Nerf This!

Andrew Ding, Mingxuan Ji, Tony Zhu

BROWN

Motivation

Novel view synthesis has been a longstanding problem in computer graphics. Traditional techniques using CNNs, feature matching and point clouds have not shown too much success in the realm. NERF attempts to take a novel approach by representing a 3D object as a radiance field: points in 3d space with different colors and opacity depending on coordinate and view direction. By using a MLP and volumetrically rendering these points, we have been able to generate images viewed from completely new viewing angles.

We have decided to implement this paper because we believe radiance fields will be the future of rendering and view synthesis in computer vision, and the NERF technique showed gains above other SOTA algorithms.

Differentiable Volumetric Rendering

Our ultimate goal is to estimate the rgb value of pixels of synthesized images from completely new viewing directions. To estimate the rgb value for one pixel, we integrate the volume density (σ) and color (***c***) of all points along the ray which is projected from the pixel in the viewing direction. Practically, we cut the ray between a manually set near bound and far bound into N line segments, and sample a point from each of the N segments. Then, we implement the discretized version of the integration function:

$$t_i \sim \mathcal{U}\left[t_n + \frac{i-1}{N}(t_f - t_n), t_n + \frac{i}{N}(t_f - t_n)\right]$$

$$\hat{C}(r) = \sum_{i=1}^{N} T_i(1 - \exp(-\sigma_i \delta_i))c_i \text{ where } \delta_i = t_{i+1} - t_i$$

$$\text{where } T_i = \exp\left(-\sum_{j=1}^{i-1} \sigma_j \delta_j\right)$$

Problem & Goal

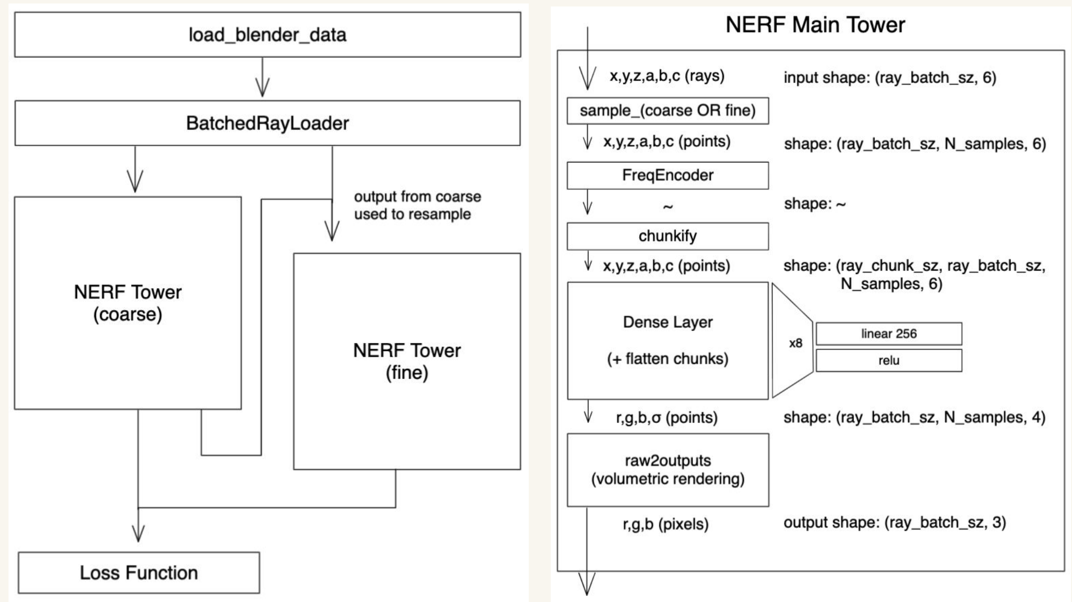
Novel View Synthesis
Given a set of training images, we want to generate views from completely different angles never seen before.

Metrics
We will be using two metrics, mean squared error loss and peak signal to noise ratio

Goal
We want to recreate the 30 PSNR score with the original models in the Realistic Synthetic 360 (Blender) dataset seen in the paper.

NERF Render Output

Model Architecture

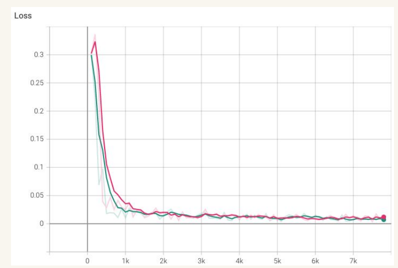


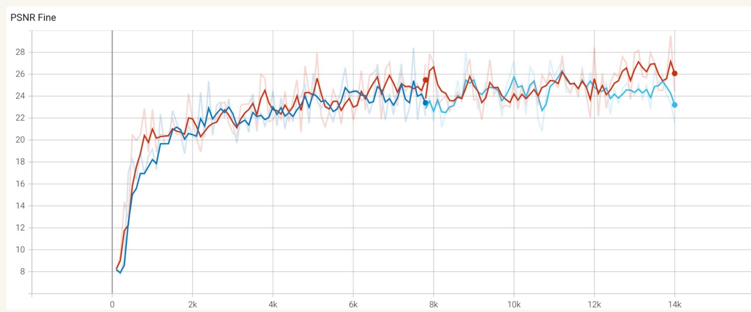
The model is complex as volumetric rendering is a part of the training loop. Depending on settings, the NERF tower can be run twice on coarse and fine hierarchical sampling settings, allowing for a speedup in the training time.

More results

Object	Epochs	Best Loss	Best PSNR
Chair	12k	4e-3	31
Drums	130k	5.5e-3	29.5
Ship	160k	3e-3	31
Lego	42k	3e-3	31
Original Paper	up to 200k	N/A	Average of 31

Table 1. Training Results





References

[1] 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, 2020.
[2] Thomas M ̈uller, 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.
[3] Alex Yu, Sara Fridovich-Keil, Matthew Tancik, Qinhong Chen, Benjamin Recht, and Angjoo Kanazawa. Plenoxels: Radiance fields without neural networks. CoRR, abs/2112.05131, 2021.

Acknowledgements

We would like to thank Srinath Sridhar and Anh, for the guidance on this project.

Code

https://github.com/tonyzhu163/nerf_this_2022/