DynaSlice: Multi-Class 4D Neural SDF Reconstruction from Dynamic Cross Sections

Thomas Walker[1], João Paulo Silva do Monte Lima[2,3], Diana Aldana Moreno[2], Oisín Nolan[4], Salvatore Esposito[1], and Tiago Novello[2]

[1] University Of Edinburgh
[2] IMPA
[3] Universidade Federal Rural de Pernambuco
[4] Eindhoven University of Technology

Abstract. Surface reconstruction from planar cross-sections is a long-standing and challenging problem with applications in medical imaging, manufacturing, and scientific visualization. In this setting, the lack of data between slicing planes results in erroneous geometric laddering artifacts with standard point-cloud reconstruction methods. While current bespoke methods provide powerful implicit neural representations (INR) for recovering static geometry from sparse slices, extending these approaches to *dynamic* and *multi-class* settings remains unexplored. In this setting, time-wise observations are also sparse, making the 4D reconstruction problem severely under-constrained and prone to spurious surfaces and temporal inconsistencies. We propose **DynaSlice**, an INR framework for reconstructing dynamic SDFs directly from temporal 2D slices. DynaSlice models evolving geometry as a *single* multichannel 4D neural SDF parameterized by a sinusoidal network (SIREN) and jointly optimized across space and time, avoiding per-frame fitting and explicit surface correspondences. We show that SIRENs naturally act as a smooth interpolator across time and space, enabling coherent surface evolution between observed slices. To further stabilize reconstruction in unobserved intervals, we introduce explicit *temporal frequency control*, which attenuates high-frequency temporal oscillations while preserving spatial detail. The multichannel formulation enables the simultaneous reconstruction of multiple structures within a unified INR, where shared network parameters capture common space–time features. Experiments on synthetic sequences and real medical data show that DynaSlice demonstrates improved geometric accuracy, and orders of magnitude faster training times than existing approaches.

Keywords: Dynamic reconstruction · Signed Distance Functions (SDF) · Implicit neural representations (INR) · Cross-sections · Medical imaging

1 Introduction

Reconstructing surfaces from planar cross-sections is a long-standing problem in computer vision and graphics, with important applications in medical imaging [12,15,19] and manufacturing [7]. Typically, CT/MRI/Ultrasound scans are segmented by a doctor or an automated algorithm to produce a set of shape boundaries on arbitrarily oriented 2D planes. Recovering a coherent 3D surface is challenging due to the sparse and non-uniform sampling of the target surface. Classical approaches which rely on interpolation [1–3,5] result in noticeable staircasing artifacts and slicing discontinuities. Similarly, out-of-the-box point-cloud reconstruction methods often fail due to the data sparsity between planes [38,43].

Recent bespoke approaches employ implicit neural representations (INRs), especially neural signed distance functions (SDFs), and demonstrated strong performance for surface reconstruction from cross-sections [38, 43]. However, these approaches focus on reconstructing *static* geometry [2,8,38,43]. Extending cross-sectional reconstruction to *dynamic* settings, where the surface evolves over time, remains an open problem. In this setting, data-sparsity along both the time and slicing axes presents a significant challenge. Cross-sectional scanning of dynamic surfaces is common in medical imaging, for example capturing breathing and cardiac cycles [48]. In these contexts, potentially arbitrary cross-section orientations occur according to the acquisition procedure, e.g. in 3D ultrasound [11] or cineMRI [4], or due to re-slicing as a post-process, where non-parallel slices lead to improved visualization and easier segmentation of the target [27, 35].

In practice, medical segmentations are often *multi-class*, including several simultaneous 4D organ sequences that must be reconstructed consistently over time. A naïve approach might be to apply existing static methods, such as CrossSDF [43] or OReX [38], independently to each organ and timestep, and interpolating the resulting fields across the temporal dimension. However, this strategy quickly becomes impractical: for clinical sequences with tens of time steps, reconstructing each organ individually requires hours of computation per frame [43], or days per-organ. Consequently, naïvely extending static SDF approaches to multi-organ, multi-frame reconstruction does not scale efficiently and motivates the need for methods that jointly model spatial and temporal correlations across all structures.

We introduce DynaSlice, a framework for reconstructing *4D surfaces* directly from dynamic 2D cross-sections (see Figure 1). Our method represents surface evolution as a multi-class 4D neural SDF defined over space and time, with individual structures encoded channel-wise in the network output. To address the challenge of sparsity, we leverage the expressive power of SIREN networks [40] while introducing an explicit temporal frequency band-limiting mechanism. Unlike standard SIRENs, which treat space and time identically, our approach separates the temporal dimension and constrains its spectral content according to the temporal sampling rate. By doing so, the network captures the appropriate level of temporal detail without destabilizing optimization or introducing oscillations, resulting in smooth and stable surface evolution.

Experiments show that DynaSlice outperforms state-of-the-art methods for dynamic surface reconstruction, producing high-fidelity 4D surfaces that are free of staircase artifacts and temporal inconsistencies.

In summary, our contributions are:

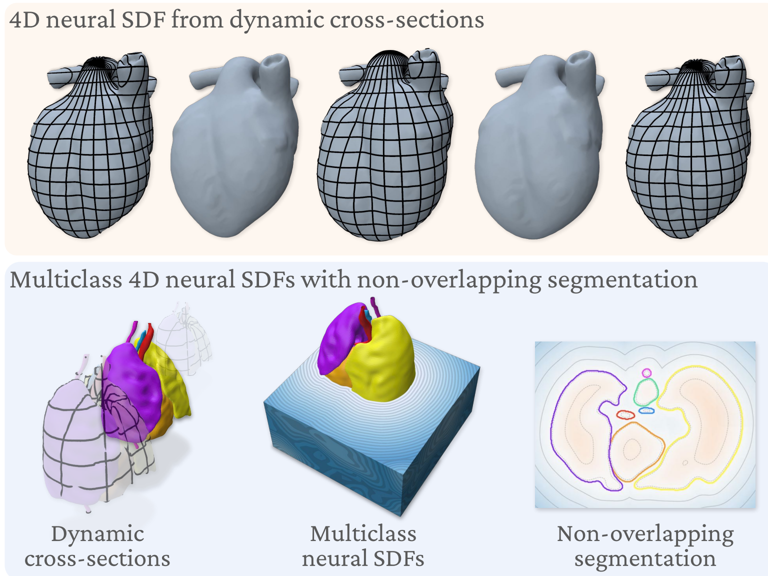

Fig. 1: We present **DynaSlice**, a robust INR framework for reconstructing 4D neural SDFs from sparse, unorganized dynamic cross-sections (top). By parameterizing the scene as a unified multi-channel neural SDF with explicit temporal frequency control, our method enables coherent surface evolution over time and produces continuous multi-structure segmentations without requiring explicit frame-to-frame correspondences (bottom).

- We propose a novel 4D neural SDF framework for reconstructing dynamic surfaces directly from temporal 2D cross-sections.
- A frequency-controlled sinusoidal representation that separately limits spatial and temporal frequencies, ensuring stable and accurate surface evolution.
- A multi-channel SDF enables the simultaneous reconstruction of multiple structures within a single implicit field, allowing efficient and accurate 3D reconstruction from multi-class medical segmentations.
- We demonstrate that DynaSlice achieves higher reconstruction fidelity compared to state-of-the-art methods, producing temporally smooth surfaces, without staircase artifacts.

2 Related Works

Neural Implicit Surface Reconstruction. Implicit neural representations (INRs) [40, 46] have emerged as a powerful method for encoding geometry as a continuous function of spatial coordinates. Early work parameterized surfaces as occupancy functions [20] or SDFs [13, 25, 30, 39, 47], enabling high-quality reconstruction from 3D supervision. Sinusoidal representation networks [40] replaced standard activations with periodic functions that act as a smooth interpolator across all input coordinates, recovering fine geometric detail that standard coordinate networks. Multi-resolution hash encodings [21] and random Fourier feature embeddings [41] provide complementary mechanisms for overcoming this bias. Neuralangelo [18] and NeuS2 [45] combine multi-resolution hash grids with neural surface rendering, achieving state-of-the-art dense reconstruction from multi-view imagery. Data-driven methods such as POCO [6] improve scalability by computing per-point latent codes via point convolutions, enabling reconstruction across both object and scene scales. Without explicit awareness of the planar acquisition geometry, coordinate networks interpret the inter-slice space between contours as freely reconstructible space, forcing the surface normals to align with the slicing planes and producing the well-known laddering artifact: reconstructed geometry collapses toward a stack of independent discs rather than a coherent volumetric surface [38, 43].

Surface Reconstruction from Cross-Sections. Surface reconstruction from planar observations is a long-standing problem, particularly in medical and geospatial imaging, where anatomy or terrain is represented through a sequence of annotated slices [1, 2, 5]. Early approaches establish correspondences between contours on neighboring planes and explicitly interpolate between them, leading to clear laddering artifacts. More recently, optimization-based formulations have recast cross-sectional reconstruction as an implicit surface-fitting problem. OReX [38] demonstrated that neural representations can infer smooth, continuous geometry directly from contour supervision, avoiding interpolation artifacts. CrossSDF [43] extended this by using hash-grid encodings and specialized loss functions to capture finer anatomical structures such as blood vessels. However, these recent approaches do not account for two prevalent modalities of medical imaging data: temporal sequences and multi-class anatomical segmentation. We propose a method that jointly models both from cross-sectional inputs.

4D Surface Reconstruction. Classical level set methods [28, 29] represent moving interfaces as the zero level set of a time-varying scalar field, but rely on hand-crafted velocity fields and dense volumetric grids that preclude learning from data. DynamicFusion [22] and Fusion4D [10] replaced these with learned warp fields over single and multi-view RGBD streams, respectively, while OccupancyFlow [23] lifted occupancy networks [20] to 4D via a continuous velocity field integrated with a neural ODE [9], recovering deforming surfaces from point cloud sequences with implicit correspondences. NeRF-based methods brought deformable reconstruction to monocular video: D-NeRF [34], Non-Rigid NeRF [42], and Nerfies [31] parametrise per-scene deformation as a learned warp into a canonical frame, with HyperNeRF [32] adding an ambient slicing surface for topological changes; CaDeX [17] further decouples shape from motion via bijective homeomorphic maps for cross-instance generalisation, and CanFields [44] handles arbitrary-length sequences of uncorrelated point clouds via a diffeomorphic motion flow in a fully unsupervised setting. NISE [26] trains on $\mathbb{R}^3 \times \mathbb{R}$ under the level set equation with a data term at t=0 and an unsupervised PDE term; Sang et al. [36] couple a modified level-set equation with Eikonal constraints to a volume-preserving velocity field, and 4Deform [37] extends this to sequences with physical and geometric regularization for free-topology interpolation. Every method above requires 3D geometry to be directly observable as point clouds, depth maps, or images; none addresses recovering a 4D surface from sparse planar cross-sections, where 2D contours on discrete planes are the only input.

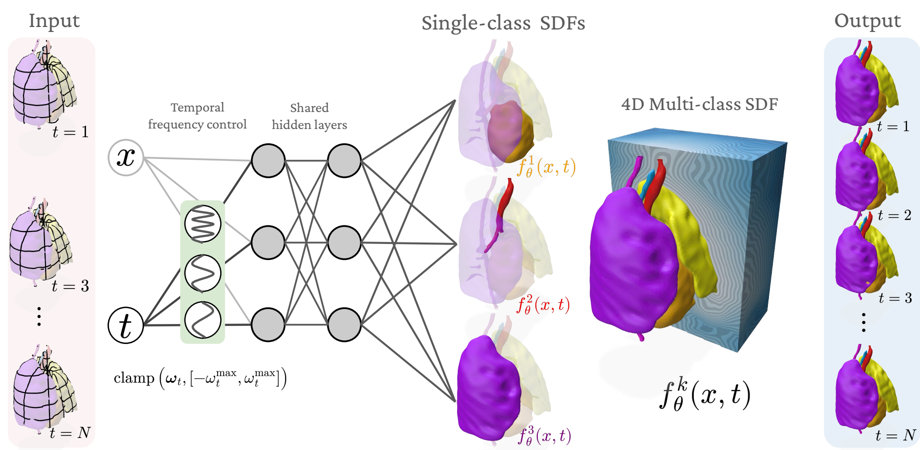

Fig. 2: Overview of DynaSlice. From sparse annotated temporal cross-sections, DynaSlice learns a *single* space–time (4D) multi-class neural SDF $f_\theta^k(\mathbf{x}, t)$ with shared hidden layers and explicit temporal frequency control (clamped ω_t). The unified model is optimized jointly over all timesteps, producing temporally coherent surface evolution and non-overlapping multi-structure reconstructions at inference time.

3 DynaSlice

DynaSlice (see Fig. 2) models evolving geometry as a continuous 4D implicit neural representation (INR) parameterized by a SIREN network [40]. We introduce several key components. First, a *temporal frequency control* module addresses the asymmetric sparsity of spatial and temporal supervision by limiting the temporal frequency spectrum during training. Second, an *orientation consistency loss* encourages correct inside–outside classification of the implicit surface. Third, a *multi-class* output layer enables the simultaneous reconstruction of multiple structures by predicting multiple SDFs within a single network.

3.1 Problem Statement

Our objective is to reconstruct K unknown evolving surfaces $\{S_t^k\}$ over time from temporal 2D cross-sections of K anatomical structures. For each timestep

t_i, we are given a collection of cross-sections $\mathcal{C}_i = \{C_i^1, \ldots, C_i^K\}$, where each C_i^k is set of cross-sections of the corresponding surface $S_{t_i}^k$.

DynaSlice represents the dynamic multi-class geometry implicitly on the zero-level set of a single multi-channel time-dependent neural SDF,

$$f : \mathbb{R}^3 \times [0, T] \to \mathbb{R}^K.$$

That is, the k-th coordinate function $f^k(\cdot, t_i)$ represents the SDF evolution of S_t^k. For clarity of exposition, we first present the network and loss details for a *single SDF class* ($K = 1$, with $f : \mathbb{R}^3 \times [0, 1] \to \mathbb{R}$) in the following two subsections before extending the formulation to the full multi-class model in Sec. 3.4.

3.2 Temporal Frequency Control

We propose learning a time-dependent INR for modeling f with explicit temporal band-limiting according to the time sampling rate (Sec. 3.2). We modify the first layer of a SIREN network [40] to control the bandlimit of each input coordinate independently, which is crucial for stabilizing space–time reconstruction from sparse temporal slices. Specifically, we parameterize f_θ as a composition of $d+1$ layers:

$$f_\theta(\mathbf{x}, t) = \mathbf{L} \circ \mathbf{S}^d \circ \cdots \circ \mathbf{S}^0(\mathbf{x}, t), \qquad \mathbf{S}^i(\mathbf{y}) = \sin\big(\mathbf{W}^i \mathbf{y} + \mathbf{b}^i\big), \tag{1}$$

where $\mathbf{W}^i$ and $\mathbf{b}^i$ are the weights and biases of the i-th layer, and $\mathbf{L}$ is an affine output layer.

Space–time frequency disentanglement. A key ingredient of DynaSlice is controlling the temporal bandlimit independently from the spatial one. We achieve this by separating the spatial and temporal frequency components in the *first* layer. Let $\boldsymbol{\omega} := \mathbf{W}^0$ and $\boldsymbol{\varphi} := \mathbf{b}^0$ be the input frequencies and phase shifts. We decompose the frequencies as $\boldsymbol{\omega} = [\boldsymbol{\omega}_{\mathbf{x}}, \boldsymbol{\omega}_t]$, such that the input embedding becomes:

$$\mathbf{S}^0(\mathbf{x}, t) = \sin\big(\boldsymbol{\omega}_{\mathbf{x}} \mathbf{x} + \boldsymbol{\omega}_t t + \boldsymbol{\varphi}\big). \tag{2}$$

Standard SIREN initialization samples all entries of $\boldsymbol{\omega}$ from the same range (e.g., $\omega_0 = 30$[5]), which is often adequate for static fitting but problematic in our setting. Because temporal supervision is sparse, large $|\boldsymbol{\omega}_t|$ magnitudes can induce oscillations between observed timesteps. In contrast, Eq. (2) exposes $\boldsymbol{\omega}_t$ explicitly, allowing us to enforce a temporal bandlimit consistent with the temporal sampling rate while keeping $\boldsymbol{\omega}_{\mathbf{x}}$ large enough to capture spatial detail.

We now show that the space–time frequency structure of the input embedding is preserved through layer composition, meaning the frequency content of *any* hidden neuron remains determined by the first-layer frequencies $\boldsymbol{\omega}$. We formalize this using the amplitude-phase expansion of SIRENs [24, 49].

[5] In SIREN, the first-layer $\mathbf{W}^0$ is initialized in $[-1, 1]$ and scaled by a frequency hyperparameter ω_0; here we absorb this scaling into $\mathbf{W}^0$ for notational simplicity.

Theorem 1. *Every hidden neuron of a SIREN admits an expansion as a sum of sinusoidal terms of the form*

$$\alpha \ \sin\Big(\langle \mathbf{k}, \boldsymbol{\omega}_{\mathbf{x}} \rangle \mathbf{x} + \langle \mathbf{k}, \boldsymbol{\omega}_t \rangle t + \lambda\Big), \tag{3}$$

where $\mathbf{k}$ *is an integer vector, the phase shift* λ *depends only on the biases, and the amplitudes* α *depend only on the hidden weights of the network.*

The proof is provided in the supplementary material. Consequently, Theorem 1 implies that the spatial and temporal frequencies appearing in any layer are *integer combinations* of the first-layer frequencies $\boldsymbol{\omega}_{\mathbf{x}}$ and $\boldsymbol{\omega}_t$, and no new frequencies outside their span are introduced by network depth. Thus, controlling $\boldsymbol{\omega}_t$ in the input layer effectively controls the temporal frequency content of the entire network, even after multiple compositions. This motivates our temporal band-limiting strategy.

Temporal frequency control. Let Δt denote the typical spacing between annotated timesteps (with time normalized to $[0, 1]$). To stabilize interpolation in unobserved intervals, we bound the temporal frequencies in $\boldsymbol{\omega}_t$ using a Nyquist threshold, i.e., $\omega_t^{\max} \propto \pi/\Delta t$. In practice, we initialize $\boldsymbol{\omega}_t$ within $[-\omega_t^{\max}, \omega_t^{\max}]$ and clamp it throughout training, while initializing $\boldsymbol{\omega}_{\mathbf{x}}$ with a standard high-capacity spatial range. Concretely, we use the modified input embedding:

$$\sin\Big(\boldsymbol{\omega}_{\mathbf{x}}\, \mathbf{x} + \mathrm{clamp}\left(\boldsymbol{\omega}_t, [-\omega_t^{\max}, \omega_t^{\max}]\right) t + \boldsymbol{\varphi}\Big), \qquad \omega_t^{\max} \propto \frac{\pi}{\Delta t}. \tag{4}$$

By Theorem 1, this explicit temporal band-limiting propagates through the network, attenuating high-frequency temporal oscillations and yielding smooth surface evolution without sacrificing spatial detail.

3.3 Loss Functional

To fit the zero-level sets of f_θ to the input collection of temporal 2D slices $\mathcal{C}_i$, we optimize θ by minimizing a sum of terms that (i) enforce data consistency on the observed cross-sections while resolving the SDF sign ambiguity via in-plane normals, and (ii) regularize the field in space–time to promote SDF validity and suppress spurious zero-crossings in weakly constrained regions.

Regularization term. Precisely, we minimize:

$$\mathcal{L}(\theta) = \sum_i \Big\{ \mathbb{E}_{\mathbf{x} \in \mathcal{C}_i} \Big[\big| f_\theta(\mathbf{x}, t_i) \big| + \lambda_n \mathscr{N}(\theta) \Big] + \lambda_d \mathscr{D}(\theta) \Big\} + \lambda_r \mathscr{R}(\theta), \tag{5}$$

where the first term enforces that the zero-level set matches the input contours. The second term enforces *orientation consistency*: at each contour point, the predicted spatial gradient $\nabla_{\mathbf{x}} f_\theta$ should align with the corresponding *in-plane 2D normal* vector $\mathbf{n}_i(\mathbf{x})$ i without forcing the field to exactly match the 2D SDF elsewhere.

$$\mathscr{N}(\theta) = \mathrm{ReLU}\big(\langle -\nabla_{\mathbf{x}} f_\theta(\mathbf{x}, t_i), \mathbf{n}_i(\mathbf{x}) \rangle\big) \tag{6}$$

This penalizes normal inversions that would otherwise flip the SDF sign locally. The third term $\mathscr{D}$ encourages f_θ to have values close to the corresponding *in-plane 2D SDF* $\mathbf{f}_i(\mathbf{x})$ at off-surface samples:

$$\mathscr{D}(\theta) = \mathbb{E}_{\mathbf{x}\in\Pi(\mathcal{C}_i)}(f_\theta(\mathbf{x}, t_i) - \mathbf{f}_i(\mathbf{x}))^2, \tag{7}$$

$\mathscr{N} = \mathrm{ReLU}(\cos(-\phi))$

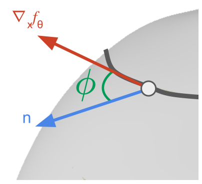

Fig. 3: Orientation consistency constraint.

where $\Pi(\mathcal{C}_i)$ is the plane that contains $\mathcal{C}_i$. Importantly. λ_d is set to be small to encourage the correct sign but allow deviations from 2D SDF labels. The regularizer $\mathscr{R}$ is evaluated on a set of space–time samples and combines two complementary constraints: (i) an Eikonal penalty that encourages $f_\theta(\cdot, t)$ to behave as a valid SDF, and (ii) a minimum-surface penalty that discourages small SDF magnitudes away from observed slices, reducing accidental zero-crossings. We define:

$$\mathscr{R}(\theta) = \mathbb{E}_{(\mathbf{x},t)}\Big[\big(\|\nabla_{\mathbf{x}} f_\theta(\mathbf{x}, t)\|_2 - 1\big)^2 + \exp\big(-\alpha\,|f_\theta(\mathbf{x}, t)|\big)\Big], \tag{8}$$

where $\alpha > 0$ controls the strength of the spurious-surface suppression. This combined regularizer stabilizes training and improves temporal interpolation by enforcing proper SDF structure while pushing the field away from zero in weakly constrained regions.

3.4 Multi-Class Model

To represent K anatomical structures simultaneously, we extend the output of the network to produce K signed distance values. The final affine layer becomes $\mathbf{W}_{d+1} \in \mathbb{R}^{K\times N_d}$ and $\mathbf{b}_{d+1} \in \mathbb{R}^K$, yielding a vector-valued function:

$$f_\theta(\mathbf{x}, t) = \big(f_\theta^1(\mathbf{x}, t), \ldots, f_\theta^K(\mathbf{x}, t)\big) \in \mathbb{R}^K, \tag{9}$$

where each component f_θ^k approximates the SDF of the k-th structure. The hidden layers are shared across all classes, allowing the network to learn a common spatiotemporal feature representation from which per-class SDFs are decoded.

Multi-Class Loss Function. Each class k is supervised independently based on its corresponding slice annotations. Given class-specific contours, the total loss is formulated as the sum of the individual class losses:

$$\mathcal{L}_{\text{total}}(\theta) = \sum_{k=1}^{K} \mathcal{L}^k(\theta), \tag{10}$$

where each term $\mathcal{L}^k$ is defined as in Eq. (5), applied exclusively to the k-th output component f_θ^k. The regularization term $\mathscr{R}(\theta)$ acts on all components to jointly enforce SDF validity and suppress spurious zero-crossings across the entire multi-class field.

Non-Overlapping Segmentation Layer. The per-class SDFs are not explicitly constrained to produce disjoint regions during training. This allows our method to fit to potentially overlapping GT segmentations present in real medical imaging data. Nonetheless, one may wish to obtain a more physically realistic non-overlapping segmentation at inference time. To do this, we assign each spatial point to the class with the smallest signed distance value. Introducing a background class with a constant $f_\theta^0 \equiv 0$, the label map is defined as:

$$\ell(\mathbf{x}, t) = \underset{k \in \{0,1,\dots,K\}}{\operatorname{argmin}} f_\theta^k(\mathbf{x}, t). \tag{11}$$

A point is assigned to structure k when f_θ^k is the most negative among all classes. The background class ensures that regions where all SDFs are positive (i.e., exterior to all target structures) are left unlabeled, seamlessly bounding the segmentation to the interior of the union of the zero-level sets. The boundaries between adjacent regions in the argmin partition naturally form surfaces where two classes tie for the minimum. Around these boundaries, the function $u(\mathbf{x}, t) = \min_k f_\theta^k(\mathbf{x}, t)$ behaves as an unsigned distance field. We provide examples of non-overlapping surfaces extracted with this method in the supplementary.

4 Experiments

We evaluate DynaSlice on single and multi-class dynamic SDF reconstruction from sparse cross-sections. We compare against adapted static INR baselines and recent cross-sectional reconstruction methods (Orex [38] and CrossSDF [43]) on both synthetic sequences and real medical data, as well as recent 4D point-cloud reconstruction method CanFields [44].

Datasets. We evaluate on (i) a synthetic benchmark of 5 dynamic single-object sequences with ground-truth 4D geometry, and (ii) 6 real multi-organ medical scans with annotated dynamic cross-sections from The Cancer Imaging Archive [14]. In total, we test on 40 separate dynamic sequences.

For synthetic sequences, each sequence consists of a ground-truth mesh evolving over 37 uniformly spaced timesteps. We subsample temporally by selecting every fourth frame to collect 10 training timesteps. We reflect the most common non-parallel segmentation setting in medical imaging [4, 11, 27, 35] and at each timestep, we extract cross-sections using (i) rotated planes via uniform rotation around the z-axis and (ii) parallel planes with uniform spacing along the z-axis (see Fig. 2 input). We use 15 rotated and 15 parallel slices per timestep for Lungs, Heart, and Diaphragm, and increase this to 35/35 for thinner structures (Vein 1 and Vein 2).

For the real medical data, each patient contains a variable number of annotated organs (5–8; see supplementary). We resample volumetric segmentations onto 20 rotated planes around the z-axis and extract parallel slices by taking every second slice along the z-axis (total count varies per organ).

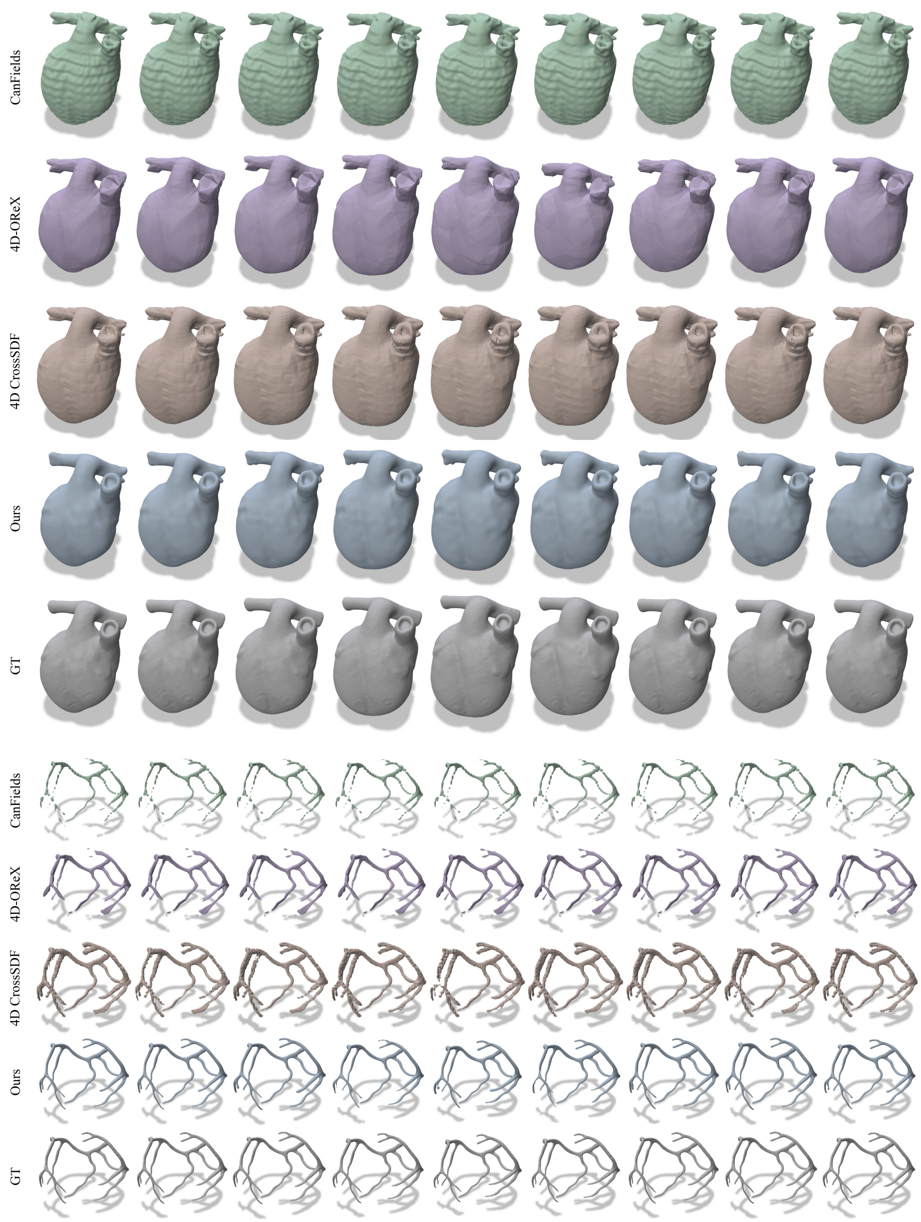

Fig. 4: Qualitative comparison on synthetic sequences. DynaSlice achieves superior reconstructions of both thick (Heart, top) and thin (Vein1, bottom) structures. Please see the supplementary material for more examples and video which better illustrate motions

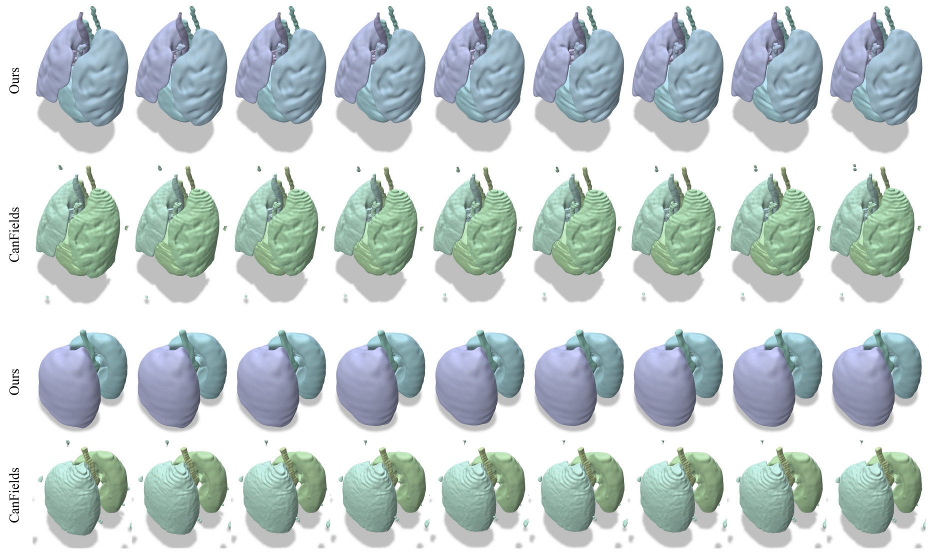

Fig. 5: Qualitative multi-structure reconstruction on real medical data. Compared to CanFields [44], DynaSlice reconstructs laddering-free, evolving surfaces with no spurious floaters, here demonstrated on both Patients 101 (top) and 110 (bottom).

Baselines We compare against recent SOTA 4D point-cloud reconstruction approach [44] to illustrate failures and artifacts resulting from point-cloud approaches.

While previous SOTA cross-sectional reconstruction works such as CrossSDF [43] and OReX [38] focus on static single-class structures, we target the more general case of time-dependent multi-class reconstruction. In this dynamic setting, a baseline of linearly interpolating the static reconstructions across time is impractical, as it requires training separate models for each class and timestep. Consequently, even a modest dataset (e.g., 10 classes over 10 timesteps) would require roughly 20 days of training on a single V100 GPU, making such an approaches computationally infeasible.

To address this, we test against a naïve temporal extension of OReX and CrossSDF. Both networks are trivially extended to receive time-coordinates as input, which are positionally encoded with the different methods respective strategies. Please see the supplementary for implementation details.

In addition to this, we compare to "4D SIREN", a naïve version of our approach without our temporal frequency control and orientation consistency loss.

Evaluation protocol. For synthetic sequences, we evaluate reconstruction quality against the ground-truth 4D meshes. Specifically, we compute Chamfer Distance (CD $\downarrow$) and Intersection over Union (IoU $\uparrow$) at all 37 ground-truth timesteps, and report results averaged over time.

For real sequences, ground-truth 3D meshes are not available. Instead, we adopt a slice-based evaluation by withholding a subset of test cross-sections across both temporal and spatial dimensions. We report 2D IoU on held-out ground truth slices under three settings:

1. **Spatial**: seen timesteps with novel 2D slice positions,
2. **Temporal**: novel timesteps with slice positions (z-values) seen during training,
3. **Spatio-temporal**: both novel timesteps and novel slice positions.

This protocol enables disentangling spatial and temporal interpolation performance.

Implementation details. DynaSlice is implemented in PyTorch [33] and trained with Adam [16]. Our networks uses $\approx 200K$ parameters and take ≈ 30 minutes to train on a single class. For training hyper-parameters and more implementation details, please see the supplementary material.

Table 1: Quantitative results for single-class reconstruction (synthetic). Comparison of reconstruction accuracy using Chamfer Distance (CD $\times 10^4$ $\downarrow$) and Intersection over Union (IoU $\uparrow$) across six anatomical structures. Our method, DynaSlice, significantly outperforms traditional 4D INR baselines in geometric fidelity while maintaining a training efficiency comparable to vanilla SIREN and orders of magnitude faster than optimization-based methods.

Method	Lungs		Diaphragm		Heart		Vein 1		Vein 2		Mean		**Time**
	CD $\downarrow$	IoU $\uparrow$	CD $\downarrow$	IoU $\uparrow$	CD $\downarrow$	IoU $\uparrow$	CD $\downarrow$	IoU $\uparrow$	CD $\downarrow$	IoU $\uparrow$	CD $\downarrow$	IoU $\uparrow$	Min. $\downarrow$
4D OReX	2.34	0.722	0.732	0.930	3.56	0.936	7.69	0.605	19.3	0.379	6.73	0.714	243
4D CrossSDF	10.9	0.279	0.978	0.928	1.82	**0.971**	5.84	0.596	7.58	0.401	5.43	0.635	319
CanFields	14.6	0.157	1.96	0.853	4.78	0.943	6.01	0.459	3.21	0.522	6.11	0.587	93
4D SIREN	6.1	0.417	2.57	0.784	12.0	0.830	2.48	0.691	2.29	0.6955	5.09	0.6836	31
DynaSlice	**1.58**	**0.809**	**0.633**	**0.941**	**1.34**	0.966	**1.06**	**0.749**	**1.28**	**0.750**	**1.19**	**0.843**	**32**

Table 2: Real multi-class (MC) dynamic reconstruction. Mean performance over all patients. We report 2D Intersection over Union (IoU $\uparrow$) for spatial, temporal, and joint spatio-temporal consistency. We report average training time per-class in minutes (m). DynaSlice achieves the best performance across all metrics.

Method	**Spatial**	**Temporal**	**Spatio-Temporal**	**Time (m)**
CanFields	0.853	0.849	0.851	90
Naïve MC (indep. SDFs)	0.874	0.867	0.861	34
Shared MC (no TC)	0.878	0.896	0.869	17
DynaSlice (Ours)	**0.936**	**0.923**	**0.921**	**17**

4.1 Quantitative and Qualitative Comparisons

In both real and synthetic settings, DynaSlice produces substantially higher quality reconstructions than competing approaches. As shown in Fig. 4 and Fig. 5, the reconstructed surfaces exhibit smooth and coherent temporal motion across the sequence. This holds for both thick anatomical structures, such as the *Heart*, and thin structures such as *Vein 1*. The 4D adaptations of both OReX and CrossSDF generate reasonable reconstructions, but their characteristic visual artifacts remain. For OReX exploding gradients the of neural indicator function on surface boundaries creates artifacts in mesh-extraction. For CrossSDF, hash-grid interpolation artifacts are present. Point-cloud-based reconstruction method CanFields produce visible laddering artifacts caused by the sparsity between cross-sectional observations. These artifacts can interfere with medical diagnosis. Our naïve baseline 4D SIREN frequently exhibit unstable optimization and struggle to capture thin or detailed geometry. These results highlight the importance of explicitly controlling input frequencies in the implicit representation, which enables DynaSlice to achieve stable optimization and improved reconstruction fidelity across diverse dynamic surface reconstruction tasks, generating smooth surfaces and stable temporal evolution without slicing artifacts on both real and synthetic datasets.

Fig. 6: Qualitative Ablation We visualize the *Diaphragm* reconstruction with naïve baseline (4D SIREN), and then add our *orientation consistency loss* (Ori.), then full DynaSlice methodology with *temporal frequency control*.

Ablation In Fig. 6 and Tab. 3 we present an ablation of our temporal frequency control and orientation consistency loss components. Even on the simplest example *Diaphragm*, a vanilla 4D SIREN network will produce holes between cross-sectional planes and suffer from unstable optimization. Adding orientation consistency loss and temporal frequency control allows the model to learn both thick structures, demonstrated both qualitatively and quantitatively.

Table 3: Ablation study on synthetic data. Impact of temporal frequency control (TC) and orientation consistency (Ori). Chamfer Distance (CD $\times 10^4$ $\downarrow$) and IoU ($\uparrow$) per structure and mean. Higher IoU is better, lower CD is better.

Setting	Comp.		Structures											
	TC	Ori	Lungs		Diaphragm		Heart		Vein 1		Vein 2		Mean	
			CD	IoU	CD	IoU	CD	IoU	CD	IoU	CD	IoU	CD	IoU
4D SIREN			6.1	0.417	2.57	0.784	12.0	0.830	2.48	0.691	2.29	0.696	5.09	0.684
+ Ori		✓	3.27	0.529	1.04	0.909	2.07	0.964	1.44	0.715	1.56	0.716	1.88	0.767
DynaSlice	✓	✓	**1.58**	**0.809**	**0.633**	**0.941**	**1.34**	**0.966**	**1.06**	**0.749**	**1.28**	**0.750**	**1.19**	**0.843**

5 Conclusion and Limitations

We introduced DynaSlice, a unified multichannel 4D neural SDF framework for dynamic surface reconstruction from sparse temporal cross-sections. By jointly modeling space and time with dedicated temporal frequency control, our approach significantly improves the stability and fidelity of reconstructed surfaces, while the shared multichannel formulation enables efficient reconstruction of multiple anatomical structures within a single model. More broadly, our work addresses a largely unexplored gap in cross-sectional reconstruction: recovering multiple temporally evolving geometries directly from sparse slices, a setting that is often computationally intractable for existing methods that rely on dense volumetric representations or per-frame reconstruction pipelines. Experiments on synthetic and real medical datasets demonstrate that DynaSlice outperforms adapted static baselines and produces stable, accurate surface evolution under sparse supervision.

Despite these advances, several limitations remain. Our approach is best suited to smooth, organically deforming structures, which align well with the continuous function representation of SIREN-based implicit networks. In contrast, industrial CT data often contains sharp edges and piecewise rigid structures that may require alternative representations or specialized modeling strategies. We also observe occasional topological flickering in very thin structures, highlighting the ongoing challenge of maintaining consistent topology under large temporal changes. Addressing these issues remains an important direction for future work. Nevertheless, we hope that this formulation opens new avenues for implicit modeling of dynamic anatomy and temporally evolving geometry from limited cross-sectional observations.

References

1. Bajaj, C.L., Coyle, E.J., Lin, K.N.: Arbitrary topology shape reconstruction from planar cross sections. Graphical Models and Image Processing **58**(6), 524–543 (1996)
2. Barequet, G., Sharir, M.: Piecewise-linear interpolation between polygonal slices. In: Annual Symposium on Computational Geometry (1994)
3. Bermano, A., Vaxman, A., Gotsman, C.: Online reconstruction of 3d objects from arbitrary cross-sections. ACM Transactions on Graphics (2011)
4. Bloomer, T.N., Plein, S., Radjenovic, A., Higgins, D.M., Jones, T.R., Ridgway, J.P., Sivananthan, M.U.: Cine mri using steady state free precession in the radial long axis orientation is a fast accurate method for obtaining volumetric data of the left ventricle. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine **14**(6), 685–692 (2001)
5. Boissonnat, J.D., Memari, P.: Shape reconstruction from unorganized cross-sections. In: Symposium on Geometry Processing (2007)
6. Boulch, A., Marlet, R.: POCO: Point convolution for surface reconstruction. In: CVPR (2022)
7. Castellan, C., Dastarac, D.: Use of computed tomography slices 3d-reconstruction as a powerful tool to improve manufacturing processes on aeroengine components. In: AIP Conference Proceedings. American Institute of Physics (2000)
8. Chen, L., Li, J., Huang, Q.: Guaranteed simply connected mesh reconstruction from an unorganized point cloud. In: The Fourteenth International Conference on Learning Representations
9. Chen, R.T.Q., Rubanova, Y., Bettencourt, J., Duvenaud, D.: Neural ordinary differential equations. In: Advances in Neural Information Processing Systems (NeurIPS). vol. 31 (2018)
10. Dou, M., Khamis, S., Degtyarev, Y., Davidson, P., Fanello, S.R., Kowdle, A., Escolano, S.O., Rhemann, C., Kim, D., Taylor, J., Kohli, P., Tankovich, V., Izadi, S.: Fusion4D: Real-time performance capture of challenging scenes. ACM Transactions on Graphics (TOG) pp. 114:1–114:13 (2016)
11. Fenster, A., Parraga, G., Bax, J.: Three-dimensional ultrasound scanning. Interface focus **1**(4), 503–519 (2011)
12. Geiger, B.: Three-dimensional modeling of human organs and its application to diagnosis and surgical planning. Ph.D. thesis, INRIA (1993)
13. Gropp, A., Yariv, L., Haim, N., Atzmon, M., Lipman, Y.: Implicit geometric regularization for learning shapes. In: International Conference on Machine Learning. pp. 3789–3799. PMLR (2020)
14. Hugo, G.D., Weiss, E., Sleeman, W.C., Balik, S., Keall, P.J., Lu, J., Williamson, J.F.: Data from 4d lung imaging of nsclc patients. (No Title) (2016)
15. Jacobs, S., Grunert, R., Mohr, F.W., Falk, V.: 3d-imaging of cardiac structures using 3d heart models for planning in heart surgery: a preliminary study. Interactive cardiovascular and thoracic surgery (2008)
16. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization (2017), `https://arxiv.org/abs/1412.6980`
17. Lei, J., Daniilidis, K.: CaDeX: Learning canonical deformation coordinate space for dynamic surface representation via neural homeomorphism. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 6624–6634 (2022)

18. Li, Z., Müller, T., Evans, A., Taylor, R.H., Unberath, M., Liu, M.Y., Lin, C.H.: Neuralangelo: High-fidelity neural surface reconstruction. In: CVPR. pp. 8456–8465 (2023)
19. Liu, L., Bajaj, C., Deasy, J.O., Low, D.A., Ju, T.: Surface reconstruction from non-parallel curve networks. In: Computer Graphics Forum (2008)
20. Mescheder, L., Oechsle, M., Niemeyer, M., Nowozin, S., Geiger, A.: Occupancy networks: Learning 3D reconstruction in function space. In: CVPR (2019)
21. Müller, T., Evans, A., Schied, C., Keller, A.: Instant neural graphics primitives with a multiresolution hash encoding. ACM Transactions on Graphics **41**(4) (2022)
22. Newcombe, R.A., Fox, D., Seitz, S.M.: DynamicFusion: Reconstruction and tracking of non-rigid scenes in real-time. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 343–352 (2015)
23. Niemeyer, M., Mescheder, L., Oechsle, M., Geiger, A.: Occupancy flow: 4D reconstruction by learning particle dynamics. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 5379–5389 (2019)
24. Novello, T., Aldana, D., Araujo, A., Velho, L.: Tuning the frequencies: Robust training for sinusoidal neural networks. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 3071–3080 (2025)
25. Novello, T., Schardong, G., Schirmer, L., Da Silva, V., Lopes, H., Velho, L.: Exploring differential geometry in neural implicits. Computers & Graphics **108**, 49–60 (2022)
26. Novello, T., da Silva, V., Schardong, G., Schirmer, L., Lopes, H., Velho, L.: Neural implicit surface evolution. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 14279–14289 (2023)
27. Orlando, N., Gillies, D.J., Gyacskov, I., Romagnoli, C., D'Souza, D., Fenster, A.: Automatic prostate segmentation using deep learning on clinically diverse 3d transrectal ultrasound images. Medical physics **47**(6), 2413–2426 (2020)
28. Osher, S., Fedkiw, R., Piechor, K.: Level set methods and dynamic implicit surfaces. Appl. Mech. Rev. **57**(3), B15–B15 (2004)
29. Osher, S., Sethian, J.A.: Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations. Journal of computational physics **79**(1), 12–49 (1988)
30. Park, J.J., Florence, P., Straub, J., Newcombe, R., Lovegrove, S.: DeepSDF: Learning continuous signed distance functions for shape representation. In: CVPR (2019)
31. Park, K., Sinha, U., Barron, J.T., Bouaziz, S., Goldman, D.B., Seitz, S.M., Martin-Brualla, R.: Nerfies: Deformable neural radiance fields. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 5865–5874 (2021)
32. Park, K., Sinha, U., Hedman, P., Barron, J.T., Bouaziz, S., Goldman, D.B., Martin-Brualla, R., Seitz, S.M.: HyperNeRF: A higher-dimensional representation for topologically varying neural radiance fields. ACM Transactions on Graphics (TOG) (2021)
33. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S.: Pytorch: An imperative style, high-performance deep learning library (2019), https://arxiv.org/abs/1912.01703
34. Pumarola, A., Corona, E., Pons-Moll, G., Moreno-Noguer, F.: D-NeRF: Neural radiance fields for dynamic scenes. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 10313–10322 (2021)

35. Qiu, W., Yuan, J., Ukwatta, E., Sun, Y., Rajchl, M., Fenster, A.: Prostate segmentation: an efficient convex optimization approach with axial symmetry using 3-d trus and mr images. IEEE transactions on medical imaging **33**(4), 947–960 (2014)
36. Sang, L., Canfes, Z., Cao, D., Bernard, F., Cremers, D.: Implicit neural surface deformation with explicit velocity fields. arXiv preprint arXiv:2501.14038 (2025)
37. Sang, L., Canfes, Z., Cao, D., Marin, R., Bernard, F., Cremers, D.: 4deform: Neural surface deformation for robust shape interpolation. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 6542–6551 (2025)
38. Sawdayee, H., Vaxman, A., Bermano, A.H.: Orex: Object reconstruction from planar cross-sections using neural fields. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 20854–20862 (2023)
39. Schirmer, L., Novello, T., da Silva, V., Schardong, G., Perazzo, D., Lopes, H., Gonçalves, N., Velho, L.: Geometric implicit neural representations for signed distance functions. Computers & Graphics **125**, 104085 (2024)
40. Sitzmann, V., Martel, J., Bergman, A., Lindell, D., Wetzstein, G.: Implicit neural representations with periodic activation functions. Advances in neural information processing systems **33**, 7462–7473 (2020)
41. Tancik, M., Srinivasan, P.P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J.T., Ng, R.: Fourier features let networks learn high frequency functions in low dimensional domains. In: Advances in Neural Information Processing Systems (NeurIPS). vol. 33 (2020)
42. Tretschk, E., Tewari, A., Golyanik, V., Zollhöfer, M., Lassner, C., Theobalt, C.: Non-rigid neural radiance fields: Reconstruction and novel view synthesis of a dynamic scene from monocular video. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 12959–12970 (2021)
43. Walker, T., Esposito, S., Rebain, D., Vaxman, A., Onken, A., Li, C., Mac Aodha, O.: Crosssdf: 3d reconstruction of thin structures from cross-sections. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 30928–30937 (2025)
44. Wang, M., Li, C., Vaxman, A.: CanFields: Consolidating diffeomorphic flows for non-rigid 4D interpolation from arbitrary-length sequences. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 28587–28598 (2025)
45. Wang, Y., Han, Q., Habermann, M., Daniilidis, K., Theobalt, C., Liu, L.: Neus2: Fast learning of neural implicit surfaces for multi-view reconstruction. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 3295–3306 (2023)
46. Xie, Y., Takikawa, T., Saito, S., Litany, O., Yan, S., Khan, N., Tombari, F., Tompkin, J., Sitzmann, V., Sridhar, S.: Neural fields in visual computing and beyond. In: Computer graphics forum. vol. 41, pp. 641–676. Wiley Online Library (2022)
47. Yang, G., Belongie, S., Hariharan, B., Koltun, V.: Geometry processing with neural fields. Advances in Neural Information Processing Systems **34**, 22483–22497 (2021)
48. Yuan, X., Liu, C., Wang, Y.: 4d myocardium reconstruction with decoupled motion and shape model. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (2023)
49. Yüce, G., Ortiz-Jiménez, G., Besbinar, B., Frossard, P.: A structured dictionary perspective on implicit neural representations. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 19228–19238 (2022)