Hyperbolic Uncertainty-Aware Few-Shot Incremental Point Cloud Segmentation

Tanuj Sur[†] Samrat Mukherjee[*] Kaizer Rahaman[‡] Subhasis Chaudhuri[*] Muhammad Haris Khan[§]
Biplab Banerjee[*]

[†]Chennai Mathematical Institute [*]Indian Institute of Technology Bombay [‡]Indian Institute of Technology Kharagpur
[§]Mohamed Bin Zayed University of Artificial Intelligence

Abstract

3D point cloud segmentation is essential across a range of applications; however, conventional methods often struggle in evolving environments, particularly when tasked with identifying novel categories under limited supervision. Few-Shot Learning (FSL) and Class Incremental Learning (CIL) have been adapted previously to address these challenges in isolation, yet the combined paradigm of Few-Shot Class Incremental Learning (FSCIL) remains largely unexplored for point cloud segmentation. To address this gap, we introduce ***Hyperbolic Ideal Prototypes Optimization*** *(*HIPO*), a novel framework that harnesses hyperbolic embeddings for FSCIL in 3D point clouds.* HIPO *employs the Poincaré Hyperbolic Sphere as its embedding space, integrating Ideal Prototypes enriched by CLIP-derived class semantics, to capture the hierarchical structure of 3D data. By enforcing orthogonality among prototypes and maximizing representational margins,* HIPO *constructs a resilient embedding space that mitigates forgetting and enables the seamless integration of new classes, thereby effectively countering overfitting. Extensive evaluations on S3DIS, ScanNetv2, and cross-dataset scenarios demonstrate* HIPO*'s strong performance, significantly surpassing existing approaches in both in-domain and cross-dataset FSCIL tasks for 3D point cloud segmentation.*

1. Introduction

Recent advancements in deep learning have spurred progress in data-efficient learning, with semi-supervised learning (SSL) [62, 69, 71] and few-shot learning (FSL) [55, 58, 66] emerging as key approaches. SSL utilizes both labeled and unlabeled data to enhance learning, while FSL enables generalization from only a few labeled samples per class. However, both approaches face challenges in real-world, dynamic environments where models must continuously adapt to new classes while retaining prior knowledge—an essential requirement for continual learning (CL) [1, 4, 8, 11, 32, 48, 56]. Few-Shot Class Incremental Learning (FSCIL) [1, 13, 24, 28, 30, 39] combines the data efficiency of FSL with the adaptability of CL, yet remains underdeveloped for complex tasks such as segmentation in unstructured 3D data (see Fig. 1).

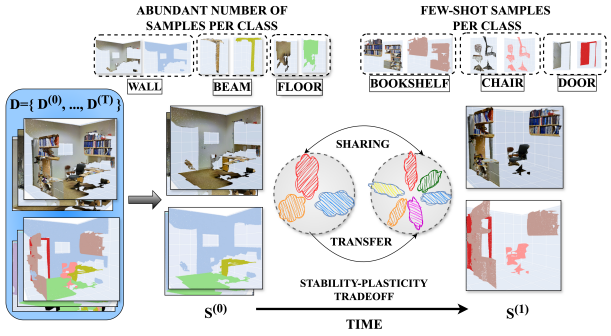


Figure 1. Our approach tackles the novel challenge of **few-shot class incremental 3D point cloud segmentation**. In each session $S^{(t)}, t \geq 0$, our model takes an RGB point cloud of an indoor scene $D^{(t)}$ as input and performs semantic segmentation. During the base training session $S^{(0)}$, sufficient examples are available for each class. In later sessions, only a limited number of labeled point cloud examples are provided for new classes, introducing data imbalance and potential overfitting and forgetting risks.

Point cloud segmentation is pivotal in applications like autonomous driving and robotics, where it enables precise object recognition and scene understanding [36, 52, 68]. This task assigns semantic labels to each point within a 3D structure, facilitating critical decision-making processes. Current supervised segmentation methods [26, 53, 65, 67] are highly reliant on extensive labeled datasets, which are costly and time-intensive to produce [34]. These methods also assume a fixed set of labels, limiting their adaptability in scenarios where new classes emerge unexpectedly. While FSCIL offers a potential solution, existing methods [1, 24, 72, 78, 79, 81, 85] struggle to represent the complex, hierarchical structures typical of point cloud data due to the limitations of Euclidean embeddings [33, 45, 46].

Capturing hierarchical relationships, such as recogniz-

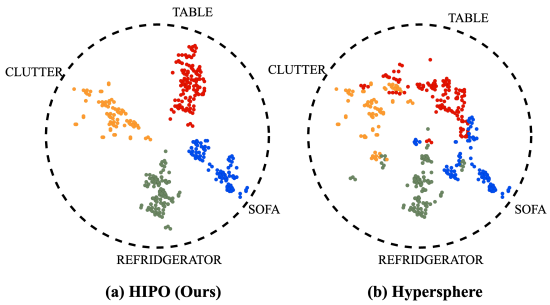
[*]Samrat Mukherjee acknowledges the support of Prime Minister Research Fellowship (PMRF).

ing subcomponents within an object, demands a framework that can accommodate intricate interrelations. This challenge is intensified in Class Incremental Learning (CIL) [82, 83], where the model must incorporate new classes while preserving previously learned representations—a balance between adaptability and memory known as the *stability-plasticity* trade-off. In FSCIL, especially under data scarcity, this trade-off is often compromised, leading to overfitting and catastrophic forgetting. Euclidean spaces exacerbate these challenges, as they lack the capacity to model evolving data scenarios [76, 81].

Hyperbolic geometry [25, 38, 42, 57, 74] provides a more suitable framework for hierarchical data [45, 84]. Its negative curvature and exponential volume growth allow it to mirror tree-like data structures, addressing the stability-plasticity trade-off [16] and supporting segmentation in unstructured contexts [27]. Unlike Euclidean spaces, hyperbolic space offers exponentially more volume [44] for accommodating incremental classes, thus reducing overlap and interference (see Fig. 2, Tab. 4). However, efficient optimization of embeddings in the hyperbolic space presents computational challenges when points are near the boundary of the Poincaré ball resulting in vanishing gradient problems [44], and without careful handling, newly added classes can still encroach upon the existing embeddings, aggravating *plasticity* (see Tab. 3). This highlights the need for an FSCIL framework that uses hyperbolic space embedding hierarchical data in low distortion while ensuring computational efficiency and managing class overlap.

Our Proposal: To address these challenges, we introduce **Hyperbolic Ideal Prototypes Optimization (HIPO)**, a novel FSCIL framework tailored for point cloud segmentation. By embedding data within the Poincaré Hyperbolic Sphere [49], HIPO leverages the unique properties of hyperbolic space to capture hierarchical structures and maintain consistent feature-classifier alignment across incremental learning sessions. Through a modified Busemann loss [7], HIPO arranges ideal prototypes orthogonally within the hyperbolic space, maximizing class separation and preserving space for new classes, which minimizes inter-class interference and optimizes the feature space for CIL.

This design takes advantage of hyperbolic space's exponential scaling, which aligns naturally with complex hierarchies found in point cloud data [3, 27, 45, 57]. Our modified Busemann loss incorporates two components: (1) a prototype alignment term to guide class embeddings towards their respective Hyperbolic Ideal Prototypes, enhancing class separability, and intra-class compactness, and (2) an uncertainty-aware term that penalizes overfitting on novel classes and preventing vanishing gradient problems. Together, these components enable HIPO to effectively address data scarcity, mitigate catastrophic forgetting, and support robust FSCIL under episodic shifts, establishing it as a powerful solution for incremental learning in 3D point cloud segmentation. Overall, our contributions are:



(a) HIPO (Ours) **(b) Hypersphere**

Figure 2. **Motivation for hyperbolic embeddings:** (a) Hyperbolic UMAP [40] of the HIPO embedding space, projected to 2D for the S3DIS dataset [2]. The hyperbolic space's negative curvature aligns with hierarchical data, enabling a more compact representation and clearer class separation than Euclidean space (**Dunn index [18] (↑) for HIPO is 0.69 while for the Euclidean space, it is 0.35**), showcasing better class compactness in our case. (b) UMAP of the classifier space in the Euclidean hypersphere

- We define and address the previously unexplored problem of FSCIL for 3D point cloud segmentation, encompassing both in-domain and cross-dataset scenarios.
- We introduce hyperbolic space as a robust embedding framework for FSCIL in 3D segmentation, effectively capturing hierarchical structures and preserving stable feature-class relationships across incremental learning sessions.
- We propose a modified Busemann loss with prototype alignment and uncertainty-aware components to ensure optimal feature-classifier alignment with minimized uncertainty in predictions.
- Through extensive experiments on benchmark datasets S3DIS [2] and ScanNetv2 [17], HIPO demonstrates superior performance and generalization, setting a new standard for FSCIL in this domain.

2. Related Works

3D Point-Cloud Semantic Segmentation: Fully supervised 3D semantic segmentation has been a well-explored area, with methods typically requiring dense point-wise labels [26, 52, 53, 65, 67, 77]. A significant advancement was PointNet [52], which directly processes raw point clouds, removing the need for voxelization but facing challenges in capturing local structures. This limitation spurred improvements like PointNet++ [53], which introduced a hierarchical approach to enhance local feature learning, and DGCNN [65], which incorporated the EdgeConv module to integrate neighborhood information for better geometric representation. More recently, models such as PointTransformer [80] have integrated self-attention mechanisms, inspired by NLP, to elevate feature expressiveness and contextual awareness.

Despite these advancements, the dependency on large labeled datasets and the assumption of a fixed class set restrict these methods in practical scenarios where new classes must be incrementally learned with minimal supervision.

(Few-Shot) Class Incremental Learning: Incremental learning methods for 3D data have predominantly focused on classification tasks rather than segmentation [14, 15, 19, 35, 75]. Approaches such as I3DOL [19] and L3DOC [35] mitigate catastrophic forgetting in 3D object classification through memory and attention mechanisms, yet they struggle as labeled supervision diminishes over time. Few-Shot Class Incremental Learning (FSCIL) [61] has emerged as a more flexible solution, enabling models to learn new classes incrementally with minimal labeled data. Representative methods like CEC [76] use graph-based task linking, while FACT [81] and ALICE [50] create adaptable feature spaces through clustering and cosine similarity. However, these techniques rely on Euclidean embeddings, which have limited capability to represent hierarchical structures intrinsic to 3D data, and frequent classifier updates introduce computational overhead [76]. Open-world semantic segmentation (OWSS) [5, 9] is a related area that focuses on dynamically detecting unknown classes at inference, though our focus on incrementally adding known classes during training sets our approach apart from OWSS.

Hyperbolic Representations for Hierarchical Data: Hyperbolic geometry has shown promise for hierarchical data representation, with foundational work by Sarkar et al. [57] and Nickel et al. [49] introducing hyperbolic neural networks and Riemannian optimization methods. Hyperbolic embeddings have demonstrated substantial benefits for capturing semantic hierarchies across applications, including image embeddings [27] and FSCIL for image classification [16]. In 3D data, recent research has shown that hyperbolic representations can model complex part-whole relationships effectively. Montanaro et al. [45] applied hyperbolic regularizers to encode hierarchical structures in 3D objects, while Hu et al. [25] leveraged hyperbolic geometry to represent latent structures in point clouds.

Our approach leverages hyperbolic embeddings to tackle FSCIL challenges in segmentation, improving adaptability, preserving hierarchies, and ensuring computational efficiency in the point cloud domain.

3. Problem Definition & Methodology

Let D represent a sequence of training datasets $\{D^{(0)}, D^{(1)}, \ldots, D^{(T)}\}$, where each $D^{(t)} = \{(P_k^{(t)}, M_k^{(t)})_{k=1}^{|D^{(t)}|}\}$ corresponds to a training session $S^{(t)}$, $t \in \{0, 1, \ldots, T\}$. The initial session dataset $D^{(0)}$ serves as the base, while T denotes the total number of incremental sessions. In each session $S^{(t)}$, $P_k^{(t)}$ represents a point cloud sample, and $M_k^{(t)}$ provides its labels within the label space $C^{(t)} = \{c_1^{(t)}, c_2^{(t)}, \ldots, c_{|C^{(t)}|}^{(t)}\}$ where $c_i^{(t)}$ is the i-th class in the incremental session t.

The base session $S^{(0)}$ typically contains an abundant number of samples per class in the label space $C^{(0)}$. However, in each subsequent incremental session $S^{(t)}, t > 0$, only a limited number of labeled point clouds $P_k^{(t)} \in \mathbb{R}^{m \times d_0}$ are provided, where m is the number of points per sample and d_0 is the feature dimension. Following previous works [1, 24, 72, 78, 79, 81, 85], we assume $|D^{(t)}| = N \times K$, where N is the number of classes, and K is the number of samples per class, following an N-way K-shot setting $\forall\, t > 0$. We assume $C^{(t)}$ to be disjoint from previous sessions, i.e., $C^{(t)} \cap C^{(t')} = \emptyset\,, \forall\, t' < t$.

During each incremental session, only the current dataset $D^{(t)}$ is accessible, with data from previous sessions unavailable for training. For evaluation in session t, the model is tested across all classes encountered up to that session, covering the label space $\bigcup_{i=0}^{t} C^{(i)}$. A solution to the FSCIL problem must jointly address uncharted catastrophic forgetting, data imbalance, and class overfitting judiciously.

3.1. Preliminaries

This section mentions the essential backgrounds needed for the proposed framework **Hyperbolic Ideal Prototypes Optimization (HIPO)**. We extensively use the concepts of the Poincaré ball model and its associated Möbius addition, exponential map, and the geodesic distance. We detail their definition and discuss their formulae in the ***Sup. Mat.***

In the Poincaré ball model of hyperbolic geometry, all the points at the boundary of the Poincaré ball represent points at infinity. We initialize a group of classifier prototypes at the boundary of the Poincaré ball termed Ideal Prototypes, following [23], for each of the n individual classes. Assuming $u_1, u_2, \cdots, u_n$ to be the Ideal Prototypes for n classes, they form the boundary of the n-dimensional Poincaré ball of curvature c, $\mathbb{B}_c^n$, following [23]:

$$\mathcal{I}_n = \{u \in \mathbb{R}^n : u_1^2 + u_2^2 + \ldots + u_n^2 = 1\} \tag{1}$$

Since Ideal Prototypes are points at an infinite geodesic distance from all other points in $\mathbb{B}_c^n$, we use the Busemann function [7] as a distance measure since computing direct infinite geodesic distance is non-trivial [23].

The pre-assigned Ideal Prototypes form an optimal geometric structure by fixing the target throughout the incremental learning. However, as already pointed out, optimizing Ideal Prototypes in the Poincaré model to address FSCIL in 3D point cloud segmentation remains unexplored.

3.2. Discussing HIPO

Our model introduces a fixed alignment strategy for feature-classifier relationships in FSCIL, tailored for the hierarchical nature of data such as 3D point clouds. Traditional Euclidean embeddings suffer from compaction [50, 76, 81]

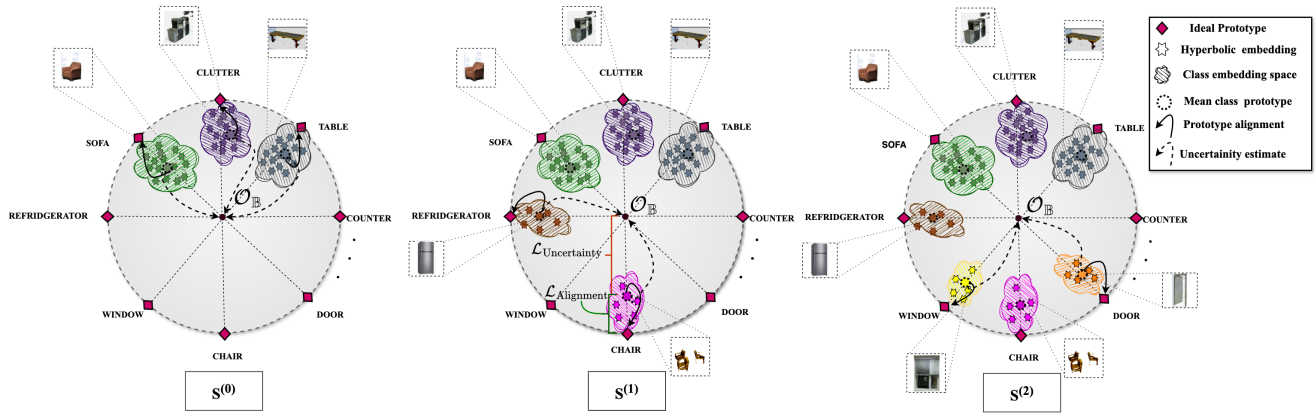


Figure 3. **HIPO Overview:** We show the changes in the Poincaré ball during the initial three sessions in the 3-2T FSCIL setting (three base classes, two incremental classes per session) for the S3DIS dataset [2], designed to easily visualize the working of HIPO. Hyperbolic Ideal Prototypes are optimally positioned to increase the margin of semantically similar classes ($\mathcal{L}_{\text{HIP}}$). As an example, the **chair** class gets positioned farther from **sofa** than from **table** or **clutter**. We align representations of the newly introduced classes with their corresponding ideal prototypes ($\mathcal{L}_{\text{Alignment}}$) regularized by a modified Busemann function ($\mathcal{L}_{\text{Uncertainty}}$). **Training is conducted using $\mathcal{L}_{\text{HIPO}}$ across sessions.**

and distortion as new classes are added, leading to overlap and forgetting. Hyperbolic space, with its negative curvature and exponential volume, provides a natural fit for preserving hierarchical structures while minimizing inter-class interference. To address catastrophic forgetting, our model anchors feature-classifier relationships within hyperbolic space, ensuring stable alignment between features and classifier prototypes. This fixed alignment prevents the drift of previously learned embeddings as new classes are introduced, preserving the original classifier prototype positions and maintaining consistency across incremental sessions. By leveraging hyperbolic geometry, our approach enables efficient class separation, robust memory retention, and adaptability, making it well-suited for FSCIL in high-dimensional, unstructured data settings. Our method has two main stages:

- Defining Ideal Prototypes: We initialize ideal prototypes for each class on the boundary of $\mathbb{B}_c^n$, positioning them mutually orthogonal and at maximal distances to enhance discriminative power, especially among classes with similar characteristics (see Sec. 3.2.1). This fixed configuration allows a dedicated space for each class, preventing the embedding drift observed in Euclidean spaces.

- Efficient Projections onto Poincaré ball $\mathbb{B}_c^n$: In Hyperbolic embedding space, mean class features are obtained by computing Fréchet [20] means over the feature embedding for the training examples belonging to the classes. Instead of directly computing class features on the Poincaré ball—an intensive task[63]—we obtain the feature vectors in Euclidean space for every class in each session and then project them into hyperbolic space using exponential map $\exp_0(.)$ [45] (see Sec. 3.2.2). The hyperbolic classifier head $\mathcal{H}(.)$ consisting of a Möbius layer $\mathcal{M}$ (also known as hyperbolic feed forward layer [22]), takes the output of the exponential layer as input and projects mean class features to $\mathbb{B}_c^n$. The Poincaré ball addresses few-shot learning using the sample mean prototype as the class representation.

Our framework consists of a point-level classification branch, combining a feature extractor $\mathcal{E}(.)$ and hyperbolic classifier head $\mathcal{H}(.)$, and is trained with two key strategies:

Feature-Classifier Alignment: For each session $t \geq 0$, we align the hyperbolic sample mean features (see Eqn. 4) with the Hyperbolic Ideal Prototypes of their respective classes using alignment loss, $\mathcal{L}_{\text{Alignment}}$. We define $\mathcal{L}_{\text{Alignment}}$ using Busemann function $\mathcal{B}$ and introduce a regularization term, quantifying uncertainty of the class features, to prevent feature collapse around the ideal prototypes, effectively reducing interference with previously seen classes (see Sec. 3.2.3) and vanishing gradient problems [44].
Point-Level Classification: The hyperbolic classifier $\mathcal{H}(.)$, consisting of a Möbius layer $\mathcal{M}$ followed by softmax [45], processes the output of $\exp_0(\mathcal{E}(.))$, with a cross-entropy loss for 3D point cloud segmentation. The joint objective combines this alignment loss $\mathcal{L}_{\text{Alignment}}$ and cross-entropy loss $\mathcal{L}_{\text{CE}}$ as follows:

$$\mathcal{L}_{\text{HIPO}} = \mathcal{L}_{\text{Alignment}} + \mathcal{L}_{\text{CE}} \tag{2}$$

This approach ensures robust feature-classifier alignment across incremental learning sessions, addressing catastrophic forgetting in 3D point cloud segmentation by preserving both old and new class representations in a rehearsal-free manner. The process is illustrated in Fig. 3, and *Algorithm 1* and a table summarizing the important variables used, mentioned in the ***Sup. Mat.***, supplement these discussions.

3.2.1. Hyperbolic ideal prototypes positioning

Building upon the goal of maintaining stable feature-classifier alignment, our framework pre-assigns classifier

prototypes on the boundary of $\mathbb{B}_c^n$ to define a structured classifier space prior to backbone training. To achieve maximal margin separation between classes, we introduce inductive biases by assigning each class a Hyperbolic Ideal Prototype, distributing these prototypes as uniformly as possible across the Poincaré ball. Inspired by Mettes *et al*. [41], we incorporate semantic class knowledge to enforce distinct separations, particularly for closely related classes, such as "chair" and "sofa". To further enhance semantic cohesion, we utilize the CLIP framework [54] due to its robust semantic representations [31, 73] in comparison to embeddings like word2vec [43], Glove [51] etc. By querying CLIP's frozen text encoder with the prompt "`A photo of a [class]`", we obtain text embeddings $S = \{s_1, \ldots, s_n\}$ for each of the n predefined classes. Leveraging the homeomorphism between hyperbolic space $\mathbb{B}_c^n$ and the hypersphere $\mathbb{S}^n$ [23], we maximize separation between nearest neighboring embeddings in $\mathbb{S}^n$ and enforce mutual orthogonality among Ideal Prototypes to mitigate conflict between new and prior knowledge. To achieve this, inspired by RankNet [6], we formulate a Hyperbolic positioning loss, $\mathcal{L}_{\text{HIP}}$, optimizing the ideal prototypes positions as:

$$\mathcal{L}_{\text{HIP}} = \frac{1}{|M|} \sum_{(i,i',i'') \in M} -\bar{W}_{ii'i''} \log W_{ii'i''} - (1 - \bar{W}_{ii'i''}) \log(1 - W_{ii'i''}) \quad (3)$$

where M represents all class triplets, and $\bar{W}_{ii'i''} = [\![\cos\theta_{\mathbf{s}_i,\mathbf{s}_{i'}} \geq \cos\theta_{\mathbf{s}_i,\mathbf{s}_{i''}}]\!]$ denotes triplet ranking order, with $W_{ii'i''} = \frac{e^{o_{ii'i''}}}{1+e^{o_{ii'i''}}}$ and $o_{ii'i''} = \cos\theta_{\mathbf{u}_i,\mathbf{u}_{i'}} - \cos\theta_{\mathbf{u}_i,\mathbf{u}_{i''}}$, where $\mathbf{u}_i, \mathbf{u}_{i'}, \mathbf{u}_{i''}$ are ideal prototypes for the corresponding classes i, $i^{'}$ and $i^{''}$ respectively.

Following Mettes *et al*. [41], $\mathcal{L}_{\text{HIP}}$ optimizes the prototypes' placement on $\mathbb{B}_c^n$ by aligning the hyperbolic prototypes with semantic ranking. This optimization establishes a classifier structure that facilitates inter-class separability and mutual orthogonality. Hyperbolic space's exponentially larger embedding capacity [44] further ensures that new classes integrate seamlessly without distorting existing embeddings, promoting robust feature organization across learning episodes.

3.2.2. Hyperbolic class prototype projection

We model the hierarchical complexity in 3D point clouds using a hyperbolic manifold representation, which preserves nuanced class relationships with minimal distortion.

At each session $t \geq 0$, we apply the feature extractor $\mathcal{E}(.;\theta)$ on input point clouds $\{(P_k^{(t)})_{k=1}^{|D^{(t)}|}\}$ to obtain feature vectors $\boldsymbol{v}$ in Euclidean space. To avoid the computational cost of calculating Fréchet means directly in hyperbolic space [63], we compute class-mean feature vectors in Euclidean space and project them onto the hyperbolic manifold $\mathbb{B}_c^n$ using $\exp_0$ and $\mathcal{M}$ as follows:

$$\begin{aligned} v_k^{(t)} &= \mathcal{E}(P_k^{(t)};\theta), \quad \forall k \in \{1, \ldots, |D^{(t)}|\} \\ \overline{v}(c_i^{(t)}) &= \frac{1}{|D^{(t)}|} \sum_{k=1}^{|D^{(t)}|} v_k^{(t)} \mathbb{1}[M_k^{(t)} = c_i^{(t)}] \\ z(c_i^{(t)}) &= \mathcal{M}(\exp_0(\overline{v}(c_i^{(t)}))) \end{aligned} \quad (4)$$

where $i = \{1, \ldots, |C^{(t)}|\}$ and θ denotes learnable parameters. Here, $z(c_i^{(t)})$ represents the hyperbolic prototype for each class $c_i^{(t)} \in C^{(t)}$ within the open ball $\mathbb{B}_c^n = \{x \in \mathbb{R}^n : c\|\boldsymbol{x}\| < 1\}$. This projection process ensures cohesive, low-distortion representations of class structures as the model adapts incrementally to new classes.

3.2.3. Uncertainty-aware Busemann learning

Previous works [12, 21, 27] highlight the utility of hyperbolic distance from any point to the origin $\mathcal{O}_{\mathbb{B}}$ in $\mathbb{B}_c^n$ as a measure of uncertainty, providing motivation to utilize the hyperbolic manifold for robust segmentation. In our approach, we align class prototypes with Ideal Prototypes placed at infinity on the hyperbolic boundary. To refine this alignment, we penalize the uncertainty of projected class features $z(c_i^{(t)})$, represented by their distance $d_{\mathbb{B}}(z(c_i^{(t)}), \mathcal{O}_{\mathbb{B}})$, thereby preventing embedding drift from $\mathcal{O}_{\mathbb{B}}$ and ensuring stability.

These Ideal Prototypes lie at an infinite geodesic distance from all points within $\mathbb{B}_c^n$, making direct prototypical learning impractical. We address this with the Busemann function $\mathcal{B}$ introduced in [7]. Ghadimi *et al*. [23] provides a closed-form expression for measuring such infinite geodesic distances. Let u_i be the Ideal Prototype for class $c_i^{(t)}$ from the set $\mathcal{I}_n$, and $\gamma_{u_i^t}$ a geodesic ray towards u_i^t, with $z_i^t = z(c_i^{(t)})$. The Busemann function $\mathcal{B}$ for a hyperbolic embedding $z_i^t \in \mathbb{B}_c^n$ relative to u_i^t is defined as:

$$\mathcal{B}_{u_i^t}(z_i^t) = \lim_{q \to \infty} \left(d_{\mathbb{B}}(\gamma_{u_i^t}(q), z_i^t) - q \right) \quad (5)$$

In the Poincaré ball $\mathbb{B}_c^n$, this simplifies to:

$$\mathcal{B}_{u_i^t}(z_i^t) = \log \frac{\|u_i^t - z_i^t\|^2}{1 - \|z_i^t\|^2}. \quad (6)$$

We also incorporate hyperbolic distance as an uncertainty measure for class prototypes in $\mathbb{B}_c^n$, which integrates into our loss function as in [23]. The alignment loss introduced in Eq. 2 is defined as:

$$\mathcal{L}_{\text{Alignment}} = \mathcal{L}_{\mathcal{B}} + \phi \mathcal{L}_{\text{Uncertainty}} \quad (7)$$

where Busemann loss $\mathcal{L}_{\mathcal{B}} = \mathcal{B}_{u_i^t}(z_i^t)$, uncertainty loss $\mathcal{L}_{\text{Uncertainty}} = d_{\mathbb{B}}(z_i^t, \mathcal{O}_{\mathbb{B}})$ and ϕ is the strength of the regularization effect of uncertainty loss.

The overall loss function in Eq. 2 can be formulated as:

$$\mathcal{L}_{\text{HIPO}} = \mathcal{L}_{\mathcal{B}} + \phi \mathcal{L}_{\text{Uncertainty}} + \mathcal{L}_{\text{CE}}$$

At test time, we employ *only* the point-level classification branch to segment 3D point clouds.

4. Experimental Evaluations

Datasets: We evaluate our method on two widely used benchmarks, S3DIS [2] and ScanNetv2 [17], to assess both in-domain and cross-dataset FSCIL performance. These datasets incorporate large-scale coloured 3D point clouds of indoor spaces which introduce semantics. See ***Sup. Mat.*** for additional dataset details. For cross-dataset FSCIL, we create a transition dataset from ScanNetv2 to S3DIS, using ten classes from ScanNetv2 as the base and introducing eight new classes from S3DIS in incremental sessions. Following a strict disjoint setup as in [10], each session includes only novel classes. The number of introduced classes varies per session to assess resilience to catastrophic forgetting, which escalates with more incremental sessions in point cloud recognition [60].

Evaluation metrics: We assess model performance using three standard metrics for incremental learning[70]: **Average Incremental Accuracy (AIA)** , which tracks average accuracy across tasks learned sequentially; **Last**, the accuracy after the final task; and **Average Forgetting Rate** ($\mathcal{F}^{(T)}_{Last}$), which measures performance decline on learned tasks after adapting to the last task. Finally, we report mean Intersection-over-Union (mIoU)[1] to evaluate segmentation performance. See ***Sup. Mat.*** for detailed metric definitions.

Architecture: Our backbone model employs PointTransformer [77], leveraging recent advances in point cloud recognition, with a Hyperbolic Classifier head $\mathcal{H}(.)$ for segmentation. We train HIPO using the Riemannian SGD optimizer, facilitated by geoopt [29] for hyperbolic operations. Training is conducted on both $D^{(0)}$ and incremental datasets $D^{(t)}, t > 0$ with a batch size of 16, AdamW optimizer [37], initial learning rate of 0.001, and weight decay of 0.0001 over 100 epochs. Architecture diagram and further hyper-parameter details are available in the ***Sup. Mat.***

Competitors: We compare our method against various baselines representing distinct FSCIL approaches. For regularization-based methods, we include **LwF with Teacher Adaptation** [59], and for optimization-based methods, we include **C-FSCIL** [24]. We also evaluate against regularization-based FSCIL methods specifically designed for 3D point cloud tasks, such as **LGKD** [72], **SDCOT** [79], and **3DPC-CISS** [70], which have shown effectiveness with the PointTransformer [77] architecture. Comparative results are provided in Tables 1 and 2. Additionally, we adapt **BiDist** [78], **CLOM** [85], **TEEN** [64], **FACT** [81], and **ALICE** [50] using the DGCNN [65] backbone for FSCIL in 3D point cloud segmentation. We report the results with the DGCNN [65] backbone in Table 5.

[1]We use accuracy and mIoU interchangably.

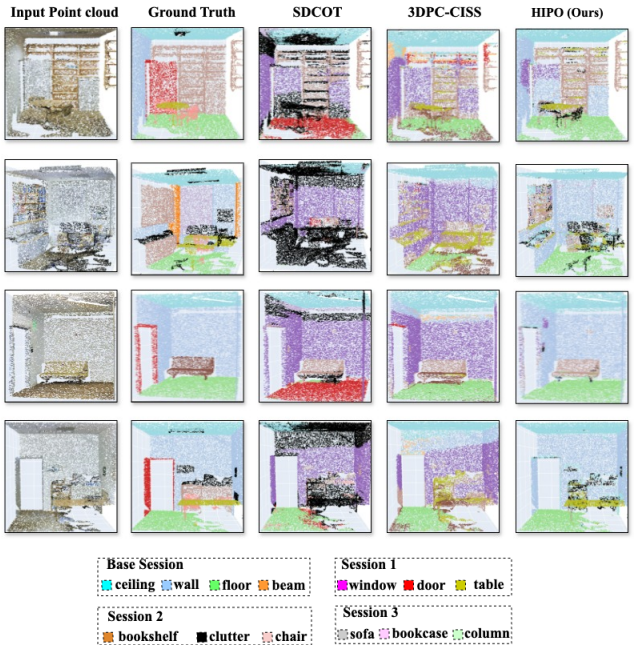


Figure 4. **Qualitative visualization of model performance:** Comparing FSCIL for 3D point cloud segmentation results of our method (HIPO) and other baselines for S3DIS dataset [2]. We observe that even with a few-shot setup, our model performs reasonably well in segmenting out small parts like walls, counters, etc., compared to other methods [70, 79].

4.1. Discussing the main results

Table 1 presents a detailed comparison of our proposed method, HIPO, against several baseline methods on the S3DIS [2] and ScanNetv2 [17] datasets across various evaluation settings. HIPO consistently outperforms the current state-of-the-art approach, 3DPC-CISS [70], achieving an average improvement of $+2.57\%$ mIoU on S3DIS and $+2.12\%$ on $\mathcal{F}^{(T)}_{Last}$. This improvement underscores the effectiveness of HIPO's hyperbolic, uncertainty-aware architecture, which mitigates catastrophic forgetting by reducing misalignment between features and prototype representations from earlier sessions. When compared to other baselines, such as LGKD [72] and C-FSCIL [24], HIPO shows substantial gains across all metrics. For instance, in the 10-3T scenario, HIPO achieves 33.40% mIoU on the **Last** metric, significantly higher than LGKD's 20.02% and C-FSCIL's 25.02%. This consistent superiority highlights HIPO's resilience against forgetting, particularly in incremental sessions with larger class increments, where other methods struggle to retain performance. Furthermore, while SDCOT [79] and LwF [59] suffer from high forgetting rates, HIPO effectively maintains stable performance across incremental tasks, with a $\mathcal{F}^{(T)}_{Last}$ rate as low as 17.70% in challenging scenarios, compared to SDCOT's 30.83%. This robustness can be attributed to HIPO's hyperbolic embed-

	4-3T			7-3T			10-3T			Average		
	Last(↑)	**AIA**(↑)	$\mathcal{F}^{(T)}_{Last}$(↓)	**Last**(↑)	**AIA**(↑)	$\mathcal{F}^{(T)}_{Last}$(↓)	**Last**(↑)	**AIA**(↑)	$\mathcal{F}^{(T)}_{Last}$ (↓)	**Last**(↑)	**AIA**(↑)	$\mathcal{F}^{(T)}_{Last}$(↓)
LwF [59]	4.33	22.05	23.62	8.60	21.80	19.80	21.30	34.05	25.50	11.41	25.30	22.97
C-FSCIL [24]	9.06	27.90	25.11	20.01	31.92	17.86	25.02	35.36	20.68	18.03	31.06	21.88
LGKD [72]	12.24	31.80	25.94	19.76	30.41	15.97	20.02	32.71	25.38	17.34	31.64	22.43
SDCOT [79]	5.30	28.42	30.83	12.00	29.46	26.20	23.80	35.30	23.00	13.70	31.06	26.67
3DPC-CISS [70]	25.70	40.50	19.73	24.20	34.50	15.45	29.60	38.20	17.20	26.50	37.73	17.46
HIPO	**30.60**	**43.87**	**17.70**	**29.40**	**36.80**	**11.10**	**33.40**	**40.25**	**13.70**	**31.13**	**40.30**	**14.16**

	10-2T			10-5T			15-5T			Average		
	Last(↑)	**AIA**(↑)	$\mathcal{F}^{(T)}_{Last}$ (↓)	**Last**(↑)	**AIA**(↑)	$\mathcal{F}^{(T)}_{Last}$ (↓)	**Last**(↑)	**AIA**(↑)	$\mathcal{F}^{(T)}_{Last}$ (↓)	**Last**(↑)	**AIA**(↑)	$\mathcal{F}^{(T)}_{Last}$ (↓)
LwF [59]	1.70	10.75	10.86	1.05	17.53	24.72	5.06	21.38	32.64	2.60	16.55	22.74
C-FSCIL [24]	7.12	15.94	10.58	11.12	24.18	19.59	11.50	23.58	**24.15**	9.25	21.23	**18.11**
LGKD [72]	6.02	15.60	11.50	9.55	25.52	22.45	11.10	23.83	25.46	8.22	21.65	19.80
SDCOT [79]	0.50	9.56	10.88	0.40	17.83	26.15	1.30	19.50	36.40	0.73	15.63	24.47
3DPC-CISS [70]	6.20	13.13	**8.32**	8.10	22.30	21.30	**12.20**	**24.95**	25.50	8.83	20.12	18.37
HIPO	**8.50**	**17.13**	10.36	**14.65**	**27.22**	**18.85**	11.90	24.45	25.10	**11.68**	**22.27**	18.12

Table 1. Comparison with literature on the **S3DIS** [2] (top table) and **ScanNetv2** [17] (bottom table) datasets, using the PointTransformer backbone [77] in a 5-shot setting. CIL mIoU (%) is reported. "M-NT" indicates that M is the number of classes in the base session $S^{(0)}$, while N is the number of classes introduced in each incremental session $S^{(t)}, t > 0$. The highest performance in each column is shown in **bold**, with the second-highest results underlined. HIPO performs superior compared to other baselines in different incremental settings.

	Average		
	Last (↑)	**AIA** (↑)	$\mathcal{F}^{(T)}_{Last}$ (↓)
LwF [59]	8.78	22.60	23.94
C-FSCIL [24]	13.55	25.02	19.72
LGKD [72]	15.63	25.56	17.51
SDCOT [79]	19.22	29.29	17.52
3DPC-CISS [70]	19.95	29.16	**16.01**
HIPO	**20.60**	**29.81**	16.05

Table 2. SOTA comparison on **ScanNetv2** → **S3DIS** dataset. We compute the average of 10-4T and 10-8T with PointTransfomer backbone [77] in 5-shot setting. HIPO achieves superior **Last** and **AIA** values with comparable $\mathcal{F}^{(T)}_{Last}$ values, highlighting better domain adaptability than other methods.

ding space, which aligns naturally with hierarchical structures, thereby enhancing inter-class separation and intra-class compactness over time.

Moreover, Fig. 4 provides qualitative comparisons, where HIPO demonstrates marked improvements over leading baselines, particularly in accurately segmenting challenging structures such as walls and counters. Further analyses and results are available in the ***Sup. Mat.***

4.2. Main ablation analysis

Cross-dataset class-incremental segmentation: To further assess the effectiveness of our proposed modules, we conduct cross-dataset incremental segmentation experiments (ScanNetv2 → S3DIS). Detailed setup information is in ***Sup. Mat.*** As shown in Table 2, the cross-dataset setting introduces a notable increase in forgetting due to domain shift. However, the results underscore HIPO's robustness in managing such domain shifts effectively, with improvements of 0.65% in **Last** and 0.52% in **AIA** metrics.

Ablation studies on penalty in Busemann loss: We provide a comprehensive ablation study on the regularization constant ϕ introduced in Eq. 7. The regularization strength is varied from $\phi = 0$ to $\phi = 1$ in intervals of 0.25 for **S3DIS** [2] (4-3T, 7-3T, 10-3T) and **ScanNetv2** [17] (10-2T, 10-5T, 15-5T). As shown in Table 3, our findings corroborate prior work [23]: without a penalty ($\phi = 0$), samples are overly drawn towards the Ideal Prototypes, resulting in overconfident predictions and lower **AIA** on test data $\mathcal{D}_{test}$ in addition to vanishing gradient problems [44]. Conversely, a high penalty ($\phi = 1$) increases inter-class confusion, also leading to lower **AIA** on $\mathcal{D}_{test}$. Our final chosen results, highlighted in **bold** in Table 3, reflect the optimal balance for ϕ.

Comparison of metrics at different curvatures: Following [47], we experimented with curvature values of $c = 0.0001$ and $c = 0.001$ to implement traditional Euclidean architectures (Poincaré Ball with zero curvature; very small Poincaré ball curvature values are chosen as curvature $c = 0$ is numerically unstable [27]) for the **S3DIS** [2] (4-3T, 7-3T, 10-3T) and **ScanNetv2** [17] (10-2T, 10-5T, 15-5T) setups. Table 4 shows that our HIPO with a Poincaré Ball curvature of $c = -1.00$ consistently outperforms its Euclidean counterpart across most settings. This finding demonstrates that Riemannian geometry leverages the hierarchical structure of point clouds more effectively than Euclidean space, suggesting further exploration in this domain.

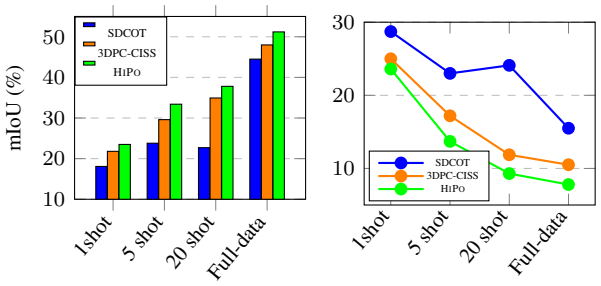
Sensitivity to the different number of training samples in each incremental episode: To assess the impact of data availability on model performance, we gradually increase the number of samples per class (shot) from one to all available examples and analyze the **Last** mIoU and $\mathcal{F}^{(T)}_{Last}$ results for different shot values in Fig. 5. Our findings indi-

Datasets	Settings	$\phi = 0.00$		$\phi = 0.25$		$\phi = 0.50$		$\phi = 0.75$		$\phi = 1.00$	
		AIA↑	$\mathcal{F}^{(T)}_{Last}$↓	**AIA**↑	$\mathcal{F}^{(T)}_{Last}$↓	**AIA**↑	$\mathcal{F}^{(T)}_{Last}$↓	**AIA**↑	$\mathcal{F}^{(T)}_{Last}$↓	**AIA**↑	$\mathcal{F}^{(T)}_{Last}$↓
S3DIS [2]	**4-3T**	42.87	17.70	43.40	17.86	43.27	18.23	**43.87**	**17.70**	43.22	18.96
	7-3T	36.13	11.60	36.46	11.20	**36.80**	**11.10**	36.70	11.10	36.06	12.10
	10-3T	39.95	14.30	**40.25**	**13.70**	40.10	14.00	40.07	14.05	40.05	14.10
ScanNetv2 [17]	**10-2T**	16.75	10.51	16.85	10.50	16.90	10.08	**17.13**	**10.36**	16.91	10.76
	10-5T	27.08	18.87	27.13	18.87	**27.22**	**18.85**	27.21	18.85	27.16	18.92
	15-5T	24.25	25.50	**24.45**	**25.10**	24.05	25.90	23.95	26.10	24.00	26.00

Table 3. Different penalty values in Busemann loss in 5-shot incremental segmentation across the **S3DIS** [2] dataset and **ScanNetV2** [17] dataset, highlighting the regularization benefits introduced by the hyperbolic distance of the projected class prototypes from the origin.

Datasets	Settings	$c = -1.00$ **(Ours)**		$c = 0.0001$		$c = 0.001$	
		AIA↑	$\mathcal{F}^{(T)}_{Last}$↓	**AIA**↑	$\mathcal{F}^{(T)}_{Last}$↓	**AIA**↑	$\mathcal{F}^{(T)}_{Last}$↓
S3DIS [2]	**4-3T**	**43.87**	**17.70**	42.95	19.93	43.23	17.38
	7-3T	36.80	11.10	36.23	11.60	**36.86**	**11.05**
	10-3T	**40.25**	**13.70**	40.20	13.80	40.22	13.75
ScanNetv2 [17]	**10-2T**	**17.13**	10.36	16.54	10.25	16.98	**10.18**
	10-5T	**27.22**	**18.85**	27.16	19.00	26.83	19.92
	15-5T	**24.45**	**25.10**	24.40	25.20	24.20	25.60

Table 4. Comparison for different curvature values in 5-shot incremental segmentation across **S3DIS** [2] and **ScanNetV2** [17] datasets. This highlights the superior stability-plasticity trade-off of hyperbolic embedding space with negative curvature, where small positive curvature approximates Euclidean space.



(a) **Last mIoU values** (b) **Forgetting** $\mathcal{F}^{(T)}_{Last}$

Figure 5. **mIoU improves and forgetting reduces with more training examples**.Our proposed method outperforms the top-2 baseline methods ([70, 79]) in terms of Last mIoU and forgetting mitigation in the 10-2T setting (Full-data: fully supervised training in incremental fashion).

cate that catastrophic forgetting is particularly severe with limited data; however, our method substantially reduces forgetting compared to other baselines. Notably, increasing the number of training examples significantly mitigates forgetting for HiPo, while it exacerbates forgetting for SDCOT.

Experiments with a different backbone: Besides PointTransformer [77], we evaluate our model with the backbone of DGCNN [65]. Unlike current SOTA FSCIL methods—**BiDist** [78], **CLOM** [85], **TEEN** [64], **FACT** [81], and **ALICE** [50]—which perform well with DGCNN but degrade with PointTransformer, our method remains effective across both architectures. As shown in Table 5, our approach achieves consistent gains with DGCNN, outperforming baselines with $+2.20\%$ mIoU on **Last**, $+0.72\%$ on **AIA**, and reducing forgetting by $+1.98\%$ on $\mathcal{F}^{(T)}_{Last}$. This demonstrates the model-agnostic nature of HiPo.

Method	Last (↑)	AIA (↑)	$\mathcal{F}^{(T)}_{Last}$ (↓)
LwF [59]	9.45	26.75	24.41
C-FSCIL [24]	20.22	30.64	22.15
LGKD [72]	18.63	29.16	15.75
SDCOT [79]	22.75	24.54	19.94
CLOM [85]	21.18	34.67	18.83
TEEN [64]	22.71	33.90	15.37
FACT [81]	24.27	26.70	21.26
3DPC-CISS [70]	22.03	24.74	19.36
BiDist [78]	27.70	39.14	15.85
HiPo	**29.90**	**39.86**	**13.87**

Table 5. Performance comparison on the S3DIS dataset [2] using the DGCNN backbone [65] in the 5-shot setting. Reported metrics include **Last**, $\mathcal{F}^{(T)}_{Last}$, and **AIA** for 4-3T and 7-3T settings. HiPo outperforms all baseline methods with DGCNN.

5. Conclusions & Future Directions

We introduce a previously unexplored problem setting along with a novel solution specifically designed for few-shot class-incremental semantic segmentation of 3D point cloud data, addressing a critical gap in the literature. Recognizing the unique challenges posed by point clouds, we identified limitations in current FSCIL methods and developed HiPo, a model that leverages hyperbolic space to embed point clouds while achieving an improved stability-plasticity trade-off. Our approach consistently outperforms state-of-the-art baselines, demonstrating that hyperbolic geometry's ability to represent hierarchical structures offers a more effective framework for FSCIL in 3D data. HiPo is particularly valuable for real-world applications like autonomous driving, industrial inspection, urban planning, and environmental monitoring, where generalization to unseen classes is essential. Furthermore, our approach enhances safety and efficiency in critical areas such as disaster response and infrastructure maintenance by enabling accurate, timely semantic segmentation of complex 3D environments. Future work will focus on managing possible outliers that may emerge without prior reference across incremental sessions, further strengthening our robustness.

References

[1] Noor Ahmed, Anna Kukleva, and Bernt Schiele. Orco: Towards better generalization via orthogonality and contrast for few-shot class-incremental learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 28762–28771, 2024. 1, 3

[2] Iro Armeni, Ozan Sener, Amir R Zamir, Helen Jiang, Ioannis Brilakis, Martin Fischer, and Silvio Savarese. 3d semantic parsing of large-scale indoor spaces. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1534–1543, 2016. 2, 4, 6, 7, 8

[3] Mina Ghadimi Atigh, Julian Schoep, Erman Acar, Nanne Van Noord, and Pascal Mettes. Hyperbolic image segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4453–4462, 2022. 2

[4] Eden Belouadah and Adrian Popescu. Il2m: Class incremental learning with dual memory. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 583–592, 2019. 1

[5] Abhijit Bendale and Terrance Boult. Towards open world recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1893–1902, 2015. 3

[6] Chris Burges, Tal Shaked, Erin Renshaw, Ari Lazier, Matt Deeds, Nicole Hamilton, and Greg Hullender. Learning to rank using gradient descent. In *Proceedings of the 22nd International Conference on Machine Learning*, page 89–96, New York, NY, USA, 2005. Association for Computing Machinery. 5

[7] Herbert Busemann. *The geometry of geodesics*. Courier Corporation, 2005. 2, 3, 5

[8] Francisco M Castro, Manuel J Marín-Jiménez, Nicolás Guil, Cordelia Schmid, and Karteek Alahari. End-to-end incremental learning. In *Proceedings of the European conference on computer vision (ECCV)*, pages 233–248, 2018. 1

[9] Jun Cen, Peng Yun, Shiwei Zhang, Junhao Cai, Di Luan, Mingqian Tang, Ming Liu, and Michael Yu Wang. Open-world semantic segmentation for lidar point clouds. In *European Conference on Computer Vision*, pages 318–334. Springer, 2022. 3

[10] Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo, Elisa Ricci, and Barbara Caputo. Modeling the background for incremental learning in semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9233–9242, 2020. 6

[11] Arslan Chaudhry, Puneet K Dokania, Thalaiyasingam Ajanthan, and Philip HS Torr. Riemannian walk for incremental learning: Understanding forgetting and intransigence. In *Proceedings of the European conference on computer vision (ECCV)*, pages 532–547, 2018. 1

[12] Bike Chen, Wei Peng, Xiaofeng Cao, and Juha Röning. Hyperbolic uncertainty aware semantic segmentation. *IEEE Transactions on Intelligent Transportation Systems*, 2023. 5

[13] Ali Cheraghian, Shafin Rahman, Pengfei Fang, Soumava Kumar Roy, Lars Petersson, and Mehrtash Harandi. Semantic-aware knowledge distillation for few-shot class-incremental learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2534–2543, 2021. 1

[14] Townim Chowdhury, Mahira Jalisha, Ali Cheraghian, and Shafin Rahman. Learning without forgetting for 3d point cloud objects. In *Advances in Computational Intelligence: 16th International Work-Conference on Artificial Neural Networks, IWANN 2021, Virtual Event, June 16–18, 2021, Proceedings, Part I 16*, pages 484–497. Springer, 2021. 3

[15] Townim Chowdhury, Ali Cheraghian, Sameera Ramasinghe, Sahar Ahmadi, Morteza Saberi, and Shafin Rahman. Few-shot class-incremental learning for 3d point cloud objects. In *European Conference on Computer Vision*, pages 204–220. Springer, 2022. 3

[16] Yawen Cui, Zitong Yu, Wei Peng, Qi Tian, and Li Liu. Rethinking few-shot class-incremental learning with open-set hypothesis in hyperbolic geometry. *IEEE Transactions on Multimedia*, 2023. 2, 3

[17] Angela Dai, Angel X Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Nießner. Scannet: Richly-annotated 3d reconstructions of indoor scenes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5828–5839, 2017. 2, 6, 7, 8

[18] David L. Davies and Donald W. Bouldin. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1(2):224–227, 1979. 2

[19] Jiahua Dong, Yang Cong, Gan Sun, Bingtao Ma, and Lichen Wang. I3dol: Incremental 3d object learning without catastrophic forgetting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 6066–6074, 2021. 3

[20] DC Dowson and BV666017 Landau. The fréchet distance between multivariate normal distributions. *Journal of multivariate analysis*, 12(3):450–455, 1982. 4

[21] Luca Franco, Paolo Mandica, Konstantinos Kallidromitis, Devin Guillory, Yu-Teng Li, Trevor Darrell, and Fabio Galasso. Hyperbolic active learning for semantic segmentation under domain shift. *arXiv preprint arXiv:2306.11180*, 2023. 5

[22] Octavian-Eugen Ganea, Gary Bécigneul, and Thomas Hofmann. Hyperbolic convolutional neural networks. In *Advances in Neural Information Processing Systems*, pages 4851–4861, 2018. 4

[23] Mina Ghadimi Atigh, Martin Keller-Ressel, and Pascal Mettes. Hyperbolic busemann learning with ideal prototypes. *Advances in Neural Information Processing Systems*, 34:103–115, 2021. 3, 5, 7

[24] Michael Hersche, Geethan Karunaratne, Giovanni Cherubini, Luca Benini, Abu Sebastian, and Abbas Rahimi. Constrained few-shot class-incremental learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9057–9067, 2022. 1, 3, 6, 7, 8

[25] Naiwen Hu, Haozhe Cheng, Yifan Xie, Pengcheng Shi, and Jihua Zhu. Hyperbolic image-and-pointcloud contrastive learning for 3d classification. *arXiv preprint arXiv:2409.15810*, 2024. 2, 3

[26] Qingyong Hu, Bo Yang, Linhai Xie, Stefano Rosa, Yulan Guo, Zhihua Wang, Niki Trigoni, and Andrew Markham.

Randla-net: Efficient semantic segmentation of large-scale point clouds. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11108–11117, 2020. 1, 2

[27] Valentin Khrulkov, Leyla Mirvakhabova, Evgeniya Ustinova, Ivan Oseledets, and Victor Lempitsky. Hyperbolic image embeddings. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 6418–6428, 2020. 2, 3, 5, 7

[28] Do-Yeon Kim, Dong-Jun Han, Jun Seo, and Jaekyun Moon. Warping the space: Weight space rotation for class-incremental few-shot learning. In *The Eleventh International Conference on Learning Representations*, 2023. 1

[29] Max Kochurov, Rasul Karimov, and Serge Kozlukov. Geoopt: Riemannian optimization in pytorch, 2020. 6

[30] Anna Kukleva, Hilde Kuehne, and Bernt Schiele. Generalized and incremental few-shot learning by explicit learning and calibration without forgetting. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 9020–9029, 2021. 1

[31] Junnan Li, Dongxu Li, Caiming Xiong, and Steven Hoi. Blip: Bootstrapping language-image pre-training for unified vision-language understanding and generation. In *International conference on machine learning*, pages 12888–12900. PMLR, 2022. 5

[32] Zhizhong Li and Derek Hoiem. Learning without forgetting. *IEEE transactions on pattern analysis and machine intelligence*, 40(12):2935–2947, 2017. 1

[33] Fangzhou Lin, Yun Yue, Songlin Hou, Xuechu Yu, Yajun Xu, Kazunori D Yamada, and Ziming Zhang. Hyperbolic chamfer distance for point cloud completion. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 14595–14606, 2023. 1

[34] Pei Lin, Zheng Cai, Jiangtao Zhang, and Lifang Ruan. A survey on point cloud segmentation: From classical methods to deep learning. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 16(3): 1–27, 2020. 1

[35] Yuyang Liu, Yang Cong, Gan Sun, Tao Zhang, Jiahua Dong, and Hongsen Liu. L3doc: Lifelong 3d object classification. *IEEE Transactions on Image Processing*, 30:7486–7498, 2021. 3

[36] Zan Liu, Zilong Zhang, Yiyi Lin, Lin Zhu, Han Jin, and Stephen Lin. Pointaugment: An auto-augmentation framework for point cloud classification. *arXiv preprint arXiv:2002.10876*, 2020. 1

[37] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017. 6

[38] Paolo Mandica, Luca Franco, Konstantinos Kallidromitis, Suzanne Petryk, and Fabio Galasso. Hyperbolic learning with multimodal large language models. *arXiv preprint arXiv:2408.05097*, 2024. 2

[39] Pratik Mazumder, Pravendra Singh, and Piyush Rai. Few-shot lifelong learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 2337–2345, 2021. 1

[40] Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018. 2

[41] Pascal Mettes, Elise Van der Pol, and Cees Snoek. Hyperspherical prototype networks. *Advances in neural information processing systems*, 32, 2019. 5

[42] Pascal Mettes, Mina Ghadimi Atigh, Martin Keller-Ressel, Jeffrey Gu, and Serena Yeung. Hyperbolic deep learning in computer vision: A survey. *International Journal of Computer Vision*, pages 1–25, 2024. 2

[43] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 26, 2013. 5

[44] Gal Mishne, Zhengchao Wan, Yusu Wang, and Sheng Yang. The numerical stability of hyperbolic representation learning. In *International Conference on Machine Learning*, pages 24925–24949. PMLR, 2023. 2, 4, 5, 7

[45] Antonio Montanaro, Diego Valsesia, and Enrico Magli. Rethinking the compositionality of point clouds through regularization in the hyperbolic space. *Advances in Neural Information Processing Systems*, 35:33741–33753, 2022. 1, 2, 3, 4

[46] Antonio Montanaro, Diego Valsesia, and Enrico Magli. Towards hyperbolic regularizers for point cloud part segmentation. In *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5, 2023. 1

[47] Gabriel Moreira, Manuel Marques, João Paulo Costeira, and Alexander Hauptmann. Hyperbolic vs euclidean embeddings in few-shot learning: Two sides of the same coin. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 2082–2090, 2024. 7

[48] Samrat Mukherjee, Tanuj Sur, Saurish Seksaria, Subhasis Chaudhuri, Gemma Roig, and Biplab Banerjee. Uidaple: Unsupervised incremental domain adaptation through adaptive prompt learning. In *ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5, 2025. 1

[49] Maximilian Nickel and Douwe Kiela. Poincaré embeddings for learning hierarchical representations. In *Advances in neural information processing systems*, pages 6338–6347, 2017. 2, 3

[50] Can Peng, Kun Zhao, Tianren Wang, Meng Li, and Brian C Lovell. Few-shot class-incremental learning from an open-set perspective. In *European Conference on Computer Vision*, pages 382–397. Springer, 2022. 3, 6, 8

[51] Jeffrey Pennington, Richard Socher, and Christopher Manning. GloVe: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543, Doha, Qatar, 2014. Association for Computational Linguistics. 5

[52] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 652–660, 2017. 1, 2

[53] Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J Guibas. Pointnet++: Deep hierarchical feature learning on

point sets in a metric space. *Advances in neural information processing systems*, 30, 2017. 1, 2

[54] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In *International conference on machine learning*, pages 8748–8763. PMLR, 2021. 5

[55] Sachin Ravi and Hugo Larochelle. Optimization as a model for few-shot learning. *arXiv preprint arXiv:1606.04474*, 2017. 1

[56] Sylvestre-Alvise Rebuffi, Alexander Kolesnikov, Georg Sperl, and Christoph H Lampert. icarl: Incremental classifier and representation learning. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pages 2001–2010, 2017. 1

[57] Rik Sarkar. Low distortion delaunay embedding of trees in hyperbolic plane. In *International symposium on graph drawing*, pages 355–366. Springer, 2011. 2, 3

[58] Yisheng Song, Ting Wang, Puyu Cai, Subrota K Mondal, and Jyoti Prakash Sahoo. A comprehensive survey of few-shot learning: Evolution, applications, challenges, and opportunities. *ACM Computing Surveys*, 55(13s):1–40, 2023. 1

[59] Filip Szatkowski, Mateusz Pyla, Marcin Przewieźlikowski, Sebastian Cygert, Bartłomiej Twardowski, and Tomasz Trzciński. Adapt your teacher: Improving knowledge distillation for exemplar-free continual learning. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 1977–1987, 2024. 6, 7, 8

[60] Yuwen Tan and Xiang Xiang. Cross-domain few-shot incremental learning for point-cloud recognition. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 2307–2316, 2024. 6

[61] Xiaoyu Tao, Xiaopeng Hong, Xinyuan Chang, Songlin Dong, Xing Wei, and Yihong Gong. Few-shot class-incremental learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 12183–12192, 2020. 3

[62] Jesper E Van Engelen and Holger H Hoos. A survey on semi-supervised learning. *Machine learning*, 109(2):373–440, 2020. 1

[63] Max van Spengler, Erwin Berkhout, and Pascal Mettes. Poincare resnet. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5419–5428, 2023. 4, 5

[64] Qi-Wei Wang, Da-Wei Zhou, Yi-Kai Zhang, De-Chuan Zhan, and Han-Jia Ye. Few-shot class-incremental learning via training-free prototype calibration. *Advances in Neural Information Processing Systems*, 36, 2024. 6, 8

[65] Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E Sarma, Michael M Bronstein, and Justin M Solomon. Dynamic graph cnn for learning on point clouds. *ACM Transactions on Graphics (tog)*, 38(5):1–12, 2019. 1, 2, 6, 8

[66] Yaqing Wang, Quanming Yao, James T Kwok, and Lionel M Ni. Generalizing from a few examples: A survey on few-shot learning. *ACM computing surveys (csur)*, 53(3):1–34, 2020. 1

[67] Wenxuan Wu, Zhongang Qi, and Li Fuxin. Pointconv: Deep convolutional networks on 3d point clouds. In *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition*, pages 9621–9630, 2019. 1, 2

[68] Yating Xu, Conghui Hu, Na Zhao, and Gim Hee Lee. Generalized few-shot point cloud segmentation via geometric words. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 21506–21515, 2023. 1

[69] Xiangli Yang, Zixing Song, Irwin King, and Zenglin Xu. A survey on deep semi-supervised learning. *IEEE Transactions on Knowledge and Data Engineering*, 35(9):8934–8954, 2022. 1

[70] Yuwei Yang, Munawar Hayat, Zhao Jin, Chao Ren, and Yinjie Lei. Geometry and uncertainty-aware 3d point cloud class-incremental semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 21759–21768, 2023. 6, 7, 8

[71] Zhilu Yang, Ying Zheng, and Jun Zhang. Learning to learn from labeled and unlabeled data with point-wise mutual information. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 5704–5711, 2019. 1

[72] Ze Yang, Ruibo Li, Evan Ling, Chi Zhang, Yiming Wang, Dezhao Huang, Keng Teck Ma, Minhoe Hur, and Guosheng Lin. Label-guided knowledge distillation for continual semantic segmentation on 2d images and 3d point clouds. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 18601–18612, 2023. 1, 3, 6, 7, 8

[73] Lewei Yao, Runhui Huang, Lu Hou, Guansong Lu, Minzhe Niu, Hang Xu, Xiaodan Liang, Zhenguo Li, Xin Jiang, and Chunjing Xu. Filip: Fine-grained interactive language-image pre-training. *arXiv preprint arXiv:2111.07783*, 2021. 5

[74] Yun Yue, Fangzhou Lin, Guanyi Mou, and Ziming Zhang. Understanding hyperbolic metric learning through hard negative sampling. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 1891–1903, 2024. 2

[75] Maciej Zamorski, Michał Stypułkowski, Konrad Karanowski, Tomasz Trzciński, and Maciej Zieba. Continual learning on 3d point clouds with random compressed rehearsal. *Computer Vision and Image Understanding*, 228: 103621, 2023. 3

[76] Chi Zhang, Nan Song, Guosheng Lin, Yun Zheng, Pan Pan, and Yinghui Xu. Few-shot incremental learning with continually evolved classifiers. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 12455–12464, 2021. 2, 3

[77] Hengshuang Zhao, Li Jiang, Jiaya Jia, Philip HS Torr, and Vladlen Koltun. Point transformer. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 16259–16268, 2021. 2, 6, 7, 8

[78] Linglan Zhao, Jing Lu, Yunlu Xu, Zhanzhan Cheng, Dashan Guo, Yi Niu, and Xiangzhong Fang. Few-shot class-incremental learning via class-aware bilateral distillation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11838–11847, 2023. 1, 3, 6, 8

[79] Na Zhao and Gim Hee Lee. Static-dynamic co-teaching for class-incremental 3d object detection. In *Proceedings of*

the AAAI Conference on Artificial Intelligence, pages 3436–3445, 2022. 1, 3, 6, 7, 8

[80] Na Zhao, Tat-Seng Chua, and Gim Hee Lee. Few-shot 3d point cloud semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8873–8882, 2021. 2

[81] Da-Wei Zhou, Fu-Yun Wang, Han-Jia Ye, Liang Ma, Shiliang Pu, and De-Chuan Zhan. Forward compatible few-shot class-incremental learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9046–9056, 2022. 1, 2, 3, 6, 8

[82] Da-Wei Zhou, Zi-Wen Cai, Han-Jia Ye, De-Chuan Zhan, and Ziwei Liu. Revisiting class-incremental learning with pre-trained models: Generalizability and adaptivity are all you need. *arXiv preprint arXiv:2303.07338*, 2023. 2

[83] Da-Wei Zhou, Qi-Wei Wang, Zhi-Hong Qi, Han-Jia Ye, De-Chuan Zhan, and Ziwei Liu. Class-incremental learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024. 2

[84] Yudong Zhu, Di Zhou, Jinghui Xiao, Xin Jiang, Xiao Chen, and Qun Liu. Hypertext: Endowing fasttext with hyperbolic geometry. *arXiv preprint arXiv:2010.16143*, 2020. 2

[85] Yixiong Zou, Shanghang Zhang, Yuhua Li, and Ruixuan Li. Margin-based few-shot class-incremental learning with class-level overfitting mitigation. *Advances in neural information processing systems*, 35:27267–27279, 2022. 1, 3, 6, 8