Few '*Zero Level Set*'-Shot Learning of Shape Signed Distance Functions in Feature Space

Amine Ouasfi Adnane Boukhayma

Inria, Univ. Rennes, CNRS, IRISA, M2S, France

Supplementary Material

Metrics

We evaluate our method and baselines using popular metrics for 3D reconstruction quality assessment. We denote here by $\mathcal{S}$ and $\hat{\mathcal{S}}$ respectively the ground-truth and the generated shape meshes. The volumetric Intersection over Union (IoU) is defined as the ratio of the the intersection of the inside volumes of the meshes to their union, and it is implemented following [3] accordingly:

$$\mathrm{IoU} = \frac{|\{x \in \Omega : x \,\mathrm{inside}\, \mathcal{S} \,\mathrm{and}\, \hat{\mathcal{S}}\}|}{|\{x \in \Omega : x \,\mathrm{inside}\, \mathcal{S} \,\mathrm{or}\, \hat{\mathcal{S}}\}|}, \tag{1}$$

where $|.|$ symbolizes the cardinality of the sets, which is approximated by sampling 100k points in the bounding volume Ω. We also report two variants of the Chamfer distance representing the two-ways nearest neighbor distance between the meshes, using averaged minimal distances for CD_1 [3] and averaged minimal squared distances for CD_2 [4]:

$$\mathrm{CD}_1 = \frac{1}{2|\mathcal{S}|} \sum_{v \in \mathcal{S}} \min_{\hat{v} \in \hat{\mathcal{S}}} \|v - \hat{v}\|_2 + \frac{1}{2|\hat{\mathcal{S}}|} \sum_{\hat{v} \in \hat{\mathcal{S}}} \min_{v \in \mathcal{S}} \|\hat{v} - v\|_2, \tag{2}$$

$$\mathrm{CD}_2 = \frac{1}{2|\mathcal{S}|} \sum_{v \in \mathcal{S}} \min_{\hat{v} \in \hat{\mathcal{S}}} \|v - \hat{v}\|_2^2 + \frac{1}{2|\hat{\mathcal{S}}|} \sum_{\hat{v} \in \hat{\mathcal{S}}} \min_{v \in \mathcal{S}} \|\hat{v} - v\|_2^2. \tag{3}$$

The metrics are also approximated here with 100k samples from the source and target meshes, where distances are computed using a KD-tree following [2]. In the remainder on the paper, we report $\mathrm{CD}_2 \times 10^{-3}$ and $\mathrm{CD}_1 \times 10^{-1}$.

Noisy input point cloud

We show in Table 1 that when dealing with noisy input point clouds (e.g. variance of 0.005), our approach (i.e. meta-learning in feature space) still outperforms standard supervised learning (e.g. IF-Nets [2]). We report in this experiment reconstruction results from training and testing on the largest class of ShapeNet [1]: table, after less than a 100 meta-learning epochs.

	IoU↑		CD_1↓		CD_2↓	
Ours w/o meta learn. (i.e. IF-Nets)	0.85	0.71	0.036	0.060	0.035	0.098
Ours	**0.87**	**0.74**	**0.034**	**0.056**	**0.029**	**0.089**

Table 1: Reconstruction on class table of ShapeNet from 3000 (left) and 300 (right) noisy points voxelized at resolution $128^{\mathbf{3}}$. ($CD_1 \times 10^{-1}$, $CD_2 \times 10^{-3}$).

Per-class results for tables 1.a & 1.b in the main submission

We show next the per-class reconstruction results on the ShapeNet [1] benchmark, from 3000 input points at voxelization resolution $128^{\mathbf{3}}$, and from 300 at resolution $32^{\mathbf{3}}$. The average reconstruction scores were reported in tables 1 & 2 in the main submission. We report IoU in table 2, CD_2 in table 3, and CD_1 in table 4. We note that IF-Nets[2] is also our method without meta-learning, and MetaSDF[5] is also our method without encoder. For MetaSDF, we report the numbers for 3000 input points.

	MetaSDF (3k pts)	IF-Nets		**Ours**	
airplane	0.64	0.71	0.90	**0.78**	**0.92**
bench	0.44	0.44	0.82	**0.59**	**0.86**
cabinet	0.65	0.66	0.81	**0.70**	**0.82**
car	0.77	0.78	0.91	**0.81**	**0.91**
chair	0.55	0.63	0.88	**0.71**	**0.90**
display	0.66	0.69	0.92	**0.81**	**0.95**
lamp	0.40	0.52	0.83	**0.54**	**0.85**
phone	0.83	0.79	0.94	**0.90**	**0.97**
rifle	0.58	0.63	0.87	**0.72**	**0.90**
sofa	0.79	0.80	0.94	**0.86**	**0.96**
speaker	0.73	0.75	0.89	**0.79**	**0.90**
table	0.55	0.56	0.85	**0.69**	**0.90**
watercraft	0.65	0.69	0.90	**0.74**	**0.92**
mean	0.63	0.67	0.88	**0.74**	**0.91**

Table 2: Reconstruction IoU (↑) on ShapeNet from 3000 points voxelized at resolution $128^{\mathbf{3}}$ (right column), and 300 points voxelized at resolution $32^{\mathbf{3}}$ (left column).

References

1. Chang, A.X., Funkhouser, T., Guibas, L., Hanrahan, P., Huang, Q., Li, Z., Savarese, S., Savva, M., Song, S., Su, H., et al.: Shapenet: An information-rich 3d model repository. arXiv preprint arXiv:1512.03012 (2015) 1, 2
2. Chibane, J., Alldieck, T., Pons-Moll, G.: Implicit functions in feature space for 3d shape reconstruction and completion. In: CVPR (2020) 1, 2
3. Mescheder, L., Oechsle, M., Niemeyer, M., Nowozin, S., Geiger, A.: Occupancy networks: Learning 3d reconstruction in function space. In: CVPR (2019) 1
4. Park, J.J., Florence, P., Straub, J., Newcombe, R., Lovegrove, S.: Deepsdf: Learning continuous signed distance functions for shape representation. In: CVPR (2019) 1

	MetaSDF (3k pts)	IF-Nets		**Ours**	
airplane	0.360	0.097	0.013	**0.067**	**0.006**
bench	0.407	0.369	0.015	**0.262**	**0.010**
cabinet	0.463	0.401	0.123	**0.234**	**0.114**
car	0.207	0.139	0.020	**0.110**	**0.019**
chair	0.657	**0.215**	0.021	0.289	**0.016**
display	0.682	0.151	0.019	**0.086**	**0.014**
lamp	2.009	**0.484**	0.027	0.655	**0.022**
phone	0.156	0.086	0.010	**0.038**	**0.008**
rifle	0.212	0.066	0.013	**0.044**	**0.006**
sofa	0.227	0.186	0.016	**0.101**	**0.014**
speaker	0.645	0.339	0.101	**0.295**	**0.084**
table	0.550	0.169	0.021	**0.111**	**0.015**
watercraft	0.394	**0.308**	0.013	0.422	**0.010**
mean	0.458	0.232	0.032	**0.209**	**0.026**

Table 3: Reconstruction $\mathrm{CD}_2 \times 10^{-3}$ (↓) on ShapeNet from 3000 points voxelized at resolution 128^3 (right column), and 300 points voxelized at resolution 32^3 (left column).

	MetaSDF (3k pts)	IF-Nets		**Ours**	
airplane	0.099	0.061	0.021	**0.047**	**0.019**
bench	0.120	0.097	0.028	**0.072**	**0.024**
cabinet	0.122	0.116	0.053	**0.083**	**0.048**
car	0.095	0.082	0.033	**0.067**	**0.032**
chair	0.156	0.096	0.033	**0.076**	**0.030**
display	0.133	0.091	0.032	**0.061**	**0.027**
lamp	0.248	**0.120**	0.030	0.121	**0.027**
phone	0.067	0.073	0.027	**0.040**	**0.021**
rifle	0.081	0.059	0.020	**0.043**	**0.016**
sofa	0.100	0.090	0.032	**0.061**	**0.028**
speaker	0.151	0.116	0.047	**0.093**	**0.045**
table	0.089	0.127	0.034	**0.063**	**0.029**
watercraft	0.119	0.094	0.026	**0.080**	**0.022**
mean	0.123	0.091	0.032	**0.070**	**0.028**

Table 4: Reconstruction $\mathrm{CD}_1 \times 10^{-1}$ (↓) on ShapeNet from 3000 points voxelized at resolution 128^3 (right column), and 300 points voxelized at resolution 32^3 (left column).

5. Sitzmann, V., Chan, E.R., Tucker, R., Snavely, N., Wetzstein, G.: Metasdf: Meta-learning signed distance functions. In: NeurIPS (2020) 2