
Inferring Neural Signed Distance Functions by Overfitting on Single Noisy Point Clouds through Finetuning Data-Driven based Priors

Anonymous Author(s)

Affiliation

Address

email

Abstract

1 It is important to estimate an accurate signed distance function (SDF) from a point
2 cloud in many computer vision applications. The latest methods learn neural SDFs
3 using either a data-driven based or an overfitting-based strategy. However, these two
4 kinds of methods are with either poor generalization or slow convergence, which
5 limits their capability under challenging scenarios like highly noisy point clouds.
6 To resolve this issue, we propose a method to prompt pros of both data-driven
7 based and overfitting-based methods for better generalization, faster inference, and
8 higher accuracy in learning neural SDFs. We introduce a novel statistical reasoning
9 algorithm in local regions which is able to finetune data-driven based priors without
10 signed distance supervision, clean point cloud, or point normals. This helps our
11 method start with a good initialization, and converge to a minimum in a much
12 faster way. Our numerical and visual comparisons with the stat-of-the-art methods
13 show our superiority over these methods in surface reconstruction and point cloud
14 denoising on widely used shape and scene benchmarks.

15

1 Introduction

16 It is an important task to estimate an implicit function from a point cloud in computer graphics,
17 computer vision, and robotics. An implicit function, such as a signed distance function (SDF),
18 describes a continuous 3D distance field to indicate the nearest distances to a surface at arbitrary
19 locations. Since point clouds are easy to obtain, they are widely used as an information source to
20 estimate SDFs, particularly without using normals that are not available for most scenarios. The
21 challenge for SDF estimation mainly comes from the difficulty of bridging the gap between the
22 discreteness of point clouds and the continuity of implicit functions.

23 Recent methods [59, 61, 28, 13, 92, 77, 55, 71] overcome this challenge using either a data-driven
24 based or an overfitting-based strategy. To map a point cloud to a signed distance field, the data-driven
25 based methods [57, 26, 34, 43, 78, 76, 21, 40, 89, 80] rely on a prior learned with signed distance
26 supervision from a large-scale dataset, while the overfitting-based methods [27, 1, 99, 2, 96, 4, 20,
27 48, 17, 85] do not need signed distance supervision and just use the point cloud to infer a signed
28 distance field. However, both of the two kinds of methods have pros and cons. The data-driven based
29 methods can do inference fast but suffers from the need of large-scale training samples and poor
30 generalization to instances that are unseen during training. Although the overfitting-based methods
31 have a better generalization ability and do not need the large-scale signed distance supervision, they
32 usually require a much longer time to converge during inference. The cons of these two kinds of
33 methods dramatically limit the performance of learning neural SDFs under challenging scenarios like
34 highly noisy point clouds. Therefore, beyond pursuing higher accuracy of SDFs, how to balance the
35 generalization ability and the convergence efficiency is also a significant issue.

36 To resolve this issue, we propose to learn an SDF from a point cloud by finetuning data-driven based
37 priors. Our key idea is to promote the advantages of both the data-driven based and the overfitting-
38 based strategy to pursue better generalization, faster inference, and higher accuracy. Our method
39 overfits a neural network on a single point cloud to approximate an SDF with a novel loss without
40 using signed distance supervision, clean point, or point normals, where the neural network was
41 pretrained as a data-driven based prior from large-scale signed distance supervision. With finetuning
42 priors, our method can generalize better on unseen instances than the data-driven based methods, and
43 also converge much more accurate SDFs in a much faster way than the overfitting-based methods.
44 Moreover, our novel loss for finetuning the data-driven based prior can conduct a statistical reasoning
45 in a local region which can recover more accurate and sharper underlying surface from noisy points.
46 We report numerical and visual comparisons with the stat-of-the-art methods and show our superiority
47 over these methods in surface reconstruction and point cloud denoising on widely used shape and
48 scene benchmarks. Our contributions are listed below,

- 49 • We introduce a method which is capable of finetuning a data-driven based prior by minimiz-
50 ing an overfitting-based loss without signed distance supervision, leading to neural SDFs
51 with better generalization, faster inference, and higher accuracy.
52 • The proposed overfitting-based loss can conduct a novel statistical reasoning in local regions,
53 which improves the accuracy of neural SDFs inferred from noisy point clouds.
54 • Our method produces the state-of-the-art results in surface reconstruction and point cloud
55 denoising on the widely used benchmarks.

56 2 Related Works

57 Learning implicit functions have achieved promising performance in various tasks [59, 61, 28, 13, 92,
58 77, 55, 71, 29, 30]. We can learn neural implicit representations from different supervision including
59 3D supervision [58, 66, 56, 16], multi-view [75, 42, 36, 98, 44, 91, 60, 39, 95, 94, 24, 83, 97, 86, 81,
60 82], and point clouds [89, 41, 57, 26]. We review works the existing methods related to point clouds
61 below.

62 2.1 Data-driven based Methods

63 In 3D supervision, many techniques utilize a data-driven approach to develop priors, applying
64 these to infer implicit models for unseen point clouds. Some strategies focus on acquiring global
65 priors [57, 26, 34, 43, 78, 76, 21, 40] at the shape level, whereas others aim to boost the generalization
66 of these priors by learning local priors [89, 80, 10, 35, 6, 49] at the component or patch level. These
67 learned priors facilitate the marching cubes algorithm [45] to reconstruct surfaces from implicit fields.
68 The effectiveness of these methods often rely on extensive datasets, but they may not generalize well
69 when facing with unseen point clouds that significantly deviate in geometry from training samples.

70 2.2 Overfitting-based Methods

71 In an effort to enhance generalization, some methods concentrate on precisely fitting neural networks
72 to single point clouds. These methods incorporate innovative constraints [27, 1, 99, 2, 96, 4, 20],
73 utilize gradients [48, 17, 85], employ differentiable Poisson solvers [67], or apply specially tailored
74 priors [49, 52] to learn either signed [48, 27, 1, 99, 2, 14, 54, 12] or unsigned distance functions [17,
75 101, 100]. Despite achieving significant advances, these approaches typically require clean point
76 clouds to accurately determine distance or occupancy fields around the point clouds.

77 2.3 Learning from Noisy Point Clouds

78 The key to accurately reconstructing surfaces on noisy point clouds is to minimize the effect of noise
79 on inferring implicit functions. PointCleanNet [70] was developed to filter out noise from point clouds
80 through a data-driven approach. GPDNet [69] incorporated graph convolution based on dynamically
81 generated neighborhood graphs to enhance noise reduction. Some other methods leveraged point
82 cloud convolution [6], alternating latent topology [87], semi-supervised strategy [103, 18], dual
83 and integrated latent [73], or neural kernel field [88, 33] to reduce noise from point clouds. On the
84 unsupervised front, TotalDenoising [9] adopts principles similar to Noise2Noise [38], utilizing a

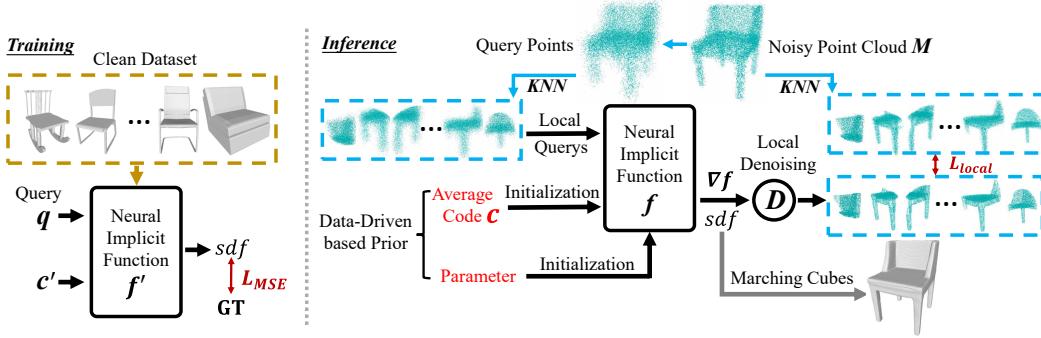


Figure 1: The overview of our method. We learn the data-driven based prior by sampling queries q to learn a neural implicit function f with a condition c under a clean dataset. During inference, we employ a novel statistical reasoning algorithm to infer a neural SDF f' for a noisy point cloud M with learned prior (average code and learned parameter).

85 spatial prior suitable for unordered point clouds. DiGS [3] employs a soft constraint for unoriented
 86 point clouds. Noise2NoiseMapping [50] leverage statistical reasoning among multiple noisy point
 87 clouds with specially designed losses. Some methods using downsample-upsample frameworks [46],
 88 gradient fields [47, 8, 15, 65, 62], convolution-free intrinsic occupancy network [64], intra-shape
 89 regularization [63], eikonal equation [93, 22], neural Galerkin [32] and neural splines [90] have been
 90 implemented to further diminish noise in point clouds. Our method falls in this category, but we aim
 91 to promote the advantages of both the data-driven based and the overfitting-based strategy to pursue
 92 better generalization, faster inference, and higher accuracy.

93 3 Method

94 **Overview.** As illustrated in Fig. 1, we aim to learn a neural SDF f from a single point cloud with
 95 noise M . Our method includes two stages, one is to learn a prior f' in a data-driven manner, the other
 96 is to infer a neural SDF f on a specific point cloud with noises M . At the first stage, we learn a prior
 97 by training a neural SDF with embeddings c'_j indicating training shapes using ground truth signed
 98 distances. At the second stage, we finetune the learned prior f' and learn an embedding c for an
 99 unseen point cloud M with noises using our proposed local noise to noise mapping. With the learned
 100 f , we can use the marching cubes algorithm [45] to extract the zero-level set as the mesh surface.

101 **Neural Signed Distance Function.** We leverage an SDF f to represent the geometry of a shape.
 102 An SDF f is an implicit function that can predict a signed distance s for an arbitrary location q , i.e.,
 103 $s = f(q)$. The latest methods usually train a neural network to approximate an SDF from signed
 104 distance supervision or infer an SDF from 3D point clouds or multi-view images. A level set is an
 105 iso-surface formed by the points with the same signed distance value. For instance, zero-level set is a
 106 special level set, which is formed by points with a signed distance of 0. On the zero-level set, the
 107 gradient $\nabla f(q)$ of the SDF f at an arbitrary location q is also the surface normal at q .

108 **Data-driven Based Prior.** As shown in Fig. 1, we employ a auto-decoding network similar to
 109 DeepSDF [66] for learning a prior and the neural SDF for single point clouds with noise. We
 110 employ a data-driven strategy to learn a prior f' first. Specifically, we learn a neural SDF f' with an
 111 embedding c as a condition of queries. For each shape, we sample queries q around the shape, and
 112 establish the signed distance supervision by recording the signed distance s to the ground truth mesh.
 113 Thus, we learn the prior by minimizing the signed distance prediction errors,

$$\min_{f, \{c_i\}} \sum_{i=1}^I \sum_{j=1}^J \|s_i^j - f'(q_j, c'_i)\|_2^2 + \alpha \sum_{i=1}^I \|c'_i\|_2^2, \quad (1)$$

114 where c'_i is a learnable condition for the i -th training shape, q_j is the j -th query that is randomly
 115 sampled around the i -th shape, and s_i^j is the ground truth signed distance. We also add a regularization
 116 term on the learned embeddings c'_i , and α is the balance weight.

117 **Signed Distance Inference.** With the learned prior f' , we infer a neural SDF f for a single point
 118 cloud with noises M . We do not require ground truth signed distances, clean point clouds, or even

119 point normal during the inference of f . Specifically, we infer f by finetuning parameters of f' with a
 120 learnable embedding \mathbf{c} indicating the single point cloud with noises. The finetuning relies on a novel
 121 statistical reasoning algorithm on local regions.

122 The advantage of our method lies in the capability of conducting the statistical reasoning in local
 123 regions. Comparing to the global reasoning method [50], our method can not only infer more accurate
 124 geometry but also significantly improve the efficiency. Our method starts from randomly sampling a
 125 local region m_n on the shape M . We randomly select one point on M , and set up its K nearest noisy
 126 points as a local region m_n . Then, we randomly sample U queries $\{\bar{q}_u\}_{u=1}^U$ around the K noisy
 127 points, and also randomly select U noisy points $\{p_v\}_{v=1}^U$ out of the K noisy points for statistically
 128 reasoning the surface in one iteration.

129 Our key idea of inferring a neural SDF f is to estimate a mean zero-level set that is consistent to all
 130 points in a local region m_n . To this end, we use the U sampled queries to represent the zero-level
 131 set in this area, and minimize its distance to the U noisy points in each iteration. Statistically, the
 132 expectation of the zero-level set should have the minimum distance to all the noisy point splitting.

133 Specifically, we first project the U sampled queries onto the zero-level set using a differentiable
 134 pulling operation [48]. For each query \bar{q}_u , we can pull it onto the zero-level set below,

$$\bar{q}'_u = \bar{q}_u - s * \nabla f(\bar{q}_u, \mathbf{c}) / |\nabla f(\bar{q}_u, \mathbf{c})|, \quad (2)$$

135 where \bar{q}'_u is the projection of \bar{q}_u on the zero-level set, $s = f(\bar{q}_u, \mathbf{c})$, $\nabla f(\bar{q}_u, \mathbf{c})$ is the gradient of f at
 136 the location \bar{q}_u , and \mathbf{c} is the learnable embedding that represents the noisy point cloud M .

137 With the pulling operation, we can transform a set of queries $\{\bar{q}_u\}_{u=1}^U$ into their projections $\{\bar{q}'_u\}$ on
 138 the zero-level set. $\{\bar{q}'_u\}$ forms a coarse zero-level set estimation, we expect this zero-level set can be
 139 consistent to various subsets of points from the noisy local regions m_n . Thus, we minimize the errors
 140 between the $\{\bar{q}'_u\}_{u=1}^U$ and a subset of points $\{p_v\}_{v=1}^U$ that are randomly selected from m_n ,

$$\min_{f, \mathbf{c}} \mathbb{E}_{m_n \sim M, \bar{q}_u \sim m_n, p_v \sim m_n} EMD(\{\bar{q}'_u\}, \{p_v\}) + \beta \|\mathbf{c}\|_2^2, \quad (3)$$

141 where we conduct this optimization by finetuning the prior f' and learning the embedding \mathbf{c} repre-
 142 senting the noisy point cloud. The expectation is over the local regions m_n that randomly sampled
 143 from the noisy point cloud M , and the subset patch p_v randomly sampled from each m_n . We follow
 144 the method [50] to use the EMD to evaluate the distance between the two sets of points, which leads
 145 the neural SDF f' to converge.

146 **Initialization.** f has the same network architecture to the prior network f' . We learn f with the
 147 parameters of f' as the initialization, which is the prior that we learned. For the embedding \mathbf{c} that
 148 represents M , we initialize \mathbf{c} as the center of the embedding space learned for priors in Eq. 1, i.e.,
 149 $\mathbf{c} = 1/I \sum_{i=1}^I \mathbf{c}'_i$. These initialization is important for the accuracy and efficiency on single noisy
 150 point clouds. This finetuning of parameters of f' also shows advantages over the auto-decoding [66]
 151 in terms of generalization and efficiency. We will justify these advantages in our experiments.

152 **Implementation Details.** We randomly select one point from noisy point cloud M as a center, and
 153 select its $K = 1000$ nearest points as a local region m_n . We also randomly sample $U = 1000$ queries
 154 around K noisy points for statistically reasoning. Specifically, We adopt a method introduced by
 155 NeuralPull [48] to sample queries around each point from the noisy point cloud. We use a Gaussian
 156 distribution centered at each point and set the standard deviation as the distance to the 51th nearest
 157 neighbor in the point cloud. We run the marching cubes for surface reconstruction at a resolution of
 158 256 for shapes, and 512 for large-scale scenes.

159 The length of the embedding \mathbf{c} or \mathbf{c}' is set to 256. We use Adam optimizer for learning a neural
 160 implicit network, which is an auto-decoder architecture similar to DeepSDF [66]. For training, we use
 161 an initial embedding learning rate of 0.0005 for updating embeddings and an auto-decoder learning
 162 rate of 0.001 for optimizing the prior network. Both learning rates are decreased by 0.5 for every
 163 500 epochs. We train the network prior network f' for 2000 epochs. For inference, we finetune the
 164 network f for each noisy point cloud in 4000 iterations with a learning rate of 0.0001.

Metrics	PSR [37]	PSG [23]	R2N2 [19]	COcc [68]	SAP [67]	OCNN [84]	IMLS [43]	POCO [7]	ALTO [87]	N2NM [50]	Ours
CD_{L1}	0.299	0.147	0.173	0.044	0.034	0.067	0.031	0.030	0.028	0.026	0.023
NC	0.772	-	0.715	0.938	0.944	0.932	0.944	0.950	0.955	0.962	0.973
F-Score	0.612	0.259	0.400	0.942	0.975	0.800	0.983	0.984	0.985	0.991	0.992

Table 1: Numerical Comparisons on ShapeNet dataset in terms of $CD_{L1} \times 10$, NC and F-Score.

165 4 Experiments and Analysis

166 We compare our method with the latest methods in terms of numerical and visual results on synthetic
167 point clouds and real scans in surface reconstruction.

168 **Datasets and Metric.** We use 8 datasets including shapes and scenes to in the evaluations. For shapes,
169 we conduct experiments under five datasets including ShapeNet [11], ABC [21], FAMOUS [21], Sur-
170 face Reconstruction Benchmark (SRB) [89] and D-FAUST [5]. For scenes, we conduct experiments
171 under three real scan datasets including 3D Scene [102], KITTI [25] and Paris-rue-Madame [72].
172 We leverage L1 Chamfer Distance (CD_{L1}), L2 Chamfer Distance (CD_{L2}) to evaluate the error
173 between the reconstructed surface and ground truth. We also use Normal Consistency (NC) [56] and
174 F-Score [79] with a threshold of 1% to evaluate the normal accuracy of the reconstructed surface. In
175 the ablation study, we also report time consumption to highlight the superiority of our data-driven
176 based prior. For KITTI and Paris-rue-Madame datasets, due to their lack of ground truth meshes, we
177 only report visual comparisons.

178 4.1 Surface Reconstruction for Shapes

179 Evaluation on ShapeNet.

180 We first report our results
181 on shapes from ShapeNet.
182 We report evaluations by
183 comparing our method with
184 the latest prior-based and
185 overfitting-based methods
186 in Tab 1. For prior-based
187 methods, we compare our
188 method with PSG [23],
189 R2N2 [19], COcc [68],
190 OCNN [84], IMLS [43],
191 POCO [7], and ALTO [87].
192 All of these methods are
193 pretrained to learn priors us-
194 ing shapes in training set
195 of ShapeNet. We also fol-
196 low these methods to use
197 the same set of training shapes to learn our prior. For overfitting-based methods, we compare
198 our method with PSR [37], SAP [67], and N2NM [50]. These methods did not need to learn a prior,
199 and have the ability of inferring neural implicit functions on each shape in the testing set. We also
200 follow these methods and report our results by finetuning our prior through overfitting on each testing
201 shape. All the shapes for testing are corrupted with noises with a variance of 0.005.

202 The comparisons in Tab. 1 indicate that our method can
203 infer much more accurate neural implicit functions than
204 the prior-based methods. The improvement comes from
205 the ability of conducting test time optimization with the
206 learned prior and inferring signed distances using the local
207 noise to noise mapping. Moreover, our local statistical reasoning not only achieves better ability of
208 recovering geometry from noisy points than overfitting-based methods but also significantly reduces
209 the time complexity during the test time overfitting procedure with our prior. Different from prior-
210 based methods, our ability of conducting test-time optimization with our local statistical reasoning
211 loss can significantly improve the generalization ability on unseen shapes. Tab. 2 shows that our
212 method can infer neural implicit functions on single shapes much faster than the overfitting-based
213 methods. We also demonstrate our advantages in visual comparisons in Fig. 2.

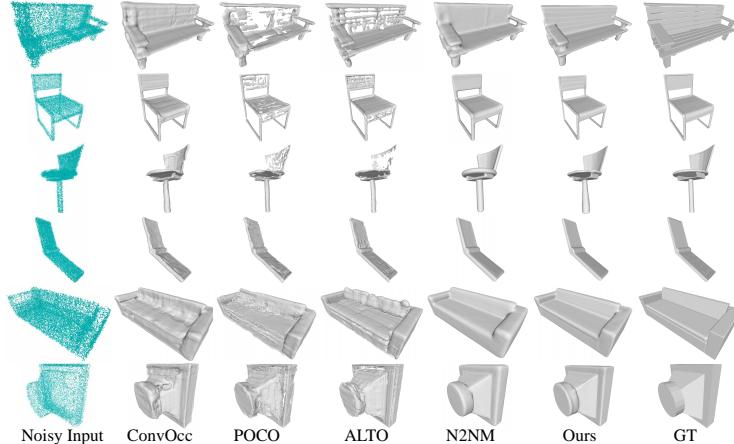


Figure 2: Comparison in surface reconstruction on ShapeNet. More visual results are provided in the appendix.

Metrics	SAP [67]	N2NM [50]	Ours
Time	14 min	46 min	5 min

Table 2: Time consumption on ShapeNet dataset with overfitting-based methods.

Dataset	PSR [37]	P2S [21]	COcc [68]	NP [48]	IMLS [43]	PCP [53]	POCO [7]	OnSurf [51]	N2NM [50]	Ours
ABC var	3.29	2.14	0.89	0.72	0.57	0.49	2.01	3.52	0.113	0.096
ABC max	3.89	2.76	1.45	1.24	0.68	0.57	2.50	4.30	0.139	0.113

Table 3: Numerical Comparisons on ABC dataset in terms of $CD_{L2} \times 100$.

Metrics	IGR [27]	Point2Mesh [31]	PSR [37]	SIREN [74]	GP [15]	ALTO [87]	Steik [93]	SAP [67]	NKSR [33]	N2NM [50]	Ours
CD_{L1}	0.178	0.116	0.232	0.123	0.086	0.089	0.079	0.076	0.069	0.067	0.055
F-Score	0.755	0.648	0.735	0.677	0.766	0.772	0.822	0.830	0.829	0.835	0.860

Table 4: Numerical Comparisons on SRB dataset in terms of $CD_{L1} \times 10$ and F-Score.

214 **Evaluation on ABC.** We also
215 report our evaluations on ABC
216 dataset in Tab. 3. We learn
217 priors from shapes in training
218 set, and finetune this prior for
219 each single shape in the testing
220 set. The numerical comparisons
221 are conducted on the testing
222 set of ABC dataset released
223 by P2S [21]. It includes two ver-
224 sions with different noise levels.
225 Similarly, we also report com-
226 parisons with prior-based methods and overfitting-based methods. With our local noise to noise
227 mapping, we achieve the best performance over all baselines. Compared to prior-based methods, such
228 as P2S [21], COcc [68], and POCO [7], our loss can infer more accurate geometry from noisy points
229 during the test time overfitting procedure. Also, the ability of finetuning the prior can also provide a
230 coarse estimation and a good start for inferring neural implicit from single noisy points. Besides the
231 accuracy, we also observe improvements on efficiency. Fig. 3 demonstrates the improvements over
232 the baselines in terms of surface completeness and edge sharpness.

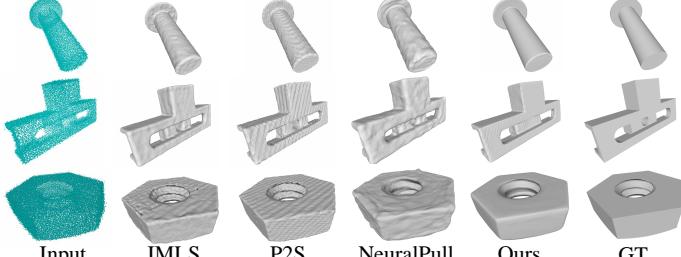


Figure 3: Comparison in surface reconstruction on ABC. More visual results are provided in the appendix.

233 **Evaluation on SRB.** We report previous
234 experiments using man-made objects in
235 ShapeNet and ABC dataset. We also re-
236 port our results on real scans on SRB
237 dataset [89]. Since there is no training sam-
238 ples on SRB, we use the prior learned from
239 the ShapeNet as the prior for real scans. Al-
240 though the shapes in ShapeNet are not sim-
241 ilar to shapes in SRB, we found the prior
242 can also work well with the scans on SRB.
243 Different from the man-made objects, real
244 scans have unknown noises. We report the
245 evaluations with the prior-based and overfitting-based methods in Tab. 4 and Fig. 4. The comparisons
246 show that our method achieves the best performance in inferring neural implicit functions. Under the
247 same experimental settings, our method is able to can infer more accurate geometry details with our
248 local noise to noise mapping.

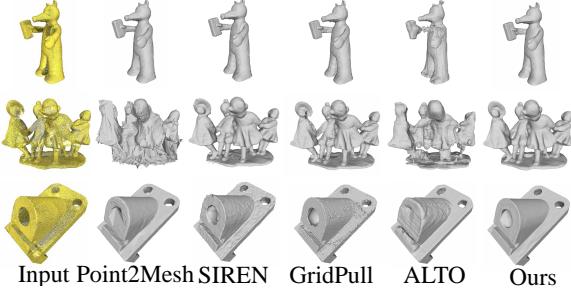


Figure 4: Comparison in surface reconstruction on SRB. More visual results are provided in the appendix.

249 **Evaluation on FAMOUS.**

250 We report evaluations
251 on more complex shapes
252 on FAMOUS dataset.
253 Similar to SRB, we also
254 use the prior learned from
255 ShapeNet. We also evaluate
256 the performance on two
257 kinds of noises in Tab. 5.
258 We can see that our method
259 can recover more geometry
260 details and achieve higher
261 accuracy and smoother
262 surfaces. We also report
263 visual comparisons in
264 Fig. 5, which also highlights our improvements in terms of accuracy, smoothness, completeness, and
265 recovered sharp edges.

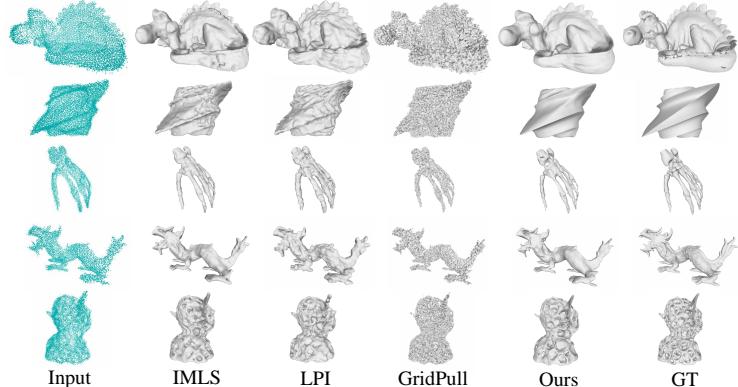


Figure 5: Comparison in surface reconstruction on FAMOUS. More visual results are provided in the appendix.

Dataset	PSR [37]	NP [48]	IMLS [43]	LPI [14]	PCP [53]	POCO [7]	OnSurf [51]	GP [15]	N2NM [50]	Ours
F-var	1.80	0.28	0.80	0.19	0.07	1.50	0.59	0.13	0.033	0.029
F-max	3.41	0.31	0.39	0.26	0.30	2.75	3.64	0.21	0.117	0.105

Table 5: Numerical Comparisons on FAMOUS dataset in terms of $CD_{L2} \times 100$.

Metrics	IGR [27]	Point2Mesh [31]	PSR [37]	SAP [67]	N2NM [50]	Ours
$CD_{L1} \times 10$	0.235	0.071	0.044	0.043	0.037	0.034
F-Score	0.805	0.855	0.966	0.966	0.966	0.973
NC	0.911	0.905	0.965	0.959	0.970	0.968

Table 6: Accuracy of reconstruction on D-FAUST dataset in terms of CD_{L1} , NC and F-Score.

266 **Evaluation on D-FAUST.** Finally, we
 267 report our results on non-rigid shapes,
 268 i.e., humans. Different from rigid
 269 shapes in the previous experiments,
 270 humans are with more complex poses.
 271 We learn prior from the training set,
 272 and finetuning the prior on unseen hu-
 273 mans with different poses. We mainly
 274 compare our method with overfitting-
 275 based methods in Tab. 6. We can see
 276 that our method achieves the best per-
 277 formance in CD, F-Score, and com-
 278 parable performance to N2NM [50]
 279 but with faster inference speed. We
 280 further show the visual comparison in
 281 Fig. 6. We can see that our method
 282 can recover more accurate geometry
 283 and poses.

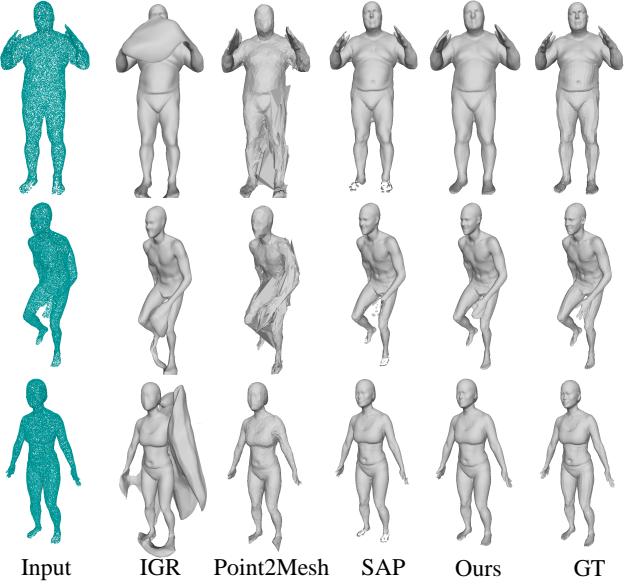


Figure 6: Comparison in surface reconstruction on D-FAUST. More visual results are provided in the appendix.

284 4.2 Surface 285 Reconstruction for Scenes

286 Since we have a limited number of
 287 scenes for training, we use the prior learned from ShapeNet as the pretrained prior in our experiments
 288 on scenes.

289 **Evaluation on 3D Scene.** We further evaluate our method in surface reconstruction for scenes in
 290 3D Scene [102]. We follow previous methods LIG [35] to randomly sample 1000 points per m^2 .
 291 We compare our method with the latest methods including COcc [68] and LIG [35], DeepLS [10],
 292 NeuralPull (NP) [48] and Noise2NoiseMapping (N2NM) [50]. For prior-based methods COcc [68]
 293 and LIG [35], we leverage their released pretrained models to produce the results, and we also provide
 294 them with the ground truth point normals. For overfitting-based methods DeepLS [10], NP [48] and
 295 N2NM [50], we overfit them to produce results with the same noisy point clouds. We follow LIG [35]
 296 to report CD_{L1} , CD_{L2} and NC for evaluation. We report the comparisons in Tab. 15. The results
 297 demonstrate that our method outperforms both kinds of methods with learned priors such as LIG [35]
 298 and overfitting-based N2NM [50]. The visual comparisons in Fig. 7 show that our method can reveal
 299 more geometry details on real scans, which justifies our capability of handling noise in point clouds.

300 **Evaluation on KITTI.** Following GridPull [15], we further evaluate our method on KITTI [25]
 301 odometry dataset (Sequence 00, frame 3000 to 4000), which contains about 13.8 million points,
 302 which are split into 15 chunks. We reconstruct each of them and concatenate them together for

Metrics	COcc [68]	LIG [35]	DeepLS [10]	NP [48]	N2NM [50]	Ours
$CD_{L2} \times 1000$	14.10	6.190	1.607	2.115	0.507	0.389
CD_{L1}	0.052	0.048	0.025	0.034	0.019	0.016
NC	0.908	0.849	0.915	0.900	0.929	0.942

Table 7: Numerical Comparisons on 3D Scene dataset in terms of CD_{L1} , CD_{L2} and NC. Detailed
 comparisons for each scene are provided in the appendix.

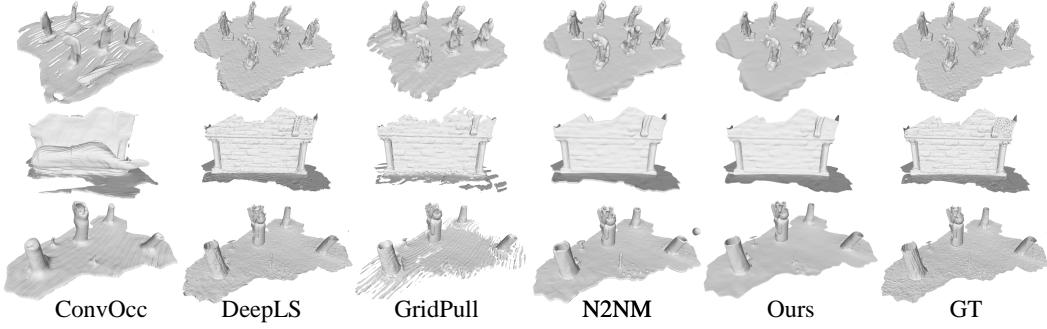


Figure 7: Comparison in surface reconstruction on 3D Scene.

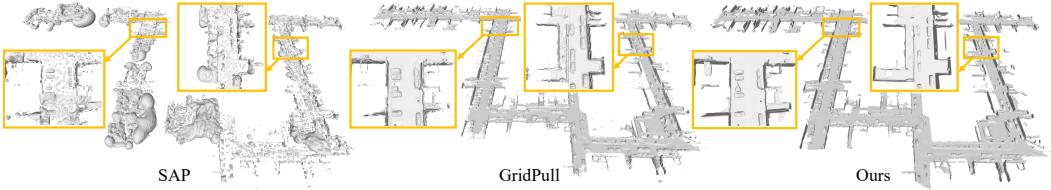


Figure 8: Comparison in surface reconstruction on KITTI.

303 visualization. We compare our method with the latest methods SAP [67] and GridPull [15]. As shown
304 in Fig. 8, our method is robust to noise in real scans, successfully generalizes to large-scale scenes,
305 and achieves visual-appealing reconstructions with more details.

306 **Evaluation on Paris-rue-Madame.** Following N2NM [50], we further evaluate our method on Paris-
307 rue-Madame [72], which contains much noise. We split the $10M$ points into 50 chunks each of which
308 is used to learn a neural implicit function. We compare our method with LIG [35] and N2NM [50].
309 For LIG [35], we produce the results for each chunk with released pretrained models. For N2NM [50],
310 we overfit all chunks until convergence. As shown in Fig. 9, we achieve better performance over
311 LIG [35] and N2NM [50] in large-scale surface reconstruction. The visual comparisons highlight our
312 advantages in reconstructing complete and detailed surfaces from noisy scene point clouds.

313 4.3 Ablation Studies

314 We conduct ablation studies on the ABC dataset [21] to justify each module of our method.

315 **Embedding Size.** We evaluate our performance on different
316 sizes of embedding c . We try several sizes $\{128, 256, 512\}$ to
317 infer the signed distance functions from a noisy point cloud.
318 The numerical comparison in Tab. 8 shows that the optimal
319 result is obtained with a size of 256. Deviations from this value, either longer or shorter dimensions,
320 leads to worse results with the current number of training samples.

321 **Prior.** We conduct experiments to ex-
322 plore the importance of data-driven
323 based prior. We first replace our
324 learned embedding c and parameter
325 with randomly initialized embedding and parameter, or only replace c with randomly initialized
326 embedding. As shown in Tab. 9, The degenerated result of “Without Prior” and “Without Embed”
327 indicates that directly inferring implicit functions without our prior or learned embedding makes it
328 difficult to accurately learn the surfaces of the noisy point clouds, and also slows the convergence.
329 Then we fix the learned parameters and only optimize the embedding c , similar to an auto-decoder.
330 The results also get worse, as shown in “Fixed Param”.
331

332 **Local Region Splitting.** We further val-
333 idate the effectiveness of local region split-
334 ting strategies. We employ three different
335 splitting strategies in Tab. 10. We first
336 split the whole space where the noisy point cloud is located uniformly into multiple voxel blocks, as
337 shown by the result of “Voxel”. The severely degenerated results indicate that this splitting strategy is
338 even worse than the global method N2NM [50], as it results in many empty voxel blocks. Then we
randomly select a point from the noisy point cloud as a center to sample all points within a radius of

Metric	128	256	512
$CD_{L2} \times 100$	0.102	0.096	0.114

Table 8: Effect of the embedding size.

Metric	Without Prior	Without Embed	Fixed Param	With Prior
$CD_{L2} \times 100$	0.108	0.103	0.144	0.096
Time	1h	12min	30min	8 min

Table 9: Effect of the prior.

Metric	Voxel	Sphere (Fixed Size)	Sphere (KNN)
$CD_{L2} \times 100$	0.314	0.101	0.096

Table 10: Effect of splitting strategies.

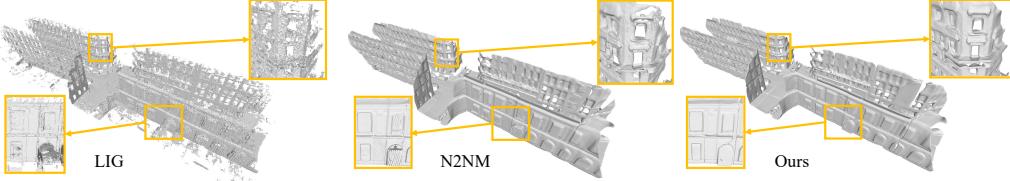


Figure 9: Comparison in surface reconstruction on Paris-rue-Madame.

339 0.1 as a local region. The result of “Sphere (Fixed Size)” slightly degenerates due to some of the
 340 spheres containing too few points. In contrast, our splitting strategy, as shown by the result of “Sphere
 341 (KNN)”, ensures that each local region has enough points to help achieve superior performance.

342 **Global and Local.** With our learned prior, we compare our
 343 performance in global and local mappings with finetuning the
 344 priors. We report results obtained with the local noise to noise
 345 mapping or the global one during the finetuning. As shown
 346 in Tab. 11, the numerical comparison shows that the global
 347 mapping struggles to infer local details from noisy point clouds. Moreover, our local prior also
 348 converges faster than the global statistical reasoning.

349 **Local Region Size.** We further validate the effec-
 350 tiveness of local region sizes (points number in a
 351 local region) in Tab. 12. We use different local re-
 352 gion sizes including {500, 1000, 3000, 5000}. The
 353 results show that 1000 is the best, both excessively large and small values of local region size lead to
 354 a decline in performance.

355 **Noise Level.** We report our performance on point clouds
 356 with different levels of noise. As shown in Tab. 13, the
 357 CD_{L2} comparison shows that our results slightly degenerate
 358 with max and extreme noise. Compared to N2NM [50], our
 359 method is more robust to noises.

360 **Sparsity.** We report the effect of the sparsity of noisy point clouds.
 361 We downsample the noisy point clouds to 50% and 25% of their
 362 original size to validate the effect of sparsity. The CD_{L2} results
 363 of in Tab. 14 indicate that our method can handle sparsity in noisy
 364 point clouds better than N2NM [50]. Since our data-driven based
 365 prior can help to learn a more complete surface and reduce the impacts brought by the sparsity.

366 **Optimization.** We visualize the optimization process in Fig. 14. We reconstruct meshes using the
 367 neural SDF f learned in different iterations. We see that the shape is updated progressively to the
 368 ground truth shapes.

Metric	Global	Local
$CD_{L2} \times 100$	0.106	0.096
Time	21 min	8 min

Table 11: Effect of local mapping.

Metric	500	1000	3000	5000
$CD_{L2} \times 100$	0.102	0.096	0.111	0.114

Table 12: Effect of local region size.

Method	middle	max	extreme
N2NM [50]	0.113	0.139	0.156
Ours	0.096	0.113	0.125

Table 13: Effect of noise level.

Method	25%	50%	100%
N2NM [50]	0.154	0.133	0.113
Ours	0.121	0.107	0.096

Table 14: Effect of sparsity.

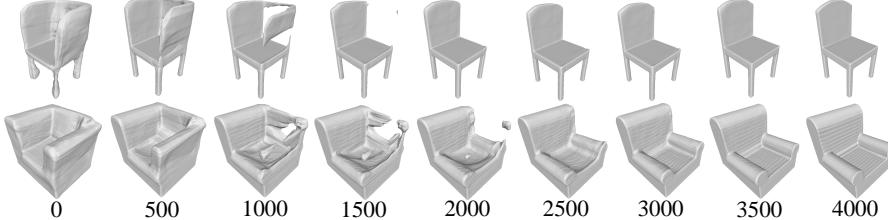


Figure 10: Optimization during inference.

369 5 Conclusion

370 We propose a method to resolve the key problem in learning SDFs for a single noisy point cloud.
 371 Our method can effectively use a data-driven based prior as an initialization, and infer a neural SDF
 372 by overfitting on a single noisy point cloud. The novel statistical reasoning successfully infers an
 373 accurate and smooth signed distance field around the single noisy point cloud with the data-driven
 374 based prior. By finetuning data-driven based priors with statistical reasoning, our method significantly
 375 improves the robustness, the scalability, the efficiency, and the accuracy in inferring SDFs from single
 376 point clouds. Our experimental results and ablations studies show our superiority and justify the
 377 effectiveness of the proposed modules.

378 **References**

- 379 [1] Matan Atzmon and Yaron Lipman. SAL: Sign agnostic learning of shapes from raw data. In *IEEE
380 Conference on Computer Vision and Pattern Recognition*, 2020.
- 381 [2] Matan Atzmon and yaron Lipman. SALD: sign agnostic learning with derivatives. In *International
382 Conference on Learning Representations*, 2021.
- 383 [3] Yizhak Ben-Shabat, Chamin Hewa Koneputugodage, and Stephen Gould. DiGS: Divergence guided
384 shape implicit neural representation for unoriented point clouds. In *IEEE Conference on Computer Vision
385 and Pattern Recognition*, 2022.
- 386 [4] Yizhak Ben-Shabat, Chamin Hewa Koneputugodage, and Stephen Gould. DiGS : Divergence guided
387 shape implicit neural representation for unoriented point clouds. *CoRR*, abs/2106.10811, 2021.
- 388 [5] Federica Bogo, Javier Romero, Gerard Pons-Moll, and Michael J. Black. Dynamic FAUST: Registering
389 human bodies in motion. In *IEEE Computer Vision and Pattern Recognition*, 2017.
- 390 [6] Alexandre Boulch and Renaud Marlet. POCO: Point convolution for surface reconstruction. In *IEEE
391 Conference on Computer Vision and Pattern Recognition*, pages 6302–6314, 2022.
- 392 [7] Alexandre Boulch and Renaud Marlet. Poco: Point convolution for surface reconstruction. In *IEEE
393 Conference on Computer Vision and Pattern Recognition*, 2022.
- 394 [8] Ruojin Cai, Guandao Yang, Hadar Averbuch-Elor, Zekun Hao, Serge Belongie, Noah Snavely, and
395 Bharath Hariharan. Learning gradient fields for shape generation. In *European Conference on Computer
396 Vision*, 2020.
- 397 [9] Pedro Hermosilla Casajus, Tobias Ritschel, and Timo Ropinski. Total denoising: Unsupervised learning
398 of 3d point cloud cleaning. In *IEEE International Conference on Computer Vision*, pages 52–60, 2019.
- 399 [10] Rohan Chabra, Jan Eric Lenssen, Eddy Ilg, Tanner Schmidt, Julian Straub, Steven Lovegrove, and
400 Richard A. Newcombe. Deep local shapes: Learning local SDF priors for detailed 3D reconstruction. In
401 *European Conference on Computer Vision*, volume 12374, pages 608–625, 2020.
- 402 [11] Angel X. Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio
403 Savarese, Manolis Savva, Shuran Song, Hao Su, Jianxiong Xiao, Li Yi, and Fisher Yu. ShapeNet: An
404 Information-Rich 3D Model Repository. Technical Report arXiv:1512.03012 [cs.GR], Stanford University
405 — Princeton University — Toyota Technological Institute at Chicago, 2015.
- 406 [12] Chao Chen, Zhizhong Han, and Yu-Shen Liu. Unsupervised inference of signed distance functions from
407 single sparse point clouds without learning priors. In *Proceedings of the IEEE/CVF Conference on
408 Computer Vision and Pattern Recognition*, pages 17712–17723, 2023.
- 409 [13] Chao Chen, Zhizhong Han, Yu-Shen Liu, and Matthias Zwicker. Unsupervised learning of fine structure
410 generation for 3D point clouds by 2D projections matching. In *IEEE International Conference on
411 Computer Vision*, 2021.
- 412 [14] Chao Chen, Yu-Shen Liu, and Zhizhong Han. Latent partition implicit with surface codes for 3d
413 representation. In *European Conference on Computer Vision*, 2022.
- 414 [15] Chao Chen, Yu-Shen Liu, and Zhizhong Han. GridPull: Towards scalability in learning implicit represen-
415 tations from 3d point clouds. In *Proceedings of the IEEE/CVF International Conference on Computer
416 Vision (ICCV)*, pages 18322–18334, 2023.
- 417 [16] Zhiqin Chen and Hao Zhang. Learning implicit fields for generative shape modeling. *IEEE Conference
418 on Computer Vision and Pattern Recognition*, 2019.
- 419 [17] Julian Chibane, Aymen Mir, and Gerard Pons-Moll. Neural unsigned distance fields for implicit function
420 learning. *arXiv*, 2010.13938, 2020.
- 421 [18] Gene Chou, Ilya Chugunov, and Felix Heide. GenSDF: Two-stage learning of generalizable signed
422 distance functions. In *Advances in Neural Information Processing Systems*, pages 24905–24919, 2022.
- 423 [19] Christopher B. Choy, Danfei Xu, JunYoung Gwak, Kevin Chen, and Silvio Savarese. 3D-r2n2: A unified
424 approach for single and multi-view 3d object reconstruction. In Bastian Leibe, Jiri Matas, Nicu Sebe, and
425 Max Welling, editors, *European Conference on Computer Vision*, volume 9912, pages 628–644, 2016.
- 426 [20] Angela Dai and Matthias Nießner. Neural Poisson: Indicator functions for neural fields. *arXiv preprint
427 arXiv:2211.14249*, 2022.

- 428 [21] Philipp Erler, Paul Guerrero, Stefan Ohrhallinger, Niloy J. Mitra, and Michael Wimmer. Points2Surf:
 429 Learning implicit surfaces from point clouds. In *European Conference on Computer Vision*, 2020.
- 430 [22] Miguel Fainstein, Viviana Siless, and Emmanuel Iarussi. DUDF: Differentiable unsigned distance fields
 431 with hyperbolic scaling. *arXiv preprint arXiv:2402.08876*, 2024.
- 432 [23] Haoqiang Fan, Hao Su, and Leonidas J. Guibas. A point set generation network for 3D object reconstruc-
 433 tion from a single image. In *2017 IEEE Conference on Computer Vision and Pattern Recognition*, pages
 434 2463–2471, 2017.
- 435 [24] Qiancheng Fu, Qingshan Xu, Yew-Soon Ong, and Wenbing Tao. Geo-Neus: Geometry-consistent neural
 436 implicit surfaces learning for multi-view reconstruction. 2022.
- 437 [25] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision
 438 benchmark suite. In *Computer Vision and Pattern Recognition*, 2012.
- 439 [26] Kyle Genova, Forrester Cole, Daniel Vlasic, Aaron Sarna, William T. Freeman, and Thomas Funkhouser.
 440 Learning shape templates with structured implicit functions. In *International Conference on Computer
 441 Vision*, 2019.
- 442 [27] Amos Gropp, Lior Yariv, Niv Haim, Matan Atzmon, and Yaron Lipman. Implicit geometric regularization
 443 for learning shapes. In *International Conference on Machine Learning*, volume 119 of *Proceedings of
 444 Machine Learning Research*, pages 3789–3799, 2020.
- 445 [28] Zhizhong Han, Chao Chen, Yu-Shen Liu, and Matthias Zwicker. DRWR: A differentiable renderer without
 446 rendering for unsupervised 3D structure learning from silhouette images. In *International Conference on
 447 Machine Learning*, 2020.
- 448 [29] Zhizhong Han, Chao Chen, Yu-Shen Liu, and Matthias Zwicker. ShapeCaptioner: Generative caption
 449 network for 3D shapes by learning a mapping from parts detected in multiple views to sentences. In *ACM
 450 International Conference on Multimedia*, 2020.
- 451 [30] Zhizhong Han, Xiyang Wang, Yu-Shen Liu, and Matthias Zwicker. Hierarchical view predictor: Unsuper-
 452 vised 3d global feature learning through hierarchical prediction among unordered views. In *Proceedings
 453 of the 29th ACM International Conference on Multimedia*, pages 3862—3871, 2021.
- 454 [31] Rana Hanocka, Gal Metzer, Raja Giryes, and Daniel Cohen-Or. Point2mesh: a self-prior for deformable
 455 meshes. *ACM Transactions on Graphics*, 39(4):126, 2020.
- 456 [32] Jiahui Huang, Hao-Xiang Chen, and Shi-Min Hu. A neural galerkin solver for accurate surface recon-
 457 struction. *ACM Trans. Graph.*, 41(6), 2022.
- 458 [33] Jiahui Huang, Zan Gojcic, Matan Atzmon, Or Litany, Sanja Fidler, and Francis Williams. Neural kernel
 459 surface reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern
 460 Recognition*, pages 4369–4379, 2023.
- 461 [34] Meng Jia and Matthew Kyan. Learning occupancy function from point clouds for surface reconstruction.
 462 *arXiv*, 2010.11378, 2020.
- 463 [35] Chiyu Jiang, Avneesh Sud, Ameesh Makadia, Jingwei Huang, Matthias Nießner, and Thomas Funkhouser.
 464 Local implicit grid representations for 3D scenes. In *IEEE Conference on Computer Vision and Pattern
 465 Recognition*, 2020.
- 466 [36] Yue Jiang, Dantong Ji, Zhizhong Han, and Matthias Zwicker. SDFDiff: Differentiable rendering of
 467 signed distance fields for 3D shape optimization. In *IEEE Conference on Computer Vision and Pattern
 468 Recognition*, 2020.
- 469 [37] Michael M. Kazhdan and Hugues Hoppe. Screened poisson surface reconstruction. *ACM Transactions
 470 on Graphics*, 32(3):29:1–29:13, 2013.
- 471 [38] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo
 472 Aila. Noise2noise: Learning image restoration without clean data. In Jennifer G. Dy and Andreas Krause,
 473 editors, *International Conference on Machine Learning*, volume 80, pages 2971–2980, 2018.
- 474 [39] Chen-Hsuan Lin, Chaoyang Wang, and Simon Lucey. SDF-SRN: Learning signed distance 3D object
 475 reconstruction from static images. In *Advances in Neural Information Processing Systems*, 2020.
- 476 [40] Cheng Lin, Changjian Li, Yuan Liu, Nenglun Chen, Yi-King Choi, and Wenping Wang. Point2Skeleton:
 477 Learning skeletal representations from point clouds. In *Proceedings of the IEEE/CVF Conference on
 478 Computer Vision and Pattern Recognition*, pages 4277–4286, 2021.

- 479 [41] Minghua Liu, Xiaoshuai Zhang, and Hao Su. Meshing point clouds with predicted intrinsic-extrinsic
480 ratio guidance. In *European Conference on Computer vision*, 2020.
- 481 [42] Shaohui Liu, Yinda Zhang, Songyou Peng, Boxin Shi, Marc Pollefeys, and Zhaopeng Cui. DIST:
482 Rendering deep implicit signed distance function with differentiable sphere tracing. In *IEEE Conference
483 on Computer Vision and Pattern Recognition*, 2020.
- 484 [43] Shi-Lin Liu, Hao-Xiang Guo, Hao Pan, Pengshuai Wang, Xin Tong, and Yang Liu. Deep implicit moving
485 least-squares functions for 3D reconstruction. In *IEEE Conference on Computer Vision and Pattern
486 Recognition*, 2021.
- 487 [44] Shichen Liu, Shunsuke Saito, Weikai Chen, and Hao Li. Learning to infer implicit surfaces without 3D
488 supervision. In *Advances in Neural Information Processing Systems*, 2019.
- 489 [45] William E. Lorensen and Harvey E. Cline. Marching cubes: A high resolution 3D surface construction
490 algorithm. *Computer Graphics*, 21(4):163–169, 1987.
- 491 [46] Shitong Luo and Wei Hu. Differentiable manifold reconstruction for point cloud denoising. In *ACM
492 International Conference on Multimedia*, pages 1330–1338. ACM, 2020.
- 493 [47] Shitong Luo and Wei Hu. Score-based point cloud denoising. In *Proceedings of the IEEE/CVF Interna-
494 tional Conference on Computer Vision*, pages 4583–4592, 2021.
- 495 [48] Baorui Ma, Zhizhong Han, Yu-Shen Liu, and Matthias Zwicker. Neural-pull: Learning signed distance
496 functions from point clouds by learning to pull space onto surfaces. In *International Conference on
497 Machine Learning*, 2021.
- 498 [49] Baorui Ma, Yu-Shen Liu, and Zhizhong Han. Reconstructing surfaces for sparse point clouds with
499 on-surface priors. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 6305–6315,
500 2022.
- 501 [50] Baorui Ma, Yu-Shen Liu, and Zhizhong Han. Learning signed distance functions from noisy 3D point
502 clouds via noise to noise mapping. In *International Conference on Machine Learning*, pages 23338–23357.
503 PMLR, 2023.
- 504 [51] Baorui Ma, Yu-Shen Liu, Matthias Zwicker, and Zhizhong Han. Reconstructing surfaces for sparse point
505 clouds with on-surface priors. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2022.
- 506 [52] Baorui Ma, Yu-Shen Liu, Matthias Zwicker, and Zhizhong Han. Surface reconstruction from point clouds
507 by learning predictive context priors. In *IEEE Conference on Computer Vision and Pattern Recognition*,
508 pages 6316–6327, 2022.
- 509 [53] Baorui Ma, Yu-Shen Liu, Matthias Zwicker, and Zhizhong Han. Surface reconstruction from point clouds
510 by learning predictive context priors. In *IEEE Conference on Computer Vision and Pattern Recognition*,
511 2022.
- 512 [54] Baorui Ma, Junsheng Zhou, Yu-Shen Liu, and Zhizhong Han. Towards better gradient consistency for
513 neural signed distance functions via level set alignment. In *Proceedings of the IEEE/CVF Conference on
514 Computer Vision and Pattern Recognition*, pages 17724–17734, 2023.
- 515 [55] Julien N. P. Martel, David B. Lindell, Connor Z. Lin, Eric R. Chan, Marco Monteiro, and Gordon
516 Wetzstein. ACORN: adaptive coordinate networks for neural scene representation. *CoRR*, abs/2105.02788,
517 2021.
- 518 [56] Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, and Andreas Geiger. Occu-
519 pancy networks: Learning 3D reconstruction in function space. In *IEEE Conference on Computer Vision
520 and Pattern Recognition*, 2019.
- 521 [57] Zhenxing Mi, Yiming Luo, and Wenbing Tao. SSRNet: Scalable 3D surface reconstruction network. In
522 *IEEE Conference on Computer Vision and Pattern Recognition*, 2020.
- 523 [58] Mateusz Michalkiewicz, Jhony K. Pontes, Dominic Jack, Mahsa Baktashmotagh, and Anders P. Eriksson.
524 Deep level sets: Implicit surface representations for 3D shape inference. *CoRR*, abs/1901.06802, 2019.
- 525 [59] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren
526 Ng. NeRF: Representing scenes as neural radiance fields for view synthesis. In *European Conference on
527 Computer Vision*, 2020.

- 528 [60] Michael Niemeyer, Lars Mescheder, Michael Oechsle, and Andreas Geiger. Differentiable volumetric
 529 rendering: Learning implicit 3D representations without 3D supervision. In *IEEE Conference on*
 530 *Computer Vision and Pattern Recognition*, 2020.
- 531 [61] Michael Oechsle, Songyou Peng, and Andreas Geiger. UNISURF: Unifying neural implicit surfaces and
 532 radiance fields for multi-view reconstruction. In *International Conference on Computer Vision*, 2021.
- 533 [62] Amine Ouasfi and Adnane Boukhayma. Few ‘zero level set’-shot learning of shape signed distance
 534 functions in feature space. In *European Conference on Computer Vision*, 2022.
- 535 [63] Amine Ouasfi and Adnane Boukhayma. Robustifying generalizable implicit shape networks with a
 536 tunable non-parametric model. In *Advances in Neural Information Processing Systems*, 2023.
- 537 [64] Amine Ouasfi and Adnane Boukhayma. Mixing-denoising generalizable occupancy networks. In
 538 *International Conference on 3D Vision*, 2024.
- 539 [65] Amine Ouasfi and Adnane Boukhayma. Unsupervised occupancy learning from sparse point cloud. In
 540 *IEEE Conference on Computer Vision and Pattern Recognition*, 2024.
- 541 [66] Jeong Joon Park, Peter Florence, Julian Straub, Richard Newcombe, and Steven Lovegrove. DeepSDF:
 542 Learning continuous signed distance functions for shape representation. In *IEEE Conference on Computer*
 543 *Vision and Pattern Recognition*, 2019.
- 544 [67] Songyou Peng, Chiyu “Max” Jiang, Yiyi Liao, Michael Niemeyer, Marc Pollefeys, and Andreas Geiger.
 545 Shape as points: A differentiable poisson solver. In *Advances in Neural Information Processing Systems*,
 546 2021.
- 547 [68] Songyou Peng, Michael Niemeyer, Lars M. Mescheder, Marc Pollefeys, and Andreas Geiger. Convolu-
 548 tional occupancy networks. In *European Conference on Computer Vision*, volume 12348, pages 523–540,
 549 2020.
- 550 [69] Francesca Pistilli, Giulia Fracastoro, Diego Valsesia, and Enrico Magli. Learning graph-convolutional
 551 representations for point cloud denoising. In *European Conference on Computer Vision*, volume 12365,
 552 pages 103–118, 2020.
- 553 [70] Marie-Julie Rakotosaona, Vittorio La Barbera, Paul Guerrero, Niloy J. Mitra, and Maks Ovsjanikov.
 554 Pointcleannet: Learning to denoise and remove outliers from dense point clouds. *Computer Graphics*
 555 *Forum*, 39(1):185–203, 2020.
- 556 [71] Konstantinos Rematas, Ricardo Martin-Brualla, and Vittorio Ferrari. Sharf: Shape-conditioned radiance
 557 fields from a single view. In *International Conference on Machine Learning*, 2021.
- 558 [72] Andrés Serna, Beatriz Marcotegui, François Goulette, and Jean-Emmanuel Deschaud. Paris-rue-madame
 559 database - A 3D mobile laser scanner dataset for benchmarking urban detection, segmentation and
 560 classification methods. In *International Conference on Pattern Recognition Applications and Methods*,
 561 pages 819–824, 2014.
- 562 [73] Jaehyeok Shim and Kyungdon Joo. DITTO: Dual and integrated latent topologies for implicit 3d
 563 reconstruction. *arXiv preprint arXiv:2403.05005*, 2024.
- 564 [74] Vincent Sitzmann, Julien N.P. Martel, Alexander W. Bergman, David B. Lindell, and Gordon Wetzstein.
 565 Implicit neural representations with periodic activation functions. In *Advances in Neural Information*
 566 *Processing Systems*, 2020.
- 567 [75] Vincent Sitzmann, Michael Zollhöfer, and Gordon Wetzstein. Scene representation networks: Continuous
 568 3D-structure-aware neural scene representations. In *Advances in Neural Information Processing Systems*,
 569 2019.
- 570 [76] Lars Mescheder Marc Pollefeys Andreas Geiger Songyou Peng, Michael Niemeyer. Convolutional
 571 occupancy networks. In *European Conference on Computer Vision*, 2020.
- 572 [77] Towaki Takikawa, Joey Litalien, Kangxue Yin, Karsten Kreis, Charles Loop, Derek Nowrouzezahrai,
 573 Alec Jacobson, Morgan McGuire, and Sanja Fidler. Neural geometric level of detail: Real-time rendering
 574 with implicit 3D shapes. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2021.
- 575 [78] Jiapeng Tang, Jiabao Lei, Dan Xu, Feiying Ma, Kui Jia, and Lei Zhang. SA-ConvONet: Sign-agnostic
 576 optimization of convolutional occupancy networks. In *Proceedings of the IEEE/CVF International*
 577 *Conference on Computer Vision*, 2021.

- 578 [79] Maxim Tatarchenko, Stephan R. Richter, Rene Ranftl, Zhuwen Li, Vladlen Koltun, and Thomas Brox.
 579 What do single-view 3D reconstruction networks learn? In *The IEEE Conference on Computer Vision*
 580 and *Pattern Recognition*, 2019.
- 581 [80] Edgar Tretschk, Ayush Tewari, Vladislav Golyanik, Michael Zollhöfer, Carsten Stoll, and Christian
 582 Theobalt. PatchNets: Patch-Based Generalizable Deep Implicit 3D Shape Representations. *European*
 583 *Conference on Computer Vision*, 2020.
- 584 [81] Delio Vicini, Sébastien Speierer, and Wenzel Jakob. Differentiable signed distance function rendering.
 585 *ACM Transactions on Graphics*, 41(4):125:1–125:18, 2022.
- 586 [82] Jiepeng Wang, Peng Wang, Xiaoxiao Long, Christian Theobalt, Taku Komura, Lingjie Liu, and Wenping
 587 Wang. NeuRIS: Neural reconstruction of indoor scenes using normal priors. In *European Conference on*
 588 *Computer Vision*, 2022.
- 589 [83] Peng Wang, Lingjie Liu, Yuan Liu, Christian Theobalt, Taku Komura, and Wenping Wang. NeuS:
 590 Learning neural implicit surfaces by volume rendering for multi-view reconstruction. In *Advances in*
 591 *Neural Information Processing Systems*, pages 27171–27183, 2021.
- 592 [84] Peng-Shuai Wang, Yang Liu, and Xin Tong. Deep octree-based cnns with output-guided skip connections
 593 for 3d shape and scene completion. In *Proceedings of the IEEE/CVF Conference on Computer Vision*
 594 and *Pattern Recognition Workshops*, pages 266–267, 2020.
- 595 [85] Ruian Wang, Zixiong Wang, Yunxiao Zhang, Shuangmin Chen, Shiqing Xin, Changhe Tu, and Wenping
 596 Wang. Aligning gradient and hessian for neural signed distance function. In *Advances in Neural*
 597 *Information Processing Systems*, volume 36, pages 63515–63528, 2023.
- 598 [86] Yiqun Wang, Ivan Skorokhodov, and Peter Wonka. HF-NeuS: Improved surface reconstruction using
 599 high-frequency details. 2022.
- 600 [87] Zhen Wang, Shijie Zhou, Jeong Joon Park, Despoina Paschalidou, Suya You, Gordon Wetzstein, Leonidas
 601 Guibas, and Achuta Kadambi. Alto: Alternating latent topologies for implicit 3d reconstruction. In
 602 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages
 603 259–270, 2023.
- 604 [88] Francis Williams, Zan Gojcic, Sameh Khamis, Denis Zorin, Joan Bruna, Sanja Fidler, and Or Litany.
 605 Neural fields as learnable kernels for 3D reconstruction. In *Proceedings of the IEEE/CVF Conference on*
 606 *Computer Vision and Pattern Recognition*, pages 18500–18510, 2022.
- 607 [89] Francis Williams, Teseo Schneider, Claudio Silva, Denis Zorin, Joan Bruna, and Daniele Panozzo.
 608 Deep geometric prior for surface reconstruction. In *IEEE Conference on Computer Vision and Pattern*
 609 *Recognition*, 2019.
- 610 [90] Francis Williams, Matthew Trager, Joan Bruna, and Denis Zorin. Neural splines: Fitting 3D surfaces with
 611 infinitely-wide neural networks. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages
 612 9949–9958, 2021.
- 613 [91] Yunjie Wu and Zhengxing Sun. DFR: differentiable function rendering for learning 3D generation from
 614 images. *Computer Graphics Forum*, 39(5):241–252, 2020.
- 615 [92] Peng Xiang, Xin Wen, Yu-Shen Liu, Yan-Pei Cao, Pengfei Wan, Wen Zheng, and Zhizhong Han.
 616 SnowflakeNet: Point cloud completion by snowflake point deconvolution with skip-transformer. In *IEEE*
 617 *International Conference on Computer Vision*, 2021.
- 618 [93] Huizong Yang, Yuxin Sun, Ganesh Sundaramoorthi, and Anthony Yezzi. StEik: Stabilizing the optimization
 619 of neural signed distance functions and finer shape representation. In *Advances in Neural Information*
 620 *Processing Systems*, 2023.
- 621 [94] Lior Yariv, Jiatao Gu, Yoni Kasten, and Yaron Lipman. Volume rendering of neural implicit surfaces. In
 622 *Advances in Neural Information Processing Systems*, 2021.
- 623 [95] Lior Yariv, Yoni Kasten, Dror Moran, Meirav Galun, Matan Atzmon, Basri Ronen, and Yaron Lipman.
 624 Multiview neural surface reconstruction by disentangling geometry and appearance. *Advances in Neural*
 625 *Information Processing Systems*, 33, 2020.
- 626 [96] Wang Yifan, Shihao Wu, Cengiz Oztireli, and Olga Sorkine-Hornung. Iso-Points: Optimizing neural
 627 implicit surfaces with hybrid representations. *CoRR*, abs/2012.06434, 2020.

- 628 [97] Zehao Yu, Songyou Peng, Michael Niemeyer, Torsten Sattler, and Andreas Geiger. MonoSDF: Exploring
629 monocular geometric cues for neural implicit surface reconstruction. *ArXiv*, abs/2022.00665, 2022.
- 630 [98] Sergey Zakharov, Wadim Kehl, Arjun Bhargava, and Adrien Gaidon. Autolabeling 3D objects with
631 differentiable rendering of sdf shape priors. In *IEEE Conference on Computer Vision and Pattern
632 Recognition*, 2020.
- 633 [99] Wenbin Zhao, Jiabao Lei, Yuxin Wen, Jianguo Zhang, and Kui Jia. Sign-agnostic implicit learning of sur-
634 face self-similarities for shape modeling and reconstruction from raw point clouds. *CoRR*, abs/2012.07498,
635 2020.
- 636 [100] Junsheng Zhou, Baorui Ma, Shujuan Li, Yu-Shen Liu, and Zhizhong Han. Learning a more continuous
637 zero level set in unsigned distance fields through level set projection. In *Proceedings of the IEEE/CVF
638 International Conference on Computer Vision*, pages 3181–3192, 2023.
- 639 [101] Junsheng Zhou, Baorui Ma, Yu-Shen Liu, Yi Fang, and Zhizhong Han. Learning consistency-aware
640 unsigned distance functions progressively from raw point clouds. In *Advances in Neural Information
641 Processing Systems*, 2022.
- 642 [102] Qian-Yi Zhou and Vladlen Koltun. Dense scene reconstruction with points of interest. *ACM Transactions
643 on Graphics*, 32(4):112:1–112:8, 2013.
- 644 [103] Runsong Zhu, Di Kang, Ka-Hei Hui, Yue Qian, Shi Qiu, Zhen Dong, Linchao Bao, Pheng-Ann Heng,
645 and Chi-Wing Fu. Ssp: Semi-signed prioritized neural fitting for surface reconstruction from unoriented
646 point clouds. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision
647 (WACV)*, pages 3769–3778, 2024.

648 A Appendix

649 A.1 Limitations

650 Our method is still limited to too large noises. For noises that corrupted shapes too much, our method
651 still produces bad results. One direction for our future work is to improve our prior, so that we could
652 have a better sense of a shape even under large noises.

653 A.2 Detailed Comparisons on 3D Scene

654 We detail our evaluations on each scene in 3D scene dataset in Tab. 15. The comparisons highlight
655 our advantages in each scene.

Name	Metrics	COcc [68]	LIG [35]	DeepLS[10]	NP [48]	N2NM [50]	Ours
Burghers	$CD_{L2} \times 1000$	27.46	3.055	0.401	1.204	0.504	0.429
	CD_{L1}	0.079	0.045	0.017	0.031	0.020	0.016
	NC	0.907	0.835	0.920	0.905	0.925	0.939
Loung	$CD_{L2} \times 1000$	9.540	9.672	6.103	1.079	0.602	0.333
	CD_{L1}	0.046	0.056	0.053	0.019	0.016	0.014
	NC	0.894	0.833	0.848	0.910	0.923	0.935
Copyroom	$CD_{L2} \times 1000$	10.97	3.610	0.609	5.795	0.442	0.389
	CD_{L1}	0.045	0.036	0.021	0.036	0.016	0.016
	NC	0.892	0.810	0.901	0.862	0.903	0.916
Stonewall	$CD_{L2} \times 1000$	20.46	5.032	0.320	0.983	0.330	0.313
	CD_{L1}	0.069	0.042	0.015	0.029	0.020	0.015
	NC	0.905	0.879	0.954	0.930	0.951	0.961
Totepole	$CD_{L2} \times 1000$	2.054	9.580	0.601	1.513	0.657	0.482
	CD_{L1}	0.021	0.062	0.017	0.054	0.023	0.020
	NC	0.943	0.887	0.950	0.893	0.945	0.957

Table 15: Numerical Comparisons on 3D Scene dataset in terms of CD_{L1} , CD_{L2} and NC.

656 A.3 More Results

657 We visualize more surface reconstruction results under ShapeNet [11], ABC [21], Surface Recon-
658 struction Benchmark (SRB) [89], FAMOUS [21] and D-FAUST [5] in Fig. 11, Fig. 12, Fig. 13,
659 Fig. 14 and Fig. 15.

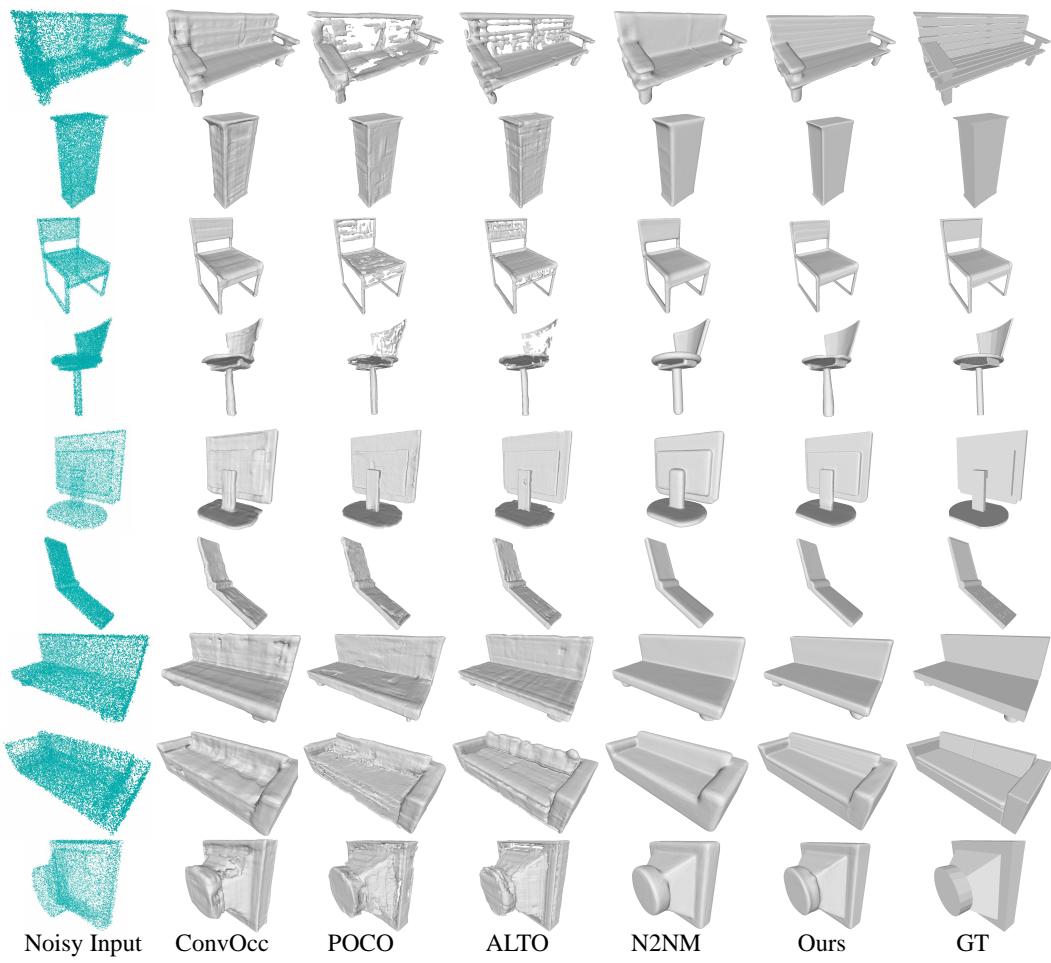


Figure 11: Comparison in surface reconstruction under ShapeNet.

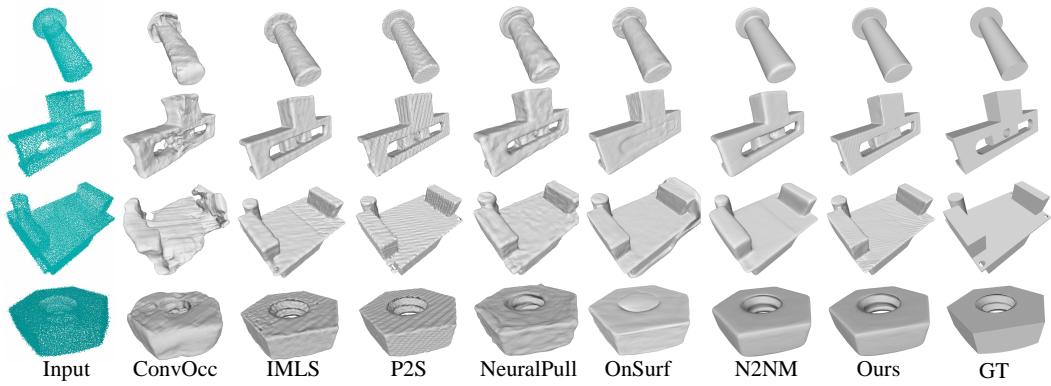


Figure 12: Comparison in surface reconstruction under ABC.

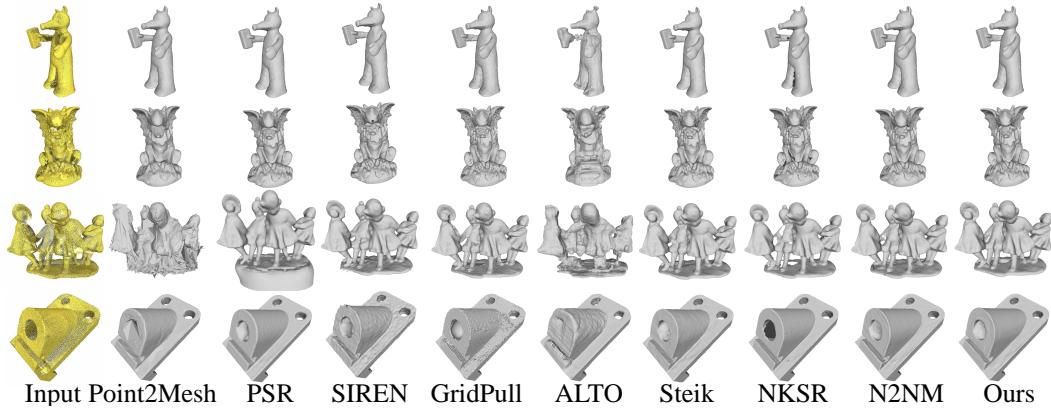


Figure 13: Comparison in surface reconstruction under SRB.

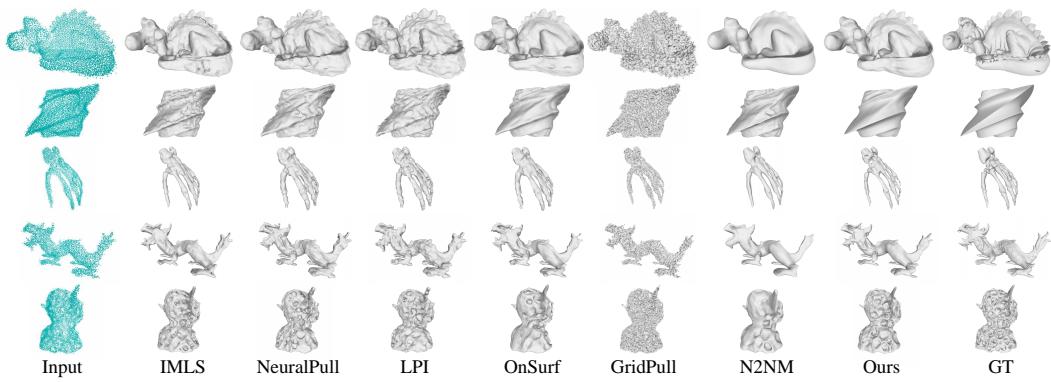


Figure 14: Comparison in surface reconstruction under FAMOUS.

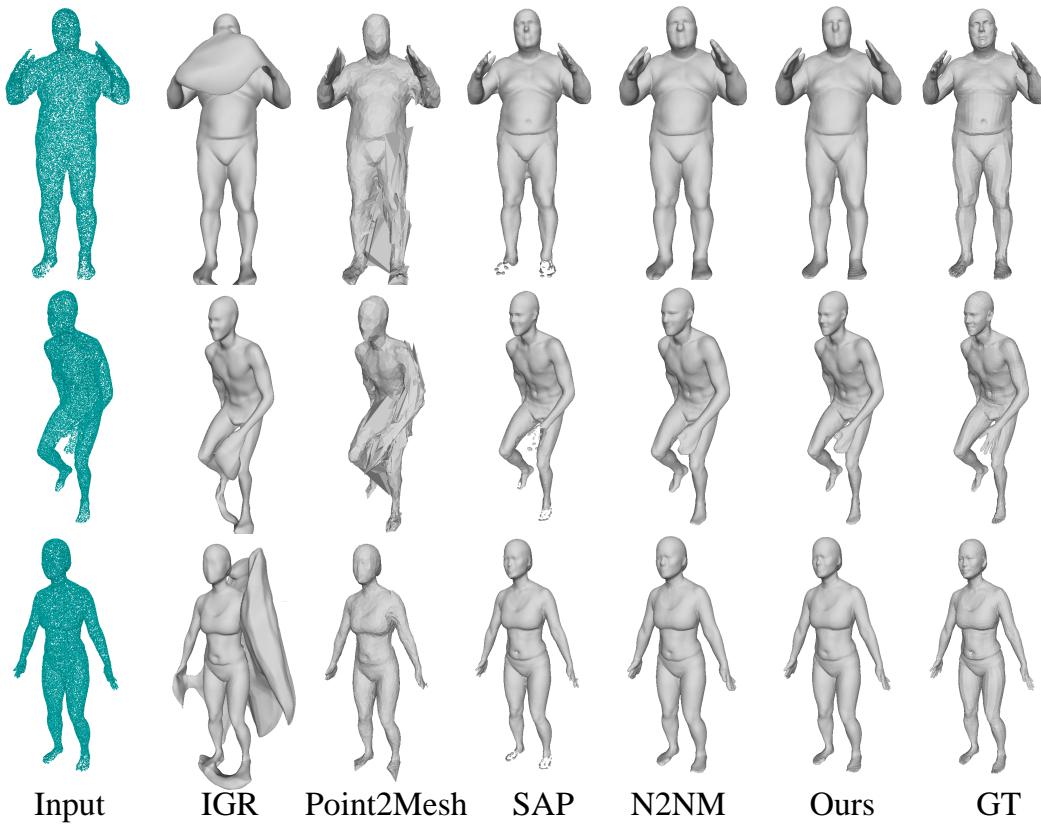


Figure 15: Comparison in surface reconstruction under D-FAUST.

660 **NeurIPS Paper Checklist**

661 **1. Claims**

662 Question: Do the main claims made in the abstract and introduction accurately reflect the
663 paper's contributions and scope?

664 Answer: [Yes]

665 Justification: Our main claims made in the abstract and introduction accurately reflect the
666 paper's contributions and scope.

667 Guidelines:

- 668 • The answer NA means that the abstract and introduction do not include the claims
669 made in the paper.
- 670 • The abstract and/or introduction should clearly state the claims made, including the
671 contributions made in the paper and important assumptions and limitations. A No or
672 NA answer to this question will not be perceived well by the reviewers.
- 673 • The claims made should match theoretical and experimental results, and reflect how
674 much the results can be expected to generalize to other settings.
- 675 • It is fine to include aspirational goals as motivation as long as it is clear that these goals
676 are not attained by the paper.

677 **2. Limitations**

678 Question: Does the paper discuss the limitations of the work performed by the authors?

679 Answer: [Yes]

680 Justification: We discuss the limitations in the Appendix A.1.

681 Guidelines:

- 682 • The answer NA means that the paper has no limitation while the answer No means that
683 the paper has limitations, but those are not discussed in the paper.
- 684 • The authors are encouraged to create a separate "Limitations" section in their paper.
- 685 • The paper should point out any strong assumptions and how robust the results are to
686 violations of these assumptions (e.g., independence assumptions, noiseless settings,
687 model well-specification, asymptotic approximations only holding locally). The authors
688 should reflect on how these assumptions might be violated in practice and what the
689 implications would be.
- 690 • The authors should reflect on the scope of the claims made, e.g., if the approach was
691 only tested on a few datasets or with a few runs. In general, empirical results often
692 depend on implicit assumptions, which should be articulated.
- 693 • The authors should reflect on the factors that influence the performance of the approach.
694 For example, a facial recognition algorithm may perform poorly when image resolution
695 is low or images are taken in low lighting. Or a speech-to-text system might not be
696 used reliably to provide closed captions for online lectures because it fails to handle
697 technical jargon.
- 698 • The authors should discuss the computational efficiency of the proposed algorithms
699 and how they scale with dataset size.
- 700 • If applicable, the authors should discuss possible limitations of their approach to
701 address problems of privacy and fairness.
- 702 • While the authors might fear that complete honesty about limitations might be used by
703 reviewers as grounds for rejection, a worse outcome might be that reviewers discover
704 limitations that aren't acknowledged in the paper. The authors should use their best
705 judgment and recognize that individual actions in favor of transparency play an impor-
706 tant role in developing norms that preserve the integrity of the community. Reviewers
707 will be specifically instructed to not penalize honesty concerning limitations.

708 **3. Theory Assumptions and Proofs**

709 Question: For each theoretical result, does the paper provide the full set of assumptions and
710 a complete (and correct) proof?

711 Answer: [NA]

712 Justification: We describe in Method one of the core contributions of the local noise-to-
713 noise mapping, and although there is no theory or theorem in it, we verify its validity and
714 reasonableness in our experiments.

715 Guidelines:

- 716 • The answer NA means that the paper does not include theoretical results.
- 717 • All the theorems, formulas, and proofs in the paper should be numbered and cross-
718 referenced.
- 719 • All assumptions should be clearly stated or referenced in the statement of any theorems.
- 720 • The proofs can either appear in the main paper or the supplemental material, but if
721 they appear in the supplemental material, the authors are encouraged to provide a short
722 proof sketch to provide intuition.
- 723 • Inversely, any informal proof provided in the core of the paper should be complemented
724 by formal proofs provided in appendix or supplemental material.
- 725 • Theorems and Lemmas that the proof relies upon should be properly referenced.

726 4. Experimental Result Reproducibility

727 Question: Does the paper fully disclose all the information needed to reproduce the main ex-
728 perimental results of the paper to the extent that it affects the main claims and/or conclusions
729 of the paper (regardless of whether the code and data are provided or not)?

730 Answer: [Yes]

731 Justification: We provide detailed information in reproducing our methods in Implementation
732 Details of Section 3.

733 Guidelines:

- 734 • The answer NA means that the paper does not include experiments.
- 735 • If the paper includes experiments, a No answer to this question will not be perceived
736 well by the reviewers: Making the paper reproducible is important, regardless of
737 whether the code and data are provided or not.
- 738 • If the contribution is a dataset and/or model, the authors should describe the steps taken
739 to make their results reproducible or verifiable.
- 740 • Depending on the contribution, reproducibility can be accomplished in various ways.
741 For example, if the contribution is a novel architecture, describing the architecture fully
742 might suffice, or if the contribution is a specific model and empirical evaluation, it may
743 be necessary to either make it possible for others to replicate the model with the same
744 dataset, or provide access to the model. In general, releasing code and data is often
745 one good way to accomplish this, but reproducibility can also be provided via detailed
746 instructions for how to replicate the results, access to a hosted model (e.g., in the case
747 of a large language model), releasing of a model checkpoint, or other means that are
748 appropriate to the research performed.
- 749 • While NeurIPS does not require releasing code, the conference does require all submis-
750 sions to provide some reasonable avenue for reproducibility, which may depend on the
751 nature of the contribution. For example
 - 752 (a) If the contribution is primarily a new algorithm, the paper should make it clear how
753 to reproduce that algorithm.
 - 754 (b) If the contribution is primarily a new model architecture, the paper should describe
755 the architecture clearly and fully.
 - 756 (c) If the contribution is a new model (e.g., a large language model), then there should
757 either be a way to access this model for reproducing the results or a way to reproduce
758 the model (e.g., with an open-source dataset or instructions for how to construct
759 the dataset).
 - 760 (d) We recognize that reproducibility may be tricky in some cases, in which case
761 authors are welcome to describe the particular way they provide for reproducibility.
762 In the case of closed-source models, it may be that access to the model is limited in
763 some way (e.g., to registered users), but it should be possible for other researchers
764 to have some path to reproducing or verifying the results.

765 5. Open access to data and code

766 Question: Does the paper provide open access to the data and code, with sufficient instruc-
767 tions to faithfully reproduce the main experimental results, as described in supplemental
768 material?

769 Answer: [Yes]

770 Justification: We provide our demonstration code as a part of our supplementary materials.
771 We will release our source code, data and sufficient instructions upon acceptance.

772 Guidelines:

- 773 • The answer NA means that paper does not include experiments requiring code.
- 774 • Please see the NeurIPS code and data submission guidelines ([https://nips.cc/
775 public/guides/CodeSubmissionPolicy](https://nips.cc/public/guides/CodeSubmissionPolicy)) for more details.
- 776 • While we encourage the release of code and data, we understand that this might not be
777 possible, so “No” is an acceptable answer. Papers cannot be rejected simply for not
778 including code, unless this is central to the contribution (e.g., for a new open-source
779 benchmark).
- 780 • The instructions should contain the exact command and environment needed to run to
781 reproduce the results. See the NeurIPS code and data submission guidelines ([https://nips.cc/
782 public/guides/CodeSubmissionPolicy](https://nips.cc/public/guides/CodeSubmissionPolicy)) for more details.
- 783 • The authors should provide instructions on data access and preparation, including how
784 to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- 785 • The authors should provide scripts to reproduce all experimental results for the new
786 proposed method and baselines. If only a subset of experiments are reproducible, they
787 should state which ones are omitted from the script and why.
- 788 • At submission time, to preserve anonymity, the authors should release anonymized
789 versions (if applicable).
- 790 • Providing as much information as possible in supplemental material (appended to the
791 paper) is recommended, but including URLs to data and code is permitted.

792 6. Experimental Setting/Details

793 Question: Does the paper specify all the training and test details (e.g., data splits, hyper-
794 parameters, how they were chosen, type of optimizer, etc.) necessary to understand the
795 results?

796 Answer: [Yes]

797 Justification: We provide all the training and test details for shapes and scenes in Section 4.

798 Guidelines:

- 799 • The answer NA means that the paper does not include experiments.
- 800 • The experimental setting should be presented in the core of the paper to a level of detail
801 that is necessary to appreciate the results and make sense of them.
- 802 • The full details can be provided either with the code, in appendix, or as supplemental
803 material.

804 7. Experiment Statistical Significance

805 Question: Does the paper report error bars suitably and correctly defined or other appropriate
806 information about the statistical significance of the experiments?

807 Answer: [No]

808 Justification: We report the average performance in terms of several metrics as the experi-
809 mental results.

810 Guidelines:

- 811 • The answer NA means that the paper does not include experiments.
- 812 • The authors should answer “Yes” if the results are accompanied by error bars, confi-
813 dence intervals, or statistical significance tests, at least for the experiments that support
814 the main claims of the paper.
- 815 • The factors of variability that the error bars are capturing should be clearly stated (for
816 example, train/test split, initialization, random drawing of some parameter, or overall
817 run with given experimental conditions).

- 818 • The method for calculating the error bars should be explained (closed form formula,
819 call to a library function, bootstrap, etc.)
820 • The assumptions made should be given (e.g., Normally distributed errors).
821 • It should be clear whether the error bar is the standard deviation or the standard error
822 of the mean.
823 • It is OK to report 1-sigma error bars, but one should state it. The authors should
824 preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis
825 of Normality of errors is not verified.
826 • For asymmetric distributions, the authors should be careful not to show in tables or
827 figures symmetric error bars that would yield results that are out of range (e.g. negative
828 error rates).
829 • If error bars are reported in tables or plots, The authors should explain in the text how
830 they were calculated and reference the corresponding figures or tables in the text.

831 **8. Experiments Compute Resources**

832 Question: For each experiment, does the paper provide sufficient information on the com-
833 puter resources (type of compute workers, memory, time of execution) needed to reproduce
834 the experiments?

835 Answer: [Yes]

836 Justification: We report our inference time with other methods in the experiments.

837 Guidelines:

- 838 • The answer NA means that the paper does not include experiments.
839 • The paper should indicate the type of compute workers CPU or GPU, internal cluster,
840 or cloud provider, including relevant memory and storage.
841 • The paper should provide the amount of compute required for each of the individual
842 experimental runs as well as estimate the total compute.
843 • The paper should disclose whether the full research project required more compute
844 than the experiments reported in the paper (e.g., preliminary or failed experiments that
845 didn't make it into the paper).

846 **9. Code Of Ethics**

847 Question: Does the research conducted in the paper conform, in every respect, with the
848 NeurIPS Code of Ethics <https://neurips.cc/public/EthicsGuidelines>?

849 Answer: [Yes]

850 Justification: The research conducted in this paper conforms in all respects to the NeurIPS
851 Code of Ethics.

852 Guidelines:

- 853 • The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
854 • If the authors answer No, they should explain the special circumstances that require a
855 deviation from the Code of Ethics.
856 • The authors should make sure to preserve anonymity (e.g., if there is a special consid-
857 eration due to laws or regulations in their jurisdiction).

858 **10. Broader Impacts**

859 Question: Does the paper discuss both potential positive societal impacts and negative
860 societal impacts of the work performed?

861 Answer: [Yes]

862 Justification: We discuss the application and potential positive impact of our method in the
863 introduction.

864 Guidelines:

- 865 • The answer NA means that there is no societal impact of the work performed.
866 • If the authors answer NA or No, they should explain why their work has no societal
867 impact or why the paper does not address societal impact.

- 868 • Examples of negative societal impacts include potential malicious or unintended uses
869 (e.g., disinformation, generating fake profiles, surveillance), fairness considerations
870 (e.g., deployment of technologies that could make decisions that unfairly impact specific
871 groups), privacy considerations, and security considerations.
872 • The conference expects that many papers will be foundational research and not tied
873 to particular applications, let alone deployments. However, if there is a direct path to
874 any negative applications, the authors should point it out. For example, it is legitimate
875 to point out that an improvement in the quality of generative models could be used to
876 generate deepfakes for disinformation. On the other hand, it is not needed to point out
877 that a generic algorithm for optimizing neural networks could enable people to train
878 models that generate Deepfakes faster.
879 • The authors should consider possible harms that could arise when the technology is
880 being used as intended and functioning correctly, harms that could arise when the
881 technology is being used as intended but gives incorrect results, and harms following
882 from (intentional or unintentional) misuse of the technology.
883 • If there are negative societal impacts, the authors could also discuss possible mitigation
884 strategies (e.g., gated release of models, providing defenses in addition to attacks,
885 mechanisms for monitoring misuse, mechanisms to monitor how a system learns from
886 feedback over time, improving the efficiency and accessibility of ML).

887 **11. Safeguards**

888 Question: Does the paper describe safeguards that have been put in place for responsible
889 release of data or models that have a high risk for misuse (e.g., pretrained language models,
890 image generators, or scraped datasets)?

891 Answer: [NA]

892 Justification: There is no such risk to the paper.

893 Guidelines:

- 894 • The answer NA means that the paper poses no such risks.
895 • Released models that have a high risk for misuse or dual-use should be released with
896 necessary safeguards to allow for controlled use of the model, for example by requiring
897 that users adhere to usage guidelines or restrictions to access the model or implementing
898 safety filters.
899 • Datasets that have been scraped from the Internet could pose safety risks. The authors
900 should describe how they avoided releasing unsafe images.
901 • We recognize that providing effective safeguards is challenging, and many papers do
902 not require this, but we encourage authors to take this into account and make a best
903 faith effort.

904 **12. Licenses for existing assets**

905 Question: Are the creators or original owners of assets (e.g., code, data, models), used in
906 the paper, properly credited and are the license and terms of use explicitly mentioned and
907 properly respected?

908 Answer: [Yes]

909 Justification: We use open-source datasets and code under their licence.

910 Guidelines:

- 911 • The answer NA means that the paper does not use existing assets.
912 • The authors should cite the original paper that produced the code package or dataset.
913 • The authors should state which version of the asset is used and, if possible, include a
914 URL.
915 • The name of the license (e.g., CC-BY 4.0) should be included for each asset.
916 • For scraped data from a particular source (e.g., website), the copyright and terms of
917 service of that source should be provided.
918 • If assets are released, the license, copyright information, and terms of use in the
919 package should be provided. For popular datasets, paperswithcode.com/datasets
920 has curated licenses for some datasets. Their licensing guide can help determine the
921 license of a dataset.

- 922 • For existing datasets that are re-packaged, both the original license and the license of
923 the derived asset (if it has changed) should be provided.
924 • If this information is not available online, the authors are encouraged to reach out to
925 the asset's creators.

926 **13. New Assets**

927 Question: Are new assets introduced in the paper well documented and is the documentation
928 provided alongside the assets?

929 Answer: [NA]

930 Justification: The paper does not release new assets.

931 Guidelines:

- 932 • The answer NA means that the paper does not release new assets.
933 • Researchers should communicate the details of the dataset/code/model as part of their
934 submissions via structured templates. This includes details about training, license,
935 limitations, etc.
936 • The paper should discuss whether and how consent was obtained from people whose
937 asset is used.
938 • At submission time, remember to anonymize your assets (if applicable). You can either
939 create an anonymized URL or include an anonymized zip file.

940 **14. Crowdsourcing and Research with Human Subjects**

941 Question: For crowdsourcing experiments and research with human subjects, does the paper
942 include the full text of instructions given to participants and screenshots, if applicable, as
943 well as details about compensation (if any)?

944 Answer: [NA]

945 Justification: The paper does not involve crowdsourcing nor research with human subjects.

946 Guidelines:

- 947 • The answer NA means that the paper does not involve crowdsourcing nor research with
948 human subjects.
949 • Including this information in the supplemental material is fine, but if the main contribu-
950 tion of the paper involves human subjects, then as much detail as possible should be
951 included in the main paper.
952 • According to the NeurIPS Code of Ethics, workers involved in data collection, curation,
953 or other labor should be paid at least the minimum wage in the country of the data
954 collector.

955 **15. Institutional Review Board (IRB) Approvals or Equivalent for Research with Human
956 Subjects**

957 Question: Does the paper describe potential risks incurred by study participants, whether
958 such risks were disclosed to the subjects, and whether Institutional Review Board (IRB)
959 approvals (or an equivalent approval/review based on the requirements of your country or
960 institution) were obtained?

961 Answer: [NA]

962 Justification: The paper does not involve crowdsourcing nor research with human subjects.

963 Guidelines:

- 964 • The answer NA means that the paper does not involve crowdsourcing nor research with
965 human subjects.
966 • Depending on the country in which research is conducted, IRB approval (or equivalent)
967 may be required for any human subjects research. If you obtained IRB approval, you
968 should clearly state this in the paper.
969 • We recognize that the procedures for this may vary significantly between institutions
970 and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the
971 guidelines for their institution.
972 • For initial submissions, do not include any information that would break anonymity (if
973 applicable), such as the institution conducting the review.