

NeRF Reconstruction Report

Aaron Miller
February 20th, 2026

Challenges During Setup and Processing

Most of my time was spent configuring and troubleshooting the software environment. I ran Ubuntu 22.04 through WSL2 on Windows and encountered several compatibility issues involving CUDA, COLMAP, and tiny-cuda-nn.

One of the largest challenges was ensuring proper GPU support across the full toolchain. Although `nvidia-smi` confirmed that the RTX 3070 Laptop GPU was visible inside WSL, COLMAP initially failed to detect a CUDA-capable device. This issue stemmed from mismatches between Conda-provided CUDA runtime libraries and the system CUDA installation. Resolving it required aligning CUDA versions, verifying driver compatibility, rebuilding COLMAP, and adjusting library paths so the correct runtime libraries were used.

In addition, building `tiny-cuda-nn` required careful matching between the installed CUDA toolkit and the PyTorch CUDA build. Several build failures were resolved by correcting version mismatches and reinstalling components to ensure consistency.

Once the environment was stable and GPU detection worked reliably, the remainder of the pipeline executed smoothly. Frame extraction, feature matching, pose estimation, and nerfstudio training completed without further issues. The Nerfstudio viewer was intuitive to use, and bounding box cropping allowed me to isolate the object cleanly before exporting results.

Assessment of Pose Estimation Quality

The recovered camera poses form a circular trajectory around the object, consistent with the real-world capture setup.

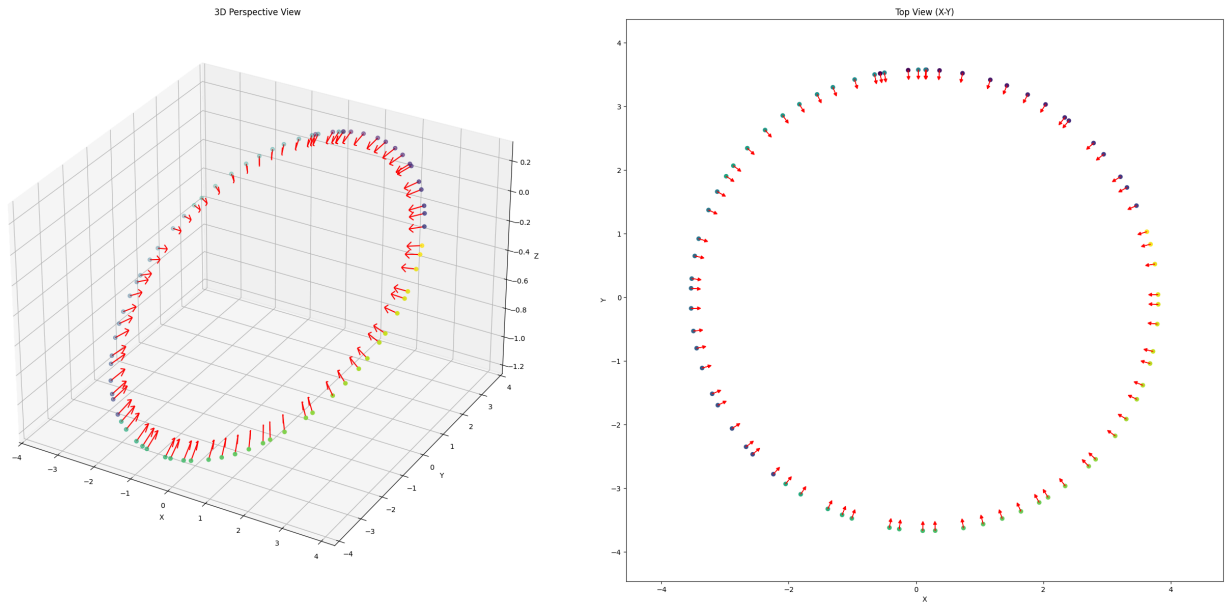


Figure 1: Perspective and top-down views of estimated camera poses.

The overall geometry appears coherent, with no major pose outliers or collapsed structure. There is some variation along the vertical (z) axis in the estimated camera path. However, in the original video, the rotating platform itself appears slightly tilted relative to the camera. This real-world tilt likely explains the observed vertical variation in the reconstruction rather than indicating a failure in pose estimation.

Overall, the pose estimates are geometrically consistent and sufficiently accurate.

Reflection on NeRF Reconstruction Results

The NeRF reconstruction is visually consistent with the source video at medium viewing distances.

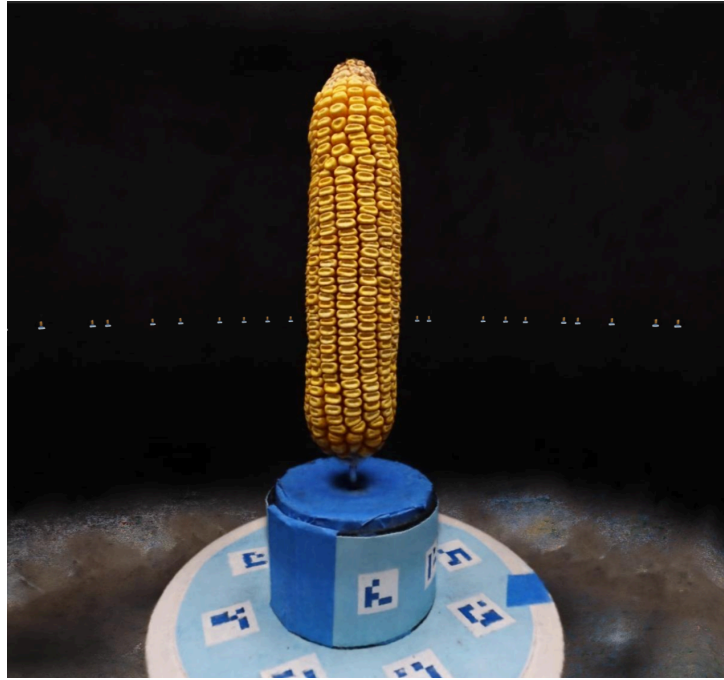


Figure 2: NeRF reconstruction from a medium viewing distance.

The global shape of the object is preserved, and there are no major structural artifacts. However, fine details degrade when inspected closely. Edges appear slightly blurred, and faint halo-like artifacts are visible from certain viewpoints, particularly from top-down angles. These limitations are likely influenced by viewpoint coverage and minor pose inaccuracies.

Training the nerfacto model took approximately 36 minutes on an RTX 3070 Laptop GPU. Training proceeded stably without divergence or instability. After training, the scene was successfully exported as a .ply point cloud.

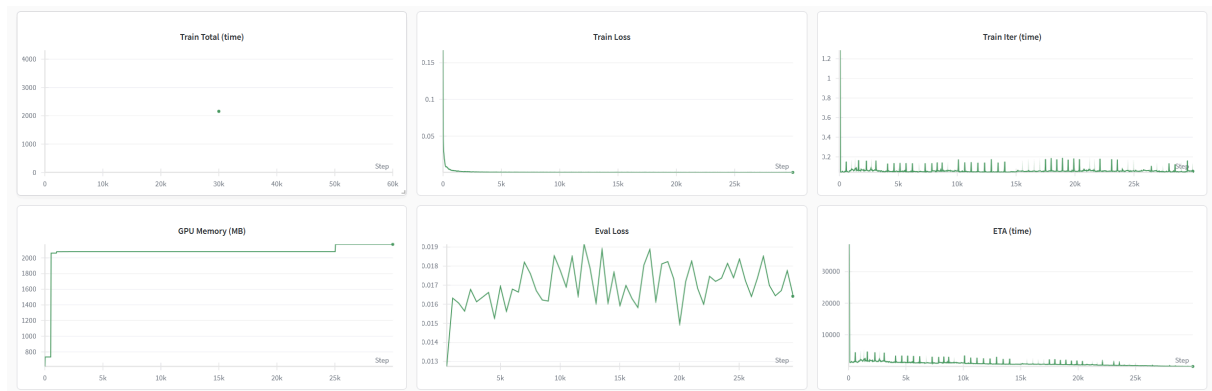


Figure 3: Training metrics and loss curves from Weights & Biases.

Conclusion

Despite substantial setup challenges related to CUDA configuration and COLMAP integration within WSL, the full pipeline was completed. The process encompasses video ingestion, camera pose estimation, NeRF training, and a final point cloud export. The resulting reconstruction is structurally accurate and visually coherent, with expected limitations in fine detail resolution.