3D Reconstruction using Implicit Function

Date 2021-01-19

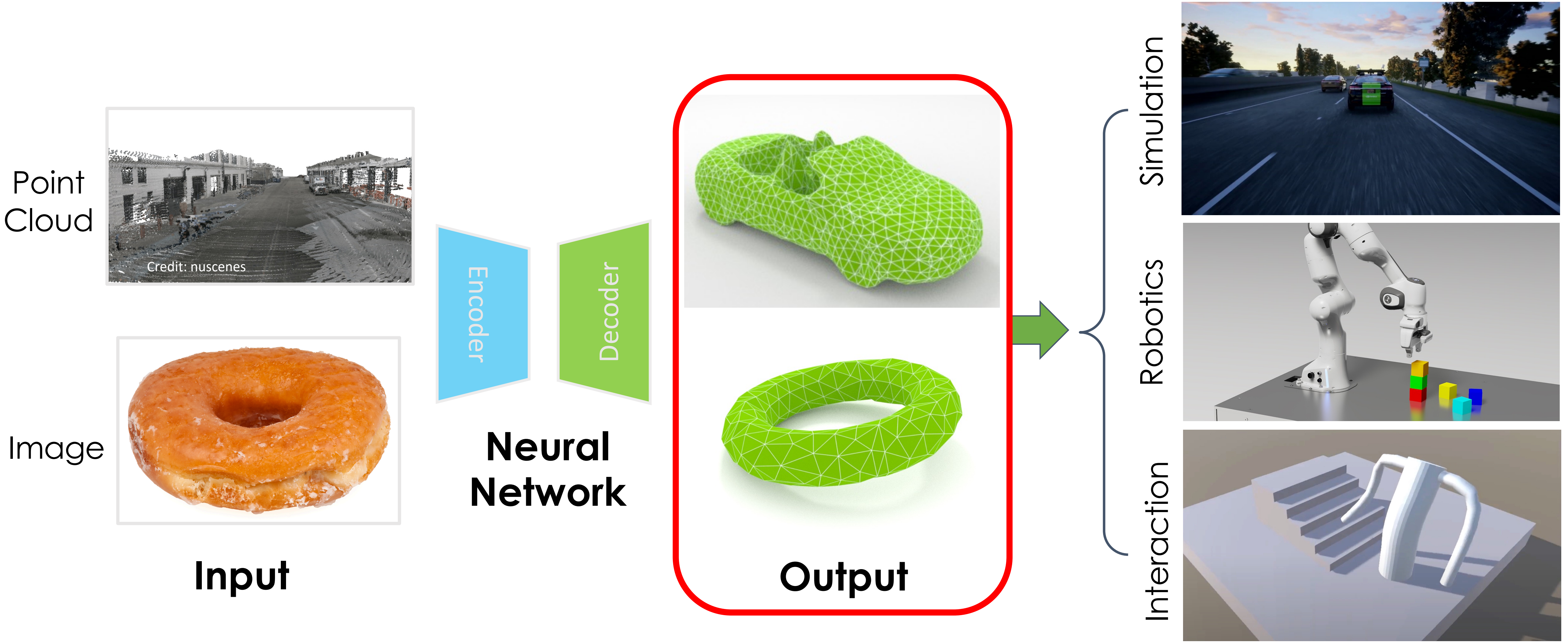
Presenter: Jun Gao

Instructor: Animesh Garg

UNIVERSITY OF TORONTO

3D Reconstruction from x

Applications



Paper List

- Occupancy Networks: Learning 3D Reconstruction in Function Space
  - CVPR 2019
- Convolutional Occupancy Networks
  - ECCV 2020
- Implicit Neural Representations with Periodic Activation Functions
  - NeurIPS 2020

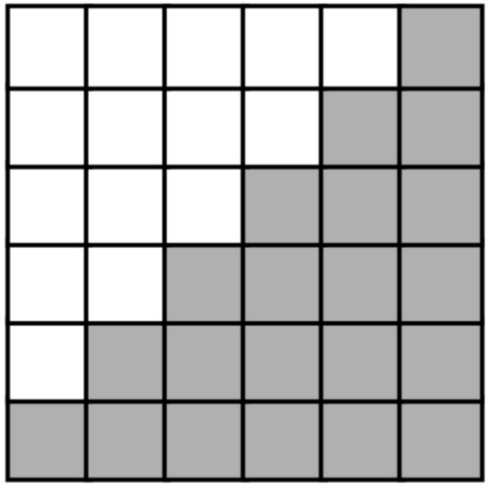
What’s a good **output** representation?

What's a good **output** representation?

Voxels:

- **Discretization** of 3D space into grid
- Easy to process with neural networks
- Cubic memory $O(n^3) \Rightarrow$ limited resolution
- Manhattan world bias

[Maturana et al., IROS 2015]



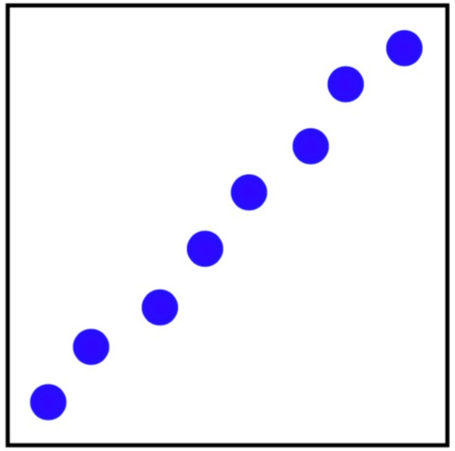
Thanks Andreas Geiger
http://www.cvlibs.net/publications/Mescheder2019CVPR_slides.pdf

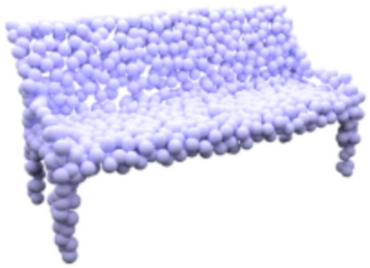
What's a good **output** representation?

Points:

- **Discretization** of surface into 3D points
- Does not model connectivity / topology
- Limited number of points
- Global shape description

[Fan et al., CVPR 2017]





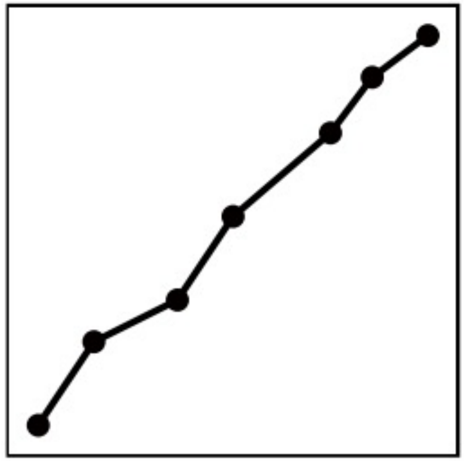
Thanks Andreas Geiger
http://www.cvlibs.net/publications/Mescheder2019CVPR_slides.pdf

What’s a good **output** representation

Meshes:

- **Discretization** into vertices and faces
- Limited number of vertices / granularity
- Requires class-specific template – or –
- Leads to self-intersections

[Groueix et al., CVPR 2018]



Thanks Andreas Geiger
http://www.cvlibs.net/publications/Mescheder2019CVPR_slides.pdf

Contribution

- Introduce a new representation for 3D geometry
- The representation can be used for reconstructing 3D geometry from various input types.
- State-of-the-art performance on 3D reconstruction

What's Implicit Function based Representation?

- Using a function to represent 3D object

Occupancy

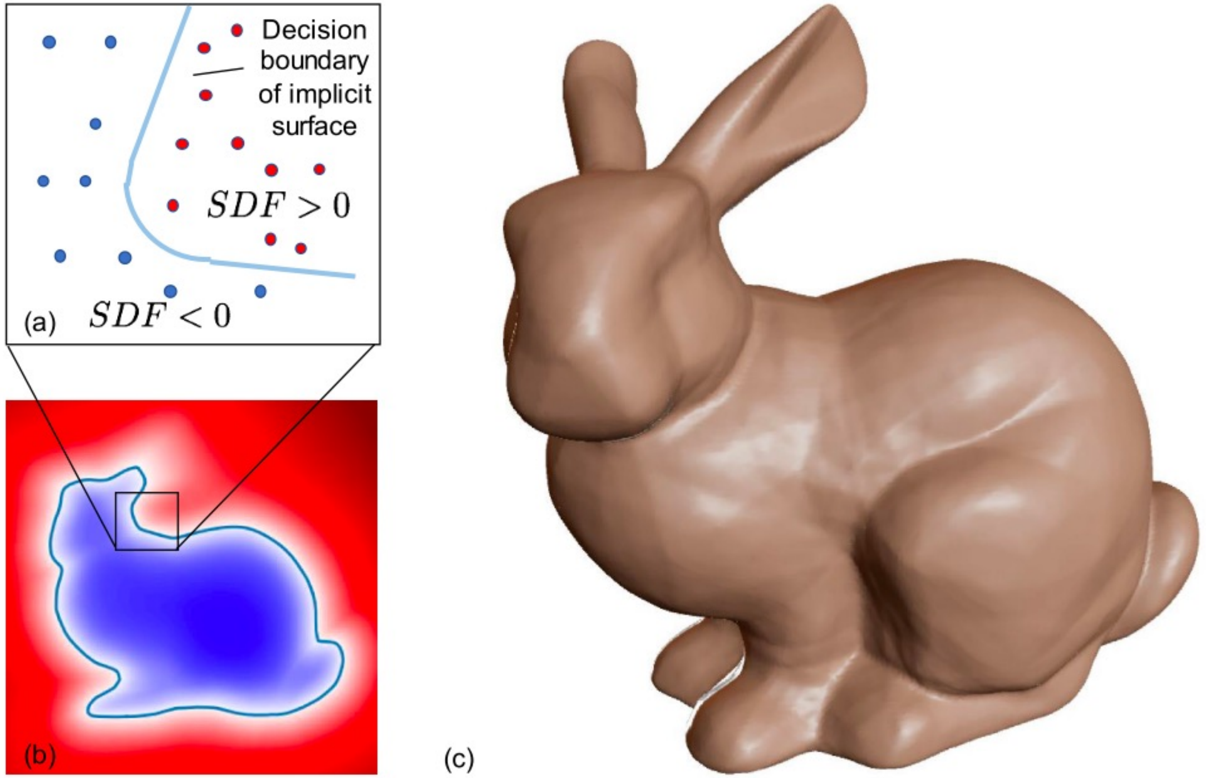
$$f_\theta : \mathbb{R}^3 \times \mathcal{X} \rightarrow [0, 1]$$

3D Location | Condition (eg, Image) | Occupancy Probability

Signed Distance Function

$$f_\theta : \mathbb{R}^3 \times \mathcal{X} \rightarrow [-D, D]$$

3D Location | Condition (eg, Image) | Signed Distance to the Surface



How to learn this function?

$$f_\theta : \mathbb{R}^3 \times \mathcal{X} \rightarrow [0, 1]$$

3D Location

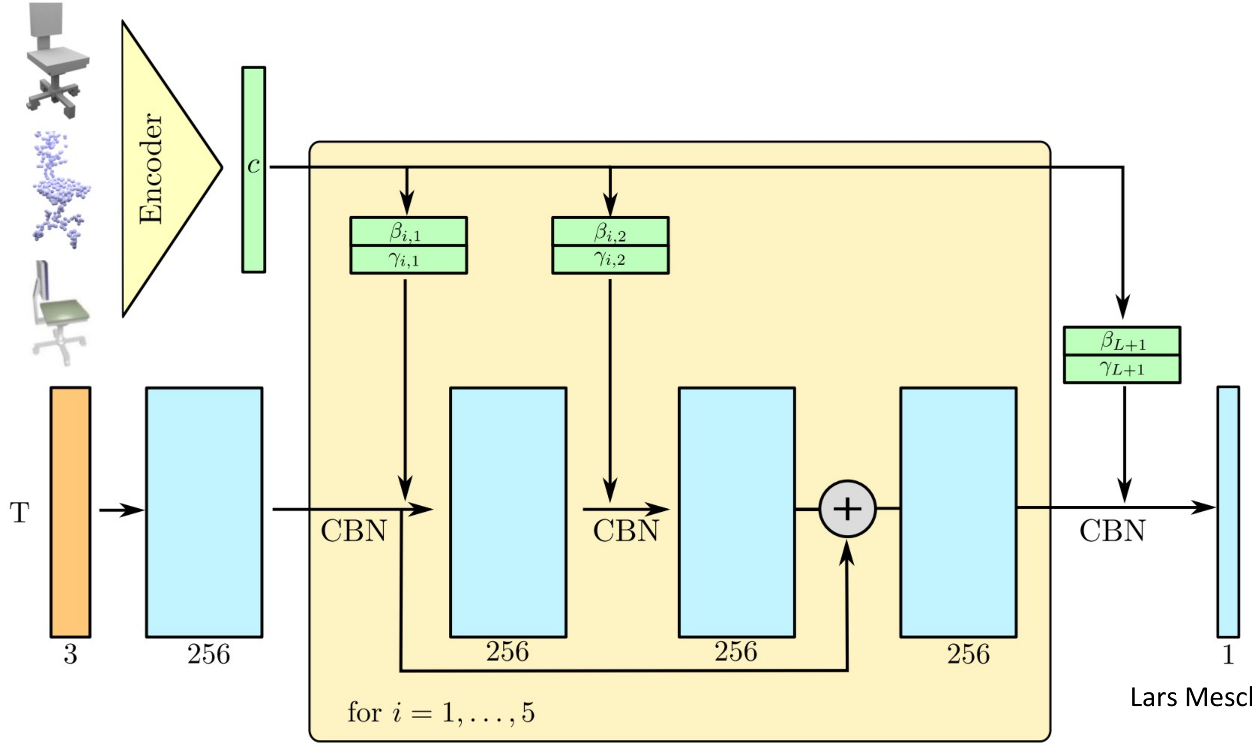
Condition (eg, Image)

Occupancy Probability

Occupancy Networks: Learning 3D Reconstruction in Function Space
Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, Andreas Geiger
CVPR 2019

Methods

- Network Architecture



Average

$$f_{out} = \gamma(c)\frac{f_{in} - \mu}{\sqrt{\sigma^2 + \epsilon}} + \beta(c)$$

Variance

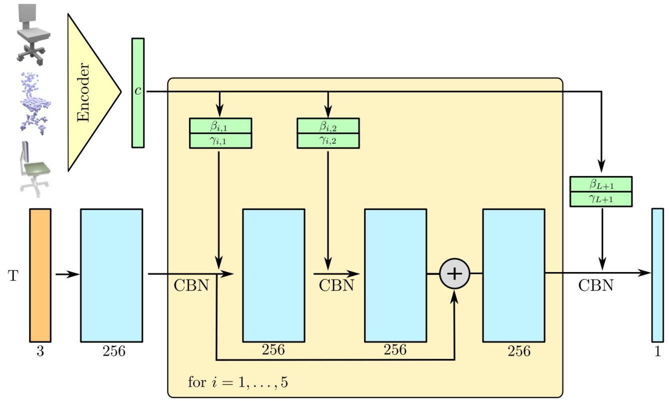
Occupancy Networks: Learning 3D Reconstruction in Function Space

Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, Andreas Geiger

CVPR 2019

Methods

- Network Architecture



- Training Objective

$$\mathcal{L}(\theta, \psi) = \sum_{j=1}^{K} \text{BCE}(f_\theta(p_{ij}, z_i), o_{ij})$$

- ► K: Randomly sampled 3D points ($K = 2048$)
- ► BCE: Cross-entropy loss

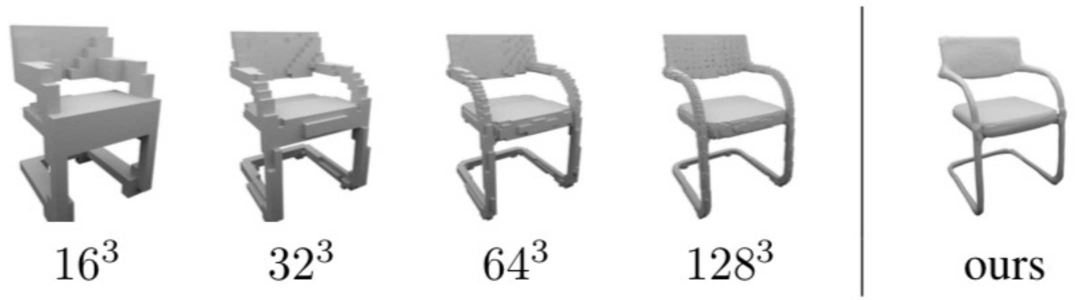
Occupancy Networks: Learning 3D Reconstruction in Function Space

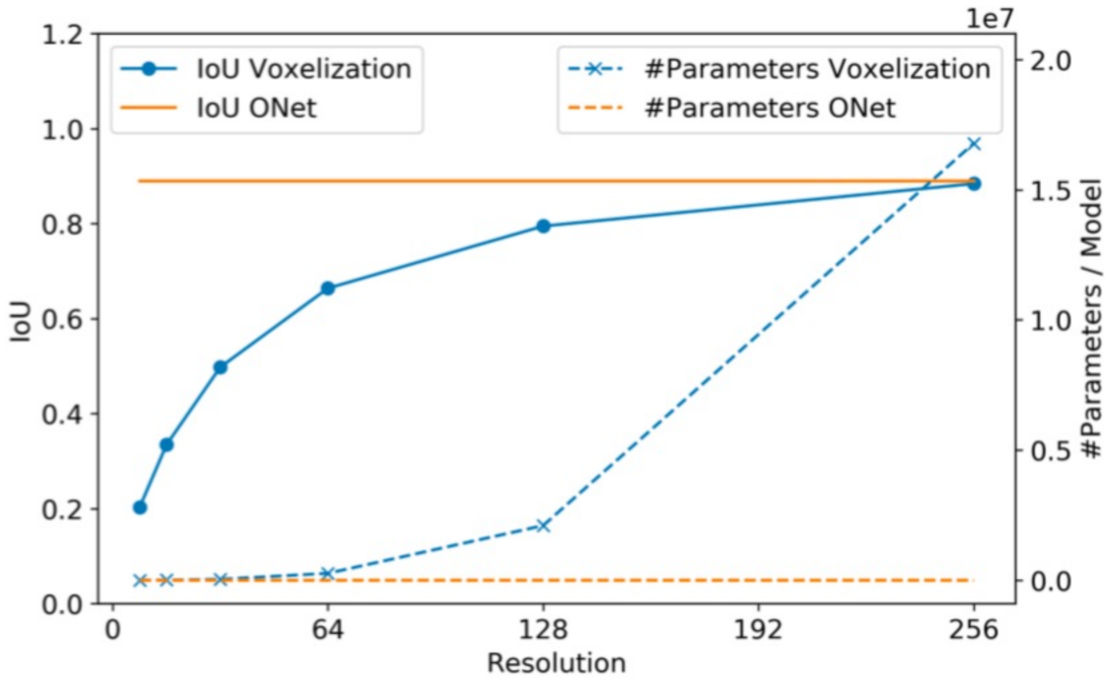
Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, Andreas Geiger

CVPR 2019

Experiments

- Optimizing for one shape
  - Discrete v.s. continuous

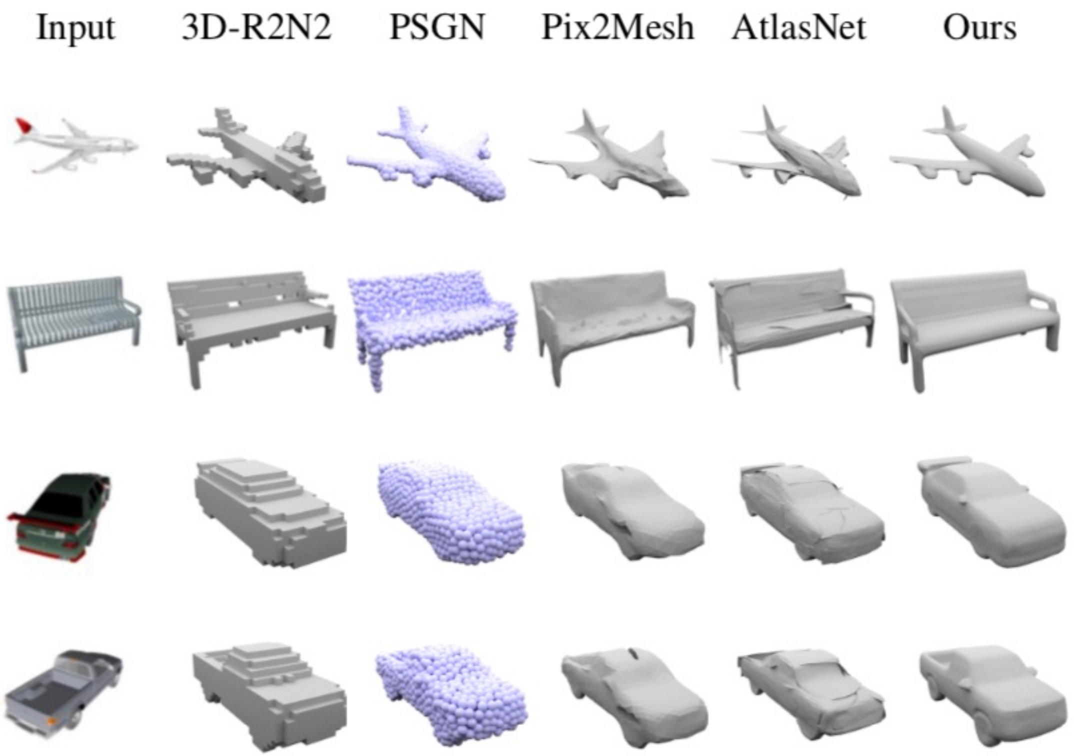




Occupancy Networks: Learning 3D Reconstruction in Function Space
Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, Andreas Geiger
CVPR 2019

Experiments

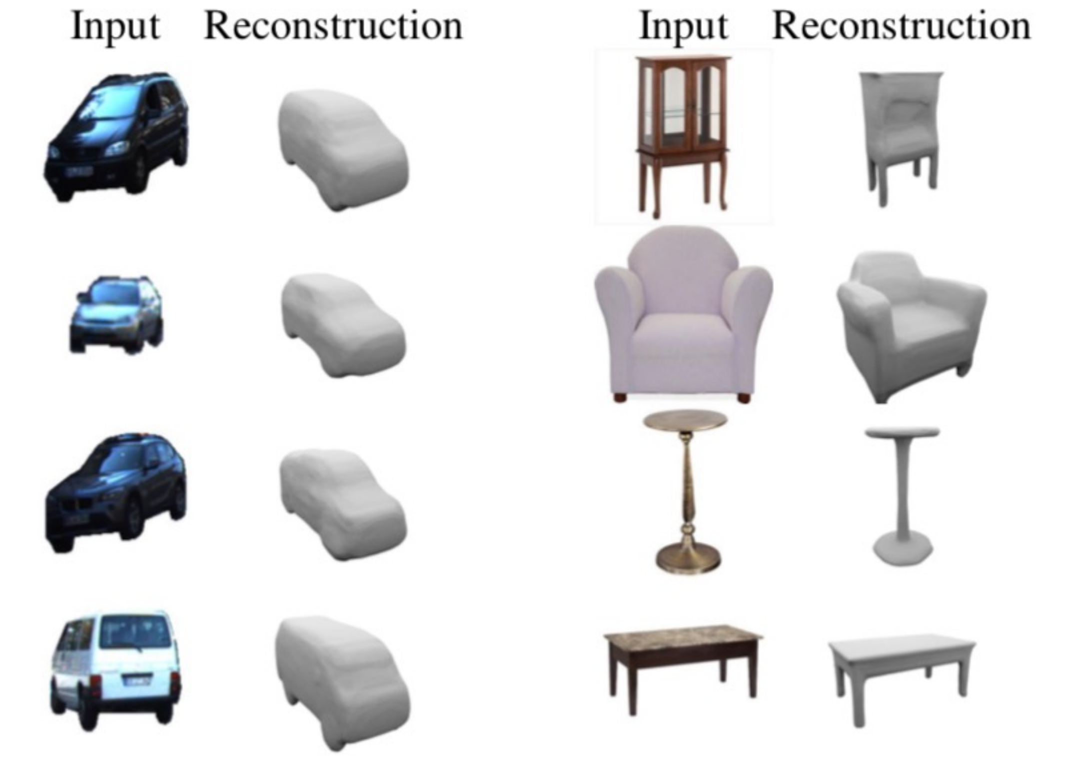
- Single Image 3D Reconstruction



Occupancy Networks: Learning 3D Reconstruction in Function Space
Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, Andreas Geiger
CVPR 2019

Experiments

- Single Image 3D Reconstruction (Testing on real images)

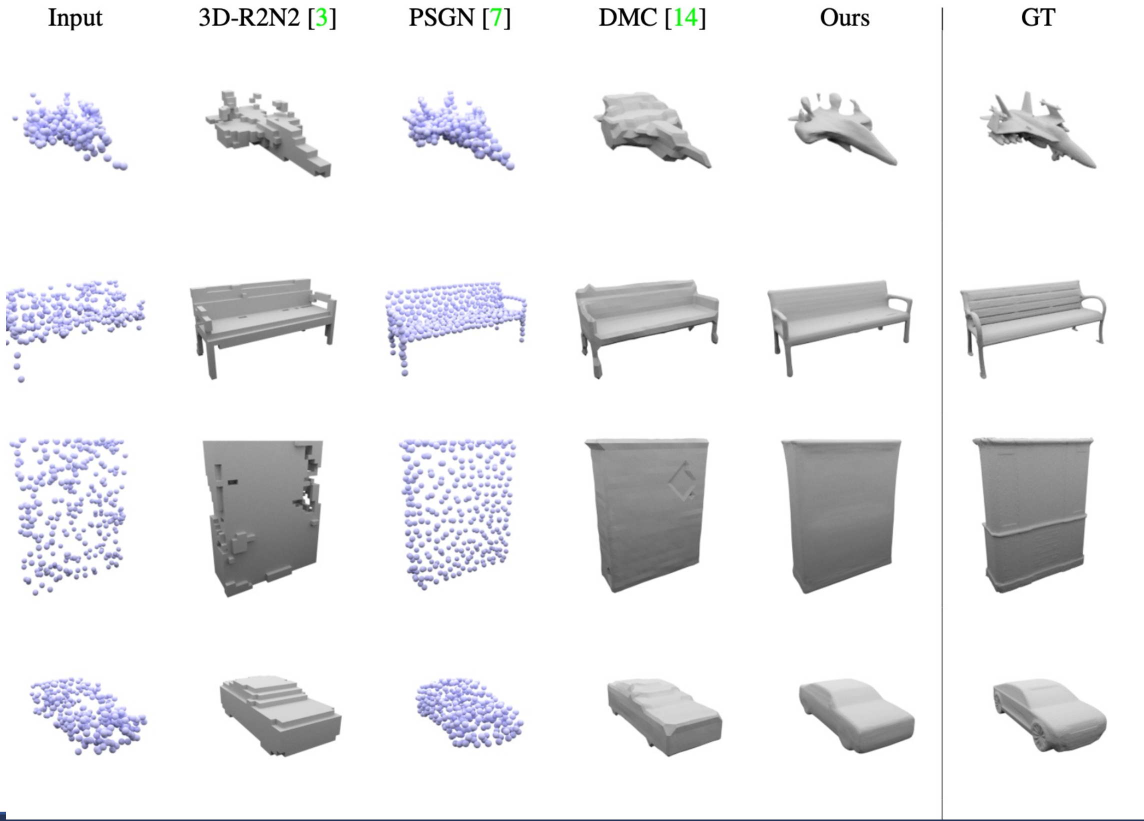


(a) KITTI (b) Online Products

Occupancy Networks: Learning 3D Reconstruction in Function Space
Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, Andreas Geiger
CVPR 2019

Experiments

- Point Cloud 3D Reconstruction



	IoU	Chamfer-L_1 †	Normal Consistency
3D-R2N2	0.565	0.169	0.719
PSGN	-	0.144	-
DMC	0.674	0.117	0.848
ONet	**0.778**	**0.079**	**0.895**

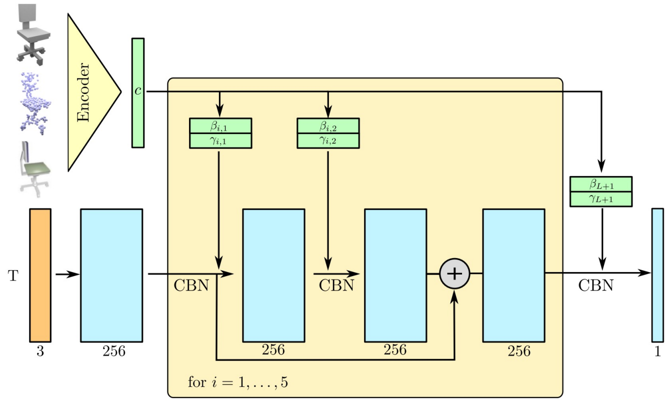
Occupancy Networks: Learning 3D Reconstruction in Function Space
Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, Andreas Geiger
CVPR 2019

Limitation

- Geometry must be extracted in post-processing step (Marching Cube)
- Extension to 4D is not straight-forward (curse of dimensionality)
- Fully connected architecture and global condition lead to oversmooth results.

Methods (Recap)

- Network Architecture



- Training Objective

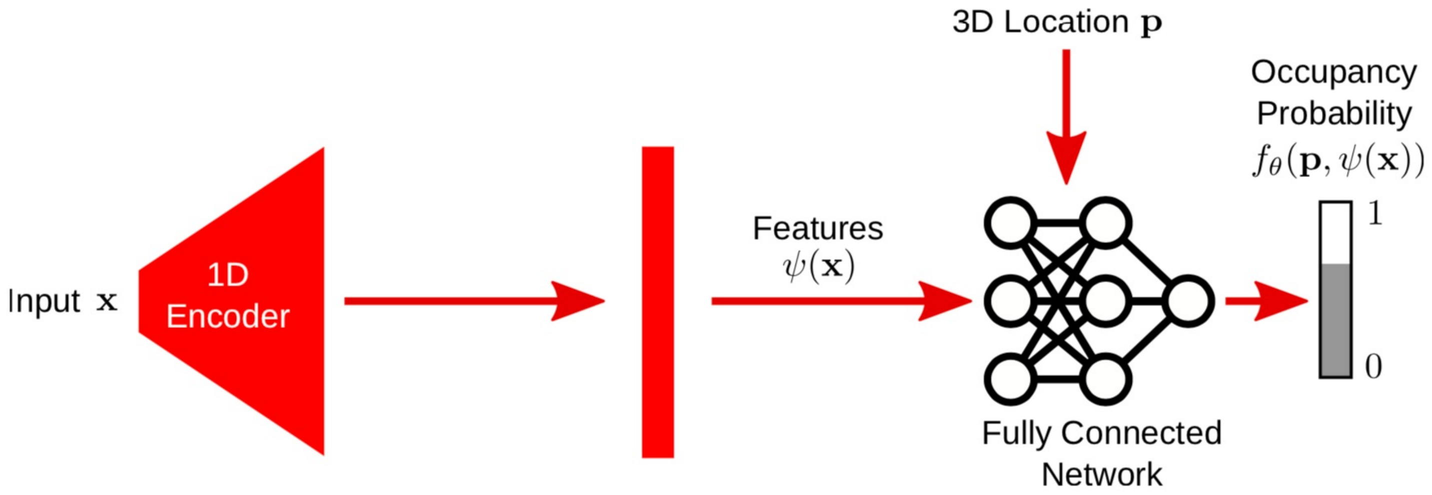
$$\mathcal{L}(\theta, \psi) = \sum_{j=1}^{K} \text{BCE}(f_\theta(p_{ij}, z_i), o_{ij})$$

- ► K: Randomly sampled 3D points $(K = 2048)$
- ► BCE: Cross-entropy loss

Occupancy Networks: Learning 3D Reconstruction in Function Space
Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, Andreas Geiger
CVPR 2019

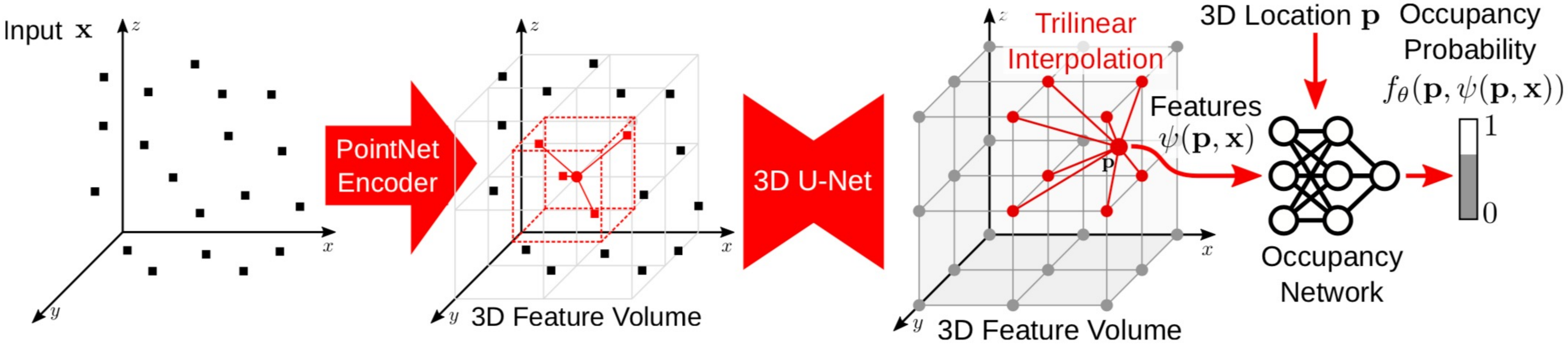
What’s missing in previous paper?

- Lack of local details
- Not scalable to scene-level reconstruction



- ► Global latent code $\Rightarrow$ no local information, overly smooth geometry
- ► Fully connected architecture $\Rightarrow$ does not exploit translation equivariance

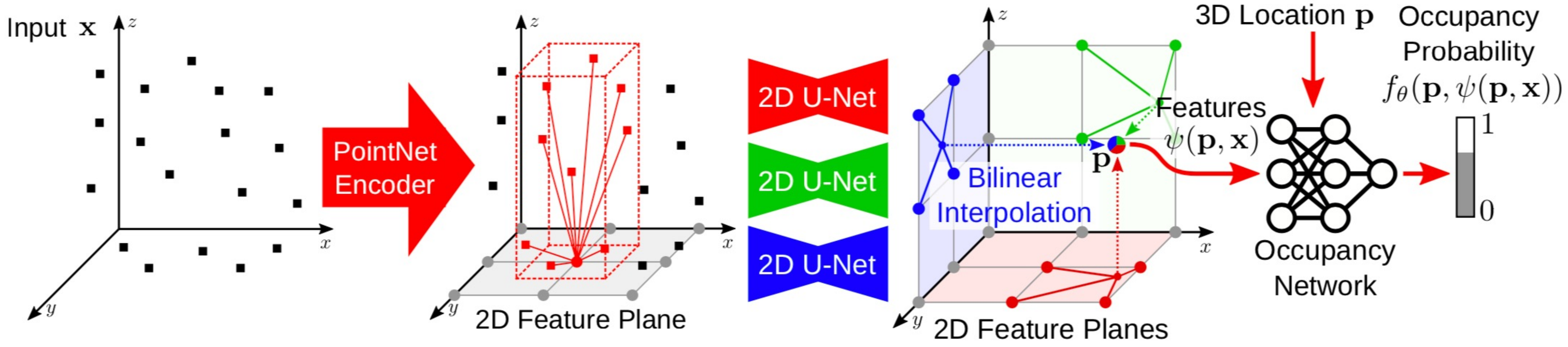
Methods



- **3D Volume Encoder:** Local PointNet processes input, volumetric feature encoding
- **3D Volume Decoder:** Processed by 3D U-Net, query features via trilinear interp.
- **Occupancy Readout:** Shallow occupancy network $f_\theta(\cdot)$

Convolutional Occupancy Networks
Songyou Peng, Michael Niemeyer, Lars Mescheder, Marc Pollefeys and Andreas Geiger
ECCV 2020

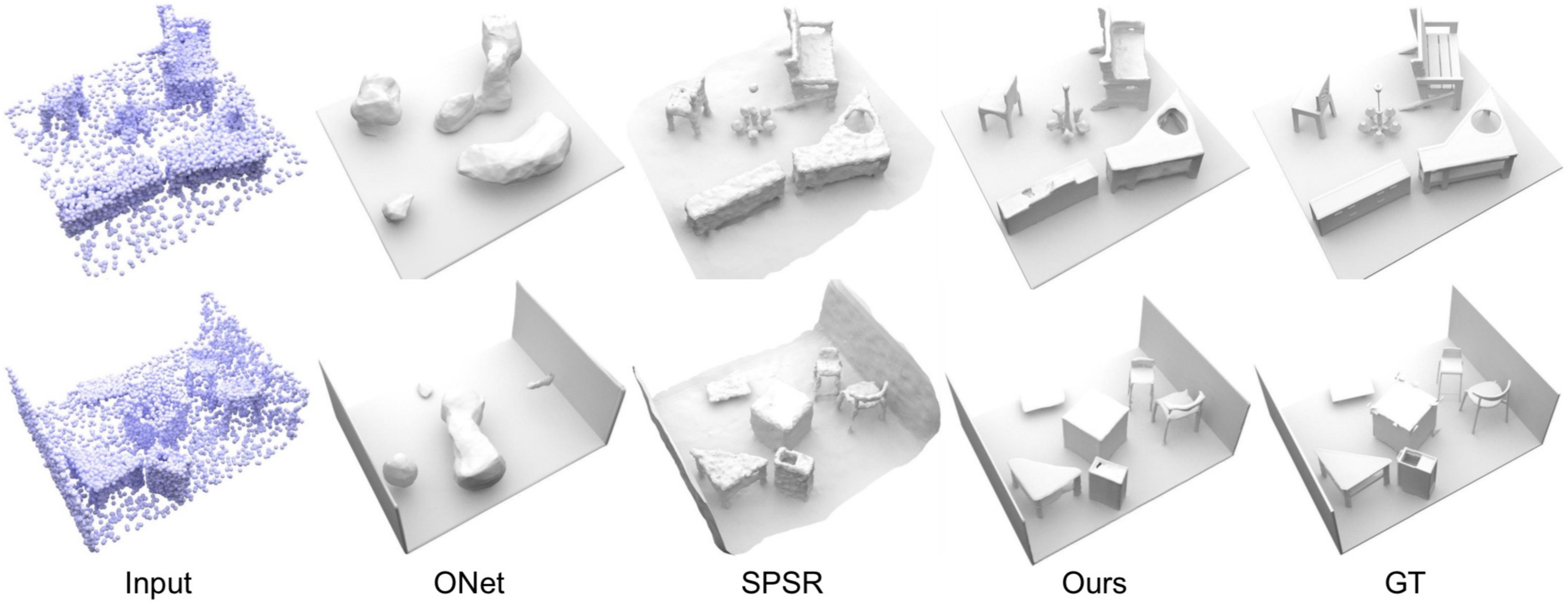
Methods



- **2D Plane Encoder:** Local PointNet processes input, project onto canonical plane
- **2D Plane Decoder:** Processed by U-Net, query features via bilinear interpolation
- **Occupancy Readout:** Shallow occupancy network $f_\theta(\cdot)$

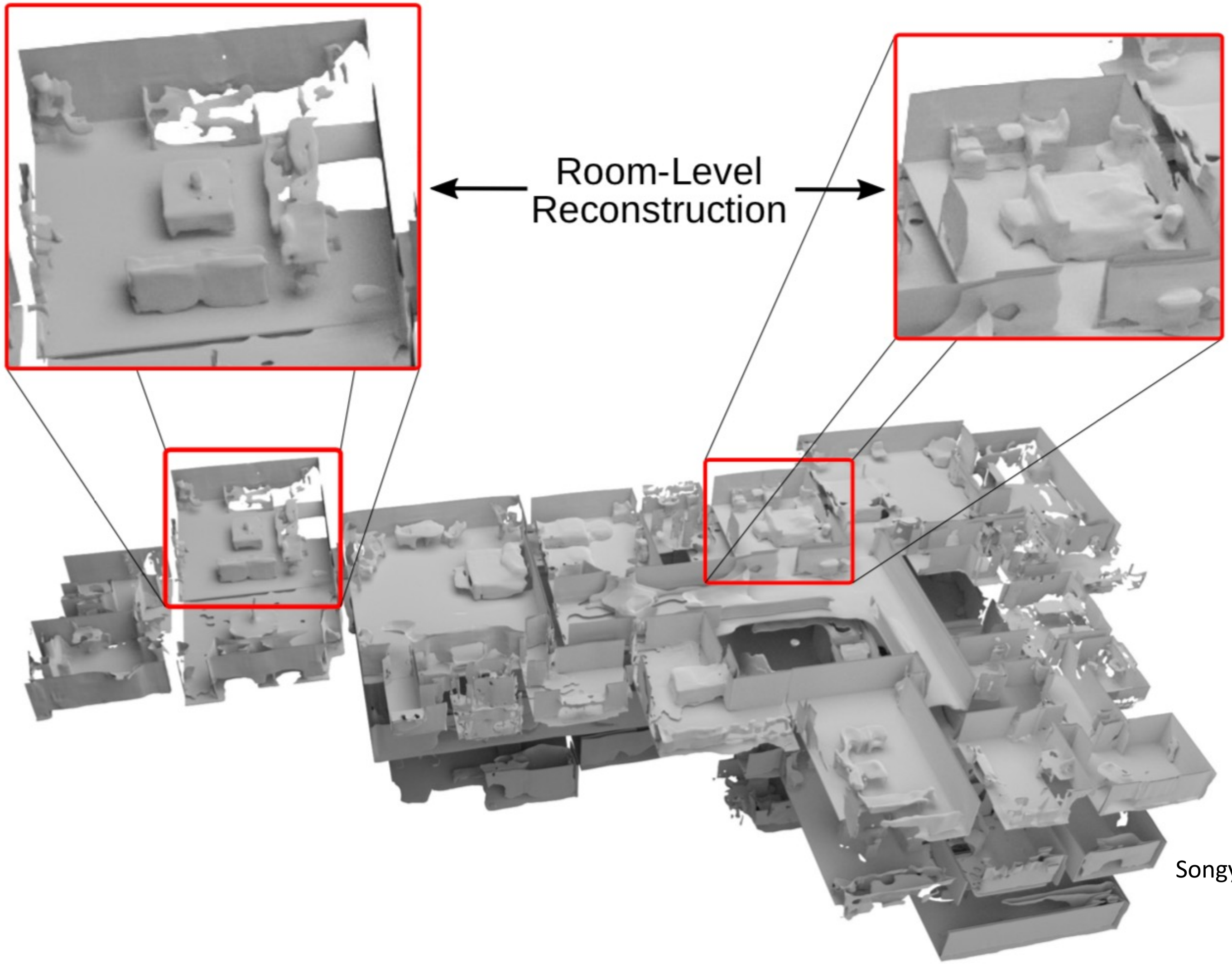
Convolutional Occupancy Networks
Songyou Peng, Michael Niemeyer, Lars Mescheder, Marc Pollefeys and Andreas Geiger
ECCV 2020

Experiments



Convolutional Occupancy Networks
Songyou Peng, Michael Niemeyer, Lars Mescheder, Marc Pollefeys and Andreas Geiger
ECCV 2020

Experiments



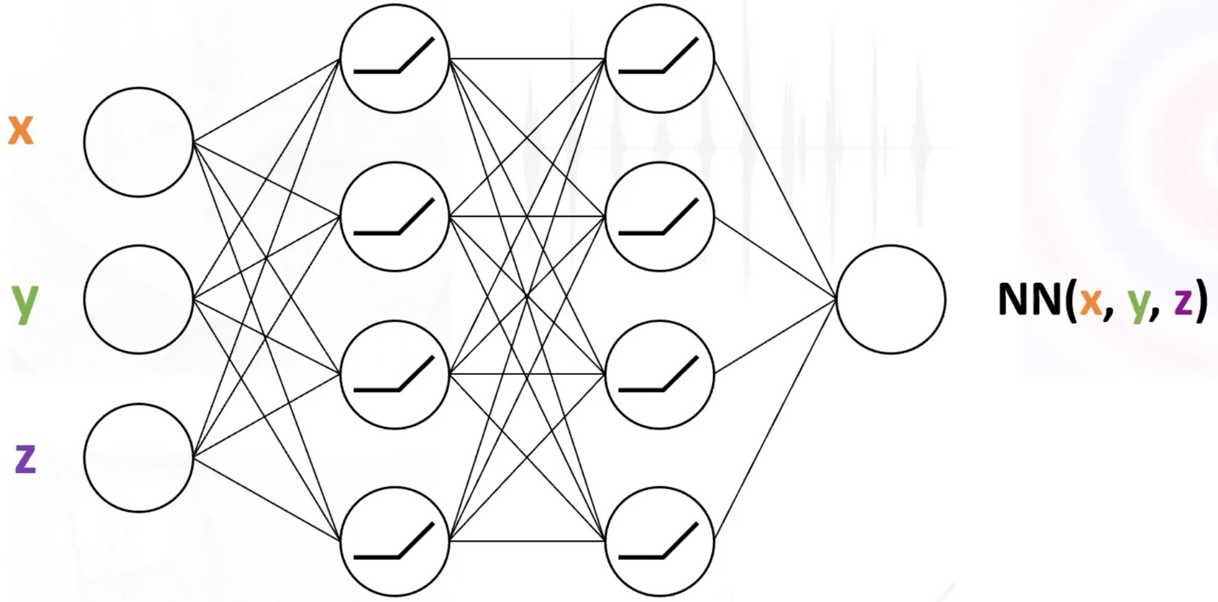
Convolutional Occupancy Networks
Songyou Peng, Michael Niemeyer, Lars Mescheder, Marc Pollefeys and Andreas Geiger
ECCV 2020

Limitation

- Still requires post-processing step (Marching Cube), and is even longer for larger scene
- Details is restricted by the intermediate voxel resolution

What's missing in previous papers?

ReLU
activation
everywhere

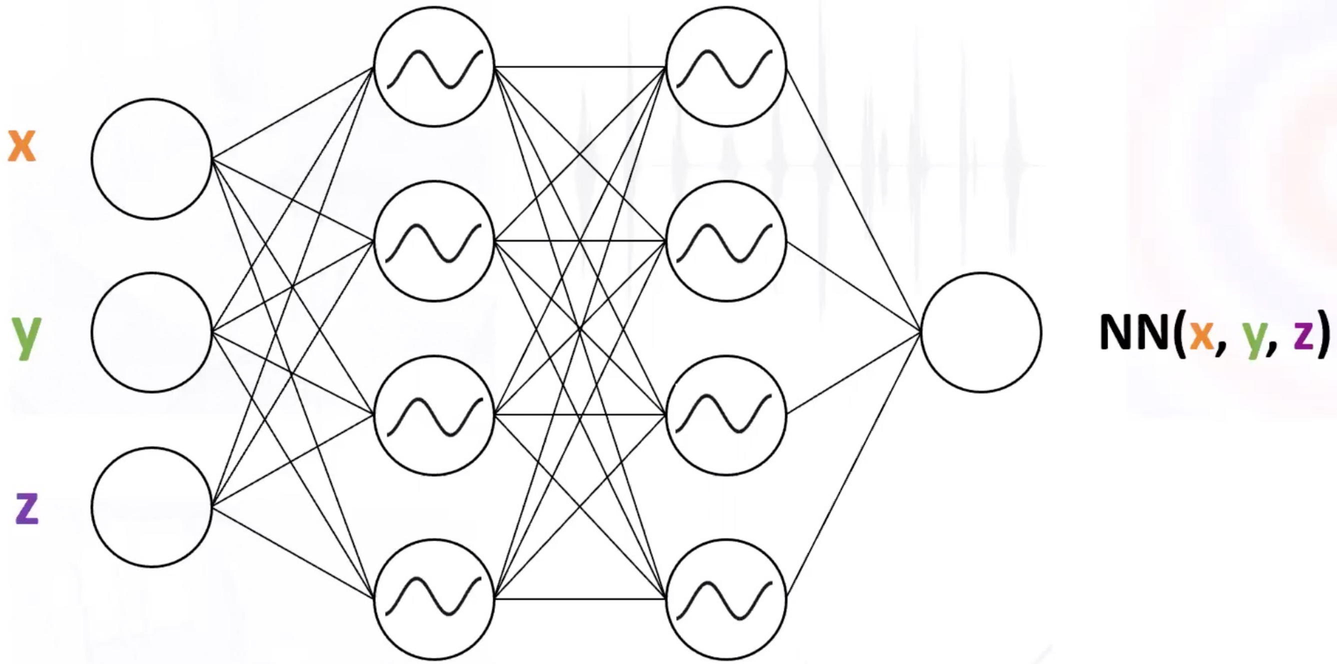


Learned function is not shift-invariant

Piecewise linear → second derivative is zero everywhere!

Implicit Neural Representations with Periodic Activation Functions
Vincent Sitzmann, Julien N. P. Martel, Alexander W. Bergman, David B. Lindell, Gordon Wetzstein
NeurIPS 2020

Method



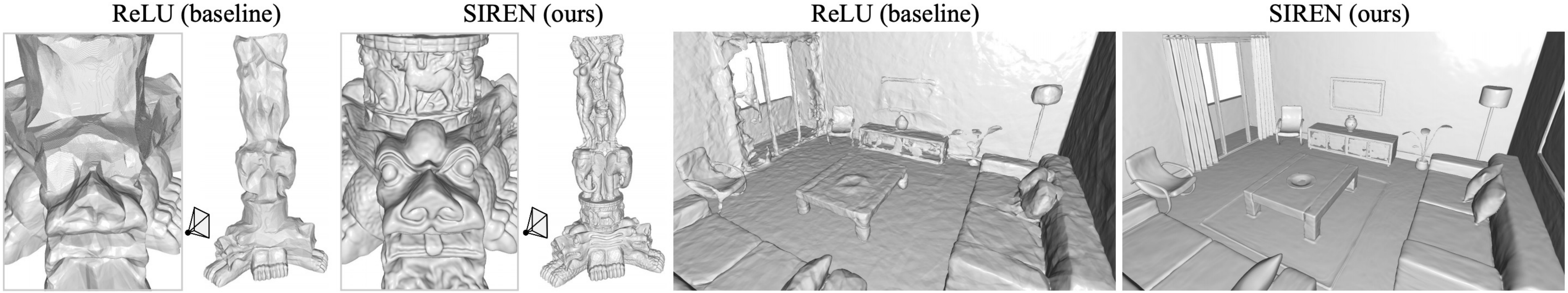
$$\Phi\left(\mathbf{x}\right) = \mathbf{W}_n \left(\phi_{n-1} \circ \phi_{n-2} \circ \ldots \circ \phi_0\right)\left(\mathbf{x}\right) + \mathbf{b}_n,$$

$$\mathbf{x}_i \mapsto \phi_i\left(\mathbf{x}_i\right) = \sin\left(\mathbf{W}_i \mathbf{x}_i + \mathbf{b}_i\right)$$

Gain some shift-invariance!
Derivative of sine is shifted sine itself!

Implicit Neural Representations with Periodic Activation Functions
Vincent Sitzmann, Julien N. P. Martel, Alexander W. Bergman, David B. Lindell, Gordon Wetzstein
NeurIPS 2020

Experiments



Implicit Neural Representations with Periodic Activation Functions
Vincent Sitzmann, Julien N. P. Martel, Alexander W. Bergman, David B. Lindell, Gordon Wetzstein
NeurIPS 2020

Summary

- Implicit representation can represent **arbitrary topology** with **geometric details**
- With **low memory footprint**
- Not restricted to specific categories
- Other research directions:
  - Finding better network architecture for implicit function
  - Extending implicit function to other domains:
    - Textures, 4D motion,
    - Nerf