文章编号：1001-9014(2025)06-0953-09　　DOI：10. 11972/j. issn. 1001-9014. 2025. 06. 013

标签高效的机载点云弱监督语义分割

梁转信[1]，　赖旭东[1*]，　颜倚天[2]

（1. 武汉大学 遥感信息工程学院，湖北 武汉 430079；
2. 清远市勘察测绘院有限公司，广东 清远 511500）

摘要：机载点云的语义分割为其下游应用提供数据基础。全监督的深度学习方法通常依赖大量标注数据，而部分弱监督方法由于标签选择的随机性，难以有效学习代表性特征。为应对上述挑战，本文提出了一种标签高效的语义分割方法，结合主动学习策略，在每个周期内基于信息熵理论主动选取最具价值的标签点更新训练集，实现模型的渐进式学习。在 LASDU 数据集和 H3D 数据集上的对比实验结果表明，在仅使用 0. 5% 和 0. 1% 标签进行训练的情况下，该方法仍能够取得优于现有对比方法的分割精度，体现了其在弱监督场景下的高效性和优越性。

关　键　词：摄影测量与遥感技术；弱监督；主动学习；信息熵；机载点云
中图分类号：P237　　**文献标识码**：A

Label-efficient weakly supervised semantic segmentation for airborne LiDAR point clouds

LIANG Zhuan-Xin[1]，　LAI Xu-Dong[1*]，　YAN Yi-Tian[2]

（1. School of Remote Sensing and Information Engineering，Wuhan University，Wuhan 430079，China；
2. Qingyuan Surveying and Mapping Institute Co. Ltd. ，Qingyuan 511500，China）

Abstract：Semantic segmentation of airborne point clouds provides essential data support for downstream applications. Fully supervised deep learning methods typically rely on large amounts of annotated data，while some weakly supervised approaches struggle to learn representative features effectively due to the randomness in label selection. To address these challenges，a label-efficient semantic segmentation method is proposed，which integrates an active learning strategy to progressively update the training set by actively selecting the most informative points based on information entropy in each learning cycle. Experimental results on the LASDU and H3D datasets show that，with only 0. 5% and 0. 1% labeled data，the proposed method outperforms existing approaches in segmentation accuracy，demonstrating its efficiency in weakly supervised conditions.

Key words：photogrammetry and remote sensing technology，weakly supervised，active learning，information entropy，airborne point clouds

引言

搭载于飞行平台（如飞机、无人机）上的激光雷达系统，能够高效获取大范围的地面点云数据。通过对机载点云进行语义分割，为每个点赋予语义类别，可为目标提取、变化检测、规划设计等下游应用提供关键支持。随着深度学习技术的发展，基于学习的点云处理方法逐步取代传统启发式算法，成为研究主流。

点云数据的语义分割是深度学习中的关键研究方向，并已发展为一项较为成熟的技术。根据训练过程中输入数据的表示方式，点云语义分割方法

收稿日期：2025-03-25，**修回日期**：2025-07-17　　**Received date**：2025-03-25，**Revised date**：2025-07-17
基金项目：国家自然科学基金重点项目城市立体形态结构化建模理论方法（42130105）；中国国家铁路集团有限公司科技研究开发计划项目（L2023G016）；清远市勘察测绘院有限公司项目（QYKC-2025-02-01-YFLX）
Foundation items：Supported by the National Natural Science Foundation of China （42130105）；the Science and Technology Research and Development Program of China State Railway Group Co. Ltd. （L2023G016）；the Qingyuan Surveying and Mapping Institute Co. , Ltd. Program （QYKC-2025-02-01-YFLX）
作者简介（Biography）：梁转信（1995—），河南商丘人，硕士，主要研究领域为点云智能处理及应用. E-mail：gisleung@163. com
[*] **通讯作者**（Corresponding author）：E-mail：laixudong@whu. edu. cn

主要分为三类：基于投影、基于体素和基于点的方法。早期研究发现，传统卷积算子无法直接应用于不规则的点云数据，因此研究者尝试将其转换为规则的图像或体素形式，以便使用卷积神经网络（Convolutional Neural Network，CNN）进行训练。尽管这种转换使得CNN能够应用于点云处理，但会造成部分三维空间信息的损失。此外，体素化方法在处理高分辨率三维数据时还面临计算开销大、内存消耗高等问题。PointNet[1]首次实现了对原始点云的直接操作，采用共享多层感知器（Multilayer Perceptron，MLP）提取逐点特征。当前，基于点的语义分割模型主要包括四类：MLP网络[2-4]、点卷积网络[5-6]、图卷积网络[7-8]和Transformer网络[9-10]。尽管结构不同，这些方法在总体策略上具有一致性：编码阶段提取点云的高维特征以捕捉局部与全局信息，解码阶段通过插值将特征映射回原始点，实现语义预测。然而，全监督方法高度依赖大规模标注数据，昂贵的人工标注成本显著限制了其在实际应用中的推广。

弱监督语义分割方法能够有效降低深度学习对标注样本的依赖，其训练数据通常具有不精确或不完整的特点。针对不完全标签问题，现有方法依据弱标签的生成策略，可分为三类。第一类利用有限数量的点标签进行训练。SQN[11]引入点特征查询网络（PFQN），通过稀疏点信号进行网络训练。实验表明，随机选取10%、1%、0.1%的点即可获得良好效果，说明密集标注存在冗余。HybridCR[12]在随机选择1%点的基础上，约束原始点与增强点之间预测结果的一致性与对比性，并引入对比损失以增强特征学习能力。然而，在大规模数据集中，即使仅标注0.1%的点，依然工作量巨大。此外，随机标注可能加剧类别不平衡问题，影响模型对关键特征的学习。第二类方法采用场景级标签，仅指示场景中出现的类别，而非对每个点进行标注。该方法虽可显著降低标注成本，但由于户外场景目标众多且类别分布不均，仅依赖场景级标签难以捕捉有效判别特征，易导致类别混淆。第三类方法为子云级标签，即标注场景中有限数量子云所包含的类别。MPRM[13]首次提出该策略，通过多注意力模块从不同层次特征中提取类别定位线索，生成伪标签以实现伪全监督训练。然而，连续采样固定半径的子云容易产生内容重复，导致标注冗余和信息冗余。此外，主动学习[14-15]策略也被用于降低人工标注成本。OCOC[15]通过为每个类别标注一个点实现低成本训练，但其仅关注类别中心，难以覆盖边界区域，限制了模型对代表性样本的学习能力。

本文旨在减少深度学习方法对数据标注的依赖，提出一种基于主动学习的弱监督语义分割方法，即便在标注率不足1%（LASDU数据集为0.5%，H3D数据集为0.1%）的极低监督条件下，所提方法仍表现出接近全监督方法的分割性能。具体而言，该方法首先从原始数据集中随机选取子云用于初始模型训练，随后，在迭代训练过程中，基于主动学习策略，在每个训练周期结束后评估未标记数据中的高损失区域，并将其选定为新的子云。在新子云中，利用信息熵理论识别高不确定性的点并标注。最后将上述子云和标签点更新至标注池以参与后续迭代训练。通过上述流程，模型的语义分割性能得以逐步提升，同时显著降低了对大规模标注数据的需求。

1 方法

1.1 方法概述

主动学习具有自适应性和高效性的特点，其核心思想是在有限的标注成本下，通过自主选择具有代表性或不确定性的样本训练网络，从而提高网络模型的泛化能力并实现高效标签。如图1所示，主动学习过程包括标注池X^l初始化及更新、模型训练、样本选择和迭代优化四个部分。首先，在初始化阶段，随机选择一小部分点云数据进行稀疏标注，构建初始标注池，同时将剩余数据归入未标注池。接着，在模型训练阶段，使用更新后的标注池继续训练或微调模型，以提高语义分割的准确性，并记录每次迭代周期的语义分割结果。然后，在样本选择阶段，根据不确定性采样、代表性采样或混合采样策略，从未标注池中选择最有价值的样本进行标注，以最大化模型性能提升，同时确保标注过程的高效性，减少冗余标注。最后，在迭代优化阶段，重复执行样本选择与模型更新步骤，直至满足预设的精度或迭代次数，实现渐进式提升，并持续优化标注效率。这种迭代式的学习方式不仅减少了标注成本，提高了标签的利用效率，还能在保证语义分割精度的同时，增强模型对复杂场景的适应性和鲁棒性。

1.2 初始化标注池

机载雷达采集的点云数据通常具有覆盖范围广、数据量庞大的特点，而弱监督训练则要求数据

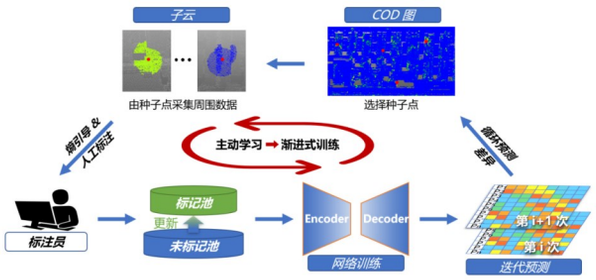

图1　基于主动学习的弱监督语义分割流程

Fig. 1　Workflow of weakly supervised semantic segmentation based on active learning

量尽可能少且具有代表性。为此,我们采用子云选择的方式对初始标注池进行有效初始化。具体而言,从大场景点云数据中随机选取种子点,并在其周围基于固定半径采集点云形成子云区域。在标注过程中,设计了一种基于 OCOC[15](One-Class One-Click)方法的标记策略:对于每个子云中的点云数据,标注员根据目视类别,在每个类别中随机选取一个点进行标注。这一策略不仅显著降低了标注量,同时确保了初始标注数据的多样性和覆盖性,为后续主动学习过程提供了高质量的初始数据支持。

1.3　训练网络模型

1.3.1　网络架构

本文选择 KPConv[16]作为骨干网络结构,其提供的三维卷积核使用空间中均匀分布的核点对核半径内的点及其邻居进行卷积。语义分割网络的输入数据来自标注池的子云、点标签和场景标签。网络层级与 KPConv 原生版本保持一致,首先进行连续的下采样和编码以提取高维度点的特征,然后逐级解码以恢复点数并学习特征,最后输出逐点类别预测。

本文提出的方法在编码阶段设计了特征融合模块,将 KPConv 卷积特征与几何特征相结合,以增强网络捕获几何结构的能力。如图 2 所示,特征融合模块由三个处理步骤组成。首先,在查询质心的 k 邻域后,使用 KPConv 核来计算局部卷积特征 f_{kpc}。局部几何特征 f_{geo} 是通过计算点云协方差的特征值得出,随后进行维度扩张,使其特征通道维度与 f_{kpc} 对齐。接下来,这两种类型的特征被拼接起来,并通过多层感知器(MLP)进一步处理,以生成最终的融合特征。最后,应用残差连接将融合特征添加到初始点特征中作为最终输出。该模块的输出可表示为:

$$F_{\mathrm{out}} = \mathrm{MLP}\left(\mathrm{Concat}\left[f_{\mathrm{kpc}}, \mathrm{Proj}\left(f_{\mathrm{geo}}, C_{\mathrm{kpc}}\right)\right]\right) + F_{\mathrm{init}} \quad (1)$$

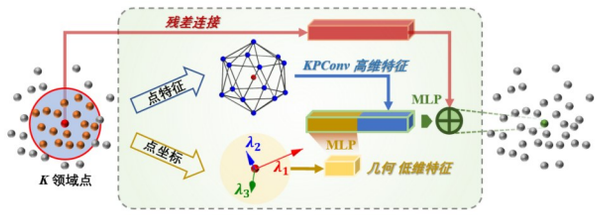

图2　特征融合模块示意图

Fig. 2　Schematic diagram of the feature fusion module

在特征融合模块中,基于由邻域点坐标构建的协方差矩阵的三个特征值计算几何特征。这些几何特征包括六个度量:线性 L_λ、平坦度 P_λ、球度 S_λ、、全方差 O_λ、各向异性 A_λ 和曲率变化 C_λ。除了上述特征外,还根据相邻点的绝对高程值计算了两个额外的属性:高程范围($R_z = z_{\max} - z_{\min}$)和高程方差 Var_z。最终的几何特征 f_{geo} 被定义为一个 8 维向量 $\left[L_\lambda, P_\lambda, S_\lambda, O_\lambda, A_\lambda, C_\lambda, R_z, Var_z\right]$。

1.3.2　损失函数

点级别约束。标注池 X^l 中点级标签提供的标记信息是最直接有效的约束。将 X^l 中标记的点集表示为 $\mathbb{P}_k = \{p_1, p_2, \cdots, p_k\}$,通过计算预测值和真实值之间的交叉熵损失为模型训练提供点级约束。考虑到类别不平衡问题,点集 $\mathbb{P}_k$ 的损失值 $\mathcal{L}_{\mathrm{point}}$ 计算如下:

$$\mathcal{L}_{\mathrm{point}} = -\frac{1}{k}\sum_{i=1}^{k}\left(w_c \cdot y_{i,c} \cdot \log\frac{\exp\left(x_{n,c}\right)}{\sum_{j=1}^{C}\exp\left(x_{n,j}\right)}\right), \quad (2)$$

式(2)中,$y_{i,c}$ 表示样本 p 的硬标签(即一种独热编码形式)的第 c 个元素的值,$\exp\left(x_{n,j}\right)$ 表示模型对样本第 j 类的预测概率。w_c 表示计算损失时每个类别的权重。如果 $\mathbb{P}_k$ 中第 j 个类别的样本数为 M_c,则 w_c 的公式为:

$$w_c = \frac{1/\sqrt{M_c}}{\sum_{j=1}^{C} 1/M_j} \quad . \quad (3)$$

子云级别约束。在标签池 X^l 中,每个子云都可以被视为一个小场景,表示为 $s_k = \{s_1, s_2, \cdots, s_k\}$。在标记阶段,子云中每个类别都被标记了一个点标签,合并这些标签可以等同于子云的场景标签,可用于在模型训练期间施加子云级约束。二进制交叉熵(BCE)用于计算场景级别的多标签损失,k 个子云的场景损失表示为:

$$\mathcal{L}_{\text{subcloud}} = -\frac{1}{k}\sum_{i=1}^{k}\sum_{j=1}^{c}\left(s_{ij}\log\hat{s}_{ij} + \left(1-s_{ij}\right)\log\left(1-\hat{s}_{ij}\right)\right) \quad (4)$$

式(4)中,s_{ij}表示第i个子云的场景标签。如果第j个类别存在于子云s_i中,则$s_{ij}=1$;否则,$s_{ij}=0$。同时,$\hat{s}_{ij}$表示模型预测的第j类在子云中存在的概率。

伪标签约束。在标注池中,大多数点属于未标记的数据,表示为$\mathbb{P}_m^{\text{u}} = \left\{p_1^{\text{u}}, p_2^{\text{u}}, \cdots, p_m^{\text{u}}\right\}$。为了缓解真实标签数量有限的问题,本文通过生成伪标签来利用这些未标记的点使模型能够进行自我监督。具体来说,使用模型的预测值作为所有未标记数据的伪标签,并计算下一次迭代中模型输出的交叉熵损失作为伪标签损失。由于较大的熵值表示伪标签的不确定性较高,因此根据其熵值计算每个伪标签的权重,以减少不确定性较高的伪标签的负面影响。伪标签的最终交叉熵损失函数表示为:

$$\mathcal{L}_{\text{pseudo}} = -\frac{1}{m}\sum_{i=1}^{m}\left(w_i^{\text{u}}\cdot\ddot{y}_{i,c}\cdot\log\frac{\exp\left(x_{n,c}\right)}{\sum_{j=1}^{C}\exp\left(x_{n,j}\right)}\right), \quad (5)$$

式(5)中,$\ddot{y}_{i,c}$表示第i个伪标签的类别,w_i^{u}表示第i个中伪标签的权重。权重w_i^{u}由以下公式计算:

$$w_i^{\text{u}} = 1 - \frac{H\left(p_i^{\text{u}}\right)}{\log C} \quad (6)$$

式(6)中,C表示类别的数量,$H\left(p_i^{\text{u}}\right)$表示在生成伪标签时基于类别的预测概率计算的熵值。

在网络模型的训练过程中,联合点级交叉熵损失$\mathcal{L}_{\text{point}}$、场景级二进制交叉熵损失$\mathcal{L}_{\text{sub-cloud}}$和伪标记损失$\mathcal{L}_{\text{proto}}$同时参与网络反向传播过程。最终网络训练的损失函数表示如下:

$$\mathcal{L} = \mathcal{L}_{\text{point}} + \mathcal{L}_{\text{subcloud}} + \mathcal{L}_{\text{proto}} + \mathcal{L}_{\text{pseudo}}. \quad (7)$$

注意,在第一次迭代中只使用了$\mathcal{L}_{\text{point}}$和$\mathcal{L}_{\text{sub-cloud}}$,从第二次迭代开始,伪标记损失$\mathcal{L}_{\text{proto}}$参与模型训练的反向传播。

1.4 主动选择样本

1.4.1 高损失区域

在更新标注池中的训练数据时,应特别关注新增样本是否能够有效提升模型性能。在深度学习任务中,模型在特定区域表现出较高的预测损失,通常意味着该类样本在当前训练集中分布不足,缺乏代表性。因此,从这些高损失区域中选取样本进行标注,能够增强模型对“薄弱”区域特征的学习能力,从而提升整体的泛化性能。准确识别并采样此类区域,并将其纳入标注池以迭代优化训练数据,是主动学习策略的核心环节之一。时间输出差异(Time Output Difference, TOD)[17]通过评估优化步骤中模型输出的差异来估计样本损失,从而有效地识别这些潜在的高损失区域。给定一个采样点t $x \in \mathbb{R}^d$,其TOD值$D_t^{\{T\}}(x)$由以下方程定义:

$$D_t^{\{T\}}(x) \triangleq \left\| f\left(x; w_{t+T}\right) - f\left(x; w_t\right)\right\|, \quad (8)$$

式(8)中,$f\left(x;\ w_t\right)$表示模型在第t次训练时对样本x的预测输出,t表示时间间隔(例如,$t>0$)。

在本文关于机载点云的研究中,使用TOD的变体—循环输出差异(Cycle Output Difference, COD),从X^u中选择子云区域进行标记。COD通过测量两次连续学习迭代之间的模型输出差异来估计样本不确定性,样本点p的COD值$D_{\text{cyclic}}^{\{T\}}\left(p\right)$表示为:

$$D_{\text{cyclic}}\left(p\right) = \left\|\hat{p}^{\{c\}} - \hat{p}^{\{c-1\}}\right\|^2 \quad (9)$$

式(9)中,$\hat{p}^{\{c\}}$表示第c次迭代训练的预测值,$\hat{p}^{\{c-1\}}$表示第$(c-1)$次迭代训练中的预测值。因此,COD值表示为两个连续迭代周期的预测输出之间的差的平方。

对于未标注池X^u中的数据,本文选择COD值最高的前b个点作为候选种子点。然而,由于点云数据通常较为密集,直接采用Top-b策略可能会导致所选种子点在空间上过于集中。这种空间聚集会引起不同子云之间较大的重叠,从而降低采样的多样性,并引入冗余数据到标注集X^l中。为缓解这一问题,本文引入局部最大值滤波策略对COD图进行重采样。具体而言,采用与子云大小一致的滑动窗口,在每个局部区域内仅保留COD值最大的点作为最终的种子点,同时将该范围内其他位置的COD值抑制(例如设为零),最后在种子点邻域内查询对应点云以构建待标注的子云样本。子云采集过程如图3所示,其中COD图中的红色点表示局部最大值,不同颜色的圆形区域表示基于固定半径范围采集的子云。

1.4.2 代表性样本

在前一步提取的子云中,识别并选择具有高不确定性的样本进行标注,通常对于提升模型性能具有关键作用。这类样本通常位于模型决策边界附近,因而对模型学习具有更高的信息增益。基于信息熵理论[18],在语义分割任务中,容易混淆的代表性样本点往往具有较高的信息熵,反映出模型对其分类的不确定性较强。为此,本文通过计算子云中每个点的预测类别概率分布的熵值$H(x)$来量化其不确定性。依据最大熵原理,从每个预测类别中选

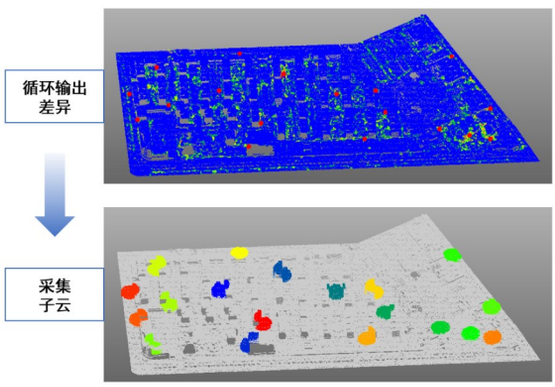

图3　基于COD的子云采集策略,不同颜色的圆形区域表示基于固定半径范围采集的子云

Fig. 3　Sub-cloud collection strategy based on COD, colored regions represent sub-clouds collected within a fixed radius

择一个具有最高熵值的点作为待标注样本,称之为“熵点”。熵值 $H(x)$ 的计算公式如下:

$$H(x)=-\sum_{i=1}^{C}P_{x=c_i}\cdot\log P_{x=c_i}\quad,\quad(10)$$

式(10)中,$P_{x=c_i}$ 表示预测该点属于类别 C_i 的概率。

在选择“熵点”时,首先利用前一轮训练得到的模型对子云中的每个点进行预测,并按照预测类别进行分组。随后,在每个类别组中选取熵值最高的点,作为对应类别下的不确定性代表点。由于熵值越高表明模型对该点的预测越不确定,这些熵点往往位于类别决策边界附近。然而,过多集中于边界的不确定性样本可能导致模型在决策边界附近过拟合,进而削弱其在其他区域的泛化能力。为此,本文引入“手动点”机制:即在子云每个语义类别中,由人工主观选择一个点进行标注。在标注实践中,这些手动点通常选自类别内部区域中视觉上更易区分的点,而非故意聚焦于边界区域。正如图 4 所示,在模拟人工标注时,通过对靠近边界的点施加选择权重惩罚机制,手动点更倾向于分布在远离决策边界的区域。

最终,结合具有代表性的熵点与手动点,语义分割网络得以在决策边界和类别中心区域同时获得有效监督,从而在提升边界识别能力的同时增强对整体类别特征的学习能力。

1.5　迭代优化

通过结合高损失区域与高不确定性样本的选择策略,实现对具有代表性的样本点的主动选取。对于无标签数据集,需对每个子云中所选样本点进行人工标注;而在已有标注的数据集上,则采用模拟人工标注的算法[15]自动为其分配标签。在此基础上,除将所选代表性样本点加入标注池外,其对应的完整子云也需一并加入。子云中其余未标注的点无需人工标注,在训练过程中为其生成伪标签和场景级标签,以开展弱监督训练。在每一轮主动学习的迭代周期中,随着标注池中训练样本的不断扩充以及代表性样本数量的持续增加,语义分割模型通过逐步迭代训练,不断优化其参数与特征表达能力,从而实现网络性能的稳步提升。

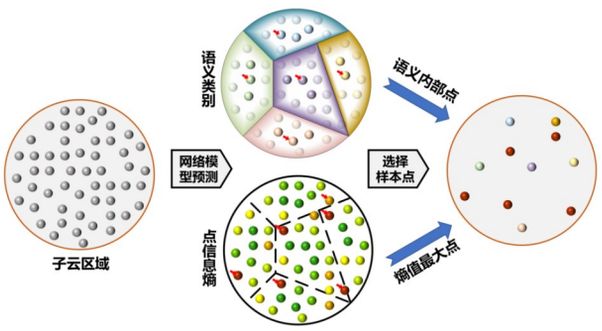

图4　子云内部样本点的主动选择过程

Fig. 4　Active selection process of sample points within sub-clouds

2　实验

2.1　数据

LASDU[19]数据集是在中国西北部黑河流域海拔约1 200 m处收集的大型ALS点云数据,平均点密度约为 3~4 pts/m^2。如图 5 所示,数据集中的点被标记为五类:地面、建筑物、树木、低植被和人工制品。注释数据集覆盖约 1 km × 1 km 的城区,拥有高密度的住宅和工业建筑。在本文的实验中,选择 Section 2 和 Section 4 作为训练集,Section 1 和 Section 3 作为测试集。

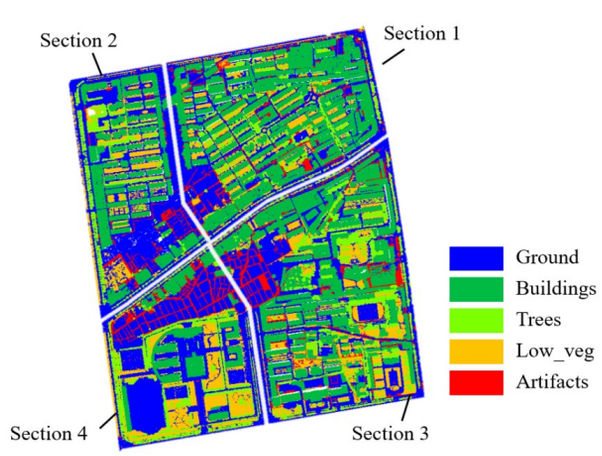

图5　LASDU点云数据集

Fig. 5　LASDU point cloud dataset

H3D[20]数据集是一个高密度的LiDAR点云数据集，其研究区域被划分为三个相互连接的子区域，分别用于训练、验证和测试。如图6所示，该数据集包含了11个语义类别，包括低植被、不透水表面、车辆、城市设施、屋顶、立面、灌木、树木、土地、垂直表面和烟囱。此外，该数据集还提供每个点的颜色信息，可作为额外的输入特征。

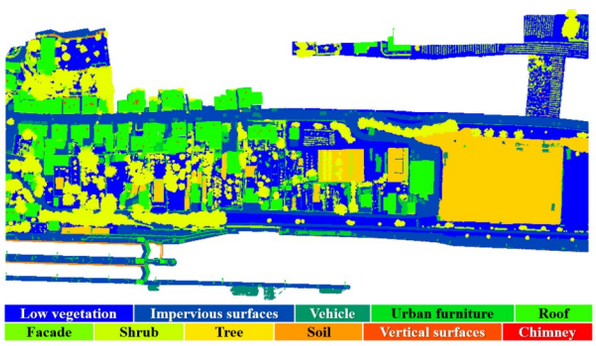

图6 H3D点云数据集

Fig. 6 H3D point cloud dataset

2.2 实验设置

在方法验证与对比实验中，针对所提出的基于主动学习的语义分割方法，共设置5轮迭代训练。每轮训练开始前，采用子云作为基本点云单元以更新标注池。在首次网络训练前，从每个场景中随机选取150个子云，构建初始标注池。每轮训练完成后，依据COD值从各场景中选取新的子云样本用于后续训练。在子云生成过程中，以采样得到的种子点为中心，提取其在XOY平面内半径10 m范围内的点集构成子云。当点云较为稀疏，导致基于半径的采样所得点数小于4 096时，则改用近邻采样以确保足够的点数。由于子云中包含的点数不固定，训练过程中采用动态批量大小策略，即将每批次的点数上限设定为102 400，而非固定子云数量，以提升训练的稳定性与效率。

为提升模型的泛化能力，训练过程中对子云数据引入了数据增强策略，具体包括围绕Z轴的随机旋转及缩放操作，以增强样本的多样性。在语义分割模型的配置方面，沿用了KPConv网络的默认训练参数和层级结构。最终，本研究在Pytorch框架下实现了所提出的弱监督语义分割方法，并在NVIDIA GeForce RTX 3060 GPU上完成了模型训练与测试。

2.3 评估指标

在点云语义分割任务中，常用评估指标包括交并比(Intersection over Union，IoU)及其均值(mean IoU，mIoU)、F1分数及其均值(mean F1，mF1)，以及总体准确率(Overall Accuracy，OA)。其中，IoU衡量预测结果与真实标签的重叠程度；F1分数为每类的精确率与召回率的调和平均，反映分类性能；OA表示整体分类的正确率，反映模型在全局数据上的表现。

本研究主要采用F1、mF1和OA作为模型性能的评估指标，其计算公式如下：

$$\begin{aligned}\mathrm{Precision}_c &= \frac{\mathrm{TP}_c}{\mathrm{TP}_c + \mathrm{FP}_c}\\ \mathrm{Recall}_c &= \frac{\mathrm{TP}_c}{\mathrm{TP}_c + \mathrm{FN}_c}\\ \mathrm{F1}_c &= \frac{\mathrm{precision}_c \times \mathrm{recall}_c}{\mathrm{precision}_c + \mathrm{recall}_c}\\ \mathrm{OA} &= \frac{\sum_{c=1}^{C} \mathrm{TP}_c}{\text{Number of all points}}\end{aligned} \quad . \tag{11}$$

2.4 结果与分析

2.4.1 LASDU数据集实验结果

从表1结果可以看出，所提出的方法在多个评估指标上均优于现有的主流弱监督方法，甚至高于经典的全监督方法PointNet++。与弱监督方法相比，在标注率仅约为0.5%的弱监督设定下，本文方法在mF1(68.0%)和OA(84.1%)两项关键指标上均取得最高值，分别优于当前最优的弱监督方法PSD(1%标注率)的67.3%和80.2%，同时在总体准确率方面也接近全监督方法KPConv(84.35%)的水平，体现了在极低标注成本下的优越性能。在Ground(88.2%)与Building(93.8%)两个类别上，本文方法分别取得了最高的F1分数，说明其在主要类别的分类效果更为稳健；在Trees(83.5%)与Low_vegetation(49.4%)类别上，表现亦处于各方法中的前列，显示出良好的场景适应性与泛化能力。尽管在Artifacts类别上未达到最高分值，但仍达到25.2%，与最高值25.4%十分接近。

相比于其他弱监督方法如SQN和OCOC，所提方法在大多数类别上取得更优结果，尤其是在Ground与Building类中表现显著提升。此外，相较于全监督方法PointNet++与KPConv，本文方法在保持极低标注比例的前提下，整体性能差距显著缩小，验证了所提出主动学习策略与弱监督框架的有效性。

在LASDU数据集上的定性实验结果如图7所示，可以观察到诸如建筑物和地面等类别已被准确

表 1　各方法在 LASDU 数据集上的定量对比实验结果

Table 1　Quantitative comparison results of different methods on the LASDU dataset

Method		F1					AvgF1 (%)	OA (%)
		Ground	Building	Trees	Low_veg.	Artifacts.		
Full Sup.	PointNet++[2]	84.02	88.97	82.68	49.32	23.59	65.72	79.70
	KPConv[16]	89.16	93.82	82.68	55.91	36.97	71.71	84.35
Weak Sup.	SQN[11] (1%)	83.4	90.5	77.9	47.6	22.5	64.4	79.0
	SQN(10%)	85.6	90.9	77	43.7	**25.4**	64.5	80.1
	PSD[21](1%)	83.6	91.7	82.4	**53.8**	25.1	67.3	80.2
	OCOC[15](1pt)	86.5	92.8	**84.6**	42	21.8	65.6	82.3
	This article(~0.5%)	**88.2**	**93.8**	83.5	49.4	25.2	**68.0**	**84.1**

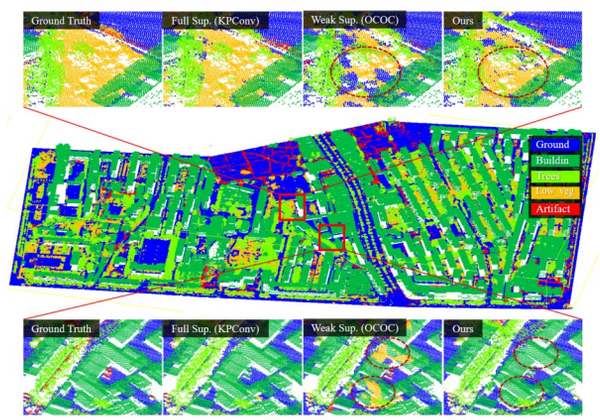

图 7　各方法在 LASDU 数据集上的定性对比实验结果

Fig. 7　Qualitative comparison results of different methods on the LASDU dataset

地分割。为了突出本方法相较于同样基于主动学习策略的 OCOC 方法的优势,本文对典型区域进行了局部放大展示。结果显示,OCOC 方法在处理高程相近的异类目标时存在混淆现象,例如低矮植被被误分类为地面,低矮建筑被误判为植被。而本方法在类似区域中有效避免了上述误判,表现出更强的类别区分能力。在仅依赖有限监督信息的情况下,其分割效果已接近全监督方法 KPConv,验证了所提方法的有效性和实用价值。

2.4.2　H3D 数据集实验结果

在弱监督条件下(仅使用约 0.1% 的标注数据),如表 2 所示,本方法在 H3D 数据集上表现出色,取得了平均 F1 分数 75.3% 和总体精度(OA) 86.8%,显著领先于其他弱监督方法。在具体类别上,本方法在多个关键类别中表现尤为突出,例如在 Chimney 类别中,F1 分数达到 87.1%,显著优于弱监督基准方法 OCOC(80.5%)。对于某些具有显著人造特征的类别(如 Impervious Surfaces、Vehicle、Urban Furniture 和 Roof),本方法在这些类别上均实现了最高精度,展现了对复杂细节特征的强大捕捉能力。此外,与全监督方法相比,本方法的平均 F1 分数(75.3%)高于 KPConv(74.9%),进一步证明了其在平衡标注成本与分类性能方面的优越性。

在 H3D 数据集上的定性实验结果如图 8 所示,主要展示了 ImpSurf、Vehicle、Urban Furniture 和 Roof 等类别的分类效果。从结果可以观察到,OCOC 在分类任务中存在一定的误分类问题,例如

表 2　各方法在 LASDU 数据集上的定量对比实验结果

Table 2　Quantitative comparison results of different methods on the LASDU dataset

Method		F1										AvgF1 (%)	OA (%)
		LowVeg	ImpSurf	Vehicle	Urb-Furn.	Roof	Facade	Shrub.	Tree	Soil	Vert-Surf	Chim-ney	
Full Sup.	PointNet++	88.3	86.9	27.9	49.6	93.6	71.1	58.3	95.0	48.9	57.7	57.3	66.8	85.5
	KPConv	88.6	87.7	80.8	63.5	94.7	78.1	61.7	95.6	33.2	71.1	68.4	74.9	87.2
Weak Sup.	SQN(1%)	80.4	79.7	34.6	54.4	92	75.2	54.2	92.6	2.4	49	64.2	61.7	80.1
	SQN(10%)	84.3	83	37	54.5	93.5	75.5	56.8	93.1	3.5	50.5	66.9	63.5	81.5
	PSD(1%)	84.5	85.7	37.8	54.4	93.9	71.9	56.7	**94.7**	**45.3**	49.5	58.9	66.7	83.7
	OCOC(1pt)	85.4	86.1	60.8	59	93.4	**80.9**	58.3	94.5	25.9	**80.6**	80.5	73.2	85.2
	This article (~0.1%)	**87.3**	**88.3**	**72.6**	**61.9**	**95.1**	80.7	**60.3**	94.5	21.1	79.1	**87.1**	**75.3**	**86.8**

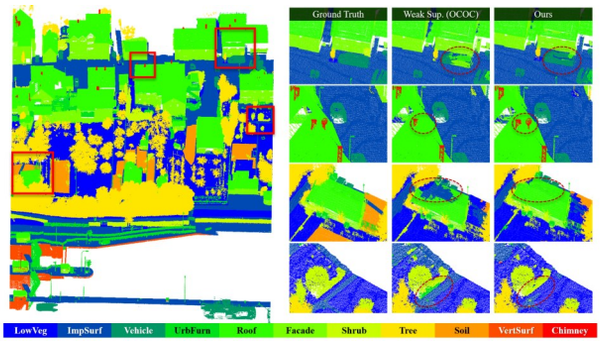

图8 各方法在H3D数据集上的定性对比实验结果

Fig. 8 Qualitative comparison results of different methods on the H3D dataset

将较高的汽车错误分类为基础设施类(Urban Furniture),以及将平坦的屋顶错误分类为不透水面(通常指道路或停车场等)。此外,对于烟囱等少样本类别,OCOC容易出现漏检或错误分类的现象。相比之下,本方法显著改善了上述问题,能够更准确地识别这些类别,展现出更高的鲁棒性和泛化能力。

3 总结

本文研究了机载点云数据的语义分割问题,特别是在弱监督条件下的挑战,并提出了一种基于主动学习策略的高效点云标注方法。通过设计子云选择策略和样本标记策略,确保有限的标注样本能够保持高代表性。实验结果表明,所提出的方法在点云分割任务中表现出显著优势,即使在极低标注率条件下,仍能获得高精度的语义分割结果,整体性能优于现有主流的弱监督方法。

尽管该方法在提高分割精度和减少标注数据需求方面取得了显著成效,但噪声数据(如人为标注错误)可能影响模型的特征学习,甚至导致过拟合。未来的研究将进一步探讨噪声对弱监督语义分割的影响,并优化鲁棒性策略,以提升模型的稳定性和泛化能力。

References

[1] Charles R Q, Su H, Kaichun M, et al. PointNet: deep learning on point sets for 3D classification and segmentation [C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 77-85.

[2] Qi C R, Yi L, Su H, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space[M]. arXiv, 2017.

[3] Qian G, Li Y, Peng H, et al. PointNeXt: revisiting PointNet++ with improved training and scaling strategies[J]. Advances in Neural Information Processing Systems, 2022, 35: 23192-23204.

[4] Ma X, Qin C, You H, et al. Rethinking network design and local geometry in point cloud: a simple residual MLP framework[M]. arXiv, 2022.

[5] Wu W, Qi Z, Li F. PointConv: deep convolutional networks on 3D point clouds[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2019). IEEE; CVF; IEEE Comp Soc, 2019: 9613-9622.

[6] Ahn P, Yang J, Yi E, et al. Projection-based point convolution for efficient point cloud segmentation[J]. IEEE Access, 2022, 10: 15348-15358.

[7] Yue C, Wang Y, Tang X, et al. DRGCNN: dynamic region graph convolutional neural network for point clouds [J]. Expert Systems with Applications, 2022, 205: 117663.

[8] Liu Y, Fan B, Xiang S, et al. Relation-shape convolutional neural network for point cloud analysis[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA: IEEE, 2019: 8887-8896.

[9] He X, Li X, Ni P, et al. Radial transformer for large-scale outdoor LiDAR point cloud semantic segmentation [J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-12.

[10] Liang Z, Lai X. Multilevel geometric feature embedding in transformer network for ALS point cloud semantic segmentation[J]. Remote Sensing, 2024, 16(18): 3386.

[11] Hu Q, Yang B, Fang G, et al. SQN: weakly-supervised semantic segmentation of large-scale 3D point clouds[C]. Computer Vision - ECCV 2022: Vol. 13687. Cham: Springer Nature Switzerland, 2022: 600-619.

[12] Li M, Xie Y, Shen Y, et al. HybridCR: weakly-supervised 3D point cloud semantic segmentation via hybrid contrastive regularization [C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022: 14910-14919.

[13] Wei J, Lin G, Yap K H, et al. Multi-path region mining for weakly supervised 3D semantic segmentation on point clouds[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE; CVF; IEEE Comp Soc, 2020: 4383-4392.

[14] Shao F, Luo Y, Liu P, et al. Active learning for point cloud semantic segmentation via spatial-structural diversity reasoning[M]. arXiv, 2022.

[15] Wang P, Yao W, Shao J. One class one click: quasi scene-level weakly supervised point cloud semantic segmentation with active learning[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2023, 204: 89-104.

[16] Thomas H, Qi C R, Deschaud J E, et al. KPConv: flexible and deformable convolution for point clouds[C]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV 2019). IEEE; IEEE Comp Soc; CVF, 2019: 6420-6429.

[17] Huang S, Wang T, Xiong H, et al. Semi-supervised active learning with temporal output discrepancy [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision. IEEE; CVF; IEEE Comp Soc, 2021: 3447-3456.

[18] Shannon C E. A mathematical theory of communication

[J]. The Bell System Technical Journal, 1948, 27(3): 379-423.

[19] Ye Z, Xu Y, Huang R, et al. LASDU: a large-scale aerial LiDAR dataset for semantic labeling in dense urban areas[J]. ISPRS International Journal of Geo-Information, 2020, 9(7): 450.

[20] Kölle M, Laupheimer D, Schmohl S, et al. The Hessigheim 3D (H3D) benchmark on semantic segmentation of high-resolution 3D point clouds and textured meshes from UAV LiDAR and Multi-View-Stereo [J]. ISPRS Open Journal of Photogrammetry and Remote Sensing, 2021, 1: 100001.

[21] Zhang Y, Qu Y, Xie Y, et al. Perturbed self-distillation: weakly supervised large-scale point cloud semantic segmentation[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 15520-15528.