A Brief Survey On Neural Surfaces

Yue Zeng
yuezeng4@illinois.edu

Chaitanya Amballa
amballa2@illinois.edu

Abstract

The field of 3D vision has seen considerable progress, especially in surface reconstruction. Recent methods utilize implicit representations, such as Signed Distance Functions (SDFs), to derive continuous 3D surfaces from discrete inputs, including point clouds and multi-view images. Unlike traditional modular approaches that separate feature extraction, surface estimation, and reconstruction, neural surface methods enable end-to-end optimization; leading to more coherent and accurate surface models. In this survey, we introduce and examine SDFs, both implicit and explicit representations through volume and surface rendering, along with the design choices that contributed to advancements in the field. We conclude by discussing some future directions and research ideas. Table 1 summarizes our survey.

1. Introduction

In recent years, the field of 3D representations has witnessed remarkable advancements, driven by the increasing demand for realistic virtual environments in various applications, ranging from graphics, VR/AR, robotics, and autonomous systems. Traditional geometric representations, such as polygonal meshes and voxel grids, have long been the backbone of 3D modeling but often suffer from limitations in expressivness, resolution, and the ability to represent complex surfaces. Consequently, this has led to the quest for better representations that leverage the power of neural networks, offering a more flexible and efficient means of capturing the intricate details of 3D structures.

One of the key problems of interest in *how do we better represent and learn these 3d structures*? One of the early advances in this domain is neural volume rendering, a technique that synthesizes novel 2D views from sparse 2D images by learning a 3D volumetric representation using neural networks, enabling the realistic visualization of 3D scenes. Volume rendering refers to the process of displaying 3D data as a 2D image, where each point in the 3D space contributes to the final pixel color based on its properties (such as density and color)

Although traditional volume rendering techniques have existed for a while, they do not easily integrate with deep learning methods. The introduction of a differentiable rendering technique mitigated this problem by allowing gradients to be backpropagated through the rendering process, enabling the integration of rendering with deep learning. Using frameworks such as NVDiffRast[12] and PyTorch3D [24], one can now optimize 3D representations directly from images, thereby facilitating the training of neural networks on large datasets without the need for explicit 3D annotations. This shift has led to a surge of interest in leveraging neural representations for various tasks, including shape modeling, surface reconstruction, and material estimation.

At the forefront of these developments is Neural Radiance Fields aka NeRFs [19] , which utilizes neural networks to model 3D scenes as continuous functions that map spatial coordinates to a color and density. By synthesizing novel views of a scene from a sparse set of input images (see [19][35]), NeRFs have demonstrated very realistic images in generating complex 3D scenes. Subsequent improvements, such as MipNeRF[1] , Instant NGP[20], and Gaussian Splats[10], have enhanced both the efficiency and quality of neural rendering, allowing for even more accurate 3D reconstructions.

While neural volume rendering and differentiable rendering have achieved notable success, accurately and efficiently representing 3D surfaces, particularly for complex topologies, is also one of the key challenges of interest. Neural surfaces provide a promising approach, utilizing implicit functions defined by neural networks to efficiently capture intricate details and handle complex topologies in 3D geometry. This survey will explore both implicit and explicit surface representation techniques, reviewing their design choices and applications, and identifying gaps in current research. We propose new directions for further exploration in the field of neural rendering. Table 1 summarized various papers of our interest, illustrating their representations, whether implicit, explicit, or hybrid, along with their respective rendering types, input modalities, and network outputs.

Paper	Representation	Rendering Type	Input Type	Network Output
Occupancy [18]	implicit	direct 3d supervision	point clouds, single images	occupancy
DeepSDF[23]	implicit	direct 3d supervision	point clouds	SDF
IGR [5]	implicit	direct 3d supervision	point clouds	SDF
VolSDF [33]	implicit	volume	2D images	(transformed) SDF
BakedSDF [34]	implicit	volume	2D images	(contracted coordinated) SDF
Lumigraph [9] [21] [32]	implicit	surface	2D images	SDF
UNISURF[22]	implicit	volume	2D images	occupancy
NeuS[26] NeuS2 [26] UNISURF[22]	implicit	volume	2D Images	SDF with *S-density*
SparseNeus [16], SceneUs [8]	implicit	volume	sparse 2D images	SDF + geometry-aware consistency
NKF [29], NKS [7]	kernels	linear combination of learned kernel basis functions with marching cubes	sparse noise point clouds	features+normals at voxels
NESI [37]	explicit	ray-traced rendering	3D shapes	occupancy
ENS [25]	explicit	neural deferred shading	multi-view images and binary masks	deformation fields + features
Efficient 3D GAN [2]	explicit-implicit with triplanes	volume with Importance Sampling	point clouds	3D features

Table 1. Comparison of Various 3D Reconstruction Papers

2. Neural Surface Representations

3D shape representations can be categorized into traditional explicit methods, like polygonal meshes and voxel grids, and neural implicit methods. Traditional methods enable real-time rendering with fast inference, but suffer from aliasing and limited smoothness in complex shapes due to their discrete nature. In contrast, neural implicit representations utilize neural networks to define shapes as continuous functions, producing watertight surfaces with intricate geometries but typically involve slower inference and complex initialization.

Sec. 2.1 discusses SDFs as a method to represent surfaces, with early works using explicit grid-based storage. Sec. 2.2 covers methods using neural networks to represent 3D surfaces implicitly, often via SDFs, combining volume rendering for high-quality reconstructions without explicit masks. Sec. 2.3 discusses methods using voxel grids and deformation fields for direct surface modeling, emphasizing efficiency, detail, and real-time applications like medical imaging.

2.1. SDFs

Level sets are a general framework that represents surfaces as the set of points where a function takes on a constant value. A signed distance function (SDF) defines the distance from a point in space to the nearest surface, with positive values indicating points outside the surface, negative values for points inside, and zero for points on the surface. The connection between SDFs and level sets is evident at the 0 level set, where the SDF equals zero, effectively defining the surface:

$$\mathcal{S} = \{\mathbf{x} \mid \phi(\mathbf{x}) = c\} \tag{1}$$

where c is a constant, and $\mathcal{S}$ represents the surface defined by the level set function.

$$SDF(\mathbf{x}) = \begin{cases} d(\mathbf{x}, \mathcal{S}) & \text{if } \mathbf{x} \text{ is outside the surface} \\ -d(\mathbf{x}, \mathcal{S}) & \text{if } \mathbf{x} \text{ is inside the surface} \\ 0 & \text{if } \mathbf{x} \text{ is on the surface} \end{cases} \tag{2}$$

where $d(\mathbf{x}, \mathcal{S})$ is the Euclidean distance from the point $\mathbf{x}$ to the nearest point on the surface $\mathcal{S}$.

Recent works has shown that representing shapes as level sets of neural networks is effective for shape analysis and reconstruction tasks.

One of the early works on using SDFs in neural rendering is presented in DeepSDF [23], where the authors propose directly regressing the SDF function using a multilayer perceptron (MLP). They showed that this method provides a more effective representation of 3D objects compared to traditional approaches like voxels, point clouds, and meshes.

IGR [5] presents a method for learning high-fidelity implicit neural representations directly from point clouds by employing implicit geometric regularization. This approach uses a loss function designed to promote the neural network's convergence to zero on the input data while preserving unit norm gradients. As a result, it generates smooth and intuitive zero-level set surfaces without the need for 3D supervision. This is accomplished through the Eikonal first-order partial differential equation, expressed as $||\nabla_x f(x; \theta) - 1||^2$.

2.2. Implicit Representations

We have seen the importance of Signed Distance Functions (SDFs) in 3D reconstruction, but the above methods primarily rely on supervision that uses 3D points as ground truth. However, obtaining these 3D points is not always feasible. In contrast, access to 2D images is often easier and more readily available. In this section, we will explore how to learn and reconstruct 3D structures by using various implicit representations.

Building on IGR [5], VolSDF [33] addresses the problem of volume rendering for implicit surfaces. While surface rendering focuses on rendering objects by modeling their surfaces explicitly (and thus concerned with what happens at the surface), volume rendering (popularized by [19]), synthesizes images by integrating information across a 3D volume to render novel views. It is non-trivial to find a proper threshold to extract surfaces from the predicted density. To alleviate this, VolSDF[33] proposes representing density as a learnable transformation of the SDF. Results show more accurate surface reconstructions compared to NeRFs [19].

NeRFs [19] [1], along with their successors like Instant NGP [20] are known to be slow both in training and inference. While MobileNeRF [3] render novel views in real-time, the underlying representation consists of a polygon soup and is not viable for applications like appearance editing and relighting. BakedSDF [34] proposes a way to render novel views in real-time for larger scenes by defining an SDF in contracted coordinate space, allowing them to extract the mesh in these contracted spaces.

UNISURF [22] aims to combine the benefits of surface rendering and volume rendering. It learns implicit surfaces via volume rendering, representing the surface with occupancy values. Additionally, UNISURF enhances reconstruction quality by reducing the sample region of volume rendering during the optimization process.

NeuS [26] also addresses a similar problem of surface reconstruction method from 2D images without relying on foreground masks (see [21] [32]). It represents surfaces as the as SDF encoded by a fully connected neural network (MLP) and introduces a new volume rendering technique to effectively minimize geometric errors during the reconstruction process. To optimize the parameters of the SDF and color field, NeuS employs a volume rendering approach that utilizes an S-density defined as

$$\phi_s(f(x)) = \frac{se^{-sx}}{(1 + e^{-sx})^2}$$

to weight contributions from different points along the viewing ray, ultimately ensuring that the intersection point with the zero-level set of the SDF has the greatest impact on the rendered output color. Results show that they consistently outperform [22].

Previous works in [9] [21] [32] rely on segmentation masks to achieve their goals. In contrast, Neuralangelo [15] eliminates the need for auxiliary information and employs a volume rendering approach that samples points along a ray and utilizes multi-resolution hash encoding (similar to InstantNGP) to generate the SDF and color.

NeuS2 [27] further improves upon [26] and [15] and can reconstruct the scene in few seconds by addressing their limitations: 8hr training time of [26] and noisy reconstructed surfaces of [15], by utilizing multi-resolution hash encodings first introduced in [20].

Instead of learning surfaces via SDFs, [17] represents the 3D surface as a level set of a SDF with a validity branch for estimating the surface existence probability at the query positions. This additional layer of representation enhances the accuracy of surface predictions in areas with sparse data or occlusions.

While the aforementioned works perform well with a reasonable number of images, [8] and [16] focus on few-shot scenarios (approximately three images) for 3D object reconstruction. They learn generalizable priors from image features using geometry encoding volumes to account for the less data. Techniques such as a multi-level geometry reasoning framework for coarse-to-fine surface recovery, multi-scale color blending scheme for improved color prediction, and a consistency-aware fine-tuning strategy to address inconsistencies arising from occlusion and noise are employed to show performant results.

One of the advantages of implicit representations is the ability to work in continuous representation. However, implicit models have been shown [18] to work well for simple objects they tend to struggle with complex scenes, where intricate geometries or varied topologies present challenges. To address these challenges, Pet-NeuS [28] proposes using local PointNet to process input features, which are then projected onto canonical planes. By doing so, the method captures local geometric variations more effectively, allowing it to handle more complex and varied scenes. In the following section, we will explore various explicit representations.

2.3. Explicit Representations

Traditional explicit methods offer some advantages by representing surfaces directly through meshes or point clouds, allowing for more efficient sampling and faster rendering times without the need for complex volumetric techniques. In contrast to implicit representations that require indirect surface modeling, explicit methods provide better control over geometric details, making them preferable in applications requiring high precision or real-time performance.

Explicit Neural Surfaces often rely on deformation fields to represent continuous geometry. These fields map a known base domain into the target shape through smooth transformations. This approach directly encodes the surface's topology, allowing for high-fidelity surface details. ENS[25] transforms a known base mesh into the target surface using neural deformation fields, enabling fast training and inference, along with high-quality mesh extraction.

Another class of methods for ENS leverages surface intersections to represent shapes explicitly. These methods rely on a set of local surfaces that intersect to form the global shape, making the representation both compact and versatile. NESI[37] represents 3D shapes via height-field surface intersections, achieving efficient geometry processing while maintaining high compression and accuracy.

Voxel-based approaches are often combined with explicit surface representations to accelerate surface reconstruction. These methods use voxel grids as the foundation for representing the geometry but introduce explicit neural mechanisms to capture surface details and ensure efficient optimization. Voxurf[30] employs voxel grids and neural networks to accelerate surface reconstruction with a two-stage training process, achieving a balance between efficiency and detail.

In the domain of medical applications, particularly for soft tissue reconstruction, several methods leverage explicit neural surface representations to enhance real-time performance and geometric accuracy. SDFPlane[13], EndoSurf[36], and LerPlane[31] all utilize explicit surface representations to address the challenges of deformable tissue modeling. SDFPlane accelerates optimization by employing a spatial-temporal structure encoder and a SDF decoder, achieving faster reconstruction with accurate geometric and visual rendering. EndoSurf reconstructs high-fidelity tissue geometry by modeling shape and texture with separate neural fields, making it suitable for RGBD endoscopic video analysis. LerPlane factorizes surgical scenes into 2D planes, significantly reducing memory usage while enabling fast and precise tissue reconstruction in real-time applications. These methods collectively advance the field of soft tissue reconstruction in surgical environments.

2.4. Applications and Other Frontiers

In recent years, neural surfaces have found diverse applications, including 3D shape reconstruction, novel view synthesis, architectural modeling, and transparent object reconstruction.

For example, [6] uses SDFs to predict manhattan floor plans from multi-view images by integrating planar constraints into implicit neural representation methods

Expanding beyond architectural contexts, [21] proposes a novel differentiable rendering formulation for learning implicit shape and texture representations, addressing the limitations of existing methods that rely on voxel- and mesh-based representations for 3D reconstruction. By leveraging implicit differentiation to derive depth gradients analytically, the proposed method enables the training of reconstruction models directly from RGB images, achieving single-view reconstructions comparable to those obtained with full 3D supervision and facilitating multi-view 3D reconstruction that produces watertight meshes. Similarly, [9] introduces a neural rendering pipeline that significantly accelerates view synthesis by optimizing an implicit surface and radiance field from posed 2D images, achieving real-time rendering rates while maintaining exceptional image quality and enabling mesh export with view-dependent texture information. Neural surface methods have also been adapted to encode finer geometric details. For instance, [14] proposes a method for learning 3D edge representations that effectively captures both lines and curves by encoding 3D edge distance and direction using unsigned distance functions derived from multi-view edge maps, along with a robust edge extraction algorithm.

For more complex geometries, [11] presents a point-guided homotopy-based optimization scheme for training SDFs on raw point clouds, addressing challenges in fitting to complex geometries by using guiding points to facilitate incremental changes towards the true shape. Additionally, the authors introduce a metric to quantify geometric differences, enhancing the optimization process and demonstrating significant improvements over existing methods, especially for difficult shapes.

Moving towards transparent object reconstruction, [4] proposes a framework for reconstructing transparent objects by explicitly incorporating physical refraction and reflection tracing into the surface reconstruction process, enabling end-to-end optimization based solely on multi-view RGB images.

Building on the recent success of neural surface techniques, [28] extends the recent NeuS method for neural surface reconstruction by introducing three key components: a tri-plane representation to enhance expressiveness, a novel learnable positional encoding to mitigate noise during reconstruction, and self-attention convolution operations for improved feature extraction from tri-plane data. These approaches collectively improve the robustness and expressiveness of neural surface models, enabling more accurate reconstructions.

In terms of computational efficiency, [2] work presents a hybrid explicit-implicit network architecture that enhances the computational efficiency and image quality of 3D GANs, enabling the unsupervised generation of high-resolution, multi-view-consistent images and high-quality 3D shapes from single-view 2D photographs. By decoupling feature generation from neural rendering and leveraging state-of-the-art 2D CNN generators like StyleGAN2, the framework achieves real-time synthesis while maintaining 3D consistency and geometry quality.

More recently, a different line of work that was motivated by the limitations of [18] emerged in reconstructing 3d shapes. Neural Kernel Fields (NKF), introduced in [29] advocates for reconstructing implicit 3D shapes using learned kernel ridge regression, which achieves state-of-the-art results for reconstructing both 3D objects and large scenes from sparse oriented points.. Building on the NKF framework, the authors in [7] present an approach that enhances scalability to large scenes via compactly supported kernel functions and memory-efficient sparse linear solvers. This method also exhibits robustness to noise through a gradient fitting solution, minimizing training requirements, and allowing effective learning from diverse datasets. As a result, the method enables rapid reconstruction of millions of points in seconds, achieving state-of-the-art performance across various reconstruction benchmarks, including single objects and complex indoor and outdoor scenes.

3. Design choices

Since surfaces are typically continuous and exhibit a strong sense of structural similarity, independently predicting opacities and SDFs for any given location may lead to non-smooth surfaces. To address this issue, various types of regularizations have been applied to enforce smoothness. One notable regularization technique that is used in [38], [33], [15], and [26] is the Eikonal loss introduced in [5]. This loss is defined to encourage the network to produce outputs that satisfy the Eikonal equation resulting in smoother surfaces (SDFs).

In certain applications, when the normals are known [5], additional regularizations are employed to smooth out the normals as well.

4. Ideas

While many methods [23] [33] [5] [26] [18] we saw earlier focus on static scenes, there is a need for efficient 3D reconstruction techniques that can handle **dynamic environments** (video) where objects and lighting conditions may change over time. Investigating methods that allow real-time updates to the reconstructed model as new images are acquired may be an interesting idea worth pursuing.

Current methods rely predominantly on image data for reconstruction. Investigating the incorporation of **additional modalities**, such as LiDAR data, and thermal images, could improve the accuracy and detail of the 3D reconstruction. Also, in situations with sparse information—where only a few images are available (see [16] [8], one can see if using the modalities can lead to better reconstruction than just using images. One can even see if we can try to do 3d reconstruction without using images say with audio alone. This approach opens up possibilities for reconstructing scenes using entirely different types of data, expanding the potential applications of 3D reconstruction techniques.

While implicit representations are gaining traction, the potential for **hybrid** approaches that combine both implicit and explicit methods (see [2]) has not been thoroughly explored. This could lead to more robust models that leverage the strengths of both representations.

Although works in [8] [16] address few-shot reconstruction, developing models that can effectively learn from even fewer images (e.g., single-view or few-pixel samples) maintaining high fidelity in complex scenes is an important challenge. This is a bigger question to answer as this may need new solutions that cannot be extended by existing works.

Like with many vision solutions, many of these methods work well in controlled environments and may struggle with noise present in real-world data, especially when images are captured from low-grade cameras. Investigating strategies to enhance **robustness** against noise in input data could be an interesting problem.

One of the key problems of interest that hasn't picked up yet is making the models **physics-grounded**. While some work have shown progress, it is crucial to be explainable. So, combining 3d reconstruction techniques with physics-based approaches may lead to better generalizations thereby allowing to enhancements in scene editing (moving a complete object), changing lighting effects, AR/VR games.

5. Conclusion

Neural Surface Reconstruction is an area that has seen notable progress, with a variety of innovative approaches leveraging both implicit and explicit methods for 3D surface representation. We learned how signed distance functions and their ability to model continuous surface variants are being used as the key to modeling implicit methods for surface reconstruction. We also explored various approaches to 3D reconstruction, highlighting the advancements in implicit, explicit, and hybrid representations. We then highlighted several applications where neural surfaces and rendering techniques have been utilized, followed by a discussion of key design choices that influence their effectiveness. We also proposed several problems of interest for further exploration, underscoring the potential of this emerging field.

References

[1] Jonathan T Barron, Ben Mildenhall, Matthew Tancik, Peter Hedman, Ricardo Martin-Brualla, and Pratul P Srinivasan. Mip-nerf: A multiscale representation for anti-aliasing neural radiance fields. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 5855–5864, 2021. 1, 3

[2] Eric R Chan, Connor Z Lin, Matthew A Chan, Koki Nagano, Boxiao Pan, Shalini De Mello, Orazio Gallo, Leonidas J Guibas, Jonathan Tremblay, Sameh Khamis, et al. Efficient geometry-aware 3d generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 16123–16133, 2022. 2, 5

[3] Zhiqin Chen, Thomas Funkhouser, Peter Hedman, and Andrea Tagliasacchi. Mobilenerf: Exploiting the polygon rasterization pipeline for efficient neural field rendering on mobile architectures, 2023. 3

[4] Weijian Deng, Dylan Campbell, Chunyi Sun, Shubham Kanitkar, Matthew E Shaffer, and Stephen Gould. Differentiable neural surface refinement for modeling transparent objects. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20268–20277, 2024. 5

[5] Amos Gropp, Lior Yariv, Niv Haim, Matan Atzmon, and Yaron Lipman. Implicit geometric regularization for learning shapes. *arXiv preprint arXiv:2002.10099*, 2020. 2, 3, 5

[6] Haoyu Guo, Sida Peng, Haotong Lin, Qianqian Wang, Guofeng Zhang, Hujun Bao, and Xiaowei Zhou. Neural 3d scene reconstruction with the manhattan-world assumption. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5511–5520, 2022. 4

[7] Jiahui Huang, Zan Gojcic, Matan Atzmon, Or Litany, Sanja Fidler, and Francis Williams. Neural kernel surface reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4369–4379, 2023. 2, 5

[8] Shi-Sheng Huang, Zixin Zou, Yichi Zhang, Yan-Pei Cao, and Ying Shan. Sc-neus: Consistent neural surface reconstruction from sparse and noisy views. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 2357–2365, 2024. 2, 3, 5

[9] Petr Kellnhofer, Lars C Jebe, Andrew Jones, Ryan Spicer, Kari Pulli, and Gordon Wetzstein. Neural lumigraph rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4287–4297, 2021. 2, 3, 4

[10] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splats: Real-time high-quality reconstruction of 3d scenes using neural networks, 2023. 1

[11] Chamin Hewa Koneputugodage, Yizhak Ben-Shabat, Dylan Campbell, and Stephen Gould. Small steps and level sets: Fitting neural surface models with point guidance. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 21456–21465, 2024. 4

[12] Samuli Laine, Janne Hellsten, Tero Karras, Yeongho Seol, Jaakko Lehtinen, and Timo Aila. Modular primitives for high-performance differentiable rendering, 2020. 1

[13] Hao Li, Jiwei Shan, and Hesheng Wang. Sdfplane: Explicit neural surface reconstruction of deformable tissues. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2024*, pages 542–552, Cham, 2024. Springer Nature Switzerland. 4

[14] Lei Li, Songyou Peng, Zehao Yu, Shaohui Liu, Rémi Pautrat, Xiaochuan Yin, and Marc Pollefeys. 3d neural edge reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 21219–21229, 2024. 4

[15] Zhaoshuo Li, Thomas Müller, Alex Evans, Russell H Taylor, Mathias Unberath, Ming-Yu Liu, and Chen-Hsuan Lin. Neuralangelo: High-fidelity neural surface reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8456–8465, 2023. 3, 5

[16] Xiaoxiao Long, Cheng Lin, Peng Wang, Taku Komura, and Wenping Wang. Sparseneus: Fast generalizable neural surface reconstruction from sparse views. In *European Conference on Computer Vision*, pages 210–227. Springer, 2022. 2, 3, 5

[17] Xiaoxu Meng, Weikai Chen, and Bo Yang. Neat: Learning neural implicit surfaces with arbitrary topologies from multi-view images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 248–258, 2023. 3

[18] Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, and Andreas Geiger. Occupancy networks: Learning 3d reconstruction in function space. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4460–4470, 2019. 2, 4, 5

[19] 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. *Communications of the ACM*, 65(1):99–106, 2021. 1, 3

[20] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM transactions on graphics (TOG)*, 41(4):1–15, 2022. 1, 3

[21] Michael Niemeyer, Lars Mescheder, Michael Oechsle, and Andreas Geiger. Differentiable volumetric rendering: Learning implicit 3d representations without 3d supervision. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 3504–3515, 2020. 2, 3, 4

[22] Michael Oechsle, Songyou Peng, and Andreas Geiger. Unisurf: Unifying neural implicit surfaces and radiance fields for multi-view reconstruction, 2021. 2, 3

[23] Jeong Joon Park, Peter Florence, Julian Straub, Richard Newcombe, and Steven Lovegrove. Deepsdf: Learning continuous signed distance functions for shape representation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 165–174, 2019. 2, 3, 5

[24] Nikhila Ravi, Jeremy Reizenstein, David Novotny, Taylor Gordon, Wan-Yen Lo, Justin Johnson, and Georgia Gkioxari. Accelerating 3d deep learning with pytorch3d. *arXiv:2007.08501*, 2020. 1

[25] Thomas Walker, Octave Mariotti, Amir Vaxman, and Hakan Bilen. Explicit neural surfaces: Learning continuous geometry with deformation fields, 2023. 2, 4

[26] Peng Wang, Lingjie Liu, Yuan Liu, Christian Theobalt, Taku Komura, and Wenping Wang. Neus: Learning neural implicit surfaces by volume rendering for multi-view reconstruction. *arXiv preprint arXiv:2106.10689*, 2021. 2, 3, 5

[27] Yiming Wang, Qin Han, Marc Habermann, Kostas Daniilidis, Christian Theobalt, and Lingjie Liu. Neus2: Fast learning of neural implicit surfaces for multi-view reconstruction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 3295–3306, 2023. 3

[28] Yiqun Wang, Ivan Skorokhodov, and Peter Wonka. Pet-neus: Positional encoding tri-planes for neural surfaces. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12598–12607, 2023. 4, 5

[29] Francis Williams, Zan Gojcic, Sameh Khamis, Denis Zorin, Joan Bruna, Sanja Fidler, and Or Litany. Neural fields as learnable kernels for 3d reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 18500–18510, 2022. 2, 5

[30] Tong Wu, Jiaqi Wang, Xingang Pan, Xudong Xu, Christian Theobalt, Ziwei Liu, and Dahua Lin. Voxurf: Voxel-based efficient and accurate neural surface reconstruction, 2023. 4

[31] Chen Yang, Kailing Wang, Yuehao Wang, Xiaokang Yang, and Wei Shen. Neural lerplane representations for fast 4d reconstruction of deformable tissues, 2023. 4

[32] Lior Yariv, Yoni Kasten, Dror Moran, Meirav Galun, Matan Atzmon, Basri Ronen, and Yaron Lipman. Multiview neural surface reconstruction by disentangling geometry and appearance. *Advances in Neural Information Processing Systems*, 33:2492–2502, 2020. 2, 3

[33] Lior Yariv, Jiatao Gu, Yoni Kasten, and Yaron Lipman. Volume rendering of neural implicit surfaces. *Advances in Neural Information Processing Systems*, 34:4805–4815, 2021. 2, 3, 5

[34] Lior Yariv, Peter Hedman, Christian Reiser, Dor Verbin, Pratul P Srinivasan, Richard Szeliski, Jonathan T Barron, and Ben Mildenhall. Bakedsdf: Meshing neural sdfs for real-time view synthesis. In *ACM SIGGRAPH 2023 Conference Proceedings*, pages 1–9, 2023. 2, 3

[35] 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*, pages 4578–4587, 2021. 1

[36] Ruyi Zha, Xuelian Cheng, Hongdong Li, Mehrtash Harandi, and Zongyuan Ge. Endosurf: Neural surface reconstruction of deformable tissues with stereo endoscope videos, 2023. 4

[37] Congyi Zhang, Jinfan Yang, Eric Hedlin, Suzuran Takikawa, Nicholas Vining, Kwang Moo Yi, Wenping Wang, and Alla Sheffer. Nesi: Shape representation via neural explicit surface intersection, 2024. 2, 4

[38] Jingyang Zhang, Yao Yao, Shiwei Li, Tian Fang, David McKinnon, Yanghai Tsin, and Long Quan. Critical regularizations for neural surface reconstruction in the wild. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6270–6279, 2022. 5