Self-supervised Neural Signed Distance Function Model For Surface Reconstruction.

Shubham Godase (MT2308)
Internal Guide: **Dr. Bhalchandra Gore**
External Guide: **Mr. Jay Kumar Faldu**, **Mr. Sudhir Paithankar** (Dassault Systèmes)
Scientific Computing, Modeling and Simulation, SPPU, Pune, India

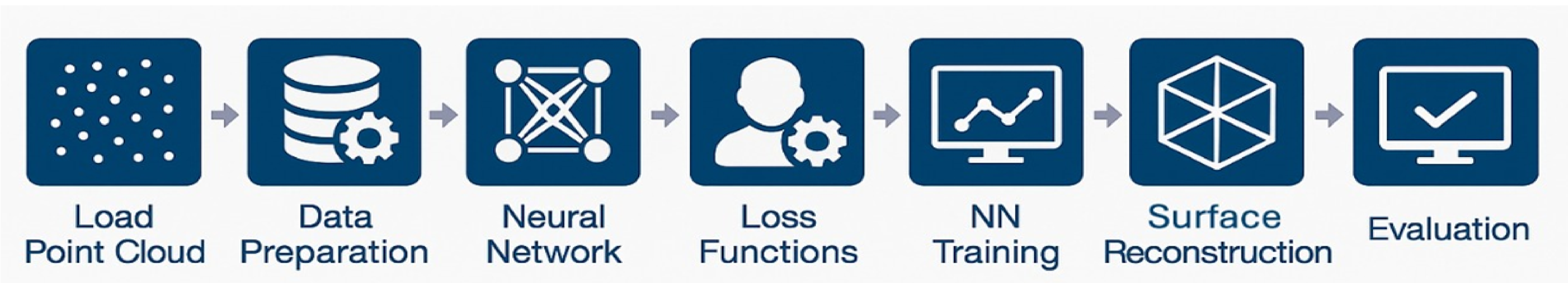
Abstract

We propose a self-supervised neural framework for 3D surface reconstruction from unstructured point clouds by learning an implicit Signed Distance Function (SDF) using Sinusoidal Representation Networks (SIREN), without requiring ground-truth distances or normals. The training process leverages geometric loss functions including Eikonal, Dirichlet (matching/non-matching) and Singular Hessian, with stability ensured through uncertainty-based loss reweighting, cosine learning rate scheduling, and gradient clipping. The predicted SDF is voxelized and meshed using Marching Cubes to produce watertight triangle surfaces. The method generalizes across complex geometries and is evaluated using Chamfer Distance, F-score, and Normal Consistency. Mesh quality is assessed via Trimesh, demonstrating applicability to CAD, medical imaging, and robotics.

Problem Statement

Reconstructing surfaces from point clouds is challenging due to missing connectivity, surface normals, and labeled distances. Traditional methods often require dense, clean data and strong supervision, which are rarely available in real-world scenarios. This work introduces a self-supervised SDF framework that learns continuous surface geometry directly from raw, sparse, and noisy point clouds without requiring ground-truth supervision or normal annotations.

Methodology



Data Description

Raw point samples are divided into three sets surface centroids P, near-surface perturbations Ω, and uniform off-surface samples Q to train the self-supervised SDF.

input point Sets:

- **P**: Input point cloud sampled directly from the 3D surface (centroid samples).
- **Ω**: Gaussian perturbations around each $p \in P$, where standard deviation σ is based on 50-nearest neighbor distance.
- **Q**: Uniformly sampled points within $[-1, 1]^3$ bounding box to prevent distant zero-level artifacts.

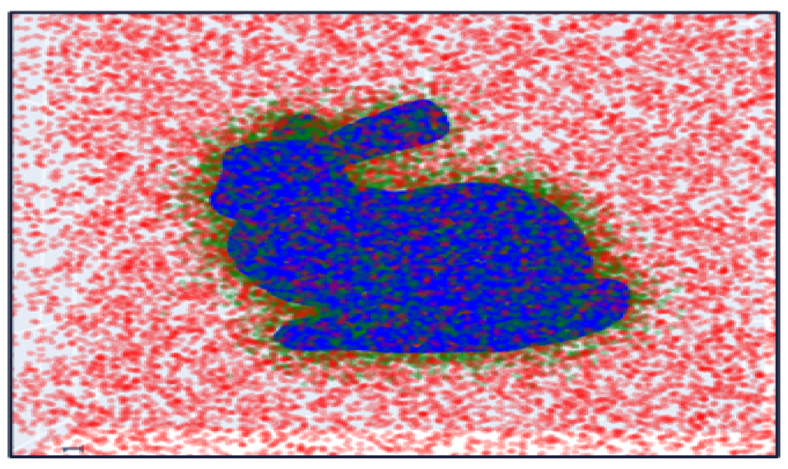


Figure 1. sample P(blue),Ω (green), Q (red)

Signed Distance Function (SDF)[3]

- Defines a scalar field $f : \mathbb{R}^3 \to \mathbb{R}$ representing the Signed Distance Function (SDF).
- Surface corresponds to the zero level set: $f(\mathbf{x}) = 0$.
- $f(\mathbf{x}) < 0$ inside the object, $f(\mathbf{x}) > 0$ outside.
- Gradient magnitude satisfies $\|\nabla f(\mathbf{x})\| = 1$ almost everywhere (Eikonal property), ensuring valid distance field behavior.

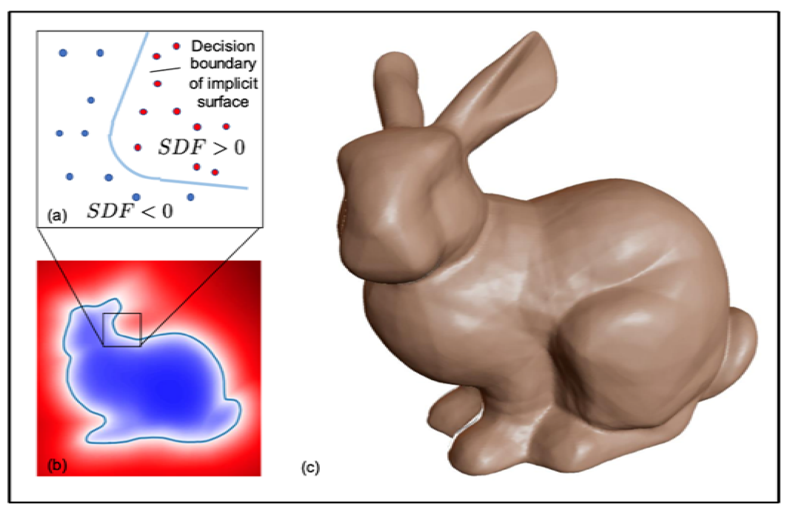


Figure 2. visual representation of SDF

Neural Network Configuration & Training:[4]

- **Architecture:** 6-layer MLP (4 hidden layers × 256 neurons), sine activations (ω=30)
- **Training Setup:** Adam optimizer, learning rate $= 1 \times 10^{-5}$, 20,000 epochs
- **Why SIREN?[4]** Periodic activations capture high-frequency geometry and enable gradient-based losses

Loss Functions

Learning an accurate Signed Distance Function (SDF) requires a combination of geometric loss terms that enforce correct distance properties, smoothness, and surface alignment without ground-truth supervision.[1]

- **Eikonal Loss (L_{Eik}):**

$$L_{Eik} = \frac{1}{|P| + |\Omega|} \int_{P \cup \Omega} |1 - \|\nabla f(\mathbf{x}; \Theta)\|| \, d\mathbf{x}$$

Ensures unit-norm gradients, enforcing true distance properties.

- **Dirichlet Matching Loss (L_{DM}):**

$$L_{DM} = \frac{1}{|P|} \int_P |f(\mathbf{p}; \Theta)| \, d\mathbf{p}$$

Forces predicted SDF to be zero on surface points.

- **Dirichlet Non-Matching Loss (L_{DNM}):**

$$L_{DNM} = \frac{1}{|Q|} \int_Q \exp\left(-\rho \, |f(\mathbf{q}; \Theta)|\right) d\mathbf{q}$$

Encourages off-surface points to stay away from zero level-set.

- **Singular Hessian Loss ($L_{SingularH}$):**

$$L_{SingularH} = \int_{Q_{near}} |\det\left(\mathbf{H}_f(\mathbf{x})\right)| \, d\mathbf{x}$$

Promotes Hessian singularity to preserve sharp features.

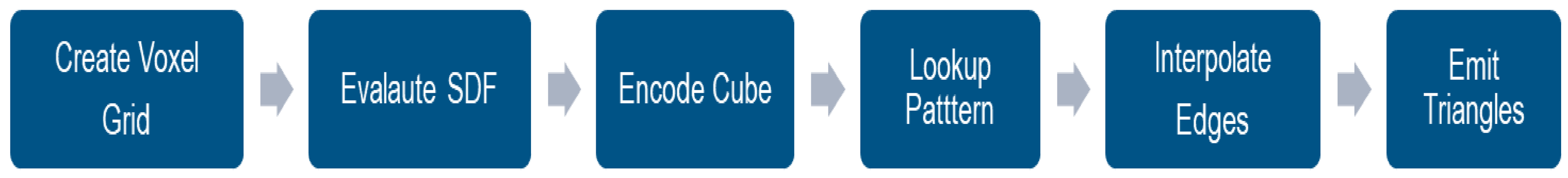
Total Loss Function:

$$L = \lambda_{\text{Eik}} L_{\text{Eik}} + \lambda_{\text{DM}} L_{\text{DM}} + \lambda_{\text{DNM}} L_{\text{DNM}} + \tau \lambda_{\text{H}} L_{\text{SingularH}}$$

where: $\lambda_{\text{Eik}} = 50$, $\lambda_{\text{DM}} = 7000$, $\lambda_{\text{DNM}} = 60$, $\lambda_{\text{H}} = 3$, and τ is an annealing factor.

Marching Cubes Surface Extraction

Marching Cubes Surface Extraction:[2] A fast voxel based algorithm that converts a continuous SDF grid into a watertight triangle mesh by marching through each cube, encoding its inside/outside pattern, then looking up and emitting the appropriate triangle configuration.



Results

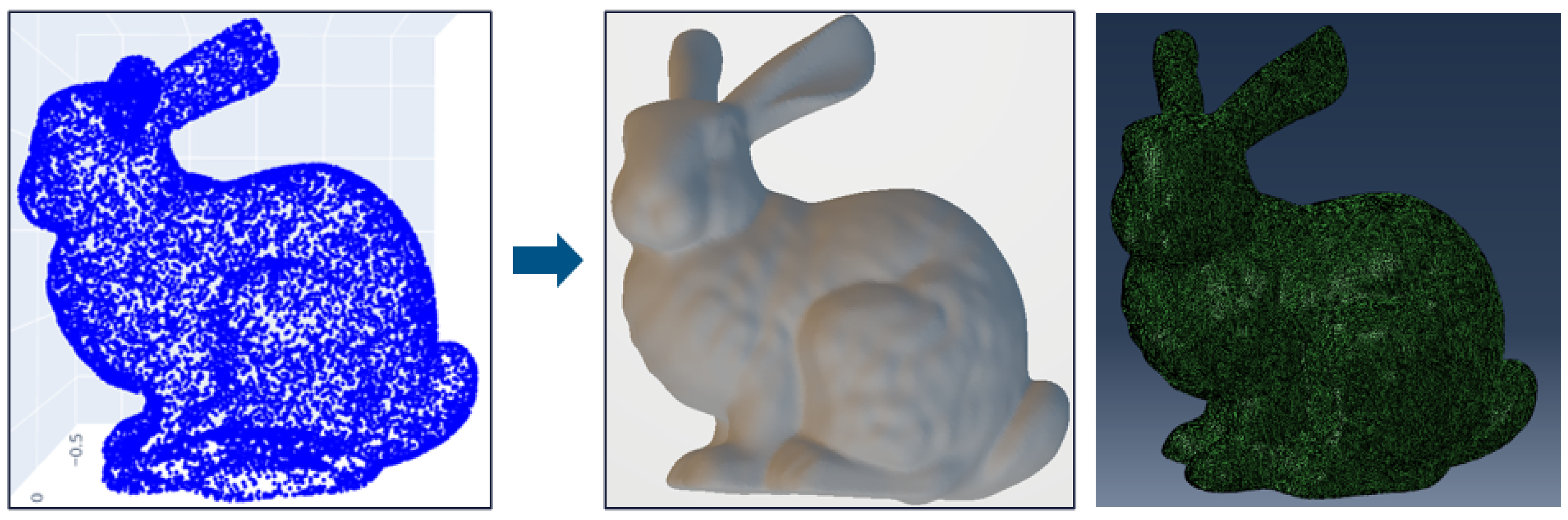


Figure 3. Bunny model:(a)point cloud, (b)reconstructed surface, (c)reconstructed trimesh

Evaluation Results

Quantitative results assess reconstructed (red) and ground-truth (blue) surfaces via Chamfer Distance (CD, scaled by 1000), F-Score (0.05 threshold), and Normal Consistency (NC), evaluated on 15,000 surface points per model.

Chamfer Distance (CD):

$$\text{CD}(P_1, P_2) = \frac{1}{2|P_1|} \sum_{p_1 \in P_1} \min_{p_2 \in P_2} \|p_1 - p_2\| + \frac{1}{2|P_2|} \sum_{p_2 \in P_2} \min_{p_1 \in P_1} \|p_2 - p_1\|$$

F-Score:

$$\text{F-Score}(t) = \frac{2 \times \text{Precision}(t) \times \text{Recall}(t)}{\text{Precision}(t) + \text{Recall}(t)}$$

Precision and Recall:

$$\text{Precision}(t) = \frac{|\{p_2 \in P_2 \mid \min_{p_1 \in P_1} \|p_2 - p_1\| < t\}|}{|P_2|}, \quad \text{Recall}(t) = \frac{|\{p_1 \in P_1 \mid \min_{p_2 \in P_2} \|p_1 - p_2\| < t\}|}{|P_1|}$$

Normal Consistency (NC):

$$\text{NC}(P_1, P_2) = \frac{1}{2|P_1|} \sum_{p_1 \in P_1} \mathbf{n}_{p_1} \cdot \mathbf{n}_{\text{NN}(p_1, P_2)} + \frac{1}{2|P_2|} \sum_{p_2 \in P_2} \mathbf{n}_{p_2} \cdot \mathbf{n}_{\text{NN}(p_2, P_1)}$$

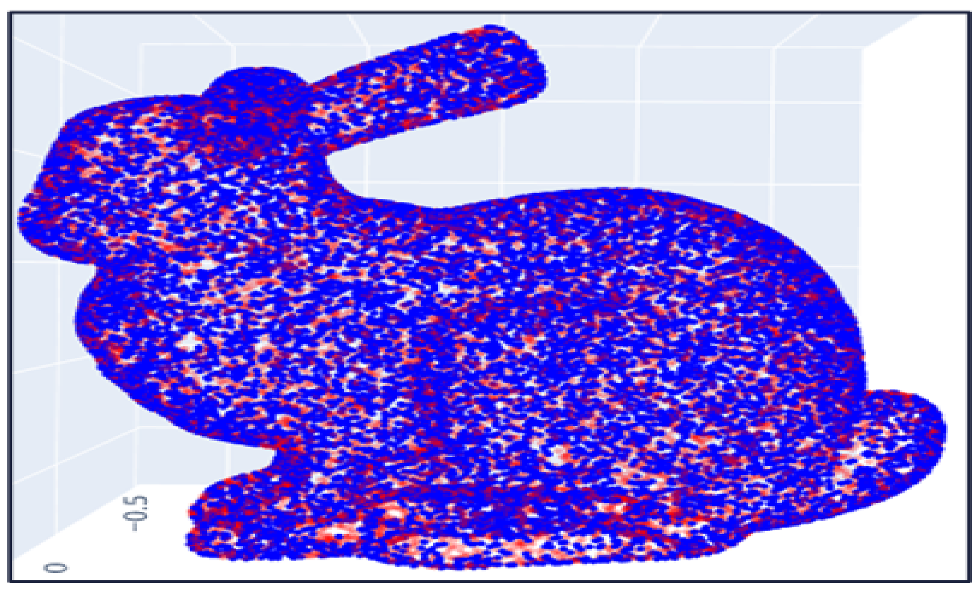


Figure 4. Predicted (Red) vs Ground Truth (Blue).

Model	CD	F-Score	NC
Bunny1	3.23	0.989	0.992
Bunny2	4.85	0.932	0.987
Sphere	1.22	0.991	0.992
Cylinder	2.03	0.987	0.988
Explicit Lug	4.89	0.940	0.992
Elephant	3.77	0.934	0.971
Petal	4.99	0.985	0.991
Sacrum Bone	4.45	0.912	0.979
Cervical Vertebrae	4.18	0.936	0.992

Table 1. Quantitative Metrics

Conclusion

The proposed self-supervised SDF framework enables high-fidelity mesh reconstruction from sparse, noisy, and unordered point clouds without ground-truth supervision. Leveraging geometric losses with SIREN and implicit representation learning, the method accurately reconstructs watertight surfaces across diverse CAD-like and organic shapes. Quantitative evaluations (CD, F-Score, NC) confirm its robustness and precision. Overall, this scalable framework provides a flexible foundation for learning-based 3D surface modeling applicable to CAD, medical imaging, and robotics.

References

[1] Q. Dong, H. Wen, R. Xu, X. Yu, J. Zhou, S. Chen, S. Xin, C. Tu, and W. Wang. Neurcross: A self-supervised neural approach for representing cross fields in quad mesh generation.

[2] W.E. Lorensen and H.E. Cline. Marching cubes: A high resolution 3d surface construction algorithm.

[3] J.J. Park, P. Florence, J. Straub, R. Newcombe, and S. Lovegrove. Deepsdf: Learning continuous signed distance functions for shape representation.

[4] V. Sitzmann, J. Martel, A. Bergman, D. Lindell, and G. Wetzstein. Siren: Implicit neural representations with periodic activation functions. https://github.com/vsitzmann/siren, 2020.