Re-Balancing of Long-Tailed Datasets Using NeRF Generation

Meher Mankikar, Deep Patel, Jimmy Zhang
Carnegie Mellon University
5000 Forbes Ave. Pittsburgh, PA 15213
mmankika@andrew.cmu.edu, dmpatel@andrew.cmu.edu, jimmyzha@andrew.cmu.edu

Abstract

In this paper, we will explore the use of NeRF generation in order to generate new instances to augment long-tailed datasets. From our preliminary tests, we see that training on an unbalanced dataset performs quite poorly. Current techniques to solve this issue include random undersampling and oversampling. However, undersampling loses useful for learning better representations in images and oversampling can overfit to the long-tail classes in training data. We aim to resolve these issues by generating new training samples that differ in view using NeRF such that the model learns better overall in comparison to prior methods.

1 Background / Literature

In the paper "NeRF : Representing Scenes as Neural Radiance Fields for View Synthesis", Mindenhall, et. al. present a method that takes in continuous 5D coordinates ((x, y, z) spatial location and (θ, ϕ) viewing direction) and outputs the volume density and RBG color of that spatial location to create a neural radiance field (NeRF). To render this NeRF, Mindenhall, et. al. sent camera rays through the scene to generate a sampled set of 3D points. They then used those points and corresponding viewing directions as inputs to a neural network that outputs a set of RGB colors and densities. Then, volume rendering techniques are used to accumulate the outputs of the neural network into a 2D image.

The paper "Improved Sampling Techniques for Learning an Imbalanced Dataset" by Lauron and Pabico discusses methods for learning using data with long tail distributions (imbalanced datasets). Issues with these kinds of datasets arise because skewed distribution of data over several classes in the dataset can lead to prediction for minority classes being largely inaccurate. The paper builds on two sampling techniques that are used to solve the issue of imbalanced datasets: random undersampling (RU) and random oversampling (RO). RU randomly removes entries from all classes except that with the smallest number of instances. This is continued until the number of instances for each class is equal. On the other hand, RO adds data by randomly replicating instances already in the dataset of all the minority classes until the number of instances for each class are equal.

2 Methods/Models

In order to evaluate the performance of utilizing NeRF to re-balance our dataset, we first need to compare the performance of this method against other popular techniques for dealing with long-tailed class distributions. To do these, we examined the performance of our model on a long-tailed dataset as our controlled experiment, and afterwards, we applied both the undersampling and oversampling techniques to our training method. We obtained a long-tailed dataset by utilizing the NeRF dataset and then performing removing samples from the dataset to create a long-tailed distribution between the classes. This distribution can be seen in the plot below. Our procedure for obtaining our baseline results are described below.

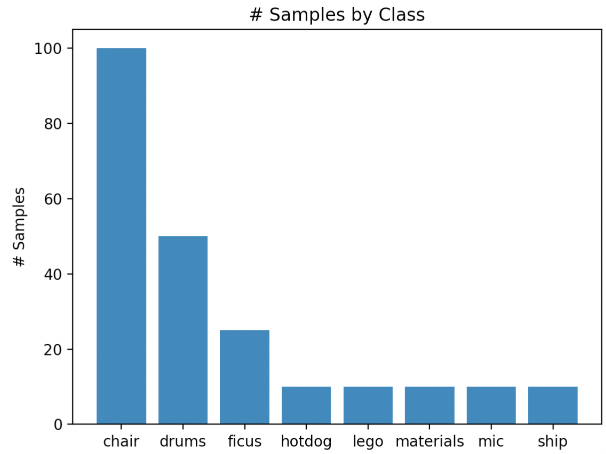

Figure 1: Long-Tailed Data Distribution for Baseline

We used a convolutional neural network with 3 convolutional blocks consisting of a convolutional operation, a ReLU activation, and finally, a Max-Pooling layer. Then, the model had a two linear-layers with 500 hidden units. We first trained it for 10 epochs using the Adam optimizer learning method with learning rate $\alpha = 0.01$ and momentum $\gamma = 0.9$. Initially, we did this on the dataset that was shown with the number of samples above. Afterwards, we trained the model again via a sub-sampled dataset where each class had the same number of samples as the smallest class from the long-tailed class distribution. Finally, we trained the model via a super-sampled dataset in which samples from smaller classes were repeated so that they had better balanced representations via repeating. We evaluate the performance of each of these models on a balanced validation set to understand how well the models perform on the long-tail classes. The results of each of the performance of these models can be seen below.

3 Preliminary Results

Figure 2: Model Performance Trained on Balanced Dataset

(a) Model Performance Trained on Undersampled Dataset

(b) Model Performance Trained on Oversampled Dataset

4 Evaluation of Preliminary Work

Our baseline model was trained on the original, highly unbalanced dataset. The validation loss does not make significant progress across the later epochs. This is because the model is overfitting on the majority classes and cannot consistently predict minority classes. With respect to the baseline, we tried two different sampling techniques to remedy the imbalance.

Performing random undersampling on the dataset, the overall validation loss decreased, but not significantly. RU may be suboptimal because too many instances from the majority class are removed, which jeopardizes the important concepts and features of the majority class. RU is likely overcompensating for the minority classes and sacrificing precision for the majority classes.

Random oversampling produced the best losses over epochs, achieving a logarithmic-like rate of decrease. RO replicates old samples from the minority classes to balance the dataset. Therefore, by keeping all the features of the majority class and increasing the number of data points, the model becomes more accurate. However, this model may be overfitting to the training data, especially in long-tail classes.

5 Future Work

To improve upon the current techniques, we realize that the subsampled and supersampled datasets either lose information on the data distribution or force overfitting to the long-tailed data, respectively. To prevent the loss of information while introducing data variety in long-tailed classes, we hope to use NeRF to generate new 2D views of the scene in each classification, and as a result, we will sample from these generated scenes to create more balanced datasets. We will evaluate our final method by comparing the validation performance of the model trained on the dataset balanced from NeRF. We will further experiment with the number of samples used to train the NeRF model for each scene to understand how well the accuracy of NeRF improves the sampling.

6 Teammates and Work Division

Format: Team Member Assigned: Task [Completion Date]

- Deep: Train NeRF model [March 20th]
- Meher: Sample new images using NeRF model [March 25th]
- Meher: Retrain model using new dataset [March 30th]
- Jimmy: Consider other metrics of evaluation including G-mean and F-Measure which consider the sensitivity of different classes. [March 20th]
- Jimmy: Evaluate performance of all models on new datasets that are not simulated [April 15th]

References

[1] Maureen Lyndel C Lauron and Jaderick P Pabico. Improved Sampling Techniques for Learning an Imbalanced Data Set. 2016.

[2] 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 Eur. Conf. Comput. Vis., 2020.