

# 基于注意力机制点卷积网络的机载 LiDAR 点云分类

王利媛, 付丽华\*

中国地质大学(武汉)数学与物理学院, 湖北 武汉 430074

**摘要** 机载 LiDAR 点云地物丰富且密度不均匀, 对其进行精准高效的分类是遥感和摄影测量领域的一项关键任务。针对其密度不均匀的特点, 引入一种密度相关的点云卷积算子——PointConv; 提出注意力机制模块, 修正提取局部信息的重要程度, 增强网络对不同点云实例的识别能力。在城市室外场景机载点云数据集 GML\_DataSetA 和三维语义标记基准数据集 ISPRS Vaihingen 上的分类结果验证了所提方法的有效性。

**关键词** 遥感; 机载 LiDAR 点云; 点云分类; 密度加权; 注意力机制

中图分类号 P237

文献标志码 A

DOI: 10.3788/LOP202259.1028007

## Airborne LiDAR Point Cloud Classification Based on Attention Mechanism Point Convolutional Network

Wang Liyuan, Fu Lihua\*

School of Mathematics and Physics, China University of Geosciences (Wuhan), Wuhan 430074, Hubei, China

**Abstract** Airborne LiDAR point cloud features are abundant, but their density is uneven. Efficient classification for the airborne LiDAR point cloud is a key task in remote sensing and photogrammetry. Because the density of the point cloud is not uniform, a density-dependent point cloud convolution operator, PointConv, was introduced to perform density weighting on the basis of traditional three-dimensional (3D) convolution. At the same time, the attention mechanism module was proposed to correct the importance of extracted local information and improve the ability of the network for identifying different point cloud instances. The effectiveness of the proposed method is demonstrated by the classification results on the GML\_DataSetA urban outdoor scene airborne point cloud dataset and the ISPRS Vaihingen 3D semantic marker reference dataset.

**Key words** remote sensing; airborne LiDAR point cloud; point cloud classification; density weighting; attentional mechanism

## 1 引言

近年来, 自动驾驶、高精地图、智慧城市、虚拟现实、增强现实等概念的提出极大地推动了大规模

三维场景语义理解和分析的发展。机载 LiDAR 扫描所获取的点云能够快速获取地形表面高精度信息, 对其进行精准分类在三维建模、城市道路提取等方面有着十分重要的现实意义。机载 LiDAR 点云

收稿日期: 2021-02-23; 修回日期: 2021-04-30; 录用日期: 2021-05-27

基金项目: 国家重点研发计划(2018YFC1503705)、湖北省教育厅科学技术研究项目(B2017597)、湖北省重点实验室开放基金(SMIL-2018-06)

通信作者: \*lihuafu@cug.edu.cn

数据量大,密度不均匀,而且分布稀疏<sup>[1]</sup>。另外,机载 LiDAR 点云所包含的地物丰富,场景复杂,导致与一般点云数据分类相比,LiDAR 点云分类更加困难。

机载 LiDAR 点云分类又称为点云语义分割。早期的研究工作集中在手动提取点云特征(如法线、曲率、平面度、点特征直方图和方向直方图等),然后选取合适的分类器进行分类。Mallet 等<sup>[2]</sup>利用支持向量机对包含空间特征、高度特征、特征值特征等 27 个分量的特征向量进行分类。He 等<sup>[3]</sup>利用迁移学习的思想,首先从 VoxNet 模型的全连接层提取特征向量,再利用 TrAdaBoost 算法进行点云分类。Kim 等<sup>[4]</sup>针对点域和特征域两组不同特征,用随机森林算法对电力线场景进行分类。但是,基于人工提取特征的方法在处理大规模场景时常常不能取得令人满意的效果<sup>[5]</sup>。

深度学习的快速发展使研究者们尝试将不规则的机载 LiDAR 点云转换为规则的二维图像或者三维体素,然后使用二维卷积神经网络(2DCNN)或 3DCNN 进行分类。对于机载 LiDAR 点云数据中的每个点,Yang 等<sup>[6]</sup>提取窗口内其邻近点的局部几何特征、全局几何特征和全波形特征,并转换成图像,使用 2DCNN 对特征图像进行分类。Yang 等<sup>[7]</sup>提取如高度、平面度、球度和偏离角方差等特征,生成多个尺度的特征图像,使用多尺度 2DCNN 实现点云分类。Huang 等<sup>[8]</sup>计算整个点云的边界框,从点云中选取一个中心点,在其局部进行体素转化,然后使用 3DCNN 完成分类。由于 3DCNN 训练需要占用大量内存和计算时间,Zhang 等<sup>[9]</sup>基于稀疏子流形卷积网络,利用 3D 数据固有的稀疏性,有效地对体素化的机载 LiDAR 点云进行语义分割。然而,由于数据类型的转化,此类方法损失了有价值的几何信息<sup>[10]</sup>,影响最终的分类精度<sup>[11]</sup>。

PointNet<sup>[12]</sup>是第一个直接在原始点云上操作的深度网络,在点云的分类和语义分割上都表现出良好的性能。Yousefhussien 等<sup>[13]</sup>修改了 PointNet,以原始三维坐标以及每个点对应的从二维地理标记图像中提取的三个光谱特征作为网络输入,增加了特征维数,实现了端到端的分类。Winiwarter 等<sup>[14]</sup>提出基于 PointNet++<sup>[15]</sup>的点云分类算法,对输入点集中的每个点提取 54 维特征向量。这些特征向量包括采用 3 种不同设置的参数  $K$  进行最近邻搜索得到的一个 6 维特征值特征向量和一个 12 维自旋图像特征向量,从而实现对点云的分类。Wen 等<sup>[16]</sup>提出一种定向约束的全卷积神经网络模型,从投射

的二维接受域中提取三维点集的局部代表特征,并引入了一种多尺度 CNN 来实现多尺度点特征学习。Li 等<sup>[17]</sup>提出一种基于几何感知集合、密集层次结构和高度注意模块的几何注意网络,实现对如屋顶、电力线等几何实例的点云分类。

为了降低机载 LiDAR 点云密度不均匀对分类的影响,本文引入一种密度相关的点云卷积算子——PointConv<sup>[18]</sup>,在此基础上构建具有下采样块和上采样块的多尺度 CNN,以实现多尺度特征学习;引入注意力机制模块,修正提取的局部信息的重要程度,可以有效提升网络的分类精度。与 PointNet++ 和 PointConv 相比,所提 point cloud convolutional network based on attention mechanism (PCNNAM) 在城市室外场景机载点云数据集 GML\_DataSetA 上取得了最好的性能。此外,在 ISPRS Vaihingen 三维语义标记基准数据集上,对于点数较少的电力线以及较难分类的建筑物表面,PCNNAM 都取得了最好的分类效果。

## 2 基于注意力机制的点卷积网络 PCNNAM

机载 LiDAR 点云地物丰富且密度不均匀。传统的三维卷积无法解决密度不均匀的问题,引入密度相关的点云卷积算子——PointConv,基于该算子构建 U-Net 网络结构,直接用于机载 LiDAR 点云分类。在此基础上添加注意力机制模块,构建 PCNNAM,增强网络对不同点云实例的识别能力。

### 2.1 密度相关的点云卷积算子——PointConv

机载 LiDAR 点云属于点云数据,具有密度不均匀的特点。其作为三维数据的一种表达形式,是由带有位置信息  $(x, y, z)$  以及其他特征信息如颜色、强度等的点所构成的集合。图 1 展示了规则二维图

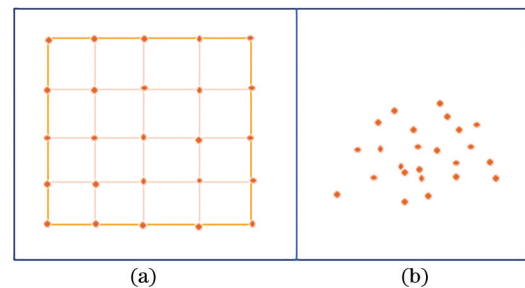


图 1 规则二维图像和点云的对比示意图。(a) 图像网格; (b) 点云

Fig. 1 Comparison schematic of regular 2D image and point cloud. (a) Image grid; (b) point cloud

像和点云的对比。相较于图 1(a)中规则网格不同位置的像素点,图 1(b)的点云分布是散乱的,点云不同点在每个局部区域的相对位置是不同的。因此,传统的三维卷积操作无法直接应用于点云。

为了使卷积操作适用于点云,Wu 等<sup>[18]</sup>提出了一种能够直接作用于点云的密度相关点云卷积算

子——PointConv。引入一种密度相关权重,其对普通的三维卷积起到重加权的作用,适用于密度不均匀的点云数据。

现有点  $p(x, y, z)$  以及其局部邻域  $G$  中点云的特征  $F(x + \Delta x, y + \Delta y, z + \Delta z)$ 。

传统的三维离散卷积为

$$\text{Conv}(W, F)_{xyz} = \sum_{(\Delta x, \Delta y, \Delta z) \in G} W(\Delta x, \Delta y, \Delta z) F(x + \Delta x, y + \Delta y, z + \Delta z), \quad (1)$$

可近似为

$$\text{Conv}(W, F)_{xyz} = \iiint_{(\Delta x, \Delta y, \Delta z) \in G} W(\Delta x, \Delta y, \Delta z) F(x + \Delta x, y + \Delta y, z + \Delta z) d\Delta x d\Delta y d\Delta z. \quad (2)$$

根据有偏样本的蒙特卡罗估计<sup>[19]</sup>:

$$\int f(x) dx = \int \frac{f(x)}{\rho(x)} \rho(x) dx \approx \frac{1}{N} \sum_{i=1}^N \frac{f(x_i)}{\rho(x_i)}, \quad (3)$$

式中:  $\rho(x_i)$  为样本  $x_i$  的概率密度;  $N$  为输入点云中点的个数。那么式(2)可离散化为

$$\text{Conv}(W, F)_{xyz} = \sum_{(\Delta x, \Delta y, \Delta z) \in G} \frac{1}{d(\Delta x, \Delta y, \Delta z)} W(\Delta x, \Delta y, \Delta z) F(x + \Delta x, y + \Delta y, z + \Delta z). \quad (4)$$

Wu 等<sup>[18]</sup>将离散化的卷积命名为 PointConv, 即

$$\text{PointConv}(W, F)_{xyz} = \sum_{(\Delta x, \Delta y, \Delta z) \in G} \frac{1}{d(\Delta x, \Delta y, \Delta z)} W(\Delta x, \Delta y, \Delta z) F(x + \Delta x, y + \Delta y, z + \Delta z), \quad (5)$$

相较于传统的三维离散卷积,这里的  $\frac{1}{d(\Delta x, \Delta y, \Delta z)}$  是局部区域  $G$  中点云密度的倒数。

图 2 展示了局部区域内的点云分布,可以看出,点云的分布是不均匀的。

中心点  $p_0$  附近的点  $p_3, p_4, p_5, p_6, p_7, p_8, p_9, p_{10}$  非常接近,相较于  $p_1, p_2$ , 密度较大。在卷积操作提取特征的过程中,和二维卷积提取规则图像的局部特征不同,对于密度较大的区域,应当赋予较小的密度权重,从而起到一个重加权的作用。因此这里的密度加权是必要的。

令  $S(\Delta x, \Delta y, \Delta z) = \frac{1}{d(\Delta x, \Delta y, \Delta z)}$ , 则经过一

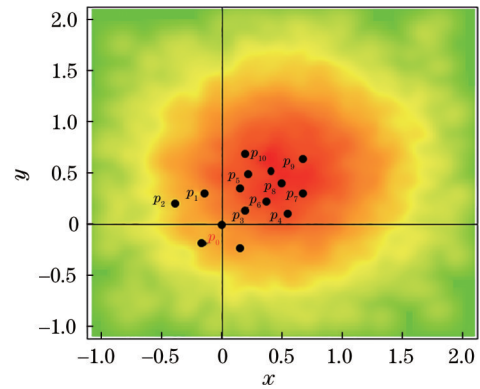


图 2 局部区域的点云示意图

Fig. 2 Schematic of point cloud in local area

次 PointConv, 输出可表示为

$$F_{\text{out}} = \text{PointConv}(W, F)_{xyz} = \sum_{(\Delta x, \Delta y, \Delta z) \in G} S(\Delta x, \Delta y, \Delta z) W(\Delta x, \Delta y, \Delta z) F(x + \Delta x, y + \Delta y, z + \Delta z) = \sum_{k=1}^K \sum_{c_{\text{in}}=1}^{C_{\text{in}}} S(k) W(k, c_{\text{in}}) F_{\text{in}}(k, c_{\text{in}}), \quad (6)$$

式中:  $k$  表示每个中心点局部邻域内的  $k$  个邻近点;  $c_{\text{in}}$  是这  $k$  个邻近点的特征。权重函数  $W(\Delta x, \Delta y, \Delta z)$  可由多层感知机从输入点云特征中直接学习得到。关于密度权重  $S(\Delta x, \Delta y, \Delta z)$ , 可以

先利用核密度估计<sup>[20]</sup>计算出每个点在点云中的密度,然后利用多层感知机从密度中学习出密度权重  $S(\Delta x, \Delta y, \Delta z)$ 。

### 2.2 网络结构

机载 LiDAR 点云具有密度不均匀的特点, PointConv 在原始三维卷积的基础上进行密度加权。为了实现多尺度特征学习,现以 PointConv 为点云卷积算子,采用下采样模块提取多尺度特征,上采样模块获取所有原始点的点特征,构建网络,将网络直接作用于机载 LiDAR 点云。整个网络结构是一个 U-Net 网络,通过上采样和下采样实现特征的传播,使用跳跃连接的方式,将具有相同分辨率的特征连接起来,再通过两个全连接层,实现对每个点的分类。整个网络结构如图 3 所示,  $m$  表示

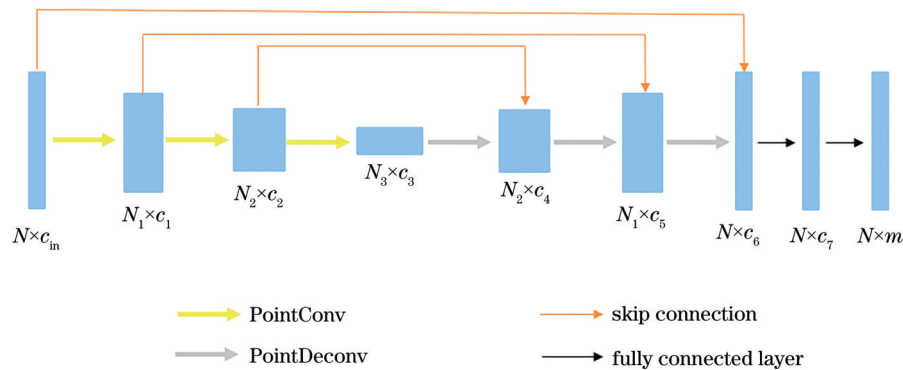


图 3 U-Net(PointConv)点云语义分割示意图

Fig. 3 U-Net(PointConv) schematic of semantic segmentation of point cloud

### 2.3 注意力机制模块

机载 LiDAR 点云含有一些几何形状丰富的点云实例,比如低矮植被、建筑物等,具有地物丰富的特点。受到近年来不同任务中注意力机制<sup>[21-23]</sup>研究的启发,为了使网络能够充分捕捉点云实例的局部细节,设计了一种基于原始点云位置信息的注意力机制模块。将其融入学习到局部信息中,对不同的局部信息赋予不同的分数,修正局部信息的重要程度。

注意力机制模块分为信息修正和信息融合两部分,网络结构示意图如图 4 所示。

图 4 中,  $N \times c_{mid}$  为提取到的局部信息,定义其为  $L(p)$ 。首先通过 PointNet 从整片点云的位置信息  $\phi(p)$  中进行全局特征提取,然后将学习到的全局特征重复  $N$  次,通过归一化 softmax 函数,将输出范围固定在区间  $[0, 1]$  内,表示网络提取对局部信息的注意力度(值越大,注意力越高),用  $S(p)$  来表示。那么  $S(p)$  可定义为

$$S(p) = \text{softmax}\{\text{MLP}[\phi(p)]\}, \quad (7)$$

式中:  $\text{MLP}(\cdot)$  为多层感知机。再将注意力度与点云提取到的局部信息对应相乘,以残差连接的方式

输入点云的类别数目。

下采样模块以原始点云作为输入,采样层通过最远点采样确定出中心点,然后通过球查询或者  $K$  近邻的方式确定邻近点,再在每个小的局部邻域内使用 PointConv 提取特征。为了获取原始点的特征,在上采样模块,首先通过反距离加权插值对输入点云进行上采样,然后以上采样后的每个点作为中心点,以和下采样同样的方式确定出局部邻域,再在局部邻域内使用 PointConv,以实现特征的传递。实际上这一过程可以看作是反卷积,称其为 PointDeconv。

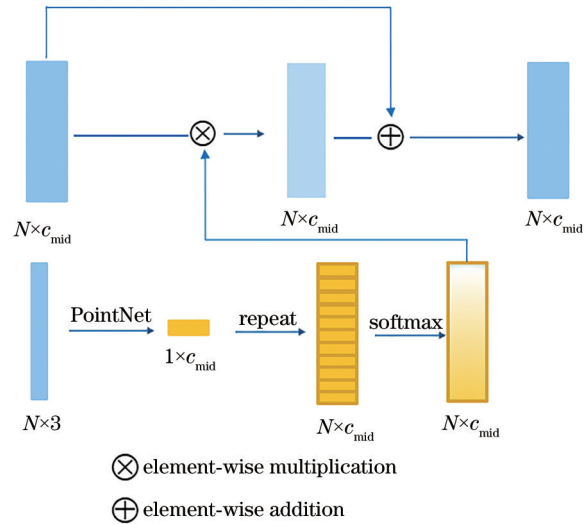


图 4 注意力机制模块结构示意图

Fig. 4 Schematic of attention mechanism module structure

与未经过注意力加权的局部信息相加,得到最终的输出结果,为

$$O(p) = L(p) \times [1 + S(p)]. \quad (8)$$

### 2.4 PCNNAM

加入注意力机制的网络结构示意图如图 5 所

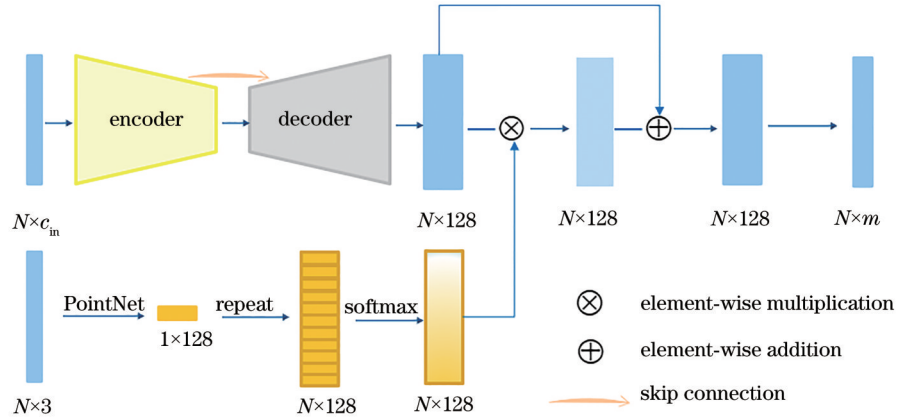


图 5 基于注意力机制的点卷积网络(PCNNAM)

Fig. 5 Point cloud convolutional network based on attention mechanism (PCNNAM)

示。通过下采样和上采样进行局部信息提取,注意力机制模块修正局部信息的重要程度,最后通过两个全连接层,得到分类结果。

实验的硬件环境为 GPU NVIDIA GTX1080 Ti 11GB 工作站。采用 Tensorflow 框架, Ubuntu 16.04 系统。设置 batch size 为 6, 初始学习率为 0.001(每隔 3000 步, 学习率减小为原来的 1/2), 采用 Adam 优化算法, 训练 1000 轮直至网络收敛。

### 3 实验部分

#### 3.1 评价指标

选择精确度、召回率以及 F1 分数来评价网络对点云分类的性能。一般来说, 精确度用于评估所有类别的总体分类精度, 定义为正确分类点在总测试点中的百分比。此外, F1 分数考虑了分类模型的精度和召回率, 定义为每个类别的精确度和召回率的调和平均值。

具体公式分别为

$$P_{\text{precision}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}}, \quad (9)$$

$$P_{\text{recall}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}}, \quad (10)$$

$$P_{\text{F}_1 \text{ score}} = 2 \times \frac{P_{\text{precision}} \times P_{\text{recall}}}{P_{\text{precision}} + P_{\text{recall}}}, \quad (11)$$

式中:  $N_{\text{TP}}$  表示本类中预测正确的个数;  $N_{\text{FP}}$  表示将其他类预测为本类的个数;  $N_{\text{FN}}$  表示本类标签预测为其他类的个数。

#### 3.2 GML\_DataSetA 点云数据集实验

##### 3.2.1 数据和数据处理

选取莫斯科国立大学图形与媒体实验室提供的城市室外场景机载点云数据集 GML\_DataSetA 对所提方法进行验证。

原始点云数据分为训练集和测试集两部分, 分别包含 1073989 和 1002668 个点, 分为地面、建筑物、车、树和低矮的植被、未知类别的噪声点, 训练集和测试集图像如图 6 所示。剔除数据集中的噪声点以及车, 选取地面、建筑物、树和低矮的植被 4 类对所提方法进行测试。各类点云分布的详细信息如表 1 所示。实验将训练集和测试集分别分成 50 m × 50 m 大小的小块, 由于数据集形状不规则、密度不均匀, 每个数据块大约包含几百到几万个点。因此, 训练网络时, 随机地在每一小块数据集上选择 8192 个点, 同时为了防止过拟合, 训练的过

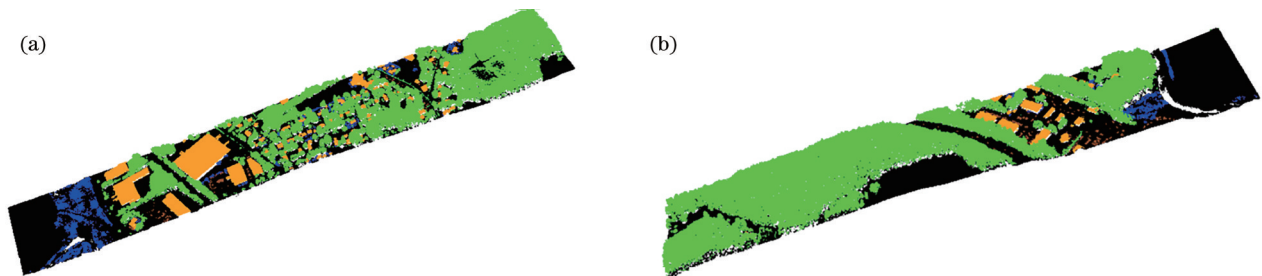


图 6 GML\_DataSetA。(a) 训练集示意图; (b) 测试集示意图

Fig. 6 GML\_DataSetA. (a) Schematic of training set; (b) schematic of test set

表1 GML\_DataSetA 训练集测试集中各类点云分布  
Table 1 All kinds of point cloud distribution in GML\_DataSetA training set and test set

Data set	Ground	Building	Tree	Low vegetation	All
Train set	557142	98244	381677	35093	1072156
Test set	439989	19592	531852	7758	999191

程中,在每一块上随机删除一些点,并设置 dropout ratio 为 12.5%。

### 3.2.2 数据加权

为了使网络更多地关注较少类别的点,训练网络时,在网络的损失函数部分添加每个类别的特定权重系数。权重的计算公式为

$$W_c = \frac{1}{\ln(\alpha + \frac{N_c}{\sum_{c=1}^C N_c})}, \quad (12)$$

式中:  $W_c$  表示第  $c$  类的权重;  $N_c$  表示第  $c$  类点的个数;  $C$  表示点的总个数;  $\alpha$  为类别平衡系数。选择交叉熵作为网络的损失函数,经过数据加权后,最终的损失函数为

$$L = \sum_{i=1}^N w_i \times \sum_{c=1}^C [y_{ic} \log p_{ic} + (1 - y_{ic}) \log (1 - p_{ic})], \quad (13)$$

式中:  $N$  表示训练时每一块中点的个数;  $p_{ic}$  和  $y_{ic}$  分别代表第  $i$  个点在第  $c$  类上的预测标签和真实标签。

### 3.2.3 实验结果

表 2 展示了 PCNNAM 对测试集分类结果的混淆矩阵。PCNNAM 在地面和树的分类上都获得了较高的 F1 分数;但对低矮植被的分类结果较差,大多数低矮植被都被错误地分成了地面和树。原因是低矮植被的点数较少,且呈现出与树相似的空间分布和拓扑特征;其次高程较小使其和地面属性更加接近。这些因素导致模型训练不足,模型得到较差的分类结果。

在数据加权部分,选取不同的类别平衡系数进

表 2 基于注意力机制的点卷积网络对测试集分类结果的混淆矩阵

Table 2 Confusion matrix of classification results in test set obtained by PCNNAM

Class	Ground	Building	Tree	Low vegetation
Ground	428997	2228	6862	1900
Building	4499	11748	2540	805
Tree	26975	5210	497374	2293
Low vegetation	2539	591	3283	1345
Precision	0.975	0.60	0.935	0.173
Recall	0.927	0.594	0.975	0.212
F1 score	0.950	0.597	0.955	0.191

行对比实验,不加权以及  $\alpha$  设置为 1.1, 1.2, 1.3, 1.4, 1.5 的网络分类结果如表 3 所示。在不使用数据加权的情况下,总体精度达到了 87.4%,但 F1 分数低于加权后的分类结果。这是由于当数据分布不均匀时,网络对数量较少的点不够敏感,从而无法很好地对其进行分类。使用数据加权后,F1 分数大部分都得到提升,并且当  $\alpha$  取 1.2 时,网络的分类精确度达到了 94%,总体 F1 分数达到了 67.3%。

表 3 不同类别平衡系数下的分类结果

Table 3 Classification results under different coefficients for class balance

Parameter	None	$\alpha=1.1$	$\alpha=1.2$	$\alpha=1.3$	$\alpha=1.4$	$\alpha=1.5$
Overall accuracy	0.874	0.860	0.940	0.912	0.911	0.788
Overall F1 score	0.487	0.487	0.673	0.559	0.597	0.468

### 3.2.4 消融实验

为了验证点云卷积算子 PointConv 密度加权及注意力机制的有效性,以直接作用于点云的 PointNet++ 为基准进行消融实验。表 4 展示了各网络在 GML\_DataSetA 测试集的分类结果。每个网络均进行  $\alpha=1.2$  的数据加权处理。可以看出,

表 4 不同网络对 GML\_DataSetA 测试集的分类结果

Table 4 Classification results of the GML\_DataSetA test set under different networks

Method	Density weighted	Attentional mechanism	F1 score				OA	Average F1 score
			Ground	Building	Tree	Low vegetation		
PointNet++	×	×	0.832	0	0.838	0	0.823	0.417
U-Net(PointConv)	✓	×	0.937	0.379	0.908	0.122	0.895	0.587
PCNNAM	✓	✓	0.950	0.597	0.955	0.191	0.940	0.673

PointNet++ 无法识别出建筑物以及低矮植被这些较少点的类别,而 U-Net(PointConv)的分类结果较之有了很大的提升,这表明了密度加权的有效性。其次,与 U-Net(PointConv)相比,PCNNAM 分类结果的总体精确度(OA)以及各类 F1 分数都有一定的提升。

从表 4 的第 5 列可以看出,PCNNAM 对建

筑物的 F1 分数相较于 U-Net(PointConv)的 F1 分数高出 0.218,这表明注意力机制模块通过修正局部信息的重要程度增强了网络的细粒度识别能力。

图 7 展示了不同模型分类结果和数据集真实分布的示意图。箭头所指部分突出显示了注意力机制模块的有效性。

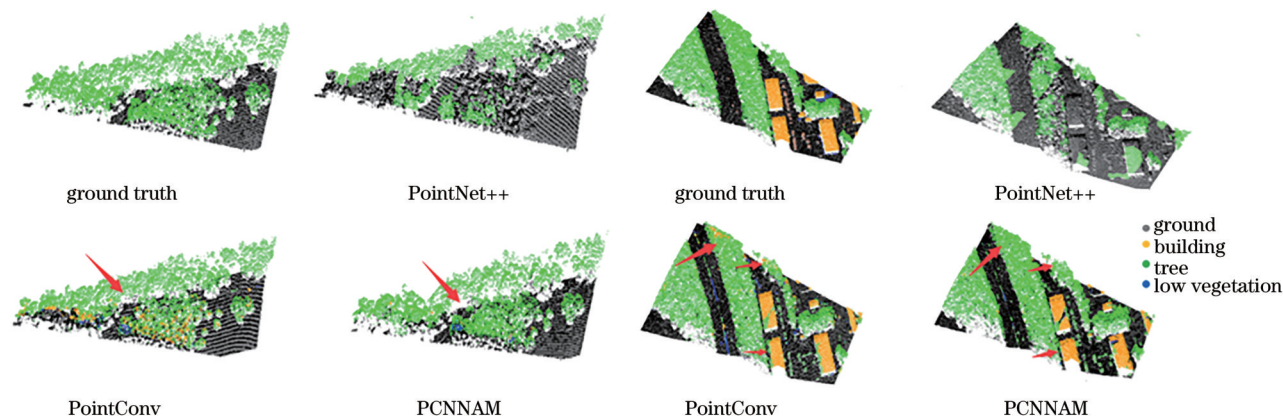


图 7 不同网络对 GML\_DataSetA 测试集的分类结果以及数据集真实分布示意图

Fig. 7 Classification results of test set under different networks and the real distribution diagram of GML\_DataSetA data set

### 3.3 在 ISPRS Vaihingen 三维语义标记基准数据集上的实验结果

为了进一步说明所提方法的有效性,在 ISPRS Vaihingen 三维语义标记基准数据集<sup>[24]</sup>进行实验。两组点云数据均于 2008 年 8 月利用 ALS50 系统在航高约 500 m、视场角 45°的条件下采集,点密度约为 4~8 个/m<sup>2</sup>。每个激光脚点包括三维坐标、反射强度、回波次数以及类别标签。该数据集将地物分为 9 个类别:电力线、低矮植被、不透水表面、车辆、栅栏、屋顶、建筑物表面、灌木、树木。训练集和测试集中各类点云分布的详细情况如表 5 所示。

由于各点的回波次数值相差不大,选取点云的  $x, y, z$  坐标及强度信息。对于训练集,随机从整块点云中选择一个 30 m × 30 m × 40 m 的长方体区域,然后从长方体中随机选择 8192 个点作为网络输入,同样采用  $\alpha=1.2$  的数据加权方式。对于测试集,将整块点云分成多个大小约为 30 m × 30 m 的小块,依次输入网络,完成对每个点语义标签的预测,实现机载 LiDAR 的点云分类。最终的总体分类精度为 81.6%,总体 F1 分数为 67.8%。各类别分类的混淆矩阵如表 6 所示。

表 7 为 PCNNAM 在 ISPRS Vaihingen 三维语

表 5 ISPRS Vaihingen 三维语义标记基准数据集中各类点云分布

Table 5 Distribution of various point clouds in the ISPRS Vaihingen 3D semantic marker benchmark data set

Category	Train set	Test set
Power line	546	600
Low vegetation	180850	98690
Impermeable surface	193723	101986
Car	4614	3708
Fence	12070	742
Roof	152045	109048
Building surface	27250	11224
Shrub	47605	24818
Tree	135173	54226
All	753876	411722

义标记基准数据集测试集上的实验结果与 ISPRS Vaihingen 三维语义标记竞赛公布的其他实验结果比较,显示了每一类的 F1 分数、总体精确度和总体 F1 分数。虽然 PCNNAM 没有取得最佳效果,但对于比较难区分的建筑物表面,PCNNAM 实现了最高的 F1 分数,这说明注意力机制模块提升了模型的细粒度识别能力,使模型充分捕捉到了点云实例的局部细节。

表 6 基于注意力机制的点卷积网络对 ISPRS Vaihingen 三维语义标记基准数据测试集分类结果的混淆矩阵  
Table 6 Confusion matrix of classification results for ISPRS Vaihingen 3D semantic marker benchmark data set obtained by PCNNAM

Category	Power line	Low vegetation	Impermeable surface	Car	Fence	Roof	Building surface	Shrub	Tree
Power line	328	1	0	0	0	76	3	2	190
Low vegetation	0	77586	9983	154	528	1137	407	6684	2211
Impermeable surface	0	7159	94183	36	9	223	82	233	61
Car	0	121	113	2439	302	94	4	625	10
Fence	0	729	107	46	1704	289	10	3759	778
Roof	124	1416	126	3	123	95357	683	1450	9765
Building surface	14	642	75	41	50	1826	4611	1596	2369
Shrub	0	3537	214	161	1073	1513	344	11364	6612
Tree	3	844	55	44	257	740	259	5388	46635
Precision	0.699	0.843	0.898	0.834	0.421	0.941	0.720	0.365	0.680
Recall	0.547	0.786	0.923	0.658	0.230	0.874	0.411	0.458	0.860
F1 score	0.614	0.834	0.911	0.736	0.297	0.907	0.523	0.406	0.759

表 7 基于注意力机制的点卷积网络结果和 ISPRS Vaihingen 三维语义标记竞赛公布的其他的实验结果比较  
Table 7 Comparison of the results of PCNNAM and other experiments published by the ISPRS Vaihingen 3D semantic markup competition

Method	F1 score									OA	Average F1 score
	Power line	Low vegetation	Impermeable surface	Car	Fence	Roof	Building surface	Shrub	Tree		
UM	0.461	0.790	0.891	0.477	0.052	0.920	0.527	0.409	0.779	0.808	0.590
WhuY3	0.371	0.814	0.901	0.634	0.239	0.934	0.475	0.399	0.780	0.823	0.616
LUH	0.596	0.775	0.911	0.731	0.340	0.942	0.563	0.466	0.831	0.816	0.684
BIJW	0.138	0.785	0.905	0.564	0.363	0.922	0.532	0.433	0.784	0.815	0.603
RIT_1	0.375	0.779	0.915	0.734	0.180	0.940	0.493	0.459	0.825	0.816	0.633
NANJ	0.620	0.888	0.912	0.667	0.407	0.936	0.426	0.559	0.826	0.852	0.693
WhuY4	0.425	0.827	0.914	0.747	0.537	0.943	0.531	0.479	0.828	0.849	0.692
U-Net (PointConv)	0.614	0.834	0.911	0.736	0.297	0.907	0.523	0.406	0.759	0.812	0.663
Propoosed method	0.589	0.809	0.902	0.740	0.319	0.911	0.587	0.435	0.768	0.8124	0.673

## 4 结 论

引入一种密度相关的点云卷积算子——PointConv, 基于其构建 U-Net 网络, 并添加注意力机制模块来实现机载 LiDAR 点云的分类。机载 LiDAR 点云数据具有密度不均匀的特点, 利用 PointConv 中的密度权重对传统三维卷积进行了重加权, 使其能够直接作用于原始的点云数据。根据机载 LiDAR 点云地物丰富的特点, 加入注意力机制模块, 修正局部信息的重要程度, 提升了模型的细粒度识别能力。在城市室外场景机载点云数据集 GML\_DataSetA 和三维语义标记基准数据集 ISPRS Vaihingen 的实验结果证明了所提方法的有效性。

## 参 考 文 献

- [1] 胡海瑛, 惠振阳, 李娜. 基于多基元特征向量融合的机载 LiDAR 点云分类[J]. 中国激光, 2020, 47(8): 0810002.  
Hu H Y, Hui Z Y, Li N. Airborne LiDAR point cloud classification based on multiple-entity eigenvector fusion [J]. Chinese Journal of Lasers, 2020, 47(8): 0810002.
- [2] Mallet C, Bretar F, Roux M, et al. Relevance assessment of full-waveform lidar data for urban area classification[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2011, 66(6): S71-S84.
- [3] He H X, Khoshelham K, Fraser C. A multiclass TrAdaBoost transfer learning algorithm for the



- classification of mobile lidar data[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 166: 118-127.
- [4] Kim H B, Sohn G. Random forests based multiple classifier system for power-line scene classification[J]. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2012: 253-258.
- [5] Weinmann M, Jutzi B, Hinz S, et al. Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2015, 105: 286-304.
- [6] Yang Z S, Jiang W S, Xu B, et al. A convolutional neural network-based 3D semantic labeling method for ALS point clouds[J]. *Remote Sensing*, 2017, 9(9): 936.
- [7] Yang Z S, Tan B, Pei H K, et al. Segmentation and multi-scale convolutional neural network-based classification of airborne laser scanner data[J]. *Sensors*, 2018, 18(10): 3347.
- [8] Huang J, You S Y. Point cloud labeling using 3D convolutional neural network[C]//2016 23rd International Conference on Pattern Recognition (ICPR), December 4-8, 2016, Cancun. New York: IEEE Press, 2016: 2670-2675.
- [9] Zhang Z X, Zhang L Q, Tong X H, et al. A multilevel point-cluster-based discriminative feature for ALS point cloud classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(6): 3309-3321.
- [10] 雷相达, 王宏涛, 赵宗泽. 基于迁移学习的小样本机载激光雷达点云分类[J]. *中国激光*, 2020, 47(11): 1110002.  
Lei X D, Wang H T, Zhao Z Z. Small sample airborne LiDAR point cloud classification based on transfer learning[J]. *Chinese Journal of Lasers*, 2020, 47(11): 1110002.
- [11] Zeng W, Gevers T. 3DContextNet: K-d tree guided hierarchical learning of point clouds using local and global contextual cues[M]//Leal-Taixé L, Roth S. *Computer vision-ECCV 2018 workshops. Lecture notes in computer science*. Cham: Springer, 2019, 11131: 314-330.
- [12] 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, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 77-85.
- [13] Youssefhusien M, Kelbe D J, Ientilucci E J, et al. A multi-scale fully convolutional network for semantic labeling of 3D point clouds[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2018, 143: 191-204.
- [14] Winiwarter L, Mandlbürger G, Schmohl S, et al. Classification of ALS point clouds using end-to-end deep learning[J]. *PGF-Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 2019, 87(3): 75-90.
- [15] 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) [2021-06-01]. <https://arxiv.org/abs/1706.02413>.
- [16] Wen C C, Yang L N, Li X, et al. Directionally constrained fully convolutional neural network for airborne LiDAR point cloud classification[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 162: 50-62.
- [17] Li W Z, Wang F D, Xia G S. A geometry-attentional network for ALS point cloud classification [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 164: 26-40.
- [18] Wu W X, Qi Z A, Li F X. PointConv: deep convolutional networks on 3D point clouds[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 9613-9622.
- [19] 陈强. 高级计量经济学及 Stata 应用[M]. 2 版. 北京: 高等教育出版社, 2014: 519-523.  
Chen Q. *Advanced econometrics and Stata applications*[M]. 2nd ed. Beijing: Higher Education Press, 2014: 519-523.
- [20] Heidenreich N B, Schindler A, Sperlich S. Bandwidth selection for kernel density estimation: a review of fully automatic selectors[J]. *ASTA Advances in Statistical Analysis*, 2013, 97(4): 403-433.
- [21] Chen L C, Yang Y, Wang J, et al. Attention to scale: scale-aware semantic image segmentation[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 3640-3649.
- [22] Hori C, Hori T, Lee T Y, et al. Attention-based multimodal fusion for video description[C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York:

IEEE Press, 2017: 4203-4212.

- [23] Wang F, Jiang M Q, Qian C, et al. Residual attention network for image classification[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI,

USA. New York: IEEE Press, 2017: 6450-6458.

- [24] Niemeyer J, Rottensteiner F, Soergel U. Contextual classification of lidar data and building object detection in urban areas[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2014, 87: 152-165.