





Contents

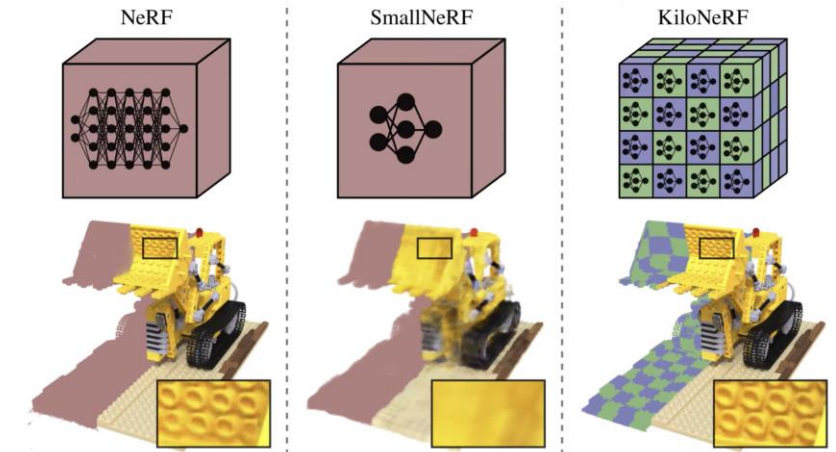
- Introduction
- Methods
 - Background: NeRF
 - KiloNeRF
 - Training with distillation
 - Sampling & Implementation details
- Results
- Conclusion



Introduction

Reiser, Christian, et al. "KiloNeRF: Speeding up Neural Radiance Fields with Thousands of Tiny MLPs." arXiv preprint arXiv:2103.13744 (2021).

- 문제 제기
 - NeRF의 가장 큰 문제점은 렌더링 속도. 실시간 합성이 되지 않음.
 - 이를 완화하고자 크기가 작은 모델을 사용하면 visual quality가 감소함.
 - 따라서, 오브젝트를 구획별로 나누고 각 구역에 작은 네트워크를 할당해서 렌더링 속도 문제를 해결하고자 함.
- 얻어지는 효과
 1. 실시간 합성
 2. View-dependent 색깔 변화, 반투명한 물체 표현 가능
- 학습 방법
 1. NeRF를 훈련시켜 teacher model로 사용
 2. Distillation: KiloNeRF가 teacher의 출력과 같은 결과를 내도록 훈련
 3. Photometric loss를 이용하여 KiloNeRF를 fine-tuning



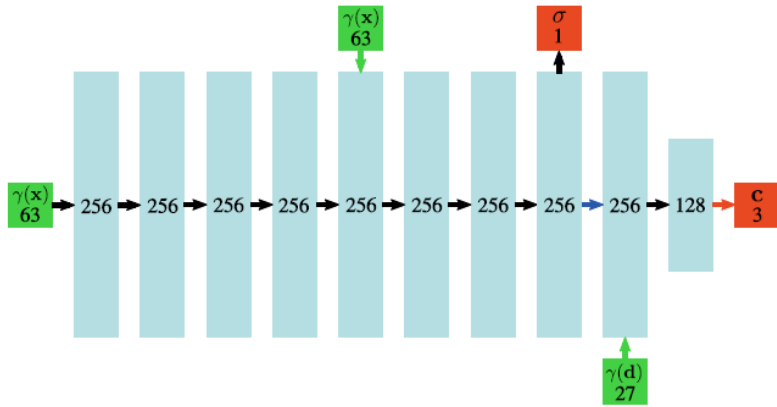


Methods

Reiser, Christian, et al. "KiloNeRF: Speeding up Neural Radiance Fields with Thousands of Tiny MLPs." arXiv preprint arXiv:2103.13744 (2021).

NeRF & KiloNeRF

NeRF: ~1056 kFLOPs



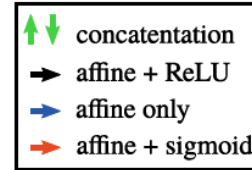
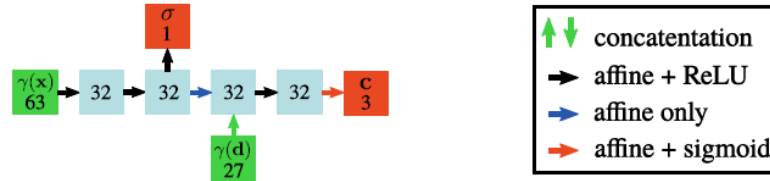
$$(\mathbf{c}_i, \sigma_i) = f_{\theta}(\mathbf{x}_i, \mathbf{d}) \quad \text{with} \quad i = 1, 2, \dots, K$$

$$\hat{\mathbf{c}} = \sum_{i=1}^K T_i \alpha_i \mathbf{c}_i$$

$$T_i = \prod_{j=1}^{i-1} (1 - \alpha_j) \quad \alpha_i = 1 - \exp(\sigma_i \delta_i)$$

$$\mathcal{L} = \frac{1}{B} \sum_b \|\mathbf{c}_b - \hat{\mathbf{c}}_b\|_2^2$$

KiloNeRF: ~12 kFLOPs

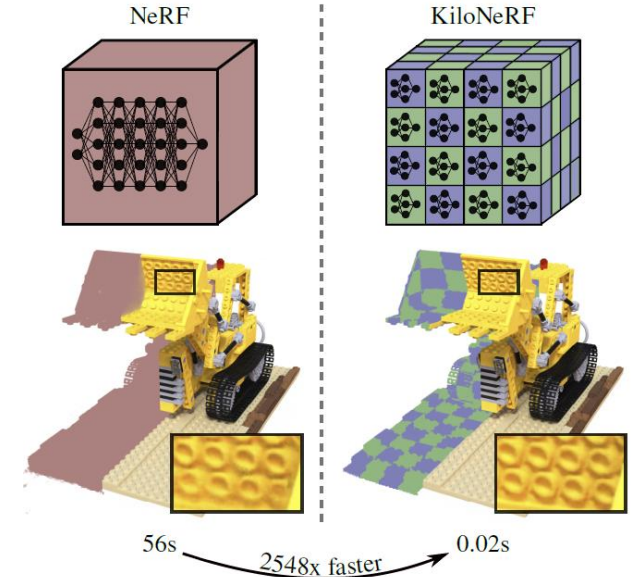


$$g(\mathbf{x}) = \lfloor (\mathbf{x} - \mathbf{b}_{min}) / ((\mathbf{b}_{max} - \mathbf{b}_{min}) / \mathbf{r}) \rfloor$$

$$(\mathbf{c}, \sigma) = f_{\theta(g(\mathbf{x}))}(\mathbf{x}, \mathbf{d})$$

네트워크 구조

- 10개의 레이어, 256 개의 히든 유닛 → 4개의 레이어, 32개의 히든 유닛
- 모든 레이어의 활성화 함수는 기본적으로 ReLU를 사용, 마지막 레이어에는 sigmoid, density를 출력하는 레이어와 다음 레이어 사이에는 활성화 함수를 적용하지 않음.





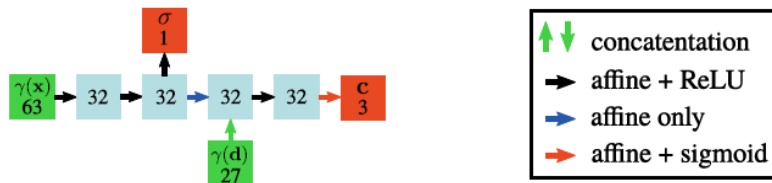
Training with distillation

KiloNeRF를 처음부터 학습시키면 artifact가 생김 → NeRF를 이용하여 파라미터를 초기화 하겠다!

- NeRF를 학습시켜 teacher로 삼고, KiloNeRF가 teacher model과 같은 결과값을 출력하도록 학습
- 각 샘플의 c 와 σ 두 가지를 비교했는데, σ 는 α 로 바뀌어서 비교
- student의 파라미터는 α 과 c 의 L2 손실함수로 최적화.
- 마지막으로 KiloNeRF는 photometric loss를 이용하여 파인 튜닝.

$$\alpha_i = 1 - \exp(\sigma_i \delta_i) \quad (4)$$

$$\mathcal{L} = \frac{1}{B} \sum_b \|\mathbf{c}_b - \hat{\mathbf{c}}_b\|_2^2 \quad (5)$$



(a) Without Distillation

(b) With Distillation



Sampling & Implementation details

- Empty Space Skipping (ESS)
 - 콘텐츠가 없는 오브젝트 포인트에 대해서는 ray를 계산하지 않겠다.
 - Occupancy grid를 상정 \rightarrow density 값을 기준으로 콘텐츠의 존재 여부를 판단하는 binary value.
 - 논문에서는 $256 \times 256 \times 256$ 의 해상도로 occupancy grid를 설정함.
 - Occupancy grid를 $3 \times 3 \times 3$ 크기의 subgrid로 나누어 해당 셀 안에 teacher model의 density가 문턱값 τ 이상인 픽셀이 하나라도 있으면 콘텐츠가 존재하는 것으로 판단.
- Early Ray Termination (ERT)
 - i 번째 샘플의 transmittance가 0에 가까워지면, 이후 $(i + 1)$ 번째 샘플 부터는 계산하지 않겠다.
 - Transmittance가 0에 가까워진다는 것은 이후 샘플의 컬러값 c 가 최종 픽셀 값에 끼치는 영향이 급격히 줄어든다는 의미.
 - ERT는 inference 때에만 적용함.

$$T_i = \prod_{j=1}^{i-1} (1 - \alpha_j) \quad (3)$$



Results

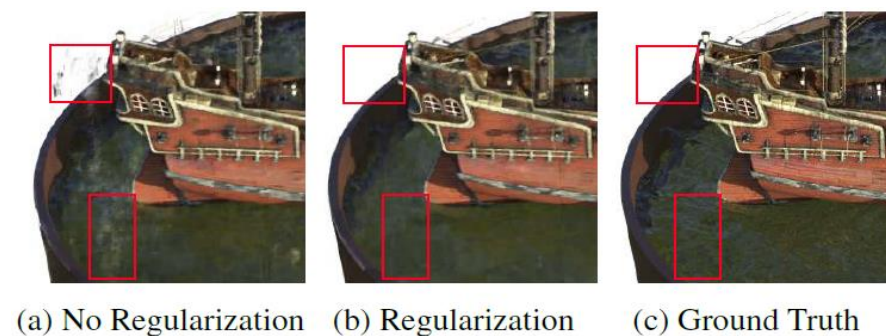
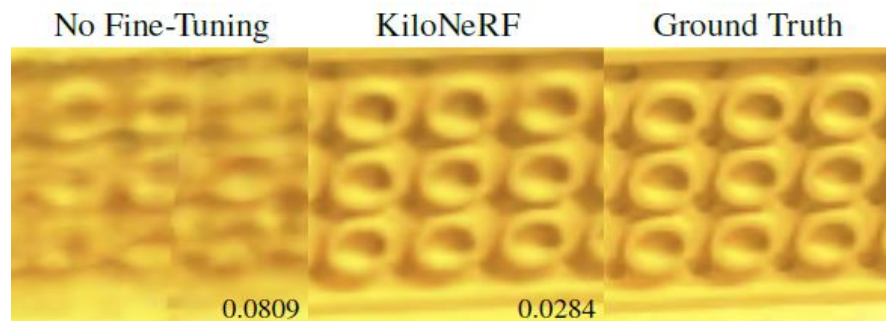
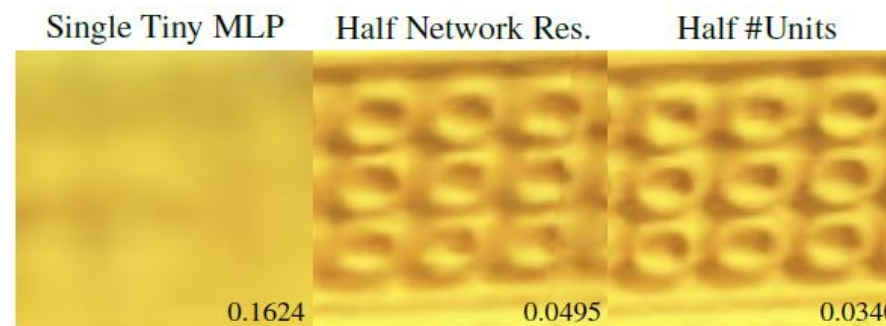
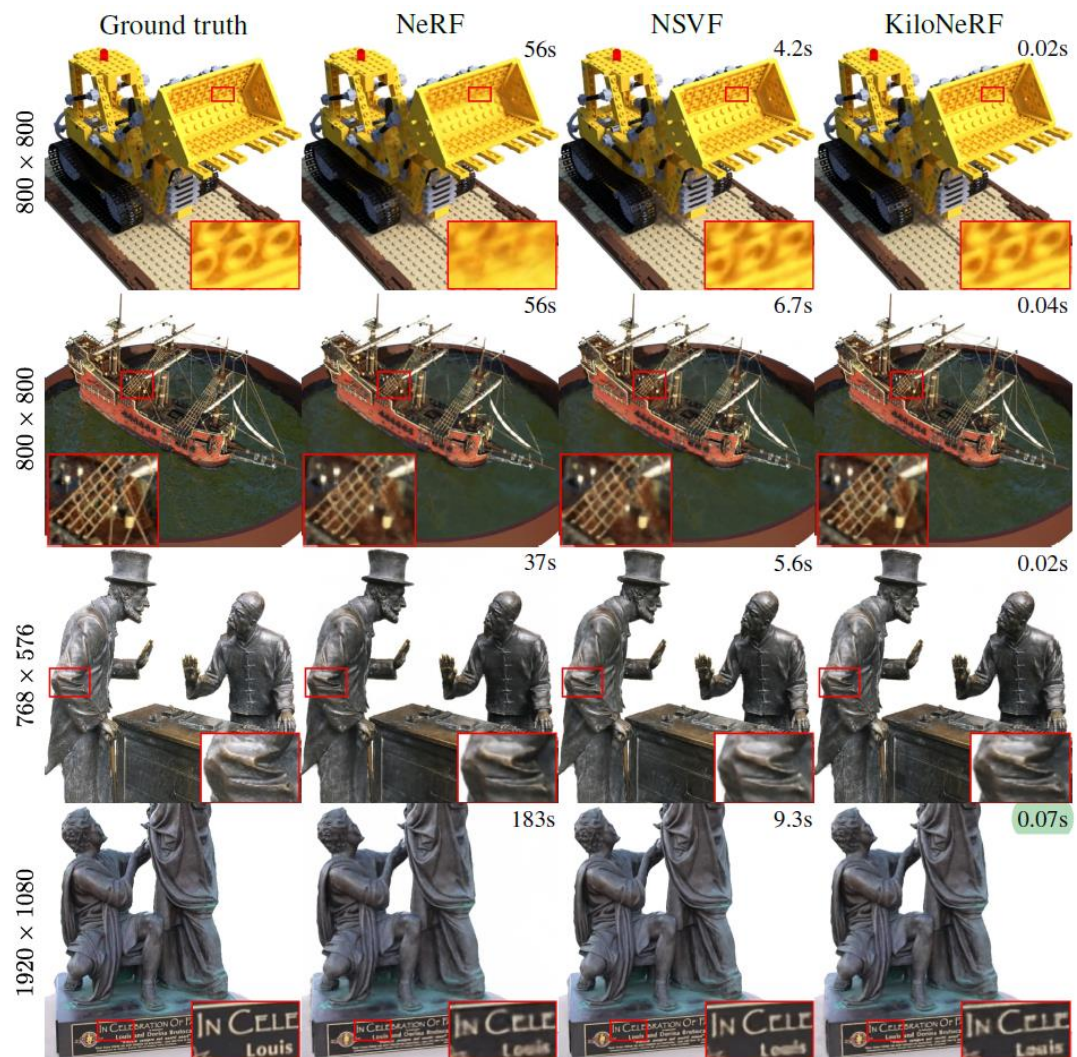
Reiser, Christian, et al. "KiloNeRF: Speeding up Neural Radiance Fields with Thousands of Tiny MLPs." arXiv preprint arXiv:2103.13744 (2021).

Resolution		BlendedMVS 768 × 576	Synthetic-NeRF 800 × 800	Synthetic-NSVF 800 × 800	Tanks & Temples 1920 × 1080
PSNR ↑	NeRF	27.29	31.01	31.55	28.32
	NSVF	26.90	31.74	35.13	28.40
	KiloNeRF	27.39	31.00	33.37	28.41
SSIM ↑	NeRF	0.91	0.95	0.95	0.90
	NSVF	0.90	0.95	0.98	0.90
	KiloNeRF	0.92	0.95	0.97	0.91
LPIPS ↓	NeRF	0.07	0.08	0.04	0.11
	NSVF	0.11	0.05	0.01	0.15
	KiloNeRF	0.06	0.03	0.02	0.09
Render time (milliseconds) ↓	NeRF	37266	56185	56185	182671
	NSVF	4398	4344	10497	15697
	KiloNeRF	30	26	26	91
Speedup over NeRF ↑	NSVF	8	13	5	12
	KiloNeRF	1258	2165	2167	2002



Results

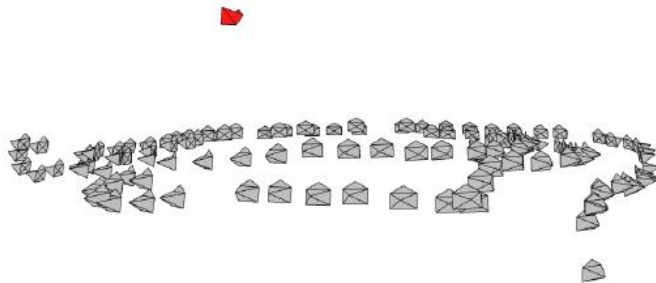
Reiser, Christian, et al. "KiloNeRF: Speeding up Neural Radiance Fields with Thousands of Tiny MLPs." arXiv preprint arXiv:2103.13744 (2021).





Conclusion

- NeRF의 느린 렌더링 속도를 보완하기 위해 더 작은 네트워크를 사용하는 방법을 제안하였다. 이 때, 오브젝트 영역을 여러 구역으로 나누고, 구역별로 더 작은 MLP를 적용했다.
- KiloNeRF는 어떤 씬에 대해서 훈련된 NeRF를 teacher model로 삼아 NeRF의 결과를 출력하는 방향으로 파라미터를 초기화 한다. 이후, photometric loss를 적용하여 정밀 훈련을 진행한다.
- NeRF와 비교해서 비슷한 수준의 visual quality를 내면서, 수 천 배 빠르게 렌더링 할 수 있다.
- 다만, NeRF의 출력을 모사하는 방향으로 초기 학습이 이루어지기 때문에, NeRF의 이미지 품질을 상한선으로 가진다는 단점이 있다.



(a) Out-of-distribution camera pose



(b) NeRF



(c) KiloNeRF