Implicit Neural Representation in Medical Imaging: A Comparative Survey

Amirali Molaei, Amirhossein Aminimehr, Armin Tavakoli, Amirhossein Kazerouni, Bobby Azad, Reza Azad, and Dorit Merhof

https://github.com/mindflow-institue/Awesome-Implicit-Neural-Representations-in-Medical-imaging

1. Introduction

Implicit Neural Representations (INRs), which are models that use neural networks to represent signals as continuous functions, have gained great attention in recent years in medical imaging tasks, especially in reconstruction.

Motivated by their practicality, we cover various aspects of INR models in medical imaging, including their application in tasks such as image reconstruction, segmentation, registration, novel view synthesis, and compression, along with the clinical importance of these models. We further discuss their utilization and advantages in different tasks and the challenges they face, along with potential future directions.

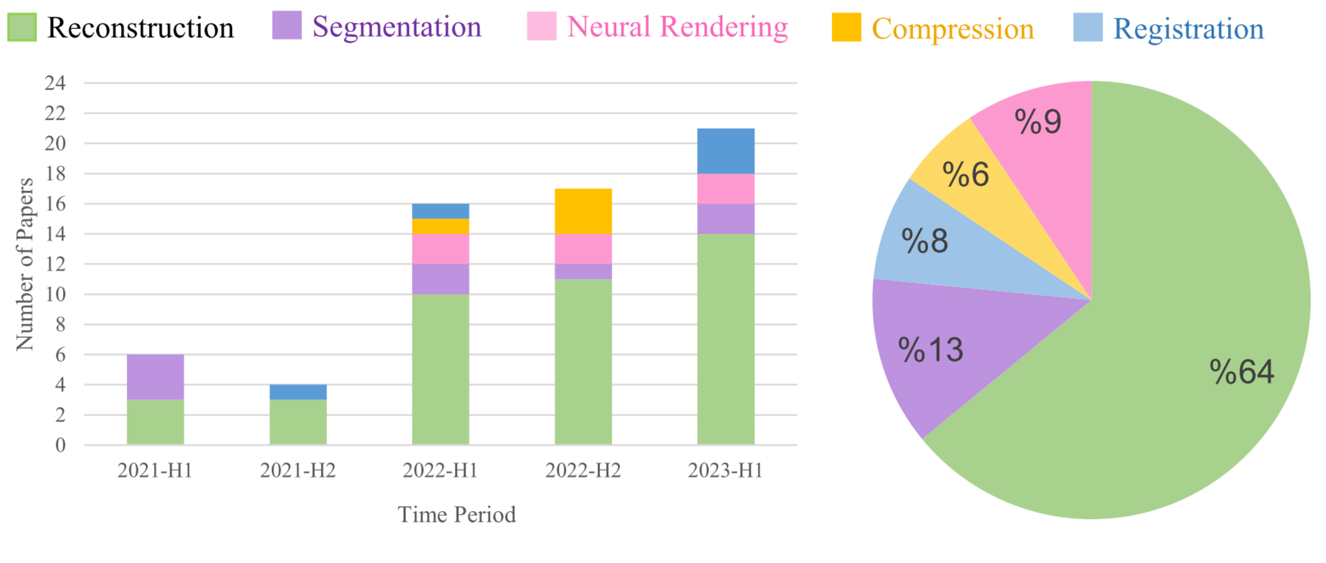

Fig. 1. The number of published papers in each medical task over the recent years. The bar charts indicate a remarkable growth of interest in applying these models to medical imaging tasks.

2. Background

Function: $spatial\ coordinates \rightarrow values$ Ex: $(x, y, z) \rightarrow RGB\ value$

Spectral bias: A problem in ReLU-based Networks where they weakly capture high-frequency in signals.

Modifications to Address Spectral Bias:

- Input Modification: Map input to higher-dimensional space by a positional encoding.
- Activation Function Modification: Use sine as the activation function (SIREN).
- Output Modification: Each node of the MLP is responsible for reconstructing a part of the signal.

Neural Radiance Fields (NeRFs): a neural volume rendering model that bridges implicit representations and novel view synthesis.

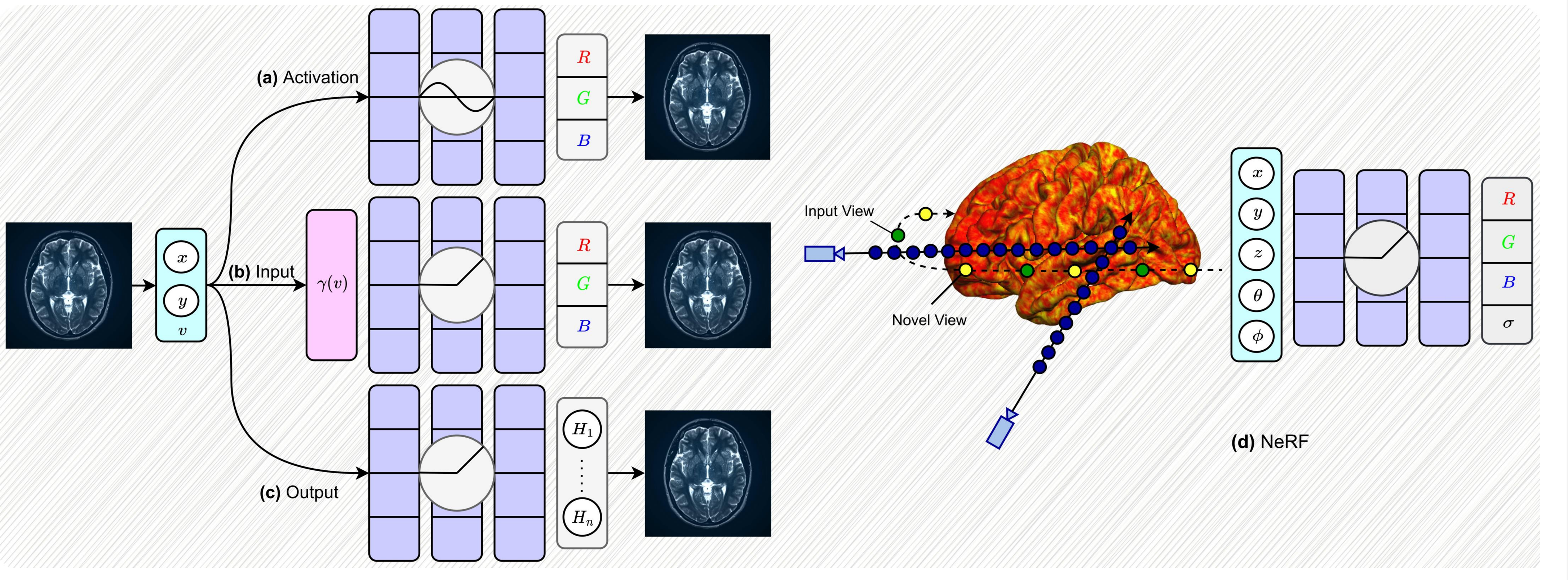

Fig. 2. The figure illustrates various modifications to alleviate the spectral bias problem in INRs, provides an overview of their underlying principles, and introduces NeRF as an additional background method

3. Clinical Importance

1. Eliminating the Need for External Annotations: INRs are valuable because they reduce the reliance on external sources like clinicians and medical professionals for ground truth annotations in medical imaging. This streamlines the annotation process, making it less time-consuming, expensive, and effort-intensive.

2. Improved Imaging Quality: INRs are particularly beneficial for tasks like super-resolution in medical imaging. They can enhance image quality and address issues like motion artifacts, blurred images, or poorly defined structures caused by patient movement during procedures such as CT scans, PET scans, MRI, and ultrasound.

3. Solving Inverse Imaging Problems: INRs enable the reconstruction of CT or MRI scans directly from sensor data, making it possible to track tissue progression over time. This aids in providing updated scans and monitoring changes in medical conditions.

4. Sparse Data Reconstruction: INRs are useful for reconstructing images from sparsely sampled data, which is crucial in applications like reducing radiation dose in CT imaging and accelerating MRI scans. This can improve diagnostic accuracy and aid medical professionals in decision-making.

5. Enhancing Robotic Surgery: Integrating INRs into robotic surgical systems enhances the perception and understanding of the surgical environment. INRs help interpret intraoperative images in real-time, providing feedback to surgeons and aiding in accurate tissue segmentation, anatomical structure localization, and surgical tool manipulation.

4. Taxonomy

We provide a taxonomy with a focus on the application of INRs in several medical imaging tasks. For each task, we review two papers with sufficient detail.

Reconstruction:	1) NeRP: Reconstructs high-quality CT and MRI images from sparsely sampled measurements by embedding a prior from an earlier scan. 2) DCTR: Calculates the loss between the new CT measurement and the measurement taken in the previous time step, then propagates the loss back to the MLP in order to remove noise from the new scan and reconstruct its CT image.
Segmentation:	1) BS-ISR: a combination of INR and CNNs to model the segmentation boundary by mapping CT slice coordinates to spline coefficients. 2) Retinal-INR: An INR model enhances the image resolution, while a Vision Transformer (ViT) extracts features from the original image
Registration:	1) mirnf: uses neural fields to represents the transformation between pair of images. 2) IDIR: uses insights from differentiable rendering to combine implicit deformable image registration model with regularization terms.
Compression:	1) TINC: Uses octree partitioning to enable visually similar blocks to share parameters within a tree-shaped neural network structure 2) SCI: Introduces adaptive partitioning to divide the data into blocks within INR’s spectrum envelop then compresses each block
Neural Rendering	1) MedNeRF: Combines NeRF and a CNN to generate CT projections from X-rays by training NeRF as the generator to output image patches and a CNN as the discriminator to refine NeRF outputs. 2) Surgical Neural Rendering: NeRF-based rendering in robotic surgery that captures non-rigid deformations and reconstruct the 3D structures of scenes.

5. Comparative Overview

1. Preference for Image Reconstruction: Image reconstruction is more popular and has received greater interest compared to others. This preference is driven by its ability to enhance resolution and reduce noise, especially in medical scenarios with uncertain imaging conditions.

2. Local Information: Some methods use CNNs to capture and encode local features and spatial relationships, which helps to create accurate and context-aware representations

3.Sparse View CT Reconstruction: The challenge of reconstructing CT images with sparse and limited measurements is greatly addressed, which is crucial for minimizing radiation exposure. Several methods use prior information or geometric relationships to improve the reconstruction process.

4.Network Type: The methods can be separated into SIREN-based and NeRF-based network designs. Most reviewed works use ReLU MLPs with Fourier mapping for input to mitigate spectral bias. NeRF-based designs are commonly used for volume rendering and view synthesis. The choice of network type depends on the specific task's objectives

6. Open Challenges

1. Computational Complexity and Training Time: Creating neural representations for individual signals demands significant memory and computational resources. INRs can be time-consuming to fit for high-dimensional data like 3D volumes, posing challenges for real-time applications. Techniques such as meta-learning and multi-scale representations aim to speed up training and optimize memory usage.

2. Scaling to Complex Signals: Representing higher-resolution or complex 3D shapes with fine detail is challenging due to the highly nonlinear nature of the mapping. Expanding the model's complexity can help but may lead to computational issues such as vanishing/exploding gradients. Researchers must balance model complexity with available computational resources, and various techniques have been developed to address this.

3. Video-Based INR: INRs perform well in video compression, allowing for parallel processing during decoding. This makes them valuable in robotic-assisted surgery where both speed and accuracy are crucial. However, modeling semantic relationships between frames in high-frequency videos presents challenges, necessitating ongoing research and development.