



# 计算机科学

COMPUTER SCIENCE

## 基于动态平衡和距离抑制的点云语义分割主动学习

曾欣然, 李天瑞, 李崇寿

### 引用本文

曾欣然, 李天瑞, 李崇寿. 基于动态平衡和距离抑制的点云语义分割主动学习[J]. 计算机科学, 2025, 52(8): 180-187.

ZENG Xinran, LI Tianrui, LI Chongshou. [Active Learning for Point Cloud Semantic Segmentation Based on Dynamic Balance and Distance Suppression](#) [J]. Computer Science, 2025, 52(8): 180-187.

---

## 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

### 标签稀疏场景下任意数据流在线学习方法

Online Capricious Data Stream Learning with Sparse Labels

计算机科学, 2025, 52(6): 139-150. <https://doi.org/10.11896/jsjcx.240300155>

### 基于最大影响力集合的主动学习方法

Active Learning Based on Maximum Influence Set

计算机科学, 2025, 52(1): 289-297. <https://doi.org/10.11896/jsjcx.231100075>

### 基于深度对比孪生网络的事件辨重方法

Deep Contrastive Siamese Network Based Repeated Event Identification

计算机科学, 2024, 51(12): 30-36. <https://doi.org/10.11896/jsjcx.240300025>

### 基于季节分解的混合神经网络的时间序列预测

Time Series Prediction of Hybrid Neural Networks Based on Seasonal Decomposition

计算机科学, 2024, 51(11A): 231200008-7. <https://doi.org/10.11896/jsjcx.231200008>

### 基于注意力机制和双分支网络的胸部疾病分类

Classification of Thoracic Diseases Based on Attention Mechanisms and Two-branch Networks

计算机科学, 2024, 51(11A): 230900116-6. <https://doi.org/10.11896/jsjcx.230900116>

# 基于动态平衡和距离抑制的点云语义分割主动学习

曾欣然 李天瑞 李崇寿

西南交通大学计算机与人工智能学院 成都 611756

(zxr35@my.swjtu.edu.cn)

**摘要** 近年来,基于深度学习的点云语义分割取得了巨大的成功,但是它在很大程度上依赖于大量密集标注的点云数据。为了降低标注成本,许多弱监督学习方法应运而生。主动学习就是弱监督学习方法的一种,它通过选择点云的一个子集进行标注来降低标注成本。但是,目前的方法对区域信息量的估计没有充分考虑区域内所有点之间的联系,并且之前的多样性选择方法需要耗费较多的时间。为了缓解这些问题,提出一种基于动态平衡和距离抑制的点云语义分割主动学习方法。该方法通过引入区域不一致性来考虑区域中所有点之间的联系,并使用动态平衡策略来调整点级不确定性和区域不一致性的重要性以衡量区域信息量。此外,设计了特征-法线距离抑制策略来选择具有代表性的区域。该策略在考虑区域之间的空间结构时使用了一种更简单的方法,通过删除邻近的相似区域来避免冗余标注,从而提高了多样性选择的效率。在 S3DIS 和 Semantic3D 数据集上的实验结果表明,所提框架展现了最先进的性能,并且有效地减少了标注成本和多样性选择时间。

**关键词:** 点云语义分割;主动学习;动态平衡;点级不确定性;区域不一致性;距离抑制

中图分类号 TP181

## Active Learning for Point Cloud Semantic Segmentation Based on Dynamic Balance and Distance Suppression

ZENG Xinran, LI Tianrui and LI Chongshou

School of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu 611756, China

**Abstract** In recent years, deep learning-based point cloud semantic segmentation has achieved remarkable success, but it heavily relies on a large amount of densely annotated point cloud data. In order to reduce the annotation cost, many weakly supervised learning methods have emerged, and active learning is one of them. It reduces the annotation cost by selecting a subset of the point cloud for annotation, but the current methods don't fully consider the connection between all points in the region when estimating the amount of regional information, and the previous diversity selection methods take a lot of time. To alleviate these issues, this paper proposes an active learning method for point cloud semantic segmentation based on dynamic balance and distance suppression. The method considers the connection between all points in the region by introducing regional inconsistency, and uses a dynamic balance strategy to adjust the importance of point-level uncertainty and regional inconsistency to measure the amount of regional information. In addition, a feature-normal distance suppression strategy is designed to select representative regions. This strategy uses a simpler method when considering the spatial structure between regions, which avoids redundant labeling by deleting adjacent similar regions, thereby improving the efficiency of diversity selection. Experimental results on the S3DIS and Semantic3D datasets show that the proposed framework demonstrates state-of-the-art performance and effectively reduces the annotation cost and diversity selection time.

**Keywords** Point cloud semantic segmentation, Active learning, Dynamic balance, Point-level uncertainty, Regional inconsistency, Distance suppression

### 1 引言

点云语义分割作为计算机视觉的重要研究方向之一,旨在为点云中的每个点分配语义标签,其在许多领域中都有广泛的应用,例如自动驾驶<sup>[1]</sup>、机器人导航<sup>[2]</sup>和增强现实<sup>[3]</sup>等。

近年来,随着深度学习技术的快速发展,许多基于深度学习的点云语义分割方法<sup>[4-9]</sup>被提出并取得了不错的性能。但是这些方法需要密集标注的点云数据,而标注点云数据是一项非常耗时且费力的工作,主要是因为三维点云数据非常密集,需要从多个视角标注且边界模糊。因此,基于深度学习的三维

到稿日期:2024-09-18 返修日期:2024-11-18

基金项目:国家自然科学基金(62202395)

This work was supported by the National Natural Science Foundation of China(62202395).

通信作者:李崇寿(lics@swjtu.edu.cn)

点云语义分割很难广泛地运用到大规模场景中。

为了减轻点云数据逐点标记的工作量,许多基于主动学习的方法应运而生。目前,基于主动学习的三维点云语义分割可以分为基于点的和基于区域的。基于点的主动学习直接选择一批信息量丰富且有差异的点来进行标注<sup>[10-11]</sup>,但是这种方法会不可避免地引入噪声。另一种基于区域的方法则使用区域作为基本查询单元,通过考虑区域包含的信息量和区域的多样性来选择区域进行标注<sup>[12-13]</sup>。然而,目前基于区域的方法在衡量区域信息量时,要么没有考虑区域内主要类点和其他类点对信息量的不同贡献<sup>[12]</sup>,要么只考虑了点级信息量<sup>[13]</sup>。此外,在进行多样性选择时,之前的方法在考虑区域之间的空间结构时需要构建完全图,这使得多样性选择需要花费较多的时间<sup>[13]</sup>。

为了解决以上问题,本文提出了一种基于动态平衡和距离抑制的点云语义分割主动学习方法(Active Learning for Point Cloud Semantic Segmentation Based on Dynamic Balance and Distance Suppression, DBDS-AL)。DBDS-AL 首先使用全局能量模型<sup>[14]</sup>将原始点云划分为纯度较高的区域作为基本查询单元,然后通过动态平衡选择过程中的点级不确定性和区域级不确定性来计算区域信息得分。这里的两种不确定性均考虑了主要类点和其他类点的不同影响。具体来说,本文认为点级不确定性与主要类点的不确定性呈正相关,与其他类点的不确定性呈负相关;而区域级不确定性则考虑预测结果相较于主要类点的离散程度。接着,选择信息量最大的区域进行标注,但是相邻的区域往往呈现语义相似性,导致冗余标注。为了使选择的区域具有代表性,提出了一种简单而快速的特征-法线距离抑制策略来删除邻域中具有相似特征和法线的冗余区域。最后,对选择的区域进行标注。

本文的主要贡献如下:

1)提出了一种动态平衡点级不确定性和区域不一致性的区域信息得分,其中点级不确定性使用加权的点不确定性计算,区域不一致性则通过计算同一区域内点预测相较于主要类点的不一致性来估计。该策略的核心是使用动态平衡在查询轮次中充分利用点级不确定性和区域不一致性,使得区域信息量的估计更加准确。

2)提出了一种简单而快速的特征-法线距离抑制策略来删除邻域中具有相似特征和法线的冗余区域,从而有效避免冗余标注。

3)本文提出的框架在室内场景数据集 S3DIS<sup>[15]</sup>和室外场景数据集 Semantic3D<sup>[16]</sup>上展示了最先进的性能,并且有效地减少了标注成本和多样性选择时间。

## 2 相关工作

### 2.1 点云语义分割

近年来,随着深度学习技术的快速发展,基于深度学习的方法逐渐取代了传统的机器学习方法,成为点云语义分割的主流技术。基于深度学习的点云语义分割方法可以有效地提取和利用点云中的特征信息,提高点云语义分割的准确性、泛化能力和效率。其大致可以分为3类:基于图像的方法、基于体素的和基于点的。

#### 2.1.1 基于图像的方法

基于图像的方法首先将原始三维点云转化为二维图像,然后利用二维深度学习模型进行语义分割,最后将二维分割结果逆投影到三维空间以得到点云中每个点的标签。最早的基于图像的方法是由 Su 等提出的 MVCNN<sup>[17]</sup>,它使用虚拟相机将原始点云渲染成多个视角的图像,然后使用 CNN 提取每个视图的特征,并将其通过池化层聚合输入到 CNN 中进行处理,以得到分割结果。此后,为了克服在复杂场景中选择足够且适当视点的难题,Boulch 等提出了 SnapNet<sup>[18]</sup>,该方法利用点云随机生成大量局部视图,即 RGB 图像和深度合成视图,然后使用全卷积网络为两个图像的像素分配语义标签,最后将像素标记反投影聚合到原始点云。但是这种方法在下采样时,会丢失边界细节。因此,Wu 等借鉴 SqueezeNet 的思想,提出了 SqueezeSeg 网络<sup>[19]</sup>,它利用球面投影将稀疏的三维点云转换为二维图像输入到基于 SqueezeNet 的 CNN 模型中进行语义分割,并使用条件随机场对语义分割结果进行优化。然而,该方法对缺点带来的噪声很敏感,因此该团队又提出了 SqueezeSegV2<sup>[20]</sup>以提高模型的鲁棒性。随后,RangeNet++<sup>[21]</sup>,SqueezeSegV3<sup>[22]</sup>,LIF-Seg<sup>[23]</sup>等方法相继被提出。但是,在将数据转化为图像的过程中,基于图像的方法会不可避免地丢失关键信息。

#### 2.1.2 基于体素的方法

基于体素的方法将不规则的原始点云转化为规则的体素,然后使用三维卷积对其进行语义分割。最早的基于体素的方法是由 Maturana 和 Scherer 提出的 VoxNet<sup>[24]</sup>,它首次将体素和三维卷积神经网络结合应用于点云的语义分割任务。为了解决 VoxNet<sup>[24]</sup>只提供粗略标记的问题,SEG-Cloud<sup>[25]</sup>使用确定性三线性插值和全连接条件随机场来实现对点的细粒度标记。鉴于上述两种方法存在体素分辨率限制和三维卷积带来的计算量大的问题,并且点云数据都很稀疏,Graham 等提出了使用稀疏卷积的点云语义分割方法 SparseConvNet<sup>[26]</sup>。此外,Choy 等提出了一种四维时空卷积神经网络 MinkowskiNet<sup>[7]</sup>,它使用广义稀疏卷积节省内存和计算量,并使用高维条件随机场实现时空一致性。此后,许多使用稀疏卷积的方法被提出,包括 SPVConv<sup>[9]</sup>,JS3C-Net<sup>[27]</sup>,PC-SCNet<sup>[28]</sup>等。然而,这些方法仍然存在信息丢失和易受体素分辨率影响的问题,并且在处理大规模点云时面临着高内存和高计算成本的挑战。

#### 2.1.3 基于点的方法

基于点的方法直接将原始点云作为输入,并使用 MLP、三维点卷积等方法来处理点云,其可以很好地利用点云的特性并降低预处理过程中的计算复杂度。PointNet 作为此类方法的开拓者,考虑了点云的无序性、置换不变性和刚性运动不变性,最终可以学到一组稀疏的关键点(即轮廓)总结输入点云<sup>[4]</sup>。为了解决 PointNet<sup>[4]</sup>无法捕获局部结构的问题,Qi 等提出了 PointNet++<sup>[5]</sup>。该网络通过引入分层神经网络和密度自适应多尺度组合法,有效地解决了局部特征提取的问题。为了更好地捕获局部区域内点之间的关系,Zhao 等提出了一种从点云局部邻域中提取上下文特征的方法 PointWeb<sup>[29]</sup>。但是上述方法都不太适用于大规模点云数据。为

了解决这一问题, Hu 等提出了 RandLANet<sup>[8]</sup>。该方法的关键是采用随机采样以降低计算复杂度和设计局部特征聚合模块来学习复杂的局部结构。前文所提到的方法均使用 MLP 处理点云, 而另一种基于点的方法则使用点卷积来处理。例如, Li 等提出了 PointCNN<sup>[30]</sup>, 设计了一种新的卷积算子(即  $\chi$ -Conv), 其可以将无序点云重新排列成隐含空间的有序点云, 以解决点云无序问题; Wu 等提出了 PointConv<sup>[31]</sup>, 它使用密度重加权卷积来近似三维连续卷积; Thomas 等提出了 KPConv<sup>[6]</sup>, 即一种用于学习点云的局部特征的核点卷积。

## 2.2 基于主动学习的点云语义分割

虽然基于深度学习的点云语义分割方法表现出色, 但是它需要足够多的标记点云数据, 而手工标注的成本往往很高。主动学习通过选择有价值的数据进行标注来训练深度学习模型, 从而降低人工标注的成本。根据选择数据的策略, 主动学习可以分为 3 类: 基于不确定性的方法、基于多样性的方法和混合方法。基于不确定性的方法选择信息量丰富(即难以预测)的样本, 基于多样性的方法选择各不相同的样本, 混合方法则同时考虑不确定性和多样性。目前, 基于主动学习的三维点云语义分割均使用混合方法, 该类方法可以分为基于点的和基于区域的。

用于点云语义分割的基于点的主动学习直接选择一批有差异的点来获取标签。HPAL 使用分层最小边缘不确定性模块和特征距离抑制模块以选择分散的查询点, 并使用师生模型来利用未标记点所包含的信息<sup>[10]</sup>。由于 HPAL<sup>[10]</sup> 计算不确定性的方式较为复杂, 因此 Lei 等提出了一种使用集成预测计算不确定性的方法<sup>[11]</sup>。然而, 这种单独查询点的方法难免会给训练带来额外噪声。用于点云语义分割的基于区域

的主动学习首先将点云划分为区域, 然后选择信息量丰富且多样化的区域来获取标签。ReDAL<sup>[12]</sup> 首先使用 VCCS 算法<sup>[32]</sup> 将大规模点云划分为子场景区域并将其作为基本查询单元; 然后使用 softmax 熵、颜色不连续性和结构复杂性这 3 个指标计算每个区域的区域信息得分; 最后执行多样性感知选择策略以选择区域。由于 ReDAL<sup>[12]</sup> 忽略了区域之间的空间几何信息, Shao 等提出了 SDDR-AL<sup>[13]</sup>, 它使用基于权重的区域不确定性估计和空间结构多样性推理来选择查询区域。但是, 上述方法的多样性选择策略的计算复杂度较高, 执行需要耗费较多的时间和计算资源。

## 3 基于动态平衡和距离抑制的点云语义分割主动学习

DBDS-AL 的总体框架如图 1 所示。DBDS-AL 在主动学习开始之前使用全局能量模型<sup>[14]</sup> 将原始点云划分为区域并将其作为基本查询单元, 同时选取少量区域进行标注, 构成标记数据集  $D_L$ 。接下来, DBDS-AL 分 5 个步骤执行主动学习: 1) 以监督的方式使用  $D_L$  训练点云语义分割模型; 2) 根据未标记数据集  $D_U$ , 使用语义分割模型预测得到的结果计算动态平衡的区域信息得分(Dynamically Balanced Regional Information Score, DBRIS); 3) 选择区域信息得分最高的区域作为候选标记集  $D_C$ , 并使用特征-法线距离抑制策略(Feature-Normal Distance Suppression Strategy, FNDS) 删除  $D_C$  中的相似区域, 得到最终的待标注区域集  $D_R$ ; 4) 使用噪声感知迭代标记策略(Noise-aware Iterative Labeling Strategy, NAIL) 标注  $D_R$ ; 5) 将标注的区域集添加到  $D_L$  中。重复上述操作, 直到标注预算耗尽或模型达到预设的性能。

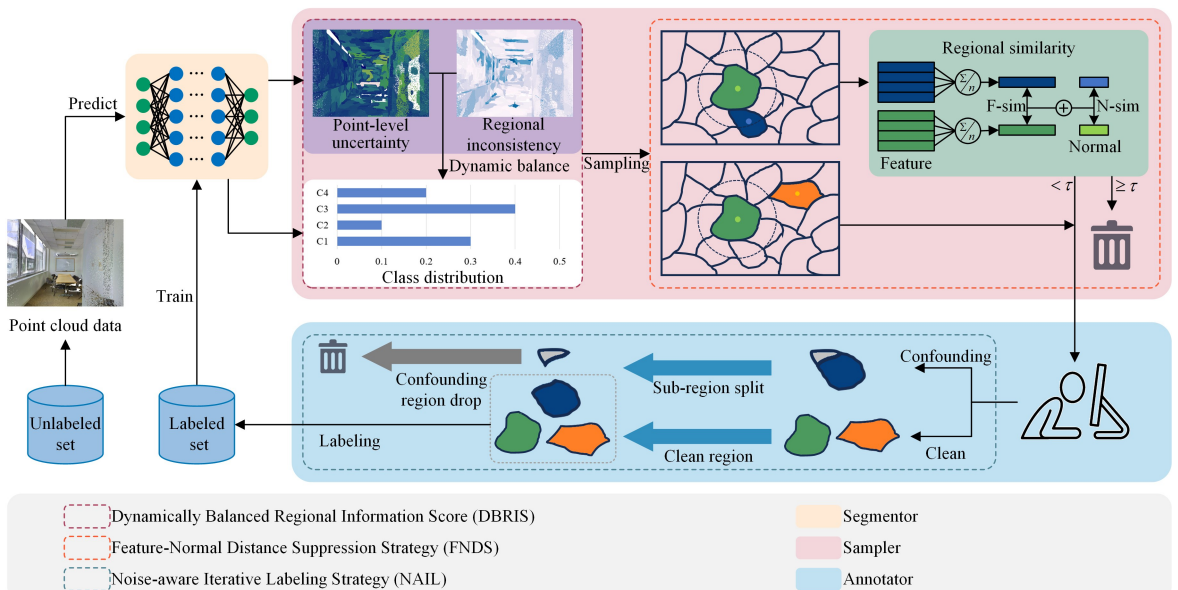


图 1 DBDS-AL 的框架图

Fig. 1 Framework diagram of DBDS-AL

### 3.1 动态平衡的区域信息得分

DBDS-AL 旨在选择最有价值的区域进行标注, 最直观的方式就是使用信息量最丰富的区域来训练模型, 但目前衡量区域信息量的方法没有充分考虑区域中所有点的

整体关联。因此, 本文提出了一种将点级信息量和区域级信息量动态结合的方法, 它既考虑了区域中主要类点和其他类点对信息量的不同影响, 又能够在训练过程中平衡两种信息量。

### 3.1.1 点级信息量

一般来说,如果模型预测的最大概率相比于其他类别的概率明显更高,则认为模型对点的预测有把握,否则模型可能难以准确预测该点的类别<sup>[39]</sup>。基于此,本文将最大概率与其他类别概率值之间的差异(BvO)作为点的不确定性度量,计算式如下:

$$u(p_i) = \frac{1 - \max(q(p_i))}{\max(q(p_i))} \quad (1)$$

其中, $q(p_i)$ 为点 $p_i$ 的预测概率。在计算得到点的不确定性之后,为了考虑区域中主要类点和其他类点对信息量的不同影响,参考SSDR-AL<sup>[13]</sup>中的点级不确定性计算方法(Wet-SU),将主要类点的权重设置为1,其他类点的权重设置为-1。因此,区域 $r$ 的点级不确定性可以通过式(2)估计。

$$PU(r) = \sum_{p_i \in M_r} 1 \times u(p_i) + \sum_{p_i \in O_r} (-1) \times u(p_i) \quad (2)$$

其中, $M_r$ 和 $O_r$ 分别表示区域 $r$ 的主要类点集和其他类点集。

### 3.1.2 区域级信息量

基于全局能量模型的假设:同一区域中的点包含相似的几何信息且属于同一类<sup>[14]</sup>,可以推测出每个区域会存在大多数点所属的主要类。因此,使用区域中所有点的预测结果和主要类点的预测结果之间的差异来反映区域整体的信息量,即使用区域不一致性来度量区域级信息量。具体计算式如下:

$$\bar{q}_{\text{dom}} = \frac{1}{|M_r|} \sum_{p_i \in M_r} q_{d(r)}(p_i) \quad (3)$$

$$\sigma_r = \sqrt{\frac{\sum_{p_i \in r} (q_{d(r)}(p_i) - \bar{q}_{\text{dom}})^2}{|r|}} \quad (4)$$

$$RIC(r) = \frac{\sigma_r}{\bar{q}_{\text{dom}}} \quad (5)$$

其中, $d(r)$ 为区域 $r$ 的主要类别, $q_{d(r)}(p_i)$ 表示点 $p_i$ 属于 $d(r)$ 的概率, $|r|$ 为区域 $r$ 包含的点的数量。式(3)用于计算区域 $r$ 的主要类点的平均概率。式(4)则用于计算区域预测结果相较于主要类点预测结果的离散程度,它类似于计算标准差,只不过本文采用的是主要类点的平均概率而不是所有点的平均概率,这使得区域不一致性的估计可以尽可能地减少其他类点带来的影响。此外,由于每个区域的 $\bar{q}_{\text{dom}}$ 有所不同,为了消除不同量纲的影响,式(5)使用 $\sigma_r$ 与 $\bar{q}_{\text{dom}}$ 的比值计算区域的相对离散程度,以得到更具可比性的指标,即区域不一致性。

### 3.1.3 动态平衡策略

为了充分利用点级不确定性和区域不一致性,受DB-DU<sup>[34]</sup>的启发,本文使用动态平衡策略调整在不同主动学习轮次中点级不确定性和区域不一致性的权重。由于点级不确定性和区域不一致性的取值范围存在差异,首先使用最大最小归一化对其进行处理,计算式如下:

$$PU_{\text{norm}}(r) = \frac{PU(r) - \min(PU > 0)}{\max(PU > 0) - \min(PU > 0)}, \text{ if } PU(r) > 0 \quad (6)$$

$$RIC_{\text{norm}}(r) = \frac{RIC(r) - \min(RIC)}{\max(RIC) - \min(RIC)} \quad (7)$$

其中, $PU$ 和 $RIC$ 分别表示所有未标记区域的点级不确定性和区域不一致性。需要注意的是点级不确定性存在负值,如果直接归一化可能会选择到混杂区域(没有有效信息的区

域),因此仅归一化 $PU$ 的正值部分。

在此基础上,使用动态平衡策略将两个指标结合起来,得到区域信息得分。动态平衡策略的动机在于随着主动学习轮次的增加,区域的预测结果会趋于一致,如果不一致,则说明该区域难以使用当前的分割网络标记准确,包含的信息量很丰富。简而言之,随着主动学习轮次的增加,区域不一致性的重要性逐步增强。因此,本文的动态平衡策略在主动学习初期将注意力放在点级不确定性上,随后逐步增加对区域不一致性的关注。动态平衡的区域信息得分的计算式如下:

$$\begin{cases} RIS(r) = \epsilon \times PU_{\text{norm}}(r) + (1 - \epsilon) \times RIC_{\text{norm}}(r) \\ \epsilon = 0.9 - (N - 1) \times 0.05 \end{cases} \quad (8)$$

其中, $\epsilon$ 为平衡系数, $N$ 是主动学习轮次。上述公式通过引入区域不一致性弥补了点级不确定性的不足。一方面,区域不一致性考虑了点预测之间的关系,可以更全面地估计区域信息量,使其更适合用于分割任务。另一方面,区域不一致性在一定程度上可以缓解点级不确定性倾向于选择大区域的情况。

最后,为了解决数据的类别不平衡问题,使用Cai等提出的类平衡获取函数(ClassBal)<sup>[35]</sup>选择区域信息量最大的未标注区域作为候选标记集。换言之,统计了区域的主要类别 $d(r)$ 的后验分布情况,并将其用于加权区域信息得分,计算式如下:

$$D_c = \arg \max_{r \in D_U} RIS(r) \times \omega(d(r)) \quad (9)$$

其中, $\omega(\cdot)$ 表示类别权重,其计算式如下:

$$\omega(c) = \exp\left(-\frac{|\{r | d(r) = c\}|}{|\{r | r \in (D_L \cup D_U)\}|}\right) \quad (10)$$

其中, $c$ 表示某一特定类别。

## 3.2 特征-法线距离抑制策略

上述方式可以选择信息量最丰富的区域进行标注,但是会不可避免地选到相似的区域,导致冗余标注。简单使用信息量的方法因为忽略了区域之间的空间结构关系,所以很可能会选到相邻的区域,而划分点云通常使用的是过分割方法<sup>[36]</sup>,相邻区域往往具有相同的标签。因此,本文提出了一种特征-法线距离抑制策略来去除相似的邻近候选区域。具体来说,遍历候选标记集 $D_c$ 中的区域。对于当前候选区域 $r_i$ ,首先计算其中心点,计算式如下:

$$\begin{cases} x_{\text{center}}^i = \frac{x_{\text{min}}^i + x_{\text{max}}^i}{2} \\ y_{\text{center}}^i = \frac{y_{\text{min}}^i + y_{\text{max}}^i}{2} \\ z_{\text{center}}^i = \frac{z_{\text{min}}^i + z_{\text{max}}^i}{2} \end{cases} \quad (11)$$

其中, $x_{\text{min}}^i, y_{\text{min}}^i, z_{\text{min}}^i$ 分别表示区域 $r_i$ 中的最小 $x, y, z$ 坐标, $x_{\text{max}}^i, y_{\text{max}}^i, z_{\text{max}}^i$ 分别表示区域 $r_i$ 中的最大 $x, y, z$ 坐标。

接下来,给定距离抑制半径 $R$ 和相似度阈值 $\tau$ ,判断区域 $r_i$ 中心点的半径 $R$ 内是否存在已被选择区域的中心点,若不存在,则直接将其添加到待标注区域集 $D_R$ 中,即:

$$D_R = D_R \cup \{r_i\}, \text{ if } \forall r_j \in D_L \cup D_R, d_{ij} \geq R \quad (12)$$

其中, $d_{ij}$ 为区域 $r_i$ 和 $r_j$ 中心点之间的欧氏距离, $D_R$ 在每轮主动学习开始时初始化为 $\emptyset$ 。此外,为了尽可能地减少 $D_R$ 和标记区域集之间以及 $D_R$ 自身的冗余,将已选择区域定义

为  $D_L$  和  $D_R$  的并集。相反,若存在邻近区域  $r_j$ ,则计算区域  $r_i$  与邻近区域之间的相似度,计算式如下:

$$\begin{cases} F\text{-sim}(r_i, r_j) = \frac{\mathbf{f}_i \cdot \mathbf{f}_j}{\|\mathbf{f}_i\| \|\mathbf{f}_j\|} \\ N\text{-sim}(r_i, r_j) = \frac{\mathbf{n}_i \cdot \mathbf{n}_j}{\|\mathbf{n}_i\| \|\mathbf{n}_j\|} \\ \text{Sim}(r_i, r_j) = \alpha \times F\text{-sim}(r_i, r_j) + \beta \times N\text{-sim}(r_i, r_j) \end{cases} \quad (13)$$

其中,  $\mathbf{f}_i$  和  $\mathbf{f}_j$  是从经过训练的分割网络中提取的区域  $r_i$  和区域  $r_j$  主要类点的特征的平均值,  $\mathbf{n}_i$  和  $\mathbf{n}_j$  是区域  $r_i$  和区域  $r_j$  中离中心点最近的点的法线,  $F\text{-sim}$  和  $N\text{-sim}$  分别为特征相似度和法线相似度。  $F\text{-sim}$  和  $N\text{-sim}$  均使用余弦相似度计算。

如果对于任意邻近区域  $r_j$ , 均满足  $\text{Sim}(r_i, r_j) < \tau$ , 即当区域  $r_i$  与任意邻近区域均不相似时, 将区域  $r_i$  添加到  $D_R$  中; 否则, 将区域  $r_i$  视为相对于已选择区域的冗余区域, 抛弃该区域。

### 3.3 噪声感知迭代标记策略

本文使用 Shao 等提出的噪声感知迭代标记策略<sup>[13]</sup>对待标注区域  $D_R$  进行标注。简而言之, 标注纯度高的区域, 如果纯度不够高, 则根据分割网络的预测结果把该区域划分为子区域, 然后依次判断子区域的纯度, 纯度高就标注, 否则就抛弃该子区域。

## 4 实验

### 4.1 数据集

为了验证 DBDS-AL 的性能, 在两个大规模点云数据集上进行了实验, 分别为室内场景数据集 S3DIS<sup>[15]</sup> 和室外场景数据集 Semantic3D<sup>[16]</sup>。

S3DIS 数据集采集于斯坦福大学的多个室内场所, 覆盖面积超过 6000 m<sup>2</sup>, 包含 6 个大型室内区域, 共 271 个房间, 每个房间为包含中等大小 3D 点的点云。它共有 11 种场景的房间, 包含 13 个类别, 超过 2.15 亿个点。本实验在 Area 1, Area 2, Area 3, Area 4, Area 6 上进行主动学习, 并在 Area 5 验证集上评估 DBDS-AL 框架的性能。

Semantic3D 数据集由静态地面激光扫描仪采集, 包含各种不同的自然场景和人造场景, 例如草地、灌木丛、教堂和车站等。它由 15 个用于训练的点云和 15 个用于在线测试的点云组成, 包含 8 个类别, 超过 40 亿个点。由于在线测试点云并未公开真实标签, 本文将 15 个用于训练的点云拆分为训练集和测试集(包括 bildstein\_station3 和 sg27\_station2), 并在

测试集上评估 DBDS-AL 框架的性能。

### 4.2 实验设置

本文使用全局能量模型<sup>[14]</sup>将原始点云划分为区域, 并使用 RandLANet<sup>[8]</sup>作为主动学习的点云语义分割模型。

在开始主动学习之前, 对 S3DIS 数据集和 Semantic3D 数据集进行网格下采样, 网格大小分别为 0.04 m 和 0.06 m。由于 Semantic3D 数据集的点密度很高, 因此对其进行了两次下采样。随后, 对数据进行划分, 在 S3DIS 数据集上生成了 468492 个区域, 并随机选取 0.5% 的区域进行标注用于初始化分割模型; 在 Semantic3D 数据集上生成了 434400 个区域, 并随机选取 0.8% 的区域进行标注用于初始化分割模型。

在每个主动学习轮次中, RandLANet 使用 Adam 优化器训练 30 个轮次。初始学习率设置为 0.01, 在 S3DIS 数据集上每个训练轮次后降低 16%, 在 Semantic3D 数据集上每个训练轮次后降低 8%。

实验共进行 5 轮主动学习, 超参数  $\alpha$  和  $\beta$  分别设置为 0.9 和 0.1。对于 S3DIS 数据集, 每轮主动学习的标注预算为 2% 的区域数, 并将距离抑制半径  $R$  设置为 0.5 m, 相似度阈值  $\tau$  设置为 0.85; 对于 Semantic3D 数据集, 每轮主动学习的标注预算为 0.5% 的区域数, 并将距离抑制半径  $R$  设置为 3 m, 相似度阈值  $\tau$  设置为 0.895。

### 4.3 对比方法

本文将所提框架与其他 5 种主动学习策略进行比较, 包括 Random + NAIL, WetSU (Entropy) + NAIL, WetSU (LC) + NAIL, WetSU (BvSB) + NAIL 和 SDR-AL。它们均使用了类平衡获取函数 (ClassBal)<sup>[35]</sup> 和噪声感知迭代标记策略 (NAIL)<sup>[13]</sup>, 并且后面 4 种策略都使用了基于权重的区域不确定性估计 (WetSU) 来聚合点的不确定性。不同之处在于, Random + NAIL 从未标记数据集中随机选择要标记的区域, WetSU (Entropy) + NAIL, WetSU (LC) + NAIL, WetSU (BvSB) + NAIL 分别使用熵<sup>[37]</sup>、最小置信度<sup>[37]</sup>、次优概率和最优概率的比值<sup>[38]</sup>来计算点的不确定性并选择信息量最高的区域。相较于 WetSU (BvSB) + NAIL, SDR-AL<sup>[13]</sup>还使用了空间结构多样性推理来选择最具信息量和代表性的区域。

### 4.4 实验结果与分析

S3DIS 数据集和 Semantic3D 数据集的实验结果分别如图 2 和图 3 所示, 图中的每个子图均以点击量衡量的固定数额的标注预算为基准。

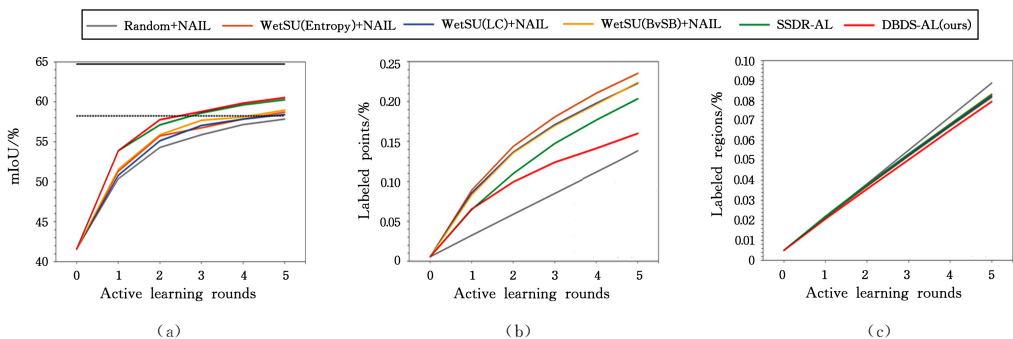


图 2 不同主动学习策略在 S3DIS 上的实验结果

Fig. 2 Experimental results of different active learning strategies on S3DIS

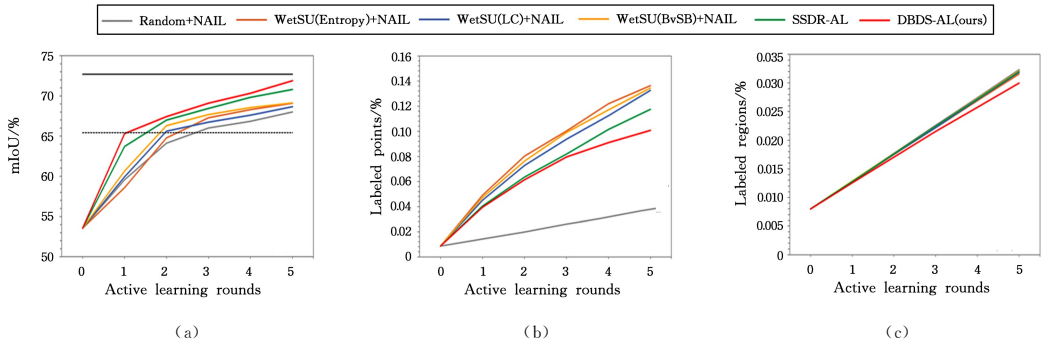


图 3 不同主动学习策略在 Semantic3D 上的实验结果

Fig. 3 Experimental results of different active learning strategies on Semantic3D

图 2(a)和图 3(a)给出了 5 轮主动学习的 mIoU,其中黑实线和虚线分别表示完全监督学习的最大性能和 90%性能。从中可以观察到,DBDS-AL 在 S3DIS 数据集上的性能优于 SSSR-AL,并且在 Semantic3D 数据集上显著优于其他主动学习策略。此外,为了验证本文方法在降低标注成本上的有效性,图 2(b)和图 3(b)给出了 5 轮主动学习标注点的百分比,图 2(c)和图 3(c)给出了 5 轮主动学习标注区域的百分比。从图 2(b)和图 3(b)中可以观察到,本文方法虽然比性能较差的 Random+NAIL 需要标注的点数多,但均比其他主动学习策略标注的点数少;而且,图 2(c)和图 3(c)表明本文的方法需要标注的区域数最少。这充分说明了本文提出的方法能够有效地降低标注成本。

表 1 列出了在 S3DIS 数据集和 Semantic3D 数据集上每轮主动学习多样性选择需要耗费的时间。鉴于上述主动学习策略只有 SSSR-AL 进行了多样性选择,本文方法仅与 SS-

DR-AL 进行比较。可以看出,本文的多样性选择方法耗费的时间是 SSSR-AL 的多样性选择方法(空间结构多样性推理)的几分之一甚至几十分之一,有效地提高了主动学习的效率。

表 1 多样性选择耗费时间的比较

Table 1 Comparison of time spent on diversity selection

Dataset	Method	Round 1	Round 2	Round 3	Round 4	Round 5
		Time	Time	Time	Time	Time
S3DIS	SSDR-AL <sup>[13]</sup>	1732.39	1468.18	1564.51	1935.54	1731.12
	DBDS-AL(ours)	55.4	49.72	49.46	48.78	49.92
Semantic-3D	SSDR-AL <sup>[13]</sup>	498.04	708.51	1218.12	1601.65	1988.31
	DBDS-AL(ours)	127.08	130.62	127.18	135.33	132.75

为了更加直观地展示分割效果,图 4 中可视化了本文方法与 Random+NAIL, WetSU(BvSB)+NAIL, SSSR-AL 的分割结果。结果表明,本文方法可以准确识别其他主动学习策略无法识别的区域,如图 4 中的红色虚线框所示。

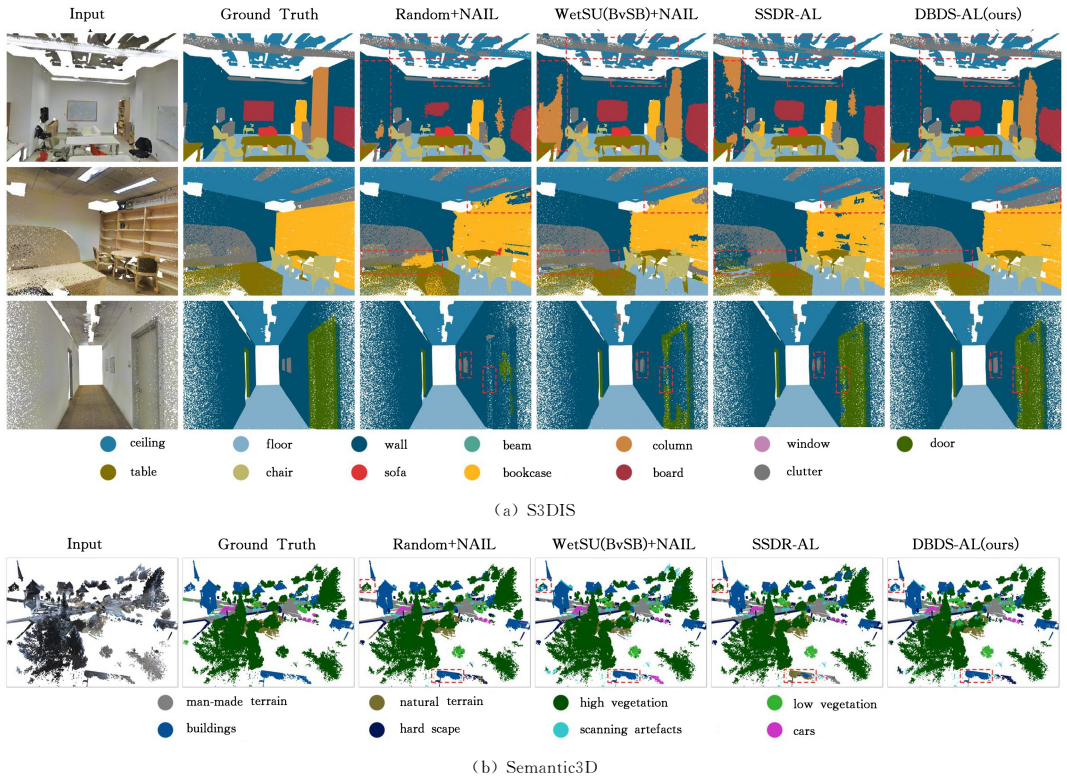


图 4 在 S3DIS 验证集和 Semantic3D 测试集上不同主动学习策略的分割结果(电子版为彩图)

Fig. 4 Segmentation results of different active learning strategies on S3DIS validation set and Semantic3D test set

## 4.5 消融实验

为了验证 BvO、动态平衡的区域信息得分和特征-法线距离抑制策略的有效性,在 S3DIS 数据集上进行了消融实验,结果如图 5 所示。

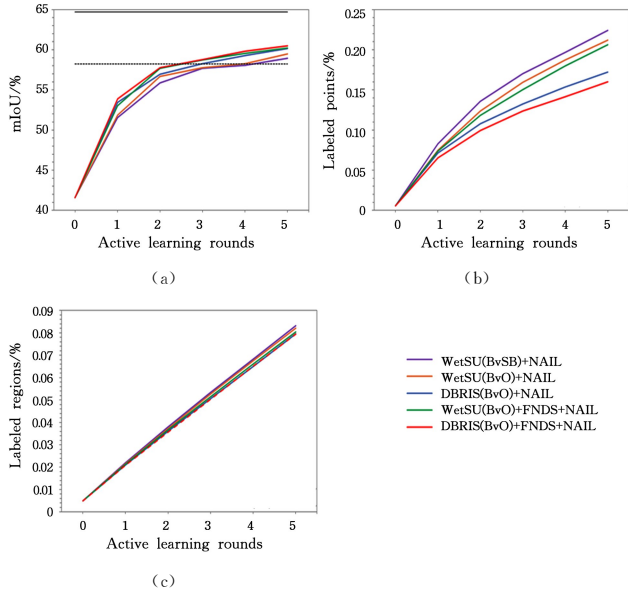


图 5 在 S3DIS Area 5 验证集上的消融实验(电子版为彩图)

Fig. 5 Ablation experiments on S3DIS Area 5 validation set

由图 5 可得出以下结论:

1) 与使用 BvSB 计算点的不确定性相比,使用 BvO 的方法使模型性能有所提升,并且对减少标注点和标注区域也有帮助,如图 5 中的紫线和橙线所示。

2) 从图 5(b) 和图 5(c) 中的橙线、蓝线和绿线的比较可知,动态平衡的区域信息得分对减少标注点和标注区域的贡献最大,这表明区域不一致性的加入有效地缓解了仅使用基于权重的区域不确定性估计趋于选择大区域的情况。此外,动态平衡的区域信息得分相较于其他不完整的信息量估计方法取得了明显的性能提升,如图 5(a) 中的紫线、橙线和蓝线所示。

3) 特征-法线距离抑制策略对主动学习性能的提升最明显,并且能够在一定程度上减少标注点和标注区域。这一结果表明,特征-法线距离抑制策略有效地减少了冗余区域,使模型能够在更具代表性的数据集上进行训练。

**结束语** 本文提出了一种基于动态平衡和距离抑制的点云语义分割主动学习,主要包含两个新设计:动态平衡的区域信息得分和特征-法线距离抑制策略。动态平衡的区域信息得分通过动态调整点级不确定性和区域不一致性的重要程度来衡量区域的信息量,特征-法线距离抑制策略通过考虑区域之间的空间结构和相似性来选择最具代表性的区域。本文方法虽然能够有效地降低标注成本,但是相较于基于点的主动学习方法仍需要标注较多的点。因此,未来的研究可以考虑将基于点和基于区域的主动学习方法相结合,以减少需要标注的点并避免引入额外的噪声。

## 参考文献

[1] PARK J, KIM C, KIM S, et al. PCSCNet: Fast 3D semantic segmentation of LiDAR point cloud for autonomous car using point

convolution and sparse convolution network[J]. *Expert Systems with Applications*, 2023, 212: 118815.

- [2] THOMAS H, AGRO B, GRIDSETH M, et al. Self-supervised learning of lidar segmentation for autonomous indoor navigation [C]// 2021 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2021: 14047-14053.
- [3] HAN L, ZHENG T, ZHU Y, et al. Live semantic 3d perception for immersive augmented reality[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2020, 26(5): 2012-2022.
- [4] 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.
- [5] QI C R, YI L, SU H, et al. Pointnet++: Deep hierarchical feature learning on point sets in a metric space[C]// Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017: 5105-5114.
- [6] THOMAS H, QI C R, DESCHAUD J E, et al. Kpconv: Flexible and deformable convolution for point clouds[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 6411-6420.
- [7] CHOY C, GWAK J Y, SAVARESE S. 4D spatio-temporal convnets: Minkowski convolutional neural networks [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 3075-3084.
- [8] HU Q, YANG B, XIE L, et al. Randa-net: Efficient semantic segmentation of large-scale point clouds[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 11108-11117.
- [9] TANG H, LIU Z, ZHAO S, et al. Searching efficient 3d architectures with sparse point-voxel convolution[C]// European Conference on Computer Vision. Cham: Springer International Publishing, 2020: 685-702.
- [10] XU Z, YUAN B, ZHAO S, et al. Hierarchical point-based active learning for semi-supervised point cloud semantic segmentation [C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023: 18098-18108.
- [11] LEI X, GUAN H, MA L, et al. DAAL-WS: A weakly-supervised method integrated with data augmentation and active learning strategies for MLS point cloud semantic segmentation[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2024, 131: 103970.
- [12] WU T H, LIU Y C, HUANG Y K, et al. Redal: Region-based and diversity-aware active learning for point cloud semantic segmentation [C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 15510-15519.
- [13] SHAO F, LUO Y, LIU P, et al. Active learning for point cloud semantic segmentation via spatial-structural diversity reasoning [C]// Proceedings of the 30th ACM International Conference on Multimedia. 2022: 2575-2585.
- [14] GUINARD S, LANDRIEU L. Weakly supervised segmentation-aided classification of urban scenes from 3D LiDAR point clouds [J]. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2017, 42: 151-157.
- [15] ARMENI I, SENER O, ZAMIR A R, et al. 3D semantic parsing

- of large-scale indoor spaces[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 1534-1543.
- [16] HACKEL T, SAVINOV N, LADICKY L, et al. Semantic3d. Net: A new large-scale point cloud classification benchmark[J]. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2017, 4: 91-98.
- [17] 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.
- [18] BOULCH A, GUERRY J, LE SAUX B, et al. SnapNet: 3D point cloud semantic labeling with 2D deep segmentation networks [J]. Computers & Graphics, 2018, 71: 189-198.
- [19] WU B, WAN A, YUE X, et al. Squeezeseg: Convolutional neural nets with recurrent crf for real-time road-object segmentation from 3d lidar point cloud[C]//2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2018: 1887-1893.
- [20] WU B, ZHOU X, ZHAO S, et al. Squeezesegv2: Improved model structure and unsupervised domain adaptation for road-object segmentation from a lidar point cloud[C]//2019 International Conference on Robotics and Automation (ICRA). IEEE, 2019: 4376-4382.
- [21] MILIOTO A, VIZZO I, BEHLEY J, et al. Rangenet++: Fast and accurate lidar semantic segmentation[C]//2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2019: 4213-4220.
- [22] XU C, WU B, WANG Z, et al. Squeezesegv3: Spatially-adaptive convolution for efficient point-cloud segmentation[C]//Computer Vision – ECCV 2020: 16th European Conference, Glasgow, UK, Part XXVIII 16. Springer International Publishing, 2020: 1-19.
- [23] ZHAO L, ZHOU H, ZHU X, et al. Lif-seg: Lidar and camera image fusion for 3d lidar semantic segmentation [J]. IEEE Transactions on Multimedia, 2023, 26: 1158-1168.
- [24] 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). IEEE, 2015: 922-928.
- [25] TCHAPMI L, CHOY C, ARMENI I, et al. Segcloud: Semantic segmentation of 3d point clouds[C]//2017 International Conference on 3D Vision (3DV). IEEE, 2017: 537-547.
- [26] GRAHAM B, ENGELCKE M, VAN DER MAATEN L. 3d semantic segmentation with submanifold sparse convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018:9224-9232.
- [27] YAN X, GAO J, LI J, et al. Sparse single sweep lidar point cloud segmentation via learning contextual shape priors from scene completion[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2021: 3101-3109.
- [28] PARK J, KIM C, KIM S, et al. PCSCNet: Fast 3D semantic segmentation of LiDAR point cloud for autonomous car using point convolution and sparse convolution network[J]. Expert Systems with Applications, 2023, 212: 118815.
- [29] ZHAO H, JIANG L, FU C W, et al. Pointweb: Enhancing local neighborhood features for point cloud processing[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 5565-5573.
- [30] LI Y, BU R, SUN M, et al. PointCNN: Convolution on x-transformed points[C]//Proceedings of the 32nd International Conference on Neural Information Processing Systems. 2018: 828-838.
- [31] WU W, QI Z, LI F. Pointconv: Deep convolutional networks on 3d point clouds[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 9621-9630.
- [32] PAPON J, ABRAMOV A, SCHOELER M, et al. Voxel cloud connectivity segmentation-supervoxels for point clouds [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2013: 2027-2034.
- [33] LI X, WANG X, CHEN X, et al. Unlabeled data selection for active learning in image classification[J]. Scientific Reports, 2024, 14(1): 424.
- [34] ZHANG S, ZHANG L, LIU Z. Active domain adaptation for semantic segmentation via dynamically balancing domainness and uncertainty [J]. Image and Vision Computing, 2024, 148: 105132.
- [35] CAI L, XU X, LIEW J H, et al. Revisiting superpixels for active learning in semantic segmentation with realistic annotation costs [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 10988-10997.
- [36] XU Y, SUN Z, HOEGNER L, et al. Instance segmentation of trees in urban areas from MLS point clouds using supervoxel contexts and graph-based optimization[C]//2018 10th IAPR Workshop on Pattern Recognition in Remote Sensing (PRRS). IEEE, 2018: 1-5.
- [37] SETTLES B, CRAVEN M. An analysis of active learning strategies for sequence labeling tasks[C]//Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing. 2008: 1070-1079.
- [38] JOSHI A J, PORIKLI F, PAPANIKOLOPOULOS N. Multi-class active learning for image classification[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2009: 2372-2379.



**ZENG Xinran**, born in 2000, postgraduate. Her main research interests include active learning and 3D point cloud semantic segmentation.



**LI Chongshou**, born in 1988, Ph.D, associate professor, is a member of CCF (No. J8308M). His main research interests include intelligent transportation, data analysis and AI.