

# 3D reconstructions of brain from MRI scans using *Neural Radiance Fields*

---

Khadija Iddrisu<sup>1</sup>, Sylwia Malec<sup>2</sup>, and Alessandro Crimi<sup>2</sup>

<sup>1</sup> African Institute for Mathematical Sciences

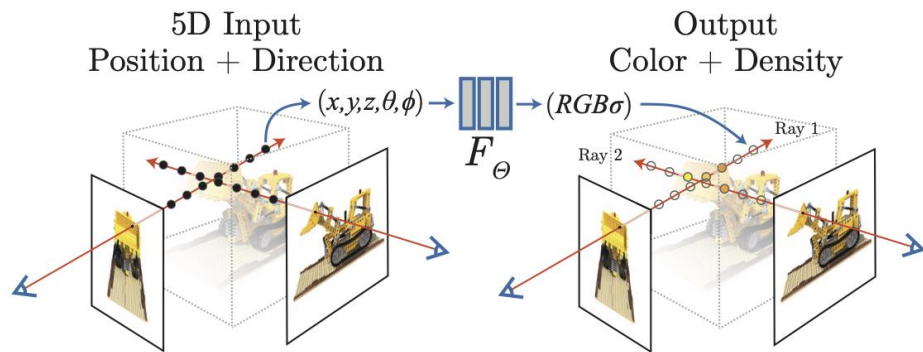
<sup>2</sup> Sano Centre For Computational Medicine

# What is NeRF?

A model for synthesizing 3D scenes from a set of 2D images.  
Encodes a scene as a continuous volumetric function using a neural network.

# How does it work?

Learns to map a 3D coordinate and viewing direction  $(x,y,z,\theta,\phi)$  to **color** and **density**.



- Requires multi-view images for training.
- Produces high-quality 3D reconstructions with fine details and realistic lighting.
- Applications: VR/AR, gaming, digital twins, and content creation.

# Challenges

- Long MRI acquisition sequences are subject to motion-related artifacts due to breathing and cardiac pulsation which can lead to inaccuracies in 3D reconstruction.
- MRI scans present several challenges due to the multiple 2D slices and variable slice thicknesses.
- NeRF faces further challenges due to the high dimensionality of MRI data, low contrast or missing information in the images.

# Related Work

- **GRAF: Generative Radiance Fields for 3D-Aware Image Synthesis;**

- ✓ Generative NeRF trained in an unsupervised manner.
- (-) Designed for synthetic datasets like faces or cars, which are less complex than medical imaging data.
- (-) Cannot handle spatial inconsistencies often present in medical imaging.

- **MedNeRF: Medical Neural Radiance Fields for CT Reconstruction**

- ✓ NeRF tailored to incorporate medical priors for accurate 3D reconstructions.
- (-) Relies on single-projection X-ray data, making the 3D reconstruction problem highly ill-posed.
- (-) Does not natively support volumetric MRI data or the need for accurate interpolation across slices.

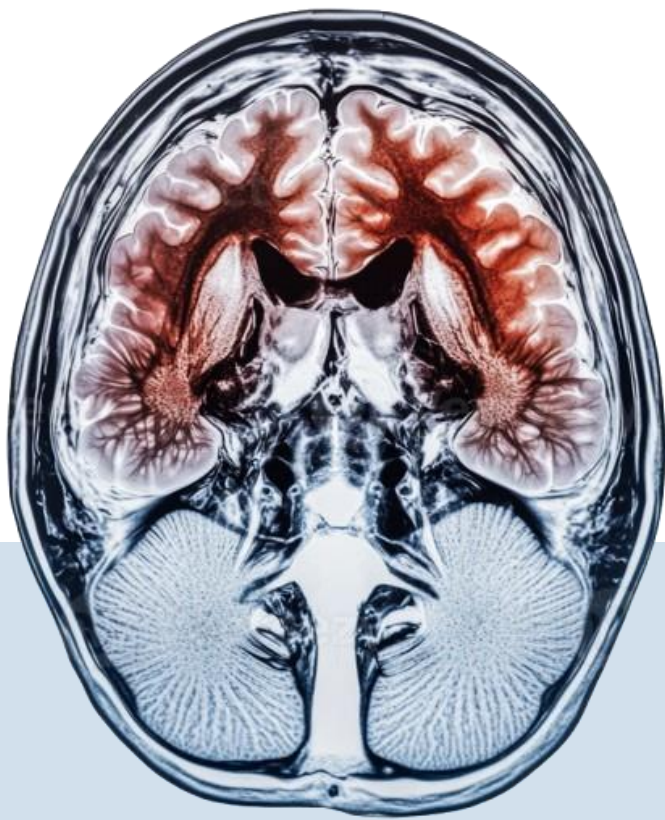


Image from freepik

## Preprocessing

to improve data quality (low contrast and poor resolution)

## Feature extraction

from input slices using a CNN architecture to provide a better input to extended-MedNeRF

## 3D reconstruction

Use of extended-MedNeRF model for 3D reconstruction

## Image registration

Aligning multiple MRI scans of the same patient can improve the quality of the data by correcting any spatial misalignments or inconsistencies, and reducing motion artifacts and distortions in the input MRI images.

## Image rescaling

Resizing the input MRI slices to a higher resolution using bicubic interpolation resulted in a significant improvement in the accuracy of the 3D reconstruction model.

$$I'(x, y) = \sum_{i=-1}^2 \sum_{j=-1}^2 h_{i+2, j+2} \cdot I(x+i, y+j)$$

$h$  is the bicubic coefficient, precomputed based on the distance between the target coordinate and the neighboring pixel locations

## Contrast enhancement

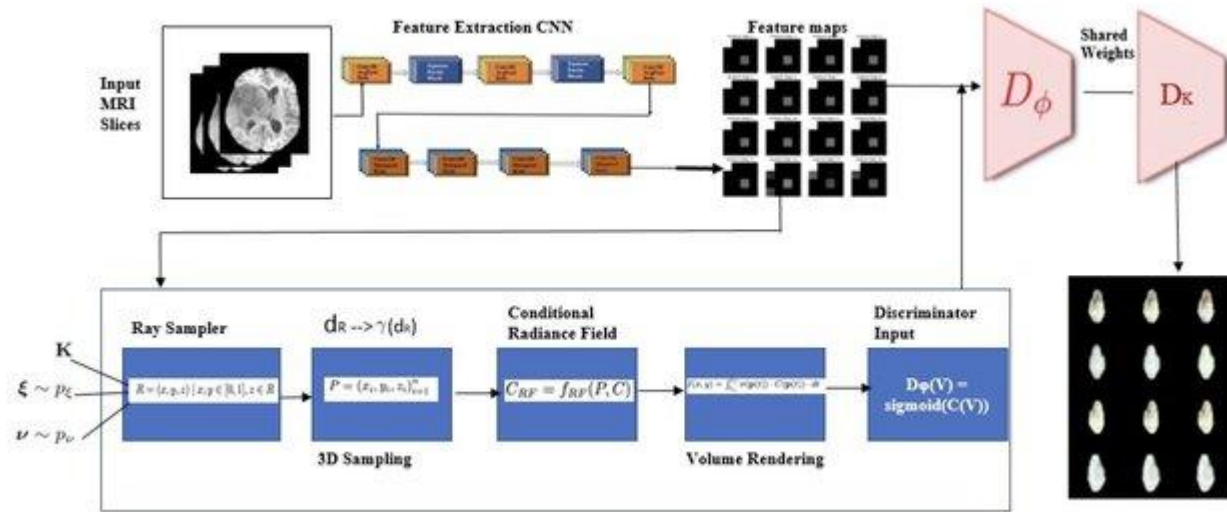
Using histogram equalization  $T(r_k)$  to adjust the pixel intensity distribution of an image and cover the entire dynamic range, improving its visual quality. The process involves calculating the cumulative distribution function  $C(r_k)$  of pixel intensities and applying a transformation to redistribute the intensities more evenly.

$$C(r_k) = \sum_{j=0}^k p_r(r_k).$$

$$T(r_k) = \left\lfloor \left( (L-1)C(r_k) + \frac{1}{2} \right) \right\rfloor$$

$p(r_k)$  denotes the probability mass function of the input image intensities.  $L$  is the number of possible intensity levels ( $L=256$  for 8-bit images)





- 3D CNN to capture patterns across slices, creating a 3D representation for each slice. This helps the model understand spatial relationships between slices.
- Flatten the 3D features of each slice into 1D vectors and combine them into a matrix for all slices.
- Squeeze-and-Excitation block to focus on the most important features.

- **Patches and Ray Sampling:**

$$P(\mathbf{u}, s) = \left\{ (sx + u, sy + v) \mid x, y \in \left\{ -\frac{K}{2}, \dots, \frac{K}{2} - 1 \right\} \right\}$$

$K \times K$  patch centered at the coordinates  $(sx, sy)$  within a 2D MRI slice constitutes the Ray Sampler. The offsets  $u$  and  $v$  vary between  $-K/2$  and  $(K/2)-1$ , representing the pixel locations within the patch. This patch is predicted by the generator  $G$ , capturing local anatomical features from the 2D slice.

- **3D Ray Correspondence:**

For each pixel in the patch, 3D rays are determined using the image coordinates  $(sx+u, sy+v)$ , the camera pose, and the intrinsic parameters.

- ➔ mapping of 2D pixel coordinates to 3D rays in the spatial domain
- ➔ enabling the model to synthesize realistic 3D representations from the MRI slices.

- **Conditional Radiance Field:**

The radiance field conditions the prediction of each pixel's color and volume density based on the 3D rays and the viewing direction.

Using a multi-layer perceptron (MLP), this model predicts an RGB color value and volume density for each point along the 3D ray, based on the latent codes for shape and appearance.



**Dataset: International Brain Tumor Segmentation (BraTS) Challenge 2021 dataset.**

**Contents:** Multi-institutional routine clinically acquired multi-parametric MRI (mpMRI) scans of glioma patients.  
Includes pathologically confirmed diagnoses.  
Contains MGMT promoter methylation status for glioblastoma cases.

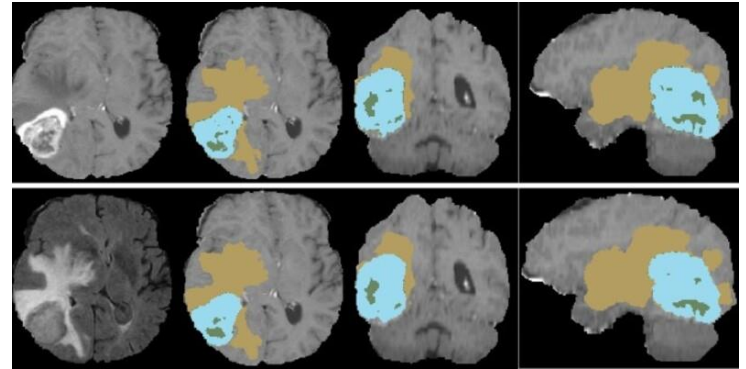
**Training Set:** 30 patients used for training.  
Remaining patients used for testing.

**Slice Selection:** 40 slices selected for each of 3 patients in initial experiments (chosen randomly to avoid bias).

**Training Parameters:** Number of Iterations: 100,000 iterations.  
Batch Size: 8.

**Views:** Only 1/5 of the views are used during training.  
72 views per object, each at 5-degree intervals.

**Rendered Views:** The model is tasked with rendering the remaining unseen views.



BraTS 2021 Task 1 Dataset from Kaggle

# Metrics

- **PSNR (Peak Signal-to-Noise Ratio)**

The amount of noise present in an image or video. It compares the original image to a compressed or distorted version of that image. PSNR is defined as the ratio between the maximum possible value of a signal and the mean squared error (MSE) of the distorted signal. Higher PSNR values indicate higher image quality.

$$\text{PSNR} = 25.01 \pm 1.17$$

- **SSIM (Structural Similarity Index)**

A metric that measures the similarity between two images. It takes into account both the structural information and the luminance and contrast information of the image. The values range from -1 to 1.

$$\text{SSIM} = 0.879 \pm 0.07$$

# Metrics

- **FID**

Measures the similarity between the generated and the real datasets. It measures the distance between the feature distributions of the generated images and the real images.

Lower FID scores indicate higher quality generated images that are more similar to real images.

FID = 160.12

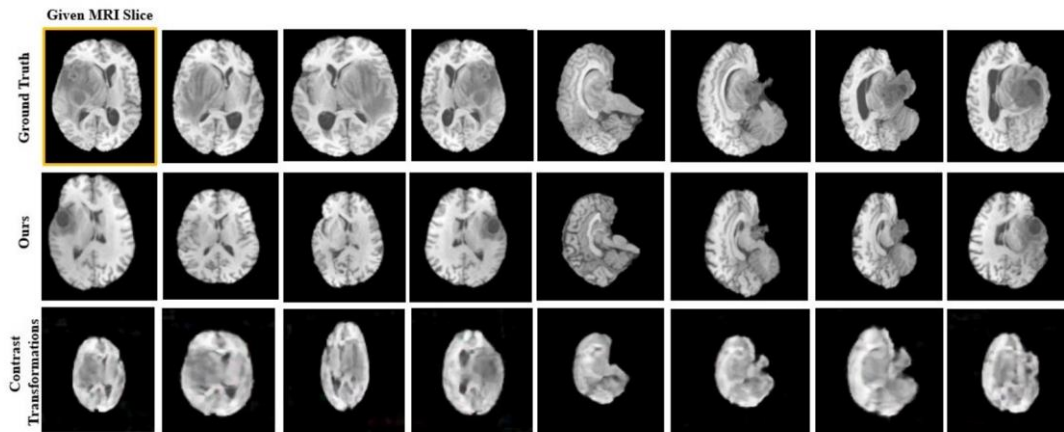
- **KID**

Measures the similarity between two datasets of images with kernel-based methods.

Lower KID scores also indicate higher quality generated images that are more similar to real images.

KID =  $0.16 \pm 0.003$

# Single-View MRI Reconstruction



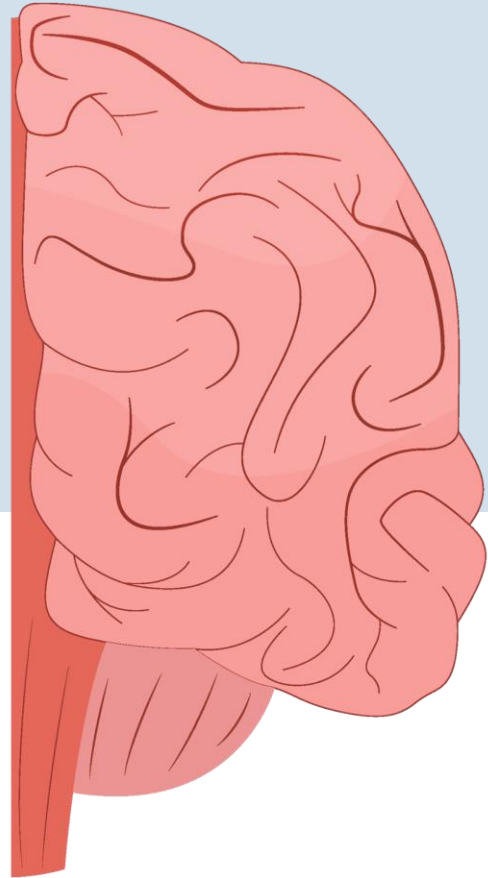
Finetune the pretrained model: Train the generator on a single slice using latent vectors for shape and appearance. Incorporate MSE loss to balance distortion and perception.

**Output:** High-quality 3D reconstructions.  
Contrast-enhanced slices for better tissue differentiation.

➔ Significantly reduces data requirements, enabling reconstructions with minimal input.

# Thanks

**Any questions?**



Please don't