

文章编号:1001-5078(2025)01-0040-06

· 激光应用技术 ·

基于特征拓展及分离池化的点云语义分割

陈崇明¹, 郁金星², 韩璐³, 王浩然³, 张佃卯⁴, 陈器⁴

(1. 国网河北省电力有限公司电力科学研究院, 河北 石家庄 050021; 2. 国网河北能源技术服务有限公司, 河北 石家庄 050021;
3. 国网河北省电力有限公司建设公司, 河北 石家庄 050030; 4. 紫光软件系统有限公司, 北京 100084)

摘要: 电力场景的三维语义信息识别是其后续精细化管理的基础和关键, 然而, 由于电力场景地物结构信息复杂、纹理多样, 为其精细化理解与识别带来了一定的困难和挑战。本文提出了一种基于改进 RandLA-Net 的电力场景点云语义分割方法, 该方法通过引入特征拓展和分离池化操作来提高模型的性能, 并在电力数据集上测试了该方法的实际效果, 将其与现有的语义分割方法进行了比较。结果表明, 该方法在准确性和效率方面具有很强的优势, 综合对比来看, 比前沿的 RandLA-Net (Random Sampling and Local Feature Aggregator Network) 提高了 2.64 和 2.9 的总体精度及平均交并比值, 验证了该方法的有效性。

关键词: 特征拓展; 分离池化; 电力场景; 点云语义分割; 自注意力机制

中图分类号: TP391.41; TN249 文献标识码: A DOI: 10.3969/j.issn.1001-5078.2025.01.006

Point cloud semantic segmentation based on feature expansion and separation pooling

CHEN Chong-ming¹, YU Jin-xing², HAN Lu³, WANG Hao-ran³, ZHANG Dian-mao⁴, CHEN Qi⁴

(1. State Grid Hebei Electric Power Research Institute, Shijiazhuang 050021, China;

2. State Grid Hebei Energy Technology Service Co., Ltd., Shijiazhuang 050021, China;

3. State Grid Hebei Construction Company, Shijiazhuang 050030, China;

4. Unis Software System Co., Ltd., Beijing 100084, China)

Abstract: The recognition of three-dimensional semantic information in power scenes is the foundation and key to its subsequent fine-grained management. However, due to the complexity of ground structure information and diverse textures in power scenes, it brings certain difficulties and challenges to its fine-grained understanding and recognition. In this paper, a semantic segmentation method based on improved RandLA Net for power scene point clouds is proposed, which improves the performance of the model by introducing feature expansion and separation pooling operations. And the actual performance of this method is tested on a power dataset and compared with existing semantic segmentation methods. The results show that this method has strong advantages in accuracy and efficiency. and in a comprehensive comparison, it improves the overall accuracy and the average intersection and merger ratio values by 2.64 and 2.9 over the cutting-edge RandLA-Net (Random Sampling and Local Feature Aggregator Network), which verifies the effectiveness of the method.

Keywords: feature extension; separation pooling; electricity scenario; point cloud semantic segmentation; self-attention mechanism

基金项目: 国网河北省电力有限公司科技项目 (No. KJ2023-033) 资助。

作者简介: 陈崇明 (1983-), 男, 硕士, 正高级工程师, 主要从事电网工程环境保护以及水土保持技术、变电站噪声监测及控制技术等方面的工作。E-mail: dmzhang226@126.com

收稿日期: 2024-04-18; 修订日期: 2024-05-07

1 引言

点云是由三维空间中的离散点组成的数据形式,广泛应用于三维场景建模和计算机视觉任务中^[1]。在电力场景目标识别中,点云语义分割作为一项重要任务,旨在将点云中的每个点分配到预定义的类别中。该任务在电力设施的自动化维护和安全监测方面具有重要意义,能够帮助自动化系统对电力设施进行智能化管理和监测^[2]。

传统的电力场景语义分割方法包括支持向量机(Support Vector Machine, SVM)^[3]、基于图论的方法^[4]、基于特征聚类的方法^[5],然而传统的非线性学习进行电力场景语义分割方法的特征往往是手工设计的,这也导致无法捕捉场景中高级的深层次显著信息,给电力场景准确高效的语义标记造成一定的局限性。

近年来,深度学习方法被引入到点云语义分割任务中,取得了显著成效^[6-7]。深度学习方法的开拓模型之一是 PointNet++^[8]。该方法是一种用于点云分割的深度学习架构,通过逐层的局部区域分组和特征聚合来提取点云的局部和全局特征,以实现目标信息识别。SO-Net(Self-Organizing Network)^[9]是一种基于集合操作的神经网络,用于对点云进行分割和特征学习。它通过对点云中的每个点的局部邻域进行排序,并使用排序后的邻域特征进行分割。DRPT(Double randomness Point Transformer)^[10]通过在 Transformer 模型卷积投影层创建新的点嵌入,利用局部邻域的动态处理在数据特征向量中持续增加全局特征属性,从而提高多目标识别中点云分类和分割的精度。PointCNN(Point Convolutional Neural Network)^[11]是一种对点云数据进行卷积操作的神经网络,能够提取点云的局部和全局信息,它的卷积操作可以作为分离池化方法的一种替代,以更好地解决互信息稀疏性的问题。

RandLA-Net^[12]是最近提出的一种用于点云分割的神经网络,它具有先进的性能和较快的速度,它利用局部感知和全局感知相结合的策略,通过对点云进行规则化采样和特征聚合来实现语义分割,取得了很好的效果,为电力场景点云的高效语义分割提供了可能。

然而,尽管 RandLA-Net 具有很多优点,但仍然面临着一些挑战。一方面,特征提取的能力有限,

难以充分捕捉点云数据中的特征和细节信息。另一方面,互信息的稀疏性也限制了模型的性能。为了解决这些问题,本文提出了一种特征拓展和分离池化的方法来改进 RandLA-Net。在改进的模型中,通过引入更多的感受野和上下文信息,更好地捕捉电力场景点云中的结构和语义信息,再通过增加网络层级及改进特征提取机制来实现。分离池化方法旨在解决互信息的稀疏性问题。传统的点云分割方法通常采用全局或局部特征池化来减少特征维度,但这样会丢失部分信息。分离池化方法通过将全局和局部的特征分开进行池化,分别保留更全局的特征和更局部的特征,并在后续进行融合,从而提高了特征的表达能力和模型性能。通过引入特征拓展和分离池化方法来改进 RandLA-Net 性能,进一步推动了电力场景点云语义分割的研究和应用。

2 方法

2.1 基本原理

本文提出的方法主要包括两个模块:特征拓展和分离池化。特征拓展模块旨在增加点云表征的区分度,提高点云表征的能力。分离池化模块旨在降低点云密度和维度,以便更有效地处理点云数据。该方法的整体网络架构如图 1 所示,其网络架构包含下采样层、上采样层及全连接层,其中下采样层由自注意力特征拓展表达模块(Self-Attentional Feature Extended Expression Module, SA-FEE)及自注意力特征分离池化模块(Self-Attentional Feature Separated Pooling Module, SA-FSP)组成。在操作中,首先,输入一个电力场景点云,先对其进行网格降采样以降低点云分辨率,D₀至 D₄为下采样层,在每一个下采样层中,点云中的每个点将通过 SA-FEE 模块及 SA-FSP 模块联合作用以聚合局部邻域点特征。逐层下采样后,每个点特征将具有更大感受野及丰富的语义和纹理特征。综合全局特征后在逐层双线性上采样至原始点云密度,并逐点预测出其所所属的语义类别,最终得到如图 1 中所示精细化电力场景点云分割结果。

在下采样过程中,使点云特征得到充分表达可有助于通过网络学习得到更具细粒度的空间结构及纹理信息。本模型通过提出 SA-FEE 模块,从空间关系及方位关系两方面对点云位置特征充分表达,

并进一步提出 SA-FSP 模块,基于自注意力机制,通过分离池化的方式有效提取得到充分表达后的数据特征。由 SA-FEE 及 SA-FSP 组成的下采样层可通

过网络准确高效学习并输出电力场景点云中更具辨识度的特征。接下来,分别介绍 SA-FEE 模块及 SA-FSP 模块。

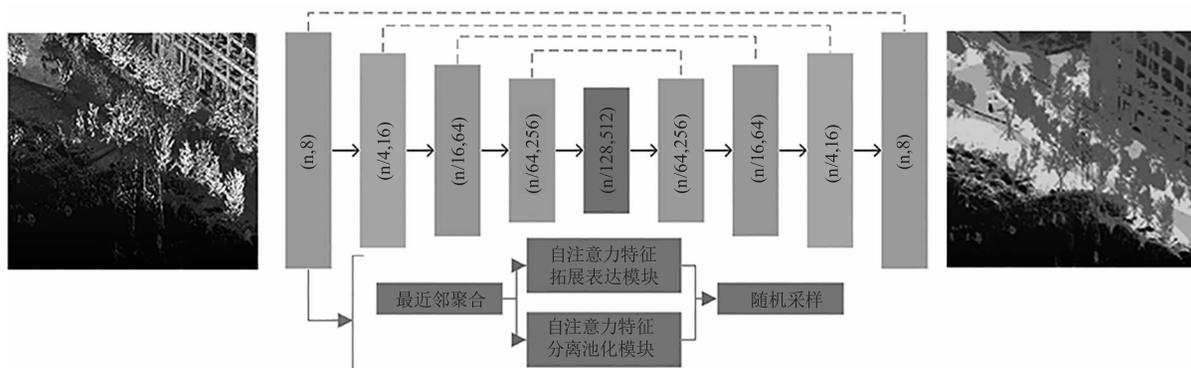


图 1 总体网络结构图

Fig. 1 Overall network structure diagram

2.2 自注意力特征拓展表达模块

对电力场景点云进行特征提取时,通常由下采样层逐层对输入场景点云持续地提取,从而输出富含大量信息的全局特征数据。然而,目前大部分点云语义分割工作,例如 PointNet 等,仅仅关注场景中每个点的位置信息,而未充分挖掘其空间关系,甚至忽略了点与点间的方位关系,导致点云特征未得到充分表达,网络特征学习能力不佳。

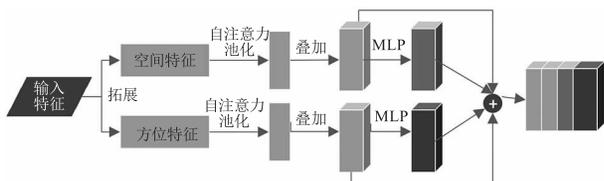


图 2 自注意力特征拓展表达模块

Fig. 2 Self-attention feature extended expression module

针对该问题,本文提出了自注意力特征拓展表达模块 SA-FEE,如图 2 所示,该模块通过将每个输入点的局部邻域位置信息拓展成空间特征及方位特征,并基于自注意力机制使其特征更具辨识度。模块主要包含两个部分:特征拓展层和自注意力特征强化层。

2.2.1 特征拓展层

在该层中,首先完成空间特征层面的拓展。已知输入第 i 点位置信息为 x_i, y_i 和 z_i ,其与第 j 个领域点的位置信息差为: x_i^r, y_i^r 和 z_i^r ,这样,两点间的欧式距离 d_i^r 可由下列式(1)解算:

$$d_i^r = \sqrt{(x_i^r)^2 + (y_i^r)^2 + (z_i^r)^2} \quad (1)$$

在解算得到两点间欧式距离的基础上,两点间的空间特征 p_i^j 可通过位置信息由公式(2)拓展获得:

$$p_i^j = x_i^r \oplus y_i^r \oplus z_i^r \oplus d_i^r \quad (2)$$

接着,由两点间位置信息也可拓展为方位特征,其中绕 Z 轴旋转不变的角度 ϕ_i^j 可由公式(3)计算所得,绕 Y 轴旋转不变的角度 ϕ_i^j 可由公式(4)计算所得:

$$\phi_i^j = \arctan\left(\frac{z_i^r}{\sqrt{(x_i^r)^2 + (y_i^r)^2}}\right) \quad (3)$$

$$\phi_i^j = \arctan\left(\frac{y_i^r}{x_i^r}\right) \quad (4)$$

最终,从方位特征层面的拓展特征 a_i^j 可由上述结果拼接所得。

$$a_i^j = \phi_i^j \oplus \phi_i^j \oplus d_i^r \quad (5)$$

2.2.2 自注意力特征强化层

经过特征拓展层所得到的特征需要做进一步的提取,自注意力特征强化层基于自注意力机制使拓展后的特征更具辨识度。首先需要对扩展后特征进行编码,计算流程如式(6)、式(7)所示。

$$f_i = g(f_{\text{raw}}, w_1, w_2, w_3) \quad (6)$$

$$f_{\text{att}} = \text{SUM_CP}(\delta(f_i) * f_i) \quad (7)$$

由上一模块提取的拓展特征 f_{raw} 作为模块的输入项,为了减少网络复杂度,首先通过基于 MLP 的线性变换函数 $g(\cdot)$ 将全局特征重映射到 $128 \times r_1$ 维、 $128 \times r_2$ 维及 $128 \times r_3$ 维,其中 w_1, w_2 和 w_3 为可学习权重参数。之后,通过 Softmax 激活函数 $\delta(\cdot)$ 计算得到自注意力值,与之前特征进行通道维度相乘操作(即“*”号),再通过 $\text{SUM_CP}(\cdot)$ 将张量先

累加再平铺到与原特征相同维度。最后, $f_{\text{raw}} f_i$ 与 f_{att} 执行相连接(即 concat)操作,得到的特征图张量经过 MLP 后进入特征通道强化层,最终转换成与输入通道一致的自注意力向量。

接着,将得到的特征图执行公式(8)操作,得到 $1 \times r$ 维特征向量 \tilde{f}_{att} 。为了方便对输入特征进行自注意力表达,需要对自注意力值进行归一化转化,因此,该部分将编码得到的 \tilde{f}_{att} 向量通过 Sigmoid 函数映射到 $0 \sim 1$ 范围内,形成通道注意力,最后采用类似残差结构的方法将该通道注意力与 f_{input} 逐通道相乘再相加,从而达到显著化表达拓展后特征中各类别语义特征的作用。

$$f_{\text{aug}} = \text{sigmoid}(f_{\text{att}}) * f_{\text{input}} + f_{\text{input}} \quad (8)$$

2.3 自注意力特征分离池化模块

在经过 SA-FEE 模块完成基于位置信息的空间特征及方位特征拓展后,所形成的点特征富含周围邻域特征信息。特征拓展的行为虽然能进一步丰富原始特征信息,但同样存在信息冗余及特征辨识度不足的问题。由此,本研究提出自注意力特征分离池化模块(Self-Attentional Feature Separated Pooling Module, SA-FSP),该模块基于自注意力机制,可使输入特征的表达更具辨识度。同时,分离池化的操作将空间几何特征及纹理与语义特征分开执行池化操作,考虑了两个特征间的表达尺度差异性,进一步使特征辨识度增加,使网络学习能力增强。

如图3所示,在接收到上一级传递的特征后,SA-FSP 模块首先将特征按空间几何及纹理信息分开,随后分别执行自注意力池化,最后两个特征累加,形成输出特征。

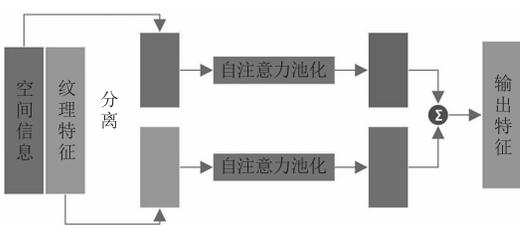


图3 自注意力特征分离池化模块

Fig. 3 Self-attentional feature separated pooling module

2.4 模型损失值计算

本文将原始点云数据输入到网络中,经过下采样层的特征提取及上采样层的特征重映射后,输出与原始点云数据各点对应的语义标签预测值,再将

预测值与输入训练的点云场景的真值标签计算加权交叉熵损失,从而得到点维度的损失(L_{pt})。具体表达如下式:

$$L_{\text{pt}} = -\frac{1}{n} \sum_{i=0}^n w_i * p(x_i) * \log(q(x_i)) \quad (9)$$

式中, n 为类别数目; w_i 为该类别样本数量占总样本的权重; $p(x_i)$ 为点 x 在 i 类的概率; $q(x_i)$ 为点 x 标签的真实分布。

3 实验与结果分析

3.1 电力场景数据集介绍

为验证模型效果,在大规模电力场景激光雷达数据集上评估了该方法,并将其与现有的方法进行了比较。本文所采用的电力场景数据集是由车载激光雷达扫描仪获取的户外电力场景点云数据集,如图4所示。该数据集根据场景情况划分为训练数据集及测试数据集,其中包含9种地物类别,依次是扫描伪影、人造路面、自然地面、植被、建筑、硬景观、车辆、电力塔、电力线。训练集和测试集共有5个文件,点云总数超过4000万。点云各类别数量统计如图5所示,图中可见该数据集类别存在较为严重的长尾分布现象,其中人造路面点超过1200万,电力塔及电力线不足10万,这对模型分割带来一定挑战。

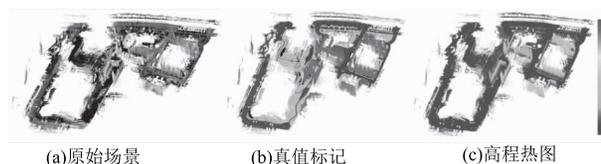


图4 电力场景数据集描述

Fig. 4 Power scenario dataset description

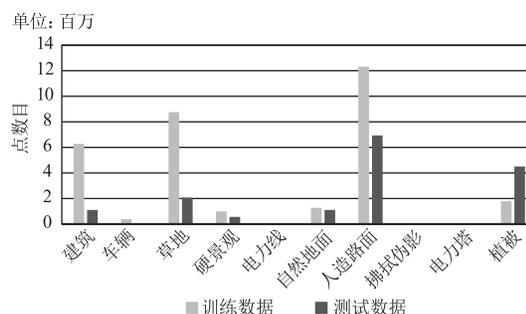


图5 数据集分类点云数量

Fig. 5 Number of point clouds by category in the dataset

3.2 评价指标

本文实验所采用的对比指标包括交并比(Intersection over Union, IoU)及总精度(Overall Accuracy, OA),其中IoU与平均IoU(mean IoU, mIoU)具体计算

如公式(10)、(12)所示。其中 n 为类别数量。 OA 的具体计算可见式(11),其中, TP (True Positive)代表预测结果类别与真值类别一致的点数量, TN (True Negative)表示本该分类正确但却被模型分类错误的点数量, FP (False Positive)代表误将不属于 Ground Truth 类别的点分到了该类的点数, FN (False Negative)代表正样本被错误分为负样本的点数。

$$IoU = \frac{TP}{TP + FP + FN} \quad (10)$$

$$OA = \frac{TP + TN}{TP + FP + FN + TN} \quad (11)$$

$$mIoU = \frac{1}{n} \sum_{i=0}^n IoU_i \quad (12)$$

3.3 实验设置

本文训练采用 100 个 Epoch,学习率初始为 0.01,随后以每 Epoch95% 的速率衰减,训练的硬件设备与系统如表 1 所示。

表 1 软硬件环境

Tab. 1 Software and hardware environments

名称	参数
CPU	Intel(R)Core(TM)i9-9900K CPU @ 3.60 GHz
GPU	GeForce GTX 2080ti
RAM	32 G
Hard disk	1 T
System	Ubuntu 18.04

表 3 测试结果以及与 RandLA-Net 结果的比较

Tab. 3 Test results and comparison with RandLA-Net results/%

方法	平均交并比/%	总精度/%	扫描伪影	人造路面	自然地面	植被	建筑	硬景观	车辆	电力塔	电力线
RandLA-Net	69.15	85.74	63.25	93.63	56.81	75.99	66.47	33.11	61.21	73.23	98.62
本文方法	72.05	88.38	68.35	94.18	61.43	80.79	77.94	33.13	62.79	74.58	95.26
结果对比	+2.90	+2.64	+5.10	+0.55	+4.62	+4.80	+11.47	+0.02	+1.58	+1.35	-3.36

本文同时选取 RandLA-Net 展开对比实验,其在 $mIoU$ 指标上的测试结果为 69.15%,比本文方法低 2.9%。在 OA 指标上的测试结果为 85.74%,比本文方法低 2.64%。从结果来看,本文方法在大多数类别的分割精度上均高于 RandLA-Net,尤其是在建筑类别分割上,本文比 RandLA-Net 提高了 11.47%。在电力类目标上,电力塔分隔精度比 RandLA-Net 高 1.35%,但由于电力线目标较细,在分割精度上稍有不足,比 RandLA-Net 低 3.36%。

为增强两种方法的对比效果,将本文方法的分

3.4 参数敏感性实验

场景点云在输入到网络前将执行网格下采样操作,以降低点云密度,同时扩大网络的空间感知范围,但过大的采样力度将造成更具细粒度的点云特征损失。因此,此节将对网格降采样力度 m 进行参数敏感性实验,以确保在实验过程中参数选择准确且恰当。

由表 2 可得,当 m 取 0.06 时网络测试结果达到最佳,对比 $m=0.04$, $mIoU$ 提高了 1.4%, OA 提高了 1.8%。

表 2 网格采样强度 m 值选取灵敏度试验

Tab. 2 Sensitivity experiment on selecting the m value of grid downsampling intensity

序号	m 的选取	总精度/%	平均交并比/%
1	0.04	86.60	70.70
2	0.06	88.40	72.10
3	0.08	63.80	41.40

3.5 测试结果及 RandLA-Net 对比

本文方法的测试结果如表 3 所示,在 $mIoU$ 指标上的测试结果为 72.05%,在 OA 指标上的测试结果为 88.38%,从各类别 IoU 结果来看,除硬景观类别外其余类别均达到 60% 以上精度,电力线及人造路面类别中更是取得了 95.26% 及 94.18% 的优秀表现。

割结果与 RandLA-Net 的分割结果进行了可视化对比,对比效果如图 6 所示。图中,左侧为地面真值的结果,中间为本文方法的结果,右侧为 RandLA-Net 的结果。在上方第一个对比场景中,对比左侧车辆以及建筑物的真值情况,本文方法较好地完成了车辆分割,RandLA-Net 在车辆的头部以及尾部出现了分割错误(标识位置),将车辆误分割为了硬景观。在中间第二个对比场景中,从图右上角及左下角的硬景观可以看出,本方法虽然在右上角的路灯中段存在一些分割错误,但总体分割效果良好,而 Rand-

LA-Net 则错误分割了左下方路灯尾部及右上方路灯顶部(标识位置)。在下方最受关注的电力场景下,本文方法对比地面真值较好地完成了分割,而 RandLA-Net 将电力塔底部区域错误分割为植被(标识位置)。

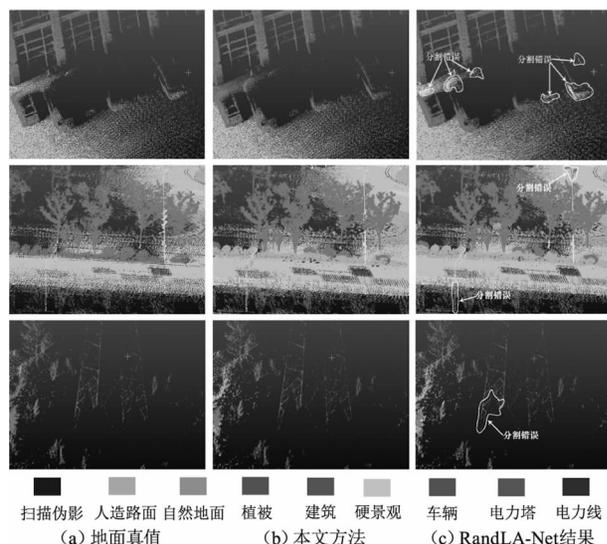


图6 结果可视化对比图

Fig. 6 Result visualization comparison chart

4 结论

本文提出了一种改进 RandLA-Net 的电力场景点云语义分割方法,该方法主要设计了特征的扩展来加强表达,并结合分离池化对电力场景进行语义分割,来充分表达电力场景点云特征的显著性及精细性。实验表明,该方法具有较好的性能,在准确性和效率方面均有所提高。此外,所提出的特征拓展和分离池化方法可以应用于其他点云分割问题中,未来主要通过进一步融合自注意力的方式来继续提升场景特征表达的灵活性及高效性。

参考文献:

- [1] Zou Juanru. Research on the application of laser point cloud data in 3D modeling technology [J]. Heilongjiang Science and Technology Information, 2015, (15): 102 - 103. (in Chinese)
邹娟茹. 激光点云数据在三维建模技术的应用研究 [J]. 黑龙江科技信息, 2015, (15): 102 - 103.
- [2] Nardinocchi C, Balsi M, Esposito S. Fully automatic point cloud analysis for powerline corridor mapping [J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, (99): 1 - 12.
- [3] Zhang C, Chen C, Hu Q, et al. Efficient segmentation and thresholding approach for power line extraction in urban areas [J]. Front. Earth Sci., 2018, 12: 102 - 113.
- [4] Hu W, Liu J, Zou Q. A robust power-line extraction method from LiDAR point clouds using graph cut based on the normalized cut (NCut) criterion [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2019, 147: 86 - 99.
- [5] Zhu Q, Tian D, Shen J, et al. (2020). Semantic segmentation of power line point cloud based on feature clustering and region growing [C] // 2020 2nd International Conference on Geomatics and System Engineering, IEEE, 2020: 689 - 693.
- [6] Meng Qiongtang, Zhao Yindi, Han Wenquan, et al. Airborne LiDAR point cloud city building change detection based on RandLA-Net J1 [J]. Remote Sensing for Natural Resource, 2022, 34(4): 113 - 121. (in Chinese)
孟琼棠, 赵银娣, 韩文泉, 等. 基于 RandLA-Net 的机载激光雷达点云城市建筑物变化检测 [J]. 自然资源遥感, 2022, 34(4): 113 - 121.
- [7] 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. (in Chinese)
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.
- [8] Qi C R, Su H, Mo K, et al. Pointnet: deep learning on point sets for 3D classification and segmentation [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 652 - 660.
- [9] Li J, Chen B M, Lee G H. So-net: Self-organizing network for point cloud analysis [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 9397 - 9406.
- [10] Ma Qinglu, Sun Xiao, Huang Xiaoxiao, et al. A transformer-based classification and segmentation approach for classifying and segmenting road field attraction clouds [J]. Laser & Infrared, 2024, 54(1): 17 - 23. (in Chinese)
马庆禄, 孙泉, 黄筱潇, 等. 基于 Transformer 的道路场景点云分类与分割方法 [J]. 激光与红外, 2024, 54(1): 17 - 23.
- [11] Li Y, Bu R, Sun M, et al. PointCNN: convolution on X-transformed points [C] // Neural Information Processing Systems, Curran Associates Inc, 2018.
- [12] Hu Q, Yang B, Xie L, et al. R and LA-Net: efficient semantic segmentation of large-scale point clouds [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11108 - 11117.