

LSENet: 局部子域感知与边缘增强动态融合的机载

LiDAR 点云语义分割网络

姬大力^{1, 2} 王永波^{1, 2} 郑南山^{1, 2} 杨敏^{1, 2}

1. 中国矿业大学 国土环境与灾害监测自然资源部重点实验室, 江苏 徐州, 221116;

2. 中国矿业大学 江苏省资源环境信息工程重点实验室, 江苏 徐州, 221116

摘 要: 机载 LiDAR 点云语义分割是三维地物精细化解译与场景理解的重要基础, 在智慧城市建设、城市三维建模、交通基础设施提取、国土空间调查和生态环境监测等领域具有重要的应用价值。针对现有深度学习点云语义分割方法在局部几何结构刻画与上下文建模之间协同不足的问题, 提出了一种新的机载 LiDAR 点云语义分割网络 LSENet(Local Subdomain-aware and Edge-enhance Dynamic Fusion Network)。首先通过局部子域自适应划分提取每个采样点细粒度空间几何特征; 其次在子域间引入特征增强的交互机制以强化局部空间感知与信息流通; 最后设计了边缘增强的分组多尺度双注意力动态融合模块, 实现了关键特征的动态加权融合, 增强了不同语义类别之间的可分性, 并提升了模型对复杂结构及边界区域的判别能力。基于 Vaihingen 3D 与 DFC2019 两个公开数据集的实验结果表明, LSENet 分别取得了 83.9% 和 97.8% 的总体精度、70.5% 和 89.1% 的平均 F1 分数, 整体性能优于部分对比方法, 验证了所构建网络在机载 LiDAR 点云语义分割中的有效性和跨场景适用性。

关键词: 机载 LiDAR 点云; 深度学习; 语义分割; 局部子域感知; 动态融合; 边缘增强

中图分类号: P237

引用格式: 姬大力, 王永波, 郑南山, 杨敏. 2026. LSENet: 局部子域感知与边缘增强动态融合的机载 LiDAR 点云语义分割网络. 遥感学报, 26(1): 7-22.

Ji D L, Wang Y B, Zheng N S, Yang M. 2026. LSENet: Local Subdomain-Aware and Edge-Enhanced Dynamic Fusion Network for Semantic Segmentation of Airborne LiDAR Point Cloud . National Remote Sensing Bulletin, 16(1):7-22[DOI: 10.11834/jrs.*****]

收稿日期: 2026-03-27; 预印本:****_**_**

基金项目: 国家自然科学基金项目 (U22A20569); 江苏省自然科学基金项目 (BK20221549)

第一作者简介: 姬大力, 2001 年生, 男, 硕士研究生, 研究方向为基于深度学习的 LiDAR 点云数据处理. E-mail: qq1339899624@163.com

通信作者简介: 王永波, 1982 年生, 男, 副教授, 研究方向为多模态点云配准; 基于深度学习的 LiDAR 点云数据处理; 基于 LiDAR 点云的地理实体表面形态特征提取、分析及其三维拓扑重建; 基于地面 LiDAR 点云与近景摄影影像相结合的数字城市混合建模等. E-mail: ybwang816@163.com

23 1 引言

24 随着激光 LiDAR 技术的不断进步，点云已成为高精度三维空间信息的重要载体，能够直接、
25 精确地刻画真实场景的几何结构与空间分布。凭借其高精度、强真实感和无投影畸变等优势，在自
26 动驾驶（Chen 等，2025, Prathiba 等，2025）、机器人感知（Nguyen 等，2024）和实景三维模型构建
27 （胡博妮等，2024）等领域得到了广泛应用。在地图快速更新、智能交通与数字化管理系统（谢东
28 辉等，2024）等应用需求不断提升的背景下，基于点云场景理解与智能应用的语义粒度逐步从“道
29 路级”向“行为级”转变。点云语义分割作为其核心任务，旨在通过为每个点分配语义类别标签，
30 从而将无序的点云数据转化为具有一致语义和几何结构的对象。

31 传统点云语义分割方法主要依赖手工设计特征，并结合支持向量机（Support Vector Machine,
32 SVM）（Hui 等，2016）、随机森林（Kim and Sohn, 2011）等机器学习模型来完成。这类方法在小
33 规模场景中表现尚可，但面对地物结构复杂、尺度差异显著及点密度不均等问题时，手工特征泛化
34 能力受限，且难以实现端到端优化，制约了其在复杂场景下的应用潜力。

35 深度学习的兴起为点云语义分割提供了更具表达能力的数据驱动范式。相较于依赖先验设计的
36 传统方法，基于深度学习的模型能够自主学习具有判别力的特征表示。根据数据输入形式不同，现
37 有方法主要包括基于投影或体素化的间接学习方法与直接处理原始点云的端到端方法。受卷积神经
38 网络（Lecun 等，1998）启发，部分研究将点云映射至二维图像或三维体素空间以适配卷积结构，
39 并通过多尺度特征融合提升性能（Zhao 等，2018）。然而投影过程不可避免地造成几何信息损失。
40 相比之下，体素化方法如 VoxNet（Maturana 等，2015）虽然在一定程度上保留了空间结构，但是体
41 素分辨率与计算复杂度之间的矛盾限制了其在大规模场景中的实用性。尽管后续的稀疏卷积技术在
42 计算效率上取得进展（Schmohl and Sörgel, 2019），但体素离散化带来的结构误差仍难以完全避免。

43 为克服上述问题，研究重心逐步转向对原始点云端到端建模的方法。PointNet（Qi 等，2017）
44 首次通过共享多层感知机（MLP）和具备置换不变性的池化操作实现无序点集的直接处理。在此基
45 础上，PointNet++（Qi 等，2017）、DGCNN（Wang 等，2019）、PointSIFT（Jiang 等，2018）与 KP-Conv
46 （Thomas 等，2019）等分别从层次化结构捕捉多尺度信息、借助动态图刻画邻域关系、引入方向

编码增强空间感知能力以及利用连续卷积逼近真实几何分布等多个角度不断完善局部几何特征表达能力。随着研究的深入，部分工作开始关注如何在增强局部判别能力的同时维持全局语义的一致性与稳定性。Huang 等 (Huang 等, 2020) 提出在深度特征学习之后引入流形嵌入与图模型优化实现全局约束, 缓解了复杂场景中远距离依赖带来的分类不稳定性。进一步地, 为在特征学习阶段直接建模局部-全局上下文关系, Zhang 等 (Zhang 等, 2022) 设计了双注意力机制并结合位置同质性模块, 实现了对局部几何特征的精细刻画与全局上下文信息的精准捕捉; Li 等 (Li 等, 2022) 通过视图解耦与局部-全局聚合机制改善信息传播受限问题; Wang 等 (Wang 等, 2024) 利用多级注意力促进层级特征交互, 从而能够同时捕获局部注意力特征和全局结构; Pan 等 (Pan 等, 2025) 通过深度注意力模块动态生成卷积权重, 在增强局部结构感知的同时融合全局空间信息; Sun 等 (Sun 等, 2026) 利用多尺度局部特征增强模块细化几何细节, 并引入智能体注意力机制在保持线性复杂度的前提下实现全局上下文的高效建模。针对交通场景的特殊性, Chen 等 (Chen 等, 2025) 结合位置编码与多上下文交互模块强化交通场景下的空间依赖建模能力。

尽管相关工作已在局部几何结构刻画与上下文建模方面取得了一定突破, 但两者之间的协同机制仍不充分: 首先点云复杂拓扑结构导致局部几何空间特征难以被精确编码; 其次注意力机制在建模长距离依赖的同时往往削弱了局部细节表达能力; 最后, 多种注意力在融合策略上往往静态分配权重, 缺乏对特征的自适应调整能力且易过度平滑边缘信息。基于上述分析, 提出一种新的点云语义分割网络 LSENet, 通过强化局部空间感知能力, 引入更具表达力的区域特征交互机制, 借助边缘增强模块与多尺度双注意力的动态融合策略, 实现对局部几何结构、上下文依赖关系及边界细节的协同建模, 从而有效提升模型在机载 LiDAR 点云语义分割中的表现。

2 研究方法

2.1 网络结构

LSENet 网络采用如图 1 类似于 U-net (Ronneberger 等, 2015) 编码器-解码器架构实现。编码器接收形状为 $N \times C$ 的原始点云数据, N 表示点数 (初始值为 4096), C 表示点云的特征维度。经最远点采样操作进行下采样, 点数目由原来的 N 变为 $N/4$ 。随后引入 LSENet 网络三大核心模块: 局

部子域空间感知的特征编码聚合模块（Local Subdomain Space-aware feature Encoding Aggregation module, LSSEA）、局部子域特征增强与交互注意力模块（Local Subdomain feature Enhancement and Interaction Attention module, LSEIA）和边缘增强的分组多尺度双注意力动态融合模块（Edge-enhanced Groupped Multi-scale Dual-Attention dynamic fusion module, EGMDA），以逐步提取并增强点云的局部结构特征和上下文语义信息。随着网络层级加深，点云数量逐层下采样为上一层的四分之一。解码器接收来自编码器数据，采用逆距离加权插值将低分辨率特征逐步恢复至原始点云分辨率，并通过跳跃连接融合对应分辨率的浅层特征，最终经全连接层输出逐点预测结果。

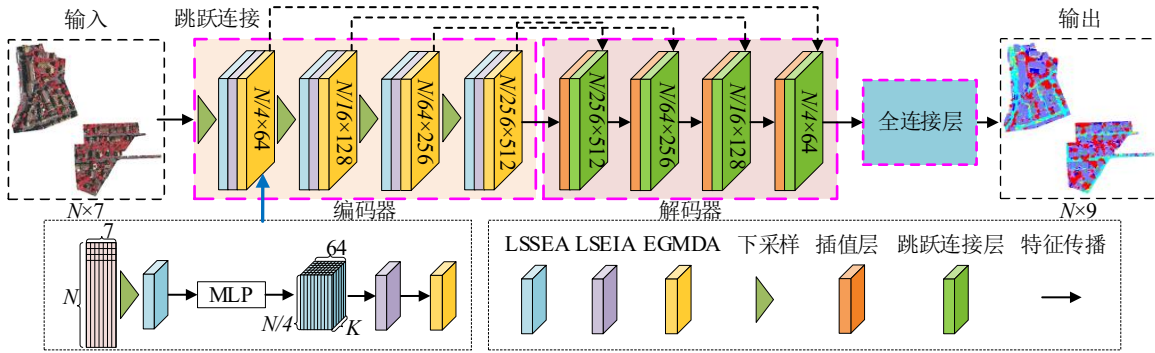


图 1 LSENet 结构

Fig. 1 Structure of LSENet

2.2 LSSEA 模块

LSSEA 旨在构建具有空间感知能力的局部结构表示。该模块首先以最远点采样选取采样点，并通过球查询获取其局部邻域；其次构建局部直角坐标系划分空间子域，融合相对位置编码、空间分布编码与几何属性编码，分别从点间相对几何关系、空间分布特性、局部几何方向性以及邻域曲面凹凸属性四个维度对点云局部结构进行精细化表达。

2.2.1 相对位置与空间感知编码

相对位置编码用于显式捕捉采样点与其邻域点之间的几何关系。设采样点为 $p_i \in \mathbb{R}^{N \times 3}$ ，特征信息表示为 f_i 。通过球查询获取的 K 个邻域点坐标 $P_i^k = \{p_i^1, p_i^2, \dots, p_i^k\}$ 及其对应特征 $F_i^k = \{f_i^1, f_i^2, \dots, f_i^k\}$ ($k=1, 2, \dots, K$)。则邻域点相对位置编码 r_i^k 定义为：

$$\mathbf{r}_i^k = \left[(p_i - p_i^k), \|p_i - p_i^k\|_2, f_i^k \right] \quad (1)$$

其中， $[\cdot]$ 表示沿特征维度的拼接操作， $\|\cdot\|_2$ 表示计算邻域点与采样点之间的欧氏距离。

空间感知编码进一步引入空间分布信息，通过对采样点及其邻域点进行主成分分析（Principal Component Analysis, PCA）构建局部直角坐标系，具体如下：

（1）构建协方差矩阵 $\mathbf{C}_i$ ：

$$\mathbf{C}_i = \frac{1}{K} \sum_{k=1}^K (p_i - p_i^k)^T (p_i - p_i^k) \quad (2)$$

（2）计算 $\mathbf{C}_i$ 的特征值与特征向量，假设特征值分别为 λ_i^1 、 λ_i^2 和 λ_i^3 且满足 $\lambda_i^1 > \lambda_i^2 > \lambda_i^3$ ，则局部直角坐标系构建时 x_i 轴、 y_i 轴、 z_i 轴分别与 λ_i^1 、 λ_i^2 、 λ_i^3 对应的特征向量相一致。

基于局部直角坐标系，根据邻近点与中心点的相对位置关系，将其划分为八个子域。采用 one-hot 编码方式对邻域点的具体分布进行表达并定义空间分布感知编码，具体表示为：

$$(\boldsymbol{\Theta}_i^k)^l = \delta(o_i^k, l) = \begin{cases} 1, & o_i^k = l \\ 0, & \text{otherwise} \end{cases}, l = 1, 2, \dots, 8 \quad (3)$$

其中， $\boldsymbol{\Theta}_i^k \in \mathbb{R}^{N \times 8}$ 为第 i 个中心点对应的第 k 个邻域点空间感知编码， o_i^k 表示该邻域点所属的子域， l 表示子域编号。

2.2.2 局部几何信息感知编码

为进一步增强对几何曲面形态的刻画能力，局部几何信息感知编码引入主方向与主曲率作为补充描述。通过在局部邻域内拟合二次曲面 $h_i(u_i, v_i) = a_i + b_i u_i + c_i v_i + d_i u_i^2 + e_i u_i v_i + t_i v_i^2$ 并构建 Weingarten 变换矩阵，利用奇异值分解获得局部表面的主曲率 μ_i^k 。主方向 $\boldsymbol{\eta}_i^k$ 对应于局部球形坐标系的 x_i 轴：

$$\mathbf{A}_i^T = - \begin{bmatrix} L_i & M_i \\ M_i & N_i \end{bmatrix} \begin{bmatrix} E_i & T_i \\ T_i & G_i \end{bmatrix}^{-1} \quad (4)$$

其中， $L_i = 2d_i / \sqrt{1 + b_i^2 + c_i^2}$ ， $M_i = e_i / \sqrt{1 + b_i^2 + c_i^2}$ ， $N_i = 2t_i / \sqrt{1 + b_i^2 + c_i^2}$ ， $E_i = 1 + b_i^2$ ， $T_i = b_i c_i$ ，

$$G_i = 1 + c_i^2。$$

将相对位置编码 r_i^k 、空间感知编码 θ_i^k 与局部几何信息编码 v_i^k 拼接后输入 MLP 进行特征提取：

$$q_i^k = \text{ReLU}\left(\text{BN}\left(W_0\left(\left[r_i^k, \theta_i^k, v_i^k\right]\right)\right)\right) \quad (5)$$

其中， $v_i^k = [\eta_i^k, \mu_i^k]$ ，ReLU 为激活函数，BN 为批归一化层， W_0 为 MLP 的权重。

2.3 LSEIA 模块

在获得子域的空间结构表示后，特征建模面临一个新的问题：子域划分虽然增强了方向感知能力，但是在一定程度上割裂了局部区域的连续性。LSEIA 模块正是围绕这一矛盾展开，通过在子域内部强化表达的同时引入跨子域信息交互机制，从而恢复局部区域的整体语义一致性。整体流程如图 2 所示。

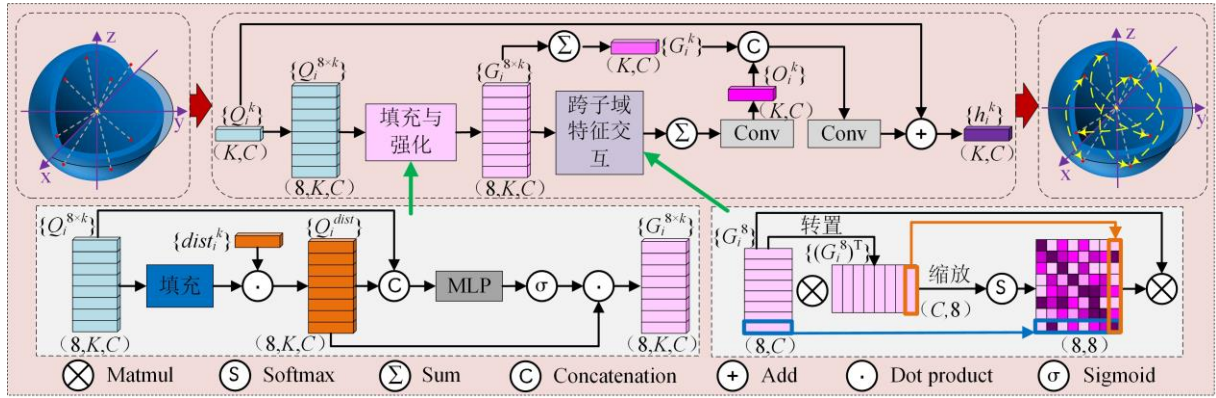


图 2 LSEIA 模块

Fig. 2 LSEIA Module

2.3.1 子域特征填充与强化

如图 2 所示，对采样点 p_i 而言，其局部邻域点的特征集合为 $Q_i^k = \{q_i^1, q_i^2, \dots, q_i^k\} (k=1, 2, \dots, K)$ 。

以各邻域点所在子域作为空间划分的依据，将输入数据的维度扩展为 $N \times 8 \times K \times C$ ，其中，8 对应八个不同的子域。在此基础上，对于第 l 个子域分支 ($l=1, 2, \dots, 8$)，仅保留落在该子域内的邻域点特征，其余点特征暂时置零。考虑到实际点云中的稀疏区域或几何边界处，部分子域内可能不存在邻域点，此时该子域分支特征全部为零，若直接将其输入后续模块，会引发梯度消失的问题。为此，引入一

种子域特征填充策略：当前子域至少存在一个邻域点时，则用该子域内邻域点的平均特征对置零位置进行填充；当前子域完全没有邻域点时，则使用整个邻域的平均特征对置零位置进行填充。填充后特征记为 $Q_i^{8 \times k}$ ，该策略保证每个子域分支均具备有效、非零的特征表示，同时获得了整个局部区域的合理上下文先验。在此基础上，引入指数形式的距离衰减权重，对邻域特征进行加权得到 Q_i^{dist} ：

$$Q_i^{dist} = Q_i^{8 \times k} \odot dist_i^k \quad (6)$$

其中， $dist_i^k = e^{-\|p_i - p_i^k\|}$ 表示第 k 个邻域点相对于中心点的距离衰减权重， $\odot$ 为逐元素乘法。

将 Q_i^{dist} 与 $Q_i^{8 \times k}$ 拼接输入 MLP 生成注意力权重，以此权重与 $Q_i^{8 \times k}$ 加权得到强化的子域特征 $G_i^{8 \times k}$ ：

$$G_i^{8 \times k} = Q_i^{8 \times k} \odot \sigma \left(W_1 \left(\text{ReLU} \left(\text{BN} \left(W_0 \left([Q_i^{dist}, Q_i^{8 \times k}] \right) \right) \right) \right) \right) \quad (7)$$

其中， σ 为 sigmoid 激活函数， W_0 、 W_1 代表 MLP 的权重。

2.3.2 跨子域特征交互

子域划分人为引入了空间隔离，造成各子域间信息相互独立。为恢复子域间的信息流动，设计了跨子域特征交互组件。与传统注意力不同的是，该组件将关注的对象由“点”提升至“子域”，从而在更高层级上实现信息整合。具体表达如下：

设每个邻域点特征为 G_i^8 ，首先经过相关度矩阵计算各子域间特征相关度，其次引入标准缩放因子 $\sqrt{C}$ 稳定训练过程，最后输入 softmax 激活函数得到归一化权重，并以此权重对 G_i^8 加权求和得到跨子域交互特征并输入卷积层进行处理得到 O_i^k ：

$$O_i^k = W \left(\sum_1^8 G_i^8 \times \text{softmax} \left(\left((G_i^8)^T \times G \right) / \sqrt{C} \right) \right) \quad (8)$$

其中， $G_i^k = \sum_1^8 G_i^{8 \times k}$ ， $W \in R^{C \times 2C}$ 代表卷积层权重。

经跨子域特征交互组件后，首先求取 8 个子域特征的总和 G_i^k 并与 O_i^k 拼接，其次输入卷积层融合并与局部邻域点的特征 Q_i^k 相加，得到跨子域交互融合特征 h_i^k ：

$$h_i^k = W \left([G_i^k, O_i^k] \right) + Q_i^k \quad (9)$$

2.4 EGMDA 模块

在基于多层特征聚合的点云处理中，随着感受野的扩大和局部特征的逐步融合，虽然点云的语言表达能力得到增强，但几何边界处的细微变化容易被全局特征所掩盖。基于此，立足于 DGCNN 图卷积思想 (Wang 等, 2019) 和通道-空间注意力机制 (Woo 等, 2018)，设计如图 3 所示的 EGMDA 模块，通过边缘信息强化、多尺度分组特征提取与动态加权融合的协同作用，实现对几何边界的精细刻画。

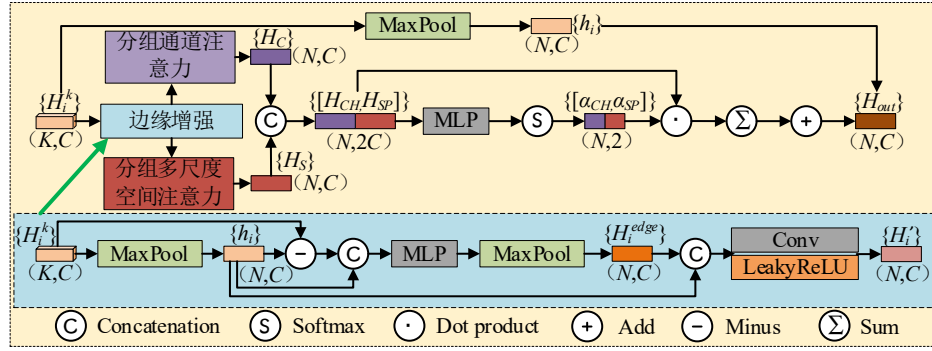


图 3 EGMDA 模块

Fig. 3 EGMDA Module

2.4.1 边缘信息增强

边缘增强策略主要利用邻域特征与采样点特征之间的差值算子进行计算，该算子以特征突变来检测边缘特征。在点云中，边缘区域特征通常偏离邻域统计，因此差分幅度可直接反映边缘特征差异。给定采样点特征 p_i 及其邻域点特征集合 $H_i^k = \{h_i^1, h_i^2, \dots, h_i^k\} \in \mathbb{R}^{N \times K \times C}$ ($k=1, 2, \dots, K$)，通过最大池化提取当前采样点的全局特征 h_i ，并采用式 (10) 提取边缘特征 H_i^{edge} ：

$$H_i^{edge} = \text{MaxPool} \left(W_1 \left(\text{ReLU} \left(\text{BN} \left(W_0 \left(e_i^k \right) \right) \right) \right) \right) \quad (10)$$

其中， $e_i^k = [(h_i^k - h_i), h_i]$ ，MaxPool 表示最大池化操作，ReLU 为激活函数，BN 为批归一化层， W_0 、 W_1 代表 MLP 的权重。

将 H_i^{edge} 与 h_i 拼接并通过 1×1 卷积和 LeakyReLU 激活函数进行处理得到边缘增强特征 H_i' ：

$$H'_i = \text{LeakyReLU}\left(W\left(\left[H_i^{\text{edge}}, h_i\right]\right)\right) \quad (11)$$

其中，LeakyReLU表示激活函数，其斜率 α 设置为0.2。

2.4.2 分组多尺度特征提取与动态融合

为提升特征表达的空间分辨率，将边缘增强后的特征 $H'_i \in \mathbb{R}^{N \times C}$ 沿通道维度均分 G 组（如图4所示，算法设置 G 为4），每组通道数为 C/G ，每组特征记为 $H_i^g \in \mathbb{R}^{N \times (C/G)}$ （ $g=1,2,\dots,G$ ）。随后，对每一组特征分别进行通道注意力与多尺度空间注意力建模，不同于单一尺度的特征提取方式，分组策略与多尺度卷积策略能够覆盖不同感受野的特征，从而更全面地描述空间结构变化。

分组通道注意力即对每组特征 H_i^g 独立进行通道注意力计算，结果记为 H_{CH}^g ，并对其进行拼接得到分组通道特征 H_{CH} ：

$$H_{CH}^g = H_i^g \cdot \sigma\left(W_1^g\left(\text{ReLU}\left(W_0^g\left(\text{AvgPool}\left(H_i^g\right)\right)\right)\right) + W_1^g\left(\text{ReLU}\left(W_0^g\left(\text{MaxPool}\left(H_i^g\right)\right)\right)\right)\right) \quad (12)$$

其中，AvgPool表示平均池化操作， $H_{CH} = [H_{CH}^1, H_{CH}^2, \dots, H_{CH}^G]$ ， W_0^g 、 W_1^g 表示每组MLP的权重。

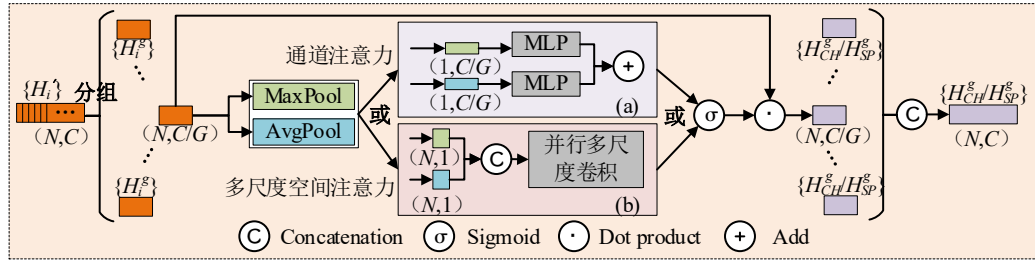


图4 分组多尺度特征提取。(a)通道注意力，(b)多尺度空间注意力。其中 G 为分组的组数（设置为4）；并行多尺度卷积对应的卷积核大小分别为 3×3 、 5×5 、 7×7 。

Fig. 4 Group Multi-scale Feature Extraction. (a)Channel Attention, (b)Multi-scale Spatial Attention. In the figure, G represents the number of group(set to 4); the convolution kernel sizes corresponding to the parallel multi-scale convolution are 3×3 , 5×5 , and 7×7 .

分组多尺度空间注意力即对每组特征 H_i^g 采用并行的 3×3 、 5×5 、 7×7 卷积核分别提取不同的感受野下的局部空间特征，再通过加权均值得到多尺度特征 H_{SP}^g ，并对其进行拼接得到分组多尺度空间特征 H_{SP} ：

$$H_{SP}^g = H_i^g \cdot \sigma \left(\frac{1}{|I|} \sum_{l \in I} f_l \left(\left[\text{AvgPool}(H_i^g), \text{MaxPool}(H_i^g) \right] \right) \right) \quad (13)$$

其中, f_l 代表卷积核为 $l \times l$ 的卷积操作, $I = \{3, 5, 7\}$, $H_{SP} = [H_{SP}^1, H_{SP}^2, \dots, H_{SP}^g]$ 。

为了自适应平衡通道与空间信息的相对重要性, 首先将 H_{CH} 与 H_{SP} 拼接后输入动态融合器得到自适应权重 $[\alpha_{CH}, \alpha_{SP}]$, 其次对两类注意力特征进行加权融合, 最后通过残差连接输出最终特征表示

H_{out} :

$$H_{out} = H_{CH} \cdot \alpha_{CH} + H_{SP} \cdot \alpha_{SP} + h_i \quad (14)$$

其中, $[\alpha_{CH}, \alpha_{SP}] = \text{softmax}(\varphi([H_{CH}, H_{SP}]))$ 。

2.5 损失函数与评价指标

考虑到实验数据集中存在类别分布不均的问题, 定义了一种带有类别权重的交叉熵损失函数:

$$L = - \sum_{u=1}^U W_u \times (y_u \log p_u + (1 - y_u) \log(1 - p_u)) \quad (15)$$

$$W_u = 1 / \ln(\lambda + N_u) / \sum_{u=1}^U N_u \quad (16)$$

其中, u 代表具体类别, U 代表第 u 类的类别总数, W_u 是指第 u 类的平衡权重, y_u 和 p_u 表示第 u 类的真实标签和预测标签为第 u 类的概率, N_u 代表第 u 类的总点数, 平衡系数 λ 的值设置为 1.2。

为全面评估 LSENet 模型的分割性能, 采用混淆矩阵、F1 指数 (F1-score) 以及总体精度 (Overall Accuracy, OA) 作为语义分类的评价指标。F1 指数以及总体精度的定义分别如下:

$$\begin{cases} Precision = TP / (TP + FP) \\ Recall = TP / (TP + FN) \end{cases} \quad \left| \begin{array}{l} F1 = 2 \times Recall \times Precision / (Recall + Precision) \\ OA = TP / (TP + FP + TN + FN) \end{array} \right. \quad (17)$$

其中, $Precision$ 代表精确率、 $Recall$ 代表召回率、 TP 代表正确分类的正样本数, FN 表示错分类为负样本的正样本数, FP 表示错分为正样本的负样本数, TN 表示正确分类的负样本数。

3 实验与分析

3.1 数据与预处理

为验证方法的有效性，在 ISPRS Vaihingen 3D 与 DFC2019 两个公开机载 LiDAR 点云数据集上分别进行对比实验评估模型性能，并基于 ISPRS Vaihingen 3D 数据集开展了消融实验评估各模块的贡献。

3.1.1 ISPRS Vaihingen 3D 数据集

该数据由国际摄影测量与遥感学会（ISPRS）提供。数据集涵盖九个语义类别：电力线、低矮植被、不透水表面、汽车、栅栏/树篱、屋顶、房屋立面、灌木和树，每个点包含坐标、RGB、强度和回波次数八种特征信息，广泛应用于点云语义分割任务研究。

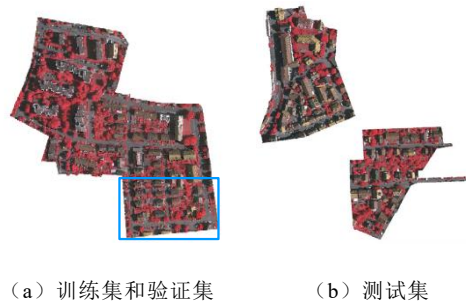


图 5 ISPRS Vaihingen 3D 数据集

Fig. 5 ISPRS Vaihingen 3D Dataset

通过官方提供的 IR-R-G 正射影像采用双线性插值将 RGB 信息映射至对应点，以坐标 $Y=5419292.0$ 为分割线，将原始训练数据（图 5(a)所示）拆分为训练集（蓝色框外部）和验证集（蓝色框内部）两个部分，图 5(b)为测试集。训练阶段数据采用块状切分策略，每块尺寸设定为 $30m \times 30m$ ，X 和 Y 方向上相邻区块重叠 $15m$ 以减弱边界效应并固定采样 4096 个点。针对点密度不均的问题，采用随机采样与替换采样相结合的方式；当区块点数不足 1024 时直接舍弃。块内点坐标采用中心化方法映射至 $[-1,1]$ ，其他属性归一化至 $[0,1]$ 。同时，引入绕 Z 轴的随机旋转以增强模型对视角变化的鲁棒性。测试阶段采用与训练阶段完全一致的数据预处理方法，以确保数据特征分布的一致性。实验仅采用坐标、RGB 和强度共七种特征作为网络输入，未使用回波次数信息。为降低随机

采样带来的预测不稳定性，对每个点进行 10 次独立预测，并通过投票策略确定最终类别标签。

3.1.2 DFC2019 数据集

该数据集由 IEEE 地球科学与遥感学会（GRSS）发布。数据集涵盖六个语义类别：地面、高植被、建筑物、水体、桥面以及未分类。每个点包含坐标、强度和回波次数五种特征信息。为了确保公平，采用 IEEE GRSS 提供的数据预处理方法，首先使用 Point Data Abstraction Library(PDAL)工具将原始点云场景进行空间裁剪并生成具有重叠区域的子块；随后以 1m 分辨率体素下采样原始点云，在降低点云密度的同时尽可能保持原始几何结构；最后将点云子块进一步切分为若干非重叠小块，每个小块最多包含 65536 个点。在训练阶段从每个小块随机采样 4096 个点输入 LSENet。同时，引入绕 Z 轴的随机旋转以增强模型对视角变化的鲁棒性。测试阶段采用与训练阶段一致的数据预处理策略。

3.2 实验环境与参数设置

表 1 展示了训练阶段超参数配置，实验基于 Windows11 操作系统，在 PyTorch2.3.1 深度学习框架下实现，模型训练与评估均在 NVIDIA TITAN RTX 显卡上完成。

表 1 实验超参数设置

Table 1 Experimental Hyperparameter Settings

超参数	参数优化器	初始学习率	批处理大小	轮次	衰减率	权重衰减	Dropout	邻域点数
值	Adam	0.005	6	300	0.95	1×10^{-4}	0.5	32

3.3 实验结果与分析

为确保对比实验的公平性和可比性，在两个数据集上均优先选择已在相应数据集上公开发表实验结果、采用相同或高度一致的评价指标、且在点云语义分割领域具有代表性的主流方法作为对比对象。由于 ISPRS Vaihingen 3D 和 DFC2019 两个数据集在场景复杂度、类别分布及数据特性上存在一定差异，因此在尽可能保留共同基线方法的基础上，分别补充了部分在相应数据集上已有公开结果的代表性方法，以更全面地评估 LSENet 的性能与泛化能力。

3.3.1 ISPRS Vaihingen 3D 数据集

为系统评估 LSENet 的性能，将其与 PointNet++ (Qi 等, 2017)、PointSIFT (Jiang 等, 2018)、KPCnv (Thomas 等, 2019)、DPE (Huang 等, 2020)、VD-LAB (Li 等, 2022)、MIA-Net (Wang 等, 2024)、HFA-Net (Liu 等, 2025) 和 DACNN (Pan 等, 2025) 八种深度学习模型在 Vaihingen 3D 数据集上进行对比实验，定量结果如表 2 所示。

表 2 LSENet 与其他深度学习模型在 Vaihingen 3D 数据集上的定量结果 (%)

Table 2 Quantitative Results of LSENet and Other Deep Learning Models on Vaihingen 3D Dataset(%)

模型	电力线	低矮植被	不透水表面	汽车	栅栏/树篱	屋顶	房屋立面	灌木	树	Avg.F1	OA
PointNet++	57.9	79.6	90.6	66.1	31.5	91.6	54.3	41.6	77.0	65.6	81.2
PointSIFT	55.7	80.7	90.9	77.8	30.5	92.5	56.9	44.4	79.6	67.7	82.2
KPCnv	63.1	82.3	91.4	72.5	25.2	94.4	60.3	44.9	81.2	68.4	83.7
DPE	68.1	86.5	99.3	75.2	19.5	91.1	44.2	39.4	72.6	66.2	83.2
VD-LAB	69.3	80.5	90.4	79.4	38.3	89.5	59.7	45.7	77.2	70.2	81.4
MIA-Net	65.8	79.5	89.7	71.1	26.2	94.0	63.5	48.1	82.8	69.0	83.3
HFA-Net	68.6	79.9	90.5	79.8	39.1	93.3	60.3	45.1	80.8	70.8	82.8
DACNN	78.3	82.1	92.9	78.1	39.7	92.8	59.8	46.1	78.7	72.1	83.7
Ours	70.7	81.7	90.1	71.7	36.7	94.5	61.9	44.0	83.3	70.5	83.9

表格中各类别下方数值对应模型在该类别上的 F1 指数，最后两列分别为模型的平均 F1 指数与总体精度 OA，其中加粗为最优结果。从各类别结果来看，LSENet 在电力线、屋顶和树木类别上的表现突出，具备显著的分割优势。其中电力线作为模型改进最明显的类别，其 F1 指数达到 70.7%，虽然较最优模型低 7.6%，但是较 Baseline 模型 (PointNet++) 提高 12.8%。该类别通常呈现细长结构且点云分布稀疏，对局部几何建模能力要求较高，LSSEA 与 LSEIA 模块能够有效挖掘无序稀疏点云的空间走向与几何方向特征，精准捕捉电力线狭长的拓扑结构并强化离散点云之间的关联；EGMDA 模块能够增强细长结构的边缘特征，并通过分组多尺度机制解耦其细粒度特征，从而有效提高了其分割精度。模型在树木类别的 F1 指数达到 83.3%，其点云呈现簇状不规则分布的特点，LSSEA 与 LSEIA 模块能够完成其不规则几何特征的全方位提取，EGMDA 模块则针对植被边缘特征进行强化，通过分组通道与多尺度空间注意力自适应聚焦树木的有效特征区域，使模型充分捕捉复杂自然地物特征。针对屋顶类别，LSENet 以 94.5% 的 F1 指数取得最优结果，LSSEA 与 LSEIA 可初步提取屋顶平面几何特征，EGMDA 模块的边缘增强机制则可辅助优化屋顶结构边界。从整体结果

来看, LSENet 在平均 F1 指数和总体精度上均保持较高水准, 总体精度较次优模型高 0.2%, 平均 F1 指数达到 70.5%, 综合性能处于对比模型的前列。表明所提出的局部子域感知与双注意力融合机制能够有效提升特征表达能力。

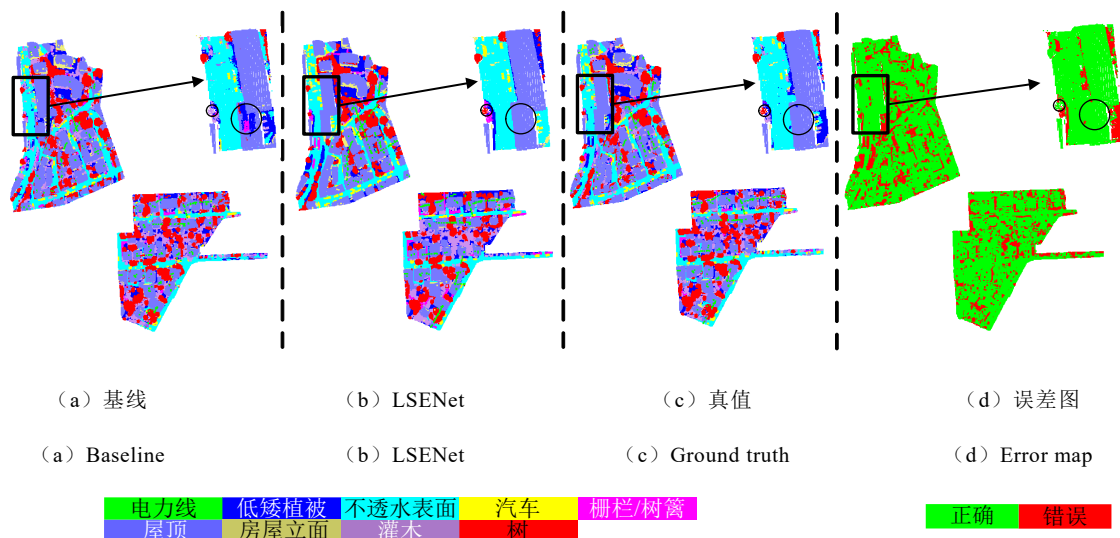


图 6 Vaihingen 3D 数据集分割可视化结果。其中黑色圆圈标注了 LSENet 相比于基线模型改进最为显著的区域, 这些区域包括屋顶以及树木与灌木的过渡地带。

Fig. 6 Vaihingen 3D Dataset Segmentation Visualization Results. The black circles indicate the areas where LSENet has shown the most significant improvement compared to the baseline model. These areas include the roof and the transition zones between trees and shrubs.

图 6 展示了以 Pointnet++为基线模型与 LSENet 在 Vaihingen 3D 数据集上的分割可视化结果。相较于基线模型, LSENet 在屋顶边界以及树木与灌木的过渡地带的预测结果更加连续且边界清晰, 有效改善了基线模型在过渡区域混淆严重的问题。图 7 则展示了分割结果的混淆矩阵, 从整体来看, 大多数类别在对角线位置具有较高数值, 说明模型具备良好的分割能力, 其中屋顶、不透水表面和树木类别分割效果最稳定。同时, 需要指出的是模型仍存在一定短板, 在栅栏/树篱、灌木类别上分割效果有待提升, 这两类地物与低矮植被、树木的空间分布、几何形态高度相似, 细粒度特征差异极小, 容易发生类别混淆。上述结果表明, LSENet 更适用于具有明显局部几何结构特征与边界信息的类别, 所提出的模块能够有效提升模型对机载 LiDAR 点云的语义理解能力。

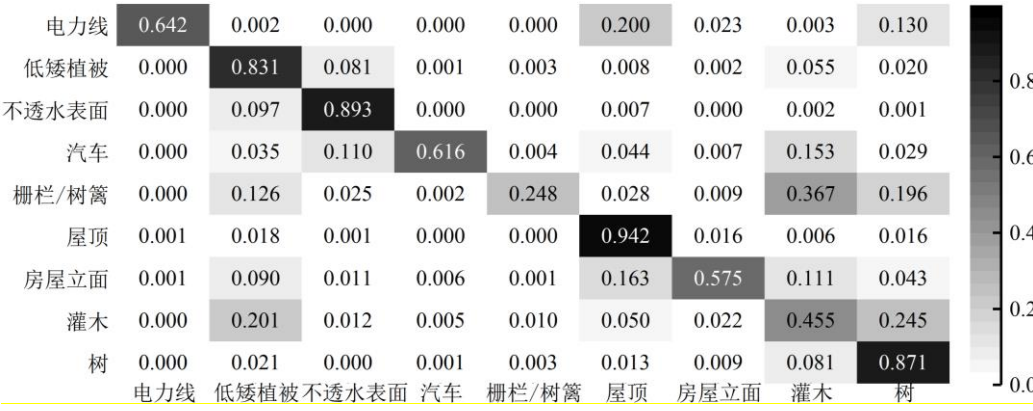


图 7 Vaihingen 3D 数据集分割混淆矩阵

Fig. 7 Vaihingen 3D Dataset Segmentation Confusion matrix

3.3.2 DFC2019 数据集

为进一步验证 LSENet 的泛化能力, 在 DFC2019 数据集上展开对比实验, 并选取 PointNet++ (Qi 等, 2017)、PointSIFT (Jiang 等, 2018)、KPConv (Thomas 等, 2019)、VD-LAB (Li 等, 2022)、GACNN (Wen 等, 2021)、RFFS-Net (Mao 等, 2022)、DAKAG-Net (Zhao 等, 2024) 和 U-agent (Sun 等, 2026) 作为对比方法, 定量结果如表 3 所示, 模型混淆矩阵与场景分割可视化结果分别如图 8、图 9 所示。

表 3 LSENet 与其他深度学习模型在 DFC2019 数据集上的定量结果 (%)

模型	地面	高植被	建筑物	水体	桥面	Avg.F1	OA
PointNet++	98.3	95.8	79.7	4.4	7.3	57.1	92.7
PointSIFT	98.6	97.0	85.5	46.4	60.4	77.6	94.0
KPConv	98.4	94.2	87.4	43.0	77.5	80.1	94.5
GACNN	99.3	96.8	91.1	42.5	84.4	82.8	95.1
VD-LAB	98	95.5	92.1	85.5	71.5	88.5	95.9
RFFS-Net	96.6	96.1	88.7	77.8	81	88	94.3
DAKAG-Net	98.9	79.9	96	83.5	86.8	92.4	97.5
U-agent	97.5	95.6	89.8	93.3	79.8	91.2	97.3
Ours	99.3	98.2	95.7	97.6	50.1	89.1	97.8

LSENet 在 DFC2019 数据集上展现出良好的泛化能力, 从整体指标看, LSENet 的总体精度达到

97.8%，为所有方法中最高值，验证了所提模型在不同场景下的良好的泛化能力。从类别表现看，LSENet 在地面、高植被和水体类别上均取得最优性能。其中，地面与高植被类别具有较明显的高程分布特征与空间连续性，因此模型能够较容易学习其稳定的几何模式，并在两类上均取得最优结果。值得一提的是，Vaihingen 3D 数据集中的树木与 DFC2019 数据集中的高植被属于同类自然地物，LSENet 在两个数据集中均取得最优结果，充分证明了网络针对此类地物特征的挖掘能力具备良好的稳定性与通用性。对于水体类别，其 F1 分数达到 97.6%，较次优结果提升了 4.3%，具有显著提升，这主要得益于 LSEIA 模块对上下文信息的有效建模，使其能够充分利用水体在空间上的连续性特征，从而提升分类准确性。建筑物类别则受益于 EGMDA 模块对边界强化与多尺度特征的动态融合能力，使模型能够更准确地区分建筑物边缘。然而，模型在桥面类别上表现相对较弱，F1 分数仅为 50.1%，显著低于 DAKAG-Net 网络的 86.8%和 GACNN 网络的 84.4%。结合图 8 所示混淆矩阵进一步分析可知，桥面点云仅有 46.2%的样本被正确分类，其余 53.8%被错误分类至其他类别。其中误分类至建筑物的比例最高，达到 20.7%，其次是地面占 15.9%，水体占 8.9%，高植被占 8.3%。

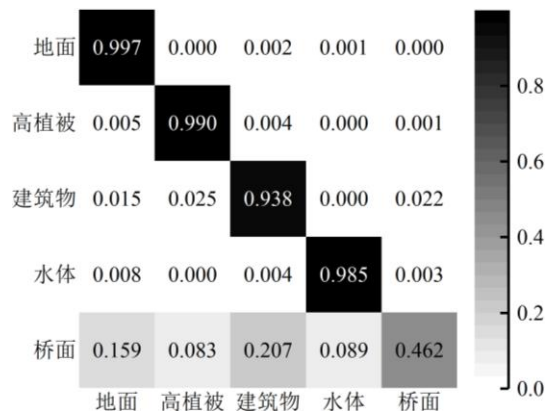


图 8 DFC2019 数据集分割混淆矩阵

Fig. 8 DFC2019 Dataset Segmentation Confusion matrix

结合图 9 所示可视化结果可知，桥面点云的误分类呈现明显的空间特性：邻近建筑物、跨越水面的桥面区域、邻近高植被以及大面积地面相连的过渡区域，大多被误预测为建筑物；而在与大面积地面连接的过渡区域，少量样本则被预测为地面。这一误分类特性主要源于桥面类别样本数量相对较少，且在高程、平面结构、与点云强度特征上与建筑物及地面具有较高相似性，因此容易产生

类别混淆。尽管存在上述局限，LSENet 整体平均 F1 分数仍然达到 89.1%，整体性能处在所有对比模型的前列，综合定量与可视化结果表明，LSENet 不仅能够有效学习规则结构样本的空间连续性特征，同时在不同场景中仍具备较好的泛化能力，但对于小样本且易混淆的类别仍存在进一步优化空间。

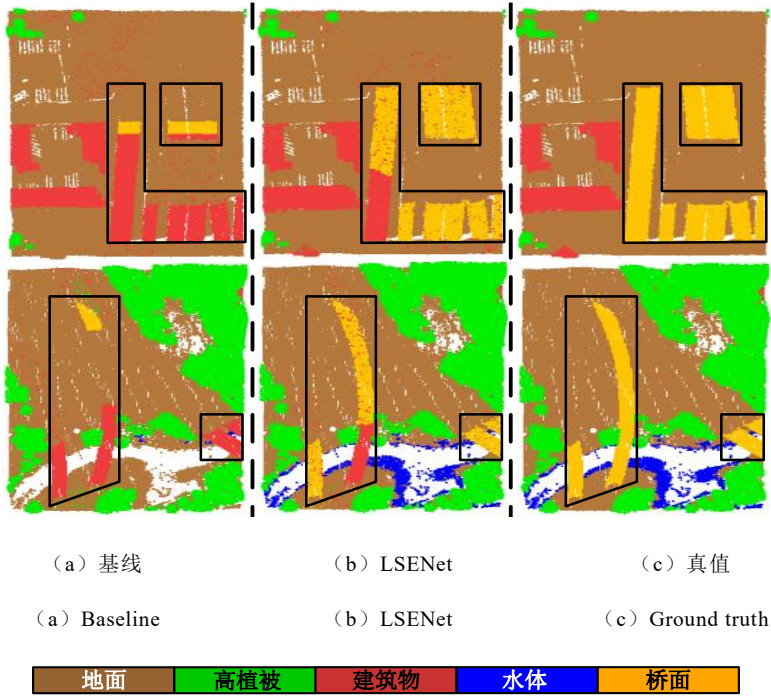


图 9 DFC2019 数据集分割可视化结果。其中黑色线框标注了 LSENet 相比于基线模型改进最为显著的区域，这些区域包括邻近建筑物、跨越水面的桥面区域、邻近高植被区域以及大面积地面相连的过渡区域。

Fig. 9 DFC2019 Dataset Segmentation Visualization Results. The black box marks the areas where LSENet has shown the most significant improvement compared to the baseline model. These areas include regions adjacent to buildings, bridge surfaces crossing water, areas adjacent to high vegetation, and transition zones connected to large ground regions.

3.4 消融实验

3.4.1 关键模块的消融

为分析 LSENet 各关键模块的实际贡献，基于 Vaihingen 3D 数据集并采用统一的 Baseline 架构，逐步引入 LSSEA、LSEIA 与 EGMDA 模块开展消融实验，相关结果如表 4 所示。结果表明，各模块均能不同程度上提升模型性能。仅引入 LSSEA 后，OA 和平均 F1 分数分别提升了 1.3%和 2.0%，说

明增强局部几何结构特征提取对点云语义分割具有积极作用。在此基础上叠加 LSEIA 模块，OA 和平均 F1 分数的提升幅度进一步扩大至 2.3%和 2.7%，表明子域间特征交互能够有效增强上下文建模能力。同时，引入 EGMDA 模块同样带来稳定提升，验证了边缘增强与多尺度注意力融合在优化特征表达方面的有效性。完整模型集成三种模块后取得最优结果，OA 达到 83.9%，平均 F1 分数达到 70.5%。最终结果表明，各模块在功能上具有良好的互补性，整体性能提升是局部几何建模、上下文交互与多尺度特征融合的协同作用的结果。

表 4 LSENet 消融实验 (%)

Table 4 Ablation Experimental Results for LSENet(%)

模型	LSSEA	LSEIA	EGMDA	OA	Avg.F1
Baseline				81.2	65.6
Net1	√			82.5	67.6
Net2	√	√		83.5	68.3
Net3	√		√	82.7	68.2
Net4		√	√	83.6	68.8
LSENet	√	√	√	83.9	70.5

3.4.2 模块内部组件的消融

为验证 EGMDA 边缘信息增强策略以及 LSEIA 子域特征填充策略的有效性，分别移除上述组件在 Vaihingen 3D 数据集上进行对比实验，实验结果如表 5 所示。

表 5 移除部分组件的定量结果 (%)

Table 5 Quantitative Results of removing some components(%)

网络	电力线	低矮植被	不透水表面	汽车	栅栏/树篱	屋顶	房屋立面	灌木	树	Avg.F1	OA
无边缘增强	64.9	82.0	89.6	71.9	31.4	94.3	61.5	43.2	82.2	69.0	83.8
无特征填充	69.4	82.1	90.4	71.4	32.1	94.2	61.2	42.8	82.0	69.5	83.7
完整网络	70.7	81.7	90.1	71.7	36.7	94.5	61.9	44.0	83.3	70.5	83.9

实验结果表明去除 EGMDA 边缘增强策略后，模型平均 F1 下降 1.5%，电力线 F1 从 70.7%降至 64.9%，树木和栅栏等边界地物精度也同步降低。结合图 10 的特征可视化热图可知，树木边缘以及电力线呈现深红色，是该策略高关注的核心区域；其次是各类地物的过渡边缘区域呈现黄色，属于

次重点关注区域，这一结果充分证明了边缘增强是细化复杂地物轮廓以及提升边界分割精度的重要辅助手段。去除 LSEIA 子域特征填充策略后，平均 F1 降至 69.5%。子域特征填充策略能够有效完善子域特征表达、弥补点云特征缺失，缺失该策略后，LSEIA 信息补全能力弱化，模型整体特征表征完整性下降，无法充分挖掘各类地物的空间关联特征，最终导致整体分割精度出现衰减。实验结果表明，EGMDA 的分组机制是电力线等细粒度目标性能提升的核心驱动因素，而边缘增强策略能够提供边界优化辅助；LSEIA 的子域特征填充策略则保障了局部特征的完整性。

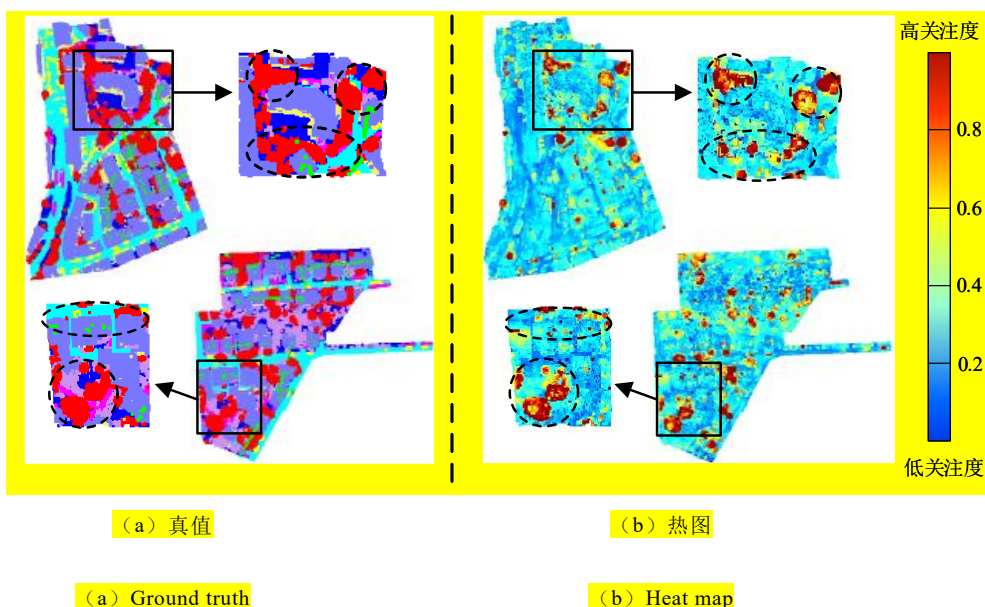


图 10 边缘增强后的特征可视化热图。图中蓝色为低关注度的点云，红色为高关注度的点云。

Fig. 10 Visualization heatmap of the enhanced features. Blue represents the points with low attention, red represents the points with high attention.

3.5 分组数量 G 的敏感度分析

EGMDA 模块的分组数量 G 决定了通道与空间特征的建模粒度，是影响模块特征学习能力的核心超参数。为探究 G 值对 LSENet 分割性能的影响，在 Vaihingen 3D 数据集上设置 $G=1$ 、2、4、8 四组对比试验，实验结果如表 6 所示。实验结果表明， G 值对模型性能影响显著，尤其是对电力线这类细粒度线状目标的影响最强。当 $G=1$ 或 $G=2$ 时，模型无法实现充分的特征解耦，电力线 F1 分数仅为 61.7% 和 60.7%。当 $G=4$ 时，模型达到整体最优性能，电力线 F1 提升至 70.7%，树木等类

别也同步取得最佳结果。这是因为 G=4 的配置能够将通道与空间特征进行精细化拆分，精准适配电力线、栅栏/树篱的狭长拓扑和植被的多尺度簇状结构。当 G 增大至 8 时，性能出现小幅下降，其主要原因可能是特征过于碎片化，破坏了地物整体特征的连续性与关联性。

表 6 不同 G 值的定量结果 (%)

Table 6 Quantitative Results of different G values(%)

G 值	电力线	低矮植被	不透水表面	汽车	栅栏/树篱	屋顶	房屋立面	灌木	树	Avg.F1	OA
G=1	61.7	81.4	90.4	70.8	29.6	94.3	62.7	40.2	82.3	68.2	83.4
G=2	60.7	82.1	90.7	68.6	32.2	94.3	60.7	41.7	82.5	68.2	83.8
G=4	70.7	81.7	90.1	71.7	36.7	94.5	61.9	44.0	83.3	70.5	83.9
G=8	60.6	81.9	90.4	75.2	31.9	94.1	62.1	42.0	82.0	68.9	83.6

3.6 网络复杂度分析

为了全面评估 LSENet 的计算效率与模型复杂度，基于表 7 所示的参数量、浮点运算量 (FLOPs) 以及单块推理时间三个指标，将 LSENet 与 PointNet++和 PointSIFT 进行对比分析。

表 7 网络模型的复杂度对比

Table 7 Complexity comparison of network models

网络	参数量/M	FLOPs/G	单块推理时间/ms
PointNet++	0.97	1.01	739.55
PointSIFT	13.56	13.70	794.32
LSENet	6.09	5.57	775.80

综合表 2 和表 7 可以看出，LSENet 的参数量为 6.09M，FLOPs 为 5.57G，单块推理时间为 775.8ms。与 PointNet++相比，LSENet 的参数量和计算量有所增加，但在分割性能上实现了显著提升。与 PointSIFT 相比，LSENet 的参数量仅为 44.9%，FLOPs 仅为 40.7%，推理时间也更低，然而，平均 F1 分数提升 2.8%，总体精度提升 1.7%。这一结果表明模型在性能与复杂度之间取得了良好的平衡。

4 结论

针对现有深度学习网络在点云语义分割中难以有效地捕捉目标实体的局部几何特征和点云的

上下文信息问题，提出了一种新的点云语义分割网络 LSENet，通过局部子域划分、子域特征交互、边缘增强与注意力动态融合机制，实现了对采样点邻域几何结构、局部上下文关系及边界信息的联合建模。具体而言，LSSEA 模块基于 PCA 的构建局部参考系，实现具备空间方向感知的几何特征编码；LSEIA 模块通过子域级特征交互增强了局部上下文建模能力；EGMDA 模块通过边缘增强与分组多尺度注意力的动态融合，提升了特征判别能力，并在一定程度上缓解了语义边界模糊问题。

基于 ISPRS Vaihingen 3D 与 DFC2019 两个公开机载 LiDAR 数据集的实验结果表明，LSENet 在总体精度与平均 F1 分数等关键指标上均取得了较优性能，验证了所提网络在标准公开基准数据集上的有效性。同时，本文模型仍存在一定局限：对于 DFC2019 数据集中桥面这类几何与强度特征与周边地物高度相似的易混淆小样本类别，模型特征区分能力不足，分割精度有待进一步提升。同时，本文实验主要基于两个经典公开数据集开展，在数据时效性、场景规模和地物类型丰富性方面仍存在拓展空间，尚需进一步验证模型在更多样化机载 LiDAR 点云场景中的适用性。

后续研究将从以下几个方面进一步展开：首先，设计更具判别力的局部几何描述符，以增强网络对复杂曲面、低密度区域以及小样本、高相似性易混淆地物的建模能力，提升模型对稀缺难分割地物的识别性能；其次，引入更大规模、多场景和更复杂地物类型的机载 LiDAR 点云数据集，系统评估方法的跨场景适用性和鲁棒性；最后，探索轻量化模块设计与动态推理策略，以降低计算开销，提升模型在大规模点云场景中的实际应用潜力。

参考文献(References)

- Chen X S, Huang L J, Ma T, Fang R Y, Shi S, Li H S. 2025. SOLVE: Synergy of Language-Vision and End-to-End Networks for Autonomous Driving. Paper presented at the 2025 Conference on Computer Vision and Pattern Recognition-CVPR-Annual. Nashville: IEEE: 12068-12077 [DOI:10.1109/cvpr52734.2025.01127]
- Chen Y M, Xia J K, Zou X C, Xiao Z T, Tang X, Zang Y F, Chen D, Shen Y Q. 2025. Contextual boundary-aware network for semantic segmentation of complex land transportation point cloud scenes. ISPRS Journal of Photogrammetry and Remote Sensing, 230: 18-31.[DOI:10.1016/j.isprsjprs.2025.09.006]
- Hu B N, Chen L, Xu B L, Bu S H, Han P C, Li K, Xia Z Y, Li N, Li K, Cao X F and Wan G. 2024. Real-time dense point cloud generation and digital model construction of surface environment based on UAV platform. National Remote Sensing Bulletin, 28(05): 1206-1221. (胡博妮, 陈霖, 徐丙立, 布树辉, 韩鹏程, 李坤, 夏震宇, 李霓, 李科, 曹雪峰, 万刚. 2024. 基于无人机平台的地表环境实时稠密点云生成与数字模型构建. 遥感学报, 28(05): 1206-1221.) [DOI:10.11834/jrs.20232597]

Huang R, Xu Y S, Hong D F, Yao W, Ghamisi P, Stilla U. 2020. Deep point embedding for urban classification using ALS point clouds: A new perspective from local to global. *ISPRS Journal of Photogrammetry and Remote Sensing*, 163: 62-81.[DOI:10.1016/j.isprsjprs.2020.02.020]

Hui Z Y, Hu Y J, Jin S G, Yevenyo Y Z. 2016. Road centerline extraction from airborne LiDAR point cloud based on hierarchical fusion and optimization. *ISPRS Journal of Photogrammetry and Remote Sensing*, 118: 22-36.[DOI:10.1016/j.isprsjprs.2016.04.003]

Jiang M, Wu Y, Zhao Z, Lu C. 2018. PointSIFT: A SIFT-like Network Module for 3D Point Cloud Semantic Segmentation. *arXiv:1807.00652*: [DOI:10.48550/arXiv.1807.00652]

Kim H B, Sohn G. 2011. RANDOM FORESTS BASED MULTIPLE CLASSIFIER SYSTEM FOR POWER-LINE SCENE CLASSIFICATION. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII-5-W12: 253-258.[DOI:10.5194/isprsarchives-XXXVIII-5-W12-253-2011]

Lecun Y, Bottou L, Bengio Y, Haffner P. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11): 2278-2324.[DOI:10.1109/5.726791]

Li J H, Weinmann M, Sun X, Diao W H, Feng Y C, Hinz S, Fu K. 2022. VD-LAB: A view-decoupled network with local-global aggregation bridge for airborne laser scanning point cloud classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 186: 19-33.[DOI:10.1016/j.isprsjprs.2022.01.012]

Liu K Q, Yuan C H, Han J W, Li W. 2025. A Dual-Branch Network Leveraging Heterogeneous Features for Semantic Segmentation of Large-Scale Point Clouds. *IEEE Transactions on Geoscience and Remote Sensing*, 63: 5706813.[DOI:10.1109/tgrs.2025.3641946]

Mao Y Q, Chen K Q, Diao W H, Sun X, Lu X N, Fu K, Weinmann M. 2022. Beyond single receptive field: A receptive field fusion-and-stratification network for airborne laser scanning point cloud classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 188: 45-61.[DOI:10.1016/j.isprsjprs.2022.03.019]

Maturana D, Scherer S, Ieee. 2015. VoxNet: A 3D Convolutional Neural Network for Real-Time Object Recognition. Paper presented at the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Hamburg, GERMANY: 922-928 [DOI:10.1109/IROS.2015.7353481]

Nguyen K, Dang T, Huber M. 2024. Real-time 3D Semantic Scene Perception for Egocentric Robots with Binocular Vision: *arXiv:2402.11872*[DOI:10.48550/arXiv.2402.11872]

Pan Y, J L, Qiu X, Li S. 2025. Research on AI Semantic Segmentation Model for Large-Scale Point Cloud Data Based on Cloud-Edge Collaboration. Paper presented at the 2025 IEEE 8th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC). 973-977 [DOI:10.1109/IAEAC65194.2025.11166233]

Prathiba S B, Kumar S K R, Anandhan D K, Kumar A S S, Selvaraj A K, Rodrigues J. 2025. Enhanced Surface Reconstruction and Semantic Segmentation of LiDAR Data in Autonomous Vehicle Perception Systems. *IEEE Transactions on Intelligent Transportation Systems*, 26(6): 7833-7842.[DOI:10.1109/tits.2025.3555367]

Qi C R, Su H, Mo K C, Guibas L J. 2017. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. Paper presented at the 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE: 77-85

[DOI:10.1109/cvpr.2017.16]

Qi C R, Yi L, Su H, Guibas L J. 2017. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. arXiv:1706.02413 [DOI:10.48550/arXiv.1706.02413]

Ronneberger O, Fischer P, Brox T. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. Paper presented at the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). Munich: 234-241 [DOI:10.1007/978-3-319-24574-4_28]

Schmohl S, Sörgel U. 2019. SUBMANIFOLD SPARSE CONVOLUTIONAL NETWORKS FOR SEMANTIC SEGMENTATION OF LARGE-SCALE ALS POINT CLOUDS. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., IV-2/W5: 77-84.[DOI:10.5194/isprs-annals-IV-2-W5-77-2019]

Sun Q, Li Y, Xiong Z, Zhang Z. 2026. U-Agent: A Point Cloud Segmentation Network Based on Local Feature Enhancement and Agent Attention for Global Modeling. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 19: 13521-13536.[DOI:10.1109/JSTARS.2026.3682857]

Thomas H, Qi C R, Deschaud J E, Marcotegui B, Goulette F, Guibas L J, Ieee. 2019. KPConv: Flexible and Deformable Convolution for Point Clouds. Paper presented at the IEEE/CVF International Conference on Computer Vision (ICCV). Seoul: 6420-6429 [DOI:10.1109/iccv.2019.00651]

Wang Y, Sun Y B, Liu Z W, Sarma S E, Bronstein M M, Solomon J M. 2019. Dynamic Graph CNN for Learning on Point Clouds. Acm Transactions on Graphics, 38(5): 1-12.[DOI:10.1145/3326362]

Wang Z Y, Chen H, Liu J, Qin J R, Sheng Y H, Yang L. 2024. Multilevel intuitive attention neural network for airborne LiDAR point cloud semantic segmentation. International Journal of Applied Earth Observation and Geoinformation, 132: 104020.[DOI:10.1016/j.jag.2024.104020]

Wen C C, Li X, Yao X J, Peng L, Chi T H. 2021. Airborne LiDAR point cloud classification with global-local graph attention convolution neural network. ISPRS Journal of Photogrammetry and Remote Sensing, 173: 181-194.[DOI:10.1016/j.isprsjprs.2021.01.007]

Woo S H, Park J, Lee J Y, Kweon I S. 2018. CBAM: Convolutional Block Attention Module. Paper presented at the 15th European Conference on Computer Vision (ECCV). Munich: 3-19 [DOI:10.1007/978-3-030-01234-2_1]

Xie D H, Li Y, Zhou K, Zhang Z X, Jin L, Yan G J, Mu X H, Li W H and Feng G Y. 2024. Information extraction and system construction of digital forest system based on multisource remote sensing data. National Remote Sensing Bulletin, 28(05): 1281-1294. (谢东辉, 李益, 周坤, 张智祥, 金霖, 阎广建, 穆西晗, 李文航, 冯功耀. 2024. 基于多源遥感数据的数字林场信息提取与系统构建. 遥感学报, 28(05): 1281-1294.) [DOI:10.11834/jrs.20242655]

Zhang K, Ye L, Xiao W, Sheng Y, Zhang S, Tao X, Zhou Y. 2022. A Dual Attention Neural Network for Airborne LiDAR Point Cloud Semantic Segmentation. IEEE Transactions on Geoscience and Remote Sensing, 60: 1-17.[DOI:10.1109/TGRS.2022.3201902]

Zhao J B, Zhou H Y, Pan F F. 2024. A Dual Attention KPConv Network Combined With Attention Gates for Semantic Segmentation of ALS Point Clouds. IEEE Transactions on Geoscience and Remote Sensing, 62: 5107914.[DOI:10.1109/tgrs.2024.3422829]

Zhao R B, Pang M Y, Wang J D. 2018. Classifying airborne LiDAR point clouds via deep features learned by a multi-scale convolutional neural network. *International Journal of Geographical Information Science*, 32(5): 960-979.[DOI:10.1080/13658816.2018.1431840]

LSENet: Local Subdomain-aware and Edge-enhance Dynamic Fusion Network for Semantic Segmentation of Airborne LiDAR Point Cloud

JI Dali^{1,2}, WANG Yongbo^{1,2}, ZHENG Nanshan^{1,2}, YANG Min^{1,2}

1. *Key Laboratory of Land Environment and Disaster Monitoring, Ministry of Natural Resources, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China;*

2. *Jiangsu Key Laboratory of Resources and Environmental Information Engineering, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China*

Abstract: [Objective] Airborne LiDAR point cloud semantic segmentation is fundamental to 3D scene understanding and has important applications in smart cities, urban modeling, infrastructure extraction, land surveying, and environmental monitoring. To address the insufficient coordination between local geometric representation and contextual modeling in existing methods, we propose LSENet (Local Subdomain-aware and Edge-enhanced Dynamic Fusion Network). LSENet extracts fine-grained geometric features through adaptive local subdomain partitioning, enhances information exchange via cross-subdomain interactions, and improves feature representation and boundary discrimination through an edge-enhanced dual-attention fusion module. Experiments on the Vaihingen 3D and DFC2019 datasets achieve Overall Accuracy (OA) values of 83.9% and 97.8%, and mean F1-scores of 70.5% and 89.1%, demonstrating its effectiveness and cross-scene applicability. [Method] LSENet adopts a U-Net-style encoder-decoder architecture with three core modules. The Local Subdomain Space-aware feature Encoding Aggregation module (LSSEA) constructs an adaptive local coordinate system using PCA and encodes relative po-

sition, spatial distribution, orientation, and curvature information for fine-grained geometric representation. The Local Subdomain feature Enhancement and Interaction Attention module (LSEIA) enhances intra-subdomain features and establishes cross-subdomain interactions to improve local contextual modeling. The Edge-enhanced Grouped Multi-scale Dual-Attention dynamic fusion module (EGMDA) integrates edge-enhanced grouped channel attention and multi-scale spatial attention through dynamic fusion, strengthening boundary representation and classification accuracy. [Result] Comparative experiments were conducted on two widely used benchmark datasets, ISPRS Vaihingen 3D and DFC2019. On Vaihingen 3D, LSENet achieved an OA of 83.9% and a mean F1-score of 70.5%, obtaining the best OA among compared methods. Significant improvements were observed in challenging categories such as power lines, roofs, and trees, demonstrating the effectiveness of the proposed modules in preserving complex structures and clear boundaries. On DFC2019, LSENet achieved an OA of 97.8% and a mean F1-score of 89.1%, ranking first in OA and demonstrating strong cross-scene generalization. Visualization results further show that LSENet produces more coherent regions and clearer boundaries than the baseline. [Conclusion] LSENet effectively integrates local geometric awareness, cross-subdomain interaction, and edge-preserving dual attention with dynamic fusion to overcome limitations of existing point cloud segmentation networks. Results on two airborne LiDAR benchmarks validate its effectiveness, robustness, and generalization capability. Future work will focus on developing more discriminative geometric descriptors, lightweight dynamic inference strategies, and evaluating the proposed method on newly released large-scale and diverse airborne LiDAR datasets to improve its applicability in real-world scenarios.

Key words: airborne LiDAR point clouds; deep learning; semantic segmentation; local subdomain awareness; dynamic fusion; edge-enhance

Supported by National Natural Science Foundation of China(U22A20569); Natural Science Foundation of Jiangsu Province (BK20221549)