

# 基于双注意力机制和多尺度特征的点云场景分割

于丽丽, 于海洋\*, 何子鑫, 陈良轩

河南理工大学自然资源部矿山时空信息与生态修复重点实验室, 河南 焦作 454000

**摘要** 针对现有的深度学习车载激光雷达(LiDAR)点云分割方法忽略高级全局单点和低级局部几何特征之间的关系导致点云分割精度不高的问题,提出了一种增强语义信息与多通道特征融合的点云场景分割算法。首先对点云信息进行补充,提取点云的归一化高程、强度值、光谱信息构造多通道点云特征,并利用网格化重采样方法建立多尺度邻域点云增强数据集,通过构建双注意力融合层实现通道维度上的特征加权标定和空间维度上的特征聚焦,加深卷积网络结构深层信息传递,挖掘点云的局部区域细粒度特征。采用国际摄影测量与遥感协会提供的数据集对所提算法进行验证,对比分析了所提算法、协会网站已公布的分类结果、主流的深度学习方法的精度值分类总体精度(OA)、综合评价指标(F1)、交并比。实验结果表明,所提算法能够达到更高的分割精度,在 Vaihingen 数据集上的平均交并比达到 52.5%。

**关键词** 遥感; LiDAR 点云; 点云分割; 双注意力机制; 多尺度特征

中图分类号 P237

文献标志码 A

doi: 10.3788/LOP202158.2428007

## Point Cloud Scene Segmentation Based on Dual Attention Mechanism and Multi-Scale Features

Yu Lili, Yu Haiyang\*, He Zixin, Chen Liangxuan

Key Laboratory of Spatio-temporal Information and Ecological Restoration of Mines, Ministry of Natural Resources, Henan University of Technology, Jiaozuo, Henan 454000, China

**Abstract** Aiming at the problem that the existing deep learning light detection and ranging (LiDAR) point cloud segmentation method ignores the relationship between high-level global single points and low-level local geometric features, which results in low point cloud segmentation accuracy, an enhanced semantic information and multi-channel feature fusion point cloud scene segmentation model is established. First, the point cloud information is supplemented, the normalized elevation, intensity value and spectral information of the point cloud are extracted to construct the multi-channel point cloud features, and the multi-scale neighborhood point cloud enhancement data set is established by using the grid resampling method. By constructing a double attention fusion layer, the feature weighted calibration in the channel dimension and feature focusing in the spatial dimension are realized, the deep information transmission of the convolution network structure is deepened, and the local regional fine-grained features of the point cloud are mined. The proposed algorithm is verified by using the data set provided by the International Association of Photogrammetry and Remote Sensing. Overall accuracy (OA) of accuracy value classification, comprehensive evaluation index (F1), and intersection ratio of the proposed algorithm, the classification results published on the association's website, and mainstream deep learning methods are compared and analyzed. Experimental results show that the proposed algorithm can achieve higher segmentation accuracy, and the average intersection union ratio on Vaihingen data set reaches 52.5%.

**Key words** remote sensing; LiDAR point cloud; point cloud segmentation; dual attention mechanism; multi-scale feature

**OCIS codes** 280.0280; 150.0150; 140.0140

收稿日期: 2020-11-11; 修回日期: 2020-12-28; 录用日期: 2021-01-20

基金项目: 国家自然科学基金(U1304402)

通信作者: \*1922797937@qq.com

## 1 引言

机载激光雷达(LiDAR)点云数据已经广泛地应用于三维城市建模、城市规划和森林资源评估等领域<sup>[1]</sup>。然而,对离散、不规则且包含噪声的LiDAR点云数据进行准确且自动化的分类是数据应用的关键,也一直是研究的热点难点。目前基于深度学习的三维点云分割方法分为两大类,即间接分割法和直接分割法<sup>[2]</sup>。间接分割法将点云数据转化为体素化网格<sup>[3]</sup>或多视图<sup>[4]</sup>,间接地从三维数据中提取点云特征实现点云分割。直接分割法保留原始点内固有的信息,直接从点云数据中提取特征信息实现分割的目的。但点云是无序且没有结构的,直接应用标准的卷积神经网络(CNN)的方法并不可行,Charles等<sup>[5]</sup>采用PointNet网络直接对无序的点云进行分割,拓展了深度学习方法在点云处理领域的应用。但PointNet网络只考虑点云单点的全局特征,忽略局部领域特征。后续研究者为了捕获点云的全局特征并将其与每个点更广泛的上下文局部特征有效结合进一步提高语义分割的精度,提出了基于相邻特征检测、图卷积、优化CNN、局部全局特征串联等方法进行点云分割。Charles等<sup>[6]</sup>提出PointNet++网络,该网络的密度自适应点层在输入采样密度时通过组合不同尺度的搜索半径解决了对点云局部特征提取的问题。Wang等<sup>[7]</sup>提出动态图卷积神经网络(DGCNN),该网络通过构建局部邻域图并利用边缘卷积(EdgeConv)考虑点的坐标与邻域点的距离获取点云的局部特征,进一步聚合生成局部几何特征。Thomas等<sup>[8]</sup>提出更灵活的可变性卷积算子的核点卷积(KPConv),KPConv通过最近距离内核点的权重对局部区域进行卷积。Zhao等<sup>[9]</sup>采用动态聚合网络(DAR-Net)生成局部卷积特征提取器和全局循环特征聚合器,综合考虑全局场景复杂性和局部几何特性。Arandjelovic等<sup>[10]</sup>提出NetVLAD模型,该模型将传统的VLAD结构嵌入到CNN中反复堆叠以捕获局部信息和场景级全局特征。上述研究对点云的分割有进一步的贡献,有助于捕获不同尺度点云之间的特征融合。但直接融合不同层特征,未对低层特征进行筛选,导致在高层特征中引入大量的干扰信息,影响分割精度。同时上述方法不能基于点云空间位置和特征通道差异,将注意力权重分配给不同的相邻点和特征通道,学习捕获区分特征进行细分,这也是场景分割的关键。

针对上述问题,本文提出一种基于CNN嵌入双注意力机制的点云语义分割模型。通过对局部特征进行丰富的上下文相关性建模,基于相邻的点云位置和通道关系生成注意力权重;增加空间注意力模块学习特征的空间相关性,采用通道注意力建立通道的相互依赖,解决仅空间相邻位置决定权值的局限性,进一步提高分割能力;融合多光谱航空影像构建多通道点云特征,提出多尺度点云数据特征训练自适应地不断地调整权重值,深层次地提取不同尺度的局部特征和全局特征,实现LiDAR点云数据分割的目的并提高分割精度。

## 2 基于注意力机制的卷积神经网络模型构建

在卷积网络模型的基础上加入空间注意力机制和通道注意力机制,在卷积层深层次地对低层特征进行加权处理,降低无用信息干扰,提高网络的语义分割精度。空间注意力机制通过点云的坐标 $x$ 、 $y$ 和归一化高程值的空间信息进行加权获得突出的显著性区域。在空间注意力机制的基础上融合通道注意力机制为点云各通道进行依赖性建模,赋予不同通道光谱信息权重,增大对分割任务有利的特征通道权重。任何两个具有相似特征的位置,无论在空间维度上的距离如何都可以相互促进注意力机制有选择地聚合不显著对象的相似特征,以突出它们的特征并避免显著特征的影响。所提算法将空间和通道两个注意力模块嵌入到卷积神经网络中,卷积神经网络结构如图1所示。

### 2.1 双重注意力机制

注意力机制可以捕获全局上下文依赖关系,在卷积神经网络中嵌入空间注意力机制和通道注意力机制可以显著提高分割性能<sup>[11]</sup>。空间注意力机制通过位置进行加权求和,选择性地聚集每个位置的特征,对相似特征进行关联,而不考虑它们之间的距离。通道注意力机制整合所有通道之间的相关性特征,有选择地强调相互依赖的通道关系。为了进一步获得全局依赖关系特征,所提算法对两个模块输出结果进行融合,自适应地集成局部依赖和全局依赖。

#### 2.1.1 空间注意力机制

在位置注意力模块中,任意位置特征的更新是通过点云所有位置上特征加权聚合实现的,权重是由两个位置的相似性决定的,不考虑两个位置的距离差异,只要它们的特征相似就可以得到较高的权

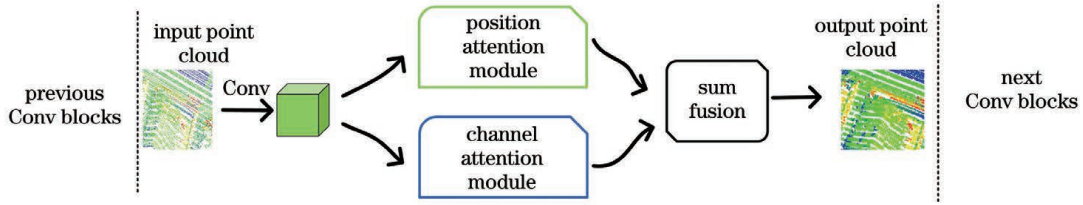


图 1 嵌入注意力机制后的卷积网络结构

Fig. 1 Convolution network structure after embedding attention mechanism

重。所提算法利用任意两点位置之间的关联,将几何特征和上下文特征结合起来增强各自的特征表达。

空间注意力机制可实现空间维度上的特征聚焦,加深网络结构深层信息传递,如图 2 所示。令  $A^{G \times N \times M}$  为输入点云特征矩阵,其中  $G$  表示组数,  $N$  和  $M$  分别表示点云数量和通道数。将  $A$  分别输入通道数为 64 的  $1 \times 1$  卷积层进行降维处理得到  $B^{G \times N \times M}$ 、 $C^{G \times N \times M}$ 。对  $B$  和  $C$  进行维度转换得到

$B^{G \times J}$  和  $C^{G \times J}$ , 其中  $J = N \times M$ 。对  $B$  进行转置并与  $C$  进行矩阵相乘,经过 Softmax 层归一化得到空间注意力矩阵  $S^{J \times J}$  系数。相似位置相关性较高表示点的依赖性关系强。

$$S_{ji} = \frac{\exp(B_i \times C_i)}{\sum_i \sum_i \exp(B_i \times C_i)}, \quad (1)$$

式中:  $S_{ji}$  表示  $i$  位置特征对  $j$  位置特征的影响。

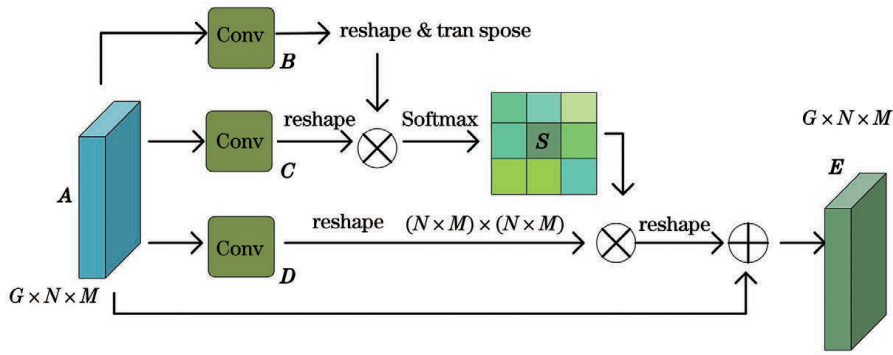


图 2 空间注意力机制模块

Fig. 2 Spatial attention mechanism module

同时,将特征  $A$  输入到卷积层得到特征矩阵  $D^{G \times N \times M}$ ,  $D^{G \times N \times M}$  经维度转换得到  $D^{G \times J}$ 。  $D$  与  $S$  的转置相乘结果为  $R^{G \times N \times M}$ , 将  $R^{G \times N \times M}$  乘以因子  $a$  并与  $A$  求和得到  $E^{G \times N \times M}$ 。

$$E_j = a \sum_{i=1}^J (S_{ji} D_i) + A_j. \quad (2)$$

$a$  的初始值为 0, 通过逐渐的学习分配更多权重, 最后输出特征增强的特征值。每个位置的结果特征

$E$  是所有位置特征和原始特征的加权和, 根据空间注意力机制选择性地聚合上下文信息提高语义一致性。

### 2.1.2 通道注意力机制

通道注意力机制使用与注意力机制类似的方法来学习任意两个通道之间的映射关系, 如图 3 所示。通道注意力机制利用通道之间的相互依赖性来更新某一个通道, 通过通道之间的关联增强通道下特定语义响应能力。与空间注意力不同的是, 通道注意

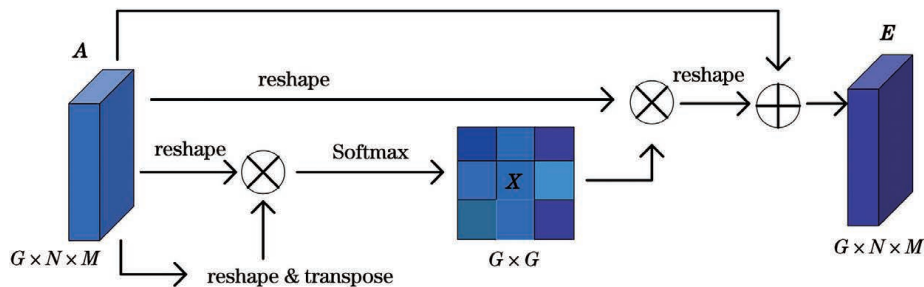


图 3 通道注意力机制

Fig. 3 Channel attention mechanism

力在一开始不进行卷积运算,避免打破通道之间的关系。直接从原始特征矩阵  $\mathbf{A}^{G \times N \times M}$  通过矩阵计算和 Softmax 函数得到通道注意力矩阵  $\mathbf{X}^{G \times G}$  系数。

$$\mathbf{X}_{ji} = \frac{\exp(\mathbf{A}_i \times \mathbf{A}_j)}{\sum_{i=1}^G \exp(\mathbf{A}_i \times \mathbf{A}_j)}, \quad (3)$$

式中:  $\mathbf{X}_{ji}$  表示  $i$  通道对  $j$  通道的影响。

同时  $\mathbf{X}$  和  $\mathbf{A}$  的转置  $\mathbf{A}^T$  进行矩阵乘法得到  $\mathbf{R}^{G \times N \times M}$ , 将结果乘一个缩放因子  $b$ , 并与  $\mathbf{A}$  求和得到输出  $\mathbf{E}^{G \times N \times M}$ 。

$$\mathbf{E}_j = b \sum_{i=1}^G (\mathbf{X}_{ji} \mathbf{A}_i) + \mathbf{A}_j. \quad (4)$$

$b$  的初始值为 0, 通过逐渐的学习分配更多权重。每个通道的最终特征是所有通道的特征和原始特征的加权和, 表明特征映射之间的长期语义依赖关系, 有助于提高特征的可分辨性。

### 2.2 基于注意力机制的卷积神经网络模型结构

卷积神经网络是一种带有卷积结构的深度神经

网络, 隐含层的卷积层和池化采样层是实现卷积网络特征提取的核心模块, 现有的三维分类算法主要采用多层感知 (MLP) 作为基本特征提取模块。所提算法首先采用 LiDAR 点云以及航空影像进行点特征的提取, 通过多尺度化实现数据增强, 再构建嵌入双注意力机制的卷积模型实现分割<sup>[12-13]</sup>。所用卷积模型主要是在 PointNet 基础上进行改进的, 由特征提取层、双注意力机制模块、最大池化层、全连接层以及分类器层组成, 网络结构如图 4 所示。该模型是端到端的直接对点云进行处理训练模型, 对于点云分割任务输入  $N$  个带有特征的点云数据, T-Net 实现数据的规范操作, 可避免数据预处理过程的信息损失; 在网络配置中通过嵌入注意力机制的 MLP 来进行点特征的提取, 提高每个点的空间特征和通道之间的相关性; 将每个点深层次的 64 维局部特征与 1024 维全局特征联合得到 1088 维联合特征; 最后通过分类器实现每个点的分类。

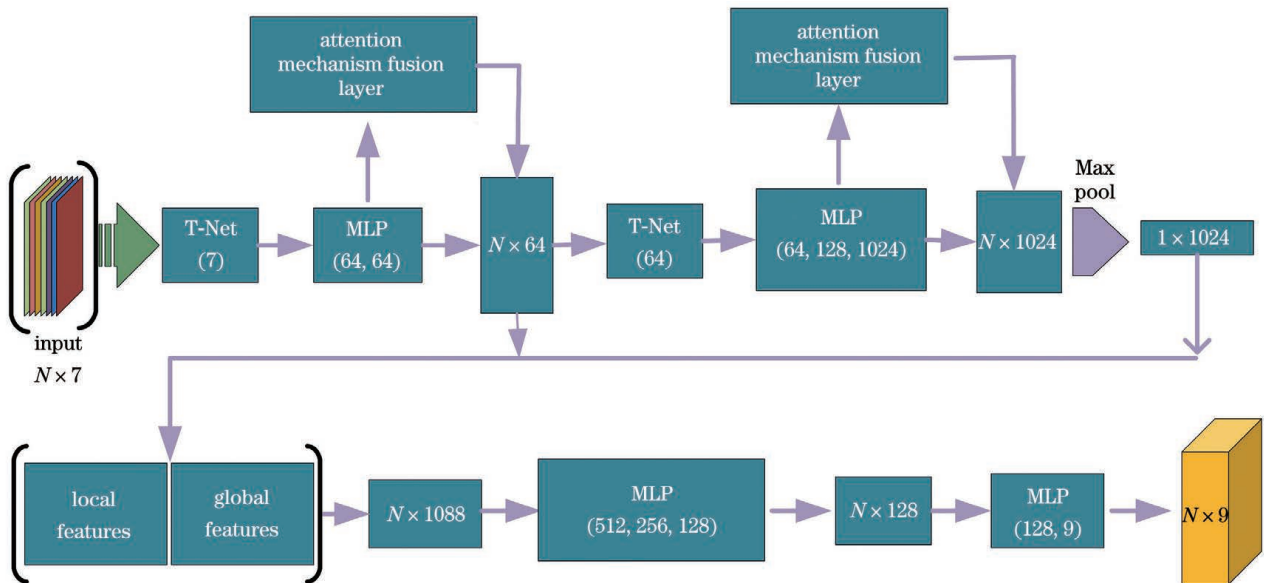


图 4 嵌入注意力机制的分割网络结构

Fig. 4 Segmentation network structure with embedded attention mechanism

## 3 多源点云特征融合

### 3.1 多源点云特征分析

实验选用  $X$ 、 $Y$ 、归一化高程值、强度值以及重采样获取相应的光谱信息作为点云的特征<sup>[14]</sup>。利用多光谱影像提取光谱作为点云的新特征, 增加单个点的特征向量维度, 获得更多的点云信息可以提高地物的精细分类能力。在模型训练过程中使用最大最小归一化处理点云中的特征值, 增加数据的内

聚性, 使数据更加密集。

#### 3.1.1 光谱特征计算

对多光谱航空影像与 LiDAR 点云进行数据融合, 通过点云数据和多光谱影像数据的配准实现点云数据的光谱信息补充, 有效地避免点云分割处理时仅有点云的位置特征出现的检错现象。将三维点云的坐标代入共线条件方程, 可以计算出每个三维点在影像上对应的像素位置  $(x, y)$ 。

$$\begin{cases} x = -f \frac{a_1(X - X_s) + b_1(Y - Y_s) + c_1(Z - Z_s)}{a_3(X - X_s) + b_3(Y - Y_s) + c_3(Z - Z_s)} \\ y = -f \frac{a_2(X - X_s) + b_2(Y - Y_s) + c_2(Z - Z_s)}{a_3(X - X_s) + b_3(Y - Y_s) + c_3(Z - Z_s)} \end{cases}, \quad (5)$$

式中:  $(X, Y, Z)$  为地面点的位置;  $f$  为相机的焦距;  $(X_s, Y_s, Z_s)$  是外方位元素的 3 个线元素;  $(a_1, b_1, c_1, a_2, b_2, c_2, a_3, b_3, c_3)$  是由外方位元素中 3 个角元素计算的旋转矩阵参数。获得像点坐标后, 所提算法采用最邻近插值法重采样计算得到对应的多光谱信息: 近红外 (NIR)、红 (R) 波段、绿 (G) 波段信息。

### 3.1.2 归一化高程值

归一化高程值是点云分割中最重要的特征, 可以有效地去除地形起伏对点云数据高程值的影响。所提算法采用高效的布料模拟滤波 (CSF) 算法<sup>[14]</sup> 对原始点云进行滤波, 得到地面点和非地面点; 采用重采样对地面点进行内插处理生成高程模型 (DEM)<sup>[15]</sup>, 每一个高程点的值减去到距离该点最近的 DEM 程值就得到了归一化高程值。

### 3.1.3 激光回波强度信息

LiDAR 系统提供激光脉冲的回波强度信息, 所

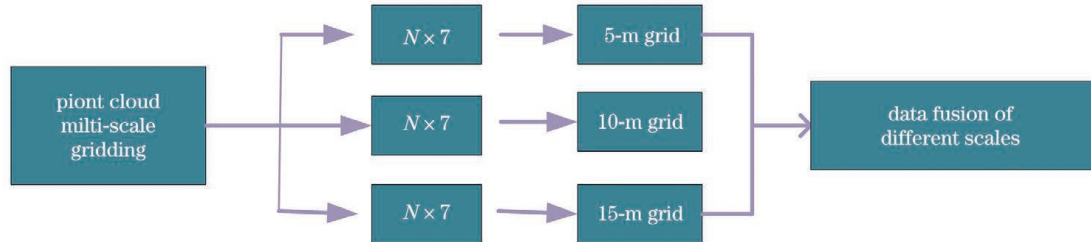


图 5 多尺度数据融合

Fig. 5 Multi-scale data fusion

## 4 实验与分析

### 4.1 实验数据集

实验使用的数据集是由国际摄影测量与遥感学会 (ISPRS) 提供的德国 Vaihingen3D 语义数据集, 该数据集是遥感领域中第一个发布的基准数据集, 包含训练数据集和测试数据集。所提算法采用 LiDAR 点云以及多光谱影像数据集进行实验, 如图 6 所示。数据地物形态复杂且不规则, 包含丰富的地理环境、城区环境及建筑物, 能够充分验证所提算法在室外大场景应用时的性能。LiDAR 点云数据密度为  $4 \text{ point}/\text{m}^3$ , 多光谱航空影像地面分辨率为  $8 \text{ cm}$ , 每张像素尺寸是  $7680 \text{ pixel} \times 13824 \text{ pixel}$ , 并提供内外方位元素。多尺度提取点的局部特征,

提算法将点云原始数据中的强度作为特征值。

### 3.2 多尺度点云特征数据融合方法

深度卷积网络需要通过训练不断地调整权重值, 而训练需要充足的样本。由于 LiDAR 点云数据的稀疏性和不均匀性, 同一目标不同位置点的密度也会不同, 为减少地物尺寸对分类结果的影响, 所提算法采用多尺度点云特征数据, 将不同尺度大小的邻域特征通过组合作为点的局部特征, 增强训练数据的样本量, 不断地优化点云之间的权重, 进一步提高对噪声和方向的鲁棒性。所提算法采用不同大小的网格对 LiDAR 点云进行分块处理<sup>[16]</sup>, 首先设定阈值对网格化之后的每个方格内点云数量进行检查, 当方格内点云数量小于保存阈值时, 则对当前方格内点云数据进行随机复制以满足点数量的要求; 如果方格内点云数量大于保存阈值则保存, 可作为实验数据集。实验中阈值参数设置为 1024, 分别选取 5, 10, 15 m 多尺度数据进行实验, 如图 5 所示。

空间注意力机制和通道注意力机制获得更多及更高维度的特征信息赋予更多的权重, 使模型做出更加准确的判断。该数据集共包含以下 9 类数据: 电力线、低矮植被、不透水表面、车辆、栅栏、屋顶、外墙、灌木、树木。

### 4.2 实验环境与网络参数设置

实验硬件配置为: 运行内存为 32 GB, 处理器为 Intel Core i9-9900K, 显卡为 RTX2080, 系统为 Windows 10。Python3.6 模型基于 Tensorflow 和 Keras 深度学习框架实现, 使用 GPU 加速。

在训练过程中网络不断进行参数更新, 但相同学习率不能适应所有参数的更新。为使不同参数计算不同的自适应学习率提高训练的精度和效率, 所提算法采取 Adam 优化方法基于数据迭代更新神

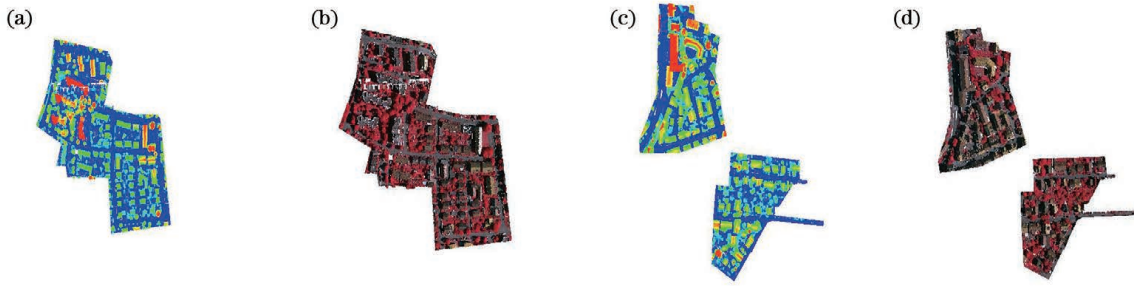


图 6 实验数据集。(a)训练数据集;(b)训练数据集对应的影像;(c)测试数据集;(d)测试数据集对应的影像  
Fig. 6 Experimental data set. (a) Training data set; (b) image corresponding to training data set; (c) test data set; (d) image corresponding to test data set

经网络权重,固定学习率为 0.001。

ReLU 和 Softmax 分别为激活函数和损失函数,ReLU 可以在一定程度上改善反向传播过程中的梯度爆炸和梯度消失的问题。

$$\text{ReLU}(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases} \quad (6)$$

Softmax 用来判断实际的输出与期望值输出的接近程度。

$$s_m = \frac{e^{f_m}}{\sum_n e^{f_n}}, \quad (7)$$

式中: $f_m, f_n$  为 Softmax 函数输入向量的第  $m, n$  个值; $s_m$  为函数输出向量  $S$  的第  $m$  个值,表示的这个样本属于第  $m$  个类别的概率。

### 4.3 分割精度评估指标

所提算法性能采用语义分割中常用的指标进行评价,一般选取总体精度(OA)、均交并比(IoU)和 F1 分数,F1 分数兼顾模型的查全率(R)和查准率(P)。

$$S_{F1} = \frac{P \times R}{P + R}, \quad (8)$$

$$R_{IoU} = \frac{N_{TP}}{N_{TP} + N_{FP} + N_{FN}}, \quad (9)$$

● power line ● low vegetation ● impervious surface ● car ● fence/hedge ● roof ● facade ● shrub ● tree

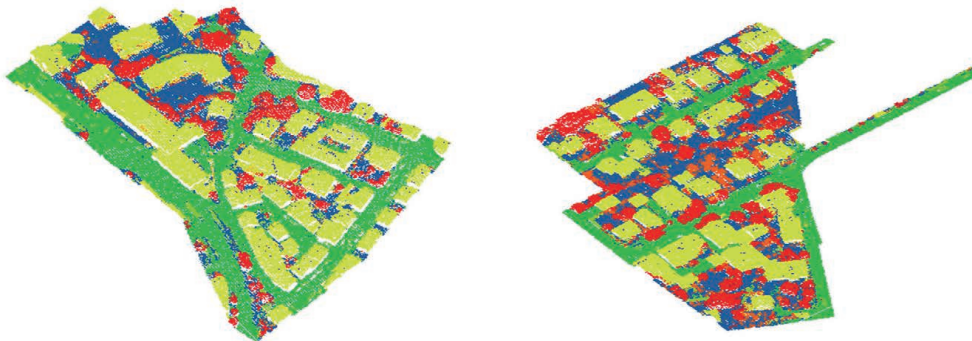


图 7 测试集分割效果可视化  
Fig. 7 Visualization of test set splitting effect

式中: $N_{TP}$  为类别是 1 的样本被系统正确判定为类别 1 的点数; $N_{FN}$  为类别是 1 的样本被系统判定为类别为 0 的点数; $N_{FP}$  为类别是 0 的样本被系统误判定为 1 的点数。

### 4.4 实验结果分析

通过实验对所提算法进行验证,分别讨论融入注意力机制和未融入注意力机制对特征分割精度的影响和所需训练时间的差别,分割的最终结果如图 7 所示。图 8 为未嵌入注意力机制的错分结果对比。从图 8 中可以看出,未融入注意力机制进行分割时,错分的结果较为明显,尤其在建筑物区域,融入注意力机制的能够更加注重局部点云特征,双通道融合特征更加丰富,错分现象得到有效的抑制。采用注意力机制的与未嵌入注意力机制的模型所需训练时间相似,表明所提算法在场景分割中的有效性。

#### 4.4.1 所提算法与 ISPRS 方法对比分析

为了与不同的研究方法进行对比,ISPRS 网络提供了不同方法实验结果。对比分析了所提算法和 ISPRS 网络提供的结果在数据集上每个类别对应的 F1 和平均 OA 值,结果如表 1 所示。

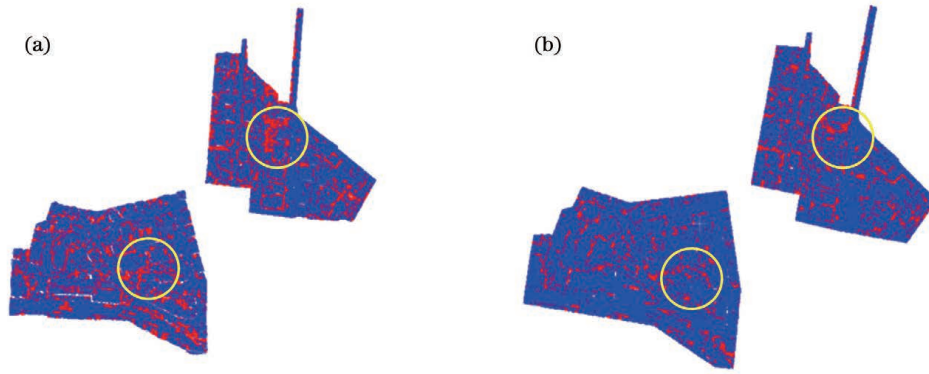


图 8 错误分割结果对比图。(a)未嵌入注意力机制;(b)嵌入注意力机制

Fig. 8 Comparison chart of misclassification results. (a) Attention mechanism not embedded; (b) embedded attention mechanism

表 1 所提算法与 ISPR 网络提供的结果对比

Table 1 Comparison of proposed algorithm and results provided by ISPR network

Method	F1 score									OA / %
	Power line	Low_ vegetation	Impervious surface	Car	Fence/hedge	Roof	Facade	Shrub	Tree	
BIJ_W	13.8	78.5	90.5	56.4	36.3	92.2	53.2	43.3	78.4	81.5
WhuY2	31.9	80.0	88.9	40.8	24.5	93.1	49.1	41.1	77.3	81.0
LUH	59.6	77.5	91.1	73.1	34.0	94.2	56.3	46.6	83.1	81.6
RIT_1	37.5	77.9	91.5	73.4	18.0	94.0	49.3	45.9	82.5	81.6
WhuY3	37.1	81.4	90.1	63.4	23.9	93.4	47.5	39.9	78.0	82.3
Proposed algorithm	55.8	81.5	94.9	73.8	45.5	94.4	30.0	43.8	83.8	86.3

从表 1 中可以看出:所提算法场景分割的总体精度 OA 达到 86.3%,相较于 RIT\_1 模型直接对点云进行分类,所提算法增强单个点和通道之间的多层次特征依赖关系,总体精度提高 4.7 个百分点;相较于 WhuY2、WhuY3 利用特征图进行分割,所提算法直接对点云进行分割且考虑点云特征之间的深层特征,分割精度得到提升;与传统的机器学习方法 LUH 进行对比,所提算法采用嵌入注意力机制深度学习模型,可以更好地提取点特征,精度提高 4.7 个百分点。

所提算法在低矮植被、不透水表面、汽车、栅栏、屋顶、树类别上的 F1 分数分别为 81.5%、94.9%、73.8%、45.5%、94.4%、83.8%,分割结果得到很大的提升;在灌木、外墙类别上的 F1 分数分别为 43.8%、30.0%,精度表现较差。灌木的错分主要来自于树,两者之间在高度和光谱反射率方面具有高度的相似性。较多的外墙被误判为屋顶,而屋顶很少判断为外墙,具有较低的召回率和较高的精度。

实验结果表明,所提双注意力机制会增加重要特征权重,权重优化特征与相关性相结合可以有效地分割点云。

#### 4.4.2 与其他深度学习分割方法对比分析

与其他深度学习分割方法进行对比分析,结果如表 2 所示。相比较直接的点云深度学习分割方法(PointNet++<sup>[6]</sup>、PointNetSIFT<sup>[17]</sup>、PointNetCNN<sup>[18]</sup>),由于注意力机制可以自适应地挖掘更具鉴别能力的细粒度局部特征,所提算法能够获得更好的特征表达,提高整体分割精度。其中场景分割的总体精度相较于 PointNetCNN 提高了 3 个百分点;F1 分数在不透水表面、栅栏、屋顶、树的分类结果上有很大的提升。表 3 为所提算法与 PointNet 网络在 Vaihingen 数据集上的 IoU 值、平均 IoU 值对比结果,所提算法的平均 IoU 值较 PointNet 网络提高 20.5% 达到较高的精度。综合实验结果表明,所提算法用到的包含光谱信息分割精度的多尺度数据,融合注意力机制能够有效地提高分类精度。

表 2 所提算法与深度学习算法对比

Table 2 Comparison of proposed algorithm and deep learning algorithms

Method	F1 score									OA /%
	Power line	Low_ vegetation	Impervious surface	Car	Fence/hedge	Roof	Facade	Shrub	Tree	
PointNet++	57.9	79.6	90.6	66.1	31.5	91.6	54.3	41.6	77.0	81.2
PointNetSIFT	55.7	80.7	90.9	77.8	30.5	92.5	5.9	44.4	79.6	82.2
PointNetCNN	61.5	82.7	91.8	75.8	35.9	92.7	57.8	49.1	78.1	83.3
Proposed algorithm	55.8	81.5	94.9	73.8	45.5	94.4	30.0	43.8	83.8	86.3

表 3 Vaihingen 上各类别 IoU

Table 3 IoU of each category on Vaihingen

unit: %

Method	Mean_IoU	Power line	Low_ vegetation	Impervious surface	Car	Fence/hedge	Roof	Facade	Shrub	Tree
PointNet	32.0	0.8	32.1	47.6	23.2	2.3	84.7	5.7	15.4	76.2
Proposed algorithm	52.5	35.6	67.7	89.5	53.4	19.7	88.6	22.2	29.5	66.4

## 5 结 论

针对点云分割问题,建立了一种端到端融合双注意力机制的深度网络分割算法用于点云结构特征学习。实验结果表明,通过双注意力机制对点云进行深层次特征的融合,对点云引入多尺度以及融合多源数据的点云特征进行提取,能够有效地避免部分有价值的信息丢失,增强目标特征并抑制无用特征信息,提高模型在复杂场景中的分割能力。在 Vaihingen 城市数据集上的测试验证了所提算法的有效性。相比于其他几种用于分类的神经网络算法,所提算法分割结果的总体精度有明显的提升,表现较为突出的包括不透水表面、栅栏、屋顶、树等点云类别。然而在实验中发现,所提算法在面对点数较少的地物全局特征时,数据处理具有一定的局限性,分割精度不够高。下一步将考虑如何在较少的点云数量条件下提升所提算法的分割精度。

## 参 考 文 献

- [1] Lai X D. Basic principle and application of airborne laser radar[M]. Beijing: Publishing House of Electronics Industry, 2010.  
赖旭东. 机载激光雷达基础原理与应用[M]. 北京: 电子工业出版社, 2010.
- [2] Jing Z W, Guan H Y, Zang Y F, et al. A survey of point cloud semantic segmentation based on deep learning[J]. Computer Science and Exploration, 2021, 15(1): 1-26.
- [3] Le T, Duan Y. PointGrid: a deep network for 3D shape understanding[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT. New York: IEEE Press, 2018: 9204-9214.
- [4] Le T, Bui G, Duan Y. A multi-view recurrent neural network for 3D mesh segmentation[J]. Computers & Graphics, 2017, 66: 103-112.
- [5] 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.
- [6] Charles R Q, Li Y, Hao S, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space[C]//Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA. [S. l.: s. n.], 2017: 5099-5108.
- [7] Wang Y, Sun Y B, Liu Z W, et al. Dynamic graph CNN for learning on point clouds[J]. ACM Transactions on Graphics, 2019, 38(5): 1-12.
- [8] Thomas H, Qi C R, Deschaud J E, et al. KPConv: flexible and deformable convolution for point clouds



- [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 6410-6419.
- [9] Zhao Z Y, Liu M, Ramani K. DAR-Net: dynamic aggregation network for semantic scene segmentation [EB/OL]. (2019-07-28) [2020-11-05]. <https://arxiv.org/abs/1907.12022>.
- [10] Arandjelovic R, Gronat P, Torii A, et al. NetVLAD: CNN architecture for weakly supervised place recognition [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 5297-5307.
- [11] Fu J, Liu J, Tian H J, et al. Dual attention network for scene segmentation [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: 3141-3149.
- [12] Liu H, Wang X L. Remote sensing image segmentation model based on attention mechanism [J]. Laser & Optoelectronics Progress, 2020, 57(4): 041015.  
刘航, 汪西莉. 基于注意力机制的遥感图像分割模型 [J]. 激光与光电子学进展, 2020, 57(4): 041015.
- [13] Dang J S, Yang J. 3D model recognition and segmentation based on multi-feature fusion [J]. Journal of Xidian University, 2020, 47(4): 149-157.  
党吉圣, 杨军. 多特征融合的三维模型识别与分割 [J]. 西安电子科技大学学报, 2020, 47(4): 149-157.
- [14] Wang H T, Lei X D, Zhao Z Z. 3D deep learning classification method for airborne LiDAR point clouds fusing spectral information [J]. Laser & Optoelectronics Progress, 2020, 57(12): 122802.  
王宏涛, 雷相达, 赵宗泽. 融合光谱信息的机载 LiDAR 点云三维深度学习分类方法 [J]. 激光与光电子学进展, 2020, 57(12): 122802.
- [15] Zhang W M, Qi J B, Wan P, et al. An easy-to-use airborne LiDAR data filtering method based on cloth simulation [J]. Remote Sensing, 2016, 8(6): 501.
- [16] Zhao Z Y, Cheng Y L, Shi X S, et al. Terrain classification of LiDAR point cloud based on multi-scale features and PointNet [J]. Laser & Optoelectronics Progress, 2019, 56(5): 052804.  
赵中阳, 程英蕾, 释小松, 等. 基于多尺度特征和 PointNet 的 LiDAR 点云地物分类方法 [J]. 激光与光电子学进展, 2019, 56(5): 052804.
- [17] 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) [2020-11-05]. <https://arxiv.org/abs/1807.00652>.
- [18] Li Y Y, Bu R, Sun M C, et al. PointCNN: convolution on X-transformed points [EB/OL]. (2018-01-23) [2020-11-05]. <https://arxiv.org/abs/1801.07791v5>.