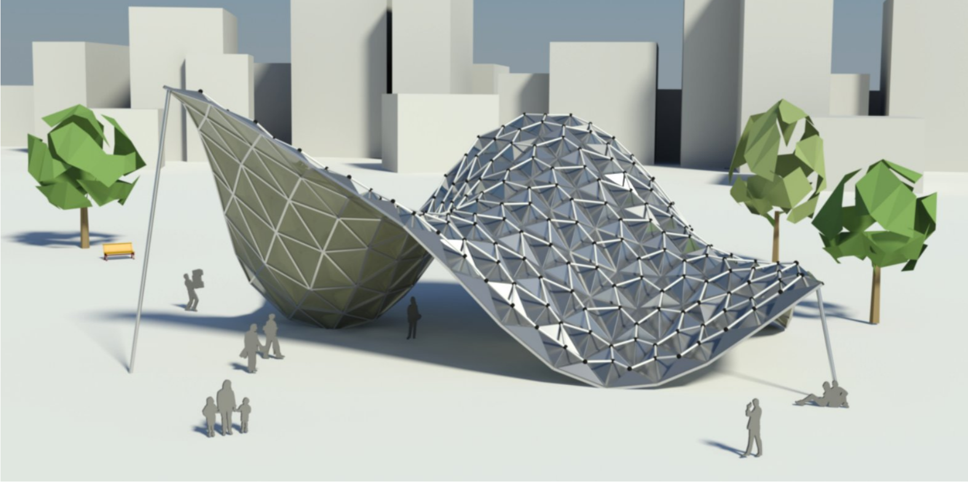


Sign Agnostic Learning with Derivatives of 3D Geometry from Raw Data

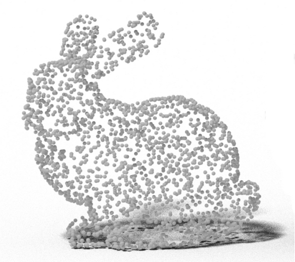
Debabrata Ghosh

Visual Computing Institute

RWTH AACHEN UNIVERSITY

Representations of 3D shapes

Discrete Representations



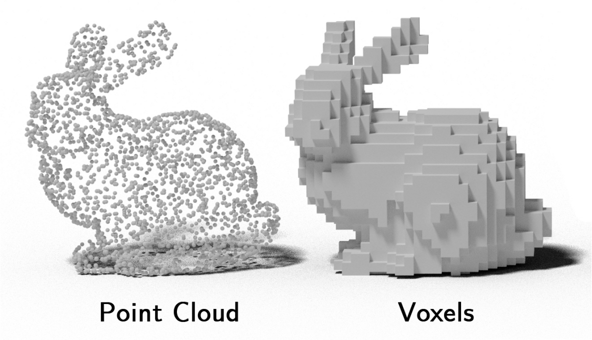
Point Cloud

Debabrata Ghosh
SALD
May 6, 2023

Visual Computing Institute
RWTH AACHEN UNIVERSITY

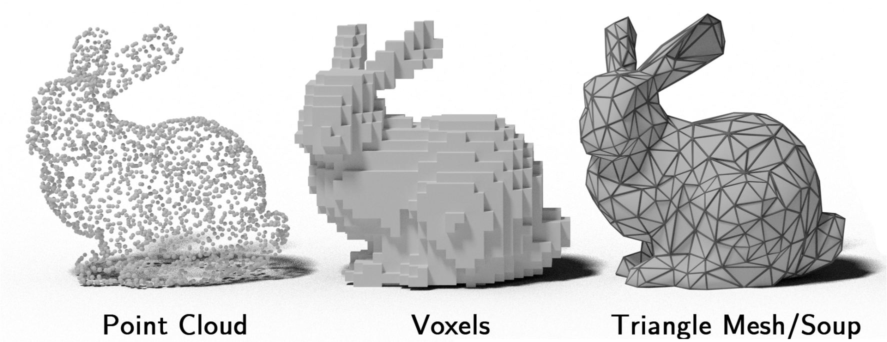
Representations of 3D shapes

Discrete Representations



Representations of 3D shapes

Discrete Representations



Representations of 3D shapes

Continuous Representations

- Parametric Representation

Debabrata Ghosh
SALD

Representations of 3D shapes

Continuous Representations

- Parametric Representation
  - $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$

Representations of 3D shapes

Continuous Representations

- Parametric Representation
  - $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$
  - Surface represented as a parametric mapping to 3D

Representations of 3D shapes

Continuous Representations

- Parametric Representation
  - $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$
  - Surface represented as a parametric mapping to 3D
- Implicit Representation

Representations of 3D shapes

Continuous Representations

- Parametric Representation
  - $f : \mathbb{R}^2 \to \mathbb{R}^3$
  - Surface represented as a parametric mapping to 3D
- Implicit Representation
  - $f : \mathbb{R}^3 \to \mathbb{R}$

Representations of 3D shapes

Continuous Representations

- Parametric Representation
  - $f : \mathbb{R}^2 \to \mathbb{R}^3$
  - Surface represented as a parametric mapping to 3D
- Implicit Representation
  - $f : \mathbb{R}^3 \to \mathbb{R}$
  - Surface represented as zero level-set: $\mathcal{S} = \left\{ \boldsymbol{x} \in \mathbb{R}^3 \mid f(\boldsymbol{x}) = 0 \right\}$

Representations of 3D shapes

Continuous Representations

- Parametric Representation
  - $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$
  - Surface represented as a parametric mapping to 3D
- Implicit Representation
  - $f : \mathbb{R}^3 \rightarrow \mathbb{R}$
  - Surface represented as zero level-set: $\mathcal{S} = \{\boldsymbol{x} \in \mathbb{R}^3 \mid f(\boldsymbol{x}) = 0\}$
  - We refer to such representation as Implicit Neural Representation (INR)

Implicit Neural Representation

... why Neural Networks?

Current works to learn INR show a recent trend of using Neural networks (NN) to model and reconstruct 3D surfaces. This stems from

Implicit Neural Representation

... why Neural Networks?

Current works to learn INR show a recent trend of using Neural networks (NN) to model and reconstruct 3D surfaces. This stems from

- flexibility and approximation power of NNs [3]

Implicit Neural Representation

... why Neural Networks?

Current works to learn INR show a recent trend of using Neural networks (NN) to model and reconstruct 3D surfaces. This stems from

- flexibility and approximation power of NNs [3]
- optimization and generalization properties of NNs

Debabrata Ghosh
SALD
May 6, 2023

Training Implicit Neural Representation

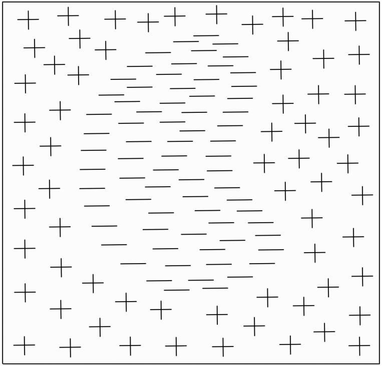
Methods and Challenges

Given raw geometric data as input, $\mathcal{X} \subset \mathbb{R}^3$, the idea is to try optimizing the weights $\boldsymbol{\theta} \in \mathbb{R}^m$ of a network $f(\boldsymbol{x}; \boldsymbol{\theta})$, where $f : \mathbb{R}^3 \times \mathbb{R}^m \to \mathbb{R}$, so that zero level-set of f represents the surface approximating $\mathcal{X}$.

Debabrata Ghosh
SALD
May 6, 2023

Training Implicit Neural Representation

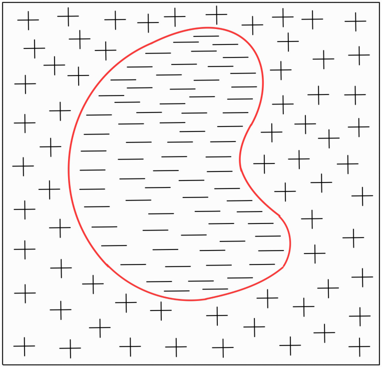
Methods and Challenges



Data for supervised learning

Debabrata Ghosh
SALD
May 6, 2023

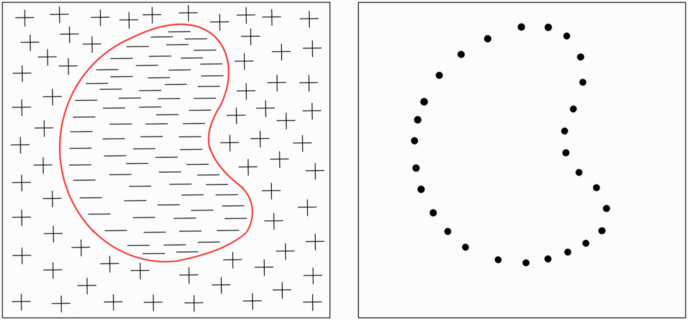
Methods and Challenges



Data for supervised learning

Training Implicit Neural Representation

Methods and Challenges

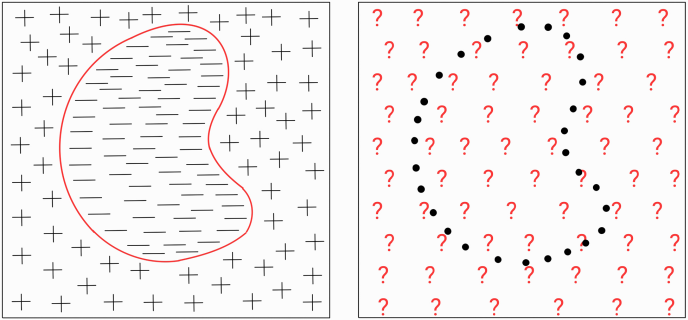


Data for supervised learning

Raw Data

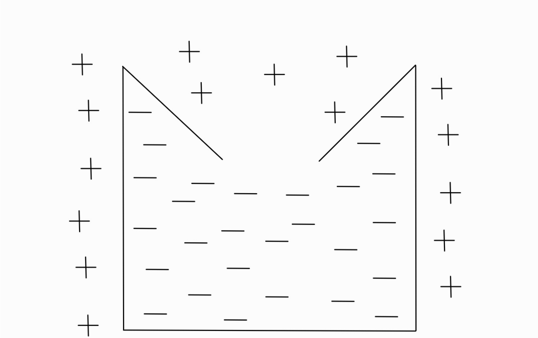
Debabrata Ghosh
SALD
May 6, 2023

Methods and Challenges



Data for supervised learning | Raw Data

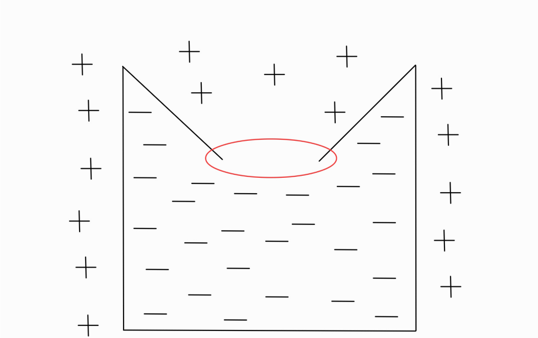
Methods and Challenges



How to learn missing parts?

Training Implicit Neural Representation

Methods and Challenges

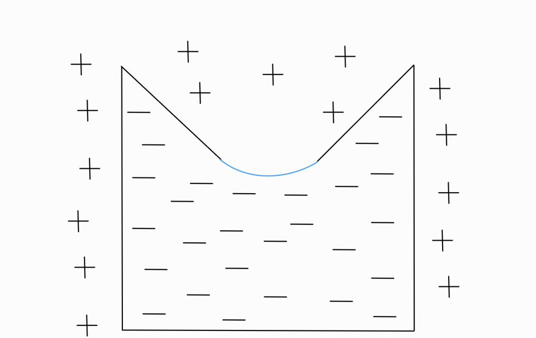


How to learn missing parts?

Debabrata Ghosh
SALD

Training Implicit Neural Representation

Methods and Challenges

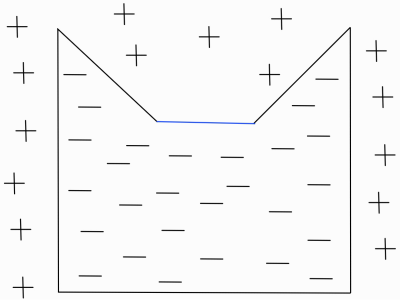


How to learn missing parts?

Debabrata Ghosh
SALD
May 6, 2023

Training Implicit Neural Representation

Methods and Challenges



How to learn missing parts?

Debabrata Ghosh
SALD

Visual Computing Institute
RWTH AACHEN UNIVERSITY

Previous Work

The division is based on the kind of 3D surface representation used.

Previous Work

The division is based on the kind of 3D surface representation used.

- Parametric Representation

Previous Work

The division is based on the kind of 3D surface representation used.

- Parametric Representation
- Implicit Representation

Previous Work

Parametric Representation

Debabrata Ghosh

Visual Computing Institute
RWTH AACHEN UNIVERSITY

Previous Work

Parametric Representation

- The most fundamental surface representation is an atlas, which is a collection of parametric charts $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ with certain coverage and transition properties [7].

Previous Work

Parametric Representation

- The most fundamental surface representation is an atlas, which is a collection of parametric charts $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ with certain coverage and transition properties [7].
- Groueix et al. reformulated this idea using a neural network to represent the surface as a union of such charts [12].

Debabrata Ghosh
SALD
May 6, 2023

Previous Work

Parametric Representation

- The most fundamental surface representation is an atlas, which is a collection of parametric charts $f : \mathbb{R}^2 \to \mathbb{R}^3$ with certain coverage and transition properties [7].
- Groueix et al. reformulated this idea using a neural network to represent the surface as a union of such charts [12].
- Williams et al. improved this construction by introducing better transitions between charts [17].

Previous Work

Implicit Representation

Previous Work

Implicit Representation

- Convolutional neural network predicting scalar values over a predefined fixed volumetric structure (e.g., grid or octree) in space [16, 18].

Debabrata Ghosh
SALD

Visual Computing Institute
RWTH AACHEN UNIVERSITY

Previous Work

Implicit Representation

- Convolutional neural network predicting scalar values over a predefined fixed volumetric structure (e.g., grid or octree) in space [16, 18].
- Multilayer Perceptron of the form $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ defining a continuous volumetric function [15, 14, 10].

Sign Agnostic Learning (SAL): Motivation

- A method for learning implicit neural representations of surfaces directly from raw $3\mathrm{D}$ data.

Background

Sign Agnostic Learning (SAL): Motivation

- A method for learning implicit neural representations of surfaces directly from raw $3\mathrm{D}$ data.
- By learning directly from raw data, e.g., inconsistently oriented point clouds or triangle soups [8] and raw scans [6] the need to obtain a ground truth signed distance representation of surfaces to be learned can be avoided.

Sign Agnostic Learning (SAL): Motivation

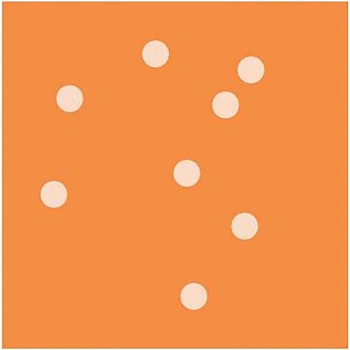
- A method for learning implicit neural representations of surfaces directly from raw $3\mathrm{D}$ data.
- By learning directly from raw data, e.g., inconsistently oriented point clouds or triangle soups [8] and raw scans [6] the need to obtain a ground truth signed distance representation of surfaces to be learned can be avoided.
- Working with complex models with inconsistent normals and/or missing parts becomes easier.

Sign Agnostic Learning (SAL): Method

Unsigned Distance Function

- Given raw geometric data, $\mathcal{X} \subset \mathbb{R}^3$, unsigned distance function can be formulated as either of:
  - standard L^2 (Euclidean) distance, $h_2(\boldsymbol{z}) = \min_{\boldsymbol{x} \in \mathcal{X}} \|\boldsymbol{z} - \boldsymbol{x}\|_2$
  - L^0 distance, $h_0(\boldsymbol{z}) = \begin{cases} 0 & \boldsymbol{z} \in \mathcal{X} \\ 1 & \boldsymbol{z} \notin \mathcal{X} \end{cases}$.

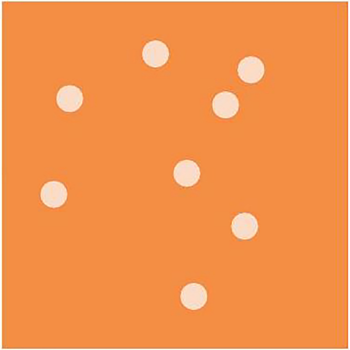
Sign Agnostic Learning (SAL): Method



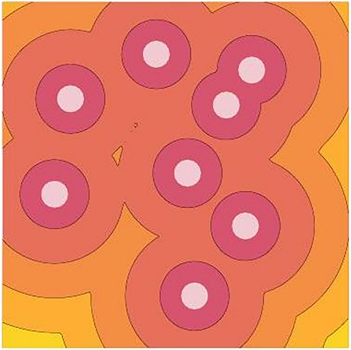
Unsigned L^0 distance (2D point cloud in gray)

Background

Sign Agnostic Learning (SAL): Method



Unsigned L^0 distance



Unsigned L^2 Distance

(2D point cloud in gray)

Background

Sign Agnostic Learning (SAL): Method

Background

Sign Agnostic Learning (SAL): Method

Unsigned Similarity Function

- $\tau : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ is a differentiable unsigned similarity function defined by the following properties:
  1. Sign agnostic: $\tau(-a, b) = \tau(a, b), \forall a \in \mathbb{R}, b \in \mathbb{R}_+$.
  2. Monotonic: $\frac{\partial \tau}{\partial a}(a, b) = g(a - b), \forall a, b \in \mathbb{R}_+$, where $g : \mathbb{R} \to \mathbb{R}$ is a monotonically increasing function with $g(0) = 0$ [1].

Sign Agnostic Learning (SAL): Method

Unsigned Similarity Function

- $\tau : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ is a differentiable unsigned similarity function defined by the following properties:
  1. Sign agnostic: $\tau(-a, b) = \tau(a, b), \forall a \in \mathbb{R}, b \in \mathbb{R}_+$.
  2. Monotonic: $\frac{\partial \tau}{\partial a}(a, b) = g(a - b), \forall a, b \in \mathbb{R}_+$, where $g : \mathbb{R} \to \mathbb{R}$ is a monotonically increasing function with $g(0) = 0$ [1].
- e.g. $\tau_\ell(a, b) = ||a| - b|^\ell$ where $\ell \geq 1$
  - τ_ℓ satisfies sign-agnostic property due to the symmetry of $|\cdot|$.
  - Since $\frac{\partial \tau}{\partial a} = \ell||a| - b|^{\ell-1}\operatorname{sign}(a - b\operatorname{sign}(a))$ it satisfies monotonicity as well.

Background

Sign Agnostic Learning (SAL): Method

SAL Loss Function

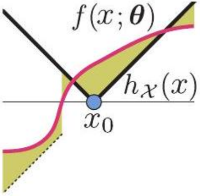
$$\mathrm{loss}(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{x}\sim\mathcal{D}}\tau\left(f(\boldsymbol{x};\boldsymbol{\theta}), h_{\mathcal{X}}(\boldsymbol{x})\right),$$

where $\mathcal{D}$ is a probability distribution defined by the input data $\mathcal{X}$; $h_{\mathcal{X}}(\boldsymbol{x})$ is some unsigned distance measure to $\mathcal{X}$; and τ is a differentiable unsigned similarity function.

Sign Agnostic Learning (SAL): Method

SAL Loss Function

- Illustrating one dimensional case: $\mathcal{X} = \{x_0\}$, $h_{\mathcal{X}}(x) = |x - x_0|$, and $\tau(a, b) = ||a| - b|$.
- When the network parameters $\boldsymbol{\theta} = \boldsymbol{\theta}^0$ are initialized properly, the minimizer $\boldsymbol{\theta}^*$ of loss defines an implicit function $f(x; \boldsymbol{\theta}^*)$ that realizes a signed version of $h_{\mathcal{X}}$; in this case $f(x; \boldsymbol{\theta}^*) = x - x_0$.



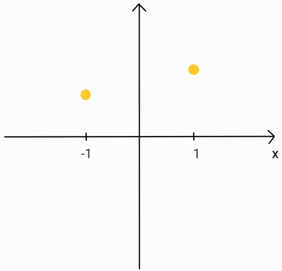
Sobolev Training

Consider a NN with one-hidden layer and the target function $f\left(x;\boldsymbol{\theta}\right)=\max\{ax,bx\}+c$ with data given for $x=\{-1,1\}$:

Background

Sobolev Training

Consider a NN with one-hidden layer and the target function $f(x;\boldsymbol{\theta}) = \max\{ax, bx\} + c$ with data given for $x = \{-1, 1\}$:

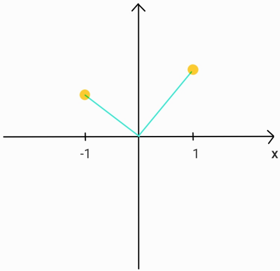


Debabrata Ghosh

Visual Computing Institute
RWTH AACHEN UNIVERSITY

Sobolev Training

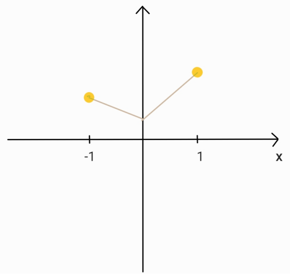
Consider a NN with one-hidden layer and the target function $f\left(x;\boldsymbol{\theta}\right)=\max\{ax,bx\}+c$ with data given for $x=\{-1,1\}$:



Background

Sobolev Training

Consider a NN with one-hidden layer and the target function $f\left(x;\boldsymbol{\theta}\right)=\max\{ax,bx\}+c$ with data given for $x=\{-1,1\}$:



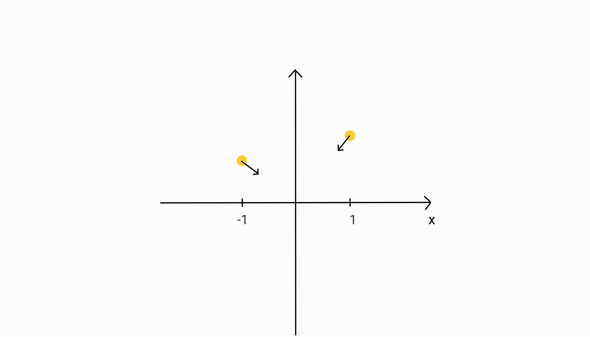
Visual Computing Institute

RWTH AACHEN UNIVERSITY

Background

Sobolev Training

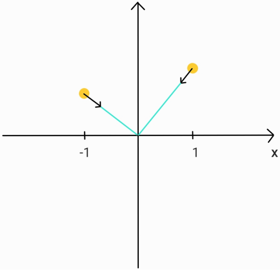
Consider a NN with one-hidden layer and the target function $f\left(x;\boldsymbol{\theta}\right) = \max\{ax, bx\} + c$ with data given for $x = \{-1, 1\}$:



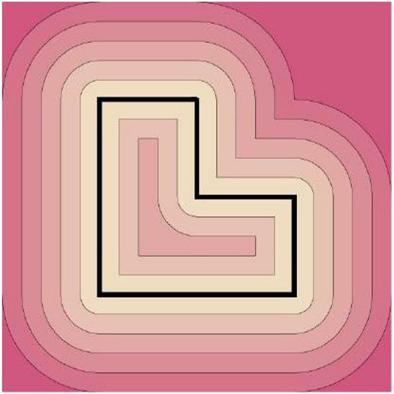
Debabrata Ghosh

Sobolev Training

Consider a NN with one-hidden layer and the target function $f\left(x;\boldsymbol{\theta}\right)=\max\{ax,bx\}+c$ with data given for $x=\{-1,1\}$:

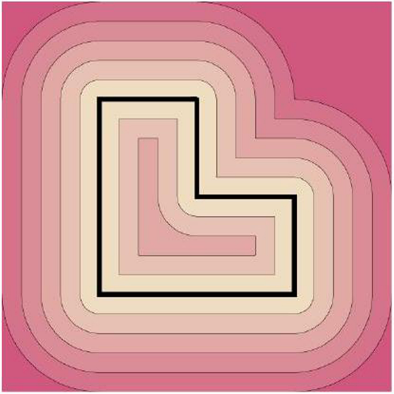


Limitations of SAL

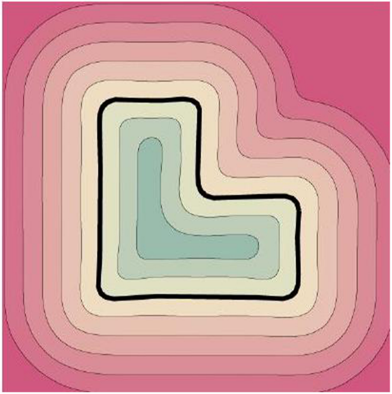


Unsigned distance

Limitations of SAL



Unsigned distance

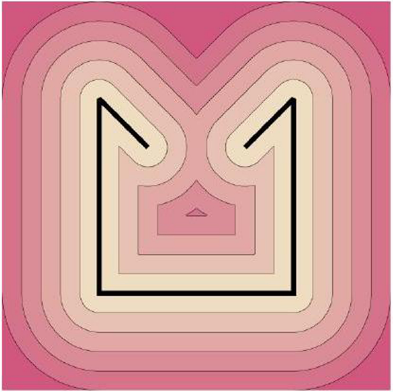


Learned from SAL

Debabrata Ghosh

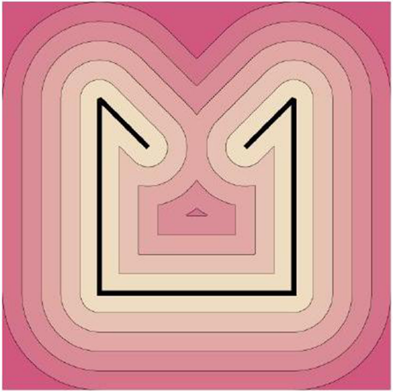
Visual Computing Institute | RWTH AACHEN UNIVERSITY

Limitations of SAL

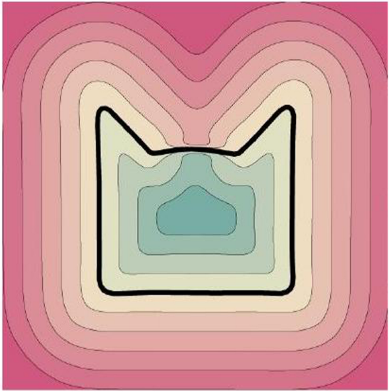


Unsigned distance

Limitations of SAL



Unsigned distance



Learned from SAL

Sign Agnostic Learning with Derivatives (SALD)

Extension of SAL Loss

$$\text{loss}(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{x}\sim\mathcal{D}}\tau(f(\boldsymbol{x};\boldsymbol{\theta}), h_{\mathcal{X}}(\boldsymbol{x})) + \lambda\mathbb{E}_{\boldsymbol{x}\sim\mathcal{D}'}\tau\left(\nabla_{\boldsymbol{x}} f(\boldsymbol{x};\boldsymbol{\theta}), \nabla_{\boldsymbol{x}} h_{\mathcal{X}}(\boldsymbol{x})\right)$$

where $\lambda > 0$ is a parameter, $\mathcal{D}'$ is a probability distribution, and $\nabla_{\boldsymbol{x}} f(\boldsymbol{x};\boldsymbol{\theta}), \nabla_{\boldsymbol{x}} h_{\mathcal{X}}(\boldsymbol{x})$ are the gradients of f, h (resp.) with respect to their input $\boldsymbol{x}$.

Debabrata Ghosh
SALD
May 6, 2023

Sign Agnostic Learning with Derivatives (SALD)

Minimal Surface Property

SALD loss possesses a minimal surface property [19] meaning the learning process using SALD loss tries to minimize the surface area of missing parts.

Debabrata Ghosh

Sign Agnostic Learning with Derivatives (SALD)

Minimal Surface Property

SALD loss possesses a minimal surface property [19] meaning the learning process using SALD loss tries to minimize the surface area of missing parts.

Theorem

For two points $x_1, x_2 \in \mathbb{R}^2$, between them a straight line is the strict global minimizer when using SALD loss.

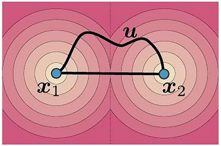
Sign Agnostic Learning with Derivatives (SALD)

Minimal Surface Property

SALD loss possesses a minimal surface property [19] meaning the learning process using SALD loss tries to minimize the surface area of missing parts.

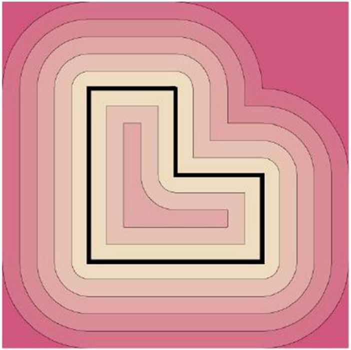
Theorem

For two points $x_1, x_2 \in \mathbb{R}^2$, between them a straight line is the strict global minimizer when using SALD loss.



Sign Agnostic Learning with Derivatives (SALD)

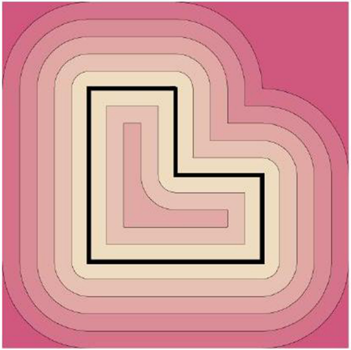
Results in 2D



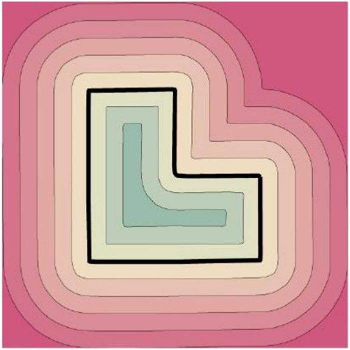
Unsigned distance

Sign Agnostic Learning with Derivatives (SALD)

Results in 2D



Unsigned distance



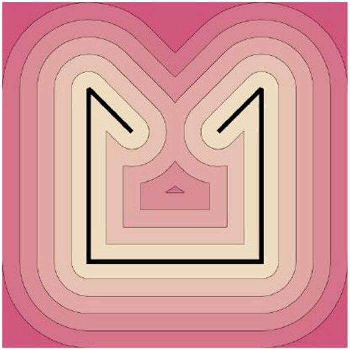
Learned from SALD

Debabrata Ghosh

Visual Computing Institute
RWTH AACHEN UNIVERSITY

Sign Agnostic Learning with Derivatives (SALD)

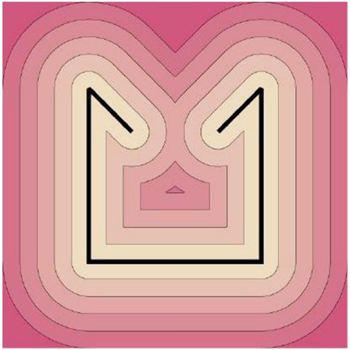
Results in 2D



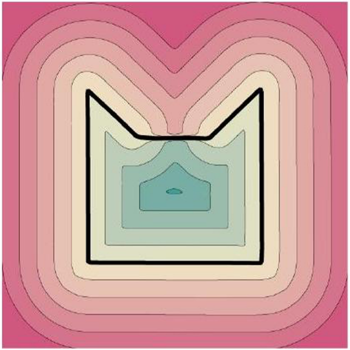
Unsigned distance

Sign Agnostic Learning with Derivatives (SALD)

Results in 2D

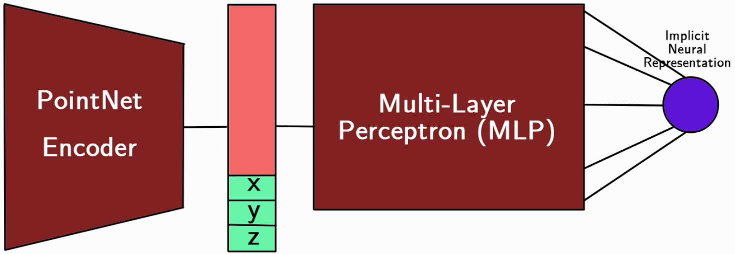


Unsigned distance



Learned from SALD

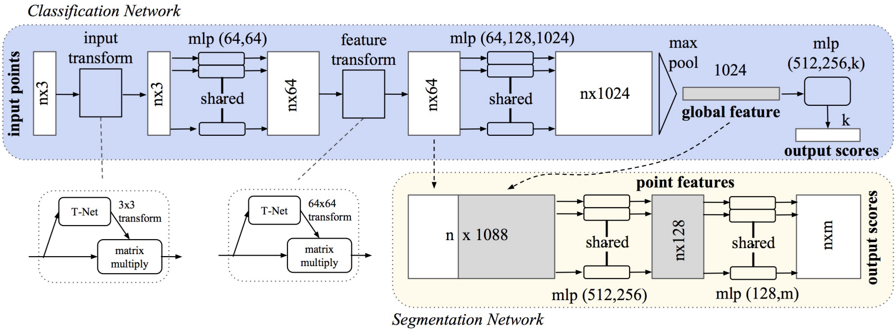
Neural Network Architecture



Variational Auto-Encoder

Debabrata Ghosh
SALD
May 6, 2023

Neural Network Architecture



PointNet Architecture

Debabrata Ghosh
SALD
May 6, 2023

Neural Network Architecture

- A modified variational encoder-decoder [13] used, where the encoder $(\boldsymbol{\mu}, \boldsymbol{\eta}) = g\left(\boldsymbol{X}; \boldsymbol{\theta}_1\right)$ is taken to be PointNet [9], $\boldsymbol{X} \in \mathbb{R}^{n \times 3}$ is an input point cloud ($n = 128^2$), $\boldsymbol{\mu} \in \mathbb{R}^{256}$ is the latent vector, and $\boldsymbol{\eta} \in \mathbb{R}^{256}$ represents a diagonal covariance matrix $\boldsymbol{\Sigma} = \mathrm{diag}(e^{\boldsymbol{\eta}})$.

Neural Network Architecture

- A modified variational encoder-decoder [13] used, where the encoder $(\boldsymbol{\mu}, \boldsymbol{\eta}) = g\,(\boldsymbol{X}; \boldsymbol{\theta}_1)$ is taken to be PointNet [9], $\boldsymbol{X} \in \mathbb{R}^{n \times 3}$ is an input point cloud ($n = 128^2$), $\boldsymbol{\mu} \in \mathbb{R}^{256}$ is the latent vector, and $\boldsymbol{\eta} \in \mathbb{R}^{256}$ represents a diagonal covariance matrix $\boldsymbol{\Sigma} = \mathrm{diag}(e^{\boldsymbol{\eta}})$.
- The encoder takes in a point cloud $\boldsymbol{X}$ and outputs a probability measure $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. The point cloud is drawn uniformly at random from the scans.

Neural Network Architecture

- A modified variational encoder-decoder [13] used, where the encoder $(\boldsymbol{\mu}, \boldsymbol{\eta}) = g\left(\boldsymbol{X}; \boldsymbol{\theta}_1\right)$ is taken to be PointNet [9], $\boldsymbol{X} \in \mathbb{R}^{n\times 3}$ is an input point cloud ($n = 128^2$), $\boldsymbol{\mu} \in \mathbb{R}^{256}$ is the latent vector, and $\boldsymbol{\eta} \in \mathbb{R}^{256}$ represents a diagonal covariance matrix $\boldsymbol{\Sigma} = \mathrm{diag}(e^{\boldsymbol{\eta}})$.
- The encoder takes in a point cloud $\boldsymbol{X}$ and outputs a probability measure $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. The point cloud is drawn uniformly at random from the scans.
- The decoder represents the implicit representation $f\left(\boldsymbol{x}; \boldsymbol{w}, \boldsymbol{\theta}_2\right)$ with the addition of a latent vector $\boldsymbol{w} \in \mathbb{R}^{256}$.

Debabrata Ghosh
SALD
May 6, 2023

Neural Network Architecture

- A modified variational encoder-decoder [13] used, where the encoder $(\boldsymbol{\mu}, \boldsymbol{\eta}) = g\,(\boldsymbol{X}; \boldsymbol{\theta}_1)$ is taken to be PointNet [9], $\boldsymbol{X} \in \mathbb{R}^{n \times 3}$ is an input point cloud ($n = 128^2$), $\boldsymbol{\mu} \in \mathbb{R}^{256}$ is the latent vector, and $\boldsymbol{\eta} \in \mathbb{R}^{256}$ represents a diagonal covariance matrix $\boldsymbol{\Sigma} = \mathrm{diag}(e^{\boldsymbol{\eta}})$.
- The encoder takes in a point cloud $\boldsymbol{X}$ and outputs a probability measure $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. The point cloud is drawn uniformly at random from the scans.
- The decoder represents the implicit representation $f\,(\boldsymbol{x}; \boldsymbol{w}, \boldsymbol{\theta}_2)$ with the addition of a latent vector $\boldsymbol{w} \in \mathbb{R}^{256}$.
- The network f is taken to be an 8-layer Multi-layer Perceptron (MLP).

Experiments and Results

Evaluation Metrics

Chamfer distance metrics to measure similarity between shapes:

$$\mathrm{d_C}\left(\mathcal{X}_1, \mathcal{X}_2\right) = \frac{1}{2}\left(\mathrm{d_{\vec{C}}}\left(\mathcal{X}_1, \mathcal{X}_2\right) + \mathrm{d_{\vec{C}}}\left(\mathcal{X}_2, \mathcal{X}_1\right)\right)$$

where

$$\mathrm{d_{\vec{C}}}\left(\mathcal{X}_1, \mathcal{X}_2\right) = \frac{1}{|\mathcal{X}_1|} \sum_{\boldsymbol{x}_1 \in \mathcal{X}_1} \min_{\boldsymbol{x}_2 \in \mathcal{X}_2} \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|$$

and the sets $\mathcal{X}_i$ are either point clouds or triangle soups.

Evaluation Metrics

Also, to measure similarity of the normals of triangle soups $\mathcal{T}_1, \mathcal{T}_2$, define:

$$\mathrm{d_N}\left(\mathcal{T}_1, \mathcal{T}_2\right) = \frac{1}{2}\left(\ \mathrm{d}_{\mathrm{N}}^{\rightarrow}\left(\mathcal{T}_1, \mathcal{T}_2\right) + \mathrm{d_N}\left(\mathcal{T}_2, \mathcal{T}_1\right)\right)$$

where

$$\mathrm{d}_{\mathrm{N}}^{\vec{T}}\left(\mathcal{T}_1, \mathcal{T}_2\right) = \frac{1}{\left|\mathcal{T}_1\right|} \sum_{\boldsymbol{x}_1 \in \mathcal{T}_1} \angle\left(\boldsymbol{n}\left(\boldsymbol{x}_1\right), \boldsymbol{n}\left(\hat{\boldsymbol{x}}_1\right)\right)$$

where $\angle(\boldsymbol{a}, \boldsymbol{b})$ is the positive angle between vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^3$, $\boldsymbol{n}\left(\boldsymbol{x}_1\right)$ denotes the face normal of a point $\boldsymbol{x}_1$ in triangle soup $\mathcal{T}_1$, and $\hat{\boldsymbol{x}}_1$ is the projection of $\boldsymbol{x}_1$ on $\mathcal{T}_2$.

Baseline

- DeepSDF [15] is chosen as a representative out of the methods that require pre-computed implicit representation for training.

Experiments and Results

Baseline

- DeepSDF [15] is chosen as a representative out of the methods that require pre-computed implicit representation for training.
- For methods that train directly on raw 3D data, comparisons were done versus SAL [1] and IGR [11].

Experiments and Results

ShapeNet

Category	Sofas		Chairs		Tables		Planes		Lamps	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
DeepSDF	**0.329**	**0.230**	**0.341**	**0.133**	0.839	**0.149**	**0.177**	0.076	**0.909**	**0.344**
SAL	0.704	0.523	0.494	0.259	**0.543**	**0.231**	0.429	0.146	4.913	1.515
SALD(VAE)	0.391	0.244	0.415	0.255	0.679	0.279	0.197	**0.062**	1.808	1.172
SALD(AD)	**0.207**	**0.147**	**0.281**	**0.157**	**0.408**	0.25	**0.098**	**0.032**	**0.506**	**0.327**

Figure: ShapeNet quantitative results with the mean and median of the Chamfer distances $(\mathrm{d_C})$ between the reconstructed $3\mathrm{D}$ surfaces and the ground truth meshes. Numbers are reported $\times 10^3$. [2]

Debabrata Ghosh
SALD
May 6, 2023

Experiments and Results

ShapeNet

Figure: ShapeNet qualitative test results. Each quadruple shows (columns from left to right): ground truth model, SAL-reconstruction, DeepSDF reconstruction, SALD reconstruction. [2]

D-Faust

	$\mathrm{d}_{\mathrm{C}}^{\rightarrow}$ (reg., recon.)		d_{N} (reg., recon.)		$\mathrm{d}_{\mathrm{C}}^{\rightarrow}$ (recon., reg.)		$\mathrm{d}_{\mathrm{N}}^{\rightarrow}$ (recon., reg.)		$\mathrm{d}_{\mathrm{C}}^{\rightarrow}$ (scan, recon.)		d_{N} (scan, recon.)	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
SAL	0.418	0.328	13.21	12.459	0.344	0.256	11.354	**10.522**	0.429	0.246	10.096	9.096
IGR	0.276	0.187	10.328	9.822	3.806	3.627	17.124	17.902	0.241	0.11	5.829	5.295
SALD	0.428	0.346	11.67	11.07	0.489	0.362	11.035	10.371	0.397	0.279	7.884	7.227

Table: D-Faust quantitative results with mean and median of the one-sided Chamfer and normal distances between registration meshes (reg), reconstructions (recon) and raw input scans (scan). The d_{C} numbers are reported $\times 10^2$. [2]

Experiments and Results

D-Faust

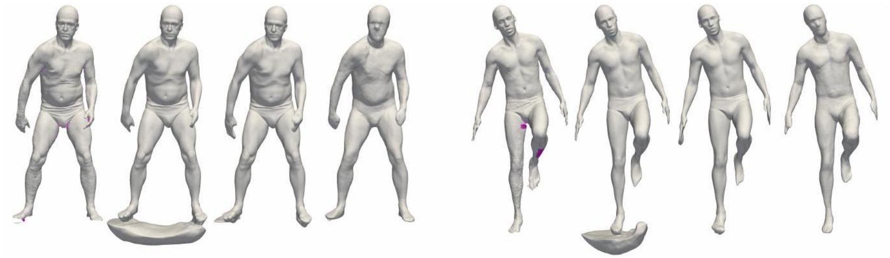


Figure: D-Faust qualitative results on test examples. Each quadruple shows (columns from left to right): raw scans (magenta depict back-faces), IGR, SALD, and SAL. [2]

Further Improvements

Shortcomings of SALD

When looking at the qualitative results of SALD we can observe some shortcomings that can be further attempted to improve.

Further Improvements

Shortcomings of SALD

Figure: Left one of each pair is the original surface and the right one is reconstructed using SALD. [2]

Further Improvements

Network Improvements

LightSAL, a novel deep convolutional architecture for learning 3D shapes based on the same technique as SAL, concentrates on efficiency in network training time and resulting model size [4].

Further Improvements

Technical Improvements

- Existing INRs require point coordinates to learn the implicit level sets of the shape and if a normal vector is available for each point, a higher fidelity representation can be learned.

Further Improvements

Technical Improvements

- Existing INRs require point coordinates to learn the implicit level sets of the shape and if a normal vector is available for each point, a higher fidelity representation can be learned.
- However, normal vectors are often not provided as raw data.

Further Improvements

Technical Improvements

- Existing INRs require point coordinates to learn the implicit level sets of the shape and if a normal vector is available for each point, a higher fidelity representation can be learned.
- However, normal vectors are often not provided as raw data.
- In DiGS, the authors propose a divergence-guided shape representation learning approach that does not require normal vectors as input [5].

Further Improvements

Technical Improvements

- Existing INRs require point coordinates to learn the implicit level sets of the shape and if a normal vector is available for each point, a higher fidelity representation can be learned.
- However, normal vectors are often not provided as raw data.
- In DiGS, the authors propose a divergence-guided shape representation learning approach that does not require normal vectors as input [5].
- Incorporating a soft constraint on the divergence of the distance function leads to smooth solutions that reliably orient gradients to match the unknown normal at each point, sometimes even better than approaches that use ground truth normal vectors directly [5].

References

[1] Matan Atzmon and Yaron Lipman. “SAL: Sign Agnostic Learning of Shapes From Raw Data”. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2020.

[2] Matan Atzmon and Yaron Lipman. “SALD: Sign Agnostic Learning with Derivatives”. In: *9th International Conference on Learning Representations, ICLR 2021*. 2021.

[3] Matan Atzmon et al. “Controlling neural level sets”. In: *Advances in Neural Information Processing Systems*. 2019, pp. 2032–2041.

[4] Abol Basher, Muhammad Sarmad, and Jani Boutellier. “LightSAL: Lightweight Sign Agnostic Learning for Implicit Surface Representation”. In: *CoRR* abs/2103.14273 (2021). arXiv: `2103.14273`.

References II

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

[6] Federica Bogo et al. "Dynamic FAUST: Registering Human Bodies in Motion". In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017, pp. 5573–5582. DOI: 10.1109/CVPR.2017.591.

[7] M.P. do Carmo. *Differential Geometry of Curves and Surfaces: Revised and Updated Second Edition*. Dover Books on Mathematics. Dover Publications, 2016. ISBN: 9780486806990.

References III

[8] Angel X. Chang et al. *ShapeNet: An Information-Rich 3D Model Repository*. 2015. arXiv: 1512.03012 [cs.GR].

[9] R. Charles et al. "PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation". In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Los Alamitos, CA, USA: IEEE Computer Society, July 2017, pp. 77–85. DOI: 10.1109/CVPR.2017.16.

[10] Zhiqin Chen and Hao Zhang. "Learning Implicit Fields for Generative Shape Modeling". In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2019).

[11] Amos Gropp et al. "Implicit Geometric Regularization for Learning Shapes". In: *Proceedings of Machine Learning and Systems 2020*. 2020, pp. 3569–3579.

References IV

[12] Thibault Groueix et al. "AtlasNet: A Papier-Mâché Approach to Learning 3D Surface Generation". In: *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*. 2018.

[13] Diederik P. Kingma and Max Welling. "Auto-Encoding Variational Bayes". In: *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*. Ed. by Yoshua Bengio and Yann LeCun. 2014.

[14] Lars Mescheder et al. "Occupancy Networks: Learning 3D Reconstruction in Function Space". In: *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*. 2019.

References V

[15] Jeong Joon Park et al. "DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2019.

[16] M. Tatarchenko, A. Dosovitskiy, and T. Brox. "Octree Generating Networks: Efficient Convolutional Architectures for High-resolution 3D Outputs". In: *IEEE International Conference on Computer Vision (ICCV)*. 2017.

[17] Francis Williams et al. "Deep geometric prior for surface reconstruction". English (US). In: IEEE Computer Society, June 2019, pp. 10122–10131. DOI: 10.1109/CVPR.2019.01037.

References VI

[18] Jiajun Wu et al. "Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling". In: *Advances in Neural Information Processing Systems*. Ed. by D. Lee et al. Vol. 29. Curran Associates, Inc., 2016.

[19] Hong-Kai Zhao, S. Osher, and R. Fedkiw. "Fast surface reconstruction using the level set method". In: *Proceedings IEEE Workshop on Variational and Level Set Methods in Computer Vision*. 2001, pp. 194–201. DOI: 10.1109/VLSM.2001.938900.