

基于 CAFF-PointNet 的机载 LiDAR 点云语义分割

赖鸣, 赵健康*, 刘传奇, 崔超, 龙海辉

上海交通大学电子信息与电气工程学院仪器科学与工程系, 上海 200240

摘要 针对机载 LiDAR 点云中几何结构复杂和不同地物尺度变化大导致小目标点云分类准确率低的问题, 本文提出了一种基于通道注意力机制进行多尺度特征融合的卷积神经网络。首先, 通过球形邻域计算点云的平面度、线性度、法向量以及本征熵等浅层几何特征, 并将其与网络提取的深层次语义特征进行融合, 增强模型对点云几何结构的感知能力; 其次, 设计基于通道注意力机制的多尺度特征融合模块, 学习特征融合的权重系数, 使网络可以自适应调整不同尺度目标的感受野大小, 实现对不同尺度信息的筛选, 进而提高小尺度目标的分类精度。实验结果表明, 与其他模型相比, 本文模型在 ISPRS 机载 LiDAR 点云上的平均 F_1 分数为 72.2%, 且对电力线和汽车类别均取得了最高的分类精度, F_1 分数分别为 64.3% 和 79.9%。

关键词 遥感; 机载激光雷达点云; 深度学习; 注意力机制; 特征融合; 语义分割

中图分类号 P237

文献标志码 A

doi: 10.3788/LOP202158.2028004

Semantic Segmentation of LiDAR Point Cloud Based on CAFF-PointNet

Lai Ming, Zhao Jiankang*, Liu Chuanqi, Cui Chao, Long Haihui

Department of Instrument Science & Engineering, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

Abstract Herein, we propose a convolutional neural network based on channel attention mechanism for multiscale feature fusion regarding the characteristics of LiDAR point clouds, such as the complex geometric structure and extreme scale variations among different categories, resulting in the issue of low classification accuracy of small targets. First, low-level features (planarity, linearity, normal vector, and eigen entropy) are calculated for each point by setting a spherical neighborhood, and they are fused with high-level features acquired by the network to improve the geometry awareness of the constructed model. Then, a multiscale feature fusion module is designed based on the channel attention mechanism to learn fusion weight coefficient so that the network can adapt to the receptive field size of different scale objects and realize different scales information filtering, which improves the classification performance of the small-scale object. According to the experiments, the average F_1 score using the ISPRS Vaihingen 3D Semantic Labeling benchmark is 72.2%. Compared with other algorithms, our model has the highest classification accuracy in the powerline and car categories with F_1 scores of 64.3% and 79.9%, respectively.

Key words remote sensing; airborne LiDAR point cloud; deep learning; attention mechanism; feature fusion; semantic segmentation

OCIS codes 280.3640; 100.4996; 100.3008

1 引言

近年来,随着无人机技术的成熟和激光雷达

(LiDAR)技术的不断发展,机载 LiDAR 已被广泛应用于城市三维建模^[1]、地形测绘^[2]以及智慧农业监测^[3]等低空遥感领域。高准确率

收稿日期: 2020-11-17; 修回日期: 2020-12-18; 录用日期: 2021-01-07

基金项目: 国家重点研发计划(2016YFC0200400)、上海商用飞机系统工程联合研究基金(2020SCASEJ)

通信作者: *zhaojiankang@sjtu.edu.cn

点云语义分割算法是生成后续测绘产品的关键,因此,车载 LiDAR 点云语义分割算法是相关领域的研究热点。点云语义分割算法通常分为基于浅层机器学习的算法和基于深度学习的算法。基于浅层机器学习的算法一般采用人工设计的特征,通过传统机器学习分类器进行分类^[4-7]。文献[4]通过点云邻域协方差矩阵提取点云的局部几何特征;文献[6-7]通过地面点滤波算法分离点云中的地面点和地物点,再提取地物点的几何特征。该类型算法的缺点是特征设计依赖于研究者的先验知识,在面对复杂场景车载 LiDAR 点云时分类效果欠佳。

基于深度学习的点云语义分割算法根据模型输入的数据类型可分为基于投影、基于体素化和基于点云三类^[8]。文献[9-10]采用投影或体素化方法将非结构化的点云转换为结构化的二维图像或三维网格,这类算法虽然解决了点云非结构化的问题,但转换过程会破坏点云的三维结构,从而不可避免地导致点云信息丢失,分类准确率有待提高。Qi 等^[11]提出了基于点云直接输入的网络 PointNet,该网络使用多层感知机独立地对每个点特征进行升维,避免了点云非结构化的问题;而且,该网络采用对称函数聚合点云特征,解决了点云的排列不变性。近年来,受 PointNet 网络的启发,研究人员提出了一系列基于点云的深度学习网络,它们大致可分为基于多层感知机、基于卷积操作、基于图卷积和基于循环神经网络(RNN)等几种类型^[8]。文献[12]提出了 PointCNN,该网络通过多层感知机学习 χ 变化矩阵,规范化点云排列顺序。文献[13-14]采用不同的卷积核权重对点云进行逐点卷积操作,同样解决了点云的排列不变性问题。文献[15-16]采用图卷积替代多层感知机来解决点云非结构化、排列不变性问题。虽然图卷积网络对局部特征的提取能力较强,但网络的复杂度较高,计算效率较低。文献[17]基于 PointNet 网络使用 RNN 引入点云的上下文信息,提高了分类准确率。为了增强 PointNet 对局部特征的提取能力,Qi 等^[18]提出了基于编码-解码结构的 PointNet++,在点云邻域迭代中使用 PointNet 提取局部特征;但该网络依然孤立地处理点云信息,无法捕捉复杂的几何结构。Jiang 等^[19]受二维图像中尺度不变特征变换(SIFT)的启发提出了 PointSIFT 网络,该网络对每个点的 8 个方向的信息进行编码,改善了局部特征的方向不变性。Li 等^[20]提出了基于几何注意力机制的 GADH 网络,并通过多层级输出的方法来解决车载 LiDAR

点云不同地物尺度变化大的问题。GADH 虽然改善了网络的几何结构感知能力,但仍然采用固定权重比的方式进行多尺度特征融合,对不同尺度特征仍然无法实现加权关注,小尺度目标的感受野无法自适应调整。

针对上述问题,本文提出了一种基于通道注意力机制进行多尺度特征融合的网络 CAFF-PointNet(PointNet based on channel attention feature fusion),用以充分挖掘点云的局部几何特征,优化多尺度特征融合的权重比,提高模型对小尺度特征的识别准确率。本文的主要贡献和创新点如下:1)设计几何结构感知模块(GAM),提取点云的几何特征,同时构建其与深层次特征的映射关系,以增强网络对局部几何结构的表征能力;2)设计注意力特征融合模块(AFF),增强层与层间特征图的信息关联,使模型可以根据点云上下文信息和局部特征自适应调整融合权重系数,实现对不同尺度特征的加权关注,增强模型对小尺度目标特征的提取能力。

2 基于注意力机制的特征融合点云语义分割模型

2.1 模型结构

本文以 PointSIFT^[19]为基础框架构建 CAFF-PointNet 模型,该模型分为几何结构感知模块和基于注意力特征融合模块的主干网络。模型的总体框架如图 1 所示。

第一部分为几何结构感知模块。本文基于点云局部邻域的协方差张量特征值和高程提取点云的局部几何特征,通过多层感知机将低层次的几何特征映射到多维特征空间中,采用 Softmax 函数进行归一化,并将归一化结果与主干网络提取的深层次特征进行加权融合,优化深层次特征的几何结构细节。

第二部分为基于注意力特征融合模块的主干网络。主干网络通过逐级降采样获得不同规模和不同尺度的局部特征,并通过逐级上采样方法恢复每个点的高维局部特征。注意力特征融合模块替代了基础网络中的拼接操作,可以对不同层级间的多尺度特征进行融合。将浅层小尺度特征和深层大尺度特征输入到注意力特征融合模块中,根据输入信息自适应调整特征感受野的大小,学习特征融合的权重系数,实现对不同尺度特征的加强优化。最后,通过迭代多个全连接层(FC)输出点云的语义概率图。

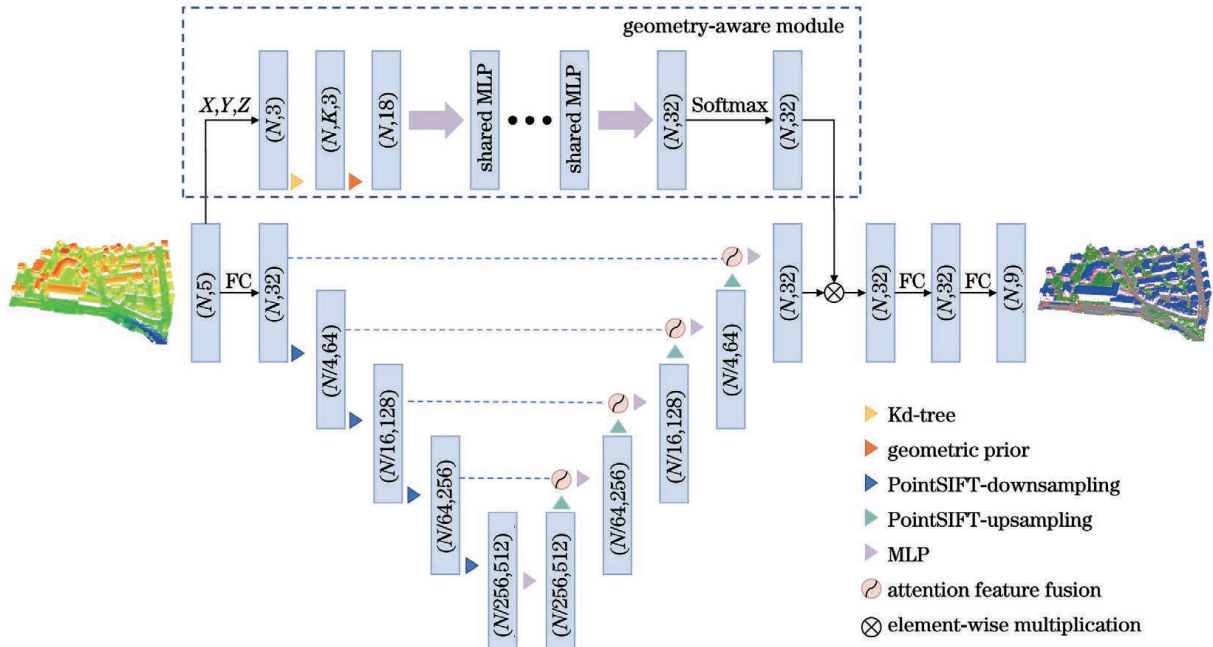


图 1 模型的总体框架图

Fig. 1 Overall framework of the model

2.2 几何结构感知模块

原始机载 LiDAR 点云数据中通常包含每个点的三维坐标、颜色、反射强度以及回波次数等信息，提取点云的局部特征是分割点云的核心步骤。PointNet++ 采用层级结构迭代的方式使用 PointNet 网络对点云的局部邻域进行特征提取。对于给定的某点 x_i 及其邻域 $x_j \in \mathcal{P}(x_i)$ ，点云局部特征的提取过程为

$$f_{\mathcal{P}} = \gamma \{ \mathcal{M}(f_{x_j}), \forall x_j \}, \forall x_j \in \mathcal{P}(x_i), \quad (1)$$

式中： f_{x_j} 是邻域点 x_j 的特征向量； \mathcal{M} 表示多层感知神经网络； γ 为最大池化函数。其中 x_i 由输入点云通过最远点采样算法获取，邻域范围为 $\mathcal{P} = \{x_j \mid \|x_j - x_i\| \leq r\}$ ， r 为球体半径。使用多层感知机独立地将邻域范围内的每个点从低维空间向高维空间进行映射，采用最大池化函数聚合邻域点的特征向量，并将特征向量赋予中心点作为该点的局部特征 $f_{\mathcal{P}}$ 。然而，该算法难以学习邻域点的相对位置关系，对局部几何结构的感知能力较弱。机载 LiDAR 点云中富含丰富的几何实例，利用点云局部几何结构的差异可以有效地区分机载 LiDAR 点云中的不同类别，如地面点、建筑物点和植被点等。Weinmann 等^[4]基于协方差特征值设计了多种手工特征，并采用传统的机器学习方法对点云进行分割。这些浅层次的手工特征包含着丰富的几何结构细节，相对于神经网络提取的深层次特征，其对复杂点

云场景的区分度较弱。本文提出的几何结构感知模块基于浅层次的几何特征 $G_{\mathcal{P}}$ ，对深度学习网络提取的深层次特征 $f_{\mathcal{P}}$ 进行优化，以增强其对点云局部几何结构的感知能力。首先计算浅层次的几何特征 $G_{\mathcal{P}}$ 。对于给定的一点 x_i ，通过 K 近邻(KNN)算法搜索其最邻近的 k 个点，构建邻域点集 $P = \{p_1, p_2, \dots, p_i, \dots, p_k\}$ ，基于该局部邻域计算邻近点协方差张量 $C_{\mathcal{P}}$ ，计算公式为

$$C_P = \frac{1}{k} \sum_{i=1}^k (p_i - \hat{p})(p_i - \hat{p})^T, \quad (2)$$

式中： \hat{p} 为邻域 \mathcal{P} 中心点的向量。由协方差张量 $C_{\mathcal{P}}$ 可以计算得到三个特征值 $\lambda_1 \geq \lambda_2 \geq \lambda_3$ ，对其进行归一化处理使 $\lambda_1 + \lambda_2 + \lambda_3 = 1$ 。基于三个特征值可以计算该邻域的线性度 L_{λ} 、平面度 P_{λ} 、球形度 S_{λ} 、全方差 O_{λ} 、特征熵 E_{λ} 、各向异性 A_{λ} 、特征值之和 Σ_{λ} ，以及曲率的变化率 C_{λ} 。它们的计算公式如表 1 所示。

除此之外，机载 LiDAR 点云中不同类别的地物点的高程分布也有显著差异。本文提取了点云局部邻域内的高程特征，如邻域内高程最大值 z_{\max} 、邻域内高程最小值 z_{\min} 、邻域内高程范围 $H_r = z_{\max} - z_{\min}$ 、当前点与邻域点高程最小值的差值 $H_b = z - z_{\min}$ 、邻域内高程的平均值 $H_m = \frac{1}{k} \sum_{i=1}^k z_i$ ，以及邻域

内高程的方差 $H_v = \sqrt{\frac{1}{k} \sum_{i=1}^k (z_i - H_m)^2}$ 。最终，浅

表 1 几何特征的计算公式

Table 1 Calculation formulae for geometric features

Geometric features	L_λ	P_λ	S_λ	O_λ	E_λ	A_λ	Σ_λ	C_λ
Formula	$\frac{\lambda_1 - \lambda_2}{\lambda_1}$	$\frac{\lambda_2 - \lambda_3}{\lambda_1}$	$\frac{\lambda_3}{\lambda_1}$	$\sqrt[3]{\lambda_1 \lambda_2 \lambda_3}$	$\frac{\lambda_1 - \lambda_3}{\lambda_1}$	$-\sum_{i=1}^3 \lambda_i \ln(\lambda_i)$	$\lambda_1 + \lambda_2 + \lambda_3$	$\frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}$

层次的几何特征 $\mathbf{G}_\varphi = [L_\lambda, P_\lambda, S_\lambda, O_\lambda, A_\lambda, E_\lambda, \Sigma_\lambda, C_\lambda, z, z_{\max}, z_{\min}, H_r, H_b, H_m, H_v, v_x, v_y, v_z]$ 。其中 $[v_x, v_y, v_z]$ 为当前点的法向量,由协方差张量最小特征值 λ_3 对应的特征向量经正交分解得到。

几何特征 \mathbf{G}_φ 为 18 维特征向量。为了与深度学习网络提取的深层次特征 \mathbf{f}_φ 进行特征维度对齐,本文使用多层感知机对 \mathbf{G}_φ 进行特征映射,然后通过 Softmax 函数在特征维度上对其进行归一化,获得 \mathbf{f}_φ 的加权系数矩阵 \mathbf{w}_φ ,最后通过逐元素乘法得到融合后的局部特征,增强网络对点云局部几何结构的感知能力。计算公式为

$$\zeta(\mathbf{f}_\varphi) = \mathbf{w}_\varphi \cdot \mathbf{f}_\varphi = \mathcal{N}\{\mathcal{M}(\mathbf{G}_\varphi)\} \cdot \mathbf{f}_\varphi, \quad (3)$$

式中: \mathcal{M} 由多层感知机、批归一化层(BN)和 ReLU 激活函数组成; \mathcal{N} 为归一化函数,本文采用 Softmax 函数进行归一化。

2.3 基于注意力机制的特征融合

机载 LiDAR 点云中不同类别地物的尺度不一,如建筑物的尺度较大,汽车的尺度较小。为了解决点云中尺度变化的问题,PointNet++ 网络采用层级结构对输入点云进行下采样的同时,逐层扩大局部特征的感受野,并在特征传播过程中对不同尺度的特征进行拼接融合。然而拼接融合的方式采用固定的权重比对特征进行融合,忽略了不同尺度特征间的差异。浅层次的小尺度特征具有丰富的局部结构细节,而深层次的大尺度特征包含了上下文信息和粗糙的语义特征。为了对不同尺度的特征进行有效的融合互补,在文献[21]和文献[22]的启发下,本文提出了基于注意力机制的特征融合模块。该特征融合模块考虑了不同尺度特征间的相关性,动态地调整融合的权重系数,在保留小尺度特征局部结构细节的同时融合上下文信息,可以提取更高分辨率的语义特征。该模块结构如图 2 所示。

对于两个特征 $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{C \times N}$, 其中 \mathbf{X} 为编码过程中网络提取的小尺度特征, \mathbf{Y} 为特征传播过程中由反距离插值得到的大尺度特征, C 和 N 分别代表特征的通道数和点云的大小。首先对 \mathbf{X} 和 \mathbf{Y} 按元素相加进行初步融合,得到特征 $\mathbf{X}' \in \mathbb{R}^{C \times N}$;然后将

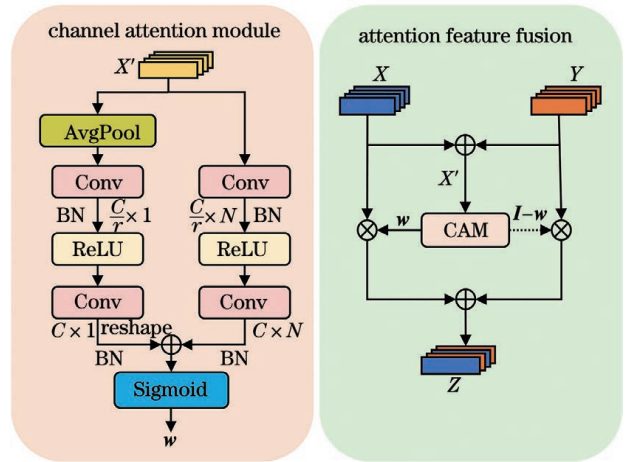


图 2 基于注意力机制的特征融合模块

特征 \mathbf{X}' 送入通道注意力模块中,分别提取全局特征 $g(\mathbf{X}')$ 和局部特征 $L(\mathbf{X}')$ 。全局特征的计算过程为 $g(\mathbf{X}') = \mathcal{B}\{W_2 \delta\{\mathcal{B}\{W_1 [\Gamma(\mathbf{X}')]\}\}\}$, (4) 局部特征的计算过程为

$$L(\mathbf{X}') = \mathcal{B}\{W_2 \delta\{\mathcal{B}[W_1(\mathbf{X}')]\}\}, \quad (5)$$

式中: $\Gamma(\mathbf{X}')$ 为全局平均池化函数; \mathcal{B} 为批归一化层; δ 为 ReLU 激活函数; W_2 和 W_1 均为全连接层。对全局特征 $g(\mathbf{X}')$ 进行广播复制,并将其与局部特征 $L(\mathbf{X}')$ 按元素相加,然后通过 Sigmoid 层进行归一化,得到特征融合的权重系数 $\mathbf{w} \in \mathbb{R}^{C \times N}$ 。计算公式为

$$\mathbf{w} = \sigma[L(\mathbf{X}') \oplus g(\mathbf{X}')], \quad (6)$$

式中: σ 为 Sigmoid 函数; \oplus 为按元素相加。经过上述操作后再将特征 \mathbf{X} 和 \mathbf{Y} 根据权重系数 \mathbf{w} 进行加权融合。融合过程为

$$\mathbf{Z} = \mathbf{w} \otimes \mathbf{X} + (\mathbf{I} - \mathbf{w}) \otimes \mathbf{Y}, \quad (7)$$

式中: \mathbf{Z} 为融合后的特征; \otimes 为按元素相乘; \mathbf{I} 为单位矩阵。

3 实验与分析

3.1 实验数据与评价指标

本文选用国际摄影测量与遥感协会(ISPRS)提供的德国 Vaihingen 地区机载 LiDAR 点云数据集^[23]进行实验。点云数据由徕卡 ALS50 系统采集,点密度为 $4 \sim 8 \text{ m}^{-2}$,每个激光点包含三维坐标、

反射强度以及回波次数等信息。数据集被人工分为地面、建筑物屋顶和树木等 9 个类别,标准分类结果由 ISPRS 提供。图 3 为数据集按照高程着色的效果图,图 3(a)为训练集点云,共有 753786 个点。图 3(b)为测试集点云,共有 411722 个点。表 2 列举了训练集和测试集中各类别的点数信息。

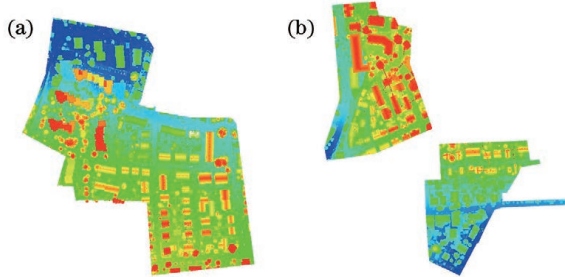


图 3 训练集和测试集数据(按高程着色)。(a)训练集;
(b)测试集

Fig. 3 Training and test sets (colored by height).

(a) Training set; (b) test set

表 2 训练集和测试集中每个类别的点数

Table 2 The number of points in each category for the training and test sets

Category	Training set	Test set
Powerline	546	600
Low vegetation	180850	98690
Impervious surfaces	193723	101986
Car	4614	3708
Fence/Hedge	12070	7422
Roof	152045	109048
Facade	27250	11224
Shrub	47605	24818
Tree	135173	54226
Total	753876	411722

为了更好地评估模型的性能,本文采用平均准确率(O_A)、查准率(P)、查全率(R)以及 F_1 分数作为评估指标。平均准确率表示测试集中正确分类的点的占比, F_1 分数为综合考虑查准率和查全率的一种评价指标。在不同类别点数差异较大的场景中,与平均准确率相比, F_1 分数对模型性能的评价更为准确。平均准确率、查准率、查全率以及 F_1 分数的计算公式为

$$O_A = \frac{T_P + T_N}{T_P + T_N + F_P + F_N}, \quad (8)$$

$$P = \frac{T_P}{T_P + F_P}, \quad (9)$$

$$R = \frac{T_P}{T_P + F_N}, \quad (10)$$

$$F_1 = \frac{2PR}{P + R}, \quad (11)$$

式中: T_P 表示某类样本中分类正确的点数; T_N 表示其他类样本中分类正确的点数; F_P 表示某类样本中分类错误的点数; F_N 表示其他类样本中分类错误的点数。

3.2 实验的实现与设置

本文使用 Pytorch 1.5.0 框架实现所提点云语义分割模型,并在 CPU 为 i7-7700k、GPU 为 NVIDIA 1080Ti 的计算机上进行训练。实验采用 Adam 优化器拟合模型,初始学习率设置为 0.001,衰减率为 0.7,动量(momentum)设置为 0.9,批大小设置为 12。为了避免网络出现过拟合,增强网络的泛化性能,在训练过程中对数据集进行数据增强操作,训练集绕 x 轴、 y 轴或 z 轴旋转一个 $0 \sim 2\pi$ 之间的随机角度,并在最后一个全连接层中加入比率为 0.5 的 Dropout 层。由于训练集中不同类别点的数量差别较大,网络在训练过程中更倾向于拟合点数较多的类别,因此点数较少的类别的分类准确率较低。鉴于此,本文采用带权重比的交叉熵损失函数监督网络训练。交叉熵损失函数的权重比因子基于训练样本中不同类别的点数比例进行调整,其计算公式为

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

式中: α 为网络参数; N_c 为类别 c 的点数; C 为训练样本的总类别数。

根据 PointNet++ 网络的参数, α 设置为 1.2。训练的最大轮回次数为 500,训练周期约 8 h。训练集损失函数曲线和测试集平均准确率曲线如图 4 所

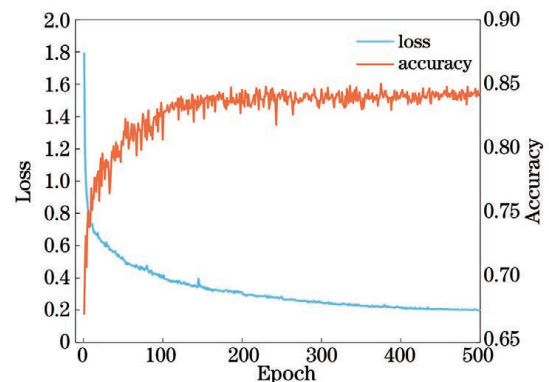


图 4 训练集损失函数曲线和测试集平均准确率曲线

Fig. 4 Loss function graph of training set and average accuracy graph of test set

示。可见,随着训练轮次的增加,训练集损失函数曲线和测试集平均准确率曲线逐渐收敛。这说明,本文提出的基于通道注意力机制的特征融合机载 LiDAR 点云语义分割模型响应快速且具有良好的收敛性。

3.3 实验结果与分析

为了验证本文模型的泛化性能,利用在训练集上得到的分类模型对测试集进行分类,分类结果如表 3 和图 5 所示。

表 3 测试集分类结果混淆矩阵

Table 3 Confusion matrix of testing set classification results

Category	Powerline	Low vegetation	Impervious surfaces	Car	Fence	Roof	Facade	Shrub	Tree
Powerline	371	2	1	0	0	156	59	0	11
Low vegetation	2	86137	8247	198	301	463	205	2764	373
Impervious surfaces	4	9139	92287	76	49	335	15	71	10
Car	1	168	234	2945	138	138	22	62	0
Fence	1	944	222	67	2188	112	153	3457	278
Roof	60	1310	99	22	23	102102	609	578	4245
Facade	61	1156	136	142	16	2288	6193	733	499
Shrub	7	6571	255	199	717	812	444	12308	3505
Tree	47	2394	23	17	139	1093	619	5060	44834
Precision / %	66.9	89.8	90.9	80.3	61.2	94.9	74.4	49.1	83.4
Recall / %	61.8	87.2	90.4	79.4	29.4	93.6	55.1	49.5	82.6
F_1 score / %	64.3	83.4	90.7	79.9	39.8	94.3	63.4	49.4	83.0

由表 3 可以看出,本文模型对低矮植被、地面、建筑物屋顶以及树木等类别的分类效果较好, F_1 分数均超过 80%,其中地面的 F_1 分数为 90.7%,建筑物屋顶的 F_1 分数为 94.3%。此外,电力线和汽车等小尺度类别也取得了不错的分类精度。对比图 5(a)和图 5(b)可以看出部分

栅栏存在误识别。结合表 3 可知,很大一部分栅栏被误识别成灌木和低矮植被。这是由于栅栏类别在训练集中的点数较少,且其高程特征与灌木、低矮植被相似,模型没有充分学习它们内部语义特征的差异,因此将部分栅栏误识别成灌木和低矮植被。

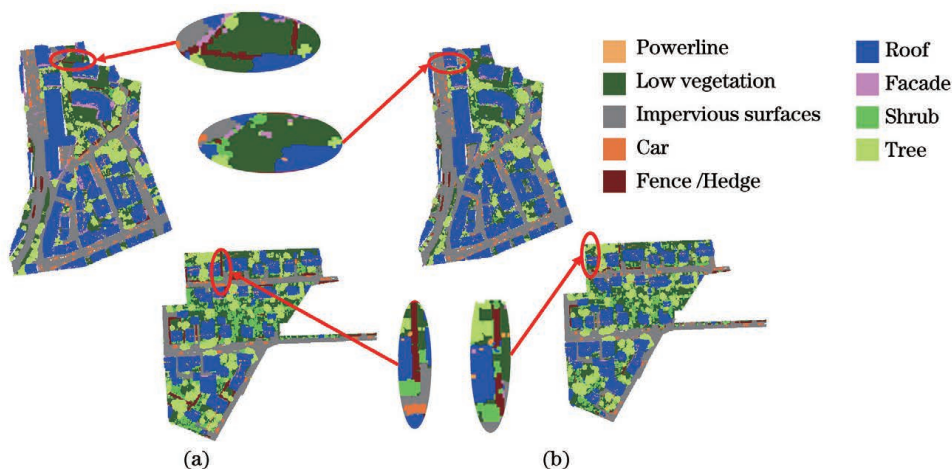


图 5 本文模型分类结果与标准分类数据的对比。(a) 标准值;(b) 本文模型

Fig. 5 Comparison of classification results between our model and ground truth. (a) Ground truth; (b) ours

为了验证本文构建的几何结构感知模块和注意力特征融合模块的有效性,在实验中构造不同的网络进行消融实验。其中基础模型基于 PointSIFT 构

建。消融实验结果如图 6 和表 4 所示。由表 4 可以看出,在基础模型上分别添加几何结构感知模块和注意力特征融合模块后,模型的平均准确率和平均

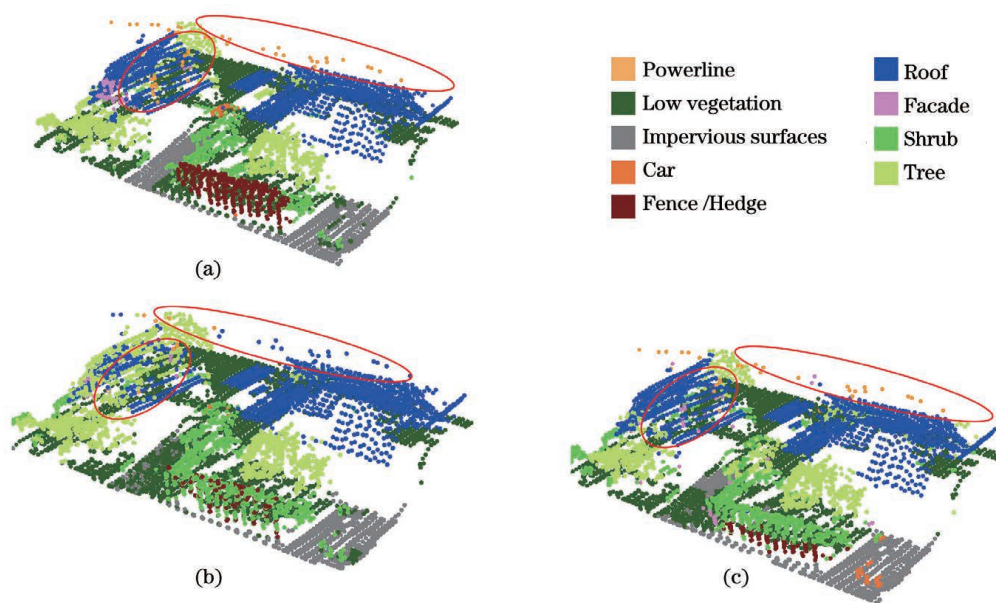


图 6 基础模型与本文模型的分类对比图。(a) 标准值；(b) 基础模型；(c) 本文模型

Fig. 6 Classification results of baseline and our model. (a) Ground truth ; (b) baseline; (c) ours

表 4 不同模型的消融实验结果

Table 4 Ablation experimental results of different models

%

Category	F_1 score			
	Baseline	Ours(with GAM)	Ours(with AFF)	Ours(final)
Powerline	55.7	56.2	61.2	64.3
Low vegetation	80.7	80.8	82.1	83.4
Impervious surfaces	90.9	91.3	89.6	90.7
Car	77.8	73.6	75.2	79.9
Fence	30.5	28.4	34.5	39.8
Roof	92.5	94.5	93.8	94.3
Facade	56.9	55.3	60.3	63.4
Shrub	44.4	43.0	47.6	49.4
Tree	79.6	81.2	80.1	83.0
OA	82.2	82.6	83.2	84.8
Average of F_1	67.7	68.3	69.5	72.2

F_1 分数均有所提高,几何结构感知模块对地面、建筑物等平面特征强的类别的分类精度较高,注意力特征融合模块能提高小尺度物体的识别准确率。本文结合几何结构感知模块和注意力特征融合模块构建的 CAFF-PointNet 网络的平均准确率和平均 F_1 分数与基础模型相比,分别提高了 2.6 个百分点和 4.5 个百分点,而且对小尺度类别(如电力线、栅栏和灌木)的分类精度有较大提高。对比图 6(b)和图 6(c)可以看出,对于在基础模型中被误识别的电力线和建筑物屋顶,本文模型能够正确分类。以上表明,本文模型能够充分挖掘低层次特征的几何结

构细节,并通过上下文几何信息和局部特征的融合,得到语义信息更加丰富的多尺度上下文特征,对大尺度类别和小尺度类别均保持了较高的分类精度。

此外,为了与其他分类模型进行对比,选取 ISPRS 官网上基于深度学习方法且排名前三的 RIT_1^[24]、WhuY4^[25]和 NANJ2^[26]模型,以及基于点云表示的 PointNet++和 PointSIFT 模型与本文模型进行对比,对比结果如表 5 所示。表中 RIT_1 在 PointNet 模型上进行扩展,设计了直接处理点云的全卷积网络,并基于地面点归一化地物点的高程,将通过遥感图像提取的点云的光谱信息作为模型的

表 5 本文模型与不同模型的定量对比

Table 5 Quantitative comparison between our model and other models

%

Category	F_1 score					
	RIT_1 ^[24]	WhuY4 ^[25]	NANJ2 ^[26]	PointNet++ ^[18]	PointSIFT ^[19]	Ours
Powerline	37.5	42.5	62	57.9	55.7	64.3
Low vegetation	77.9	82.7	88.8	79.6	80.7	83.4
Impervious surfaces	91.5	91.4	91.2	90.6	90.9	90.7
Car	73.4	74.7	66.7	66.1	77.8	79.9
Fence	18	53.7	40.7	31.5	30.5	39.8
Roof	94	94.3	93.6	91.6	92.5	94.3
Facade	49.3	53.1	42.6	54.3	56.9	63.4
Shrub	45.9	47.9	55.9	41.6	44.4	49.4
Tree	82.5	82.8	82.6	77	79.6	83
OA	81.6	84.9	85.2	81.2	82.2	84.8
Average of F_1	63.3	69.2	69.3	65.6	67.7	72.2

输入,实现端到端的分类。WhuY4 将点云的低层次特征(如归一化高程、平面度、直线度和曲率等)转换为二维图像,通过多尺度卷积神经网络实现点云语义分割。NANJ2 根据 RGB、高程、强度和粗糙度等特征生成多尺度二维特征图,设计卷积神经网络学习点云的深层次特征,最后使用 Softmax 回归分类器(SRC)生成数据的分类结果。

由表 5 可知:本文模型的平均 F_1 分数为 72.2%,比其他模型中的最高者(NANJ2)高 2.9 个百分点。本文模型的平均准确率为 84.8%,与

NANJ2 相比低 0.4 个百分点。这是由于 NANJ2 模型使用了额外的 RGB 信息作为输入,而本文模型的输入仅为点云的三维坐标、反射强度和回波次数。对比各类别的 F_1 分数可知,本文模型在电力线、汽车、建筑物屋顶、墙立面以及树木类别上取得了最高的分类精度,并且在低矮植被和地面类别上仍然保持了较高的分类精度。为了直观地对比不同模型的分类效果,本文给出 RIT_1、WhuY4、NANJ2 和本文模型在某区域上的分类对比图,如图 7 所示。

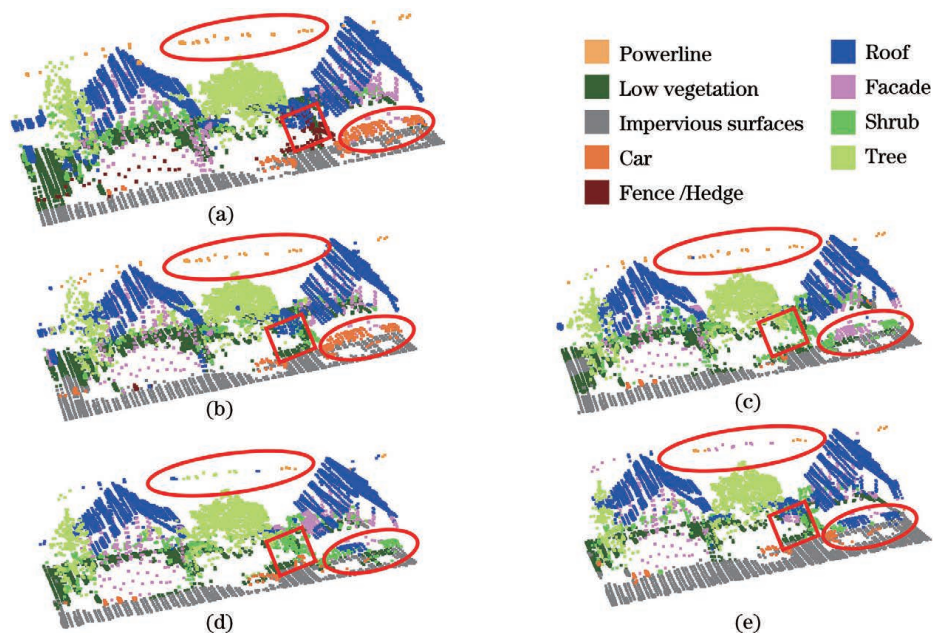


图 7 不同模型分类效果的对比。(a) 标准值;(b) 本文模型;(c) NANJ2;(d) RIT_1;(e) WhuY4

Fig. 7 Comparison of classification results of different models. (a) Ground truth; (b) ours; (c) NANJ2; (d) RIT_1; (e) WhuY4

对比图 7(a)~(e)可以看出,在椭圆形区域,其他模型均出现了误识别的情况(将汽车点云误识别为灌木、建筑物屋顶等类别),而本文模型能正确地识别出汽车点云。这是由于该处的汽车点云与建筑物屋顶距离较近,且它们具有类似的几何结构,模型倾向于将其拟合成大尺度目标(建筑物屋顶),所以容易出现误识别;而本文模型通过学习特征融合的权重系数来调整特征的感受野,减少了无关信息的干扰,提高了小尺度目标(汽车点云)的识别准确率。在矩形区域,各模型对栅栏类别均存在误识别的情况,这可能是由于此处各类别点云距离较近,几何结构复杂,且栅栏类别在训练集中的点数较少,模型无法很好地学习语义特征的差异,导致分类准确率较低。

4 结 论

本文提出了一种基于通道注意力机制的特征融合机载 LiDAR 点云语义分割模型 CAFF-PointNet。针对机载 LiDAR 点云几何结构复杂的特点,设计几何结构感知模块,充分挖掘点云局部邻域的几何结构。为了提高机载 LiDAR 点云中小尺度目标的分类精度,设计注意力机制特征融合模块,使模型针对输入的多尺度特征能够自适应地调整不同感受野特征的权重系数,充分利用上下文信息和局部特征,减少冗余信息对模型的干扰。在 ISPRS 机载 LiDAR 点云数据集上对本文模型进行了验证,结果表明,与其他同类模型相比,本文模型的总体分类准确率得到了提升,特别是对小尺度类别(如电力线和汽车)取得了最高的分类精度。

然而,本文模型对栅栏、灌木等类别的分类精度仍有待提高,这是因为本文模型受训练集中栅栏、灌木等类别点数过少的限制,同时,对于距离较近或几何结构类似的点云,本文模型仍然难以捕捉其内部语义特征的差异。接下来,拟考虑融合点云的 RGB 信息,并构建 RGB 特征和几何特征的融合模块,以实现更高精度的语义分割。

参 考 文 献

- [1] Wang R S, Peethambaran J, Chen D. LiDAR point clouds to 3-D urban models: a review [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(2): 606-627.
- [2] Liu X Y. Airborne LiDAR for DEM generation: some critical issues[J]. Progress in Physical Geography: Earth and Environment, 2008, 32(1): 31-49.
- [3] Liu T, Su W, Wang C, et al. A method of estimating maize LAI using airborne LiDAR data[J]. Journal of China Agricultural University, 2016, 21(3): 104-111.
刘婷, 苏伟, 王成, 等. 基于机载 LiDAR 数据的玉米叶面积指数反演[J]. 中国农业大学学报, 2016, 21(3): 104-111.
- [4] Weinmann M, Jutzi B, Mallet C. Feature relevance assessment for the semantic interpretation of 3D point cloud data[J]. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2013, II-5/W2: 313-318.
- [5] Zhang J X, Lin X G, Ning X G. SVM-based classification of segmented airborne LiDAR point clouds in urban areas [J]. Remote Sensing, 2013, 5(8): 3749-3775.
- [6] Shi X S, Cheng Y L, Zhao Z Y, et al. Point cloud classification algorithm based on IPTD and SVM[J]. Laser & Optoelectronics Progress, 2019, 56(16): 161002.
释小松, 程英蕾, 赵中阳, 等. 基于三角网滤波和支持向量机的点云分类算法 [J]. 激光与光电子学进展, 2019, 56(16):161002.
- [7] Xue D D, Cheng Y L, Shi X S, et al. Point clouds classification algorithm based on cloth filtering algorithm and improved random forest[J]. Laser & Optoelectronics Progress, 2020, 57(22): 221017.
薛豆豆, 程英蕾, 释小松, 等. 综合布料滤波与改进随机森林的点云分类算法 [J]. 激光与光电子学进展, 2020, 57(22):221017.
- [8] Guo Y L, Wang H Y, Hu Q Y, et al. Deep learning for 3D point clouds: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 5434: 1-27.
- [9] 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, Mexico. New York: IEEE Press, 2016: 2670-2675.
- [10] Lawin F J, Danelljan M, Tosteberg P, et al. Deep projective 3D semantic segmentation [M]//Felsberg M, Heyden A, Krüger N. Computer analysis of images and patterns. Lecture notes in computer science. Cham: Springer, 2017, 10424: 95-107.
- [11] Qi C R, 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] Li Y Y, Bu R, Sun M C, et al. PointCNN:

- convolution on X-transformed points [C] // Advances in Neural Information Processing Systems, December 3-8, 2018, Montréal, Canada. San Diego: NIPS, 2018: 820-830.
- [13] 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.
- [14] Boulch A. ConvPoint: continuous convolutions for point cloud processing [J]. Computers & Graphics, 2020, 88: 24-34.
- [15] Hou X D, Yu X X, Liu H P. 3D point cloud classification and segmentation model based on graph convolutional network [J]. Laser & Optoelectronics Progress, 2020, 57(18): 181019.
侯向丹, 于习欣, 刘洪普. 基于图卷积网络的三维点云分类分割模型 [J]. 激光与光电子学进展, 2020, 57(18): 181019.
- [16] Phan A V, Nguyen M L, Nguyen Y L H, et al. DGCNN: a convolutional neural network over large-scale labeled graphs [J]. Neural Networks, 2018, 108: 533-543.
- [17] Huang Q G, Wang W Y, Neumann U. Recurrent slice networks for 3D segmentation of point clouds [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 2626-2635.
- [18] Qi C R, Yi L, Su H, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space [C] // Advances in Neural Information Processing Systems, December 3-8, 2017, Long Beach, CA, USA. Canda. Red Hook: Curran Associates Inc., 2017: 5099-5108.
- [19] 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) [2020-11-10]. <https://arxiv.org/abs/1807.00652>.
- [20] 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.
- [21] Oktay O, Schlemper J, Folgoc L L, et al. Attention U-Net: learning where to look for the pancreas [EB/OL]. (2018-08-11) [2020-11-10]. <https://arxiv.org/abs/1804.03999>.
- [22] Dai Y M, Gieseke F, Oehmcke S, et al. Attentional feature fusion [EB/OL]. (2020-09-29) [2020-11-10]. <https://arxiv.org/abs/2009.14082>.
- [23] 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.
- [24] Yousefhussein 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.
- [25] 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.
- [26] Zhao R B, Pang M Y, Wang J D. Classifying airborne LiDAR point clouds via deep features learned by a multi-scale convolutional neural network [J]. International Journal of Geographical Information Science, 2018, 32(5): 960-979.