

文章编号 1004-924X(2024)24-3658-16

整合边缘卷积与全局-局部自注意力的 机载点云分类

涂静敏¹, 严进¹, 李礼², 姚剑^{1,2}, 李婕^{1*}, 康妍斐³

(1. 湖北工业大学 太阳能高效利用及储能运行控制湖北省重点实验室, 湖北 武汉 430068;

2. 武汉大学 遥感信息工程学院, 湖北 武汉 430079;

3. 武汉市勘察设计有限公司, 湖北 武汉 430020)

摘要: 激光点云分类是实现三维场景理解的基础。针对机载点云大场景分类中存在的特征表达不足、样本类别不均衡的问题, 本文提出一种整合边缘卷积与全局-局部自注意力的机载点云分类方法。首先, 以 U-net 为网络框架, 融合 Point Transformer 与边缘卷积模块, 使得模型能够关注到复杂地物边界和纹理信息, 获得表达能力更好的局部几何特征。其次, 创新性地提出一种融合全局上下文信息和局部结构特征的自注意力机制, 全局自注意力模块倾向于整个输入序列的信息, 而局部自注意力模块则更侧重于局部区域的细节。两种机制结合增强了对长距离依赖关系和局部结构的捕捉, 同时使得模型能够兼顾少数类别的关键特征, 在一定程度上降低样本类别不均衡对分类精度的影响, 有助于提高模型对复杂地物关系的分类能力。最后, 在公开的 ISPRS-3D 数据集和 WHU-Urban3D 数据集上对本文所提出的方法进行验证, 实验结果表明, 该方法在两个数据集上的分类精度分别为 82.5% 和 87.4%, 优于 PointNet++, Stratified Transformer 等经典网络及 ISPRS 3D 官网竞赛网络, 可有效提升机载点云分类精度。

关键词: 机载激光雷达; 点云分类; 边缘卷积; 全局-局部自注意力; U-net

中图分类号: TN957.52 **文献标识码:** A **doi:** 10.37188/OPE.20243224.3658

Airborne point cloud classification integrating edge convolution and global-local self-attention

TU Jingmin¹, YAN Jin¹, LI Li², YAO Jian^{1,2}, LI Jie^{1*}, KANG Yanfei³

(1. Hubei Collaborative Innovation Center for High-efficiency Utilization of Solar Energy,

Hubei University of Technology, Wuhan 430068, China;

2. School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China;

3. Wuhan Survey and Design Co., LTD., Wuhan 430020, China)

* Corresponding author, E-mail: jielonline@hbut.edu.cn

Abstract: Laser point cloud classification is the basis for 3D scene understanding. In order to solve the problems of insufficient feature expression and unbalanced sample categories in the classification of large scenes of airborne point clouds, this paper proposed an airborne point cloud classification method that inte-

收稿日期: 2024-07-08; 修订日期: 2024-09-21.

基金项目: 国家自然科学基金项目 (No. 42301515); 智能光电系统感知及应用四川省高校重点实验室开放基金 (No. ZNGD2308); 湖北工业大学博士科研启动基金 (No. XJ2021004501); 实景地理环境安徽省重点实验室开放课题 (No. 2022PGE008); 深圳市科技重大专项 (No. KJZD20230923114611023)

grates edge convolution and global-local self-attention. Firstly, U-net was used as the network framework to integrate Point Transformer and edge convolution modules, so that the model could pay attention to complex ground object boundaries and texture information, and obtain local geometric features with better expressive ability. Secondly, a self-attention mechanism that integrates global context information and local structural features was innovatively proposed, and the global self-attention module tended to the information of the entire input sequence, while the local self-attention module pays more attention to the details of the local region. The combination of the two mechanisms enhanced the capture of long-distance dependence and local structure, and at the same time, the model could take into account the key features of a few categories, reduced the impact of sample category imbalance on the classification accuracy to a certain extent, and helped to improve the classification ability of the model for complex ground object relationships. Finally, the proposed method was verified on the public ISPRS-3D dataset and WHU-Urban3D dataset, and the experimental results show that the classification accuracy of the proposed method on the two datasets is 82.5% and 87.4%, respectively, which is better than that of the classical networks such as PointNet++ and Stratified Transformer and the ISPRS 3D official website competition network, which can effectively improve the classification accuracy of airborne point clouds.

Key words: airborne lidar; point cloud classification; graph convolution; global-local self-attention; U-net

1 引 言

机载激光雷达(Light Laser Detection and Ranging, LiDAR)利用激光束扫描地表,结合定位系统获取高分辨率三维地图数据,具有快速获取大范围、高精度地面信息的优势^[1-3]。分类是通过特定的算法和处理流程,将数据按照不同的特征和属性进行分类,而机载点云分类是三维场景理解的基础,在地形测绘^[4]、城市规划^[5]、森林监测^[6]、电力线检测^[7]和道路检测^[8]等领域起着重要作用。然而,机载点云分类在处理大规模点云数据时,面临着数据噪声、不完整性、不均匀采样,地物复杂等问题^[9-11],因此机载点云分类仍具有进一步探索的意义。

当前,机载点云分类方法多通过传统机器学习方法和深度学习方法实现。传统机器学习方法在点云分类中常采用基于手工提取特征的方法。例如Zhang等^[12]提出的支持向量机(Support Vector Machine, SVM),He等^[13]提出的随机森林(Random Forest, RF)以及Munoz等^[14]提出的马尔可夫随机场(Markov Random Field, MRF)和Niemeyer等^[15]提出的条件随机场(Conditional Random Field, CRF)等。该类方法首先对点云进行特征提取,包括几何特征(如法向量、曲率)、统计特征(如点密度)、局部特征(如形状上下文

描述子),再利用这些特征训练分类器,实现对点云中不同类别的地物或物体的分类和分割。传统机器学习方法的挑战在于需要手动设计特征提取方法,且对点云数据的非结构性和高维性处理较为复杂,容易受到噪声和采样不均匀等问题的影响,导致分类精度不高。

与传统机器学习方法相比,深度学习方法通过神经网络结构自动学习数据的特征表达,无需手动设计特征提取器,能够处理大规模高维数据并实现端到端的学习和预测^[16]。在早期点云分类的研究上,多以基于体素或基于投影为主。基于体素的方法例如Rabbani等^[17]提出使用光滑度约束方法对点云进行分割,通过对点云数据进行体素化表示,实现了对点云中光滑度的建模和约束,从而实现了有效的分割。基于投影的方法例如Chen等^[18]提出了一种多视角三维物体检测网络,首先通过将点云数据投影到多个视角,再利用深度学习方法实现了对三维物体的语义分割和检测。然而无论是基于体素或投影方法都存在信息丢失、计算复杂度高、不规则形状难以处理等缺点,在处理大规模、复杂点云数据中存在一定的局限性。随着PointNet^[19]和PointNet++^[20]模型的出现,越来越多的学者关注到基于点的端到端的深度学习方法,由此也涌现了一系列的改进方法。例如Li等^[21]提出了一种

3D Multiscale PointNet的方法,通过多尺度特征学习和聚合,改进了PointNet++的性能,在点云处理任务中取得了更好的效果。受自注意力机制的启发,Zhao等^[22]提出的Point Transformer模型基于自注意力机制实现网络模型的构建。该模型通过设计自注意力层,结合位置编码构建局部Transformer block,实现语义分割,部件分割和分类等任务。然而,计算机视觉领域的上述基于点的深度学习模型只注重探索局部几何结构^[23],而忽略了点与点之间的全局上下文关系。此外,上述深度学习方法大多对室内场景点云进行分类处理,但从离散、规模较大、地物复杂、类别分布不均衡且包含噪声的机载点云数据中准确分类各种典型地物目标的难度较大,成为近年来相关领域的研究热点之一^[24]。

综上所述,本文提出一种整合边缘卷积的全局-局部自注意力机制神经网络(Edge Convolution And Global-Local Self Attention Neural Network, E-GLAT)。其优势在于:(1)所提出的网络可以以端到端的方式进行训练,整体网络参考U-net结构设计,融合Point Transformer,增加边缘卷积模块(EdgeConv Block),根据邻域点的距离信息构造空间图以提取点云的相邻关系特征,使得模型能够关注到复杂地物边界和纹理信息,

从而提高整体分类准确性;(2)设计全局-局部自注意力机制模块(Global-Local Self Attention Block, GLAT),在特征提取阶段引入全局-局部自注意力机制,使得模型更加重视少数类别的关键特征,在一定程度上可以解决类别分布不均衡的问题,从而提高少数类别的分类准确性;(3)本文在ISPRS-3D数据集和WHU_ALS数据集上设计三组对比实验:消融实验、深度学习经典网络对比实验、ISPRS官网竞赛对比实验,实验结果进一步证明本文所提出的整合边缘卷积的全局-局部自注意力机制神经网络(E-GLAT)的有效性。

2 整合边缘卷积与全局-局部自注意力的机载点云分类方法

2.1 整体网络结构

为了实现机载点云的有效分类,本文设计的整合边缘卷积与全局-局部自注意力网络主要由三部分组成:边缘卷积模块、全局自注意力模块和局部自注意力模块,整体网络设计如图1所示。首先,通过边缘卷积对机载点云进行特征提取,修正局部信息的重要程度;其次,使用Transformer的自注意力机制来有效地提取点云数据

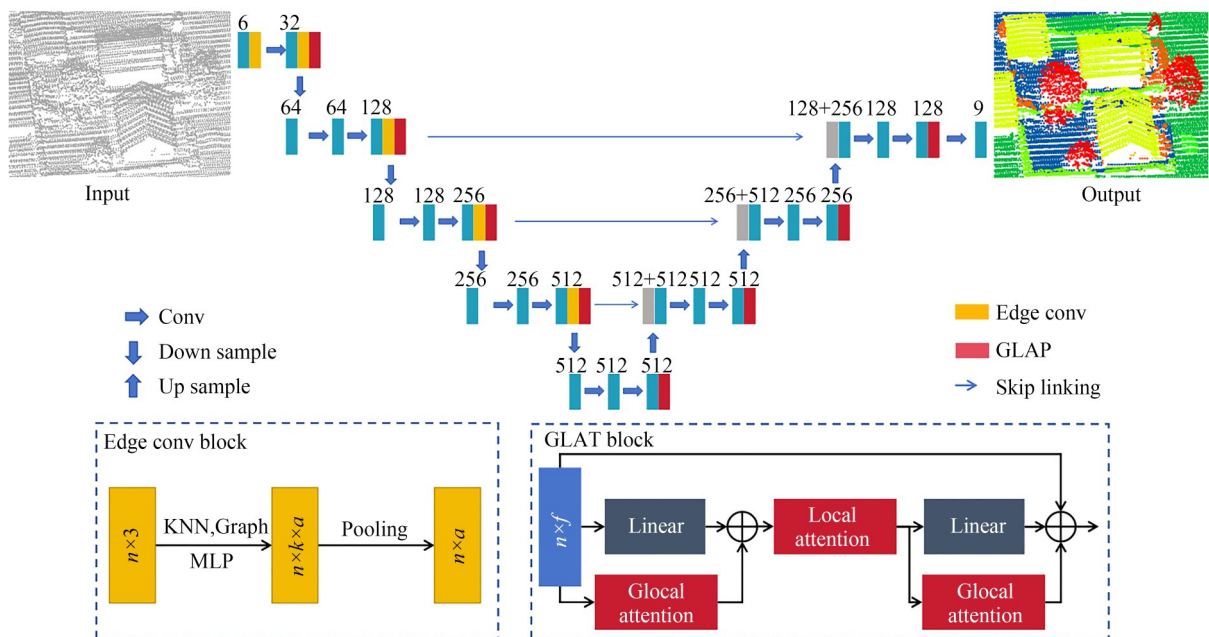


图1 整体网络结构

Fig. 1 Overall network structure

中的全局和局部信息,并建立点与点之间的关联性;最后通过全连接层,将学习到的特征映射到具体的类别,从而得到分类结果。

整体网络结构采用 U-net 网络设计。在传统下采样阶段,由于降低了点云数据的分辨率,导致信息的丢失,因此,本文考虑在下采样时增加边缘卷积模块捕获局部特征。此外,本文设计一个 GLAT 模块对点云的全局特征和局部特征进行有效的提取,该模块设计包括特征嵌入、位置编码、局部自注意力、全局自注意力和特征融合等步骤。在每个上采样阶段,点集大小逐步增加到 128, 512, 2 048, 8 192 个点。为了在下采样阶段中合并低级别信息,本文在下采样和上采样阶段之间添加跳跃连接,将来自下采样阶段的低级别特征与上采样阶段具有相同分辨率大小的特征连接起来。最后一个上采样层的点特征被输入到一个全连接层,从而输出所有输入点的语义标签。

2.2 边缘卷积模块

在点云分类任务中,一个常见的挑战是缺乏对点与点之间拓扑关系的有效描述。这意味着可能会失去点云中重要的局部结构信息,而这些

信息对于分类物体、场景或执行其他任务至关重要。为了解决这个难题,本文引入边缘卷积模块(EdgeConv)。如图 2 所示,边缘卷积能够有效捕捉点与点之间的连接和相对位置关系来增强点云数据的局部特征表达能力。通过边缘卷积,不仅可以顾及每个点本身的特征,还可以充分利用周围邻近点的信息,从而提高后续任务的分类效果。

边缘卷积操作包含边缘特征提取、深度特征提取以及局部特征提取。具体步骤如下:

(1) 边缘特征提取。首先,对输入的每个点 $P_i (i=0, 1, \dots, N)$ 周围的 K 个邻近点 $P_j (j=0, 1, \dots, k)$, 构成 KNN 图。其次,通过计算中心点(即 P_i)与其周围邻近点 P_j 的特征差异来捕捉局部结构信息。这些特征差异被组合到中心点的特征中,形成了边缘特征 e_{ij} 。得到的边缘特征中包含了每个点的局部信息,过程如图 2(a) 所示。边缘特征 e_{ij} 的表达式为:

$$e_{ij} = [f(P_i), f(P_j - P_i)], \quad (1)$$

其中: $f(P_i)$ 为中心点的特征; $f(P_j - P_i)$ 为邻域节点特征减中心点的特征。

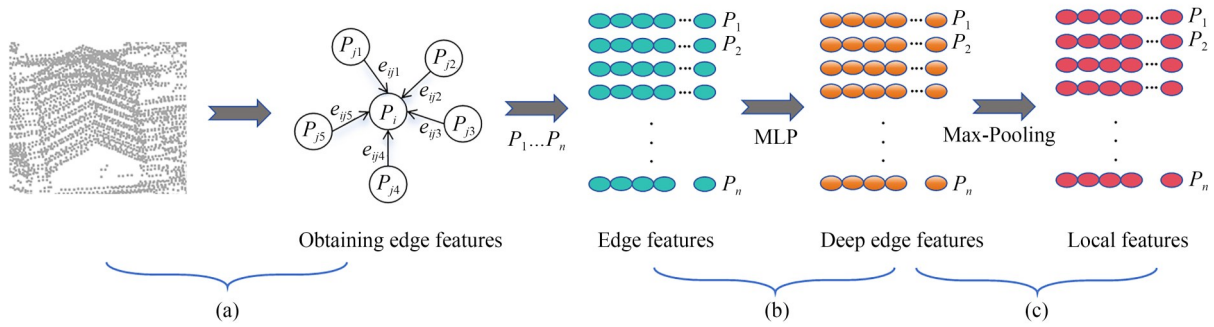


图 2 局部特征提取流程图

Fig. 2 Flow chart of local feature extraction

(2) 深度特征提取。使用 MLP 对边缘特征进行深度特征提取,旨在进一步丰富和提取深层特征,从而增强特征的表达能力,过程如图 2(b) 所示。提取的深度边缘特征 e'_{ij} 的表达式为:

$$e'_{ij} = h[f(P_i), f(P_j - P_i)], \quad (2)$$

其中, h 为多层感知机。

(3) 局部特征提取。最后,通过最大池化操作(Maxpooling),从每个点的深度边缘特征中提

取出最显著的特征,形成了点云的局部特征。这些局部特征具有更深层次的信息,如图 2(c) 所示。局部特征的 P'_i 表达式为:

$$P'_i = \text{Maxpooling}(i, j \in E) h[f(x_i), f(P_j - P_i)]. \quad (3)$$

通过以上操作,能够有效地提取到点云的局部特征,从而在后续的点云分类任务中发挥重要作用。

2.3 GLAT

全局-局部自注意力机制模块(Global-Local Self Attention Block, GLAT)的设计结合了两种自注意力机制的优势。全局自注意力机制通过捕捉点云中长距离的依赖关系,为模型提供了丰富的全局上下文信息,有助于模型识别和理解整个场景的宏观结构。同时局部自注意力模块专注于捕捉点云的局部结构特征,它关注于每个点周围的细节和邻近区域,从而使得模型能够细致地理解局部的几何形状和特征。

具体实现如图 3 所示,首先,将输入的点云特征信息通过一个线性层 linear 进行初步特征变换和全局自注意力层捕捉整个点云的全局上下文信息,并将两者信息叠加得到包含全局信息的特征。其次,包含全局信息的特征被送入局部自注意力层得到进一步丰富的局部特征信息。再次,将得到的局部特征信息经过线性层映射到更高的维度和全局自注意力层进行处理,以再次融入全局上下文信息。最后,将原始输入特征与变换后的特征相加得到最终的输出特征。这种融合不仅保留了两种机制各自的优势,还增强了模型对不同尺度特征的表达能力,从而有效地实现机载点云分类任务。

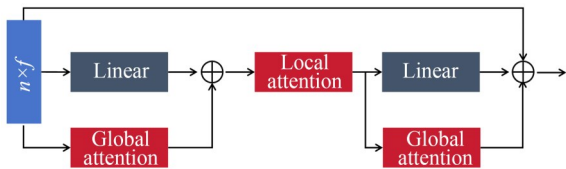


图 3 GLAT 模块示意图

Fig. 3 Schematic diagram of the GLAT module

2.3.1 Local Attention

点云数据本质上是由一系列不规则地分布在三维空间中的点集合 $P = \{P_i\} (i = 0, 1, \dots, N)$ 组成。注意力机制在处理点云数据时能够适应点云的不规则性和灵活性,有效地捕获点与点之间的关系,从而更好地处理点云分类、分割等任务^[29]。因此,本文的局部自注意力模块采用基于半径的球查询方法,即在每个数据点周围的局部邻域内应用自注意力网络。对于空间中的任意一点 $P_i \in P$ 在给定半径范围内取 K 个点组成点云子集 $P_j (j = 0, 1, \dots, k)$,其局部邻域保证了固定

的区域尺度,从而使局部区域特征在空间上更具泛化性。 P_i 的输出特征可以表示为 y_i ,其计算方式如式(4)所示:

$$y_i = \sum_{x_j \in \mathcal{X}} \rho \left(\Upsilon \left(\varphi(P_i) - \psi(P_j) + \theta(P_i - P_j) \right) \right) \odot \left(\alpha(P_j) + \theta(P_i - P_j) \right), \quad (4)$$

其中: φ, ψ 和 α 是逐点特征变换(线性投影), ρ 是归一化函数 Softmax; Υ 是 MLP 映射函数;位置编码函数 θ 是具有两个线性层和一个 ReLU 非线性的 MLP。Local Attention 层设计如图 4 所示。

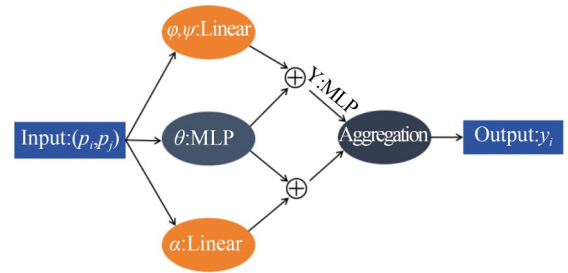


图 4 Local Attention 模块示意图

Fig. 4 Schematic diagram of the local attention module

2.3.2 Global attention

Point Transformer 网络在 K 邻域上计算自注意力,即一种特定类型的自注意力模式-局部自注意力^[25]。局部自注意力只关注 K 个最近的邻域点,缺少点与点之间的全局上下文关系。因此,本文在 Point Transformer 网络基础上增加全局自注意力。全局自注意力在本文中以对称的方式实现,在自注意力(Self-attention)机制中,查询、键和值都是从同一个输入序列生成的,计算注意力权重后将其加权应用于值,确保了注意力计算的对称性。这样的实现方式使得每个元素能够基于整个输入序列的信息进行调整和更新,从而有效捕捉长距离的依赖关系。

对于任意一点 $P_i \in P$,其全局自注意力的关键是选择一些全局标记点 $P_g \in P$,使得信息能够在未被局部自注意力关注的点之间更好地流动。全局自注意力在局部自注意力基础上补充所有点之间的全局上下文信息。全局标记点 P_g 的选取规则如下:首先,依次计算空间中任意一点 P_i 到其余所有点 $P_r \in P (r \neq i, r = 0, 1, \dots, N)$ 之间的平方距离 $D(i, r)$,以获得维度为 $N \times N$ 的距离矩阵,其可以表示为:

$$D(i, r) = \sum_{i=1}^N \sum_{r=1}^N (x_i - x_r)^2 + (y_i - y_r)^2 + (z_i - z_r)^2. \quad (5)$$

其次,对距离矩阵 $D(i, r)$ 中的每一行元素按照从小到大的顺序进行排序,并返回排序后的索引 D' 。选择索引 D' 每一行前 g 个元素的索引得到距离每个点最近的 g 个邻居点的索引,从而选出 g 个全局标记点 P_g 。选出的全局标记点 P_g 基于整个点云的相对距离排序,而不是局限于局部邻域,有助于网络中所有点的信息流动,从而有效地提取所有点之间的全局上下文信息。最后,使用减法关系并将位置编码 θ 添加到注意力向量 $\mathbf{r}$ 和转换后的特征 α ,输出特征可以表示为:

$$y'_i = \sum_{x_i \in \mathcal{X}} \rho \left(\mathbf{r} \left(\varphi(P_i) - \psi(P_g) + \theta(P_i - P_g) \right) \right) \odot \left(\alpha(P_i) + \theta(P_i - P_g) \right), \quad (6)$$

其中: φ, ψ 和 α 是逐点特征变换(线性投影), ρ 是归一化函数 Softmax; $\mathbf{r}$ 是MLP映射函数;位置编码函数 θ 是具有两个线性层和一个ReLU非线性的MLP。Global Attention层设计如图5所示。

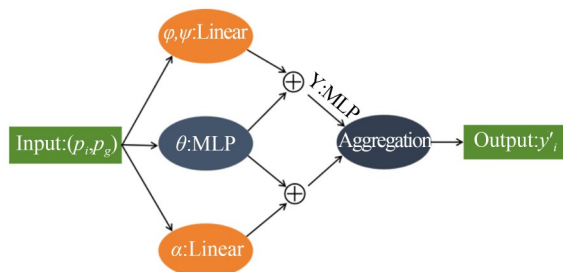


图5 Global Attention模块示意图

Fig. 5 Schematic diagram of the Global Attention module

3 实验

3.1 评价指标

根据ISPRS三维激光测距竞赛的标准评价指标,本文采用精确率、召回率、F1得分和总体准确率(OA)来评价点云分类的性能^[26]。OA(指定为测试集中正确分类的点的百分比)检查所有类别的分类准确度。F1分数基于分类模型的精确度和召回率来衡量每个类别的分类性能,当每个类别中的点数存在较大差异时,F1分数是比OA更好的选择。OA、精确度、召回率和F1分数指定如下:

$$OA = \frac{TP + TN}{TP + TN + FP + FN}, \quad (7)$$

$$precision = \frac{TP}{TP + FP}, \quad (8)$$

$$recall = \frac{TP}{TP + FN}, \quad (9)$$

$$F1 = 2 * \frac{precision * recall}{precision + recall}, \quad (10)$$

其中:TP为预测该类别正确的样本数量,FP为预测该类别错误的样本数量,TN为预测非该类别正确的样本数量,FN为该类别没有被预测出来的样本数量。

3.2 实验结果

3.2.1 数据集

为验证本文所提出模型的有效性,拟采用国际摄影测量与遥感协会(ISPRS)提供的Vaihingen地区的LiDAR数据集(ISPRS-3D数据集)^[27]和武汉大学发布的WHU-Urban3D数据集^[28]对所提方法进行验证。

3.2.1.1 ISPRS-3D数据集

ISPRS-3D数据集由来自德国Vaihingen地区的Leica ALS 50系统采集的机载激光扫描数据组成。LiDAR数据被分类为九个语义类(Niemeyer等人,2014):电力线(p_l),低矮植被(l_v),不透水层(i_s),车辆(car),栅栏(f_h),屋顶(r_f),外墙(f_e),灌木(s_b),树(t_e)。ISPRS 3D标注网站上提供了三个不同区域的场景:一个具有753 876个点的场景用作训练集,另外两个具有411 722个点的场景用作测试集。每个点云数据提供了空间XYZ坐标、强度、返回数、返回数和语义标签。如图6所示数据集中的三个场景,并且参考了每个类别的渲染颜色。

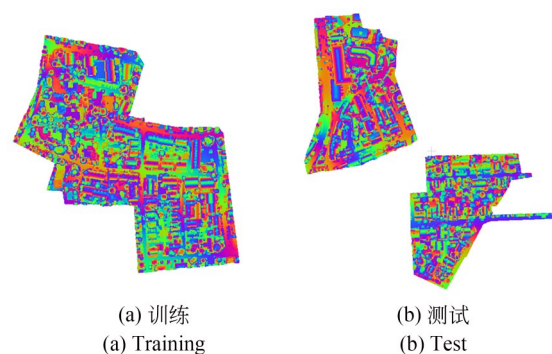


图6 ISPRS-3D数据集

Fig. 6 ISPRS-3D dataset

3. 2. 1. 2 WHU-Urban3D 数据集

WHU-Urban3D 数据集在中国繁华的城市大都市采集,覆盖面积超过 $3.6 \times 10^6 \text{ m}^2$ 。该数据被分类为九个语义类:建筑(b_d),树(t_e),植被(v_g),低矮植被(l_v),路灯(l_g),电力线(e_c),地面(g_d),其他(o_t),车辆(v_h)。如图 7 所示,整个区域的 ALS 点云被划分为数千个 $200 \text{ m} \times 200 \text{ m}$ 块,其中 80 多个块被选中进行标注。而本文选取了数据集的 35 个场景,其中 25 个场景被拆分为训练,10 个场景用于测试。



图 7 WHU-Urban3D 数据集
Fig. 7 WHU-Urban3D dataset

3. 2. 2 实验细节

点云的原始 3D 坐标和可选特征(本文采用从机载 LiDAR 数据获取的强度值)输入到本文的 E-GLAT 模型中。然而,由于 GPU 内存有限,整个训练点集不能直接输入到网络中。为了解决以上问题,把整个场景分成一组 $30 \text{ m} \times 30 \text{ m} \times 40 \text{ m}$ 的长方体区域。在模型训练阶段,为了提高模型的鲁棒性和防止过拟合,首先随机选取几个长方体区域,从每个长方体区域中随机选取 8 192 个点,并随机丢弃其中 12.5% 的点。在模型测试阶段,来自测试集中的分割长方体区域的所有点都被馈送到训练模型中,以实现点对点分类。如图 8 所示,以 ISPRS-3D 数据集为例,由于每个长方体区域的点的数量不同,在场景边缘出现的分裂小破碎区域被合并到周围的较大区域中,以确保每个长方体区域的完整性。该模型基于 PyTorch 框架实现,使用 Adam 优化器,学习率

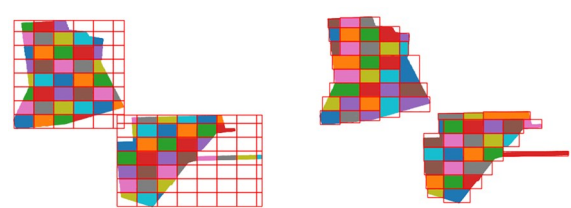


图 8 ISPRS-3D 数据预处理
Fig. 8 ISPRS-3D data preprocessing

为 0.01,每 30 步将学习率减小为原来的 0.1 倍,训练 100 轮直至网络收敛。批量大小设置为 8。

3. 2. 3 实验结果与分析

为了进一步验证本文所提出网络结构的有效性,本文进行了参数设置实验,同时在 ISPRS-3D 数据集和 WHU-Urban3D 数据集上设计了多组实验,并从视觉效果和定量定性两个维度对实验结果进行分析,实验结果表明,本文所设计的边缘卷积模块和全局-局部自注意力模块有助于机载点云分类精度的提高。

3. 2. 3. 1 参数设置

为了控制输入点集的密度,本文将训练块中采样点的数量控制在 8 192 个点。以采样点集中任意一点为圆心, r 为搜索半径所确定区域内的点即为 K 的候选点。因此, K 和 r 的合理取值使得模型能够更可靠和稳定地表示所捕获的特征。本文讨论当采样点为 8 192 时,邻近点 K 和搜索半径 r 的取值,以确定它们的最佳配置。如表 1 所示,以总体准确率(OA)和平均 F1 评分作为评价分类性能的关键指标。由实验可知,初始取值为 $K=15, r=0.1$ 时,总体准确率为 81.8%,平均 F1 得分为 65.2%;随着邻近点数 K 和搜索半径 r 的增大总体准确率和平均 F1 得分均有所提高,当 $K=20, r=0.2$ 时模型获得更稳定的特征,达到了最好的分类结果。注意,邻近点数量和搜索半径过大时可能对分类性能产生负面影响。当 K, r 分别取值为 20 和 0.2 时,我们的模型达到了最佳性能,总体准确率为 82.5%,平均 F1 得分为 66.2%。

表 1 参数设置

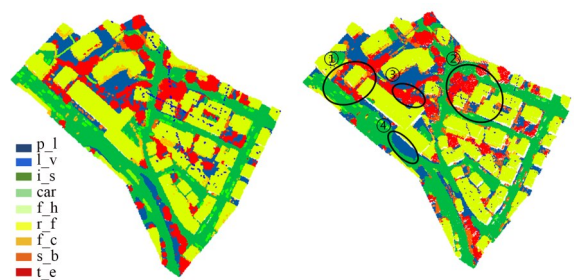
Tab. 1 Parameter settings

K	r	OA	Avg.F1
15	0.1	81.8	65.2
15	0.2	82.1	65.7
15	0.3	81.7	64.9
20	0.1	82.2	65.9
20	0.2	82.5	66.2
20	0.3	82.0	65.7
25	0.1	81.9	65.4
25	0.2	82.3	66.0
25	0.3	82.1	65.8

3.2.3.2 视觉效果分析

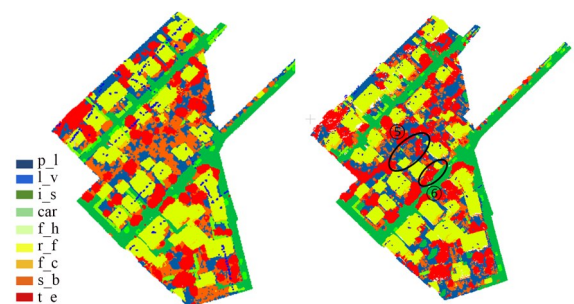
ISPRS-3D数据集原始数据真实标签及本方法的分类结果如图9(a)和图9(b)所示,对比发现,由于本文提出的边缘卷积模块和全局-局部自注意力机制模块共同作用于点云数据的特征提取,边缘卷积模块专注于捕捉点云中的局部空间结构和边界信息,而全局-局部自注意力机制模块进一步增强模型对整个点云场景的全局上下文理解。这种结合使得网络能够有效地解析和提取复杂的地物特征,从而对大部分数据点进行了正确预测,例如①和②处,将房屋,灌木,车,和不透水表面进行了正确分割,其中屋顶和不透水表面的分割效果较好,F1分数超过了90%。但由于不同尺度的地物混杂且拓扑特征相似,对部分地物的区分存在一定欠缺,例如③处将低矮植被错分为灌木,④处将不透水表面错分为低矮植被,⑤和⑥处将灌木错分为树和低矮植被。

WHU-Urban3D数据集原始数据真实标签及本方法的分类结果如图10(a)和图10(b)所示,对比发现本方法对大部分数据点进行了正确预测,例如①和②处,将建筑,树,植被,车和地面进



(a) 真实标签(左)和本方法的分类结果(右)

(a) Ground truth (left) and classification results of this method (right)

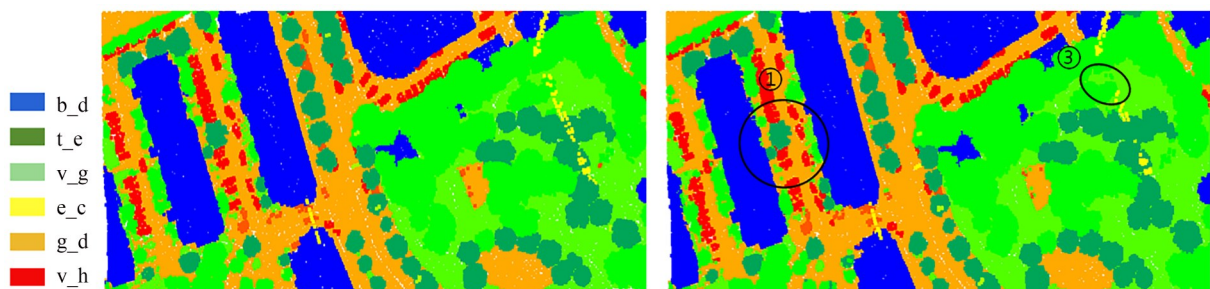


(b) 真实标签(左)和本方法的分类结果(右)

(b) Ground truth (left) and classification results of this method (right)

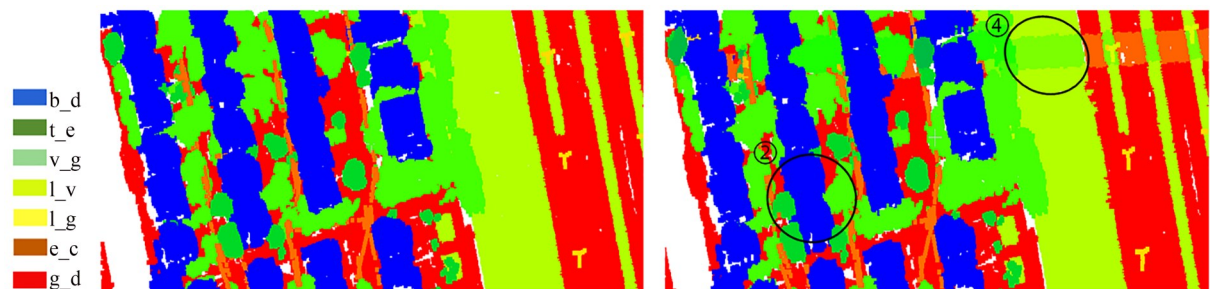
图9 ISPRS-3D数据集分类结果

Fig. 9 Classification results of ISPRS-3D dataset



(a) 真实标签(左)和本方法的分类结果(右)

(a) Ground truth (left) and classification results of this method (right)



(b) 真实标签(左)和本方法的分类结果(右)

(b) Ground truth (left) and classification results of this method (right)

图10 WHU-Urban3D数据集分类结果

Fig. 10 Classification results of WHU-Urban3D dataset

行了正确分割,其中建筑和地面的分割效果较好,F1分数超过了90%。证明了本文所提出的整合边缘卷积与全局-局部自注意力方法的有效性。但由于有些地物的稀疏性和相似性,对部分地物的区分存在一定欠缺,例如③处电力线分割效果一般,④处将地面错分为植被。

3. 2. 3. 3 定量分析

为了量化本方法的分类效果,给出了分类结果的混淆矩阵并计算其精确度、召回率和F1分数。ISPRS-3D数据集分类结果如表2所示。可以发现,本文所提出的整合边缘卷积与全局-局部自注意力方法的分类效果较好。例如在不透水表面和

屋顶类别的F1分数超过了90%。在低矮植被和树类别的F1分数达到80%左右。而对电力线,车辆,外墙和灌木类别的F1分数在40%~70%之间。由于有些类别在整体数据中所占比例较小,以及与其他类别的界限不明显,导致模型难以有效地进行分类,例如在栅栏类别上的分类效果一般,F1分数低于40%。从混淆矩阵可以发现,分类效果一般的外墙主要受屋顶和树的影响,而树对电力线和灌木类别的区分均造成了较大影响,具体表现为较多的电力线地物点和灌木地物点被错分为树。

WHU-Urban3D数据集分类结果如表3所示。可以发现,本文对大多数类别获得了较好的

表 2 ISPRS-3D数据集分类结果的混淆矩阵及评价指标

Tab. 2 Confusion matrix and evaluation index of the classification results of ISPRS-3D (%)

Class	p_l	l_v	i_s	car	f_h	r_f	f_e	s_b	t_e
p_l	33.5	0.3	0.0	0.0	0.0	31.8	11.7	0.2	22.6
l_v	0.0	73.0	9.8	0.2	0.4	1.3	0.4	6.4	8.5
i_s	0.0	3.6	94.9	0.3	0.0	0.3	0.2	0.7	0.1
car	0.0	11.2	14.0	58.2	3.7	2.4	0.6	8.6	1.3
f_h	0.0	12.1	1.6	0.6	24.5	1.2	0.5	48.1	11.4
r_f	0.1	1.4	0.0	0.0	0.0	93.7	1.7	1.4	1.8
f_e	0.2	6.2	0.5	1.6	0.2	23.1	50.1	5.8	12.3
s_b	0.0	15.6	0.4	0.5	1.3	3.2	1.9	53.2	24.0
t_e	0.0	4.7	0.0	0.1	0.2	1.7	1.3	8.7	83.4
Precision	69.6	85.1	88.5	77.6	50.5	92.9	70.9	43.6	77.4
Recall	33.5	73.0	94.9	58.2	24.5	93.7	50.1	53.2	83.4
F1 score	45.9	78.6	91.6	66.5	33.0	93.3	58.7	47.9	80.3

表 3 WHU-Urban3D数据集分类结果的混淆矩阵及评价指标

Tab. 3 Confusion matrix and evaluation index of the classification results of WHU-Urban3D (%)

Class	b_d	t_e	v_g	l_v	l_g	e_c	g_d	o_t	v_h
b_d	93.2	2.7	3.4	0.3	0.1	0.2	0.0	0.2	0.0
t_e	5.2	79.0	9.3	4.7	0.4	0.3	0.0	0.3	0.9
v_g	3.3	2.6	88.5	1.3	0.0	0.0	0.5	2.3	1.6
l_v	4.2	5.2	8.9	74.1	0.0	0.0	0.2	5.1	2.3
l_g	10.6	17.8	8.1	0.2	51.7	7.5	0.0	3.9	0.3
e_c	8.1	4.2	1.3	0.0	6.9	77.0	0.0	2.6	0.0
g_d	0.4	0.1	2.7	1.6	0.2	0.1	93.3	1.3	0.4
o_t	9.8	5.3	7.4	6.0	6.7	8.2	1.9	51.4	3.4
v_h	2.5	1.7	4.3	12.7	1.3	0.0	11.1	7.4	59.0
Precision	92.8	92.2	86.1	74.3	81.3	77.8	92.5	60.8	78.1
Recall	93.2	79.0	88.5	74.1	51.7	77.0	93.3	51.4	59.0
F1 score	93.0	85.1	87.3	74.2	63.2	77.4	92.9	55.7	67.2

分类效果,例如在建筑和地面类别的 F1 分数超过了 90%。在树和植被类别的 F1 分数超过了 80%。而在低矮植被和电力线的 F1 分数超过了 70%。验证了本文所提出的整合边缘卷积与全局-局部自注意力方法的有效性。同时在数据集中,存在一些类别的样本数量相对较少且分布稀疏,例如路灯等。这些类别的地物在形状和轮廓上与其他类别相似,导致模型难以对这些类别有效地进行分类,例如对路灯,其他和车辆的分类效果一般,F1 分数在 50%~70% 之间。从混淆矩阵可以发现,分类效果一般的车辆主要受低矮植被和地面的影响,而建筑对路灯和电力线类别的区分均造成了较大影响,具体表现为较多的路灯地物点和电力线地物点被错分为建筑。

3.2.4 对比实验

为验证整合边缘卷积与全局-局部自注意力方法的可行性和有效性,本文设计三组方案进行对比分析与分类结果评价:(1)消融实验:与仅包含局部自注意力的 Point Transformer,添加了全局自注意力的 Point Transformer,以及添加了边

缘卷积的 Point Transformer 进行对比,以验证本方法的改进效果;(2)不同方法对比实验:将本方法的分类结果与其他深度学习网络模型进行对比,以验证本方法的先进性。(3)ISPRS 3D 标记竞赛对比实验:将本方法的分类结果与 ISPRS 数据集官网上其他针对城市大场景下的激光点云数据研究成果进行对比,以验证本方法对这类数据的适用性。

3.2.4.1 消融实验

ISPRS-3D 数据集分类结果如表 4 所示,可以发现相较于原始 Point Transformer,引入全局自注意力机制后,增强了模型对整个点云场景的宏观理解,而边缘卷积的加入使得模型在微观层面上对点云的边界和邻近关系具有更加细致的感知能力。两者融合有效地提高了模型对复杂地物的表征能力,点云分类的整体分类精度由 81.1% 上升到 82.5%,提高了 1.4%;除此之外,所有地物类别分类精度均有一定程度的提高,其中少数类别电力线、汽车和栅栏的 F1 分数提高较为明显,分别提高了 4.5%,7.7% 和 8.6%,这进一步验证了本文提出方法的有效性。

表 4 ISPRS-3D 数据集消融实验
Tab. 4 Ablation experiments on ISPRS-3D dataset (%)

Class	F1 score									OA
	p_l	l_v	i_s	car	f_h	r_f	f_e	s_b	t_e	
PT	41.4	76.9	89.1	58.8	24.4	91.6	57.9	41.6	77.9	81.1
PT-GA	61.3	79.0	90.0	71.6	25.9	92.6	61.8	36.9	76.9	81.6
PT-EC	51.7	80.5	91.2	58.0	19.9	90.8	60.6	40.1	74.9	81.4
Ours(PT-GA-EC)	45.9	78.6	91.6	66.5	33.0	93.3	58.7	47.9	80.3	82.5

WHU-Urban3D 数据集分类结果如表 5 所示。结果表明,本文所提出的整合边缘卷积与全局-局部自注意力网络结构比仅使用局部空间自

注意力机制的 Point Transformer 获得了更好的分类性能,点云分类的整体分类精度由 84.0% 上升到 87.4%,提高了 3.4%;同时在对所有地物

表 5 WHU-Urban3D 数据集消融实验
Tab. 5 Ablation experiments on the WHU-Urban3D dataset (%)

Method	F1 score									OA
	b_d	t_e	v_g	l_v	l_g	e_c	g_d	o_t	v_h	
PT	91.4	79.9	84.7	71.6	52.3	73.8	89.9	44.0	58.2	84.0
PT-GA	92.8	82.5	86.9	75.6	58.3	76.7	93.1	52.3	60.2	85.9
PT-EC	92.2	82.9	86.2	72.9	54.6	75.7	91.3	48.4	68.0	85.5
Ours(PT-GA-E)	93.0	85.1	87.3	74.2	63.2	77.4	92.9	55.7	67.2	87.4

类别进行分类时,分类精度均有一定提高。其中少数类别树木、路灯、车辆以及其他类的分类效果较好,F1 分数分别提高了 5.2%,10.9%,9%和 11.7%,这进一步验证了本文提出方法的有效性。

3.2.4.2 不同方法对比实验

本文将提出的模型与其他深度学习模型进行了比较,具体包括 PointNet^[19],PointNet++^[20],DGCNN^[30],RandLANet^[35],PointMLP^[36]和 Stratified Transformer^[37]。ISPRS-3D 数据集分类结果如表 6 中所示,显示了这些模型每个类别的 F1 分数,OA 和平均 F1 分数。与 PointNet 网络相比,本方法的总体精度和 F1 分数均值分别提升了 16.8% 和 24.3%。因为 PointNet 作为一种直接处理原始点云数据的深度学习网络,虽然在全局特征提取方面表现出色,但在捕捉局部信息方面存在一些局限性,因此分类效果一般。而本文提出的边缘卷积模块通过建立点与点之间的拓扑关系,有效地提取点云数据中的局部特征,从而得到了更好的分类效果。与 PointNet++ 网络相比,本方法的总体精度和 F1 分数均值分别提升了 1.5% 和 0.8%。与 DGCNN 网络相比,本方法的总体精度和 F1 分数均值分别提升了 1.4% 和 1.3%。虽然这两种网络都加强了局部特征的提取,分类效果较好,但本文提出的全局自注意力机制模块能够使得信息在所有点之间流动,有效地提取了全局上下文信息,因此取得了更好的分类效果。与 RandLANet 网络相比,本方法的总体精度和 F1 分数均值分别提升了 1.0% 和 0.6%。RandLANet 网络是一种用于大规模点云数据的语义分割网络,它

通过结合随机点采样和局部特征聚合模块,保留了点云数据复杂的几何细节。尽管随机采样提高了效率,但可能会丢弃关键特征从而影响分类效果。与 PointMLP 网络相比,本方法的总体精度和 F1 分数均值分别提升了 1.1% 和 0.8%。PointMLP 提出了一种纯粹基于残差多层感知器(MLP)的方法来处理点云数据。这种方法的核心思想是重新思考网络设计和局部几何结构在点云分析中的作用,通过简化网络结构来提高效率和性能。虽然简化的网络设计提高了效率,但可能会忽略一些对点云分析重要的局部几何细节。与 Stratified Transformer 网络相比,本方法的总体精度和 F1 分数均值分别提升了 0.9% 和 0.5%。Stratified Transformer 网络提出了一种分层策略,对于每个查询点,密集地采样附近的点并稀疏地采样远距离的点作为其关键,从而在低计算成本下有效扩大感受野并捕捉长距离上下文来提高分类性能。而本文提出的全局-局部自注意力机制模块既能捕捉点云中长距离的依赖关系,还能关注于每个点周围的细节和邻近区域,从而获得了更好的分类效果。此外,本文所提出的网络在不透水表面,屋顶,外墙,灌木和树类别中实现了最高的分类精度。这验证了本文提出方法的有效性。

为了直观地展现实验结果,本文选择两个近期具有代表性且分类精度较好的模型 PointMLP 和 Stratified Transformer 进行可视化对比分析。如图 11 所示,在 ISPRS-3D 数据集上,从整体效果而言,三种方法都能够实现机载点云的有效分类,准确度均能达到 80% 以上,但从细节而言,如

表 6 不同方法在 ISPRS-3D 数据集上的分类结果
Tab. 6 Classification results of different methods on ISPRS-3D dataset (%)

Method	F1 score									OA	Avg. F1
	p_l	l_v	i_s	car	f_h	r_f	f_e	s_b	t_e		
PointNet	52.6	70.0	83.3	11.2	7.5	74.8	7.8	24.6	45.4	65.7	41.9
PointNet++	54.7	80.5	90.3	72.4	35.6	89.5	50.2	42.3	73.1	81.0	65.4
DGCNN	45.2	80.6	89.6	77.1	29.6	90.0	55.3	41.2	74.8	81.1	64.9
RandLANet	48.7	78.2	87.3	68.1	37.2	91.1	57.8	45.3	77.4	81.5	65.6
PointMLP	54.8	78.3	90.8	69.8	32.2	89.4	56.1	41.5	76.3	81.4	65.4
Stratified transformer	51.8	80.9	90.5	66.7	35.1	90.2	56.4	44.2	76.0	81.6	65.7
Ours	45.9	78.6	91.6	66.5	33.0	93.3	58.7	47.9	80.3	82.5	66.2

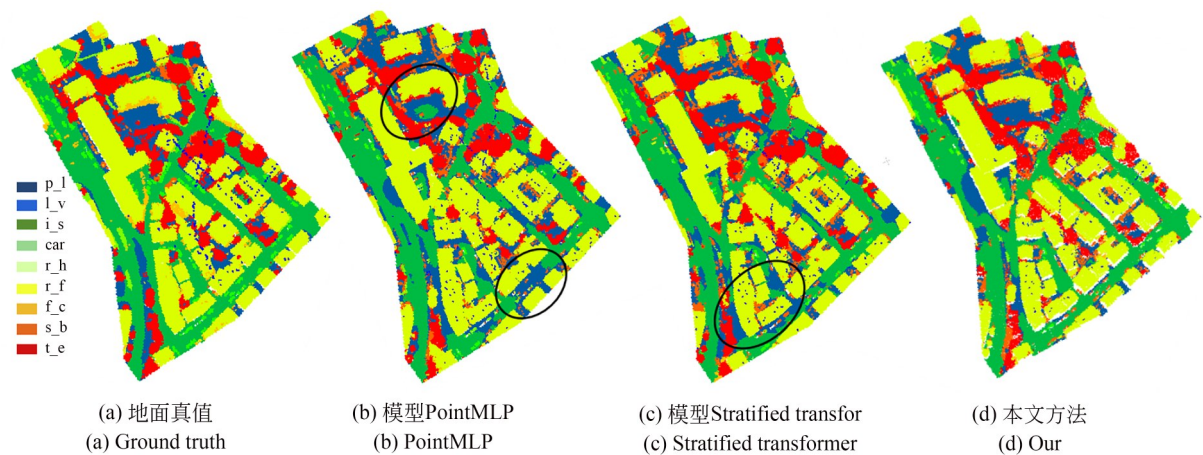


图 11 ISPRS-3D 数据集不同模型下分类结果图

Fig. 11 Classification results of different models in ISPRS-3D dataset

图椭圆框所示,PointMLP 对低矮植被的分类效果较差,而与 Stratified Transformer 相比,在一些地物混杂且拓扑特征相似的区域,例如建筑物,不透水表面,低矮植被和灌木混合的区域,本文方法分类效果更好。

WHU-Urban3D 数据集分类结果如表 7 所

示,和 PointNet 相比,本方法对所有地物的 F1 评分均有所提升;在建筑,树,低矮植被,地面和其他类的 F1 评分实现了最高的分类精度。此外,本文所提出方法在总体精度 OA 和平均 F1 分数方面均优于其他网络,这验证了本文提出方法的有效性。

表 7 不同方法在 WHU-Urban3D 数据集上的分类结果

Tab. 7 Classification results of different methods on the WHU-Urban3D dataset unit (%)

Method	F1 score									OA	Avg. F1
	b_d	t_e	v_g	l_v	l_g	e_c	g_d	o_t	v_h		
PointNet	81.6	54.3	73.8	43.5	51.7	67.8	88.5	31.3	46.2	70.5	59.8
PointNet++	90.1	75.3	83.5	71.3	62.7	82.5	91.3	46.4	71.1	85.3	74.9
DGCNN	91.3	77.8	87.4	70.8	60.1	74.5	91.7	47.2	77.4	85.6	75.3
RandLANet	91.7	79.3	88.1	71.8	63.5	79.7	91.5	53.7	66.9	86.2	76.2
PointMLP	92.4	80.8	86.2	73.1	62.7	75.3	90.3	54.5	68.2	86.0	75.9
Stratified transformer	92.6	81.2	87.5	72.6	64.1	76.5	91.2	53.8	68.5	86.5	76.4
Ours	93.0	85.1	87.3	74.2	63.2	77.4	92.9	55.7	67.2	87.4	77.3

如图 12 所示,在 WHU-Urban3D 数据集上,PointMLP,Stratified Transformer 和本方法都能对大部分数据点进行正确预测。但从细节而言,如图椭圆框所示,PointMLP 和 Stratified Transformer 网络均存在将树木错分为植被和将地面错分为植被,而本文所提方法分类效果较好。证明了本文所提出的整合边缘卷积与全局-局部自注意力方法的有效性。

3. 2. 4. 3 ISPRS-3D 标记竞赛对比实验

为了进一步评估本文所提出的 E-GLAT 模型,将其与 ISPRS 3D 标记竞赛中的其他 5 个模型进行比较,具体包括 UM^[31],WHUY2,WHUY3^[32],BIJ_W^[33],RIT_1^[34]。表 8 中显示了这些模型每个类别的 F1 分数,OA 和平均 F1 分数。可以发现所提出的方法在 OA 和平均 F1 得分方面优于所有其他方法。具体而言,OA 和平均 F1 分数分别比其他方法的最佳值高 1.7% 和

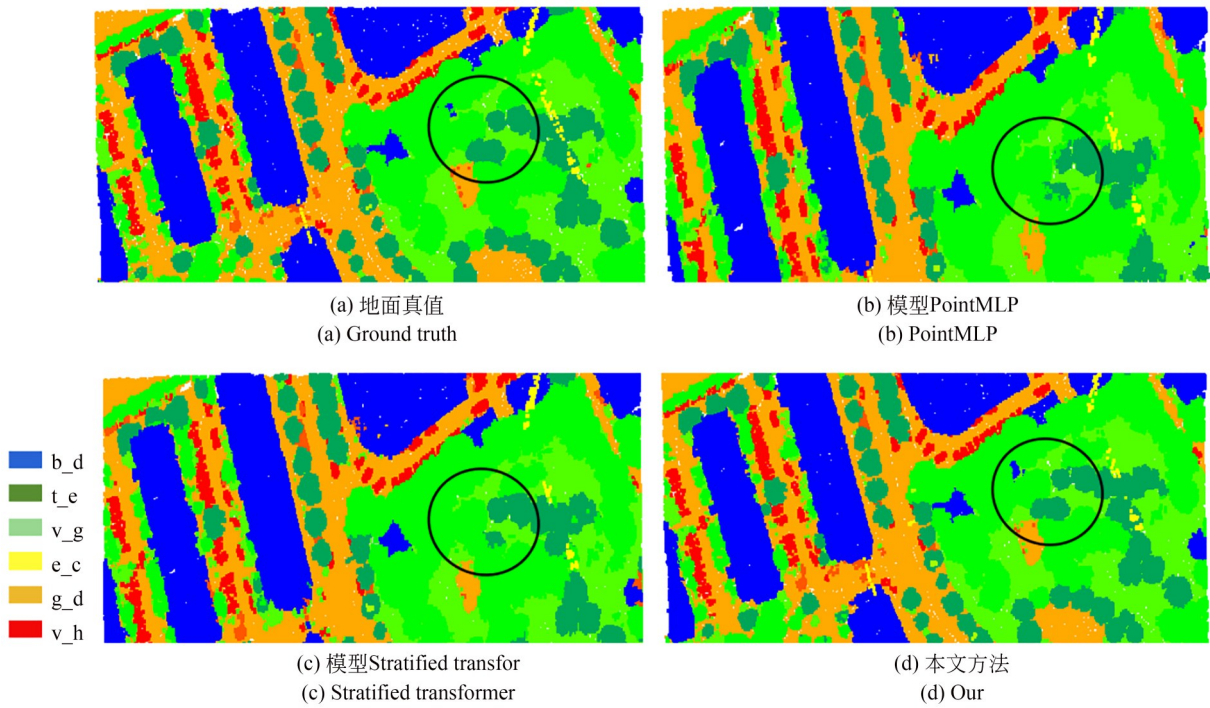


图 12 WHU-Urban3D 数据集不同模型下分类结果图
Fig. 12 Classification results of different models in WHU-Urban3D dataset

表 8 本方法与现有方法在 ISPRS-3D 数据集上的分类结果

Tab. 8 Classification results of the proposed method and the existing methods on the ISPRS 3D (%)

Method	F1 score									OA	Avg. F1
	p_l	l_v	i_s	car	f_h	r_f	f_e	s_b	t_e		
UM	46.1	79.0	89.1	47.7	5.2	92.2	52.7	40.9	77.9	80.8	59.0
WhuY2	31.9	80.0	88.9	40.8	24.5	93.1	49.4	41.1	77.3	81.0	58.6
WhuY3	37.1	81.4	90.1	63.4	23.9	93.4	47.5	39.9	78.0	82.3	61.6
BIJ_W	13.8	78.5	90.5	56.4	36.3	92.2	53.2	43.3	78.4	81.5	60.3
RIT_l	37.5	77.9	91.5	73.4	18.0	94.0	49.3	45.9	82.5	81.6	63.3
Ours	45.9	78.6	91.6	66.5	33.0	93.3	58.7	47.9	80.3	82.5	66.2

7.6%。证明了本文所提出的整合边缘卷积与全局-局部自注意力方法的有效性。

4 结 论

针对从规模较大且地物复杂的机载点云数据中准确分类各种地物目标的难题,提出了一种整合边缘卷积的全局-局部自注意力机制神经网络(E-GLAT)。首先,本文设计边缘卷积模块,利用邻域点的距离信息构建空间图,帮助模型更好地理解物体的轮廓和形状,从而增强点云数据

的局部特征表达能力。其次,设计全局-局部自注意力模块,全局自注意力模块倾向于整个输入序列的信息,而局部自注意力模块则更侧重于局部区域的细节。两种机制结合增强了对长距离依赖关系和局部结构的捕捉,有助于提高模型对复杂地物关系的分类能力。最后,为了充分评估本文所提出的 E-GLAT 模型的有效性,在 ISPRS-3D 数据集和 WHU 数据集上进行广泛的实验,并从定量和定性两方面与众多方法进行比较。实验结果表明,本文所提出的模型优于当前大多数流行的点云分类模型,能够获得较好分类

效果。此外,本文引入了边缘卷积和自注意力机制模块,当训练数据量有限时,由于模型的复杂性增加,可能会出现过拟合现象。同时边缘卷积和自注意力机制模块也依赖于超参数的优化和

调整,找到最佳的超参数设置需要一定的实验和调整。在未来的研究中,我们将尝试深入研究小样本、弱监督下的机载点云分类,努力在更严谨的理论框架下进一步提高分类效果。

参考文献:

- [1] LUO B, YANG J, SONG S, *et al.* Target classification of similar spatial characteristics in complex urban areas by using multispectral LiDAR[J]. *Remote Sensing*, 2022, 14(1): 238.
- [2] ZHANG P X, WANG Q, GAO R J, *et al.* Ground point cloud segmentation based on local threshold adaptive method[J]. *Opt. Precision Eng.*, 2023, 31(17): 2564-2572. (in Chinese)
- [3] 徐健,刘慧,沈跃,等. 基于改进 PointNet++ 模型的苗圃树木点云分类与分割[J]. *中国激光*, 2024, 51(8): 193-203.
XU J, LIU H, SHEN Y, *et al.* Point clouds classification and segmentation for nursery trees based on improved PointNet++ Model[J]. *Chinese Journal of Lasers*, 2024, 51(8): 193-203. (in Chinese)
- [4] KACPER J, WIOLETA P. Point cloud does matter. Selected issues of using airborne LiDAR elevation data in geomorphometric studies of rugged sandstone terrain under forest-Case study from Central Europe[J]. *Geomorphology*, 2022, 412: 108316.
- [5] ZHOU K, LINDENBERGH R, GORTE B, *et al.* LiDAR-guided dense matching for detecting changes and updating of buildings in Airborne LiDAR data[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 162: 200-213.
- [6] DAI Z, HE R, WANG H T, *et al.* Adaptive single tree extraction method based on fusion of airborne lidar and vegetation index[J]. *Opt. Precision Eng.*, 2023, 31(22): 3331-3344. (in Chinese)
- [7] 龚建宗. 机载激光雷达在山区输电线路设计中的应用[J]. *测绘与空间地理信息*, 2024, 47(1): 170-172.
GONG J Z. Application of airborne LiDAR data in transmission line design in mountainous area[J]. *Geomatics & Spatial Information Technology*, 2024, 47(1): 170-172. (in Chinese)
- [8] HORVÁTH E, POZNA C, UNGER M. Real-time LIDAR-based urban road and sidewalk detection for autonomous vehicles[J]. *Sensors*, 2021, 22(1): 194.
- [9] 李峙含,花海洋,张浩. 基于对比学习的高光谱和 LiDAR 地物分类方法[J]. *激光与光电子学进展*, 2023, 60(22): 3788/LOP230540.
LI Z H, HUA H Y, ZHANG H. Classification based on hyperspectral image and LiDAR data with contrastive learning[J]. *Laser & Optoelectronics Progress*, 2023, 60(22): 3788/LOP230540. (in Chinese)
- [10] 涂静敏,沈阳,李婕,等. 体素与点混合增长的机载点云屋顶平面分割[J]. *中国激光*, 2024, 51(22): 2210002.
TU J M, SHEN Y, LI J, *et al.* On-board point cloud roof plane segmentation for voxel and point hybrid growth[J]. *Chinese Journal of Lasers*, 2024, 51(22): 2210002. (in Chinese)
- [11] 李婕,李青清,李礼,等. 基于深度学习的机载点云屋顶平面提取算法[J]. *激光技术*, 2024, 48(5): 628.
LI J, LI Q Q, LI L, *et al.* An airborne point cloud roof plane extraction algorithm based on deep learning[J]. *Laser Technology*, 2024, 48(5): 628.
- [12] ZHANG J X, LIN X G, NING X. SVM-based classification of segmented airborne LiDAR point clouds in urban areas[J]. *Remote Sens*, 2013, 5: 3749-3775.
- [13] HE M Y, CHENG Y L, LIAO X J, *et al.* Building extraction algorithm by fusing spectral and geometrical features[J]. *Laser & Optoelectronics Progress*, 2018, 55(4): 042803.
- [14] MUNOZ D, VANDAPEL N, HEBERT M. On-board contextual classification of 3-D point clouds with learned high-order markov random fields[C]. 2009 *IEEE International Conference on Robotics and Automation*. Kobe, Japan. IEEE, 2009: 2009-2016.
- [15] NIEMEYER J, ROTTENSTEINER F, SOERGEL U. Contextual classification of lidar data and building object detection in urban areas[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2014, 87: 152-165.

- [16] 耿笑笑. LiDAR点云多尺度注意力语义分割方法研究[D]. 武汉: 武汉大学, 2022.
- GENG X X. *Research on LiDAR Point Cloud Multi-scale and Attentive Semantic Segmentation* [D]. Wuhan: Wuhan University, 2022. (in Chinese)
- [17] RABBANI T, VAN DEN HEUVEL F, VOSSELMANN G. Segmentation of point clouds using smoothness constraint[J]. *International archives of photogrammetry, remote sensing and spatial information sciences*, 2006, 36(5): 248-253.
- [18] CHEN X Z, MA H M, WAN J, et al. Multi-view 3D object detection network for autonomous driving [C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI, USA. IEEE, 2017: 6526-6534.
- [19] CHARLES R Q, HAO S, MO K C, et al. PointNet: deep learning on point sets for 3d classification and segmentation[C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI, USA. IEEE, 2017: 77-85.
- [20] QI C, YI L, SU H, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space[J]. *Advances in Neural Information Processing Systems*, 2017: 5105-5114.
- [21] LI D W, SHI G L, WU Y H, et al. Multi-scale neighborhood feature extraction and aggregation for point cloud segmentation[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2021, 31(6): 2175-2191.
- [22] ZHAO H S, JIANG L, JIA J Y, et al. Point transformer [C]. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*. Montreal, QC, Canada. IEEE, 2021: 16239-16248.
- [23] 徐田野, 丁海勇. 基于融合图卷积的深度学习点云分类方法[J]. *激光与光电子学进展*, 2022, 59(2): 0228005.
- XU T Y, DING H Y. Deep learning point cloud classification method based on fusion graph convolution [J]. *Laser & Optoelectronics Progress*, 2022, 59(2): 0228005. (in Chinese)
- [24] 白斌. 机载LiDAR在大比例尺地形图测绘中的应用分析[J]. *测绘与空间地理信息*, 2024, 47(4): 186-188, 192.
- BAI B. Application analysis of airborne LiDAR in large-scale topographic mapping[J]. *Geomatics & Spatial Information Technology*, 2024, 47(4): 186-188, 192. (in Chinese)
- [25] ZAHEER M, GURUGANESH G, DUBEY K A, et al. Big bird: Transformers for longer sequences[J]. *Advances in neural information processing systems*, 2020, 33: 17283-17297.
- [26] LI J H, WEINMANN M, SUN X, et al. VD-LAB: a view-decoupled network with local-global aggregation bridge for airborne laser scanning point cloud classification[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2022, 186: 19-33.
- [27] NIEMEYER J, ROTTENSTEINER F, SOERGEL U. Contextual classification of lidar data and building object detection in urban areas[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2014, 87: 152-165.
- [28] HAN X F, LIU C, ZHOU Y Z, et al. WHU-Urban3D: an urban scene LiDAR point cloud dataset for semantic instance segmentation[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2024, 209: 500-513.
- [29] 刘盛, 曹益烽, 黄文豪, 等. 融合稀疏注意力和实例增强的雷达点云分割[J]. *中国图象图形学报*, 2023, 28(2): 483-494.
- LIU S, CAO Y F, HUANG W H, et al. LiDAR point cloud semantic segmentation combined with sparse attention and instance enhancement [J]. *Journal of Image and Graphics*, 2023, 28(2): 483-494. (in Chinese)
- [30] PHAN A V, LE NGUYEN M, NGUYEN Y, et al. DGCNN: a convolutional neural network over large-scale labeled graphs [J]. *Neural Networks*, 2018, 108: 533-543.
- [31] HORVAT D, ŽALIK B, MONGUS D. Context-dependent detection of non-linearly distributed points for vegetation classification in airborne LiDAR[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, 116: 1-14.
- [32] YANG Z S, JIANG W S, XU B, et al. A convolutional neural network-based 3D semantic labeling method for ALS point clouds[J]. *Remote Sensing*, 2017, 9(9): 936.
- [33] 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*. Salt Lake City, UT, USA. IEEE, 2018: 2589-2597.
- [34] YOUSEFHUSSEN M, KELBE D J, IENTILUCCI E J, *et al.* A multi-scale fully convolutional network for semantic labeling of 3D point clouds[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2018, 143: 191-204.
- [35] HU Q Y, YANG B, XIE L H, *et al.* RandLANet: efficient semantic segmentation of large-scale point clouds[C]. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (CVPR). Seattle, WA, USA. IEEE, 2020: 11105-11114.
- [36] MA X, QIN C, YOU H X, *et al.* Rethinking Network Design and Local Geometry in Point Cloud: A Simple Residual Mlp Framework [EB/OL]. 2022; arXiv: 2202.07123. <http://arxiv.org/abs/2202.07123>
- [37] LAI X, LIU J H, JIANG L, *et al.* Stratified transformer for 3D point cloud segmentation[C]. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (CVPR). New Orleans, LA, USA. IEEE, 2022: 8490-8499.

作者简介:



涂静敏(1991—),女,福建南平人,博士,讲师,2016年于中国地质大学(武汉)获得硕士学位,2021年于武汉大学遥感信息工程学院获得博士学位,主要研究方向为:计算机视觉、三维点云数据处理研究。E-mail: jingmin.tu@hbut.edu.cn

通讯作者:



李婕(1984—),女,湖北宜昌人,博士,副教授,2013年于武汉大学电子信息学院获得博士学位,主要研究方向为:计算机视觉、智慧农业。E-mail: jielonline@hbut.edu.cn