[A4-004] 딥러닝 코딩 실습

Lecture 03: Implicit Neural Representations

Hak Gu Kim

hakgukim@cau.ac.kr

Immersive Reality & Intelligent Systems Lab (IRIS LAB)

Graduate School of Advanced Imaging Science, Multimedia & Film (GSAIM)

Chung-Ang University (CAU)

19 Jan. 2023

Immersive Reality &
Intelligent Systems Lab.

Topic

- Implicit Neural Representations
  — Occupancy Networks
  — INRs with Periodic Activation Functions

Explicit Representation of Signals

- Traditionally, discrete representations for signals are used

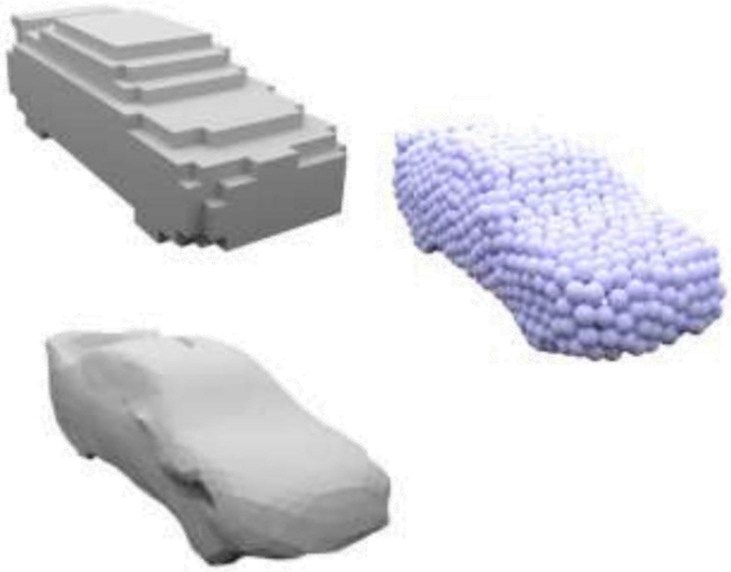
1D Signal: Audio

Samples of
sound wave

2D Signal: Images

Pixels

3D Signal: Shapes



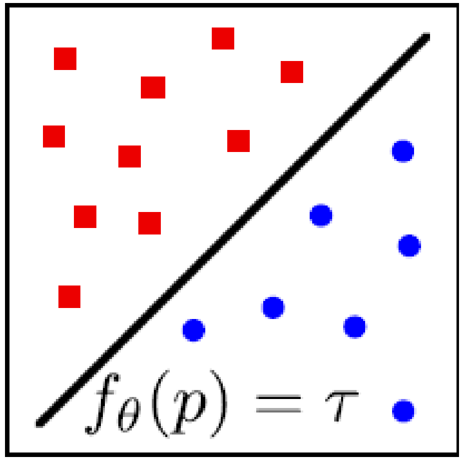
Voxels / Mesh /
Point clouds

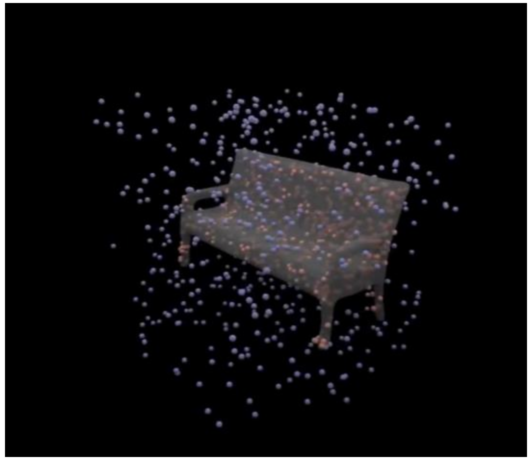
Implicit Representation of Signals

- In implicit neural representations (INRs), signals are parameterized by neural networks
- Example of implicit neural representations for 3D shapes

— Do not represent 3D shape explicitly

— Consider surface implicitly as decision boundary





L. Mescheder et al., Occupancy Networks: Learning 3D Reconstruction in Function Space, **CVPR**, 2019
J. J. Park et al., DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation, **CVPR**, 2019

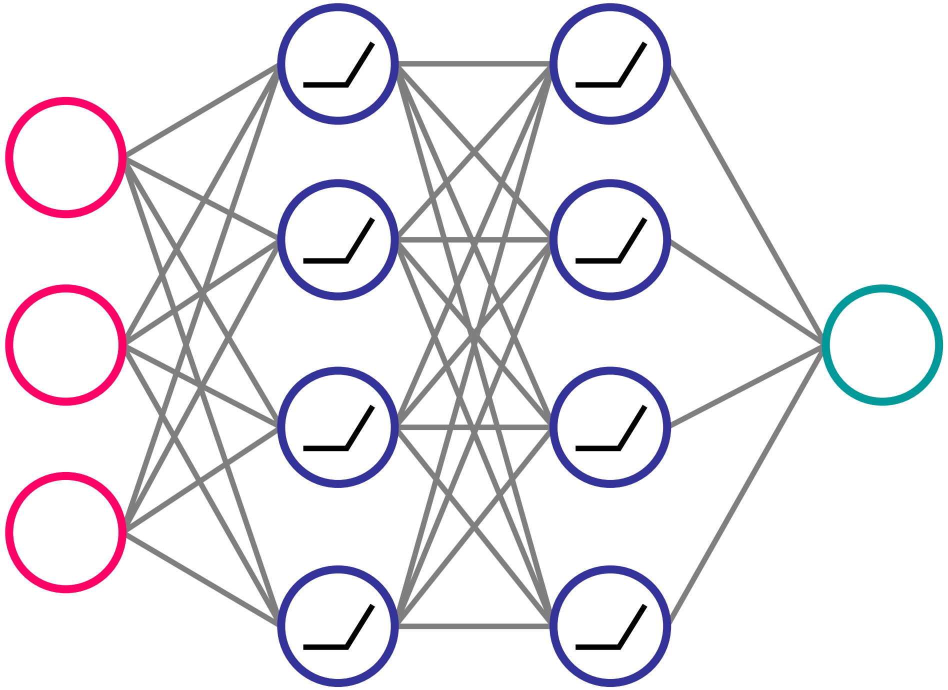
Implicit Representation of Signals

- Benefits of Implicit Neural Representations (INRs)

— Agnostic to grid resolution

— Differentiations computed automatically

Input:
Coordinates
$\mathbf{x} \in \mathbb{R}^3$



Output:
RGB or Signed distance
$\Phi(\mathbf{x}) = s$

L. Mescheder et al., Occupancy Networks: Learning 3D Reconstruction in Function Space, **CVPR**, 2019
J. J. Park et al., DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation, **CVPR**, 2019

Occupancy Network

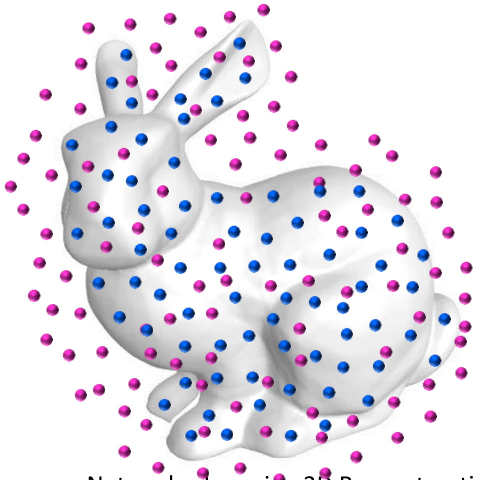
- Occupancy Network
  — The occupancy network f_θ takes a pair $(\mathbf{p}, \mathbf{x})$ as input and outputs a real number which represents the probability of occupancy:

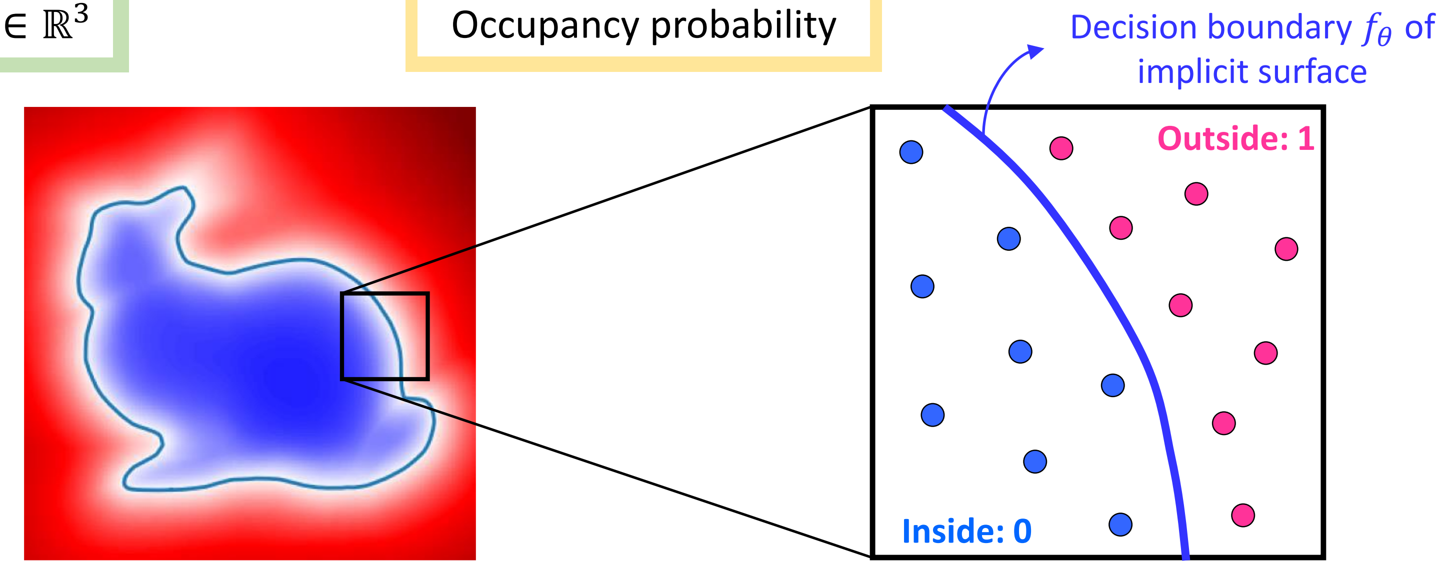
Observation: $\mathbf{x} \in \mathcal{X}$

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

3D point: $\mathbf{p} \in \mathbb{R}^3$

Occupancy probability





L. Mescheder et al., Occupancy Networks: Learning 3D Reconstruction in Function Space, **CVPR**, 2019

Occupancy Network: **Architecture**

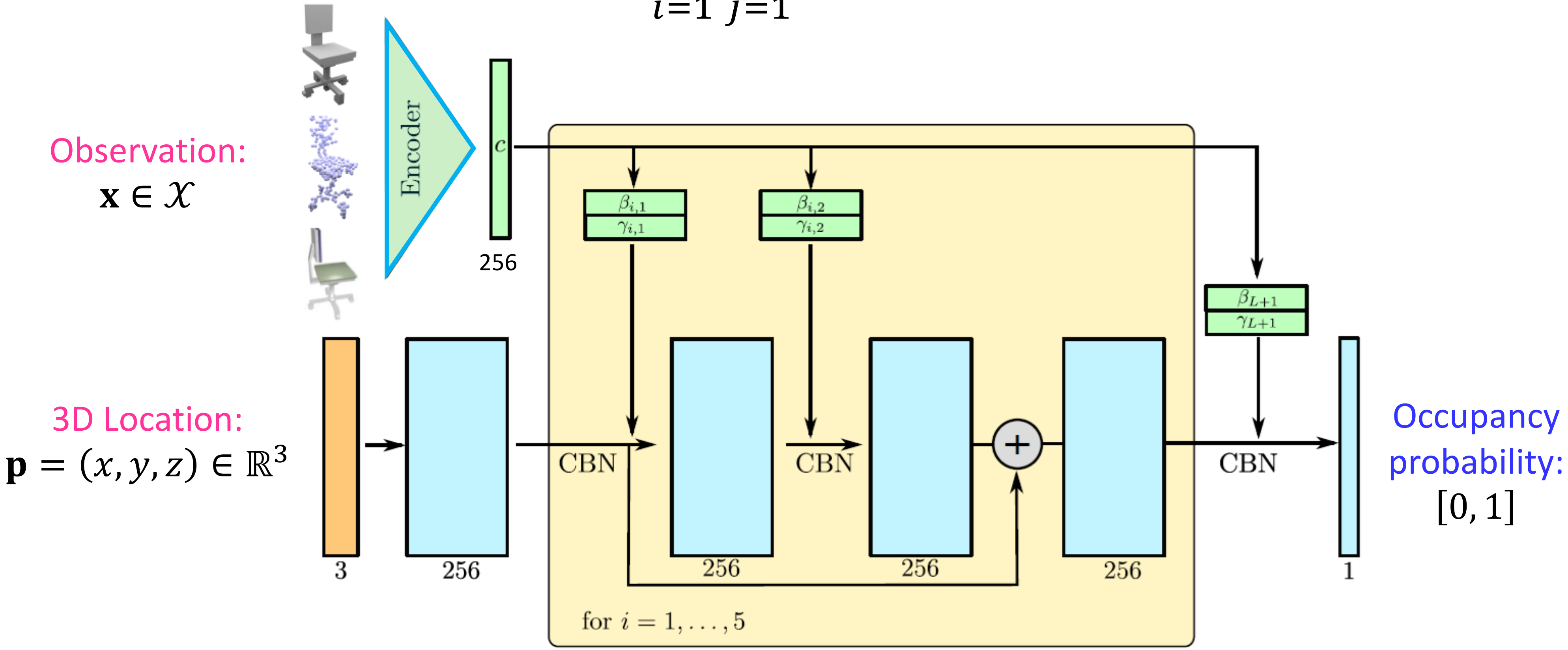
- Model & Loss Function

$$\mathcal{L}(\theta) = \frac{1}{|\mathcal{B}|}\sum_{i=1}^{|\mathcal{B}|}\sum_{j=1}^{K} BCE\left(f_\theta\left(\mathbf{p}_{ij}, \mathbf{x}_i\right), o_{ij}\right)$$

$|\mathcal{B}|$: #. of batch

K : #. of randomly sampled points

BCE : Cross-entropy loss

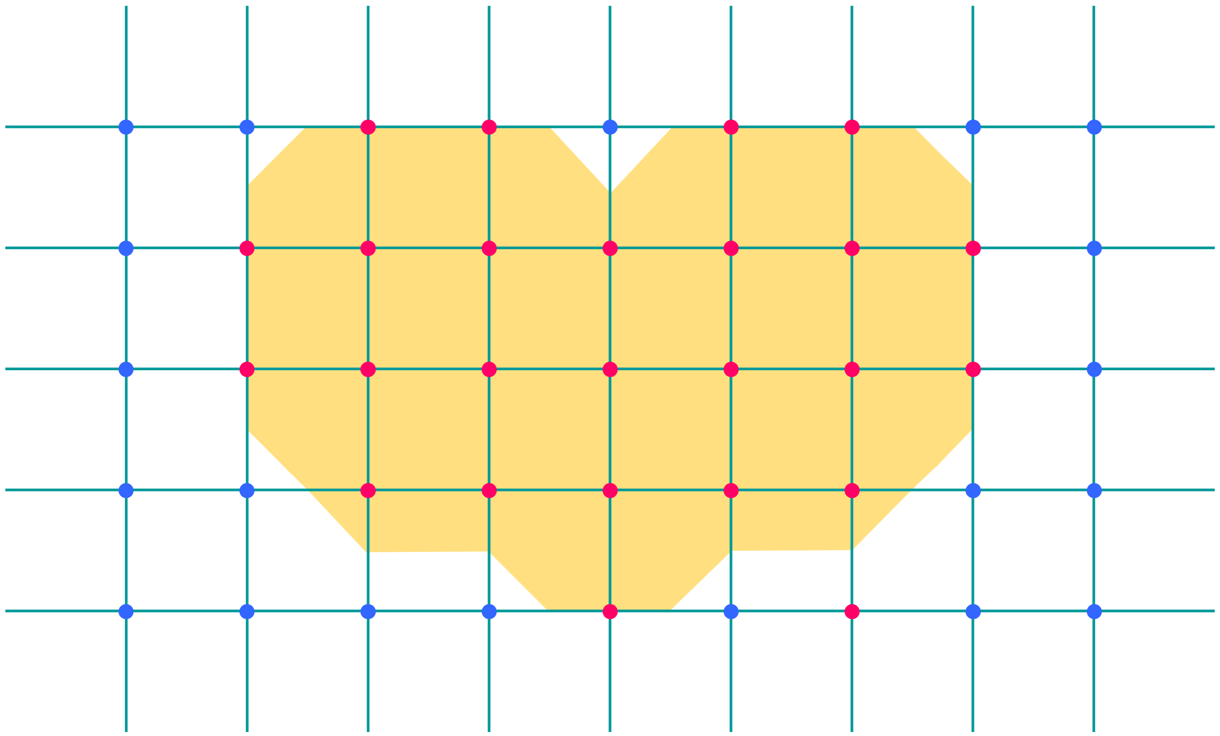


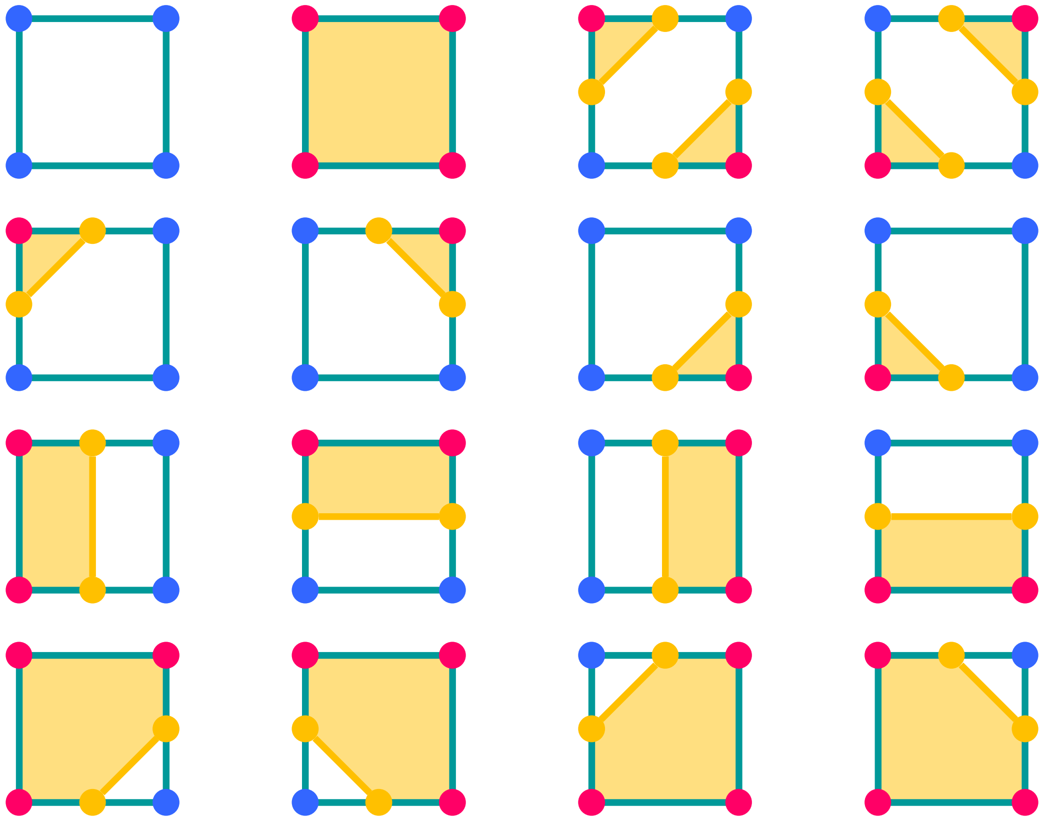
L. Mescheder et al., Occupancy Networks: Learning 3D Reconstruction in Function Space, **CVPR**, 2019

Fully Connected Residual Blocks

Marching Cube for 3D Visualization

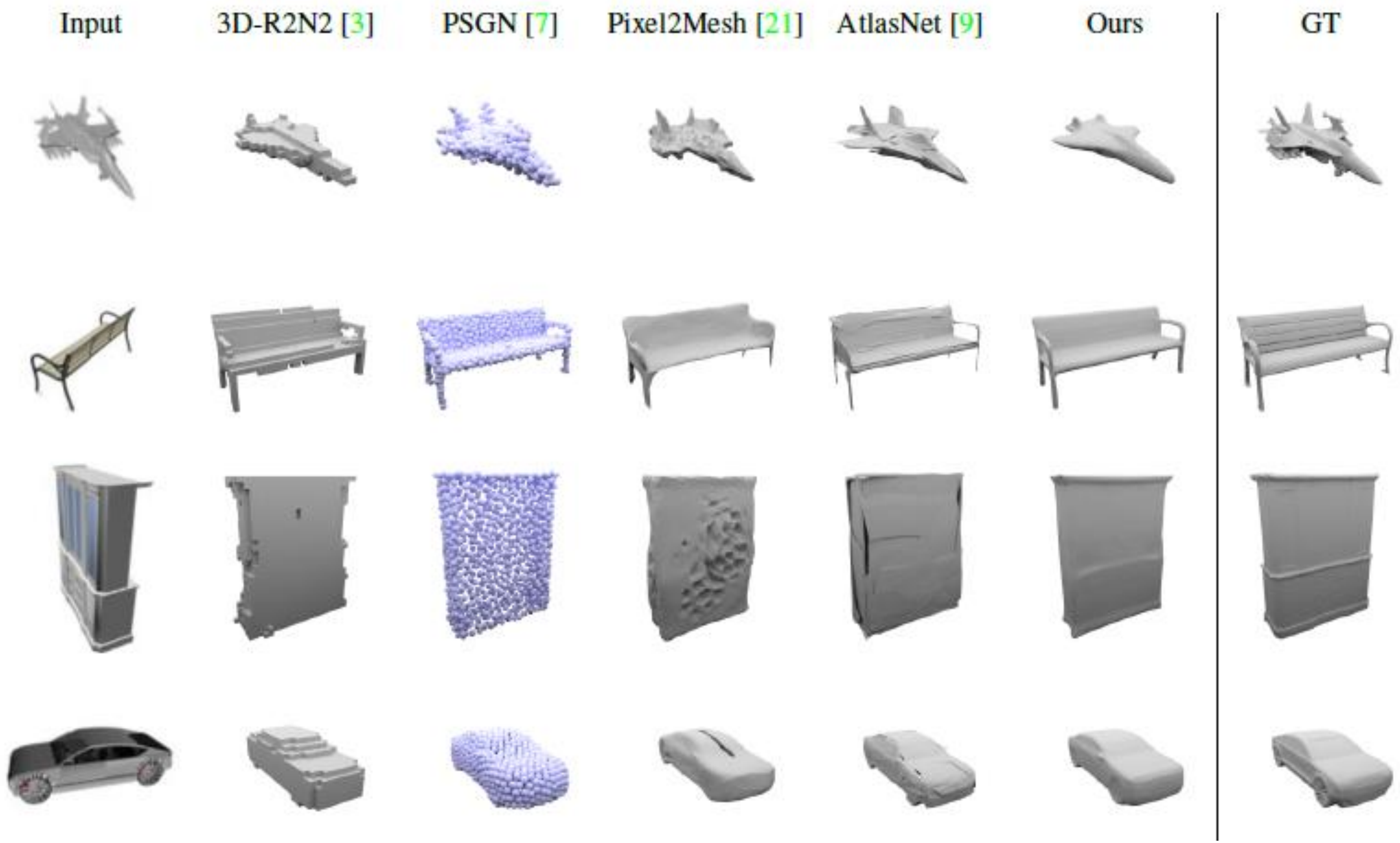
- Marching Cubes
  — A computer graphics algorithm for extracting a polygonal mesh of an isosurface from a three-dimensional discrete scalar field (e.g., point clouds)



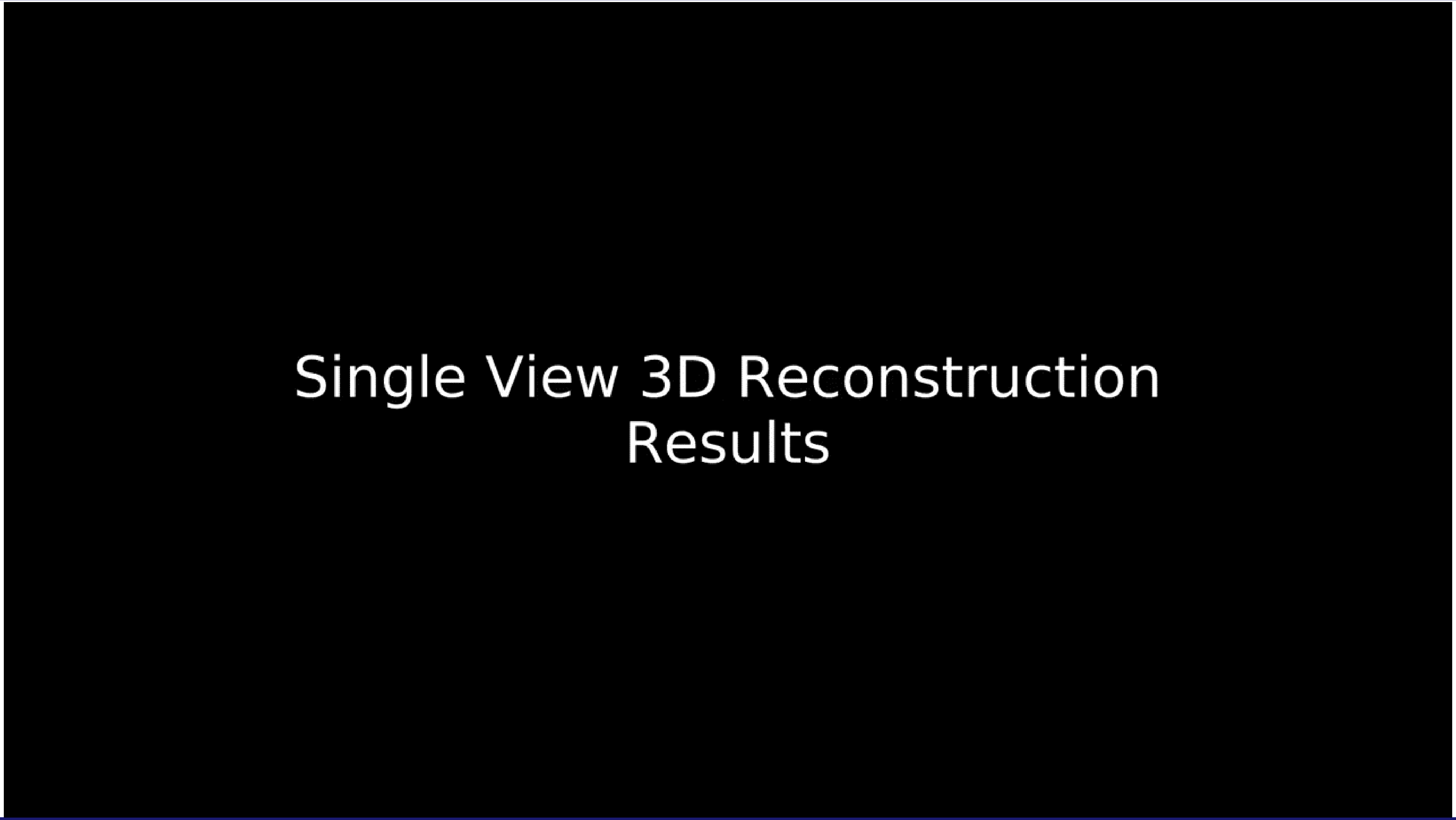


https://en.wikipedia.org/wiki/Marching_cubes

Results of Occupancy Network



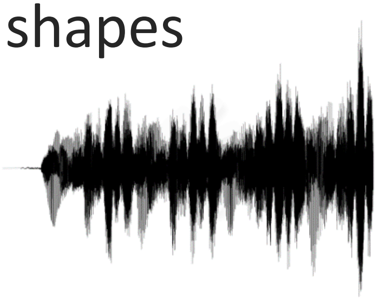
Results of Occupancy Network

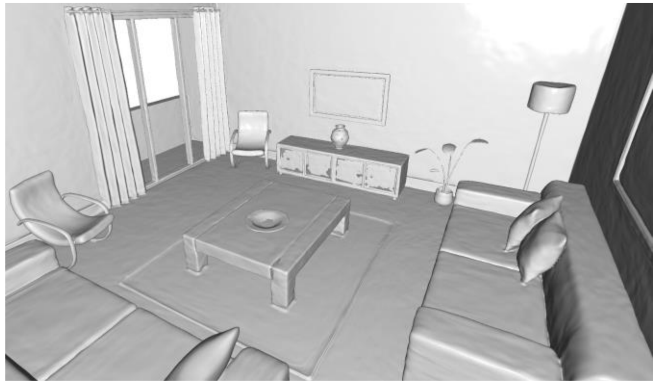


Problems of INRs

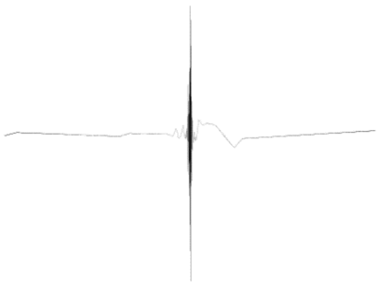
- Recent implicit neural representations build on ReLU-based MLPs lack the capacity to represent fine details in the underlying signals
    - Missing high frequencies in sound wave
    - Blurry images
    - Distorted 3D shapes

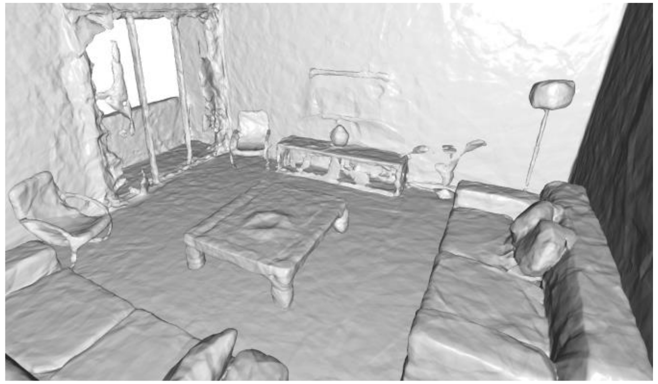
Original Signals





Signals represented by INRs (MLP w/ ReLU)

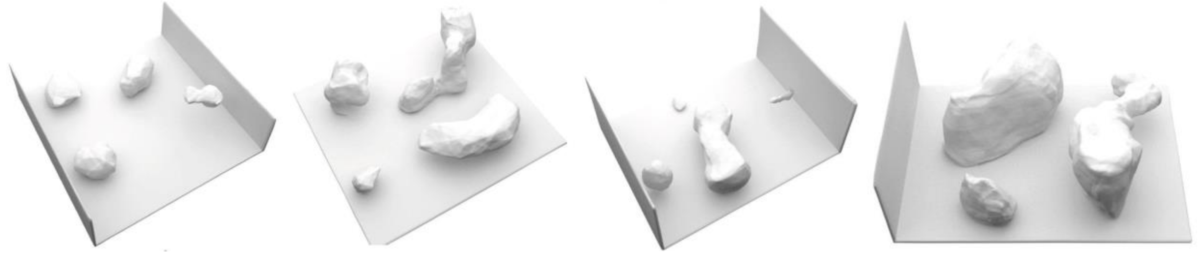




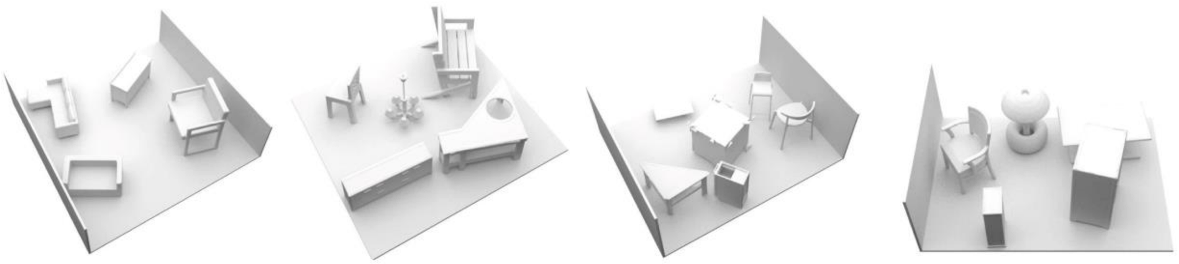
https://www.vincentsitzmann.com/siren/

Problems of INRs

- Recent implicit neural representations build on ReLU-based MLPs lack the capacity to represent fine details in the underlying signals
  - Missing high frequencies in sound wave
  - Blurry images
  - Distorted 3D shapes



ONet



Ground-Truth Mesh

ONet

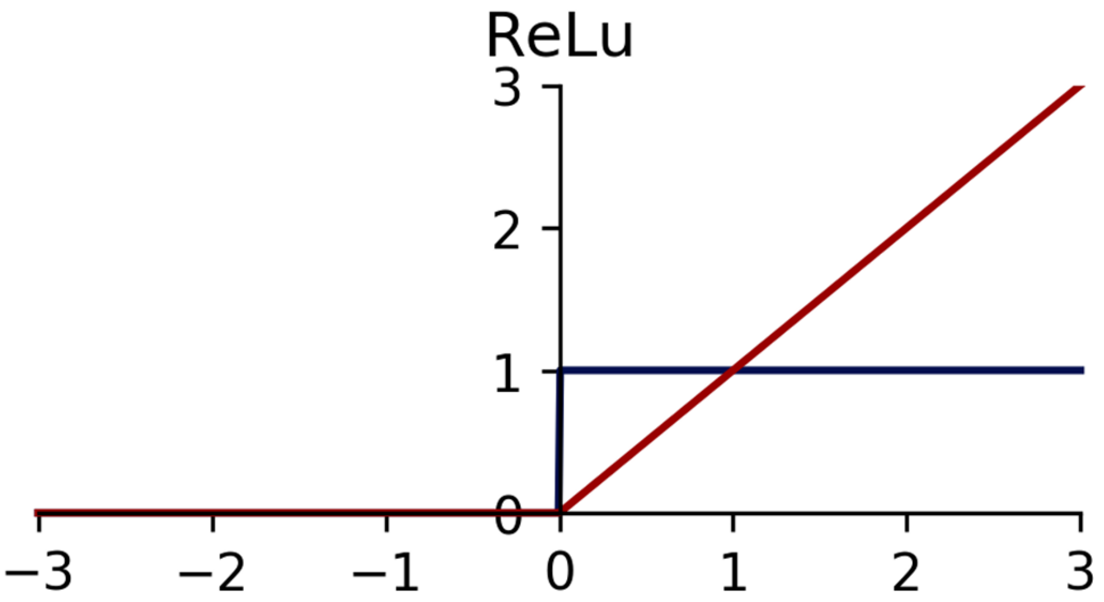
Ground-Truth Mesh

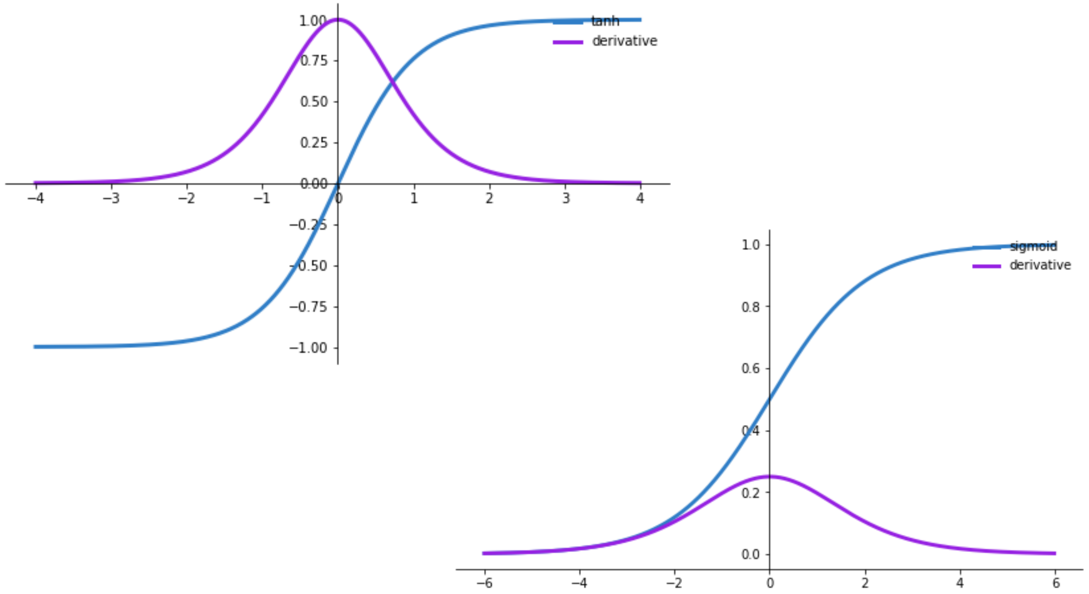
Limitations of **ReLU**

- Characteristics of ReLU

— Being piecewise linear → The second derivative is zero

— Incapable of modeling information contained in higher-order derivatives of natural signals





https://arxiv.org/abs/2006.13311
https://www.analyticsvidhya.com/blog/2021/04/activation-functions-and-their-derivatives-a-quick-complete-guide/

Sinusoidal Representation Networks (SIREN)

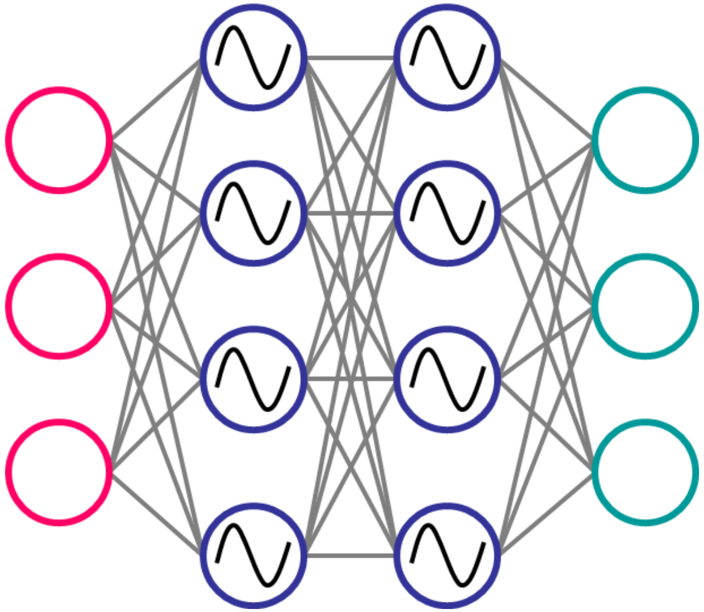
- Contributions of SIREN

— A continuous INR using periodic activation functions (sine) that fits complicated signals as well as their derivatives robustly

— An initialization scheme for training and validation that distributions of these representations can be learned using hypernetworks

— Demonstration of a wide range of applications:

  - image, video, and audio representation
  - 3D shape reconstruction
  - solving first/second-order differential equations



V. Sitzmann et al., Implicit Neural Representations with Periodic Activation Functions, **NeurIPS**, 2020

Problem Formulation

- Implicit problem formulation takes as in put the spatial coordinates $\mathbf{x} \in \mathbb{R}^m$ and, optionally, derivatives of Φ with respect to these coordinates

$$F(\mathbf{x}, \Phi, \nabla_x \Phi, \nabla_x^2 \Phi, \cdots) = 0, \qquad \Phi : \mathbf{x} \mapsto \Phi(\mathbf{x})$$

- Our goal is to find a class of functions $\Phi(\mathbf{x})$ that satisfies the relation F

Find $\Phi(\mathbf{x})$

subject to $\mathcal{C}_m(a(\mathbf{x}), \Phi(\mathbf{x}), \nabla_x \Phi(\mathbf{x}), \cdots) = 0, \;\; \forall \mathbf{x} \in \Omega_m, \;\; m = 1, \cdots, M$

V. Sitzmann et al., Implicit Neural Representations with Periodic Activation Functions, **NeurIPS**, 2020

SIREN: **Model Architecture**

- Periodic Activations for INRs

— MLPs with Sine activations

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

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

- ϕ_i : $\mathbb{R}^{M_i} \mapsto \mathbb{R}^{N_i}$ is the i-th layer of the networks
- $\mathbf{W}_i$: $\mathbb{R}^{M_i \times N_i}$ is the i-th weight matrix
- $\mathbf{b}_i$: $\mathbb{R}^{N_i}$ is the i-th bias vector
- $\mathbf{x}_i$: $\mathbb{R}^{M_i}$ is the i-th input vector

V. Sitzmann et al., Implicit Neural Representations with Periodic Activation Functions, **NeurIPS**, 2020

SIREN: **Initialization Scheme**

- Key idea of Initialization in SIREN

— It is to preserve the distribution of activations through the network so that the final output at initialization does not depend on the number of layers

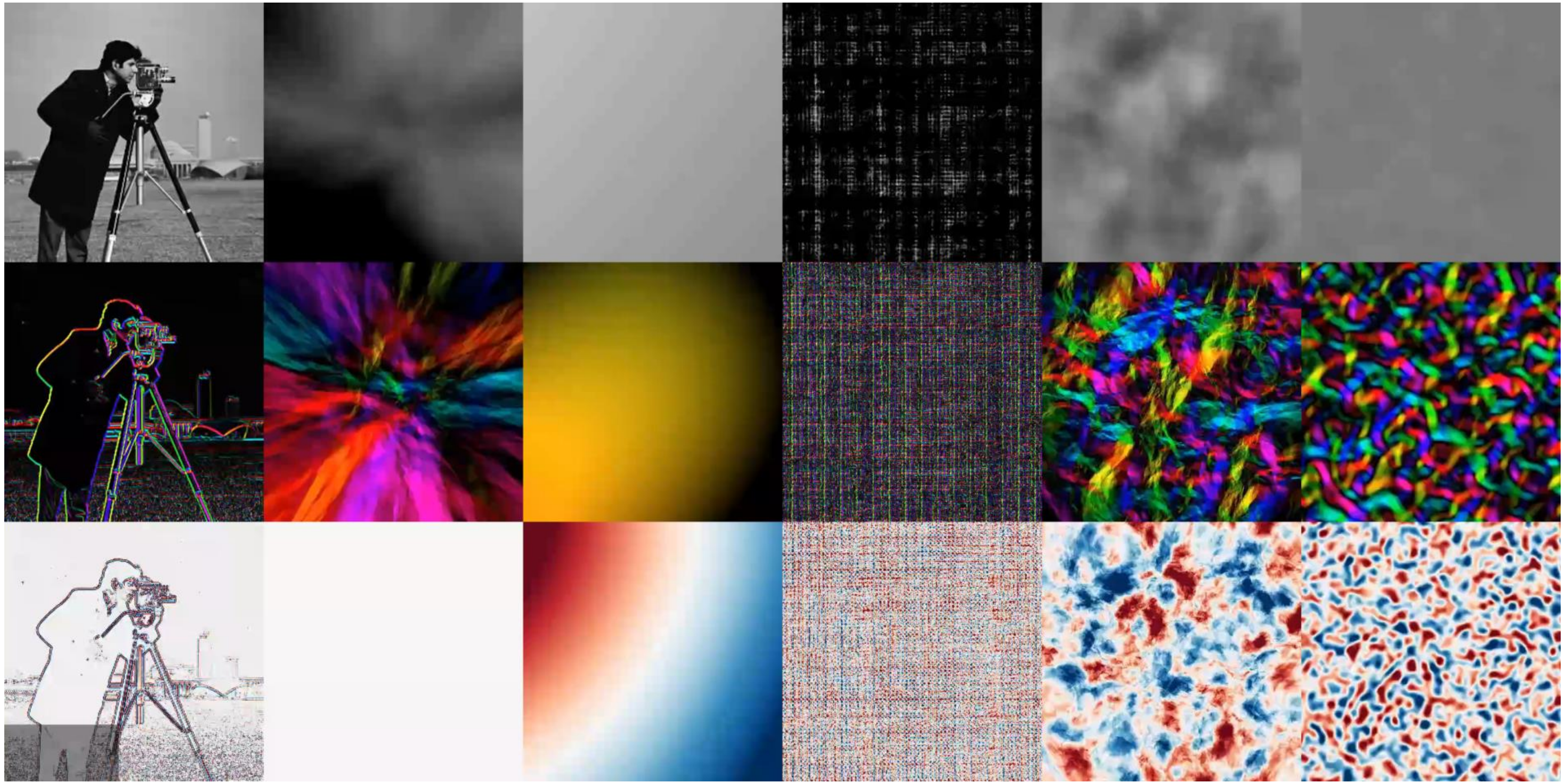
— Without carefully chose uniformly distributed weights, SIREN doesn't perform well

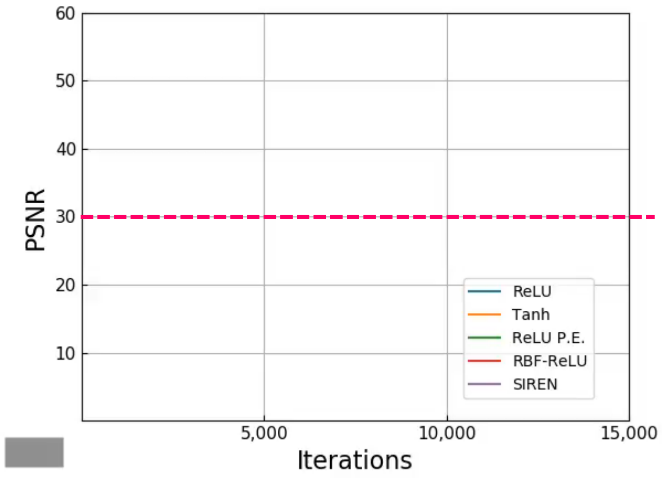
- Assuming that $\mathbf{x} \sim \mathcal{U}(-1,1)$
- For the 1st layer: Initialize the weights $\mathbf{W}$ of the 1st layer such that $\sin(\omega_0 \cdot \mathbf{W}\mathbf{x} + \mathbf{b})$ spans multiple periods over $[-1,1]$
- For other layers: Initialize the weights $\mathbf{W}$ according to $\mathcal{U}\left(-\sqrt{6/n}, \sqrt{6/n}\right)$

V. Sitzmann et al., Implicit Neural Representations with Periodic Activation Functions, **NeurIPS**, 2020

Experiments: Image Representation

- Mapping 2D Pixel Coordinates to a Color





https://www.vincentsitzmann.com/siren/

Experiments: Video Representation

- SIREN with pixel coordinates together with a time coordinate to parameterize a video

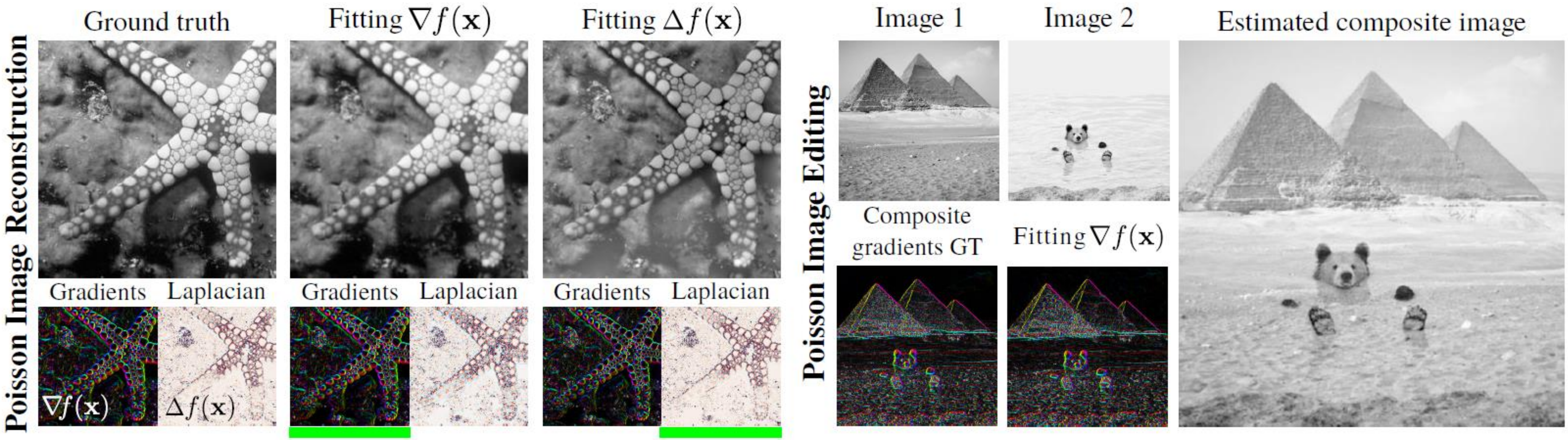
$$\mathcal{L} = \sum_i \|\Phi(\mathbf{x}_i) - f(\mathbf{x}_i)\|^2$$

https://www.vincentsitzmann.com/siren/

Experiments: Poisson Image Reconstruction

- SIREN for solving Poisson equation to reconstruct images from their derivatives

$$\mathcal{L}_{\text{grad}} = \int_{\Omega} \|\nabla_{\mathbf{x}}\Phi(\mathbf{x}) - \nabla_{\mathbf{x}} f(\mathbf{x})\| d\mathbf{x}$$

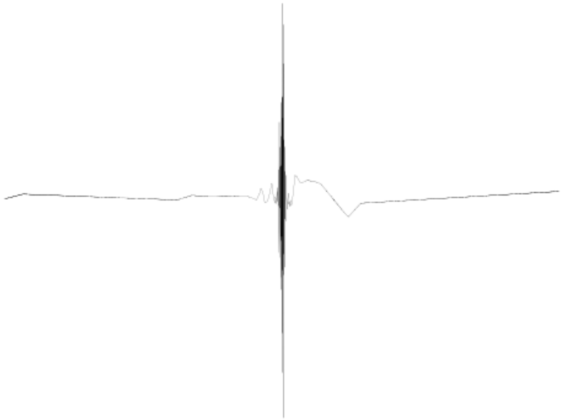


Experiments: Audio Representation

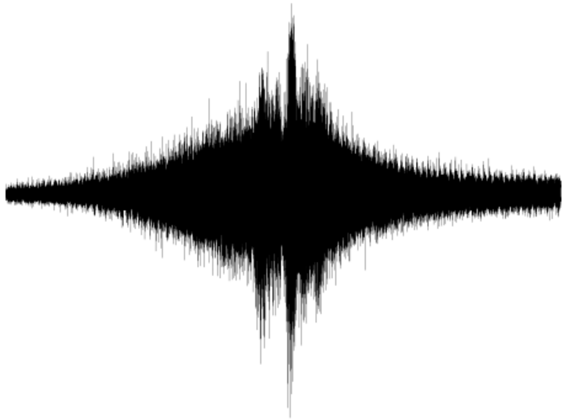
- SIREN with a single, time-coordinate input and scalar output may parameterize audio signals
  — It succeeds in reproducing the audio signal for music

$$\mathcal{L} = \int_{\Omega} \|\Phi(\mathbf{x}) - f(t)\| dt$$

Ground-Truth



MLP + ReLU



MLP + ReLU
w/ Positional Encoding

SIREN

https://www.vincentsitzmann.com/siren/

Experiments: 3D Scene Representation

- SIREN for fitting differentiable signed distance functions (SDFs) to represent 3D scene

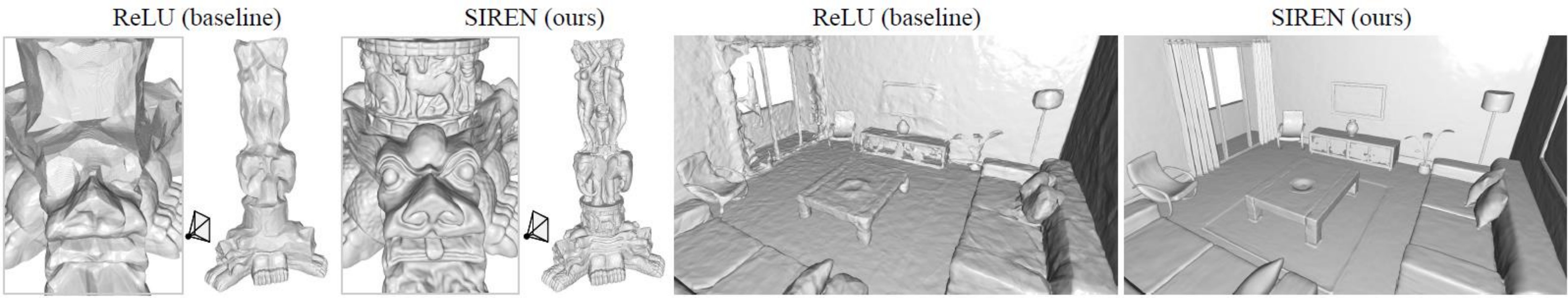
When $\mathbf{x}$ is on the surface,
SDF to be 1

When $\mathbf{x}$ is not on the surface,
penalize it

$$\mathcal{L}_{\mathrm{sdf}} = \int_{\Omega} \| |\nabla_{\mathbf{x}} \Phi(\mathbf{x})| - 1 \| d\mathbf{x} + \int_{\Omega_0} \|\Phi(\mathbf{x})\| + (1 - \langle \nabla_{\mathbf{x}} \Phi(\mathbf{x}), \mathbf{n}(\mathbf{x}) \rangle) d\mathbf{x} + \int_{\Omega \setminus \Omega_0} \psi\big(\Phi(\mathbf{x})\big) d\mathbf{x}$$

Gradient of SDF to be 1

Gradient of SDF to be
the normal vector of the surface



Experiments: 3D Scene Representation

Practice: INRs

- Practice

— Implicit Neural Representations with Periodic Activation Functions

  - https://github.com/vsitzmann/siren