Efficient 3D Scene Modeling with Instant-NGP, Nerfacto, and TensoRF: A Case Study on the ARISPE Meteorite

CHINMAY AMRUTKAR
chinmay.amrutkar@asu.edu
Arizona State University

Abstract

This work explores 3D scene reconstruction using Neural Radiance Fields (NeRF) frameworks on a real-world dataset captured at Arizona State University (ASU). A dataset of 55 images of the ARISPE meteorite specimen was collected through a comprehensive 360-degree capture. Camera poses were estimated using Agisoft Metashape and appropriately formatted for NeRF pipelines. We employed three state-of-the-art reconstruction methods — Instant-NGP, Nerfacto, and TensoRF — to reconstruct the scene. Each method was trained and evaluated based on training efficiency, qualitative reconstruction fidelity, and resilience to background noise. Instant-NGP achieved extremely fast results but suffered from fine detail loss; Nerfacto balanced training time and reconstruction quality effectively, producing highly detailed outputs; TensoRF exhibited greater noise artifacts despite longer training times. Our comparative study highlights the strengths and limitations of each approach in practical reconstruction tasks, providing insights into trade-offs between speed, quality, and robustness in NeRF-based pipelines.

1. Introduction

Reconstructing realistic 3D models of real-world objects from 2D images has been a long-standing challenge in computer vision and graphics. Traditional techniques such as multi-view stereo (MVS) and structure-from-motion (SfM) have been widely adopted, but they often struggle with fine detail preservation, varying lighting conditions, and background clutter. Recent advancements in Neural Radiance Fields (NeRF) have revolutionized the 3D reconstruction landscape by learning continuous volumetric scene representations directly from sparse 2D image sets.

In this work, we explore and compare the capabilities of three modern NeRF-based reconstruction techniques — **Instant-NGP**, **Nerfacto**, and **TensoRF** — applied to a real-world dataset. The chosen scene for reconstruction is the "ARISPE" meteorite, a 122-kilogram iron specimen historically used as an anvil, currently displayed inside the Gallery of Scientific Exploration at Arizona State University (ASU).

The reconstruction pipeline begins with collecting a comprehensive image dataset through a 360-degree photographic sweep around the meteorite, followed by precise camera pose estimation using Agisoft Metashape. To accommodate the input requirements of different NeRF pipelines, the extracted camera parameters were converted into a standard `transforms.json` format.

Subsequently, each reconstruction method was trained on the same dataset, enabling a fair qualitative comparison in terms of training time, output fidelity, fine detail preservation, and resilience to artifacts. Instant-NGP offered extremely fast training but showed some limitations in capturing high-frequency details. Nerfacto demonstrated superior reconstruction quality, particularly for thin structures like printed text, while TensoRF, despite its longer training duration, encountered noticeable background noise artifacts.

Through this study, we aim to provide practical insights into the trade-offs between training speed and reconstruction quality across different NeRF variants, highlighting their respective strengths, limitations, and suitability for real-world 3D scene reconstruction tasks.

2. Dataset Collection and Preprocessing

The foundation of successful Neural Radiance Field (NeRF) reconstructions lies in acquiring a high-quality and diverse image dataset that captures the target object from multiple perspectives. For this study, we focused on the "ARISPE" meteorite specimen at Arizona State University (ASU), which provides a rich and detailed surface structure suitable for 3D modeling.

2.1. Data Collection

A total of **55 high-resolution images** were captured using a mobile phone camera, performing a complete **360-degree sweep** around the meteorite. Care was taken to ensure that the images were collected under uniform lighting conditions and covered a wide variety of azimuth and elevation angles. The camera was positioned at different heights and distances to capture not only the main body but also finer surface textures and nearby contextual elements such as the exhibit plate and surrounding environment.

Maintaining significant overlap between consecutive images was prioritized, which is critical for successful camera pose estimation and high-fidelity surface reconstruction. The resulting dataset provided comprehensive coverage of the object, ensuring sufficient parallax and view diversity to support neural volumetric learning. A sample from the collected dataset is shown in Figure 1, highlighting the meteorite specimen as captured during data acquisition.

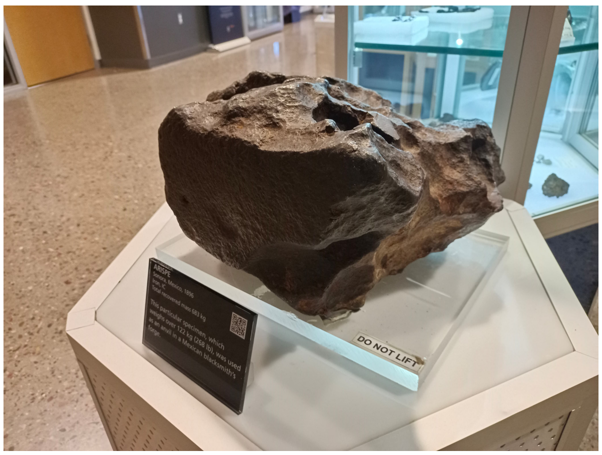


Figure 1. Sample image captured during dataset collection showing the ARISPE meteorite at ASU.

2.2. Camera Pose Estimation and Data Preprocessing

To proceed with NeRF-based reconstruction, accurate camera poses (intrinsics and extrinsics) are required for each captured image. We employed **Agisoft Metashape Professional**, a robust structure-from-motion (SfM) software suite, for this purpose. The following preprocessing steps were conducted:

- **Feature Matching and Alignment:** Feature points were detected and matched across images, enabling estimation of relative camera positions and orientations.
- **Sparse Point Cloud Generation:** During camera alignment, Metashape simultaneously generated a sparse 3D point cloud representing the underlying scene geometry.
- **Pose Export:** The calculated camera parameters were exported in `.xml` format.

To make the dataset compatible with NeRF Studio and other frameworks, the Metashape-exported camera parameters were converted into a standardized `transforms.json` file using a custom parser script. This conversion extracted each camera's intrinsic matrix, rotation matrix, and translation vector and structured them according to NeRF Studio's input format.

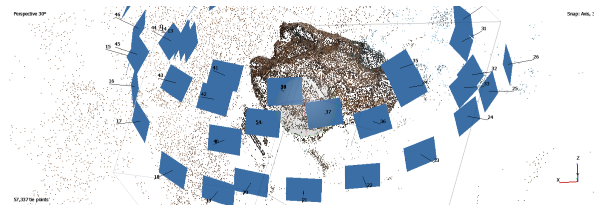


Figure 2. Sparse point cloud reconstruction and camera poses visualized in Agisoft Metashape.

An example visualization of the generated sparse point cloud and estimated camera trajectories is shown in Figure 2, demonstrating the 360-degree coverage achieved during data collection. By completing this preprocessing pipeline, we ensured that the dataset was fully prepared for training various NeRF models, providing consistent camera calibration and input formatting essential for robust 3D reconstruction.

3. 3D Reconstruction with Instant-NGP

The first reconstruction method evaluated was **Instant Neural Graphics Primitives (Instant-NGP)**, an efficient and highly accelerated neural scene representation technique. Instant-NGP achieves near real-time training speeds by leveraging a multiresolution hash grid encoding, enabling rapid volumetric scene modeling with significantly lower memory requirements compared to traditional NeRF approaches.

3.1. Training Setup

The processed dataset, including 55 images and the corresponding `transforms.json` file, was used for training. The Instant-NGP executable (`instant-ngp.exe`) was launched, and training commenced with default parameters optimized for scene reconstruction tasks.

Remarkably, training completed in approximately **1–2 minutes** using a consumer-grade **NVIDIA RTX 3050 GPU**, highlighting Instant-NGP's unparalleled speed even on modest hardware. During the early iterations, major structural elements of the ARISPE meteorite quickly emerged, and convergence was achieved without the need for extensive hyperparameter tuning.

3.2. Qualitative Results and Observations

The reconstructed 3D model generated by Instant-NGP successfully captured the global geometry, shape, and volumetric presence of the meteorite. The large dents, curvature, and overall surface roughness were reproduced with moderate fidelity, resulting in a visually plausible volumetric representation of the object.

However, Instant-NGP exhibited certain limitations:

- Fine surface textures, such as micro-dents and grain patterns, were slightly smoothed out in the reconstruction.
- The black information plate adjacent to the meteorite, containing the text ”ARISPE,” appeared heavily blurred and was not readable.
- Minor background artifacts and floating noise particles were observed around the reconstructed object, possibly resulting from incomplete scene isolation in the input images.

An example of the reconstructed output is shown in Figure 3, illustrating both the strengths and limitations of Instant-NGP for real-world object reconstruction tasks.

Instant-NGP proved to be exceptionally fast and effective for reconstructing the broader structural elements of the ARISPE meteorite. Nevertheless, it struggled to recover fine high-frequency details and thin structures such as text. These observations are consistent with the known trade-offs of Instant-NGP, which prioritizes rapid convergence and efficiency over perfect photometric accuracy or surface detail recovery. While suitable for applications requiring fast prototyping or rough visualizations, Instant-NGP may not be ideal when fine detail fidelity and textural sharpness are critical requirements.

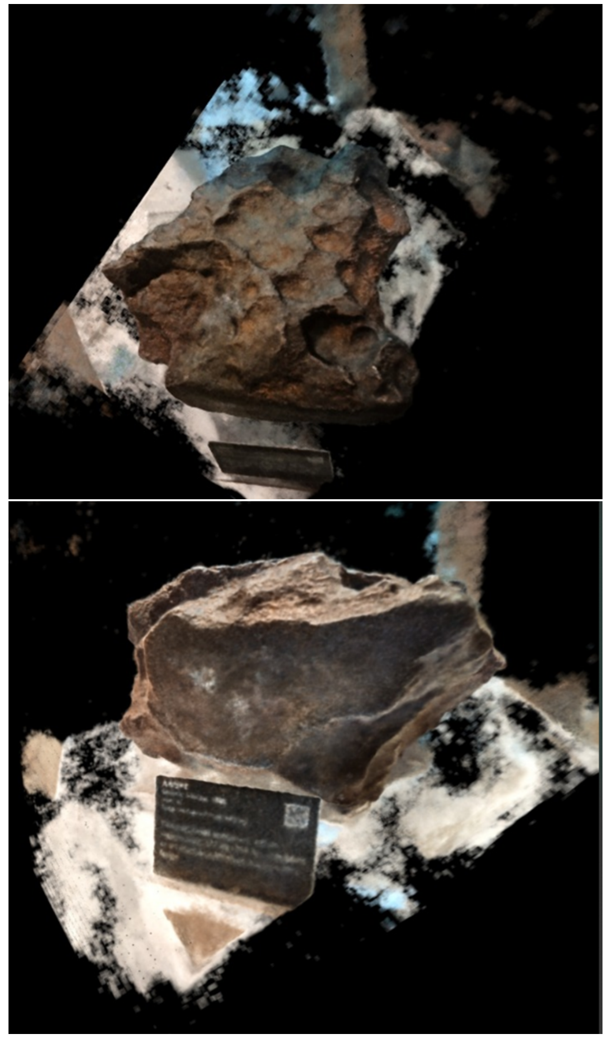


Figure 3. Rendered 3D output of the ARISPE meteorite using Instant-NGP.

4. 3D Reconstruction with Nerfacto

The second reconstruction method applied was **Nerfacto**, an enhanced neural rendering model provided within the NeRF Studio framework. Nerfacto is designed to strike a balance between training efficiency and reconstruction quality by incorporating improved proposal sampling, feature compression, and scene regularization strategies compared to conventional NeRF variants.

4.1. Training Setup

The Nerfacto model was trained using the same 55-image dataset and `transforms.json` file produced during preprocessing. The environment was carefully set up on a **Python 3.8** Conda environment, with **PyTorch 2.0.1** and **CUDA 11.7** support to ensure compatibility with the training pipeline. The training process took approximately **14–15 minutes** to complete on the **NVIDIA RTX 3050 GPU**. Despite the increased time compared to Instant-NGP, the results yielded by Nerfacto exhibited notably superior quality, validating the investment in additional training duration.

4.2. Qualitative Results and Observations

Nerfacto successfully reconstructed both the global and fine-grained features of the ARISPE meteorite with remarkable detail. As depicted in Figure 4, the following improvements were observed over Instant-NGP:

- Surface textures, including pits, dents, and roughness, were faithfully captured, resulting in a realistic surface geometry.
- The black information plate adjacent to the meteorite was reconstructed with high sharpness; the **”ARISPE”**

text and accompanying descriptions were fully readable.

- The background environment, including nearby glass enclosures and floor textures, was modeled with significantly greater clarity and consistency.

Figure 4. Rendered 3D output of the ARISPE meteorite using Nerfacto (captured at 2048 texture resolution).

The Nerfacto model demonstrated a substantial qualitative improvement over Instant-NGP, particularly in recovering fine features and maintaining scene coherence. The sharpness of the reconstructed text and the minimal background noise artifacts indicate that Nerfacto effectively leveraged multi-view consistency and hierarchical volumetric sampling to achieve high-fidelity scene modeling. While the increased training time (approximately 14–15 minutes) represents a trade-off compared to Instant-NGP's speed, the significantly improved output quality justifies this additional computational investment, especially for applications where detail and realism are paramount.

5. 3D Reconstruction with TensoRF

The third and final reconstruction method explored was **TensoRF**, a tensor factorization-based NeRF approach integrated within the NeRF Studio framework. TensoRF is designed to optimize volumetric radiance fields efficiently by leveraging low-rank tensor decompositions, offering potential benefits in memory consumption and faster convergence for large-scale 3D scenes.

5.1. Training Setup

The same dataset and preprocessing pipeline, including the 55 captured images and associated `transforms.json` file, were employed to ensure consistency across all reconstruction methods. The training was conducted within the same Conda environment previously prepared for Nerfacto. The TensoRF model required approximately **92 minutes** of training time on the **NVIDIA RTX 3050 GPU**, which was substantially longer compared to both Instant-NGP and Nerfacto.

5.2. Qualitative Results and Observations

The 3D reconstruction results obtained with TensoRF are presented in Figure 5. At a glance, the major structural features of the ARISPE meteorite were captured. However, several quality issues were also apparent:

- Significant noise particles were observed throughout the reconstructed volume, particularly in the background regions, detracting from the visual coherence of the scene.

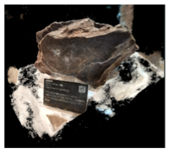
- The black plate containing the "ARISPE" description exhibited heavy distortion, and the text was largely unreadable.

Despite its theoretical advantages in memory efficiency and tensor-based volumetric representation, TensoRF underperformed compared to both Instant-NGP and Nerfacto for this specific dataset. One major contributing factor is hypothesized to be the complex and cluttered background in the real-world captures, which introduced high-frequency noise patterns. TensoRF's volumetric interpolation may have inadvertently overfit to these background artifacts, leading to degraded reconstruction quality. Moreover, the relatively low number of views (55 images) combined with small object-specific features, such as fine surface text and sharp boundaries, likely challenged TensoRF's decomposition model, which favors smoothness over discontinuities. In summary, while TensoRF successfully reconstructed the broader shape and mass distribution of the meteorite, its inability to recover fine textures and its vulnerability to noise artifacts limited its effectiveness for this particular application.

6. Comparative Analysis of Reconstruction Methods

To provide a comprehensive evaluation of the three NeRF-based reconstruction techniques — **Instant-NGP**,

Table 1. Qualitative Comparison of Different NeRF Methods

Parameters	**Instant-NGP**	**Nerfacto**	**TensoRF**
Training Time	2 minutes	16 minutes	92 minutes
Reconstruction Quality	Moderate	High	Average
Observations	Moderate surface detail Noisy background artifacts Fine text (ARISPE plate) blurred	Highly detailed reconstruction Clean background Fine text fully readable	Detailed structure reconstruction Significant noise particles Text unreadable due to artifacts
Sample Image			

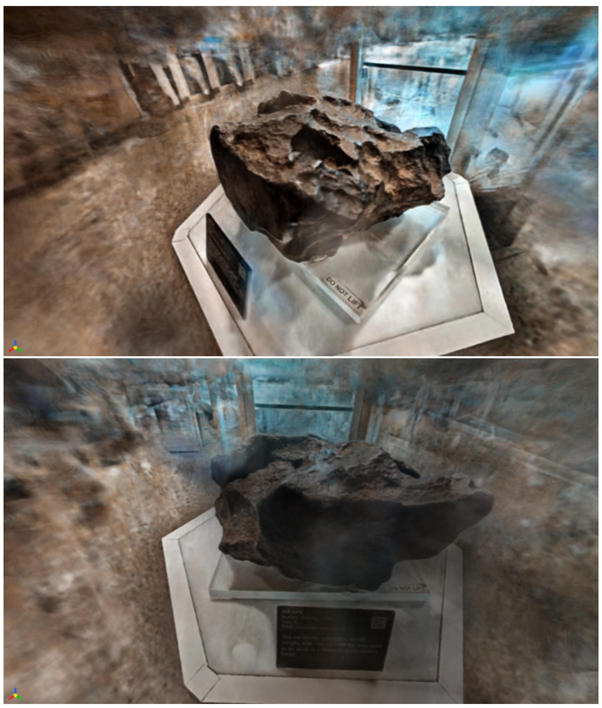


Figure 5. Rendered 3D output of the ARISPE meteorite using TensoRF (captured at 2048 texture resolution).

Nerfacto, and **TensoRF** — we conducted a detailed comparison across several qualitative and computational criteria. Each method was assessed based on training time, reconstruction quality, artifact presence, and the ability to recover fine-grained structural details.

Instant-NGP demonstrated exceptional training efficiency, completing the optimization process in approximately **2 minutes** on the NVIDIA RTX 3050 GPU. This rapid convergence highlights the benefits of the multiresolution hash grid encoding approach, which enables fast scene learning even on mid-tier hardware. Nerfacto required a moderately longer training time of approximately **16 minutes**, while TensoRF exhibited the longest training duration, taking about **92 minutes** under identical conditions. In terms of reconstruction quality, Nerfacto consistently outperformed both Instant-NGP and TensoRF. The 3D output from Nerfacto preserved sharp surface textures, produced clean background reconstructions, and successfully recovered fine features such as the printed "ARISPE" label on the adjacent black information plate. In contrast, Instant-NGP, despite achieving a visually coherent overall shape, exhibited slightly blurred surface textures and notable background noise artifacts. Particularly, small high-frequency features such as printed text appeared distorted or unreadable in the Instant-NGP outputs, likely due to its prioritization of training speed over fine-grain detail modeling. TensoRF, although theoretically designed to offer memory and computational efficiency through tensor decomposition, did not achieve superior qualitative results in this case. While it reconstructed the broader structure of the meteorite, the final output contained a substantial amount of background noise and floating artifacts. Moreover, finer details were poorly preserved, and the surrounding environment exhibited significant visual distortion. These issues suggest that TensoRF's factorization approach might be more sensitive to background complexity and inconsistencies in real-world datasets when compared to the more robust Nerfacto pipeline.

A summary of the comparative results is provided in Table 1, which outlines the training times, reconstruction quality assessments, major observations, and visual samples for each method. Overall, while Instant-NGP offered the advantage of extremely fast training and produced reasonable global scene reconstructions, Nerfacto achieved a substantially higher fidelity representation at the cost of slightly longer computation time. TensoRF, despite its theoretical benefits, was hindered by noise amplification and less effective fine-detail recovery, leading to lower visual quality. In conclusion, for the reconstruction of the ARISPE meteorite, **Nerfacto emerged as the most reliable method**, balancing training efficiency with superior reconstruction quality. Instant-NGP remains an attractive choice for quick approximations when training time is critical, while TensoRF may require more careful dataset preparation or further fine-tuning to achieve comparable results in real-world scenes. This analysis underscores the importance of selecting reconstruction strategies based on both computational constraints and the desired level of visual fidelity.

7. Conclusion

This project demonstrated the full workflow of capturing, preprocessing, and reconstructing a real-world 3D scene using multiple state-of-the-art NeRF techniques. By selecting the ARISPE meteorite as the reconstruction target, we were able to explore the challenges of real-world dataset preparation, camera pose estimation, and neural scene modeling in a practical and structured manner. A comprehensive dataset was captured, leveraging a 360-degree imaging setup to ensure diverse viewpoint coverage, and preprocessing was performed using Agisoft Metashape to obtain accurate camera poses. Three distinct NeRF-based reconstruction methods were then evaluated: Instant-NGP, Nerfacto, and TensoRF. Each method exhibited unique strengths and weaknesses. Instant-NGP offered exceptional speed, enabling rapid scene reconstruction within minutes, but at the cost of moderate background artifacts and blurred fine details. Nerfacto, though requiring a slightly longer training time, delivered the highest reconstruction fidelity, accurately recovering surface textures, fine text, and background structures. TensoRF, despite theoretical advantages in memory efficiency, struggled with noise artifacts and less reliable detail recovery when applied to this complex, cluttered real-world dataset. Through detailed qualitative comparison, it was concluded that Nerfacto provided the best balance between training efficiency and output quality, making it the most effective method for the given task. This work not only highlights the practical capabilities of modern NeRF pipelines but also emphasizes the critical role of dataset quality, camera calibration, and method selection in achieving robust and realistic 3D reconstructions. Future work may focus on improving background isolation during data capture and experimenting with hybrid or adaptive training strategies to further enhance fine-detail recovery in challenging real-world scenes.

References

- NeRF Studio Quickstart Guide: https://docs.nerf.studio/quickstart/installation.html
- Instant-NGP GitHub Repository: https://github.com/NVlabs/instant-ngp
- NeRF Studio Official Documentation: https://docs.nerf.studio/
- TensoRF GitHub Repository: https://github.com/apchenstu/TensoRF
- Agisoft Metashape Professional: https://www.agisoft.com/
- Agi2Nerf XML to transforms.json Script: https://github.com/EnricoAhlers/agi2nerf