引文格式：王子明. 面向点云语义分割的多特征融合PointNet++网络[J]. 地理空间信息, 2023, 21(5): 25-28. doi:10.3969/j.issn.1672-4623.2023.05.007

面向点云语义分割的多特征融合PointNet++网络

王子明[1]
（1. 江苏省地质工程勘察院，江苏 南京 211102）

摘　要： PointNet和PointNet++方法以最大池化为聚合函数使得深度神经网络可直接分类无序点云，得到了较高的分类精度，但对点云空间相关性局部特征提取能力不足，制约了点云语义分割精度的提升。针对该问题，设计了一种面向点云语义分割的多特征融合PointNet++网络，在PointNet++网络中加入一个特征编码器，并以最小信息熵法计算的最优邻域来计算人工特征作为特征编码器的输入。在ISPRS提供的Vaihingen区域三维点云分类标准数据集上进行语义分割实验对比，结果表明多特征融合PointNet++网络语义分割精度比PointNet和PointNet++分别提高了4.3%和3.2%。

关键词： LiDAR点云；PointNet；点云分类；语义分割
中图分类号： P237　　**文献标志码：** B　　**文章编号：** 1672-4623（2023）05-0025-04

Multi-feature Fusion PointNet++ Network for Point Cloud Semantic Segmentation

WANG Ziming[1]
(1. Geo-engineering Investigation Institute of Jiangsu Province, Nanjing 211102, China)

Abstract: The PointNet and PointNet++ methods use the maximum pooling as the aggregation function, so that the deep neural network can directly classify the unordered point cloud, and obtain high classification accuracy. However, PointNet and PointNet++ methods have insufficient ability to extract local features of point cloud spatial correlation, which restricts the accuracy of point cloud semantic segmentation. Aiming at this problem, we designed a multi-feature fusion PointNet++ network for point cloud semantic segmentation. We added a feature encoder to the PointNet++ network, and used the optimal neighborhood calculated by the minimum information entropy method to calculate artificial features as the input of feature encoder. We compared the semantic segmentation experiments on the standard dataset of 3D point cloud classification in the Vaihingen area provided by ISPRS. The classification accuracy of multi-feature fusion PointNet++ network is 4.3% and 3.2% higher than that of PointNet and PointNet++ respectively.

Key words: LiDAR point cloud, PointNet, point cloud classification, semantic segmentation

机载激光雷达（LiDAR）具有全天时、全天候、高精度、高效率的优点[1]，且随着硬件的发展，其成本在不断降低，应用越来越普及。然而，LiDAR点云缺乏光谱和纹理信息，给基于点云的地物提取、分割分类带来了挑战[2]。近年来，越来越多的学者采用深度学习方法进行遥感数据解译，尤其对二维栅格影像的分类已取得优异结果，其分类精度明显优于传统基于专家设计特征的分类方法[3]；而三维点云的不规则性和无序性，使得基于深度学习的点云分类远远难于图像分类。鉴于此，学者们对不规则点云进行体素化，利用三维卷积神经网络对点云进行分类[4]或通过投影生成大量视角的二维影像再利用传统的二维卷积神经网络进行分类。这些方法使点云规则化，便于神经网络处理，但损失了细节信息，增加了点云数据量，且三维卷积神经网络对计算资源的要求较高；二维投影则损失了点云部分三维信息，制约了分类精度的提升。为避免对点云进行规则化，PointNet[5]网络于2017年被提出，以最大池化为聚合函数解决了点云无序性的问题，使得PointNet可直接分类输入的无序点云；但其无法获得点云的局部特征，因此对于点云语义分割的表现不佳。针对该问题，PointNet++[6]网络被提出，通过提取点云多尺度特征提高了点云语义分割的精度。虽然PointNet++提取的深度特征精度优于传统特征，但其学习能力仍难以满足广域复杂场景的点云分类问题。因此，本文通过融合传统特征与PointNet++特征，提出了一种多特征融合PointNet++网络，在PointNet++网络中加入一个特征编码器，并以最小信息熵法计算的最优邻域来计算人工特征作为特征编码器的输入。

收稿日期：2022-03-18；修回日期：2022-06-07。
项目来源：山东省自然科学基金资助项目（ZR2021QD074）。
作者简介：王子明（1985—），高级工程师，主要研究方向为测绘工程技术与遥感技术应用，E-mail：wuhandaxuewang2003@126.com。

1 特征提取与描述

点云特征是点云分割、分类[7]的基础，不仅包括单点本身的三维坐标信息、强度信息、回波信息、高度信息等，还包括与该点关联的邻域属性（高度、法向量、平面度等）。点云高度是用于地物识别的重要特征。点云的 Z 坐标仅代表点高程，由于地形起伏的影响，地物高程无法反映地物高度，因此需通过地面滤波算法[8]获取地面点，再将每个点的高程减去最近地面高程，即可得到该点的高度特征。令某点 i 的高程为 H_i，离它最近点的高程为 H_i^{nearest}，则点 i 的高度特征记为：

$$h_i = H_i - H_i^{\text{nearest}} \tag{1}$$

点云许多邻域特征的计算需以协方差矩阵为基础，令某点为 p_i，通过KD-tree[9]或八叉树[10]计算其 K 个邻域点，记为 $N(p_i)$，令 $N(p_i)$ 构造的协方差矩阵为 $\boldsymbol{C}_i$，则有：

$$\boldsymbol{C}_i = \sum_{p \in N(p_i)} (p - o_i)(p - o_i)^{\mathrm{T}} \tag{2}$$

$$o_i = \frac{1}{n} \sum_{p \in N(p_i)} p \tag{3}$$

计算 $\boldsymbol{C}_i$ 的特征值，并依照大小排序记为 λ_0、λ_1、$\lambda_2(\lambda_0 > \lambda_1 > \lambda_2)$，其对应的3个特征向量记为 e_0、e_1、e_2，其中 e_0 为最优拟合平面的法向量，e_1、e_2 则是拟合平面在点 p_i 处的两个伸展方向（图1）。点云法向量可反映目标点云局部切平面的性质，是非常重要的曲面属性特征。匀质区域的点法向量的方向一致或变化平缓；反之，则变化剧烈。

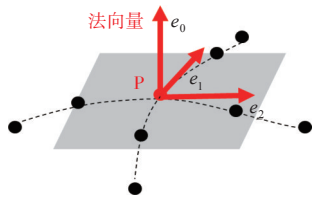


图1 点云特征向量

特征值可用于计算点云的各向同性 O_λ、各向异性 A_λ、平面度 P_λ、球面度 S_λ、线性度 L_λ[11]，其公式如下：

$$O_\lambda = (\lambda_1\lambda_2\lambda_3)^{1/3} \tag{4}$$

$$A_\lambda = \frac{\lambda_1 - \lambda_3}{\lambda_1} \tag{5}$$

$$P_\lambda = \frac{\lambda_2 - \lambda_3}{\lambda_1} \tag{6}$$

$$S_\lambda = \frac{\lambda_3}{\lambda_1} \tag{7}$$

$$L_\lambda = \frac{\lambda_1 - \lambda_2}{\lambda_1} \tag{8}$$

计算协方差矩阵需设定点的邻域点数量 K，K 代表了邻域大小，而点云不同邻域大小呈现的特征是不同的。由于机载LiDAR点云的密度不均匀，高密度位置点的 K 个邻域点表征的邻域距离将小于低密度位置点。通过设定距离而不是点的数量来计算邻近点可解决该问题，但难以找到一个合理的距离来提取所有地物、所有构件的局部特征。鉴于此，本文通过信息熵[12]来确定最佳邻域大小。信息熵可量化对信息的度量，利用点 p_i 的 K 个邻域点计算的协方差矩阵和协方差矩阵的特征值，将点 p_i 的信息熵记为 $E(K)$，则最优邻域 $\hat{K}$ 为：

$$\hat{K} = \arg\min_{K \in \aleph} E(K) = \arg\min_{K \in \aleph} \left(-\sum_{i=0}^{2} \lambda_i \ln \lambda_i\right) \tag{9}$$

式中，$\aleph$ 为 $[K_{\min}, K_{\max}]$ 中的整数。

最后得到特征组合为：

$$F = [h_i, O_\lambda, A_\lambda, P_\lambda, S_\lambda, L_\lambda] \tag{10}$$

2 网络框架

2.1 PointNet网络

PointNet[5]是一个可直接输入原始点云的端到端点云分类网络。PointNet语义分割网络框架见图2（m 为类别数量），PointNet通过输入变换模块和特征变换模块实现了点云分类的旋转不变性，使得点云特征在特征空间上对齐。共享多层感知机（MLP）利用一个共享权重网络计算每个点的特征，从点云三维坐标逐步提取1 024维的高维特征，再通过最大池化得到输入点云的全局特征，解决了点云无序性的问题，且对点缺失、插入等误差具有较好的鲁棒性。

为得到点云语义分割结果，需将特征变换模块输出的点云中的每个点后加上1 024维的全局特征，以融合点的坐标特征和全局特征；再通过共享MLP得到每个点属于每个类别的分数。

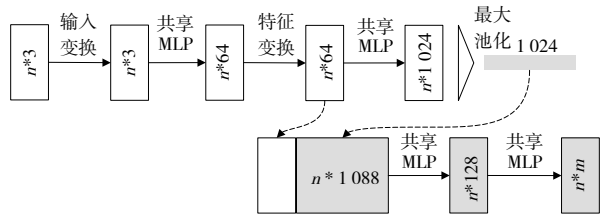


图2 PointNet语义分割网络框架

2.2 抽象集合模块

PointNet没有考虑点云的局部特征，因此语义分割精度不足。针对该问题，PointNet++[6]采用多层次结

构学习点云多尺度的深度特征。PointNet++设计了一种抽象集合模块（图3），由采样层、分组层和PointNet层组成，采样层通过最远点下采样法从 n 个输入点中选择 n' 个种子点；分组层利用球查询法将点集合分为 n' 个区域；PointNet层利用PointNet网络提取每个区域的特征。PointNet++网络通过将多个抽象集合模块级联组合，逐步提取更大尺度的特征，从浅层特征提取到深层特征。

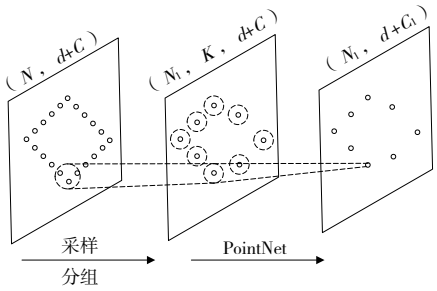


图3　PointNet++网络抽象集合模块

2.3　本文网络框架

PointNet++虽然通过抽象集合模块提取了多尺度点云特征，但未充分挖掘点与点之间的空间相关性。因此，本文提出了一种多特征融合PointNet++网络，在PointNet++语义分割网络中加入一个人工特征分支，为网络补充点云的空间关系特征。本文提出的网络框架（图4）包括坐标编码器、特征编码器和解码器3个分支，坐标编码器和解码器组合即原始的PointNet++网络，解码器利用跳跃连接将坐标编码器中对应尺度的点云特征合并，并通过插值逐步上采样点云至原始分辨率以实现语义分割；特征编码器分支加入了人工设计的点云空间相关性特征，经过两层共享MLP后，通过跳跃连接合并到解码器对应分辨率的点云中。

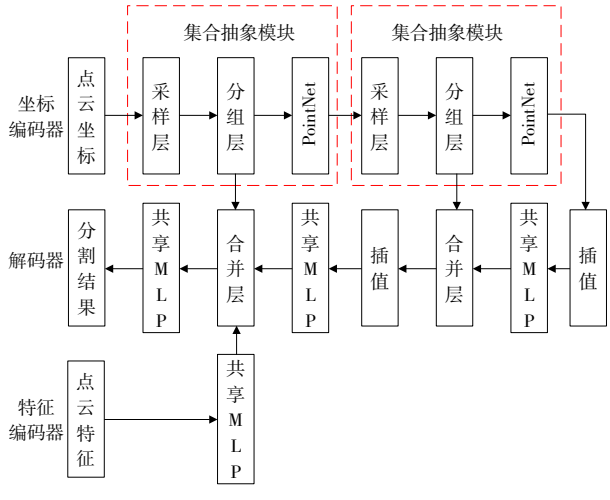


图4　本文网络框架

2.4　精度评价

网络模型训练完成后，需通过测试数据进行精度评定。本文采用的评价指标为F1分数（F_1）和总体精度（OA），F_1 结合了准确率和召回率，可综合衡量分类器对某类地物的识别能力；OA 可衡量分类器对整个场景所有地物的分类能力。令像素总数为 N，y_i、$\hat{y}_l$ 分别为第 i 个像素的真实值和预测值，则 OA 的计算公式为：

$$OA\left(y_i,\hat{y}_l\right)=\frac{1}{N}\sum_{i=0}^{N-1}I(\hat{y}_l,y_i) \tag{11}$$

$$I\left(\hat{y}_l,y_i\right)=\begin{cases}1 & \text{if } \hat{y}_l=y_i\\ 0 & \text{if } \hat{y}_l\neq y_i\end{cases} \tag{12}$$

准确率（pre）是某一类别识别正确的样本占所有被识别为该类别样本总数的比例；召回率（rec）是某一类识别正确的样本占该类别总样本个数的比例，那么 F_1 的计算公式为：

$$F_1=\frac{2\times\text{pre}\times\text{rec}}{\text{pre}+\text{rec}} \tag{13}$$

3　实验结果与分析

3.1　实验数据与方法

实验采用ISPRS提供的德国Vaihingen区域LiDAR点云语义分割数据集[13]。点云数据如图5所示，包含电力线、树、矮植被、灌木、不透水层、汽车、栅栏、屋顶、立面9个类别。为了证明本文方法的优越性，将本文方法与传统PointNet和PointNet++方法进行对比。实验的3种网络均以Tensorflow 2.6为框架，显卡为两块Geoforce 2080Ti。本文提出的网络输入点个数为1 024，其他超参数与PointNet++网络相同[6]。

3.2　实验结果

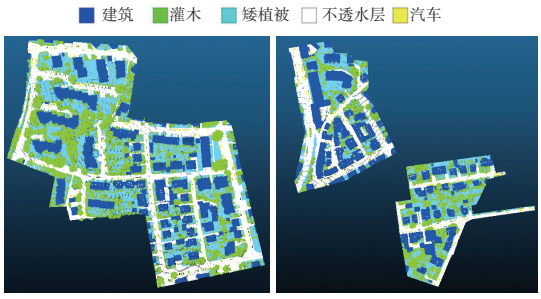
3种网络的语义分割精度见表1，可以看出，PointNet++的 OA 比PointNet提升了1.1%，说明PointNet++网络通过多个抽象集合模块级联组合，并结合跳跃连接成功融合了多尺度的点云特征，语义分割精度优于PointNet；本文方法的 OA 比PoinNet++提升了3.2%，除立面和灌木精度稍低外，其他类别的 F_1 都得到了特征，说明本文方法成功融合了人工特征与PointNet++特征，提高了地物识别精度。虽然PointNet++相较于PointNet实现了局部特征提取，但仍不能很好地对局部信息建模，因此人工特征可补充PointNet++深度特征，提高语义分割精度。

本文方法与PointNet++的分类结果错误如图6所示，可以看出，两种方法的错误区域更多集中在树和矮植被，这与表1的精度结论一致；PointNet++分类错误区域比本文方法多且明显，图6a中黄色椭圆内为

道路及其旁边的树与矮植被区域，PointNet++识别错误较多，树和矮植被的点云分布复杂，需要更多的空间关系特征辅助分类，因此结合人工特征后可提高识别精度。

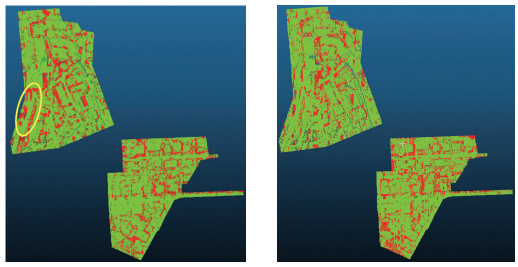
表1 3种网络的语义分割精度/%

方法模型	F_1									OA
	不透水层	屋顶	矮植被	树	汽车	栅栏	立面	灌木	电力线	
PointNet	88.1	89.1	67.1	73.2	51.5	41.1	53.5	51.1	41.5	76.8
PointNet++	89.0	89.9	68.2	74.9	57.9	42.3	51.8	50.9	38.5	77.9
本文方法	90.1	90.8	73.5	79.9	61.5	49.8	50.3	51.8	40.1	81.1



a 训练数据 b 测试数据

图5 ISPRS LiDAR点云语义分割数据集



a PointNet++ b 本文方法

图6 两种方法分类错误图

注：绿色为分类正确点云，红色为分类错误点云。

4 结 语

LiDAR的应用越来越普及，但其点云缺乏光谱和纹理信息，给基于点云的地物提取、分割分类带来了挑战。三维点云的不规则性和无序性，使得基于深度学习的点云分类远远难于图像分类。针对该问题，PointNet和PointNet++以最大池化为聚合函数解决了点云无序性问题，使其可直接分类输入的无序点云，但不能很好地提取点云空间关系特征，对局部特征的建模能力不足，因此对于点云语义分割的表现不佳。鉴于此，本文通过融合传统特征与PointNet++特征，设计了一种面向点云语义分割的多特征融合PointNet++网络，在PointNet++网络中加入一个特征编码器，并以最小信息熵法计算的最优邻域来计算人工特征作为特征编码器的输入。在ISPRS提供的Vaihingen区域三维点云分类标准数据集上进行语义分割实验对比发现，本文方法的语义分割精度比PointNet和PointNet++分别提高了4.3%和3.2%。。

参考文献

[1] 刘润东，宋袁龙，彭桂辉. 一种机载LiDAR点云条带重叠度自动检查算法[J]. 地理空间信息，2018，16(11):42-43

[2] Tinkham W T，HUANG H，Smith A，et al. A Comparison of Two Open Source LiDAR Surface Classification Algorithms[J]. Remote Sensing，2011，3(3):638-649

[3] Cwab C，XIANG L C，Xy A，et al. Airborne LiDAR Point Cloud Classification with Global-local Graph Attention Convolution Neural Network[J]. ISPRS Journal of Photogrammetry and Remote Sensing，2021，173:181-194

[4] 汪文琪，李宗春，付永健，等. 融合点、体素和对象特征的多基元点云分类[J]. 地球信息科学学报，2022，24(2):365-377

[5] 付洪波，李德元，张志强，等. 基于PointNet的点云数据处理及识别技术应用研究[J]. 测绘与空间地理信息，2021，44(12):194-195

[6] 姚钺，任明武. 基于PointNet++改进的点云特征提取与分类网络架构[J]. 计算机与数字工程，2021，49(10):2 052-2 056

[7] 朱琴，杨英宝，张宁宁. 利用LiDAR和航空影像的屋顶边缘提取及优化[J]. 地理空间信息，2018，16(4):36-39

[8] 阎跃观，陈中章，孙阳，等. 密集低矮植被区LiDAR点云地面滤波算法[J]. 测绘通报，2021(7):1-5

[9] 姚定忠，何军，刘祎. 一种基于kd树的实时大规模地形可视化算法[J]. 科学技术与工程，2012，12(2):338-341

[10] 杨振发，万刚，李锋，等. 多分辨率LOD的海量点云显示技术研究[J]. 地理空间信息，2016，14(10):22-25

[11] 叶刚. 融合影像信息的LiDAR点云多特征分类方法[J]. 地理空间信息，2018，16(6):70-75

[12] 谢妙，林泳昌，朱晓姝. 一种基于信息熵的自适应k值KNN二分类方法[J]. 合肥工业大学学报(自然科学版)，2021，44(11):1 483-1 486

[13] Niemeyer J，Rottensteiner F，Sörgel 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