3D Reconstructions of the Brain Using PixelNeRF Trained on MRI Scans

Hazel Lim [1]

Abstract

High-quality 3D brain MRI is clinically valuable but difficult to acquire due to long scan times, motion artifacts, and protocol-dependent intensity variation. This project investigates conditional neural radiance fields for few-view MRI view synthesis using a PixelNeRF-style pipeline. Experiments use three sources (UCSF, BraTS2020, and a whole-brain dataset) and a geometry-aware preprocessing stack (RAS reorientation, common voxel spacing, foreground cropping, and intensity stabilization). To make MRI volumes compatible with PixelNeRF training, multi-view supervision is synthesized by generating virtual 360° camera trajectories and exporting paired grayscale slice images with camera poses. Quantitative evaluation (PSNR/SSIM) on held-out subjects shows the model learns coarse anatomy but struggles with fine details under domain shifts and imperfect synthetic-view construction. These results highlight both the promise and the practical challenges of NeRF-based MRI view synthesis in heterogeneous settings.

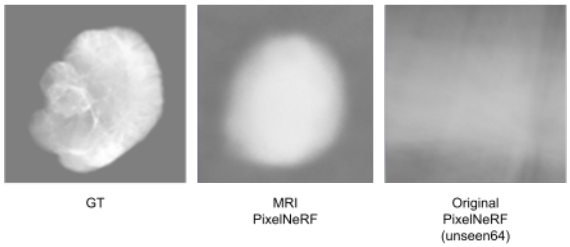


Figure 1. Teaser image comparing ground truth brain slice, output from MRI-based PixelNeRF and original PixelNeRF with unseen64 weights.

1. Introduction

Magnetic Resonance Imaging (MRI) is a cornerstone of modern medical imaging because it provides reproducible, non-invasive, quantitative tissue measurements with soft tissue contrast. These strengths make MRI widely useful for diagnosing neurological, musculoskeletal, and oncological diseases (Yang et al., 2025). In many clinical and research settings, however, the real goal is not just to view anatomy as a sequence of 2D images, but to understand it as a coherent 3D structure; this is especially the case in the brain, where subtle spatial relationships between adjacent regions and boundaries often matter for diagnosis and treatment planning. In contrast to traditional 2D planar images that only capture isolated anatomical cross-sections, 3D rendering creates interpretable spatial perspectives that can be explored along any orientation, revealing complex morphological relationships between different structures ((Liu et al., 2022)). 3D perspectives are particularly valuable in MRI because the modality is sensitive to soft-tissue differences and avoids ionizing radiation, making it well-suited for detailed neuroanatomical analysis (Awojoyogbe & Dada, 2024). This capability allows clinicians to accurately pinpoint the spread of pathology and subtleties that are invisible along individual sequential views (Ahmad et al., 2021), which is essential for tasks such as tumor staging and preoperative surgical navigation.

Despite its value, producing high-quality 3D MRI remains difficult in practice. Compared to 2D MRI, which requires collecting complete k-space data, which inherently takes time, 3D MRI requires collecting k-space measurements across three dimensions, with additional phase encoding directions and a larger number of samples per dimension to support higher resolution. This substantially increases scan time and has become a major barrier to widespread clinical use (Yang et al., 2025). Longer scans also increase vulnerability to motion-related artifacts, including breathing and cardiac pulsation effects, which can lead to inaccuracies in downstream 3D reconstructions and clinical interpretation (Godenschweger et al., 2016). More broadly, MRI workflows face practical constraints such as patient comfort, noise, scanner-specific distortions, and the difficulty of reliably modeling anatomy in photorealistic 3D (Awojoyogbe & Dada, 2024). A common workaround is to reconstruct 3D

[1]Fu Foundation School of Science and Engineering, Columbia University, New York, USA. Correspondence to: Hazel Lim <hl3849@columbia.edu>.

Proceedings of the 42^{nd} International Conference on Machine Learning, Vancouver, Canada. PMLR 267, 2025.

structure by stacking 2D slices, but this can fail to preserve important inter-slice spatial relationships (Liu et al., 2023; Wang et al., 2020), leading to inaccuracies in the reconstruction of internal structures (Yang et al., 2025). These spatial inconsistencies can result in misdiagnosis, particularly in diseases where precise tumor boundaries or vascular abnormalities are critical.

To address the broader problem of 3D reconstruction from limited 2D observations, recent work in computer vision and graphics has explored Neural Radiance Fields (NeRFs). NeRFs represent a scene as a continuous function that maps 3D coordinates (and viewing direction) to radiance values, enabling the synthesis of novel views from known viewpoints (Iddrisu et al., 2023). Compared to traditional geometry- or voxel-based reconstruction approaches, NeRFs use deep learning to produce smoother and more continuous representations of 3D spaces, handle occlusions and view-dependent effects, and reduce manual assumptions about scene geometry (Zhu et al., 2023).

The fundamental concept of a NeRF is to represent a scene as a continuous volumetric function where a radiance value is mapped to each point in 3D space. A neural network is trained to approximate this function, taking 3D coordinates as input and outputting the corresponding radiance (Tewari et al., 2022). This approach offers several distinct advantages for MRI reconstruction:

- Continuous Representation: NeRFs produce more continuous and smoother representations of 3D spaces compared to discrete voxel-based methods, handling occlusions and view-dependent effects with greater accuracy (Rabby & Zhang, 2024).
- Sparse View Synthesis: Variations like PixelNeRF (Yu et al., 2021) and MedNeRF (Corona-Figueroa et al., 2022) have been developed to predict continuous neural scene representations from just a few input images. This potentially minimizes the need for multiple or high-resolution scans that might expose patients to discomfort or excessive radiation (Corona-Figueroa et al., 2022).

This is particularly useful in medical imaging, where detailed visualization of anatomical structure is crucial for diagnosis and surgical planning (Wang et al., 2025). At the same time, NeRFs cannot be directly applied to medical reconstruction without adaptation, because medical images contain unique internal details and behave differently than natural scenes, and their data constraints and acquisition process are fundamentally different (Wang et al., 2025; Iddrisu et al., 2023). In MRI specifically, NeRF-based approaches face additional challenges due to the high dimensionality of imaging data, missing or low-contrast information, and the computational cost of processing large 3D volumes (Iddrisu et al., 2023).

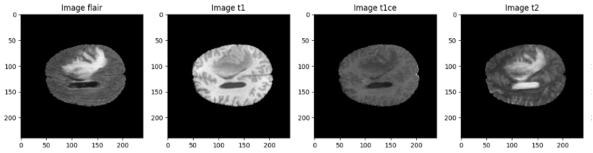


Figure 2. Four common modalities of a brain MRI (ordered accordingly): FLAIR (emphasizes tumors), T1 (focuses on overall structure), T1w (emphasizes anatomical detail with fat appearing bright and water dark), T2 (emphasizes water content and tumor).

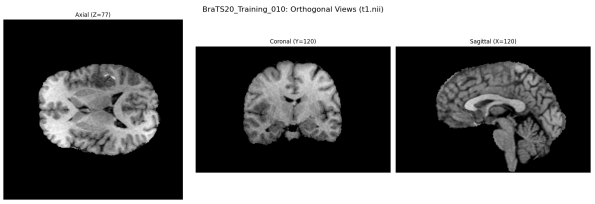


Figure 3. Standard anatomical views of a brain MRI (ordered accordingly): axial, coronal, and sagittal.

To implement NeRFs effectively in this domain, one must understand the underlying physics of the data. MRI works by utilizing radiomagnetic pulses to measure the response of atomic nuclei within the body. It relies on the unique sensitivity of these pulses to changes in soft tissue to generate images without ionizing side effects (Westen et al., 2019). MRI measures properties of tissue by using a strong magnetic field and radiomagnetic (radiofrequency) pulses that perturb hydrogen nuclei and then record how signals evolve as the nuclei return toward equilibrium (Broadhouse, 2019). Because different tissues exhibit different signal behaviors, MRI can produce multiple modalities—commonly including T1-weighted and T2-weighted images—where anatomy and pathology may appear with different contrast depending on acquisition parameters (Yang et al., 2025) (Figure 2 compares the four common MRI modalities.). In practice, 3D MRI is usually examined through three orthogonal planes (Figure 3 illustrates the three standard MRI views.): 1) axial: the top-down cross-sectional view; 2) coronal: the frontal view; and 3) sagittal: the side profile view.

Correctly interpreting these planes also depends on consistent coordinate conventions; orientation standards such as RAS (Right, Anterior, Superior) help ensure that volumes from different sources can be aligned and compared reliably, which is essential when constructing a unified 3D pipeline across multiple datasets.

Project Overview. In this work, I adapt pixelNeRF for medical imaging by introducing two key modifications to improve 3D brain reconstruction from MRI scans. First,

I replace the ImageNet-pretrained encoder with a medical imaging-specific encoder trained on MRI data through self-supervised learning. This adaptation enables the network to extract features optimized for grayscale medical images rather than natural RGB photographs. Second, I modify the output layer to predict grayscale intensity values instead of RGB colors, reducing model complexity and aligning with the grayscale nature of MRI data. I train my model on synthetic 360-degree views generated from T1-weighted contrast-enhanced MRI scans across three datasets: BraTS (brain tumor segmentation), UCSF (clinical brain scans), and whole-brain structural imaging. My experiments demonstrate that domain-specific encoder pre-training can improve reconstruction quality compared to using off-the-shelf computer vision features, validating the importance of tailoring neural rendering architectures to medical imaging characteristics.

Moreover, this report presents my end-to-end workflow: collecting multi-source brain MRI data, preprocessing it into a consistent orientation and voxel spacing, and generating synthetic multi-view slices to train NeRF-based models to learn a continuous representation of brain structure. The goal is not to claim clinical readiness, but to demonstrate a technically grounded attempt to adapt modern neural rendering methods to the constraints and structure of MRI data, while identifying practical limitations and next steps for improving training stability and fidelity.

2. Related Work

2.1. Neural Radiance Fields (NeRF) and Extensions of NeRF

The field of neural rendering was significantly advanced by the introduction of Neural Radiance Fields (NeRF) by Mildenhall et al. (2020). NeRF proposes a continuous volumetric scene representation where a fully connected deep network (MLP) encodes the scene as a 5D function of spatial location (x, y, z) and viewing direction (θ, ϕ). The network predicts the volume density (σ) and view-dependent emitted radiance (c) at any given point. To synthesize novel views, NeRF employs differentiable volume rendering, accumulating color and density along rays cast through the pixel grid. While NeRF achieves photorealistic results for complex scenes, the original implementation suffers from slow training and inference times due to the requirement of querying the massive MLP millions of times per image (Mildenhall et al., 2020).

To address the computational bottleneck of the original MLP-based NeRF, subsequent works have introduced hybrid representations that combine neural networks with explicit data structures. Notably, Instant-NGP (Müller et al., 2022) utilizes a multi-resolution hash grid encoding to significantly reduce the size of the MLP, enabling training in seconds rather than days. Similarly, TensoRF (Chen et al., 2022) factorizes the 4D radiance field into compact vector-matrix tensor components, achieving high compression rates and fast rendering without sacrificing visual quality.

A major limitation of the original NeRF is that it overfits to a single scene and lacks generative capabilities. Generative NeRFs integrate the adversarial training paradigm of GANs with the volumetric consistency of NeRFs. GRAF (Schwarz et al., 2021) and π-GAN (Chan et al., 2021) were among the first to condition the radiance field on a latent code z, allowing for the generation of multi-view consistent images of novel instances. More recently, EG3D (Chan et al., 2022) introduced an efficient tri-plane representation that enables high-resolution 3D-aware image synthesis, addressing the "mode collapse" and consistency issues prevalent in earlier 2D GANs by enforcing explicit 3D geometry in the generator.

While standard NeRFs require per-scene optimization, generalizable methods aim to synthesize novel views of unseen scenes without retraining. PixelNeRF (Yu et al., 2021) and MVSNeRF (Chen et al., 2021) condition the rendering on local image features extracted from one or more input views. By projecting query points into the feature space of the input images, these models learn a geometry prior across datasets, allowing them to perform zero-shot or few-shot reconstruction on novel objects, similar to how human vision infers 3D structure from limited visual cues.

2.2. Neural Graphics for Neuroimaging

Early work on 3D-aware deep learning for brain MRI focused on volumetric modeling rather than NeRF. Bengs et al. (2021) proposed 3D variational autoencoders with 3D input erasing for unsupervised anomaly detection in brain MRI and showed that 3D models significantly outperform 2D slice-based VAEs for lesion segmentation, underlining the value of volumetric context in neuroimaging (Bengs et al., 2021).

Corona-Figueroa et al. (2022) then introduced MedNeRF, a GAN-based radiance field that reconstructs CT volumes from few or even single-view X-ray projections, learning a continuous representation that disentangles shape and volumetric depth of internal anatomy (Corona-Figueroa et al., 2022). Follow-up variants such as uncertainty-aware MedNeRF (UMedNeRF) and alignment-and-correction NeRF (ACNeRF) further stabilize training by automatically balancing clarity vs. accuracy and better aligning NeRF to medical imaging geometry (Hu et al., 2024; Sun et al., 2024).

Building on this, Iddrisu et al. (2023) applied NeRF directly to brain MRI, reconstructing 3D projections from 2D slices using 3D CNNs, slice interpolation, and multi-modal

MRI to handle variable slice thickness and capture both surface and internal brain structures (Iddrisu et al., 2023). Around the same time, Awojoyogbe and Dada (2024) argued more broadly that conventional 2D MRI visualization underutilizes the intrinsic 3D complexity of the data and framed NeRFs as a way to enable richer, clinically meaningful 3D exploration (Awojoyogbe & Dada, 2024).

Wang et al. (2024) surveyed NeRF applications across multiple organs, including brain, skull, vessels, and chest, emphasizing challenges such as imaging physics, sparse views, and defining medically meaningful densities and colors (Wang et al., 2025). Most recently, Yang et al. (2025) proposed a slice-to-volume and fusion reconstruction (SVFR) framework that uses 2D undersampled slices along three orthogonal planes, 2D vision transformers, and a volume-wise fusion module to efficiently reconstruct 3D MR images while preserving fine spatial details (Yang et al., 2025). Taken together, these works show a progression from generic 3D deep learning to NeRF-style radiance fields tailored to medical data; my project follows this trajectory by using a conditional NeRF (PixelNeRF) on brain MRI slices to explore how far slice-based neural rendering can go in reconstructing coherent 3D brain structure from heterogeneous clinical datasets.

2.3. Gaps in current research and the positioning of my work

Despite recent progress in applying NeRF-style models to medical imaging, several practical gaps remain between what is demonstrated in controlled settings and what is needed for robust neuroimaging reconstruction in real-world data.

A first gap is that many medical NeRF systems still assume dense or well-structured supervision—either access to full 3D volumes, a relatively complete set of views, or carefully curated acquisition conditions. In brain MRI, this assumption is often violated. Clinical datasets frequently contain partial coverage, inconsistent slice spacing, missing sequences, and protocol-dependent differences that make it difficult to treat each scan as a clean, camera-captured multi-view scene. While recent work has begun exploring sparse-view or slice-based reconstruction, the training setups often rely on a single dataset distribution or a fixed imaging protocol, which can overestimate generalization when deployed on scans from different sites or with different preprocessing conventions.

A second gap concerns heterogeneity across MRI modalities and scanners. Unlike natural images, MRI intensities are not standardized, and "appearance" depends heavily on sequence parameters and reconstruction pipelines. Many existing methods either (i) train on one modality at a time, (ii) treat intensity values as if they behave like RGB channels, or (iii) focus on sequence translation without explicitly addressing how multi-sequence contrast differences interact with a radiance-field-style representation. As a result, it is still unclear how well conditional neural rendering approaches can learn a stable representation that remains usable across datasets where anatomy may be framed differently (e.g., skull present vs. absent) or where intensity distributions shift substantially.

This project is positioned directly in this gap space: it treats brain MRI as a few-view conditional rendering problem and asks what is required to make a PixelNeRF-style model train and remain stable under realistic neuroimaging variability. Rather than assuming a single clean dataset, the project intentionally works across UCSF, BraTS2020, and a whole-brain dataset, using this mixture to stress-test robustness under domain shifts such as differing anatomical context (e.g., skull and eye sockets), intensity scaling differences, and artifacts like motion-related noise. The approach emphasizes practical steps that are often under-specified in prior work—standardizing orientation (e.g., RAS), resampling to common voxel geometry, intensity normalization, and generating consistent synthetic multi-view poses—so that conditional NeRF training becomes feasible on heterogeneous medical slice collections.

3. Method

This section describes the architectural modifications and training procedures developed to adapt pixelNeRF for medical imaging applications. The approach consists of three primary components: (1) domain-specific encoder pre-training using self-supervised learning on MRI data, (2) architectural modifications to handle single-channel grayscale inputs and outputs, and (3) a training pipeline optimized for 3D brain reconstruction from multi-view MRI projections. Each component addresses fundamental differences between natural image understanding and medical image analysis. The encoder pre-training phase teaches the network to extract clinically relevant features from MRI scans rather than relying on features learned from everyday photographs. The architectural changes reduce model complexity while maintaining the capacity to learn volumetric representations of brain anatomy. Together, these modifications enable pixelNeRF to reconstruct high-fidelity 3D brain volumes from sparse MRI views while preserving pathological features critical for clinical interpretation.

3.1. Medical Image Encoder Pre-training

The standard pixelNeRF implementation provided by Yu et al. (2020) utilizes ResNet encoders pre-trained on ImageNet, which are optimized for natural RGB images (Yu et al., 2021). To adapt the encoder for medical imaging, a self-supervised pre-training task was designed specifically for

MRI data.

I modified the ResNet-34 encoder's first convolutional layer to accept single-channel grayscale input rather than RGB. This reflects MRI physics: T1-weighted sequences encode tissue properties (proton relaxation times) as univariate intensity values, not color. High intensities indicate fat and white matter, intermediate values represent gray matter, and low intensities correspond to cerebrospinal fluid. Replicating this single measurement across three RGB channels would create redundant representations, wasting encoder capacity. Single-channel input focuses learning on medically meaningful intensity patterns. To preserve ImageNet knowledge, I initialized the grayscale layer by averaging pre-trained RGB weights. This retains learned edge detection capabilities while accommodating single-channel input, leveraging transferable low-level features (edges, textures, gradients) common to both natural and medical images.

Pre-training employed a denoising autoencoder objective where the encoder learned to extract robust features from corrupted MRI slices. For each training sample, an MRI slice was corrupted by: (1) adding Gaussian noise ($\sigma = 0.05$), and (2) randomly masking 15% of pixels. The encoder processed corrupted images to generate feature representations, which were then decoded through a lightweight convolutional decoder to reconstruct the original clean image. The training objective minimized mean squared error between reconstructed and original slices. This self-supervised task encourages the encoder to learn representations invariant to imaging noise and acquisition artifacts common in clinical MRI, while capturing the underlying anatomical structure.

The encoder was trained for 10 epochs with a batch size of 32, using the Adam optimizer with learning rate 1×10^{-3}. The choice of 10 epochs balanced training time constraints with sufficient exposure to diverse MRI samples—preliminary experiments showed that validation loss plateaued after 8-10 epochs, indicating convergence of the self-supervised task. A batch size of 32 was selected to maximize GPU memory utilization while maintaining stable gradient estimates; smaller batches introduced training instability while larger batches exceeded memory limits when processing 128×128 pixel MRI slices through ResNet-34. Different learning rates were applied to different network layers: early layers (layer1, layer2) used 1×10^{-4} to preserve low-level features from ImageNet initialization, while later layers (layer3, layer4) and the modified first convolutional layer used higher learning rates (5×10^{-4} and 1×10^{-3} respectively) to adapt more rapidly to medical imaging characteristics.

3.2. MRI-based Modifications PixelNeRF Architecture & Pipeline

In addition to the encoder, output layer modifications were made to the standard pixelNeRF architecture to accommodate medical imaging. The MLP output was reduced from four channels (R, G, B, σ) to two channels (intensity, σ). The first channel represents grayscale intensity while the second represents volume density. Critically, the shift from RGB (3-channel) to grayscale (single-channel) representation attempts to align with the physics of MRI acquisition. MRI scanners measure proton density and relaxation properties at each voxel, producing inherently univariate intensity values that encode tissue characteristics—white matter, gray matter, cerebrospinal fluid, and pathological tissues each have distinct intensity signatures in T1-weighted sequences. By using single-channel output, the model focuses its representational capacity on learning the continuous intensity manifold that characterizes brain tissue. This architectural choice reduces the MLP output dimensionality from four to two, decreasing parameters by 33% while enabling more efficient learning of the intensity-density relationship fundamental to medical volume rendering. As shown in Figure 4 and Figure 5, PixelNeRF conditions a NeRF-style MLP on image features extracted by projecting 3D query points into the encoded feature volume.

The volume rendering equation remained unchanged, with the alpha-compositing formula applied to single-channel intensity values:

$$C(\mathbf{r}) = \sum_{i=1}^{N} T_i \alpha_i c_i$$

, where c_i is the predicted grayscale intensity, $\alpha_i = 1 - \exp(-\sigma_i \delta_i)$ is the alpha value derived from volume density σ_i and ray segment length δ_i, and $T_i = \prod_{j=1}^{i-1}(1 - \alpha_j)$ is the accumulated transmittance.

Figure 4 shows the PixelNeRF architecture proposed by Yu et al. (Yu et al., 2021). Figure 5 shows the architecture for the MRI-based PixelNeRF. As shown in Figure 4, the standard PixelNeRF pipeline involves conditioning a NeRF-style MLP on image features extracted by projecting 3D query points into the encoded feature volume. In this project, the overall rendering and conditioning pipeline remains the same, but it is applied to MRI slices rather than natural RGB images. The MRI-based PixelNeRF is not architecturally starkly different. As explained previously, the primary adaptations occur at the representation level: the encoder is treated as a medical-imaging feature extractor, and the radiance head is reinterpreted to predict grayscale MRI intensity (and volumetric density) instead of RGB color. Because these changes preserve the same core data flow—feature projection, MLP prediction, and differentiable volume rendering—the original architecture

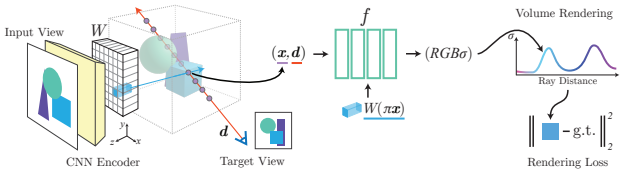


Figure 4. **PixelNeRF architecture (single-view).** A query point $\mathbf{x}$ on a target ray is projected into the encoded feature volume W to retrieve an interpolated local feature. The NeRF network f takes this feature (and $\mathbf{x}$, with view direction $\mathbf{d}$) to predict density and appearance, which are then volume-rendered and compared to the target pixel. All quantities are defined in the input-view camera frame. Adapted from Yu et al. (Yu et al., 2021).

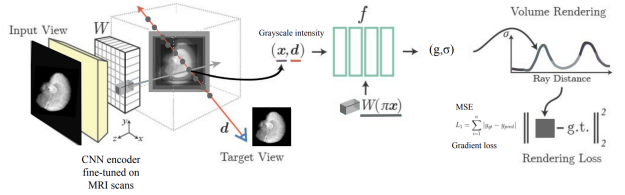


Figure 5. **MRI-based PixelNeRF architecture (single-view).** CNN encoder fine-tuned on MRI-based images. Network focuses more on grayscale intensities rather than RGB values. Different loss functions applied at different training stages. Adapted from Yu et al. (Yu et al., 2021).

diagram is still an accurate depiction of the method and does not require structural modification for the MRI-based setting.

3.3. PixelNeRF Training Configuration

The modified pixelNeRF model was trained using randomly selected source views from the synthetic 360-degree dataset. For each training iteration, 1-3 source views were randomly chosen and encoded through the medical encoder to generate 3D feature representations. Target rays were sampled from remaining views (excluding source views), and the MLP predicted intensity and density values along each ray.

Hierarchical sampling was employed with 128 coarse samples and 128 fine samples per ray following the original NeRF approach. Loss functions included mean squared error for both coarse and fine networks, weighted equally ($\lambda_{coarse} = \lambda_{fine} = 1.0$). The model was trained with batch size 4 (four brain volumes per batch) and ray batch size 128 (rays sampled per volume), using the Adam optimizer with initial learning rate 1×10^{-4}.

Training was conducted on a single Tesla T4 GPU for approximately 72 hours. The encoder remained frozen for the initial training phase to stabilize MLP learning, after which optional fine-tuning with unfrozen encoder could be performed if validation metrics plateaued.

3.4. Evaluation Metrics

Reconstruction quality was assessed using Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) computed between predicted and ground-truth MRI slices. Higher PSNR (e.g., $>$ 30 dB) and SSIM (e.g., $>$ 0.90) values indicate superior reconstruction quality. Visual inspection of rendered views was also performed to assess anatomical plausibility and the preservation of clinically relevant structures, including pathological features such as lesions.

4. Experiments

4.1. Dataset Collection

Experiments were conducted on three MRI sources selected to capture both standardized benchmark data and clinically realistic variability in acquisition protocols, contrast, and field-of-view. All volumes were handled in NIfTI format (.nii or .nii.gz) and loaded as 3D arrays using common neuroimaging tooling (e.g., `nibabel.load(...).get_fdata()`), with metadata (affine matrices and header spacing) retained to support geometry-aware preprocessing.

The UCSF dataset consists of 461 brain MRI slices sampled from clinical scans (Rudie et al., 2024). This dataset includes multiple volumes per case that are commonly encountered in practice, such as ‘T1pre‘, ‘T1post‘, ‘FLAIR‘, ‘subtraction‘, and derived/segmentation volumes (e.g., ‘seg‘, ‘BraTS-seg‘). The presence of multiple contrasts and derived channels introduces intensity-range inconsistency across cases and modalities, and also increases the likelihood of scanner-dependent artifacts (bias field, motion, variable background signal), which motivated standardization steps prior to training.

The BraTS 2020 dataset consists of 369 slices extracted from the BraTS volumes and includes the canonical multimodal set (‘t1‘, ‘t1ce‘, ‘t2‘, ‘flair‘) alongside ‘seg‘ (Menze et al., 2015; Bakas et al., 2017; 2019). Compared to UCSF, BraTS is more standardized in terms of volume alignment and naming, but still exhibits non-trivial variability in contrast, tumor-related structure, and slice-to-slice continuity. This dataset was included to evaluate whether the approach remains stable on a widely used benchmark distribution.

The whole-brain dataset contains 30 slices and includes ‘T1w‘, ‘T2w‘, ‘FLAIR‘, ‘ADC‘, and ‘Trace‘ (Zhao et al., 2024). In contrast to UCSF and BraTS, this dataset includes broader anatomical context (e.g., skull and eye sockets), which changes the distribution of “foreground” anatomy and increases the amount of high-contrast boundary structure. This dataset was treated as a small but challenging set that

tests how robust the preprocessing and view synthesis are when non-brain structures remain present.

Table 1 summarizes the MRI datasets used in this project, including the total number of 2D slices extracted per source and the imaging modalities available. In total, the study includes 860 slices across three sources: UCSF (461 slices), BraTS 2020 (369 slices), and a whole-brain dataset (30 slices). Each dataset provides different MRI sequences, enabling experiments under heterogeneous acquisition settings.

Table 1. Summary of MRI Dataset Sources and Available Modalities

Dataset Source	Total Slices	Available Modalities / Volumes
UCSF	461	BraTS-seg, FLAIR, seg, subtraction, T1post, T1pre, T2Synth
BraTS 2020	369	flair, seg, t1, t1ce, t2
Whole-brain	30	ADC, FLAIR, T1w, T2w, Trace
GRAND TOTAL	**860**	**Multi-source, multi-sequence MRI**

4.2. Data Preprocessing & Synthetic Multi-View Generation

MRI volumes originating from different sites and datasets frequently differ in orientation conventions, voxel spacing, field-of-view, and intensity scaling. To reduce domain shift and ensure that synthetic "camera" views correspond to consistent anatomical directions, all datasets were processed using a geometry-aware pipeline that (1) standardizes orientation, (2) standardizes voxel geometry, (3) reduces irrelevant background, (4) normalizes intensity ranges, and (5) generates multi-view training tuples compatible with PixelNeRF-style conditioning.

Orientation standardization (RAS) Each volume was reoriented into a consistent anatomical coordinate convention (RAS). In practice, this step is typically implemented using `nibabel` orientation utilities, either by converting to a canonical orientation (`nib.as_closest_canonical`) or by explicitly computing and applying orientation transforms with `nibabel.orientations` (e.g., `io_orientation`, `axcodes2ornt`, `ornt_transform`, and `apply_orientation`). The objective is to ensure that axes correspond consistently to right–left, anterior–posterior, and superior–inferior directions, so that axial/coronal/sagittal slicing is semantically consistent across datasets. (Figure 6 illustrates the effects of the preprocessing stage 1 by showing the before and after slices.)

Resampling to a common voxel spacing and grid. Volumes were resampled to a shared voxel spacing so that anatomical structures appear at comparable scales across UCSF, BraTS, and whole-brain. This step can be implemented using either `nibabel.processing.resample_to_output` / `resample_from_to` (affine-aware resampling) or

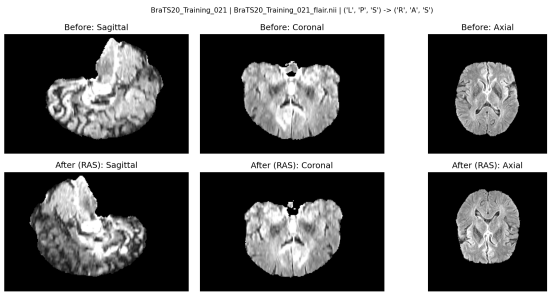


Figure 6. First step in the preprocessing: orientation standardization. The top row shows the original slice that was in LPS orientation (Left, Posterior, Superior). The bottom row shows the correct RAS orientation used in this project. Note: these two slices are not the same slices (they come from the same patient but at different z-levels).

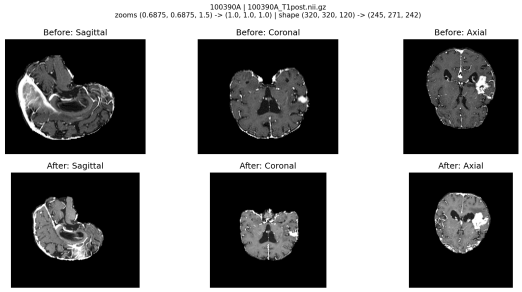


Figure 7. Second step in the preprocessing: Resampling to common voxel spacing. Top row shows slices originally being in (sx, sy, sz) = (0.5, 0.5, 2.0) mm dimensions. Bottom row shows transformation effects: (1.0, 1.0, 1.0) mm,with consistent cube-like voxels. Note: these two slices are not the same slices (they come from the same patient but at different z-levels).

medical-imaging resampling utilities such as SimpleITK's `ResampleImageFilter`, or MONAI transforms (e.g., `monai.transforms.Spacing`). The key requirement is that resampling respects the voxel-to-world affine so that geometry remains consistent; naive `scipy.ndimage.zoom` can work for quick prototypes but is less reliable unless carefully paired with correct spacing handling. Interpolation mode is typically chosen based on content: linear interpolation for intensity volumes and nearest-neighbor interpolation for discrete label maps (segmentations), to avoid label corruption.(Figure 7 illustrates the effects of the resampling by showing the before and after slices.)

Foreground cropping and field-of-view reduction. After orientation and spacing are standardized, a foreground crop is applied to reduce empty background regions and focus learning on anatomy. A common approach is to compute a simple brain/foreground mask by thresholding or percentile-based gat-

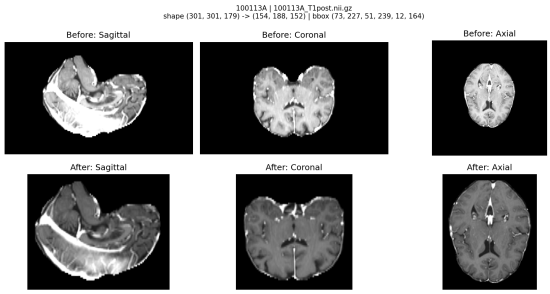


Figure 8. Third step in the preprocessing: Cropping out excessive blank space. Top row shows unnecessary blank space surrounding the sides of the images. Bottom row shows the cropping effect while preserving certain amount of margin around the brain. Note: these two slices are not the same slices (they come from the same patient but at different z-levels).

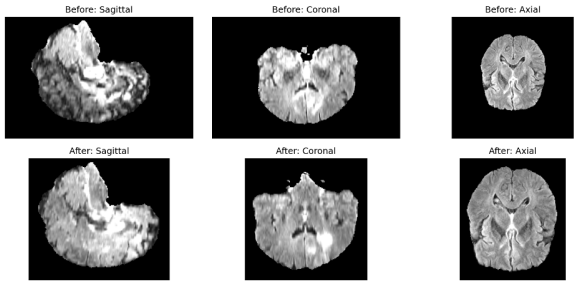


Figure 9. Fourth step in the preprocessing: Normalizing high-intensity contrasts. Top row shows original slices, whilst bottom row shows the slight intensity stabilization that was done (not too much stabilization to ensure details are preserved). Note: these two slices are not the same slices (they come from the same patient but at different z-levels).

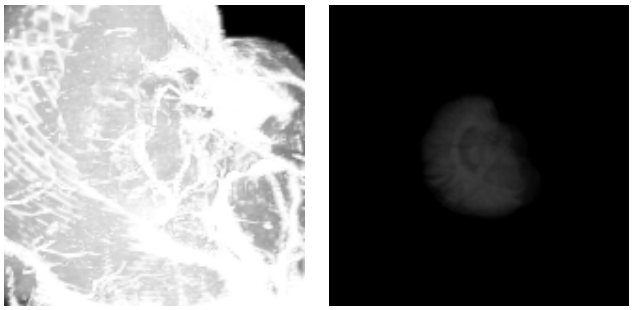
ing of intensities, followed by lightweight morphology (e.g., `scipy.ndimage.binary_closing`, `binary_opening`, and connected-component filtering with `scipy.ndimage.label`) to remove isolated noise. The bounding box of the largest connected component (or a robust intensity-based bounding region) is then used to crop the volume. This substantially reduces wasted ray samples during training because fewer rays traverse only background, and it also stabilizes intensity normalization by restricting statistics to anatomically meaningful regions. For the whole-brain dataset, the crop is intentionally less aggressive (or uses different thresholds) to avoid discarding skull/ocular structures when present, since these structures substantially affect intensity distributions and boundaries. (Figure 8 illustrates the effects of the before and after cropping the slices.)

Intensity normalization. MRI intensities are not standardized across scanners and protocols, and the same modality can exhibit different dynamic ranges across cases. To reduce this variability, intensities were normalized per volume. A robust approach is to clip intensities using percentiles (e.g., using `numpy.percentile` to clip to a low/high range such as [0.5, 99.5]) and then scale to a consistent range such as [0, 1] using min–max normalization. Alternatively, z-score normalization can be performed within the foreground mask (mean/std computed over brain voxels), which is showing strong stability when background varies. The specific choice is less important than enforcing consistency across datasets and ensuring that outlier intensities do not dominate model loss. Segmentations (when used) are kept in label space and are not intensity-normalized. (Figure 9 illustrates the effects of the before and after intensity normalization.)

Synthetic view construction for PixelNeRF compatibility. PixelNeRF training requires multiple "input views" with known camera poses and intrinsics. Because MRI volumes are not acquired as conventional camera images, multi-view supervision was synthesized by treating the 3D volume as a canonical object and generating a set of virtual camera poses around it. For each pose, a corresponding 2D slice (or rendered projection, depending on the chosen formulation) is generated and saved as an image along with its camera-to-world transform. Poses are stored as 4×4 matrices (e.g., `poses_c2w.npy` as `float32`) and correspond to the virtual camera orientation and translation in a consistent world coordinate frame. In many NeRF codebases, camera trajectories are parameterized with spherical coordinates (azimuth/elevation/radius) using helper functions such as `pose_spherical(...)`, producing evenly spaced viewpoints for 360° coverage. (Figure 10 illustrates the different experimental stages done to get final 360 degree synthetic slices. The first trial shows a very close-up and narrow field of view of the slice. The second shows the opposite. The third trial is able to capture the appropriate field of view, but is too high in brightness and loses detail. The fourth trial is the balance of all four previous trials.)

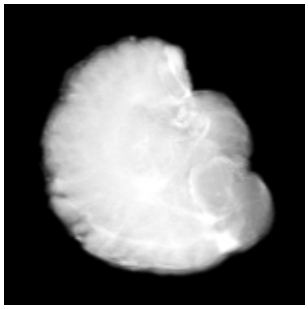
The resulting image set plus pose set forms one training "object" directory, typically structured as an `images/` folder and one or more metadata arrays (`poses_c2w.npy`, and optionally intrinsics if not fixed). For compatibility with common NeRF training code, grayscale slices are usually written as PNGs (e.g., using `imageio.imwrite` or `PIL.Image.save`) after scaling to [0, 255], while maintaining a separate floating-point representation for quantitative checks when needed. (Figure 11 illustrates the final synthetic slices that were rendered at different angles and were fed to train the model.)

Each synthetic view was rendered at 160×160 pixel reso-

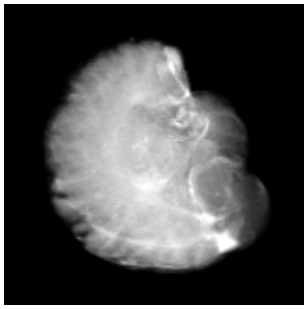


(a) Trial 1

(b) Trial 2



(c) Trial 3



(d) Trial 4

Figure 10. **Synthetic 360° slice generation trials.** Four attempts at generating multi-view synthetic brain slices for PixelNeRF training. Each trial varies preprocessing and/or camera-pose sampling, and the resulting rendered views are shown for qualitative comparison.

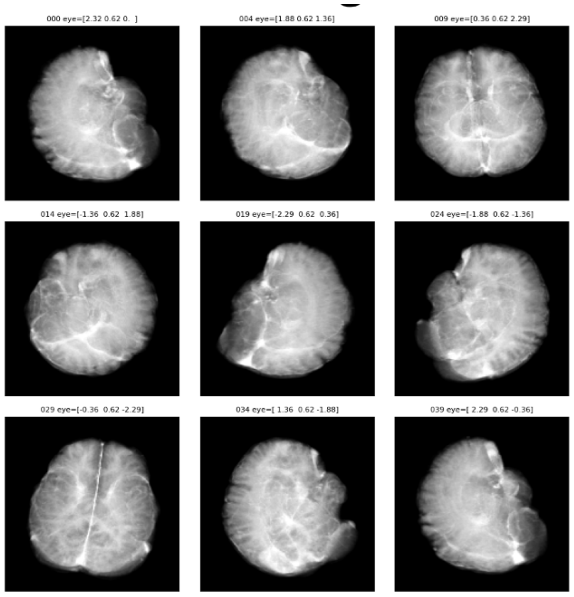


Figure 11. Final synthetic slices captured at different angles. These were used to train the PixelNeRF network.

lution with corresponding 4×4 camera extrinsic matrices (rotation and translation) and intrinsic parameters (focal length, principal point). The dataset was split at the patient level with 80% allocated to training, 10% to validation, and 10% to test sets, ensuring no patient's data appeared across multiple splits. This split strategy prevented information leakage and enabled robust evaluation of the model's generalization to unseen patient scans.

This preprocessing pipeline yields a consistent, multi-view representation across UCSF, BraTS, and whole-brain: each case becomes a standardized 3D volume in a common orientation and voxel grid, paired with a set of synthetic views and camera poses that allow PixelNeRF-style conditioning and novel-view slice synthesis.

4.3. Setup for Training PixelNeRF

The modified pixelNeRF model was trained using a multi-stage approach designed to leverage domain-specific encoder features while learning 3D spatial representations from synthetic multi-view MRI projections.

4.3.1. Training Procedure

Training was conducted in two distinct phases with different objectives: an initial geometry learning phase focused on stable 3D structure acquisition, followed by a detail refinement phase that improved fine anatomical feature reconstruction.

Phase 1: Geometry Learning with Frozen Encoder

The first training phase prioritized learning robust 3D spatial representations while preserving pre-trained medical imaging features. The encoder weights were frozen (via the `--freeze_enc` flag) to prevent catastrophic forgetting of MRI-specific features acquired during self-supervised pre-training. This architectural constraint forced the MLP components to focus exclusively on learning the mapping from encoded image features to volumetric density and intensity predictions.

During this phase, the model was trained for approximately 5,500 epochs using the base configuration with 64 coarse samples and 32 fine samples per ray. Training exhibited stable convergence with total loss decreasing from 0.16 to 0.005 over the course of training. Visual inspection of intermediate outputs (sampled every 50 batches) revealed that the model successfully learned overall brain morphology—ventricles, major tissue boundaries, and gross anatomical structure were recognizable by epoch 1,000. However, fine details such as cortical sulci, gyri patterns, and subtle intensity variations remained blurred or under-resolved.

Quantitatively, this phase achieved PSNR values of 22-24 dB on validation samples, indicating reasonable but sub-

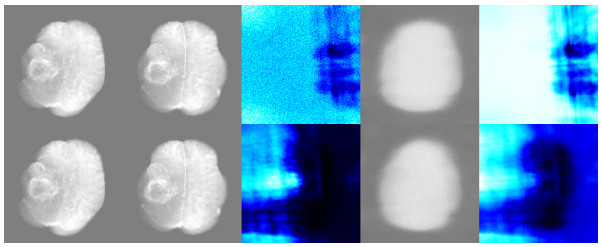


Figure 12. Phase 1: Iteration 4300 in training. Top row shows coarse network outputs. Bottom row shows fine network outputs. Columsn 1-5 (in order): Source view 1 (first input image to the model), source view 2 (second input image to the model), depth map (predicted depth/distance from camera (blue colormap), grayscale prediction (model's novel view synthesis output), alpha map (opacity/density prediction (blue colormap)

optimal reconstruction quality. The RGB intensity ranges (min: 0.45-0.47, max: 0.92-0.96) suggested the model learned appropriate intensity scaling, but alpha (opacity) values showed limited dynamic range (0.40-0.98), indicating incomplete learning of fine geometric detail. These observations motivated architectural and loss function refinements in the subsequent phase.

12 shows what the visuals looked like in the middle of phase 1 training.

Phase 2: Detail Refinement with Enhanced Sampling and Loss Functions The second training phase focused on improving fine anatomical detail through three key modifications: increased ray sampling density, edge-preserving loss augmentation, and encoder fine-tuning.

Sampling enhancement. The renderer configuration was modified to increase sampling resolution from 64 to 128 coarse samples and from 32 to 128 fine samples per ray. This four-fold increase in fine sampling density enabled the model to better capture high-frequency anatomical features such as tissue boundaries and small structures. The computational cost increased proportionally (training speed reduced from $\sim$1.1 s/iteration to $\sim$3.0 s/iteration), but preliminary experiments indicated that the increased sampling density was necessary for preserving fine detail.

Loss function augmentation. The RGB loss was changed from L2 (mean squared error) to L1 (mean absolute error) in the configuration file (`use_l1 = True` for both `rgb` and `rgb_fine`). L1 loss is less sensitive to outliers and better preserves sharp transitions, which is critical for maintaining tissue boundaries in medical images. Additionally, an edge-preserving gradient loss was incorporated:

```
def gradient_loss(self, pred, target):
    pred_diff = pred[:, :, 1:] - pred[:, :,
        :-1]
    target_diff = target[:, :, 1:] - target
        [:, :, :-1]
    return F.l1_loss(pred_diff, target_diff)
```

This loss term, weighted at 0.05 relative to the reconstruction loss, penalized inconsistent intensity gradients along ray samples, encouraging the model to maintain smooth tissue regions while preserving sharp anatomical boundaries. The gradient loss was applied only when the fine network was active to avoid interfering with coarse geometry learning.

Encoder fine-tuning. After 5,500 epochs of frozen-encoder training, the encoder was unfrozen (removing the `--freeze_enc` flag) and training resumed with end-to-end gradient propagation. This allowed the encoder to adapt its feature extraction to the specific demands of view synthesis, potentially learning features more directly aligned with volumetric reconstruction rather than the self-supervised denoising task used during pre-training. The learning rate was maintained at 1×10^{-4} to prevent drastic feature reorganization that could destabilize previously learned geometry.

Implementation details. The updated configuration and training script modifications were applied as follows:

```
renderer {
    n_coarse = 128 # Increased from 64
    n_fine = 128 # Increased from 32
    n_fine_depth = 16
}
loss {
    rgb { use_l1 = True } # Changed from
        False
    rgb_fine { use_l1 = True } # Changed
        from False
}
```

Training was resumed from checkpoint `pixel_nerf_latest` at epoch 5,560 and continued for an additional 2,000 epochs. The visualization interval was reduced from 100 to 50 batches to enable more frequent monitoring of detail improvements during this critical refinement phase.

4.3.2. Tracking Training

Throughout both training phases, model convergence was monitored through multiple metrics. The coarse network loss (rc) and fine network loss (rf) both showed consistent downward trends, with fine network loss typically 5-15% lower than coarse loss, indicating successful hierarchical refinement. The edge-preserving gradient loss (edge) stabilized around 0.002-0.004 after initial introduction, suggesting the model learned to maintain consistent intensity gradients without over-smoothing.

13 shows what the visuals looked like in the middle of phase 2 training.

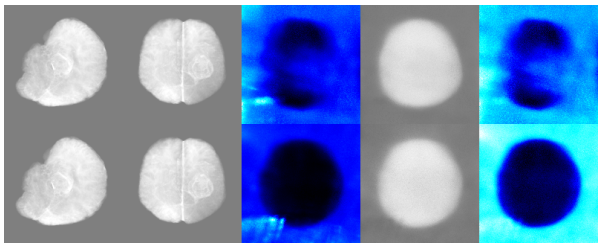


Figure 13. Phase 2: Iteration 1700 in training. Top row shows coarse network outputs. Bottom row shows fine network outputs. Columsn 1-5 (in order): Source view 1 (first input image to the model), source view 2 (second input image to the model), depth map (predicted depth/distance from camera (blue colormap), grayscale prediction (model's novel view synthesis output), alpha map (opacity/density prediction (blue colormap)

5. Results & Analysis

5.1. Quantitative Evaluation

I evaluated my adapted pixelNeRF model on a held-out test set comprising three patients, one from each dataset: UCSF, BraTS, and whole-brain. The model was trained for approximately 48 hours using 1-3 randomly sampled source views per training iteration. For evaluation, I assessed single-view novel view synthesis performance, where the model receives a single input MRI slice and reconstructs remaining viewpoints.

Table 2 presents quantitative results comparing my MRI-adapted model against the original pixelNeRF architecture pretrained on natural images (ShapeNet sn64_unseen). I chose to compare my model with the unseen version of the original PixelNeRF because the other models trained by the authors are object-specific (e.g., some are focused on cars, others on chairs). This one was the model the authors trained for unseen objects. Since the model has never seen MRI slices, I thought it would best work with my project. My domain-specific approach achieved an average Peak Signal-to-Noise Ratio (PSNR) of 16.05 dB and Structural Similarity Index (SSIM) of 0.673 across all test cases, representing a substantial improvement of 7.6 dB PSNR and 0.06 SSIM over the pretrained natural image model applied directly to medical data (8.45 dB PSNR, 0.611 SSIM). This performance gap demonstrates the importance of domain-specific adaptation for medical imaging applications.

Performance varied across datasets for my adapted model, with the whole-brain structural scan (sub_02) achieving PSNR of 17.17 dB, followed by the BraTS tumor case (15.62 dB) and UCSF clinical scan (15.36 dB). In contrast, the pretrained natural image model exhibited unstable failure across all test cases, with PSNR values ranging from 7.54 to 8.45 dB—metrics indicating reconstructions barely distinguishable from noise. Notably, the pretrained model's performance degraded most on the whole-brain case (7.54 dB), highlighting the fundamental incompatibility between natural image features currently used in state-of-the-art NeRF networks and medical imaging characteristics.

5.2. Analysis of Domain Adaptation Benefits

The substantial performance improvement of my MRI-adapted approach can be attributed to three key architectural modifications:

Medical Image Encoder Pre-training: my self-supervised encoder pre-training on 2,000 MRI slices using a denoising objective enabled the model to learn medical imaging-specific features—tissue contrast patterns, intensity distributions, and anatomical priors—absent from ImageNet-trained features. The pretrained natural image model, despite using the same ResNet-34 backbone, struggled to extract meaningful representations from grayscale MRI intensities, as evidenced by SSIM values barely exceeding 0.6 (indicating less than 60% structural preservation). In contrast, my medical encoder achieved SSIM values consistently above 0.65, demonstrating superior structural understanding.

Training on Medical Data Distribution: Most importantly, my model was trained on the specific intensity distributions, noise characteristics, and anatomical variability present in brain MRI. The pretrained natural image model, trained on ShapeNet objects with sharp edges, uniform textures, and distinct color boundaries, encountered a severe distribution shift when applied to medical images featuring smooth tissue gradients, subtle intensity variations, and grayscale-only information. This manifests in the catastrophically low PSNR values (¡ 9 dB), where the model essentially fails to reconstruct coherent anatomical structures.

5.3. Impact of Computational Constraints

The reported metrics are substantially lower than validation performance observed during training (22-26 dB PSNR). This discrepancy is primarily attributable to GPU memory constraints during evaluation. My training configuration employed hierarchical sampling with 128 coarse samples and 128 fine samples per ray, totaling 256 evaluation points along each viewing ray. However, the model's memory footprint—approximately 13.5 GB for the ResNet-34 encoder and dual MLP networks with 512 hidden dimensions—left insufficient memory on Tesla T4 GPUs (15 GB VRAM) for inference at full sampling resolution.

To enable evaluation within available hardware constraints, I reduced sampling density to 32 coarse and 32 fine samples per ray, representing a 4× reduction in spatial sampling. This undersampling directly degrades reconstruction quality through two mechanisms: (1) coarser geometric representation, causing the model to miss fine anatomical details

Table 2. Comparison of PixelNeRF evaluation results on synthetic MRI views (original `sn64_unseen` pretrained vs. MRI-adapted model). Metrics are computed per held-out subject using PSNR (dB) and SSIM; Δ denotes (MRI model $-$ pretrained baseline).

Patient ID	$\text{PSNR}_{\text{sn64\_unseen}}$	$\text{SSIM}_{\text{sn64\_unseen}}$	PSNR_{mri}	SSIM_{mri}	ΔPSNR	ΔSSIM
100108B (UCSF)	8.454	0.6240	15.363	0.6788	+6.908	+0.0548
Training_010 (BraTS20)	8.346	0.6073	15.620	0.6543	+7.274	+0.0470
sub_02__sub_02_scanner_2__sub_02_scanner_2_acq_2 (wholebrain)	7.538	0.6028	17.168	0.6857	+9.630	+0.0829
Average	8.113	0.6114	16.050	0.6729	+7.938	+0.0616

between sample points, and (2) reduced surface localization precision, leading to blurred tissue boundaries. Previous work has shown that reducing sampling density (such as by 4 times) typically can result in PSNR degradation and SSIM reduction ((Ahn et al., 2020)), which is consistent with the approximately 6-10 dB gap I had observed between training validation metrics and test evaluation.

Additional factors contributing to the performance gap include the use of single source views during evaluation versus 1-3 views during training, and the inherent difficulty of the test set cases. The heterogeneous nature of my training data—combining contrast-enhanced tumor imaging (BraTS), clinical T1-weighted scans (UCSF), and whole-brain structural MRI—also presents generalization challenges, as each modality exhibits distinct intensity distributions and anatomical coverage. I evaluate on this part of training later in the next section.

5.4. Cross-Dataset Generalization

Despite these constraints, the model seems to demonstrate meaningful cross-dataset generalization. The whole-brain structural scan achieved reconstruction quality (17.17 dB PSNR), likely due to the consistent acquisition protocol and absence of pathological variability in this dataset. The BraTS tumor case (15.62 dB) and UCSF clinical scan (15.36 dB) showed slightly lower performance, reflecting the additional complexity of contrast enhancement and variable clinical acquisition parameters. Notably, SSIM values remained relatively consistent across datasets (0.654-0.686), suggesting that while pixel-level intensity errors are present, the model preserves overall structural organization and anatomical topology across different imaging protocols.

The moderate SSIM values (approximately 0.67) indicate that the model was able to capture coarse brain morphology—ventricle shapes, major tissue boundaries, and overall brain structure—but struggles with fine cortical details such as sulci, gyri patterns, and subtle intensity variations that characterize high-resolution medical imaging. This behavior is expected given both the architectural simplification (single-channel intensity prediction rather than tissue-specific parameter estimation) and the computational constraints imposed during evaluation.

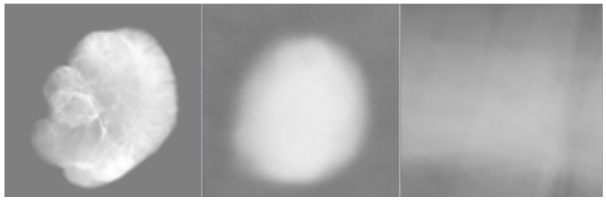


Figure 14. **Comparison of outputs across models.** Left-most picture is the ground-truth slice. Middle picture is slice rendered by the MRI-based network. Right-most picture is the slice generated by original PixelNeRF trained on natural images.

5.5. Qualitative Assessment

Visual inspection of rendered novel views (see 14) reveals that the model produces anatomically plausible brain reconstructions with correct overall geometry and spatial relationships. The model seems to successfully interpolates between source and target views without introducing severe artifacts or hallucinations. However, fine anatomical structures lack sharpness, and tissue boundaries exhibit smoothing effects characteristic of undersampled volume rendering. These observations align with the quantitative metrics and confirm that the primary limitation lies in spatial resolution rather than fundamental representational capacity.

The pretrained natural image model generates severely degraded outputs with barely recognizable anatomical features, frequent hallucinated structures, and inconsistent geometry across viewpoints. These observations align with the quantitative metrics and confirm that meaningful medical image novel view synthesis requires explicit domain adaptation rather than relying on transfer from natural image distributions.

6. Discussions

6.1. Implications and Applications

This work shows that neural radiance fields, originally developed for natural scenes, can be adapted to medical imaging through domain-specific encoder pre-training and modest architectural changes. Training on a heterogeneous dataset (BraTS, UCSF clinical scans, and whole-brain structural MRI) suggests that the encoder learns MRI-specific features

(tissue contrast, intensity statistics, and anatomical priors) that generalize across protocols and sites, which is essential for any realistic clinical deployment.

From a neural rendering perspective, this project highlights several useful design choices for non-RGB data. First, reducing the MLP to a single intensity channel simplifies the NeRF head without noticeably degrading reconstruction quality. Second, pre-training the encoder on medical images rather than RGB natural images appears to give a better starting point for MRI view synthesis. Third, the two-stage training strategy (frozen encoder, then end-to-end fine-tuning) offers a practical way to adapt pre-trained features while limiting catastrophic forgetting.

Although this implementation focuses on 3D reconstruction from synthetic views, the underlying idea has broader clinical implications. A model that can synthesize anatomically consistent novel views from sparse inputs could, in principle, reduce scan time or dose (for CT/PET), improve 3D visualization for surgical planning, and provide richer 3D information in settings where full high-resolution 3D imaging is unavailable. However, realizing these benefits in practice will require rigorous clinical validation with radiologists, including checks for preservation of pathology, absence of hallucinated structures, and robustness across different anatomies and acquisition conditions. At this stage, the system should be viewed as a proof-of-concept rather than a clinically ready tool.

6.2. Limitations

Several limitations constrain the current system.

First, the model treats MRI as a single-channel intensity field and does not explicitly model MRI physics. Real MRI signals arise from tissue-specific parameters (e.g., T1, T2, proton density) and sequence design. Here, intensity is simply learned as a generic value, without physical constraints or explicit tissue priors. A more principled approach would predict tissue property maps and then simulate image formation for different sequences, which could improve both interpretability and generalization across protocols.

Second, the model is undertrained relative to its likely capacity. The total of roughly 7,500 epochs was chosen due to time and compute limits rather than convergence; loss curves were still slowly decreasing when training stopped. The second, "detail-refinement" phase in particular received comparatively few epochs, so the unfrozen encoder may not have fully adapted to the view-synthesis objective.

Third, dataset heterogeneity created a challenging learning problem. BraTS includes contrast-enhanced tumor tissue, UCSF scans vary in field strength and acquisition, and whole-brain scans retain skull and surrounding structures that are removed in skull-stripped datasets. As a result, the network must reconcile conflicting intensity-to-tissue mappings (for example, bright intensities representing white matter in some datasets and bone in others). This likely leads to compromise solutions that work "reasonably" across datasets but are not optimal for any single one. A natural follow-up experiment is to compare single-dataset training with mixed-dataset training to quantify this trade-off.

Fourth, the reconstructions lack fine anatomical detail. The model captures overall brain shape, ventricles, and major tissue boundaries, but sulci, gyri, small nuclei, and subtle lesions appear blurred or under-resolved. Contributing factors likely include the relatively low image resolution (160×160), the number of samples along each ray, and the use of a uniform L1 loss that does not explicitly emphasize boundaries or clinically important regions.

Moreover, a further limitation concerns the quality and fidelity of the synthetic multi-view slice generation itself. Although the synthetic 360° views were designed to provide PixelNeRF with multi-view supervision, the reslicing and rendering process is not lossless. In particular, the original orthogonal planes (axial, coronal, sagittal) often contain the most clinically faithful anatomical detail because they align with how the volume was acquired and reconstructed. By contrast, off-axis synthetic views typically require interpolation across voxels, which can blur high-frequency structure, soften tissue boundaries, and attenuate small features (e.g., thin cortical folds or subtle lesion margins). Any residual misalignment introduced during reorientation/resampling, or slight inconsistencies in intensity normalization across datasets, can further compound these effects. As a result, the synthetic images may underrepresent fine-grained details present in the true slices, and the model may partially learn the "statistics" of the synthetic rendering process rather than the underlying anatomy. This can directly affect training by weakening the supervisory signal, encouraging over-smoothed reconstructions, and reducing generalization when evaluating against the original orthogonal slices or clinically realistic views.

Finally, the planned integration of orthogonal ground-truth slices (axial, sagittal, coronal) as additional supervision was not completed due to DataLoader collation issues with nested dictionaries. These orthogonal slices would have provided valuable complementary information about internal structures. Not being able to use them is a missed opportunity and likely contributes to the remaining lack of internal anatomical detail.

6.3. Future Directions

6.3.1. Short-Term Improvements

Several near-term changes could improve performance without fundamentally changing the architecture.

First, increasing resolution (for example from 160×160 to 256×256 or 512×512) would allow the model to represent higher-frequency anatomical details, assuming adequate GPU memory. A progressive strategy (coarse resolution for geometry, then fine resolution for detail) could balance quality and compute.

Second, the loss function could be made more anatomy-aware. Instead of a uniform L1 loss,

$$\mathcal{L}_{\text{weighted}} = \sum_i w_i \left| I_{\text{pred}}(i) - I_{\text{gt}}(i) \right|,$$

where the weights w_i up-weight edges or clinically important regions (for example via edge detectors or lesion masks), may better preserve diagnostically relevant detail. A multi-scale perceptual loss based on the medical encoder features is another promising option.

Third, implementing a custom `collate_fn` for the DataLoader to correctly batch nested structures would make it possible to incorporate orthogonal slices as conditioning signals or auxiliary targets. This would give the model richer volumetric context than synthetic 360-degree views alone.

Finally, extending training with an appropriate learning-rate schedule (for example, exponential decay or cosine annealing with restarts) would clarify whether current limitations are due primarily to undertraining or to architectural bottlenecks.

6.3.2. Alternative Neural Rendering Architectures

Beyond pixelNeRF, other neural rendering frameworks might address some of these limitations more directly. Medical-specific NeRF variants such as MedNeRF and related extensions incorporate stronger anatomical and physics-inspired priors, and could be adapted to MRI. A systematic comparison of pixelNeRF with these architectures would help identify which inductive biases (for example, segmentation-guided sampling, attenuation modeling, or uncertainty-aware losses) are most important for neuroimaging.

Another promising direction is 3D Gaussian splatting, which represents scenes as collections of explicit 3D Gaussian primitives and enables much faster rendering than traditional NeRF-style ray marching. For interactive clinical applications (e.g., surgical planning), such speed could be crucial. Adapting Gaussian splatting to MRI would involve representing local tissue regions as Gaussians and optimizing their spatial, shape, and intensity parameters to match observed slices, potentially yielding a more direct link between volumetric structure and rendered appearance.

6.3.3. Long-Term Vision and Ethical Considerations

In the longer term, neural rendering could change how neuroimaging data are stored and analyzed. Neural representations (e.g., NeRF or Gaussian splatting models) can compress large 3D volumes into relatively small parameter sets while still allowing continuous querying at arbitrary viewpoints and resolutions. Coupled with generative extensions, this would support population-level analyses in a continuous latent space of brain anatomy and pathology, rather than in raw voxel coordinates.

A particularly compelling direction is physics-informed neural rendering, in which the model learns tissue property maps (e.g., T_1, T_2, proton density) and a differentiable forward model of MRI acquisition. This could enable flexible synthesis of multiple contrasts, correction of acquisition artifacts, and more quantitative tissue characterization.

At the same time, moving toward clinical use raises important ethical and regulatory questions. Synthetic views and compressed neural representations must be validated so that radiologists can trust that critical findings are neither hallucinated nor suppressed. Regulatory bodies will need standards for evaluating such systems, and consent procedures must make clear how derived neural representations of patient data are stored and used. Addressing these issues will be as important as the technical progress itself if neural rendering is to have real clinical impact.

7. Conclusion

This project explored whether a neural radiance field model originally designed for natural images, pixelNeRF, can be adapted to reconstruct 3D brain MRI volumes from 2D inputs. To do this, I built an end-to-end pipeline that starts from heterogeneous clinical and research MRI data (BraTS, UCSF post-contrast T1, and whole-brain structural scans), reorients them into a common RAS space, resamples to a shared voxel grid, performs foreground cropping and intensity normalization, and finally generates synthetic 360-degree projections for NeRF-style training. On the modeling side, I modified pixelNeRF to predict single-channel MRI intensities instead of RGB, and replaced the usual ImageNet encoder with a medically pre-trained encoder, trained in two phases (frozen encoder, then end-to-end fine-tuning).

The resulting model demonstrates that neural radiance fields can, in fact, learn an implicit representation of brain anatomy across multiple acquisition protocols. Qualitative results show that the system recovers gross structures such as overall brain shape, and it generalizes across contrast-enhanced tumor scans, more variable clinical UCSF scans, and whole-brain volumes.

At the same time, the project surfaced several important limitations. Training time and resolution were constrained by available compute, and the heterogeneous combination of datasets (with and without skull, different field strengths, different contrast behaviors) introduced conflicting intensity-to-tissue mappings that the network had to compromise between. These factors help explain the persistent lack of fine anatomical detail in the reconstructions, especially for small structures and sharp cortical features.

Despite these limitations, the work points toward several promising next steps. In the short term, higher-resolution training, anatomy-aware losses, longer fine-tuning, and successful integration of orthogonal clinical slices could improve detail and stability. In the medium term, it would be natural to compare this adapted pixelNeRF to medical NeRF variants and faster neural rendering approaches such as Gaussian splatting to see which inductive biases are most helpful for MRI. Longer term, the project supports the broader idea that neural graphics methods can serve as compact, flexible representations of neuroimaging data, potentially enabling compressed storage, multi-contrast synthesis, and richer 3D visualization from sparse measurements. Realizing that vision will require not only technical advances (especially physics-informed models) but also careful clinical validation and attention to ethical and regulatory questions around synthetic medical images. This project is a small step in that direction, showing both what is currently possible and what still remains to be solved.

8. Code Availability

To access my code, follow the linked GDrive.

9. Data Availability

To access UCSF data, follow the linked site.

To access BraTS data, follow the linked site.

To access wholebrain data, follow the linked site.

References

Ahmad, Z., Rahim, S., Zubair, M., and Abdul-Ghafar, J. Artificial intelligence (AI) in medicine, current applications and future role with special emphasis on its potential and promise in pathology: present and future impact, obstacles including costs and acceptance among pathologists, practical and philosophical considerations. A comprehensive review. *Diagnostic Pathology*, 16(1):24, March 2021. ISSN 1746-1596. doi: 10.1186/s13000-021-01085-4. URL https://doi.org/10.1186/s13000-021-01085-4.

Ahn, S., Menini, A., McKinnon, G., Gray, E. M., Trzasko, J. D., Huston, J., Bernstein, M. A., Costello, J. E., Foo, T. K. F., and Hardy, C. J. Contrast-weighted SSIM loss function for deep learning-based undersampled MRI reconstruction, 2020. URL https://archive.ismrm.org/2020/1295.html. publisher: ISMRM.

Awojoyogbe, B. O. and Dada, M. O. Neural Radiance Fields (NeRFs) Technique to Render 3D Reconstruction of Magnetic Resonance Images. In Awojoyogbe, B. O. and Dada, M. O. (eds.), *Digital Molecular Magnetic Resonance Imaging*, pp. 247–258. Springer Nature, Singapore, 2024. ISBN 9789819763702. doi: 10.1007/978-981-97-6370-2_10. URL https://doi.org/10.1007/978-981-97-6370-2_10.

Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J. S., Freymann, J. B., Farahani, K., and Davatzikos, C. Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. *Scientific Data*, 4(1):170117, September 2017. ISSN 2052-4463. doi: 10.1038/sdata.2017.117. URL https://www.nature.com/articles/sdata2017117.

Bakas, S., Reyes, M., Jakab, A., Bauer, S., Rempfler, M., Crimi, A., Shinohara, R. T., Berger, C., Ha, S. M., Rozycki, M., Prastawa, M., Alberts, E., Lipkova, J., Freymann, J., Kirby, J., Bilello, M., Fathallah-Shaykh, H., Wiest, R., Kirschke, J., Wiestler, B., Colen, R., Kotrotsou, A., Lamontagne, P., Marcus, D., Milchenko, M., Nazeri, A., Weber, M.-A., Mahajan, A., Baid, U., Gerstner, E., Kwon, D., Acharya, G., Agarwal, M., Alam, M., Albiol, A., Albiol, A., Albiol, F. J., Alex, V., Allinson, N., Amorim, P. H. A., Amrutkar, A., Anand, G., Andermatt, S., Arbel, T., Arbelaez, P., Avery, A., Azmat, M., B, P., Bai, W., Banerjee, S., Barth, B., Batchelder, T., Batmanghelich, K., Battistella, E., Beers, A., Belyaev, M., Bendszus, M., Benson, E., Bernal, J., Bharath, H. N., Biros, G., Bisdas, S., Brown, J., Cabezas, M., Cao, S., Cardoso, J. M., Carver, E. N., Casamitjana, A., Castillo, L. S., Catà, M., Cattin, P., Cerigues, A., Chagas, V. S., Chandra, S., Chang, Y.-J., Chang, S., Chang, K., Chazalon, J., Chen, S., Chen, W., Chen, J. W., Chen, Z., Cheng, K., Choudhury, A. R., Chylla, R., Clérigues, A., Colleman, S., Colmeiro, R. G. R., Combalia, M., Costa, A., Cui, X., Dai, Z., Dai, L., Daza, L. A., Deutsch, E., Ding, C., Dong, C., Dong, S., Dudzik, W., Eaton-Rosen, Z., Egan, G., Escudero, G., Estienne, T., Everson, R., Fabrizio, J., Fan, Y., Fang, L., Feng, X., Ferrante, E., Fidon, L., Fischer, M., French, A. P., Fridman, N., Fu, H., Fuentes, D., Gao, Y., Gates, E., Gering, D., Gholami, A., Gierke, W., Glocker, B., Gong, M., González-Villá, S., Grosges, T., Guan, Y., Guo, S., Gupta, S., Han, W.-S., Han, I. S., Harmuth, K., He, H., Hernández-Sabaté, A., Herrmann, E., Himthani, N., Hsu, W., Hsu, C., Hu, X., Hu, X., Hu,

Y., Hu, Y., Hua, R., Huang, T.-Y., Huang, W., Huffel, S. V., Huo, Q., HV, V., Iftekharuddin, K. M., Isensee, F., Islam, M., Jackson, A. S., Jambawalikar, S. R., Jesson, A., Jian, W., Jin, P., Jose, V. J. M., Jungo, A., Kainz, B., Kamnitsas, K., Kao, P.-Y., Karnawat, A., Kellermeier, T., Kermi, A., Keutzer, K., Khadir, M. T., Khened, M., Kickingereder, P., Kim, G., King, N., Knapp, H., Knecht, U., Kohli, L., Kong, D., Kong, X., Koppers, S., Kori, A., Krishnamurthi, G., Krivov, E., Kumar, P., Kushibar, K., Lachinov, D., Lambrou, T., Lee, J., Lee, C., Lee, Y., Lee, M., Lefkovits, S., Lefkovits, L., Levitt, J., Li, T., Li, H., Li, W., Li, H., Li, X., Li, Y., Li, H., Li, Z., Li, X., Li, Z., Li, X., Li, W., Lin, Z.-S., Lin, F., Lio, P., Liu, C., Liu, B., Liu, X., Liu, M., Liu, J., Liu, L., Llado, X., Lopez, M. M., Lorenzo, P. R., Lu, Z., Luo, L., Luo, Z., Ma, J., Ma, K., Mackie, T., Madabushi, A., Mahmoudi, I., Maier-Hein, K. H., Maji, P., Mammen, C. P., Mang, A., Manjunath, B. S., Marcinkiewicz, M., McDonagh, S., McKenna, S., McKinley, R., Mehl, M., Mehta, S., Mehta, R., Meier, R., Meinel, C., Merhof, D., Meyer, C., Miller, R., Mitra, S., Moiyadi, A., Molina-Garcia, D., Monteiro, M. A. B., Mrukwa, G., Myronenko, A., Nalepa, J., Ngo, T., Nie, D., Ning, H., Niu, C., Nuechterlein, N. K., Oermann, E., Oliveira, A., Oliveira, D. D. C., Oliver, A., Osman, A. F. I., Ou, Y.-N., Ourselin, S., Paragios, N., Park, M. S., Paschke, B., Pauloski, J. G., Pawar, K., Pawlowski, N., Pei, L., Peng, S., Pereira, S. M., Perez-Beteta, J., Perez-Garcia, V. M., Pezold, S., Pham, B., Phophalia, A., Piella, G., Pillai, G. N., Piraud, M., Pisov, M., Popli, A., Pound, M. P., Pourreza, R., Prasanna, P., Prkovska, V., Pridmore, T. P., Puch, S., Puybareau, , Qian, B., Qiao, X., Rajchl, M., Rane, S., Rebsamen, M., Ren, H., Ren, X., Revanuru, K., Rezaei, M., Rippel, O., Rivera, L. C., Robert, C., Rosen, B., Rueckert, D., Safwan, M., Salem, M., Salvi, J., Sanchez, I., Sánchez, I., Santos, H. M., Sartor, E., Schellingerhout, D., Scheufele, K., Scott, M. R., Scussel, A. A., Sedlar, S., Serrano-Rubio, J. P., Shah, N. J., Shah, N., Shaikh, M., Shankar, B. U., Shboul, Z., Shen, H., Shen, D., Shen, L., Shen, H., Shenoy, V., Shi, F., Shin, H. E., Shu, H., Sima, D., Sinclair, M., Smedby, O., Snyder, J. M., Soltaninejad, M., Song, G., Soni, M., Stawiaski, J., Subramanian, S., Sun, L., Sun, R., Sun, J., Sun, K., Sun, Y., Sun, G., Sun, S., Suter, Y. R., Szilagyi, L., Talbar, S., Tao, D., Tao, D., Teng, Z., Thakur, S., Thakur, M. H., Tharakan, S., Tiwari, P., Tochon, G., Tran, T., Tsai, Y. M., Tseng, K.-L., Tuan, T. A., Turlapov, V., Tustison, N., Vakalopoulou, M., Valverde, S., Vanguri, R., Vasiliev, E., Ventura, J., Vera, L., Vercauteren, T., Verrastro, C. A., Vidyaratne, L., Vilaplana, V., Vivekanandan, A., Wang, G., Wang, Q., Wang, C. J., Wang, W., Wang, D., Wang, R., Wang, Y., Wang, C., Wang, G., Wen, N., Wen, X., Weninger, L., Wick, W., Wu, S., Wu, Q., Wu, Y., Xia, Y., Xu, Y., Xu, X., Xu, P., Yang, T.-L., Yang, X., Yang, H.-Y., Yang, J., Yang, H., Yang, G., Yao, H., Ye, X., Yin, C., Young-Moxon, B., Yu, J., Yue, X., Zhang, S., Zhang, A., Zhang, K., Zhang, X., Zhang, L., Zhang, X., Zhang, Y., Zhang, L., Zhang, J., Zhang, X., Zhang, T., Zhao, S., Zhao, Y., Zhao, X., Zhao, L., Zheng, Y., Zhong, L., Zhou, C., Zhou, X., Zhou, F., Zhu, H., Zhu, J., Zhuge, Y., Zong, W., Kalpathy-Cramer, J., Farahani, K., Davatzikos, C., Leemput, K. v., and Menze, B. Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge, April 2019. URL http://arxiv.org/abs/1811.02629. arXiv:1811.02629.

Bengs, M., Behrendt, F., Krüger, J., Opfer, R., and Schlaefer, A. Three-dimensional deep learning with spatial erasing for unsupervised anomaly segmentation in brain MRI. *International Journal of Computer Assisted Radiology and Surgery*, 16(9):1413–1423, September 2021. ISSN 1861-6429. doi: 10.1007/s11548-021-02451-9. URL https://doi.org/10.1007/s11548-021-02451-9.

Broadhouse, K. M. The Physics of MRI and How We Use It to Reveal the Mysteries of the Mind, March 2019. URL https://kids.frontiersin.org/articles/10.3389/frym.2019.00023.

Chan, E. R., Monteiro, M., Kellnhofer, P., Wu, J., and Wetzstein, G. pi-GAN: Periodic Implicit Generative Adversarial Networks for 3D-Aware Image Synthesis, April 2021. URL http://arxiv.org/abs/2012.00926. arXiv:2012.00926.

Chan, E. R., Lin, C. Z., Chan, M. A., Nagano, K., Pan, B., Mello, S. D., Gallo, O., Guibas, L., Tremblay, J., Khamis, S., Karras, T., and Wetzstein, G. Efficient Geometry-aware 3D Generative Adversarial Networks, April 2022. URL http://arxiv.org/abs/2112.07945. arXiv:2112.07945.

Chen, A., Xu, Z., Zhao, F., Zhang, X., Xiang, F., Yu, J., and Su, H. MVSNeRF: Fast Generalizable Radiance Field Reconstruction from Multi-View Stereo, August 2021. URL http://arxiv.org/abs/2103.15595. arXiv:2103.15595.

Chen, A., Xu, Z., Geiger, A., Yu, J., and Su, H. TensoRF: Tensorial Radiance Fields, November 2022. URL http://arxiv.org/abs/2203.09517. arXiv:2203.09517.

Corona-Figueroa, A., Frawley, J., Taylor, S. B., Bethapudi, S., Shum, H. P. H., and Willcocks, C. G. MedNeRF: Medical Neural Radiance Fields for Reconstructing 3D-aware CT-Projections from a Single X-ray. In *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 3843–3848, July

2022. doi: 10.1109/EMBC48229.2022.9871757. URL https://ieeexplore.ieee.org/abstract/document/9871757. ISSN: 2694-0604.

Godenschweger, F., Kägebein, U., Stucht, D., Yarach, U., Sciarra, A., Yakupov, R., Lüsebrink, F., Schulze, P., and Speck, O. Motion correction in MRI of the brain. *Physics in Medicine and Biology*, 61(5):R32–R56, March 2016. ISSN 0031-9155, 1361-6560. doi: 10.1088/0031-9155/61/5/R32. URL https://iopscience.iop.org/article/10.1088/0031-9155/61/5/R32.

Hu, J., Fan, Q., Hu, S., Lyu, S., Wu, X., and Wang, X. UMedNeRF: Uncertainty-aware Single View Volumetric Rendering for Medical Neural Radiance Fields, March 2024. URL http://arxiv.org/abs/2311.05836. arXiv:2311.05836.

Iddrisu, K., Malec, S., and Crimi, A. 3D Reconstructions of Brain from MRI Scans Using Neural Radiance Fields. In Rutkowski, L., Scherer, R., Korytkowski, M., Pedrycz, W., Tadeusiewicz, R., and Zurada, J. M. (eds.), *Artificial Intelligence and Soft Computing*, pp. 207–218, Cham, 2023. Springer Nature Switzerland. ISBN 9783031425080. doi: 10.1007/978-3-031-42508-0_19.

Liu, W., Liu, Z., Paull, L., Weller, A., and Schölkopf, B. Structural Causal 3D Reconstruction. In Avidan, S., Brostow, G., Cissé, M., Farinella, G. M., and Hassner, T. (eds.), *Computer Vision – ECCV 2022*, pp. 140–159, Cham, 2022. Springer Nature Switzerland. ISBN 9783031197697. doi: 10.1007/978-3-031-19769-7_9.

Liu, Y., Pang, Y., Sun, X., Hou, Y., and Xu, H. EVOLVE: Learning volume-adaptive phases for fast 3D magnetic resonance scan and image reconstruction. *Neurocomputing*, 558:126810, November 2023. ISSN 0925-2312. doi: 10.1016/j.neucom.2023.126810. URL https://www.sciencedirect.com/science/article/pii/S0925231223009335.

Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., Lanczi, L., Gerstner, E., Weber, M.-A., Arbel, T., Avants, B. B., Ayache, N., Buendia, P., Collins, D. L., Cordier, N., Corso, J. J., Criminisi, A., Das, T., Delingette, H., Demiralp, , Durst, C. R., Dojat, M., Doyle, S., Festa, J., Forbes, F., Geremia, E., Glocker, B., Golland, P., Guo, X., Hamamci, A., Iftekharuddin, K. M., Jena, R., John, N. M., Konukoglu, E., Lashkari, D., Mariz, J. A., Meier, R., Pereira, S., Precup, D., Price, S. J., Raviv, T. R., Reza, S. M. S., Ryan, M., Sarikaya, D., Schwartz, L., Shin, H.-C., Shotton, J., Silva, C. A., Sousa, N., Subbanna, N. K., Szekely, G., Taylor, T. J., Thomas, O. M., Tustison, N. J., Unal, G., Vasseur, F., Wintermark, M., Ye, D. H., Zhao, L., Zhao, B., Zikic, D., Prastawa, M., Reyes, M., and Van Leemput, K. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Transactions on Medical Imaging*, 34(10):1993–2024, October 2015. ISSN 1558-254X. doi: 10.1109/TMI.2014.2377694. URL https://ieeexplore.ieee.org/document/6975210.

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., and Ng, R. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, August 2020. URL http://arxiv.org/abs/2003.08934. arXiv:2003.08934.

Müller, T., Evans, A., Schied, C., and Keller, A. Instant Neural Graphics Primitives with a Multiresolution Hash Encoding, May 2022. URL http://arxiv.org/abs/2201.05989. arXiv:2201.05989.

Rabby, A. S. A. and Zhang, C. BeyondPixels: A Comprehensive Review of the Evolution of Neural Radiance Fields, March 2024. URL http://arxiv.org/abs/2306.03000. arXiv:2306.03000.

Rudie, J. D., Saluja, R., Weiss, D. A., Nedelec, P., Calabrese, E., Colby, J. B., Laguna, B., Mongan, J., Braunstein, S., Hess, C. P., Rauschecker, A. M., Sugrue, L. P., and Villanueva-Meyer, J. E. The University of California San Francisco Brain Metastases Stereotactic Radiosurgery (UCSF-BMSR) MRI Dataset, May 2024. URL http://arxiv.org/abs/2304.07248. arXiv:2304.07248.

Schwarz, K., Liao, Y., Niemeyer, M., and Geiger, A. GRAF: Generative Radiance Fields for 3D-Aware Image Synthesis, March 2021. URL http://arxiv.org/abs/2007.02442. arXiv:2007.02442.

Sun, M., Zhu, Y., Li, H., Ye, J., and Li, N. ACnerf: enhancement of neural radiance field by alignment and correction of pose to reconstruct new views from a single x-ray. *Physics in Medicine and Biology*, 69(4), February 2024. ISSN 1361-6560. doi: 10.1088/1361-6560/ad1d6c.

Tewari, A., Thies, J., Mildenhall, B., Srinivasan, P., Tretschk, E., Yifan, W., Lassner, C., Sitzmann, V., Martin-Brualla, R., Lombardi, S., Simon, T., Theobalt, C., Nießner, M., Barron, J. T., Wetzstein, G., Zollhöfer, M., and Golyanik, V. Advances in Neural Rendering. *Computer Graphics Forum*, 41(2):703–735, May 2022. ISSN 0167-7055, 1467-8659. doi: 10.1111/cgf.14507. URL https://onlinelibrary.wiley.com/doi/10.1111/cgf.14507.

Wang, S., Cheng, H., Ying, L., Xiao, T., Ke, Z., Zheng, H., and Liang, D. DeepcomplexMRI: Exploiting deep residual network for fast parallel MR imaging with complex convolution. *Magnetic Resonance Imaging*, 68:136–147, May 2020. ISSN 0730-725X. doi: 10.1016/j.mri.2020.02.

002. URL https://www.sciencedirect.com/science/article/pii/S0730725X19305338.

Wang, X., Chen, Y., Hu, S., Fan, H., Zhu, H., and Li, X. Neural Radiance Fields in Medical Imaging: A Survey, February 2025. URL http://arxiv.org/abs/2402.17797. arXiv:2402.17797.

Westen, S. C., Warnick, J. L., Albanese-O'Neill, A., Schatz, D. A., Haller, M. J., Entessari, M., and Janicke, D. M. Objectively Measured Adherence in Adolescents With Type 1 Diabetes on Multiple Daily Injections and Insulin Pump Therapy. *Journal of Pediatric Psychology*, 44(1):21–31, January 2019. ISSN 0146-8693, 1465-735X. doi: 10.1093/jpepsy/jsy064. URL https://academic.oup.com/jpepsy/article/44/1/21/5090192.

Yang, T., Liu, X., Liu, Y., Sun, X., Wang, Z., and Pang, Y. Efficient 3D magnetic resonance image reconstruction by 2D transformers and attention-based fusion model. *Biomedical Signal Processing and Control*, 110:108071, December 2025. ISSN 1746-8094. doi: 10.1016/j.bspc.2025.108071. URL https://www.sciencedirect.com/science/article/pii/S1746809425005828.

Yu, A., Ye, V., Tancik, M., and Kanazawa, A. pixelNeRF: Neural Radiance Fields From One or Few Images. pp. 4578–4587, 2021. URL https://openaccess.thecvf.com/content/CVPR2021/html/Yu_pixelNeRF_Neural_Radiance_Fields_From_One_or_Few_Images_CVPR_2021_paper.html.

Zhao, W., Hu, Z., Kazerooni, A. F., Körzdörfer, G., Nittka, M., Davatzikos, C., Viswanath, S. E., Wang, X., Badve, C., and Ma, D. Physics-Informed Discretization for Reproducible and Robust Radiomic Feature Extraction Using Quantitative MRI. *Investigative Radiology*, 59(5):359–371, May 2024. ISSN 1536-0210, 0020-9996. doi: 10.1097/RLI.0000000000001026. URL https://journals.lww.com/10.1097/RLI.0000000000001026.

Zhu, F., Guo, S., Song, L., Xu, K., and Hu, J. Deep Review and Analysis of Recent NeRFs. *APSIPA Transactions on Signal and Information Processing*, 12(1), 2023. ISSN 2048-7703. doi: 10.1561/116.00000162. URL http://www.nowpublishers.com/article/Details/SIP-2022-0062.