Dual Manifold Regularization Steered Robust Representation Learning for Point Cloud Analysis

Jian Bi, Qianliang Wu, Jianjun Qian, Lei Luo*, Jian Yang*

PCA Lab, Key Lab of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education
School of Computer Science and Engineering, Nanjing University of Science and Technology, China
{jianbi, wuqianliang, csjqian, cslluo, csjyang}@njust.edu.cn

Abstract

With the rapid advancement of 3D scanning technology, point clouds have become a crucial data type in computer vision and machine learning. However, learning robust representations for point clouds remains a significant challenge due to their irregularity and sparsity. In this paper, we propose a novel Dual Manifold Regularization (DMR) framework that makes full use of the properties of positive and negative curvature in manifolds to improve the representation of point clouds. Specifically, we leverage DMR based on hyperbolic and hyperspherical manifolds to address the limitations of traditional single-manifold regularization techniques, including inadequate generalization ability and adaptability to data diversity, as well as the difficulty of capturing complex relationships between data. To begin, we utilize the tree-like structure of the hyperbolic manifold to model the part-whole hierarchical relationships within point clouds. This allows for a more comprehensive representation of the data, improving the model's capability to understand complex shapes. Additionally, we construct positive samples through topological consistency augmentation and employ contrastive learning techniques in the hyperspherical manifold to capture more discriminative features within the data. Our experimental results show that our method outperforms traditional supervised learning and single-manifold regularization techniques in point cloud analysis. Specifically, for shape classification, DMR achieves a new State-Of-The-Art (SOTA) performance with 94.8% Overall Accuracy (OA) on ModelNet40 and 90.7% OA on ScanObjectNN, surpassing the recent SOTA model without increasing the baseline parameters.

Introduction

Currently, research in point cloud analysis primarily focuses on model innovation which explores various ways to improve performance. For instance, increasing network depth, employing more complex multi-layer perceptron (MLP) structures (Ma et al. 2022; Hu et al. 2023a; Zhang et al. 2023; Lu et al. 2023; Yao et al. 2023; Wu et al. 2025, 2023), or incorporating convolutional networks (Wang et al. 2019; Li et al. 2018; Xu et al. 2018; Lin, Huang, and Wang 2020, 2021; Wu, Qi, and Fuxin 2019; Thomas et al. 2019, 2024; Chen et al. 2023) and transformer models (Zhao et al. 2021; Guo et al. 2021) to enhance the model's representation capability. However, unlike images (Hu et al. 2023b), despite these innovations significantly enhancing the model's processing ability, they often fail to consider the intrinsic hierarchical structure and overall topological relationships within point clouds. Understanding these aspects is beneficial for grasping the spatial distribution and geometric properties of 3D point clouds. It is well-known that regularization not only encourages the model to focus more on the intrinsic structure and features in the data but also helps prevent model overfitting. Therefore, it becomes crucial to incorporate regularization methods (Montanaro, Valsesia, and Magli 2022) to address the above-mentioned shortcomings.

HyCoRe (Montanaro, Valsesia, and Magli 2022) proposes a hyperbolic regularization to attempt to improve the model's representation capabilities for point cloud. However, there are some critical theoretical issues in this method that limit the model's performance. These issues mainly include (1) an incorrect assumption of a common ancestor

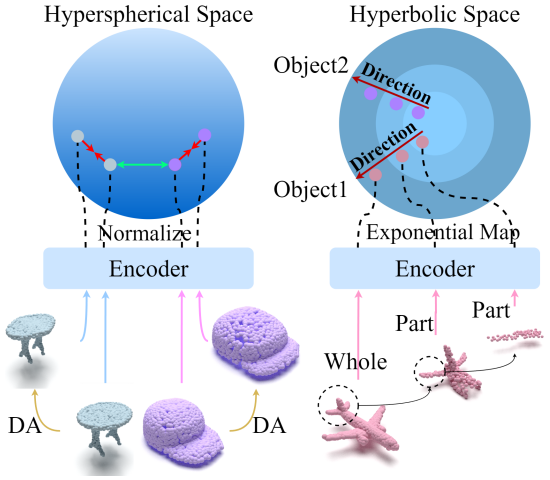

Figure 1: Illustration of regularization on the hyperspherical and hyperbolic manifolds. On the left, red arrows denote positive pairs tending to attract each other, while the green arrow denotes negative pairs tending to repel each other. On the right, parts of different sizes are located at various levels, and parts from distinct objects have different directions.

*corresponding authors

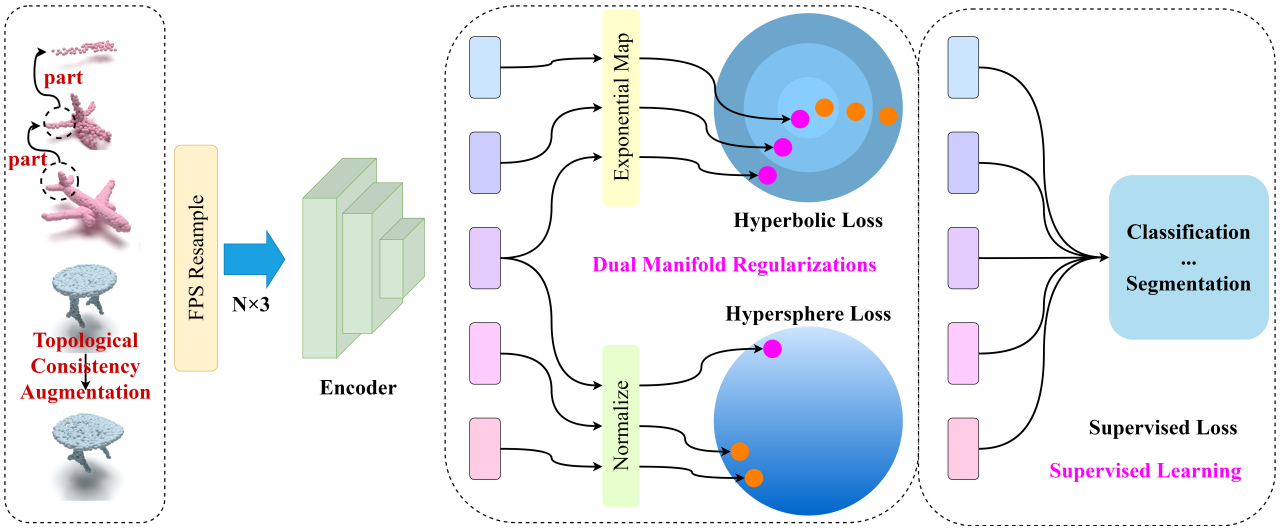

Figure 2: The overall framework of DMR begins with farthest point sampling for resampling, ensuring that the number of points in the object and parts remains consistent. After enhancing the samples through TCA, all the whole and partial inputs are fed into the encoder. The DMR is then applied to improve the model's overall performance.

among different point clouds, which leads to insufficient loss performance based on this assumption; (2) The absence of a common ancestor makes it challenging for the model to leverage hyperbolic distance-based triplet loss to learn discriminative features between categories; and (3) using the inverse of the number of point clouds as a distance threshold in hyperbolic space is inappropriate due to the exponential nature of distance relationships in hyperbolic space.

To address the aforementioned issues, we propose a novel hyperbolic regularization constraint. We discard the common ancestor assumption and instead focus solely on hierarchical relationships within the point cloud. Specifically, we consider the entire point cloud and its parts as positive pairs, utilizing the hyperbolic manifold to represent their hierarchical relationships. In this representation, the whole is located near the edges while the parts are closer to the center. Additionally, we treat different parts as negative pairs, ensuring that they are directionally distant from each other, as illustrated on the left side of Figure 1. This method allows us to effectively capture the subtle differences in the intrinsic hierarchical structure of different point clouds.

However, the absence of a common ancestor among different point clouds in hyperbolic space results in similar hyperbolic distances between different whole point clouds. This poses a challenge in distinguishing between categories. To tackle this problem, we propose a novel method: Topological Consistency Augmentation (TCA). TCA focuses on learning sample pairs that preserve topological consistency and geometric diversity, which facilitates capturing subtle class differences in the hyperspherical manifold.

We ultimately design a dual manifold regularization, as shown in Figure 1. In hyperbolic space, we can learn a more robust point cloud representation for capturing intrinsic hierarchical structure; while in hypersphere space, the model can learn discriminative geometric features for categories classification. The contributions are summarized as follows:

- We derive novel hierarchical loss and contrastive loss to jointly learn robust representations of point clouds within a hyperbolic manifold.
- We propose an innovative topological consistency augmentation for point clouds to generate positive samples that are both topologically consistent and geometrically diverse, thereby learning more discriminative representations between classes in the hyperspherical manifold.
- We are the first to propose collaborative learning of point cloud representations across manifolds with different curvatures. Extensive experiments to demonstrate the superiority of our proposed method over all baselines.

Preliminaries

Hyperbolic Geometry

A hyperbolic space (H_m) is a complete, connected Riemannian manifold with constant negative sectional curvature. These special manifolds are all isometric to each other with isometries defined as $O^+(m,1)$. Among these isometries, there are five common models that previous studies often work on (Cannon et al. 1997), where the Poincaré Ball in n dimensions D_c^n is a hyperbolic space with $c = -1$. Since the Poincaré ball can maintain numerical stability during gradient-based learning, we choose the Poincaré Ball (D_c^n, g_c^D) as our basic model (Ganea, Bécigneul, and Hofmann 2018; Nickel and Kiela 2017; Tifrea, Bécigneul, and Ganea 2018), where n is the dimension size, c denotes a constant negative curvature, and g^D represents the Riemannian metric. The manifold D_c^n is formulated as follows:

$$D_c^n = \{x \in R^n : c\|x\|^2 < 1, c \geq 0\}. \quad (1)$$

We assume $c > 0$, such that D_c^n corresponds to a ball of radius $r = 1/\sqrt{c}$ in Euclidean space. The geodesic distance

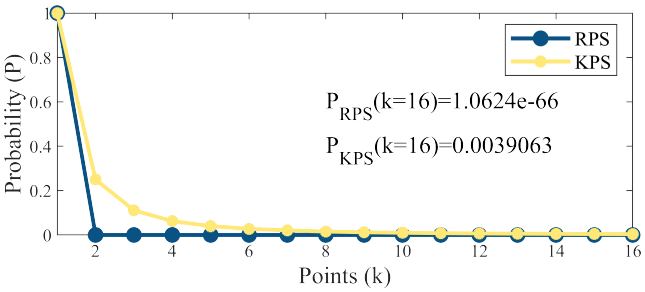

Figure 3: Probability of common ancestor of different objects under two sampling algorithms.

between two points $p, q \in D^n$ is specified as:

$$d_{D_c^n}(p,q) = cosh^{-1}\left(1 + 2\frac{\|p-q\|^2}{(1-\|p\|^2)(1-\|q\|^2)}\right). \quad (2)$$

With the origin $q = 0$, Eq. (2) could be simplified to:

$$d_{D_c^n}(0,p) = 2tanh^{-1}(\|p\|). \quad (3)$$

Unlike Euclidean geometry, this particular space has the property of inferring the underlying hierarchical structure of a dataset in its embedding (Atigh et al. 2022; Li et al. 2022).

To facilitate the transformation of features from Euclidean to hyperbolic spaces, a differentiable bijective operator, exponential map, is employed. We add the hyperbolic network layer at the end of the original deep learning model to map the input features from R^n to the hyperbolic manifold D_c^n via the exponential map, which is given by:

$$z = exp_c(x) := tanh(\sqrt{c}\|x\|)\frac{x}{\sqrt{c}\|x\|}. \quad (4)$$

In this module, the exponential map gives us a way to map the output of a network, which is in the Euclidean space, to the Poincaré Ball. One useful intuition (Sala et al. 2018) to help understand the advantage of the hyperbolic space is given two points $p, q \in D_c^n$ $s.t.$ $\|p\| = \|q\|$ and $p \neq q$,

$$d_{D_c^n}(p,q) \to d_{D_c^n}(p,o) + d_{D_c^n}(o,q), \quad (5)$$

$$as \quad \|p\| = \|q\| \to r. \quad (6)$$

This property basically reflects the fact that the shortest path in a tree is the path through the earliest common ancestor, and it is reproduced in the Poincaré when the points are both close to the boundary.

Hyperspherical Geometry

Unlike hyperbolic manifold, the hypersphere (S^{n-1}) is a sphere in n-dimensional Euclidean space with positive curvature and consists of all points that are at a distance equal to a constant r (i.e., radius) from a fixed point (i.e., the center of the sphere), where $n-1$ denotes the dimension of the hypersphere.

Hyperspherical geometry provides an efficient way to capture the intrinsic structure and relationships of the data in a high-dimensional space. We map the input features x from R^n to the hyperspherical manifold S^{n-1} through normalization, which is given by:

$$z = Normalize(x) = \frac{x}{\|x\|}. \quad (7)$$

Cosine similarity evaluates the similarity of two vectors by measuring the cosine of the angle between them. In hyperspherical geometry, this measure is very applicable because the length of a vector may change with increasing dimensionality, but the angle remains the same. The similarity between any p, q on the unit hypersphere is as follows:

$$d_{cosine\ similarity}(p,q) = cos(\theta) = \frac{p \cdot q}{\|p\|\|q\|}, \quad (8)$$

where d is the cosine similarity between two points, p and q are two unit vectors, θ is the angle between p, q, and $p \cdot q$ is dot product.

Method

In this section, we will first introduce our motivation analysis and the advantages of our dual manifold regularization. Next, we will provide a detailed explanation of our regularization methods for point clouds in different curvature manifolds, including hyperbolic and hypersphere constraints. Our ultimate goal is to improve the model's ability to learn the geometric representation of point clouds through the implementation of regularization constraints. The framework of the DMR method is illustrated in Figure 2. **The proof of all theories can be found in the supplementary materials.**

Motivation Analysis

We conducted an examination of the hyperbolic regularization for point clouds proposed in HyCoRe (Montanaro, Valsesia, and Magli 2022). We discovered several inconsistencies in theory and practice. In particular, our result in **Theorem 1** suggests that in hyperbolic space, point clouds are intrinsically related hierarchically, with a nearly zero probability of different point clouds sharing a common ancestor. Therefore, using shared ancestry as the basis for deriving a loss function for network training is inappropriate.

Additionally, the description of the loss for hyperbolic hierarchy in HyCoRe is inaccurate. In Eq. (9), using only the reciprocal of the number of point clouds as the threshold between parts and the whole is theoretically incorrect, as distances between different levels in hyperbolic space follow an exponential relationship.

$$Rhier(z_w, z_p) = max(0, -\|z_w\|_h + \|z_p\|_h + \frac{\gamma}{N_p}), \quad (9)$$

where z_w and z_p are the hyperbolic representation of the whole and a part from the same point cloud. γ is a hyperparameter and N_p is the number of points in the part.

Furthermore, we found that while the hierarchical properties of hyperbolic space offer better representations, it struggles to learn discriminative information between categories. Specifically, a distance-based triplet (Weinberger and Saul 2009) loss fails to separate point clouds of different categories. This is due to the absence of a common ancestor,

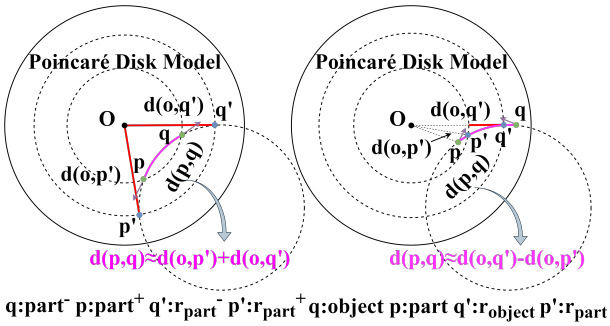

Figure 4: Hierarchical loss and contrastive loss.

causing distances between points in hyperbolic space to depend on their paths to the central point, which can result in two visually similar points having a maximum distance (**The visualization is in the supplementary material**).

To tackle these challenges, we propose a novel Dual Manifold Regularization (DMR). By utilizing the hyperbolic manifold, DMR can uncover the inherent hierarchical relationships within point clouds, from parts to the whole. Additionally, our method leverages the hyperspherical manifold to improve the discriminative abilities of our representations across different categories.

Whole-Part Hierarchy in 3D Point Clouds

There have been studies exploring hierarchical relationships between whole and parts as early as in the field of language (Nickel and Kiela 2017), images (Ge et al. 2023) and point cloud (Montanaro, Valsesia, and Magli 2022) and the use of hierarchical relationships to understand point cloud is not new. This gives inspiration to the understanding of the point cloud to learn this hierarchical relationship by dividing the point cloud into parts of different sizes. HyCoRe (Montanaro, Valsesia, and Magli 2022) treats widgets consisting of small parts (e.g., simple structures like disks, squares, triangles) as universal ancestors of more complex shapes included in many different objects. These structures gradually become more object or class-specific as they are combined into more complex parts. However, probabilistic analysis shows that the probability of two different objects sharing an ancestor is close to 0. Therefore, we posit that this intuitive hierarchical relationship exists only within the same object. We provide the related theorem and assume that the number of common ancestor parts between two objects is N/K, where N is the number of points in the object, k is the number of points in the part, $C(N,k)$ represents the number of combinations of k points sampled from N points, and $C(N,1)$ denotes sampling samples from N combinations containing k points one of which contains k points. The theorem is as follows:

Theorem 1. Let the two object point clouds contain N points, part point cloud (common ancestor, random point sampled from object) contains k points, $0 < k < N$. Then the maximum probability that a part sampled from an object is a common ancestor of the two objects by random point sampling is $P_{RPS} = \frac{(N/k)^2}{C(N,k)^2}$.

Theorem 2. Let the two object point clouds contain N points, part point cloud (common ancestor, k nearest neighbor point sampled from object) contains k points, $0 < k < N$. Then the maximum probability that a part sampled from an object is a common ancestor of the two objects by k nearest neighbor point sampling is $P_{KPS} = \frac{(N/k)^2}{C(N,1)^2}$.

We also present the probability of common ancestry for both sampling cases, as shown in Figure 3, where the P_{RPS} is $1.0624e^{-66}$ and the P_{KPS} probability is 0.0039 for $k = 16$. This means that the probability of different objects occurring during training is almost 0. Consequently, in the hyperbolic manifold, intrinsic hierarchical relationships are generally meaningful only within the same whole-part structure rather than across different objects.

After the above analysis, we propose to investigate a part-whole hierarchy within an object. Formally, given a whole point cloud $P = \{p_1, p_2, ..., p_N\}$. We sample a part by setting $P_i = \{p_1, p_2, ..., p_i\}$, where $0 < i < N$ is the subset of P. In a hyperbolic manifold, when i is smaller, the embedded part is closer to the center.

Hierarchical Contrastive Learning

In the Poincaré disk model, the projected position of a point cloud is determined by the number of points it contains. Generally, point clouds with a higher number of points are situated closer to the boundary of the disk, while those with fewer points are closer to the center.

Suppose we have an object P containing N points, and its part P_i containing M points, where $M < N$. According to the Poincaré disk model, the hierarchy of a point cloud can be represented by its position in hyperbolic space. For point clouds P and P_i, we can use the following formula to describe their position in a hyperbolic space with radius $r = 1$. Location of the object point cloud P:

$$r_P = d_{D_c^n}(0, N/(N+1)). \tag{10}$$

Location of the part point cloud P_i:

$$r_{P_i} = d_{D_c^n}(0, M/(N+1)). \tag{11}$$

In this case, r_P is close to the boundary and r_{P_i} is close to the center. $N+1$ guarantees that $r < 1$.

Therefore, projecting the object and its part embeddings learned through the encoder network and Exponential mapping f_h onto the Poincaré disk requires adhering to the radius constraint. This measure is crucial for preserving the hierarchy in hyperbolic space. We propose the theorem of **Distance-Preserving** to formalize this distance constraint.

Theorem 3 (Distance-Preserving) Let the object point cloud P contain N points, part point cloud P_i contains M points, and $P_i \subseteq P$. Then, $d_{D_c^n}(f_h(P), f_h(P_i)) = d_{D_c^n}(0, N/N+1) - d_{D_c^n}(0, M/N+1)$.

Therefore, we can derive the hierarchical relationship between the part and the object and use it to construct the hierarchical loss.

$$L_h(P, P_i) = |d_{D_c^n}(f_h(P), f_h(P_i)) - (r_P - r_{P_i})|, \tag{12}$$

where $0 < i < N$. Figure 4 (right) illustrates an approximate representation of L_h. This diagram visually demonstrates the hierarchical structure and spatial relationships between whole-part in the Poincaré disk model.

To increase the discrimination of different objects, we need to position them in distinct directions. Therefore, we propose a new contrastive loss to enhance the model's representation performance as follows:

$$L_c(P_i^+, P_i^-) = |d_{D_c^n}(f_h(P_i^+), f_h(P_i^-)) - (r_{P_i^+} + r_{P_i^-})|, \quad (13)$$

where $0 < i < N$. An approximate diagram of L_c is shown in Figure 4 (left). P_i^+, P_i^- sample from different objects.

Since only one part of the object is sampled for each training, ultimately, the complete regularization loss in hyperbolic space can be written as:

$$HLoss = L_h + L_c. \quad (14)$$

The hyperbolic manifold aids in exploring the hierarchical relationships within the point cloud, but has limitations in extracting discriminative features between classes. To address this issue, we combine contrastive learning (Chen et al. 2021) on a hyperspherical manifold to capture unique information across categories. Recognizing that existing augmentation methods can disrupt topological consistency, we introduce a novel Topological Consistency Augmentation (TCA) method. TCA generate positive samples maintain topological consistent topological and geometric diversity, allowing the model to learn more robust and discriminative point cloud representations.

Topological Consistency Augmentation

TCA aims to preserve the topological structure of the original point cloud by homeomorphic mapping and perturb the local structure using a Sine function to simulate the distortion and deformation of objects in the real world, thereby expanding the geometric diversity of point clouds. Figure 5 illustrates how TCA augments the point clouds.

Homeomorphic mapping (Derrick 1973) is an important mathematical tool used to describe the equivalence relations between topological spaces. It keeps the proximity of points in space unchanged and makes topological Spaces have the same topological properties. These properties make homeomorphic mapping widely used in topology, geometry, physics and other fields.

Definition 1. (Homeomorphic Mapping) Given two topological spaces X, Y, and given a mapping $f : X \to Y$. f is a homeomorphic mapping of two spaces when it is satisfied that f is a bijection and f and f^{-1} are continuous, denoted as $X \cong Y$.

Definition 2. (Local Homeomorphic Mapping) Let $f : X \to Y$ is a mapping between two topological spaces X and Y. If for every point x in X, exists a neighborhood U of x such that $f(U)$ is an open set in X and $f_U : U \to f(U)$ is a homeomorphic mapping, then f is a local homeomorphic mapping.

Homeomorphic and Local Homeomorphic Mapping have similar mathematical properties, both of which require the mapping and its inverse mapping to be continuous and maintain the topological structure of the space(Armstrong 2013).

Proposition 1. (Topological consistency) If f is a homeomorphic mapping from X to Y, then X and Y have the same topological properties.

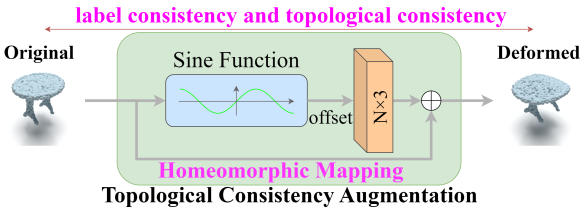

Figure 5: An overview of our TCA. We use TCA to get the augmented point cloud. TCA can be adapted for a variety of tasks due to label consistency and topological consistency.

Proposition 2. (Reflexivity, symmetry, and transitivity) The homeomorphism relation is an equivalence relation, and therefore it has reflexivity (any topological space is homeomorphism to itself), symmetry (if $X \cong Y$, then $Y \cong X$), and transitivity (if $Y \cong Z$, then $X \cong Z$).

We have chosen to use the Sine function as our mapping method to more accurately simulate the distortion and deformation of an object's surface in the real world. The Sine function's inherent periodic nature allows us to adjust the number of regions that are deformed.

Theorem 4. (Homeomorphic Mapping Based on Sine Function) Given two topological spaces X, Y, and given a mapping $f : X \to Y = X + Asin(\omega X + \varphi)$, if $-1 \leq A\omega \leq 1$, then f is a homeomorphic mapping, else f is a local homeomorphic mapping.

Since f is a homeomorphic mapping, we can use it to augment the point cloud and ensure topological consistency. Given a set of points $P = \{p_1, p_2, ..., p_N\}$, where N represents the number of points in the Euclidean space (x, y, z). SinPoint applies homeomorphic mapping and the resulting augmented point cloud $P^{'}$ is given as follows:

$$P^{'} = P + Asin(\omega P + \varphi), \quad (15)$$

where $Asin(\omega P + \varphi)$ is displacement field of P. We need to adjust A and ω to produce more diverse point clouds. In this paper, we set $A \sim U(-a, a)$ and $\omega \sim U(-w, w)$ to obey the uniform distribution, which can produce more samples conforming to the laws of the real world and make the distribution of samples more uniform.

Similarity Contrastive Learning

Contrastive learning is to learn effective feature representations by comparing different samples. When implementing contrastive learning between positive and negative sample pairs, we typically aim to minimize the differences between positive pairs (making them more similar) while maximizing the differences between negative pairs (making them more dissimilar). The triplet loss (Weinberger and Saul 2009) encourages the model to learn representations where the distance $d_{S^n}(Q_i, Q_i^{'})$ between an anchor Q_i and its positive sample $Q_i^{'}$ is less than the distance $d_{S^n}(Q_i, Q_j)$ between the anchor Q_i and its negative sample n by a certain margin δ. Mathematically, this can be expressed as:

$$SLoss = max\left(0, d_{S^n}(Q_i, Q_i^{'}) - d_{S^n}(Q_i, Q_j) + \delta\right), \quad (16)$$

where $SLoss$ represents the triplet loss, $d_{S^n}(\cdot,\cdot)$ denotes the cosine similarity, and δ is the pre-set margin, ensuring that the similarity between positive sample pairs is significantly higher than that between negative sample pairs.

We employed topological consistency augmentation method to generate positive samples with topological consistency, which exhibit sufficient geometric diversity. This diversity ensures that the model can effectively capture the geometric information of different category data, thereby optimizing data representation and significantly enhancing the model's generalization ability.

Loss Functions

The DMR is included in the final loss in this way:

$$Loss = L_{supervised} + \alpha SLoss + \beta HLoss, \qquad (17)$$

where $L_{supervised}$ is a supervised loss in various tasks *(e.g., cross-entropy in classfication)*, $SLoss$ and $HLoss$ are derived from dual manifold regularization.

Experiments

We present experimental results of our method for object point cloud analysis, including object classification and shape part segmentation. Our evaluations utilize three datasets: ModelNet40 (Wu et al. 2015), ScanObjectNN (Uy et al. 2019), and ShapeNetPart (Yi et al. 2016) to demonstrate our DMR effectiveness. More additional ablation experiments and code are in the supplementary materials.

Classification on ModelNet40

ModelNet40 (Wu et al. 2015) is a commonly used point cloud classification dataset, which has 40 object categories containing 9843 training models and 2468 test models. We use pointnet++ (Qi et al. 2017b), DGCNN (Wang et al. 2019), and PointMLP (Ma et al. 2022) as the backbone. The class mean accuracy (mAcc) and the overall accuracy (OA) were used to measure the performance of the model.

Comparisons with SOTA models. We report well-known and state-of-the-art supervised models (Qi et al. 2017b; Wu, Qi, and Fuxin 2019; Thomas et al. 2019; Wang et al. 2019; Guo et al. 2021; Qiu, Anwar, and Barnes 2021; Xu et al. 2021; Zhao et al. 2021; Qian et al. 2022; Xiang et al. 2021; Ma et al. 2022; Ran, Liu, and Wang 2022; Yao et al. 2023; Montanaro, Valsesia, and Magli 2022). The experimental results are summarized in Table 1. Our DMR achieves an Overall Accuracy (OA) of 94.6%, surpassing the SOTA model HGNet's 94.5% without a voting strategy. With the voting strategy incorporated, our method achieves a higher OA of 94.8%, demonstrating significant improvement over the baseline network PointMLP by more than 0.7%, without increasing parameters. This establishes our method as achieving new SOTA performance on ModelNet40.

Ablation study of modules. The results shown in Table 2 reveal that hyperspherical regularization alone yields performance comparable to HyCoRe (Montanaro, Valsesia, and Magli 2022). However, when using only hyperbolic regularization, our method significantly outperforms HyCoRe.

Method	mAcc	OA	Training
PointNet++	-	90.7	supervised
PointCNN	88.1	92.5	supervised
PointConv	-	92.5	supervised
KPConv	-	92.9	supervised
DGCNN	90.2	92.9	supervised
PCT	-	93.2	supervised
DeepGCN	90.9	93.6	supervised
GBNet	91.0	93.8	supervised
GDANet	-	93.8	supervised
Point Trans.	90.6	93.7	supervised
PointNeXt-S	90.9	93.7	supervised
CurveNet	-	93.8	supervised
PointMLP	91.3	94.1	supervised
RepSurf-U	91.4	94.4	supervised
HGNet	91.9	94.5	supervised
PointMLP +HyCoRe	91.7	94.3	supervised +regularized
PointMLP +DMR (Ours)	**91.9**	**94.6**	supervised +regularized
PointMLP +DMR* (Ours)	**92.0**	**94.8**	supervised +regularized

Table 1: 3D shape classification on ModelNet40

Method	mAcc	OA
DGCNN	90.2	92.2
DGCNN+HyCoRe	91.0 (+0.8)	93.7 (+1.5)
DGCNN+SLoss (Ours)	**91.8 (+1.6)**	**93.7 (+1.5)**
DGCNN+HLoss (Ours)	**91.1 (+0.9)**	**94.0 (+1.8)**
DGCNN+DMR (Ours)	**91.9 (+1.7)**	**94.3 (+2.1)**

Table 2: Ablation study of modules using DGCNN

Moreover, combining both hyperbolic and hyperspherical regularization further enhances the model's performance. This indicates that in the point cloud representation learning, dual manifold regularization is more effective than using single-manifold regularization.

Comparisons with Various Baselines. We validated the performance of our method across various classic models (Qi et al. 2017b; Wang et al. 2019; Ma et al. 2022) and compared it with different approaches, including the regularization method HyCoRe (Montanaro, Valsesia, and Magli 2022) and the attention-based method GAM (Hu et al. 2023a), as shown in Table 3. The experimental results demonstrate that our method consistently improves performance across different backbone networks without introducing any additional parameters. Specifically, we improved the classification performance of the DGCNN model on ModelNet40 to 94.3%, surpassing PointMLP's 94.1%. Even with the larger parameter PointMLP model, our method still achieved a 0.5% performance boost. This demonstrates that the proposed method leverages novel ideas, complementary to what is exploited by existing architectures, and it is thus able to boost the performance even of state-of-the-art meth-

Method	mAcc	OA	Training
DGCNN	90.2	92.2	supervised
DGCNN+HyCoRe	91.0	93.7	+Regularized
DGCNN+GAM	90.5	93.3	+Module
DGCNN+DMR	91.9	94.3	+Regularized
DGCNN+DMR*	**92.2**	**94.4**	+Regularized
PointNet++	89.1	90.5	supervised
PointNet++ +HyCoRe	90.3	91.1	+Regularized
PointNet++ +GAM	91.5	92.8	+Module
PointNet++ +DMR	91.2	93.8	+Regularized
PointNet++ +DMR*	**91.5**	**94.1**	+Regularized
PointMLP	91.3	94.1	supervised
PointMLP+HyCoRe	91.7	94.3	+Regularized
PointMLP+GAM	91.5	94.2	+Module
PointMLP+DMR	91.9	94.6	+Regularized
PointMLP+DMR*	**92.0**	**94.8**	+Regularized

Table 3: Comparisons with Various Baselines and Training

Method	mAcc	OA	Training
PointNet++	75.4	77.9	supervised
DGCNN	73.6	78.1	supervised
PointCNN	75.1	78.5	supervised
GBNet	77.8	80.5	supervised
MVTN	-	82.8	supervised
RepSurf-U	81.3	84.3	supervised
PointMLP	83.9	85.4	supervised
PointNorm	85.6	86.8	supervised
PointNeXt-S	85.8	87.7	supervised
PointMetaBase-S	86.9	88.2	supervised
SPoTr	86.8	88.6	supervised
PointVector	86.8	88.6	supervised
HGNet	87.5	89.2	supervised
KPNeXt-L	88.1	89.3	supervised
DeLA	89.3	90.4	supervised
X-3D (DeLA)	89.9	**90.7**	supervised
PointMLP +HyCoRe	85.9	87.2	supervised +regularized
PointMLP +DMR (Ours)	**89.0**	**89.5**	supervised +regularized
PointMLP +DMR* (Ours)	**89.8**	**90.2**	supervised +regularized
DeLA +DMR (Ours)	**90.0**	**90.7**	supervised +regularized

Table 4: 3D shape classification on ScanObjectNN

ods. It is also remarkable that an older, yet still popular, architecture like DGCNN and PointNet++ is able to outperform complex and sophisticated models such as the PCT, PointMLP and PointNeXt, when regularized by DMR.

Classification on ScanObjectNN

ScanObjectNN (Uy et al. 2019) is a real-world scan dataset that contains 15,000 truly scanned objects, grouped into 15 classes, with 2,902 unique object instances. ScanObjectNN presents significant challenges to existing point cloud analysis methods due to occlusion and noise.

Table 4 shows that our DMR achieves state-of-the-art performance in both mAcc and OA on ScanObjectNN (Uy et al. 2019). It has reached 90.7% in OA and 90.0%. Meanwhile, DMR is 4.8% higher than pointMLP (Ma et al. 2022) in OA after voting, which also achieves a level comparable to SOTA methods (Qi et al. 2017b; Wang et al. 2019; Li et al. 2018; Qiu, Anwar, and Barnes 2021; Ran, Liu, and Wang 2022; Ma et al. 2022; Zheng et al. 2023; Qian et al. 2022; Lin et al. 2023; Park et al. 2023; Yao et al. 2023; Thomas et al. 2024; Chen et al. 2023; Sun et al. 2024). It is worth noting that DMR far surpasses HyCoRe to reach 3.0%, which also further verifies that our method is theoretically correct.

Method	Cls. mIoU	Ins. mIoU
PointNet++	81.9	85.1
DGCNN	80.9	85.1
KPConv	85.1	86.4
PointTransformer	83.7	86.6
PointMLP	84.6	86.1
StratifiedFormer	85.1	86.6
PointNeXt-S	84.6	86.7
PointMetaBase-S	84.4	86.7
X-3D (PointMetaBase-S)	**85.1**	**87.0**
PointMetaBase-S+DMR	**85.3**	**87.1**

Table 5: 3D Part segmentation on ShapeNetPart

Shape Part Segmentation on ShapeNetPart

We further validate our DMR on the ShapeNetPart (Yi et al. 2016) dataset for the 3D shape part segmentation. The dataset collects 16,881 shape models across 16 categories. Most objects in the dataset are labeled with fewer than 6 parts and contain 50 different parts. We replace baseline with PointMetaBase-S (Lin et al. 2023) to test the generality of our method. Table 5 summarizes DMR has achieved the best performance with 87.1% Ins. mIoU, an increase of 0.4% over PointMetaBase-S (Lin et al. 2023) and SOTA methods (Wang et al. 2019; Qi et al. 2017a,b; Thomas et al. 2019; Zhao et al. 2021; Ma et al. 2022; Lai et al. 2022; Qian et al. 2022; Lin et al. 2023; Sun et al. 2024). The segmentation experiment also proves the effectiveness of introducing DMR into the shape segmentation task of point cloud.

Conclusion

In this paper, we propose a novel dual manifold regularization (DMR) method to enhance point cloud analysis. This approach leverages the hyperbolic manifold to learn hierarchical relationships within point clouds and utilizes the hyperspherical manifold to learn more discriminative features between different categories. Extensive experimental results demonstrate that this method enables the learning of more robust point cloud representations and significantly improves model generalization performance. This method is highly significant for advancing point cloud analysis and provides new directions for future research.

Acknowledgments

This work was supported by the National Science Fund of China under Grant Nos. 62361166670, U24A20330 and 62276135.

References

Armstrong, M. A. 2013. *Basic topology*. Springer Science & Business Media.

Atigh, M. G.; Schoep, J.; Acar, E.; Van Noord, N.; and Mettes, P. 2022. Hyperbolic image segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4453–4462.

Cannon, J. W.; Floyd, W. J.; Kenyon, R.; Parry, W. R.; et al. 1997. Hyperbolic geometry. *Flavors of geometry*, 31(59-115): 2.

Chen, B.; Xia, Y.; Zang, Y.; Wang, C.; and Li, J. 2023. Decoupled local aggregation for point cloud learning.

Chen, S.; Niu, G.; Gong, C.; Li, J.; Yang, J.; and Sugiyama, M. 2021. Large-margin contrastive learning with distance polarization regularizer. In *International Conference on Machine Learning*, 1673–1683. PMLR.

Derrick, W. 1973. A condition under which a mapping is a homeomorphism. *The American Mathematical Monthly*, 80(5): 554–555.

Ganea, O.; Bécigneul, G.; and Hofmann, T. 2018. Hyperbolic neural networks. *Advances in Neural Information Processing Systems*, 31.

Ge, S.; Mishra, S.; Kornblith, S.; Li, C.-L.; and Jacobs, D. 2023. Hyperbolic contrastive learning for visual representations beyond objects. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 6840–6849.

Guo, M.-H.; Cai, J.-X.; Liu, Z.-N.; Mu, T.-J.; Martin, R. R.; and Hu, S.-M. 2021. Pct: Point cloud transformer. *Computational Visual Media*, 7: 187–199.

Hu, H.; Wang, F.; Zhang, Z.; Wang, Y.; Hu, L.; and Zhang, Y. 2023a. GAM: Gradient attention module of optimization for point clouds analysis. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 835–843.

Hu, X.; Zhong, B.; Liang, Q.; Zhang, S.; Li, N.; Li, X.; and Ji, R. 2023b. Transformer tracking via frequency fusion. *IEEE Transactions on Circuits and Systems for Video Technology*, 34(2): 1020–1031.

Lai, X.; Liu, J.; Jiang, L.; Wang, L.; Zhao, H.; Liu, S.; Qi, X.; and Jia, J. 2022. Stratified transformer for 3d point cloud segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 8500–8509.

Li, L.; Zhou, T.; Wang, W.; Li, J.; and Yang, Y. 2022. Deep hierarchical semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1246–1257.

Li, Y.; Bu, R.; Sun, M.; Wu, W.; Di, X.; and Chen, B. 2018. Pointcnn: Convolution on x-transformed points. *Advances in Neural Information Processing Systems*, 31.

Lin, H.; Zheng, X.; Li, L.; Chao, F.; Wang, S.; Wang, Y.; Tian, Y.; and Ji, R. 2023. Meta architecture for point cloud analysis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 17682–17691.

Lin, Z.-H.; Huang, S.-Y.; and Wang, Y.-C. F. 2020. Convolution in the cloud: Learning deformable kernels in 3d graph convolution networks for point cloud analysis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1800–1809.

Lin, Z.-H.; Huang, S.-Y.; and Wang, Y.-C. F. 2021. Learning of 3d graph convolution networks for point cloud analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(8): 4212–4224.

Lu, T.; Liu, C.; Chen, Y.; Wu, G.; and Wang, L. 2023. APP-Net: Auxiliary-Point-Based Push and Pull Operations for Efficient Point Cloud Recognition. *IEEE Transactions on Image Processing*, 32: 6500–6513.

Ma, X.; Qin, C.; You, H.; Ran, H.; and Fu, Y. 2022. Rethinking Network Design and Local Geometry in Point Cloud: A Simple Residual MLP Framework. In *International Conference on Learning Representations*.

Montanaro, A.; Valsesia, D.; and Magli, E. 2022. Rethinking the compositionality of point clouds through regularization in the hyperbolic space. *Advances in Neural Information Processing Systems*, 35: 33741–33753.

Nickel, M.; and Kiela, D. 2017. Poincaré embeddings for learning hierarchical representations. *Advances in Neural Information Processing Systems*, 30.

Park, J.; Lee, S.; Kim, S.; Xiong, Y.; and Kim, H. J. 2023. Self-positioning point-based transformer for point cloud understanding. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 21814–21823.

Qi, C. R.; Su, H.; Mo, K.; and Guibas, L. J. 2017a. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 652–660.

Qi, C. R.; Yi, L.; Su, H.; and Guibas, L. J. 2017b. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *Advances in Neural Information Processing Systems*, 30.

Qian, G.; Li, Y.; Peng, H.; Mai, J.; Hammoud, H.; Elhoseiny, M.; and Ghanem, B. 2022. Pointnext: Revisiting pointnet++ with improved training and scaling strategies. *Advances in Neural Information Processing Systems*, 35: 23192–23204.

Qiu, S.; Anwar, S.; and Barnes, N. 2021. Geometric back-projection network for point cloud classification. *IEEE Transactions on Multimedia*, 24: 1943–1955.

Ran, H.; Liu, J.; and Wang, C. 2022. Surface representation for point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 18942–18952.

Sala, F.; De Sa, C.; Gu, A.; and Ré, C. 2018. Representation tradeoffs for hyperbolic embeddings. In *International Conference on Machine Learning*, 4460–4469. PMLR.

Sun, S.; Rao, Y.; Lu, J.; and Yan, H. 2024. X-3D: Explicit 3D Structure Modeling for Point Cloud Recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5074–5083.

Thomas, H.; Qi, C. R.; Deschaud, J.-E.; Marcotegui, B.; Goulette, F.; and Guibas, L. J. 2019. Kpconv: Flexible and deformable convolution for point clouds. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 6411–6420.

Thomas, H.; Tsai, Y.-H. H.; Barfoot, T. D.; and Zhang, J. 2024. KPConvX: Modernizing Kernel Point Convolution with Kernel Attention. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5525–5535.

Tifrea, A.; Bécigneul, G.; and Ganea, O.-E. 2018. Poincaré glove: Hyperbolic word embeddings.

Uy, M. A.; Pham, Q.-H.; Hua, B.-S.; Nguyen, T.; and Yeung, S.-K. 2019. Revisiting point cloud classification: A new benchmark dataset and classification model on real-world data. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 1588–1597.

Wang, Y.; Sun, Y.; Liu, Z.; Sarma, S. E.; Bronstein, M. M.; and Solomon, J. M. 2019. Dynamic graph cnn for learning on point clouds. *ACM Transactions on Graphics (tog)*, 38(5): 1–12.

Weinberger, K. Q.; and Saul, L. K. 2009. Distance metric learning for large margin nearest neighbor classification. *Journal of Machine Learning Research*, 10(2).

Wu, Q.; Jiang, H.; Ding, Y.; Luo, L.; Xie, J.; and Yang, J. 2023. Diff-PCR: Diffusion-Based Correspondence Searching in Doubly Stochastic Matrix Space for Point Cloud Registration. *arXiv preprint arXiv:2401.00436*.

Wu, Q.; Jiang, H.; Luo, L.; Li, J.; Ding, Y.; Xie, J.; and Yang, J. 2025. Diff-Reg: Diffusion Model in Doubly Stochastic Matrix Space for Registration Problem. In *European Conference on Computer Vision*, 160–178. Springer.

Wu, W.; Qi, Z.; and Fuxin, L. 2019. Pointconv: Deep convolutional networks on 3d point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 9621–9630.

Wu, Z.; Song, S.; Khosla, A.; Yu, F.; Zhang, L.; Tang, X.; and Xiao, J. 2015. 3d shapenets: A deep representation for volumetric shapes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1912–1920.

Xiang, T.; Zhang, C.; Song, Y.; Yu, J.; and Cai, W. 2021. Walk in the cloud: Learning curves for point clouds shape analysis. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 915–924.

Xu, M.; Zhang, J.; Zhou, Z.; Xu, M.; Qi, X.; and Qiao, Y. 2021. Learning geometry-disentangled representation for complementary understanding of 3d object point cloud. In *Proceedings of the AAAI conference on artificial intelligence*, 3056–3064.

Xu, Y.; Fan, T.; Xu, M.; Zeng, L.; and Qiao, Y. 2018. Spidercnn: Deep learning on point sets with parameterized convolutional filters. In *Proceedings of the European Conference on Computer Vision*, 87–102.

Yao, T.; Li, Y.; Pan, Y.; and Mei, T. 2023. Hgnet: Learning hierarchical geometry from points, edges, and surfaces. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 21846–21855.

Yi, L.; Kim, V. G.; Ceylan, D.; Shen, I.-C.; Yan, M.; Su, H.; Lu, C.; Huang, Q.; Sheffer, A.; and Guibas, L. 2016. A scalable active framework for region annotation in 3d shape collections. *ACM Transactions on Graphics (ToG)*, 35(6): 1–12.

Zhang, R.; Wang, L.; Wang, Y.; Gao, P.; Li, H.; and Shi, J. 2023. Starting From Non-Parametric Networks for 3D Point Cloud Analysis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5344–5353.

Zhao, H.; Jiang, L.; Jia, J.; Torr, P. H.; and Koltun, V. 2021. Point transformer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 16259–16268.

Zheng, S.; Pan, J.; Lu, C.; and Gupta, G. 2023. Pointnorm: Dual normalization is all you need for point cloud analysis. In *2023 International Joint Conference on Neural Networks*, 1–8. IEEE.