

---

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

---

Chao Chen<sup>1</sup> Yu-Shen Liu<sup>1\*</sup> Zhizhong Han<sup>2</sup>

<sup>1</sup>School of Software, Tsinghua University, Beijing, China

<sup>2</sup>Department of Computer Science, Wayne State University, Detroit, USA  
chenchao19@tsinghua.org.cn liuyushen@mails.tsinghua.edu.cn  
h312h@wayne.edu

## Abstract

It is important to estimate an accurate signed distance function (SDF) from a point cloud in many computer vision applications. The latest methods learn neural SDFs using either a data-driven based or an overfitting-based strategy. However, these two kinds of methods are with either poor generalization or slow convergence, which limits their capability under challenging scenarios like highly noisy point clouds. To resolve this issue, we propose a method to promote pros of both data-driven based and overfitting-based methods for better generalization, faster inference, and higher accuracy in learning neural SDFs. We introduce a novel statistical reasoning algorithm in local regions which is able to finetune data-driven based priors without signed distance supervision, clean point cloud, or point normals. This helps our method start with a good initialization, and converge to a minimum in a much faster way. Our numerical and visual comparisons with the state-of-the-art methods show our superiority over these methods in surface reconstruction and point cloud denoising on widely used shape and scene benchmarks. The code is available at <https://github.com/chenchao15/LocalN2NM>.

## 1 Introduction

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

Recent methods [62, 64, 29, 14, 95, 80, 58, 74] overcome this challenge using either a data-driven based or an overfitting-based strategy. To map a point cloud to a signed distance field, the data-driven based methods [60, 27, 36, 45, 81, 79, 22, 42, 92, 83] rely on a prior learned with signed distance supervision from a large-scale dataset, while the overfitting-based methods [28, 1, 102, 2, 99, 4, 21, 50, 18, 88] do not need signed distance supervision and just use the point cloud to infer a signed distance field. However, both of the two kinds of methods have pros and cons. The data-driven based methods can do inference fast but suffers from the need of large-scale training samples and poor

---

\*The corresponding author is Yu-Shen Liu. This work was supported by National Key R&D Program of China (2022YFC3800600), and the National Natural Science Foundation of China (62272263, 62072268), and in part by Tsinghua-Kuaishou Institute of Future Media Data.

generalization to instances that are unseen during training. Although the overfitting-based methods have a better generalization ability and do not need the large-scale signed distance supervision, they usually require a much longer time to converge during inference. The cons of these two kinds of methods dramatically limit the performance of learning neural SDFs under challenging scenarios like highly noisy point clouds. Therefore, beyond pursuing higher accuracy of SDFs, how to balance the generalization ability and the convergence efficiency is also a significant issue.

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

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

## 2 Related Works

Learning implicit functions has achieved promising performance in various tasks [62, 64, 29, 14, 95, 80, 58, 74, 30, 31, 33]. We can learn neural implicit representations from different supervision including 3D supervision [61, 69, 59, 17], multi-view images [78, 44, 38, 101, 46, 94, 63, 41, 98, 97, 25, 86, 100, 89, 84, 85], and point clouds [92, 43, 60, 27]. We briefly review the existing methods related to point clouds below.

### 2.1 Data-driven based Methods

In 3D supervision, many techniques utilize a data-driven approach to learning priors, and then apply these learned priors to infer implicit models for unseen point clouds. Some strategies focus on acquiring global priors [60, 27, 36, 45, 81, 79, 22, 42] at the shape level, whereas others aim to boost the generalization of these priors by learning local priors [92, 83, 11, 37, 6, 51] at the component or patch level. These learned priors facilitate the marching cubes algorithm [47] to reconstruct surfaces from implicit fields. The effectiveness of these methods often rely on extensive datasets, but they may not generalize well when facing with unseen point clouds that significantly deviate in geometry from training samples.

### 2.2 Overfitting-based Methods

In an effort to enhance generalization, some methods concentrate on precisely fitting neural networks to single point clouds. These methods incorporate innovative constraints [28, 1, 102, 2, 99, 4, 21], utilize gradients [50, 18, 88], employ differentiable Poisson solvers [70], or apply specially tailored priors [51, 54] to learn either signed [50, 28, 1, 102, 2, 15, 56, 13] or unsigned distance functions [18, 104, 103]. Despite achieving significant advances, these approaches typically require clean point clouds to accurately determine distance or occupancy fields around the point clouds.

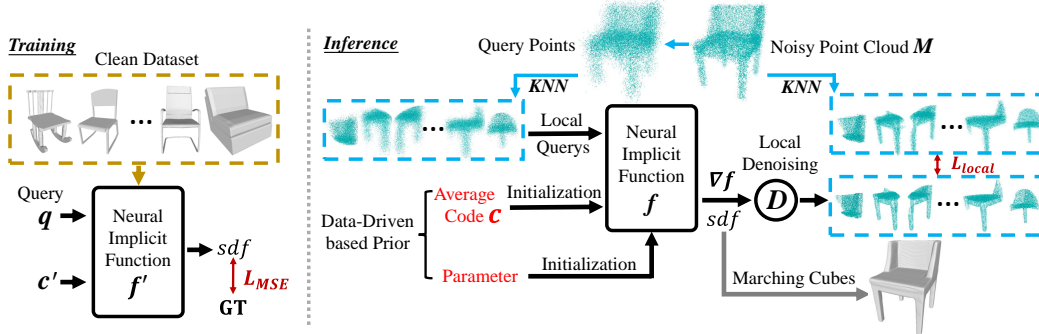


Figure 1: The overview of our method. We learn the data-driven based prior by learning a neural implicit function  $f'$  with a condition  $c'$  on a clean dataset. During inference, we employ a novel statistical reasoning algorithm to infer a neural SDF  $f$  for a noisy point cloud  $M$  with learned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

### 2.3 Learning from Noisy Point Clouds

The key to accurately reconstructing surfaces on noisy point clouds is to minimize the effect of noise in inferring implicit functions. PointCleanNet [73] was developed to filter out noise from point clouds through a data-driven approach. GPDNet [72] incorporated graph convolution based on dynamically generated neighborhood graphs to enhance noise reduction. Some other methods leveraged point cloud convolution [6], alternating latent topology [90, 57], semi-supervised strategy [106, 19], dual and integrated latent [76], or neural kernel field [91, 35] to reduce noise from point clouds. On the unsupervised front, TotalDenoising [10] adopts principles similar to Noise2Noise [40], utilizing a spatial prior suitable for unordered point clouds. DiGS [3] employs a soft constraint for unoriented point clouds. Noise2NoiseMapping [52] leverage statistical reasoning among multiple noisy point clouds with specially designed losses. Some methods using downsample-upsample frameworks [48], gradient fields [49, 9, 16, 68, 65], convolution-free intrinsic occupancy network [67], intra-shape regularization [66], eikonal equation [96, 23], neural Galerkin [34] and neural splines [93] have been implemented to further diminish noise in point clouds. Our method falls in this category, but we aim to promote the advantages of both the data-driven based and the overfitting-based strategy to pursue better generalization, faster inference, and higher accuracy.

## 3 Method

**Overview.** We aim to infer a neural SDF  $f$  from a single point cloud with noises  $M$ . Our method includes two stages shown in Fig. 1, one is to learn a prior  $f'$  in a data-driven manner, the other is to infer a neural SDF  $f$  on unseen noisy point cloud  $M$ . At the first stage, we learn a prior by training a neural SDF using ground truth signed distances of clean meshes indicated by embeddings  $c'_j$ . At the second stage, we finetune the learned prior  $f'$  to infer a neural SDF  $f$  of  $M$  using our proposed local noise to noise mapping, where the embedding  $c$  indicating  $M$  is also learned. We can use the marching cubes algorithm [47] to extract the zero-level set of  $f$  as the mesh surface of  $M$ .

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

**Data-driven Based Prior.** As shown in Fig. 1, we employ an auto-decoder similar to DeepSDF [69] for learning a prior  $f'$  in a data-driven manner and inferring a neural SDF  $f$  for single point clouds with noises, respectively. We employ a data-driven strategy to learn a prior  $f'$  from clean meshes first. Specifically, we learn  $f'$  with an embedding  $c'_j$  as a condition of queries. For each shape, we sample queries  $q$  around a shape represented by  $c'_j$ , and establish the signed distance supervision by recording the signed distance  $s$  to the ground truth mesh. Thus, we learn the prior  $f'$  by minimizing the prediction errors to the ground truth signed distances,

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

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

**Signed Distance Inference.** With the learned prior  $f'$ , we infer a neural SDF  $f$  for a single point cloud with noises  $M$ . We do not require ground truth signed distances, clean point clouds, or even point normal during the inference of  $f$ . Specifically, we infer  $f$  by finetuning parameters of  $f'$  with a learnable embedding  $c$  indicating the single point cloud with noises. The finetuning relies on a novel statistical reasoning algorithm on local regions.

The advantage of our method lies in the capability of conducting the statistical reasoning in local regions. Comparing to the global reasoning method [52], our method is able to not only infer more accurate geometry but also significantly improve the efficiency. Our method starts from randomly sampling a local region  $m_n$  on the shape  $M$ . We randomly select one point on  $M$  as the center of  $m_n$ , and set up its  $K$  nearest noisy points as a local region  $m_n$ . Then, we randomly sample  $U$  queries  $\{\bar{q}_u\}_{u=1}^U$  around  $m_n$ , and also randomly select  $U$  noisy points  $\{p_v\}_{v=1}^U$  out of  $m_n$  for statistically reasoning the surface in each iteration.

Our key idea of inferring a neural SDF  $f$  is to estimate a mean zero-level set that is consistent to all points in the local region  $m_n$ . To this end, we use the  $U$  sampled queries  $\{\bar{q}_u\}$  to represent the zero-level set in this area using  $f$ , and minimize the distances of the  $U$  noisy points  $\{p_v\}$  to the zero-level set in each iteration. Statistically, the expectation of the zero-level set should have the minimum distance to all the noisy point splitting in region  $m_n$ .

Specifically, we first project the  $U$  sampled queries  $\{\bar{q}_u\}$  onto the zero-level set of  $f$  using a differentiable pulling operation [50]. For each query  $\bar{q}_u$ , its projection on the zero-level set is,

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

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

With the pulling operation, we can use projections  $\{\bar{q}'_u\}$  of queries  $\{\bar{q}_u\}$  to approximate the zero-level set in region  $m_n$ . With a coarse zero-level set estimation, we expect this zero-level set can be consistent to various subsets of noises  $\{p_v\}$  sampled from  $m_n$ . Thus, we minimize the errors between the  $\{\bar{q}'_u\}_{u=1}^U$  and a subset of points  $\{p_v\}_{v=1}^U$  on area  $m_n$  in each optimization iteration,

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

where we learn  $f$  through finetuning the prior  $f'$  and learning the embedding  $c$  representing the noisy point cloud  $M$ . The expectation is over the local regions  $m_n$  that randomly sampled from the noisy point cloud  $M$ , and the subset patch  $p_v$  randomly sampled from each  $m_n$ . We follow the method [52] to use the EMD to evaluate the distance between the two sets of points, which leads the neural SDF  $f$  to converge on the specific noisy point cloud  $M$ .

**Initialization.** The network architecture of  $f$  is the same to the one of prior  $f'$ . We learn  $f$  with the parameters of  $f'$  as the initialization, representing the prior that we learned. For the embedding  $c$  that represents  $M$ , we initialize  $c$  as the center of the embedding space learned by the prior  $f'$  in Eq. 1, i.e.,  $c = 1/I \sum_{i=1}^I c'_i$ . This initialization is important for the accuracy and efficiency of learning  $f$  for single noisy point cloud  $M$ . This finetuning of parameters of  $f'$  also shows advantages over the auto-decoding [69] in terms of generalization and efficiency. We will justify these advantages in our experiments.

**Implementation Details.** We randomly select one point from noisy point cloud  $M$  as a center, and select its  $K = 1000$  nearest points to form a local region  $m_n$ . We also randomly sample  $U = 1000$  queries around the  $K$  noisy points for statistically reasoning. Specifically, we adopt a method

Metrics	PSR [39]	PSG [24]	R2N2 [20]	COcc [71]	SAP [70]	OCNN [87]	IMLS [45]	POCO [7]	ALTO [90]	N2NM [52]	Ours
$CD_{L1}$	0.299	0.147	0.173	0.044	0.034	0.067	0.031	0.030	0.028	0.026	<b>0.023</b>
NC	0.772	-	0.715	0.938	0.944	0.932	0.944	0.950	0.955	0.962	<b>0.973</b>
F-Score	0.612	0.259	0.400	0.942	0.975	0.800	0.983	0.984	0.985	0.991	<b>0.992</b>

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

introduced by NeuralPull [50] to sample queries around each one of the  $K$  noisy points. We use a Gaussian distribution centered at each point and set the standard deviation as the distance to the 51th nearest neighbor in the point cloud. We run the marching cubes for surface reconstruction at a resolution of 256 for shapes, and 512 for large-scale scenes.

The length of the embedding  $c$  or  $c'$  is set to 256. We use Adam optimizer for learning a neural implicit network, which is an auto-decoder similar to DeepSDF [69]. For training, we use an initial embedding learning rate of 0.0005 for updating embeddings and an auto-decoder learning rate of 0.001 for optimizing the prior network. Both learning rates are decreased by 0.5 for every 500 epochs. We train the prior network  $f'$  for 2000 epochs. For inference, we finetune the network  $f'$  for each noisy point cloud in 4000 iterations with a learning rate of 0.0001.

## 4 Experiments and Analysis

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

**Datasets and Metric.** We use eight datasets including shapes and scenes in the evaluations. For shapes, we conduct experiments under five datasets including ShapeNet [12], ABC [22], FA-MOUS [22], Surface Reconstruction Benchmark (SRB) [92] and D-FAUST [5]. For scenes, we conduct experiments under three real scan datasets including 3D Scene [105], KITTI [26], Paris-rue-Madame [75], and nuScenes [8]. We leverage L1 Chamfer Distance ( $CD_{L1}$ ), L2 Chamfer Distance ( $CD_{L2}$ ) to evaluate the error between the reconstructed surface and ground truth. We also use Normal Consistency (NC) [59] and F-Score [82] with a threshold of 1% to evaluate the normal accuracy of the reconstructed surface. In the ablation study, we also report time consumption to highlight the superiority of our data-driven based prior. For KITTI and Paris-rue-Madame datasets, due to their lack of ground truth meshes, we only report visual comparisons.

### 4.1 Surface Reconstruction for Shapes

#### Evaluation on ShapeNet.

We first report our results on shapes from ShapeNet. We report evaluations by comparing our method with the latest prior-based and overfitting-based methods in Tab 1. For prior-based methods, we compare our method with PSG [24], R2N2 [20], COcc [71], OCNN [87], IMLS [45], POCO [7], and ALTO [90]. All of these methods are pretrained to learn priors using shapes with noises in training set of ShapeNet. We also follow these meth-

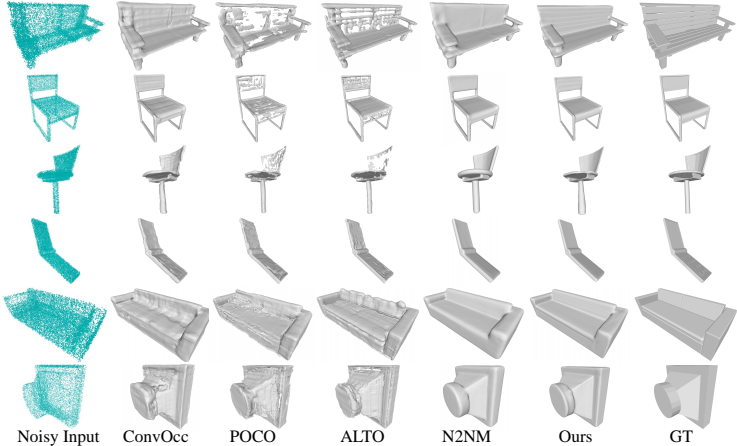


Figure 2: Comparison in surface reconstruction on ShapeNet. More visual results are provided in the appendix.

ods to use the same set of training shapes to learn our prior. For overfitting-based methods, we compare our method with PSR [39], SAP [70], and N2NM [52]. These methods did not need to learn a prior, and have the ability of inferring neural implicit functions on each shape in the testing set. We also follow these methods and report our results by finetuning our prior through overfitting on each testing shape. All the shapes for testing are corrupted with noises with a variance of 0.005.

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

Metrics	SAP [70]	N2NM [52]	Ours
Time	14 min	46 min	5 min

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

**Evaluation on ABC.** We also report our evaluations on ABC dataset in Tab. 3. We learn priors from shapes in training set, and finetune this prior for each single shape in the testing set. The numerical comparisons are conducted on the testing set of ABC dataset released by P2S [22]. It includes two versions with different

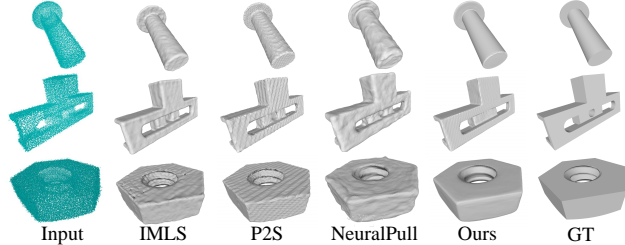


Figure 3: Comparison in surface reconstruction on ABC. More visual results are provided in the appendix.

noise levels. Similarly, we also report comparisons with prior-based methods and overfitting-based methods. With our local noise to noise mapping, we achieve the best performance over all baselines. Compared to prior-based methods, such as P2S [22], COcc [71], and POCO [7], our loss can infer more accurate geometry during the test time overfitting procedure. Also, the ability of finetuning the prior can also provide a coarse estimation and a good start for inferring neural implicit from single noisy points. Besides the accuracy, we also observe improvements on efficiency. Fig. 3 demonstrates the improvements over the baselines in terms of surface completeness and edge sharpness.

**Evaluation on SRB.** We report previous experiments using man-made objects in ShapeNet and ABC dataset, We also report our results on real scans on SRB dataset [92]. Since there is no training samples on SRB, we use the prior learned from the ShapeNet as the prior for real scans. Although the shapes in ShapeNet are not similar to shapes in SRB, we found the prior can also work well with the scans on SRB. Different from the man-made objects, real scans have unknown noises. We report the evaluations with the prior-based and overfitting-based methods in Tab. 4 and Fig. 4. The comparisons show that our method achieves the best performance in implicit surface reconstruction. Under the same experimental settings, our method can infer more accurate geometry details with our local noise to noise mapping.

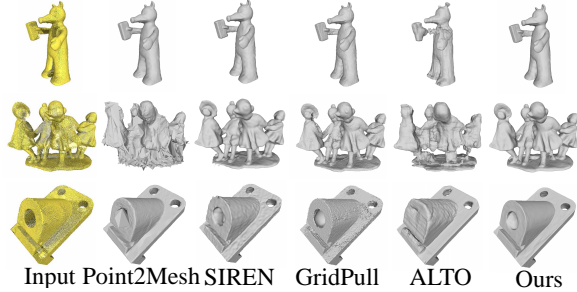


Figure 4: Comparison in surface reconstruction on SRB. More visual results are provided in the appendix.

**Evaluation on FAMOUS.** We report evaluations on more complex shapes on FAMOUS dataset. Similar to SRB, we also use the prior learned from ShapeNet. We evaluate the performance on two kinds of noises in Tab. 5. We can see that our method can recover more geometry details and achieve higher accuracy and smoother surfaces. We also report visual comparisons in Fig. 5, which also highlights our improvements in

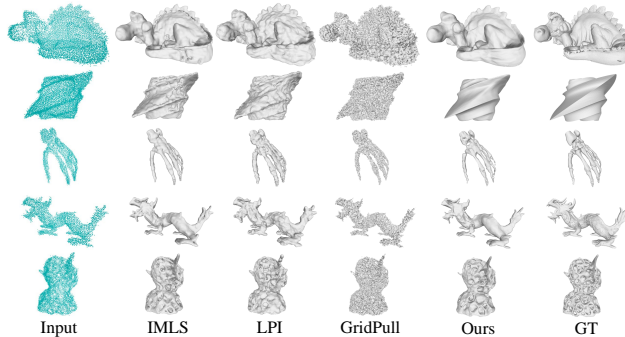


Figure 5: Comparison in surface reconstruction on FAMOUS. More visual results are provided in the appendix.

Dataset	PSR [39]	P2S [22]	COcc [71]	NP [50]	IMLS [45]	PCP [55]	POCO [7]	OnSurf [53]	N2NM [52]	Ours
ABC var	3.29	2.14	0.89	0.72	0.57	0.49	2.01	3.52	0.113	<b>0.096</b>
ABC max	3.89	2.76	1.45	1.24	0.68	0.57	2.50	4.30	0.139	<b>0.113</b>

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

Metrics	IGR [28]	Point2Mesh [32]	PSR [39]	SIREN [77]	GP [16]	ALTO [90]	Steik [96]	SAP [70]	NKSR [35]	N2NM [52]	Ours
$CD_{L1}$	0.178	0.116	0.232	0.123	0.086	0.089	0.079	0.076	0.069	0.067	<b>0.055</b>
F-Score	0.755	0.648	0.735	0.677	0.766	0.772	0.822	0.830	0.829	0.835	<b>0.860</b>

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

terms of accuracy, smoothness, completeness, and recovered sharp edges.

**Evaluation on D-FAUST.** Finally, we report our results on non-rigid shapes, i.e., humans. Different from rigid shapes in the previous experiments, humans are with more complex poses. We learn a prior from the training set, and finetuning the prior on unseen humans with different poses. We mainly compare our method with overfitting-based methods in Tab. 6. We can see that our method achieves the best performance in CD, F-Score, and comparable performance to N2NM [52] but with faster inference speed. We further show the visual comparison in Fig. 6. We can see that our method can recover more accurate geometry and poses.

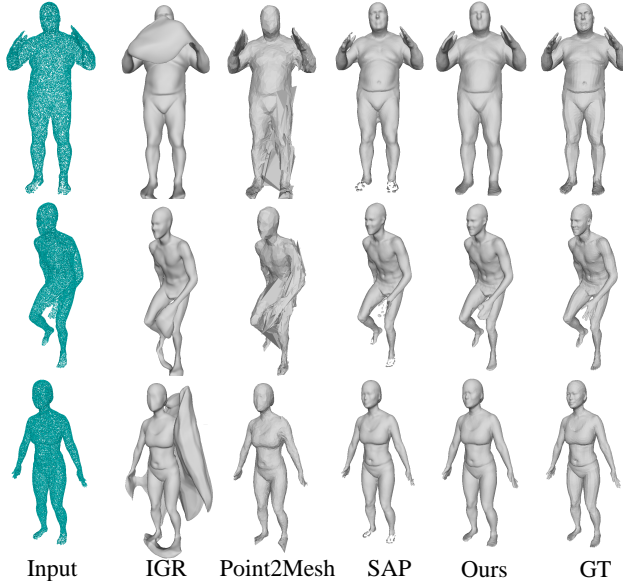


Figure 6: Comparison in surface reconstruction on D-FAUST. More visual results are provided in the appendix.

## 4.2 Surface Reconstruction for Scenes

Since we have a limited number of scenes for training, we use the prior learned from ShapeNet as the pretrained prior in our experiments on scenes. Specifically, we conduct experiments on four different scene datasets: 3D Scene [105], KITTI [26], Paris-rue-Madame [75] and nuScenes [8], where the results on nuScenes are reported in the appendix.

**Evaluation on 3D Scene.** We further evaluate our method in surface reconstruction for scenes in 3D Scene [105]. We follow previous methods LIG [37] to randomly sample 1000 points per  $m^2$ . We compare our method with the latest methods including COcc [71] and LIG [37], DeepLS [11], NeuralPull (NP) [50] and Noise2NoiseMapping (N2NM) [52]. For prior-based methods COcc [71] and LIG [37], we leverage their released pretrained models to produce the results, and we also provide them with the ground truth point normals. For overfitting-based methods DeepLS [11], NP [50] and N2NM [52], we overfit them to produce results with the same noisy point clouds. We follow LIG [37] to report  $CD_{L1}$ ,  $CD_{L2}$  and NC for evaluation. We report the comparisons in Tab. 7. The results demonstrate that our method outperforms both kinds of methods with learned priors such as LIG [37] and overfitting-based N2NM [52]. The visual comparisons in Fig. 7 show that our method can reveal more geometry details on real scans, which justifies our capability of handling noise in point clouds.

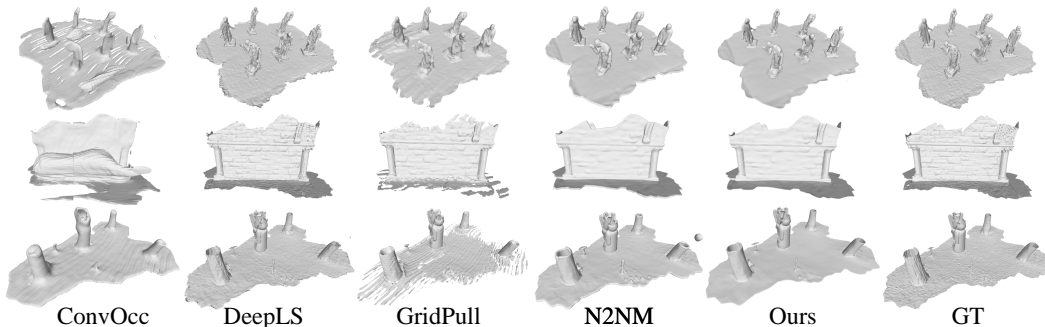


Figure 7: Comparison in surface reconstruction on 3D Scene.

Dataset	PSR [39]	NP [50]	IMLS [45]	LPI [15]	PCP [55]	POCO [7]	OnSurf [53]	GP [16]	N2NM [52]	Ours
F-var	1.80	0.28	0.80	0.19	0.07	1.50	0.59	0.13	0.033	<b>0.029</b>
F-max	3.41	0.31	0.39	0.26	0.30	2.75	3.64	0.21	0.117	<b>0.105</b>

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

Metrics	IGR [28]	Point2Mesh [32]	PSR [39]	SAP [70]	N2NM [52]	Ours
$CD_{L1} \times 10$	0.235	0.071	0.044	0.043	0.037	<b>0.034</b>
F-Score	0.805	0.855	0.966	0.966	0.966	<b>0.973</b>
NC	0.911	0.905	0.965	0.959	<b>0.970</b>	0.968

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

**Evaluation on KITTI.** Following GridPull [16], we further evaluate our method on KITTI [26] odometry dataset (Sequence 00, frame 3000 to 4000), which contains about 13.8 million points, which are split into 15 chunks. We reconstruct each of them and concatenate them together for visualization. We compare our method with the latest methods SAP [70] and GridPull [16]. As shown in Fig. 8, our method is robust to noise in real scans, successfully generalizes to large-scale scenes, and achieves visual-appealing reconstructions with more details.

**Evaluation on Paris-rue-Madame.** Following N2NM [52], we further evaluate our method on Paris-rue-Madame [75], which contains much noises. We split the 10 million points into 50 chunks each of which is used to learn a neural implicit function. We compare our method with LIG [37] and N2NM [52]. For LIG [37], we produce the results for each chunk with released pretrained models. For N2NM [52], we overfit on all chunks until convergence. As shown in Fig. 9, we achieve better performance over LIG [37] and N2NM [52] in large-scale surface reconstruction, which highlight our advantages in reconstructing complete and detailed surfaces from noisy scene point clouds.

### 4.3 Ablation Studies

We conduct ablation studies on the ABC dataset [22] to justify each module of our method.

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

Metric	128	256	512
$CD_{L2} \times 100$	0.102	<b>0.096</b>	0.114

Table 8: Effect of the embedding size.

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

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

Table 9: Effect of the prior.

**Local Region Splitting.** We further validate the effectiveness of local region splitting strategies. We employ three different splitting strategies in Tab. 10. We first

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

Table 10: Effect of splitting strategies.

split the whole space where the noisy point cloud is located uniformly into multiple voxel blocks, as shown by the result of “Voxel”. The severely degenerated results indicate that this splitting strategy is even worse than the global method N2NM [52], 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 0.1 as a local region. The result of “Sphere (Fixed Size)” slightly degenerates due to some of the spheres containing too few points. In contrast, our splitting strategy, as shown by the result of “Sphere (KNN)”, ensures that each local region has enough points to help achieve superior performance.

Metrics	COcc [71]	LIG [37]	DeepLS[11]	NP [50]	N2NM [52]	Ours
$CD_{L2} \times 1000$	14.10	6.190	1.607	2.115	0.507	<b>0.389</b>
$CD_{L1}$	0.052	0.048	0.025	0.034	0.019	<b>0.016</b>
NC	0.908	0.849	0.915	0.900	0.929	<b>0.942</b>

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.



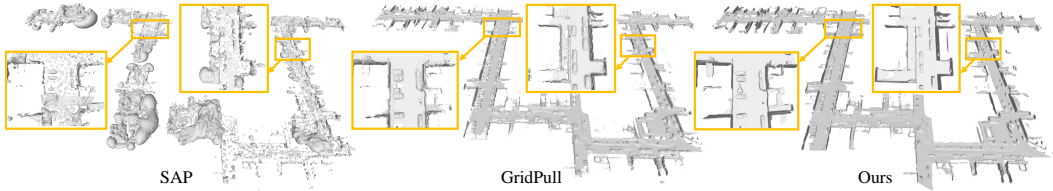


Figure 8: Comparison in surface reconstruction on KITTI.

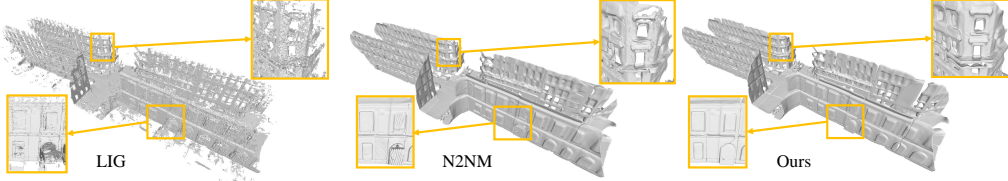


Figure 9: Comparison in surface reconstruction on Paris-rue-Madame.

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

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

Table 11: Effect of local mapping.

**Local Region Size.** We further validate the effectiveness of local region sizes (points number in a local region) in Tab. 12. We use different local region sizes including  $\{500, 1000, 3000, 5000\}$ . The results show that 1000 is the best.

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

Table 12: Effect of local region size.

**SDF initialization.** We further validate the effectiveness of different SDF initializations in Tab. 13 and Fig. 10, including random initialization, geometry initialization [1], initialization to a simple square shape, and ours.

Metric	Random	Square	Sphere (SAL)	Ours
Time	8.3min	7.1min	5.5min	<b>5.0min</b>

Table 13: The effect of SDF initialization.

We can see our prior can reconstruct more accurate surfaces from single noisy point clouds in much shorter time than any other initializations.



Figure 10: Comparison with different SDF initializations.



Figure 11: Visual results with nonuniform noises.

**Noise Type.** We report our performance with various noise types, i.e., impulse noise, quantization noise, Laplacian noise, and Gaussian noise. Visual comparison in Fig. 12 justifies that we can also handle other types of noise quite well. Moreover, we also tried more challenging cases with nonuniform noises which do not have a zero expectation across a shape, like a shape with only a half of points having noises or a shape with several patches having noises. The result in Fig. 11 shows that our method can also handle nonuniform noises well.

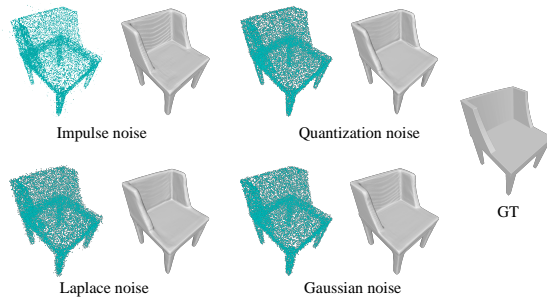


Figure 12: Visual results with different noise types.

**Noise Level.** We report our performance on point clouds with different levels of noise. As shown in Tab. 14, the noise levels of middle and max come from the ABC dataset [22]. The middle indicates noises

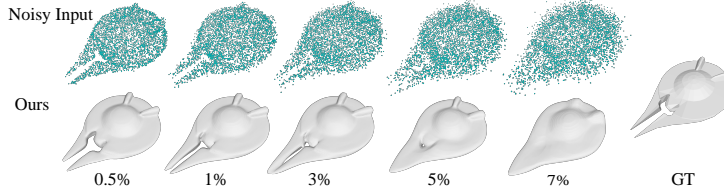


Figure 13: Visual comparison with different noise levels.

with a variance of  $0.01L$ , where  $L$  is the longest edge of the bounding box. The max indicates noises with a variance of  $0.05L$ . Our extreme noise comes with a variance of  $0.07L$ .

The  $CD_{L2}$  comparison shows that our results slightly degenerate with max and extreme noise, but still outperform N2NM [52]. The visual results in Fig. 13 indicates that our method is more robust to noises even when the noise variance is as large as 7%.

Method	Middle	Max	Extreme
N2NM [52]	0.113	0.139	0.156
Ours	<b>0.096</b>	<b>0.113</b>	<b>0.125</b>

Table 14: Effect of noise level.

**Sparsity.** We report the effect of the sparsity of noisy point clouds. We downsample the noisy point clouds to 25% and 50% of their original size to validate the impact of sparsity. The  $CD_{L2}$  results in Tab. 15 and visual comparisons in Fig. 14 indicate that our method can handle sparsity in noisy point clouds better than N2NM [52]. Since our data-driven based prior can help to learn a more complete surface and reduce the impacts brought by the sparsity.

Method	25%	50%	100%
N2NM [52]	0.154	0.133	0.113
Ours	<b>0.121</b>	<b>0.107</b>	<b>0.096</b>

Table 15: Effect of sparsity.

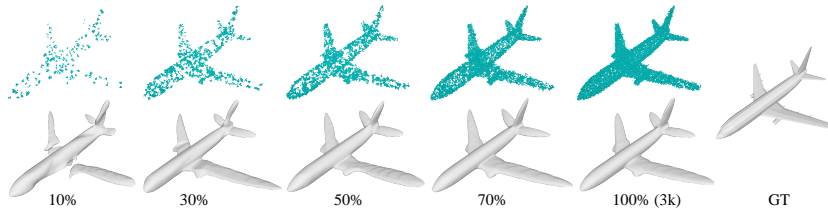


Figure 14: Visual comparison with different point numbers.

**Time Consumption.** Since our method can handle sparsity and require less time as the point number decreases, we conduct an experiment with downsampled noisy points in Tab. 16. Fig. 14 indicates that we can work well on much fewer points, and also provide an alternative of improving efficiency.

Metric	10%	30%	50%	70%	100% (3k)
Time	3.1min	3.6min	4.0min	4.5min	5.0min

Table 16: The comparison of time consumption with different point numbers.

**Optimization.** We visualize the optimization process in Fig. 15. We reconstruct meshes using the neural SDF  $f$  learned in different iterations. We

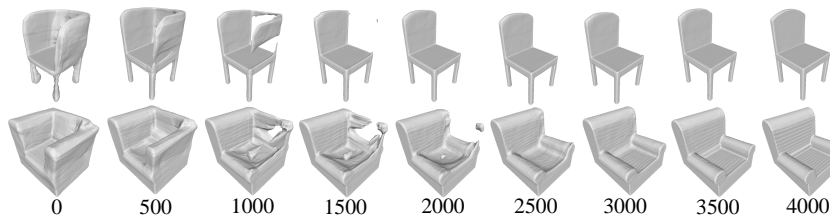


Figure 15: Optimization during inference.

see that the shape is updated progressively to the ground truth shapes.

## 5 Conclusion

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

## References

- [1] Matan Atzmon and Yaron Lipman. SAL: Sign agnostic learning of shapes from raw data. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2020.
- [2] Matan Atzmon and yaron Lipman. SALD: sign agnostic learning with derivatives. In *International Conference on Learning Representations*, 2021.
- [3] Yizhak Ben-Shabat, Chamin Hewa Koneputugodage, and Stephen Gould. DiGS: Divergence guided shape implicit neural representation for unoriented point clouds. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2022.
- [4] Yizhak Ben-Shabat, Chamin Hewa Koneputugodage, and Stephen Gould. DiGS : Divergence guided shape implicit neural representation for unoriented point clouds. *CoRR*, abs/2106.10811, 2021.
- [5] Federica Bogo, Javier Romero, Gerard Pons-Moll, and Michael J. Black. Dynamic FAUST: Registering human bodies in motion. In *IEEE Computer Vision and Pattern Recognition*, 2017.
- [6] Alexandre Boulch and Renaud Marlet. POCO: Point convolution for surface reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6302–6314, June 2022.
- [7] Alexandre Boulch and Renaud Marlet. Poco: Point convolution for surface reconstruction. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2022.
- [8] Holger Caesar, Varun Bankiti, Alex H. Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, and Oscar Beijbom. nuscenes: A multimodal dataset for autonomous driving. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020.
- [9] Ruojin Cai, Guandao Yang, Hadar Averbuch-Elor, Zekun Hao, Serge Belongie, Noah Snavely, and Bharath Hariharan. Learning gradient fields for shape generation. In *European Conference on Computer Vision*, 2020.
- [10] Pedro Hermosilla Casajus, Tobias Ritschel, and Timo Ropinski. Total denoising: Unsupervised learning of 3d point cloud cleaning. In *IEEE International Conference on Computer Vision*, pages 52–60, 2019.
- [11] Rohan Chabra, Jan Eric Lenssen, Eddy Ilg, Tanner Schmidt, Julian Straub, Steven Lovegrove, and Richard A. Newcombe. Deep local shapes: Learning local SDF priors for detailed 3D reconstruction. In *European Conference on Computer Vision*, volume 12374, pages 608–625, 2020.
- [12] Angel X. Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, Jianxiong Xiao, Li Yi, and Fisher Yu. ShapeNet: An Information-Rich 3D Model Repository. Technical Report arXiv:1512.03012 [cs.GR], Stanford University — Princeton University — Toyota Technological Institute at Chicago, 2015.
- [13] Chao Chen, Zhizhong Han, and Yu-Shen Liu. Unsupervised inference of signed distance functions from single sparse point clouds without learning priors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17712–17723, 2023.
- [14] Chao Chen, Zhizhong Han, Yu-Shen Liu, and Matthias Zwicker. Unsupervised learning of fine structure generation for 3D point clouds by 2D projections matching. In *IEEE International Conference on Computer Vision*, 2021.
- [15] Chao Chen, Yu-Shen Liu, and Zhizhong Han. Latent partition implicit with surface codes for 3d representation. In *European Conference on Computer Vision*, 2022.
- [16] Chao Chen, Yu-Shen Liu, and Zhizhong Han. GridPull: Towards scalability in learning implicit representations from 3d point clouds. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 18322–18334, 2023.
- [17] Zhiqin Chen and Hao Zhang. Learning implicit fields for generative shape modeling. *IEEE Conference on Computer Vision and Pattern Recognition*, 2019.
- [18] Julian Chibane, Aymen Mir, and Gerard Pons-Moll. Neural unsigned distance fields for implicit function learning. *arXiv*, 2010.13938, 2020.
- [19] Gene Chou, Ilya Chugunov, and Felix Heide. GenSDF: Two-stage learning of generalizable signed distance functions. In *Advances in Neural Information Processing Systems*, pages 24905–24919, 2022.
- [20] Christopher B. Choy, Danfei Xu, JunYoung Gwak, Kevin Chen, and Silvio Savarese. 3D-r2n2: A unified approach for single and multi-view 3d object reconstruction. In Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling, editors, *European Conference on Computer Vision*, volume 9912, pages 628–644, 2016.
- [21] Angela Dai and Matthias Nießner. Neural Poisson: Indicator functions for neural fields. *arXiv preprint arXiv:2211.14249*, 2022.
- [22] Philipp Erler, Paul Guerrero, Stefan Ohrhallinger, Niloy J. Mitra, and Michael Wimmer. Points2Surf: Learning implicit surfaces from point clouds. In *European Conference on Computer Vision*, 2020.

- [23] Miguel Fainstein, Viviana Siless, and Emmanuel Iarussi. DUDF: Differentiable unsigned distance fields with hyperbolic scaling. *arXiv preprint arXiv:2402.08876*, 2024.
- [24] Haoqiang Fan, Hao Su, and Leonidas J. Guibas. A point set generation network for 3D object reconstruction from a single image. In *2017 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2463–2471, 2017.
- [25] Qiancheng Fu, Qingshan Xu, Yew-Soon Ong, and Wenbing Tao. Geo-Neus: Geometry-consistent neural implicit surfaces learning for multi-view reconstruction. 2022.
- [26] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *Computer Vision and Pattern Recognition*, 2012.
- [27] Kyle Genova, Forrester Cole, Daniel Vlasic, Aaron Sarna, William T. Freeman, and Thomas Funkhouser. Learning shape templates with structured implicit functions. In *International Conference on Computer Vision*, 2019.
- [28] Amos Gropp, Lior Yariv, Niv Haim, Matan Atzmon, and Yaron Lipman. Implicit geometric regularization for learning shapes. In *International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 3789–3799, 2020.
- [29] Zhizhong Han, Chao Chen, Yu-Shen Liu, and Matthias Zwicker. DRWR: A differentiable renderer without rendering for unsupervised 3D structure learning from silhouette images. In *International Conference on Machine Learning*, 2020.
- [30] Zhizhong Han, Chao Chen, Yu-Shen Liu, and Matthias Zwicker. ShapeCaptioner: Generative caption network for 3D shapes by learning a mapping from parts detected in multiple views to sentences. In *ACM International Conference on Multimedia*, 2020.
- [31] Zhizhong Han, Xiyang Wang, Yu-Shen Liu, and Matthias Zwicker. Hierarchical view predictor: Unsupervised 3d global feature learning through hierarchical prediction among unordered views. In *Proceedings of the 29th ACM International Conference on Multimedia*, pages 3862–3871, 2021.
- [32] Rana Hanocka, Gal Metzer, Raja Giryes, and Daniel Cohen-Or. Point2mesh: a self-prior for deformable meshes. *ACM Transactions on Graphics*, 39(4):126, 2020.
- [33] Pengchong Hu and Zhizhong Han. Learning neural implicit through volume rendering with attentive depth fusion priors. In *Advances in Neural Information Processing Systems*, 2023.
- [34] Jiahui Huang, Hao-Xiang Chen, and Shi-Min Hu. A neural galerkin solver for accurate surface reconstruction. *ACM Trans. Graph.*, 41(6), 2022.
- [35] Jiahui Huang, Zan Gojcic, Matan Atzmon, Or Litany, Sanja Fidler, and Francis Williams. Neural kernel surface reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4369–4379, 2023.
- [36] Meng Jia and Matthew Kyan. Learning occupancy function from point clouds for surface reconstruction. *arXiv*, 2010.11378, 2020.
- [37] Chiyu Jiang, Avneesh Sud, Ameesh Makadia, Jingwei Huang, Matthias Nießner, and Thomas Funkhouser. Local implicit grid representations for 3D scenes. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2020.
- [38] Yue Jiang, Dantong Ji, Zhizhong Han, and Matthias Zwicker. SDFDiff: Differentiable rendering of signed distance fields for 3D shape optimization. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2020.
- [39] Michael M. Kazhdan and Hugues Hoppe. Screened poisson surface reconstruction. *ACM Transactions on Graphics*, 32(3):29:1–29:13, 2013.
- [40] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. Noise2noise: Learning image restoration without clean data. In Jennifer G. Dy and Andreas Krause, editors, *International Conference on Machine Learning*, volume 80, pages 2971–2980, 2018.
- [41] Chen-Hsuan Lin, Chaoyang Wang, and Simon Lucey. SDF-SRN: Learning signed distance 3D object reconstruction from static images. In *Advances in Neural Information Processing Systems*, 2020.
- [42] Cheng Lin, Changjian Li, Yuan Liu, Nengjun Chen, Yi-King Choi, and Wenping Wang. Point2Skeleton: Learning skeletal representations from point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4277–4286, 2021.
- [43] Minghua Liu, Xiaoshuai Zhang, and Hao Su. Meshing point clouds with predicted intrinsic-extrinsic ratio guidance. In *European Conference on Computer vision*, 2020.
- [44] Shaohui Liu, Yinda Zhang, Songyou Peng, Boxin Shi, Marc Pollefeys, and Zhaopeng Cui. DIST: Rendering deep implicit signed distance function with differentiable sphere tracing. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2020.

- [45] Shi-Lin Liu, Hao-Xiang Guo, Hao Pan, Pengshuai Wang, Xin Tong, and Yang Liu. Deep implicit moving least-squares functions for 3D reconstruction. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2021.
- [46] Shichen Liu, Shunsuke Saito, Weikai Chen, and Hao Li. Learning to infer implicit surfaces without 3D supervision. In *Advances in Neural Information Processing Systems*, 2019.
- [47] William E. Lorensen and Harvey E. Cline. Marching cubes: A high resolution 3D surface construction algorithm. *Computer Graphics*, 21(4):163–169, 1987.
- [48] Shitong Luo and Wei Hu. Differentiable manifold reconstruction for point cloud denoising. In *ACM International Conference on Multimedia*, pages 1330–1338. ACM, 2020.
- [49] Shitong Luo and Wei Hu. Score-based point cloud denoising. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4583–4592, 2021.
- [50] Baorui Ma, Zhizhong Han, Yu-Shen Liu, and Matthias Zwicker. Neural-pull: Learning signed distance functions from point clouds by learning to pull space onto surfaces. In *International Conference on Machine Learning*, 2021.
- [51] Baorui Ma, Yu-Shen Liu, and Zhizhong Han. Reconstructing surfaces for sparse point clouds with on-surface priors. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 6305–6315, 2022.
- [52] Baorui Ma, Yu-Shen Liu, and Zhizhong Han. Learning signed distance functions from noisy 3D point clouds via noise to noise mapping. In *International Conference on Machine Learning*, pages 23338–23357. PMLR, 2023.
- [53] Baorui Ma, Yu-Shen Liu, Matthias Zwicker, and Zhizhong Han. Reconstructing surfaces for sparse point clouds with on-surface priors. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2022.
- [54] Baorui Ma, Yu-Shen Liu, Matthias Zwicker, and Zhizhong Han. Surface reconstruction from point clouds by learning predictive context priors. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 6316–6327, 2022.
- [55] Baorui Ma, Yu-Shen Liu, Matthias Zwicker, and Zhizhong Han. Surface reconstruction from point clouds by learning predictive context priors. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2022.
- [56] Baorui Ma, Junsheng Zhou, Yu-Shen Liu, and Zhizhong Han. Towards better gradient consistency for neural signed distance functions via level set alignment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17724–17734, 2023.
- [57] Aihua Mao, Biao Yan, Zijing Ma, and Ying He. Denoising point clouds in latent space via graph convolution and invertible neural network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5768–5777, June 2024.
- [58] Julien N. P. Martel, David B. Lindell, Connor Z. Lin, Eric R. Chan, Marco Monteiro, and Gordon Wetzstein. ACORN: adaptive coordinate networks for neural scene representation. *CoRR*, abs/2105.02788, 2021.
- [59] Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, and Andreas Geiger. Occupancy networks: Learning 3D reconstruction in function space. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2019.
- [60] Zhenxing Mi, Yiming Luo, and Wenbing Tao. SSRNet: Scalable 3D surface reconstruction network. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2020.
- [61] Mateusz Michalkiewicz, Jhony K. Pontes, Dominic Jack, Mahsa Baktashmotlagh, and Anders P. Eriksson. Deep level sets: Implicit surface representations for 3D shape inference. *CoRR*, abs/1901.06802, 2019.
- [62] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. NeRF: Representing scenes as neural radiance fields for view synthesis. In *European Conference on Computer Vision*, 2020.
- [63] Michael Niemeyer, Lars Mescheder, Michael Oechsle, and Andreas Geiger. Differentiable volumetric rendering: Learning implicit 3D representations without 3D supervision. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2020.
- [64] Michael Oechsle, Songyou Peng, and Andreas Geiger. UNISURF: Unifying neural implicit surfaces and radiance fields for multi-view reconstruction. In *International Conference on Computer Vision*, 2021.
- [65] Amine Ouasfi and Adnane Boukhayma. Few ‘zero level set’-shot learning of shape signed distance functions in feature space. In *European Conference on Computer Vision*, 2022.
- [66] Amine Ouasfi and Adnane Boukhayma. Robustifying generalizable implicit shape networks with a tunable non-parametric model. In *Advances in Neural Information Processing Systems*, 2023.

- [67] Amine Ouasfi and Adnane Boukhayma. Mixing-denoising generalizable occupancy networks. In *International Conference on 3D Vision*, 2024.
- [68] Amine Ouasfi and Adnane Boukhayma. Unsupervised occupancy learning from sparse point cloud. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2024.
- [69] Jeong Joon Park, Peter Florence, Julian Straub, Richard Newcombe, and Steven Lovegrove. DeepSDF: Learning continuous signed distance functions for shape representation. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2019.
- [70] Songyou Peng, Chiyu "Max" Jiang, Yiyi Liao, Michael Niemeyer, Marc Pollefeys, and Andreas Geiger. Shape as points: A differentiable poisson solver. In *Advances in Neural Information Processing Systems*, 2021.
- [71] Songyou Peng, Michael Niemeyer, Lars M. Mescheder, Marc Pollefeys, and Andreas Geiger. Convolutional occupancy networks. In *European Conference on Computer Vision*, volume 12348, pages 523–540, 2020.
- [72] Francesca Pistilli, Giulia Fracastoro, Diego Valsesia, and Enrico Magli. Learning graph-convolutional representations for point cloud denoising. In *European Conference on Computer Vision*, volume 12365, pages 103–118, 2020.
- [73] Marie-Julie Rakotosaona, Vittorio La Barbera, Paul Guerrero, Niloy J. Mitra, and Maks Ovsjanikov. Pointcleannet: Learning to denoise and remove outliers from dense point clouds. *Computer Graphics Forum*, 39(1):185–203, 2020.
- [74] Konstantinos Rematas, Ricardo Martin-Brualla, and Vittorio Ferrari. Sharf: Shape-conditioned radiance fields from a single view. In *International Conference on Machine Learning*, 2021.
- [75] Andrés Serna, Beatriz Marcotegui, François Goulette, and Jean-Emmanuel Deschaud. Paris-rue-madame database - A 3D mobile laser scanner dataset for benchmarking urban detection, segmentation and classification methods. In *International Conference on Pattern Recognition Applications and Methods*, pages 819–824, 2014.
- [76] Jaehyeok Shim and Kyungdon Joo. DITTO: Dual and integrated latent topologies for implicit 3d reconstruction. *arXiv preprint arXiv:2403.05005*, 2024.
- [77] Vincent Sitzmann, Julien N.P. Martel, Alexander W. Bergman, David B. Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. In *Advances in Neural Information Processing Systems*, 2020.
- [78] Vincent Sitzmann, Michael Zollhöfer, and Gordon Wetzstein. Scene representation networks: Continuous 3D-structure-aware neural scene representations. In *Advances in Neural Information Processing Systems*, 2019.
- [79] Peng Songyou, Niemeyer Michael, Mescheder Lars, Pollefeys Marc, and Geiger Andreas. Convolutional occupancy networks. In *European Conference on Computer Vision*, 2020.
- [80] Towaki Takikawa, Joey Litalien, Kangxue Yin, Karsten Kreis, Charles Loop, Derek Nowrouzezahrai, Alec Jacobson, Morgan McGuire, and Sanja Fidler. Neural geometric level of detail: Real-time rendering with implicit 3D shapes. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2021.
- [81] Jiapeng Tang, Jiabao Lei, Dan Xu, Feiying Ma, Kui Jia, and Lei Zhang. SA-ConvONet: Sign-agnostic optimization of convolutional occupancy networks. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021.
- [82] Maxim Tatarchenko, Stephan R. Richter, Rene Ranftl, Zhuwen Li, Vladlen Koltun, and Thomas Brox. What do single-view 3D reconstruction networks learn? In *The IEEE Conference on Computer Vision and Pattern Recognition*, 2019.
- [83] Edgar Tretschk, Ayush Tewari, Vladislav Golyanik, Michael Zollhöfer, Carsten Stoll, and Christian Theobalt. PatchNets: Patch-Based Generalizable Deep Implicit 3D Shape Representations. *European Conference on Computer Vision*, 2020.
- [84] Delio Vicini, Sébastien Speierer, and Wenzel Jakob. Differentiable signed distance function rendering. *ACM Transactions on Graphics*, 41(4):125:1–125:18, 2022.
- [85] Jiepeng Wang, Peng Wang, Xiaoxiao Long, Christian Theobalt, Taku Komura, Lingjie Liu, and Wenping Wang. NeuRIS: Neural reconstruction of indoor scenes using normal priors. In *European Conference on Computer Vision*, 2022.
- [86] Peng Wang, Lingjie Liu, Yuan Liu, Christian Theobalt, Taku Komura, and Wenping Wang. NeuS: Learning neural implicit surfaces by volume rendering for multi-view reconstruction. In *Advances in Neural Information Processing Systems*, pages 27171–27183, 2021.
- [87] Peng-Shuai Wang, Yang Liu, and Xin Tong. Deep octree-based cnns with output-guided skip connections for 3d shape and scene completion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 266–267, 2020.

- [88] Ruian Wang, Zixiong Wang, Yunxiao Zhang, Shuangmin Chen, Shiqing Xin, Changhe Tu, and Wenping Wang. Aligning gradient and hessian for neural signed distance function. In *Advances in Neural Information Processing Systems*, volume 36, pages 63515–63528, 2023.
- [89] Yiqun Wang, Ivan Skorokhodov, and Peter Wonka. HF-NeuS: Improved surface reconstruction using high-frequency details. 2022.
- [90] Zhen Wang, Shijie Zhou, Jeong Joon Park, Despoina Paschalidou, Suyu You, Gordon Wetzstein, Leonidas Guibas, and Achuta Kadambi. Alto: Alternating latent topologies for implicit 3d reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 259–270, 2023.
- [91] Francis Williams, Zan Gojcic, Sameh Khamis, Denis Zorin, Joan Bruna, Sanja Fidler, and Or Litany. Neural fields as learnable kernels for 3D reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 18500–18510, 2022.
- [92] Francis Williams, Teseo Schneider, Claudio Silva, Denis Zorin, Joan Bruna, and Daniele Panozzo. Deep geometric prior for surface reconstruction. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2019.
- [93] Francis Williams, Matthew Trager, Joan Bruna, and Denis Zorin. Neural splines: Fitting 3D surfaces with infinitely-wide neural networks. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 9949–9958, 2021.
- [94] Yunjie Wu and Zhengxing Sun. DFR: differentiable function rendering for learning 3D generation from images. *Computer Graphics Forum*, 39(5):241–252, 2020.
- [95] Peng Xiang, Xin Wen, Yu-Shen Liu, Yan-Pei Cao, Pengfei Wan, Wen Zheng, and Zhizhong Han. SnowflakeNet: Point cloud completion by snowflake point deconvolution with skip-transformer. In *IEEE International Conference on Computer Vision*, 2021.
- [96] Huizong Yang, Yuxin Sun, Ganesh Sundaramoorthi, and Anthony Yezzi. StEik: Stabilizing the optimization of neural signed distance functions and finer shape representation. In *Advances in Neural Information Processing Systems*, 2023.
- [97] Lior Yariv, Jiatao Gu, Yoni Kasten, and Yaron Lipman. Volume rendering of neural implicit surfaces. In *Advances in Neural Information Processing Systems*, 2021.
- [98] Lior Yariv, Yoni Kasten, Dror Moran, Meirav Galun, Matan Atzmon, Basri Ronen, and Yaron Lipman. Multiview neural surface reconstruction by disentangling geometry and appearance. *Advances in Neural Information Processing Systems*, 33, 2020.
- [99] Wang Yifan, Shihao Wu, Cengiz Oztireli, and Olga Sorkine-Hornung. Iso-Points: Optimizing neural implicit surfaces with hybrid representations. *CoRR*, abs/2012.06434, 2020.
- [100] Zehao Yu, Songyou Peng, Michael Niemeyer, Torsten Sattler, and Andreas Geiger. MonoSDF: Exploring monocular geometric cues for neural implicit surface reconstruction. 2022.
- [101] Sergey Zakharov, Wadim Kehl, Arjun Bhargava, and Adrien Gaidon. Autolabeling 3D objects with differentiable rendering of sdf shape priors. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2020.
- [102] Wenbin Zhao, Jiabao Lei, Yuxin Wen, Jianguo Zhang, and Kui Jia. Sign-agnostic implicit learning of surface self-similarities for shape modeling and reconstruction from raw point clouds. *CoRR*, abs/2012.07498, 2020.
- [103] Junsheng Zhou, Baorui Ma, Shujuan Li, Yu-Shen Liu, and Zhizhong Han. Learning a more continuous zero level set in unsigned distance fields through level set projection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 3181–3192, 2023.
- [104] Junsheng Zhou, Baorui Ma, Yu-Shen Liu, Yi Fang, and Zhizhong Han. Learning consistency-aware unsigned distance functions progressively from raw point clouds. In *Advances in Neural Information Processing Systems*, 2022.
- [105] Qian-Yi Zhou and Vladlen Koltun. Dense scene reconstruction with points of interest. *ACM Transactions on Graphics*, 32(4):112:1–112:8, 2013.
- [106] Runsong Zhu, Di Kang, Ka-Hei Hui, Yue Qian, Shi Qiu, Zhen Dong, Linchao Bao, Pheng-Ann Heng, and Chi-Wing Fu. Ssp: Semi-signed prioritized neural fitting for surface reconstruction from unoriented point clouds. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 3769–3778, 2024.

## A Appendix

### A.1 Limitations

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

### A.2 Detailed Comparisons on 3D Scene

We detail our evaluations on each scene in 3D scene dataset in Tab. 17. The comparisons highlight our advantages in each scene.

Name	Metrics	COcc [71]	LIG [37]	DeepLS[11]	NP [50]	N2NM [52]	Ours
Burghers	$CD_{L2} \times 1000$	27.46	3.055	<b>0.401</b>	1.204	0.504	0.429
	$CD_{L1}$	0.079	0.045	0.017	0.031	0.020	<b>0.016</b>
	NC	0.907	0.835	0.920	0.905	0.925	<b>0.939</b>
Lounge	$CD_{L2} \times 1000$	9.540	9.672	6.103	1.079	0.602	<b>0.333</b>
	$CD_{L1}$	0.046	0.056	0.053	0.019	0.016	<b>0.014</b>
	NC	0.894	0.833	0.848	0.910	0.923	<b>0.935</b>
Copyroom	$CD_{L2} \times 1000$	10.97	3.610	0.609	5.795	0.442	<b>0.389</b>
	$CD_{L1}$	0.045	0.036	0.021	0.036	<b>0.016</b>	<b>0.016</b>
	NC	0.892	0.810	0.901	0.862	0.903	<b>0.916</b>
Stonewall	$CD_{L2} \times 1000$	20.46	5.032	0.320	0.983	0.330	<b>0.313</b>
	$CD_{L1}$	0.069	0.042	0.015	0.029	0.020	<b>0.015</b>
	NC	0.905	0.879	0.954	0.930	0.951	<b>0.961</b>
Totepole	$CD_{L2} \times 1000$	2.054	9.580	0.601	1.513	0.657	<b>0.482</b>
	$CD_{L1}$	0.021	0.062	<b>0.017</b>	0.054	0.023	0.020
	NC	0.943	0.887	0.950	0.893	0.945	<b>0.957</b>

Table 17: Numerical Comparisons on 3D Scene dataset in terms of  $CD_{L1}$ ,  $CD_{L2}$  and NC.

### A.3 More Results

We visualize more surface reconstruction results under ShapeNet [12], ABC [22], Surface Reconstruction Benchmark (SRB) [92], FAMOUS [22], D-FAUST [5] and nuScenes [8] in Fig. 16, Fig. 17, Fig. 18, Fig. 19, Fig. 20 and Fig. 21.



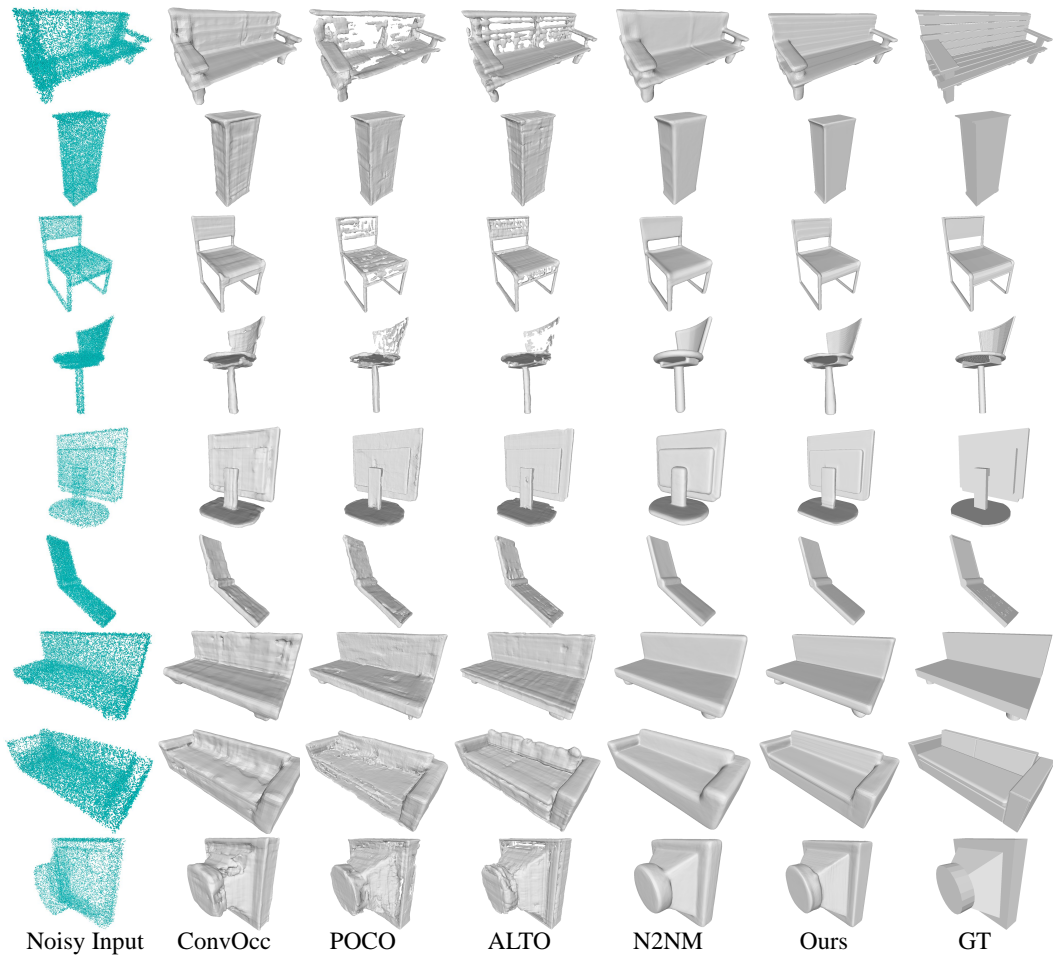


Figure 16: Comparison in surface reconstruction on ShapeNet.

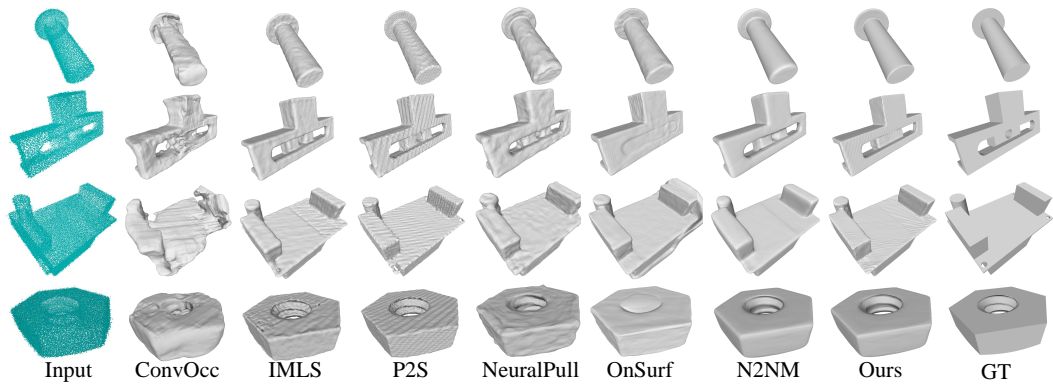


Figure 17: Comparison in surface reconstruction on ABC.

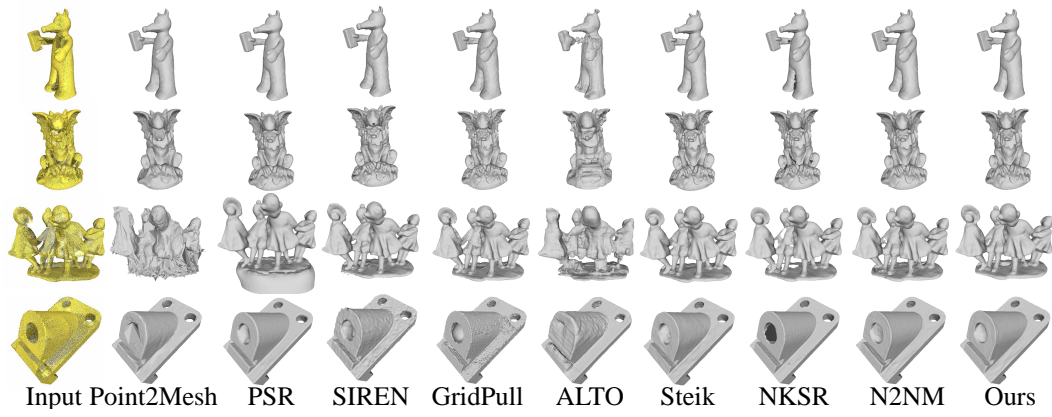


Figure 18: Comparison in surface reconstruction on SRB.

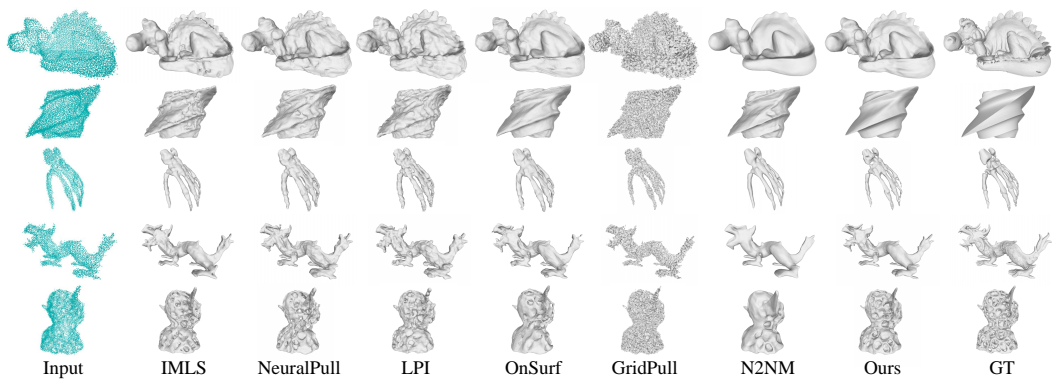


Figure 19: Comparison in surface reconstruction on FAMOUS.

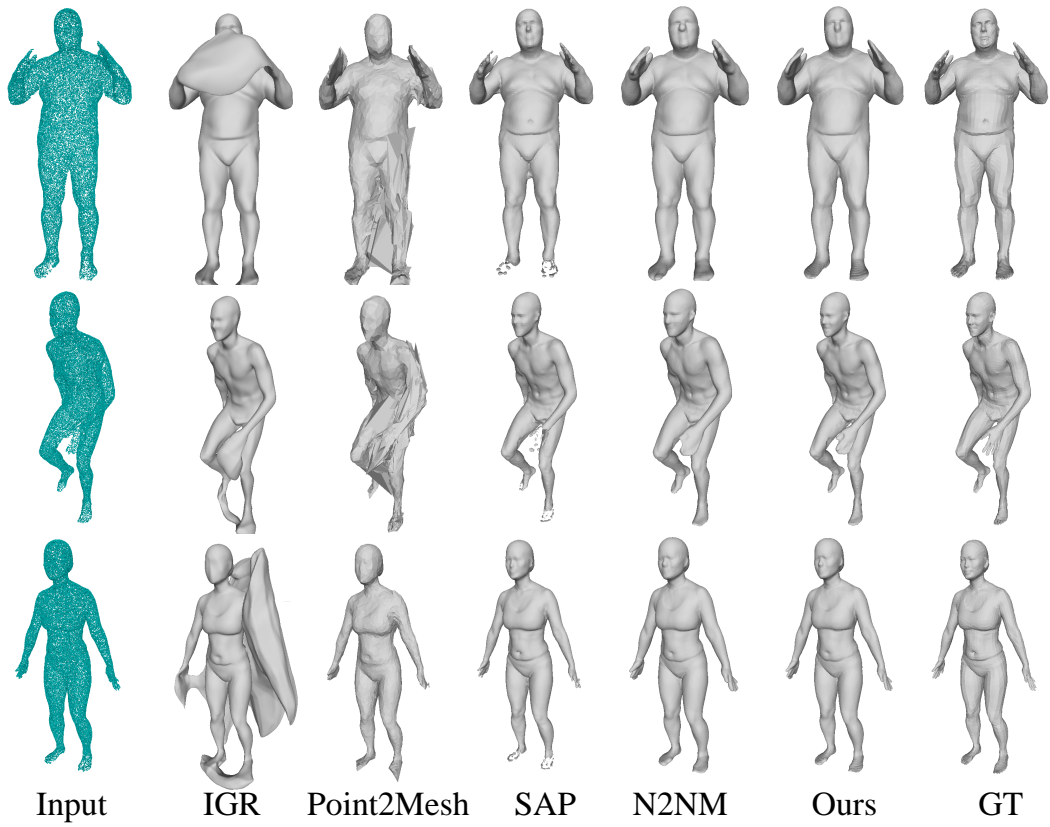


Figure 20: Comparison in surface reconstruction on D-FAUST.

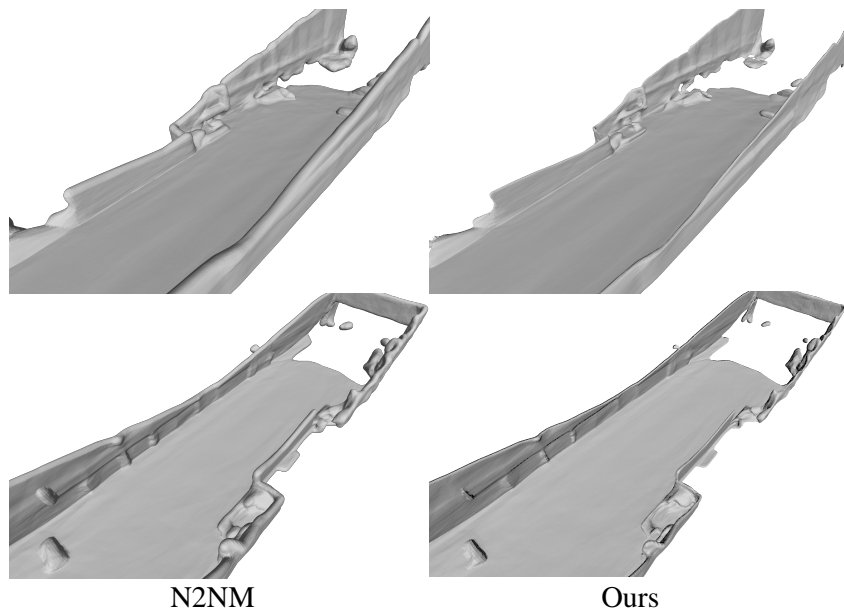


Figure 21: Comparison in surface reconstruction on nuScenes.

## NeurIPS Paper Checklist

### 1. Claims

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

Answer: [Yes]

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

Guidelines:

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

### 2. Limitations

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

Answer: [Yes]

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

Guidelines:

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

### 3. Theory Assumptions and Proofs

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

Answer: [NA]

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

Guidelines:

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

#### 4. Experimental Result Reproducibility

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

Answer: [Yes]

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

Guidelines:

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

#### 5. Open access to data and code

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [Yes]

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

Guidelines:

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

## 6. Experimental Setting/Details

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

Answer: [Yes]

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

Guidelines:

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

## 7. Experiment Statistical Significance

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

Answer: [No]

Justification: We report the average performance in terms of several metrics as the experimental results.

Guidelines:

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

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

## 8. Experiments Compute Resources

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

Answer: [Yes]

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

Guidelines:

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

## 9. Code Of Ethics

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

Answer: [Yes]

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

Guidelines:

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

## 10. Broader Impacts

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

Answer: [Yes]

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

Guidelines:

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

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

#### 11. Safeguards

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

Answer: [NA]

Justification: There is no such risk to the paper.

Guidelines:

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

#### 12. Licenses for existing assets

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

Answer: [Yes]

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

Guidelines:

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



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

### 13. **New Assets**

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

Answer: [NA]

Justification: The paper does not release new assets.

Guidelines:

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

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

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

Answer: [NA]

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

Guidelines:

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

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

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

Answer: [NA]

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

Guidelines:

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