Name: Aditya Mehta
Project Title: Generative Adversarial NeRF (GANeRF)
Advisor: Felix Heide
Class: COS IW 01: Extreme Neural Rendering

Motivation and Goal:

Advancements in Neural Radiance Fields (NeRFs) have opened the path for high quality 3D scene representations to be made from a series of photos collected at different camera poses around a target object. Applications of NeRFs include contributions to self-driving cars, scene recreations for film and sports, as well as scientific data collection. However, the use of NeRFs comes with a steep drawback—a high volume of training images from various precise camera angles are required to create a functioning model.

The aim of this project is to pair a Generative Adversarial Network (GAN) with NeRFs, by attaching a discriminator trained on publicly available datasets pertaining to the images in question in order to improve the quality of scene reconstructions for fewer provided training images. By doing so, the generated images from different camera poses from NeRF will bear more resemblance to collected photographs, thereby resulting in a clearer scene representation.

Results from this research have the potential to significantly decrease the number of training images needing to be collected, which will hopefully make NeRFs easily available to the general public without the need for such accurate camera poses as is currently required.

Problem Background and Related Work:

Most of the research put into NeRFs are centered around decreasing the training time, or removing certain necessary parameters of the NeRFs such as limited camera poses or scene movement with the knowledge that more images will be expected to improve the model in any case. Previous research has worked with utilizing GANs in conjunction with NeRFs in order to remove the need for camera poses in NeRFs (GNeRFs) which showed promising results. Similar research has also gone into reducing the time for NeRF computations (NVIDIA) among other limiting factors. Work has even gone into doing the reverse of this project: generating training data for other computer vision neural networks using NeRFs (Neural-Sim). However not much research has been done in utilizing public datasets to improve the accuracy of the scene representations largely because NeRFs are quite new and public datasets are largely limited in images per single object.

Approach:

The approach of utilizing a GAN to augment the learning of the NeRF will utilize images in the public domain to reinforce that scene reconstructions provide realistic images from

intermediary views. By tuning NeRFs to create accurate intermediate images, this greatly increases the quality of scene representations, since at the current moment, parts of objects obscured or hidden in the majority of images will appear distorted or fuzzy. In making the intermediate images more accurate to what a captured photograph would look like, this tunes the distorted or fuzzy sections of the scene reconstruction to be more accurate. This can then be applied to any iteration of NeRF to improve accuracy as public datasets expand.

Plan:

1. I will first begin by analyzing potential options for NeRFs and datasets to utilize in conjunction with a GAN as well as familiarizing myself with using the NeRF. This is because for general NeRFs, I may need to collect data myself, or find specific datasets from which to use in training the GANeRF, while for specific NeRFs, the training data might not have an associated large dataset for general images given their testing images.
2. I'll then investigate how to construct the discriminator and how to select well-chosen intermediary camera poses (rather than centered where the majority of training images were taken from). This investigation is because I have never built a GAN from scratch before, and as such I anticipate this being the most challenging part of the research.
3. I'll conduct a series of tests for the best way to utilize the dataset in improving the discriminator. These would aim to test the number of dataset images to use, potential filters for the dataset images, as well as reducing the number of training images for the NeRF generative part of the GANeRF.
4. If time permitted or along the process, I would like to attempt to make the generative NeRF section swappable for other NeRF models and datasets in order to generalize this model to work with all NeRF models.

Evaluation:

For evaluation purposes, I will maintain the discriminator from training the GANeRF and test it against the counterpart NeRF and compare the results of the discriminator on a series of test camera poses. By doing so, I can quantify the improvement of the GANeRF over its counterpart NeRF in terms of its ability to fool the discriminator. This will tell me the accuracy of the GANeRF and whether or not a substantial improvement has been made.