

# DFE3D: 双重特征增强的三维点云类增量学习<sup>①</sup>



孙昊<sup>1</sup>, 帅惠<sup>2</sup>, 许翔<sup>3</sup>, 刘青山<sup>4</sup>

<sup>1</sup>(南京信息工程大学 计算机学院, 南京 210044)

<sup>2</sup>(南京邮电大学 计算机学院、软件学院、网络空间安全学院, 南京 210023)

<sup>3</sup>(南京航空航天大学, 南京 211106)

<sup>4</sup>(南京邮电大学, 南京 210023)

通信作者: 刘青山, E-mail: [qslu@njupt.edu.cn](mailto:qslu@njupt.edu.cn)

**摘要:** 随着点云采集技术的发展和三维应用需求的增加, 实际场景要求针对流动数据持续动态地更新点云分析网络. 对此, 提出了双重特征增强的三维点云类增量学习方法, 通过增量学习使点云目标分类技术能够适应新数据中不断出现新类别目标的场景. 该方法通过对点云数据特性和旧类信息的研究分别提出了差异性局部增强模块和知识注入网络, 以缓解类增量学习中的新类偏好问题. 具体而言, 差异性局部增强模块通过感知丰富的局部语义, 表征出三维点云物体中不同的局部结构特性. 随后, 根据目标中每个局部结构的全局信息获得各个局部的重要性权重, 强化对差异性局部特征的感知, 从而提高新旧类特征差异性. 另外, 知识注入网络将旧模型中的旧知识注入新模型的特征学习过程中, 增强后的混合特征能够更有效缓解旧类信息不足导致的新类偏好加剧现象. 在三维点云数据集 ModelNet40, ScanObjectNN, ScanNet, ShapeNet 上的实验表明, 该方法与现有最优方法相比, 在 4 个数据集上的平均增量准确率有 2.03%、2.18%、1.65%、1.28% 提升.

**关键词:** 三维点云目标分类; 类增量学习; 差异性局部增强; 知识注入; 灾难性遗忘; 点云表征

引用格式: 孙昊, 帅惠, 许翔, 刘青山. DFE3D: 双重特征增强的三维点云类增量学习. 计算机系统应用, 2024, 33(8): 132-144. <http://www.c-s-a.org.cn/1003-3254/9574.html>

## DFE3D: Class-incremental Learning for 3D Point Cloud with Dual Feature Enhancement

SUN Hao<sup>1</sup>, SHUAI Hui<sup>2</sup>, XU Xiang<sup>3</sup>, LIU Qing-Shan<sup>4</sup>

<sup>1</sup>(School of Computer Science, Nanjing University of Information Science and Technology, Nanjing 210044, China)

<sup>2</sup>(School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

<sup>3</sup>(Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

<sup>4</sup>(Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

**Abstract:** As point cloud acquisition technology develops and the demand for 3D applications increases, real-world scenarios require continuous and dynamic updating of the point cloud analysis network with streaming data. This study proposes a dual feature enhancement for the class-incremental 3D point cloud object learning method, which adapts point cloud object classification to scenarios where new category objects keep emerging in newly acquired data through incremental learning. This study proposes a discriminative local enhancement module and knowledge injection network respectively to alleviate new class bias problems in class-incremental learning by studying the characteristics of point cloud data and old class information. Specifically, the discriminative local enhancement module characterizes the various local structural characteristics of 3D point cloud objects by perceiving expressive local features. Subsequently, the importance weights of each local structure are obtained based on the global information of each local structure, enhancing the perception of differential local features and improving the differentiation of new and old class features. Furthermore, the knowledge injection network injects old knowledge from the old model into the feature learning process of the new

① 基金项目: 科技创新 2030—“新一代人工智能”重大项目 (2021ZD0112200)

收稿时间: 2024-02-05; 修改时间: 2024-03-05; 采用时间: 2024-03-11; csa 在线出版时间: 2024-06-28

CNKI 网络首发时间: 2024-07-02

model. The enhanced hybrid features can effectively mitigate the increased new class bias caused by the lack of old class information. Under the incremental learning experimental settings of the 3D point cloud datasets ModelNet40, ScanObjectNN, ScanNet, and ShapeNet, extensive experiments show that compared with existing state-of-art methods, the method in this study has an average incremental accuracy improvement of 2.03%, 2.18%, 1.65%, and 1.28% on the four datasets.

**Key words:** 3D point cloud object classification; class-incremental learning; discriminative local enhancement; knowledge injection; catastrophic forgetting; point cloud representation

三维点云目标分类技术是构建三维世界感知框架的基础,在自动驾驶、智能机器人、增强现实等诸多领域得到广泛运用.大多数基于深度学习的点云目标分类技术<sup>[1-5]</sup>都要在获得所有训练数据后进行离线训练,要求数据集满足独立同分布的条件.然而,由于三维点云数据的迭代增长和三维应用需求的不断扩展,现实场景中数据通常是按次序分批到达,并且不同批次数据的分布存在差异.对此,常规方法需要在每个批次数据到达时,联合所有批次的数据重新训练,以提升模型的能力,但是联合训练面临着巨大的计算资源消耗.此外,物理内存大小的限制、数据隐私安全的制约等因素会导致过去批次的完整数据不可重用.因此,如何在数据访问受限且分批到达的场景下完成训练任务是现实应用中亟待解决的问题.

为了缓解过去批次数据访问受限造成的灾难性遗忘<sup>[6]</sup>问题,增量学习设计渐进式学习的人工神经网络去模拟人类认知系统<sup>[7]</sup>.具体而言,就是通过模型结构和训练策略的设计,防止神经网络在内容不同的新批次数据上学习新知识时忘记旧数据上学习的旧知识.在自然图像理解领域,大量的增量学习工作<sup>[8-11]</sup>已经取得了巨大的成功.近些年,三维点云增量学习也逐渐得到关注.L3DOC<sup>[12]</sup>首次将点云目标分类和增量学习相结合,将深度网络逐层解耦为负责旧知识的点知识部分和负责新知识的任务特异部分,特定的更新策略使模型学习达到较好的可塑性-稳定性的平衡<sup>[13]</sup>.其后,一些工作关注增量学习中更加困难的类增量学习问题,即新批次中仅包含新类别数据,其过程如图1所示.其中,I3DOL<sup>[14]</sup>使用局部质心重构模块自适应地选择特征学习过程中局部感知的区域,以缓解点云类增量学习过程中的表征不规则问题.RCR<sup>[15]</sup>提出了点云随机压缩回放,仅用极小的内存代价保留了用于回放的旧样例集,并引入重建损失作为正则项约束模型优化.

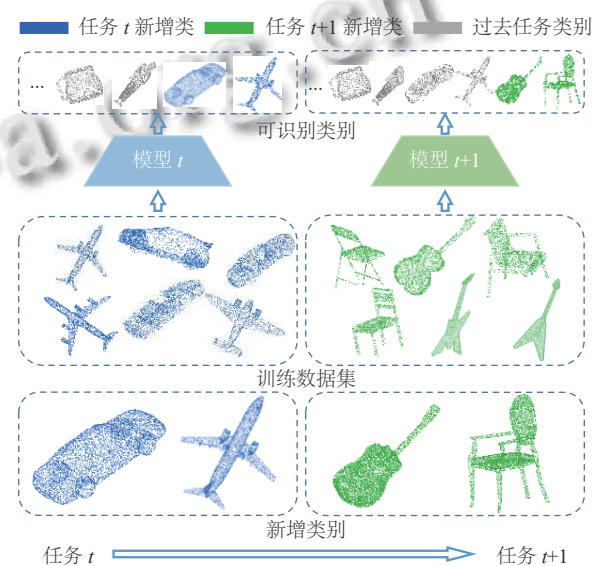


图1 点云目标分类的类增量学习过程示意图

尽管上述方法通过设计复杂的点云网络结构和旧样例回放方法推动了点云类增量学习的发展,但仍存在不足.如图2, iCaRL<sup>[8]</sup>和 I3DOL 在最后增量阶段对旧类目标分类错误,产生遗忘现象的原因是类别增量后的学习阶段本质上是一个新旧类样本不平衡的学习过程.

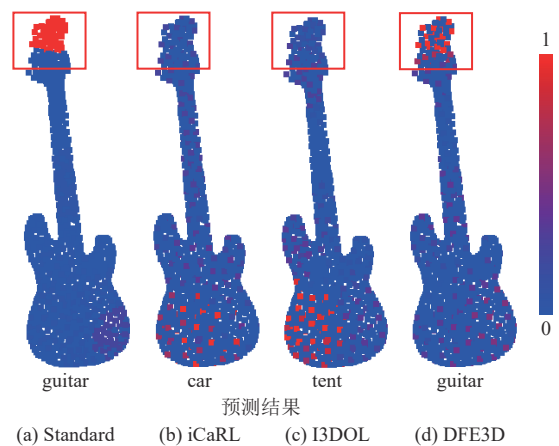


图2 不同方法的关注点示例图

该过程中新类训练样本远多于旧样例,使模型过拟合于新类特征,而对旧类特征的关注不足,最终造成新旧类特征缺少区分度,模型易于将旧类混淆为新类。

因此,要缓解类增量学习中的灾难性遗忘问题,关键在于缓解对新类的学习偏好问题。首先,是要增强点云特征学习过程中对局部差异性特征的辨别能力。由于点云数据仅包含三维坐标,没有纹理等信息,因此点云分类主要依赖于局部结构信息。例如,图2中吉他琴头等有代表性的局部结构包含差异性特征,有利于类别区分。而平面等广泛存在于各类目标中的局部结构则对应类别信息较少的相似性特征。然而,不平衡的样本数量意味着旧类即使分类错误带来的惩罚也极小,使新类与旧类特征趋向相似性局部结构。所以,为了捕获有区分度的特征,应当去关注类别表征明显的差异性局部结构。其次,旧类信息的不足直接导致了偏好问题加剧,尽管通过回放旧样例集可以直接回顾到旧类知识,但是有限的缓存大小使各个旧类的可重用样本数量被不断压缩,导致各类的旧信息不断减少,并且这种非结构化知识的质量依赖于精心设计的旧样例集选择策略,所以此单一来源的旧知识无法满足长期的增量学习过程。而旧知识不仅存在于旧样例集中,也以更加结构化、抽象化的形式存在于过去阶段的旧模型<sup>[10]</sup>之中。所以,提炼旧模型中的结构化知识进行多源知识补充是增加旧类信息的可行方案。

针对上述问题,本文提出了双重特征增强的三维点云类增量学习方法(dual feature enhancement for class-incremental 3D point cloud object learning, DFE3D)。

DFE3D设计了对于不同局部结构进行区别关注的差异性局部增强模块(discriminative local enhancement, DLE),通过增强差异性局部特征的感知能力,缓解新旧类特征区分度不足的问题。并且,在回放的基础上设计了知识注入网络(knowledge injection network, KIN),将旧模型中的旧知识以高维特征的形式直接融合到新阶段模型中,以缓解旧类信息不足造成的新类偏好加剧。具体而言,DLE分别编码点云数据各局部的特征信息与空间信息,通过对两个层次的信息进行融合获得点云数据中丰富的局部语义特征。同时,通过各局部在全局中的高维特征信息自适应地引入了各局部结构的重要性权重,强化对差异性结构的感知并减小相似性结构引入的冗余信息干扰,扩大不同类别目标间的特征区分度。随后,KIN将新增模型分支作为适应新批次数据的可塑性分支,同时将旧模型作为稳定性

分支,自适应地生成旧知识注入可塑性分支中,形成增强后的混合知识特征来缓解知识来源单一的特征造成的旧类信息短缺问题,进一步缓解新类的偏好现象。本文设计了多个三维点云类增量学习的实验方案,并且与通用类增量学习方法和三维点云类增量学习方法进行对比,验证了DFE3D的有效性。消融实验也验证了DFE3D中两个模块的有效性。本文的主要贡献可以概括为以下3点。

- 针对数据访问受限且分批次到达的现实场景下的三维点云目标分类任务,提出了双重特征增强的三维点云类增量学习方法,通过增加差异性局部结构的辨别度和利用多源旧类知识来缓解点云类增量学习过程中产生的灾难性遗忘。

- 设计了差异性局部增强模块去感知差异性局部结构信息并且减少冗余的相似性局部结构信息干扰,缓解增量学习过程中的点云新旧类特征辨识度不足的问题。

- 设计了知识注入网络,在新模型分支作为可塑性分支的基础上,引入了旧模型构成的稳定性分支进行旧知识的结构化注入,通过多源知识补充缓解了旧类信息不足而加剧的新类偏好问题。

## 1 相关工作

### 1.1 点云目标分类

点云目标分类任务的一个挑战在于点云数据固有的置换不变性和不规则性。对此,点云目标分类方法大体可以分为以下3类:第1类基于投影<sup>[1,16]</sup>,这类方法将三维点云投影成多视角的二维图像,然后使用二维卷积网络去分析不同视角间的特征关联性。第2类方法基于体素<sup>[2,17]</sup>,这类方法体素化三维点云,然后使用三维卷积处理规则的体素表示。一些方法<sup>[18,19]</sup>使用稀疏卷积来缓解体素化后由于点云稀疏性引起的数据利用低效问题。然而,基于投影或体素的方法引入了中间表征,在点云数据处理的过程中不可避免会遭受信息损失。因此,第3类基于点的方法在近些年发展迅速,这类方法直接对原始点云数据进行处理。PointNet<sup>[3]</sup>首先提出使用共享的多层感知机(multi-layer perceptron, MLP)单独对每个点的特征处理,随后聚合形成全局特征。由于点云数据的局部结构语义丰富,PointNet++<sup>[4]</sup>在PointNet的基础上设计了分层结构,并对点云局部结构进行逐层编码。之后工作为了进一步探索点云局部结构中的语义特征,设计了基于图<sup>[20,21]</sup>和基于卷积<sup>[5,22]</sup>的方法提

升点云局部特征提取的能力. 近年来, 一些受 Transformer 启发的方法<sup>[23,24]</sup>将注意力机制使用到局部或全局中点关系的建模中, 成为基于点的方法发展的趋势. 尽管这些方法在获得所有数据时表现优秀, 但是在多任务或多阶段的现实场景任务中性能会快速下降, 产生灾难性遗忘.

## 1.2 类增量学习

类增量学习严格定义新批次的数据类别与之前所有批次数据类别不重叠, 在增量学习 3 大场景<sup>[7]</sup>中最具有挑战性, 近年来受到广泛关注. 类增量学习方法可以分为无记忆和基于回放的方法.

无记忆的方法不需要在增量学习过程中维护一个额外的旧样例集. 一些方法<sup>[9,25]</sup>引入额外的正则项对模型参数和超参数的更新过程进行限制, 从而防止旧知识的快速遗忘. 另外一些方法<sup>[26,27]</sup>通过生成模型或预训练模型取代内存缓冲区, 将生成信息或提示信息用于新阶段的学习. 这些无记忆的方法受到的限制小, 但是由于多阶段增量时正则约束力逐渐减弱、生成模型训练困难和提示设计困难等原因, 表现逊色于基于回放的方法.

基于回放的方法保留当前阶段之前所有批次数据的子集作为旧样例集, 在后续的增量过程重用旧样例集来帮助网络回忆旧类知识, 从而缓解灾难性遗忘. iCaRL<sup>[8]</sup>首先提出使用适量的内存保留旧样例集, 随后使用知识蒸馏<sup>[28]</sup>来利用过去模型约束新类的特征学习, 并且使用最近类均值分类器替代全连接层分类器和利用羊群算法更新旧样例集, 随后的方法在此工作流程上进行了改进. 文献<sup>[29]</sup>提出了新旧类数据不平衡会引起新类偏好问题, 并重新训练一个分类器权重修正器来缓解. 之后, 一些工作<sup>[30-33]</sup>针对缓解新类偏好问题进行了研究. 另外, 一些工作<sup>[34-36]</sup>针对旧样例集的更新进行了研究. GSS<sup>[35]</sup>以参数梯度为特征, 通过最小化固定角约束提高旧样例集多样性. GCR<sup>[36]</sup>根据模型参数关于过去数据上的梯度维护一个核心集. 一些工作<sup>[10,37,38]</sup>通过采用不同的知识蒸馏变体, 利用旧模型参数中的旧知识约束新模型优化. GeoDL<sup>[37]</sup>利用不同特征构造低维流形, 并最小化测地路径上流形间的差异. CSCCT<sup>[38]</sup>提出跨空间聚类和控制迁移分别扩大类间差异和约束类间的特征传输. 这些方法尝试设计通用的类增量学习方法, 可以简单并有效地迁移到不同模态下的数据上. 但是, 由于点云目标通常存在噪声和不规则性, 这

些方法无法感知有利于防止灾难性遗忘的几何特征. 同时, 基于动态架构<sup>[39]</sup>和 Transformer 的方法<sup>[40]</sup>发展迅速, 但是由于点云领域的网络设计与二维领域存在差异, 该类方法难以迁移到点云类增量学习领域.

因此, 点云类增量学习近年来逐渐受到关注. I3DOL<sup>[14]</sup>方法使用局部质心重构模块和注意力机制动态地感知局部特征, 并通过数据统计信息来缓解偏好问题. 但是该方法依赖质心的重构质量, 并且统计信息生成的全局协调项会受到达类别顺序影响, 在增量过程中表现不稳定. RCR<sup>[15]</sup>方法保留降采样后的点云作为旧样例集, 并且使用重建损失来作为正则项, 防止潜在特征对新数据的过度拟合. 但是, RCR 方法的性能受限于重建分支的表现, 并且需要保留所有样本的降采样数据, 在数据严格保密的场景下无法应用. 区别于先前的方法, 本文提出的 DFE3D 方法不需要额外保存统计信息和所有的点云数据, 仅通过双重的特征增强策略来最大限度地平衡类增量学习过程中的可塑性-稳定性.

## 2 方法

### 2.1 问题定义和方法概要

本文研究点云类增量学习问题. 假设点云数据分批到达, 即需要按次序处理一系列独立的点云任务  $\mathcal{T}_t$ ,  $t \in \{1, \dots, T\}$ . 阶段  $t$  的任务  $\mathcal{T}_t$  包含一组训练数据  $\mathcal{D}_t = \{(x_i^t, y_i^t)\}_i^{N_t}$ , 其中  $x_i^t \in \mathbb{R}^{n \times 3}$  是输入的原始点云坐标数据,  $n$  是每个点云数据包含的点数,  $y_i^t \in \mathcal{Y}_t$  是该点云数据对应的类别标签,  $N_t$  为新批次数据量. 不同任务的数据集样本类别不重复, 即  $\mathcal{Y}_t \cap \mathcal{Y}_i = \emptyset$ ,  $i \in \{1, \dots, t-1\}$ . 同时, 基于回放的方法保留少量过去阶段的数据, 所以在阶段  $t$ , 还维护着整个增量阶段内存大小固定的空间用来保存之前增量任务  $\{\mathcal{T}_1, \dots, \mathcal{T}_{t-1}\}$  中的少量训练数据, 作为旧样例集  $\mathcal{M}_t$ . 旧样例集中的数据将与当前任务的新数据混合为  $\mathcal{D}_t \cup \mathcal{M}_t$  用于  $t$  阶段的训练过程. 最终目标是训练一个点云模型, 其预测结果在整个标签空间  $\{\mathcal{Y}_i\}_{i=1}^t$  上都表现良好.

本文提出了双重特征增强的三维点云类增量学习方法 (DFE3D), 其结构如图 3. 首先使用点云编码器处理原始点云输入, 得到编码后的特征信息和空间信息. 随后, 差异性局部增强模块 (第 2.2 节) 的局部邻域特征提取操作表征出语义丰富的局部特征, 全局感

知的重要性评估操作对差异性结构和相似性结构进行区别感知, 强化对差异性结构的关注. 编码器和差异性局部增强模块共同构成该网络的特征提取部分  $\Phi_t$ . 最后, 知识注入网络 (第 2.3 节) 从旧模型特征层

$\Phi_{t-1}$  构成的稳定性分支中提取出旧模型知识注入由当前阶段新增模型  $\Phi_t$  构成的可塑性分支中, 形成混合知识特征  $h_{mix}$ , 送入分类器  $C$  中进行预测, 以缓解单一知识特征引起的灾难性遗忘.

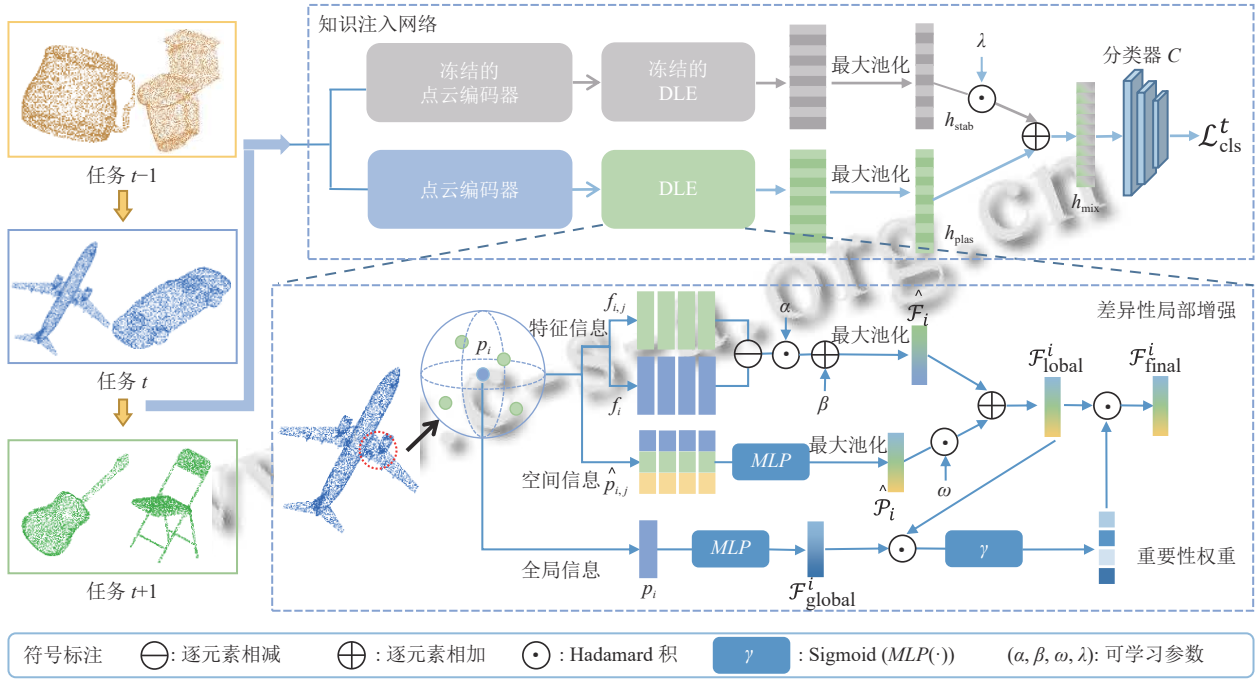


图3 DFE3D方法总览图

### 2.2 差异性局部增强

常规点云目标分类方法<sup>[4,5]</sup>通过复杂的点对设计提取局部语义, 但忽视了各局部结构的差异化感知, 受新旧类数据不均衡导致的新类偏好问题影响较大, 在类增量学习场景中会迅速丧失对旧类的特征感知能力. 对此, 为了区别感知差异性局部特征和冗余的相似性局部特征, 以使点云局部特征学习适应点云类增量学习场景, 本文提出了差异性局部增强模块 (DLE), 其主要思路为: 局部邻域特征提取操作分别处理局部的特征信息和空间信息, 融合以获得丰富的局部特征. 全局感知的重要性评估操作根据每个局部在全局中的特征自适应地生成重要性权重, 根据每个局部对于点云目标的重要性以差异化地关注不同局部特征.

首先局部邻域特征提取操作分别从每个局部结构的局部空间信息与特征信息中表征出每个局部的几何关系与特征关系, 将其自适应地融合, 从而明确点云数据中每个局部结构的特性. 假设编码器的输出  $\mathcal{P} = \{p_i \in \mathbb{R}^3\}_{i=1}^n$  提供坐标信息, 以及相应的特征信息  $\mathcal{F} = \{f_i \in \mathbb{R}^c\}_{i=1}^n$ . 根据三维空间内的坐标信息, 使用 K 近邻 (K-nearest

neighbor, KNN) 算法为  $\mathcal{P}$  中每个点  $p_i$  找到  $K$  个最近邻点组成的局部邻域  $\mathcal{N}_{p_i} = \{p_{i,j} \in \mathbb{R}^3\}_{j=1}^k$ .

对于局部空间信息, 为了获得更丰富的空间信息特征, 根据给定的锚点坐标  $p_i$  及其邻近点坐标  $p_{i,j}$ , 将原始空间信息整合为空间信息集  $\hat{p}_{i,j}$ . 随后, 将空间信息集经过 MLP 处理获得深层空间信息, 并将输出维度对齐到特征信息的相同维度以便两者的信息融合. 该操作过程如下:

$$\hat{\mathcal{P}}_i = \text{MaxPooling}(\text{MLP}(\hat{p}_{i,1}, \hat{p}_{i,2}, \dots, \hat{p}_{i,k})) \quad (1)$$

其中,  $\hat{\mathcal{P}}_i \in \mathbb{R}^c$  为局部邻域  $\mathcal{N}_{p_i}$  的空间局部特征. 其中空间信息集为:

$$\hat{p}_{i,j} = \text{concat}(\|p_i - p_{i,j}\|, p_i, p_{i,j}, p_i - p_{i,j}) \quad (2)$$

其中,  $\hat{p}_{i,j} \in \mathbb{R}^{10}$  将锚点坐标  $p_i$ , 近邻点坐标  $p_{i,j}$ , 两者间的坐标残差  $p_i - p_{i,j}$  以及两者间的 L2 距离通过级联操作进行整合.

对于局部特征信息, 通过利用锚点  $p_i$  的特征  $f_i$  与其邻近点特征  $\{f_{i,j}\}_{j=1}^k$  的边向量  $\hat{f}_{i,j} = f_{i,j} - f_i, j = 1, 2, \dots, k$ ,

使用一组跨任务的可学习参数  $\alpha \in \mathbb{R}^c$  和  $\beta \in \mathbb{R}^c$  对特征进行一个简单映射, 并且通过最大池化层得到编码后局部特征信息. 公式如下:

$$\widehat{\mathcal{F}}_i = \text{MaxPooling}(\alpha \odot (\hat{f}_{i,1}, \hat{f}_{i,2}, \dots, \hat{f}_{i,k}) + \beta) \quad (3)$$

其中,  $\widehat{\mathcal{F}}_i \in \mathbb{R}^c$  为局部邻域  $N_{p_i}$  跨任务映射后的局部特征信息,  $\odot$  为 Hadamard 积,  $c$  为特征通道数.

为了减少信息间的差异, 使空间信息与特征信息可以更好地融合, 引入一个可学习的参数在通道维度放缩空间信息, 操作如下:

$$\mathcal{F}_{\text{local}}^i = \omega \odot \widehat{\mathcal{P}}_i + \widehat{\mathcal{F}}_i \quad (4)$$

其中,  $\mathcal{F}_{\text{local}}^i \in \mathbb{R}^c$  是以锚点  $p_i$  为中心的局部特征,  $\omega \in \mathbb{R}^c$  为调整空间信息的可学习参数.

尽管局部邻域特征提取操作提取出输入点云中丰富的局部特征, 但是该过程是逐点的特征增强, 没有下采样阶段, 存在冗余的局部结构信息干扰类别辨识. 为了突出点云差异性局部结构, 使用全局感知的重要性评估部分通过每个局部的全局信息评估每个局部对于该目标类别分辨的重要性, 并通过生成的重要性权重自适应地给予不同局部结构差异性的关注.

由于  $p_i$  为每个局部在全局坐标系中的位置信息, 可以反映目标的全局形状信息<sup>[21]</sup>. 所以, 通过编码锚点的坐标信息以获得全局信息, 操作如下:

$$\mathcal{F}_{\text{global}}^i = \text{MLP}(p_i) \quad (5)$$

其中,  $\mathcal{F}_{\text{global}}^i \in \mathbb{R}^c$  为锚点  $p_i$  的全局特征编码.

随后, 将全局特征以逐通道的注意力形式融合到之前提取出的局部特征中, 形成全局感知后每个局部的特征, 融合后的特征送入一个映射函数得到逐通道的重要性权重, 根据映射后的重要性权重, 调整对各局部特征的关注意度. 具体操作如下:

$$\mathcal{F}_{\text{final}}^i = \gamma(\mathcal{F}_{\text{local}}^i \odot \mathcal{F}_{\text{global}}^i) \odot \mathcal{F}_{\text{local}}^i \quad (6)$$

其中,  $\mathcal{F}_{\text{final}}^i$  为差异性局部增强后的  $p_i$  点的特征,  $\gamma$  为一个由 *MLP* 和 *Sigmoid* 构成的映射函数, 将特征映射为重要性权重. *DLE* 模块的结构如图 3 所示.

### 2.3 知识注入网络

*DLE* 加强了对点云实例中差异性局部结构的感知, 一定程度上缓解了点云特征中的新类拟合现象, 但仍存在旧类信息不足的问题. 单分支网络在类增量学习过程中依赖旧样例集进行旧类特征学习. 但是由于

缓存大小有限, 各旧类保存的样本数量随着增量过程不断被压缩, 对新类特征的拟合逐渐加剧, 所以单一来源的旧信息很难指导多阶段的增量过程.

为了解决这个问题, 知识注入网络 (*KIN*) 引入了过去阶段旧模型来指导当前阶段的特征学习. *KIN* 将原单分支网络扩展为双分支网络, 由可塑性分支和稳定性分支共同构成. 其中, 可塑性分支即为新增模型分支, 负责充分学习当前训练数据的特征信息, 最大限度地拟合当前可用数据. 稳定性分支由过去阶段学习到的旧模型构成, 自适应地提供旧模型中结构化的旧知识. 最后, 通过将稳定性分支的旧知识注入当前的可塑性分支中, 形成增强后的双重知识特征送入分类器进行学习和预测.

由于过去阶段的知识已经作为高度抽象的高维信息存在于过去阶段的旧模型中, 随意地改动都会让结构化信息受到破坏. 在阶段  $t$  的训练之前, 将  $t-1$  阶段训练得到的模型特征提取器  $\Phi_{t-1}$  保留并冻结为稳定性特征回放器  $\Phi'_{t-1}$ , 以最小程度地减少旧信息损失. 对于每个输入  $x \in \mathcal{D}_t \cup \mathcal{M}_t$ ,  $\Phi'_{t-1}$  对  $x$  前向计算, 将送入分类器前特征  $h_{\text{stab}}$  作为稳定性旧知识. 同时, 当前阶段模型的可塑性分支  $\Phi_t$  由  $\Phi_{t-1}$  的参数初始化得到, 对输入  $x$  也进行前向计算, 同样将  $h_{\text{plas}}$  作为可塑性特征, 公式如下:

$$h_{\text{stab}} = \Phi'_{t-1}(x); h_{\text{plas}} = \Phi_t(x) \quad (7)$$

其中,  $h_{\text{stab}} \in \mathbb{R}^d$  为稳定性分支输出特征,  $h_{\text{plas}} \in \mathbb{R}^d$  为可塑性分支输出特征.

为了在注入旧知识后获得更加紧凑的特征表示, 并最小化对稳定性旧知识的影响, *KIN* 引入了一个稳定性特征修正系数, 动态地调整稳定性分支. 具体公式如下:

$$h_{\text{mix}} = \lambda \odot h_{\text{stab}} + h_{\text{plas}} \quad (8)$$

其中,  $\lambda \in \mathbb{R}^d$  是一个可学习的参数,  $\odot$  为 Hadamard 积,  $h_{\text{mix}} \in \mathbb{R}^d$  是知识注入后的混合知识特征. *KIN* 的结构如图 3 所示.

### 2.4 损失函数与算法流程

针对输入的点云  $x$ , 经过 *DLE* 和 *KIN* 双重增强之后, 获得展平后的特征  $h_{\text{mix}}$ , 经过分类器 *C* 得到类别预测结果  $c$ . 最终, 本方法训练时优化目标如下:

$$\mathcal{L}_{\text{cls}}^t(x, y) = - \sum_{k=1}^{K_t} \delta_{k=y} \log(c_k(x)) \quad (9)$$

其中,  $K_t$  是阶段  $t$  为止所有的已知类别,  $\delta_{y=k}$  是一个指示函数,  $c_k$  是第  $k$  类的预测概率,  $y$  为输入点云  $x$  对应的类别标签. 同时, 由于 DFE3D 方法在各增量阶段的优化流程一致, 本文展示了阶段  $t$  优化流程, 具体见算法 1.

算法 1. DFE3D 算法的训练流程 (第  $t$  阶段)

输入: 训练数据  $\mathcal{D}_t$ , 旧样例集  $M_t$ , 现有模型  $(\Phi_t, C)$ , 过去模型特征提取器  $\Phi_{t-1}$ , 可学习参数  $\lambda$ .

// 编码器和差异性局部增强模块共同构成  $\Phi$ ;

输出:  $\theta_t \in \{\Phi_t, C, \lambda\}$ .

1) 初始化:  $\Phi_t \leftarrow \Phi_{t-1}$ , 随机初始化  $C$ ;

2) 参数冻结:  $\Phi_{t-1} \leftarrow \Phi_{t-1}$

3) for epochs do

for  $(x, y)$  in  $\mathcal{D}_t \cup M_t$  do

$h_{\text{mix}} = \lambda \odot \Phi_{t-1}(x) + \Phi_t(x)$

$c = C(h_{\text{mix}})$

计算分类损失  $\mathcal{L}_{\text{cls}}^t$

End for

End for

4) 使用随机采样策略更新获得旧样例集  $M_{t+1}$

### 3 实验结果与分析

为了验证双重特征增强的三维点云类增量学习方法 (DFE3D) 的有效性, 本文分别在两个合成点云分类数据集 ModelNet40<sup>[41]</sup>、Shape-Net<sup>[42]</sup>, 两个真实室内场景下的点云分类数据集 ScanObjectNN<sup>[43]</sup>、ScanNet<sup>[44]</sup> 上与通用类增量学习方法和三维点云类增量学习方法进行对比. 此外, 本文的消融实验评估了差异性局部增强模块 (DLE) 和知识注入网络 (KIN) 的有效性.

#### 3.1 实验设置

##### 3.1.1 对比方法

由于点云类增量学习相关工作较少, 并且一些在二维图像上进行实验的类增量学习方法并不依赖于图像特性, 旨在设计通用的类增量学习方法, 可以简单有效的迁移到各种模态下. 所以, 本文将 DFE3D 方法与通用类增量学习方法和三维点云类增量学习方法进行对比. 本文选择了 9 个先进的基于回放的通用类增量学习方法, 包括 iCaRL<sup>[8]</sup>、BiC<sup>[29]</sup>、WA<sup>[30]</sup>、GDumb<sup>[34]</sup>、LUCIR<sup>[31]</sup>、SS-IL<sup>[32]</sup>、GeoDL<sup>[37]</sup>、CSCCT<sup>[38]</sup> 和 CafeBoost<sup>[33]</sup>. 同时, 点云类增量学习方法中, 考虑到数据回放策略的公平性, 本文选择 I3DOL<sup>[14]</sup> 进行对比. 对于每个方法的旧样例集采样方法, iCaRL、LUCIR、GeoDL、CSCCT 和 CafeBoost 方法选择羊群采样策略, 而其他方法采用

随机采样策略.

##### 3.1.2 数据集

本文实验选用两个合成点云数据集和两个室内场景下的点云数据集. 其中, ModelNet40 是通过三维 CAD 模型的表面进行随机点采样得到合成数据集, 包含了 40 个常见类别, 有 9843 个训练样本和 2468 个测试样本. ShapeNet 也是采样于三维 CAD 模型的合成数据集, 包含了 55 个常见类别, 有 35708 个训练样本、5158 个测试样本和 10261 个验证样本. 现实场景下的数据集 ScanObjectNN 包含了 15 个常见类, 有 11416 个训练样本, 2882 个测试样本, 本文实验使用的变体是 PB\_T50\_RS. ScanNet 数据集包含了 17 个类, 共 12060 个训练样本, 3416 个测试样本.

##### 3.1.3 实验场景与评价指标

参考类增量学习中最常用实验场景 CIFAR100-B0<sup>[8]</sup>, 本文将在以下 4 个实验场景下进行测试: 1) ModelNet40-B0: 将 40 个类别平均划分为 10 个增量阶段, 每个阶段新增 4 个新类别的数据, 固定保留样本上限为 800 个样本. 2) ScanObjectNN-B0: 将 15 个类平均划分为 5 个增量阶段, 每个阶段新增 3 个新类别的数据, 固定保留样本上限为 600. 3) ScanNet-B0: 将 17 类划分为 6 个增量阶段, 每个阶段新增 3 个新类别的数据, 最后一个增量阶段新增 2 类, 固定保留样本上限为 600. 4) ShapeNet-B0: 将 55 个类划分为 9 个增量步骤, 每个阶段新增 6 个新类别的数据, 最后 1 个增量阶段新增 7 类, 固定保留样本上限为 1100.

实验中评价指标为 Top-1 平均增量准确率<sup>[7]</sup>, 其定义为所有增量阶段的 Top-1 增量准确率的平均值, 以反映方法的综合增量性能表现. 其中, 每个阶段的 Top-1 增量准确率是在包含当前阶段所有已知类别的点云测试数据集上测得的 Top-1 分类准确率, 反映不同阶段模型的增量性能, 具体公式如下:

$$Acc_{\text{avg}} = \frac{\sum_i^T Acc_i}{T} \quad (10)$$

其中,  $Acc_{\text{avg}}$  是 Top-1 平均增量准确率,  $Acc_i$  为第  $i$  个增量阶段的 Top-1 增量准确率,  $T$  为总增量阶段数. 本文对每个方法进行了 3 个随机 seed 的实验, 选用每个增量阶段准确率以及最终平均增量准确率的均值作为最终结果, 并且在对比实验和消融实验中额外报告方差

结果,以公平地比较不同方法的整体增量性能。

### 3.1.4 实验细节

本文的实验部署平台为 PyTorch, 运算平台为搭载 NVIDIA GeForce RTX 3090 GPU 的服务器。本文所有方法的编码器使用 PointNet<sup>[3]</sup>网络。训练过程优化器选择 Adam, 初始学习率为 0.01, batchsize 为 32, 每个增量阶段训练 120 轮, 并在 58 和 96 轮时, 将学习率降为之前的 0.1 倍。

### 3.2 定量对比实验结果

表 1-表 4 分别报告了 ModelNet40-B0、ScanObjectNN-B0、ScanNet-B0、ShapeNet-B0 上的结果。实验显示各阶段已知类别数, 并列出各阶段的增量准确率, 加粗代表最优效果,  $\Delta$  代表 DFE3D 方法相较于其他方法的提升。从实验结果可见, 在 4 个实验场景下,

DFE3D 的平均增量准确率分别优于通用类增量学习方法 2.76%、2.18%、1.65%、2.05%, 表明了 DFE3D 方法在点云类增量学习中相较于通用的类增量学习方法更具有优势。同时, DFE3D 方法在 4 个实验场景下分别优于点云类增量学习方法 I3DOL 2.03%、2.26%、1.95%、1.28%, 表明了 DFE3D 方法对现有有点云类增量学习方法的不足之处有所改进。另外, 在 ModelNet40-B0、ScanObjectNN-B0、ScanNet-B0 实验场景下, 由各阶段增量准确率可见, 本文方法在增量初期表现一般的情况下, 在后续的增量过程中性能基本都优于对比方法。并且, 在最后一个增量阶段结束时, 在完整测试集上的性能同样优于所有对比的方法。多方面实验验证了 DFE3D 方法在点云类增量学习过程中更加有效地缓解了灾难性遗忘。

表 1 ModelNet40-B0 实验场景上的定量对比实验和消融实验结果 (%)

方法	每个增量阶段已知类别数										Acc <sub>avg</sub>	$\Delta$
	4	8	12	16	20	24	28	32	36	40		
iCaRL <sup>[8]</sup>	98.77	94.70	89.24	82.32	76.95	71.93	69.35	65.14	62.51	57.62	76.85 ± 0.31	(+5.93)
BiC <sup>[29]</sup>	99.26	94.38	89.16	81.75	77.63	74.05	70.48	68.54	67.53	64.49	78.73 ± 0.55	(+4.05)
WA <sup>[30]</sup>	99.01	94.57	88.62	79.56	77.39	73.81	70.20	67.80	66.49	63.55	78.10 ± 1.23	(+4.68)
LUCIR <sup>[31]</sup>	99.01	94.18	89.33	81.51	79.03	74.20	71.24	66.71	66.57	62.07	78.39 ± 1.49	(+4.39)
GDumb <sup>[34]</sup>	99.14	95.36	89.20	81.01	77.94	74.07	71.37	67.48	66.13	62.59	78.43 ± 0.75	(+4.35)
SS-IL <sup>[32]</sup>	99.01	93.53	87.64	81.24	75.16	73.08	68.67	66.89	66.19	62.11	77.35 ± 0.38	(+5.43)
GeoDL <sup>[37]</sup>	99.01	93.14	88.89	81.68	79.96	75.13	72.56	68.95	67.59	63.08	79.00 ± 0.29	(+3.78)
I3DOL <sup>[14]</sup>	99.26	95.49	90.71	83.03	81.72	77.34	74.05	71.07	68.88	65.95	80.75 ± 0.16	(+2.03)
CSCCT <sup>[38]</sup>	99.35	<b>95.73</b>	90.78	83.24	79.11	75.48	70.60	68.63	66.67	62.07	79.17 ± 0.76	(+3.61)
CafeBoost <sup>[33]</sup>	<b>99.37</b>	95.12	90.86	83.02	79.32	77.08	73.98	70.16	67.90	63.43	80.02 ± 0.43	(+2.76)
DFE3D w/o DLE	99.01	94.64	88.53	79.56	76.73	73.79	71.19	67.58	65.84	61.52	77.84 ± 0.82	(+4.94)
DFE3D w/o KIN	98.26	93.84	88.96	83.41	79.87	75.90	71.69	70.73	69.06	65.48	79.72 ± 0.59	(+3.06)
DFE3D	99.17	94.70	<b>91.85</b>	<b>86.00</b>	<b>83.22</b>	<b>80.94</b>	<b>76.84</b>	<b>74.32</b>	<b>72.63</b>	<b>68.14</b>	<b>82.78 ± 0.37</b>	—

表 2 ScanObjectNN-B0 实验场景上的定量对比实验和消融实验结果 (%)

方法	每个增量阶段已知类别数					Acc <sub>avg</sub>	$\Delta$
	3	6	9	12	15		
iCaRL <sup>[8]</sup>	88.93	79.89	68.58	59.88	51.63	69.78 ± 0.58	(+2.18)
BiC <sup>[29]</sup>	89.02	74.36	68.40	60.26	53.25	69.06 ± 0.44	(+2.90)
WA <sup>[30]</sup>	89.44	73.60	68.34	58.15	51.75	68.25 ± 0.32	(+3.71)
LUCIR <sup>[31]</sup>	89.65	75.32	67.31	57.77	47.92	67.60 ± 1.13	(+4.36)
GDumb <sup>[34]</sup>	89.65	72.80	66.35	57.67	48.79	67.05 ± 0.98	(+4.91)
SS-IL <sup>[32]</sup>	89.57	72.59	69.38	60.34	50.87	68.55 ± 0.75	(+3.41)
GeoDL <sup>[37]</sup>	89.82	75.58	68.33	58.96	52.46	69.03 ± 1.48	(+2.93)
I3DOL <sup>[14]</sup>	<b>90.29</b>	77.74	70.03	60.06	50.37	69.70 ± 0.56	(+2.26)
CSCCT <sup>[38]</sup>	89.98	77.81	67.89	57.62	48.69	68.40 ± 1.03	(+3.56)
CafeBoost <sup>[33]</sup>	89.26	77.68	68.60	58.79	51.77	69.22 ± 0.89	(+2.74)
DFE3D w/o DLE	89.40	72.90	67.06	59.07	50.74	67.83 ± 0.30	(+4.13)
DFE3D w/o KIN	90.05	79.47	65.10	56.45	50.20	68.25 ± 0.90	(+3.71)
DFE3D	89.51	<b>81.78</b>	<b>71.22</b>	<b>61.45</b>	<b>55.82</b>	<b>71.96 ± 0.39</b>	—



### 3.3 消融实验结果

为了验证 DFE3D 方法中 DLE 模块和 KIN 的有效性, 本文在 ModelNet40-B0、ScanObjectNN-B0、ScanNet-B0、ShapeNet-B0 上进行了消融实验, 如表 1-表 4 所示, DFE3D w/o DLE 和 DFE3D w/o KIN 分别代表 DFE3D 方法在没有 DLE 模块和 KIN 时的性能表现. 与 DFE3D 方法相比, DFE3D w/o DLE 缺少了增量学习过程中针对点云数据本身特性的设计, 未差异化地对待点云局部特征, 导致其在 4 个实验场景下的性能分别下降 4.94%、4.13%、2.48%、

3.50%, 反映了 DLE 模块在类增量学习中对于 DFE3D 方法特征感知部分的重要性. 即便 DFE3D w/o DLE 中存在旧模型中旧知识的特征注入, 但性能下降仍然严重, 反映出知识注入网络最终的作用受稳定性分支所提供旧知识质量的影响. DFE3D w/o KIN 是仅利用 DLE 模块的单分支网络用于点云类增量学习过程, 由于缺少额外的模型旧知识的注入, 最终在各个实验场景下的性能下降 3.06%、3.71%、1.84%、2.38%, 反映了 KIN 的设计引入的多源知识对于点云类增量学习的有效性.

表 3 ScanNet-B0 实验场景上的定量对比实验和消融实验结果 (%)

方法	每个增量阶段已知类别数						Acc <sub>avg</sub>	Δ
	3	6	9	12	15	17		
iCaRL <sup>[8]</sup>	94.36	83.65	75.23	70.47	68.04	64.33	76.01 ± 0.32	(+5.87)
BiC <sup>[29]</sup>	94.47	83.23	77.55	74.41	71.62	70.16	78.57 ± 0.84	(+3.31)
WA <sup>[30]</sup>	94.63	83.47	77.86	74.70	71.51	69.84	78.67 ± 0.30	(+3.21)
LUCIR <sup>[31]</sup>	94.97	85.44	78.63	75.93	71.12	66.31	78.73 ± 0.90	(+3.15)
GDumb <sup>[34]</sup>	94.55	84.80	78.40	74.15	69.16	67.80	78.14 ± 1.69	(+3.74)
SS-IL <sup>[32]</sup>	94.60	84.56	78.64	75.11	70.96	65.49	78.22 ± 1.70	(+3.66)
GeoDL <sup>[37]</sup>	95.06	86.15	80.39	76.28	71.73	70.07	79.94 ± 0.19	(+1.94)
I3DOL <sup>[14]</sup>	95.30	84.58	80.12	77.05	73.26	69.27	79.93 ± 1.34	(+1.95)
CSCCT <sup>[38]</sup>	95.58	85.80	79.24	77.04	72.44	65.52	79.27 ± 0.31	(+2.61)
CafeBoost <sup>[33]</sup>	95.30	86.54	80.41	76.80	72.48	69.85	80.23 ± 1.31	(+1.65)
DFE3D w/o DLE	94.27	85.46	81.05	74.99	72.11	68.53	79.40 ± 0.30	(+2.48)
DFE3D w/o KIN	96.24	86.95	78.75	76.42	74.77	67.13	80.04 ± 0.86	(+1.84)
DFE3D	<b>96.30</b>	<b>89.16</b>	<b>82.67</b>	<b>77.30</b>	<b>75.48</b>	<b>70.34</b>	<b>81.88 ± 0.82</b>	—

表 4 ShapeNet-B0 实验场景上的定量对比实验和消融实验结果 (%)

方法	每个增量阶段已知类别数									Acc <sub>avg</sub>	Δ
	6	12	18	24	30	36	42	48	55		
iCaRL <sup>[8]</sup>	92.10	86.89	77.62	72.05	68.64	63.87	60.96	56.99	52.56	70.18 ± 1.78	(+5.58)
BiC <sup>[29]</sup>	93.83	87.11	77.75	74.07	70.66	68.73	65.73	63.91	57.31	73.23 ± 0.72	(+2.53)
WA <sup>[30]</sup>	93.96	87.07	77.12	74.90	71.30	68.43	65.69	63.46	56.76	73.19 ± 0.42	(+2.57)
LUCIR <sup>[31]</sup>	92.94	87.20	79.21	75.44	71.87	66.06	61.86	58.38	48.39	71.26 ± 0.30	(+4.50)
GDumb <sup>[34]</sup>	93.51	86.01	76.83	72.09	68.18	65.67	63.09	58.51	50.48	70.49 ± 0.97	(+5.27)
SS-IL <sup>[32]</sup>	93.19	87.70	79.42	76.61	72.91	67.40	63.90	62.39	52.45	72.89 ± 0.34	(+2.87)
GeoDL <sup>[37]</sup>	92.77	87.22	79.20	77.03	74.30	71.03	65.61	63.18	53.05	73.71 ± 0.21	(+2.05)
I3DOL <sup>[14]</sup>	<b>94.71</b>	86.79	78.84	76.86	73.28	<b>71.53</b>	<b>67.98</b>	64.69	55.60	74.48 ± 0.31	(+1.28)
CSCCT <sup>[38]</sup>	93.24	88.09	79.96	75.58	71.64	67.49	63.07	58.18	51.29	72.06 ± 1.30	(+3.70)
CafeBoost <sup>[33]</sup>	93.24	88.40	80.68	76.28	72.72	69.89	65.85	61.90	53.54	73.61 ± 1.87	(+2.15)
DFE3D w/o DLE	93.65	86.21	77.67	74.17	70.35	68.38	63.93	62.34	53.64	72.26 ± 0.36	(+3.50)
DFE3D w/o KIN	93.12	87.51	79.42	76.04	72.12	69.79	65.98	63.31	53.09	73.38 ± 0.79	(+2.38)
DFE3D	94.38	<b>89.93</b>	<b>82.47</b>	<b>78.70</b>	<b>75.23</b>	70.31	67.27	<b>65.33</b>	<b>58.27</b>	<b>75.76 ± 0.37</b>	—

### 3.4 增量任务数量的定量分析

为了验证 DFE3D 方法在增量任务数量不同时的性能表现, 本文在 3 个实验场景下进行实验, 并在图 4 中展示结果. 其中, 图 4(a) ModelNet40-B0-20 steps 将 ModelNet40 数据集划分为 20 个任务, 每个任务新增

2 类数据; 图 4(b) ScanObjectNN-B0-7 steps 将 ScanObjectNN 数据集划分为 7 个任务, 每个任务新增 2 类数据, 最后任务新增 3 类数据; 图 4(c) ShapeNet-B0-18 steps 将 ShapeNet 数据集划分为 18 个任务, 每个任务新增 3 类, 最后任务新增 4 类数据.

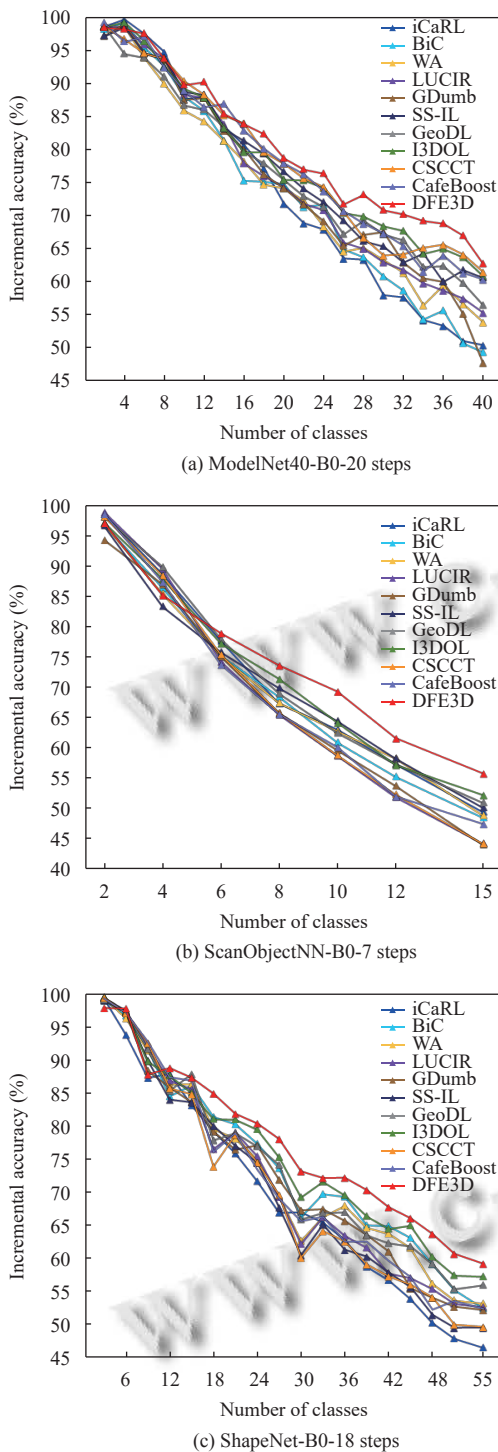


图4 不同增量任务数量时的定量分析

如图4(a)所示, DFE3D方法在前8个增量任务下性能表现一般, 但是后续的增量任务中, 各个对比方法由于难以平衡多阶段增量过程中的可塑性和稳定性, 与DFE3D方法的差异逐渐显现. DFE3D的增量准确率均优于对比方法, 性能曲线也保持平稳, 且最后阶段

的性能指标也保持最优. 验证了即使在面对更多的增量任务时, DFE3D方法依旧能很好地缓解点云类增量学习过程中的灾难性遗忘. 同时, 在另外两个实验场景图4(b)和图4(c)中也存在着和图4(a)一样的现象, DFE3D都能够完成多阶段点云类增量学习任务, 进一步验证了DFE3D方法在更多任务增量情景下的广泛适用性.

### 3.5 不同的旧样例集容量大小的定量分析

为了进一步研究旧样例集大小对DFE3D方法的影响, 本实验在ScanObjectNN-B0实验场景下, 将旧样例集容量大小从600个实例分别调整为300和900个实例, 并进行实验测试. 如图5所示, 各方法在点云类增量学习过程中对旧样例集容量都较为敏感, 在容量从900下降到300时, 各方法自身的平均增量准确率下降6.52%~9.64%, 说明回放策略对于点云类增量学习的重要性. 尽管如此, DFE3D在旧样例集容量为300和900时, 对比其他方法仍然有较好的性能表现, 分别在平均增量准确率上提升了1.28%和1.76%. 验证了DFE3D方法对于过去旧信息利用的合理性.

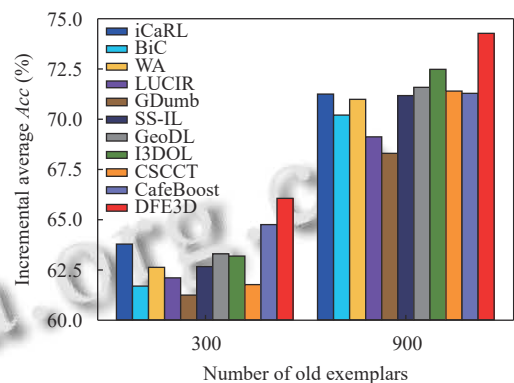


图5 不同旧样例集大小的定量分析

### 3.6 不同的增量类别顺序的定量分析

为了进一步验证DFE3D在增量过程中的鲁棒性, 本实验测试按次序到达的数据的类别变化对各方法的影响. 在ScanObjectNN-B0实验场景下, 将15个类别的到达顺序随机打乱, 并且重新划分为5个增量任务. 如表5所示, 加粗代表最优结果,  $\Delta$ 代表各方法在最优次序下与最差次序下结果的差值, 在3个不同的增量类别顺序下, DFE3D方法表现均优于所有对比方法, 最终3个次序下平均增量准确率的均值也高于最优对比方法1.46%. 值得注意的是, LUCIR方法在最优顺序下的性能指标与最差增量顺序下的性能指标相差4.54%,

其余各个方法都存在一定的差距,验证了不同的增量类别顺序对点云类增量学习存在着一定的影响,并且这个影响会随着方法的鲁棒性不同而改变。DFE3D方法即使在表现最差的次序3下仍然保持着72.38%的平均增量准确率,与最优次序下的性能表现仅相差1.84%,充分验证了DFE3D方法对于不同增量类别顺序的鲁棒性。

表5 ScanObjectNN-B0 实验场景上不同增量类别顺序下的对比实验 (%)

方法	不同的增量次序时的 $Acc_{avg}$			$Acc_{avg}$	$\Delta$
	次序1	次序2	次序3		
iCaRL <sup>[8]</sup>	71.39	71.54	68.71	70.55	2.83
BiC <sup>[29]</sup>	71.36	70.08	68.13	69.86	3.23
WA <sup>[30]</sup>	71.84	70.02	69.03	70.30	2.81
LUCIR <sup>[31]</sup>	70.68	70.02	66.14	68.95	4.54
GDumb <sup>[34]</sup>	69.42	68.21	67.12	68.25	2.30
SS-IL <sup>[32]</sup>	71.64	70.17	68.51	70.11	3.13
GeoDL <sup>[37]</sup>	72.19	70.96	70.02	71.06	2.17
I3DOL <sup>[14]</sup>	72.48	71.92	70.05	71.48	2.43
CSCCT <sup>[38]</sup>	72.54	69.83	68.49	70.29	4.05
CafeBoost <sup>[33]</sup>	73.24	71.53	70.56	71.78	2.68
DFE3D	<b>74.22</b>	<b>73.13</b>	<b>72.38</b>	<b>73.24</b>	1.84

#### 4 结论与展望

本文提出了双重特征增强的三维点云类增量学习方法,使点云目标分类技术能够适应新类数据分批到达的现实场景。DFE3D方法的差异性局部增强模块提取点云数据局部的特征信息与空间信息以获得丰富的局部结构特征,并且通过全局感知的重要性评估操作突出差异性局部区域,扩大了新旧类特征间的区分度。同时,知识注入分支将额外的旧模型知识结构化地注入模型的可塑性分支中,以混合知识特征缓解单一知识来源造成旧类信息不足问题。多个数据集上的对比实验证明了本文方法的有效性,消融实验验证了该方法两个模块的有效性。实验还证明了该方法在更多增量任务时的稳定性,对旧样例集容量和增量类别顺序的鲁棒性。然而,本文专注于类增量场景下的三维点云目标分类任务,仍然存在着任务场景上的局限。在未来的工作中,我们将进一步探索增量式的点云目标检测、语义分割等三维场景理解任务。

#### 参考文献

1 Hamdi A, Giancola S, Ghanem B. MVTN: Multi-view

transformation network for 3D shape recognition. Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021. 1–11.

2 Riegler G, Osman Ulusoy A, Geiger A. OctNet: Learning deep 3D representations at high resolutions. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 6620–6629.

3 Qi CR, Su H, Mo KC, *et al.* PointNet: Deep learning on point sets for 3D classification and segmentation. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 77–85.

4 Qi CR, Yi L, Su H, *et al.* PointNet++: Deep hierarchical feature learning on point sets in a metric space. Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 5105–5114.

5 许翔, 帅惠, 刘青山. 基于卦限卷积神经网络的3D点云分析. 自动化学报, 2021, 47(12): 2791–2800.

6 McCloskey M, Cohen NJ. Catastrophic interference in connectionist networks: The sequential learning problem. Psychology of Learning and Motivation, 1989, 24: 109–165.

7 周大蔚, 汪福运, 叶翰嘉, 等. 基于深度学习的类别增量学习算法综述. 计算机学报, 2023, 46(8): 1577–1605.

8 Rebuffi SA, Kolesnikov A, Sperl G, *et al.* iCaRL: Incremental classifier and representation learning. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 5533–5542.

9 Kirkpatrick J, Pascanu R, Rabinowitz N, *et al.* Overcoming catastrophic forgetting in neural networks. Proceedings of the National Academy of Sciences of the United States of America, 2017, 114(13): 3521–3526.

10 Douillard A, Cord M, Ollion C, *et al.* PODNet: Pooled outputs distillation for small-tasks incremental learning. Proceedings of the 16th European Conference on Computer Vision. Glasgow: Springer, 2020. 86–102.

11 张子颖, 王敏. 基于Faster R-CNN和增量学习的车辆目标检测. 计算机系统应用, 2020, 29(2): 181–186. [doi: 10.15888/j.cnki.csa.007259]

12 Liu YY, Cong Y, Sun G, *et al.* L3DOC: Lifelong 3D object classification. IEEE Transactions on Image Processing, 2021, 30: 7486–7498. [doi: 10.1109/TIP.2021.3106799]

13 Grossberg S. Consciousness CLEARs the mind. Neural Networks, 2007, 20(9): 1040–1053. [doi: 10.1016/j.neunet.2007.09.014]

14 Dong JH, Cong Y, Sun G, *et al.* I3DOL: Incremental 3D

- object learning without catastrophic forgetting. Proceedings of the 35th AAAI Conference on Artificial Intelligence. AAAI, 2021. 6066–6074.
- 15 Zamorski M, Stypułkowski M, Karanowski K, *et al.* Continual learning on 3D point clouds with random compressed rehearsal. *Computer Vision and Image Understanding*, 2023, 228: 103621. [doi: [10.1016/j.cviu.2023.103621](https://doi.org/10.1016/j.cviu.2023.103621)]
- 16 Wei X, Yu RX, Sun J. Learning view-based graph convolutional network for multi-view 3D shape analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(6): 7525–7541. [doi: [10.1109/TPAMI.2022.3221785](https://doi.org/10.1109/TPAMI.2022.3221785)]
- 17 Wang PS, Liu Y, Guo YX, *et al.* O-CNN: Octree-based convolutional neural networks for 3D shape analysis. *ACM Transactions on Graphics*, 2017, 36(4): 72.
- 18 Graham B, Engelcke M, van der Maaten L. 3D Semantic segmentation with submanifold sparse convolutional networks. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 9224–9232.
- 19 Choy C, Gwak JY, Savarese S. 4D Spatio-temporal ConvNets: Minkowski convolutional neural networks. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 3070–3079.
- 20 孟繁林, 何晓曦, 刘应浒, 等. 基于自注意力机制的点云分类分割. *计算机系统应用*, 2024, 33(1): 177–184. [doi: [10.15888/j.cnki.csa.009341](https://doi.org/10.15888/j.cnki.csa.009341)]
- 21 Wang Y, Sun YB, Liu ZW, *et al.* Dynamic graph CNN for learning on point clouds. *ACM Transactions on Graphics*, 2019, 38(5): 146.
- 22 Xu MT, Ding RY, Zhao HS, *et al.* PAConv: Position adaptive convolution with dynamic kernel assembling on point clouds. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 3172–3181.
- 23 Zhao HS, Jiang L, Jia JY, *et al.* Point Transformer. Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021. 16239–16248.
- 24 Liu YH, Tian B, Lv YS, *et al.* Point cloud classification using content-based Transformer via clustering in feature space. *IEEE/CAA Journal of Automatica Sinica*, 2024, 11(1): 231–239. [doi: [10.1109/JAS.2023.123432](https://doi.org/10.1109/JAS.2023.123432)]
- 25 Dhar P, Singh RV, Peng KC, *et al.* Learning without memorizing. Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 5133–5141.
- 26 Gao QK, Zhao C, Ghanem B, *et al.* R-DFCIL: Relation-guided representation learning for data-free class incremental learning. Proceedings of the 17th European Conference on Computer Vision. Tel Aviv: Springer, 2022. 423–439.
- 27 Smith JS, Karlinsky L, Gutta V, *et al.* CODA-Prompt: COntinual decomposed attention-based prompting for rehearsal-free continual learning. Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023. 11909–11919.
- 28 Hinton G, Vinyals O, Dean J. Distilling the knowledge in a neural network. arXiv:1503.02531, 2015.
- 29 Wu Y, Chen YP, Wang LJ, *et al.* Large scale incremental learning. Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 374–382.
- 30 Zhao BW, Xiao X, Gan GJ, *et al.* Maintaining discrimination and fairness in class incremental learning. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 13205–13214.
- 31 Hou SH, Pan XY, Loy CC, *et al.* Learning a unified classifier incrementally via rebalancing. Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 831–839.
- 32 Ahn H, Kwak J, Lim S, *et al.* SS-IL: Separated Softmax for incremental learning. Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021. 824–833.
- 33 Qiu BL, Li HL, Wen HT, *et al.* CafeBoost: Causal feature boost to eliminate task-induced bias for class incremental learning. Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023. 16016–16025.
- 34 Prabhu A, Torr PHS, Dokania PK. GDumb: A simple approach that questions our progress in continual learning. Proceedings of the 16th European Conference on Computer Vision. Glasgow: Springer, 2020. 524–540.
- 35 Aljundi R, Lin M, Goujaud B, *et al.* Gradient based sample selection for online continual learning. Proceedings of the 33rd International Conference on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2019. 1058.
- 36 Tiwari R, Killamsetty K, Iyer R, *et al.* GCR: Gradient coreset based replay buffer selection for continual learning. Proceedings of the 2022 IEEE/CVF Conference on Computer

- Vision and Pattern Recognition. New Orleans: IEEE, 2022. 99–108.
- 37 Simon C, Koniusz P, Harandi M. On learning the geodesic path for incremental learning. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 1591–1600.
- 38 Ashok A, Joseph KJ, Balasubramanian VN. Class-incremental learning with cross-space clustering and controlled transfer. Proceedings of the 17th European Conference on Computer Vision. Tel Aviv: Springer, 2022. 105–122.
- 39 Hu ZY, Li YS, Lyu JC, *et al.* Dense network expansion for class incremental learning. Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023. 11858–11867.
- 40 Gao XY, He YH, Dong SL, *et al.* DKT: Diverse knowledge transfer Transformer for class incremental learning. Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023. 24236–24245.
- 41 Wu ZR, Song SR, Khosla A, *et al.* 3D ShapeNets: A deep representation for volumetric shapes. Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015. 1912–1920.
- 42 Yi L, Kim VG, Ceylan D, *et al.* A scalable active framework for region annotation in 3D shape collections. ACM Transactions on Graphics, 2016, 35(6): 210.
- 43 Uy MA, Pham QH, Hua BS, *et al.* Revisiting point cloud classification: A new benchmark dataset and classification model on real-world data. Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019. 1588–1597.
- 44 Dai A, Chang AX, Savva M, *et al.* ScanNet: Richly-annotated 3D reconstructions of indoor scenes. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 2432–2443.

(校对责编: 张重毅)