计算机系统应用 ISSN 1003-3254, CODEN CSAOBN
[doi: 10.15888/j.cnki.csa.010093] [CSTR: 32024.14.csa.010093]
©中国科学院软件研究所版权所有.
E-mail: csa@iscas.ac.cn
http://www.c-s-a.org.cn
Tel: +86-10-62661041

基于 RandLA-CGNet 的大规模室内点云语义分割①

王建超[1], 王浩雨[1], 苏　鹤[2], 王震洲[1], 张　丹[1]

[1](河北科技大学 信息科学与工程学院, 石家庄 050018)
[2](河北工业大学 电工产品可靠性技术省部共建协同创新中心, 天津 300401)
通信作者: 苏　鹤, E-mail: suhe2280@163.com

摘　要: 随着数字孪生虚拟现实技术的应用越来越广泛, 针对大规模室内建筑点云语义分割中整体精度有限、小物体识别精度低及边界分割模糊等问题, 提出一种大规模室内点云语义分割的方法 RandLA-CGNet. 在编码层中构建局部-全局上下文融合 (local-global context fusion, LGCF) 模块, 在保留局部邻域信息的同时融入整体上下文语义; 在解码层设计范数门控通道特征 (norm-gated channel feature, NGCF) 模块, 通过对网络特征图的通道维度进行自适应重标定, 增强有用信息、抑制冗余噪声, 增强对细节和边界的敏感性, 提高模型的精细化识别能力; 最后采用融合型损失函数 (focused cross-entropy loss, FCE loss), 在保证模型对大多数样本稳定收敛和整体精度的同时, 增加对难分样本与少数类样本的关注, 从而提升模型在边界区域和稀有类别上的分割性能. 实验结果表明, 本文提出的模型在 S3DIS 数据集上经六折交叉验证 *OA*、*mAcc* 和 *mIoU* 分别提升至 88.8%、83.4% 和 71.9%, 较基准模型分别提高 0.8%、1.4% 和 1.9%. 与主流算法相比, 较 LG-Net 分别提升 0.5%、1.0% 和 1.1%, 总体精度以及平均交并比较 FGC-AF 提升 0.2% 和 0.7%; RandLA-CGNet 在保持整体性能优势的同时, 对小物体以及边界细节分割的 *IoU* 提升了 1%–6%, 有效提升对低频类别与复杂边界的识别能力, 为点云语义分割任务中少样本类别与细节边界的精准建模提供有效解决方案.

关键词: 点云语义分割; RandLA-Net; 小物体识别; 边界分割; 低频类别

引用格式: 王建超,王浩雨,苏鹤,王震洲,张丹.基于 RandLA-CGNet 的大规模室内点云语义分割.计算机系统应用. http://www.c-s-a.org.cn/1003-3254/10093.html

RandLA-CGNet for Large-scale Indoor Point Cloud Semantic Segmentation

WANG Jian-Chao[1], WANG Hao-Yu[1], SU He[2], WANG Zhen-Zhou[1], ZHANG Dan[1]

[1](School of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang 050018, China)
[2](Provincial and Ministerial Co-construction Collaborative Innovation Center on Reliability Technology of Electrical Products, Hebei University of Technology, Tianjin 300401, China)

Abstract: As the digital twin VR technology is increasingly widely applied, a method named RandLA-CGNet for large-scale indoor point cloud semantic segmentation is proposed to solve the problems such as the limited overall accuracy, low recognition accuracy for small objects, and blurred boundary segmentation in point cloud semantic segmentation of large-scale indoor buildings. In the encoder layer, a local-global context fusion (LGCF) module is constructed, preserving local neighborhood information while incorporating global contextual semantics. In the decoder layer, a norm-gated channel feature (NGCF) module is designed, which performs the adaptive recalibration of feature maps along the channel dimension to enhance useful information and suppress redundant noise, thereby enhancing sensitivity to details and boundaries, and improving the model's refined recognition capability. Finally, focused cross-entropy loss (FCE loss), a

① 基金项目: 国家重点研发计划 (2024YFD2402205); 河北省高等学校科学技术研究项目 (QN2025371)
收稿时间: 2025-07-28; 修改时间: 2025-09-19; 采用时间: 2025-10-14; csa 在线出版时间: 2025-12-29

hybrid loss function, is adopted to ensure stable convergence for the majority of samples and maintain overall accuracy. Additionally, this function increases the focus on hard samples and minority class samples, thereby enhancing the model's segmentation performance in boundary regions and for rare classes. Experimental results show that the proposed model on the S3DIS dataset by employing 6-fold cross-validation increases *OA*, *mAcc*, and *mIoU* to 88.8%, 83.4%, and 71.9% respectively, an improvement of 0.8%, 1.4%, and 1.9% respectively compared with the baseline models. Compared to mainstream algorithms, it increases LG-Net by 0.5%, 1.0%, and 1.1% respectively, with the overall accuracy and mean intersection of union (*IoU*) 0.2% and 0.7% higher than FGC-AF respectively. While maintaining overall performance advantages, RandLA-CGNet improves the *IoU* for small objects and boundary detail segmentation by 1%–6%, significantly enhancing the recognition capability for low-frequency classes and complex boundaries. Finally, an effective solution is provided for the precise modeling of few-sample classes and detail boundaries in point cloud semantic segmentation tasks.

Key words: point cloud semantic segmentation; RandLA-Net; small object recognition; boundary segmentation; low-frequency category

1 引言

点云语义分割因其在机器人[1]、遥感[2]与自动驾驶[3]等领域的广泛应用而备受关注. 作为最贴近传感器原始数据的三维表示, 点云以其紧凑格式承载了丰富的空间分布信息, 为场景理解提供了深层次的几何线索. 尽管深度学习在图像特征提取方面已取得卓越成果, 但点云固有的无序性与不规则性使得经典的卷积神经网络难以直接应用. 早期研究[4–8]通常将点云投影为体素、深度图或多视图图像, 并借助成熟的二维卷积网络进行处理. 然而, 这类策略往往计算成本高昂, 且不可避免地损失了原始三维结构的细节信息.

为应对这一挑战, PointNet[9]首次提出直接处理原始点云的方法, 利用共享的多层感知器逐点提取特征, 并通过全局最大池化进行特征聚合, 实现了端到端的点云学习. 随后, PointNet++、DGCNN、PointCNN[10–12]等方法相继引入邻域划分、图卷积或注意力机制, 显著提升了对小规模点云局部几何结构的建模能力. 但是, 在包含百万级点的大规模场景分割任务中, 这些基于邻域或图操作的模型计算开销急剧增加, 难以高效扩展.

针对大规模点云语义分割的效率与精度问题, 近期研究提出了多种改进方案. 例如, Zhao 等人[13]提出的 LG-Net 通过学习自适应融合权重并行提取局部与全局特征, 能在多层下采样中平衡细节与语义信息; 但其并行双路结构在深层网络中反复执行平均池化和全连接操作, 导致参数量和计算量显著增加, 难以部署在资源受限设备或延迟敏感场景. Chen 等人[14]提出的 PointMM 在多空间尺度引入多头注意力池化, 显著改善了小物体的分割效果, 但其在超大规模场景下的推理效率尚未充分验证, 且多尺度采样与多头注意力堆叠带来显著的内存和计算负担. 由 Chen 等人[15]提出的 FGC-AF 将点云特征组织为图结构, 利用图卷积捕捉复杂的非欧几里得邻域关系, 并通过注意力融合自适应加权不同层特征, 其图构建还融合了几何距离、法线、曲率等高阶信息; 然而, FGC-AF 需要为每批数据重建图邻接关系并执行多次图卷积, 其时间复杂度和内存开销约为标准局部特征聚合方法的 3–4 倍. Zhao 等人[16]构建的 BADENet 结合边界感知双边卷积与增强损失, 有效提升稀疏类别的识别精度, 却引入了更复杂的网络结构. Fan 等人[17]提出的 SCF-Net 利用可分离卷积与空间通道交互大幅降低了计算量, 但在处理极细边界时仍存在平滑过渡问题. Hu 等人[18]提出的 RandLA-Net 凭借随机下采样策略和轻量级局部特征聚合模块, 在速度与内存效率方面取得了卓越平衡; 然而, 该网络缺乏对全局长距离依赖性的显式建模, 导致在处理大范围平面区域与小体量目标时难以兼顾精度.

面对上述模型在效率、精度、细节保留或结构复杂度等方面的局限, 本文提出一种基于 RandLA-Net 架构的网络 RandLA-CGNet. 该网络通过融合局部-全局上下文信息并引入通道自适应重标定机制, 在维持轻量化的前提下, 提升了大规模室内点云分割任务的整体精度、小物体与低频类别的识别能力以及边界细节

的还原效果.

2 RandLA-Net 改进

2.1 RandLA-CGNet 模型结构

RandLA-Net 是一种面向大规模三维点云的网络, 通过随机下采样与轻量级局部特征聚合 (LFA) 实现了对百万级点云的高效训练与实时推理, 但仅依赖邻域聚合往往缺乏全局上下文信息, 对大平面与小体量目标交界处的分割精度不足, 且对少样本类别和复杂边界的细节恢复能力有限. 为此, 本文构建 RandLA-CGNet 模型. 首先, 在 RandLA-Net 的编码层, 设计局部-全局上下文融合 (local-global context fusion, LGCF) 模块, 优化替代最后 3 层下采样阶段的局部特征聚合 LFA 模块, 使其在对局部邻域特征进行聚合的同时, 通过动态比率上下文注意力 (dynamic ratio context attention, DRCA) 注入整体语义信息, 从而在保证对小体量结构细节建模能力的同时, 增强了特征的全局一致性. 在解码层的每一次跳跃连接后, 均引入范数门控通道特征 (norm-gated channel feature, NGCF) 模块, 通过对通道级别的全局描述和可学习门控因子的自适应重标定, 动态放大关键特征通道并抑制冗余噪声, 提升了模型对细长结构与复杂边界的敏感性. 逐点分类头部保留了原有结构, 仅在融合不同尺度特征后采用 1×1 卷积实现多分辨率预测, 从而兼顾大范围平面和小物体目标的分割精度. 最后损失函数采用融合型损失函数 (focused cross-entropy loss, FCE loss), 在保持模型对多数易分点快速收敛和整体准确率的同时, 对难分样本与少数类样本给予更高梯度权重, 进一步提升模型在边界区域及低频类别上的分割性能. 改进后的 RandLA-Net 结构如图 1 所示.

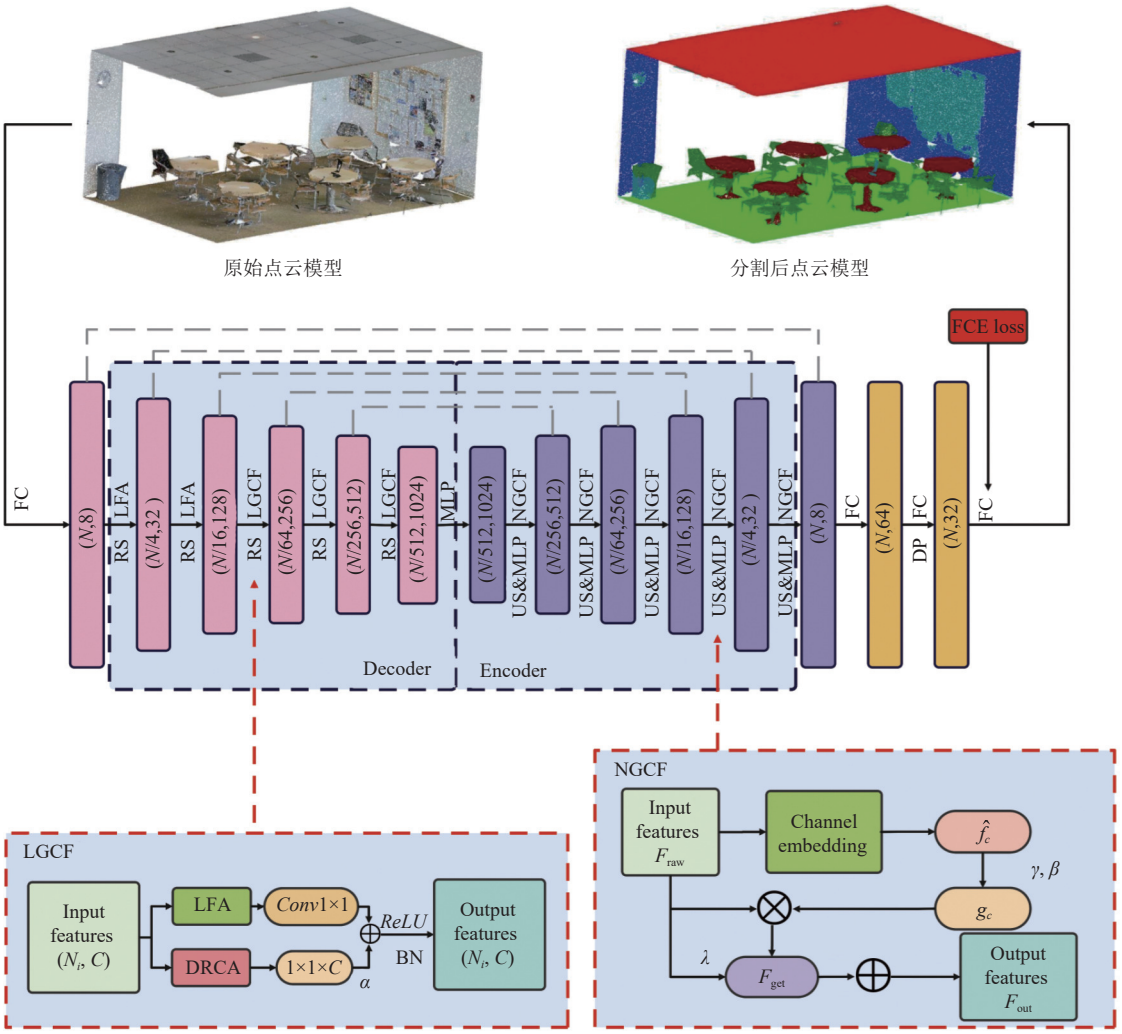

图 1 RandLA-CGNet 网络结构图

2.2 局部-全局上下文融合模块

在大规模室内建筑点云语义分割任务中, 既要捕捉梁柱、家具等小尺度结构的几何细节, 又要理解整体平面结构的全局语义, 对于模型的表达能力和泛化性能提出了极高要求. 鉴于此, 本文在 LFA[18]模块基础上, 依据 FGC-AFNet (feature graph convolution network with attentive fusion) [15]中的全局上下文注意力机制 (global context attention, GCA), 提出了一种局部-全局上下文融合模块 LGCF. 该模块将点云局部邻域特征聚合与全局语义关注并行融合, 并以可学习的通道融合系数平衡两路信息, 实现跨尺度语义与细节的深度耦合, 从而增强了网络对结构细节与少样本目标的感知能力.

由于在 RandLA-Net 等分层网络中, 浅层特征更侧重于局部几何细节, 其通道数较少且感受野有限, 此时引入复杂的全局注意力机制可能干扰对细微结构的捕捉, 且收益有限. 深层网络特征更具语义抽象性, 但原有 LFA 模块缺乏显式全局建模能力, 此时引入 LGCF 的全局分支能有效地注入跨层次的整体语义信息, 增强特征的全局一致性; 经过多次下采样后, 深层点云数据点数显著减少, 在此处引入全局注意力机制对计算和内存的开销影响最小. 因而本设计采用在编码器后 3 个下采样层中以 LGCF 替换原有 LFA.

具体而言, 设某一层下采样后输出点数为 N, 每点特征维度为 C, LGCF 模块内部首先执行局部特征聚合操作. 对于第 i 个采样点 p_i, 通过预先构建的邻域索引 $j\in N\{i\}$获取其 K 个邻居点. 随后, 对每对 (p_i, p_j) 计算相对位置向量 $\Delta p_{ij}=p_i-p_j$ 及距离 $d_{ij}=\|\Delta p_{ij}\|$, 并将这些空间编码与邻点特征向量 $f_j\in R^C$ 拼接后输入一个小型多层感知机 (MLP), 得到维度为 C'的中间表示; 在此基础上, 模块进一步通过具有 Softmax 的注意力机制生成权重 s_{ij}, 并对所有邻域特征加权求和, 最终经一次 1×1 卷积映射和残差连接, 输出局部聚合特征:

$$F_{\text{local},i}=LeakyReLU\left(Conv1\times1\left(\sum_{j\in N(i)}^{K}s_{ij}h_{ij}\right)+f_i\right)\in R^C \tag{1}$$

其中, K 为邻域大小, h_{ij} 为 MLP 输出, 且$h_{ij}\in R^{C'}$, C'通常取 $C/2$.

与此同时, LGCF 引入了 DRCA 分支, 在原始 GCA 设计中, reduction_ratio 是一个手动设定的超参数, 决定了两级全连接层的宽度, 也就是中间隐藏向量的维度为C/r, 再升回C. 这种方法虽然简单有效, 但在不同深度的特征图中, 合适的压缩比也会不同: 浅层特征通道维度较小, 信息密度更高, 过度压缩可能丢失关键局部细节; 深层特征通道维度更大, 包含更多全局语义, 压缩比更高以减少计算开销同时保留主要信息. 基于此, 本文在 GCA 的基础上设计了动态比率上下文注意力 DRCA 模块. 通过在模块内部定义一个可训练向量 $\hat{r}$, 网络可在训练中自学习每个通道的压缩程度, 这样既不会因为过度压缩导致浅层特征丢失局部细节, 也可以在减少计算开销的同时保留深层特征的主要信息.

DRCA 模块的核心改进在于引入了一个可训练向量 $\hat{r}$, 其维度与输入特征通道数 C 相同, 为每个通道赋予独立的自适应压缩参数. 该向量所有元素初始化为一个基准值r_0 (本研究设置为 16), 通过网络训练过程中的梯度反向传播进行更新. 压缩比的实际控制通过 *Sigmoid* 函数实现, 中间层维度由$C/Sigmoid(\hat{r})$动态计算. *Sigmoid* 输出值趋近于 1 表示该通道信息重要, 压缩程度低; 趋近于 0 则表示该通道冗余度高, 压缩程度高. 网络能够根据全局损失信号的梯度反向传播, 自动学习并更新向量$\hat{r}$, 从而为不同深度、不同重要性的特征通道分配合适的压缩比率. 对于细节丰富的浅层特征, 网络倾向于学习较大的$\hat{r}$, 使得$Sigmoid(\hat{r})$接近 1, 压缩程度较低, 以保留更多局部结构信息; 对于高层语义特征, 网络可学习较小的$\hat{r}$, 使$Sigmoid(\hat{r})$趋近于 0, 实现较高压缩, 在保留主要信息的同时提升计算效率. 避免了固定压缩超参数可能带来的浅层细节丢失或深层冗余计算问题, 平衡过程由全局损失函数驱动自动优化. DRCA 的结构如图 2 所示.

首先对所有 N 个局部特征按通道执行平均池化:

$$\overline{f_c}=\frac{1}{N}\sum_{i=1}^{N}F_{\text{local},i,c},\ c=1,\cdots,C \tag{2}$$

并将$\overline{f_c}\in R^C$通过两层可训练的全连接网络 (第 1 层降维至$C/Sigmoid(\hat{r})$, 第 2 层恢复至 C). 随后逐通道地与局部特征相乘, 将整体语义信号注入到每个点的特征表示中, 最后得到:

$$F_{\text{global},i,c}=F_{\text{local},i,c}\times a_c \tag{3}$$

完成局部和全局两路信息提取后, LGCF 引入一组可训练的融合系数向量$\alpha\in R^C$, 在通道维度自适应地

平衡两者贡献, 具体融合公式为:

$$F_{\text{out},i,c} = F_{\text{local},i,c} + \alpha_c F_{\text{global},i,c} \tag{4}$$

在训练之初, 融合系数α_c通常初始化为 0.5, 表示对局部与全局特征等比例关注; 随着网络迭代, α 将根据损失反馈自动收敛到最佳值, 实现对不同通道信息比重的动态调节. 融合之后, 为保持数值稳定性并引入非线性, 模块对$F_{\text{out},i}$进行批量归一化和 *LeakyReLU* 激活, 输出最终特征:

$$\widetilde{F_{\text{out},i}} = LeakyReLU\left(BatchNorm\left(F_{\text{out},i}\right)\right)$$

LGCF 结构如图 3 所示.

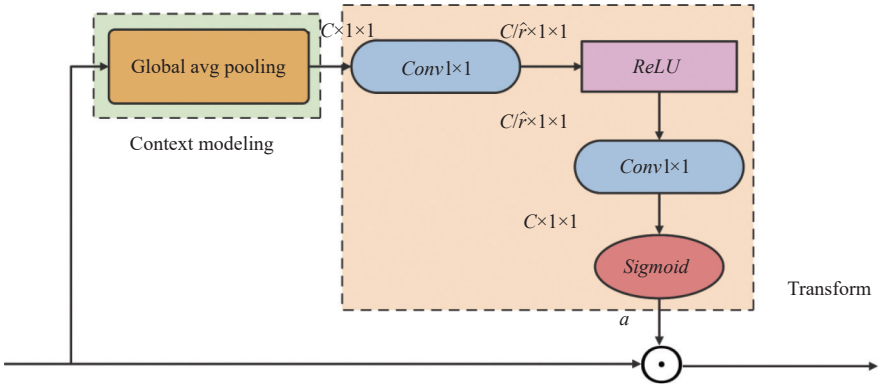

图 2 DRCA 结构图

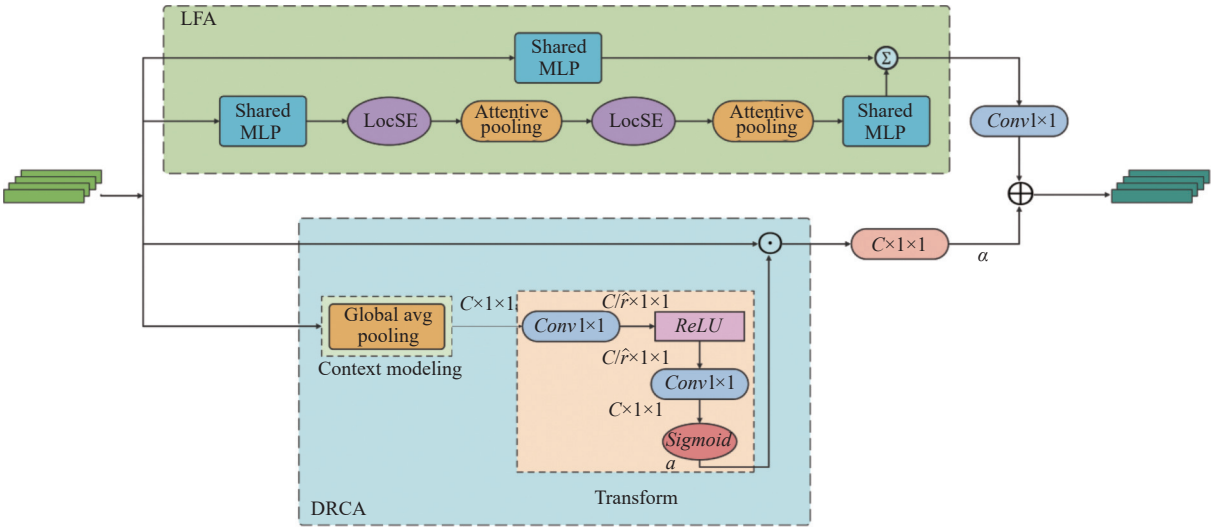

图 3 LGCF 结构图

将 LGCF 模块集成到 RandLA-Net 编码器的后 3 层下采样后, 可替换原网络仅依赖 LFA 的单一路径输出, 从而在深层特征中获得更丰富的上下文信息. 该设计的计算和参数开销极小: 全局池化复杂度为O(NC), 两层全连接复杂度为O(C^2/r), 融合系数复杂度为O(C), 总体 FLOPs 增幅不足 5%; 参数量增幅通常小于基础网络的 5%. 由于 LGCF 仅在深层小规模点集处生效, 对整体推理速度影响微乎其微.

2.3 范数门控通道特征模块

在 RandLA-Net 解码器的每一层跳跃连接和反卷积操作后, 本文引入门控通道转换 (gated channel transformation, GCT)[19]模块的核心思想, 在此基础上设计了 NGCF 模块, 旨在对解码特征的通道维度进行自适应重标定, 通过引入可学习的门控参数γ、β, 并采用残差融合的方式将其输出与原始解码特征以固定权重 0.5 混合, 以实现对通道维度上有用信息的动态放大和冗余噪声的抑制, 同时保持网络训练的稳定性. 原始 GCT 通常只采用 L_2 范数来计算每个通道的统计量, 本文设计的模块为适配不同特征分布, 支持 L_1 与 L_2 范数静态选择来计算通道响应. 经实验验证, 采用 L_2 范数能获得更稳定优异的性能, 故在所有实验中统一采用 L_2 范数以平衡性能与效率. NGCF 结构如图 4 所示.

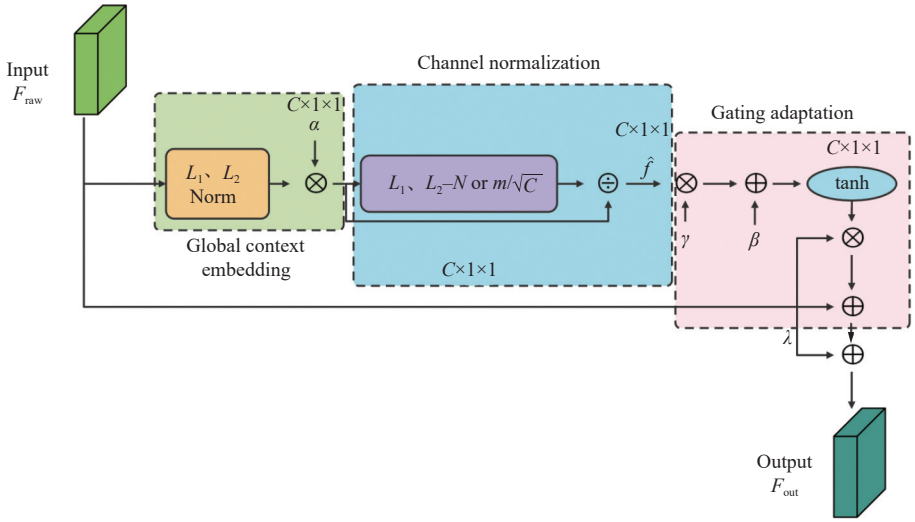

图 4 NGCF 结构图

具体而言, 设某解码层产生的原始特征为$F_{\text{raw}} \in R^{B\times N\times C}$, 其中 B 表示批大小、N 表示点数、C 表示通道数. NGCF 模块首先对F_{raw}在点 (空间) 维度上进行范数聚合, 从而获取每个通道的全局响应:

$$L_1 : f_c = \sum_{i=1}^{N} \left| F_{\text{raw},i,c} \right|,\ c = 1,\cdots,C \tag{5}$$

$$L_2 : f_c = \sqrt{\sum_{i=1}^{N} F_{\text{raw},i,c}^2 + \varepsilon} \tag{6}$$

接着, 将$s \in R^C$归一化为:

$$\hat{f}_c = \frac{f_c}{\dfrac{1}{C}\displaystyle\sum_{k=1}^{C} f_k} \tag{7}$$

以消除不同通道尺度差异.

随后, 通过两组可训练参数$\gamma,\beta \in R^C$构建门控函数:

$$g_c = 1 + \tanh\left(\gamma_c \hat{f}_c + \beta_c\right) \tag{8}$$

“+1”设计确保当$\gamma_c \hat{f}_c + \beta_c \approx 0$时仍能保持通道响应; tanh 激活函数既能对极端值进行压制, 也能保持梯度稳定. 通道重标定向量$g \in R^C$生成后, 即刻与原始特征逐通道相乘, 得到中间特征:

$$F_{\text{ng},i,c} = F_{\text{raw},i,c} \times g_c,\ i = 1,\cdots,N \tag{9}$$

其中, $F_{\text{ng},i,c}$为经过范数门控通道特征模块 (NGCF) 的门控重标定操作后生成的中间特征.

为了将变换成果平滑地融入解码特征, 并防止单纯通道重标定对特征分布产生过大扰动, 本文采用了残差融合的方式:

$$F_{\text{out}} = F_{\text{raw}} + \lambda F_{\text{ng}},\ \lambda = 0.5 \tag{10}$$

在此处固定融合权重$\lambda = 0.5$, 一方面简化了参数学习的复杂度, 避免过多可训练超参数引入的训练不稳定; 另一方面, 通过弱化 NGCF 分支对最终特征的影响, 可保留原有解码特征的主干信息, 降低对少量噪声通道抑制或放大过度的风险, 从而在提升模型判别力的同时, 确保训练过程平滑收敛. 最后, 对融合后的特征F_{out}做批量归一化 (*BatchNorm*) 和 *LeakyReLU* 激活以进一步稳定数值和引入非线性:

$$\widetilde{F_{\text{out}}} = LeakyReLU\left(BatchNorm\left(F_{\text{out}}\right)\right) \tag{11}$$

NGCF 模块的参数开销仅包含 $2C$ 个可训练的门控参数γ,β, 其计算量主要来自一次范数聚合$O(BNC)$与一次逐通道映射$O(BC)$. 残差融合过程不引入额外参数, 仅进行一次张量加权和, 因此整体 FLOPs 增幅通常低于 3%. 将 NGCF 部署到每个解码层的残差路径上, 不仅能在通道维度上动态突出解码特征中对边界和小物体更加敏感的通道, 还能通过残差保留原始解码信息, 避免对解码管线造成过度干扰.

2.4 损失函数设计

在 RandLA-Net 的原始架构中, 采用交叉熵损失函数对点云进行分类, 这在多数场景下表现优异. 但在室内建筑点云数据集中, 类别分布往往高度不平衡: 梁、柱等结构的点数远少于墙面、地板等主要类别. 交叉熵损失在训练过程中会偏向样本量大的类别, 从而使少数类别得到的训练信号不足, 影响其分割效果. 为了

解决这一问题, 本文基于焦点损失 (focal loss)[20], 构建了融合型损失函数 FCE loss, 焦点损失通过对易分类样本的梯度贡献进行衰减, 自动将更多注意力集中到难分类或少数类别样本上, 从而有效缓解数据不平衡带来的负面影响. 同时将焦点损失与交叉熵损失联合使用, 在保持整体收敛速度和准确度的同时, 提升对少数类别点云的分割性能. 交叉熵损失函数为:

$$L=\begin{cases}-\ln(p), & t=1\\ -\ln(1-p), & t=0\end{cases} \tag{12}$$

焦点损失函数为:

$$L=\begin{cases}-\alpha(1-P)^{\gamma}\ln(p), & t=1\\ -(1-\alpha)p^{\gamma}\ln(1-p), & t=0\end{cases} \tag{13}$$

其中, p 表示模型对正类的预测概率, t 表示真值标签 ($t\in\{0,1\}$), α、γ 为调节因子. 为兼顾整体收敛稳定性与对难分、少数样本的关注, 本文采用加权求和方式构建融合型损失函数:

$$L_{\text{fce}}=\lambda L_{\text{ce}}+(1-\lambda)L_{\text{fl}} \tag{14}$$

其中, 融合权重λ设置为 0.5. 这种线性组合方式, 避免引入复杂的动态调度机制, 保证训练过程的稳定性.

当模型对真实类别 t 的预测概率 p 提升时, $(1-p)^{\gamma}$ 会迅速减小, 这意味着那些已经被正确分类的“易样本”对整体损失的贡献会被自动削弱; 而预测不准的“难样本”由于p^{γ}或$(1-p)^{\gamma}$保持较大值, 则在总损失中占据更高比重. 超参数 α 则用于调节正负样本的权重, 当 α 较大时, 正类样本($t=1$)的梯度贡献更显著, 反之则强化负类($t=0$)的学习; 而 γ 控制了对易分样本的衰减程度, γ 越大, 简单样本被忽略得越彻底. 将这种焦点损失函数应用于室内三维点云分割, 不仅能够使网络更多关注难以区分的细节和少数类, 还能缓解类别不平衡带来的偏置, 从而显著提升建筑场景中梁、柱、家具等少样本结构的分割质量, 有效减少模型在多类别点云分割时的不确定性.

3 实验结果分析

3.1 实验环境及参数配置

实验环境均使用的是 Ubuntu 18.04 操作系统, CPU 为 Intel Core i7-13700K, 16 核 24 线程, 主频为 3.4 GHz, 运行内存 32 GB, GPU 为 NVIDIA GeForce RTX 3090, 24 GB 显存. 优化器选用 Adam 算法, 初始学习率设置为 0.01, 衰减率设置为 0.95. 详细设置如表 1 所示.

表 1 环境详细设置

Type	Version	Type	Value
GPU	RTX 3090	Epoch	100
Python	3.6.9	Batch	6
TensorFlow	1.15	Learning rate	1×10^{-2}
CUDA	11.1	Momentum	0.95

3.2 数据集介绍

本文研究的目的是提升点云语义分割模型整体精度, 以及将局部特征与全局语义相关联, 增强小物体和边界特征表达, 缓解类别不平衡. 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 S3DIS (Stanford large-scale 3D in-door spaces dataset) 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 6 维向量形式呈现, 包括空间位置 (xyz) 与颜色属性 (RGB).

3.3 评估指标

为了准确评估 RandLA-CGNet 的有效性, 本研究采用了多个指标进行比较分析. 语义分割性能一般采用总体精度 (overall accuracy, OA)、类平均精度 (mean class accuracy, $mAcc$) 和平均交并比 (mean intersection over union, $mIoU$) 进行评估. OA 是数据集中分割正确的样本数与总样本数之比; $mAcc$ 是数据集中每个类别的分割准确率求和再平均的结果; 交并比 (IoU) 是模型对某一类别预测结果和真实值的交集与并集的比值, 其用于测量真实值和预测值之间的相关度, 相关度越高, 该值越高; $mIoU$ 则是所有类别 IoU 的平均值, 可定义为:

$$OA=\frac{1}{N}\sum_{i=1}^{C}TP_i \tag{15}$$

$$mAcc=\frac{1}{C}\sum_{i=1}^{C}\frac{TP_i}{TP_i+FN_i} \tag{16}$$

$$mIoU=\frac{1}{C}\sum_{i=1}^{C}\frac{TP_i}{TP_i+FP_i+FN_i} \tag{17}$$

其中, TP_i表示类别 i 中被正确分类的点数, FP_i表示其他类被错误分类为 i 的点数, FN_i表示类别 i 中被错误分类的点数, C 表示语义分割的类别数量, N 表示测试点云中所有点的数量.

3.4 消融实验

为验证所提出改进模块对模型性能的提升效果, 本文在 S3DIS 数据集的 Area5 上以 RandLA-Net 为基准模型设计了 8 组消融实验. 首先, 将编码层中的局部特征聚合 LFA 模块替换为局部-全局上下文融合 LGCF 模块; 其次, 在解码层引入范数门控通道特征模块 NGCF; 最终, 构建融合型损失函数 FCE loss. 通过逐步叠加各改进模块并进行对比测试, 实验结果 (见表 2) 表明, LGCF 模块、NGCF 模块以及 FCE loss 损失函数均能有效提升模型性能, 完整组合方案实现了最佳性能. 本文表格中加粗标注均表示最优结果.

实验结果表明, 相较于基线模型 1, 通过 LGCF 模块使模型 2 的 *OA*、*mAcc* 和 *mIoU* 分别提升 0.9%、1.5% 和 1.8%, 除了木板的 *IoU* 略有下降外, 其他类别的 *IoU* 均有上升; 模型 3 将 NGCF 模块集成到解码器中, *OA*、*mAcc* 和 *mIoU* 分别提升 1.1%、1.7% 和 2.3%; 模型 4 采用交叉熵损失函数与焦点损失函数的组合形式来缓解不均衡数据对模型的负面影响, 推动 *OA*、*mAcc* 和 *mIoU* 分别提升 0.2%、0.2% 和 0.3%, 其中柱的 *IoU* 提升 10.7%. 对模型 5–7 进行两两组合后发现, 引入 LGCF 能够在保持局部细节的同时提升全局的识别精度, 同时 *mIoU* 保持在约 64.3%; 而当 NGCF 和 FCE loss 组合后, 尽管 *OA* 较仅引入 NGCF 模块时有所下降, 但窗、椅子、沙发、书架等类别均较基准模型有所提升. 在 LGCF 和 FCE loss 组合模型中, *OA* 指标相较于模型 2 (仅 LGCF) 略有下降, 而少数类别 (如 sofa) 性能显著提升. 这表明 FCE loss 有效地将优化重心转向难分的少数类别, 其与 LGCF 的关系是功能协同而非相互抑制. 最终模型在集成所有模块后实现了性能的全面最优.

表 2 消融实验结果 (%)

Method	LGCF	NGCF	FCE loss	*OA*	*mAcc*	*mIoU*	*IoU*											
							ceiling	floor	wall	beam	column	window	door	table	chair	sofa	bookcase	board	clutter
1 (base)	—	—	—	87.6	70.3	62.3	92.8	97.6	80.8	0	17.4	61.1	37.1	77.9	85.8	65.2	71.0	69.6	53.5
2	√	—	—	88.5	71.8	64.1	93.1	**98.0**	82.8	0	24.6	**67.4**	47.2	78.9	85.9	62.3	72.8	66.2	54.4
3	—	√	—	88.7	72.0	64.6	**94.5**	97.7	82.8	0.1	22.4	62.8	49.5	**79.5**	86.4	69.8	73.4	67.1	54.2
4	—	—	√	87.8	70.5	62.6	92.6	97.7	80.2	0	**28.1**	60.7	37.5	75.9	87.3	60.3	72.2	69.0	52.7
5	√	√	—	89.0	72.1	64.3	94.0	97.5	**83.6**	0	25.8	63.4	46.7	77.7	86.7	64.9	72.7	68.0	54.4
6	√	—	√	88.0	73.4	64.6	92.4	97.6	82.0	0.1	27.1	62.0	**55.1**	75.0	85.6	74.4	67.5	**70.7**	50.7
7	—	√	√	88.1	71.5	63.8	93.9	97.9	82.6	0	20.0	66.7	38.1	77.0	86.9	67.2	**74.4**	68.1	**57.2**
8	√	√	√	**89.1**	**73.6**	**65.9**	93.6	97.6	83.3	0.1	25.5	66.4	44.9	79.4	**89.6**	**78.5**	73.2	69.0	55.5

所有模型经过有序组合后, 相较于基准模型 1, *OA*、*mAcc* 和 *mIoU* 均有明显提升. 最终, 结合所有改进策略的 RandLA-CGNet 模型, *OA*、*mAcc* 和 *mIoU* 有效提升 1.5%、3.3% 和 3.6%, 各项指标均优于基线模型.

为进一步探究 NGCF 模块中融合权重 λ 的设置对模型性能的影响, 本文设计了详细的消融实验 (见表 3). 实验结果表明, λ=0.5 时 *OA*、*mAcc* 和 *mIoU* 分别达到 88.7%、72.0% 和 64.6%, 是一个性能最优的平衡点. 当 λ 取值较低 (如 0.25) 时, NGCF 增强特征融入不足, 关键通道显著性提升受限, *mIoU* 较 λ=0.5 时下降 1.3%; 而当 λ 取值较高 (如 0.75 或 1.0) 时, NGCF 输出的重构特征对原始特征分布造成过度扰动, 模型整体性能较 λ=0.5 时下降明显. λ=0.5 时平衡了原始信息与增强信号, 既保留原始特征的核心信息, 又能充分发挥 NGCF 对关键通道的增强作用, 尤其对敏感稀疏类别, 有效平衡了特征扰动与信息增益, 且结果稳定.

表 3 NGCF 模块不同融合权重 λ 的消融实验 (%)

λ值	*OA*	*mAcc*	*mIoU*	*IoU*												
				ceiling	floor	wall	beam	column	window	door	table	chair	sofa	bookcase	board	clutter
0	87.6	70.3	62.3	92.8	97.6	80.8	0	17.4	61.1	37.1	77.9	85.8	65.2	71.0	**69.6**	53.5
0.25	88.0	71.0	63.3	93.4	97.6	80.9	0	20.1	59.7	46.8	78.4	85.7	69.8	71.2	67.2	52.4
0.5	**88.7**	**72.0**	**64.6**	**94.5**	**97.7**	**82.8**	0.1	**22.4**	**62.8**	**49.5**	**79.5**	**86.4**	**69.8**	**73.4**	67.1	**54.2**
0.75	87.8	71.6	62.8	91.5	97.6	80.8	0	21.8	62.1	41.3	77.5	86.0	69.4	68.0	68.6	51.4
1.0	86.9	71.3	62.6	91.0	97.5	80.4	0	21.1	61.5	42.8	78.7	84.9	69.2	69.8	67.7	49.6

3.5 对比实验

为验证本文提出的算法在室内大规模点云语义分割中的优越性, 分别使用六折交叉验证和单独将 Area5 作为测试区域对本文算法与当前主流算法进行对比.

表 4 为不同模型使用六折交叉验证在 S3DIS 数据集的实验结果, 六折交叉验证是将数据集分为 6 份, 每一份分别作为测试集进行 6 次的测试和训练, 可以充分评估模型的泛化能力. 本文模型的总体精度比 RandLA-Net 高出 0.8%, 类平均精度较 RandLA-Net 提高 1.4%, 在平均交并比上比 RandLA-Net 提高 1.9%. 更细粒度地看, 在核心类别上也普遍优于 RandLA-Net: 如对 wall (墙壁) 的 *IoU* 从 80.6% 提升至 81.5%, beam (梁) 从 62.4% 提升至 63.3%、clutter (杂物) 从 60.1% 提升至 61.9%, 在整体性能与类别级精度上全面超越 PointNet (*OA* 从 78.6% 提升至 88.8%、*mIoU* 从 47.6% 提升至 71.9%)、RSNet (*mAcc* 从 66.5% 提升至 83.4%、*mIoU* 从 56.5% 提升至 71.9%) 等早期方法. 对于边界敏感类 window (窗)、door (门)、table (桌), 单类别分别较 KPConv、SPG 有略微下降, 但与其他算法相比, 均有明显提升, 窗的 *IoU* 较 PointNet 提升 18.1%, 门的 *IoU* 较 DGCNN 提高 20.7%, 桌的 *IoU* 较 RSNet 的 *IoU* 提升 13.1%. 对于 beam (梁) 和 column (柱) 等稀疏类别, 较 DGCNN、PointWeb、ShellNet 等模型提升 2.9%–16.3% 不等, 虽然在 floor (地板)、wall (墙壁) 等类别中不及 PointNetCNN、KPConv 等模型, 但是均达到了当前先进水平. 此外, 虽然本文方法在整体分类精度上略逊于 PointMM 的 90.4%, 但在平均交并比上仍较之提升 1.2%, 并且在关键评估指标上全面超越 SCF-Net、LG-Net 和 FGC-AF 等先进模型. Wall (墙壁)、beam (梁)、clutter (杂物) 等核心类别实现性能提升, 验证了本文模型对复杂室内场景分割的有效性.

表 5 为将 Area5 作为测试区域时的实验结果. 使用 Area5 作为测试集, 由于 Area5 和其他区域有所不同, 包含的物体存在差异, 可以更好地评估模型的泛化性能. 可以看出本文模型的总体精度比 RandLA-Net 高出 1.5%, 类平均精度较 RandLA-Net 提高 3.3%, 在平均交并比上较 RandLA-Net 提高 3.6%. 在大类别与易混淆类别中 本文模型对于 wall (墙壁) 的交并比低于 MinkowskiNet 的 86.2%, 但较 RandLA-Net 提升 2.5%; 对于小物体与边界敏感类别, 本文模型较其他主流模型取得领先: 在 sofa (沙发) 类别上的 *IoU* 达到 78.5%, 较 RandLA-Net 的 65.2% 提升 13.3%; 在 window (窗) 类别上的 *IoU* 达到 66.4%, 较 RandLA-Net 的 61.1% 提升 5.3 %, 较 PCT 提升 4.8%. 与 LG-Net 和 BADENet 相比, 本方法的总体精度分别提高 1.0% 和 0.4%, 平均交并比则分别提升 1.0% 和 3.3%, 同时在类平均精度上较 BADENet 提升了 9.3%, 实现了对两者的超越. Area5 中的实验结果, 更验证了本文模型在点云分割跨区域泛化、小物体识别、边界精细化中的有效性.

3.6 可视化分析

为了能够直观分析点云语义分割结果, 本文将点云语义分割结果可视化, 如图 5 所示, 从上到下依次为原始点云、带有真实分割标签的分割结果、由主流点云处理模型 PointNet++分割后的结果、RandLA-Net 模型分割后的结果以及本文方法分割后的结果, 选取数据集中的 3 个场景进行分析.

表 4 六折交叉验证在 S3DIS 数据集上的对比实验 (%)

Model	*OA*	*mAcc*	*mIoU*	*IoU*												
				ceiling	floor	wall	beam	column	window	door	table	chair	sofa	bookcase	board	clutter
PointNet[9]	78.6	66.2	47.6	88.0	88.7	69.3	42.4	23.1	47.5	51.6	54.1	42.0	9.6	38.2	29.4	35.2
RSNet[21]	—	66.5	56.5	92.5	92.8	78.6	32.8	34.4	51.6	68.1	59.7	60.1	16.4	50.2	44.9	52.0
SPG[22]	86.4	73.0	62.1	89.9	95.1	76.4	62.8	47.1	55.3	68.4	**73.5**	69.2	63.2	45.9	8.7	52.9
PointCNN[12]	88.1	75.6	65.4	94.8	97.3	75.8	63.3	51.7	58.4	57.2	71.6	69.1	39.1	61.2	52.2	58.6
PointWeb[23]	87.3	76.2	66.7	93.5	94.2	80.8	52.4	41.3	64.9	68.1	71.4	67.1	50.3	62.7	62.2	58.5
ShellNet[24]	87.1	—	66.8	90.2	93.6	79.9	60.4	44.1	64.9	52.9	71.6	84.7	53.8	64.6	48.6	59.4
KPConv[25]	—	79.1	70.6	93.6	92.4	**83.1**	63.9	54.3	66.1	**76.6**	57.8	64.0	**69.3**	**74.9**	61.3	60.3
SCF-Net[17]	88.4	82.7	71.6	93.3	96.4	80.9	64.9	47.4	64.5	70.1	71.4	81.6	67.2	64.4	67.5	60.9
DGCNN[11]	84.5	—	55.5	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
LG-Net[13]	88.3	82.4	70.8	93.7	96.4	81.3	**65.2**	51.8	66.2	69.7	69.1	75.1	63.9	63.5	66.0	58.4
PointMM[14]	90.4	—	70.7	**95.4**	**97.5**	81.1	59.5	38.8	**66.5**	73.9	73.0	**84.0**	53.5	68.1	58.6	**69.5**
FGC-AF[15]	88.6	—	71.2	93.5	96.6	82.0	63.7	**53.9**	65.3	70.7	71.3	73.1	59.6	65.4	**68.8**	61.5
RandLA-Net[18]	88.0	82.0	70.0	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
Ours	**88.8**	**83.4**	**71.9**	93.7	97.0	81.5	63.3	48.5	65.6	71.4	72.8	78.3	66.0	67.4	66.7	61.9

表 5 S3DIS 数据集中 Area5 的对比实验 (%)

Model	*OA*	*mAcc*	*mIoU*	*IoU*												
				ceiling	floor	wall	beam	column	window	door	table	chair	sofa	bookcase	board	clutter
PointNet[9]	—	49.0	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
PointNet++[10]	84.0	63.6	55.2	90.5	98.1	75.8	0	6.8	57.6	21.5	69.6	79.3	50.7	62.5	58.8	45.7
PointCNN[12]	85.9	63.9	57.3	92.3	98.2	79.4	0	17.6	22.8	62.1	74.4	80.6	31.7	66.7	62.1	56.7
SPGraph[22]	86.4	66.5	58.0	89.4	96.9	78.1	0	42.8	48.9	61.6	**84.7**	75.4	69.8	52.6	2.1	52.2
PCT[26]	—	67.7	61.3	92.5	98.4	80.6	0	19.4	61.6	48.0	76.6	85.2	46.2	67.7	67.9	52.3
MinkowskiNet[27]	—	71.7	65.4	91.8	**98.7**	**86.2**	0	34.1	48.9	**62.4**	81.6	**89.8**	47.2	**74.9**	**74.4**	**58.6**
DGCNN[11]	83.3	—	47.6	92.8	97.5	74.9	0	11.8	50.7	23.7	69.6	66.4	48.7	8.6	31.6	42.1
LG-Net[13]	88.1	—	64.9	—	—	—	—	—	—	—	—	—	—	—	—	—
BADENet[16]	88.7	64.3	62.6	—	—	—	—	—	—	—	—	—	—	—	—	—
RandLA-Net	87.6	70.3	62.3	92.8	97.6	80.8	0	17.4	61.1	37.1	77.9	85.8	65.2	71.0	69.6	53.5
Ours	**89.1**	**73.6**	**65.9**	**93.6**	97.6	83.3	**0.1**	25.5	**66.4**	44.9	79.4	89.6	**78.5**	73.2	69.0	55.5

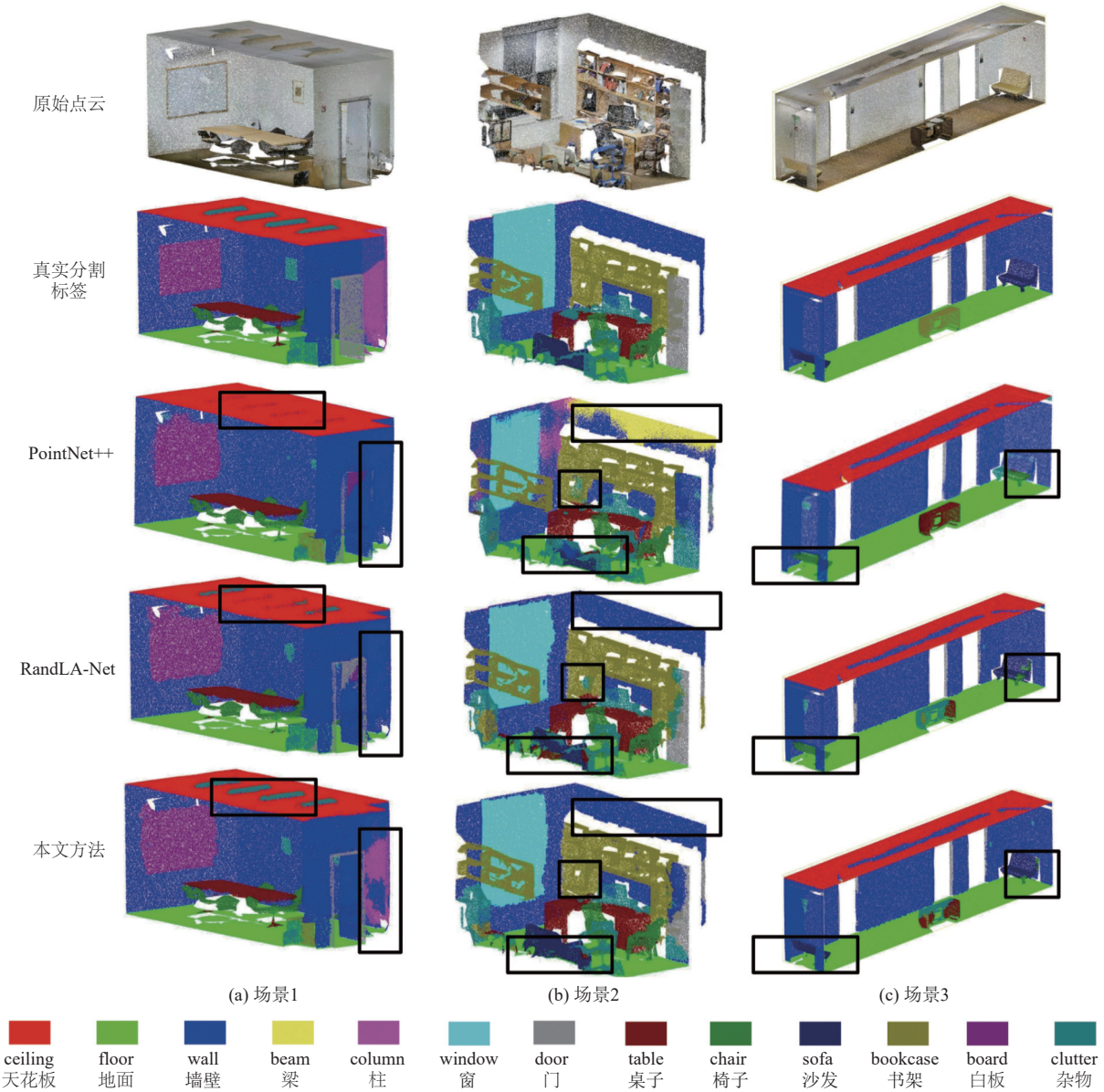

图 5 分割结果对比图

在图 5(a) 场景 1 中, PointNet++在天花板与杂物 (clutter) 交界处表现出明显的“漏分”和“误分”, 其特征表达过程中缺乏稀疏下采样策略, 不仅分割边界过于粗糙, 有时还将杂物丢到背景或随机标错. 相比之下, RandLA-Net 通过局部特征聚合取得更平滑的结果, 却依然将杂物误归为天花板, 这主要源于其仅依赖局部特征聚合, 缺乏对整体场景的全局语义感知, 因而难以判断“天花板”何时结束、杂物何时开始. 为此, 本文在编码器后半段引入 LGCF. LGCF 的全局分支通过对所有点云特征进行平均池化与通道重标定, 生成跨场景的语义注意力向量, 从而在通道维度上增强天花板和平面结构的表征, 抑制对杂物的误激活, 使分割边界更加清晰锐利. 此外, PointNet++与 RandLA-Net 在右侧柱子 (column) 处常将其误分为墙壁 (wall), 柱体底座与地面 (floor) 分界模糊. 在本文提出的模型中, FCE loss 在训练时为这些“难分”点分配更高的梯度权重, 从而有效提升模型对柱子等低频类别的识别能力, 减少误分, 恢复柱体完整轮廓.

在图 5(b) 场景 2 中, PointNet++对书架 (bookcase) 与沙发 (sofa) 等小物体结构的刻画最为薄弱, 经常出现“空洞”或将其误标为椅子 (chair)、桌 (table), 甚至将大面积墙体分割为梁与柱体; RandLA-Net 虽改善了局部几何的表达, 但在书架顶部、沙发扶手等细节依然漏分, 这些小物体点云往往样本稀少、局部形状易混淆. LGCF 的局部分支依旧负责精细几何编码, 而解码层的 NGCF 模块通过范数自适应重标定通道, 增强对细粒度结构 (如沙发脚) 的响应, 再配合焦点损失对这些难分样本加权, 本文模型能够精确地分割书架、沙发与地面的分界.

在图 5(c) 场景 3 中, 由于模型在处理外观相似、空间相邻的家具类别时缺乏足够的判别力. PointNet++和 RandLA-Net 在相似外观的家具类别 (如沙发和椅子) 上均表现出混淆: 左侧沙发误标为椅子, 右侧沙发部分区域又被误归为地面. 本文提出的 NGCF 模块在解码阶段对通道维度上的特征进行“门控”滤波, 动态赋予与沙发特征最相关的通道更高响应, 并在残差融合时保留原始解码信息, 从而有效消除沙发-椅子-地面三者之间的判别盲区, 提升对沙发面积及边界的分割准确性.

上述典型场景中的对比结果充分证明本文所提模型在细节分割、边界检测及少样本类别识别上的明显提升, 进一步验证了 LGCF、NGCF 与 FCE loss 三者协同作用的有效性.

4 结论与展望

本文针对 RandLA-Net 在大规模室内建筑点云分割中对小物体与边界细节建模不足的问题, 设计了局部-全局上下文融合 (LGCF) 和范数门控通道特征 (NGCF) 两大模块, 同时采用 FCE loss 来解决类别不平衡. LGCF 在编码器后 3 层下采样后并行融合局部特征聚合和全局通道注意力, 既强化邻域几何细节, 又注入跨层次的整体语义; NGCF 则在每次解码层跳跃连接后, 以可学习的门控参数和残差融合的方式, 对通道维度进行动态重标定, 有效突出对边界与稀有类别有贡献的通道并抑制噪声通道. 联合采用融合型损失函数, 使模型在兼顾多数易分点快速收敛的同时, 更多关注难分和少数类样本. 经 S3DIS 数据集六折交叉验证, 本文方法相比原始 RandLA-Net 在整体准确率 (*OA*) 上由 88.0% 提升至 88.8%, 类平均精度 (*mAcc*) 由 82.0% 提升至 83.4%, 平均交并比 (*mIoU*) 由 70.0% 提升至 71.9%, 其中梁、柱、板材等小体量类别 *IoU* 平均提升 2.3%, 门框与桌椅边界等细节处提升 2.7%; 与此同时, 参数量仅增加约 4%, 增幅不足 5%, 兼顾了精度与效率.

综上所述, 本研究提出的改进算法为点云语义分割任务中对低频类别和细节边界的准确建模提供了有效解决方案. 未来将进一步探索多模态传感融合, 以及跨场景自适应迁移等方向, 在更丰富的室内环境中实现更通用、更高效的三维语义分割解决方案.

参考文献

1 Zhang YT, Wang ZY, Wang X, *et al*. A 3D pickup estimation method for industrial parallel robots based on point cloud simplification and registration. The International Journal of Advanced Manufacturing Technology, 2024, 133(11): 5175–5195.

2 Gao Y, Xia SB, Wang C, *et al*. Semantic segmentation of airborne LiDAR point clouds with noisy labels. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 5706614.

3 Abbasi R, Bashir AK, Rehman A, *et al*. 3D Lidar point cloud segmentation for automated driving. IEEE Intelligent Transportation Systems Magazine, 2025, 17(1): 8–29. [doi: 10.1109/MITS.2023.3325854]

4 Lawin FJ, Danelljan M, Tosteberg P, *et al*. Deep projective 3D semantic segmentation. Proceedings of the 17th International Conference on Computer Analysis of Images and Patterns. Ystad: Springer, 2017. 95–107.

5 Boulch A, Le Saux B, Audebert N. Unstructured point cloud semantic labeling using deep segmentation networks. Proceedings of the 2017 Workshop on 3D Object Retrieval. Lyon: Eurographics Association, 2017. 17–24.
6 Tatarchenko M, Park J, Koltun V, *et al*. Tangent convolutions for dense prediction in 3D. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 3887–3896.
7 Wu BC, Wan A, Yue XY, *et al*. SqueezeSeg: Convolutional neural nets with recurrent CRF for real-time road-object segmentation from 3D LiDAR point cloud. Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA). Brisbane: IEEE, 2018. 1887–1893.
8 Milioto A, Vizzo I, Behley J, *et al*. RangeNet++: Fast and accurate LiDAR semantic segmentation. Proceedings of the 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Macau: IEEE, 2019. 4213–4220.
9 Charles R.Qi, Su H, Mo K, *et al*. PointNet: Deep learning on point sets for 3D classification and segmentation. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 77–85.
10 Charles R.Qi, Yi L, Su H, *et al*. PointNet++: Deep hierarchical feature learning on point sets in a metric space. Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 5105–5114.
11 Wang Y, Sun YB, Liu ZW, *et al*. Dynamic graph CNN for learning on point clouds. ACM Transactions on Graphics, 2019, 38(5): 146.
12 Li YY, Bu R, Sun MC, *et al*. PointCNN: Convolution on *X*-transformed points. Proceedings of the 32nd International Conference on Neural Information Processing Systems. Montréal: Curran Associates Inc., 2018. 828–838.
13 Zhao YQ, Ma XY, Hu B, *et al*. A large-scale point cloud semantic segmentation network via local dual features and global correlations. Computers & Graphics, 2023, 111: 133–144.
14 Chen RX, Wu J, Luo Y, *et al*. PointMM: Point cloud semantic segmentation CNN under multi-spatial feature encoding and multi-head attention pooling. Remote Sensing, 2024, 16(7): 1246. [doi: 10.3390/rs16071246]
15 Chen J, Chen YP, Wang C. Feature graph convolution network with attentive fusion for large-scale point clouds semantic segmentation. IEEE Geoscience and Remote Sensing Letters, 2023, 20: 6501405.
16 Zhao J, Lu J, Zhou J, *et al*. Boundary-aware dual edge convolution network for indoor point cloud semantic segmentation. Computers and Electrical Engineering, 2024, 116: 109219. [doi: 10.1016/j.compeleceng.2024.109219]
17 Fan SQ, Dong QL, Zhu FH, *et al*. SCF-Net: Learning spatial contextual features for large-scale point cloud segmentation. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 14499–14508.
18 Hu QY, Yang B, Xie LH, *et al*. Learning semantic segmentation of large-scale point clouds with random sampling. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(11): 8338–8354.
19 Yang ZX, Zhu LC, Wu Y, *et al*. Gated channel transformation for visual recognition. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 11791–11800.
20 Lin TY, Goyal P, Girshick R, *et al*. Focal loss for dense object detection. Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017. 2999–3007.
21 Huang QH, Wang WY, Neumann U. Recurrent slice networks for 3D segmentation of point clouds. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 2626–2635.
22 Landrieu L, Simonovsky M. Large-scale point cloud semantic segmentation with superpoint graphs. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 4558–4567.
23 Zhao HS, Jiang L, Fu CW, *et al*. PointWeb: Enhancing local neighborhood features for point cloud processing. Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 5560–5568.
24 Zhang ZY, Hua BS, Yeung SK. ShellNet: Efficient point cloud convolutional neural networks using concentric shells statistics. Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019. 1607–1616.
25 Thomas H, Charles R.Qi, Deschaud JE, *et al*. KPConv: Flexible and deformable convolution for point clouds. Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019. 6410–6419.
26 Guo MH, Cai JX, Liu ZN, *et al*. PCT: Point cloud transformer. Computational Visual Media, 2021, 7(2): 187–199. [doi: 10.1007/s41095-021-0229-5]
27 Choy C, Gwak J, Savarese S. 4D spatio-temporal ConvNets: Minkowski convolutional neural networks. Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 3070–3079.

(校对责编: 李慧鑫)