

NEURAL RADIANCE FIELDS IN MEDICAL IMAGING: CHALLENGES AND NEXT STEPS

Xin Wang^{1,*}, Shu Hu², Heng Fan³, Hongtu Zhu⁴, Xin Li¹

¹ University at Albany, State University of New York {xwang56, xli48}@albany.edu

² Purdue University hu968@purdue.edu

³ University of North Texas heng.fan@unt.edu

⁴ University of North Carolina at Chapel Hill htzhu@email.unc.edu

ABSTRACT

Neural Radiance Fields (NeRF), as a pioneering technique in computer vision, offer great potential to revolutionize medical imaging by synthesizing three-dimensional representations from the projected two-dimensional image data. However, they face unique challenges when applied to medical applications. This paper presents a comprehensive examination of applications of NeRFs in medical imaging, highlighting four imminent challenges, including fundamental imaging principles, inner structure requirement, object boundary definition, and color density significance. We discuss current methods on different organs and discuss related limitations. We also review several datasets and evaluation metrics and propose several promising directions for future research.

Index Terms— Neural Radiance Fields (NeRF), Medical Image, Deep Learning

1. INTRODUCTION

Neural Radiance Fields (NeRFs) significantly advance traditional 3D reconstruction methods in several key aspects [1]. Unlike conventional techniques that often rely on geometric or voxel-based approaches, NeRFs utilize deep learning to create highly detailed and photorealistic 3D models from a set of 2D images [2]. Compared with traditional reconstruction methods, this approach allows for rendering complex scenes with intricate light and shadow interplay, which traditional methods struggle to capture accurately. NeRFs excel in handling occlusions and view-dependent effects, producing more continuous and smoother representations of 3D spaces. Additionally, NeRFs require fewer manual interventions and assumptions about scene geometry, leading to more automated and scalable 3D reconstructions [3]. This capability is particularly advantageous in medical imaging, where precise and detailed visualization of anatomical structures is crucial, offering enhanced diagnostic and surgical planning tools over traditional methods [4].

By employing deep learning techniques to synthesize highly detailed and photorealistic 3D models from sets of 2D

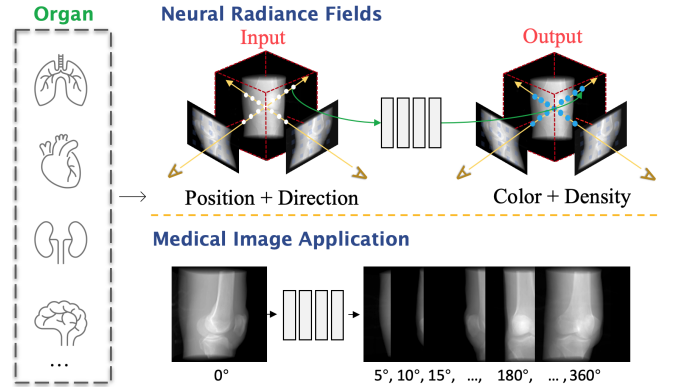


Fig. 1. Overview of this work, discuss NeRF for medical imaging of different organs. **Top:** NeRF model [1]. **Bottom:** An application of NeRF in medical images, the goal is to render the 3D volume from a single view [5, 6].

medical images, such as CT scans or MRIs, NeRFs provide an unprecedented level of detail and accuracy in visualizing complex anatomical structures [1]. This technique surpasses traditional 3D reconstruction methods by effectively handling intricate light interactions and view-dependent phenomena, thus providing clearer, more continuous representations of internal body structures [7]. Furthermore, NeRFs can potentially minimize the need for multiple or high-resolution scans that expose patients to excessive ionizing radiation [5]. Although the NeRFs-based method has yielded surprising results in traditional 3D image reconstruction, it cannot be directly applied to medical image reconstruction due to the unique internal details of medical images [8, 3].

This paper aims to highlight the significant motivation and challenges associated with implementing NeRFs in medical imaging. We have systematically categorized existing methods based on organ types, acknowledging that the distinct characteristics of various organ structures, such as sparsity, overlap, and visibility, pose considerable hurdles to achieving comprehensive morphological reconstructions. These obstacles often lead to incomplete 3D models, potentially omitting crucial details (see Figure 1 for an example). Then, we ex-

* Corresponding author.

plore the future research opportunities that could contribute to overcoming these obstacles and advancing the field.

The key contributions of this work are outlined as follows:

- We outline the key motivation and challenges associated with applying NeRFs to medical imaging, identifying the main hurdles that must be overcome.
- We categorize and review the existing NeRF methodologies based on specific organs, highlighting their limitations and the unique challenges of different anatomical structures.
- Furthermore, we explore potential future research avenues, emphasizing how using NeRFs in medical imaging can revolutionize diagnostic processes, surgical planning, and medical education. This represents a significant advancement in medical visualization and patient care, promising to enhance accuracy and efficacy of treatments.

The remaining content of this paper is organized as follows. After introducing the motivation and challenges of NeRFs in medical imaging in Section 2, we describe the medical image NeRF applications for different organs and discuss the limitations of the current methods in Section 3. In Sections 4, we summarize the public datasets for NeRFs in medical images. We discuss the future directions in Section 5 and conclude the paper in Section 6.

2. MOTIVATION AND CHALLENGES

2.1. Motivation

Neural Radiance Fields (NeRFs) in medical imaging present innovative objectives to enhance imaging techniques in healthcare, addressing several key clinical needs.

(1) Minimizing Radiation Exposure. NeRFs in medical imaging can potentially reduce the necessity for multiple or high-resolution scans that subject patients to high levels of ionizing radiation. By generating 3D representations from a sparse collection of 2D images, these methods can decrease the frequency of exposures, lowering radiation doses and diminishing the risks associated with radiation exposure. This technique not only reduces the risk of radiation-related health conditions but also alleviates issues related to prolonged patient immobility during imaging sessions.

(2) Reducing Time Consumption. Compiling multiple scans into a coherent 3D representation is time-intensive, posing significant challenges in time-sensitive medical situations. NeRFs in medical imaging streamline this process by enabling the quick creation of 3D models from a limited number of images, thereby accelerating data collection and image reconstruction. This rapid processing is particularly advantageous in emergency medicine settings, where efficiency can directly impact patient care and outcomes.

(3) Lowering Imaging Costs. Achieving detailed 3D images with conventional CT scans requires many X-ray projections, each with a fine slice thickness, which can be costly. NeRFs

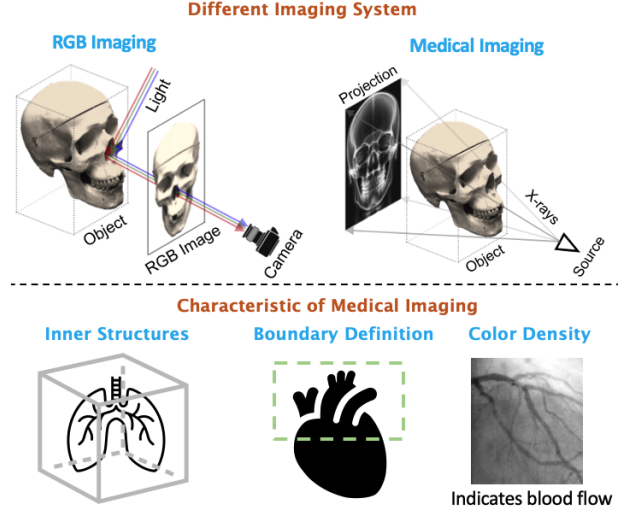


Fig. 2. Overview of the challenges of NeRFs in medical images. *Top:* The fundamental imaging principles of RGB image and medical image are different, an example of X-rays imaging is from [9]. *Bottom:* (left) Medical images necessitate detailed inner structure visualization. (mid) often suffer from poor object boundary definition, and (right) exhibit the varying significance of color density.

in medical imaging address this issue by efficiently producing detailed 3D models from smaller images, potentially reducing the financial burden. The capability of NeRFs in medical imaging to deliver precise volumetric reconstructions with fewer data inputs suggests a decrease in the reliance on expensive imaging resources, offering a cost-effective alternative to traditional imaging methods.

(4) Reduce Motion-related Artifacts. Medical imaging reconstructions typically involve multiple scan acquisitions, causing lengthy sessions that necessitate patient stillness, thereby increasing the risk of motion artifacts. Specifically, Magnetic Resonance Imaging (MRI) sequences are highly susceptible to motion-related artifacts caused by breathing and cardiac pulsation, potentially leading to inaccuracies in 3D reconstruction. Such inaccuracies are critical in clinical settings, where precise anatomical measurements are essential for accurate treatment planning and patient monitoring.

2.2. Challenges

Despite promising performance in reconstructing 3D natural images, Neural Radiance Fields (NeRFs) face several challenges when applied to medical imaging, primarily due to the unique characteristics and requirements of medical image data. An overview of these challenges is shown in Figure 2.

(1) Different Imaging Principles. Fundamental imaging principles in consumer electronics and medical imaging are different. For example, X-ray while color-based (RGB) NeRF methods have seen considerable development. Yet, their direct application to X-ray imaging may yield suboptimal outcomes due to inherent differences between visible light and

X-ray modalities. Unlike visible light imaging, which captures external features by reflecting off an object’s surface, X-ray imaging penetrates the object and attenuates, creating images that primarily reveal internal structures. These internal details are crucial for 3D X-ray reconstruction, highlighting the need for tailored NeRF algorithms to effectively accommodate unique characteristics of X-ray imaging [9].

(2) Need for Detailed Inner Structures. Medical images often require high levels of detail to represent intricate inner structures of the body accurately. NeRFs, which synthesize 3D models from 2D images, must be capable of capturing these details to be clinically useful. However, accurately rendering the complex internal anatomy, such as the fine distinctions between different types of tissues or the subtle anomalies indicative of early-stage diseases, can be challenging. This is due to the limitations in the resolution and the interpolation capabilities of current NeRF models, which may struggle to replicate the exact internal textures and structures needed for precise medical diagnosis and treatment planning.

(3) Ambiguity of Object Boundary. Medical images frequently run into ambiguity with object boundaries between different types of tissues or between pathological and healthy tissues. This ambiguity in boundary delineation poses a significant challenge for NeRFs, as these models rely on clear distinctions in visual data to reconstruct 3D spaces accurately. In medical imaging, where precise boundary identification is crucial for tasks such as tumor localization or the measurement of organ volumes, NeRFs must overcome this hurdle to ensure the reliability and accuracy of generated 3D models.

(4) Significance of Color Density. In medical imaging, variations in color density can convey critical information, such as the presence of blood flow (See Figure 2), variations in tissue density, or the accumulation of contrast agents. These subtle variations are essential for diagnosing conditions and planning treatments. NeRFs, which inherently model the light transport in a scene, must be adapted to interpret and accurately reproduce these color density variations in the synthesized 3D images. The challenge lies in the fact that NeRFs were initially designed for visual scenes where lighting plays a significant role in the appearance of objects, rather than for medical scenarios where color density variations are more closely related to organ properties and physiological conditions.

3. METHODS

The unique properties of various organ structures introduce significant challenges to the reconstruction process, such as sparsity, overlap, and visibility issues, which can hinder the creation of complete morphological models, potentially resulting in missing details in the 3D reconstructions.

A prime example is the vessel structure, especially the challenge of vessel sparsity, where the thin nature of vessels

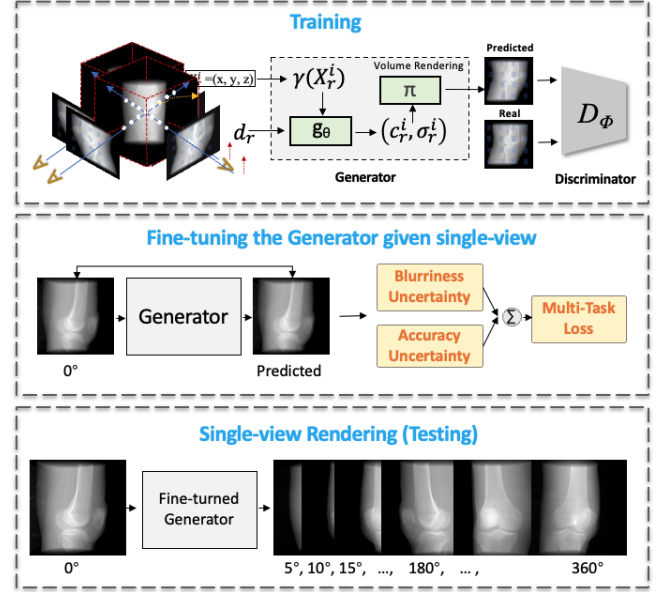


Fig. 3. Over view of *Uncertainty-aware MedNeRF* [6]. **Top:** Network structures and training phrase [5]. **Mid:** Fine-tuning the generator for balancing the *Blurriness* and *Accuracy*. **Bottom:** Rendering the whole 3D volume (each view taken at a 5-degree interval) using a single view slide.

requires a thorough sampling methodology for accurate depiction. Consequently, this section delves into a comprehensive classification based on the application of NeRFs across different organs. Our goal is to acquaint researchers with the essential operations and functionalities of NeRF models in medical imaging, addressing the intricacies involved in capturing the full complexity of organ structures.

3.1. NeRFs for Different Organs

Knee. Knee injuries and degenerative conditions, such as osteoarthritis, are very common [10]. Improved imaging and 3D reconstruction techniques can greatly aid in the diagnosis, treatment planning, and monitoring of such conditions, making the knee a highly relevant subject for research and application of NeRF technology. Furthermore, the knee is a common site for surgical interventions, including arthroscopy, ligament reconstruction, and total knee arthroplasty. NeRF-generated 3D models can enhance preoperative planning and simulation, providing surgeons with detailed visualizations of the anatomy, which can improve surgical outcomes.

Recently, MedNeRF [5], a GAN-based Raditonal Filed, successfully untangles the surface shape, volumetric depth, and internal anatomical structures from 2D images by learning a function that assigns a radiance value to each pixel. MedNeRF is a Deep Learning model capable of reconstructing CT projections from limited or single-view X-ray images. Leveraging a novel architecture based on neural radi-

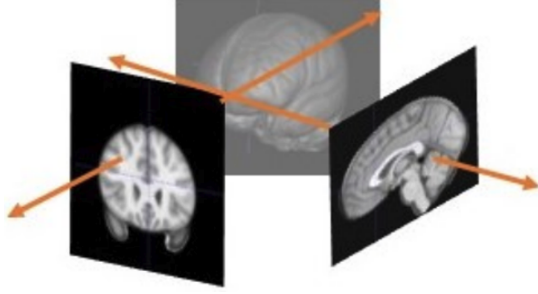


Fig. 4. Overview of extended-MedNeRF with 2D Slices of Brain MRI scans in [11].

ance fields, this model learns a continuous representation of CT scans. It effectively disentangles the shape and volumetric depth of both surface and internal anatomical structures from 2D images. However, due to the nature of GAN network architecture for image generation and using a single X-ray for CT projection reconstruction during image generation, the network may overfit to the single X-ray, leading to an imbalance between image clarity and accuracy. Thus, during the single-view volumetric rendering stage, MedNeRF fine-tuned the generator using the input view by jointly optimizing the reconstruction and the generator. The model showcases high-fidelity renderings, both qualitatively and quantitatively, and is compared favorably against other recent methods based on radiance fields, highlighting its effectiveness in reducing radiation exposure while maintaining imaging quality.

However, the hyperparameters in MedNeRF for balancing two losses (clarity and accuracy) are typically configured manually and subject to observation, and heavily depend on human expertise, which can pose challenges in achieving the optimal trade-off between image blurriness and accuracy. Manual adjustments of these loss weights make it hard to achieve the right balance between image blurriness and accuracy. To address this challenge, Uncertainty-aware MedNeRF (UMedNeRF) [6] is presented to automatically learn hyperparameters for clarity and precision balancing problems. As shown in Figure 3, the key to improving the quality of the generated images lies in simultaneously ensuring image resolution and accuracy. In that case, it becomes a multi-task learning problem. UMedNeRF uses an uncertainty-aware conditional radiance field that simultaneously weighs multiple loss functions automatically in the multi-task settings and generates the predicted patch for single-view volumetric rendering.

More recently, alignment and correction NeRF (ACNeRF) [12] has been proposed to improve MedNeRF in two key areas: (1) enhancing the match between NeRF models and medical imagery by leveraging the models' generalization capabilities more effectively, and (2) refining the reasoning process to better accommodate variations in rendering ranges. ACNeRF is designed to reconstruct projections of various poses from the projection of a single pose, it em-

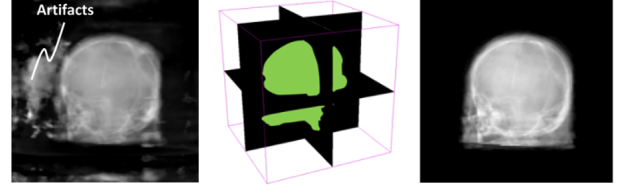


Fig. 5. Overview of Masked NeRF [13] **Left:** Skull projections from NeRF. **Right:** Skull projections from Masked NeRF scene representations.

plays latent codes to linearly influence positional encoding, facilitating image alignment while ensuring the 3D structure remains intact by preventing generator parameter changes from causing disruptions. During the inference phase, a pose judgment network aids in correcting poses, thereby optimizing the model's rendering views. For rendering narrow-range projections, ACNeRF incorporates a novel refinement approach that uses frequency-domain masks to fine-tune the generator.

Brain. Brain Magnetic Resonance Imaging (MRI) is pivotal in medical diagnostics and research because it allows for the detailed visualization of brain structures, aiding in the diagnosis, monitoring, and treatment planning of neurological conditions. Utilizing advanced imaging techniques, e.g., NeRFs, can enhance the quality and accuracy of brain images, and thus improve the detection of abnormalities, such as tumors, lesions, and degenerative diseases, providing crucial information for clinical decision-making and patient care. Besides, NeRFs can contribute to the development of more precise and less invasive diagnostic methods, improving patient outcomes and advancing the understanding of brain anatomy and function.

Specifically, the advent of NeRFs offers a promising solution, enabling the generation of precise 3D reconstructions from MRI images with minimal user input. Leveraging 3D convolutional neural networks, a recent work [11] reconstructs 3D projections from 2D MRI slices, overcoming challenges such as variable slice thickness and integrating multiple MRI modalities for enhanced robustness (see Figure 4). This technique facilitates accurate slice interpolation, capturing the shape and volumetric depth of both surface and internal structures. It also improves the comprehensiveness and reliability of 3D anatomical reconstructions, offering medical professionals the ability to visualize and analyze complex structures with fewer scans, potentially decreasing imaging times and mitigating motion-related complications.

Skulls. Skull Cone Beam Computed Tomography (CBCT) scans are the gold standard for diagnosing cranial fractures. They provide high-resolution images of the bone, allowing clinicians to identify even minor fractures that could be critical to a patient's treatment plan. In the realm of image-guided, minimally invasive medical procedures, accurate pose estimation plays a pivotal role, utilizing X-ray projections to pre-

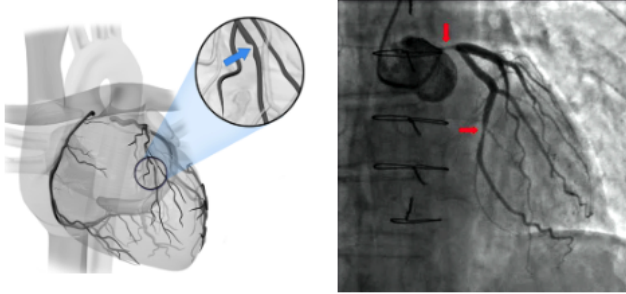


Fig. 6. Overview of Coronary Angiography [14].

cisely navigate to targets within 3D space.

An unconstrained NeRF model applied to skull CBCT reconstructions exhibited artifacts beyond the reconstructed anatomical boundaries [13]. These artifacts compensate for overfitting to a constrained set of ground-truth X-ray images, resulting in implausible visualizations for novel poses. To mitigate these issues, a Masked-NeRF (mNeRF) approach is introduced by leveraging a 3D head mask derived from CBCT scans to spatially constrain NeRF’s extrapolations. The mask, deliberately expanded by 3mm for comprehensive coverage, ensures that densities beyond its perimeter are negated, facilitating seamless integration within the deep learning framework for gradient-based optimization and processing (See Figure 5).

Heart Vessel. Imaging of heart vessels is crucial for diagnosing various heart conditions, including coronary artery disease (CAD), aneurysms, and blockages [15]. It allows doctors to assess the health of the heart’s blood vessels and identify areas of reduced blood flow. An example of Coronary Angiography is shown in Figure 6.

Conversely, the inherent characteristics of vessel structures present additional challenges, such as sparsity, overlap, and visibility, which obstruct the comprehensive reconstruction of vascular morphologies, potentially leading to omissions in the 3D models. These challenges include (1) vessel sparsity, which refers to the slender nature of vessels necessitating a comprehensive sampling approach for accurate reconstruction, (2) vessel overlap, concerning the concealment of crucial vessel segments in 2D projections due to occlusions, and (3) vessel visibility issues, where parts of the vessel may be obscured due to contrast variability or the superimposition of background structures.

A recent study [16] is dedicated to addressing the challenges inherent in the 3D reconstruction of coronary angiography images and exploring the applicability of the Neural Radiance Fields (NeRF) in overcoming these obstacles. Existing methods for the 3D reconstruction of coronary arteries often fall short, hampered by issues such as vessel morphology characteristics including sparsity, overlap, and distinguishing between foreground and background, alongside constraints posed by sparse view and limited angle X-ray pro-

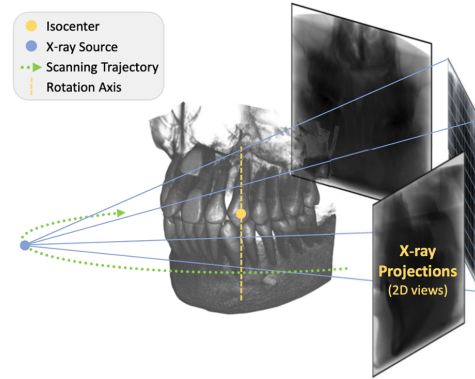


Fig. 7. Overview of Sparse-view Neural Attenuation Fields (SNAF) [18]. An X-ray source systematically traverses around the region of interest, producing sparse projections captured by a flat-panel detector positioned on the opposite side.

jections. NeRF, with its promising outcomes in medical imaging (X-ray) applications, could address these challenges, although its effectiveness in handling specific nuances of X-ray angiography and vessel morphology has yet to be fully demonstrated. This work delves into the potential and limitations of employing NeRF for 3D reconstruction from X-ray angiography, conducting thorough experimental analyses to assess how angiographic challenges impact reconstruction quality. The results indicate NeRF’s promising capabilities for 3D reconstruction in the context of sparse and limited angle X-ray projections, while also pinpointing specific limitations, such as background overlap, that necessitate further research. Note that the method for large vessel Aorta is introduced in [17].

Teeth. Teeth Cone Beam Computed Tomography (CBCT) images enable the detection and diagnosis of various dental conditions such as cavities, tooth decay, impacted teeth, and gum disease. They provide detailed views that are not visible to the naked eye, allowing for early treatment.

Neural Attenuation Fields (NAF) [17] innovatively models the attenuation coefficients as a continuous function across 3D space, utilizing a dense neural network for parameterization. NAF synthesizes projections and refines the network by reducing the discrepancy between actual and generated projections. A novel learning-based encoder, incorporating hash coding, enhances the model’s ability to delineate high-frequency details more effectively than traditional frequency-domain encoders. This improvement is attributed to the encoder’s ability to leverage the inherent smoothness and sparsity of human anatomical structures, leading to superior performance and efficiency in capturing intricate organ details.

Furthermore, efforts to reconstruct high-quality CBCT images from sparse-view 2D projections have been ongoing, yet current leading methods often yield images marred by artifacts and missing fine details (See Figure 7). Sparse-view Neural Attenuation Fields (SNAF) [18] is proposed for sparse-view CBCT reconstruction. The view augmentation

strategy addresses the data insufficiency challenges posed by sparse input views.

Foot. Foot CT images are crucial for diagnosing a wide range of conditions, from fractures and dislocations to degenerative diseases like arthritis and conditions such as plantar fasciitis. Developing imaging techniques that can offer clear, detailed views of the foot’s internal structures can significantly improve diagnostic accuracy and treatment planning [17].

Abdomen. Abdominal CT imaging is crucial for diagnosing diseases such as cancers, inflammatory conditions, and vascular disorders. It helps in detecting tumors, assessing organ damage, identifying blockages or abnormalities in the intestines, and evaluating the extent of diseases (staging), which is essential for treatment planning [17].

Chest. Chest CT imaging stands as a cornerstone of medical diagnostics due to its ability to non-invasively reveal critical information about major organs, its role in a wide array of clinical applications, and its continual evolution alongside imaging technology advancements. The chest houses critical organs such as the lungs, heart, and major blood vessels. Imaging these structures can help diagnose a wide range of conditions, from respiratory infections and lung cancer to heart disease and vascular disorders. Chest images provide valuable insights into the anatomy and potential abnormalities within the thoracic cavity. The works in the previous section [5, 6, 12, 17] also are evaluated using the chest CT images.

3.2. Limitations

Although impressive progress has been made in the performance of NeRF in medical images, there are several concerns over the current methods.

Dataset Quality and Diversity. NeRFs require a substantial amount of paired data to train, specifically diverse and high-quality images from various angles and perspectives, see Section 4 for more details. In medical imaging, obtaining such paired datasets can be difficult due to privacy concerns, the rarity of certain conditions, and the technical and ethical challenges associated with collecting and sharing patient data.

Limited Resolution and Detail. While NeRFs can produce impressive 3D reconstructions, capturing the intricate details necessary for medical diagnosis can be challenging. The resolution and fidelity of NeRF-generated images may not yet match the standards required to visualize the fine anatomical structures or pathological changes critical for accurate medical assessments [19].

Handling of Transparent and Reflective Surfaces. Medical images often contain structures that are semi-transparent or have varying degrees of reflectivity, such as soft tissues versus bones. NeRFs, originally designed for opaque objects in natural light settings, may struggle to accurately model these properties, leading to less accurate reconstructions.

Boundary Ambiguity and Artifacts. The accurate delineation of boundaries between different tissues or pathological versus healthy areas is critical in medical imaging. NeRFs may produce images with ambiguous boundaries or artifacts, especially when dealing with complex or overlapping structures, potentially leading to misinterpretations.

Generalization Across Different Imaging Modalities. Medical imaging encompasses a wide range of modalities (e.g., MRI, CT, and X-ray), each with its own unique characteristics. NeRFs developed for one type of imaging may not generalize well on another without significant adjustments or retraining, limiting their versatility [20].

Interpretability and Clinical Validation. For NeRFs to be adopted in clinical practice, their outputs must be interpretable by medical professionals and validated against established medical imaging standards [21]. The black-box nature of deep learning models, including NeRFs, can make it challenging to understand how they generate their outputs, complicating clinical validation and trust.

4. DATASETS AND EVALUATION METRICS

Public datasets play a pivotal role in the advancement of Neural Radiance Fields (NeRFs) in medical imaging, offering a wealth of diverse and annotated data essential for the development and validation of algorithms.

4.1. Overview

Training Medical Image NeRFs needs datasets comprising paired 2D (e.g. X-ray) and 3D (e.g. CT, MRI) scans. However, assembling a real dataset of such pairs is virtually impractical due to the rarity of obtaining X-ray pairs and CT (or MRI) scans captured simultaneously in clinical practice [22]. Thus, the Reconstructed Radiographs (DRR) technique is employed for simulating and generating X-ray images. This method entails the creation of virtual X-ray projections using three-dimensional medical image data, such as CT or MRI. By generating DRRs, we can avoid the need of paired X-rays and corresponding CT reconstructions, which would otherwise subject patients to increased radiation exposure. The overview relation between Digitally Reconstructed Radiograph (DRR) and medical imaging NeRFs is shown in Figure 8.

4.2. Digitally Reconstructed Radiographs (DRR)

Digitally Reconstructed Radiographs (DRRs) are synthetic X-ray images generated from three-dimensional (3D) computed tomography (CT) data using digital image processing techniques. This technology allows for the creation of X-ray-like images without the need for additional radiation exposure to the patient. DRRs are produced by simulating the path of

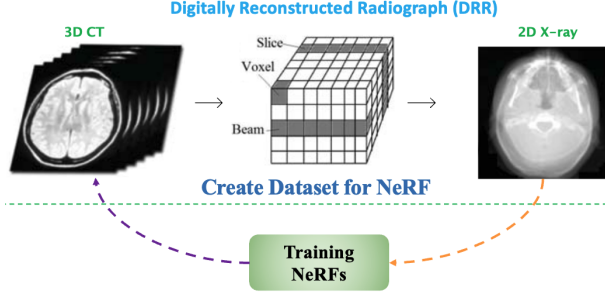


Fig. 8. The relation between Digitally Reconstructed Radiograph (DRR) and medical imaging NeRFs [23]. **Left:** leverages original CT scan slices, **Mid:** perceived as a three-dimensional voxel array, **Right:** to generate a DRR. This direct volume rendering scheme efficiently transforms the volumetric data from a CT scan into a visually comprehensive DRR, offering a synthesized view similar to traditional X-ray imagery.

X-rays through a 3D volume dataset, mimicking the attenuation and absorption properties of X-rays as they pass through different tissues and structures within the body.

The primary utility of DRRs lies in their application within various medical fields, including radiation oncology, orthopedics, and interventional radiology, among others. In radiation therapy planning, for instance, DRRs are used to accurately target treatment areas, allowing for precise alignment of the therapeutic radiation beam with the tumor while sparing surrounding healthy tissue. Similarly, in orthopedics, DRRs assist in pre-surgical planning and the alignment of prosthetic implants by providing clear images of bone structures.

Furthermore, DRR technology facilitates the comparison and verification of patient positioning during treatment, ensuring that the actual patient alignment matches the planned treatment setup. By leveraging existing CT datasets to generate these radiograph-like images, DRRs eliminate the need for additional diagnostic X-ray examinations, thereby reducing the cumulative radiation dose to the patient and enhancing the efficiency of medical imaging workflows. More DRR generation examples can be found in [23].

4.3. Public Dataset

MedNeRF Dataset [5] (Github [24]). The dataset includes 20 CT chest scans as in [25] and [26], in addition to 5 CT knee scans documented in [27] and [28]. These scans represented a diverse array of patients, encompassing various contrast modalities and encapsulating both typical and atypical anatomical structures. The DRR generation process involves the emulation of the rotation of a radiation source and imaging panel around the vertical axis, resulting in the creation of DRRs with a resolution of 128×128 at five-degree intervals. Consequently, each object produced a total of 72 distinct DRRs.

NAF Dataset [17] (Github [29]). This publicly available CT datasets of human organs, covering the chest, jaw, foot, and abdomen. The chest dataset is sourced from the LIDC-IDRI collection, while datasets for the jaw, foot, and abdomen are obtained from the Open Scientific Visualization Datasets. Given that these datasets solely provide volumetric CT scans, the projections are produced using the tomographic toolbox TIGRE.

PerX2CT [22] (Github [30]). To reconstruct CT images from X-rays, a dataset comprising paired X-ray and CT scans is essential. However, assembling a real dataset of such pairs is virtually impractical due to the rarity of obtaining X-ray pairs and CT scans captured simultaneously in clinical practice. Consequently, existing research has turned to digitally reconstructed radiograph (DRR) technology as a solution, enabling the generation of synthetic X-rays from actual CT volumes to simulate the required paired data for studies, the dataset is constructed following [31].

4.4. Evaluation Metrics

Various methods have been developed to evaluate the effectiveness of NeRFs in medical imaging applications.

3D Projection Reconstruction. To assess the efficacy of the model in reconstructing CT projections, two principal image quality metrics: Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) are used [5]. PSNR focuses on the accuracy of luminance, quantifying the variance between the original and the model-generated images. Meanwhile, SSIM offers a holistic evaluation, taking into account factors such as luminance, contrast, and structural integrity.

Utilizing these metrics is essential for a precise evaluation of the model’s performance in CT projection reconstruction, providing a reliable basis for the research outcomes. These metrics are widely recognized for their effectiveness in measuring the quality of image reconstructions and CT projections, underscoring their importance in the analysis.

2D Rendering. Another evaluation of the model focuses on the task of 2D rendering, utilizing the Frechet Inception Distance (FID) and Kernel Inception Distance (KID) metrics for a comprehensive comparison of image quality, where lower scores signify superior quality [5].

When benchmarking against prevalent 2D rendering methods such as pixelNeRF [32] and GIRAFFE [33, 34] on the datasets, these methods fall short, failing to reach the performance levels of the GRAF baseline or the model. Conversely, the models [5, 6] demonstrate superior performance, outperforming baseline models in terms of both FID and KID metrics across all tested datasets.

5. FUTURE DIRECTIONS

Besides continuing to address the aforementioned limitations, we also envision a few important directions of NeRFs in medical imaging that will receive more attention in the future.

Enhanced Detail Resolution. Future developments should focus on improving the resolution and detail that NeRFs can capture, especially for internal structures within the medical images. Advanced algorithms could be designed to more effectively learn fine-grained details from limited or sparse imaging data, potentially through the integration of higher-resolution imaging techniques or the development of more sophisticated methods.

Improved Boundary Delineation. Addressing the issue of poor boundary definition in medical images will be crucial. Research could explore the use of hybrid models that combine NeRFs with segmentation networks or the incorporation of edge-detection mechanisms within the NeRF architecture to better distinguish between different types of tissues and pathological versus healthy areas.

Accurate Color and Density Representation. Since color density variations convey vital information in medical images, future NeRF models could benefit from enhancements in accurately modeling and reproducing these variations. This could involve adapting NeRFs to more precisely interpret and generate images based on the organ properties and physiological conditions that these color densities represent, perhaps by integrating additional data sources or modalities that provide complementary information.

Reducing Computational Demands. The computational intensity of training and deploying NeRF models, especially for high-resolution medical imaging data, poses a significant challenge. Future research should aim to optimize these models for greater efficiency, potentially through algorithmic improvements, the use of more efficient neural network architectures, or leveraging hardware accelerations.

Customization and Personalization. Tailoring NeRFs to individual patient data for personalized medicine could greatly enhance diagnostic accuracy and treatment planning. Future directions might include developing adaptable NeRF models that can quickly learn from new patient data or incorporating dynamic adjustments based on real-time imaging during surgical procedures.

Integration with Diagnostic. Exploring the integration of NeRF-generated 3D models into diagnostic workflows, surgical planning, and even real-time image guidance during procedures could have profound implications for patient care. This includes the development of tools and interfaces that allow clinicians to interact with and manipulate 3D reconstructions for better planning and execution of treatments.

Integration with new Advanced Techniques. The future of NeRFs in medical imaging is poised for significant advance-

ments, particularly with the integration of new techniques.

For example, NeRFs could integrate with 3D Gaussian Splatting [35] for real-time radiance field rendering, which promises to address some of the current limitations of NeRFs, such as the high computational cost and the need for faster rendering times without compromising the quality of the 3D reconstructions. 3D Gaussian Splatting, by efficiently approximating the continuous radiance fields with discrete samples, can greatly accelerate the rendering process. This method, when applied to NeRFs in medical imaging, could facilitate real-time visualization of complex anatomical structures, making it invaluable for surgical planning and intra-operative navigation. The computational intensity of NeRFs currently limits their application in scenarios where quick decision-making is critical, such as emergency medicine or real-time surgical assistance. Incorporating 3D Gaussian Splatting could substantially reduce the computational load, enabling the deployment of NeRF-based applications on more accessible hardware, including mobile devices used in remote diagnostics.

Besides, the integration of NeRFs with Reinforcement Learning (RL) [36] in medical imaging represents a forward-looking approach that can significantly improve the efficiency, accuracy, and personalization of medical imaging and related procedures. By learning from interactions between data and environments, RL models can optimize NeRFs to meet specific needs of medical imaging, addressing challenges such as scan time, patient comfort, and the complexity of interpreting medical images [37]. This synergy has potential to transform medical diagnostics and treatment planning, making healthcare more efficient and effective for patients.

Furthermore, as mentioned before, medical images often come with noise and artifacts due to various reasons like patient movement, suboptimal equipment settings, or inherent limitations of imaging modalities. Robust machine learning (RML) algorithms are designed to work effectively even in the presence of such noise [38, 39], ensuring consistent diagnostic performance. Generally, the need for RML in NeRFs becomes increasingly critical in high-stakes domains like medical applications, where the inherent risks significantly heighten the importance of such robustness.

Last but not least, the Foundation Models (FMs) [40] offer a promising approach to address some of the unique challenges faced in medical image analysis and reconstruction [41]. The integration of NeRFs with FMs in medical imaging could lead to significant advancements in accuracy, efficiency, and generalizability across different medical imaging modalities and clinical applications. FMs are trained on vast datasets, enabling them to learn a wide range of features and patterns [42]. When applied to medical imaging NeRFs, these models can help overcome the challenge of data variability across different medical conditions and imaging modalities. The depth and complexity of FMs allow them to capture high-level abstractions and intricate details from medical images

[43]. When used to generate or refine NeRFs, these models can potentially improve the accuracy of 3D reconstructions, enhance the detection of diseases, and provide more precise anatomical models for simulation purposes.

6. CONCLUSION

In conclusion, the exploration of Neural Radiance Fields (NeRF) in medical imaging illuminates a promising pathway toward enhancing diagnostic precision, treatment planning, and early disease detection. This paper has dissected the multifaceted challenges inherent to applying NeRFs to medical image data, such as adapting to the distinct imaging principles of medical modalities, capturing the intricate internal structures necessary for accurate diagnostics, defining object boundaries with precision, and interpreting color density variations critical for medical analysis. Through a comprehensive examination of current methodologies and an insightful discussion on the limitations and potential of NeRFs, we emphasize the necessity for innovative solutions that can navigate these obstacles.

As we venture into future advancements, the integration of NeRFs with emerging technologies and methodologies holds the promise to significantly propel medical imaging forward, emphasizing the critical need for continued research, collaborative efforts, and the development of novel approaches to fully harness the capabilities of NeRFs in revolutionizing medical imaging.

Acknowledgments. Copyrights of the figures belong to their own authors and/or holders (sources are cited in the captions). The author sincerely thanks all readers for taking the time to read the manuscript. If any related work has been overlooked, please do not hesitate to contact us. The feedback will be greatly appreciated and help us considerably improve this work.

Xin Wang is supported by University at Albany Start-up Grant.

7. REFERENCES

- [1] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng, “Nerf: Representing scenes as neural radiance fields for view synthesis,” 2021, vol. 65, ACM New York, NY, USA.
- [2] Fang Zhu, Shuai Guo, Li Song, Ke Xu, Jiayu Hu, et al., “Deep review and analysis of recent nerfs,” *APSIPA Transactions on Signal and Information Processing*, vol. 12, no. 1, 2023.
- [3] AKM Rabby and Chengcui Zhang, “Beyondpixels: A comprehensive review of the evolution of neural radiance fields,” *arXiv preprint arXiv:2306.03000*, 2023.
- [4] Pechin Lo, Bram Van Ginneken, Joseph M Reinhardt, et al., “Extraction of airways from ct (exact’09),” *IEEE Transactions on Medical Imaging*, vol. 31, no. 11, 2012.
- [5] Abril Corona-Figueroa, Jonathan Frawley, et al., “Mednerf: Medical neural radiance fields for reconstructing 3d-aware ct-projections from a single x-ray,” in *EMBC. IEEE*, 2022.
- [6] Jing Hu, Qinrui Fan, Shu Hu, Siwei Lyu, Xi Wu, and Xin Wang, “Umednerf: Uncertainty-aware single view volumetric rendering for medical neural radiance fields,” *IEEE ISBI*, 2024.
- [7] Weinan Song, Haoxin Zheng, Jiawei Yang, Chengwen Liang, and Lei He, “Oral-nexf: 3d oral reconstruction with neural x-ray field from panoramic imaging,” *arXiv preprint arXiv:2303.12123*, 2023.
- [8] Amirali Molaei, Amirhossein Aminimehr, Armin Tavakoli, Amirhossein Kazerouni, Bobby Azad, Reza Azad, and Dorit Merhof, “Implicit neural representation in medical imaging: A comparative survey,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 2381–2391.
- [9] Yuanhao Cai, Jiahao Wang, Alan Yuille, Zongwei Zhou, and Angtian Wang, “Structure-aware sparse-view x-ray 3d reconstruction,” *arXiv preprint arXiv:2311.10959*, 2023.
- [10] Nicholas Bien, Pranav Rajpurkar, Robyn L Ball, Jeremy Irvin, Allison Park, Erik Jones, Michael Bereket, Bhavik N Patel, Kristen W Yeom, Katie Shpanskaya, et al., “Deep-learning-assisted diagnosis for knee magnetic resonance imaging: development and retrospective validation of mrnet,” *PLoS medicine*, vol. 15, no. 11, pp. e1002699, 2018.
- [11] Khadija Iddrisu, Sylwia Malec, and Alessandro Crimi, “3d reconstructions of brain from mri scans using neural radiance fields,” in *International Conference on Artificial Intelligence and Soft Computing*. Springer, 2023, pp. 207–218.
- [12] Mengcheng Sun, Yu Zhu, Hangyu Li, Jiongyao Ye, and Nan Li, “Acnerf: enhancement of neural radiance field by alignment and correction of pose to reconstruct new views from a single x-ray,” *Physics in Medicine & Biology*, vol. 69, no. 4, pp. 045016, 2024.
- [13] Chaochao Zhou, Syed Hasib Akhter Faruqui, Abhinav Patel, Ramez N Abdalla, Michael C Hurley, Ali Shaibani, Matthew B Potts, Babak S Jahromi, Leon Cho, Sameer A Ansari, et al., “Robust single-view cone-beam x-ray pose estimation with neural tuned tomography (nett) and masked neural radiance fields (mnerf),” *arXiv preprint arXiv:2308.00214*, 2023.
- [14] Wikipedia contributors, “Coronary catheterization — wikipedia, the free encyclopedia,” 2023, [Online; accessed 2023-02-05].
- [15] Xin Wang and Hongtu Zhu, “Artificial intelligence in image-based cardiovascular disease analysis: A comprehensive survey and future outlook,” *arXiv preprint arXiv:2402.03394*, 2024.
- [16] Kirsten WH Maas, Nicola Pezzotti, Amy JE Vermeer, Danny Ruijters, and Anna Vilanova, “Nerf for 3d reconstruction from x-ray angiography: Possibilities and limitations,” in *VCBM 2023: Eurographics Workshop on Visual Computing for Biology and Medicine*. Eurographics Association, 2023, pp. 29–40.
- [17] Ruyi Zha, Yanhao Zhang, and Hongdong Li, “Naf: neural attenuation fields for sparse-view cbct reconstruction,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2022, pp. 442–452.

- [18] Yu Fang, Lanzhu Mei, Changjian Li, Yuan Liu, Wenping Wang, Zhiming Cui, and Dinggang Shen, “Snaf: Sparse-view cbct reconstruction with neural attenuation fields,” *arXiv preprint arXiv:2211.17048*, 2022.
- [19] Zixuan Chen, Lingxiao Yang, Jian-Huang Lai, and Xiaohua Xie, “Cunerf: Cube-based neural radiance field for zero-shot medical image arbitrary-scale super resolution,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 21185–21195.
- [20] Magdalena Wysocki, Mohammad Farid Azampour, Christine Eilers, Benjamin Busam, Mehrdad Salehi, and Nassir Navab, “Ultra-nerf: neural radiance fields for ultrasound imaging,” in *Medical Imaging with Deep Learning*. PMLR, 2024, pp. 382–401.
- [21] Yuehao Wang, Yonghao Long, Siu Hin Fan, and Qi Dou, “Neural rendering for stereo 3d reconstruction of deformable tissues in robotic surgery,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2022, pp. 431–441.
- [22] Daeun Kyung, Kyungmin Jo, Jaegul Choo, Joonseok Lee, and Edward Choi, “Perspective projection-based 3d ct reconstruction from biplanar x-rays,” in *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp. 1–5.
- [23] J Montúfar, Marcelo Romero, V Muñoz-Jiménez, Rogelio J Scougall-Vilchis, and Brissa Jiménez, “Perspective and orthogonal cbct/ct digitally reconstructed radiographs compared to conventional cephalograms,” in *Conf. Biomedical Engineering and Sciences—BIOENG’16*, 2018, pp. 41–45.
- [24] Link, “Mednerf dataset,” <https://github.com/abrilcf/mednerf>, 2024.
- [25] E. B. Tsai, S. Lungren, M. P. Hershman, M. Roshkovan, L. Colak, et al., “Medical imaging data resource center (midrc) - rsna international covid open research database (ricord) release 1b - chest ct covid,” 2021.
- [26] K. Clark, B. Vendt, K. Smith, J. Freymann, et al., “The cancer imaging archive (tcia): Maintaining and operating a public information repository,” 2013, vol. 26.
- [27] Michael D. Harris, Adam J. Cyr, Azhar A. Ali, et al., “A combined experimental and computational approach to subject-specific analysis of knee joint laxity,” 2016, vol. 138.
- [28] Azhar A. Ali, Sami S. Shalhoub, Adam J. Cyr, et al., “Validation of predicted patellofemoral mechanics in a finite element model of the healthy and cruciate-deficient knee,” 2016.
- [29] Link, “NAF CBCT: A Repository for Sparse-View CBCT Reconstruction Using Neural Adjacency Fields,” <https://github.com/Ruyi-Zha/nafebct>.
- [30] Link, “PerX2CT: A GitHub Repository for Cross-Modal Medical Imaging Conversion,” <https://github.com/dek924/PerX2CT>.
- [31] Xingde Ying, Heng Guo, Kai Ma, Jian Wu, Zhengxin Weng, and Yefeng Zheng, “X2ct-gan: reconstructing ct from biplanar x-rays with generative adversarial networks,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 10619–10628.
- [32] Alex Yu, Vickie Ye, Matthew Tancik, and Angjoo Kanazawa, “pixelnerf: Neural radiance fields from one or few images,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021.
- [33] Katja Schwarz, Yiyi Liao, Michael Niemeyer, and Andreas Geiger, “Graf: Generative radiance fields for 3d-aware image synthesis,” 2020, vol. 33.
- [34] Michael Niemeyer and Andreas Geiger, “Giraffe: Representing scenes as compositional generative neural feature fields,” in *CVPR*, 2021.
- [35] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis, “3d gaussian splatting for real-time radiance field rendering,” *ACM Transactions on Graphics*, vol. 42, no. 4, 2023.
- [36] Xin Wang, Ziwei Luo, Jing Hu, Chengming Feng, Shu Hu, Bin Zhu, Xi Wu, Xin Li, and Siwei Lyu, “Deep reinforcement learning for image-to-image translation,” *arXiv preprint arXiv:2309.13672*, 2023.
- [37] Danny Driess, Ingmar Schubert, Pete Florence, Yunzhu Li, and Marc Toussaint, “Reinforcement learning with neural radiance fields,” *Advances in Neural Information Processing Systems*, vol. 35, pp. 16931–16945, 2022.
- [38] Shu Hu, Yiming Ying, Xin Wang, and Siwei Lyu, “Sum of ranked range loss for supervised learning,” *JMLR*, vol. 23, no. 112, pp. 1–44, 2022.
- [39] Shu Hu, Xin Wang, and Siwei Lyu, “Rank-based decomposable losses in machine learning: A survey,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023.
- [40] Rishi Bommasani, Drew A Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S Bernstein, Jeannette Bohg, Antoine Bosselut, Emma Brunskill, et al., “On the opportunities and risks of foundation models,” *arXiv preprint arXiv:2108.07258*, 2021.
- [41] Yunkun Zhang, Jin Gao, Zheling Tan, Lingfeng Zhou, Kexin Ding, Mu Zhou, Shaoting Zhang, and Dequan Wang, “Data-centric foundation models in computational healthcare: A survey,” *arXiv preprint arXiv:2401.02458*, 2024.
- [42] Yukun Zhou, Mark A Chia, Siegfried K Wagner, Murat S Ayhan, Dominic J Williamson, Robbert R Struyven, Timing Liu, Moucheng Xu, Mateo G Lozano, Peter Woodward-Court, et al., “A foundation model for generalizable disease detection from retinal images,” *Nature*, 2023.
- [43] Shaoting Zhang and Dimitris Metaxas, “On the challenges and perspectives of foundation models for medical image analysis,” *arXiv preprint arXiv:2306.05705*, 2023.