

基于光电拓扑和全局特征的室内3D点云分割

朱晓红¹, 刘洋², 赵云¹, 纳智敏¹

(1. 云南电网有限责任公司曲靖供电局, 云南 曲靖 655000; 2. 西安交通大学 数学与统计学院, 陕西 西安 710049)

摘要: 三维点云语义分割作为机器人迈向认知智能的关键技术, 目前存在边缘特征分类精度低的问题。这类问题主要是由于语义分割网络在提取点的深层特征时忽略了邻域点间深层的特征关联, 在特征采样时忽略了全局特征的细粒度表达能力所导致。基于此, 本文提出的方法在特征提取阶段采用拓扑感知机制, 帮助网络更大程度地捕获邻域点特征之间的相似性关联; 在语义信息生成阶段引入基于 U-Net 架构的全局特征增强模块, 利用与特征提取阶段相同维度的特征张量实现对于上采样特征的增强, 帮助网络聚焦细粒度的局部信息, 从而实现更精确的语义判别能力。本文的方法在 S3DIS 数据集的 Area5 测试集上达到了 86.7% 的准确率与 60.7% 的平均交并比, 在 ShapeNet 数据集上达到了 85.1% 的平均交并比。实验结果表明, 与对比的经典方法相比, 本文提出的点云语义分割算法取得了更优的语义分割效果。

关键词: 三维点云语义分割; 局部图机制; 相似性关联; 全局特征增强; 细粒度局部信息

中图分类号: TN958.98; TP391.41 文献标识码: A DOI: 10.3969/j.issn.1001-5078.2026.02.023

Segmentation of indoor 3D point cloud based on optoelectronic topology and global features

ZHU Xiao-hong¹, LIU Yang², ZHAO Yun¹, NA Zhi-min¹

(1. Qujing Power Supply Bureau of Yunnan Power Grid Co. LTD, Qujing 655000, China;

2. College of Mathematics & Statistics, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: 3D point cloud semantic segmentation, as a critical technology for robots to attain cognitive intelligence, currently faces the challenge of low classification accuracy in edge features. This issue primarily arises from two limitations in existing semantic segmentation networks: the neglect of deep feature correlations among neighboring points during feature extraction, and the insufficient fine-grained representation capability of global features during feature sampling. To address these issues, the proposed method introduces a topology-aware mechanism during the feature extraction stage, enabling the network to capture similarity relationships between neighborhood point features to a greater extent. In the semantic information generation stage, a global feature enhancement module based on the U-net architecture is developed, which utilizes feature tensors with identical dimensionality to those from the feature extraction stage to reinforce upsampled features. This dual-strategy approach allows the network to focus on fine-grained local information while preserving global consistency, thereby achieving enhanced semantic discrimination. Experimental results demonstrate that the method achieves an accuracy of 86.7% accuracy and a mean Intersection over Union (mIoU) of

基金项目: 云南电网公司科技项目计划“配电网线路夜间巡视与隐患检测关键技术研究与应用”(No. YNKJXM20220070)资助。

作者简介: 朱晓红(1978-), 男, 硕士, 高级工程师, 从事电力系统运行与维护工作。E-mail: xiang_099@163.com

收稿日期: 2025-01-04; 修订日期: 2025-03-28

60.7 % on the Area 5 test set of the S3DIS dataset, along with an mIoU of 85.1 % on the ShapeNet dataset. compared to classical methods, the proposed point cloud semantic segmentation algorithm achieves superior semantic segmentation performance.

Keywords: 3D point cloud semantic segmentation; local graph mechanism; similarity associations; global feature enhancement; fine-grained local information

1 引言

近年来,随着 3D 传感器(激光雷达传感器、深度相机)的高速发展,三维点云数据被广泛应用于自动驾驶、大场景建模、室内导航以及机器人认知等领域^[1-3]。相比于传统 2D 图像,三维点云可以更全面地表达物体几何形状,并在复杂场景下保留丰富的场景语义信息^[4]。语义分割任务已经在 2D 图像领域取得了显著发展,研究重点逐步转向 3D 点云^[5-18]。然而,点云的无序性、不规则性及稀疏性使其高精度分割面临诸多挑战。

目前主流的点云语义分割方法可分为两大类:基于规则化表示^[5-11]和基于原始点云方法^[12-20]。规则化表示方法通常将点云数据转化为体素或图像,并通过标准 2D/3D 卷积分析提取语义信息。VoxelNet^[5]最早引入体素表示,利用 3D 卷积提取特征,但存在分辨率受限、边界伪影严重及计算成本高的问题^[12]。为解决高分辨率损失,SegCloud^[6]采用三线性插值和条件随机场提升一致性。OctNet^[7]通过八叉树存储特征,将计算集中于密集区域,有效降低内存开销。TangentConv^[8]使用切线卷积在曲面几何上进行语义分割,但容易丢失细粒度结构。

SqueezeSeg^[9]将点云投影至球面,通过 CRF 优化分割,后续 SqueezeSegV2^[10]增强了抗噪性能。

另一类方法处理原始点云,PointNet^[12]开创性地使用多层感知机提取全局特征,但缺乏局部特征提取,局部信息利用不足。PointNet++^[13]通过分层采样、邻域特征聚合改进局部感知能力,提升鲁棒性。PointCNN^[14]提出 X-transformation 对点云排序规范化,进而应用卷积操作。KPCConv^[15]设计灵活的核点卷积,增强局部特征表达能力。

总体而言,基于规则化方法虽具备一定优势,但存在细节特征与边缘信息丢失问题;而基于原始点云的方法在局部特征建模、细粒度结构提取方面仍有提升空间,如何平衡全局与局部特征、提升效率与精度,仍是研究热点。

2 本文方法

本文提出了一种端到端的网络架构解决点云语义分割问题,如图 1 所示。在特征提取和下采样阶段,采用基于拓扑感知的注意力机制,融合邻域点的局部特征,提升细节和边缘信息。上采样阶段采用 U 型网络架构^[21],利用下采样阶段的全局特征补偿上采样,进而实现更精确的语义信息判别。

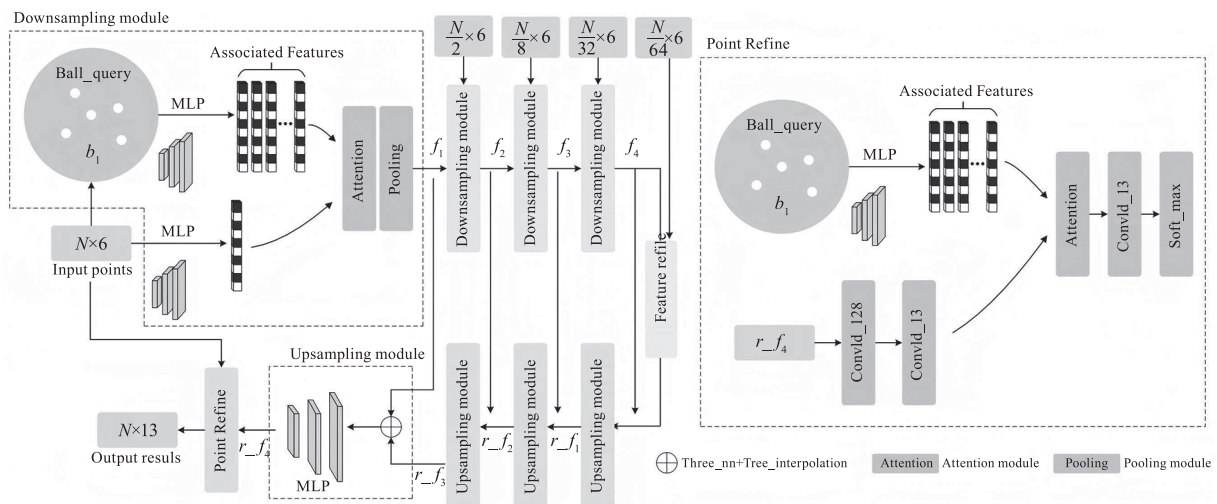


图 1 基于拓扑感知和全局特征增强的点云语义分割算法网络结构

Fig. 1 Network structure of the proposed point cloud semantic segmentation algorithm based on local graph and global feature enhancement

2.1 基于拓扑感知的下采样模块

基于拓扑感知的下采样模块包括拓扑生成、基于注意力的特征融合、最大池化三个结构。

2.1.1 拓扑生成模块

点云中的点集关系可分为两种情况,如图2和图3所示:图2中,选中区域内的点属于同一类别,空间信息和颜色信息强相关,构建拓扑图时,相同类别的点被聚类,增强中心点的语义特征;图3中,选中区域包含不同类别的点,局部拓扑结构整合不同类别的信息。

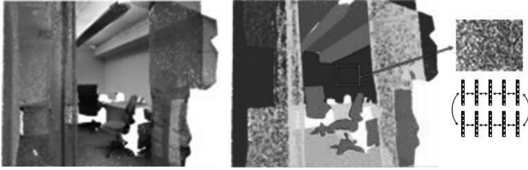


图2 S3DIS数据集中同类别点聚合示意图

Fig. 2 Aggregation of points from the same category on S3DIS

三维点云数据具有强邻域相关性,在三维空间中,一定距离下包含 n 个点的点集可以被看成一个整体 $P = \{P_0, P_1, P_2, \dots, P_{n-1}\}$, 其中 P_0 为中心点, $\{P_1, P_2, \dots, P_{n-1}\}$ 表示在指定距离 d 下, P_0 与 P_0 距离在 d 之内的点的集合。

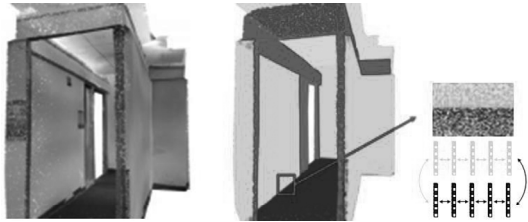


图3 S3DIS数据集中不同类别点聚合示意图

Fig. 3 Aggregation of points from different category on S3DIS

本模块采用球查询 (Ball query, BQ) 寻找近邻点, BQ 的伪代码如下:

算法3 球查询

输入: 球查询半径 $radius$, 每个邻域球中要采样的点

$nsample, shape$ 为 $(B, N, 3)$ 的输入点集合 xyz

输出: $shape$ 为 $(B, N, nsample)$ 的采样点的索引矩阵

1: 获取输入点集合的大小 B 和 N , 以及输出采样点集合的大小为 $nsample$ 。

2: for $xyz_{i,j,k} \in xyz$ do

3: 初始化 $n_{i,j,k} = 0$

4: 计算 $xyz_{i,j,k}$ 与 xyz 中其余点的距离 d

5: if $d \leq radius$ then

6: 将该点记录为 $xyz_{i,j,k}$ 的邻域点

7: $n_{i,j,k}++$; else

9: continue

10: end if

11: end for

12: for $xyz_{i,j,k} \in xyz$ do

13: if $n_{i,j,k} > nsample$ then

14: 随机采样至 $nsample$ 个索引

15: else if $n_{i,j,k} = nsample$ then

16: continue 17: else

18: 将最后一个索引重复 $nsample - n_{i,j,k}$ 次

19: end if

20: end for

对输入维度为 $(B, N, 3)$ 的点云, BQ 得到了维度为 (B, N, S) 的索引张量, 其中 B 为 $batch$, N 为点云的数量, S 为邻域球中点的数量。

2.1.2 基于注意力的特征融合模块

在点云数据中, 不同类别物体的边界难以区分, 在中心点局部邻域内, 相同或不同类别的点对于中心点特征的提取均具有积极作用, 采用基于注意力的方法融合特征并生成特征张量, 网络的架构图如图4所示。

Graph generated by ball query

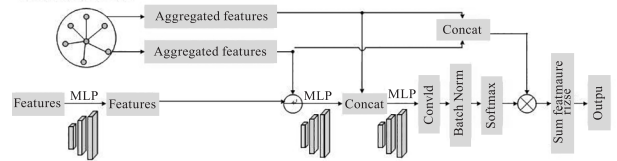


图4 基于注意力的特征融合模块的整体工作流程图

Fig. 4 The overall process of the attention-based

feature fusion module

在注意力模块中, 采用拓扑生成的局部图以及当前点云中的点特征作为模块的输入。在拓扑生成模块中, 首先得到邻居节点的索引矩阵 M , 然后基于该索引矩阵进一步聚合输入点云与特征向量, 分别得到聚合点云 A_p 和聚合特征 A_f 。 A_f 与原始输入特征进行差分, 得到局部拓扑内每个邻域特征相对于顶点 V_i 特征的偏移量, 然后经过 MLP 计算, 得到偏移特征的潜在表示 Q_L , 如式所示:

$$Q_L = Y(A_f - f) \quad (1)$$

式中, f 表示原始输入特征; Y 表示 MLP 的变换函数。

Q_L 与 A_p 在特征维度进行 Concat 拼接后, 经过权重编码阶段, 得到权重系数:

$$W = \text{softmax}(BN(\text{Conv}(\delta(\text{Concat}(Q_L, A_p)))))) \quad (2)$$

式中, BN 表示批归一化操作 (Batch Normalization, BN); δ 表示 MLP 变换函数。 softmax 函数将不同特征点的权重映射到 $(0, 1)$, 并使所有元素的和为 1。得到的聚合特征表示如下:

$$f_{\text{output}} = \text{sum}(W \cdot \text{Concat}(A_p, A_f)) \quad (3)$$

其中, sum 表示在邻域点所处维度上求和。

2.1.3 池化降采样模块

在拓扑生成和特征融合模块中, 网络仅构建了邻域特征的局部关联, 并通过注意力机制聚合特征, 但缺乏全局高阶特征提取能力。受 PointNet++ 启发, 本文设计了一种同步下采样架构, 在不同尺度特征中学习局部上下文信息。点云下采样模块通过最远点采样 (FPS) 从输入点云中选择子集, 并研究了最大池化、最小池化和平均池化三种池化方式, 具体结果将在消融试验中讨论。

2.2 基于 U-Net 架构的全局特征增强

通过 PointNet++ 中的三线性插值方法融合下采样阶段以及上采样过程中的高维中间特征, 然后采用 MLP 进行深度特征映射, 三线性插值定义如下:

$$f_i = \frac{\sum_{j=1}^M \omega_j(p_i) \cdot f_j}{\sum_{j=1}^M \omega_j(p_i)} \quad (4)$$

$$\omega_j(p_i) = \begin{cases} \frac{1}{\|p_i - p_j\|_2}, & \text{if } p_j \in N_3(p_i) \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

其中, f_i 是点 p_i 处插入的特征; M 为插值的邻域点的数量; $\omega_j(p_i)$ 为点 p_j 与点 p_i 相关性权重函数; $N_3(p_i)$ 表示与 p_i 相距最近的三个点的集合。

在下采样过程中的特征提取、特征融合以及点云下采样步骤均基于拓扑生成块, 其共享的子图索引在上采样过程中被用来指引点云语义信息的重建。一个上采样模块的完整数据流如图 5 所示。

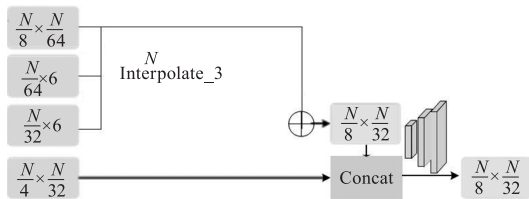


图 5 上采样模块工作流程图

Fig. 5 The overall process of the up-sampling module

3 实验与分析

3.1 数据集与评价指标

本文选取 S3DIS 室内数据集^[22]作为实验数据以评估本文提出的方法和对比方法在场景语义分割任务中的表现, 该数据集由斯坦福大学开发, 由 RGB-D 传感器采集, 包含了 3 个建筑物内的 6 个大型室内区域, 272 个房间, 共包括 13 个语义标签。对于数据集中的每一个点, 数据集中都给出了包括空间坐标, RGB 颜色以及所属类别的信息。本文选取六个大型室内区域中的 Area5 作为测试集, 其余五个区域作为训练集。

在部件分割任务中, 本文选取 ShapeNet 数据集^[23], 该数据集包括了 16 个对象, 16881 个 CAD 模型, 每个对象都包含了 2 ~ 5 个部件, 共计 50 个部件。

本文选取整体准确率 (Overall Accuracy, OA) 和平均交并比 (Mean Intersection over Union, mIoU) 来评估模型性能。其中, OA 表示预测正确的点的比例, mIoU 则表示在每一类中, 模型的预测和真实值的交集与并集比值的平均值。OA 和 mIoU 如公式 (6)、(7) 所示:

$$OA = \frac{\sum_{i=1}^c TP}{\sum_{i=1}^c (TP + FN)} \quad (6)$$

$$mIoU = \frac{1}{c} \cdot \sum_{i=1}^c \frac{TP}{TP + FN + FP} \quad (7)$$

其中, TP 表示实际为真, 预测也为真的样本数; FN 是实际为真, 预测为假的样本数; FP 是实际为假, 预测为真的样本数; c 表示标签数量。

3.2 实验设置

实验环境基于 Ubuntu18.04.2 操作系统, Intel (R) Xeon (R) E5-2680 v3 CPU, GeForce RTX 2080Ti GPU, PyTorch 深度学习框架。采用 Adam 为网络参数优化器, 初始学习率为 0.001, 动量设置为 0.9, 衰减率设为 0.0001, batch 设为 16, epoch 设为 150。每个点云中采样 2048 个点作为输入, 输入的 shape 为 $(16, 2048, 6)$ 。图拓扑采样中, 选取邻域点数量 S 为 16。

3.3 实验结果分析

3.3.1 语义分割

为了测试本文的方法在 S3DIS 数据集上的表现,

选取 PointNet^[12]、PointNet++^[13]、TangentConv^[8]、SPG^[24]、PointCNN^[14]、PCCN^[25]、AMFF-DGCNN^[26] 作为对比方法。所有的方法都在划分好的 S3DIS 数据集集中进行训练和测试以保证对比的公平性。实验结果如表 1 所示,可以看出,本文的方法在 S3DIS 数据集上的 OA 和 mIoU 分别为 86.7% 和 60.7%, 相比于 PointNet++, 分别提升了 3% 和 9%, 相比于其他对比方法, 也有较为显著地提升。在单个类别的分割表现中, 本文的方法在天花板、墙壁、椅子、地板上达到了最优的性能, 在窗户、桌子、书橱、杂物等类别中依旧表现突出。

本文在 Area5 测试集中随机选择了三个场景进行可视化验证。场景一(图 6 第一行)中, 本文方法在窗户识别上表现更精确, 但对柱子识别与原始语义信息存在差距。网络能够提取墙角处的墙柱信息, 但墙边的墙柱被错误判定为窗户。场景二(图 6 第二行)中, 本文方法相比 PointNet++ 在窗户和墙壁分割上更为精确, 尤其在复杂场景中, PointNet++ 将墙壁误判为杂物。场景三(图 6 第三行)中, PointNet++ 将书橱的大部分点误识别为门, 而本文方法能够准确分割出书橱的结构。

综上, 本文方法在三个场景中的分割效果都较为接近真实图像, 无论是大物体还是小物体的分割,

均能提供更精确的结果。这表明了本文所提出的拓扑感知和全局特征增强模块在提升点云语义分割任务中具有更精确的效果, 可极大程度地增强语义分割任务的精度。

3.3.2 部件分割

选取 PointNet^[12]、Kd-Net^[27]、GAPNet^[28]、SCN^[29]、RSNet^[30]、DGCNN^[31] 作为对比方法来测试本文的方法在部件分割任务中的性能。为了对比的公平性, 所选方法都使用相同的训练集和测试集进行训练和验证。选取每个类别的 mIoU 以及平均 mIoU 作为性能评价指标, 实验结果如表 2 所示。

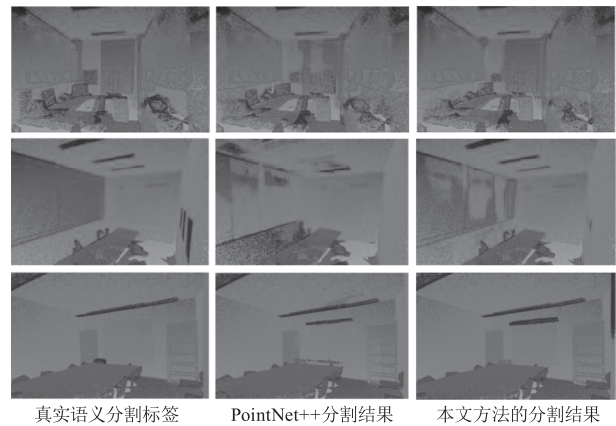


图 6 S3DIS 数据集上语义分割结果图

Fig. 6 Results of the semantic segmentation on S3DIS dataset

表 1 不同方法在 S3DIS 数据集上的分割精度对比(选择 Area5 作为测试集)

Tab. 1 Comparison of the segmentation accuracy employing different methods on S3DIS dataset (Area5 for testing set)

Methods	OA	mIoU	ceiling	floor	wall	beam	column	window	door	table	chair	sofa	bookcase	board	clutter
PointNet AMFF -	79.3	41.1	88.8	97.3	69.8	0.1	3.9	46.3	10.8	59.0	52.6	5.9	40.3	26.4	33.2
DGCNN	84.6	51.6	92.8	87.8	78.1	0.0	24.9	51.9	31.1	69.1	73.9	16.2	52.7	39.5	42.2
PointNet++	83.7	51.7	89.5	97.7	76.2	0.0	0.0	61.0	30.3	70.5	62.2	33.2	60.5	46.3	43.6
TangentConv	82.5	52.6	90.5	97.7	74.0	0.0	20.7	39.0	31.3	77.5	69.4	57.3	38.5	48.8	39.8
SPG	84.1	55.8	92.2	95.0	71.9	0.0	15.0	46.5	60.9	65.1	69.5	56.8	38.2	6.86	51.3
PointCNN	85.9	57.3	92.3	98.2	79.4	0.0	17.6	22.8	62.1	74.4	80.6	31.7	66.7	62.1	56.7
PCCN	-	58.3	92.3	96.2	75.9	0.3	6.0	69.5	63.5	66.9	65.6	47.3	68.9	59.1	46.2
Ours	86.7	60.7	93.1	97.9	79.4	0.0	16.5	62.6	46.4	74.9	81.1	49.9	68.1	68.4	51.3
Methods	OA	mIoU	ceiling	floor	wall	beam	column	window	door	table	chair	sofa	bookcase	board	clutter

表 2 不同方法在 ShapeNet 数据集上的部件分割实验结果对比

Tab. 2 Comparison of the component segmentation results on ShapeNet dataset using different methods

model	mean	bag	cap	car	chair	earphone	guitar	knife	lamp	laptop	motor	mug	pistol	rocket	skateboard	table
PointNet	83.7	78.7	82.5	74.9	89.6	73.0	91.5	85.9	80.8	95.3	65.2	93.0	81.2	57.9	72.8	80.6
Kd-Net	82.3	74.6	74.3	70.3	88.6	73.5	90.2	87.2	81.0	94.9	57.4	86.7	78.1	51.8	69.9	80.3
GAPNet	84.7	84.1	88.8	78.1	90.7	70.1	91.0	87.3	83.1	96.2	65.9	95.0	81.7	60.7	74.9	80.8
SCN	84.6	80.8	83.5	79.3	90.5	69.8	91.7	86.5	82.9	96.0	69.2	93.8	82.5	62.9	74.4	80.8
RSNet	84.9	86.4	84.1	78.2	90.4	69.3	91.4	87.0	83.5	95.4	66.0	92.6	81.8	56.1	75.8	82.2
DGCNN	84.7	85.7	84.8	78.2	90.5	75.4	91.1	87.2	82.4	95.6	62.9	94.5	80.7	63.6	75.4	81.8
Ours	85.1	80.6	85.7	79.1	90.9	69.5	90.8	87.6	83.8	95.4	67.5	92.8	81.2	57.1	74.9	82.6

可以看出,本文的方法取得了最优的部件分割效果,相比于 PointNet 和 DGCNN 的平均 mIoU 分别提高了 1.4% 和 0.4%。同时,本文的方法在椅子、刀具、灯、桌子这四个类别上取得了最优的语义分割结果,而在手机、吉他、电脑、手枪和滑板这五个类别上的分割表现优于大部分方法。

3.3.3 消融试验

为了进一步探究模型的性能,在 S3DIS 原始测试集 Area5 中进行消融试验。对于拓扑生成模块,本文在局部图拓扑生成阶段采用了不同数量的点进行消融试验。结果如表 3 所示。

表 3 S3DIS 数据集中的消融实验语义分割结果

Tab. 3 Semantic segmentation results of ablation experiments on the S3DIS dataset

Experimental Settings	OA	mIoU
global graph($S=8$)	81.8	55.0
global graph($S=16$)	83.2	56.8
Unet + global graph($S=1$)	82.7	55.6
Unet + global graph($S=4$)	86.1	57.6
Unet + global graph($S=8$)	86.2	59.8
Unet + global graph($S=16$)	86.7	60.7

从表中可以看出,当降低邻域点的数量时,OA 和 mIoU 都有下降的趋势,当 S 取 1 时,mIoU 下降了 5.1%。而去除了 U-Net 架构的网络,分割性能也呈现下降趋势,当 S 取 16 和 8 时,mIoU 分别下降了 3.9% 和 4.8%,OA 分别下降了 3.5% 和 4.4%。消融试验的实验结果进一步证明了本文的方法在语义分割任务中的优势。

同时,本文还基于池化模式的选择进行了消融试验,分别使用最大池化、最小池化和平均池化作为

下采样策略处理数据,结果如表 4 所示。

最大池化在本文语义分割任务中取得了最优效果,因为它能够更好地保留边缘特征信息。相比之下,最小池化和平均池化未能有效保留边缘信息,因此最大池化更符合边缘信号特征,表现最佳。

表 4 S3DIS 数据集中不同池化策略的语义分割结果

Tab. 4 Semantic segmentation results of different down-sampling strategies on the S3DIS dataset

Down-sampling strategies	OA	mIoU
Max pooling	86.7	60.7
Min pooling	86.2	59.0
Ave pooling	86.4	59.5

4 结 语

本文提出了一种基于拓扑感知和全局特征增强的语义分割算法,旨在解决边缘特征提取困难和上采样过程中细粒度特征表达不足的问题。通过构建输入点云的邻域图拓扑,逐步下采样提取全局特征,并在上采样阶段使用 Unet 架构增强全局特征。通过 ShapeNet 和 S3DIS 数据集验证,本文方法在分割性能上优于现有深度学习方法。未来将研究语义特征的深层次关联,设计更先进的网络应对复杂点云场景。

参考文献:

- [1] Guo Y L, Wang H Y, Hu Q Y, et al. Deep learning for 3D point clouds: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(12): 4338 - 4364.
- [2] 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 (CVPR). Honolulu:CVF,2017: 652 – 660.
- [3] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39 (4):640 – 651.
- [4] Zhang R, Li G Y, Wunderlich T, et al. A survey on deep learning-based precise boundary recovery of semantic segmentation for images and point clouds [J]. International Journal of Applied Earth Observation and Geoinformation, 2021, 102:102411.
- [5] 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). Piscataway, NJ:IEEE,2015:922 – 928.
- [6] Tehapmi L, Choy C, Armeni I, et al. SEGCloud: semantic segmentation of 3D point clouds [C]//2017 International Conference on 3D Vision (3DV). Piscataway, NJ:IEEE, 2017:537 – 547.
- [7] Riegler G, Ulusoy A O, Geiger A. OctNet: learning deep 3D representations at high resolutions [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ:IEEE,2017:6620 – 6629.
- [8] Tatarchenko M, Park J, Koltun V, et al. Tangent convolutions for dense prediction in 3D [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE,2018:3887 – 3896.
- [9] Wu B C, Wan A, Yue X Y, 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). Piscataway, NJ:IEEE,2018:1887 – 1893.
- [10] Wu B C, Zhou X Y, Zhao S C, 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). Piscataway, NJ: IEEE, 2019: 4376 – 4382.
- [11] Kumar K S C, Al-Stouhi S. Multi-scale voxel class balanced ASPP for LIDAR pointcloud semantic segmentation [C]//2021 IEEE Winter Conference on Applications of Computer Vision Workshops (WACVW). Piscataway, NJ: IEEE,2021:117 – 124.
- [12] 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 (CVPR). Honolulu:CVF,2017: 652 – 660.
- [13] 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. Cambridge: Curran Associates Inc. ,2017:5105 – 5114.
- [14] LI Y Y, BU R, SUN M C, et al. PointCNN: convolution on x-transformed points [C]//Advances in Neural Information Processing Systems. Cambridge: MIT Press,2018,31: 828 – 838.
- [15] 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). Piscataway, NJ:IEEE,2019:6410 – 6419.
- [16] Zhao H S, Jiang L, Fu C W, et al. PointWeb: enhancing local neighborhood features for point cloud processing [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ:IEEE, 2019:5560 – 5568.
- [17] Wang L, Huang Y C, Hou Y L, et al. Graph attention convolution for point cloud semantic segmentation [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ:IEEE,2019:10288 – 10297.
- [18] Zhang Z Y, Hua B S, Chen W, et al. Global context aware convolutions for 3D point cloud understanding [C]//2020 International Conference on 3D Vision (3DV). Piscataway, NJ:IEEE,2020:210 – 219.
- [19] Huang J, You S Y. Point cloud labeling using 3D convolutional neural network [C]//2016 23rd International Conference on Pattern Recognition (ICPR). Piscataway, NJ: IEEE,2016:2670 – 2675.
- [20] Dricot A, Pereira F, Ascenso J. Rate-distortion driven adaptive partitioning for octree-based point cloud geometry coding [C]//2018 25th IEEE International Conference on Image Processing (ICIP). Piscataway, NJ: IEEE, 2018: 2969 – 2973.
- [21] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation [C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Cham:Springer,2015:234 – 241.
- [22] Armeni I, Sener O, Zamir A R, et al. 3D semantic parsing

- of large-scale indoor spaces [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ:IEEE,2016;1534 – 1543.
- [23] Yi L, Kim V G, Ceylan D, et al. A scalable active framework for region annotation in 3D shape collections [J]. ACM Transactions on Graphics, 2016, 35(6):210.
- [24] Landrieu L, Simonovsky M. Large-scale point cloud semantic segmentation with superpoint graphs [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE,2018;4558 – 4567.
- [25] Wang S L, Suo S, Ma W C, et al. Deep parametric continuous convolutional neural networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE,2018;2589 – 2597.
- [26] Hao W, Wang H X, Wang Y. Semantic segmentation of three-dimensional point cloud based on spatial attention and shape feature [J]. Laser & Optoelectronics Progress, 2022, 59(8):0828004. (in Chinese)
郝雯,王红霄,汪洋. 结合空间注意力与形状特征的三维点云语义分割 [J]. 激光与光电子学进展, 2022, 59(8):0828004.
- [27] Klokov R, Lempitsky V. Escape from Cells: deep Kd-networks for the recognition of 3D point cloud models [C]//2017 IEEE International Conference on Computer Vision (ICCV). Piscataway, NJ:IEEE,2017;863 – 872.
- [28] Chen C, Fragonara L Z, Tsourdos A. GAPointNet: graph attention based point neural network for exploiting local feature of point cloud [J]. Neurocomputing, 2021, 438:122 – 132.
- [29] Xie S N, Liu S N, Chen Z Y, et al. Attentional ShapeContextNet for point cloud recognition [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE,2018;4606 – 4615.
- [30] Huang Q G, Wang W Y, Neumann U. Recurrent slice networks for 3D segmentation of point clouds [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE,2018;2626 – 2635.
- [31] Wang Y, Sun Y B, Liu Z W, et al. Dynamic graph CNN for learning on point clouds [J]. ACM Transactions on Graphics, 2019, 38(5):146.