

一种考虑法矢量的点云语义分割网络

尚鹏飞^{1,2}, 陈义^{1,2*}, 吕伟嘉^{1,2}, 郑芳^{1,2}, 王杰龙^{1,2}¹同济大学测绘与地理信息学院, 上海 200092;²现代工程测量国家测绘地理信息局重点实验室, 上海 200092

摘要 在基于深度学习的点云分类任务中, PointNet 模型直接以点云的三维坐标作为输入, 但其存在着对形状不规则物体的分类效果不理想的问题。在 PointNet 模型的基础上增加一个法矢量估计的模块, 提出一种考虑点云法矢量的语义分割网络。其中, 法矢量估计方法采用的是主成分分析方法。实验结果表明, 改进模型的总体准确率、平均类别准确度和平均类别交互比相较于原始模型分别提升了 2.3 个百分点、7.1 个百分点和 3.9 个百分点。13 个语义类别中有 10 个类别的分类效果得到提升, 其中对沙发和柱状物的分类准确度分别提升了 45.6 个百分点和 42.2 个百分点, 平均类别交互比分别提升了 19.8 个百分点和 25.0 个百分点。结果表明, 考虑法矢量的 PointNet 网络能够在一定程度上提升网络的整体性能, 对沙发和柱状物的分类效果有显著提升。

关键词 图像处理; 点云; 深度学习; 语义分割; 法矢量; 主成分分析

中图分类号 TP391 文献标志码 A

DOI: 10.3788/LOP202259.1610011

Point-Cloud Semantic Segmentation Network Considering Normals

Shang Pengfei^{1,2}, Chen Yi^{1,2*}, Lv Weijia^{1,2}, Zheng Fang^{1,2}, Wang Jielong^{1,2}¹College of Surveying and Geo-Informatics, Tongji University, Shanghai 200092, China;²Key Laboratory of Modern Engineering Surveying of National Administration of Surveying, Mapping and Geoinformation, Shanghai 200092, China

Abstract In deep learning-based point-cloud semantic classification, PointNet considers the three-dimensional coordinates of the point cloud as a direct input, however, the classification of irregular shape objects is a challenge. In this study, we propose a semantic segmentation network considering the normals of point cloud by adding a normal estimation module on PointNet. We estimate the normals using a principal component analysis method. Compared with the original model, the overall accuracy, mean per-class accuracy, and mean per-class intersection-over-union of the improved model are improved by 2.3 percentage points, 7.1 percentage points, and 3.9 percentage points respectively. Among the 13 semantic classes, the classification accuracy for 10 classes is improved, of which the classification accuracy of sofa and column is improved by 45.6 percentage points and 42.2 percentage points, respectively, and the mean per-class intersection-over-union is improved by 19.8 percentage points and 25.0 percentage points, respectively. Results show that the semantic segmentation network considering normals can improve the overall performance of the network to a certain extent and can significantly improve the classification effect of sofa and column.

Key words image processing; point cloud; deep learning; semantic segmentation; normals; principal component analysis

1 引言

点云是继矢量地图和影像数据之后的第三类重要的时空数据源, 对三维空间的精细化描述具有无可替代的重要作用, 近年来被更多地应用在自动驾驶、机器人技术等场景理解实践中^[1-2]。深度学习近年来常被应用在计算机视觉中, 但是应用在点云中还处于起步

阶段, 在语义分割^[3-4]和目标检测^[5-6]任务中面临众多挑战。点云的语义分割指的是按照不同的语义类别附上不同的语义标签将点云场景分隔开, 通过深度学习的方法实现语义分割, 按照输入对象的不同分为基于点云变体的方法和基于原始点云的方法, 其中基于点云变体的方法包括多视图方法^[7-11]和体素化方法^[12-16]。但是多视图方法会受到投影角度的限制, 体素化方法

收稿日期: 2021-05-18; 修回日期: 2021-07-13; 录用日期: 2021-07-20

通信作者: *chenyi@tongji.edu.cn

则会增加量化误差。两种方法都会对点云的三维结构信息造成一定的损失。为了避免这种情况,文献[17]提出了一种将三维原始点云作为输入的语义分割网络 PointNet,对每个点做特征提取,实现语义分类。以 PointNet 框架为基础又陆续提出了 PointNet++^[18]、Mo-Net^[19]、PATs^[20]、PointWeb^[21]、PointASNL^[22] 等网络。PointNet 框架中对点云的预处理包括输入转换和归一化,缺少点云几何属性的计算。法向量作为点云必不可少的属性之一,其估计方法在点云处理中具有重要的位置,许多表面重建算法需要精确估计的法向量^[23]。在常见的法向量估计方法中,主成分分析(PCA)方法因其算法简单高效、稳定性强而被广泛使用^[24]。按照 PCA 的思想,又陆续提出了加入权重^[25-26]、自适应邻域^[27]、集成 PCA^[28] 的方法。但是改进的方法没有改变 PCA 方法是低通滤波器的本质,在增加计算复杂度的情况下对估计效果的提升是有限的。所以本文直接采用原始 PCA 方法估计法向量。采用经典的 PointNet 模型作为语义分割网络,在此基础上内嵌一个 PCA 模块实现点云法向量估计,探究法向量信息在点云语义分割中的作用。

2 基本原理

2.1 法向量估计

从离散点云中不能得到真实的法向量信息,所以一般通过划定一定范围作为点云中某一个点的局部邻域,将邻域中的点作为对象,进行平面或者曲面拟合,将平面或者曲面的法向量作为当前点的法向量^[29]。PCA 能够高效地实现这一过程^[30]。PCA 的思想是矩阵的奇异值分解(SVD),一个矩阵 \mathbf{A} 可以表示为

$$\mathbf{A} = \mathbf{U} \mathbf{S} \mathbf{V}^T, \quad (1)$$

式中: \mathbf{U} 是左奇异向量矩阵; \mathbf{V} 是右奇异向量矩阵,两者是正交阵。若 \mathbf{A} 的秩 $R(\mathbf{A})$ 满足 $R(\mathbf{A}) = p \leq \min(n, t)$, 则奇异值矩阵为

$$\mathbf{S} = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \ddots & & 0 \\ \vdots & 0 & \sigma_p & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}, \quad (2)$$

式中, $\sigma_i (i = 1, 2, \dots, p)$ 表示 \mathbf{A} 的非零奇异值,满足 $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p$, 对应的特征向量即左奇异特征向量 $\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$, 其中 $\mathbf{u}_i (i = 1, 2, \dots, n)$ 是 n 维的列向量。

在点云法向量估计中,每个点的 k 个最邻近点构成邻域 $N = \{(x_i, y_i, z_i) | i = 1, 2, \dots, k\}$, 其中 x, y, z 表示最近邻点的三维坐标,表示成矩阵 \mathbf{X} 为

$$\mathbf{X} = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots \\ x_k & y_k & z_k \end{bmatrix}, \quad (3)$$

构建其协方差阵 $\mathbf{X}^T \mathbf{X}$, 对 $\mathbf{X}^T \mathbf{X}$ 进行 PCA。奇异值分解将给定的一组变量线性变换为另一组不相关的变量,并且使变换后的第一变量的方差最大,即第一主成分,其他分量的方差依次递减^[31]。在这里,设想一个邻域中的点云,由于空间平面的第三主成分为垂直于空间平面的向量,所以局部拟合平面的法向量就是 $\mathbf{X}^T \mathbf{X}$ 的第三主成分,即 $\mathbf{X}^T \mathbf{X}$ 的最小奇异值对应的特征向量。不妨令式(1)中的 $\mathbf{A} = \mathbf{X}^T \mathbf{X}$, 则当前邻域估计的法向量就是最小奇异值 σ_p 对应的特征向量 \mathbf{u}_p ^[32]。对点云中的每个点构建邻域,重复上面的操作,计算每个点的法向量。为了加快检索速度,将点云数据构建成 K 维(KD)树结构,每个点取最邻近的 K 个点作为邻域点。实验表明,当 $K = 10$ 时模型的性能达到最佳,所以本文中的实验采用这一数值。

2.2 语义分割网络

由于点云数据的无序性和不规则性,很难应用标准卷积进行处理^[33]。PointNet 是第一个提出的直接以点云作为输入的深度学习网络,图 1 是其 vanilla 框架的基本组成,对应的数学原理如下式所示:

$$f(\{x_1, \dots, x_n\}) \approx \gamma g[h(x_1), \dots, h(x_n)], \quad (4)$$

式中: h 是多层感知机(MLP),将每个点的特征由 3 维扩展到更高维; g 是作为对称函数的最大池化层,将高维特征映射到 1 维空间; γ 是分类器,用来输出分类结果,PointNet 采用 MLP 作为分类器; f 是一个集合函数,根据广义近似原理,只要点集的特征维度扩展得足够高,理论上可以任意地估计该映射函数。另外,集合函数的输出是输入的统计结果,所以输入中的小扰动不会对输出造成大的影响,在一定程度上保证了模型的稳定性。

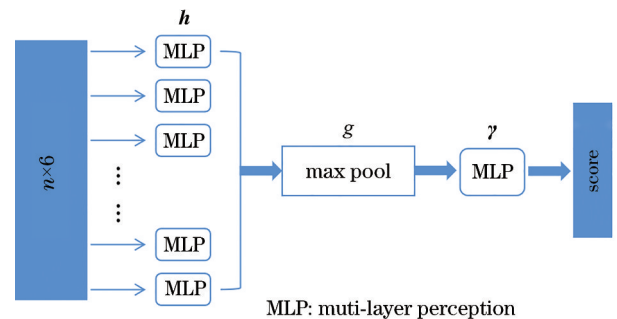


图 1 PointNet (vanilla) 框架

Fig. 1 Framework of PointNet (vanilla)

在 PointNet 的语义分割模型中,每个点的输入特征维度是 6 维(XYZ 坐标以及 RGB 颜色信息),在预处理过程中得到每个点在各自房间中的归一化坐标,特

征维度扩展至 9 维。对每个点采用 MLP 实现特征提取,得到高维的局部特征。这一过程是对每一个点进行特征提取,也就不存在标准卷积所要求的规则性问题。最大池化层将高维特征映射至 1 维,提取点云的全局特征。同时,最大池化层作为对称函数解决了点云的无序性问题。PointNet 的语义分割网络在 vanilla 框架的基础上增加了局部特征和全局特征的连接操作,丰富了每个点的特征。通过分类器,把点云的特征映射在分类空间,得到每个类别的分数。空间维度是由语义类别的个数决定的。本文的研究对象是 PointNet 的语义分割部分,为了方便表达,后文中统一用 PointNet 来表示 PointNet 的语义分割网络。

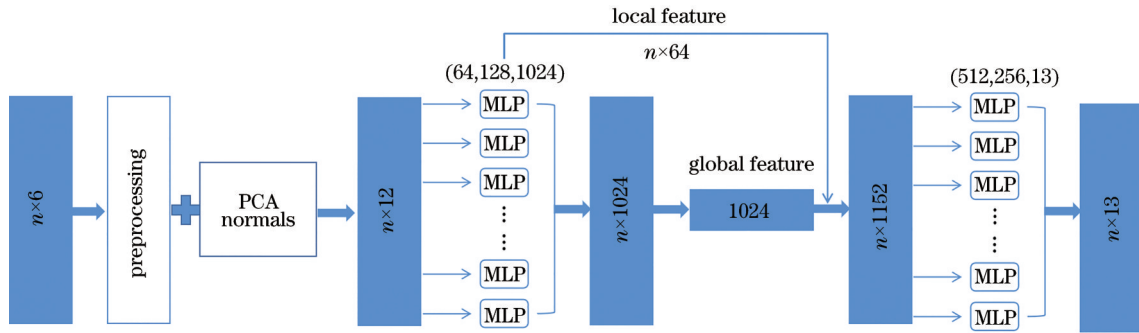


图 2 PPCAN 模型框架(改自文献[7])

Fig. 2 Framework of PPCAN (modified from Ref. [7])

采用 S3DIS 数据集^[34]作为实验数据,该数据集的场景包括 6 个区域共 272 个房间,其中包括会议室、复印室、办公区、走廊、礼堂、门厅、厕所等功能区,场景总面积达到了 6020 m²,点云数据中点的个数超过了 2 亿个,每个点提供了坐标和颜色信息。其中共有 13 种语义类别,天花板、地板、墙面、柱子、梁、窗户、门属于结构元素,桌子、椅子、书柜、沙发、板状物属于常见的物品和家具,不属于以上类别划为其他类别。

为了探究法向量在语义分割中所发挥的作用,设计了模型对比实验。一种不考虑法向量的 PointNet 模型,另一种是加入 PCA 法向量估计的 PPCAN 模型。S3DIS 数据集的前 5 个区域作为训练集,第 6 个区域作为测试集,训练集和测试集中点的数量分别占到数据集的 86% 和 14%。PointNet 模型在训练阶段并不是全部点都参与,而是将随机采样点用于训练模型。在测试阶段,PointNet 是对所有测试集的点均进行类别预测。在本文中,考虑到时间成本和硬件配置,在预测阶段同样采用了随机采样的方法,在预测阶段只预测了采样点的类别。为了验证模型的鲁棒性,对训练好的模型进行了多次测试。不同空间特征的几何结构的法向量分布有不同的特点,为了探究法向量对每一种类别的影响,本文统计分析了 13 种类别的分类结果。

为了得到量化的评价结果,本文采用在语义分割中常用的总体准确率(OA)、平均类别准确度(MA)以及平均类别交并比(mIoU)^[35-39]作为评价指标,计算公

3 实验分析

3.1 实验设计

本文实验的操作系统是 Windows 系统,计算机配置 8 核 i7-2600 CPU 和 8G 内存,深度学习环境基于 Anaconda 和 tensorflow1.15 搭建,编程语言是 python3.7。语义分割网络的核心代码在作者的 GitHub 上获得。在此基础上,在点云数据预处理模块中增加法向量估计功能,基于 python pcl 模块实现 PCA 法向量估计,改进后的语义分割网络本文暂且称作 PPCAN(PointNet with PCA estimated normals)模型,模型框架如图 2 所示。

式如下:

$$R_{OA} = \frac{S_{TP}}{N}, \quad (5)$$

$$A_{MA} = \frac{1}{13} \sum_{i=1}^{13} \frac{S_{TPi}}{N_i}, \quad (6)$$

$$A_{mIoU} = \frac{1}{13} \sum_{i=1}^{13} R_{IoUi}, \quad (7)$$

$$R_{IoUi} = \frac{S_{TPi}}{S_{TPi} + S_{FNi} + S_{FPi}}, \quad (8)$$

式中: R_{OA} 表示分类正确的点的个数占总点数的比例; A_{MA} 表示每种类别准确率的平均数; A_{mIoU} 表示全部类别 IoU 的平均值; R_{IoUi} 表示标签真实值与预测值的交集和并集的比值; S_{TP} 表示分类正确即真正; S_{FP} 表示错把其他类别分类成当前类别即假正; S_{FN} 表示错把当前类别分类成其他类别即假负; N 表示点云总点数;13 表示 13 种语义类别。

3.2 结果分析

两种模型的语义分割实验结果表明,相比于 PointNet 模型,PPCAN 模型在训练阶段(图 3)的起始精度略高,随着训练进行,两种模型的最终精度趋于一致。说明点云的法向量属性在训练初始阶段会给特征提取过程提供一定的帮助,这种帮助随着训练时间的增加逐渐减弱,直至可以忽略不计。这说明法向量所提供的信息可以通过训练达到对模型同样的贡献。

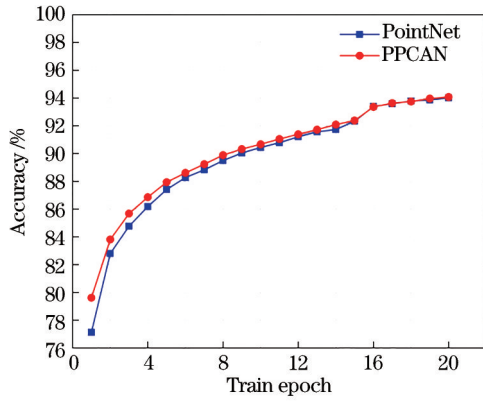


图 3 训练阶段准确率对比

Fig. 3 Comparison of accuracy in training epoch

比较 PointNet 和 PPCAN 模型的实验效果。两种模型分别对区域 6 的点云进行多次语义类别预测,将多次预测结果取平均,统计 3 种评价指标如表 1 所示。PPCAN 模型的 OA、MA 和 mIoU 均优于 PointNet 模型,分别提升了 2.3 个百分点、7.1 个百分点和 3.9 个百分点,网络的整体性能得到一定程度的提升。

鉴于不同空间几何体的法矢量的分布特征不尽相同,所有法矢量对于不同类别物体的预测贡献是不同的。表 2 中,统计了 PointNet 和 PPCAN 模型对 13 种类别的预测结果的准确率,其中沙发和柱状物

表 1 基于 S3DIS 的语义分割结果

Table 1 Results of semantic segmentation on S3DIS unit: %

Model	OA	MA	mIoU
PointNet	82.9	67.4	57.1
PPCAN	85.2	74.5	61.0

的准确率提升超过了 40 个百分点,分别提升 45.6 个百分点和 42.2 个百分点。如表 3 所示,对应的 IoU 分别提升了 19.8 个百分点和 25.0 个百分点,达到 30.1% 和 47.4%。图 4 中斜线处点云表示柱状物,剩余点云表示其余 12 种类别,比较图 4(a) 和图 4(b),图 4(b) 中分类正确的点的个数要明显多于图 4(a),这表明 PPCAN 对柱状物的分类效果要优于 PointNet 模型。PPCAN 在沙发类别上的优势则更加明显,图 5 中斜线处点云表示沙发类别,在图 5(a) 中 PointNet 只正确分类了极少的沙发点云,而 PPCAN 则正确分类了绝大多数的沙发类别的点云。究其原因,沙发的空间形态不规则,少有平整的平面或者特征鲜明的曲面,可以说沙发点云的法矢量是有区别于其他语义类别的;相比于沙发,柱状物的空间形态更加规则,不论是圆柱还是棱柱,其法矢量的分布总是近似以一条竖线为中心向外发散,辨识度较高。不难看出,沙发和柱状物的点云法矢量的分布特点鲜明,容易被网络所辨认,所以更容易被准确预测。

表 2 13 种类别的准确率比较

Table 2 Comparison of accuracy among 13 classes

unit: %

Model	Ceiling	Floor	Wall	Beam	Column	Window	Door	Table	Chair	Sofa	Bookcase	Board	Clutter
PointNet	96.1	98.5	89.2	55.8	23.5	65.8	90.6	76.6	66.4	20.1	64.1	65.3	64.1
PPCAN	97.8	97.7	87.3	71.1	65.7	86.3	83.6	79.4	67.4	65.7	70.2	27.4	70.0

表 3 13 种类别的 IoU 比较

Table 3 Comparison of IoU among 13 classes

unit: %

Model	Ceiling	Floor	Wall	Beam	Column	Window	Door	Table	Chair	Sofa	Bookcase	Board	Clutter
PointNet	91.6	96.0	69.1	47.5	22.4	61.2	70.0	64.0	55.9	10.3	48.9	55.3	50.7
PPCAN	93.9	97.0	73.7	62.8	47.4	57.3	69.5	67.5	58.2	30.1	53.2	26.6	55.3

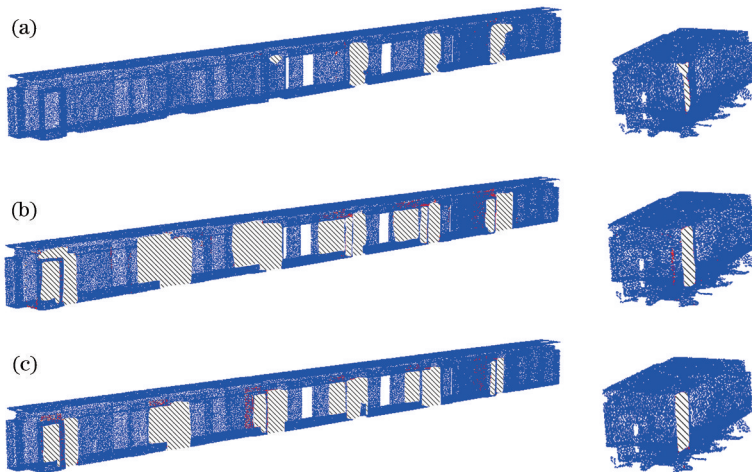


图 4 柱状物分类效果图。(a) PointNet; (b) PPCAN; (c) 真实类别

Fig. 4 Semantic segmentation results of column. (a) PointNet; (b) PPCAN; (c) ground truth

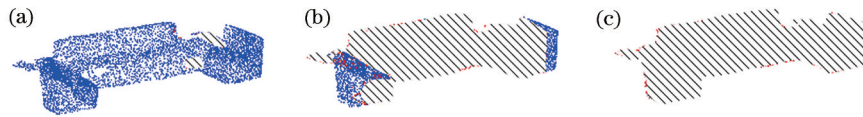


图 5 沙发分类效果图。(a) PointNet; (b) PPCAN; (c) 真实类别

Fig. 5 Semantic segmentation results of sofa. (a) PointNet; (b) PPCAN; (c) ground truth

S3DIS 数据集中的 13 种类别中,在加入法矢量之后,有 10 种类别的 IoU 是升高的(如图 6)。窗户、门和板的 IoU 是降低的,因为这 3 种类别的点云具有相似的空间形态,法矢量的起点近似在同一平面上,指向近

似相同的方向。所以在加入法矢量这一属性之后,这 3 种类别特别是板的分类精度有大幅度的下降。如图 7 所示,很多标签是板的点云被错误地分类成墙壁、窗户等类别。

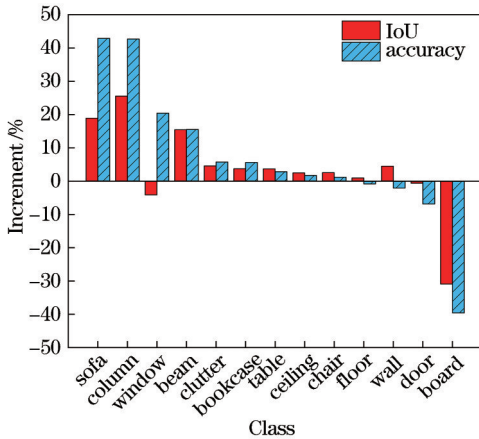


图 6 PPCAN 相较于 PointNet 的 IoU 和分类精度增量

Fig. 6 Increment of IoU and accuracy of PPACAN compared with PointNet

图 8 中展示了 PointNet 和 PPCAN 模型各自在 S3DIS 数据集上的 8 次预测结果,分别统计 13 种类别交并比和准确率。直观地,交并比和准确率越高的语义类别,8 次实验结果在图上的分布更加集中。反之,交并比和准确率较低的语义类别,实验结果在图上的分布更加分散,而且结果的分布与交并比的关系更加紧密。在图 8(a)和图 8(b)中,天花板和地板的准确率和交并比均超过 90%,8 次实验结果在图上是高度重合的,在图 8(b)中天花板和地板的交并比的极差分别是 0.22% 和 0.08%,在图 8(a)中对应的极差分别是 0.24% 和 0.10%;交并比在 40% 至 80% 之间的类别在图 8(a)和图 8(b)中分布是逐渐分散的,交并比的极差大多分布在 0.32% 至 1.60% 的范围内。分类效果较差的[图 8(a)沙发和柱状物;图 8(b)板和沙发]类别的分布是比较明显分散的,交并比的极差均在 1.5% 左右。这表明模型分类的准确性与稳健性是相互关联的。

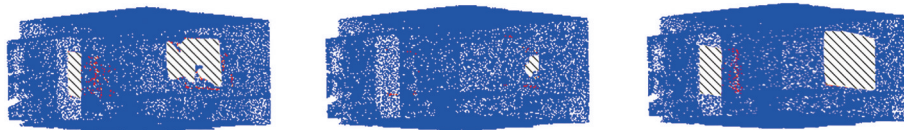


图 7 板分类效果图。(a) PointNet; (b) PPCAN; (c) 真实类别

Fig. 7 Semantic segmentation results of board. (a) PointNet; (b) PPCAN; (c) ground truth

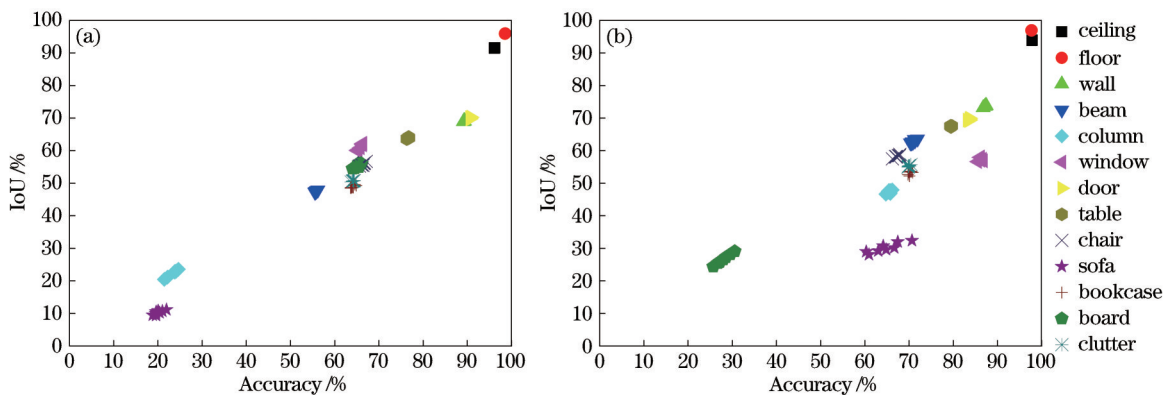


图 8 PointNet 和 PPCAN 两种模型的 8 次预测精度结果。(a) PointNet; (b) PPCAN

Fig. 8 Results of 8 times predication on both PointNet and PPCAN. (a) PointNet; (b) PPCAN

在图 9 中,在统计了 PPCAN 模型的交并比的误差特性,横轴是 PPCAN 模型的 8 组预测结果的每个类别的平均交并比,纵轴表示的是交并比的极差[图 9(a)]

和标准差[图 9(b)],可以看出 PPCAN 模型对某一种类别的预测交并比的稳健性与交并比本身数值的大小是成某种反比例关系的,即当对某一类别的预测准确

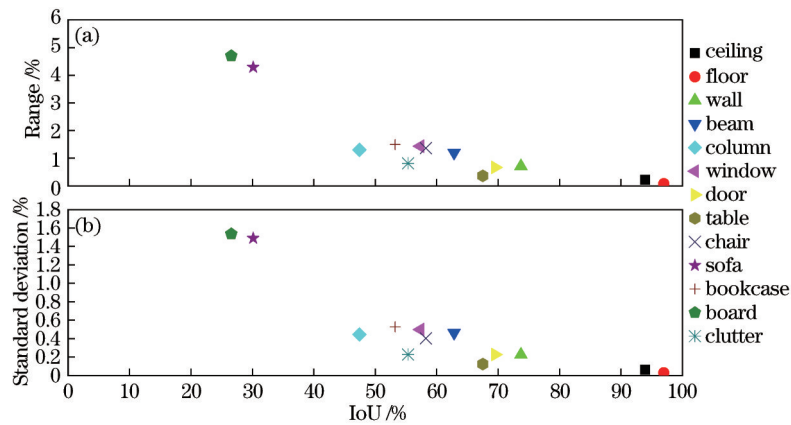


图9 PPCAN模型交并比的误差特性。(a)极差;(b)标准差

Fig. 9 Error characteristics of IoU with PPCAN. (a) Range; (b) standard deviation

率较高的时候,有理由相信这是一个稳健性较高的结果。但同时发现,交并比数值最小的类别,它的交并比标准差也只是1.5%左右,对于数据量庞大的点云语义分割来说也是相对稳定的结果。

4 结 论

针对PointNet语义分割网络没有考虑点云的几何属性的问题,在原始PointNet语义分割网络的基础上增加了法向量估计功能,采用PCA估计点云的法向量,提出了一种考虑法向量的语义分割模型PPCAN。结合实验结果,得出了以下结论:首先特征维度在基于深度学习的点云分类中是一个很重要的因素,更丰富的信息能够有效地提升网络的整体性能,提出的PPCAN模型的OA、MA和mIoU均优于PointNet模型,分别提升了2.3个百分点、7.1个百分点和3.9个百分点;但是属性信息对网络的贡献不是一视同仁的,新增的属性信息会更多地影响对其敏感类别,使该类别变得更加具有辨识度,有利于网络的准确预测。本文中法向量信息能够显著提升沙发和柱状物的语义分割效果,两种类别的交并比分别提升了19.8个百分点和25.0个百分点;同时模型对不同类别的预测的稳健性也是各不相同的,更容易被预测的类别比如天花板、地板等的稳健性更高。

该研究存在的不足是只考虑了法向量信息,没有考虑点云的其他空间属性,比如曲率、点云密度等,同时,未来需要更多的点云空间属性以及更丰富的数据集来验证本文得出的结论。

参 考 文 献

- [1] 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.
- [2] 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, 2021, 43(12): 4338-4364.
- [3] 张爱武,刘路路,张希珍.道路三维点云多特征卷积神经网络语义分割方法[J].中国激光,2020,47(4):0410001.
Zhang A W, Liu L L, Zhang X Z. Multi-feature 3D road point cloud semantic segmentation method based on convolutional neural network[J]. Chinese Journal of Lasers, 2020, 47(4): 0410001.
- [4] 雷相达,王宏涛,赵宗泽.整合迁移学习与全卷积网络的小样本机载激光雷达点云分类[J].中国激光,2021,48(16):1610001.
Lei X D, Wang H T, Zhao Z Z. Small-sample airborne LiDAR point cloud classification based on transfer learning and fully convolutional network[J]. Chinese Journal of Lasers, 2021, 48(16): 1610001.
- [5] 楼晨凤,张湧,刘亚.基于互补梯度增强的红外线列扫描图像小目标检测[J].光学学报,2021,41(21):2104001.
Lou C F, Zhang Y, Liu Y. Small target detection of infrared linear array image based on complemented gradient enhancement[J]. Acta Optica Sinica, 2021, 41(21): 2104001.
- [6] 赵亮,胡杰,刘汉,等.基于语义分割的深度学习激光点云三维目标检测[J].中国激光,2021,48(17):1710004.
Zhao L, Hu J, Liu H, et al. Deep learning based on semantic segmentation for three-dimensional object detection from point clouds[J]. Chinese Journal of Lasers, 2021, 48(17): 1710004.
- [7] 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, December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 945-953.
- [8] Yu T, Meng J J, Yuan J S. Multi-view harmonized bilinear network for 3D object recognition[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: 186-194.
- [9] Yang Z, Wang L W. Learning relationships for multi-

- view 3D object recognition[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 7504-7513.
- [10] Qi C R, Su H, Nießner M, et al. Volumetric and multi-view CNNs for object classification on 3D data[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 5648-5656.
- [11] 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, June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 1912-1920.
- [12] 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.
- [13] Riegler G, Ulusoy A O, Geiger A. OctNet: learning deep 3D representations at high resolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6620-6629.
- [14] Wang P S, Liu Y, Guo Y X, et al. O-CNN: octree-based convolutional neural networks for 3D shape analysis [J]. ACM Transactions on Graphics, 2017, 36(4): 72.
- [15] 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, USA. New York: IEEE Press, 2018: 9204-9214.
- [16] Wang D Z, Posner I. Voting for voting in online point cloud object detection[EB/OL]. [2021-03-06]. <http://www.roboticsproceedings.org/rss11/p35.pdf>.
- [17] 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, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 77-85.
- [18] Qi C R, Li Y, Hao S, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space [C]//2017 In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17), December 4-9, 2017, Long Beach, CA, USA. Red Hook: Curran Associates Inc., 2017: 5105-5114.
- [19] Joseph-Rivlin M, Zvirin A, Kimmel R. Moment⁺et: flavor the moments in learning to classify shapes[C]//2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), October 27-28, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 4085-4094.
- [20] Yang J C, Zhang Q, Ni B B, et al. Modeling point clouds with self-attention and Gumbel subset sampling [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: 3318-3327.
- [21] Zhao H S, Jiang L, Fu C W, et al. PointWeb: enhancing local neighborhood features for point cloud processing [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: 5560-5568.
- [22] Yan X, Zheng C D, Li Z, et al. PointASNL: robust point clouds processing using nonlocal neural networks with adaptive sampling[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 5588-5597.
- [23] 李宝,程志全,党岗,等.三维点云法向量估计综述[J].计算机工程与应用,2010,46(23):1-7.
- Li B, Cheng Z Q, Dang G, et al. Survey on normal estimation for 3D point clouds[J]. Computer Engineering and Applications, 2010, 46(23): 1-7.
- [24] 陈西江,章光,花向红.于法向量夹角信息熵的点云简化算法[J].中国激光,2015,42(8):0814003.
- Chen X J, Zhang G, Hua X H. Point cloud simplification based on the information entropy of normal vector angle [J]. Chinese Journal of Lasers, 2015, 42(8): 0814003.
- [25] Gross M, Pfister H. Point-based graphics[M]//The Morgan Kaufmann series in computer graphics. San Francisco: Morgan Kaufmann Publishers Inc., 2007.
- [26] Pauly M, Keiser R, Kobbelt L P, et al. Shape modeling with point-sampled geometry[J]. ACM Transactions on Graphics, 2003, 22(3): 641-650.
- [27] Mitra N J, Nguyen A N, Guibas L. Estimating surface normals in noisy point cloud data[J]. International Journal of Computational Geometry & Applications, 2004, 14 (4n05): 261-276.
- [28] Yoon M, Lee Y J, Lee S, et al. Surface and normal ensembles for surface reconstruction[J]. Computer-Aided Design, 2007, 39(5): 408-420.
- [29] Li B, Schnabel R, Klein R, et al. Robust normal estimation for point clouds with sharp features[J]. Computers & Graphics, 2010, 34(2): 94-106.
- [30] Hoppe H, DeRose T, Duchamp T, et al. Surface reconstruction from unorganized points[J]. ACM SIGGRAPH Computer Graphics, 1992, 26(2): 71-78.
- [31] 王松,夏绍玮.一种鲁棒主成分分析(PCA)算法[J].系统工程理论与实践,1998,18(1):9.
- Wang S, Xia S W. A robust principal component analysis (PCA) algorithm[J]. Systems Engineering-Theory & Practice, 1998, 18(1): 9.
- [32] Xie Y X, Tian J J, Zhu X X. Linking points with labels in 3D: a review of point cloud semantic segmentation[J]. IEEE Geoscience and Remote Sensing Magazine, 2020, 8(4): 38-59.
- [33] Li Y, Bu R, Sun M, et al. PointCNN: convolution on X-transformed points[C]//2018 Conference and Workshop on Neural Information Processing Systems (NIPS), December 3-8, 2018, Montréal, Canada, USA. [S.l.: s.n.], 2018:

- 823-838.
- [34] Armeni I, Sener O, Zamir A R, et al. 3D semantic parsing of large-scale indoor spaces[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1534-1543.
- [35] 汪文琪, 李宗春, 付永健, 等. 基于改进多规则区域生长的点云多要素分割[J]. 光学学报, 2021, 41(5): 0528001.
Wang W Q, Li Z C, Fu Y J, et al. Multi-factor segmentation of point cloud based on improved multi-rule region growing[J]. Acta Optica Sinica, 2021, 41(5): 0528001.
- [36] 刘金锦, 李浩军. 基于点云切片改进法的不规则物体体积测量[J]. 光学学报, 2021, 41(23): 2312003.
Liu J J, Li H J. Volume measurement of irregular objects based on improved point cloud slicing method[J]. Acta Optica Sinica, 2021, 41(23): 2312003.
- [37] Behley J, Garbade M, Milioto A, et al. Semantic KITTI: a dataset for semantic scene understanding of LiDAR sequences[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 9296-9306.
- [38] Roynard X, Deschaud J E, Goulette F. Paris-Lille-3D: a large and high-quality ground-truth urban point cloud dataset for automatic segmentation and classification[J]. The International Journal of Robotics Research, 2018, 37(6): 545-557.
- [39] Tan W K, Qin N N, Ma L F, et al. Toronto-3D: a large-scale mobile LiDAR dataset for semantic segmentation of urban roadways[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 14-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 797-806.