When Pulling Fails: Understanding and Alleviating SDF Collapse in Sparse Freehand Ultrasound Reconstruction

Jiuan Chen[1] , **Song Lai**[1,2] , **Mingyang Zhao**[3,4] and **Gaofeng Meng**[1,4,5]
[1]Centre for Artificial Intelligence and Robotics,
Hong Kong Institute of Science & Innovation, Chinese Academy of Sciences
[2]Department of Computer Science, City University of Hong Kong
[3]State Key Laboratory of Mathematical Sciences,
Academy of Mathematics and Systems Science, Chinese Academy of Sciences
[4]University of Chinese Academy of Sciences
[5]Institute of Automation, Chinese Academy of Sciences
jiuanchen1119@gmail.com, songlai2-c@my.cityu.edu.hk,
zhaomingyang@amss.ac.cn, gaofeng.meng@cair-cas.org.hk

Abstract

Despite being a cost-effective modality for volumetric imaging, freehand three-dimensional (3D) ultrasound produces inherently sparse data due to the significant elevational gaps left by tracked 2D sweeps. This sparsity poses a unique challenge for Implicit Neural Representations (INRs). While successful in other domains, INRs applied here tend to fail as the learned signed distance fields (SDF) collapse toward zero inside the object, leading to the loss of concavities and the incorrect closure of anatomical gaps. Our analysis identifies the root cause as a statistical bias in gradient-based sampling objectives, showing that symmetric volumetric sampling mathematically drives the expected SDF value to zero. To rectify this, we present a geometry-aware framework that explicitly anchors the non-negative half-space. Our method utilizes a boundary-directed exterior sampling strategy to ensure non-negative constraints in empty areas, complemented by an ellipsoid-based adversarial mechanism to regularize the global field distribution. Experiments on multiple anatomical datasets demonstrate that our approach mitigates field collapse and improves geometric fidelity and topological consistency on most metrics. Code is available at https://github.com/jiuanchen/Pulling-Fails-SDF.

1 Introduction

Ultrasonography (US) remains widely used due to its safety, portability, and real-time capabilities [Fiorentino *et al.*, 2023]. However, standard 2D US is intrinsically slice-based and operator-dependent. To obtain comprehensive 3D geometric characterization, clinicians typically rely on freehand sweeps, where tracked 2D B-scans are reconstructed into a volume [Prager *et al.*, 2010; Huang and Zeng, 2017]. Unlike dense laser scans, these acquisitions are inherently sparse and non-uniform, often containing significant information gaps between slices due to varying sweep speeds and visibility constraints [Adriaans *et al.*, 2024]. These sparse slice settings pose severe challenges for surface reconstruction algorithms, which must bridge inter-slice regions without over-smoothing genuine anatomical details.

Recent advances in Implicit Neural Representations (INRs) have shifted the paradigm from discrete voxelization to continuous function learning [Molaei *et al.*, 2023]. Specifically, methods leveraging the property that the gradient of a signed distance function (SDF) points towards the nearest surface—often termed pulling-based learning—have shown great promise [Ma *et al.*, 2021]. In the ultrasound domain, recent works like FUNSR [Chen *et al.*, 2024b] and RoCoSDF [Chen *et al.*, 2024a] have successfully adapted this concept to handle noisy, unorganized US point clouds. Ideally, these methods optimize the network by utilizing the gradient to project query points onto the zero level-set.

However, existing formulations implicitly assume dense surface sampling or sufficient boundary coverage. When applied to the sparse volumetric point clouds derived from segmented ultrasound masks, these methods exhibit a critical failure mode: catastrophic over-smoothing and topological degradation. As illustrated in Figure 1, standard approaches fail to preserve anatomical concavities and often collapse the internal field, resulting in unstable optimization and inaccurate geometry.

In this paper, we revisit the theoretical foundations of gradient-based SDF learning to understand this limitation. We demonstrate that the failure is not merely a consequence of data sparsity but stems from a mathematical bias in the training objective. When query points are sampled symmetrically around volumetric data rather than strictly near a surface manifold, the optimization objective minimizes the residual variance by driving the SDF magnitude to zero everywhere, rather than establishing a valid distance field. This *interior contraction* phenomenon causes the implicit field to degener-

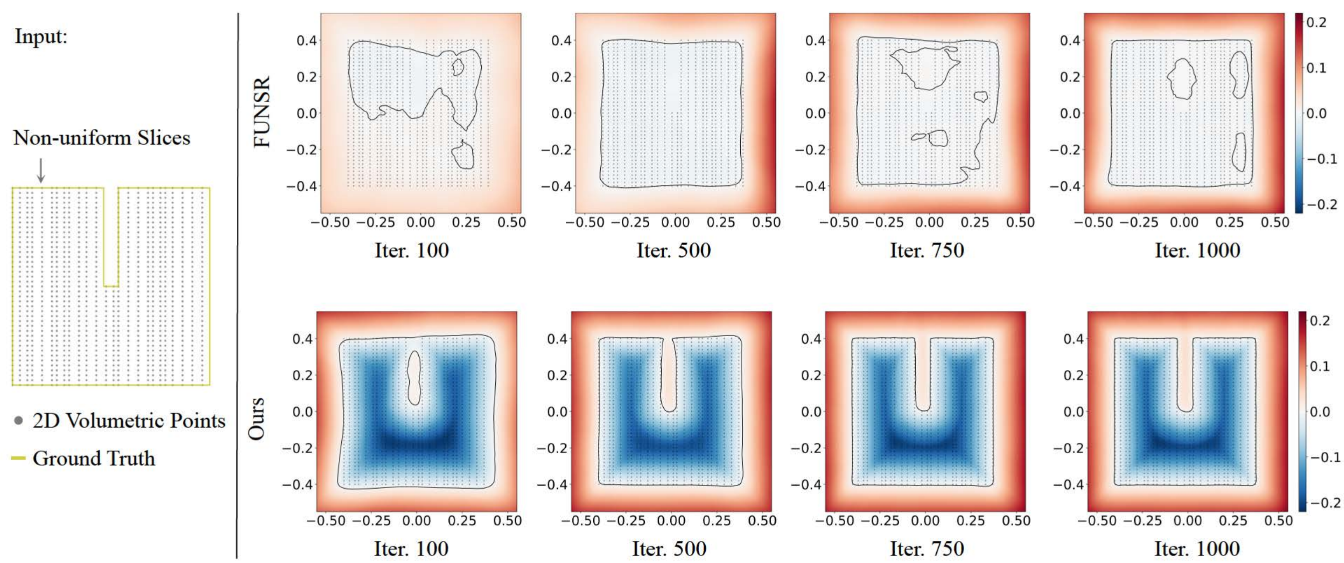


Figure 1: 2D slice-based reconstruction. Left: ground truth and sparse slice-like volumetric points. Right: FUNSR collapses the concavity and shows unstable evolution, whereas our method preserves the recessed region and learns a more structured SDF.

ate, filling essential anatomical voids.

To address this, we propose a robust reconstruction framework that counteracts interior contraction by explicitly modeling the exterior space and the global field distribution. Our approach incorporates a boundary-directed sampling strategy that identifies unoccupied space to enforce non-negative field constraints, effectively anchoring the optimization. Furthermore, we introduce an adversarial regularization scheme based on randomized ellipsoids to approximate valid SDF distributions, ensuring the network learns a geometrically consistent field even in regions devoid of data.

Our main contributions are summarized as follows:

- We reveal that symmetric volumetric sampling in gradient-based SDF learning induces an interior-contraction bias that drives SDF values toward zero.
- We propose a geometry-aware reconstruction framework combining boundary-directed exterior sampling with an ellipsoid-based adversarial regularizer.
- We validate the approach on sparse slice-based anatomical datasets and show improved geometric fidelity and topological consistency.

2 Related Works

In this section, we review prior literature relevant to our approach, focusing on implicit neural representations and their application to ultrasound surface reconstruction.

2.1 Implicit Neural Representations

Implicit neural representations (INRs) have emerged as a powerful paradigm for shape modeling, representing geometry as the zero level-set of a continuous function, typically a SDF, parameterized by a neural network. Early supervised approaches, such as DeepSDF [Park *et al.*, 2019], required watertight meshes for training, limiting their applicability in raw data scenarios. To address this, unsupervised methods were developed to learn directly from point clouds. IGR [Gropp *et al.*, 2020] and DiGS [Ben-Shabat *et al.*, 2022] utilize the Eikonal equation to regularize the gradient norm, yet they often struggle with convergence speed and fine detail preservation.

Alternatively, gradient-based projection methods, exemplified by NeuralPull [Ma *et al.*, 2021], interpret the SDF gradient as a projection direction. By minimizing the distance between a query point and its projection onto the nearest surface point, these methods achieve faster convergence and higher fidelity. However, these formulations implicitly assume that the input point cloud represents a surface manifold. As demonstrated in Section 3, when applied to volumetric point clouds—common in medical segmentation masks—this objective introduces a statistical bias that drives the field to collapse.

2.2 Ultrasound Surface Reconstruction

Traditional freehand 3D ultrasound reconstruction typically involves compounding 2D B-scans into a voxel grid followed by iso-surface extraction algorithms like Marching Cubes [Lorensen and Cline, 1987]. While robust, these discrete methods suffer from resolution quantization and artifacts due to interpolation. Mesh-based approaches, such as Poisson Surface Reconstruction [Kazhdan and Hoppe, 2013], offer continuous surfaces but are highly sensitive to the non-uniform sampling and noise inherent in ultrasound data.

Recently, INRs have been adapted for ultrasound recon-

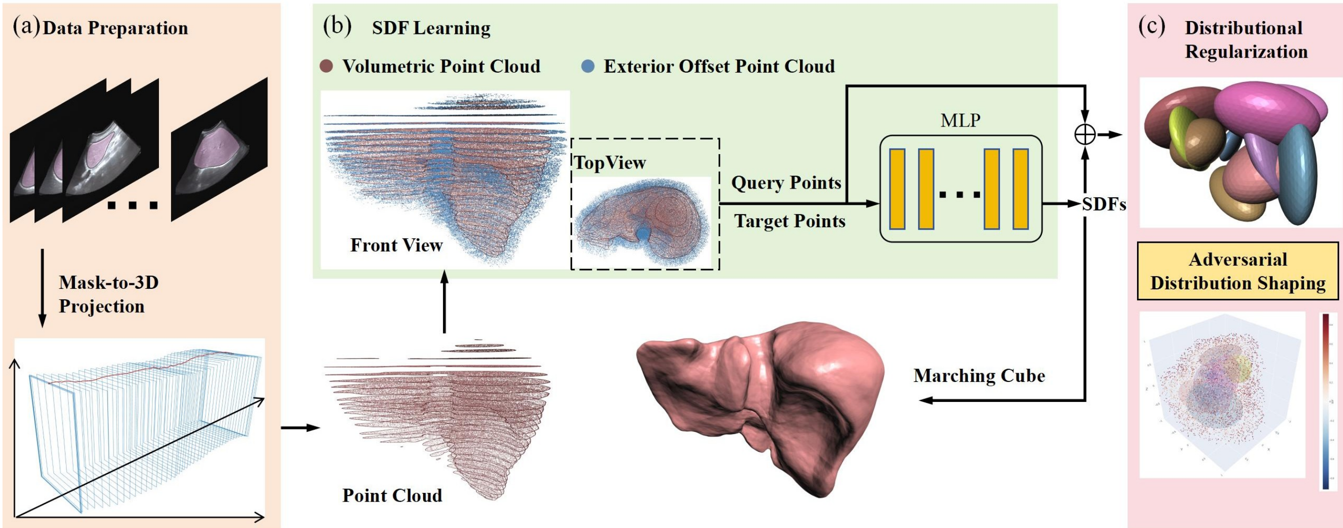


Figure 2: Overview of the proposed framework. (a) Data preparation: sparse 3D ultrasound slices are converted into a volumetric point cloud (red), and boundary-directed sampling generates exterior offset points (blue). (b) SDF learning: front and top views are orthogonal projections of the same point cloud; query points are perturbed samples and target points are their nearest neighbors in $\mathcal{P}$. (c) Distributional regularization: the EA module matches off-surface SDF-value distributions, and the final surface is extracted by Marching Cubes.

struction. FUNSR [Chen *et al.*, 2024b] extended the NeuralPull framework to volumetric ultrasound data by incorporating sign-consistency losses and on-surface adversarial training. RoCoSDF [Chen *et al.*, 2024a] further improved this by leveraging multi-view constraints. Despite these advancements, these methods still rely on the fundamental gradient-projection objective. Consequently, they remain susceptible to the interior contraction pathology when data is sparse or slice-based, leading to topological errors.

3 Mechanism of Field Collapse

To understand the root cause of the reconstruction failures observed in sparse ultrasound scenarios, we analyze one local update of gradient-based SDF optimization. Let $\mathcal{P} = \{\mathbf{p}_i\}$ denote the input point cloud, $\mathbf{x}$ a generic spatial coordinate, $\mathbf{q}$ a sampled query point, and $\mathbf{t} = \mathrm{NN}_{\mathcal{P}}(\mathbf{q})$ its nearest target point in $\mathcal{P}$. The network predicts an SDF $f_\theta : \mathbb{R}^3 \to \mathbb{R}$, and $\mathbf{n} = \nabla f_\theta(\mathbf{q})$ denotes the local gradient. Superscript $*$ denotes the scalar minimizer of the local least-squares subproblem, not the globally optimized network. All equations are written for a single query point, while the training losses are averaged over mini-batches. In segmented medical volumes, $\mathcal{P}$ is often a *volumetric* point cloud filling the mask volume rather than a *surface* point cloud lying only on the boundary. This analysis is therefore local and first-order; it exposes the bias of the projection loss under volumetric sampling, but does not claim a global convergence proof for the full non-convex training dynamics.

3.1 The Pulling Paradigm

Gradient-projection methods, such as NeuralPull [Ma *et al.*, 2021] and FUNSR [Chen *et al.*, 2024b], learn f_θ by minimizing the distance between a query point $\mathbf{q}$ and its first-order projection $\mathbf{q}'$ onto the zero level-set. Following NeuralPull, we use the Newton projection for a scalar level-set:

$$\mathbf{q}' = \mathbf{q} - f_\theta(\mathbf{q})\frac{\nabla f_\theta(\mathbf{q})}{\|\nabla f_\theta(\mathbf{q})\|_2^2}. \tag{1}$$

The squared-norm denominator follows by imposing the linearized constraint $f_\theta(\mathbf{q}) + \nabla f_\theta(\mathbf{q})^\top(\mathbf{q}' - \mathbf{q}) = 0$. When the field is Eikonal-normalized with $\|\nabla f_\theta\|_2 \approx 1$, this update reduces to the normalized-gradient form used in related formulations. The training objective minimizes the L_2 distance between this projected point $\mathbf{q}'$ and the nearest ground truth point $\mathbf{t} \in \mathcal{P}$:

$$\mathcal{L}_{sdf} = \|\mathbf{q}' - \mathbf{t}\|_2^2. \tag{2}$$

Let $\mathbf{d} = \mathbf{q} - \mathbf{t}$ and $\mathbf{n} = \nabla f_\theta(\mathbf{q})$. Substituting Eq. (1) into the residual $\mathbf{r} = \mathbf{q}' - \mathbf{t}$ gives

$$\mathbf{r}(s) = \mathbf{d} - \frac{s}{\|\mathbf{n}\|_2^2}\mathbf{n}, \quad s = f_\theta(\mathbf{q}). \tag{3}$$

Under the local linearization above, $\mathbf{t}$ and $\mathbf{n}$ are treated as fixed when solving for the scalar s. Differentiating $\|\mathbf{r}(s)\|_2^2$ gives $2(s - \mathbf{d}^\top\mathbf{n})/\|\mathbf{n}\|_2^2 = 0$ for $\|\mathbf{n}\|_2 > 0$, hence

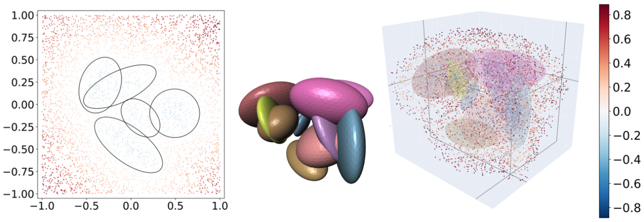
$$f^*(\mathbf{q}) = (\mathbf{q} - \mathbf{t})^\top\mathbf{n}. \tag{4}$$

Thus, the local projection loss fits the signed displacement of $\mathbf{q}$ along the current gradient direction.

3.2 Interior Contraction

The critical issue arises from how the query points $\mathbf{q}$ are sampled relative to the target points $\mathbf{t}$. In standard implementations, queries are generated by adding random Gaussian noise ε to the input points: $\mathbf{q} = \mathbf{t} + \varepsilon$, where $\varepsilon \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$.

When the input $\mathcal{P}$ is a surface point cloud, the nearest neighbor $\mathbf{t}$ is on the manifold, and the displacement ε approximates the distance to the surface. However, when $\mathcal{P}$ is



(a) 2D ellipsoids and SDFs (b) 3D ellipsoids (c) 3D SDFS

Figure 3: Synthetic SDF generation using randomized ellipsoids. By aggregating diverse primitive shapes, we approximate the statistical distribution of valid signed distance fields, providing a robust prior for adversarial learning.

a volumetric point cloud—common in slice-based ultrasound data—the target $\mathbf{t}$ lies inside the object volume.

Substituting $\mathbf{q} = \mathbf{t} + \boldsymbol{\varepsilon}$ into Eq. (4), the target SDF value becomes:

$$f^*(\mathbf{q}) = \boldsymbol{\varepsilon} \cdot \mathbf{n}. \tag{5}$$

Taking the expectation over the symmetric perturbation distribution, conditional on the local target $\mathbf{t}$ and gradient $\mathbf{n}$, gives

$$\mathbb{E}[f^*(\mathbf{q}) \mid \mathbf{t}, \mathbf{n}] = \mathbb{E}[\boldsymbol{\varepsilon}]^\top \mathbf{n} = 0. \tag{6}$$

Gaussianity is not essential here; the first-order bias follows from zero-mean symmetric perturbations. If $\mathbf{n}$ varies with $\mathbf{q}$, Eq. (6) should be read as a local approximation around the current update.

This result indicates that, for volumetric targets away from the boundary, repeated local projection updates are biased toward small SDF magnitudes. Consequently, the zero level-set is not specifically anchored to the anatomical boundary and can drift toward the object interior. We term this local failure mode *interior contraction*.

This explains the instability and over-smoothing observed in Figure 1. The network attempts to satisfy the conflicting requirements of maintaining a non-zero gradient (to define a direction) while minimizing the SDF magnitude everywhere to satisfy Eq. (6). This leads to a collapsed field where anatomical gaps are closed and concavities are filled, as the model defaults to a trivial solution of near-zero values throughout the volume. To reconstruct accurate surfaces from sparse volumetric data, it is therefore necessary to break this symmetry and explicitly anchor the non-negative exterior space.

4 Methodology

To address the interior contraction identified in Section 3, we introduce a geometry-aware reconstruction framework. Our approach anchors the implicit field by explicitly modeling the exterior space and regularizing the global distribution of SDF values. The overall architecture is depicted in Figure 2.

4.1 Boundary-Directed Exterior Sampling (ES)

The primary driver of field collapse is the lack of constraints in the unoccupied space surrounding the sparse volume. To enforce valid positive SDF values, we must identify exterior regions without relying on pre-computed watertight masks, which are unavailable in raw point cloud scenarios. We propose a boundary-directed sampling strategy based on local point distribution statistics.

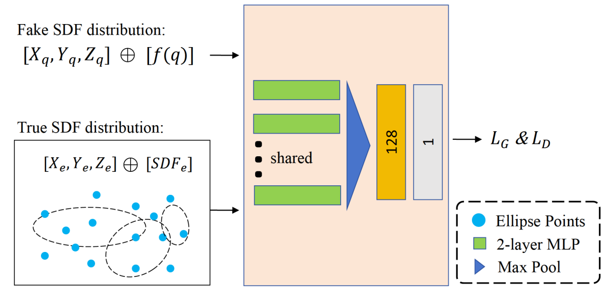


Figure 4: Permutation-invariant EA discriminator. Given query coordinates $\mathcal{Q}$ and their SDF values, it distinguishes the predicted distribution $f_\theta(\mathcal{Q})$ from the analytic ellipsoid distribution $f_E(\mathcal{Q})$, thereby regularizing off-surface field statistics.

For a point $\mathbf{p}$ in the volumetric cloud $\mathcal{P}$, we compute the local neighborhood imbalance vector $\mathbf{g}(\mathbf{p})$:

$$\mathbf{g}(\mathbf{p}) = \frac{1}{k} \sum_{\mathbf{x} \in \mathcal{N}_k(\mathbf{p})} (\mathbf{x} - \mathbf{p}), \tag{7}$$

where $\mathcal{N}_k(\mathbf{p}) \subset \mathcal{P}$ denotes the k-nearest neighbors of $\mathbf{p}$. The magnitude $\|\mathbf{g}(\mathbf{p})\|_2$ serves as a boundary saliency score: interior points have nearly isotropic neighborhoods and small imbalance, while boundary points have larger imbalance vectors pointing inward.

We identify boundary candidates whose saliency exceeds a density-adaptive threshold τ. From these candidates, we project outward along $\hat{\mathbf{n}}_b = -\mathbf{g}(\mathbf{p})/\|\mathbf{g}(\mathbf{p})\|_2$ to generate a set of exterior points $\mathcal{E}$. The offset distance is limited by the local neighbor spacing, so $\mathcal{E}$ remains in unoccupied space rather than crossing back into $\mathcal{P}$.

4.2 Non-Negative Field Constraints

With the exterior set $\mathcal{E}$ established, we introduce a constraint to counteract the statistical tendency of the SDF to drift toward zero. Since $\mathcal{E}$ lies predominantly in the exterior space of the anatomical structure, the predicted SDF values $f_\theta(\mathbf{e})$ for all $\mathbf{e} \in \mathcal{E}$ must be positive.

We enforce this via a one-sided penalty function. For a batch of exterior queries $\mathcal{B}_{ext} \subset \mathcal{E}$:

$$\mathcal{L}_{ext} = \frac{1}{|\mathcal{B}_{ext}|} \sum_{\mathbf{e} \in \mathcal{B}_{ext}} \text{Softplus}(-f_\theta(\mathbf{e})). \tag{8}$$

This loss penalizes negative predictions in the exterior region, effectively anchoring the zero level-set to the boundary between $\mathcal{P}$ and $\mathcal{E}$ and preventing the interior contraction described in Eq. (6).

4.3 Distributional Regularization via Ellipsoid Adversarial (EA)

While $\mathcal{L}_{ext}$ constrains the sign of the field, it does not ensure that the SDF gradients and magnitudes are geometrically plausible throughout the volume. Standard adversarial

Dataset	Method	$\mathrm{CD_{L1}} \downarrow$	$\mathrm{CD_{L2}} \downarrow$	HD95 $\downarrow$	HD $\downarrow$	NC $\uparrow$	Dice2D $\uparrow$	Dice3D $\uparrow$
Liver	DiGS [Ben-Shabat *et al.*, 2022]	0.0848 (0.0180)	0.0106 (0.0041)	0.204 (0.043)	0.360 (0.057)	0.791 (0.037)	0.770 (0.094)	0.831 (0.100)
	StEik [Yang *et al.*, 2023]	0.0481 (0.0154)	0.0038 (0.0030)	0.107 (0.041)	0.278 (0.072)	0.888 (0.032)	0.860 (0.054)	0.926 (0.039)
	NeuralPull [Ma *et al.*, 2021]	0.0362 (0.0463)	0.0067 (0.0245)	0.096 (0.169)	0.225 (0.177)	0.943 (0.047)	0.635 (0.167)	0.689 (0.164)
	NeuralTPS [Chen *et al.*, 2025]	0.0592 (0.0176)	0.0054 (0.0038)	0.149 (0.067)	0.309 (0.083)	0.907 (0.028)	0.775 (0.118)	0.850 (0.117)
	FUNSR [Chen *et al.*, 2024b]	0.1028 (0.0723)	0.0213 (0.0347)	0.246 (0.208)	0.406 (0.227)	0.848 (0.079)	0.675 (0.196)	0.764 (0.224)
	Ours	**0.0179** (0.0021)	**0.0003** (0.0003)	**0.023** (0.006)	**0.142** (0.078)	**0.972** (0.008)	**0.920** (0.049)	**0.956** (0.041)
Spleen	DiGS [Ben-Shabat *et al.*, 2022]	0.0672 (0.0260)	0.0079 (0.0055)	0.178 (0.054)	0.336 (0.084)	0.856 (0.045)	0.834 (0.082)	0.868 (0.084)
	StEik [Yang *et al.*, 2023]	0.0309 (0.0074)	0.0015 (0.0009)	0.063 (0.020)	0.223 (0.097)	0.948 (0.017)	**0.933** (0.024)	**0.963** (0.011)
	NeuralPull [Ma *et al.*, 2021]	0.0210 (0.0129)	0.0015 (0.0036)	0.052 (0.075)	0.188 (0.130)	0.981 (0.022)	0.659 (0.170)	0.698 (0.186)
	NeuralTPS [Chen *et al.*, 2025]	0.0435 (0.0135)	0.0026 (0.0024)	0.081 (0.055)	0.233 (0.093)	0.962 (0.024)	0.748 (0.154)	0.799 (0.153)
	FUNSR [Chen *et al.*, 2024b]	0.1087 (0.0240)	0.0720 (0.0986)	0.214 (0.175)	0.328 (0.387)	0.925 (0.081)	0.785 (0.198)	0.873 (0.212)
	Ours	**0.0151** (0.0014)	**0.0002** (0.0001)	**0.015** (0.004)	**0.057** (0.022)	**0.992** (0.004)	0.866 (0.074)	0.916 (0.054)
Vertebrae	DiGS [Ben-Shabat *et al.*, 2022]	0.0615 (0.0643)	0.0079 (0.0165)	0.129 (0.066)	0.287 (0.068)	0.863 (0.086)	0.789 (0.077)	0.878 (0.057)
	StEik [Yang *et al.*, 2023]	0.0569 (0.0643)	0.0072 (0.0157)	0.127 (0.072)	0.283 (0.091)	0.875 (0.098)	0.771 (0.084)	0.834 (0.130)
	NeuralPull [Ma *et al.*, 2021]	0.0419 (0.0695)	0.0053 (0.0169)	0.073 (0.076)	0.206 (0.085)	0.937 (0.103)	0.626 (0.216)	0.699 (0.247)
	NeuralTPS [Chen *et al.*, 2025]	0.0683 (0.0641)	0.0083 (0.0171)	0.150 (0.070)	0.313 (0.085)	0.902 (0.094)	0.643 (0.129)	0.719 (0.202)
	FUNSR [Chen *et al.*, 2024b]	0.0836 (0.0590)	0.0098 (0.0162)	0.156 (0.078)	0.269 (0.110)	0.787 (0.080)	0.646 (0.126)	0.742 (0.131)
	Ours	**0.0376** (0.0704)	**0.0047** (0.0167)	**0.044** (0.079)	**0.098** (0.092)	**0.949** (0.107)	**0.847** (0.070)	**0.941** (0.031)
Vertebral Column	DiGS [Ben-Shabat *et al.*, 2022]	**0.0201** (0.0056)	0.0005 (0.0003)	0.032 (0.010)	0.129 (0.042)	0.870 (0.024)	**0.869** (0.025)	0.905 (0.013)
	StEik [Yang *et al.*, 2023]	0.0202 (0.0043)	0.0005 (0.0002)	0.034 (0.012)	0.139 (0.036)	0.866 (0.022)	0.866 (0.020)	0.903 (0.012)
	NeuralPull [Ma *et al.*, 2021]	0.0236 (0.0052)	0.0010 (0.0006)	0.066 (0.021)	0.187 (0.057)	0.865 (0.023)	0.831 (0.043)	0.872 (0.040)
	NeuralTPS [Chen *et al.*, 2025]	0.0502 (0.0088)	0.0032 (0.0016)	0.120 (0.031)	0.255 (0.061)	0.773 (0.037)	0.729 (0.035)	0.771 (0.038)
	FUNSR [Chen *et al.*, 2024b]	0.0409 (0.0131)	0.0021 (0.0014)	0.087 (0.026)	0.189 (0.056)	0.782 (0.039)	0.720 (0.085)	0.777 (0.070)
	Ours	0.0208 (0.0040)	**0.0004** (0.0002)	**0.032** (0.007)	**0.118** (0.048)	**0.924** (0.006)	0.865 (0.013)	**0.906** (0.006)

Table 1: Quantitative comparison with representative SDF/INR reconstruction methods. We report surface reconstruction metrics (CD, HD) and volumetric metrics (Dice2D, Dice3D) on four anatomical datasets with sparse sampling ($k = 25$). Mean and standard deviation are reported.

approaches typically only discriminate values on the surface (where $f(x) = 0$), which fails to constrain the global field behavior.

To regularize the full 3D field, we propose an Ellipsoid-based Adversarial learning scheme that matches the distribution of the predicted SDF to a prior distribution derived from valid geometric shapes. Since the true SDF of the anatomy is unknown, we approximate the space of valid SDFs using an ensemble of randomized ellipsoids.

As shown in Figure 3, we generate synthetic ellipsoid fields $\mathcal{S}_E$ by analytically computing the SDF f_E of ellipsoids with randomized semi-axes and rotations. For a query set $\mathcal{Q} = \{\mathbf{x}_j\}_{j=1}^{M}$, this yields a reference SDF vector $\mathbf{v}_E = [f_E(\mathbf{x}_1), \ldots, f_E(\mathbf{x}_M)]$.

We employ a DeepSets-based discriminator D_ϕ, illustrated in Figure 4, which operates on sets of point-value pairs. In Figure 4, $\mathbf{X}_q, \mathbf{Y}_q, \mathbf{Z}_q$ denote the coordinate vectors of $\mathcal{Q}$, while $f_\theta(\mathcal{Q})$ and $f_E(\mathcal{Q})$ denote predicted and ellipsoid SDF values, respectively. The discriminator distinguishes $\mathbf{v}_E$ from the network prediction $\mathbf{v}_\theta = [f_\theta(\mathbf{x}_1), \ldots, f_\theta(\mathbf{x}_M)]$. The generator objective encourages the learned field to exhibit valid off-surface SDF statistics, rather than constraining only the zero level-set.

To reduce saturation and keep the discriminator as a regularizer rather than the dominant training signal, we use a small loss weight and the Least-Squares GAN objective:

$$\mathcal{L}_{adv} = \frac{1}{2}\mathbb{E}\left[(D_\phi(\mathbf{v}) - 1)^2\right]. \tag{9}$$

4.4 Optimization Objective

The final training objective combines the surface fitting loss, the exterior constraint, and the distributional regularization:

$$\mathcal{L} = \lambda_{sdf}\mathcal{L}_{sdf} + \lambda_{ext}\mathcal{L}_{ext} + \lambda_{adv}\mathcal{L}_{adv}, \tag{10}$$

where $\mathcal{L}_{sdf}$ is the standard gradient-projection loss. The adversarial term is used as a weak regularizer rather than the primary fitting objective. Unless specified otherwise, we use an 8-layer MLP with 256 hidden units, Adam with learning rate 10^{-4}, $\lambda_{sdf} = 1.0$, $\lambda_{ext} = 0.5$, $\lambda_{adv} = 0.1$, $k = 50$, and $\tau = 0.05$.

5 Experiments

5.1 Experimental Setup

Datasets and Preparation. We evaluate our method on 80 anatomical meshes sourced from MedShapeNet [Li *et al.*, 2025] and CTSpine1K [Deng *et al.*, 2021] covering four categories: *Liver*, *Spleen*, *Vertebrae*, and *Vertebral Column*. These structures range from smooth deformable organs to complex multi-component topologies. All meshes are normalized to the unit cube $[-1, 1]$. To simulate sparse free-hand ultrasound acquisition, we slice the ground-truth meshes along the longitudinal axis. While dense voxelized masks are generated for evaluation, the input point clouds $\mathcal{P}$ are derived by subsampling these slices with a sparsity factor of $k = 25$, meaning we retain every 25-th slice along the longitudinal slice index. This creates a challenging setting with significant inter-slice gaps, mimicking clinical data scarcity.

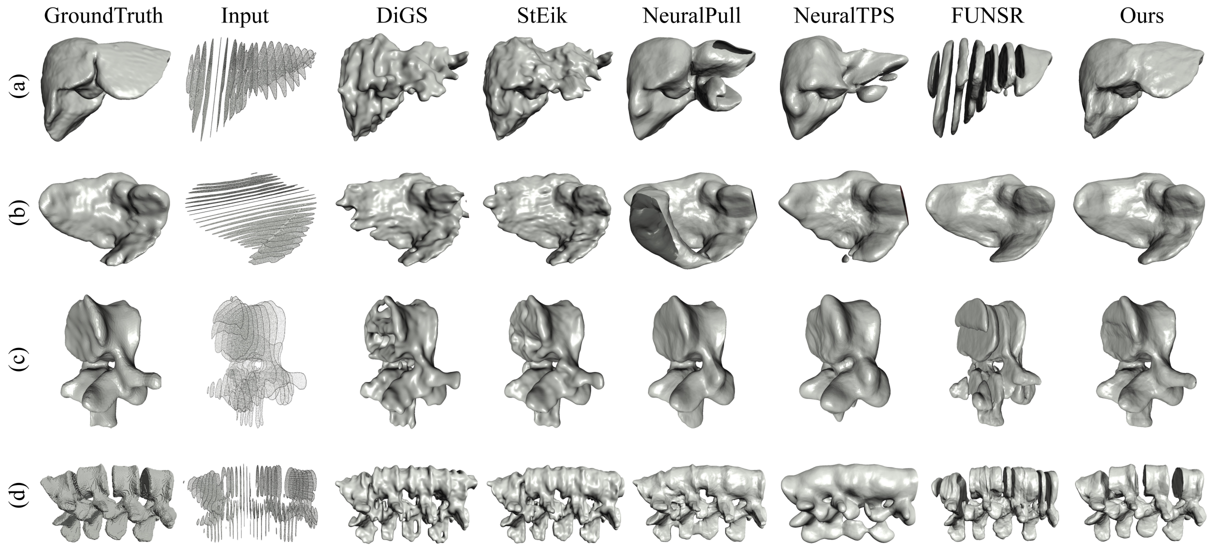


Figure 5: Qualitative comparison on sparse ultrasound reconstruction. From top to bottom: Liver, Spleen, Vertebrae, and Vertebral Column.

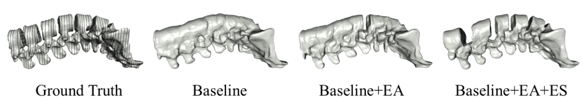


Figure 6: Visual ablation comparison. Left: The Baseline fails to bridge gaps, resulting in a fragmented surface. Middle: Adding EA improves the coarse reconstruction but may leave rough artifacts. Right: The full method combines EA and ES to produce a smooth, continuous, and geometrically accurate reconstruction.

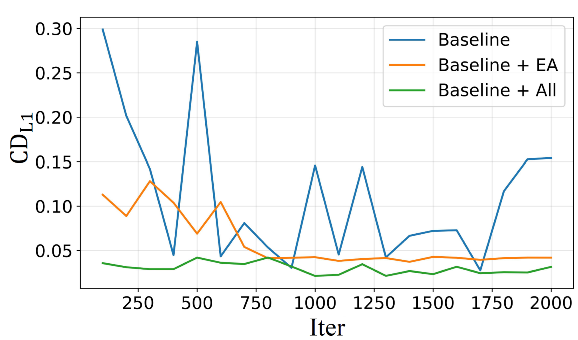


Figure 7: Convergence curves. We plot the Chamfer Distance (CD_{L1}) on the validation set during training.

Baselines. We compare against five representative implicit reconstruction methods: Eikonal-based **DiGS** [Ben-Shabat *et al.*, 2022] and **StEik** [Yang *et al.*, 2023]; gradient-projection-based **NeuralPull** [Ma *et al.*, 2021] and **FUNSR** [Chen *et al.*, 2024b]; and the thin-plate spline approach **NeuralTPS** [Chen *et al.*, 2025]. We focus on SDF/INR baselines because our analysis targets projection and Eikonal-style SDF objectives; classical grid compounding and Poisson reconstruction are discussed as reconstruction paradigms but do not optimize a directly comparable neural SDF from sparse volumetric masks.

Evaluation Metrics. Geometric fidelity is assessed using L1/L2 Chamfer Distance ($CD_{L1/L2}$), Hausdorff Distance (HD/HD95), and Normal Consistency (NC). Crucially, to quantify the field collapse phenomenon where surfaces shrink into the interior, we evaluate volumetric accuracy using the **3D Dice Score**, computed by voxelizing the reconstructed implicit field against the dense ground-truth volume.

5.2 Baseline Comparison

Quantitative Analysis. Table 1 presents a comparison across four anatomical datasets. Our method obtains the best or tied-best result on 24 of 28 reported metrics, but the exceptions are informative. On Spleen, StEik achieves higher Dice2D/Dice3D, consistent with the spleen's relatively smooth and convex geometry where strong Eikonal smoothing can favor volumetric overlap. In contrast, our method achieves lower surface errors and higher normal consistency on the same dataset, indicating better boundary fidelity. On Vertebral Column, our CD_{L1} and Dice2D are slightly below DiGS/StEik, but the gaps are small relative to the reported standard deviations; meanwhile, our method achieves the best CD_{L2}, HD, NC, and Dice3D. These results suggest that the proposed exterior anchoring is most beneficial for sparse and topologically ambiguous shapes, while not uniformly dominating every metric on simpler or already well-regularized cases.

Qualitative Analysis. The visual comparisons in Figure 5 corroborate the quantitative results. Baseline methods frequently produce shrunken geometries, characterized by the erroneous closure of anatomical gaps (e.g., in Vertebrae) and the erosion of thin structures (e.g., in Spleen). These artifacts are direct consequences of the unconstrained zero-level set drifting into the object interior. Our approach instead preserves the anatomical envelope and separates adjacent structures, validating our anti-collapse constraints on sparse, non-uniform data.

Method	CD_{L1} ↓	CD_{L2} ↓	HD95 ↓	HD ↓	NC ↑	Dice2D ↑	Dice3D ↑
Baseline	0.0409 (0.0131)	0.0021 (0.0014)	0.087 (0.026)	0.189 (0.056)	0.782 (0.039)	0.720 (0.085)	0.777 (0.070)
+ EA	0.0240 (0.0059)	0.0007 (0.0005)	0.047 (0.020)	0.157 (0.053)	0.899 (0.021)	0.839 (0.041)	0.880 (0.037)
+ EA + ES (**Ours**)	**0.0208** (0.0040)	**0.0004** (0.0002)	**0.032** (0.007)	**0.118** (0.048)	**0.924** (0.006)	**0.865** (0.013)	**0.906** (0.006)

Table 2: Quantitative ablation study of key components on the Vertebral Column dataset. Mean and standard deviation are reported.

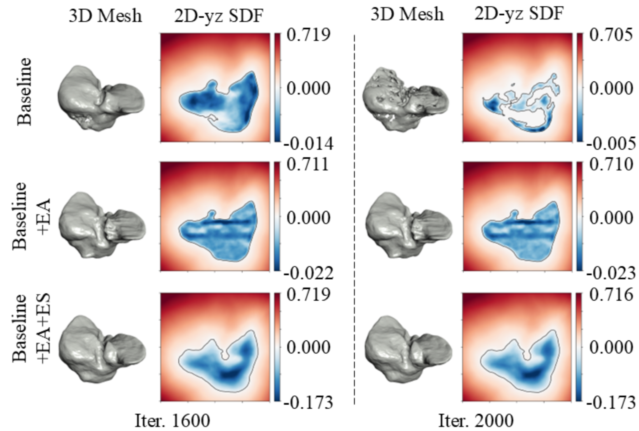


Figure 8: Evolution of the SDF landscape. We visualize the 2D-yz cross-section of the predicted SDF during training.

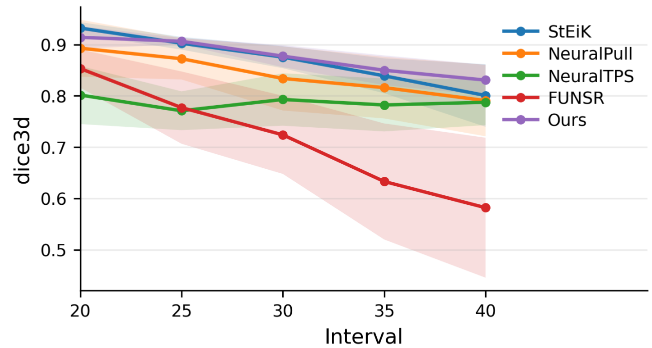


Figure 9: Robustness to sparsity. Dice3D scores across increasing slice intervals ($k \in [20, 40]$).

5.3 Ablation Study

To investigate the contribution of individual components, we conduct a stepwise ablation study on the FUNSR baseline. The results are summarized in Table 2.

Impact of Ellipsoid Adversarial (EA). The baseline model exhibits poor geometric recovery due to the inherent ambiguity of the interior field. The Ellipsoid Adversarial (EA) module regularizes the distribution of the implicit field against a synthetic geometric prior, suppressing local irregularities and improving surface smoothness.

Impact of Exterior Sampling (ES). The full model further incorporates Boundary-Directed Exterior Sampling (ES), which explicitly constrains the non-negative half-space and prevents the zero-level set from collapsing. This improves volumetric overlap and establishes a more valid coarse topology, as reflected by the final row in Table 2.

5.4 Performance Analysis

Topological Evolution. Figure 6 visually dissects the impact of our modules. The Baseline fails to bridge the inter-slice gaps, resulting in a fragmented and porous surface. The EA variant improves the coarse reconstruction but still leaves surface artifacts. The full model combines EA with $\mathcal{L}_{ext}$, yielding a continuous and anatomically plausible surface.

Convergence Stability. We further analyze the training dynamics in Figure 7. The Baseline suffers from slow and unstable convergence. In contrast, the introduction of Exterior Sampling provides a clear gradient signal from the unoccupied space, leading to a rapid and monotonic decrease in error within the first few epochs. The full method maintains this stability while achieving a lower final error floor, confirming that our geometry-aware constraints improve optimization behavior. This curve analyzes convergence quality rather than runtime or memory overhead.

Field Distribution Analysis. To further validate our theoretical analysis of interior contraction, we visualize the evolution of the learned SDF in Figure 8. As predicted in Eq. (6), the baseline method exhibits a clear tendency towards field collapse: the high-confidence positive regions (red) recede, and the zero level-set shrinks into the anatomical interior, failing to capture the concavity. Conversely, our method maintains a robust positive field in the unoccupied space. This effectively pulls the zero level-set back to the boundary, preventing the closure of the concavity and ensuring topological correctness.

Robustness to Data Sparsity. We further evaluate model resilience by increasing the slice interval k from 20 to 40, simulating extreme data scarcity. As shown in Figure 9, standard gradient-based methods suffer a catastrophic performance drop as the supporting interior points diminish, whereas our method exhibits stronger stability and maintains high volumetric accuracy.

6 Conclusion

We identified SDF collapse in sparse freehand ultrasound reconstruction as an interior-contraction bias from symmetric volumetric sampling. The proposed boundary-directed exterior sampling and ellipsoid-based adversarial regularization reduce collapse and improve most geometric and topological metrics on controlled sparse-slice anatomical datasets.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (Grant No. 62376267), the Beijing Natural Science Foundation (Grant No. L253019), and the innoHK project.

Contribution Statement

Jiuan Chen and Song Lai contributed equally. Gaofeng Meng is the corresponding author.

References

[Adriaans *et al.*, 2024] Chrissy A Adriaans, Mark Wijkhuizen, Lennard M van Karnenbeek, Freija Geldof, and Behdad Dashtbozorg. Trackerless 3D Freehand Ultrasound Reconstruction: A review. *Applied Sciences*, 14(17):7991, 2024.

[Ben-Shabat *et al.*, 2022] Yizhak Ben-Shabat, Chamin Hewa Koneputugodage, and Stephen Gould. DiGS: Divergence guided shape implicit neural representation for unoriented point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 19323–19332, 2022.

[Chen *et al.*, 2024a] Hongbo Chen, Yuchong Gao, Shuhang Zhang, Jiangjie Wu, Yuexin Ma, and Rui Zheng. RoCoSDF: Row–column scanned neural signed distance fields for freehand 3D ultrasound imaging shape reconstruction. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2024*, volume LNCS 15004, pages 721–731. Springer Nature Switzerland, 2024.

[Chen *et al.*, 2024b] Hongbo Chen, Logiraj Kumaralingam, Shuhang Zhang, Sheng Song, Fayi Zhang, Haibin Zhang, Thanh-Tu Pham, Kumaradevan Punithakumar, Edmond HM Lou, Yuyao Zhang, et al. Neural implicit surface reconstruction of freehand 3D ultrasound volume with geometric constraints. *Medical Image Analysis*, 98:103305, 2024.

[Chen *et al.*, 2025] Chao Chen, Yu-Shen Liu, and Zhizhong Han. NeuralTPS: Learning signed distance functions without priors from single sparse point clouds. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(1):565–582, 2025.

[Deng *et al.*, 2021] Yang Deng, Ce Wang, Yuan Hui, Qian Li, Jun Li, Shiwei Luo, Mengke Sun, Quan Quan, Shuxin Yang, You Hao, et al. CTSpine1K: A large-scale dataset for spinal vertebrae segmentation in computed tomography. *arXiv preprint arXiv:2105.14711*, 2021.

[Fiorentino *et al.*, 2023] Maria Chiara Fiorentino, Francesca Pia Villani, Mariachiara Di Cosmo, Emanuele Frontoni, and Sara Moccia. A review on deep-learning algorithms for fetal ultrasound-image analysis. *Medical Image Analysis*, 83:102629, 2023.

[Gropp *et al.*, 2020] Amos Gropp, Lior Yariv, Niv Haim, Matan Atzmon, and Yaron Lipman. Implicit geometric regularization for learning shapes. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 3789–3799. PMLR, 2020.

[Huang and Zeng, 2017] Qinghua Huang and Zhaozheng Zeng. A review on real-time 3D ultrasound imaging technology. *BioMed Research International*, 2017(1):6027029, 2017.

[Kazhdan and Hoppe, 2013] Michael Kazhdan and Hugues Hoppe. Screened poisson surface reconstruction. *ACM Transactions on Graphics*, 32(3):29:1–29:13, 2013.

[Li *et al.*, 2025] Jianning Li, Zongwei Zhou, Jiancheng Yang, Antonio Pepe, Christina Gsaxner, Gijs Luijten, Chongyu Qu, Tiezheng Zhang, Xiaoxi Chen, Wenxuan Li, et al. MedShapeNet–a large-scale dataset of 3D medical shapes for computer vision. *Biomedical Engineering/Biomedizinische Technik*, 70(1):71–90, 2025.

[Lorensen and Cline, 1987] William E. Lorensen and Harvey E. Cline. Marching cubes: A high resolution 3D surface construction algorithm. In *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '87)*, pages 163–169, New York, NY, USA, 1987. Association for Computing Machinery.

[Ma *et al.*, 2021] Baorui Ma, Zhizhong Han, Yu-Shen Liu, and Matthias Zwicker. Neural-Pull: Learning signed distance function from point clouds by learning to pull space onto surface. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 7246–7257. PMLR, 2021.

[Molaei *et al.*, 2023] 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 (ICCV) Workshops*, pages 2381–2391, 2023.

[Park *et al.*, 2019] 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.

[Prager *et al.*, 2010] Richard W. Prager, Umer Z. Ijaz, A. H. Gee, and Graham M. Treece. Three-dimensional ultrasound imaging. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 224(2):193–223, 2010.

[Yang *et al.*, 2023] Huizong Yang, Yuxin Sun, Ganesh Sundaramoorthi, and Anthony Yezzi. StEik: Stabilizing the optimization of neural signed distance functions and finer shape representation. In *Advances in Neural Information Processing Systems*, volume 36, 2023.