International Journal of Multidisciplinary Research and Growth Evaluation.

# LLM for Differentiable Surface Sampling for Masked Modeling on Point Clouds

**Zeyu Wang [1*], Wenjian Sun [2], Zong Cheng Chu [3], Yiqian Zhang [4], Zhizhong Wu [5]**
[1] University of California Los Angeles, USA
[2] YanTai university, China
[3] ByteDance, USA
[4] State Key Laboratory of Biotherapy, West China Hospital, China
[5] University of California Berkeley, USA

* Corresponding Author: **Zeyu Wang**

## Article Info

**ISSN (online):** 2582-7138
**Volume:** 05
**Issue:** 05
**September-October** 2024
**Received:** 20-07-2024
**Accepted:** 22-08-2024
**Page No:** 440-447

**Abstract**
We present MaskPOINT, a novel scheme of masked-wise pretrained models for point cloud self-supervised learning, addressing the challenges posed by 3D understanding, including the ambiguous underlying geometry and irregular topology information. In fact, surface-based geodesic topology pro- vides strong cues for 3D semantic analysis and geometric modeling, but such connectivity information is lost in point clouds and hardly considered. In view of this, we formulate this task into an point sampling problem and man- ifold understanding problem, and develop two techniques, a differentiable topology sampling module and manifold contrastive learning, to enable it. Specifically, the *Topology Sampling* samples a point cloud into point cloud geodesic field given a query point. Then, the *Manifold Understanding* enables the point cloud dived into topology strands based on manifold contrastive learning. Then, we randomly mask out some strands of input topology and feed them into the vanilla Transformers. The pre-training objective is to recover the underlying geometry and irregular topology information the masked. We evaluate our pretrained models across several downstream tasks, including point clouds classification, 3D segmentation, few-shot, registra- tion and demonstrate competitive results while representations learned by Mask Point owning strong interpretability.

**DOI:** https://doi.org/10.54660/.IJMRGE.2024.5.5.440-447

**Keywords:** communicative language teaching, grammar, language teaching

## 1. Introduction

Mask-based pre-training technology attracts many researchers by being influential in various ML scenarios, including robustness [2, 11, 12, 13, 14, 15, 16, 17], regularizer [3, 7, 22, 21, 22, 23, 34] and neural network interpretability [32, 33, 34, 5, 7]. In terms of self-supervised learning (SSL), the key enabler of mask-based SSL is to recover the mask context based on the pretext task. Thanks to Transformer, most of mask-based SSL framework is based on that strong backbone, because of its tokenlizer ability and captures long-range dependence. Starting from masked language modeling, BERT and its variant achieve state-of-the-art performance across nearly all NLP downstream tasks by predicting masked tokens during pretraining. Notably, recent image self-supervised studies start to explore a novel mask-wise reconstruction methods, which inspired by Masked Language Modeling (e.g, BERT), namely mask image modeling. In general mask image modeling is to formulate the image as visual token that can be better described by the ViT transformer. In general, prior arts attempt to figure out a problem, that is how to leverage original image features as a self-supervised signal facilitating tokenization image [50, 44, 33, 29, 27].

**Fig 1:** Mask Topology pre-training pipeline

Despite, tokenize produce achieves promising results in 2D vision [**?**] and language community, it is under-explored in the 3D vision and graphics field yet. Several attempts have been made for self-supervised representation learning on point clouds tasks, such as reconstruction [] and generation. Technically, point cloud pre-training model can be roughly categorized into two branches, contrastive learning and transformer-friendly pre-training. For the first line, the 3D point contrastive learning either contrast among various point clouds or different projected views of the input point cloud for representation learning. On the other hand, exiting transformer-based point cloud models biases local feature aggregation and neighbor embedding, where betrays the tokenize 3D points and hardly servers as transformer-friendly pre-training. Until recently, such positive initiatives in image and language are hardly replicated in 3D point cloud. That shares us a hint *what makes masked 3D point cloud problem such tricky*. We try to figure out that from the following observations:

1. **Geometric Vocabulary.** Different from language [Devlin et al. researches omit point cloud surface topology that possesses crucial self-supervised contents. We advocate, a geometric vocabulary that hold geometric features and surface topology indeed refine self-supervised 3D point pre-training
2. **3D Sampling.** The other essential difference between the 3D point cloud and already explored self-supervised modal is sampling process. Both language and image rarely toward sampling to satisfy downstream tasks, opposite full information offers more high-frequency rich semantics. In contrast, the raw 3D point data is redundant or sparse posing a major challenge to obtain vital features. One widely solution is to employ Farthest Point Sampling (FPS) that selecting a subset of points in which points are farthest away from each other to describe global point cloud. However, classic sampling never value the downstream task, while recent work showed that learning a task-specific sampling are helpful. Yet, the proposed techniques overlook the point topology task-specific for self-supervised and even not differentiable.

To this end, we present **Mask**ed **Topology** sampling for self-supervised 3D geometric pre-training to jointly analyze the 3D point-wise masked modeling from the previous two non-trivial points. Driven by previous analysis, we first dig the deeper into the fundamental problem of point clouds and present a fully *differen- tiable topology sampling*. Superior to OcCo [Wang et al. 2021a], the current state-of-the-art describes an encoder-decoder architecture to reconstruct the occluded viewpoints, via a non-differentiable generation.
In terms of better tokenize *geometric vocabulary*, we follow the popular study[**?**] that predicts the vanilla 3D point typology proper- ties (e.g, normal, geodesic, curvature) instead of directly predicting point patches that largely empowers the SSL point to learn 3D silhouette.

## 2. Related Work
### 2.1. Transformers
The Transformer [Vaswani et al. 2017] backbone been immensely dominated NLP society, such as [Brown et al. 2020; Qian et al. 2021]. Thanks to its minimal inductive and long-distance capturing feature, Transformer starts to empower 2D vision as a strong backbone, starting from Dosovitskiy. et al. Latterly, kind of strong visual Transformers backbone like Swin-ViT [Liu et al. 2022c, 2021] and ViT surpass ConvNets among various tasks, such as detection [Kim et al. 2021], instance segmentation[Wang et al. 2021b], image in painting [Yu et al. 2021b]. Despite there is a trend of unifying transformer in the field of NLP and 2D vision, the development of transformer in 3D vision falls behind the expected. Luckily, there is some of preliminary studies initiating the Transformers-based point cloud backbone. For instance, PCT [41, 43, 44, 45, 46, 47, 48] employs a coordinate- based embedding layer while altering the self-attention mechanism with an offset-attention optimization process that can be better similarly formulated as a Laplace process. Besides, Point Transformer also modifies the Transformer layer, which enables invariant to permutation that smoothly suits to point cloud processing. Driven by previous efforts, Point-BERT [Yu et al. 2022] success in encouraging standard Transformers to 3D point cloud learning, with minimal inductive bias. Despite this, we argue that prior art [Yu et al. 2022] and some non-peer review work[Liu et al. 2022a; Pang et al. 2022] naively tokenize point cloud into transformer-friendly parches, yet makes the transformer backbone out of the surface manifold in point cloud. By contract, we seek to introduce the surface and topology feature of point cloud into the transformer backbone to better understand global point cloud structure.

### 2.2. Sampling methods for Points Cloud
Point cloud sampling serves as a fundamental technology of the learning-based point cloud process, which refines the raw inputs while reducing computational costs for downstream tasks. Widely- adopted point cloud sampling methods include RS, FPS, and IDIS. A handful of recent works began to explore advanced and so- phisticated sampling schemes. Nonetheless, despite the remarkable progress in point cloud sampling, these methods are task-agnostic and rather universal, lacking awareness of the important features which a particular task may require. Recently, Dovrat et al. pro- posed a learnable, data-driven sampling strategy by imposing a specific task loss to enforce the sampler in learning specific related features for a particular task. Later, Lang et al extended the learning approach by introducing a differentiable relaxation to minimise the training and inference accuracy gap for the sampling operation. Nevertheless, by introducing an additional task loss, sampling ultimately becomes constrained and prone to over fitting on a specific task model instead of the task itself. Additionally, this also requires significant extra training to fit a particular goal.

### 2.3. Self-supervised Learning
Mask-wise Pretraining. Mask-based pretraining was initially pro- posed in natural language processing (NLP) tasks. Bidirectional Encoder Representations from Transformers (BERT) [37] enables representation learning from unlabelled

texts in a self-supervised manner, in which the Masked Language Modeling (MLM) makes the major contribution. Following BERT, a number of studies researched mask-wise pretraining in linguistic models [60, 61, 62, 63, 64, 65]. Because of the shift from convolution to Transformer backbones in computer vi- sion tasks, BEiT first adapted the MLM strategy to image backbone pretraining. BEiT tokenizes the 2D image into different visual tokens, after which, it randomly masks some cells and feeds the corrupted feature map into the backbone. The model is trained to recover the tokens of the masked cells. Lately, Masked Auto-Encoder (MAE) [55] further pushed the limits of mask-wise pretraining. MAE directly train the backbone to recover the par- tially masked images in an end-to-end manner. In 3D vision tasks, Point-BERT [Yu et al. 2022] was the first mask-wise pre-training strategy. In this work, we .. d

SSL for 3D. Self-supervised learning (SSL) has gained tremendous visibility in computer vision tasks, which enables representation learning without labels [70, 71, 72, 73, 74, 75, 76]. [?] proposed rotation-based SSL in 3D point cloud. [?] presented jigsaw puzzle-based SSL in the 3D point cloud domain. Auto-encoding are also of better qualities demonstrated that robustness improve- ments can be gained with SSL in 3D adversarial training. [**?**] proposed OcCo, a completion-based pretraining in 3D point clouds to better capture the geometric feature of 3D data. PointContrast lever- ages contrastive loss on different persepctive of 3D scenes to train the backbone models [71, 45, 44]. As mentioned earlier, Point-BERT [Yu et al. 2022] was the first study on adapting mask-wise pretraining. In this study, we present a more sophisticated mask-wise.

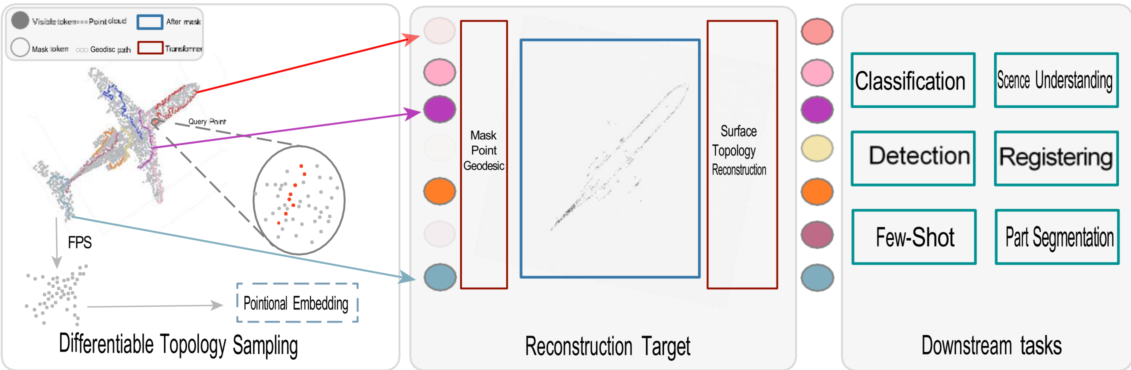


**Fig 2:** Mask POINT pre-training pipeline

## 3. Masked Topology

The overall framework of our Mask POINT is illustrated in Fig. 5.

Given masked and embedded point patches, we learn the transformer- based encoder and decoder to reconstruct the underlying masked manifold by simultaneously predicting the 3d parallel vectors field. In the following subsections, we introduce the main components in detail Mask POINT system.

### Differentiable Topology Sampling

We formulate construct surface-based geodesic topology of point cloud as a point sampling problem. Recent efforts [Lang et al. 2020] explore the differentiable relaxation for point cloud sampling toward addressing loss-relevant tasks. But, they ignore the fundamentally understand the topology of the point could. In the following, we will give a brief description of Sample Net and clarify our differences to it.

**Sample Net** Given a set of point clouds which Sample Net targets at sampling a smaller and concentrate subset of points, which normally equips a fix-defined number of sampled points so that the downstream task working on the subsets can be optimized. To make the whole process differentiable, SampleNet define a small point group for each, then the small point group project onto raw point clouds $xn$ to obtain final input to the down- stream task. However, learn task-specific sampling (i.e., perhaps suitable for not a surface-mind task such as re- construction), while we take a step deeper and aim to propose an surface-friendly differentiable sampler.

**Topology Sampling** The goal of the topology sampling is to develop a learning-based sampling module $f\theta$ with trainable parameters, which takes in a point cloud with m points and outputs a geodesic manifold between two points . It takes as input the anchor point, also its neighborhood and with the current path descriptor $\gamma t$ and predicts at every time step $t$ the next point in the path sequence using an differentiable heuristic manner:

In other word, we can formulate the problem as how to obtain a topology priors point representation, and surface coverage path descriptor with a reasonable heuristic point selector. To exploit topology of the 3D point cloud, we leverage DGCNN [Wang et al. 2019] rather than Point Net to encoder the raw manifold feature. Actually, assuming two points on two distinct fingertips that are considered close in the Euclidean space, these points are actually distant on the surface curve and consequently in the graph feature space. However, naive DGCNN [Wang et al. 2019] expressively describe neighborhood information using a single aggregation function. Sim- ilar to Sharp et. al [2020], we propose a graph node convolution operation for node feature update.

Where is the feature of point $xi$ in layer, $\theta n$ denotes the parameters of $n$-th layer and N$i$ is the neighbors of point. Differ- ent from naive solution, the proposed relative feature formulation enables point clouds decomposition from the shape topology.

In addition, we need to proper encode the path direction to avoid predictions that are not aligned with the current predicted path. In particular, loops around nodes along with misleading direction vectors can derail the path from the target point. To avoid that, we The key point is to define how to include the next point in the sequence via a differentiable process. We leverage the transition probabilityto make this fully differentiable:

$$\mathcal{P}\left(\mathbf{p}^{t+1} = \mathbf{x}_j \mid x^t, Q\right) = \text{softmax}\left(\text{MLP}\left(\left[r_j \| f_j \| \gamma^t\right]\right)\right)$$

where $Q$ is the point set that $x$ lie in and MLP denotes a multilayer perceptron. In this paper we used a 5 layer MLP accompanied with ReLU activations.

**Neural Architecture**

Token Embedding. Before forwarding the visible point patches Pvis to the encoder, we first embed them via token embedding. Following, we instantiate the token embedding with a lightweight Point Net, which is composed of multi-layer perceptrons (MLP) and a max pooling f

**Reconstruction Target**

## 4. Experiments

### 4.1. Pipeline Setups

### 4.2. Downstream Tasks

In this subsection, we verify our methods via the five main point cloud downstream task and report the results on the widely used benchmarks. Besides the previous studies [Yu et al. 2022] visits tasks, we firstly explore the possibility of the point cloud 3D registration. On the hand, 3D registration servers as an essential task while requires strong prior knowledge of 3D surfaces, that is a ideal down- stream task to confirm our assumption.

Object Classification. We conduct our object classification experiments on Model Net40 [?], which includes 12,311 CAD models from 40 categories. 9,843 models are used for training while 2,468 models for testing. We evaluate our methods ranging the point cloud resolutions from 1024, 4096, and 8192 to three levels. And all the levels of the point cloud is used a uniform sampling from CAD mesh faces.

*Real Word Dataset.* We next conduct experiments using the real- world scan dataset Scan Object NN Table 2 shows the results. Our approach achieves SOTA performance on all three splits. On the hardest PB s.

**Table 1:** Performance of different methods on point cloud datasets

| Methods | OBJ-BG | OBJ-ONLY | PB-T50-RS |
|---|---|---|---|
| Point Net | 73.3 | 79.2 | 68.0 |
| Spider CNN | 77.1 | 79.5 | 73.7 |
| Point Net++ | 82.3 | 84.3 | 77.9 |
| Point CNN | 86.1 | 85.5 | 78.5 |
| DGCNN | 82.8 | 86.2 | 78.1 |
| BGA-DGCNN | 81.1 | 82.9 | 79.7 |
| BGA-PN++ | 77.6 | 78.9 | 80.2 |
| Transformer | 79.86 | 80.55 | 77.24 |
| Transformer-OcCo | 84.85 | 85.54 | 78.79 |
| Point-BERT [Yu et al. 2022] | 87.43 | 88.12 | 83.07 |
| Mask Point [Liu et al. 2022b] | 89.35 | 88.10 | 84.3 |
| Mask POINT (ours) | 89.23 | 87.22 | 85.60 |

*Part Segmentation.* We visualize the feature distributions using tSNE. Fig. shows the features learned by Mask POINT after self- supervised training on the ShapeNet dataset.

**Table 2:** Performance of different methods on MioU

| Method | Mean inst. mIoU |
|---|---|
| Point Net++ | 85.1 |
| Point CNN | 86.1 |
| DGCNN | 85.2 |
| KP Conv deform | 86.4 |
| KP Conv rigid | 86.2 |
| GDA Net | 86.5 |
| Point Transformer [Zhao et al. 2021] | 86.6 |
| Point Voxel Transformer | 86.5 |
| Point-BERT [Yu et al. 2022] | 84.8 |
| Mask POINT (ours) | 85.6 |
| Mask Point[Liu et al. 2022b] | 86.1 |

*Scenes Understanding.* We conduct the semantic segmentation on the Stanford 3D Indoor Scene Dataset (S3DIS), which contains 6 large-scale indoor areas with points shared by 13 classes.

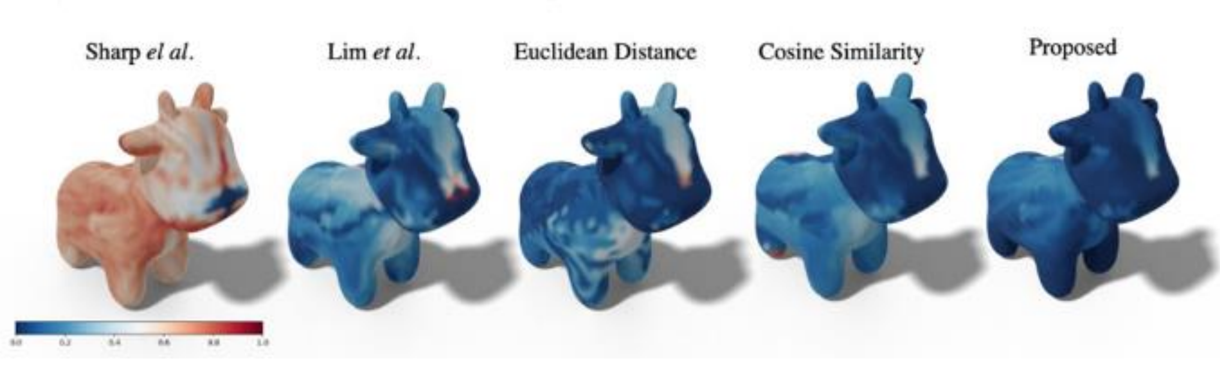


**Fig 3:** Mask POINT geodisc

**Table 3:** Performance of different methods in various indicators of image4x100

| Methods | Input | OA | mAcc | mIoU |
|---|---|---|---|---|
| Point Net | Image 4x100 | 83.4 | 49.0 | 41.1 |
| Point CNN | Image 4x100 | 85.9 | 63.9 | 57.3 |
| KP Conv | Image 4x100 | 86.2 | 72.8 | 67.1 |
| Transformer | Image 4x100 | 86.8 | 68.6 | 60.0 |
| Transferring features protocol | | | | |
| Point-MAE [Pang et al. 2022] | Image 4x100 | 87.4 | 71.3 | 61.0 |
| Mask Point (Ours) | Image 4x100 | 88.3 | 74.1 | 62.1 |
| Non-linear classification protocol | | | | |
| Point-MAE[Pang et al. 2022] | Image 4x100 | 85.3 | 65.4 | 56.1 |
| MaskPOINT (Ours) | Image 4x100 | 83.5 | 45.8 | 81.5 |

*Object Detection.* Our most closely related work, Point-BERT showed experiments only on object-level classification and segmentation tasks. In this paper, we evaluate a model's pretrained representation on a more challenging scene-level downstream task: 3D object detection on ScanNetV2.

**Table 4:** AP performance of different methods under Geo input

| Methods | Methods | Pretrained Input | $AP_{25}$ | $AP_{50}$ |
|---|---|---|---|---|
| STRL | ✓ | Geo | 59.5 | 38.4 |
| Implicit Auto encoder | ✓ | Geo | 61.5 | 39.8 |
| Random Rooms | ✓ | Geo | 61.3 | 36.2 |
| Point Contrast | ✓ | Geo | 59.2 | 38.0 |
| Depth Contrast | ✓ | Geo | 61.3 | 39.0 |
| Depth Contrast | ✓ | Geo + RGB | 64.0 | 42.9 |
| 3DETR | ✓ | Geo | 62.1 | 37.9 |
| Point-BERT | ✓ | Geo | 61.0 | 38.3 |
| Mask Point | ✓ | Geo | 63.4 | 40.6 |
| Mask Point (12-layers) | ✓ | Geo | 64.2 | 42.0 |
| Mask POINT (ours) | ✓ | Geo | 61.7 | 42.8 |

*Registration.* We further conduct comparison evaluation on real- world indoor scenes: 7Scenes.

**Table 5:** Four evaluation metrics for different models

| | 7 Scenes | | | |
|---|---|---|---|---|
| **Model** | **MAE(t)** | **MAE(r)** | **MIE(t)** | **MAE(r)** |
| ICP | 2.4022 | 0.0699 | 4.4832 | 0.1410 |
| Go-ICP | 0.7339 | 0.0259 | 1.3927 | 0.0522 |
| S4PCS | 1.3964 | 0.0547 | 2.7792 | 0.1103 |
| FGR | 2.2477 | 0.0808 | 4.1850 | 0.1573 |
| RANSAC | 1.2349 | 0.0429 | 2.3167 | 0.0839 |
| IDAM | 4.4153 | 0.1385 | 8.6178 | 0.2756 |
| RPM-Net | 0.3267 | 0.0125 | 0.6277 | 0.0246 |
| FMR | 1.1085 | 0.0398 | 2.1323 | 0.0786 |
| Point-MAE[Liu et al. 2022a] | 0.2374 | 0.0005 | 0.3987 | 0.0010 |
| Ours | 0.0492 | 0.0023 | 0.0897 | 0.0049 |

*Few-shot Learning.* We conduct the experiments of few shot learning on the Scan Object NN dataset under the "n-way, m-shot" setting, where n is the number of randomly sampled classes and m is the number of samples in each class. The n$x$m samples are adopted for training, while we randomly sample 20 unseen objects from each class for testing. We report the results of each setting with 10 independent experiments. Results with $n$ = 5, 10 and $m$ = 10, 20 are presented in Tab.

**Table 6:** Four evaluation metrics for our models

| | 5-way | | 10-way | |
|---|---|---|---|---|
| | **10 -shot** | **20 -shot** | **10 -shot** | **20-shot** |
| Transformer | 51.9 ± 8.3 | 61.6 ± 8.5 | 38.5 ± 5.9 | 45.5 ± 3.9 |
| Transferring features protocol | | | | |
| Point-MAE | 63.9 ± 7.0 | 77.0 ± 5.2 | 53.6 ± 5.4 | 61.6 ± 2.7 |
| Mask POINT (Ours) | 65.3 ± 4.9 | 77.4 ± 5.2 | 53.8 ± 5.3 | 55.4 ± 2.1 |
| Linear classification protocol | | | | |
| Point-MAE | 48.3 ± 7.8 | 56.0 ± 6.2 | 39.2 ± 10.1 | 42.1 ± 1.7 |
| Mask POINT(Ours) | 51.0 ± 8.2 | 59.8 ± 7.9 | 41.7 ± 9.2 | 61.6 ± 2.7 |
| Non-linear classification protocol | | | | |
| Point-MAE | 56.4 ± 6.8 | 67.2 ± 6.5 | 44.3 ± 6.2 | 50.8 ± 3.6 |
| Mask POINT(Ours) | 60.8 ± 6.6 | 68.3 ± 6.7 | 46.6 ± 6.4 | 61.6 ± 2.7 |

### 4.3. Analyses and Discussions
**Analyses and Discussions**

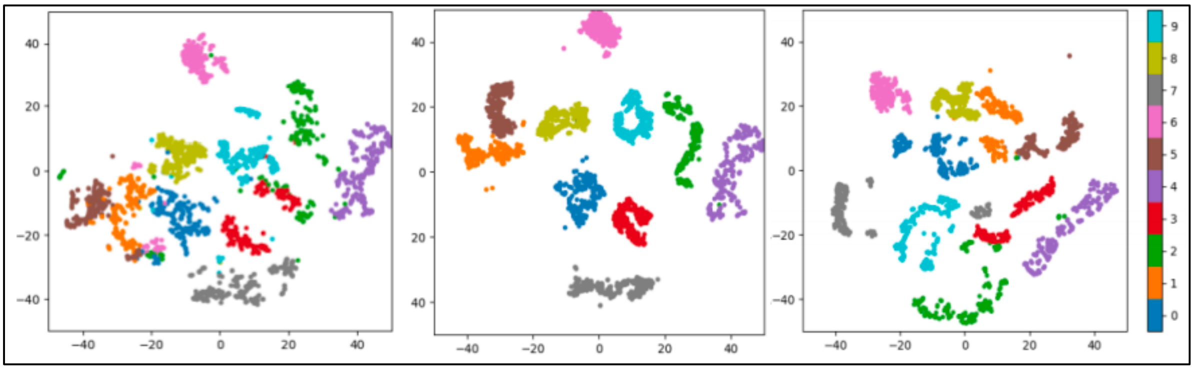


**Fig 4:** Point cloud clustering graph

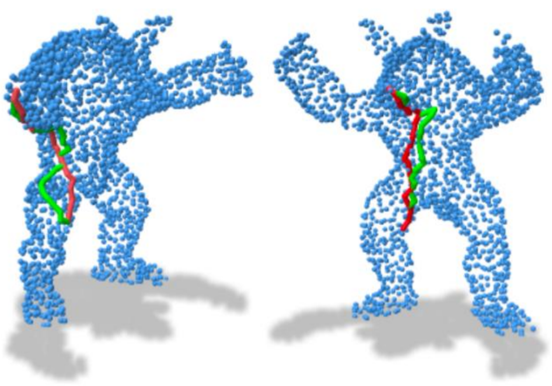

**Fig 5:** Point cloud result graph

As an attempt to evaluate the performance of the proposed method in extreme cases where the explicit 3D shape topology is absent, we attempted to estimate shortest paths on unstructured point clouds.

## References
1. Yang R. CaseGPT: a case reasoning framework based on language models and retrieval-augmented generation. arXiv preprint arXiv:2407.07913; c2024.
2. Chen M. Investigating the influence of interannual precipitation variability on terrestrial ecosystem productivity [Doctoral dissertation]. Massachusetts Institute of Technology; c2023.
3. Chen M, Chen Y, Zhang Q. A review of energy consumption in the acquisition of bio-feedstock for microalgae biofuel production. Sustainability. 2021;13(16):8873.
4. Zhang X, Soe AN, Dong S, Chen M, Wu M, Htwe T. Urban resilience through green roofing: a literature review on dual environmental benefits. In: E3S Web of Conferences. EDP Sciences. 2024;536:01023.
5. Chen M, Chen Y, Zhang Q. Assessing global carbon sequestration and bioenergy potential from microalgae cultivation on marginal lands leveraging machine learning. Science of The Total Environment. 2024;948:174462.
6. Chen M. Annual precipitation forecast of Guangzhou based on genetic algorithm and backpropagation neural network (GA-BP). In: International Conference on Algorithms, High Performance Computing, and Artificial Intelligence (AHPCAI 2021). SPIE. 2021;12156:182-186.
7. Dong S, Xu T, Chen M. Solar radiation characteristics in Shanghai. In: Journal of Physics: Conference Series. IOP Publishing. 2022;2351(1):012016.
8. Gao H, Li Y, Long K, Yang M, Shen Y. A survey for foundation models in autonomous driving. arXiv preprint arXiv:2402.01105; c2024.
9. Yao J, et al. Ndc-scene: boost monocular 3D semantic scene completion in normalized device coordinates space. In: 2023 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE Computer Society; c2023.
10. Yao J, et al. QE-BEV: query evolution for bird's eye view object detection in varied contexts. ACM Multimedia; c2024.
11. Pan X, et al. Harmonic NeRF: geometry-informed synthetic view augmentation for 3D scene reconstruction in driving scenarios. ACM Multimedia; c2024.
12. Liu Y, Yang H, Wu C. Unveiling patterns: a study on semi-supervised classification of strip surface defects. IEEE Access. 2023;11:119933-46.
13. Restrepo D, Wu C, Vásquez-Venegas C, Nakayama LF, Celi LA, López DM. DF-DM: a foundational process model for multimodal data fusion in the artificial intelligence era. Research Square; c2024.
14. Moukheiber D, Restrepo D, Cajas SA, Montoya MPA, Celi LA, Kuo KT, et al. A multimodal framework for extraction and fusion of satellite images and public health data. Scientific Data. 2024;11(1):634.
15. Pan Q, Cao H, Zhu Y, Liu J, Li B. Contextual client selection for efficient federated learning over edge devices. IEEE Transactions on Mobile Computing; c2023.
16. Cao H, Pan Q, Zhu Y, Liu J. Birds of a feather help: context-aware client selection for federated learning. In: Proc. Int. Workshop Trustable Verifiable Auditable Federated Learning Conjunction AAAI; 2022:1-8.
17. Shimizu S, et al. Boron neutron capture therapy for recurrent glioblastoma multiforme: imaging evaluation of a case with long-term local control and survival. Cureus. 2023;15(1).
18. Niitsu H, et al. Tumor response on diagnostic imaging after proton beam therapy for hepatocellular carcinoma. Cancers. 2024;16(2):357.

19. Wang Y, Ban X, Wang H, Li X, Wang Z, Wu D, et al. Particle filter vehicles tracking by fusing multiple features. IEEE Access. 2019;7:133694-706.
20. Ma B, Ma B, Gao M, Wang Z, Ban X, Huang H, Wu W. Deep learning-based automatic inpainting for material microscopic images. Journal of Microscopy. 2021;281(3):177-89.
21. Jiang L, Yu C, Wu Z, Wang Y. Advanced AI framework for enhanced detection and assessment of abdominal trauma: integrating 3D segmentation with 2D CNN and RNN models. arXiv preprint arXiv:2407.16165. 2024.
22. Wu X, Wu Y, Li X, Ye Z, Gu X, Wu Z, et al. Application of adaptive machine learning systems in heterogeneous data environments. Global Academic Frontiers. 2024;2(3):37-50.
23. Zhu Z, Wang Z, Wu Z, Zhang Y, Bo S. Adversarial for sequential recommendation walking in the multi-latent space. Applied Science and Biotechnology Journal for Advanced Research. 2024;3(4):1-9.
24. Jiang L, Yu C, Wu Z, Wang Y. Advanced AI framework for enhanced detection and assessment of abdominal trauma: integrating 3D segmentation with 2D CNN and RNN models. arXiv preprint arXiv:2407.16165; c2024.
25. Yan H, Wang Z, Xu Z, Wang Z, Wu Z, Lyu R. Research on image super-resolution reconstruction mechanism based on convolutional neural network. arXiv preprint arXiv:2407.13211; c2024.
26. Wang L, Cheng Y, Gu X, Wu Z. Design and optimization of big data and machine learning-based risk monitoring system in financial markets. arXiv preprint arXiv:2407.19352; c2024.
27. Wang Z, et al. Research on autonomous robots navigation based on reinforcement learning. arXiv preprint arXiv:2407.02539; c2024.
28. Yan H, et al. Research on image super-resolution reconstruction mechanism based on convolutional neural network. arXiv preprint arXiv:2407.13211; c2024.
29. Li W, et al. An intelligent electronic lock for remote-control system based on the internet of things. Journal of Physics: Conference Series. IOP Publishing; 2018:1069(1).
30. Xu Q, et al. Applications of explainable AI in natural language processing. Global Academic Frontiers. 2024;2(3):51-64.
31. Zhong Y, et al. Deep learning solutions for pneumonia detection: performance comparison of custom and transfer learning models. medRxiv. 2024;2024-06.
32. Zhu A, et al. Exploiting diffusion prior for out-of-distribution detection. arXiv preprint arXiv:2406.11105; c2024.
33. Li K, et al. Exploring the impact of quantum computing on machine learning performance; c2024.
34. Gu W, et al. Predicting stock prices with FinBERT-LSTM: integrating news sentiment analysis. arXiv preprint arXiv:2407.16150; c2024.
35. Tao Y. SQBA: sequential query-based blackbox attack. In: Proceedings of the Fifth International Conference on Artificial Intelligence and Computer Science (AICS 2023). SPIE; 2023:721-729.
36. Tao Y. Meta-learning enabled adversarial defense. In: 2023 IEEE International Conference on Sensors, Electronics and Computer Engineering (ICSECE). IEEE; 2023:1326-1330.
37. Lin Z, et al. Text sentiment detection and classification based on integrated learning algorithm. Applied Science and Engineering Journal for Advanced Research. 2024;3(3):27-33.
38. Lin Z, et al. Neural radiance fields convert 2D to 3D texture. Applied Science and Biotechnology Journal for Advanced Research. 2024;3(3):40-44.
39. Tan C, et al. Editable neural radiance fields convert 2D to 3D furniture texture. International Journal of Engineering and Management Research. 2024;14(3):62-65.
40. Dang B, Zhao W, Li Y, Ma D, Yu Q, Zhu EY. Real-time pill identification for the visually impaired using deep learning. arXiv preprint arXiv:2405.05983; c2024.
41. Dang B, Ma D, Li S, Qi Z, Zhu E. Deep learning-based snore sound analysis for the detection of night-time breathing disorders. Applied and Computational Engineering. 2024;76:109-114. doi:10.54254/2755-2721/76/20240574.
42. Li S, Dong X, Ma D, Dang B, Zang H, Gong Y. Utilizing the LightGBM algorithm for operator user credit assessment research. Applied and Computational Engineering. 2024;75(1):36-47. doi:10.54254/2755-2721/75/20240503.
43. Xiao M, Bo S, Wu Z. Multiple greedy quasi-Newton methods for saddle point problems. arXiv preprint arXiv:2408.00241; c2024.
44. Bo S, Xiao M. Dynamic risk measurement by EVT based on stochastic volatility models via MCMC. arXiv preprint arXiv:2201.09434; c2022.
45. Wang R, Behandish M. Surrogate modeling for physical systems with preserved properties and adjustable tradeoffs. arXiv preprint arXiv:2202.01139; c2022.
46. Wang R, Shapiro V. Topological semantics for lumped parameter systems modeling. Advanced Engineering Informatics. 2019;42:100958.
47. Wang R, Behandish M. Surrogate modeling for physical systems with preserved properties and adjustable tradeoffs. arXiv preprint arXiv:2202.01139; c2022.
48. Liu M, Wang Q, Lu W. Peridynamic simulation of brittle-ice crushed by a vertical structure. International Journal of Naval Architecture and Ocean Engineering. 2017;9(2):209-218.
49. Liu M, Oswald J. Coarse-grained molecular modeling of the microphase structure of polyurea elastomer. Polymer. 2019;176:1-10.
50. Zan Y, et al. Research on real-time simulation system of ship motion based on Simulink. The Open Mechanical Engineering Journal; 2014:8(1).
51. Cui X, et al. A hybrid wavelet-based adaptive immersed boundary finite-difference lattice Boltzmann method for two-dimensional fluid–structure interaction. Journal of Computational Physics. 2017;333:24-48.
52. Cui X, et al. A coupled volume penalization-thermal lattice Boltzmann method for thermal flows. International Journal of Heat and Mass Transfer. 2018;127:253-266.
53. Cui X, et al. A coupled two-relaxation-time lattice Boltzmann-volume penalization method for flows past obstacles. Mathematics and Computers in Simulation. 2022;198:85-105.
54. Guo K, Cui X, Liu M. A coupled lattice Boltzmann-volume penalization for flows past fixed solid obstacles with local mesh refinement. Mathematical Problems in Engineering. 2018;2018(1):6732082.

55. Liu M, Ye J, Oswald J. Coarse-grained molecular simulation of the role of curing rates on the structure and strength of polyurea. Computational Materials Science. 2023;230:112428.
56. Bo, Shi, and Minheng Xiao. "Dynamic Risk Measurement by EVT based on Stochastic Volatility models via MCMC." arXiv preprint arXiv:2201.09434 (2022).
57. Zhu L, et al. FEGAN: a feature-oriented enhanced GAN for enhancing thermal image super-resolution. IEEE Signal Processing Letters; c2024.
58. Qin H. Revolutionizing cryptocurrency operations: the role of domain-specific large language models (LLMs). International Journal of Computer Trends and Technology. 2024;72(6):101-113.
59. Qin H, Li Z. A study on enhancing government efficiency and public trust: the transformative role of artificial intelligence and large language models. International Journal of Engineering and Management Research. 2024;14(3):57-61. doi:10.5281/zenodo.12619360.
60. Qin H, Li Z. Precision in practice: enhancing healthcare with domain-specific language models. Applied Science and Engineering Journal for Advanced Research. 2024;3(4):28-33. doi:10.5281/zenodo.13253336.
61. Zhu A, et al. Cross-task multi-branch vision transformer for facial expression and mask wearing classification. Journal of Computer Technology and Applied Mathematics. 2024;1(1):46-53.
62. Li K, et al. Utilizing deep learning to optimize software development processes. Journal of Computer Technology and Applied Mathematics. 2024;1(1):70-76.
63. Li K, et al. The application of augmented reality (AR) in remote work and education. Journal of Computer Technology and Applied Mathematics. 2024;1(1):33-39.
64. Hong B, et al. The application of artificial intelligence technology in assembly techniques within the industrial sector. Journal of Artificial Intelligence General Science (JAIGS). 2024;5(1):1-12.
65. Xiao, Minheng, and Shi Bo. "Electroencephalogram Emotion Recognition via AUC Maximization." arXiv preprint arXiv:2408.08979 (2024).
66. Dai S, et al. AI-based NLP section discusses the application and effect of bag-of-words models and TF-IDF in NLP tasks. Journal of Artificial Intelligence General Science (JAIGS). 2024;5(1):13-21.
67. Zhao P, et al. Task allocation planning based on hierarchical task network for national economic mobilization. Journal of Artificial Intelligence General Science (JAIGS). 2024;5(1):22-31.
68. Li Z, et al. Stock market analysis and prediction using LSTM: a case study on technology stocks. Innovations in Applied Engineering and Technology; 2023:1-6.
69. Mo Y, et al. Large language model (LLM) AI text generation detection based on transformer deep learning algorithm. International Journal of Engineering and Management Research. 2024;14(2):154-159.
70. Li S, Mo Y, Li Z. Automated pneumonia detection in chest x-ray images using deep learning model. Innovations in Applied Engineering and Technology; 2022:1-6.
71. Mo Y, et al. Password complexity prediction based on Roberta algorithm. Applied Science and Engineering Journal for Advanced Research. 2024;3(3):1-5.
72. Song J, et al. A comprehensive evaluation and comparison of enhanced learning methods. Academic Journal of Science and Technology. 2024;10(3):167-171.
73. Liu T, et al. Spam detection and classification based on DistilBERT deep learning algorithm. Applied Science and Engineering Journal for Advanced Research. 2024;3(3):6-10.
74. Dai S, et al. The cloud-based design of unmanned constant temperature food delivery trolley in the context of artificial intelligence. Journal of Computer Technology and Applied Mathematics. 2024;1(1):6-12.
75. Mo Y, et al. Make Scale Invariant Feature Transform “fly” with CUDA. International Journal of Engineering and Management Research. 2024;14(3):38-45.
76. He S, et al. Lidar and monocular sensor fusion depth estimation. Applied Science and Engineering Journal for Advanced Research. 2024;3(3):20-26.
77. Liu J, et al. Unraveling large language models: from evolution to ethical implications—introduction to large language models. World Scientific Research Journal. 2024;10(5):97-102.
78. Mo Y, Zhang Y, Li H, Wang H, Yan X. Prediction of heart failure patients based on multiple machine learning algorithms. Applied and Computational Engineering. 2024;75:1-7. doi:10.54254/2755-2721/75/20240498.
79. Liu, Minghao, Qing Wang, and Wei Lu. “Peridynamic simulation of brittle-ice crushed by a vertical structure.” International Journal of Naval Architecture and Ocean Engineering 9.2 (2017): 209-218.
80. Xiao, Minheng, Shi Bo, and Zhizhong Wu. "Multiple greedy quasi-newton methods for saddle point problems." arXiv preprint arXiv:2408.00241 (2024).
81. Bo, Shi. "Application of K-means clustering algorithm in evaluation and statistical analysis of internet financial transaction data." arXiv preprint arXiv:2202.03146 (2022).
82. Bo, Shi, and Minheng Xiao. "Root Cause Attribution of Delivery Risks via Causal Discovery with Reinforcement Learning." arXiv preprint arXiv:2408.05860 (2024).