

# 基于改进 PointNet++ 的 Lidar 点云分割模型

张驰, 王志杰\*, 吴昊, 陈动

南京林业大学土木工程学院, 江苏 南京 210037

**摘要** 针对 PointNet++ 在特征提取阶段未能深层挖掘 Lidar 点云的语义特征及其在特征聚合阶段采用最大池化聚合导致特征丢失, 进而导致点云分割精度下降的问题, 通过改进 PointNet++ 的特征提取及特征聚合模块, 提出一种基于特征偏差值和注意力机制的点云分割模型。首先, 利用球形采样获取不同的局部邻域, 并采用 K 最近邻 (KNN) 算法筛选邻域点, 计算不同邻域的特征偏差值, 获取点云的深层语义信息, 增强模型对不同局部邻域的识别能力; 其次, 利用基于注意力机制的特征聚合模块代替 PointNet++ 中的最大池化模块, 在聚合特征阶段学习不同特征的权重, 从而提高模型对不同结构信息的筛选能力, 增强模型的分割性能; 最后, 为了进一步优化模型架构, 在全连接层中加入残差模块, 共享权重, 避免参数冗余, 提升模型性能。基于 ISPRS 提供的 Vaihingen 数据集与斯坦福的 S3DIS 数据集进行实验验证, 实验结果表明, 所提模型总体精度达到 86.69%, 较 PointNet++ 提高了 5.49 个百分点, 同时平均 F1 得分达到了 73.97%, 较 PointNet++ 提高了 8.30 个百分点。在 S3DIS 数据集上的实验结果表明, 与 PointNet++、RandLA-Net 和 ConvPoint 等主流模型相比, 所提模型结果较 PointNet++ 也有提升, 即相较于 PointNet++ 的分割结果, 改进后的模型能够充分提取点云的语义特征, 有效提高模型分割精度。

**关键词** 点云分割; PointNet++; Lidar; 特征偏差值; 注意力机制; 特征融合; 残差结构

中图分类号 TP391.41

文献标志码 A

DOI: 10.3788/LOP231106

## Lidar Point Cloud Segmentation Model Based on Improved PointNet++

Zhang Chi, Wang Zhijie\*, Wu Hao, Chen Dong

College of Civil Engineering, Nanjing Forestry University, Nanjing 210037, Jiangsu, China

**Abstract** It is known that PointNet++ cannot deeply explore the semantic features of Lidar point clouds during feature extraction and loses features due to the use of maximum pooling during feature aggregation. These problems result in a decrease in point cloud segmentation accuracy. To address these problems, this study proposes a point cloud segmentation model based on feature deviation values and attention mechanisms by improving the feature extraction and feature aggregation modules of PointNet++. First, different local neighborhoods are obtained using spherical sampling, and then neighborhood points are selected using the K-nearest neighbor (KNN) method to calculate the feature deviation values of different neighborhoods. This enhances the model's recognition ability for different local neighborhoods and obtains deep semantic information of the point cloud. Second, an attention-based feature aggregation module is used to replace the maximum pooling module in PointNet++ to learn the weights of different features during feature aggregation. This improves the model's ability to filter information from different structures and enhances the segmentation performance of the model. To further optimize the model architecture, a residual module is added to the fully connected layer to avoid parameter redundancy and improve the model performance through weight sharing. Experimental results are validated on the Vaihingen and S3DIS datasets provided by ISPRS and Stanford University, respectively, and are compared with the results of the experiments and mainstream models provided by ISPRS. The overall accuracy (OA) and average F1 score reach 86.69% and 73.97%, respectively, which are 5.49 percentage points and 8.30 percentage points higher than those of PointNet++, respectively. The experimental results on the S3DIS dataset are compared with those of PointNet++, RandLA-Net, and ConvPoint, and show a clear improvement over those of PointNet++. The experiments show that the improved model can fully extract the semantic features of point clouds and effectively improve the segmentation accuracy of the model as compared with the segmentation results of PointNet++.

**Key words** point cloud segmentation; PointNet++; Lidar; feature deviation value; attention mechanism; feature fusion; residual structure

收稿日期: 2023-04-17; 修回日期: 2023-05-11; 录用日期: 2023-05-19; 网络首发日期: 2023-05-29

基金项目: 大场景机载点云建筑语义建模(41971415)

通信作者: \*stampabc@126.com

# 1 引言

随着星载、机载、地基遥感知传感器不断发展,点云获取越来越容易,其在自动驾驶<sup>[1]</sup>、机器人测绘<sup>[2]</sup>、三维重建<sup>[3]</sup>等领域得到广泛应用<sup>[4]</sup>。点云的有效分类是应用的前提,为此,学界开展了大量研究工作以期实现更高效的点云分割方法,从而更好地识别和分析点云。深度学习的出现,促进了计算机视觉领域在点云分类分割中的发展。然而点云是一种不规则且无序的几何数据,具有规则架构的卷积神经网络在点云分类中效率较低,因此,如何利用神经网络的优势将其应用在点云分类领域是当前研究热点。目前针对点云的处理方法可以分为以下几类:基于体素的方法、基于投影的方法、基于点的方法和基于图的方法。

基于体素的方法将点云转化为规则的格网数据,解决了点云无序和不规则的问题。其中,具有代表性的是 Maturana 等<sup>[5]</sup>提出的 VoxNet, VoxNet 将点云采样处理为体素网格后,利用 3D 卷积神经网络(CNN)对其进行分类。Wu<sup>[6]</sup>提出 3DShapeNet,将点云的几何形状转换为体素网格中的联合分布,在一定程度上解决了点云的非结构化问题。但是,这种处理方法会导致数据稀疏、计算量增加和复杂度提高等问题。为了解决这些问题,学者们利用 kd-tree<sup>[7]</sup>和 Octree<sup>[8]</sup>两种结构来处理点云,这在一定程度上改善了体素化后的问题。另外,基于投影的方法也被一些学者尝试,如 Su 等<sup>[9]</sup>提出的 MV 卷积神经网络基于多视角获取的 2D 影像获取多视角的投影视图,然后利用 2D 卷积神经网络提取特征。Wang 等<sup>[10]</sup>设计了一种循环聚类 and 池化的模块用以从相似的视图中提取并汇集信息。然而,这种方法会受到几何信息丢失的限制,导致点云的语义信息缺失。基于点的方法成为直接处理点云的有效手段。PointNet<sup>[11]</sup>通过共享的多层感知器(MLP)直接处理无序的点云,并使用最大池化来聚合各类特征的最大值以实现置换不变性。PointNet++<sup>[12]</sup>则通过采样和分组来提取点云的局

部特征,并逐渐扩大感受野来获取更多的特征信息。这两种方法都取得了可喜的成果。此外,基于图的方法主要关注点与点之间的关系,通过构建局部图并利用可学习的内核提取卷积运算中的局部信息来处理点云。EdgeConv<sup>[13]</sup>通过描述点与其邻域之间关系的边缘特征构建局部图,并很好地保留了点之间的关系。3D-GCN<sup>[14]</sup>利用全局池化和多分辨率池化全局和局部特征,PointGCN<sup>[15]</sup>则利用每个点的  $k$  个最近邻点构建局部图。这些方法都取得了较好的效果,并且是当前点云分割领域的主流方法之一。

上述方法在提取点云的局部特征时,利用性能优异的特征提取器筛选了点云局部的复杂特征,但在提取的过程中,忽略了局部区域本身所提供的几何特征。在聚合局部特征时,多数采用最大池化聚合,最大池化会导致点云特征丢失,进而影响分割精度。为了解决这些问题,本文提出一种改进的 PointNet++<sup>[12]</sup>模型,通过设计轻量级特征提取器来提取点云局部结合特征,从而增强点云的局部特征,同时引入注意力机制来聚合点云的局部特征,并给不同的特征赋予不同的权重以增强模型的场景分割能力。核心思想是结合局部几何特征和复杂特征,以增强点云的局部特征。实验结果表明,所提模型能够有效地提高点云场景分割的精度。

## 2 研究方法

### 2.1 基于特征偏差值和注意力机制的模型构建

所提方法的分割流程主要包括特征提取、特征聚合、场景分割,如图 1 所示。首先,在特征提取阶段,采用基于特征偏差值的提取方式计算不同局部邻域的特征偏差值,进而获取不同邻域的几何特征;其次,在聚合特征阶段,采用基于注意力机制的池化模块聚合特征用以场景分割,解决最大池化的聚合方式导致特征丢失影响模型分割精度的问题;此外,在模型的设计过程中,采用残差多层感知器代替多层感知器,残差多层感知器的使用可避免随着模型深度增加导致的参数冗余,进一步优化模型结构。

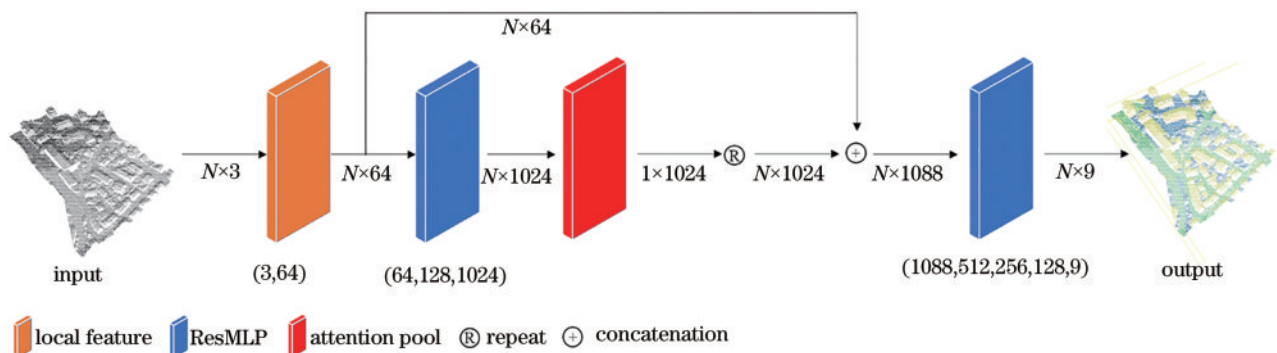


图 1 模型架构

Fig. 1 Schematic diagram of the model

## 2.2 基于特征偏差值的特征提取模块

### 2.2.1 基于特征偏差值的特征提取模块设计

在 Lidar 点云分割任务中的几何特征提取阶段,对于输入的点云  $P = \{p_i \in \mathbb{R}^3, i = 1, 2, \dots, N\}$ , PointNet++<sup>[12]</sup> 通过最远点采样算法降采样点云,之后利用 K 最近邻 (KNN) 算法筛选邻域点并用最大池化聚合特征。对于 PointNet++<sup>[12]</sup> 特征提取部分进行改进的方法较多:杨晓文等<sup>[16]</sup>提出一种改进的 KNN 算法来提取点云的局部特征,在提取局部特征时考虑了点云分布的密度差异;赖鸣<sup>[17]</sup>等提出的 CAFF-PointNet 在局部特征提取阶段,通过计算平面度、线性度、法向量等浅层几何特征,并与网络提取到的语义特征一起强化模型的几何感知能力。

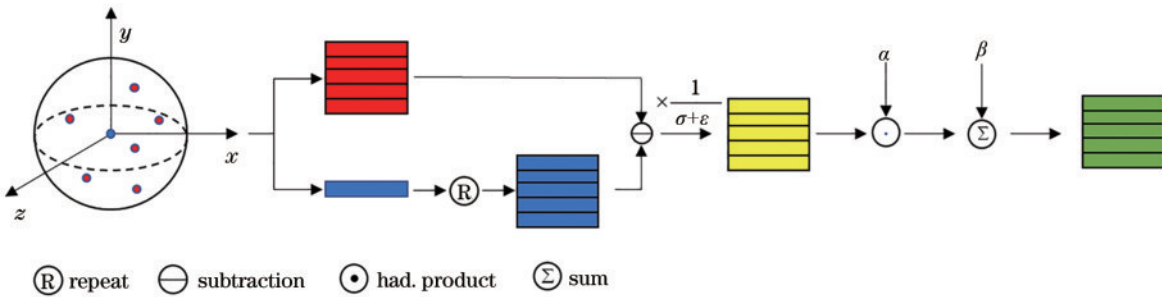


图 2 基于特征偏差值的特征提取模块

Fig. 2 Feature extraction module based on feature deviation values

首先,以采样点为球心构建球形邻域,并在球形邻域内筛选出  $k$  个邻近点,进而计算采样点到各邻近点的欧氏距离:

$$d = \|f_{i,j} - f_i\|, \quad (1)$$

式中:  $d$  表示欧氏距离;  $f_j$  表示第  $j$  个采样点;  $f_{i,j}$  表示  $f_i$  邻域内的第  $j$  个邻近点。

之后,计算邻域的特征偏差值:

$$\tau = \sqrt{\frac{1}{k \times n \times d} \sum_{i=1}^n \sum_{j=1}^k d}, \quad (2)$$

式中:  $\tau$  表示特征偏差值;  $k$  表示邻域内邻近点个数;  $n$  表示采样点的个数。

最后,得到采样点所在区域的特征值:

$$\{f_{i,j}\} = \alpha \odot \frac{\{f_{i,j}\} - f_i}{\tau + \epsilon} + \beta, \quad (3)$$

式中:  $\alpha$ 、 $\beta$  是两个自学习的参数;  $\epsilon$  是一个用以数值稳定性的小数<sup>[18]</sup>。

通过上述方法,在提取局部特征时,由于局部区域不同的几何结构,将求得不同的特征偏差值,通过计算特征偏差值,深层挖掘点云语义信息,从而增强点云的局部特征。

### 2.3 残差多层感知器

PointNet<sup>[11]</sup> 的核心思想为利用多层感知器和最大池化构建一个一般函数以解决点云的置换不变性问

题,对于 Lidar 点云的场景分割任务,由于点云包含着复杂的几何结构,为了区别不同的几何结构,挖掘深层的语义特征,本研究借鉴群组归一化<sup>[18]</sup>在处理特征时计算特征偏差值的思路,构建了基于特征偏差值的几何特征提取器,采用球形采样构建不同的局部邻域,对每一个邻域,计算其特征偏差值进而求得几何特征。

### 2.2.2 基于特征偏差值特征提取模块的实现

通过球形采样和 KNN 来获取点云局部区域的特征,是提取几何特征所采用的常见方法。由于点云不同区域的几何结构差异较大,因此 KNN 难以充分挖掘局部邻域的几何特征,为了解决该问题,设计了基于特征偏差值的局部几何特征提取器,如图 2 所示。

题,对于输入的点云  $P = \{p_i \in \mathbb{R}^3, i = 1, 2, \dots, N\}$ ,通过多层感知器实现特征的升维,进而采用最大池化从每组特征中选出最大值,实现特征聚合。通过重复上述方法扩大感受野,PointNet<sup>[11]</sup> 可以获取点云的全局特征,进而实现对点云的分割分类。PointNet++<sup>[12]</sup> 针对 PointNet<sup>[11]</sup> 缺少局部特征的问题,利用球形采样和 KNN 提取点云的局部特征,并在特征提取后,利用最大池化聚合特征,从而实现对点云局部区域的信息提取。

多层感知器在模型中的作用为提取由局部特征提取器所得到的点云局部特征,可以描述为

$$g_i = \max[\Phi(f_{i,j}) | j = 1, 2, 3, \dots, k], \quad (4)$$

式中:  $g$  表示最大池化提取到的全局特征;  $\max$  是最大池化;  $\Phi$  是多层感知器。

但是,由于点云信息的密度大,多层感知器在提取点云特征时,传统的多层感知器提取的点云特征每一层的参数都是独立的,需要大量的参数来描述特征之间的关系,这会导致模型的参数数量非常庞大。为了解决这个问题,可以采用残差多层感知器<sup>[19]</sup>,其能够共享权重使得所有层都使用相同的权重矩阵来计算输入和输出之间的关系,从而减小模型的参数数量。与 PointNet++ 相比,其参数数量仅有 79.21%,残差块具体如图 3 所示,模型参数数量如表 1 所示。



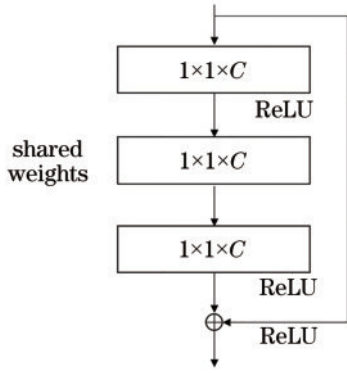


图 3 残差块示意图

Fig. 3 Schematic diagram of the residual block

表 1 模型参数数量

Table 1 Number of model parameters

Model	Number of model parameters /million
PointNet++	1.78
Proposed model	1.41

## 2.4 基于注意力机制的池特征聚合模块

### 2.4.1 注意力池化模块的设计

PointNet++<sup>[12]</sup>采用最大池化聚合特征导致部分特征丢失,进而影响点云分割的性能,而注意力机制可以从大量的输入特征中学习更关键的特征,减小不重要特征的注意力分数,甚至可以过滤掉一些不相关的特征,在多个特征中提取更重要的特征。

所提模型采用注意力池化来聚合点云的局部特征,注意力池化模块可以从特征集中计算不同特征的权重,保留有用的特征,最终得到聚合特征。首先,注意力池化模块利用全连接层整合特征,之后通过 Softmax 函数获取注意力得分。最后,将这些特征加权后聚合得到点云的聚合特征,如图 4 所示。

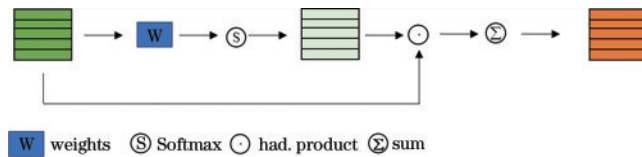


图 4 基于注意力机制的池特征聚合模块

Fig. 4 Pool feature aggregation module based on attention mechanism

### 2.4.2 注意力池化模块的实现

采用注意力机制对特征集  $p_{i,j}$  进行聚合,通过设计一个共享函数来学习特征集内每一个特征的注意力得分:

$$Q^j = \lambda(p_{i,j}, \varphi) \quad (5)$$

式中:  $\lambda$  是计算特征得分的函数;  $\varphi$  是一个可学习的参数;  $Q$  是一个  $j$  维的数组。  $\lambda$  一般由多层感知器和 Softmax 函数组成。

聚合后的特征为特征组内特征与其权重之积的和,相比最大值聚合,基于注意力机制的特征聚合能够为场景分割提供更多的特征信息。

$$\hat{p}_{i,j} = \sum (p_{i,j} \cdot q_j), q \in Q. \quad (6)$$

## 3 实验及结果分析

### 3.1 实验数据及评价指标

实验数据采用国际摄影测量与遥感学会 (ISPRS) 提供的德国 Vaihingen3D 数据集,该数据集提供了 9 类不同的数据,包括电力线、低矮植被、地面、汽车、篱笆、屋顶、立面、灌木、树木。其中,训练集共有 753876 个点,测试集有 411722 个点,表 2 展示了训练集和测试集中每个类别的点数,图 5 展示了测试集和训练集。

表 2 训练和测试数据集中每个对象类别中的点数

Table 2 Number of points in each object class in the training and test datasets

Category	Training set	Test set
Power line	546	600
Low_vegetation	180850	98690
Impervious surface	193732	101986
Car	4614	3708
Fence/hedge	12070	7422
Roof	152045	109048
Facade	27250	11224
Shrub	47605	24818
Tree	135173	54226
Total	753876	411722

评价指标采用总准确率 ( $R_{OA}$ )、查准率 ( $R_{precision}$ )、查全率 ( $R_{recall}$ ) 和 F1 得分。  $R_{OA}$  的计算公式如下:

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

式中:  $N_{TP}$ 、 $N_{TN}$ 、 $N_{FP}$ 、 $N_{FN}$  分别代表真阳性、真阴性、假阳性、假阴性的数量。

F1 得分根据每个类别的查准率和查全率计算,当某些类别数量不平衡的情况下, F1 可能高于  $R_{OA}$ 。

$$s_{F1} = 2 \times \frac{R_{precision} \times R_{recall}}{R_{precision} + R_{recall}}, \quad (8)$$

$$R_{precision} = N_{TP} / (N_{TP} + N_{FP}), \quad (9)$$

$$R_{recall} = N_{TP} / (N_{TP} + N_{FN}). \quad (10)$$

### 3.2 实验环境与网络参数设置

实验硬件配置如下:运行内存为 16 GB,处理器为 Intel(R) Xeon(R) CPU @ 2.30 GHz,显卡为 Tesla P100。网络参数设置如下:使用随机梯度下降 (SGD) 提高训练精度和效率,初始学习率 (learning rate) 设置为 0.1,最小学习率设置为 0.005,采用余弦退火算法动态调整学习率,优化器采用 SGD,全连接层 dropout 保留参数为 0.5。

余弦退火算法是一种动态调整学习率的技术,从一个较大的学习率开始,迅速将其降低值接近 0 的

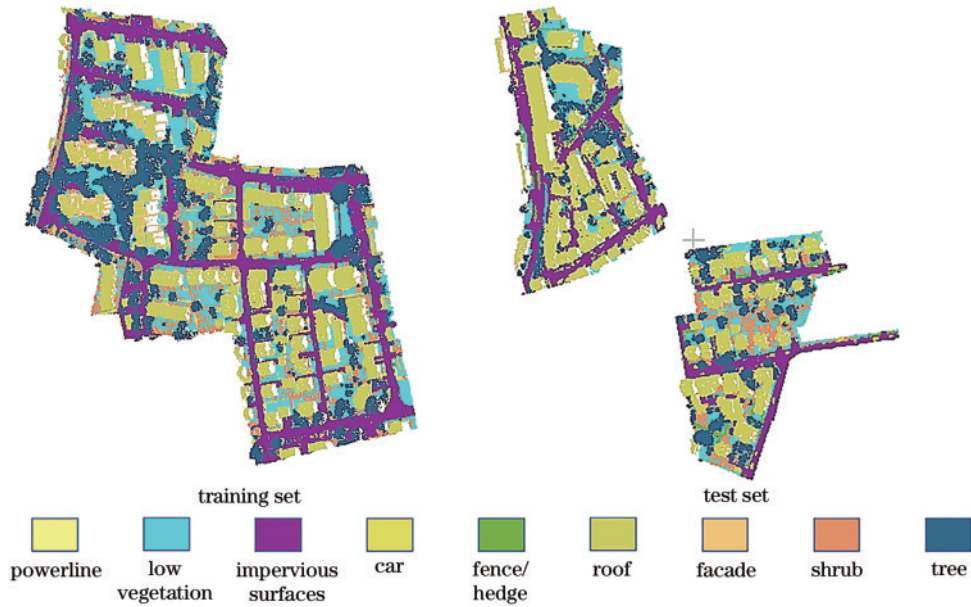


图 5 Vaihingen城市三维语义数据集  
Fig. 5 Vaihingen city 3D semantic dataset

值,然后再次提高学习率。由于模型在刚开始训练阶段其权重是随机的,此时利用一个较大的学习率,模型可能会出现振荡现象,在较小的学习效率预热下模型将趋于稳定,余弦退火策略有利于加快模型的收敛速度。

### 3.3 实验结果分析

为了验证模型在 Lidar 点云分割任务中的性能,在训练集上对模型进行训练,并在测试集上验证模型的分割效果,分割结果可视化如图 6 所示,混淆矩阵如表 3 所示。

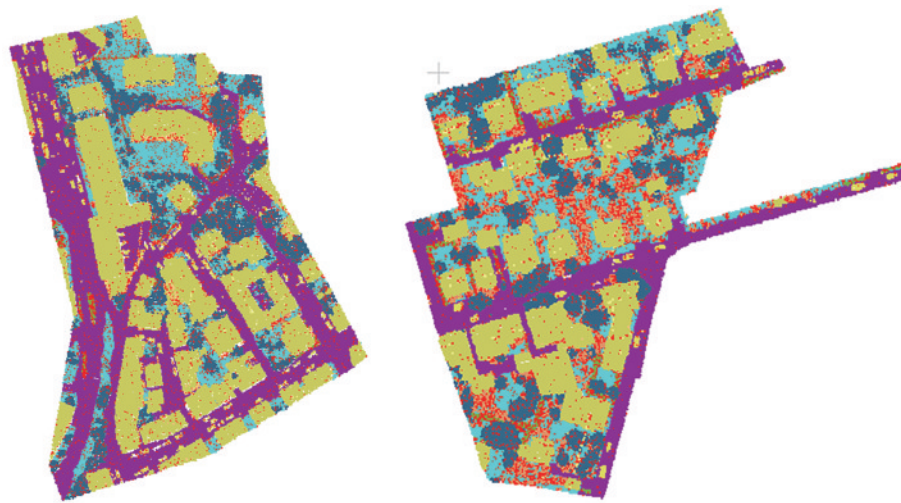


图 6 测试集分割结果可视化  
Fig. 6 Visualization of test set segmentation results

低矮植被、地面、屋顶、树木的 F1 得分均超过了 80%,地面的 F1 分数为 92.08%,屋顶的 F1 分数为 96.67%,对于小尺度类别的物体,如汽车也取得了较好的分割精度,F1 得分为 83.28%。结合图 6 和表 3 可知,分类错误主要集中在篱笆和灌木,图 7 展示了部分错误分割的结果与标准样本的对比,错误分割的原因是篱笆和灌木在结构上近似且篱笆的训练点数较少,模型没有充分学习篱笆的特征,导致大部分篱笆被模型识别为灌木,部分高程较低的点则被识别为低矮植被。

### 3.4 消融实验

为了验证所提特征提取模块和注意力池化模块对模型的有效性,设计不同的模型架构进行消融实验。实验结果如表 4 所示,其中,proposed-1 表示几何提取的消融,proposed-2 表示池化模块的消融。几何提取模块和注意力池化模块的引入,提高了模型的平均准确率和平均 F1 得分。

分别用 KNN 和最大池化代替改进的基于特征偏差值的特征提取模块和基于注意力机制的特征聚合模

表 3 测试集分类结果混淆矩阵

Table 3 Confusion matrix of classification results for the test set

Category	Power line	Low_vegetation	Impervious surface	Car	Fence/hedge	Roof	Facade	Shrub	Tree
Power line	401	0	0	0	1	161	23	0	14
Low_vegetation	0	82761	8563	357	517	298	402	3957	1835
Impervious surface	0	5973	95649	69	6	217	13	51	8
Car	0	215	134	3157	12	0	7	183	0
Fence/hedge	0	1071	155	58	2544	217	164	3186	27
Roof	178	812	89	12	1	106589	1297	0	70
Facade	1	281	1017	5	17	1984	6239	657	1043
Shrub	8	4892	128	149	379	1134	214	11954	5960
Tree	17	1083	20	67	127	871	835	3549	47657
$R_{precision} / \%$	66.28	85.24	90.44	81.49	70.59	95.62	67.86	50.79	84.18
$R_{recall} / \%$	66.83	83.86	93.79	85.14	34.28	97.75	55.49	48.17	87.89
$s_{F1} / \%$	66.56	84.55	92.08	83.28	46.15	96.67	61.05	49.44	85.99

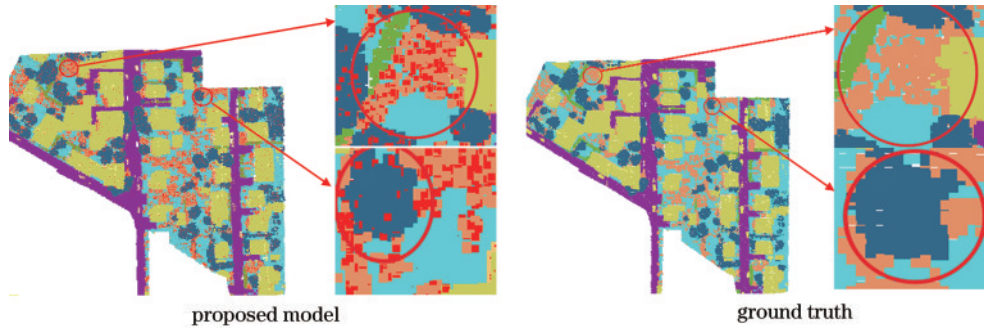


图 7 部分错误分割示意图

Fig. 7 Schematic diagrams of partial error segmentation

表 4 消融实验结果

Table 4 Results of ablation experiments

Category	$s_{F1} / \%$		
	proposed-1	proposed-2	proposed
Power line	58.32	63.14	66.56
Low_vegetation	79.07	79.63	84.55
Impervious surface	87.47	87.86	92.08
Car	81.71	82.15	83.28
Fence/hedge	37.75	36.58	46.15
Roof	92.49	92.89	96.67
Facade	41.27	42.60	61.05
Shrub	42.82	42.85	49.44
Tree	84.1	84.12	85.99
$R_{OA} / \%$	81.32	81.73	86.69
Average $s_{F1} / \%$	67.22	67.98	73.97

块,用以验证改进后的模块的有效性。proposed-1 展示了 KNN 代替基于特征偏差值的特征提取模块的分割结果,较 PointNet 提高了 0.12 个百分点。proposed-2 展示了最大池化代替基于注意力机制的特征聚合模块的分割结果,较 PointNet++ 提高 0.53 个百分点。基于特征偏差值的几何特征提取器可以有效获取点云的深层语义信息,强化模型对不同类别的物体的识别能力,同时,注意力机制通过学习不同特征的权重参数,

能够获取更多有效的特征,提升了模型在小尺度模型如汽车等的准确率。所提基于 PointNet++ 的改进模型能够有效实现点云的场景分割,提取点云的深层几何特征并通过几何信息和局部特征的注意力融合,得到丰富的语义信息,对点云场景中的大尺度和小尺度目标均能较好实现较好的分割。

### 3.5 与现有方法的对比

将所提模型的分割结果与 ISPRS 提供的实验结果以及其他深度学习分割方法进行了对比分析,结果如表 5 所示。在 ISPRS 提供的结果中,RIT\_1 模型基于地面点的高程值,引入点云的光谱信息作为模型的输入,是一种基于点的点云分割模型。NANJ2 由多层次的多特征构建特征图,通过卷积网络提取点云的深层语义信息,实现点云的场景分割。所提基于 PointNet++ 的改进点云分割模型的  $R_{OA}$  较 NANJ2 提高了 1.49 个百分点。NANJ2 通过引入 RGB、高程、强度和粗糙度等特征在灌木的分类中取得了最高的 F1 分数。而所提模型利用点云的三维坐标作为模型的输入,对比各类分割结果可以看出,其在篱笆、汽车等小尺度模型取得了不错的分割结果。

同时,选择 3 种直接分割点云的深度学习方法 PointNet++<sup>[12]</sup>、PointNetSIFT<sup>[20]</sup>、PointNetCNN<sup>[21]</sup> 进



表 5 与现有模型的对比  
Table 5 Comparison with existing models

Method	$s_{F1} / \%$									OA / %
	Power line	Low_vegetation	Impervious surface	Car	Fence/hedge	Roof	Facade	Shrub	Tree	
RIT_1	37.5	77.9	91.5	73.4	18.0	94.0	49.3	45.9	82.5	81.6
BJI_W	13.8	78.5	90.5	56.4	36.3	92.2	53.2	43.3	78.4	81.5
WhuY3	37.1	81.4	90.1	63.4	23.9	93.4	47.5	39.9	78.0	82.3
NANJ2	62	88.8	91.2	66.7	40.7	93.6	42.6	55.9	82.6	85.2
PointNet++[12]	57.9	79.6	90.6	66.1	31.5	91.6	54.3	41.6	77.9	81.2
PointNetSIFT[20]	55.7	80.7	90.9	77.8	30.5	92.5	56.9	44.4	79.6	82.2
PointCNN[21]	61.5	82.7	91.8	75.8	35.9	92.7	57.8	49.1	78.1	83.3
Proposed method	66.56	84.55	92.08	83.28	46.15	96.67	61.05	49.44	85.99	86.69

行对比,所提模型由于引入了注意力机制,在聚合特征时能够自适应提取更好的分割点云的细粒度局部特征,且在几何特征的提取部分,引入特征偏差值,可以更好地地区分不同局部邻域结构的差异,相比于PointNetCNN<sup>[21]</sup>,所提模型精度提高了3.39个百分点。综合结果表明,所提模型通过改进几何提取方式和特征聚合方式,可以有效提高复杂场景下的点云分割精度。

### 3.6 室内场景分割实验

为了进一步检验模型的场景分割能力,在S3DIS大场景室内点云数据集上进行实验,仅利用点云坐标作为数据输入,表6为实验结果。可以看出,所提模型在室内场景分割中也取得较好的结果,但与RandLA-Net<sup>[22]</sup>、ConvPoint<sup>[23]</sup>等仍有差距,这是因为RandLA-Net<sup>[22]</sup>使用了随机采样的方法,并构建了由局部空间编码器、注意力池化模块及扩张残差模块组成的局部特征聚合模块为模型提供丰富的语义特征,其在空间编码器部分,利用相对位置编码,链接中心点与邻近点及坐标差值与欧氏距离获得了采样点的冗余特征,为了解决特征冗余在增强特征的同时带来的特征数量增多的问题,利用了注意力池化来聚合特征以及扩展残差块增强模型的计算效率。与所提模型相比,RandLA-Net<sup>[22]</sup>的特征更加丰富同时在提取局部特征时引入注意力池化进一步强化了局部特征。ConvPoint<sup>[23]</sup>的技术核心在于如何更好地利用卷积来处理点云,为了解决这一问题,ConvPoint<sup>[23]</sup>采用了一种离散卷积神经网络,利用离散公式聚合点云的空间域和特征域,为了确保卷积核对输入尺度的鲁棒性,对输入的点云作了归一化处理,降低了模型对点云大小

表 6 S3DIS数据集上的分割结果

Table 6 Segmentation results on S3DIS dataset

Method	$R_{OA} / \%$
PointNet	78.6
PointNet++	81.0
RandLA-Net	88.0
ConvPoint	87.9
ConvPoint-NoColor	85.2
Proposed method	83.1

的敏感性。与所提模型相比,RandLA-Net<sup>[22]</sup>的特征更加丰富同时在提取局部特征时引入注意力池化进一步强化了局部特征。

## 4 结 论

本研究针对PointNet++<sup>[12]</sup>在Lidar点云场景分割任务中未能深层挖掘点云的局部几何特征及在聚合特征时特征丢失的问题,设计了一种基于特征偏差值和注意力机制的Lidar点云语义分割模型,并在ISPRS提供的Vaihingen数据集以及斯坦福的S3DIS上验证了模型的有效性。

对于复杂的点云场景,基于特征偏差值的几何特征提取模块能够有效提取点云的深层几何特征,充分挖掘点云的语义信息。此外,为了避免在场景分割任务中出现特征丢失的问题,在聚合特征时引入注意力机制,对于输入的特征,能够自适应地学习不同特征的重要性,充分利用提取到的几何信息,减少冗余特征的干扰,避免部分有价值信息的丢失,提高模型在复杂场景中的分割能力。

所提模型能够有效完成Lidar点云的场景分割任务,与ISPRS提供的实验及PointNet++<sup>[12]</sup>、PointNetSIFT<sup>[20]</sup>、PointNetCNN<sup>[21]</sup>的结果对比中,所提模型取得了最高的平均精度,同时在小尺度物体如汽车、篱笆等也取得了最高的F1得分,这说明所提改进模型可以有效挖掘深层信息,有效实现Lidar点云的场景分割。

所提模型在处理篱笆、灌木等结构近似且邻近的点云时,存在错误识别部分样本的问题,接下来可从引入RGB特征并考虑如何增强点云在小样本点数量时的分割精度方面开展进一步研究。

## 参 考 文 献

- [1] 叶语同,李必军,付黎明.智能驾驶中点云目标快速检测与跟踪[J].武汉大学学报(信息科学版),2019,44(1):139-144,152.  
Ye Y T, Li B J, Fu L M. Fast object detection and tracking in laser data for autonomous driving[J]. Geomatics and Information Science of Wuhan University, 2019, 44(1):

- 139-144, 152.
- [2] 黄鹤, 佟国峰, 夏亮, 等. SLAM 技术及其在测绘领域中的应用[J]. 测绘通报, 2018(3): 18-24.  
Huang H, Tong G F, Xia L, et al. SLAM technology and its application in surveying and mapping[J]. Bulletin of Surveying and Mapping, 2018(3): 18-24.
- [3] Yang B S, Huang R G, Li J P, et al. Automated reconstruction of building LoDs from airborne LiDAR point clouds using an improved morphological scale space [J]. Remote Sensing, 2016, 9(1): 14.
- [4] 杨必胜, 董震. 点云智能研究进展与趋势[J]. 测绘学报, 2019, 48(12): 1575-1585.  
Yang B S, Dong Z. Progress and perspective of point cloud intelligence[J]. Acta Geodaetica et Cartographica Sinica, 2019, 48(12): 1575-1585.
- [5] Maturana D, Scherer S. VoxNet: a 3D Convolutional Neural Network for real-time object recognition[C]//2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), September 28-October 2, 2015, Hamburg, Germany. New York: IEEE Press, 2015: 922-928.
- [6] Wu Z R, Song S R, Khosla A, et al. 3D ShapeNets: a deep representation for volumetric shapes[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 1912-1920.
- [7] Klovov R, Lempitsky V. Escape from cells: deep kd-networks for the recognition of 3D point cloud models [C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 863-872.
- [8] Tatarchenko M, Dosovitskiy A, Brox T. Octree generating networks: efficient convolutional architectures for high-resolution 3D outputs[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2107-2115.
- [9] Su H, Maji S, Kalogerakis E, et al. Multi-view convolutional neural networks for 3D shape recognition [C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2016: 945-953.
- [10] Wang C, Pelillo M, Siddiqi K. Dominant set clustering and pooling for multi-view 3D object recognition[C]//Proceedings of the British Machine Vision Conference 2017, September, 2017, London, UK. London: British Machine Vision Association, 2017: 1-12.
- [11] Charles R Q, Hao S, Mo K C, et al. PointNet: deep learning on point sets for 3D classification and segmentation [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 77-85.
- [12] Qi C R, Yi L, Su H, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space[EB/OL]. (2017-06-07)[2023-02-03]. <https://arxiv.org/abs/1706.02413>.
- [13] Wang Y, Sun Y B, Liu Z W, et al. Dynamic graph CNN for learning on point clouds[J]. ACM Transactions on Graphics, 38(5): 146.
- [14] Lin Z H, Huang S Y, Wang Y C F. Learning of 3D graph convolution networks for point cloud analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(8): 4212-4224.
- [15] Zhang Y X, Rabbat M. A graph-CNN for 3D point cloud classification[C]//2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), April 15-20, 2018, Calgary, AB, Canada. New York: IEEE Press, 2018: 6279-6283.
- [16] 杨晓文, 王爱兵, 韩燮, 等. 基于KNN-PointNet的点云语义分割[J]. 激光与光电子学进展, 2021, 58(24): 2410013.  
Yang X W, Wang A B, Han X, et al. Point cloud semantic segmentation based on KNN-PointNet[J]. Laser & Optoelectronics Progress, 2021, 58(24): 2410013.
- [17] 赖鸣, 赵健康, 刘传奇, 等. 基于CAFF-PointNet的机载LiDAR点云语义分割[J]. 激光与光电子学进展, 2021, 58(20): 2028004.  
Lai M, Zhao J K, Liu C Q, et al. Semantic segmentation of LiDAR point cloud based on CAFF-PointNet[J]. Laser & Optoelectronics Progress, 2021, 58(20): 2028004.
- [18] Wu Y X, He K M. Group normalization[J]. International Journal of Computer Vision, 2020, 128(3): 742-755.
- [19] Touvron H, Bojanowski P, Caron M, et al. ResMLP: feedforward networks for image classification with data-efficient training[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(4): 5314-5321.
- [20] Jiang M Y, Wu Y R, Zhao T Q, et al. PointSIFT: a SIFT-like network module for 3D point cloud semantic segmentation[EB/OL]. (2018-07-02)[2023-02-03]. <https://arxiv.org/abs/1807.00652>.
- [21] Li Y, Bu R, Sun M, et al. PointCNN: convolution on X-transformed points[EB/OL]. (2018-01-23)[2023-02-06]. <https://arxiv.org/abs/1801.07791>.
- [22] Hu Q Y, Yang B, Xie L H, et al. RandLA-net: efficient semantic segmentation of large-scale point clouds[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 11105-11114.
- [23] Boulch A. ConvPoint: continuous convolutions for point cloud processing[EB/OL]. (2019-04-04)[2023-02-03]. <http://arxiv.org/abs/1904.02375>.