SRMCV Project: Extending the Diffeomorphic Neural Reconstruction Approach - Week 5 Report

1 Overview

During our previous work, we found out that our network was unable to reconstruct input scenes with non-trivial backgrounds. Before using a NeRF to obtain the background color, we tested if this idea could work by using the ground truth rendered background as input to our rendering function. Unfortunately, this alone did not fix the training results on this scene, in particular, the ears were not reconstructed correctly.

In order to find out whether the problem was with our choice of background or with the network in general, we experimented with different background setups.

1.1 Simpler background

In our previous work, the ever-present background behind the bunny's ears seems to have resulted in their reconstruction being deformed. Thus, our first thought was to set up the background in a way that there are at least some viewpoints from which the ears have no cubes directly behind them. This should hopefully have a positive effect on their reconstruction.

Our results can be seen in Figure 1, where on the left are the ground truth images of the full scene consisting of meshes of cubes as the background and the bunny mesh, and on the right the reconstructed mesh. We see that this time, where there is no background at all behind the ears, the ears are well reconstructed. However, there is a significant surface of cubes behind the lower part of the head, which once again causes undesired deformations in the mesh. We would attribute this once again to the color of the background being identical and will explore other adaptations to the background in the following.

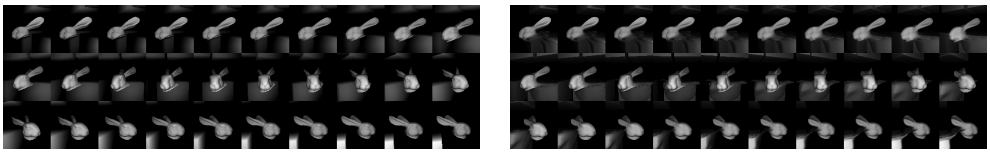

Figure 1: On the left, the ground truth images of the full scene consisting of meshes of cubes as the background and the bunny mesh. On the right, the reconstructed mesh.

1.2 View-points from the back of the bunny

Another possible way to improve the reconstructed ears is to add viewpoints behind the bunny's head and perform a full rotation. To do this, we placed the cubes a little further away from the bunny to allow the camera to make a full rotation around the head. Our results can be seen in Figure 2, where on the left are the ground truth images of the full scene consisting of meshes of cubes as background and the bunny mesh, and on the right is the reconstructed mesh.

This time we see that with a full rotation around the head, there is a significant improvement in the reconstruction of all parts of the bunny's head, including the ears. Thus, the previously observed

artifacts are successfully fixed with this new setup. Of course, the fact that the cubes are also further away and thus less strongly lit may also contribute to this result.

One thing to note here is that the network struggles to strongly separate the ears, even when provided views from behind the bunny. Possible fixes for this could be reducing the velocity loss imposed on the network outputs, increasing the number of timesteps or making use of learning rate decay during training. This needs further investigation.

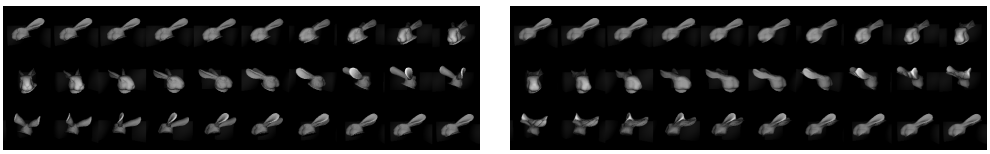

Figure 2: On the left, the ground truth images of the full scene consisting of meshes of cubes as the background and the bunny mesh. On the right, the reconstructed mesh.

1.3 Colored background

As the results in subsection 1.1 and subsection 1.2 show, there is still room for improvement. So far, the bunny and the background cubes had the same color, white. This can make it difficult for the optimization to distinguish between the background and the bunny itself as the foreground. So we decided to try rendering the scene with colored cubes in the background and the white bunny mesh as the foreground. This should give the optimizer a better chance to distinguish between foreground and background.

After setting up the dataset, we perform training by once again providing the ground truth background to the render function. As the results in Figure 3 show, the colored background helps to improve the reconstruction results a lot. The model is now able to clearly distinguish between the background and the bunny, resulting in a better reconstruction of the ears.

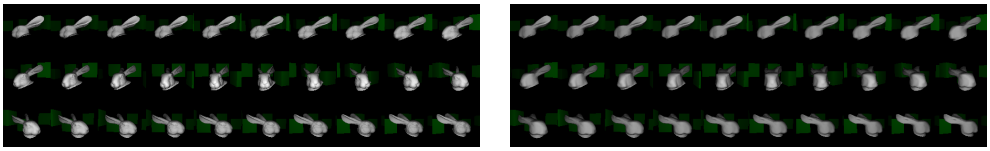

Figure 3: On the left, ground truth images with colored background. On the right, the reconstructed mesh.

1.4 Dataset Creation and Data Loading

Increasing the number of possible datasets over the weeks, by adding different view-points, a possible background consisting of different cube arrangements as well as colors, has left our dataset folder very cluttered. This can also lead to a lot of unnecessary bugs for reasons such as a wrong L0 value, which varies depending on the dataset or a mismatched set of camera positions.

In order to make the training with our codebase more intuitive we cleaned up large parts of it. This includes a script for dataset creation that simply expects a set of camera viewpoints (e.g. exported from Blender) and a mesh and can then automatically prepare all the datasets necessary for training using all our different configurations. For training we can also now easily swap between datasets by changing the appropriate values in our config file.

2 Next Steps

From our experiments with different backgrounds, we conclude that our reconstructions improve significantly with a background setup that allows to distinguish between the bunny's ears and the background.

Therefore, we were able to show that our method works when provided with the correct background image so a possible next step would be to use a NeRF to train the background images at the same time as the mesh reconstruction.

3 Task Assignment

Vasiliki Papadouli:

- Implementation of the new backgrounds using Blender software and creation of the new datasets.
- Training experiments of simpler background and full rotation and observing the results.

Alexander Fuchs:

- Reworked dataset creation and dataloader
- Training and debugging for rotation and color dataset
- Fixed validation set renders
- Clean-up

Zehranaz Canfes:

- Validation dataset bugfix.
- Modify rendering and dataset creation code to render textured backgrounds.
- Create dataset of colored cubes and the whole scene with colored background.
- Train with colored background.
- Clean up some parts of the code.