SuperSDF: Sparse SDF Super-Resolution for Surface Extraction

SAGAR PANWAR*, Inria, France
NISSIM MARUANI*, Inria, France
CÉLINE LOSCOS, L Research, France
MATHIEU DESBRUN, Inria – Ecole Polytechnique, France
PIERRE ALLIEZ, Inria, France

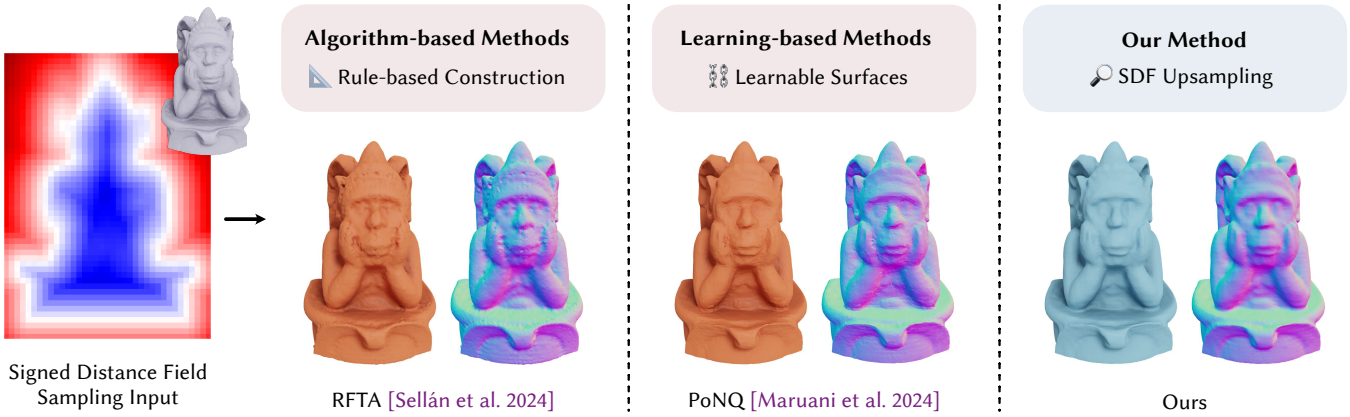


Fig. 1. **Super-resolution SDF for mesh generation** Given an input *SDF sampling* (left, shown via cross-section with false-color visualization and original 3D shape overlaid), algorithm-based methods such as Reach For the Arcs exploit analytic properties of the SDF to extract a surface mesh. Existing data-driven methods such as PoNQ learn priors based on specialized, trainable mesh representations to improve reconstruction quality. We instead propose SuperSDF, a learning-based SDF super-resolution method requiring no auxiliary mesh representation (right). Once extracted with, e.g., standard Marching Cubes [Lorensen and Cline 1987], our reconstructed meshes achieve superior surface accuracy compared to all prior methods. Unencumbered by SDF properties or learnable mesh representations, the neural network is free to operate directly in SDF space without additional representational constraints and can focus exclusively on learning high-resolution SDF geometry from data through direct sparse upsampling — thus yielding faster and higher-quality surface generation.

Signed Distance Fields (SDFs) are a powerful volumetric representation for 3D geometry. Recent advances in surface generation from SDFs increasingly rely on learnable surface representations and direct mesh supervision. In this work, we challenge this trend and show that high-quality surface reconstruction can be achieved by learning to refine the volumetric signal itself. We present SuperSDF, a learning-based approach for sparse SDF super-resolution that operates directly in SDF space, without auxiliary surface representations or mesh-level supervision. Using a sparse voxel neural network restricted to a narrow band near the surface, our method predicts high-resolution signed-distance values from coarse inputs in a scalable and resolution-agnostic manner. Standard isosurface extraction algorithms can then process the resulting super-resolved SDFs to produce accurate, detailed surface meshes. Our results show that learning-based SDF upsampling alone is sufficient to recover fine geometric details missed by classical interpolation and prior reconstruction methods. Compared to state-of-the-art machine learning approaches, our method generates higher-fidelity surfaces at a fraction of the computational cost and scales to volumetric resolutions that were previously impractical.

CCS Concepts: • **Computing methodologies** → **Shape inference**; **Shape representations**; **Volumetric models**.

Additional Key Words and Phrases: Signed Distance Fields, Surface Extraction, Super-Resolution, Sub-Voxel Accuracy.

ACM Reference Format:
Sagar Panwar, Nissim Maruani, Céline Loscos, Mathieu Desbrun, and Pierre Alliez. 2026. SuperSDF: Sparse SDF Super-Resolution for Surface Extraction. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers (SIGGRAPH Conference Papers '26), July 19–23, 2026, Los Angeles, CA, USA.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3799902.3811176

1 Introduction

While Signed Distance Fields (SDFs) offer a flexible and powerful shape representation, their dependence on structured grid sampling inherently limits their geometric fidelity, especially near fine details and sharp edges. Since downstream applications often require explicit surface meshes, the conversion of SDFs into high-quality, detail-preserving surfaces presents a significant challenge.

*Both authors contributed equally to this research.

Authors' Contact Information: Sagar Panwar, sagarpanwar249@gmail.com, Inria, Sophia-Antipolis, France; Nissim Maruani, nissim.maruani@inria.fr, Inria, Palaiseau, France; Céline Loscos, celine.loscos.pro@gmail.com, L Research, Sophi-Antipolis, France; Mathieu Desbrun, mathieu.desbrun@inria.fr, Inria – Ecole Polytechnique, Palaiseau, France; Pierre Alliez, pierre.alliez@inria.fr, Inria, Sophia-Antipolis, France.

SIGGRAPH Conference Papers '26, Los Angeles, CA, USA

Classical isosurface extraction methods such as Marching Cubes [Lorensen and Cline 1987] and its many variants offer robust and efficient ways to generate a surface from an SDF. Nevertheless, their reliance on local polynomial interpolation limits their ability to recover sub-voxel geometric details. Recent optimization-based approaches such as [Sellán et al. 2024] have attempted to overcome this limitation by exploiting the geometric properties of SDFs, but often incur significant computational overhead or introduce artifacts that degrade the quality of the extracted surface. Learning-based methods [Chen et al. 2022; Chen and Zhang 2021; Maruani et al. 2023, 2024] have recently offered a radically different and more potent approach to surface generation from SDF by introducing custom learnable surface representations and directly supervising the reconstructed geometry. While effective at learning how to inject predicted details in the final output, these methods often require specialized architectures and complex training pipelines.

In this work, we argue that existing learning methods unnecessarily entangle surface representation design with the learning process itself. We propose instead SuperSDF, a learning-based approach that focuses exclusively on *upsampling* signed distance fields. Our method employs a voxel-based neural network that efficiently predicts high-resolution SDFs from low-resolution inputs, without any direct supervision on the reconstructed surface. To improve scalability, upsampling is exclusively performed on a narrow band near the surface, leveraging a state-of-the-art sparse voxel grid framework [Williams et al. 2024]. By decoupling SDF super-resolution from the subsequent surface extraction, our approach remains compatible with standard isosurfacing methods while substantially improving reconstruction quality. We demonstrate that learning-based upsampling alone is sufficient to recover fine geometric details with modest memory requirements and computational cost. As a result, our method scales to high resolutions (see Fig. 2) that remain out of reach for prior learning-based reconstruction methods.

2 Background: Signed Distance Fields

The concept of Signed Distance Field (SDF) was originally introduced by Blinn [Blinn 1982] for the modeling and rendering of implicit surfaces. Early work demonstrated that a volumetric 3D SDF representation provides an effective means of converting unorganized point clouds into coherent surfaces [Hoppe et al. 1992]. Since then, SDFs have become a foundational tool across numerous areas of computer graphics, owing to their concise volumetric encoding and efficient spatial queries. In particular, modern surface reconstruction methods continue to rely on SDFs as a central data structure, which has proven especially advantageous in the context of learning-based reconstruction algorithms [Chen et al. 2023; Li et al. 2023; Peng et al. 2023; Wang et al. 2021; Wu et al. 2023; Zhang et al. 2022]. Furthermore, an SDF Φ possesses several exploitable geometric properties By construction, Φ satisfies the Eikonal equation, implying that its gradient has unit magnitude, $\|\nabla\Phi(x)\| = 1$, at every point where Φ is differentiable (i.e., throughout the domain except on the medial axis). Geometrically, the absolute value $|\Phi(x)|$ represents the radius of the largest open ball centered at x that remains free of surface points and is tangent to the boundary at its closest location(s). Moreover, the gradient $\nabla\Phi(x)$ indicates the direction toward the nearest point on the surface. As a consequence, the mapping $x - \Phi(x)\nabla\Phi(x)$ yields the orthogonal projection of x onto the zero-level set. These classical properties form the basis of most contemporary SDF-based surface reconstruction techniques.

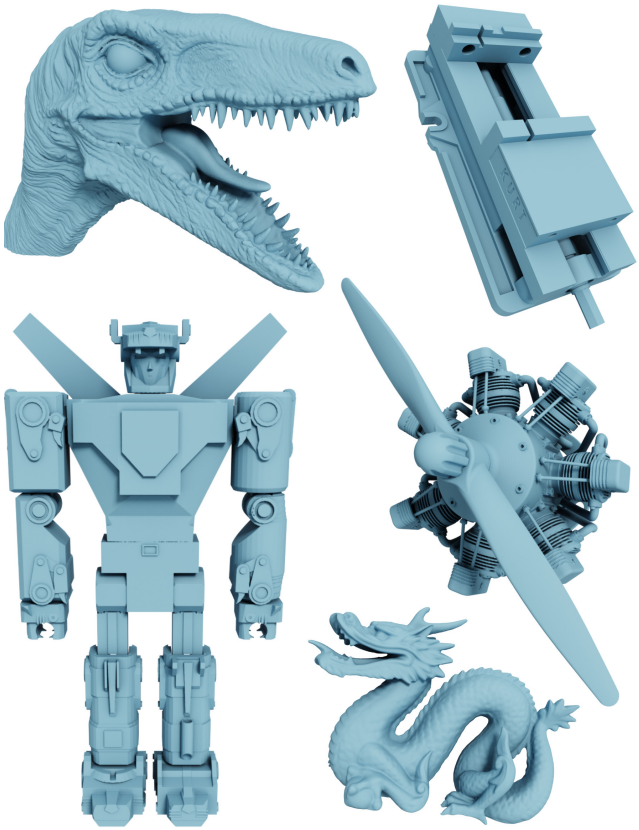


Fig. 2. **Outputs from high-resolution inputs.** Due to our efficient sparse implementation, our approach can generate high-quality surfaces of mechanical or free-from shapes at a resolution of 2048^3 from 512^3 SDF inputs.

3 Related Work

3.1 Surface Reconstruction from Signed Distance Fields

Traditional Methods. Reconstructing a surface from a signed distance field is a classical problem in geometry processing. The most common approach to isosurface extraction is the Marching Cubes (MC) algorithm [Lorensen and Cline 1987]. Given an SDF sampled over a regular grid, it iterates over the scalar field grid and generates triangles based on isosurface-crossings along edges, producing a closed and manifold surface. However, this approach is purely primal: vertices are constrained to grid edges, limiting its ability to capture sharp geometric features. To improve feature preservation, Dual Contouring [Ju et al. 2002] leverages the approximated SDF gradient to place a single vertex per isosurface-intersecting cell so as to best capture the local surface. Subsequent extensions addressed manifoldness and sharp feature reconstruction [Schaefer et al. 2007; Schaefer and Warren 2004] as well. Although all these methods are efficient in terms of memory usage and computational cost, they rely on local pre-canned polynomial interpolation of distances or distance gradients, which often leads to the loss of fine, sub-voxel

geometric details for complex geometries. More recently, a new class of optimization-based approaches has emerged. Exploiting the observation that each SDF sampled value corresponds to the center of an empty sphere, Reach for the Arcs [Sellán et al. 2024] predicts candidate tangent points and iteratively reconstructs the surface using Screened Poisson reconstruction [Kazhdan and Hoppe 2013]. Building on this idea, Kohlbrenner and Alexa [2025] further refine the formulation by explicitly constructing a set of Maximal Empty Spheres. While these approaches effectively recover details that a pure MC extraction would miss, they exhibit limitations in terms of surface smoothness and computational efficiency.

Learning-based Methods. In 2021, Chen and Zhang [2021] proposed a different approach: rather than exploiting the geometrical properties of the SDF, they trained a Convolutional Neural Network (CNN) to predict a surface from a volumetric SDF input, learning from a dataset of known watertight shapes. By restricting the network's receptive field to small, local volumetric SDF patches, their method demonstrated strong generalization to previously unseen object categories. Subsequent work further refined this approach by producing surface representations with a reduced number of vertices [Chen et al. 2022]. Addressing the same problem, VoroMesh [Maruani et al. 2023] and PoNQ [Maruani et al. 2024] propose neural surface representations based on Voronoi diagrams and Delaunay triangulations respectively, enabling more accurate and flexible surface extraction. Other approaches seek to integrate classical geometric principles into learning-based frameworks: Self-supervised Dual Contouring [Sundararaman et al. 2024] reinjects information from the input SDF during training, while FlexiCubes [Shen et al. 2023] and SparseFlex [He et al. 2025] introduce dense and sparse learnable formulations inspired by Dual Marching Cubes. Collectively, these methods demonstrate that learning-based priors can substantially improve surface reconstruction. However, their primary contributions lie in the design of *custom learnable surface representations*. In contrast, we argue that the effectiveness of neural networks can be leveraged for upsampling alone, without any direct supervision on the reconstructed surface itself.

3.2 Upsampling and Super-resolution

Signal Upsampling. Upsampling low-resolution signals is a fundamental problem in geometry and signal processing. In the context of signed distance fields, Frisken et al. [Frisken et al. 2000] showed that trilinear interpolation yields satisfactory results only in regions with smooth variations. When strict interpolation constraints are relaxed, approximating schemes such as B-spline interpolation [Thévenaz et al. 2002] can further improve reconstruction quality. While computationally efficient and widely adopted, these approaches rely on polynomial interpolation rather than leveraging data-driven priors to better inform the upsampling process.

Learning-based Super-resolution. In the image domain, Dong et al. [2015] were among the first to demonstrate the effectiveness of CNNs for super-resolution. Numerous follow-up works have since been proposed; among them, the Local Implicit Image Function [Chen et al. 2021] is particularly relevant to our work as it explicitly assumes continuity of the underlying image signal. In 3D, super-resolution techniques have been extensively explored in the context of medical imaging — see, for instance, the survey by Li et al. [2021]. More recent generative approaches employ multi-scale formulations [Ren et al. 2024; Shim et al. 2023; Xiang et al. 2025a,b], which can be interpreted as a form of super-resolution. However, none of these methods are directly applicable to SDF reconstruction. In contrast, we introduce a sparse voxel-based, lightweight neural network that is explicitly trained for the task of SDF upsampling.

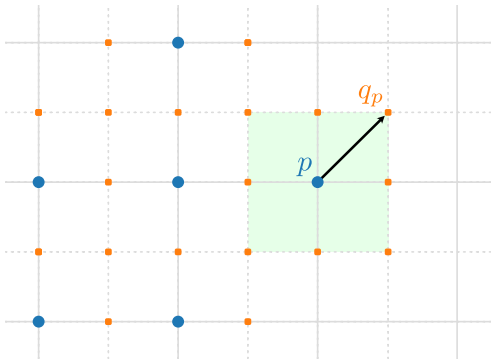


Fig. 3. **Sparse grid upsampling.** Given an input sparse voxel grid $\mathcal{I}_N$ (blue, with one active voxel highlighted in green), we first generate a sparse subdivided grid (orange). Our neural network ϕ_θ then predicts the SDF values of all the newly introduced grid nodes following the sparsity condition.

4 Method

This section presents the core components of our method. Sec. 4.1 describes the input SDF pre-processing step, which generates a sparse SDF representation. The network architecture is detailed in Sec. 4.2. Finally, training and inference are presented in Secs. 4.3 and 4.4, respectively.

4.1 Input SDF Preprocessing

From Dense to Sparse. The input to our method is a 3D regular sampling of the signed distance field. Following NDC [Chen and Zhang 2021], we first normalize each shape to lie within the half unit ball. We then construct a uniform grid $\mathcal{G}_N \subset [-0.5, 0.5]^3$ with resolution N^3 and evaluate the SDF value at every voxel center p:

$$\Phi(\mathcal{G}_N) \coloneqq \{\Phi(p)\}_{p \in \mathcal{G}_N} \tag{1}$$

(see our Supplemental Material for added comments about initial normalization). Naively applying operations on dense volumetric grids incurs prohibitive cubic complexity. As leveraged in prior work [Chen et al. 2022; Chen and Zhang 2021; Maruani et al. 2023, 2024], the most relevant signed distance values to infer the surface geometry are concentrated in the vicinity of the zero level set. Accordingly, we filter out samples with large absolute SDF values and retain only those in close proximity to the surface. This approach effectively reduces the dense grid to a sparse grid. More precisely, we apply a mask using a dynamic threshold that is inversely proportional to the grid resolution:

$$\mathcal{I}_N \coloneqq \left\{ p \in \mathcal{G}_N \;\middle|\; |\Phi(p)| < \tfrac{3}{2N} \right\}. \tag{2}$$

Our next goal is to predict an upsampling of the sparse SDF voxel grid that we denote $\{\mathcal{I}_N, \Phi(\mathcal{I}_N)\}$.

Sparse Grid Refinement. In order to maintain a sparse pipeline, we focus our SDF predictions to regions near the active voxels $\mathcal{I}_N$ of the input grid. Specifically, for each active voxel center p, we generate a set of candidate query locations $q_p \in \mathbb{R}^3$ on a finer-resolution grid $(K \times N)^3$ within its local neighborhood (Fig. 3). Each query location must meet two conditions: (i) alignment with the finer-resolution grid, and (ii) inclusion within the associated voxel cell, that is:

$$q_p \in \mathcal{G}_{K\times N} \quad \text{and} \quad \|q_p - p\|_\infty \leq \frac{1}{2N} \tag{3}$$

This formulation ensures that the input (ground-truth) SDF values are preserved in the finer grid. We now have all the necessary components to construct our super-resolution network.

4.2 Neural Network Architecture

Architecture. Among the super-resolution architectures reviewed in Sec. 3, we adopt the method of Chen et al. [2021], as it explicitly assumes continuity of the underlying signal, an assumption that holds for signed distance fields. Following their formulation, our network ϕ_θ is parametrized to predict SDF values at continuous query locations $q_p \in \mathbb{R}^3$ within each voxel (Fig. 4). In practice, we concatenate each query location q_p with the sparsified input $\Phi(p)$ to form 4D voxel features subsequently processed by the network.

Our architecture is based on a sparse 3D variant of the Residual U-Net [Zhang et al. 2018], featuring ResNet-style residual blocks [He et al. 2016] and skip connections. Consistent with prior work [Chen et al. 2022; Chen and Zhang 2021; Maruani et al. 2023, 2024], we employ a restricted receptive field to ensure generalization to unseen shapes. Our U-Net backbone comprises four downsampling stages, a bottleneck stage, and four upsampling stages, with two residual blocks per stage for a total of 21M trainable parameters — see our Supplemental Material for a schematic description of our network.

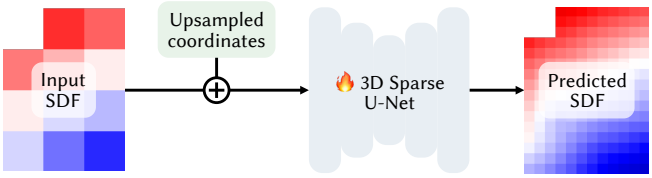


Fig. 4. **Overview of our SDF super-resolution approach.** SuperSDF uses a sparse, scalable voxel pipeline to learn SDF super-resolution to generate high-quality surfaces from coarse SDFs, without learnable surface representations or mesh supervision involved.

Scale-agnostic Rescaling. Our network takes as input the coarse SDF values together with their corresponding fine query locations. Both quantities are scale-dependent, with magnitudes that vary according to the voxel resolution. To ensure generalization to arbitrary scales, we apply a scale-agnostic normalization. We rescale the SDF Φ to $\tilde{\Phi}$ and query locations q_p to $\tilde{q}_p$ so as to lie within fixed ranges:

$$\tilde{\Phi}(I_N) \coloneqq 2N\Phi(I_N) \in [-3,\, 3] \tag{4}$$

$$\tilde{q}_p \coloneqq 2N(q_p - p) \in [-1, 1]^3. \tag{5}$$

4.3 Training

Data preparation. Following prior work, we train our model on the first split of the ABC dataset [Koch et al. 2019], which consists of approximately 5.3K CAD models. After normalization, we precompute the SDF values at resolutions of 32^3 and 128^3. We train our network exclusively for 4× upsampling, as this configuration yields the best empirical performance (see Sec. 5.3 for further discussion). Note that owing to our scale-agnostic normalization, the trained model can be directly applied to finer-resolution input grids at inference time without requiring additional training.

Loss Function. During training, we sample a random query position q_p for each voxel (see Fig. 3), and minimize the mean squared error (MSE) between the predicted and ground-truth SDF values through a loss of the form:

$$\mathcal{L}(\theta) \coloneqq \frac{1}{|\mathcal{I}_N|} \sum_{p\in\mathcal{I}_N} \left\|\phi_\theta\big(\tilde{\Phi}(\mathcal{I}_N), \tilde{q}_p\big) - \Phi(q_p)\right\|_2^2. \tag{6}$$

Additional details. We train our model for 300 epochs with a batch size of 16 and AdamW (`torch.optim.AdamW`) for a learning rate of 10^{-3}. Training takes approximately 3.5 hours on a single GPU.

4.4 Inference

Final SDF prediction. For each active voxel, our method predicts an SDF value at a given query location q. During inference, this can lead to multiple predictions for the same spatial point, since boundary locations may be shared by neighboring voxels. Specifically, a query point may lie on a face shared by two voxels, an edge shared by four voxels, or a corner shared by eight voxels. To obtain a smooth and stable estimate (see Chen et al. [2021]), we simply average these redundant predictions. If we denote by $\mathcal{V}(q) \coloneqq \{p_i \mid \exists q_{p_i} = q\}$ the multiple prediction at q, the final SDF value at q is computed as:

$$\Phi_\theta(q) \coloneqq \frac{1}{|\mathcal{V}(q)|} \sum_{p_i\in\mathcal{V}(q)} \phi_\theta\big(\tilde{\Phi}(\mathcal{I}_N), \tilde{q}_{p_i}\big) \tag{7}$$

Mesh Extraction. The final surface mesh is obtained by applying a standard isosurface extraction method, such as Marching Cubes [Lorensen and Cline 1987], to the generated high-resolution SDF grid. Unlike previous learning-based methods, our approach does not require any supervision on the extracted surface itself, which streamlines the overall pipeline, yet delivers state-of-the-art performance as we demonstrate next.

5 Experiments

We now present a series of experiments organized into three parts: Sec. 5.1 evaluates the quality of SDF upsampling, Sec. 5.2 compares our reconstructed surfaces with those produced by state-of-the-art methods, while Sec. 5.3 details additional experiments and applications of our method. Following standard practice, we evaluate all methods on a subset of the Thingi10k dataset [Zhou and Jacobson 2016] consisting of 30 watertight shapes selected by VoroMesh [Maruani et al. 2023]. For a fair comparison, none of the learning-based baselines are trained on this dataset. Both the upsampling and reconstruction pipelines, as well as the additional experimental setups, are implemented in Python using PyTorch and the sparse fVDB framework [Williams et al. 2024]. We trained our network on a workstation with an NVIDIA RTX A6000 GPU (49 GB VRAM); at inference time, we used a consumer-grade NVIDIA GeForce RTX 3080 GPU (10 GB VRAM). We encourage the reader

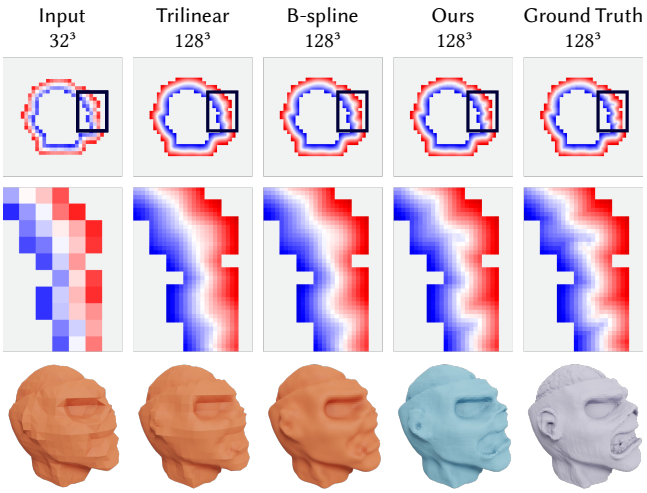


Fig. 5. **Comparing upsampling methods** From a sparse, coarse SDF of zombie head (a), trilinear (b) and spline-based (b) interpolations smooth the SDF well, but distort the zero isovalue due to the arbitrariness of these polynomial-based upsampling. In contrast, our data-driven method produces a super-resolved SDF that captures fine details and aligns with the ground truth far more accurately.

Table 1. **Statistics on SDF upsampling**. Compared to polynomial interpolation of the SDF, SuperSDF improves the MSE and MAE on Thingi30, for input sizes 32^3 (top), 64^3 (middle), and 128^3 (bottom).

Method	Input Size	MSE ↓	MAE ↓
Trilinear	32^3	7.19×10^{-6}	1.77×10^{-3}
B-Spline	32^3	3.10×10^{-6}	1.14×10^{-3}
Ours	32^3	$\mathbf{1.82 \times 10^{-6}}$	$\mathbf{8.51 \times 10^{-4}}$
Trilinear	64^3	1.11×10^{-6}	6.38×10^{-4}
B-Spline	64^3	4.59×10^{-7}	4.01×10^{-4}
Ours	64^3	$\mathbf{2.55 \times 10^{-7}}$	$\mathbf{3.15 \times 10^{-4}}$
Trilinear	128^3	1.51×10^{-7}	2.10×10^{-4}
B-Spline	128^3	5.34×10^{-8}	1.23×10^{-4}
Ours	128^3	$\mathbf{3.00 \times 10^{-8}}$	$\mathbf{1.14 \times 10^{-4}}$

to consult our supplementary material (including a video summary and sample meshes) for a better assessment of our method.

5.1 SDF Upsampling

Baselines. To evaluate the effectiveness of our learning-based SDF upsampling, we compare it against two classical baselines for an upsampling of a factor ×4: trilinear interpolation [Frisken et al. 2000] and cubic B-splines of order 3 [Thévenaz et al. 2002], implemented in SciPy [Virtanen et al. 2020].

Metrics. We evaluate all methods on the same active voxels near the isosurface (see Eq. (2) and Fig. 5). For trilinear interpolation and B-spline upsampling, the output is first computed from the dense input and then masked. We evaluate reconstruction accuracy using the mean absolute error (MAE) and mean squared error (MSE), corresponding to the mean ℓ_1 and ℓ_2 distances between voxel-wise predicted and ground-truth SDF values.

Results. Qualitatively, our learned priors help SuperSDF recover thin structures far better than traditional interpolation schemes. As illustrated in Fig. 5, starting from a relatively coarse input, the predicted SDF more closely matches the ground-truth values and enables the recovery of fine (subvoxel) geometric details, whereas trilinear and B-spline interpolation tend to smooth out high-frequency structures by blindly applying a polynomial interpolation. Quantitatively, our method consistently outperforms both trilinear interpolation and B-spline interpolation in terms of MAE and MSE across all evaluated resolutions (32^3, 64^3, and 128^3, see Tab. 1). These results highlight the benefit of learned priors for SDF upsampling, a point we further substantiate in the following section.

5.2 Surface Reconstruction

Baselines. We compare output mesh quality against several methods: (i) Marching Cubes (MC) [Lorensen and Cline 1987]; (ii) Marching Cubes applied to SDFs upsampled via trilinear interpolation or cubic B-splines (see Section 5.1); (iii) Maximal Empty Spheres (MES) [Kohlbrenner and Alexa 2025]; (iv) Reach for the Arcs (RFTA) [Sellán et al. 2024]; and (v) PoNQ [Maruani et al. 2024]. All methods are evaluated on identical input SDFs, and all baselines are executed using the official implementations released by the respective authors. For MES, results are reported only up to a resolution of 32^3, as computation times become prohibitive for certain shapes in our dataset (see Tab. 2). This limitation does not compromise the comparison since lower resolutions pose the greatest challenge for accurate mesh reconstruction.

Metrics. We follow the evaluation protocol of prior works [Chen et al. 2022; Chen and Zhang 2021; Maruani et al. 2023, 2024]. Specifically, we uniformly sample 10^5 points from both the ground-truth and reconstructed meshes and compute the Chamfer Distance (CD), F1 score, and Normal Consistency (NC) to offer several geometric scores. The use of multiple metrics enables a comprehensive assessment of surface quality under complementary criteria.

Results. Qualitatively, our approach achieves excellent surface generation, with a much reduced number of visual artifacts — see Figs. 1, 5, 7, 8, 12, and 13. In contrast to Marching Cubes, which is

Table 2. **Average Timings on Thingi30**. While some prior methods incur large runtimes, our approach outputs meshes in seconds, with an output quality far superior to direct extraction methods such as Marching Cubes.

Method	Input Size	Mean	Min	Max
MES	32^3	185.464	38.121	1475.907
RFTA	32^3	25.0348	19.146	28.784
PoNQ	32^3	0.974	**0.513**	1.519
Ours+MC	32^3	**0.764**	0.670	**1.154**
RFTA	64^3	106.346	81.519	155.856
PoNQ	64^3	5.873	2.977	8.476
Ours+MC	64^3	**1.125**	**0.857**	**1.449**
RFTA	128^3	1095.909	628.926	1785.074
PoNQ	128^3	56.663	22.711	85.210
Ours+MC	128^3	**2.778**	**1.740**	**3.971**

Table 3. **Surface reconstruction on Thingi30.** Even combined with just Marching Cubes for surface extraction, our method still outperforms prior work in terms of surface (CD, F1) and Normal (NC) accuracy.

Method	Input Size	CD ↓ ($\times 10^{-5}$)	F1 ↑	NC ↑
MC	32^3	9.124	0.606	0.905
Trilinear+MC	32^3	8.622	0.632	0.918
B-Spline+MC	32^3	6.514	0.743	0.937
MES	32^3	9.528	0.454	0.925
RFTA	32^3	3.339	0.698	0.935
PoNQ	32^3	1.363	0.810	0.942
Ours+MC	32^3	**1.351**	**0.820**	**0.950**
MC	64^3	1.166	0.854	0.952
Trilinear+MC	64^3	1.086	0.868	0.961
B-Spline+MC	64^3	0.815	0.912	0.973
RFTA	64^3	1.180	0.869	0.963
PoNQ	64^3	0.775	0.924	0.971
Ours+MC	64^3	**0.711**	**0.925**	**0.977**
MC	128^3	0.677	0.933	0.978
Trilinear+MC	128^3	0.667	0.935	0.982
B-Spline+MC	128^3	0.643	0.938	0.988
RFTA	128^3	0.926	0.906	0.975
PoNQ	128^3	0.640	0.939	0.984
Ours+MC	128^3	**0.638**	**0.939**	**0.988**

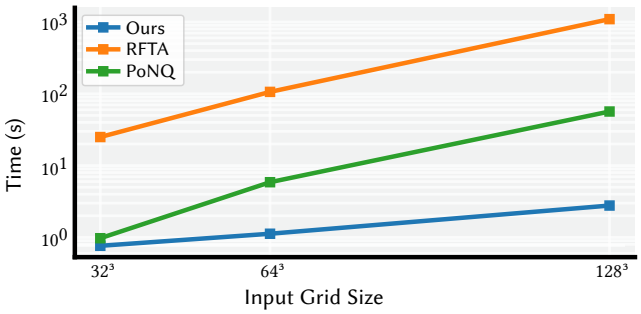
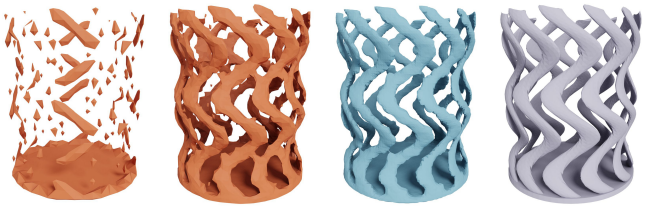


Fig. 6. **Timings.** In contrast to RFTA [Sellán et al. 2024] that relies on optimization or PoNQ [Maruani et al. 2024] that requires an ad-hoc mesh extraction procedure, our method efficiently produces highly detailed meshes with competitive runtimes, e.g., 2.8 seconds on average for a 128^3 input grid.

inherently constrained to place edge crossings on the input grid, our super-resolution formulation performs subdivision of the input domain, enabling the recovery of geometric details at subvoxel scale (Fig. 8). Compared to traditional methods, we provide a more accurate estimate of the surface as demonstrated in Fig. 7. Quantitatively, our method consistently outperforms all baselines across all metrics and resolutions (see Tab. 3). Improvements are observed in Normal Consistency, Chamfer Distance, and F1 score, with the largest gains appearing at low resolutions. This trend is expected: as the input resolution increases, the SDF becomes progressively less ambiguous, narrowing the performance gap among all methods. These results substantiate the central claim of our work that high-quality surface extraction from SDF can be achieved without *any* surface supervision or learnable surface representation. Beyond

(a) MES (b) RFTA (c) Ours (d) Ground Truth

Fig. 7. **Inspecting results.** From a coarse SDF of the ground truth mesh (d), Maximal Empty Sphere (a) fails to capture many details of the original shape. Reach for the Arcs (b) produces a detailed-enriched result, but its normal map shows significant artifacts. In contrast, both our geometry and its normal field (c) are closer to the original shape.



(a) March. Cubes (b) PoNQ (c) Ours (d) Ground Truth

Fig. 8. **Geometry with Fine Details.** A coarse SDF of this shape cannot be directly extracted via MC without entirely losing its geometry and topology. While PoNQ recovers fine structures quite well, it produces numerous visible artifacts compared to SuperSDF.

reconstruction quality, our streamlined framework also achieves competitive runtime performance. For example, while RFTA can require up to 30 minutes to process some objects at a resolution of 128^3, our method completes the same task in at most 4 seconds. As a result, SuperSDF scales to substantially larger inputs than prior approaches (see Fig. 2), requiring only a few minutes at 512^3 resolution. Additional timings are provided in Fig. 6 and Tab. 2. Compared to state-of-the-art methods for surface generation from SDF, SuperSDF reduces runtimes from several minutes to a few seconds.

5.3 Additional experiments

In this section, we present additional experiments that highlight interesting applications of our method.

Upscaling factor impact analysis. While we train and evaluate our model using a ×4 upscaling factor, we can use other upscaling factors at test time given our continuous formulation. Fig. 9 reports results on the horse head model for factors $\{\times 1, \times 2, \times 4, \times 8, \times 16\}$ with a 32^3 input. The most significant gains are observed when increasing the resolution to ×4, which motivates the choice of this

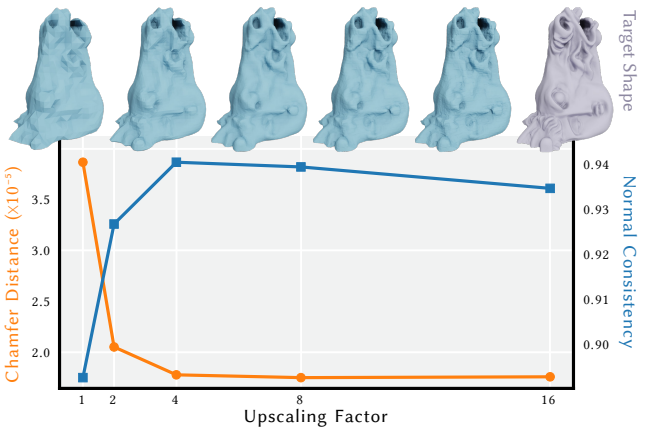


Fig. 9. **Evaluations for various upscaling factors.** Despite supporting arbitrary upscaling factors, our method (here for a horse head model using a 32^3 SDF input) performs most reliably at ×4 upscaling when evaluated based on chamfer distance (orange) and normal consistency (blue).

factor throughout the paper. Beyond ×4, scores of Normal Consistency and Chamfer Distance slightly degrade. We attribute this behavior to (a) the fact that these higher resolutions are not encountered during training and (b) the inherent loss of information and possible ambiguities imposed by too coarse input SDFs.

Offset. Although most of this work addresses zero-isosurface extraction, our method is equally applicable to offset surfaces $\{x \in \mathbb{R}^3 \mid \Phi(x) = \delta\}$ (e.g., Zint et al. [2023]). Direct inference on offsets is challenging, as our network is trained solely on a thin band straddling the zero-isosurface. However, by uniformly shifting the input SDF by δ before entering the pipeline, we effectively re-purpose the trained network for offset extraction, see Fig. 10.

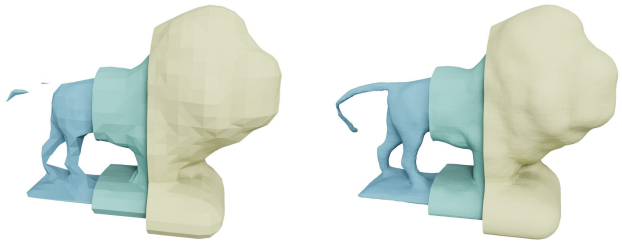
Dual contouring. Although we primarily rely on Marching Cubes due to the availability of an efficient GPU-based implementation in fVDB [Williams et al. 2024], alternative mesh extraction techniques are also directly applicable. In particular, voxel-wise SDF gradients computed via finite differences enable the application of the Dual Contouring algorithm, which produces meshes with improved edge sharpness (see Fig. 11).

6 Discussion

SuperSDF, our super-resolution framework, delivers state-of-the-art performance for surface generation from SDFs. Several avenues, however, remain open for further exploration.

To keep training costs manageable, we relied on ground-truth SDFs and a fixed ×4 upsampling factor. It might also be interesting to handle approximate SDFs, especially those generated by a neural network. Exposing our model at training time to noisy SDFs, and to a wider range of upsampling factors, may improve SuperSDF in terms of robustness and generalizability to inputs of lower quality and/or resolution.

Although our backbone architecture proves to be effective, the overall approach is model-agnostic. Incorporating more recent generative techniques, such as flow matching [Lipman et al. 2023], may further improve geometric accuracy and produce sharper reconstructions, particularly at high upscaling ratios.



(a) Marching Cubes (b) Ours

Fig. 10. **Offset surfaces.** By adding an offset δ to an input SDF grid $\Phi(\mathcal{I}_N)$, our learning-based upsampling approach can be directly used to compute offset surfaces. Here, we start from a coarse SDF of a lion model and extract meshes for three offsets ($\delta \in \{0, 0.3, 0.6\}$), using Marching Cubes applied either to the original coarse grid (a) or to our super-resolved SDF grid (b).

Finally, although we employ a uniform upscaling factor across the entire grid, further efficiency gains could be achieved by *adaptively* subdividing the input based on a network-predicted estimate of the underlying surface complexity.

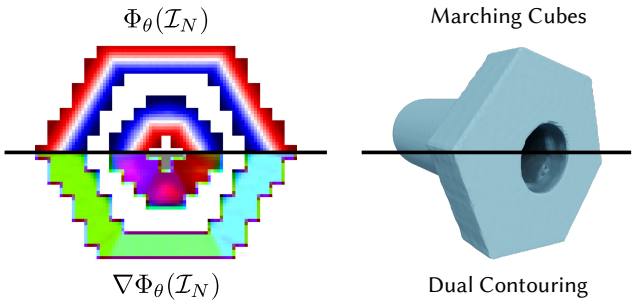


Fig. 11. **Alternative Mesh Extraction.** While our experiments all rely on Marching Cubes (top) due to the availability of efficient GPU implementations, we can also estimate the gradient $\nabla\Phi(\mathcal{I}_N)$ of the super-resolved SDF and generate a mesh via Dual Contouring (bottom) for instance.

7 Conclusion

We have introduced SuperSDF, a learning-based framework for sparse signed distance field super-resolution that departs from the prevailing trend of designing increasingly elaborate learnable surface representations. By focusing exclusively on SDF upsampling, our method demonstrates that high-quality surface reconstruction does not require direct supervision on meshes, nor the introduction of specialized geometric priors beyond those implicitly captured by the network during training. The resulting pipeline is conceptually simple, computationally efficient, and empirically effective: it scales to high resolutions and recovers geometric details consistently better than current SOTA methods across a range of metrics. Moreover, SuperSDF substantially accelerates inference, reducing runtime from minutes to seconds.

Stepping back, our approach highlights a novel perspective that is applicable beyond SDF reconstruction: learning-based 3D pipelines need not always rely on explicitly trained surface representations to create detailed 3D shapes. This decoupling suggests that future advances in geometric deep learning may lie not in reinventing isosurface extraction, but in faithfully refining the underlying volumetric signals that underpin them.

Acknowledgments

All meshes in this paper are from Thingi10K [Zhou and Jacobson 2016]. This work was supported by the French government through the 3IA Côte d'Azur Investments in the project managed by the National Research Agency (ANR-23-IACL-0001), a Choose France Inria chair, the MediTwin consortium (funded as part of France 2030), and by Synopsis and Adobe Research.

References

James F. Blinn. 1982. A Generalization of Algebraic Surface Drawing. *ACM Trans. Graph.* 1, 3 (July 1982), 235–256. doi:10.1145/357306.357310

Yinbo Chen, Sifei Liu, and Xiaolong Wang. 2021. Learning continuous image representation with local implicit image function. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 8628–8638.

Zhang Chen, Zhong Li, Liangchen Song, Lele Chen, Jingyi Yu, Junsong Yuan, and Yi Xu. 2023. NeuRBF: A Neural Fields Representation with Adaptive Radial Basis Functions. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE, Paris, France, 4159–4171. doi:10.1109/ICCV51070.2023.00386

Zhiqin Chen, Andrea Tagliasacchi, Thomas Funkhouser, and Hao Zhang. 2022. Neural dual contouring. *ACM Transactions on Graphics* 41, 4 (July 2022), 1–13. doi:10.1145/3528223.3530108

Zhiqin Chen and Hao Zhang. 2021. Neural marching cubes. *ACM Transactions on Graphics* 40, 6 (Dec. 2021), 1–15. doi:10.1145/3478513.3480518

Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. 2015. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence* 38, 2 (2015), 295–307. Publisher: IEEE.

Sarah F Frisken, Ronald N Perry, Alyn P Rockwood, and Thouis R Jones. 2000. Adaptively sampled distance fields: A general representation of shape for computer graphics. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. 249–254.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.

Xianglong He, Zi-Xin Zou, Chia-Hao Chen, Yuan-Chen Guo, Ding Liang, Chun Yuan, Wanli Ouyang, Yan-Pei Cao, and Yangguang Li. 2025. SparseFlex: High-Resolution and Arbitrary-Topology 3D Shape Modeling. doi:10.48550/arXiv.2503.21732 arXiv:2503.21732 [cs].

Hugues Hoppe, Tony DeRose, Tom Duchamp, John McDonald, and Werner Stuetzle. 1992. Surface reconstruction from unorganized points. In *Proceedings of the 19th annual conference on computer graphics and interactive techniques*. 71–78.

Tao Ju, Frank Losasso, Scott Schaefer, and Joe Warren. 2002. Dual contouring of hermite data. In *Proceedings of the 29th annual conference on Computer graphics and interactive techniques (SIGGRAPH '02)*. Association for Computing Machinery, New York, NY, USA, 339–346. doi:10.1145/566570.566586

Michael Kazhdan and Hugues Hoppe. 2013. Screened poisson surface reconstruction. *ACM Transactions on Graphics* 32, 3 (June 2013), 1–13. doi:10.1145/2487228.2487237

Sebastian Koch, Albert Matveev, Zhongshi Jiang, Francis Williams, Alexey Artemov, Evgeny Burnaev, Marc Alexa, Denis Zorin, and Daniele Panozzo. 2019. Abc: A big cad model dataset for geometric deep learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 9601–9611.

Maximilian Kohlbrenner and Marc Alexa. 2025. A Polyhedral Construction of Empty Spheres in Discrete Distance Fields. In *Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers*. 1–10.

Yufei Li, Bruno Sixou, and Francois Peyrin. 2021. A review of the deep learning methods for medical images super resolution problems. *Irbm* 42, 2 (2021), 120–133. Publisher: Elsevier.

Zhaoshuo Li, Thomas Müller, Alex Evans, Russell H. Taylor, Mathias Unberath, Ming-Yu Liu, and Chen-Hsuan Lin. 2023. Neuralangelo: High-Fidelity Neural Surface Reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 8456–8465.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. 2023. Flow Matching for Generative Modeling. doi:10.48550/arXiv.2210.02747 arXiv:2210.02747 [cs].

William E. Lorensen and Harvey E. Cline. 1987. Marching cubes: A high resolution 3D surface construction algorithm. *ACM SIGGRAPH Computer Graphics* 21, 4 (Aug. 1987), 163–169. doi:10.1145/37402.37422

Nissim Maruani, Roman Klokov, Maks Ovsjanikov, Pierre Alliez, and Mathieu Desbrun. 2023. VoroMesh: Learning Watertight Surface Meshes with Voronoi Diagrams. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. 14565–14574. https://openaccess.thecvf.com/content/ICCV2023/html/Maruani_VoroMesh_Learning_Watertight_Surface_Meshes_with_Voronoi_Diagrams_ICCV_2023_paper.html

Nissim Maruani, Maks Ovsjanikov, Pierre Alliez, and Mathieu Desbrun. 2024. PoNQ: A Neural QEM-Based Mesh Representation. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Seattle, WA, USA, 3647–3657. doi:10.1109/CVPR52733.2024.00350

Rui Peng, Xiaodong Gu, Luyang Tang, Shihe Shen, Fanqi Yu, and Ronggang Wang. 2023. GenS: Generalizable Neural Surface Reconstruction from Multi-View Images. In *Thirty-seventh Conference on Neural Information Processing Systems (NeurIPS)*.

Xuanchi Ren, Jiahui Huang, Xiaohui Zeng, Ken Museth, Sanja Fidler, and Francis Williams. 2024. XCube: Large-Scale 3D Generative Modeling using Sparse Voxel Hierarchies. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Seattle, WA, USA, 4209–4219. doi:10.1109/CVPR52733.2024.00403

Scott Schaefer, Tao Ju, and Joe Warren. 2007. Manifold Dual Contouring. *IEEE Transactions on Visualization and Computer Graphics* 13, 3 (May 2007), 610–619. doi:10.1109/TVCG.2007.1012

Scott Schaefer and Joe Warren. 2004. Dual marching cubes: Primal contouring of dual grids. IEEE, 70–76.

Silvia Sellán, Yingying Ren, Christopher Batty, and Oded Stein. 2024. Reach for the Arcs: Reconstructing Surfaces from SDFs via Tangent Points. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers*. ACM, Denver CO USA, 1–12. doi:10.1145/3641519.3657419

Tianchang Shen, Jacob Munkberg, Jon Hasselgren, Kangxue Yin, Zian Wang, Wenzheng Chen, Zan Gojcic, Sanja Fidler, Nicholas Sharp, and Jun Gao. 2023. Flexible Isosurface Extraction for Gradient-Based Mesh Optimization. *ACM Transactions on Graphics* 42, 4 (Aug. 2023), 1–16. doi:10.1145/3592430

Jaehyeok Shim, Changwoo Kang, and Kyungdon Joo. 2023. Diffusion-Based Signed Distance Fields for 3D Shape Generation. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Vancouver, BC, Canada, 20887–20897. doi:10.1109/CVPR52729.2023.02001

Ramana Sundararaman, Roman Klokov, and Maks Ovsjanikov. 2024. Self-Supervised Dual Contouring. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 4681–4691.

Philippe Thévenaz, Thierry Blu, and Michael Unser. 2002. Interpolation revisited [medical images application]. *IEEE Transactions on medical imaging* 19, 7 (2002), 739–758.

Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C. J. Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, and Paul van Mulbregt. 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* 17, 3 (March 2020), 261–272. doi:10.1038/s41592-019-0686-2 Number: 3 Publisher: Nature Publishing Group.

Peng Wang, Lingjie Liu, Yuan Liu, Christian Theobalt, Taku Komura, and Wenping Wang. 2021. NeuS: Learning Neural Implicit Surfaces by Volume Rendering for Multi-view Reconstruction. In *Advances in Neural Information Processing Systems*, M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan (Eds.), Vol. 34. Curran Associates, Inc., 27171–27183. https://proceedings.neurips.cc/paper_files/paper/2021/file/e41e164f7485ec4a28741a2d0ea41c74-Paper.pdf

Francis Williams, Jiahui Huang, Jonathan Swartz, Gergely Klar, Vijay Thakkar, Matthew Cong, Xuanchi Ren, Ruilong Li, Clement Fuji-Tsang, Sanja Fidler, et al. 2024. fvdb: A deep-learning framework for sparse, large scale, and high performance spatial intelligence. *ACM Transactions on Graphics (TOG)* 43, 4 (2024), 1–15.

Tong Wu, Jiaqi Wang, Xingang Pan, Xudong Xu, Christian Theobalt, Ziwei Liu, and Dahua Lin. 2023. Voxurf: Voxel-based Efficient and Accurate Neural Surface Reconstruction. In *International Conference on Learning Representations (ICLR)*.

Jianfeng Xiang, Xiaoxue Chen, Sicheng Xu, Ruicheng Wang, Zelong Lv, Yu Deng, Hongyuan Zhu, Yue Dong, Hao Zhao, Nicholas Jing Yuan, and Jiaolong Yang. 2025a. Native and Compact Structured Latents for 3D Generation. doi:10.48550/arXiv.2512.14692 arXiv:2512.14692 [cs].

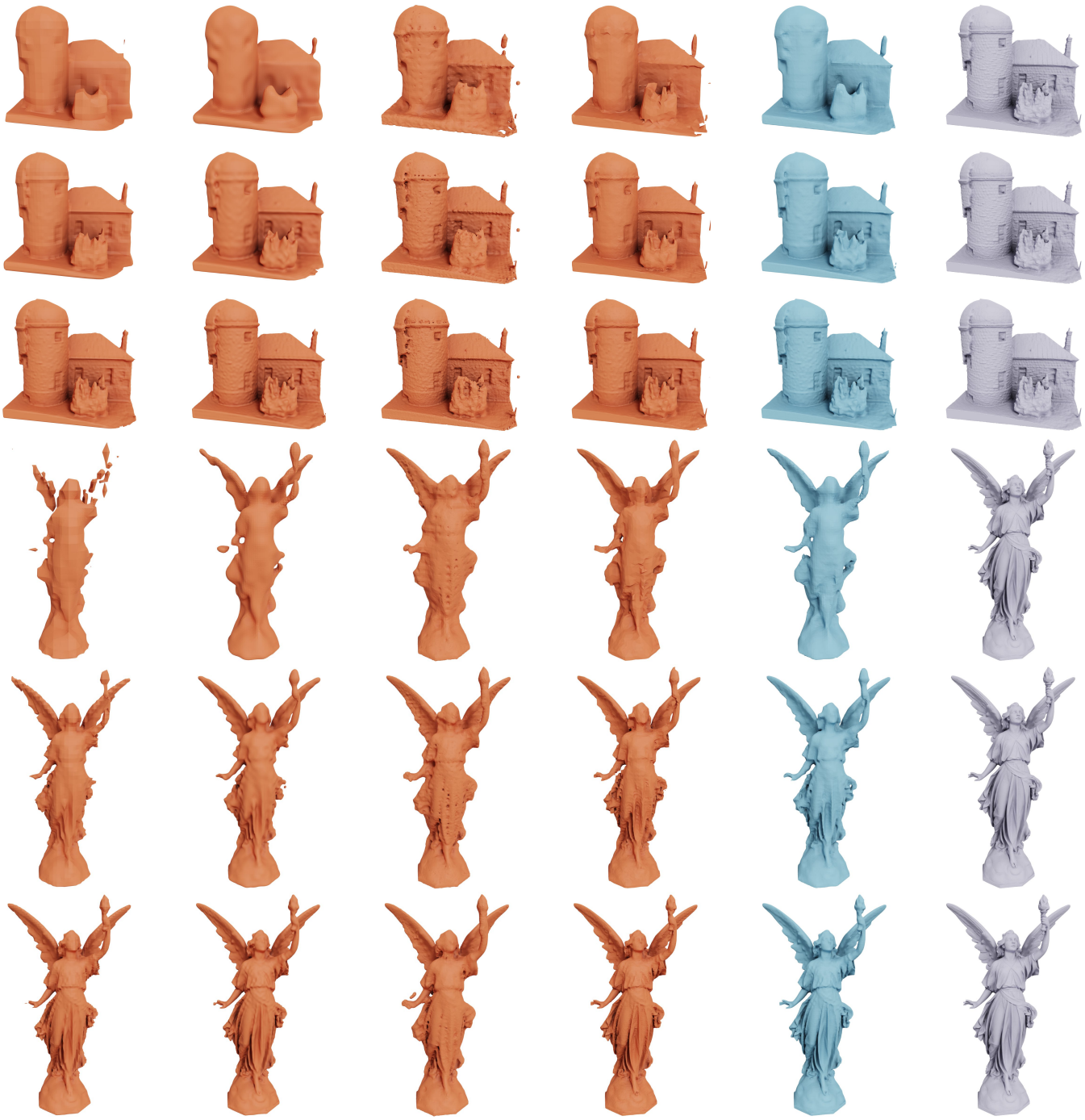
Jianfeng Xiang, Zelong Lv, Sicheng Xu, Yu Deng, Ruicheng Wang, Bowen Zhang, Dong Chen, Xin Tong, and Jiaolong Yang. 2025b. Structured 3d latents for scalable and versatile 3d generation. 21469–21480.

Jingyang Zhang, Yao Yao, Shiwei Li, Tian Fang, David McKinnon, Yanghai Tsin, and Long Quan. 2022. Critical Regularizations for Neural Surface Reconstruction in the Wild. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 6270–6279.

Zhengxin Zhang, Qingjie Liu, and Yunhong Wang. 2018. Road extraction by deep residual u-net. *IEEE Geoscience and Remote Sensing Letters* 15, 5 (2018), 749–753.

Qingnan Zhou and Alec Jacobson. 2016. Thingi10k: A dataset of 10,000 3d-printing models. *arXiv preprint arXiv:1605.04797* (2016).

Daniel Zint, Nissim Maruani, Mael Rouxel-Labbé, and Pierre Alliez. 2023. Feature-Preserving Offset Mesh Generation from Topology-Adapted Octrees, Vol. 42. Wiley Online Library, e14906. Issue 5.



(a) Trilinear (b) B-Spline (c) RFTA (d) PoNQ (e) Ours (f) Ground Truth

Fig. 12. **Results I**. For two different models, we show the meshes generated via different approaches (from simple trilinear and B-spline interpolation, to RFTA, PoNQ, and ours) based on inputs of resolution 32^3, 64^3 and 128^3 respectively, to evaluate their quality at different scales.

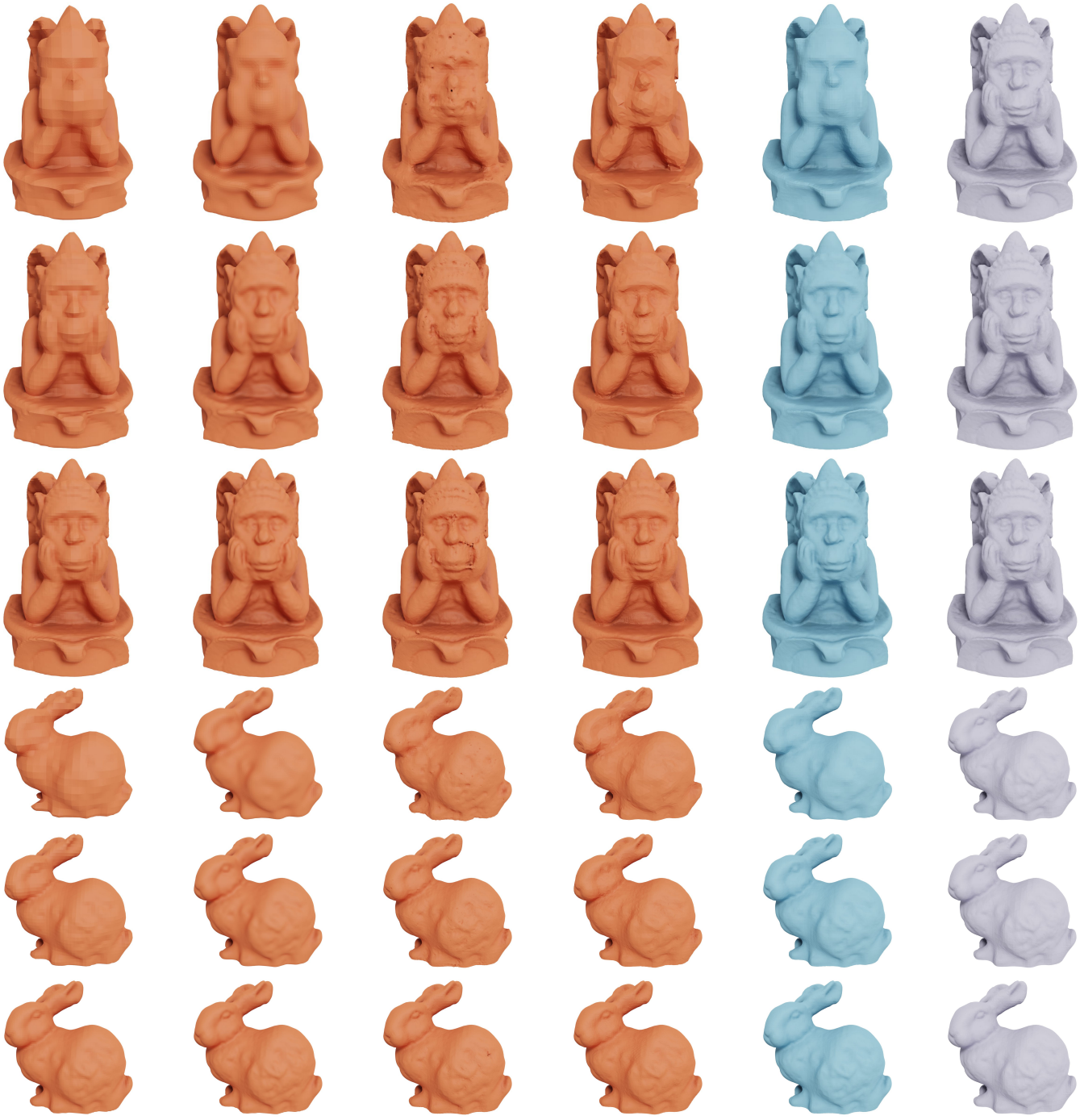


Fig. 13. **Results II**. For two different models, we show the meshes generated via different approaches (from simple trilinear and B-spline interpolation, to RFTA, PoNQ, and ours) based on inputs of resolution 32^3, 64^3 and 128^3 respectively, to evaluate their quality at different scales.

SuperSDF: Sparse SDF Super-Resolution for Surface Extraction
Supplemental Material

SAGAR PANWAR*, Inria, France
NISSIM MARUANI*, Inria, France
CÉLINE LOSCOS, L Research, France
MATHIEU DESBRUN, Inria – Ecole Polytechnique, France
PIERRE ALLIEZ, Inria, France

We provide additional details and tests in this supplementary material to further evaluate our SuperSDF approach.

1 Additional Experiments

We begin with experiments to further validate our method's robustness and performance.

1.1 Robustness to noise

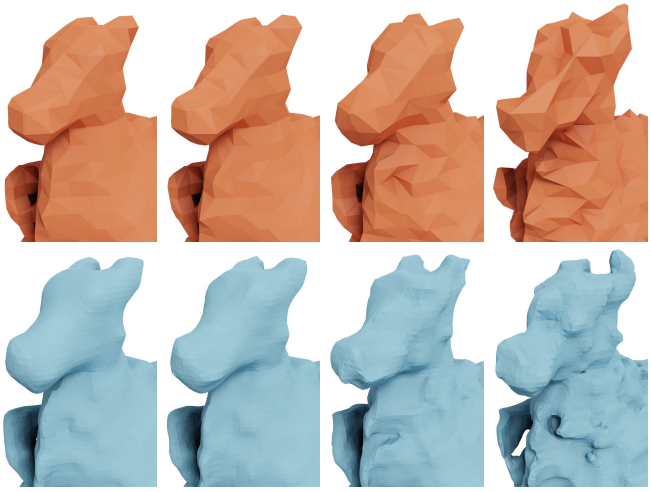
While our network was not trained on noisy SDF inputs, we found that, due to its reliance on convolutions, it acts as a natural low-pass filter — thus handling noise remarkably well. To quantify this property, we added a random uniform noise to the SDF values of models in Thingi30 [Maruani et al. 2023], with a magnitude equal to a percentage of the voxel's half-width, and found that our pipeline consistently outperforms Marching Cubes. We report qualitative results in Fig. 1 and Tab. 1.

Table 1. **F1 scores for varying levels of noise.** We report the F1 score on Thingi30 at 32^3 resolution, for a uniform noise with a magnitude defined by a percentage of the half-voxel width.

Method	0% Noise	20% Noise	50% Noise	100% Noise
MC	0.605	0.569	0.430	0.283
Ours+MC	**0.820**	**0.754**	**0.487**	**0.286**

1.2 Iterated Super-Resolution

We note that SuperSDF can be applied to its own outputs. Compared to the direct baseline ($32^3 \rightarrow 128^3$), iterated super-resolution ($32^3 \rightarrow 128^3 \rightarrow 512^3$) leads to small, yet consistent improvements on all metrics as evidenced by Tab. 2.



(a) 0% Noise (b) 20% Noise (c) 50% Noise (d) 100% Noise

Fig. 1. **Robustness to noise.** Compared to Marching Cubes (top), our method (bottom) still performs well when uniform noise is added to the input SDF, for a magnitude of up to a voxel's half-width.

Table 2. **Iterated Super-Resolution on Thingi30.**

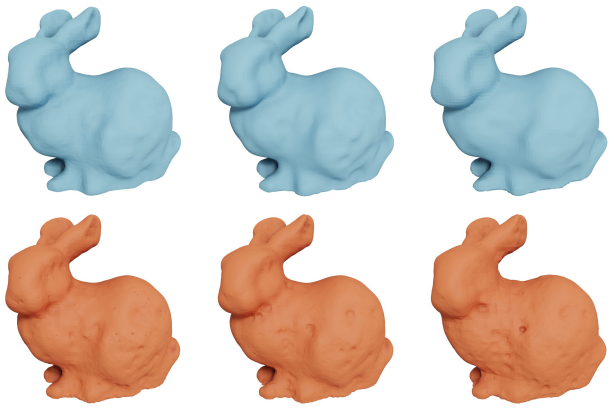
Method	Grid Size	CD ↓ ($\times 10^{-5}$)	F1 ↑	NC ↑
Ours+MC	$32^3 \rightarrow 128^3$	1.351	0.820	0.950
Ours+MC	$32^3 \rightarrow 128^3 \rightarrow 512^3$	**1.317**	**0.824**	**0.952**

1.3 Scaling

When reconstructing a shape from a signed distance field, the scale of the shape relative to the unit grid dictates the final mesh density and, consequently, the reconstruction quality. We chose to scale each object within the half unit ball like in [Chen and Zhang 2021; Maruani et al. 2024]. Reach For The Arcs (RFTA) [Sellán et al. 2024], on the other hand, uses a different scaling based on the object bounding box. We compared the results of RFTA for different scaling procedures in Fig. 2 and Tab. 3, and found that, while down-scaling the object obviously degrades performance, using a box scaling rather than a sphere scaling does not noticeably affect the quantitative or qualitative results. We stress that, for a given scaling, all methods have access to the exact same information, ensuring a fair and consistent comparison.

*Both authors contributed equally to this research.

Authors' Contact Information: Sagar Panwar, sagarpanwar249@gmail.com, Inria, Sophia-Antipolis, France; Nissim Maruani, nissim.maruani@inria.fr, Inria, Palaiseau, France; Céline Loscos, celine.loscos.pro@gmail.com, L Research, Sophi-Antipolis, France; Mathieu Desbrun, mathieu.desbrun@inria.fr, Inria – Ecole Polytechnique, Palaiseau, France; Pierre Alliez, pierre.alliez@inria.fr, Inria, Sophia-Antipolis, France.



(a) Ball scaling (.5) (b) Box scaling (.75) (c) Box scaling (.5)

Fig. 2. **RFTA and scaling.** Comparisons of various scaling procedures for SuperSDF (top row) and Reach For The Arcs [Sellán et al. 2024] (bottom row); see also statistics in Tab. 3.

Table 3. **Testing different box scalings of RFTA**. As expected, scaling down the object inside a fixed bounding box leads to worse predictions.

Scale	Method	Input Size	CD ↓ ($\times 10^{-5}$)	F1 ↑	NC ↑
0.25	RFTA	32^3	107.293	0.189	0.837
0.25	RFTA	64^3	40.498	0.295	0.869
0.25	RFTA	128^3	20.440	0.428	0.880
0.25	Ours	32^3	13.826	0.425	0.871
0.25	Ours	64^3	2.652	0.702	0.928
0.25	Ours	128^3	0.892	0.891	0.966
0.50	RFTA	32^3	14.943	0.467	0.903
0.50	RFTA	64^3	3.864	0.694	0.939
0.50	RFTA	128^3	1.898	0.818	0.953
0.50	Ours	32^3	2.653	0.701	0.929
0.50	Ours	64^3	0.895	0.891	0.966
0.50	Ours	128^3	0.656	0.936	0.984
0.75	RFTA	32^3	3.576	0.712	0.936
0.75	RFTA	64^3	1.322	0.866	0.962
0.75	RFTA	128^3	0.918	0.907	0.975
0.75	Ours	32^3	1.252	0.834	0.953
0.75	Ours	64^3	0.707	0.927	0.978
0.75	Ours	128^3	0.636	0.939	0.989

2 Evaluation Dataset

Similar to previous learning-based methods [Maruani et al. 2023, 2024], we use the publicly available Thingi30 dataset, corresponding to a subset [1] of the Thingi10K dataset, for evaluation in our paper. Additional examples of reconstruction of shapes with thin features are given in Fig. 4.

[1]Thingi30 contains the meshes from Thingi10K with ID 44234, 47984, 53159, 64444, 64764, 68380, 68381, 72870, 72960, 73075, 75496, 75655, 75656, 75662, 75665, 76277, 77245, 78671, 79241, 90889, 92763, 92880, 95444, 252119, 313444, 316358, 354371, 398259, 441708, and 527631.

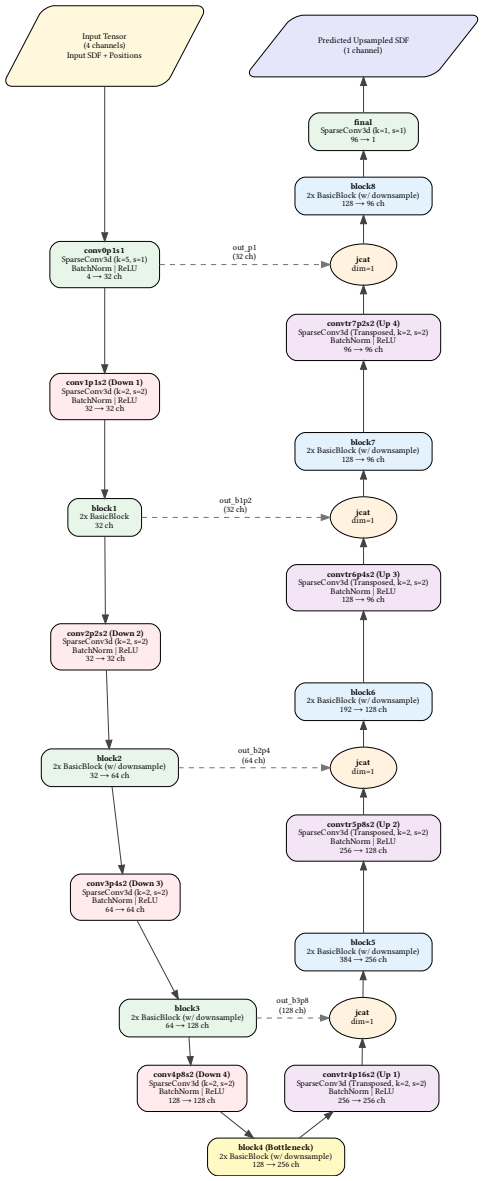


Fig. 3. Detailed architecture of our U-Net-based SuperSDF.

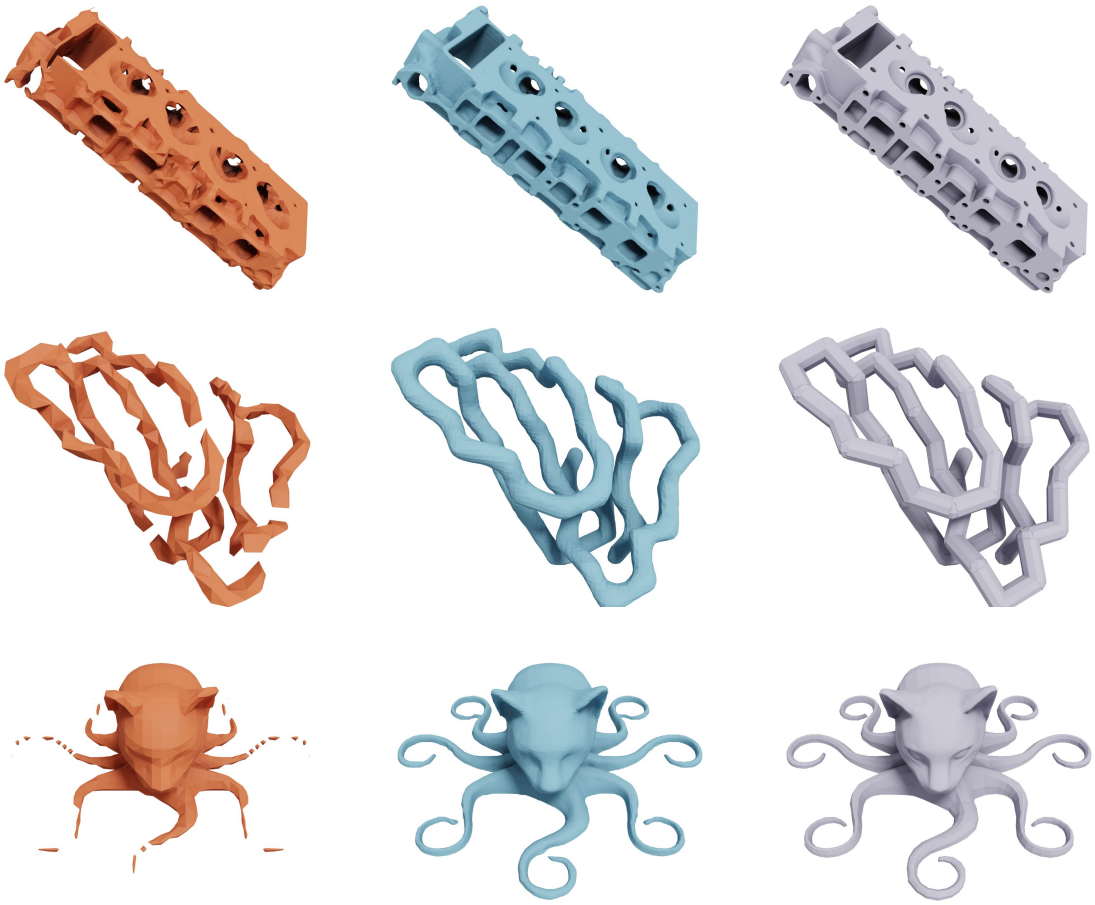
3 Architecture

Finally, we provide in Fig. 3 a thorough description of the architecture used in SuperSDF.

References

Zhiqin Chen and Hao Zhang. 2021. Neural marching cubes. *ACM Transactions on Graphics* 40, 6 (Dec. 2021), 1–15. doi:10.1145/3478513.3480518

Nissim Maruani, Roman Klokov, Maks Ovsjanikov, Pierre Alliez, and Mathieu Desbrun. 2023. VoroMesh: Learning Watertight Surface Meshes with Voronoi Diagrams. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. 14565–14574. https://openaccess.thecvf.com/content/ICCV2023/html/Maruani_VoroMesh_Learning_Watertight_Surface_Meshes_with_Voronoi_Diagrams_ICCV_2023_paper.html



(a) March. Cubes (b) Ours (c) Ground Truth

Fig. 4. **Reconstruction of fine details.** Evaluation of SuperSDF on three models of Thingi10 with thin structures.

Nissim Maruani, Maks Ovsjanikov, Pierre Alliez, and Mathieu Desbrun. 2024. PoNQ: A Neural QEM-Based Mesh Representation. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Seattle, WA, USA, 3647–3657. doi:10.1109/CVPR52733.2024.00350

Silvia Sellán, Yingying Ren, Christopher Batty, and Oded Stein. 2024. Reach for the Arcs: Reconstructing Surfaces from SDFs via Tangent Points. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers*. ACM, Denver CO USA, 1–12. doi:10.1145/3641519.3657419