

2025 届研究生硕士学位论文

分类号: _____

学校代码: _____ 10269

密 级: _____

学 号: _____ 51255902087



華東師範大學

East China Normal University

硕士学位论文

MASTER'S DISSERTATION

基于双曲流形映射的点云持续学习方法研究

院 系:	软件工程学院
专 业:	软件工程
研 究 方 向:	持续学习
指 导 教 师:	魏宪
学位申请人:	俞津

2025 年 3 月

Dissertation for Master's Degree in 2025

University Code: 10269

Student ID: 51255902087

EAST CHINA NORMAL UNIVERSITY

Research of Continual Learning Method Based on Hyperbolic Manifold Mapping for Point Cloud

Department:	Software Engineering Institution
Major:	Software Engineering
Research Direction:	Continual Learning
Supervisor:	Xian Wei
Candidate:	Jin Yu

March, 2025

华东师范大学学位论文原创性声明

郑重声明：本人呈交的学位论文《基于双曲流形映射的点云持续学习方法研究》，是在华东师范大学攻读硕士/博士（请勾选）学位期间，在导师的指导下进行的研究工作及取得的研究成果。除文中已经注明引用的内容外，本论文不包含其他个人已经发表或撰写过的研究成果。对本文的研究做出重要贡献的个人和集体，均已在文中作了明确说明并表示谢意。

作者签名：_____

日期： 年 月 日

华东师范大学学位论文著作权使用声明

《基于双曲流形映射的点云持续学习方法研究》系本人在华东师范大学攻读学位期间在导师指导下完成的硕士/博士（请勾选）学位论文，本论文的研究成果归华东师范大学所有。本人同意华东师范大学根据相关规定保留和使用此学位论文，并向主管部门和相关机构如国家图书馆、中信所和“知网”送交学位论文的印刷版和电子版；允许学位论文进入华东师范大学图书馆及数据库被查阅、借阅；同意学校将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于（请勾选）

- ☐ 1. 经华东师范大学相关部门审查核定的“内部”或“涉密”学位论文*，
于 年 月 日解密，解密后适用上述授权。
- ☐ 2. 不保密，适用上述授权。

导师签名：_____

本人签名：_____

年 月 日

* “涉密”学位论文应是已经华东师范大学学位评定委员会办公室或保密委员会审定过的学位论文（需附获批的《华东师范大学研究生申请学位论文“涉密”审批表》方为有效），未经上述部门审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权）。

摘 要

点云数据在自动驾驶、机器人视觉、三维重建等领域具有广泛的应用，其高维非结构化特性使得深度学习模型在处理点云任务时面临诸多挑战。特别是在持续学习场景下，模型需要在不断接受新任务的同时保持对旧知识的记忆，但由于点云数据的动态性和复杂性，传统方法在增量学习过程中灾难性遗忘依旧明显。现有的大多数持续学习方法依赖于欧氏空间建模，难以有效捕捉点云数据的层级化特征，影响模型的长期学习能力和泛化能力。针对这一问题，本文围绕双曲流形映射，以重放为核心思想提出了两种点云持续学习方法，提升模型的稳定性和适应性。

首先，本文研究了基于双曲流形映射的缓存重放方法，该方法利用双曲几何的紧致性优化点云数据特征，并通过贪心采样策略优先选择更关键的重放缓存样本，以提高模型对旧任务的记忆能力。实验结果表明，该方法在有限存储条件下能够有效减少遗忘效应，提升点云分类任务的持续学习性能。

其次，为了应对部分存储受限场景，本文进一步提出基于生成式模型的双曲映射点云持续学习方法，该方法结合扩散模型，在双曲空间中生成符合流形分布的点云样本，以此代替传统的缓存重放，从而降低存储需求。同时，本文设计了在双曲空间下扩散模型的全过程知识蒸馏策略及生成样本评分筛选机制，从而进一步优化持续学习任务。实验结果显示，该方法下的模型表现出良好的持续生成能力，并在多个主流数据集 (ModelNet40、ShapeNet、ScanNet) 的分类任务上均取得了优异表现。

本文的研究表明，双曲流形映射在点云持续学习中的有效性，且提出的两种方案在不同环境下各有优势，并为后续研究提供了新的方向。

关键词：点云持续学习、双曲流形映射、灾难性遗忘、缓存重放、生成式重放

ABSTRACT

Point cloud data has extensive applications in autonomous driving, robotic vision, and 3D reconstruction. However, its high-dimensional and unstructured nature poses significant challenges for deep learning models in handling point cloud tasks. Particularly in the continual learning scenario, models must continuously acquire new tasks while retaining knowledge from previous ones. However, due to the dynamic and complex nature of point cloud data, traditional methods still suffer from catastrophic forgetting in incremental learning. Most existing continual learning approaches rely on Euclidean space modeling, which struggles to effectively capture the hierarchical structure of point cloud data, thereby limiting the model's long-term learning capability and generalization ability. To address this issue, this paper proposes two point cloud continual learning methods based on hyperbolic manifold mapping, with replay as the core idea, to enhance model stability and adaptability.

Firstly, this paper investigates a cache-based replay method using hyperbolic manifold mapping. By leveraging the compactness of hyperbolic geometry, this method optimizes point cloud feature representation and employs a greedy sampling strategy to prioritize more critical replay buffer samples, thereby improving the model's ability to retain knowledge from previous tasks. Experimental results demonstrate that this method effectively mitigates forgetting under limited storage conditions and enhances the continual learning performance in point cloud classification tasks.

Secondly, to address scenarios with severe storage constraints, this paper further proposes a generative model-based hyperbolic mapping continual learning method for point clouds. By integrating a diffusion model, this approach generates point cloud samples in hyperbolic space that conform to the underlying manifold distribution, replacing traditional cache-based replay and thus reducing storage requirements. Meanwhile, we design a full-process knowledge distillation strategy for the diffusion model in hyperbolic space,

along with a sample scoring and selection mechanism for generated data, thereby further optimizing the continual learning task. Experimental results show that models trained with this method exhibit strong continual generation capabilities and achieve superior performance on classification tasks across multiple benchmark datasets, including ModelNet40, ShapeNet, and ScanNet.

This study demonstrates the effectiveness of hyperbolic manifold mapping in continual learning for point cloud data. The two proposed approaches each exhibit advantages in different environments and offer new directions for future research.

Keywords: *continual learning of point cloud, hyperbolic manifold mapping, catastrophic forgetting, cache-based replay, generative replay*

目 录

摘要	i
Abstract	iii
图目录	vii
表目录	ix
第一章 绪 论	1
1.1 研究背景	1
1.2 挑战	2
1.3 研究意义	3
1.4 本文主要工作	4
1.5 本文组织结构	5
第二章 相关工作综述	7
2.1 持续学习	7
2.1.1 基于重放的方法	8
2.1.2 基于正则化的方法	11
2.1.3 基于模型架构的方法	12
2.1.4 基于优化方式的方法	13
2.1.5 基于表示学习的方法	13
2.2 点云与持续学习	14
2.2.1 点云的深度学习发展	14
2.2.2 点云的持续学习方法	16
第三章 基于双曲流形映射的点云持续学习缓存重放方法研究	19
3.1 持续学习任务设置	20
3.2 双曲空间投影	20

3.3	双曲流形缓存重放	22
3.4	实验及分析	25
3.4.1	实验设置	25
3.4.2	数据集介绍	26
3.4.3	结果分析	27
3.5	章节小结	29
第四章	基于生成式模型的双曲映射点云持续学习方法研究	31
4.1	基于双曲空间的扩散模型点云生成	32
4.1.1	传统扩散模型	32
4.1.2	点云的扩散模型生成	34
4.2	基于双曲空间的生成式重放点云持续学习	36
4.2.1	持续学习任务设置	36
4.2.2	评估指标	36
4.2.3	分类器和评分机制	39
4.2.4	生成模型在双曲空间下的全过程知识蒸馏	40
4.3	实验及分析	43
4.3.1	实验设置	43
4.3.2	数据集介绍	44
4.3.3	结果分析	44
4.4	章节小结	51
第五章	总结与展望	53
5.1	研究工作总结	53
5.2	研究工作展望	54
参考文献	57
致谢	69
攻读硕士学位期间科研情况	71

图目录

图 1.1 本文组织架构图。	4
图 2.1 持续学习相关研究。	8
图 3.1 双曲流形扩张淘汰算法图解。	23
图 3.2 HyMR 模型整体架构。	25
图 3.3 ModelNet 数据集预览。	27
图 3.4 ShapeNet 数据集预览。	28
图 3.5 类增量学习设置下，在完成 t 个任务后的整体平均准确率，训练和验证期间，每个任务的类别信息是未知的。图中展示了 HyMR(双曲映射) 与其他基于欧式空间的方法在 ModelNet40 和 ShapeNet 数据集上训练 20 轮任务后的准确率变化。	30
图 3.6 任务增量学习设置下，完成 t 个任务后的整体平均准确率，训练和验证期间，每个任务的类别信息是已知的。图中展示了 HyMR(双曲映射) 与其他基于欧式空间的方法在 ModelNet40 和 ShapeNet 数据集上训练 20 轮任务后的准确率变化。	30
图 4.1 HyperDiffIL 模型架构。	42
图 4.2 ScanNet 数据集预览。	44
图 4.3 类增量学习设置下，在完成若干个任务后在 ModelNet40 数据集上的整体平均准确率。	45
图 4.4 类增量学习设置下，在完成若干个任务后在 ShapeNet 数据集上的整体平均准确率。	46
图 4.5 类增量学习设置下，在完成若干个任务后在 ScanNet 数据集上的整体平均准确率。	47

图 4.6 经过四轮单类别增量任务后, 不同用于生成模型的持续学习方法的生成结果可视化对比。	48
图 4.7 消融实验-双曲映射和分类器引导在 ModelNet40 数据集上完成若干个任务后类增量任务的整体平均准确率。	49
图 4.8 消融实验-双曲映射和分类器引导在 ShapeNet 数据集上完成若干个任务后类增量任务的整体平均准确率。	50

表目录

表 3.1 在类别增量学习中, HyMR(双曲映射) 与其他在欧式空间的不同方法 在 ModelNet40、ShapeNet 数据集上经过若干轮训练后的平均准确率 (%).	28
表 3.2 在任务增量学习中, HyMR(双曲映射) 与其他在欧式空间的不同方法 在 ModelNet40、ShapeNet 数据集上经过若干轮训练后的平均准确率 (%).	29
表 3.3 类增量任务在 ModelNet40 和 ShapeNet 数据集上, 对于不同消融方 案的 HyMR 方法, 经过 10 轮和 20 轮任务训练后的最终结果.	29
表 4.1 1-NNA 结果分析.	38
表 4.2 在类别增量学习中, 不同方法在 ModelNet40、ShapeNet 以及 ScanNet 数据集上经过若干轮训练后的平均准确率 (%).	45
表 4.3 经过四轮单类别增量任务后, 不同用于生成模型的持续学习方法的 生成能力指标对比.	49
表 4.4 在 ModelNet40 和 ShapeNet 数据集上完成若干轮任务后类增量任务 的整体平均准确率受全过程蒸馏和样本评分机制的影响.	51
表 4.5 各种消融方案对 HyperDiffIL 生成器性能的影响.	52

第一章 绪 论

1.1 研究背景

近年来，随着信息技术的迅猛发展和大数据时代的到来，各行业对智能化解决方案的需求持续增长，推动了人工智能技术的快速普及和深入研究。进而，人工智能技术在多个领域取得了令人瞩目的成就，从人脸识别、自动驾驶到医疗诊断、自然语言处理等 [1, 2, 3, 4]，无不体现了其强大的潜力和广泛的应用前景。在这一背景下，深度学习作为人工智能的核心技术之一，因其出色的自动特征提取能力和强大的数据建模能力，受到了学术界和工业界的高度关注。

深度学习模型依托其多层神经网络结构，能够从海量数据中自动提取层次化的特征表示，为解决复杂的任务提供了高效的方法。例如，在 ImageNet 图像识别挑战赛中，深度学习技术使得分类准确率从 2010 年的仅有 71.8% 到 2015 年提出残差网络 (ResNet) 后的 96.4% [5, 6]，这一优势使得深度学习在计算机视觉、语音识别和自然语言处理等领域取得了突破性进展。通过深度学习技术，语音识别的准确率已接近人类水平，根据微软研究院的报告，语音识别系统的词错误率从 21 世纪初的 15% 下降至远低于 10% [7]。自然语言处理领域也实现了从语法分析到语义理解的质变，比如谷歌的 BERT 模型在多项语言理解任务中达到了前所未有的高性能 [8]。与此同时，人工智能的研究领域也呈现出不断细化和扩展的趋势，学者们开始探索更适合复杂场景和任务的新型学习方法，以应对多样化数据带来的挑战，如在自动驾驶领域，基于多传感器融合的多模态端到端自动驾驶也成为了最近的热点之一 [9]，并且不断有新的突破，这进一步验证了深度学习在实际应用中的潜力。深度学习的崛起不仅拓宽了人工智能研究的广度，还极大提升了技术的实际应用价值，成为推动智能化社会发展的重要基石。

在各种数据类型中，3D 点云 (Point cloud) 数据作为一种常见的数据形式，近年来在计算机视觉和机器学习领域受到广泛关注。随着自动驾驶 [10]、机器人导航 [11] 和虚拟现实 [8] 等技术的发展，基于点云的深度学习方法得到了广泛应用。不过，值得注意的是这些实际应用场景普遍具有动态性强、任务更新频繁的特点。例

如,在自动驾驶系统中,车辆可能会运行于不同的城市、气候或昼夜环境中,每种环境下感知系统所需处理的场景类别和数据分布都可能发生变化,系统十分依赖于对多任务的持续学习能力 [12];服务型机器人在执行不同任务或被部署至新空间时,也需快速识别新的物体或场景布局 [13]。在这些现实场景中,训练数据往往以分阶段、小批量的形式不断到达,难以一次性收集完备。这就要求模型能够在有限数据和资源条件下不断适应新任务,同时避免遗忘已有知识,这些都体现了点云深度学习模型在部署阶段对长期学习能力的迫切需求。

然而,现有模型在学习新任务时往往会由于不断迭代导致旧任务知识的遗忘,这种现象被称为“灾难性遗忘”(catastrophic forgetting)[14]。其次,模型通常需要依赖大规模标注数据,而在许多场景中,标注成本高昂且数据无法一次性获得,往往是一批一批地获得。最后,随着任务复杂度的增加,模型参数、旧数据的存储和计算资源的需求也呈指数级增长,这给部署带来了额外挑战。

为了解决这些问题,持续学习 (Continual Learning) 成为近年来研究的热点方向。持续学习的核心目标是在面对动态数据流时,使模型能够逐步吸收新知识,同时有效保留和利用旧任务的知识。与传统的深度学习方法不同,持续学习注重模型的“稳定-可塑性平衡”(stability-plasticity balance),既要保持对旧任务的稳定性,又要具备对新任务的适应能力。通过引入记忆重放、参数隔离和知识蒸馏等策略,持续学习为解决灾难性遗忘问题提供了有效途径,为深度学习技术在动态环境下的应用开辟了新方向。

1.2 挑战

但是,目前持续学习绝大多数的研究主要聚焦于图像等二维数据 [15, 16, 17, 18],而对点云这样的三维数据的持续学习研究,相较于其他模态数据,仅仅处于起步阶段 [19, 20],相对较少。

此外,3D 点云分类是点云数据分类和识别中的一项重要任务,其主要挑战在于点云数据的不规则性和稀疏性。随着研究的不断深入,一系列有效的三维点云分类算法相继提出 [21, 22]。尽管这些方法在三维点云分类任务中取得了显著进展,但它们仍然存在一定的局限性。首先,目前大多数方法仅仅依赖于欧式空间进行

特征学习，而欧式空间在处理层级树状结构的数据时表现有限，难以有效地捕捉点云数据的复杂拓扑关系。此外，现有方法在局部特征表达、全局信息融合及对不规则点云的适应性方面仍有改进空间。因此，需要探索更加适合三维点云特征表示的空间，一些非欧空间的表征可能更为合适，例如双曲空间 (Hyperbolic Space)。

1.3 研究意义

本研究的核心目标是探索更加适合三维点云数据的表示空间，并提升模型的持续学习能力，从而缓解 3D 数据的灾难性遗忘问题。传统的深度学习方法在点云分类任务中取得了重要进展，但主要基于欧式空间进行特征学习，无法充分表征点云数据的层级结构，进而限制了模型的泛化能力和长期学习能力。考虑到三维点云天然具有树状层级结构，而双曲空间因其指数增长的嵌入能力能够更有效地表示层级数据，本研究尝试将双曲空间引入点云的持续学习框架，以优化点云的特征表达和知识保留机制。具体而言，本研究的意义体现在以下几个方面：

(1) 提升点云数据的特征表示能力：

由于双曲空间能够以低失真方式嵌入层级数据，本研究探索其在点云分类任务中的应用，有望提高模型对复杂点云拓扑结构的捕捉能力，从而提升分类性能和鲁棒性。

(2) 缓解点云数据灾难性遗忘，为三维视觉持续学习的研究提供思路：

传统的点云分类模型在增量学习过程中容易遗忘先前学习的任务。本研究通过结合双曲空间嵌入策略和知识蒸馏等持续学习方法，使模型能够在学习新任务的同时更好地保留旧任务知识，从而提高模型的长期学习能力。

(3) 降低计算和存储开销：

在持续学习中，使用传统方法存储大量旧任务数据会带来较高的计算和存储成本。双曲空间的紧凑表示特性能够减少存储需求，同时结合生成模型或重放机制，有助于在有限存储资源下实现高效的点云学习。

(4) 促进点云深度学习的实际应用：

在自动驾驶、机器人导航和智能监控等应用场景中，点云数据的分布和任务需求是动态变化的。本研究提出的方法可以增强模型对新环境的适应能力，提高

其在真实世界中的部署效果，为三维视觉任务提供更强的泛化能力。

1.4 本文主要工作

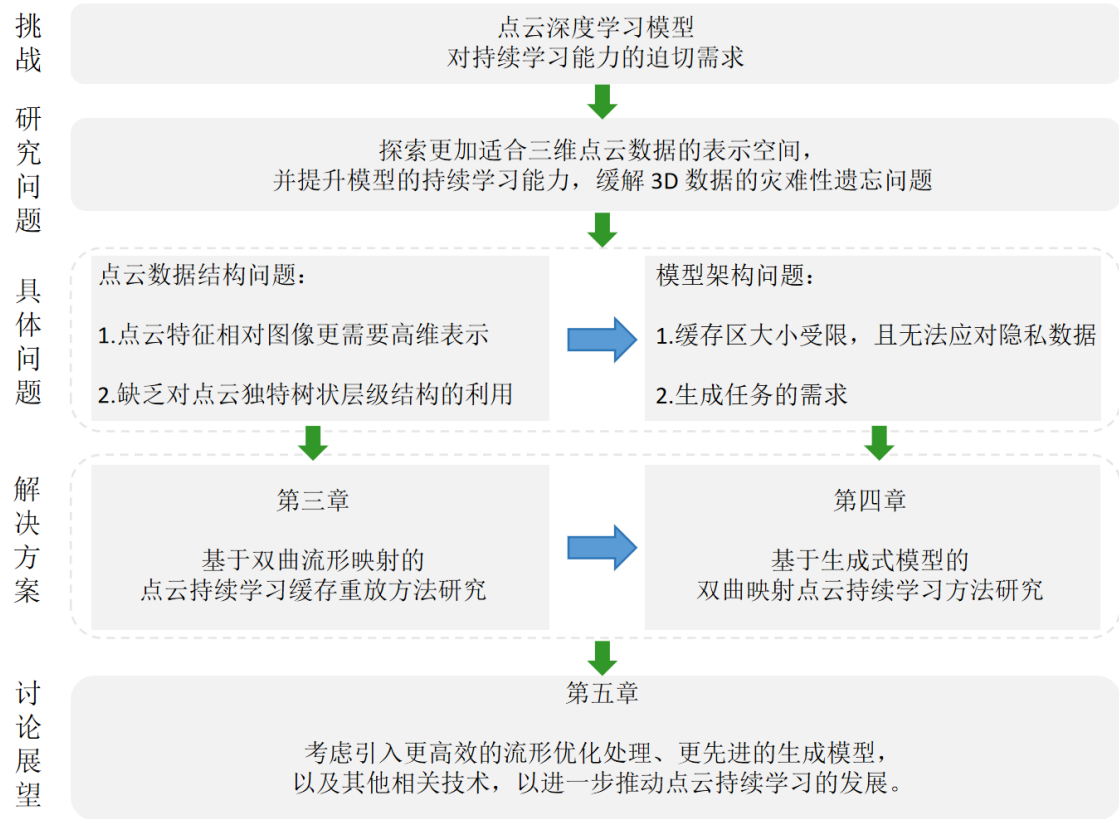


图 1.1 本文组织架构图。

针对点云数据在持续学习场景下的灾难性遗忘问题, 本文围绕双曲流形映射提出了两种点云持续学习方法, 以提升模型的稳定性和适应性, 主要贡献如下: 本文首先研究了基于双曲流形映射的缓存重放策略, 该方法利用双曲几何的紧致性优化点云数据特征, 并通过贪心采样策略优先选择更关键的重放缓存样本, 以提高模型对旧任务的记忆能力, 我们通过实验验证了该方法在主流点云数据集 ModelNet40、ShapeNet 上的有效性, 结果表明该方法能够在仅保留部分真实样本的情况下显著降低遗忘效应, 提升点云分类任务的持续学习性能; 其次, 为了进一步降低对缓存存储的依赖, 本文提出了一种基于扩散模型的双曲映射点云生成式重放方法, 该方法结合扩散模型 (Diffusion Model), 在双曲空间中生成符合流形分布的点云样

本，并结合全过程知识蒸馏策略，提升模型的抗遗忘能力，通过在 ModelNet40、ShapeNet、ScanNet 数据集上的实验，验证了该方法在无需直接存储旧样本的情况下，依然能有效缓解灾难性遗忘，显著提升点云分类和生成任务的持续学习性能。而通过实验，我们发现提出的两种方法应对不同环境各有优势，有助于推动点云持续学习领域后续的研究。

在创新性方面，可以将本课题的创新点总结如下：首先，点云的持续学习方法已有不少，但少有对点云数据的层级特征做研究，本文创新性引入双曲流形映射，将点云数据投影至双曲空间，以更紧凑的方式表示点云结构，提高特征提取能力；其次，重放方法中缓存区的选取策略也是一个影响性能的关键因素，因此我们设计并优化了一种基于贪心采样的缓存管理策略，在有限存储条件下优先保留对模型决策影响最大的样本，从而提升模型的长期学习能力；而针对一些存储受限的环境，我们引入扩散模型，并设计了一种双曲空间下的点云扩散生成框架，利用扩散过程生成高质量的点云样本，以替代传统缓存重放，减少存储需求；最后，对于这一生成式重放框架，我们提出双曲空间下的全过程知识蒸馏策略以及样本评分筛选机制，进一步提升模型的持续学习能力。

1.5 本文组织结构

本文共分为五个章节，如图1.1所示，具体安排如下：

第一章为绪论，首先介绍研究背景，讨论三维点云数据在计算机视觉与机器学习领域的应用价值，以及当前深度学习方法在点云分类任务中的挑战。随后，分析点云数据在持续学习中的特殊性，并探讨当前主流的重放策略与正则化策略所面临的局限性。最后，概述本文的研究内容、创新点及文章整体框架。

第二章为一些相关工作的研究进展，主要介绍与本文研究相关的基础理论和技术，包括持续学习的基本概念、主流的持续学习策略(如重放与正则化)及其应用局限，同时简单综述点云的深度学习和持续学习发展。

第三章主要介绍基于双曲流形映射的点云持续学习缓存重放方法研究，重点研究缓存重放策略在点云数据上的应用。首先，分析缓存重放策略的适用性，并指出其在知识存储和重放阶段面临的挑战。随后，从几何视角出发，基于双曲流形映

射构建新的数据选择和重放机制，以提升点云数据在重放过程中的表达能力。最后，实验部分分析验证方法的有效性。

第四章主要介绍基于生成式模型的双曲映射点云持续学习重放方法研究，探讨基于生成式模型的点云重放策略，以应对存储受限场景下的持续学习问题。首先，分析之前工作中重放方法的特点，并指出其局限性。接着，结合双曲空间的特性和知名的扩散模型，提出一种新的基于生成式模型的点云重放方法，使得模型能够在无需直接存储旧样本的情况下，生成符合双曲流形分布的点云数据，同时具备良好的泛化性。实验部分也对比了该方法与传统缓存重放策略的优劣，并评估其在长任务序列中的表现。

第五章为总结与展望，总结本文在点云持续学习领域的研究成果，涵盖双曲流形映射在缓存重放、生成式重放的应用，并分析其在不同场景下的优势与不足。最后，对未来可能的研究方向进行展望，包括更高效的流形优化技术、更先进的生成模型，以及可以考虑更适用于点云数据的 Prompt 设计，以进一步推动点云持续学习的发展。

第二章 相关工作综述

2.1 持续学习

学习是人类适应和生存的核心能力之一，使个体能够不断积累、更新并灵活应用各种知识，以应对外部环境的变化和挑战。人类的学习过程是长期且动态的，我们不仅能够在不同阶段获取新的知识，还能在此基础上进行推理、归纳和创新。然而，传统的机器学习方法往往是基于静态数据集进行训练，模型在训练完成后无法继续学习新知识，因而缺乏适应不断变化环境的能力。这种局限性使得模型在面对新的任务或数据分布时，往往需要重新训练，既增加了计算成本，也难以满足实际应用中的实时学习需求。因此，实现具备长期适应性和知识积累能力的人工智能系统，成为了当前机器学习研究的重要方向之一，这也促使了持续学习的兴起。

持续学习是一种旨在使人工智能系统能够在不断接收新任务或新数据的过程中，保持稳定性能的机器学习范式。其核心目标是让模型在学习新知识的同时，仍然能够有效保留并利用先前积累的知识，而不会因新任务的训练而遗忘旧任务的知识。在现实世界中，数据通常是连续且非独立同分布 (Non-IID) 的，模型需要具备处理新数据的能力，同时防止历史知识的丢失。这与人类的认知过程相似：我们在学习新技能或新概念时，往往能够在已有知识的基础上进行拓展，而不会完全遗忘之前掌握的内容。因此，持续学习被认为是迈向更强大、类人人工智能系统的重要一步。

然而，在神经网络中实现类似人类的持续学习能力并非易事。神经网络在按顺序学习多个任务时，往往会遭遇严重的知识干扰问题，即模型在学习新任务时，其权重更新可能会破坏对旧任务的记忆，导致性能大幅下降。这种被称为灾难性遗忘的现象，是持续学习研究中亟待解决的核心挑战之一。灾难性遗忘的根本原因在于，深度学习模型通常依赖梯度下降进行优化，当模型在新任务上进行训练时，模型的参数会调整以适应新数据，而这可能会由于冲突，导致覆盖先前学习到的知识。因此，如何在新旧任务之间取得平衡，使模型能够在持续学习的过程中兼顾稳定性和可塑性，成为持续学习研究的重要议题。

近年来，随着持续学习研究的不断推进，众多方法相继被提出并验证了其有效性。从整体思路来看，这些方法主要可归纳为五大类：(1) 基于重放的方法，通过存储或近似重建旧任务的数据，使模型在学习新知识的同时能够“复习”先前的信息，从而缓解灾难性遗忘；(2) 基于正则化的方法，该方法通过学习并继承部分旧模型的关键参数，以减少对过去知识的遗忘；(3) 基于模型架构的方法，通过优化神经网络的结构设计，使模型能够适应新任务的同时，尽可能减少对旧任务知识的干扰；(4) 基于表示学习的方法，通过构建鲁棒且分布合理的特征表示，使模型在新任务中能够更有效地复用过往知识，提升跨任务的迁移能力；(5) 基于优化方式的方法，通过调整优化策略(如梯度计算、参数更新机制等)，增强模型在持续学习过程中的稳定性和适应性。这些方法在实际应用的过程中，往往会相互交融结合使用，本小结将依次列举介绍这五类方法的核心思想、典型实现方式及其代表性研究成果。

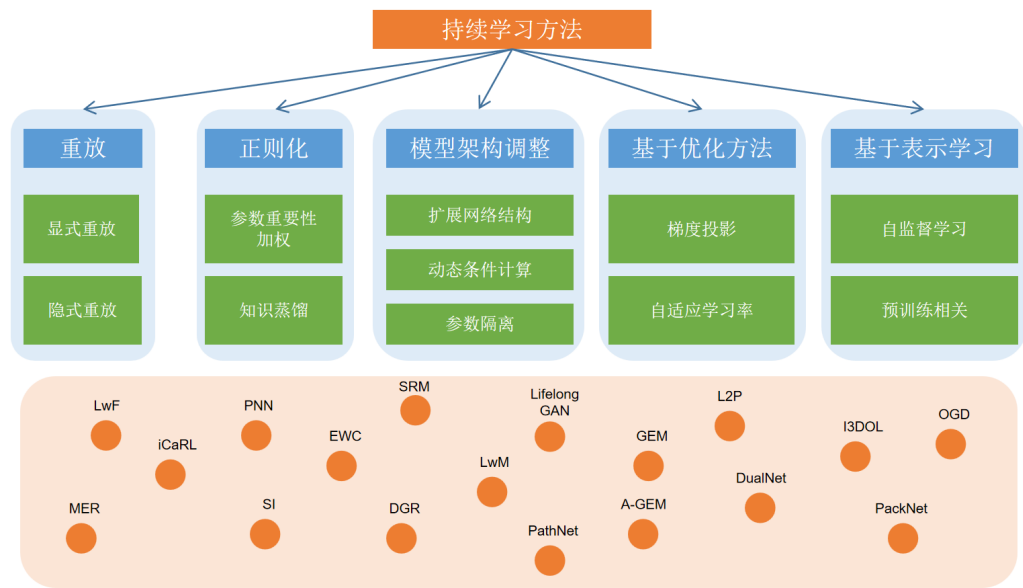


图 2.1 持续学习相关研究。

2.1.1 基于重放的方法

基于重放的方法核心思想是通过直接或间接地提供旧任务的信息，使模型在新旧任务之间保持平衡，研究者们通常面临存储成本、隐私、过拟合以及训练开销

的问题。根据重放方式的不同,主要可以分为显式重放 (Explicit Replay) 和隐式重放 (Implicit Replay)。

显式重放通常依赖一个存储模块,直接保存部分旧任务数据,并在训练新任务时将其混合训练,这一方式通常也被称为经验重放 (Experience Replay)。是持续学习中最常用的重放策略之一,其核心思想是开辟一块小的内存缓存区来存储一部分旧任务的数据,并在新任务训练时混合这些旧数据进行学习,以缓解灾难性遗忘。在这一过程中,由于存储空间的有限性,研究者们需要考虑如何高效利用缓存,针对性地选取、保留、更新旧样本,以强化模型对旧知识的回顾。经验重放最早由 Lin 于 1992 年在强化学习 (Reinforcement Learning, RL) 领域提出,用于稳定 Q-learning 训练过程 [23]。这一策略的基本做法是存储智能体的过去经验,并在训练时从存储池中随机抽取样本进行更新,以提高学习的稳定性。随着研究的深入,多个改进方案相继提出,以应对传统经验重放方法面临的挑战。

首先,经验重放中可以通过对旧样本进行约束,降低优化过程中对旧知识的干扰。例如, GEM[24] 是一种基于优化的经验重放方法,旨在通过约束梯度更新来缓解持续学习中的灾难性遗忘。在训练新任务时, GEM 会计算当前任务的梯度和旧任务的梯度,如果新任务的梯度方向会损害旧任务的性能,则通过投影技术调整梯度方向,使其既能学习新知识,又不会过度干扰旧知识。然而,由于其计算开销较高,在大规模数据集上可能不够高效。而后续基于此的改进方法 A-GEM[25] 则通过用一个“平均梯度”近似所有旧任务的梯度,并采用更简单的梯度投影策略来降低计算复杂度,解决了 GEM 的痛点。MER[26] 同样在经验重放的过程中对梯度计算方面做了优化处理,采用了元学习框架,通过多步优化让新任务的梯度方向更符合旧任务,从而提高知识迁移能力,并减少新任务对旧任务的干扰。相较于标准经验重放方法如 GEM, MER 通过元学习优化梯度,利用梯度对齐技术,提高了重放样本的利用率,使得新旧任务的学习更加兼容,有效减少灾难性遗忘,即使在较小的存储限制下,也能保持较好的性能。在缓存区处理方面, MIR[27] 通过优先选择那些受新任务影响最大的样本进行重放,即只重放最易遗忘的样本,而不是随机选择样本,提高了重放数据的利用率,从而最小化灾难性遗忘。

经验重放也可以和正则化方法结合起来,调整模型的参数,对抗灾难性遗忘。

例如, iCaRL[28] 采用知识蒸馏 (Knowledge Distillation, KD) 方法, 基于 KL 散度的蒸馏损失, 约束新模型的输出, 使其尽可能接近旧模型的预测, 以减少旧任务知识的丢失。LUCIR[29] 同样通过知识蒸馏的方法来保持旧类别知识, 并在此基础上提出采用基于余弦相似度的分类器 (最后一层), 从而减少类别不平衡影响。在神经网络最后一层做文章的还有 SS-IL[30], 通过伪标签 (Pseudo Labels) 机制来维持旧类别的知识, 并设计了一个平衡损失函数 (Balanced Loss), 对新老类别的训练数据进行均衡优化, 防止模型过度偏向新类别。近几年较为新颖的方法则有 ELI[31] 和 FOSTER[32] 等工作。ELI 提出了一个基于能量的潜在对齐器, 学习潜在表征的能量流形, 使先前的任务潜伏期具有较低的能量, 而当前的任务潜伏期具有较高的能量值, 通过由新旧模型学习到的能量流形来对抗在持续学习过程中发生的表征转移。FOSTER 则首先通过动态扩展新模块来拟合目标与旧模型之间的残差, 然后通过蒸馏策略删除冗余参数和特征维度, 以保持单一主干模型, 从而使模型能够自适应地学习新类别。

另一方面, 隐式重放则利用生成模型或特征存储机制, 近似重建旧任务的关键信息, 以减少存储开销和数据隐私风险, 常见的有生成式重放 (Generative Replay) 以及特征重放 (Feature Replay)。

生成式重放的核心思想是通过生成模型来产生旧任务的数据, 从而在学习新任务时模拟回顾旧任务的过程, 从而避免遗忘。生成式重放同样面临着一系列挑战, 诸如生成模型的稳定性和质量、生成模型与分类器的联合训练以及生成模型的复杂性和计算开销等问题。这一研究起步较晚, 但也有不少早期的工作为此奠定了基础。DGR[33] 是生成式重放方法的开创性工作之一, 通过引入变分自编码器 (VAE) 作为生成模型, 在学习新任务时通过生成旧任务数据来保持旧任务知识, 解放了显式存储旧任务数据需要的空间。由于自编码器生成的样本具备良好的稳定性, 可以精确生成需要的旧知识, 因此在诸多方法中都得到了广泛的应用, 如 SRM[34]、EEC[35] 等。而相比之下, 生成对抗模型 (GAN)[36] 凭借其出色的生成质量, 能够生成高细粒度的旧样本, 在复杂任务下更具优势, 同样备受研究者的青睐, 在 GFR[37]、MeRGAN[38] 等方法中发挥着出色的表现。此外, 研究者们还提出了类似 L-VAEGAN[39] 这样结合了自编码器的精确以及生成对抗模型的质量

的混合处理方法,各取所长来对抗灾难性遗忘。近年来,随着扩散模型 [40] 的不断发展,因其在图像生成、数据合成和概率建模方面的强大能力,逐渐成为生成式重放的新方向。扩散模型通过逐步向数据添加噪声并学习去噪的过程,能够生成更高质量、更稳定且更具多样性的样本,克服了 GAN 不稳定易模式崩塌和 VAE 生成模糊质量低的问题,在持续学习生成式重放领域中具有强大的潜力。因此,越来越多的研究探索如何将扩散模型引入持续学习,以提升生成式重放的有效性,提出了诸如 DDGR[41]、DiffClass[42] 等相关工作。

特征重放作为隐式重放的另一方向,作为一种更高效的重放方式,逐渐受到关注。该方法的核心思想是在神经网络的特征空间而非输入空间进行重放,即存储或合成旧任务的中间层特征,而不是原始数据。这种策略能够缓解存储压力,同时提升跨任务的知识迁移能力。如 K Zhu[43] 提出一个融合主枝扩展和侧枝更新的结构重组策略以维护旧特征,并在特征层面对新旧模型进行知识蒸馏,又如 FeTrIL[44] 使用了一个固定特征提取器,通过学习旧知识的特征,在新任务中重放旧特征,从而达到特征层面“复习”的目的。

2.1.2 基于正则化的方法

基于正则化的方法通过在优化过程中添加额外的约束,来减少模型在学习新任务时对旧任务知识的遗忘,通常不需要存储旧数据或生成样本,而是通过不同形式的正则化策略,在参数空间或损失函数上施加约束,以维持模型对旧任务的记忆能力。

参数重要性加权是正则化持续学习方法的早期研究主要聚焦方向,即引入一个惩罚损失函数,根据每个参数的变化和贡献来降低参数的剧变导致的灾难性遗忘。例如 EWC[17] 使用费舍尔信息矩阵来衡量参数的重要性,对关键参数施加二次惩罚; SI[16] 通过累积优化路径中的梯度变化来估计参数的重要性,并在损失函数中添加约束; RWalk[45] 则是结合了 EWC 和 SI 的优点,引入黎曼测度来更精确地衡量参数对旧任务的重要性,并动态调整参数更新策略; MAS[46] 以无监督和在线的方式利用数据在输出层的变化敏感性来评估参数的重要性,并据此进行正则化。

知识蒸馏方法同样深受研究者的追捧,通过对新任务的训练施加旧模型的软目标或特征表示,使新模型在学习新任务时仍能保持对旧任务的记忆。具体来说,方法通常采用教师-学生模型,将旧任务模型视为教师,新任务需要训练的模型视为学生,进行知识蒸馏,从而对抗灾难性遗忘。例如,LwF[15]使用旧模型的输出层的结果作为软目标,对新任务进行蒸馏约束,以保持模型对旧任务的认知;LwM[47]引入注意力蒸馏损失,并证明惩罚分类器注意力图的变化有助于在添加新的类别时保留基类的信息;LifelongGAN[48]同样采用知识蒸馏将学习到的知识从以前的网络转移到新网络,并在条件生成领域获得了出色的表现。

2.1.3 基于模型架构的方法

基于模型架构的方法通过对神经网络结构进行动态调整,以适应新任务的学习,并最大程度地减少对旧任务知识的遗忘。这类方法的核心思想是在模型架构层面引入任务适应性机制,例如使用扩展网络结构、动态条件计算和参数隔离等方法,从而有效应对灾难性遗忘问题。

持续学习模型可以拆分为任务间共享的部分和每个任务独立的部分,而扩展网络结构指的就是扩展和优化这些独立于任务的部分,比如RCM[49]将标准神经网络架构的卷积重新参数化为不可训练的共享部分(滤波器组)和任务特定部分(调制器),从而消除任务之间的不利相互作用;又如PNN[50]通过逐步增加新的神经网络模块来适应新任务的学习,同时通过冻结旧任务的网络权重来保持对旧任务的知识。动态计算则是指通过引入任务条件分支或门控机制,使模型在不同任务中采用不同的计算路径,如PathNet[51]则巧妙地设计了多个可并行训练的候选模块,为每个任务选择不同的路径,让不同任务只激活部分模块,降低了旧任务的干扰。随着任务的增加,对模型架构的修改往往会导致网络规模的不断扩大,而参数隔离则是利用任务特定的掩码、注意力机制或模块化设计,确保新任务的训练不会干扰旧任务,比较典型的代表是PackNet[52],通过网络剪枝为不同任务分配专属参数,从而避免了网络规模的无限增长。

2.1.4 基于优化方式的方法

基于优化方式的方法应对灾难性遗忘的方式相较之前在损失函数层面的操作有所不同，这类方法往往通过直接调整优化过程，使模型在学习新任务时最大程度保留旧任务的知识，从而减少灾难性遗忘。

梯度投影 (Gradient Projection) 和自适应学习率 (Adaptive Learning Rate) 是两种常用的优化方式的实现方法。梯度投影会在更新参数前，将梯度投影到一个特定子空间，以确保新任务的梯度更新不会破坏旧任务的知识，例如，OGD[53] 的投影子空间中，神经网络对前一个任务的输出不会改变，并且投射的梯度仍然在学习新任务的有用方向上，有效地利用了神经网络的高容量；而之前提到的 GEM[24] 和 A-GEM[25] 也同样采用了梯度投影的方法。自适应学习率则是根据参数在旧任务中的重要性，自适应地调整学习率，以减少对重要参数的影响，前文介绍的 RWalk[45] 方法同样利用到了这一点。

2.1.5 基于表示学习的方法

基于表示学习的方法通过学习稳健且分布良好的特征表示，优化模型的特征空间，使模型在任务之间具有更好的可迁移性，在适应新任务的同时，尽可能保持对旧任务知识的记忆。一般而言，可以通过自监督学习和与预训练相关的手段来实现表示学习的方法，从而降低灾难性遗忘的影响。

自监督学习是一种无监督学习的特殊形式，它通过让模型自主生成监督信号来进行训练，在持续学习领域中，一般采用引入对比损失的方法来实现。例如 Co2L[54] 通过构造正样本 (相同类别) 和负样本 (不同类别)，拉近正样本对的表示并推开负样本对的表示，利用对比损失来学习每个任务，并采用自监督损失来提取新旧模型之间的知识，从而使模型增强了对旧任务特征的保持能力，提高了对新任务的适应性。又如，LUMP[55] 在当前任务和以前的任务实例之间进行插值，以减轻无监督表示的灾难性遗忘。再如，DualNet[56] 十分巧妙地提出了一个包含快慢两种学习系统的框架用于表示学习，二者互补推进持续学习。

预训练 (Pretraining) 是一种深度学习方法，一般指在大规模数据集上训练模型，使其学习通用特征表示，然后在特定的下游任务上进行微调 (Fine-tuning)。预训练

的目标是让模型在无监督或自监督的环境下掌握数据的结构，从而在后续任务中表现更好。在持续学习中，在预训练模型的基础上对下游任务进行持续学习或者持续学习预训练模型，对抵抗灾难性遗忘都有巨大的帮助。例如，Beyza Ermis 等[57]提出基于著名的 Transformer[58]，在预训练的基础上扩展模型，从而无需重新训练模型，在保持良好的预测性能的同时，也兼顾了推理的时间。又如，L2P[59]利用预训练的嵌入层 (Embedding Layer) 和 Transformer 编码器，帮助模型从维护的提示词 (Prompt) 池中选取对应的提示词，以优化提示以指导模型预测；而由于提示词池中的内容作为一个很小的可学习参数，解放了重放方法中需要使用的缓存区所需紧张的内存空间，但最终获得了与重放方法相媲美的结果，因此作为一个新兴方向也吸引着大量研究者的目光。

2.2 点云与持续学习

随着近年来深度学习在三维数据处理领域的广泛应用，点云作为三维空间中物体和场景的主要表示形式，逐渐成为研究热点。由于点云数据的非结构化、稀疏性和不规则性，传统的卷积神经网络 (CNN)[60] 难以直接应用，这促使研究者探索专门针对点云的深度学习方法。近年来，点云深度学习取得了显著进展，大幅提升了点云特征提取和分类的能力。然而，点云深度学习在现实世界应用中也面临灾难性遗忘的挑战，其传统的深度学习模型在面对动态环境，尤其是自动驾驶和机器人感知等方向，同样需要持续学习的能力。又由于点云不同于传统图像等数据的特殊的结构，对其在线的增量和自适应学习方法提出了新的挑战。因此，本小结将会简单介绍一下点云的深度学习方面的发展和点云的一些典型持续学习方法，并列举介绍其核心思想和应用实现。

2.2.1 点云的深度学习发展

在 3D 视觉领域，点云深度学习主要用于解析和理解三维数据。根据具体应用场景，点云学习任务可以大致分为分类、分割、目标检测与跟踪三大类型。

首先是分类领域，即对点云对象进行分类，例如区分汽车、飞机、桌子等不同类别，这在计算机视觉、计算机图形学和工业设计中具有重要应用。从较为早期的

基于体素 (Voxel-based) 和多视图 (Multi-view) 的方法, 再到后来的直接操作点云数据的方法, 对点云分类领域都有着巨大的贡献。

基于体素的方法通常将点云体素化为三维网格, 然后在此之上利用 3D 卷积神经网络进行分类, 如 VoxNet[61] 通过将体积占用网格表示与监督式 3D 卷积神经网络相结合来进行点云分类; OctNet[62] 利用一组不平衡的八叉树结构分层划分一个点云, 每个叶节点存储一个池化特征表示, 从而支持深度和高分辨率的 3D 卷积神经网络; 3DmFV[63] 用紧凑且计算高效的方式将由输入点云转换的三维网格的离散结构与 Fisher 向量的连续泛化相结合, 最终通过传统的卷积神经网络架构学习全局表示, 在分类任务上获得了优秀的性能。

由于 3D 卷积神经网络相较 2D 在点云处理上计算开销较大, 而又由于后者在图像处理上具备优势, 因此核心思想为将点云投影为多个二维视图, 然后利用 2D 卷积神经网络进行特征提取, 最后融合不同视角的信息以完成分类任务的多视图方法在点云任务上也取得了良好的效果。例如, H Su 等 [64] 提出了一种新颖的卷积神经网络架构, 它将来自 3D 形状的多个视图的信息组合成一个紧凑的形状描述符, 从而提供更好的识别性能; Z Yang 等 [65] 提出了一个关系网络来有效地连接不同视点的相应区域, 强化单个视图图像的信息, 并整合这些视图以获得有区别的三维物体表示, 从而在三维物体识别和检索任务上发挥着出色的表现。

直接操作点云数据, 即直接在原始点云数据上进行特征提取和学习, 是近年来点云深度学习研究的一个重要方向, 避免了体素化导致的空间分辨率损失以及多视图方法丢失的几何信息。在这一方向上诞生了许多比较典型和著名的方法, 如 PointNet[21] 采用多层感知器 (MLP) 直接对每个点进行特征提取, 并通过最大池化 (Max-Pooling) 聚合全局信息, 并利用 T-Net 进行仿射变换, 解决了旋转不变性的难题; PointNet++[21] 在 PointNet 基础上, 增加了层级化特征提取, 通过局部区域采样 +KNN 算法进行特征聚合, 解决了 PointNet 忽略局部几何信息的问题; DGCNN[66] 采用 KNN 计算局部邻域, 并构造动态图, 通过图卷积提取局部特征; Point Transformer[67] 采用自注意力机制 (Self-Attention) 直接在点云上进行特征学习, 使模型能够捕捉远程依赖关系。

其次是分割领域, 目的是将点云数据划分为不同的类别或实例, 从而识别并

理解场景中的对象。根据分割粒度的不同,点云分割可以分为语义分割 (Semantic Segmentation)、实例分割 (Instance Segmentation)、部分分割 (Part Segmentation)。语义分割是指为每个点赋予一个语义类别,如“地面”、“行人”、“车辆”等,有诸如 A Boulch[68]、FJ Lawin[69] 等提出的基于投影的方法,也有如 J Huang[70]、L Tchapmi[71] 等提出的将点云转换为密集/稀疏的离散表示的方法,还有如之前提到的 PointNet、PointNet++ 等也可以用于语义分割,这些方法都推进了点云语义分割的不断发展。实例分割不仅需要进行语义分类,还要区分同一类别中的不同实例,如两辆相邻的汽车应该被分为不同的实例,在这一方向上,主要有基于建议 (Proposal-based) 的方法,将实例分割问题转换为三维目标检测和实例掩模预测两个子任务,如 Y Shi[72]、B Yang[65] 等人的工作;也有无建议 (Proposal-free) 的方法,不设置目标检测模块,而是通常将实例分割作为语义分割后的后续聚类步骤,例如 SGPN[73]、DRG[74] 等。而部分分割主要用于 3D 物体解析,目的是将对象的不同组成部分 (如汽车的车轮、座椅等) 分割出来,不仅需要进行语义分割,同一对象中相同组成部分的数量可能有所不同,也需要实例分割,在这一方向上,CoSegNet[75]、VoxSegNet[76] 等方法取得了出色的表现。

最后是目标检测与跟踪领域,这类任务可以拆解为两步,先是从点云数据中检测目标物体,并输出它们的 3D 边界框,然后在连续帧点云数据中跟踪目标物体,即在时间序列上建立物体的轨迹,确保在动态环境 (如自动驾驶) 下能够持续识别同一物体。J Ku[77]、M Liang[78]、S Shi[79]、S Giancola[80] 等研究者们的工作在点云的目标检测与追踪任务上做出了不小的贡献。

2.2.2 点云的持续学习方法

由于传统点云学习方法通常依赖静态、离线的大规模数据集进行训练,在面对动态环境、增量数据流、类别扩展等持续学习场景时,容易受到灾难性遗忘的影响,因此点云的持续学习领域的研究同样重要。相对于传统图像的持续学习,点云数据在持续学习中有着特殊难点,如点云表示的不变性、分布漂移 (Domain Shift)、模型容量限制、特殊结构等,因此这一方面的研究较少,但依然有着出色的工作取得了不小的成绩,并为后来者开辟新思路提供帮助。

T Chowdhury 等 [81] 使用对象类的语义词向量改进了蒸馏过程, 为 3D 数据的持续学习建立了新的基准结果。M Sadowski 等 [82] 提出一种自主分支构建 (ABC) 方法, 使用重建损失将新的训练示例映射到模型的正确分支, 填补了点云生成领域持续学习的空白。M Zamorski 等 [20] 巧妙地利用了点云的特殊结构可以压缩的性质, 通过保存一组高度压缩的过去数据, 使用重放和正则等方法, 从而对抗灾难性遗忘, 降低了顺序任务不断迭代导致的精度损失。X Wang 等 [83] 等方差作为额外的先验知识注入到网络中, 有效地提炼到先前任务的几何方差信息, 从而克服了几个主流点云数据集中与姿势无关 (旋转不变性) 的场景的挑战。T Chowdhury 等 [84] 通过利用正交基向量, 使用一组预定义的规则描述 3D 对象, 使点云支持少量样本类增量学习任务。

传统的点云持续学习方法大多基于欧式空间 (Euclidean Space) 进行建模, 即假设点云数据分布在标准的三维坐标系中, 并采用多层感知机、卷积神经网络、Transformer 或图神经网络等模型直接在点云坐标或其转换后的特征上进行学习。然而, 点云数据本质上是高维流形 (Manifold) 上的非欧几里得数据, 直接在欧式空间中建模可能导致信息丢失、几何结构扭曲、表征能力受限, 从而影响模型的泛化能力、迁移能力和持续学习能力。因此, 如何突破欧式表示的局限性, 探索更鲁棒的点云表示学习方法, 逐渐成为点云持续学习的关键研究方向。

第三章 基于双曲流形映射的点云持续学习缓存重放方法研究

近年来,持续学习在深度学习领域受到了越来越多的关注。为了解决灾难性遗忘这一挑战,研究人员提出了多种应对策略,例如基于重放的策略 [85]、基于正则化的策略 [16] 以及基于参数隔离的策略 [86]。值得注意的是,这些策略在设计上具有正交性,因此在某些场景中可以进行组合,从而充分发挥各自的优势,取得更优异的效果。

在这些方法中,重放策略作为一种简单而有效的手段得到了广泛应用。重放策略通过存储部分原始样本或使用生成模型生成伪样本,在学习新任务的同时,通过重放先前任务的样本来缓解遗忘问题。通过周期性地回顾过往的经验和信息,该方法能够加强学习和记忆,将知识和技能更持久地存储于长期记忆中。然而,重放方法可能在样本选择过程中倾向于特定类型的样本,从而可能引入样本选择偏差,影响模型的泛化能力。

本章节聚焦于三维对象即点云的持续学习任务,针对现有方法的局限性提出了改进方法。三维点云数据具有显著的分层树状结构,每个点云由多个较小部分组合而成。利用这一特性,我们能够无缝地结合全局信息和局部信息,从而提升持续学习的能力。然而,传统的欧几里得空间 (Euclidean Space) 由于其体积随半径多项式增长的特性,无法很好地表示树状数据,从而导致表示能力的不足。幸运的是,由于负曲率的特性,双曲空间 (Hyperbolic Space) 能够以低失真嵌入分层结构 [87],显著增强对三维数据特征的表示能力。

本章节的主要贡献如下:

(1) 优化并改进了基于贪心采样策略的重放缓存区,扩展其至三维点云数据,提升了其在持续学习中的适用性。

(2) 引入双曲空间映射,利用双曲空间更好表示点云特征信息,并在此基础上提出了一种结合双曲空间全局与局部信息的知识蒸馏策略,帮助模型在持续学习任务中有效克服灾难性遗忘问题。

3.1 持续学习任务设置

首先，我们将简要概述持续学习的任务设置，然后详细阐述我们提出的方法的具体内容。

我们先定义一组持续学习任务序列，记为 T ，包含总共 N 个任务，即

$$T = (T_1, T_2, \dots, T_N)$$

对于每个任务 t ，其对应的数据集为

$$D_t = ((x_t^1, y_t^1), (x_t^2, y_t^2), \dots, (x_t^n, y_t^n))$$

其中 x 表示样本， y 表示标签。我们规定在学习过程中，模型只能使用容量有限的缓存区 B 来存储少量来自先前任务的样本。

在评估标准方面，我们采用平均整体任务准确率 (ACC) 作为判别依据，其定义为

$$ACC = \frac{1}{T} \sum_i^T a_i \quad (3.1)$$

其中 a_i 表示任务 i 的准确率。

3.2 双曲空间投影

双曲空间 (Hyperbolic Space)[88] 是一种具有负曲率的非欧几里得空间。与欧几里得空间中的平面或球面不同，双曲空间呈现“马鞍形”负曲率。在双曲空间中，直线是最短路径，但直线之间的角度概念与欧几里得空间不同——直线之间的角度是双曲角。此外，在双曲空间中，两条平行线会越来越接近，但永不相交。因此，双曲空间在许多应用中表现出了优异性能，尤其是在表示和建模层次化结构数据方面，例如语言表示学习、图嵌入和神经网络中的层次建模等领域。

目前已有不少的工作 [89, 90, 91, 92] 指出，将点云数据映射到双曲空间，相比欧式空间能够更好地获取其特征信息，这促使我们考虑利用双曲投影来更好地处理点云数据。

在本章节，我们采用广泛使用的庞加莱圆盘模型 (Poincaré Disk Model) 来描述双曲空间。双曲空间的负曲率由以下用于描述空间度量的公式决定：

$$G_R = (\lambda_x^c)^2 g_E = \frac{2}{1 + \|x\|} g_E \quad (3.2)$$

其中， λ 是保角因子 (Conformal Factor)， g_E 是欧几里得度量张量， c 是庞加莱圆盘的曲率。双曲空间不同于定义在欧几里得空间中的向量空间，但许多算子可以通过莫比乌斯向量空间 (Möbius Vector Spaces) 扩展到双曲空间。

庞加莱球是一种曲率为 $c = -1$ 的双曲空间，其距离定义为：

$$d_D(x, y) = \cosh^{-1} \left(1 + \frac{2\|x - y\|^2}{(1 - \|x\|^2)(1 - \|y\|^2)} \right) \quad (3.3)$$

由于庞加莱球是一个黎曼流形，对于球面上的每一点，可以使用指数映射 (Exponential Map) 和对数映射 (Logarithmic Map) 将切向量映射到流形上，或将流形上的点映射回切空间：

$$\exp_x^c(w) = x \oplus_c \left(\tanh \left(\sqrt{c} \frac{\lambda_x^c \|w\|}{2} \right) \frac{w}{\sqrt{c} \|w\|} \right) \quad (3.4)$$

$$\log_x^c(y) = \frac{2}{\sqrt{c} \lambda_x^c} \operatorname{arctanh}(\sqrt{c} \|-x \oplus_c y\|) \frac{-x \oplus_c y}{\|-x \oplus_c y\|} \quad (3.5)$$

其中， w 是流形在点 x 的切空间中的切向量。

点云作为一种具有树状分层结构的数据类型，非常适合映射到双曲空间中。每个点云由较小的部分向上生长而成。因此，我们旨在将点云数据表示在双曲空间中，利用双曲空间在表示树状分层结构方面的卓越能力，更好地理解点云的特征信息 [89]。具体而言，在庞加莱圆盘上，我们将完整点云定位在靠近圆盘边缘的位置，而将点云结构的部分定位在靠近圆盘中心的位置。此外，我们确保不同类别的点云彼此保持较远的距离，而同一类别的点云保持相对较近的距离。为实现这一目标，模型主要受以下两个正则化公式的约束：

$$L_{pos} = \max(0, -\|z_{whole}^+\| + \|z_{part}^+\| + \delta) \quad (3.6)$$

$$L_{neg} = \max(0, d(z_{whole}^+, z_{part}^+) - d(z_{whole}^+, z_{part}^-) + \delta) \quad (3.7)$$

其中， δ 表示偏移量， z_{whole}^+ 表示完整的正样本点云， z_{part}^+ 表示正样本点云的部分， z_{part}^- 表示负样本点云的部分。通过这种方式，我们能够更好地理解点云的局部特征。

3.3 双曲流形缓存重放

最近的研究 [93] 表明，重放方法是一种应对灾难性遗忘的简单且有效的手段。然而，重放方法在平衡存储容量与性能以及样本选择方面面临挑战，此外，如何对这些缓存内容进行重放也是研究的重点问题之一，我们的方法将围绕解决这些问题展开。

重放方法设计

首先，我们在模型架构中引入一个双曲映射特征提取层，通过指数映射将点云特征从欧式空间映射到双曲空间，并对提取出的特征进行拼接等进一步的处理。

同时，如前所述，点云数据呈现出明显的层次化树状结构，每个复杂的点云由简单的部分组成。换句话说，点云数据具有显著的局部结构特征。然而，仅考虑局部特征并不够全面，因为全局信息同样至关重要。幸运的是，我们的网络设计能够同时满足对局部和全局信息的需求。在池化之前，我们可以在双曲空间中捕获点云数据的局部特征，而在池化之后，我们可以获得点云数据的全局特征。因此，我们通过公式3.3从教师-学生模型中蒸馏出这两种特征，并最终将这些结果融合以应对灾难性遗忘的问题。这种综合利用局部和全局特征的方法有望使模型能够更好地理解点云数据，并促进不同任务之间的迁移学习和持续学习，从而更好地适应不同场景中的数据变化。因此，我们的损失函数定义如下：

$$\begin{aligned} Loss = & L_D(f_s(x; \theta), y) + L_M(f_s(x; \theta), y) \\ & + L_{distance}(\phi_s(x; \theta), \phi_t(x; \theta)) + L_{pos} + L_{neg} \end{aligned} \quad (3.8)$$

其中， L_D 是对应当前数据集的交叉熵， L_M 是对应重放缓存区中存储的数据集的交叉熵， $L_{distance}$ 表示 TS 蒸馏模型之间的距离 (包括整体结构和部分结构的距离)。

重放缓存设计

重放缓存区的大小是有限的，一旦达到其上限，随着新任务的加入，不可避免地会出现缓存替换问题。然而，简单的随机替换方法往往效果较差，因此选择高质量的旧数据变得至关重要。我们希望缓存的旧知识能够充分代表其对应的旧任务，

但在实践中，这一点很难实现。因此，我们需要精心挑选具有代表性的数据作为重放的候选内容。

我们的方法是通过将数据分布投影到一个流形球面上，让该球面的半径随着新任务的加入不断扩展。而由于数据特征分布在双曲空间上，基于双曲空间的质心和半径需要依赖迭代的梯度下降算法，需要耗费不少计算资源，因此我们对特征投影回欧式空间构造流形。

对于新样本，我们将其流形中心与当前流形直径进行比较。如果其距离超过当前半径，将其包含在缓存区中是有利的。至于被替换的旧样本，考虑到计算成本以及衡量旧样本对缓存区贡献的难度，我们选择采用随机替换策略。

此外，由于我们在重放方法中结合了局部和全局特征，因此分别为局部特征和全局特征建立独立的流形球面。我们基于归一化权重为每个特征独立选择样本，以充分利用数据的多样性和代表性。这种综合策略不仅有效解决了缓存替换问题，还增强了模型的泛化能力和任务的鲁棒性。我们的选择策略如算法1所示。

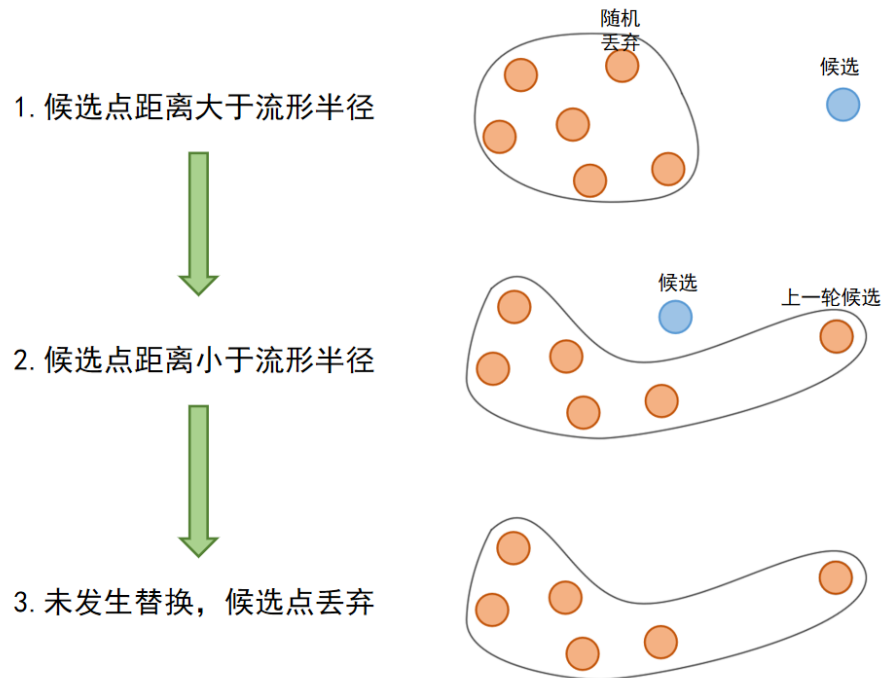


图 3.1 双曲流形扩张淘汰算法图解。

算法 1 双曲流形扩张淘汰算法

Input: Φ_s, D_t, \mathcal{M} , 缓存区大小上限 $\text{buffer_size}, n, \alpha_1, \alpha_2$

Output: \mathcal{M}

```

1:  $n \leftarrow 0$ 
2: for 所有  $(x, y)$  属于  $D_t$  do
3:   if  $|\mathcal{M}| < \text{buffer\_size}$  then
4:      $\mathcal{M}.\text{append}(x, y)$ 
5:   else
6:      $f_{local}^{\mathcal{M}}, f_{global}^{\mathcal{M}} \leftarrow \log_0^c(\Phi_s(\mathcal{M}))$ 
7:      $f_{local}^x, f_{global}^x \leftarrow \log_0^c(\Phi_s(x))$ 
8:      $c_{local}, d_{local} \leftarrow$  计算流形质心和半径( $f_{local}^{\mathcal{M}}$ )
9:      $c_{global}, d_{global} \leftarrow$  计算流形质心和半径( $f_{global}^{\mathcal{M}}$ )
10:    计算是否比最远样本更偏离中心:
11:     $\delta \leftarrow \alpha_1(\text{dist}(c_{local}, f_{local}^x) - d_{local})$ 
12:       $+ \alpha_2(\text{dist}(c_{global}, f_{global}^x) - d_{global})$ 
13:    if  $\delta > 0$  then
14:       $i \leftarrow \text{randint}(0, \text{buffer\_size})$ 
15:       $\mathcal{M}[i] \leftarrow (x, y)$ 
16:    else
17:       $i \leftarrow \text{randint}(0, n)$ 
18:      if  $i < \text{buffer\_size}$  then
19:         $\mathcal{M}[i] \leftarrow (x, y)$ 
20:      end if
21:    end if
22:  end if
23:   $n \leftarrow n + 1$ 
24: end for
25: return  $\mathcal{M}$ 

```

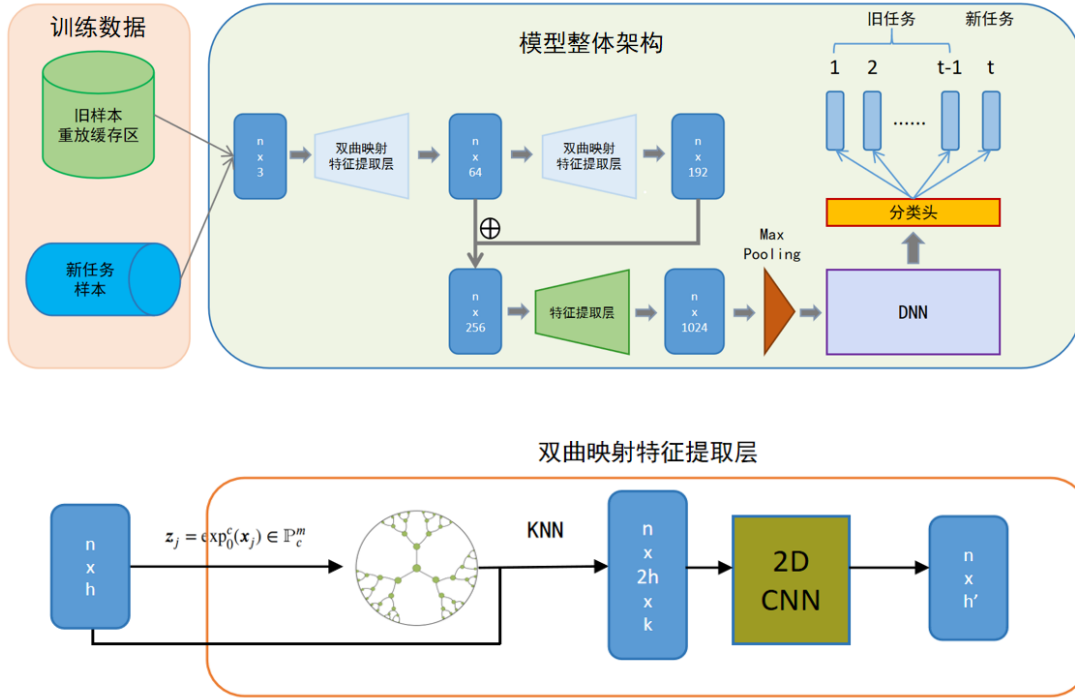


图 3.2 HyMR 模型整体架构。

3.4 实验及分析

我们将我们的方案取名为双曲流形重放 (HyMR)，整体架构如图3.2所示，并在不同数据集上使用 HyMR 进行多任务持续学习，然后将其与最新的以及常见的点云持续学习方法进行比较。

3.4.1 实验设置

我们在两个广泛使用的点云数据集 ModelNet[94] 和 ShapeNet[95] 上评估了算法的有效性。对于所有数据集，我们对每个点云数据统一采样 2048 个点，并进行共 20 轮的类增量任务。在初始任务中，我们选择样本数量最多的两个类别，而在随后的每一轮中，依次加入两个新类别。用于重放的缓存区大小设置为总样本数的 5%，约为 500 个样本。

我们使用 PyTorch 库实现了 HyMR 的所有模块，并在两块 NVIDIA TESLA V100 GPU 上进行训练。对于模型的骨干网络 (backbone)，我们选择了在点云分类中常用的 PointNet，并在此基础上引入了一个双曲映射模块。在双曲映射特征提取

层中,我们引入了一个双曲映射层,将特征投影到双曲空间中,然后使用卷积核大小为 1 的卷积网络层进行特征提取,最后将获得的特征进行拼接。

对于分类器,我们使用了三层全连接层(512-256-类别数)。在分类头中,我们通过更改最终输出层来匹配类别数量的增长。具体来说,每当出现新任务时,模型会冻结之前的输出层,并为新类别创建一个专用的输出层。这种方法有助于将模型大小保持在合理范围内。学习率统一设置为 0.003,批量大小设置为 16,每个任务最多训练 50 个周期 (epochs)。

我们将提出的方法与几种较新的点云持续学习相关方法进行了比较,并以顺序微调 (sequential fine-tuning) 作为对照。需要指出的是,部分对比方法 (如 iCaRL[28]、LwF[15] 等) 最初设计用于图像分类任务,其核心思想分别为样本重放、知识蒸馏和正则化约束。在迁移到点云任务中时,我们对其进行了必要的调整,使其适用于点云数据的输入格式与特征表示。例如,对于 iCaRL,我们在提取点云特征后执行最近邻分类策略;对于 LwF,我们保留其损失函数的核心机制,并将主干网络替换为适用于点云处理的基础结构 (PointNet[21])。尽管这些方法在点云场景中并非原生设计,但通过适当的适配,仍具备一定的代表性,能够作为有效的 baseline 进行性能比较。我们在所有数据上进行了五次独立实验,并计算了平均结果。

3.4.2 数据集介绍

ModelNet40 是一个广泛用于三维点云分类任务的基准数据集,最早用于评估 3D 形状识别任务。该数据集是 ModelNet 数据集 [94] 的一个子集,专门包含了 40 个常见物体类别,因此得名 ModelNet40。该数据集包含 40 个不同类别的物体,每个类别涵盖各种日常物体,如桌子、椅子、飞机、汽车等,涵盖了现实世界中常见的三维结构。数据集中的每个对象均由 3D CAD 模型转换而来,并采用点云或网格的形式进行表示,为研究人员提供了一个标准化的测试环境,以验证模型在三维形状识别任务上的泛化能力。

ShapeNet 是一个大规模的三维模型数据库,专为计算机视觉、计算机图形学和深度学习研究而构建。该数据集涵盖了大量来自不同类别的三维物体,广泛涉及日常生活中的物品,如家具、交通工具和家用电器等。所有模型均采用标准化的

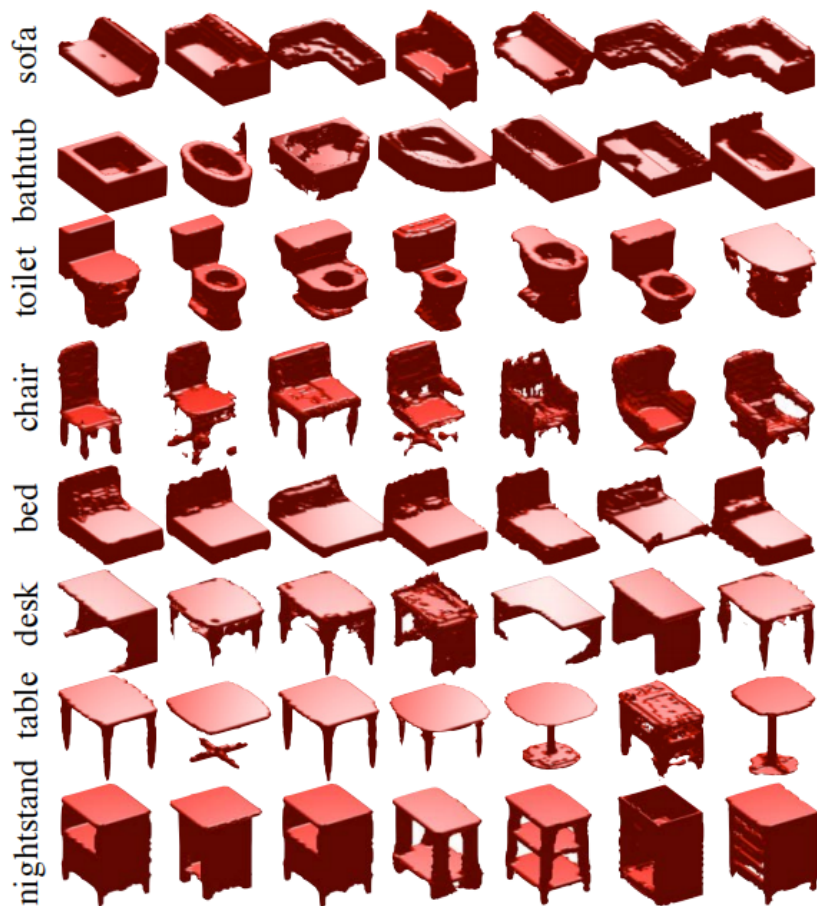


图 3.3 ModelNet 数据集预览。

格式，保证了数据的一致性，并在几何、拓扑结构和语义层面进行了细致的标注。这些标注信息不仅包括物体的类别，还涵盖了部分分割、对齐信息、物理尺寸和对称性等特征，使得 ShapeNet 在语义理解、形状生成和三维重建等任务中具有极高的应用价值。由于其规模庞大、类别丰富且标注详尽，ShapeNet 已成为三维视觉和计算机图形学领域最具代表性的数据集之一，为学术界和工业界的研究提供了强有力的支持。

3.4.3 结果分析

在表 3.3 中，我们展示了在 ShapeNet 和 ModelNet 数据集上的类别增量消融实验结果，比较了不同缓存区大小和是否使用双曲映射的情况。结果表明，HyMR 的性能在一定程度上依赖于缓存区的大小，并且采用双曲空间映射显著增强了模型



图 3.4 ShapeNet 数据集预览。

表 3.1 在类别增量学习中，HyMR(双曲映射) 与其他在欧式空间的不同方法在 ModelNet40、ShapeNet 数据集上经过若干轮训练后的平均准确率 (%)。

数据集	任务数	顺序微调	iCaRL[28]	LwF[15]	RPS-Net[96]	I3DOL[19]	HyMR(Ours)
ModelNet40	10	13.3	71.8	62.3	86.7	89.7	90.3
	20	2.1	57.2	58.3	31.5	61.5	78.0
ShapeNet	10	9.7	72.9	72.7	86.9	87.8	89.8
	20	1.3	36.4	48.3	69.5	74.1	80.8

对点云特征的理解能力。

在图 3.5、3.6和表3.1、3.2中，我们展示了在 ShapeNet 和 ModelNet 数据集上，随着任务轮数增加，HyMR 方法与其他方法在类别增量和任务增量设置下的实验结果。可以观察到，HyMR 相较于基线方法表现出更优越的性能，并且随着任务轮数的增加，其效果更加显著。

表 3.2 在任务增量学习中, HyMR(双曲映射) 与其他在欧式空间的不同方法在 ModelNet40、ShapeNet 数据集上经过若干轮训练后的平均准确率 (%)。

数据集	任务数	iCaRL[28]	LwF[15]	EWC[17]	HyMR(Ours)
ModelNet40	10	89.9	92.8	92.2	98.6
	20	88.2	91.7	87.7	98.3
ShapeNet	10	93.3	91.9	83.3	98.6
	20	82.0	75.3	70.3	98.3

表 3.3 类增量任务在 ModelNet40 和 ShapeNet 数据集上, 对于不同消融方案的 HyMR 方法, 经过 10 轮和 20 轮任务训练后的最终结果。

数据集	消融方案	准确率 (10 轮任务)	准确率 (20 轮任务)
ModelNet40	HyMR	90.3	78.0
	传统蒸馏	90.9	77.7
	缓存区随机淘汰	88.3	76.8
	无双曲映射	90.6	77.9
	缓存区 100 上限	85.7	70.3
	缓存区 250 上限	86.9	73.5
ShapeNet	HyMR	89.8	80.8
	传统蒸馏	89.6	80.1
	缓存区随机淘汰	87.6	77.7
	无双曲映射	88.5	80.3
	缓存区 100 上限	84.4	69.9
	缓存区 250 上限	87.4	74.1

3.5 章节小结

在本章节的工作中, 我们提出了一种针对点云数据持续学习的新方法。该方法通过将特征映射到双曲空间, 并结合全局与局部视角, 利用重放策略来解决问题。我们认为, 利用双曲空间出色的层次化表示能力, 并全面考虑数据特征, 可以有效缓解点云分类任务中的灾难性遗忘问题。然而, 该方法在某些方面仍然存在一些问题, 例如在选择缓存区时计算直径和距离的计算成本较高, 并且过于依赖缓存区。在第四章中, 我们将继续优化方法, 并寻求更优的解决方案。

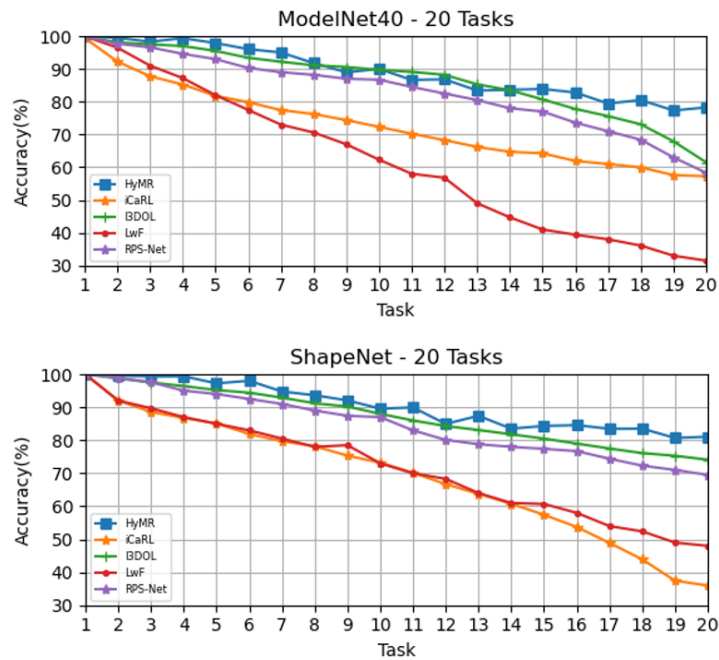


图 3.5 类增量学习设置下，在完成 t 个任务后的整体平均准确率，训练和验证期间，每个任务的类别信息是未知的。图中展示了 HyMR(双曲映射) 与其他基于欧氏空间的方法在 ModelNet40 和 ShapeNet 数据集上训练 20 轮任务后的准确率变化。

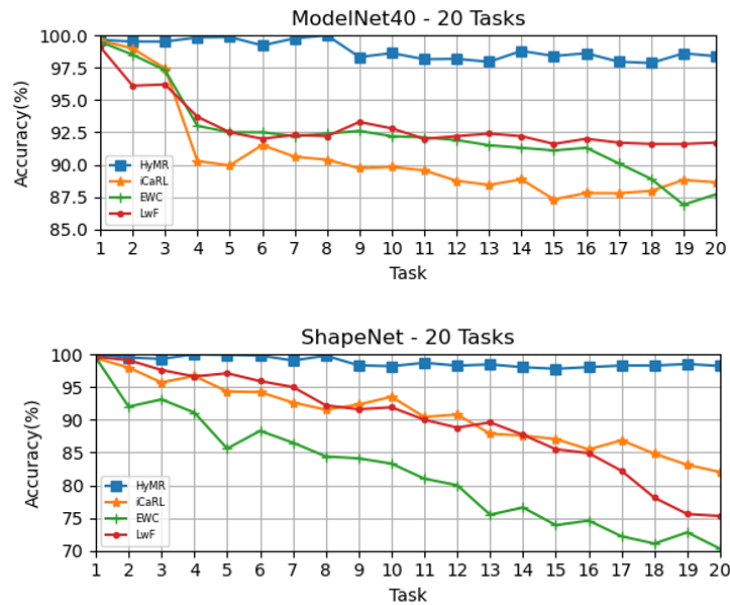


图 3.6 任务增量学习设置下，完成 t 个任务后的整体平均准确率，训练和验证期间，每个任务的类别信息是已知的。图中展示了 HyMR(双曲映射) 与其他基于欧氏空间的方法在 ModelNet40 和 ShapeNet 数据集上训练 20 轮任务后的准确率变化。

第四章 基于生成式模型的双曲映射点云持续学习方法研究

在上一章节中，我们提出了一种基于双曲流形映射的点云持续学习缓存重放方法，通过优化贪心采样策略，提升了重放缓存区在点云数据上的适用性，并结合双曲空间的全局与局部信息，改善了模型的表示能力。然而，我们发现基于缓存的重放方法仍然存在一些局限性：缓存策略在采样过程中可能倾向于特定类型的样本，影响模型的泛化能力；缓存区的大小有限，无法完整保留所有历史任务的数据，导致旧任务的样本多样性受限；某些任务（如医疗、金融）可能无法直接存储原始数据，限制了缓存重放的应用范围；而且，目前点云生成持续学习的研究内容较少，传统的方法迁移到点云上后灾难性遗忘依旧明显。

同时，即使在之前工作中我们已经采用了优化的贪心采样策略和双曲映射提升了缓存重放效果，但从实验的结果中也能看出，旧任务表现下降仍然明显，尤其在任务数增加时效果迅速劣化，并且模型对早期任务的性能仍然随任务数增加持续下降，下降速度和缓存大小成反比例关系。

根据上述情况，我们意识到仅依赖缓存机制的持续学习方法在点云场景下存在天然的适用性边界。一方面，缓存区的有限容量难以适应点云数据的大规模与高维度特性；另一方面，在对隐私敏感或数据共享受限的场景中，缓存机制本身的适用性也受到制约。于是，我们尝试从“存储真实样本”的被动方案，转向“主动生成高质量样本”的生成式方法，期望以更高效、更灵活的方式缓解灾难性遗忘问题，使之成为一种理论上更普适的补充甚至替代方案，因为它不依赖保存真实样本，而是从模型中提取知识用于数据重构。

因此，本章提出了一种基于生成模型的双曲映射点云持续学习重放方法，用扩散模型 (Diffusion Model) 替代传统的重放缓存区，生成高质量、多样化的点云数据，并结合双曲空间的表示能力，提高点云数据的重放质量和持续学习能力。具体来说，本章节的主要贡献如下：

(1) 提出了一种基于扩散模型的点云生成式重放方法 (HyperDiffIL)，突破了传统缓存区方法的存储限制，提高了重放样本的多样性，并且支持分类和生成任务。

(2) 结合双曲空间进行特征映射和优化, 提升了点云数据的表示能力, 并在此基础上设计一种生成样本筛选机制, 使生成样本更准确。

(3) 设计了一种适用于双曲空间下扩散模型特征的全过程知识蒸馏策略, 确保生成样本的质量, 同时减少灾难性遗忘, 提高模型的长期学习能力。

4.1 基于双曲空间的扩散模型点云生成

4.1.1 传统扩散模型

去噪扩散概率模型 (Denoising Diffusion Probabilistic Model, DDPM)[40] 是一种基于马尔可夫链的生成方法, 它通过正向扩散过程将数据逐步扰动为纯噪声, 然后通过反向去噪恢复到原始数据分布。在生成任务上, 这一反向过程的精确控制使得扩散模型能够在生成任务中表现出卓越的性能, 尤其是在面对高维度、复杂数据时, 能够有效地处理数据的细粒度结构, 避免了传统生成模型中常见的模式崩塌问题。通过这种方式, 扩散模型不仅能够生成高质量的样本, 还能够任务中进行有效的特征学习, 展现出强大的生成能力和鲁棒性。

前向扩散过程

扩散模型的前向过程通过逐步向数据添加噪声, 使其分布收敛到标准正态分布。设 $\mathbf{x}_0 \sim q(\mathbf{x}_0)$ 为从真实数据分布中采样的数据, 正向扩散 (从 $t-1$ 步扩散到 t 步) 的过程定义如下:

$$q(\mathbf{x}_{1:T}|\mathbf{x}_0) = \prod_{t=1}^T q(\mathbf{x}_t|\mathbf{x}_{t-1}) \quad (4.1)$$

我们逐步向数据添加噪声, 使其逐渐变得无序, 最终趋近于一个标准高斯分布。这个过程可以用以下的马尔可夫链建模:

$$q(\mathbf{x}_t|\mathbf{x}_{t-1}) = N(\mathbf{x}_t; \sqrt{1 - \beta_t}\mathbf{x}_{t-1}, \beta_t\mathbf{I}) \quad (4.2)$$

其中, β_t 是一个超参数, 称为噪声调度 (variance schedule), 决定了每一步扩散的噪声大小, $N(\mu, \Sigma)$ 表示均值为 μ , 协方差矩阵为 Σ 的正态分布。通过重参数化, 我们可以很快地得到任意时刻 t 处的点云分布:

$$q(\mathbf{x}_t|\mathbf{x}_0) = N(\mathbf{x}_t; \sqrt{\bar{\alpha}_t}\mathbf{x}_0, (1 - \bar{\alpha}_t)\mathbf{I}) \quad (4.3)$$

其中,

$$\alpha_t = 1 - \beta_t, \quad \bar{\alpha}_t = \prod_{s=1}^t \alpha_s \quad (4.4)$$

于是, 在具体实践中, 我们可以直接通过一步采样得到 \mathbf{x}_t :

$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t} \mathbf{x}_0 + \sqrt{1 - \bar{\alpha}_t} \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim N(\mathbf{0}, \mathbf{I}) \quad (4.5)$$

反向去噪过程

为了从噪声恢复出原始数据, 我们需要学习一个反向去噪模型 $p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)$, 即:

$$p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) = N(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \boldsymbol{\Sigma}_\theta(\mathbf{x}_t, t)) \quad (4.6)$$

其中, $\boldsymbol{\mu}_\theta(\mathbf{x}_t, t)$ 是可学习的均值函数 (由神经网络估计), $\boldsymbol{\Sigma}_\theta(\mathbf{x}_t, t)$ 是可学习的协方差矩阵。利用贝叶斯定理, 我们可以计算真实的后验分布:

$$q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) = N(\mathbf{x}_{t-1}; \tilde{\boldsymbol{\mu}}_t(\mathbf{x}_t, \mathbf{x}_0), \tilde{\beta}_t \mathbf{I}) \quad (4.7)$$

$$\tilde{\boldsymbol{\mu}}_t(\mathbf{x}_t, \mathbf{x}_0) = \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1 - \bar{\alpha}_t} \mathbf{x}_0 + \frac{\sqrt{\alpha_t}(1 - \bar{\alpha}_{t-1})}{1 - \bar{\alpha}_t} \mathbf{x}_t \quad (4.8)$$

$$\tilde{\beta}_t = \frac{(1 - \bar{\alpha}_{t-1})}{1 - \bar{\alpha}_t} \beta_t \quad (4.9)$$

而在具体应用中, 我们可以用神经网络 $\epsilon_\theta(\mathbf{x}_t, t)$ 来模拟 \mathbf{x}_0 :

$$\hat{\mathbf{x}}_0 = \frac{\mathbf{x}_t - \sqrt{1 - \bar{\alpha}_t} \epsilon_\theta(\mathbf{x}_t, t)}{\sqrt{\bar{\alpha}_t}} \quad (4.10)$$

最终, 去噪步骤为:

$$\mathbf{x}_{t-1} = \boldsymbol{\mu}_\theta(\mathbf{x}_t, t) + \sqrt{\tilde{\beta}_t} \mathbf{z}, \quad \mathbf{z} \sim N(\mathbf{0}, \mathbf{I}) \quad (4.11)$$

训练目标

扩散模型的训练目标通常是最大化对数似然函数, 这通常是一个难以直接计算的目标。因此, 采用变分下界 (Variational Lower Bound, VLB) 来优化, 目标是最小化变分下界:

$$L_{\text{VLB}} = \mathbb{E}_q \left[\sum_{t=1}^T D_{\text{KL}}(q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) \parallel p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)) - \log p_\theta(\mathbf{x}_0) \right] \quad (4.12)$$

而为了简化训练过程，扩散模型通常采用一种更直接的训练目标，即通过最小化模型在各时间步 t 上输出的噪声估计与真实噪声之间的误差。简化后可以得到最终的训练目标为：

$$L_{\text{simple}} = \mathbb{E}_{t, \mathbf{x}_0, \epsilon} [\|\epsilon - \epsilon_\theta(\mathbf{x}_t, t)\|^2] \quad (4.13)$$

4.1.2 点云的扩散模型生成

设定点云数据集为：

$$\mathbf{X}_0 = \{\mathbf{x}_i\}_{i=1}^N, \quad \mathbf{x}_i \in \mathbb{R}^3 \quad (4.14)$$

传统扩散模型主要应用于图像生成，但如第三章所述，由于点云数据的非欧几里得结构，直接在欧几里得空间进行扩散可能无法很好地捕捉点云的层次化信息。因此，我们考虑将点云特征投影到双曲空间进行扩散和反向过程，并最终在三维欧几里得空间生成点云，以获取更好的效果。这里和第三章一样，我们选择庞加莱圆盘：

$$\mathbb{D}^n = \{\mathbf{z} \in \mathbb{R}^n \mid \|\mathbf{z}\| < 1\} \quad (4.15)$$

其中，庞加莱度量定义为：

$$ds^2 = \frac{4d\mathbf{z}^2}{(1 - \|\mathbf{z}\|^2)^2} \quad (4.16)$$

为了将点云投影到双曲空间，我们采用指数映射：

$$\mathbf{z}_i = \exp_0(\mathbf{x}_i) = \tanh(\|\mathbf{x}_i\|) \frac{\mathbf{x}_i}{\|\mathbf{x}_i\|} \quad (4.17)$$

从而得到点云在双曲空间的表示：

$$\mathbf{Z}_0 = \{\mathbf{z}_i\}_{i=1}^N, \quad \mathbf{z}_i \in \mathbb{D}^2 \quad (4.18)$$

双曲空间中的正向扩散

在传统的扩散模型中，数据点在欧几里得空间内逐步添加噪声，以转换为标准高斯分布。然而，在双曲空间中，数据点的运动需要符合双曲几何。因此，我们定义双曲空间内的扩散过程：

$$q(\mathbf{z}_t | \mathbf{z}_{t-1}) = N_{\mathbb{D}}(\mathbf{z}_t; \sqrt{1 - \beta_t} \mathbf{z}_{t-1}, \beta_t I) \quad (4.19)$$

其中 β_t 是噪声调度参数, 决定每一步扩散的噪声大小。由于双曲空间具有负曲率, 点之间的加法需要使用莫比乌斯加法 (Möbius Addition):

$$\mathbf{a} \oplus_{\mathbb{D}} \mathbf{b} = \frac{(1 + 2\langle \mathbf{a}, \mathbf{b} \rangle + \|\mathbf{b}\|^2)\mathbf{a} + (1 - \|\mathbf{a}\|^2)\mathbf{b}}{1 + 2\langle \mathbf{a}, \mathbf{b} \rangle + \|\mathbf{a}\|^2\|\mathbf{b}\|^2} \quad (4.20)$$

于是, 正向扩散的计算变为:

$$\mathbf{z}_t = (1 - \beta_t)\mathbf{z}_{t-1} \oplus_{\mathbb{D}} \sqrt{\beta_t}\mathbf{n}, \quad \mathbf{n} \sim N(0, I) \quad (4.21)$$

这个扩散过程确保了点云在双曲空间中的分布遵循双曲几何结构。

双曲空间中的反向推导

和传统扩散模型一样, 为了从噪声恢复点云, 我们需要学习一个适用于双曲空间的反向去噪模型。而由于我们需要对生成的类别有所要求, 即保证生成点云属于目标类别 y , 我们在反向去噪过程中引入了分类器 (下文持续学习任务中介绍) 引导。设类别分类器为:

$$C_\phi(\mathbf{z}) \in \mathbb{R}^K \quad (4.22)$$

其中, $C_\phi(\mathbf{z})$ 是 PointNet 或类似神经网络的输出, 表示点云属于 K 个类别的概率分布。分类器引导去噪的目标是修改去噪方向, 使得生成点云朝向目标类别优化。因此, 我们使用分类器梯度引导 (Classifier Guidance) 的方式:

$$\mathbf{g}_\phi = \nabla_{\mathbf{z}} \log p(y|\mathbf{z}) \quad (4.23)$$

其中, \mathbf{g}_ϕ 表示类别 y 对应的分类器梯度。最终的引导去噪方向为:

$$\mathbf{z}_{t-1} = \mu_\theta(\mathbf{z}_t, t) + \lambda \mathbf{g}_\phi + \sqrt{\tilde{\beta}_t}\mathbf{z}, \quad \mathbf{z} \sim N(0, I) \quad (4.24)$$

其中, λ 为控制分类器引导强度的超参数。

而在双曲空间中完成去噪后, 我们需要将点云投影回三维欧几里得空间, 由庞加莱球模型的对数映射, 我们可以得到所需的生成样本:

$$\mathbf{x}_i = \frac{\mathbf{z}_i}{\|\mathbf{z}_i\|} \tanh^{-1}(\|\mathbf{z}_i\|) \quad (4.25)$$

双曲扩散生成的训练目标

最终训练目标类似公式仍基于变分下界 (VLB)，但由于引入类别分类器梯度引导，因此根据公式4.12和公式4.13，我们可以得到最终的 VLB 和简化的损失函数：

$$L_{\text{VLB}} = \mathbb{E}_q \left[\sum_{t=1}^T D_{\text{KL}}(q(\mathbf{z}_{t-1} | \mathbf{z}_t, \mathbf{z}_0) \parallel p_{\theta}(\mathbf{z}_{t-1} | \mathbf{z}_t)) - \log p_{\theta}(\mathbf{z}_0) \right] \quad (4.26)$$

$$L_{\text{simple}} = \mathbb{E}_{t, \mathbf{z}_0, \epsilon} [\|\epsilon - \epsilon_{\theta}(\mathbf{z}_t, t) + \lambda \nabla_{\mathbf{z}_t} \log p(y | \mathbf{z}_t)\|^2] \quad (4.27)$$

4.2 基于双曲空间的生成式重放点云持续学习

我们设计的模型架构包含在第三章工作的基础上改进的做下游任务的分类器和作为生成器的扩散模型两大部分组成。

4.2.1 持续学习任务设置

同第三章的工作类似，在点云分类的持续学习任务中，模型需要在一系列任务 T 上进行学习，同时避免灾难性遗忘。设定一个任务序列：

$$T = \{T_1, T_2, \dots, T_N\} \quad (4.28)$$

每个任务 T_t 由数据集 D_t 组成：

$$D_t = \{(\mathbf{x}_1^t, y_1^t), (\mathbf{x}_2^t, y_2^t), \dots, (\mathbf{x}_n^t, y_n^t)\} \quad (4.29)$$

其中， \mathbf{x}^t 代表点云样本， y^t 代表其类别标签。

在持续学习过程中，模型仅能访问当前任务的数据 D_t ，而无法直接访问过去任务的数据 $D_{1:t-1}$ ，这会导致模型在学习新任务时遗忘旧任务的知识。为了解决这一问题，我们引入本章介绍的基于双曲空间的生成式重放方法。

4.2.2 评估指标

为了衡量模型的持续学习性能 (包括分类器的分类能力以及生成器的生产能力)，我们采用以下主要评估指标：

整体准确率

模型在 T 轮任务后, 对所有已学习任务的准确率计算为:

$$ACC = \frac{1}{T} \sum_{i=1}^T a_i \quad (4.30)$$

其中 a_i 是模型在第 i 任务上的测试准确率。

JSD

Jensen-Shannon Divergence(JSD) 衡量真实点云分布 P_r 和生成点云分布 P_g 之间的相似度, 基于 KL 散度计算:

$$JSD(P_r, P_g) = \frac{1}{2} D_{KL}(P_r \| M) + \frac{1}{2} D_{KL}(P_g \| M) \quad (4.31)$$

$$D_{KL}(P \| Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)} \quad (4.32)$$

其中, $M = \frac{1}{2}(P_r + P_g)$ 为两个分布的平均值, $D_{KL}(P \| Q)$ 是 KL 散度。JSD 越小, 表示生成的点云分布与真实数据分布越接近, 生成质量越高; 反之, 则代表生成点云与真实点云严重偏离。

MMD

Minimum Matching Distance(MMD) 衡量真实点云分布 P_r 和生成点云分布 P_g 之间的最小匹配距离, 用于评估形状相似性:

$$MMD(P_r, P_g) = \frac{1}{|P_r|} \sum_{x \in P_r} \min_{y \in P_g} d(x, y) \quad (4.33)$$

其中, $d(x, y)$ 表示两点云样本的距离, 可使用推土机 (EMD) 距离计算:

$$d_{EMD}(P_r, P_g) = \min_{\phi: P_r \rightarrow P_g} \sum_{x \in P_r} \|x - \phi(x)\| \quad (4.34)$$

MMD 越小, 表示生成样本与真实样本更相似, 生成模型质量更高。

COV

Coverage(COV) 衡量生成点云覆盖了多少真实点云的分布, 用于评估生成多样:

$$\text{COV}(P_r, P_g) = \frac{|\{x \in P_r \mid \exists y \in P_g, d(x, y) < \tau\}|}{|P_r|} \quad (4.35)$$

其中, τ 是一个匹配阈值, 用于判断生成样本是否匹配真实样本, $d(x, y)$ 同样可使用 EMD 计算。COV 越高, 表示生成样本分布覆盖了更多的真实数据, 多样性更好。

1-NNA

1-Nearest Neighbor Accuracy(1-NNA) 衡量生成点云和真实点云在特征空间的区分度, 它基于最近邻分类, 计算一个样本的最近邻是否来自相同的分布, 评估多样性和过拟合情况。首先计算所有样本的最近邻 (1-NN):

$$\text{NN}(F(X)) = \arg \min_{Y \in P_F \setminus \{X\}} d(F(X), F(Y)) \quad (4.36)$$

其中, $F(X), F(Y)$ 表示点云的特征, $P_F = P_r \cup P_g$ 是所有样本的特征集合。然后计算 1-NNA 评分:

$$\text{1-NNA}(S_g, S_r) = \frac{\sum_{X \in S_g} \mathbb{I}[\text{NN}(X) \in S_g] + \sum_{Y \in S_r} \mathbb{I}[\text{NN}(Y) \in S_r]}{|S_g| + |S_r|} \quad (4.37)$$

其中, \mathbb{I} 是指示函数:

$$\mathbb{I}(x) = \begin{cases} 1, & \text{如果最近邻样本属于相同分布} \\ 0, & \text{如果最近邻样本属于不同分布} \end{cases} \quad (4.38)$$

表 4.1 1-NNA 结果分析。

1-NNA 值	解释
$\approx 50\%$	生成点云与真实点云不可区分, 表示分布高度相似, 质量较好。
$> 50\%$	生成点云过拟合, 意味着生成样本与真实数据的特征差异较大。
$< 50\%$	生成点云质量较低, 可能缺乏多样性, 或偏离真实点云分布。

4.2.3 分类器和评分机制

对于分类器的选取，我们沿用章节三中的适合双曲空间的基于 PointNet 的点云分类器，其核心结构包括局部特征提取层、全局特征聚合层以及分类头层。分类器的目标是学习从点云 \mathbf{x} 到类别标签 y 的映射：

$$p_{\phi}(y|\mathbf{x}) = \text{softmax}(C_{\phi}(\mathbf{x})) \quad (4.39)$$

其中， C_{ϕ} 代表分类器的网络参数， softmax 用于计算类别概率。

但与工作一不同的是，在这里我们不使用传统的旧样本缓存区，而是使用上面介绍的基于双曲空间的点云扩散模型生成旧样本，并通过分类器来引导扩散模型生成高质量的旧样本。在每轮任务 t 训练时，分类器 C_{ϕ} 会对扩散模型生成的旧样本进行监督，目标是优化生成器，使其生成的点云更符合旧任务的类别分布。

对于扩散生成过程 $p_{\theta}(\mathbf{z}_{t-1}|\mathbf{z}_t)$ ，如公式4.27所示，我们在反向去噪过程的损失函数中引入分类器梯度。

生成样本质量评估与筛选

虽然扩散模型本身在反向推导过程中添加了噪声，但生成样本实际上还一定程度上由分类器引导生成的，生成样本的质量仍然可能受到分类器自身参数的影响。为了降低这一方面的影响，同时平衡生成效率与重放效果，我们设计了一种基于评分机制的筛选机制。该算法首先对生成的旧样本进行质量评估，然后根据预设的评分标准筛选出具有较高代表性的样本，从而过滤掉那些质量较低、代表性差的样本。这样做不仅提高了生成样本的整体质量，还能有效避免在重放过程中因大量低质量旧样本数据引发分类器过拟合的问题，从而增强模型在新旧任务之间的迁移和适应能力。

对于扩散模型生成的旧类别样本 \mathbf{z}_{gen} ，我们使用分类器的交叉熵损失来衡量其质量：

$$L_{\text{CE}}(C_{\phi}(\mathbf{z}), y) = - \sum_i y_i \log p(y_i|\mathbf{z}) \quad (4.40)$$

其中， $C_{\phi}(\mathbf{z})$ 是分类器的输出，表示类别预测概率， $p(y_i|\mathbf{z})$ 是生成样本属于类别 y 的概率。由于交叉熵损失越小，表明分类器越确信该样本属于正确类别，因此我们

对交叉熵取负作为评分值：

$$S(\mathbf{z}) = -L_{\text{CE}}(C_\phi(\mathbf{z}), y) \quad (4.41)$$

如算法2所示，对于每个旧类别，我们首先使用分类器引导的扩散模型生成 n 个样本，并计算其分类评分。重放样本的数量上限为 M ，当样本集未满足时，直接加入生成样本；当样本集已满足时，对比当前样本评分与样本集中的最高评分，若当前样本评分更高，则直接丢弃，否则以一定概率 ρ 随机替换样本集中的样本，反之，替换评分最高的样本。

基于这一算法，对于每个旧任务类别，我们可以获得一组代表性比较强的生成样本，将其与新任务的样本结合，得到当前任务下的训练数据集。利用这一集合，我们对当前的生成器和分类器进行学习和“复习”。

4.2.4 生成模型在双曲空间下的全过程知识蒸馏

为了防止模型在持续学习中遗忘旧任务的去噪能力，我们不只是通过生成本来引导模型，同时考虑进行知识蒸馏，让学生模型 ϵ_{θ_i} 学习前一任务的教师模型 $\epsilon_{\theta_{i-1}}$ 的知识。

传统知识蒸馏

在标准知识蒸馏中，学生模型通过减小与教师模型的分布距离来学习，通常采用 KL 散度：

$$L_{\text{KD}} = D_{\text{KL}}(p_\theta(\mathbf{z}) \| p_{\theta^*}(\mathbf{z})) \quad (4.42)$$

然而，该方法仅对最终输出进行匹配，无法保证扩散模型在整个去噪过程中的稳定性；而扩散模型如公式4.24所示，是由若干个时间步反向生成，仅仅对最终输出进行蒸馏并不全面，需要考虑全过程都有覆盖的蒸馏方法。

双曲空间下的全过程蒸馏

首先，在欧几里得空间中，我们使用 L_2 范数进行蒸馏：

$$L_{\text{distill, Euclidean}} = \mathbb{E}_{\mathbf{z}_t} \left[\|\epsilon_{\theta_{t-1}}(\mathbf{z}_t, t) - \epsilon_{\theta_t}(\mathbf{z}_t, t)\|^2 \right] \quad (4.43)$$

算法 2 基于样本评分的筛选算法

```

1: 初始化: 生成样本数量  $n$ , 每个旧任务类别生成样本上限  $M$ , 替换概率  $\rho$ , 评分函数  $S(\mathbf{z})$ , 当前类别重放样本集  $\mathbf{Z}_{\text{replay}} = \emptyset$ 
2: for 每个旧类别  $y$  do
3:   生成  $n$  个样本  $\mathbf{Z}_{\text{gen}} = \{\mathbf{z}_{\text{gen},i}\}_{i=1}^n$ 
4:   计算每个样本的评分  $S(\mathbf{z})$ 
5:   for 每个样本  $\mathbf{z} \in \mathbf{Z}_{\text{gen}}$  do
6:     if  $|\mathbf{Z}_{\text{replay}}| < M$  then
7:       直接加入样本  $\mathbf{Z}_{\text{replay}} \leftarrow \mathbf{Z}_{\text{replay}} \cup \{\mathbf{z}\}$ 
8:     else
9:       获取样本集中最高评分  $S_{\max} = \max_{\mathbf{z}' \in \mathbf{Z}_{\text{replay}}} S(\mathbf{z}')$ 
10:      if  $S(\mathbf{z}) > S_{\max}$  then
11:        丢弃当前样本
12:      else
13:        以概率  $\rho$  替换样本集中的一个随机样本
14:        以  $1 - \rho$  的概率替换评分最高的样本:
15:         $\mathbf{Z}_{\text{replay}} \leftarrow \mathbf{Z}_{\text{replay}} \setminus \{\arg \max_{\mathbf{z}' \in \mathbf{Z}_{\text{replay}}} S(\mathbf{z}')\} \cup \{\mathbf{z}\}$ 
16:      end if
17:    end if
18:  end for
19: end for
20: 返回: 过滤后的重放样本集  $\mathbf{Z}_{\text{replay}}$ 

```

但在双曲空间中, 数据点的几何关系更加复杂, 因此不能简单使用欧氏距离, 而应该采用双曲空间度量。在庞加莱球模型中, 两个点 $\mathbf{z}_1, \mathbf{z}_2$ 之间的双曲距离定义为:

$$d_{\mathbb{D}}(\mathbf{z}_1, \mathbf{z}_2) = \text{arcosh} \left(1 + 2 \frac{\|\mathbf{z}_1 - \mathbf{z}_2\|^2}{(1 - \|\mathbf{z}_1\|^2)(1 - \|\mathbf{z}_2\|^2)} \right) \quad (4.44)$$

在全过程蒸馏中, 我们希望让新任务模型的去噪预测 ϵ_{θ_t} 与旧任务模型的去噪预测 $\epsilon_{\theta_{t-1}}$ 在双曲度量下尽量趋于一致, 因此可以得到蒸馏损失:

$$L_{\text{distill, hyperbolic}} = \mathbb{E}_{\mathbf{z}_t} \left[d_{\mathbb{D}}(\epsilon_{\theta_{t-1}}(\mathbf{z}_t, t), \epsilon_{\theta_t}(\mathbf{z}_t, t)) \right] \quad (4.45)$$

而为了使得这一蒸馏操作对整个反向推导过程都有覆盖, 同时兼顾效率, 我们设计了相应的全过程蒸馏算法, 如算法3所示。

最终，我们的生成模型的损失函数定义如下：

$$L = L_{\text{simple}, D} + \lambda L_{\text{simple}, M} + \mu L_{\text{distill, hyperbolic}} \quad (4.46)$$

其中， $L_{\text{simple}, D}$ 代表当前任务样本的损失， $L_{\text{simple}, M}$ 代表重放样本的损失， λ 和 μ 代表重放和蒸馏系数。

算法 3 双曲空间下的全过程蒸馏

- 1: 旧任务模型 $\epsilon_{\theta_{t-1}}$ ，新任务模型 ϵ_{θ_t} ，时间步数 T
 - 2: **for** 每个重放样本 $x_0 \sim \mathbf{Z}_{\text{replay}}$ **do**
 - 3: 从时间步长 $T, \dots, 1$ 中随机采样覆盖步 t
 - 4: 生成加噪样本: $x_t = \sqrt{\alpha_t}x_0 + \sqrt{1 - \alpha_t}\epsilon$
 - 5: 计算去噪预测: $\epsilon_{\theta_t}(x_t, t) = f_{\theta_t}(x_t, t)\epsilon_{\theta_{t-1}}(x_t, t) = f_{\theta_{t-1}}(x_t, t)$
 - 6: 计算蒸馏损失 $L_{\text{distill, hyperbolic}}$
 - 7: 计算去噪损失 L_{simple}
 - 8: 计算总损失 L
 - 9: 更新模型参数
 - 10: **end for**
-

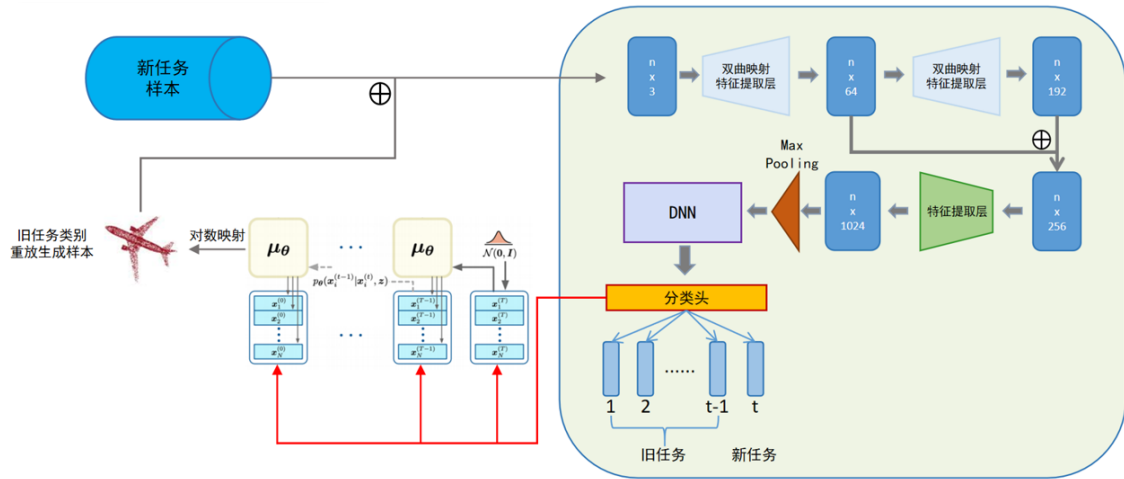


图 4.1 HyperDiffIL 模型架构。

4.3 实验及分析

为了和第三章工作进行对比实验以评价模型的分类精度，我们在实验部分的大部分超参数尽可能维持一致，并且所有实验同样在两块 NVIDIA Tesla V100 GPU 上进行，每块 GPU 具有 32GB 显存，代码主要使用 PyTorch 库实现。并且，由于本工作采用了生成式重放，因此我们还对生成模型自身的持续学习能力与目前主流的一些方法做了对比。

4.3.1 实验设置

如前文所述，本章主要聚焦于类增量任务，同样采用 ShapeNet 和 ModelNet 两个广泛用于点云分类任务的数据集每个任务涉及两个新类别，整个训练过程包括 20 轮任务。每个点云样本使用 FPS 采样 2048 个点作为输入，并在训练时进行数据增强，包括随机旋转、缩放和高斯噪声扰动，并最后做归一化处理。在此基础上，我们引入 ScanNet[97] 数据集，由于其提供的类别数量较少，因此在保证其他设置相同的情况下，总共进行 8 轮任务的训练。而由于 ScanNet 不是专门用于 3D 物体分类，所以我们通过提取一些具有较多点的类别的对象级点云，每个点云通过上、下采样的方式维持同样的 2048 个点，并转化为类似 ModelNet 的格式进行预处理。所有数据集的训练集、验证集和测试集比例设置为 70%、10% 和 20%，进行类增量分类任务的实验。

另一方面，由于 HyperDiffIL 主要采用了基于生成式重放的方式来对抗灾难性遗忘，性能也相当依赖于生成模型的生成能力，因此我们也对生成器的性能与常见的一些相关方法在 ShapeNet 和 ModelNet40 上进行了对比实验。

为了防止遗忘，每个旧类别的重放样本数量设为 128，择优选取率为 50%，即每轮任务前，扩散模型针对每个旧类别生成 256 个候选样本，最终根据评分算法选取 128 个样本加入重放集合。学习率设置为 0.003，使用 Adam 优化器，批量大小由于显存限制设置为 8，每个任务训练的迭代轮数上限设置为 50 (Epoch)。

框架方面，对于生成器，我们采用在双曲空间下基于类 U-Net 结构的扩散模型；对于分类器，我们在网络输入处添加了一个双曲投影层，并支持模型由第二层进行输入，同时舍弃了双曲特征投影层中保留的欧式空间信息，调整了神经元的

分配，以更好适配对生成模型的引导。

4.3.2 数据集介绍



图 4.2 ScanNet 数据集预览。

ModelNet 和 ShapeNet 数据集已于第三章中简单介绍，这里我们主要介绍 ScanNet 数据集。ScanNet[97] 是一个专门用于三维室内场景理解的大规模数据集，旨在推动计算机视觉和人工智能领域在三维感知方面的发展。该数据集通过 RGB-D 传感器采集真实世界的室内环境，涵盖了各种复杂的室内布局，如卧室、客厅、办公室和教室等，并提供了丰富的标注信息，目前被广泛用于三维目标检测、语义分割、实例分割、场景重建、对象识别等任务，并成为许多先进算法的实验基准。

4.3.3 结果分析

对比实验

如图4.3、4.4、4.5以及表4.2所示，我们选取了第三章的模型 HyMR，以及实验结果当中较为突出的两个模型 I3DOL 和 RPS-Net 与我们的 HyperDiffIL 方法进行对比实验，测试了不同方法在 ModelNet40、ShapeNet 和 ScanNet 数据集上的持续

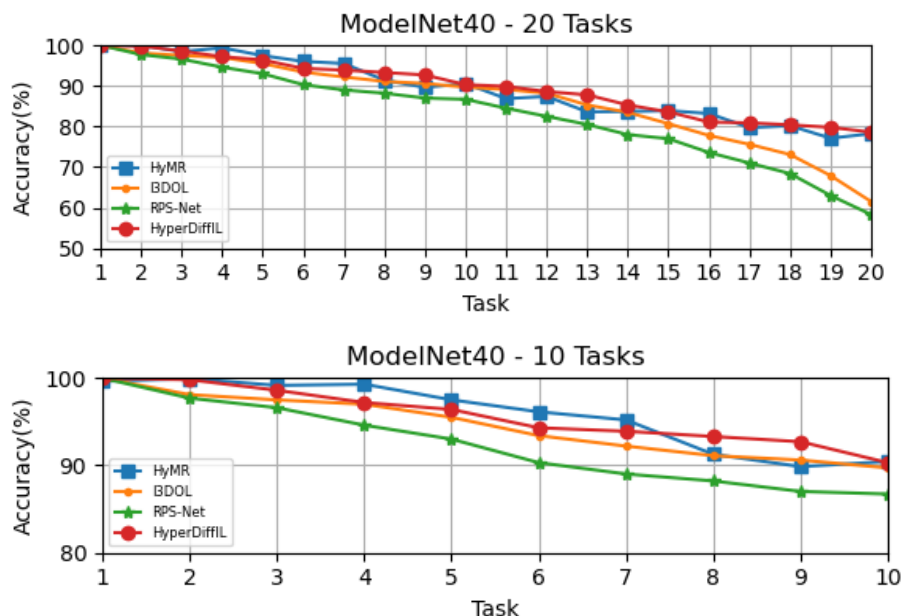


图 4.3 类增量学习设置下，在完成若干个任务后在 ModelNet40 数据集上的整体平均准确率。

表 4.2 在类别增量学习中，不同方法在 ModelNet40、ShapeNet 以及 ScanNet 数据集上经过若干轮训练后的平均准确率 (%)。

数据集	任务数	顺序微调	RPS-Net[96]	I3DOL[19]	HyMR(Ours-缓存式)	HyperDiffIL(Ours-生成式)
ModelNet40	10	13.3	86.7	89.7	90.3	90.2
	20	2.1	58.3	61.5	78.0	78.6
ShapeNet	10	9.7	86.9	87.8	89.8	90.7
	20	1.3	69.5	74.1	80.8	80.9
ScanNet	8	1.8	53.3	56.8	65.9	64.4

学习分类性能，可以看到所有方法的准确率随着任务数增加逐渐下降，表现出灾难性遗忘的现象。

在 ModelNet40 数据集上，HyperDiffIL 和 HyMR 方法在整个训练过程中准确率曲线较为接近，并且保持了较高的准确率，尤其是 HyperDiffIL，在经过 20 轮任务后，最终准确率达到 78.6%，明显优于 RPS-Net(58.3%) 和 I3DOL(61.5%)，说明 HyperDiffIL 在长时间序列任务上的抗遗忘能力更强。虽然 HyMR 方法在经过完整任务训练后的最终准确率略低于 HyperDiffIL，但前 10 任务的总体表现区别不大，甚至略优。在 ShapeNet 数据集上，总体趋势和 ModelNet40 区别不大，并且在

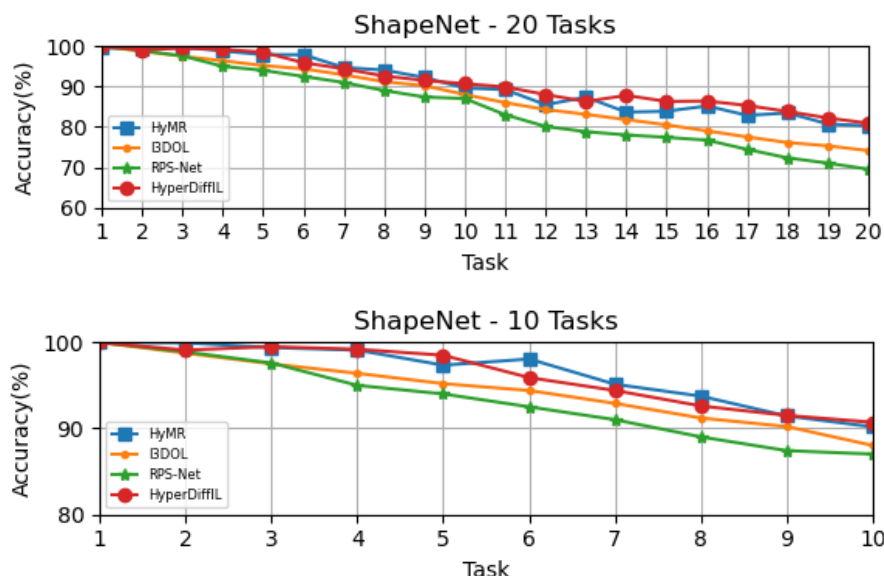


图 4.4 类增量学习设置下,在完成若干个任务后在 ShapeNet 数据集上的整体平均准确率。

经历 10 轮任务后, HyperDiffIL 方法的平均准确率取得了最好的成绩,但是从前 10 轮任务结果趋势来看, HyMR 依旧保持更平滑的水准;而在经过 20 轮任务训练后, HyperDiffIL 优势又重新回归。因此, HyperDiffIL 方法在 ModelNet 和 ShapeNet 上的长任务序列表现最佳,随着任务增多,虽然生成式样本的质量可能下降,但由于不像 HyMR 受限于重放缓存区, HyperDiffIL 凭借多样化的重放内容,仍然保持更好的抗遗忘能力。

而在 ScanNet 数据集上,由于其本身为 3D 真实场景数据,主要用于语义分割,应用到分类任务需要重构,重构后的质量可能不如前两个结构化的合成数据集,因此所有方法在 ScanNet 上经历相同数量任务后的准确率都比 ShapeNet 和 ModelNet40 低。而从整体趋势和最终准确率来看, HyMR 均表现最佳,说明针对这种短任务序列的场景, HyMR 这样精确的重放方法依旧保持优势。

在生成能力的评估上,我们在 ShapeNet 上按类别汽车、椅子、飞机、桌子的顺序,进行单类别任务的类增量训练。我们选取 HyperCloud[98] 作为基准模型顺序微调,并使用非持续学习的联合训练(即将所有类别一次性训练)作为参照。为了进一步验证我们提出方法的有效性,我们还选取了近年来常见的、可应用于点

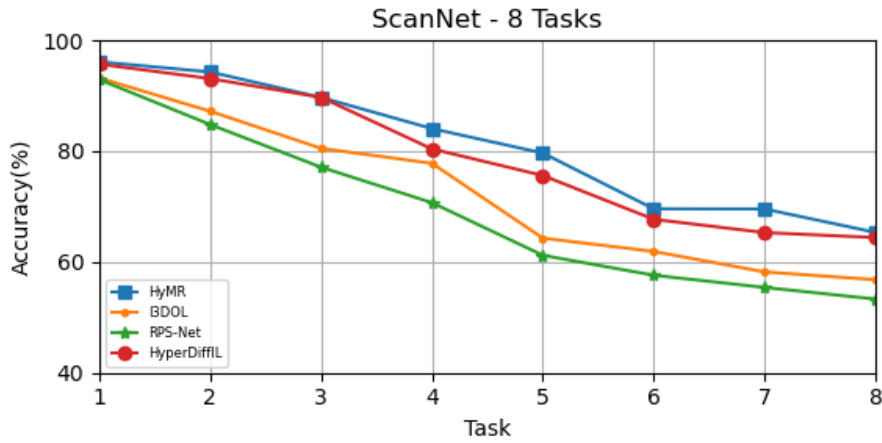


图 4.5 类增量学习设置下，在完成若干个任务后在 ScanNet 数据集上的整体平均准确率。

云生成任务的持续学习方法进行比较。此外，我们加入 L_2 正则化以及固定缓存区大小 (500) 的精准重放 (Exact replay) 作为对比。

在经历了四轮任务训练后，我们获得了如表4.3及图4.6所示的用于评估模型生成能力的各项指标数据及部分方法的生成点云可视化展示。其中，HyperDiffIL 的平均 JSD 值为 0.103，仅次于 PNN，并且与联合训练后的性能相差不大，这表明 HyperDiffIL 生成的点云分布与真实点云分布较为接近，优于大多数持续学习方法；在最小匹配距离上表现不佳，仍有一定优化空间，可能由于生成器基于双曲空间的扩散推导，而忽略了欧氏空间下的几何匹配能力；而在覆盖率和最近邻准确率分析中，HyperDiffIL 在所有持续学习任务方法中表现最佳，生成器更好地覆盖了真实的点云分布，并且与真实数据的可分性更低。

消融实验

为了分析 HyperDiffIL 关键模块对分类性能和生成质量的影响。我们设计了一系列消融实验，通过去除模型中的核心组件或者修改相关参数值，以验证每个模块在生成质量、多样性和抗遗忘能力方面的贡献。消融实验主要围绕双曲空间投影、分类器引导扩散、全过程知识蒸馏以及样本评分机制这四个模块进行，分类任务数据集采用针对类别相关任务质量较高的 ShapeNet 和 ModelNet40，生成任务数据集采用和之前一致的 ShapeNet。

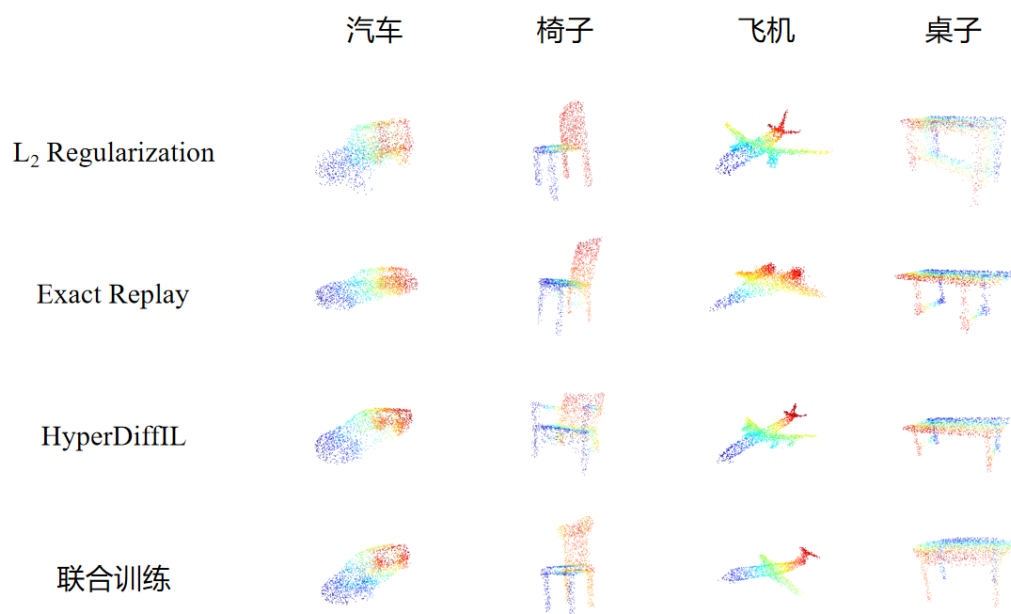


图 4.6 经过四轮单类别增量任务后，不同用于生成模型的持续学习方法的生成结果可视化对比。

对于双曲空间投影，我们取消了模型框架中的双曲指数映射模块以及分类器中的双曲特征提取层，并在欧氏空间中进行生成器中的扩散过程，遵循公式4.13作为训练目标，进行分类和生成任务。针对分类器引导扩散模块，我们在每次新任务重放过程中，取消了公式4.27引入的梯度损失，让生成器随机生成旧样本重放，生成的类别由分类器自主判断，保证生成样本总数与原方法一致。而对生成器的全过程知识蒸馏，我们采取不蒸馏和仅对最终步进行蒸馏作为对照。最后，由于样本评分对重放样本做了筛选，并且设置了选取上限，因此我们依次进行了不评分、不同样本淘汰率、不同样本类别数上限的对比。

首先如图4.7和4.8所示，我们将完整模型和去除双曲映射或分类器模块的模型进行了对比。在去除双曲空间投影后，处理前期任务时的整体平均准确率没有太大的影响，但是随着中期任务和后期任务的到来，欧式空间难以很好表征点云结构的劣势显现出来，双曲空间在长期任务中能够更好地表示点云特征，提升抗遗忘能力。而在去除了分类器的引导模块后，由于任务训练初始阶段的类别数目并不多，

表 4.3 经过四轮单类别增量任务后, 不同用于生成模型的持续学习方法的生成能力指标对比。

方法	JSD($\rightarrow 0$)	MMD($\rightarrow 0$)	COV%($\rightarrow 100\%$)	1-NNA%($\rightarrow 50\%$)
HyperCloud[98]	0.353	16.7	22.7	80.5
EWC[17]	0.288	13.9	27.0	88.6
Generative replay[33]	0.192	15.1	48.4	72.3
L_2 regularization	0.286	14.4	26.3	93.9
Exact replay	0.181	11.6	39.8	77.1
PNN[82]	0.095	12.2	48.4	83.8
HyperDiffIL(Ours)	0.103	13.7	49.1	69.8
联合训练	0.080	10.6	53.5	65.0

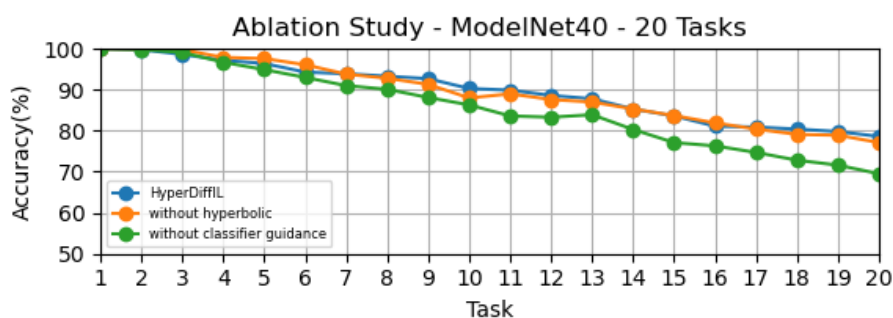


图 4.7 消融实验-双曲映射和分类器引导在 ModelNet40 数据集上完成若干个任务后类增量任务的整体平均准确率。

单靠生成模型随机生成和分类器识别类别的形式足以应对, 但随着类别数目的增多, 在没有分类器引导的情况下, 生成器容易出现一定的生成类别倾向, 导致对另外的类别识别性能下降, 并且错误分类概率增加, 从而出现较大的准确率下滑。

全过程蒸馏和评分机制模块对模型的影响则如表4.4所示。当把生成模型的蒸馏模块去除后, 扩散模型在新任务中仅仅依靠自身生成样本重放来对抗遗忘, 而生成的内容又很大程度上影响了生成器自身和分类器的效果, 因此准确率出现了下降, 尤其是在经历全部 20 轮任务后, 掉点较为明显; 只进行最终步蒸馏, 依旧能保持部分性能, 但仍不及全过程蒸馏, 说明去噪过程中所有时间步的知识一致性对生成质量至关重要。不对生成样本进行质量筛选, 所有样本均用于重放, 最终的

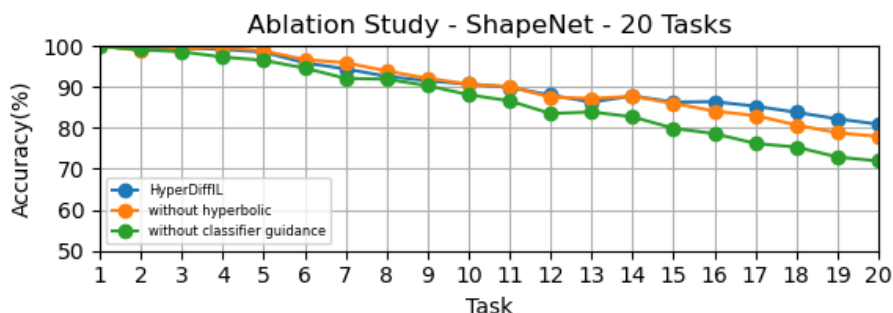


图 4.8 消融实验-双曲映射和分类器引导在 ShapeNet 数据集上完成若干个任务后类增量任务的整体平均准确率。

结果虽然变化不大，但仍有略微下跌，这说明样本评分机制虽然影响较小，但仍能提供一定的质量提升；而从调整了淘汰率和样本上限之后的结果可以看出，合理的样本筛选可以提升旧任务样本的质量，当提升淘汰率，在一定范围内可以提高整体的准确率，提升样本上限和传统缓存区方法一样，同样意味着更多的旧知识的回顾，但是过高的淘汰率和样本上限意味着更长的生成时间和更多的缓存，影响效率的同时还容易陷入低质量样本的陷阱。

由于生成质量是生成式重放模型性能的主要影响源，我们分析了 HyperDiffIL 的扩散模型生成器在不同消融方案下的生成能力，得到如表4.5所示的结果。从各项指标随着不同方案的变化可以看出，双曲空间投影对整体生成质量 (JSD 和 COV) 有积极作用，但对几何匹配 (MMD) 的影响相对较小，可能是由于实验中采用的是最终生成的在欧式空间下的点云进行评估所致，因此双曲映射的主要作用在于增强点云层级结构的表示能力，而不是直接影响点云几何形状匹配；去除蒸馏后，生成样本偏离真实数据分布，导致 1-NN 分类器更容易区分生成样本和真实样本，而针对扩散模型的全过程蒸馏有助于生成样本保持一致性，提高旧类别样本的再现质量，但若仅在最终步骤进行蒸馏虽有一定效果，但不足以完全避免生成样本的退化；评分机制可以提升多样性，但对生成分布影响较小，并且过高的淘汰率在陷入低质量样本的怪圈后，可能导致模式崩溃；而更大的单类别样本上限可以明显提升生成样本的多样性，并减少几何匹配误差，但同样也意味着更高的时间消耗导致的效率低下。

表 4.4 在 ModelNet40 和 ShapeNet 数据集上完成若干轮任务后类增量任务的整体平均准确率受全过程蒸馏和样本评分机制的影响。

数据集	消融方案	准确率 (10 轮任务)	准确率 (20 轮任务)
ModelNet40	HyperDiffIL	90.3	78.6
	不蒸馏	87.63	68.2
	仅最终步蒸馏	91.2	76.0
	不评分筛选	90.1	77.9
	25% 淘汰率	88.7	74.1
	75% 淘汰率	90.8	80.1
	64 样本上限	90.2	78.8
	256 样本上限	91.4	80.3
ShapeNet	HyperDiffIL	90.7	80.9
	不蒸馏	84.4	74.5
	仅最终步蒸馏	90.0	80.9
	不评分筛选	90.3	80.7
	25% 淘汰率	90.0	81.1
	75% 淘汰率	90.8	80.3
	64 样本上限	89.1	74.7
	256 样本上限	90.5	82.2

4.4 章节小结

本章提出了一种基于生成式重放和双曲映射的点云持续学习方法 HyperDiffIL, 结合扩散模型的高质量生成能力以提升旧任务知识的保持效果, 并且避免了传统缓存区方法的存储限制, 并在数据多样性、隐私保护等方面具有适用性。我们的实验结果表明, HyperDiffIL 在点云分类任务中表现良好, 特别是在长任务序列下的分类精度和生成质量方面具有明显优势, 能够达到与 HyMR 相当甚至更优的分类准确率, 但在较短任务序列中, 其分类性能略低于 HyMR 这样基于缓存区的精准重放方法。生成质量评估方面, HyperDiffIL 在部分指标上表现优越, 表明其生成

表 4.5 各种消融方案对 HyperDiffIL 生成器性能的影响。

消融方案	JSD($\rightarrow 0$)	MMD($\rightarrow 0$)	COV%($\rightarrow 100\%$)	1-NNA%($\rightarrow 50\%$)
HyperDiffIL	0.103	13.7	49.1	69.8
无双曲映射	0.111	13.2	48.7	69.9
不蒸馏	0.188	14.3	44.6	71.6
仅最终步蒸馏	0.129	13.9	50.3	72.8
不评分筛选	0.101	13.5	47.6	72.8
25% 淘汰率	0.109	13.9	48.8	70.3
75% 淘汰率	0.125	14.3	51.7	68.4
64 样本上限	0.131	15.6	47.8	73.9
256 样本上限	0.107	13.6	52.3	72.4
联合训练	0.080	10.6	53.5	65.0

样本与真实数据分布较为接近，且具有较高的多样性。此外，我们进行的一系列消融实验证明了双曲空间投影、分类器引导扩散、全过程知识蒸馏、样本评分机制等核心模块对模型的贡献。当然，HyperDiffIL 依旧存在一些缺陷，比如扩散模型计算成本较高，并且针对生成器的抗遗忘控制较少。在未来，我们的研究可以进一步优化生成模型、增强长序列任务的抗遗忘能力，并拓展至多模态任务，以提高其适用性和实际应用价值。

第五章 总结与展望

5.1 研究工作总结

近年来,三维点云数据在计算机视觉和智能感知领域的应用日益广泛。点云数据可以高效地表达三维空间中的物体结构,在自动驾驶、机器人视觉、虚拟现实等方向有着广泛的应用。然而,由于点云数据的不规则性和高维特性,传统的数据处理方法往往难以直接适用。在增量学习的场景下,如何在学习新任务时有效保持旧任务的信息,是当前研究面临的重要挑战。

在此背景下,持续学习成为研究的重点方向之一,目的是使模型能够在不断更新的任务环境中保持知识的完整性,而不会因为新任务的加入而导致旧任务的遗忘。现有的持续学习方法主要依赖于欧式空间进行数据表示,在点云数据处理中存在一定的局限性。故针对此问题,本文深入研究了基于双曲流形映射的方法,以适应点云数据的层级化特征,提高模型在长期学习任务中的稳定性。

首先,针对缓存重放策略,本文提出了一种适用于点云的基于双曲流形的存储管理优化方法。该方法利用双曲几何的特性,使得点云数据在存储时能够进行更紧凑的表示,同时提升重放过程中样本的代表性。通过引入贪心采样策略,本文进一步优化了数据存储选择,使得关键样本的存储优先级更高。实验结果表明,该方法在 ModelNet40、ShapeNet 这些主流合成数据集上能够有效减少遗忘效应,并提高平均分类精度。该策略的优势在于存储效率高,能够在有限的存储空间内保持较优的重放效果。然而,其性能仍然受到存储容量以及计算效率的限制,在长任务序列的情况下可能出现样本不足以及效率低下的问题。

另一方面,由于考虑到缓存重放策略在存储资源受限时的适用性问题,本文进一步探讨了另一种可行的解决方案,即生成式重放方法。在缓存重放策略的基础上,本文提出了一种适用于点云的基于双曲流形的生成式重放方法,以减少对存储空间的依赖。模型能够在双曲空间中生成符合流形分布的点云样本,从而提供高质量的知识重放,而无需存储大量旧任务数据。实验结果表明,虽然生成模型的训练和推理计算量较大,但在长线的任务处理及生成质量上时表现更为出色,并

展现了较强的泛化能力。不过由于其训练逻辑和分类器之间存在强依赖关系，可能面临部分细节丢失的问题，依旧无法完全克服灾难性遗忘。

综上所述，本文围绕点云数据的持续学习问题，结合缓存重放和生成式重放策略，探索了双曲流形映射在增量学习中的应用。实验结果表明，提出的两种方法分别适用于短线、非存储受限的任务和长线、存储受限的任务，在任务中都能够有效降低灾难性遗忘，提高模型的持续学习能力，为点云数据的持续学习提供了一种新的思路。

5.2 研究工作展望

虽然本文的方法在点云持续学习领域取得了一定的进展，但仍有许多值得进一步研究的方向。当前的持续学习方法主要侧重于减少遗忘，而在如何提高知识的前向传递能力方面仍有很大探索空间。未来的研究可以进一步考虑如何更充分地利用点云数据的几何特性，使得新任务的学习可以从旧任务中有效迁移知识，像真正的教师-学生一样，获得举一反三的能力，从而提升整体的学习效率而不是被动地依赖旧模型的正则化和重放。另一方面，当前的双曲流形映射方法主要依赖于传统的梯度优化策略，而针对流形空间的优化技术仍有改进空间。未来可以考虑自适应流形优化方法，如基于黎曼优化的元学习方法，以提高模型对不同任务的适应性。

此外，从第三章的工作可以看出，现有的方法在任务增量学习场景下表现较好，但在类增量学习场景中仍然存在较大的挑战。由于点云数据的类别间表征存在重叠，导致新任务的加入可能影响旧任务的分类效果。因此，未来的研究可以结合自监督学习或能量建模方法，以减少类别间的表征冲突，提高类增量学习的适应能力。

在生成式重放方面，尽管本文的方法减少了对存储空间的依赖，但生成器和分类器之间的强依赖关系导致性能难以进一步提升，并且扩散模型仍然面临计算开销较大的问题。未来的研究可以探索更高效的生成架构，如利用 DDIM[99] 跨步生成，优化扩散模型的计算复杂度，提高生成样本的质量；也可以考虑降低生成器和分类器之间的依赖，如自监督生成、多视角生成等。同时，可以结合图神经网络

(GNN) 或神经隐式表示方法, 以进一步增强点云数据的生成保真度。

在存储限制方面, 本文提出的方法主要依赖于重放缓存区以及生成模型, 这两者对于存储都有一定的规模要求。而近年来, Prompt 机制在自然语言处理和计算机视觉领域取得了一定的研究进展。已经有针对引入 prompt 机制的持续学习工作 [59] 被提出, 由于 prompt 本身只是一个提示词向量, 占用内存空间很小, 因此可以用来解放缓存区内存。未来的研究可以探索如何为点云数据设计合适的 Prompt 机制, 并结合双曲流形映射方法, 以提升跨任务的知识迁移能力。

最后, 点云持续学习方法的应用场景仍然值得进一步拓展。目前的研究主要关注点云分类任务, 而在点云分割、目标检测等更复杂的任务中, 持续学习仍然面临较大挑战。未来还可以考虑如何将双曲流形映射拓展到这些任务, 以提升模型在实际应用中的适应能力。此外, 持续学习在自动驾驶、机器人感知、医学影像分析等领域同样具有重要应用价值, 未来可以结合多模态学习方法, 探索点云数据与图像、文本等多源信息的联合持续学习策略, 从而将研究的成果应用到这些更加具体的场景上。

总的来说, 点云持续学习仍然是一个值得深入探索的研究方向。本文的研究虽然在一定程度上缓解了灾难性遗忘问题, 但仍然存在诸多可以优化的方面。未来的研究可以围绕更高效的存储管理、生成式建模、流形优化以及跨模态学习等方面展开, 以进一步提升点云数据持续学习的能力, 为智能感知系统的发展提供有力支持。

参考文献

- [1] JIANG E. A review of the comparative studies on traditional and intelligent face recognition methods[C] // 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL). 2020 : 11 – 15.
- [2] GRIGORESCU S, TRASNEA B, COCIAS T, et al. A survey of deep learning techniques for autonomous driving[J]. Journal of field robotics, 2020, 37(3) : 362 – 386.
- [3] GOU F, LIU J, XIAO C, et al. Research on artificial-intelligence-assisted medicine: A survey on medical artificial intelligence[J]. Diagnostics, 2024, 14(14) : 1472.
- [4] LOPEZ M M, KALITA J. Deep Learning applied to NLP[J]. arXiv preprint arXiv:1703.03091, 2017.
- [5] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[J]. Advances in neural information processing systems, 2012, 25.
- [6] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C] // Proceedings of the IEEE conference on computer vision and pattern recognition. 2016 : 770 – 778.
- [7] XIONG W, WU L, ALLEVA F, et al. The Microsoft 2017 conversational speech recognition system[C] // 2018 IEEE international conference on acoustics, speech and signal processing (ICASSP). 2018 : 5934 – 5938.
- [8] BEKELE M K, PIERDICCA R, FRONTONI E, et al. A survey of augmented, virtual, and mixed reality for cultural heritage[J]. Journal on Computing and Cultural Heritage (JOCCH), 2018, 11(2) : 1 – 36.

- [9] XIAO Y, CODEVILLA F, GURRAM A, et al. Multimodal end-to-end autonomous driving[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 23(1): 537–547.
- [10] SHI W, RAJKUMAR R. Point-gnn: Graph neural network for 3d object detection in a point cloud[C] // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020 : 1711–1719.
- [11] RUSU R B, MARTON Z C, BLODOW N, et al. Towards 3D point cloud based object maps for household environments[J]. Robotics and Autonomous Systems, 2008, 56(11): 927–941.
- [12] CHEN L, WU P, CHITTA K, et al. End-to-end autonomous driving: Challenges and frontiers[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024.
- [13] SHE Q, FENG F, HAO X, et al. Openloris-object: A dataset and benchmark towards lifelong object recognition[J]. arXiv preprint arXiv:1911.06487, 2019.
- [14] MCCLOSKEY M, COHEN N J. Catastrophic interference in connectionist networks: The sequential learning problem[G] // Psychology of learning and motivation : Vol 24. [S.l.]: Elsevier, 1989 : 109–165.
- [15] LI Z, HOIEM D. Learning without forgetting[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 40(12): 2935–2947.
- [16] ZENKE F, POOLE B, GANGULI S. Continual learning through synaptic intelligence[C] // International conference on machine learning. 2017 : 3987–3995.
- [17] KIRKPATRICK J, PASCANU R, RABINOWITZ N, et al. Overcoming catastrophic forgetting in neural networks[J]. Proceedings of the national academy of sciences, 2017, 114(13): 3521–3526.

- [18] SERRA J, SURIS D, MIRON M, et al. Overcoming catastrophic forgetting with hard attention to the task[C] // International conference on machine learning. 2018 : 4548–4557.
- [19] DONG J, CONG Y, SUN G, et al. I3dol: Incremental 3d object learning without catastrophic forgetting[C] // Proceedings of the AAAI Conference on Artificial Intelligence : Vol 35. 2021 : 6066–6074.
- [20] ZAMORSKI M, STYPUŁKOWSKI M, KARANOWSKI K, et al. Continual learning on 3D point clouds with random compressed rehearsal[J]. Computer Vision and Image Understanding, 2023, 228 : 103621.
- [21] QI C R, SU H, MO K, et al. Pointnet: Deep learning on point sets for 3d classification and segmentation[C] // Proceedings of the IEEE conference on computer vision and pattern recognition. 2017 : 652–660.
- [22] QI C R, YI L, SU H, et al. Pointnet++: Deep hierarchical feature learning on point sets in a metric space[J]. Advances in neural information processing systems, 2017, 30.
- [23] LIN L-J, MITCHELL T M. Memory approaches to reinforcement learning in non-Markovian domains[M]. [S.l.] : Citeseer, 1992.
- [24] LOPEZ-PAZ D, RANZATO M. Gradient episodic memory for continual learning[J]. Advances in neural information processing systems, 2017, 30.
- [25] CHAUDHRY A, RANZATO M, ROHRBACH M, et al. Efficient lifelong learning with a-gem[J]. arXiv preprint arXiv:1812.00420, 2018.
- [26] RIEMER M, CASES I, AJEMIAN R, et al. Learning to learn without forgetting by maximizing transfer and minimizing interference[J]. arXiv preprint arXiv:1810.11910, 2018.

- [27] ALJUNDI R, BELILOVSKY E, TUYTELAARS T, et al. Online continual learning with maximal interfered retrieval[J]. Advances in neural information processing systems, 2019, 32.
- [28] REBUFFI S-A, KOLESNIKOV A, SPERL G, et al. icarl: Incremental classifier and representation learning[C] // Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. 2017 : 2001 – 2010.
- [29] HOU S, PAN X, LOY C C, et al. Learning a unified classifier incrementally via rebalancing[C] // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019 : 831 – 839.
- [30] AHN H, KWAK J, LIM S, et al. Ss-il: Separated softmax for incremental learning[C] // Proceedings of the IEEE/CVF International conference on computer vision. 2021 : 844 – 853.
- [31] JOSEPH K, KHAN S, KHAN F S, et al. Energy-based latent aligner for incremental learning[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022 : 7452 – 7461.
- [32] WANG F-Y, ZHOU D-W, YE H-J, et al. Foster: Feature boosting and compression for class-incremental learning[C] // European conference on computer vision. 2022 : 398 – 414.
- [33] SHIN H, LEE J K, KIM J, et al. Continual learning with deep generative replay[J]. Advances in neural information processing systems, 2017, 30.
- [34] RIEMER M, KLINGER T, BOUNEFFOUF D, et al. Scalable recollections for continual lifelong learning[C] // Proceedings of the AAAI conference on artificial intelligence : Vol 33. 2019 : 1352 – 1359.
- [35] AYUB A, WAGNER A R. Eec: Learning to encode and regenerate images for continual learning[J]. arXiv preprint arXiv:2101.04904, 2021.

- [36] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[J]. Advances in neural information processing systems, 2014, 27.
- [37] LIU X, WU C, MENTA M, et al. Generative feature replay for class-incremental learning[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. 2020 : 226 – 227.
- [38] WU C, HERRANZ L, LIU X, et al. Memory replay gans: Learning to generate new categories without forgetting[J]. Advances in neural information processing systems, 2018, 31.
- [39] YE F, BORS A G. Learning latent representations across multiple data domains using lifelong VAEGAN[C] // Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XX 16. 2020 : 777 – 795.
- [40] HO J, JAIN A, ABBEEL P. Denoising diffusion probabilistic models[J]. Advances in neural information processing systems, 2020, 33 : 6840 – 6851.
- [41] GAO R, LIU W. Ddgr: Continual learning with deep diffusion-based generative replay[C] // International Conference on Machine Learning. 2023 : 10744 – 10763.
- [42] MENG Z, ZHANG J, YANG C, et al. Diffclass: Diffusion-based class incremental learning[C] // European Conference on Computer Vision. 2024 : 142 – 159.
- [43] ZHU K, ZHAI W, CAO Y, et al. Self-sustaining representation expansion for non-exemplar class-incremental learning[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022 : 9296 – 9305.
- [44] PETIT G, POPESCU A, SCHINDLER H, et al. Fetril: Feature translation for exemplar-free class-incremental learning[C] // Proceedings of the IEEE/CVF winter conference on applications of computer vision. 2023 : 3911 – 3920.

- [45] CHAUDHRY A, DOKANIA P K, AJANTHAN T, et al. Riemannian walk for incremental learning: Understanding forgetting and intransigence[C] // Proceedings of the European conference on computer vision (ECCV). 2018 : 532–547.
- [46] ALJUNDI R, BABILONI F, ELHOSEINY M, et al. Memory aware synapses: Learning what (not) to forget[C] // Proceedings of the European conference on computer vision (ECCV). 2018 : 139–154.
- [47] DHAR P, SINGH R V, PENG K-C, et al. Learning without memorizing[C] // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019 : 5138–5146.
- [48] ZHAI M, CHEN L, TUNG F, et al. Lifelong gan: Continual learning for conditional image generation[C] // Proceedings of the IEEE/CVF international conference on computer vision. 2019 : 2759–2768.
- [49] KANAKIS M, BRUGGEMANN D, SAHA S, et al. Reparameterizing convolutions for incremental multi-task learning without task interference[C] // Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XX 16. 2020 : 689–707.
- [50] RUSU A A, RABINOWITZ N C, DESJARDINS G, et al. Progressive neural networks[J]. arXiv preprint arXiv:1606.04671, 2016.
- [51] FERNANDO C, BANARSE D, BLUNDELL C, et al. Pathnet: Evolution channels gradient descent in super neural networks[J]. arXiv preprint arXiv:1701.08734, 2017.
- [52] MALLYA A, LAZEBNIK S. Packnet: Adding multiple tasks to a single network by iterative pruning[C] // Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. 2018 : 7765–7773.

- [53] FARAJTABAR M, AZIZAN N, MOTT A, et al. Orthogonal gradient descent for continual learning[C] // International Conference on Artificial Intelligence and Statistics. 2020 : 3762–3773.
- [54] CHA H, LEE J, SHIN J. Co2l: Contrastive continual learning[C] // Proceedings of the IEEE/CVF International conference on computer vision. 2021 : 9516–9525.
- [55] MADAAN D, YOON J, LI Y, et al. Representational continuity for unsupervised continual learning[J]. arXiv preprint arXiv:2110.06976, 2021.
- [56] PHAM Q, LIU C, HOI S. Dualnet: Continual learning, fast and slow[J]. Advances in Neural Information Processing Systems, 2021, 34 : 16131–16144.
- [57] ERMIS B, ZAPPELLA G, WISTUBA M, et al. Memory efficient continual learning with transformers[J]. Advances in Neural Information Processing Systems, 2022, 35 : 10629–10642.
- [58] VASWANI A. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017.
- [59] WANG Z, ZHANG Z, LEE C-Y, et al. Learning to prompt for continual learning[C] // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022 : 139–149.
- [60] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11) : 2278–2324.
- [61] MATURANA D, SCHERER S. Voxnet: A 3d convolutional neural network for real-time object recognition[C] // 2015 IEEE/RSJ international conference on intelligent robots and systems (IROS). 2015 : 922–928.
- [62] RIEGLER G, OSMAN ULUSOY A, GEIGER A. Octnet: Learning deep 3d representations at high resolutions[C] // Proceedings of the IEEE conference on computer vision and pattern recognition. 2017 : 3577–3586.

- [63] BEN-SHABAT Y, LINDENBAUM M, FISCHER A. 3d point cloud classification and segmentation using 3d modified fisher vector representation for convolutional neural networks[J]. arXiv preprint arXiv:1711.08241, 2017.
- [64] SU H, MAJI S, KALOGERAKIS E, et al. Multi-view convolutional neural networks for 3d shape recognition[C] // Proceedings of the IEEE international conference on computer vision. 2015 : 945 – 953.
- [65] YANG Z, WANG L. Learning relationships for multi-view 3D object recognition[C] // Proceedings of the IEEE/CVF international conference on computer vision. 2019 : 7505 – 7514.
- [66] WANG Y, SUN Y, LIU Z, et al. Dynamic graph cnn for learning on point clouds[J]. ACM Transactions on Graphics (tog), 2019, 38(5) : 1 – 12.
- [67] ZHAO H, JIANG L, JIA J, et al. Point transformer[C] // Proceedings of the IEEE/CVF international conference on computer vision. 2021 : 16259 – 16268.
- [68] BOULCH A, LE SAUX B, AUDEBERT N, et al. Unstructured point cloud semantic labeling using deep segmentation networks.[J]. 3dor@ eurographics, 2017, 3 : 1 – 8.
- [69] LAWIN F J, DANELLJAN M, TOSTEBERG P, et al. Deep projective 3D semantic segmentation[C] // Computer Analysis of Images and Patterns: 17th International Conference, CAIP 2017, Ystad, Sweden, August 22-24, 2017, Proceedings, Part I 17. 2017 : 95 – 107.
- [70] HUANG J, YOU S. Point cloud labeling using 3d convolutional neural network[C] // 2016 23rd International Conference on Pattern Recognition (ICPR). 2016 : 2670 – 2675.
- [71] TCHAPMI L, CHOY C, ARMENI I, et al. Segcloud: Semantic segmentation of 3d point clouds[C] // 2017 international conference on 3D vision (3DV). 2017 : 537 – 547.

- [72] SHI Y, CHANG A X, WU Z, et al. Hierarchy denoising recursive autoencoders for 3D scene layout prediction[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019 : 1771 – 1780.
- [73] WANG W, YU R, HUANG Q, et al. Sgpn: Similarity group proposal network for 3d point cloud instance segmentation[C] // Proceedings of the IEEE conference on computer vision and pattern recognition. 2018 : 2569 – 2578.
- [74] HU S-M, CAI J-X, LAI Y-K. Semantic labeling and instance segmentation of 3D point clouds using patch context analysis and multiscale processing[J]. IEEE transactions on visualization and computer graphics, 2018, 26(7) : 2485 – 2498.
- [75] ZHU C, XU K, CHAUDHURI S, et al. CoSegNet: Deep co-segmentation of 3D shapes with group consistency loss[J]. arXiv preprint arXiv:1903.10297, 2019, 3.
- [76] WANG Z, LU F. Voxsegnet: Volumetric cnns for semantic part segmentation of 3d shapes[J]. IEEE transactions on visualization and computer graphics, 2019, 26(9) : 2919 – 2930.
- [77] KU J, MOZIFIAN M, LEE J, et al. Joint 3d proposal generation and object detection from view aggregation[C] // 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2018 : 1 – 8.
- [78] LIANG M, YANG B, WANG S, et al. Deep continuous fusion for multi-sensor 3d object detection[C] // Proceedings of the European conference on computer vision (ECCV). 2018 : 641 – 656.
- [79] SHI S, WANG X, LI H. Pointtrnn: 3d object proposal generation and detection from point cloud[C] // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019 : 770 – 779.

- [80] GIANCOLA S, ZARZAR J, GHANEM B. Leveraging shape completion for 3d siamese tracking[C] // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019 : 1359–1368.
- [81] CHOWDHURY T, JALISHA M, CHERAGHIAN A, et al. Learning without forgetting for 3d point cloud objects[C] // Advances in Computational Intelligence: 16th International Work-Conference on Artificial Neural Networks, IWANN 2021, Virtual Event, June 16–18, 2021, Proceedings, Part I 16. 2021 : 484–497.
- [82] SADOWSKI M, PICZAK K J, SPUREK P, et al. Continual Learning of 3D Point Cloud Generators[C] // Neural Information Processing: 28th International Conference, ICONIP 2021, Sanur, Bali, Indonesia, December 8–12, 2021, Proceedings, Part I 28. 2021 : 330–341.
- [83] WANG X, WEI X. Continual Learning for Pose-Agnostic Object Recognition in 3D Point Clouds[J]. arXiv preprint arXiv:2209.04840, 2022.
- [84] CHOWDHURY T, CHERAGHIAN A, RAMASINGHE S, et al. Few-shot class-incremental learning for 3d point cloud objects[C] // European Conference on Computer Vision. 2022 : 204–220.
- [85] ROLNICK D, AHUJA A, SCHWARZ J, et al. Experience replay for continual learning[J]. Advances in neural information processing systems, 2019, 32.
- [86] LI X, ZHOU Y, WU T, et al. Learn to grow: A continual structure learning framework for overcoming catastrophic forgetting[C] // International conference on machine learning. 2019 : 3925–3934.
- [87] SARKAR R. Low distortion delaunay embedding of trees in hyperbolic plane[C] // International symposium on graph drawing. 2011 : 355–366.

- [88] PENG W, VARANKA T, MOSTAFA A, et al. Hyperbolic deep neural networks: A survey[J]. IEEE Transactions on pattern analysis and machine intelligence, 2021, 44(12): 10023–10044.
- [89] MONTANARO A, VALSESIA D, MAGLI E. Rethinking the compositionality of point clouds through regularization in the hyperbolic space[J]. Advances in Neural Information Processing Systems, 2022, 35: 33741–33753.
- [90] LIU Y, ZHANG P, HE Z, et al. Hyperbolic Contrastive Learning for Hierarchical 3D Point Cloud Embedding[J]. arXiv preprint arXiv:2501.02285, 2025.
- [91] METTES P, GHADIMI ATIGH M, KELLER-RESSEL M, et al. Hyperbolic deep learning in computer vision: A survey[J]. International Journal of Computer Vision, 2024, 132(9): 3484–3508.
- [92] HU N, CHENG H, XIE Y, et al. Hyperbolic Image-and-Pointcloud Contrastive Learning for 3D Classification[C] // 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2024: 4973–4979.
- [93] DE LANGE M, ALJUNDI R, MASANA M, et al. A continual learning survey: Defying forgetting in classification tasks[J]. IEEE transactions on pattern analysis and machine intelligence, 2021, 44(7): 3366–3385.
- [94] WU Z, SONG S, KHOSLA A, et al. 3d shapenets: A deep representation for volumetric shapes[C] // Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 1912–1920.
- [95] CHANG A X, FUNKHOUSER T, GUIBAS L, et al. Shapenet: An information-rich 3d model repository[J]. arXiv preprint arXiv:1512.03012, 2015.
- [96] RAJASEGARAN J, HAYAT M, KHAN S H, et al. Random path selection for continual learning[J]. Advances in neural information processing systems, 2019, 32.

-
- [97] DAI A, CHANG A X, SAVVA M, et al. Scannet: Richly-annotated 3d reconstructions of indoor scenes[C] // Proceedings of the IEEE conference on computer vision and pattern recognition. 2017 : 5828 – 5839.
- [98] SPUREK P, WINCZOWSKI S, TABOR J, et al. Hypernetwork approach to generating point clouds[J]. arXiv preprint arXiv:2003.00802, 2020.
- [99] SONG J, MENG C, ERMON S. Denoising diffusion implicit models[J]. arXiv preprint arXiv:2010.02502, 2020.

致 谢

时光荏苒，硕士生涯即将落下帷幕。这段旅程充满挑战与收获，我的成长离不开众多师长、家人和朋友的支持与帮助，在此深表感激。

首先，衷心感谢我的导师魏宪老师，不仅在学术上给予了我悉心指导，引领我进入几何深度学习领域，更在科研方法、论文撰写、实验分析等方面提供了宝贵的建议。导师的严谨治学态度和独到学术见解，使我在科研道路上少走了许多弯路，也让我逐步培养了独立思考与研究的能力。在此，我由衷地感谢导师的辛勤付出和无私教诲。

其次，感谢我的父母，他们始终是我坚强的后盾，给予我无微不至的关怀和坚定的支持，使我能专注于学术研究，在面对挑战时保持信心。

同时，我要感谢我的师兄徐梓豪，以及温凌峰、程继均、王昊等同门，在实验、论文写作和学术讨论中给予的帮助，使我得以不断进步。实验室的良好氛围也让我受益匪浅，倍感温暖。

此外，我要特别感谢我的室友李捷、王鑫轲、段昌峰，给予了我许多关心与帮助。无论是学习上的交流，还是日常琐事的互相照应，大家共同营造了一个温馨而融洽的生活环境，让我在紧张的科研之余能够放松心情、调整状态。与你们的相处时光，是我硕士生活中一段温暖而美好的回忆。

最后，我还要感谢我的女朋友孙苡玮，在这段求学的时光里，你始终陪伴左右，无论是科研压力大的时候，还是遇到挫折和迷茫时，你总能给予我安慰与鼓励。你的支持让我在繁忙的研究生活中感受到温暖，也让我更加坚定地朝着目标努力前行。

硕士求学虽近尾声，但这段经历将成为我人生中的宝贵财富。我将铭记所得，勇敢前行，迎接未来的挑战。

俞津

二〇二五年三月于普陀校区

攻读硕士学位期间科研情况

■ 学术论文

■ 以第一作者身份完成的学术论文

[1] Yu J, Xu Z, Li Z, et al. **Continual Learning of 3D Point Cloud with Hyperbolic Manifold Replay**[C]//2024 IEEE 36th International Conference on Tools with Artificial Intelligence (ICTAI). IEEE, 2024: 206-212.

