2025 年 第 25 卷 第 6 期
2025, 25(6): 02461 -08

科 学 技 术 与 工 程
Science Technology and Engineering

ISSN 1671—1815
CN 11—4688/T

DOI:10. 12404/j. issn. 1671-1815. 2403205
引用格式:李昊宇, 廖维张. 基于改进 RandLA-Net 的建筑构件点云提取方法[J]. 科学技术与工程, 2025, 25(6): 2461-2468.
Li Haoyu, Liao Weizhang. Building component point cloud extraction method based on improved RandLA-Net[J]. Science Technology and Engineering, 2025, 25(6): 2461-2468.

建筑科学

# 基于改进 RandLA-Net 的建筑构件点云提取方法

李昊宇[1,2], 廖维张[1,2*]
(1. 北京建筑大学工程结构与新材料北京市高等学校工程研究中心, 北京 100044;
2. 北京建筑大学北京未来城市设计高精尖创新中心, 北京 100044)

**摘 要** 点云数据在建筑逆向建模、三维重建乃至施工进程等方面均有巨大优势。采集建筑结构点云时通常包含海量点云,并且如梁、柱等构件的点云数据至关重要。现有的三维点云语义分割方法对大规模点云进行处理时存在局部特征提取不充分、识别精度有待提升等问题,提出了一种改进 RandLA-Net 深度学习网络的建筑关键构件的大规模点云语义分割方法,通过在局部空间编码部分增加坐标注意力模块提高分割结果的鲁棒性;构建了通道注意力模块增强模型的特征判断能力;引入了焦点损失函数训练网络,有效解决了建筑点云场景内类别不平衡的问题,实现了对建筑结构点云数据的快速处理和对建筑关键构件的有效提取。最后通过实验进行性能对比分析。试验结果表明,改进模型对大规模点云进行语义分割相较于传统 RandLA-Net 模型在整体准确率和局部构件提取准确率上均有提升,证实了本文方法具有更强的性能和应用价值。

**关键词** 三维点云; 建筑工程点云; 深度学习; 点云语义分割; 注意力机制
**中图法分类号** TU17; **文献标志码** A

# Building Component Point Cloud Extraction Method Based on Improved RandLA-Net

LI Hao-yu[1,2], LIAO Wei-zhang[1,2*]
(1. Beijing Higher Education Engineering Research Center for Engineering Structures and New Materials, Beijing University of Architecture, Beijing 100044, China; 2. Beijing High Precision Innovation Center for Future Urban Design, Beijing Architecture University, Beijing 100044, China)

[**Abstract**] The significant advantages of point cloud data are presented in domains such as architectural reverse modeling, 3D reconstruction, and construction progress monitoring. Vast amounts of data are typically involved in the collection of point clouds for architectural structures, with the point clouds of components like beams and columns being particularly crucial. The challenges faced by current semantic segmentation methods for 3D point clouds when processing large-scale data include insufficient extraction of local features and suboptimal recognition accuracy. An enhanced approach for the semantic segmentation of large-scale point clouds of key architectural components using the RandLA-Net deep learning network was proposed. In this regard, the robustness of segmentation results was improved by incorporating a coordinate attention module in the local spatial encoding section. Furthermore, an extended channel attention module has been developed to strengthen the model's capability in feature discernment, and a focal loss function has been introduced to effectively train the network, while addressing class imbalance issues within architectural point cloud scenes. Consequently, the efficient processing of architectural structure point cloud data and the extraction of key components are enabled. The performance comparisons and analyses conducted through experiments demonstrate that the original RandLA-Net model is outperformed by our model in terms of overall accuracy and component extraction precision in semantic segmentation of large-scale point clouds, thereby confirming the enhanced performance and practical value of the proposed method.

[**Keywords**] 3D point cloud; construction engineering point cloud; deep learning; point cloud semantic segmentation; attention mechanism

**收稿日期**: 2024-04-29; **修订日期**: 2024-12-16
**基金项目**: 国家自然科学基金重点项目(52130809);北京建筑大学校级教研重点项目(Y2106)
**第一作者**: 李昊宇(1999—),男,汉族,北京人,硕士研究生。研究方向:三维点云获取与处理。E-mail:18513830765@163. com。
* **通信作者**: 廖维张(1978—),男,汉族,浙江苍南人,博士,教授。研究方向:数字化检测与监测、智能建造关键技术。E-mail:liaoweizhang@bucea. edu. cn。

随着建筑行业的发展,建筑结构逐渐复杂并且规模也更加庞大,仅靠图纸难以表达出建筑物内部的空间结构关系。使用传统测绘手段对建筑测量时常会有效率低、精确度低等问题。相比传统技术,点云技术在建筑逆向建模、施工进程等方面均展现出了巨大优势。例如,在老旧建筑的拆除或改造时,通常会存在建筑设计图纸及竣工图纸存在严重丢失或与建筑现状不匹配的情况;同时,工程竣工时的地形现状、地上与地下各种建筑物等也是项目施工情况与验收的重要依据[1]。另外,在利用图纸传递工程信息时,各方提出的修改意见不能及时反映到图纸当中,从而极易造成工程信息丢失[2]。随着建筑信息化和智能化技术的不断发展,激光扫描、倾斜摄影等技术,拓宽了数据采集途径,使其可满足各种场景下的数据需求[2]。使用三维激光扫描技术可以快速且准确地将建筑结构的几何信息进行记录,点云不仅可以记录扫描对象的位置坐标信息,还可记录其颜色等属性,如今已逐渐成为建筑领域数字化发展的支撑技术[3-7]。然而,由于建筑工程点云通常包含海量点云,并且存在场景中类别不平衡的问题,导致现有的方法无法做到对建筑工程点云数据的关键信息如梁、柱等进行准确地提取,对于逆向建模等后续操作存在不利影响。

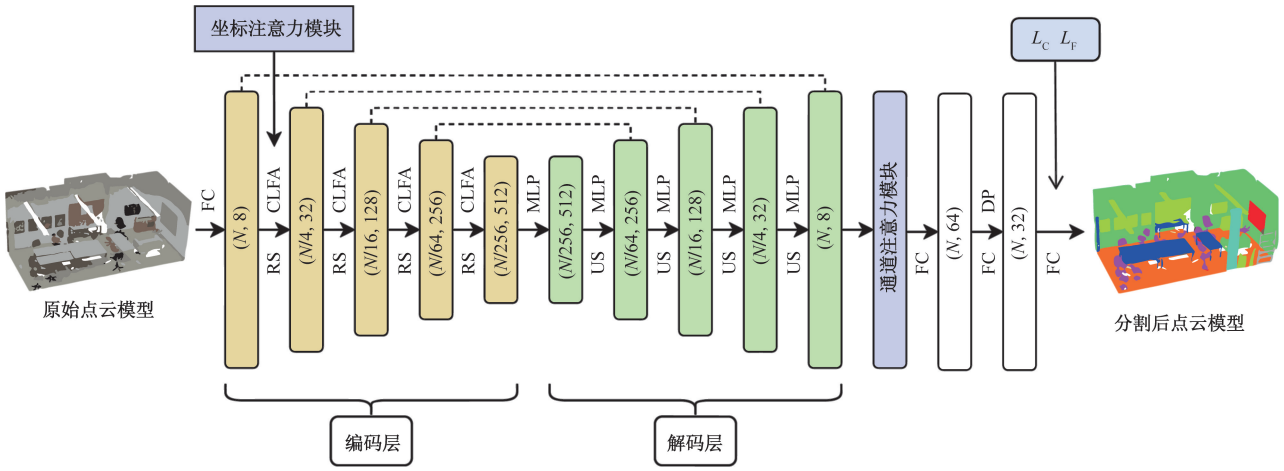
随着深度学习理论的逐步完善,三维激光点云数据在物体自动识别、提取及分割方面取得了显著进展[7]。语义分割在点云处理中扮演着核心角色,为建筑工程点云数据的提取开辟了创新性的技术路径。近年来,中外学者在三维场景的语义理解方面取得了显著的研究成果。早期在点云数据中使用深度学习技术的研究者们采取了一种方法,即将点云映射到多个二维视图上,这样做可以把无规律的点云数据转换成有规律的图像数据。通过这种方法展现三维形态,并运用已经成熟的二维卷积技术来从多视图图像中提取信息并进行分类[8]。Su 等[9]将卷积等常规处理技术应用于投影的二维图像,以实现点云数据的语义分割。而随着卷积神经网络在图像语义分割方面出现突出表现,同时体素和图像在数据结构上的类似性,研究人员开始把点云数据转换成体积化的(体素化)数据,以此构建三维神经网络模型。Maturana 等[10]提出了基于体素数据的点云语义分割模型,与传统的二维图像处理技术不同,该方法是在多维环境中进行操作,通过使用三维卷积核提取体素数据特征,来解决数据无序的问题,并且能够维持数据的多维特性,从而有效地进行点云的语义分割[8]。然而,上述两类方法均未能充分利用点云特性,导致部分信息流失。研究人员从三维数据源出发,直接从点云中提取特征信息[11],与间接方法相比,能够降低信息损失,并在点云分割任务中展现出优异表现。Qi 等[12]为该类方法的先驱,提出了 PointNet 模型,该网络能够直接对点云数据进行分类与分割任务处理。

目前,现有方法在建筑室内场景中难以对建筑构件进行有效提取。鉴于 RandLA-Net[13]在室内大规模点云数据集 S3DIS(stanford large-scale 3D indoor spaces dataset)[14]取得了优异的效果,且在处理大规模点云数据方面更具优势,因此本文研究使用 RandLA-Net 作为主干网络,采集建筑结构所形成的点云数据的特点,得到可对如墙、柱、梁等重要建筑构件进行精确提取的点云语义分割模型。并使用三维点云数据集 S3DIS 进行实验,对实验结果进行评估并可视化。

## 1 研究方法

现有方法在处理三维点云时,缺乏对于现实世界三维信息的特定特性足够的重视。特别是在建筑领域,从建筑工程中提取的点云数据常展现出显著的类别失衡现象。一些主流类别在场景中占比过高,导致深度学习模型在训练时偏向于优化这些多数类别的识别效果。然而,对于那些只占比数量少的类别,由于训练数据的限制,模型难以掌握足够的特征。例如,墙、地板、梁、柱几乎是所有建筑室内场景中不可或缺的元素,提取这些建筑构件对于逆向建模也至关重要。但是在实际场景中,墙的数量却远超梁的数量,在三维点云数据集 S3DIS 中,“墙”的点数大约为 $5\times10^6$,这一数量显著高于“梁”标签所对应的点数。此类失衡可能导致模型在测试阶段产生偏差,倾向于将物体识别为占据较多点数的类别。另外,针对大场景空间区域,为提升网络解释力,必须融合大量上下文信息进入局部特征,以此形成综合的上下文联系。当处理的数据量增加并且这些数据展现出更广泛的特征差异时,模型能够从给定的空间范围内提取到的相关信息量也会随之提升。本文研究基于 RandLA-Net,在主干网络上增加了通道注意力模块,以缓解分割时出现类间模糊的问题;在局部特征聚合部分增加了坐标注意力(coordinate attention, CA)模块,以更好地聚合特征,形成融合了坐标注意力的局部特征聚合模块(coordinate local feature aggregation,CLFA);引入了焦点损失函数 $L_F$,提升关键建筑构件分割精度,解决建筑场景内类别不平衡的问题。网络的整体结构如图 1 所示。

本文模型首先将点云划分为不同的局部区域,




FC 为全连接层;RS 为随机采样;MLP 为多层感知机;US 为上采样;DP 为 Dropout 结构

图 1　整体网络架构

Fig. 1　Overall network architecture

然后对每个局部区域内的点进行特征聚合,网络按照($N\rightarrow N/4\rightarrow N/16\rightarrow N/64\rightarrow N/256$)的顺序对点云进行稀释(其中 $N$ 为点云数量),同时,每个点的特征维度按(8→32→128→256→512)的顺序增加以保留更多信息。之后将主干模型输出的特征图输入通道注意力模块,得到扩展特征图,随后通过融合扩展特征图与主干网络产生的特征图,进而生成一个融合特征图,并引入焦点损失 $L_F$ 联合交叉熵损失 $L_C$ 来训练网络,得到最终的语义分割结果。最后,通过全连接层对点云场景进行分割,得到最终的语义分割结果。

### 1.1　通道注意力模块

建筑结构内部点云常会出现类间模糊的情况,例如,墙壁和白板这两种物体具有不同的语义标签,但因其具有相似的外形结构,导致区分类别时存在困难。故需要提高模型的判断特征能力,更好地区分这些特征。通道注意力通过建模通道之间的相互依赖关系,可以突出信息量大的特征通道,抑制冗余的特征通道,从而提升网络的特征表达能力。现有方法大多采用降维的方式来实现通道注意力。Wang 等[15]设计了 ECA(efficient channel attention)模型,通过实验证明,降低通道维度会对最终输出产生不利影响,导致损失一些有价值的信息[16]。鉴于此,本文研究提出了一种无降维的通道注意力模块。具体来说,在主干网络提取出尺寸为 $N\times C$ 的特征图 A 后,本文研究直接在原始通道维度上进行全局平均池化操作。采用一维卷积对经过全局平均池化的通道进行处理,其中卷积核大小通过函数自适应调整[13]。内核大小 $k$ 代表局部跨通道交互的范围。对于给定的输入特征图的通道数量 $C$,可以将这个 $C$ 值映射到一个适当的卷积核大小 $k$。卷积核大小 $k$ 可以自适应地确定为

$$k=\psi(C)=\left|\frac{\log_2 C}{\gamma}+\frac{b}{\gamma}\right|_{\text{odd}} \tag{1}$$

式(1)中:$|x|_{\text{odd}}$ 为最接近 $x$ 的奇数;$b=1$;$\gamma=2$;$C$ 为通道数量。

随后,为了揭示相邻通道间的互动关系,本文研究引入交叉信道算子,该步骤通过核大小为 $k$ 的 $1\times1$ 卷积层来实现。接着,跨信道交互可表示为

$$H_j=\sigma\left(\boldsymbol{W}\sum_{i=1}^{M}\frac{1}{N}A_i\right) \tag{2}$$

式(2)中:$\boldsymbol{W}$ 为 $k\times M$ 的参数矩阵;$A$ 为主干网络输出的特征图,下标 $i$ 代表通道 $i$;$N$ 为特征图尺寸;$M$ 为通道维数;$k=5$;$\sigma$ 为一个 Sigmoid 函数。

通过 ECA 设计方法降低通道维度引起的信息丢失,并提升邻近通道间的互动性,如图 2 所示。可以对各个点的特征进行更新,并充分利用点之间的空间结构信息,从而克服了传统方法难以处理复杂场景的局限性,缓解分割时出现边缘混淆的问题,实现更为精准的物体边界分割。

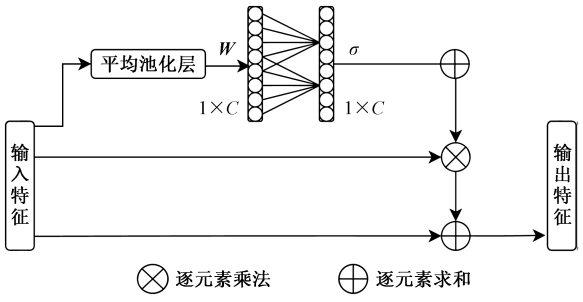


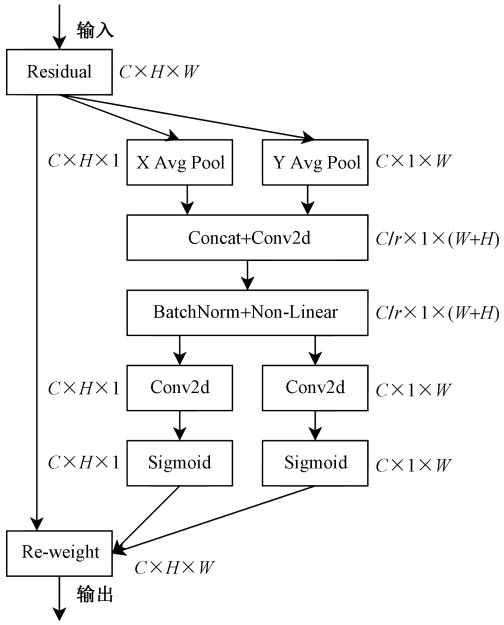

图 2　通道注意力模块

Fig. 2　Expanded channel attention module

### 1.2　坐标注意力模块

在对建筑结构进行扫描时,点云数据往往包含

众多干扰元素,从而影响关键部件的精确度。为了提升关键部分在复杂环境中的特征表现,并增强其注意力程度,本研究在编码阶段引入坐标注意力(coordinate attention, CA)[17-18]模块,将广泛的上下文信息融合至局部特征之中,从而构建出丰富的上下文关联。通过将位置信息融入通道注意力中,实现对构件特征的精确识别,从而提升算法性能[14]。CA 结构如图 3 所示。




r 为缩放系数;X Avg Pool 代表一维水平全局池化;
Y Avg Pool 代表一维垂直全局池化

图 3 坐标注意力模块结构

Fig. 3 Structure of the coordinate attention Module

在特征输入 CA 模块之后,分别沿垂直和水平方向对下式进行全局池化处理。通过式(3),将空间信息实现全局编码,以便注意力块准确地利用特征图的位置信息捕获远程交互。

$$z_c = \frac{1}{HW}\sum_{i=1}^{H}\sum_{j=1}^{W} x_c(i,j) \tag{3}$$

式(3)中:$z_c$ 为第 $c$ 个通道的输出;$H$ 为池化核的高度;$W$ 为池化核的宽度;$i$、$j$ 分别为通道 $i$ 和通道 $j$。

针对特定输入 $X$,利用坐标注意力机制可以在池化核心的两个空间区域$(H, 1)$和$(1, W)$内,分别沿水平和垂直方向对各通道进行平均池化[14]。因此,第 $c$ 个通道在高度 $h$ 处的输出公式为

$$z_c^h(h) = \frac{1}{W}\sum_{0\leqslant i\leqslant W} x_c(h,i) \tag{4}$$

沿垂直方向进行分解后,宽度为 $w$ 的第 $c$ 个通道输出为

$$z_c^w(w) = \frac{1}{H}\sum_{0\leqslant i\leqslant H} x_c(j,w) \tag{5}$$

式中:$z_c^h(h)$ 和 $z_c^w(w)$ 分别表示高度为 $h$ 和宽度为 $w$ 在第 $c$ 个通道上的具有方向感知的输出特征。

针对两个沿不同方向聚合特征的池化过程,其会形成两个含有空间信息的特征图。通过这种设计,注意力机制可以在多个维度上理解远距离的数据关联性,而在其他维度上则维持了对位置的精确识别。这种能力使得模型能够更有效地识别并追踪其需要关注的对象。紧接着,坐标注意力机制通过 Contact 操作将两个特征图连接起来,之后进行 1×1 卷积,并进行归一化操作[17],过程为

$$\boldsymbol{f} = \delta(F_1[z^h, z^w]) \tag{6}$$

式(6)中:$[z^h, z^w]$ 为空间维度上的串联操作;$\delta$ 为非线性激活函数;$\boldsymbol{f}$ 为两个不同方向(水平和垂直)上编码空间信息的特征图。

接着,将函数 $\boldsymbol{f}$ 按照空间维度分解为两个独立的张量 $\boldsymbol{f}^w$ 和 $\boldsymbol{f}^h$。在后续的处理过程中,两个尺寸均为 1×1 的卷积操作 $F_w$ 和 $F_h$ 被应用于分别将 $\boldsymbol{f}^h$ 和 $\boldsymbol{f}^w$ 转换为与输入 $\boldsymbol{X}$ 具有相同通道数的张量 $\boldsymbol{g}^h$ 和 $\boldsymbol{g}^w$[17]。该过程可以分为式(7)和式(8)两个部分进行总结。

$$\boldsymbol{g}^h = \sigma[F_h(\boldsymbol{f}^h)] \tag{7}$$

$$\boldsymbol{g}^w = \sigma[F_w(\boldsymbol{f}^w)] \tag{8}$$

式中:$\sigma$ 为 Sigmoid 函数。

在此后,输出 $\boldsymbol{g}^h$ 和 $\boldsymbol{g}^w$ 被扩展并分别用作注意力权重。最终,经过 CA 模块处理后的图像特征 $y$ 可以表示为

$$\boldsymbol{y}_c(i,j) = \boldsymbol{x}_c(i,j)\boldsymbol{g}_c^h(i)\boldsymbol{g}_c^w(j) \tag{9}$$

### 1.3 焦点损失函数

在 RandLA-Net 网络架构中损失函数为交叉熵损失函数,该函数在分类问题上有着出色的表现。但是在建筑室内点云数据中,常会出现类别分布不平衡的问题,待提取的目标如梁、柱等点云数量远少于如墙、地板等其他类别数据,使用交叉熵损失函数,由于其更倾向于优化多数类别,会导致少数类别难以得到充分训练。在处理数据集中存在的显著不平衡问题时,焦点损失函数(focal loss)[19-20]方法展现了其显著的效用。通过减少那些易于分类的样本在损失计算中的权重,有效地缓解了不均衡数据对模型性能的负面影响。当样本数量在不同类别间存在差异时,焦点损失函数展现出较优的处理能力。鉴于此,本文研究在引入了焦点损失函数联合训练模型,旨在减少样本数量不平衡而造成的少数类别样本分割效果较差的现象,交叉熵损失函数[13]为

$$L = \begin{cases} -\ln y', & y = 1 \\ -\ln(1-y'), & y = 0 \end{cases} \tag{10}$$

焦点损失函数[19]为

$$L = \begin{cases} -\alpha(1-y')^{\gamma}\ln y', & y=1 \\ -(1-\alpha)y'^{\gamma}\ln_e(1-y'), & y=0 \end{cases} \quad (11)$$

式(11)中:$y'$ 为模型的预测类别概率;$y$ 为真实样本;$\alpha$、$\gamma$ 为调节因子,参数 $\alpha$ 的值为 0.3,参数 $\gamma$ 的值则为 1.5。

当模型对于正确类别的预测概率上升时,$(1-y')\gamma$ 的数值会相应下降,代表正确预测的实例在总损失中所占的比重变小;而预测错误的实例则在总损失中占据了更大的比重。通过调整参数 α 来使得不同类别的样本在模型中的比例更为均衡,并且通过参数 $\gamma$ 来减少简单样本对模型训练的影响,使模型更加集中于那些难以分类的实例。应用焦点损失函数来进行这种平衡,能够有效提高模型在处理室内三维点云数据时的分割效果。在点云分割任务中引入焦点损失函数,对于处理建筑工程数据具有显著优化效果,并能缓解语义分割模型的类别间不确定性问题。

## 2　实验及结果分析

### 2.1　实验环境和评价指标

本次实验环境为:实验显卡 NVIDIA GeForce 4080 Super,显存为 16 GB,内存为 32 GB 操作系统是 Ubuntu 18.04,使用的深度学习框架为 tensorflow 2.6,CUDA 11.4,基于 Python3.6 构建网络模型,实验使用 Adam 优化器,初始学习率为 0.001,临近点数量 $K$ 为 16,训练批次为 8,衰减率为 0.9。网络的训练和推理都在 NVIDIA GeForce 4080 Super 上进行。本文研究采纳了 3 个关键的性能评估指标来衡量图像分割任务的效果,分别是整体精度(OA)、平均精度($m_{\mathrm{Acc}}$)以及平均交并比($m_{\mathrm{IoU}}$)[7]。

### 2.2　实验数据集介绍

本文研究的目的是研究如何提升点云语义分割模型整体精度;以及解决建筑重要构件如柱、梁等识别精确度低的问题。为了评估模型在处理建筑结构点云数据方面的效能,本文研究选取了 S3DIS 数据集进行实验。数据集 S3DIS 是由斯坦福大学发布的一个大规模室内三维空间数据集。该数据集包含了多个室内环境的点云数据和语义标签,用于室内场景理解与三维语义分割的研究。构成 6 个室内空间区域(总共 272 个房间),涵盖 13 种语义标签。每一点均以六维向量形式呈现,包括空间位置($xyz$)与颜色属性(RGB)。

### 2.3　实验结果

在 S3DIS 数据集的实验中使用 6 倍交叉验证进行实验结果的评估。语义分割结果根据平均类别交并比、平均类别准确度和总体准确度与近些年来经典的深度学习的方法进行对比,如表 1 所示,最优的结果已加粗显示。试验结果表明,使用本研究提出的技术与 RandLA-Net 相比,整体识别精度有 1.6% 的增长,平均识别精度增加了 3.7%,同时平均交并比也有 4.8% 的提升,这证明了本文改进措施的有效性。将本文方法与其他多种网络模型进行对比分析后,结果显示本文提出的点云语义分割技术在整体准确性和平均交并比方面均超过了其他模型,虽然总体准确度与 PointTransformer 相差 0.6%,但也已经基本达到目前模型的先进水平。

**表 1　S3DIS 数据集语义分割结果**

**Table 1　Semantic Segmentation Results on the S3DIS Dataset**

| 方法 | OA/% | $m_{\mathrm{Acc}}$/% | $m_{\mathrm{IoU}}$/% |
|---|---|---|---|
| PointNet[12] | 78.6 | 66.2 | 47.6 |
| SPG[21] | 86.4 | 73.0 | 62.1 |
| DGCNN[22] | 84.5 | — | 55.5 |
| KPConv[23] | 81.0 | — | 70.6 |
| PointTransformer[24] | 90.2 | 81.9 | 73.5 |
| RandLA-Net[13] | 88.0 | 81.0 | 70.0 |
| 本文方法 | 89.6 | 84.7 | 74.8 |

对比 S3DIS 数据集的实例分割平均精度,对于室内场景中高频出现的重要构件类别,例如,天花板、地板墙壁等类别虽然略逊色于 PointTransformer,相比原模型,本文方法均有所度提升。本文方法在对于低频出现但相对重要的构件如梁、柱等类别实现了最佳性能,本文方法相比原模型提升幅度较大在“梁”的识别精度上提升了 15.8%;在“柱”的识别精度上提升了 10.6%,均达到当前先进水平。此外例如窗户、桌子、沙发以及杂物等,本文方法也实现了比原模型更好的分割结果。为了能够直观分析点云语义分割的结果,本文研究将点云语义分割结果可视化,如图 4 所示,从上至下依次为原始点云、带有真实分割标签的分割结果、由 RandLA-Net 模型分割后的结果以及本文方法分割后的结果,选取数据集中 3 个场景进行分析。通过点云语义分割可视化结果可以看出,RandLA-Net 在柱、梁部分的识别效果不精确,观察图 4 中的 3 组结果可以看出,RandLA-Net 将原始点云中的柱均识别为墙。如图 4(a)标注所示,RandLA-Net 将原始点云中的部分梁识别为墙;并且,原模型对于场景内同时存在黑板和墙两个相似类别时会出现边界混淆的现象,无法准确分割墙和黑板。而本文方法能够有效解决场景内类别识别错误、边界混淆等问题,如图 4(b)和图 4(c)所示。在梁、柱、墙部分的识别效果已经非常接近真实标签,对于柱和梁的识别提取精确度远超原模型,进一步说明了本文方法的有效性。

表 2 S3DIS 数据集实例分割平均精度

Table 2 Average precision of instance segmentation on the S3DIS dataset

| 方法 | 平均精度/% | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 屋顶 | 地板 | 墙 | 梁 | 柱 | 窗户 | 门 | 桌子 | 椅子 | 沙发 | 书柜 | 板 | 其他 |
| PointNet[12] | 88.0 | 88.7 | 69.3 | 42.4 | 23.3 | 47.5 | 51.6 | 54.1 | 42.0 | 9.6 | 38.4 | 29.4 | 35.2 |
| SPG[21] | 89.9 | 95.1 | 76.6 | 62.8 | 47.1 | 55.3 | 68.4 | 73.5 | 69.2 | 63.2 | 45.9 | 8.7 | 52.9 |
| DGCNN[22] | 93.2 | 95.9 | 72.8 | 54.6 | 32.2 | 56.2 | 50.7 | 62.9 | 63.4 | 22.7 | 38.2 | 32.5 | 46.8 |
| KPConv[23] | 93.6 | 92.5 | 83.1 | 63.9 | 54.3 | 66.1 | 76.6 | 57.8 | 64.0 | 69.3 | 74.9 | 65.7 | 64.8 |
| PointTransformer[24] | 94.3 | 97.5 | 81.7 | 55.6 | 58.2 | 66.2 | 78.2 | 77.6 | 74.1 | 69.5 | 71.2 | 66.5 | 64.8 |
| RandLA-Net[13] | 93.1 | 96.1 | 80.6 | 62.4 | 48.0 | 64.4 | 69.4 | 69.4 | 76.4 | 60.0 | 64.2 | 65.9 | 60.1 |
| 本文方法 | 94.0 | 96.8 | 82.1 | 78.2 | 58.6 | 65.5 | 70.9 | 77.7 | 74.8 | 71.5 | 68.3 | 66.7 | 65.0 |

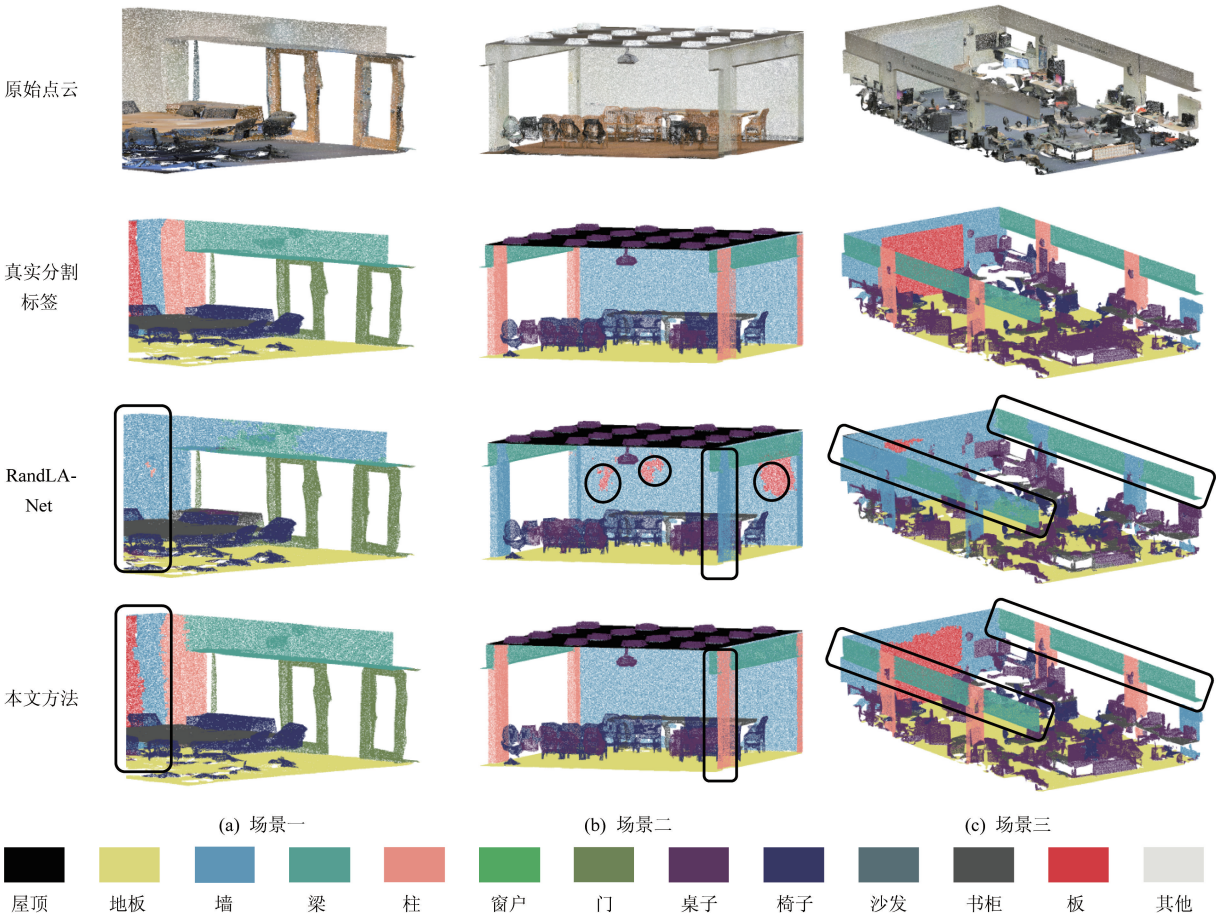



图 4 S3DIS 数据集上的语义分割可视化效果图

Fig. 4 Visualized semantic segmentation results on the S3DIS datase

## 2.4 消融实验

为了验证通道注意力模块(表 5 中记为 ECA)、CA 模块和焦点损失函数的有效性,对模型进行消融实验,消融试验分为 5 组,分别是不做任何改变的原模型、单独加入了 3 个模块以及同时加入了 3 个模块的本文方法。所有消融实验都在 S3DIS 数据集上选取相同实验设置,采用六倍交叉验证法进行对比研究,消融试验结果如表 5 所示。

通过消融试验可得,本文研究中加入的三部分对原网络在评价指标上均有一定程度上的提升,其中通过引入通道注意力模块提高深度学习模型在判断特征时的区分能力,对分割效果的提升较为明显。本文提出的改进方法够有效实现点云的场景

表 3 不同模块在 S3DIS 数据集上的消融试验

Table 3 Ablation experiments of different modules on the S3DIS dataset

| 方法 | OA/% | $m_{Acc}$/% | $m_{IoU}$/% |
|---|---|---|---|
| RandLA-Net | 87.9 | 81.0 | 69.7 |
| RandLA-Net + FL | 88.3 | 81.9 | 70.8 |
| RandLA-Net + CA | 87.8 | 81.8 | 71.5 |
| RandLA-Net + ECA | 89.0 | 83.6 | 72.2 |
| 本文方法 | 89.6 | 84.7 | 74.8 |

分割,提取点云的深层几何特征得到丰富的语义信息,对建筑室内点云场景实现较好的分割。

## 3　结论

针对建筑构件点云难以精确提取的问题,通过分析采集建筑结构所形成的点云数据的特点,基于 RandLA-Net 模型提出了一种优化的建筑构件点云提取模型。通过在 S3DIS 数据集上进行试验,得出以下主要结论。

(1)加入了无降维的通道注意力模块,加深了网络对特征的学习,提高了深度学习模型在判断特征时的区分能力,实现了更为精确的分割结果。

(2)在通过引入坐标注意力模块,将广泛的上下文信息融合至局部特征之中,从而构建出丰富的上下文关联,增强了模型在环境复杂时提取关键数据方面的性能。

(3)通过加入焦点损失函数,有效解决了现有模型在处理室内点云数据时,因类别不平衡而造成的边缘混淆和分类错误的问题。

实验结果表明,本文方法相对于原模型在多项评价指标上均有提升,并且在建筑构件识别提取精度上有明显提升。在 S3DIS 数据集试验结果表明,本文方法在整体识别精度增长了 1.6%,平均识别精度增长了 3.7%,平均交并比提升了 4.8%;在提取屋顶、地板、墙、梁、柱时,平均分割精度分别提升了 1.1%、0.7%、1.5%、15.8%、10.6%。在建筑室内场景中,本文方法能够对建筑构件实现更为精确的分割提取,并且在整体分割精度上有所提升。该方法可为建筑逆向建模提供更精确的信息,为工程进展的把控和工期评估提供重要参考与依据。

### 参　考　文　献


[1] 徐敬海, 卜兰, 杜东升, 等. 建筑物 BIM 与实景三维模型融合方法研究[J]. 建筑结构学报, 2021, 42(10): 215-222.
Xu Jinghai, Bu Lan, Du Dongsheng, et al. Research on the fusion method of building information modeling (BIM) and reality-based 3D models[J]. Journal of Building Structures, 2021, 42(10): 215-222.

[2] 刘界鹏, 崔娜, 周绪红, 等. 基于三维激光扫描的房屋尺寸质量智能化检测方法[J]. 建筑科学与工程学报, 2022, 39(4): 71-80, 3-4.
Liu Jiepeng, Cui Na, Zhou Xuhong, et al. Intelligent detection method for house dimension quality based on 3D laser scanning[J]. Journal of Architecture, Civil Engineering, and Environment, 2022, 39(4): 71-80, 3-4.

[3] 伍根, 熊小龙. 基于三维激光扫描测绘技术的 BIM 逆向建筑建模方法研究[J]. 城市勘测, 2023(5): 33-37.
Wu Gen, Xiong Xiaolong. Research on the reverse building modeling method of bim based on 3D laser scanning surveying technology [J]. Urban Surveying, 2023(5): 33-37.

[4] 林楷奇, 郑俊浩, 陆新征. 数字孪生技术在土木工程中的应用: 综述与展望[J]. 哈尔滨工业大学学报, 2024, 56(1): 1-16.
Lin Kaiqi, Zheng Junhao, Lu Xinzheng. Application of digital twin technology in civil engineering: review and outlook[J]. Journal of Harbin Institute of Technology, 2024, 56(1): 1-16.

[5] 陈冠华, 李博, 朱铮涛. 基于改进 PointNet 的空调散热器 V 形槽 3D 点云分割算法[J]. 科学技术与工程, 2024, 24(5): 1963-1971.
Chen Guanhua, Li Bo, Zhu Zhengtao, et al. A 3D point cloud segmentation algorithm for air conditioner radiator V-shaped grooves based on an improved PointNet[J]. Science Technology and Engineering, 2024, 24(5): 1963-1971.

[6] 张帆, 孙楚津, 覃思中, 等. 基于 BIM 和深度学习点云分割的施工检查方法模拟研究[J]. 工程力学, 2024, 41(2): 194-201.
Zhang Fan, Sun Chujin, Qin Sizhong, et al. Simulation study on construction inspection method based on BIM and deep learning point cloud segmentation[J]. Engineering Mechanics, 2024, 41(2): 194-201.

[7] 王艺娴, 胡雨凡, 孔庆群, 等. 三维点云语义分割: 现状与挑战[J]. 工程科学学报, 2023, 45(10): 1653-1665.
Wang Yixian, Hu Yufan, Kong Qingqun, et al. Semantic segmentation of 3D point clouds: current status and challenges[J]. Journal of Engineering Science, 2023, 45(10): 1653-1665.

[8] 双丰, 黄兴文, 李勇, 等. 基于深度学习的大规模点云语义分割方法综述[J]. 测绘科学, 2023, 48(2): 195-209.
Shuang Feng, Huang Xingwen, Li Yong, et al. A review of large-scale point cloud semantic segmentation methods based on deep learning[J]. Science of Surveying and Mapping, 2023, 48(2): 195-209.

[9] 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. Santiago: IEEE, 2015: 945-953.

[10] Maturana D, Scherer S. Voxnet: a 3D convolutional neural network for real-time object recognition[C]// IEEE/RSJ International Conference on Intelligent Robots and Systems. Hamburg: IEEE, 2015: 922-928.

[11] 卢健, 贾旭瑞, 周健, 等. 基于深度学习的三维点云分割综述[J]. 控制与决策, 2023, 38(3): 595-611.
Lu Jian, Jia Xurui, Zhou Jian, et al. A review of three-dimensional point cloud segmentation based on deeplearning[J]. Control and Decision, 2023, 38(3): 595-611.

[12] Qi C R, Su H, Mo K, et al. Pointnet: deep learning on point sets for 3d classification and segmentation[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017: 652-660.

[13] Hu Q, Yang B. Randla-net: efficient semanticsegmentation of large-scale point clouds[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 11108-11117.

[14] Armeni I, Sener O, Zamir A, et al. 3D semantic parsing of large-scale indoor spaces[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Washington: IEEE,

2016: 1534-1543.

[15] Wang Q, Wu B, Zhu P, et al. Eca-net: efficient channel attention for deep convolutional neural networks[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2020: 11534-11542.

[16] 杨青松, 郝如江, 范亚飞, 等. 一种改进轻量化神经网络的齿轮箱故障诊断方法[J]. 科学技术与工程, 2024, 24(7): 2699-2705.
Yang Qingsong, Hao Rujiang, Fan Yafei, et al. A gearbox fault diagnosis method based on improved lightweight neural network[J]. Science Technology and Engineering, 2024, 24(7): 2699-2705.

[17] Hou Q, Zhou D, Feng J. Coordinate attention for efficient mobile network design[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 13713-13722.

[18] 李启明, 阙祖航. 基于改进 YOLOv5 的 X 射线图像危险品检测[J]. 科学技术与工程, 2023, 23(4): 1598-1606.
Li Qiming, Que Zuhang. Detection of hazardous items in X-ray images based on an improved YOLOv5[J]. Science Technology and Engineering, 2023, 23(4): 1598-1606.

[19] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.

[20] 刘晋川, 黎向锋, 刘安旭, 等. 改进 RetinaNet 的无人机小目标检测[J]. 科学技术与工程, 2023, 23(1): 274-282.
Liu Jinchuan, Li Xiangfeng, Liu Anxu, et al. Small object detection for drones using an improved RetinaNet[J]. Science Technology and Engineering, 2023, 23(1): 274-282.

[21] Xu Q, Zhou Y, Wang W, et al. SPG: unsupervised domain adaptation for 3D object detection *via* semantic point generation[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. New York: IEEE, 2021: 15446-15456.

[22] 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): 1-12.

[23] 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. Seoul: IEEE, 2019: 6410-6419.

[24] Zhao H, Jiang L, Jia J, et al. Point transformer[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. New York: IEEE, 2021: 16259-16268.