

引文格式:汤圣君,张韵婕,李晓明,等.超体素随机森林与LSTM神经网络联合优化的室内点云高精度分类方法[J].武汉大学学报(信息科学版),2023,48(4):525-533.DOI:10.13203/j.whugis20220125



Citation: TANG Shengjun, ZHANG Yunjie, LI Xiaoming, et al. A High-Precision Indoor Point Cloud Classification Method Jointly Optimized by Super Voxel Random Forest and LSTM Neural Network[J]. Geomatics and Information Science of Wuhan University, 2023, 48(4):525-533.DOI:10.13203/j.whugis20220125

超体素随机森林与LSTM神经网络联合优化的室内点云高精度分类方法

汤圣君¹ 张韵婕² 李晓明¹ 姚萌萌¹ 叶致煌¹
李亚鑫³ 郭仁忠¹ 王伟玺¹

1 深圳大学建筑与城市规划学院智慧城市研究院&自然资源部城市自然资源监测与仿真重点实验室,广东 深圳,518061

2 香港城市大学建筑与土木工程系,香港,999077

3 香港理工大学深圳研究院,广东 深圳,518061

摘要:针对现有三维点云数据分割分类方法存在分类目标内部不一致的问题,提出一种超体素随机森林与长短期记忆神经网络(long short-term memory,LSTM)联合优化的室内点云高精度分类方法。该方法根据超体素结构具备内部特征一致性的特点,对原始点云进行超体素划分,并以超体素为基本单元进行多元特征计算,搭建室内点云超体素随机森林分类模型,实现点云数据的粗分类。在此基础上,引入LSTM对粗分类的超体素邻域连接关系进行神经网络模型训练与预测,实现超体素粗分类结果的优化。基于开放数据集对所提分类方法进行有效性和精度验证,结果显示,该方法在公开数据集中对13类要素的分类精度可达到83.2%;与经典的深度学习框架相比,该方法在小样本训练时可以达到更优的分类精度。

关键词:超体素;随机森林;LSTM;点云分类;室内建模

中图分类号:P208;P237

文献标识码:A

收稿日期:2022-03-09

DOI:10.13203/j.whugis20220125

文章编号:1671-8860(2023)04-0525-09

A High-Precision Indoor Point Cloud Classification Method Jointly Optimized by Super Voxel Random Forest and LSTM Neural Network

TANG Shengjun¹ ZHANG Yunjie² LI Xiaoming¹ YAO Mengmeng¹ YE Zhihuang¹
LI Yaxin³ GUO Renzhong¹ WANG Weixi¹

1 School of Architecture and Urban Planning ,Shenzhen University & Key Laboratory of Urban Natural Resources Monitoring and Simulation, MNR, Shenzhen 518061, China

2 Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong 999077, China
3 Shenzhen Research Institute , Hong Kong Polytechnic University , Shenzhen 518061, China

Abstract: Objectives: To address the problem of internal inconsistency of classification targets in existing three dimensional(3D) point cloud data segmentation and classification methods. we propose a high-precision classification method for indoor point cloud jointly optimized by super voxel random forest and long short-term memory (LSTM) neural network. Methods: The method takes into account that the super voxel structure has the characteristics of internal feature consistency, divides the original point cloud into super voxels, and uses super voxels as the basic unit for multivariate feature calculation to build a super voxel ran-

基金项目:深圳市科技计划(JCYJ20210324093012033);广东省自然科学基金(2121A1515012574);自然资源部城市国土资源监测与仿真重点实验室开放基金(KF-2021-06-125);2020年佛山市促进高校科技成果服务产业发展扶持项目(2020DZXX04);国家自然科学基金(71901147, 41901329, 41971354, 41971341);深圳大学新入职教师项目(2019056)。

第一作者:汤圣君,博士,副研究员,主要从事城市三维建模、计算机视觉等方面的研究。shengjuntang@szu.edu.cn

通讯作者:王伟玺,博士,副教授。wangwx@szu.edu.cn

dom forest classification model for indoor point cloud to achieve coarse classification of point cloud data. On this basis, LSTM is introduced to train and predict the neural network model for the hyper voxel neighborhood connectivity of coarse classification to achieve the optimization of hyper voxel coarse classification results. The validity and accuracy of the proposed classification method are verified based on the open dataset.

Results: The results show that the classification accuracy of the proposed classification method can reach 83.2% for 13 types of elements in the open dataset. The training data of the LSTM optimization network proposed in this paper used only the label information of region 1 for model training, while other deep learning frameworks used regions 1–5 for model training, so from the perspective of training data requirements, the point cloud data classification framework proposed in this paper can achieve a relatively better prediction result with a small portion of the training data set. The super voxel-based LSTM optimization method approach has high classification accuracy on objects with obvious set features such as ceiling, floor and wall, however, it is inferior to the deep learning algorithm RandLA-Net in classifying objects with complex structures such as chair, sofa and bookcase. **Conclusions:** In this paper, we consider the association characteristics between different types of elements embedded in the connection relations among super voxels, and introduce LSTM to train and predict the model for the coarse classification of super voxel neighborhood connection relations to achieve the optimization of coarse classification results of super voxels. The proposed method can achieve better classification accuracy when trained with small samples compared with the classical deep learning framework.

Key words: super voxel; random forest; LSTM; point cloud classification; indoor modeling

随着室内空间应用越来越多,室内三维数据语义分割成为众多研究学者研究的热点^[1],它是支撑各类智能应用的关键,例如室内导航^[2]、室内机器人^[3]和增强现实^[4]等。点云数据语义信息提取是从杂乱无序的点云中识别与提取要素的过程^[5],核心是利用分割算法将整个场景无组织的点云数据划分得到一系列的点云集合,使得每个集合内的点云包含具有相同语义和感知信息的数据,每个点云集合对应场景内某类型的实体,使得点云具有对象化的语义信息^[6]。目前,国内外研究人员为了提高室内点云数据的分割精度与处理速度做了大量的工作,但依然存在两个重要的挑战^[7-8]:(1)原始点云数据是杂乱、稀疏和非结构化的,且存在数据采集不完整、密度不均匀以及噪声等问题^[9-10],导致点云数据分割算法难以泛化到不同的场景^[11];(2)现阶段点云分割算法主要是根据颜色和几何特征对点云数据进行分类,这类算法依赖大量的训练数据进行模型学习,而室内空间物体结构复杂且多样,当前的算法容易出现适用性低、稳定性差等问题^[11]。

目前室内三维点云语义分割研究主要包含基于多视图的点云分类、基于体素网格的点云分类以及基于原始三维点云的分类算法3种类型。基于多视图的点云分类算法是将三维点云数据根据三维成像原理从不同角度投影为二维影像,进而基于成熟的二维影像分割算法进行场景的

语义分割^[12]。这类算法可以使用成熟的高精度的预训练二维卷积神经网络来初始化多视图模型参数,大幅降低神经网络训练的难度,同时可以避免三维空间中空心物体、非流体几何等三维几何问题的影响。例如 MVCNN (multi-view convolutional neural networks for 3D shape recognition)^[12]是首个多视图三维物体识别的深度神经网络,通过视图信息共享的卷积神经网络对每一个视图都进行二维影像特征提取,然后基于多视图的最大池化层将提取到的多视图特征信息融合为一个全局三维物体特征信息来实现全局三维物体特征分类。GVCNN(group-view convolutional neural networks for 3D shape recognition)^[13]与 Dominant^[14]框架是在 MVCNN 基础上的改进,利用分组的方法对多视图特征进行融合,再进一步利用视图间的相似性提高识别准确率。然而,这类方法需要强大的 GPU(graphics processing unit)进行数据训练,且无法顾及三维空间中的所有特征。因此,部分研究针对三维点云体素表达从三维空间对场景进行特征学习。与点云和多边形面片不同,每个体素在立体网格中都有一个规则排列的索引。该方法将二维卷积神经网络拓展为三维卷积神经网络,可以直接应用于三维体素卷积。RotationNet^[15]方法融合物体识别和视角估测两个目标函数进行语义识别神经网络的搭建,并且增加各视图的信息作为隐含变量

参与神经网络的训练学习。3D-ShapeNets^[16]是第一个采用该思路的神经网络模型,该网络以立体体素网格上的二值变量(体素是否存在于物体)的空间分布来表达一个三维形状。VoxNet^[17]利用浅层三维卷积神经网络处理体素化的三维点云数据。ORION(orientation-boosted voxel-nets)方法^[18]是在 VoxNet 的基础上增加了对物体旋转方位进行估算的子目标,增加该子目标可以提高语义识别的准确率。然而,体素的处理时间和存储占用资源会根据其分辨率大小呈现立方次幂增长,大部分早期研究的方法只能用于低分辨率和浅层的神经网络学习。因此,OctNet 方法^[19]提出对三维立体网格使用不平衡八叉树进行划分,以解决三维立体网格中有效体素的稀疏问题,该方法可用于更高分辨率和更深的神经网络训练。

上述方法在特征计算过程中依然存在特征损失问题,近年来,大量学者研究如何利用原始点云进行特征学习和语义分类。PointNet 算法^[11]是第一个基于三维点云的神经网络模型,该网络首先使用多层感知机(multilayer perceptron, MLP)学习每个点的特征,然后利用对称函数获得全局的物体描述子。PointNet++^[20]在 PointNet 的基础上加入了层次化的特征提取结构。它提出将整个点云分割为数个局部分组的集合抽象层,这个集合抽象层与卷积神经网络中的卷积层作用相似,通过融合数个集合抽象层最后输出特征的感受野。与 PointNet++ 的思路不同,KCNet^[21]则使用图池化层和核相关性来挖掘点云中的局部特征信息。与 KCNet 的目的相似,Kd-Net^[22]首先基于输入点云构建一个 Kd(K-dimensional)树,然后从叶节点到根节点自底向上进行层次化的特征信息提取。然而,由于室内结构复杂程度高,数据本身容易存在数据遮挡,且训练数据集难以获取等问题,目前的室内三维点云语义分割方法训练时间长,难以达到理想的分类精度。

针对上述问题,本文提出了一种超体素随机森林与长短时记忆(long short-term memory,LSTM)神经网络联合优化的室内点云高精度分类方法。该方法基于超体素具备内部特征一致性的特点,对原始点云进行超体素划分,并以超体素为基本单元进行特征计算,搭建室内点云超体素随机森林分类模型,实现点云数据的粗分类。在此基础上,引入 LSTM^[23]对粗分类的超体素邻域连接关

系进行神经网络模型训练与预测,实现超体素粗分类结果的优化。

1 超体素随机森林室内点云粗分类

超体素是由一组三维空间上相似点聚类组成的集合,能够表达三维物体表面的一些特征信息与属性信息,点云聚合成超体素后从无序变为有序,超体素之间具有邻接关系,且超体素本身也使得海量混乱的数据变得更加容易处理。因此,本文提出的超体素随机森林的室内点云粗分类方法的最小分类单元为超体素,这样可以有效提升点云数据分类的效率并提升点云邻域搜索的准确度。超体素随机森林与 LSTM 网络联合优化的室内点云分割方法流程如图 1 所示,在粗分类阶段,原始点云通过超体素聚类得到超体素中心点,同时计算超体素的多维度特征,并将其用于随机森林模型的训练,粗分类的过程主要包含随机化、决策树生成以及投票表决分类等步骤。

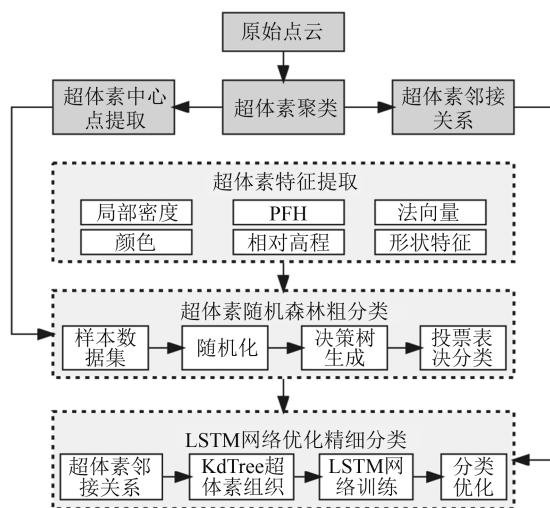


图 1 超体素随机森林与 LSTM 网络联合优化的室内点云分割方法流程

Fig.1 Flowchart of Indoor Point Cloud Segmentation Method Jointly Optimized by Super Voxel Random Forest and LSTM Network

1.1 超体素特征计算

本文中涉及的超体素特征主要包含 6 种类型,分别为局部密度特征、点特征直方图(point-feature histogram, PFH)特征、法向量特征、颜色信息、相对高程特征以及形状特征。本文方法是以超体素作为基本分类单元进行分类的,提取的是每个超体素中心点的特征信息。

局部密度特征为从一个点到最近的 k 个邻域

点的平均距离。因此,对于超级体素中的每个中心点,通过构建 Kd Tree(K-dimensional tree)和 FLANN (fast library for approximate nearest neighbors)算法实现邻域点的快速检索,进而通过计算两对相邻点之间的平均欧氏距离来获得点的局部密度特征。

PFH 特征是通过参数化查询点与邻域点之间的空间差异,形成一个多维直方图对点的 k 邻域几何属性进行描述。具体而言,其基于点与其 k 邻域之间的关系以及法向量描述样本的几何特征。本文每个中心点的 PFH 特征是通过对原始点云建立 Kd Tree,再利用 k 邻近搜索计算得到。

点云中每个点的法向量都代表点所在曲面的方向,可以准确描述平面和曲面信息。本文通过平面拟合的方式计算超体素的法向量信息,并计算其与垂直方向的夹角作为随机森林特征。

颜色特征室内环境下,大部分的分类目标具有颜色一致性,因此 RGB(red, green, blue)颜色在室内点云分割过程中具有重要的作用。本文将超体素作为基本的分类单元,因此每个超体素的颜色信息是由超级体素内各点的 RGB 平均值决定的。

相对高程特征超体素的相对高程特征是通过超体素中心点高度与地面平面高程差获得。考虑到同一楼层中地面的高度不同或者不同楼层的高度不一致问题,本文首先对目标楼层采用平面分割方式获取地面的平面信息,以该信息作为当前位置超体素高程的起始信息,再计算相对高程特征。

形状特征参数首先通过点云局部主成分分析法(principal component analysis,PCA)分解得到 Eigen 特征值,再基于特征值组合计算得到对应的特征信息。传统的 Eigen 特征值是基于 k 邻近搜索获取的局部点云计算得到,为了获取更准确的领域点云,本文以超体素本身作为当前超体素中心点的领域信息,将超体素内部的点云用于 Eigen 特征值的计算,经过特征分解后得到 3 个特征值,分别是 $\lambda_1, \lambda_2, \lambda_3$,其中 3 个特征值从大到小依序排列,即 $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ 。在此基础上,根据形状特征计算方法,分别对超体素的曲率 C_e 、线性度 L_e 、平面度 P_e 、分散度 S_e 以及异向性 A_e 等信息进行计算,计算如下:

$$\left\{ \begin{array}{l} C_e = \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3} \\ L_e = \frac{\lambda_1 - \lambda_2}{\lambda_1} \\ P_e = \frac{\lambda_2 - \lambda_3}{\lambda_1} \\ S_e = \frac{\lambda_1}{\lambda_3} \\ A_e = \frac{\lambda_3 - \lambda_1}{\lambda_3} \end{array} \right. \quad (1)$$

1.2 超体素随机森林模型构建

本文的随机森林模型构建以超体素作为训练和预测的基本单元,与传统的随机森林构建方法相似,超体素随机森林由 N 个决策树 $\{h(X, \theta_n), n=1, 2, \dots, N\}$ 作为初始分类器组成,通过集成学习的方式得到最终的组合分类器。随机森林会统计每一个决策树分类结果,并投票决定输出分类。其中, $\{\theta_n | n=1, 2, \dots, N\}$ 属于随机变量序列,由随机森林中 Bagging 策略和特征子空间策略决定。首先,Bagging 策略是从原始数据集中随机抽样 N 个与原数据集大小相同的训练样本 $\{T_n | n=1, 2, \dots, N\}$ (每次抽取约 63% 的样本),并且为每个训练样本集 T_n 训练一个决策树。然后,特征子空间策略是从数据特征中抽选一个特征子集合,并从中选择最优特征分割节点,从而对决策树中的每个节点进行分裂细化。最后,随机森林是通过集成多个决策树分类器的组合分类器,最终通过分类器投票决定分类结果。其分类的基本流程如下:

1)用 bootstrap 抽样法从原始样本集中随机抽取 K 个训练样本集。

2)对 K 个训练样本集分别构建决策树模型,得到 K 种分类结果。具体而言,每一个决策树会从输入变量的 M 个特征中选取 N 个特征,一般 $N = \sqrt{M}$,再使用信息熵(entropy)与基尼指数(Gini)作为节点分裂标准,计算如下:

$$E_{\text{entropy}}(D) = - \sum_{i=1}^n p_i \log_2 p_i \quad (2)$$

$$G_{\text{Gini}}(D) = 1 - \sum_{i=1}^n p_i^2 \quad (3)$$

式中, n 表示训练数据集 D 包含的类别个数; p_i 表示训练数据属于某类别的概率。

3)根据 K 种分类结果,采用投票表决的方式确定最终分类结果。

2 LSTM 神经网络优化的室内点云精细分类

与原始点云数据不同,点云经过超体素聚类后可以获取超体素与超体素之间的连接关系,其连接关系中蕴含了不同类型要素之间的关联特性,例如桌面的超体素与桌面杂物超体素具有一定的关联性,通过关联性的判断可以避免将杂物错误分割为椅子等物件。因此,本文在超体素随机森林粗分类结果的基础上,提出 LSTM 神经网络优化的超体素关联序列建模方法,将室内三维点云的分类结果优化。基于 LSTM 网络对超体素空间连接关系建模的核心原因是 LSTM 对序列数据建模过程中,其内部结构中的横向神经元会贯穿系列数据,神经元的状态信息可以在整个链上顺序传递,且只存在线性交互,因此在该链状神经元上的信息就可以大致保持不变,从而保留长期信息。这使得其在长短期的信息提取上有显著的优势,因此 LSTM 常运用于长文本分类、时间序列数据预测等具备明显先后顺序且具备关联性的数据中,在这些数据中通常能够获得比以往传统时间序列预测方法更加优秀的结果。

本文以 keras 深度学习框架为基础进行超体素 LSTM 神经网络的搭建,模型搭建方式如图 2 所示,具体步骤如下:

1) 室内场景超体素聚类后,场景被分割为若干块,且超体素与超体素之间具有连接关系。

2) 对场景中划分的每个超体素邻域信息进行迭代搜索,通过 Kd Tree 搜索当前超体素的周围体素,结合其特征信息,按距离大小组合为空间序列集合。

3) 对超体素 LSTM 网络训练,设计了具有 3 层 LSTM 层加一层全连接层的模型,LSTM 层均采用 tanh 作为激活函数,最终进入全连接层,并使用 softmax 作为激活函数实现场景的多目标分类,其中采用随机方式对神经网络参数做初始化,rmsprop 作为优化器。训练过程中,模型训练 batch_size 设置为 128,迭代次数(epoch)为 80 次。考虑到训练数据存在不同类型的超体素不均问题,因此根据训练集中各类别的数量,计算类别权重并加入训练过程。

4) 将随机森林获取的超体素分类结果以及超体素的邻接关系作为输入,并基于超体素 LSTM 网络训练模型对现有的分类结果进行优化,最终得到优化后的超体素模型。

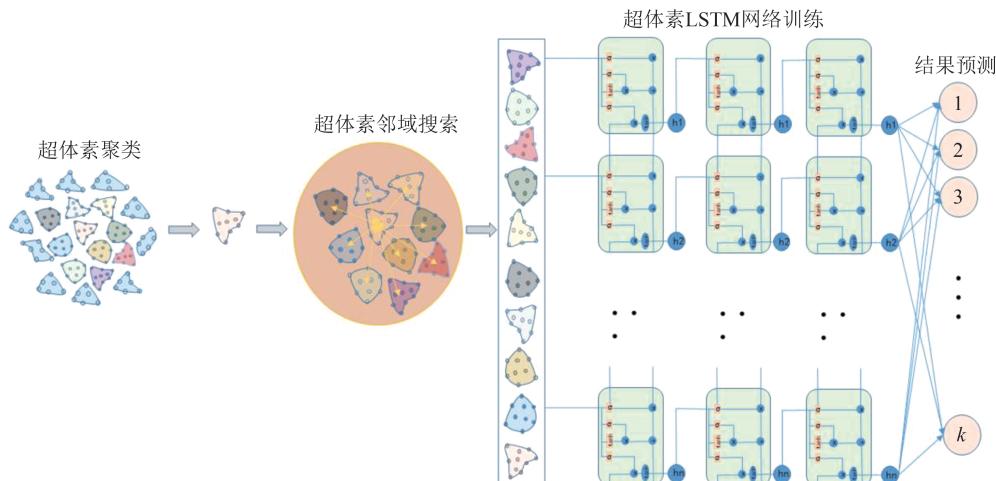


图 2 LSTM 网络优化的室内点云精细分类方法

Fig.2 Classification of Indoor Point Clouds of LSTM Optimization

3 实验分析

3.1 数据预处理

本文实验过程中所用点云数据集为斯坦福大学的公开数据集(S3DIS),如图 3 所示。S3DIS 数据集是斯坦福大学开发的带有像素级语义标注的语义数据集,S3DIS 数据集分为 6 个区域,共

包含 272 个场景,可分为 11 类场景,点云包含 14 个类别的标签信息,如天花板、楼梯、沙发、桌子、板、书柜、门、墙、地板、梁、窗户、椅子、杂物、柱等。本文选取区域 1~5 作为训练数据,区域 6 作为测试区域进行精度评估。在进行语义分割前,对区域的原始点云进行了超体素分割,其参数包含 voxel_resolution,seed_resolution,color_importance、

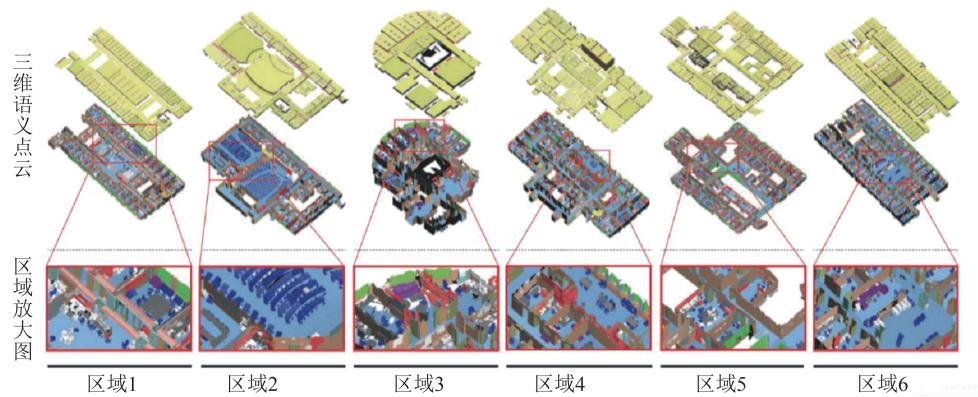


图3 S3DIS数据集

Fig.3 Datasets of S3DIS

spatial_importance 和 normal_importance, 分别取值 0.008、0.175、0.3、0.5 和 1.0。

3.2 实验结果与分析

3.2.1 超体素分割结果分析

合理的超体素参数设置能够避免出现错误分割的情形,对使用以上参数的S3DIS数据集的超体素分割精度进行统计,结果如图4所示。由图4可知,超体素整体分割精度为94.5%左右,对椅子、地板、书柜、沙发的分割精度可达到97%以上。

图5为超体素各类别误分割占比,从图5可以看出,超体素中被错误分类的部分主要与墙、天花板、门、书架等相关性高,这与室内点云的分布状况有密切的联系,对于接近墙体、天花板、桌子部分的点云数据,存在数量稀疏、结构不完整的情况,而超体素分割算法仅仅考虑其向量、颜色、形状信息等特征对点云进行分割难以避免、误分割的情况。在超体素的分割结果的基础上,利用本文方法对改进点云进行语义分类实验。

3.2.2 LSTM神经网络训练

为获取最优的训练结果,对不同 epoch 取值情况下的模型收敛效果进行了统计,结果如图6所示。由图6可知,对于本文提出的LSTM网络其 epoch 参数设置为 80 左右最佳,可以有效避免模型过拟合的问题。

3.2.3 点云分类结果分析

采用平均交并比(mean intersection over union, mIoU)与平均精度(mean accuracy, mAcc)作为点云分类的精度评价指标,mIoU表示数据分类的真实值与预测值的交集和并集之比,mAcc 表示分类的真实值与预测值的交集和真实值之比。假设共有 $k+1$ 个类别(包括背景类),则有:

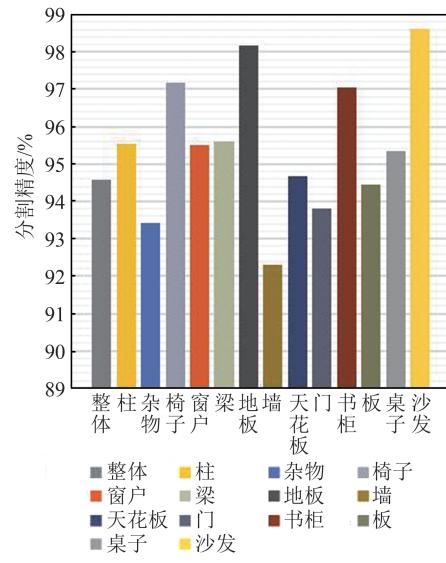


图4 超体素分割精度

Fig.4 Accuracy of Super Voxel

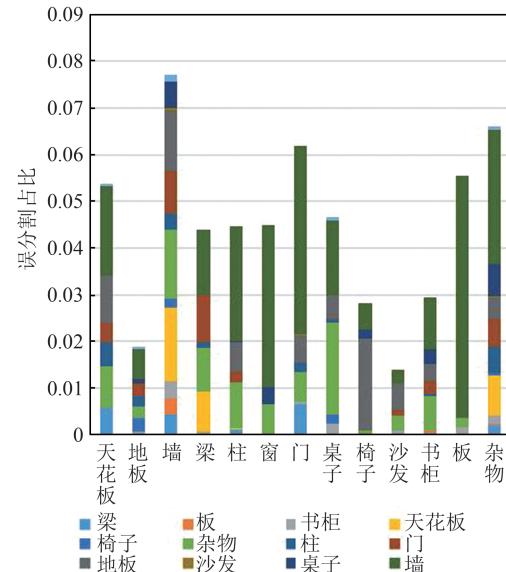


图5 超体素各类别误分割占比

Fig.5 Percentage of Mis-segmentation in Each Category of Super Voxels

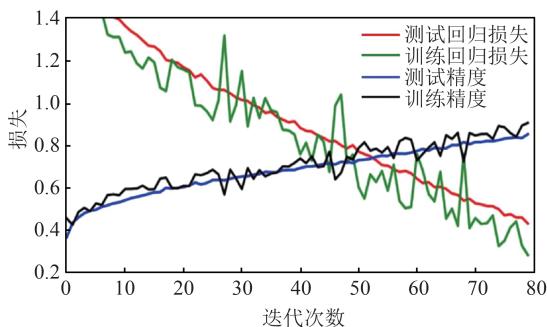


图6 损失与迭代次数的关系

Fig.6 Relations Between Loss and Epoch

$$N_{mIoU} = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}} \quad (4)$$

$$N_{mAcc} = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{i=0}^k p_{ii}} \quad (5)$$

式中, p_{ij} 是将 i 类预测为 j 类的点数; p_{ii} 表示真实值为 i 、预测值为 i 的点数; p_{ji} 表示真实值为 j 、预测值为 i 的点数。

本文选取了5种常用的点云分类框架,包含基于原始点云的随机森林(random forest, RF)分类, PointCNN、PVCNN++、PointNet++ 和 RandLA-Net, 并对其分类结果进行了对比分析, 表1列出了不同分类方法的分类精度情况。

从表1可知, 基于原始点云的随机分类方法精度最低, 其 mIoU 仅达到 8.7%, mAcc 为 24.3%, 在所有分类算法中精度最低。而原始的超体素随机森林粗分类方法可以达到 72.4% 的 mAcc 精度, 这表明超体素的预处理可以有效提升点云数据的分类精度。在粗分类的基础上, LSTM 优化后的 mIoU 可以达到 46.3%, 其 mAcc 为 83.2%, 与深度学习框架 PointCNN、PVCNN++ 和 RandLA-Net 获取的精度相近。

本文提出的 LSTM 优化网络的训练数据仅采用区域 1 的标签信息进行模型训练, 而其他深度学习框架采用区域 1~5 进行了模型训练, 因此从训练数据需求的角度来说, 本文提出的点云数据分类框架可以在小部分训练数据集的情况下达到一个相对较优的预测结果。

表2 中列出了超体素 LSTM 优化方法与 RandLA-Net 方法在不同类型物体分类精度上的对比。

由表2可以看出, 基于超体素 LSTM 优化方法在天花板、地板和墙等具有明显集合特征的物体上都具有较高的分类精度, 然而在椅子、沙发和书柜等复杂结构的物体分类上精度逊色于深度学习算法 RandLA-Net。因此在具体应用中, 超体素 LSTM 优化方法更加适用于对结构信息的提取, 基于深度学习的算法更加适用于对室内家具信息的提取。

3.2.4 时间效率分析

本文采用硬件设备对数据进行训练与预测, 不同方法在数据训练和预测过程中所需时间消耗如表3所示。

由表3可以看出, 由于深度学习在训练过程中涉及的参数相对较多, 因此花费在训练和预测上的时间普遍较高, 大多在 4 h 以上, 其中由于 RandLA-Net 框架在训练过程中采用了随机降采样的方法对点云数据进行一定的降采样, 因此其训练时间和预测时间都有所下降。相比而言, 本文提出的超体速 LSTM 优化方法在训练和预测效率上都有很大的优势, 这主要是由于机器学习方法模型参数相对较少, 在采用较少训练数据的情况下依然可以获得可靠的预测结果, 这将有利于大规模三维点云数据的语义分割和目标识别应用。

表1 不同方法的分类精度对比/%

Tab.1 Comparison of Classification Accuracy of Different Methods/%

精度	超体素随机森林粗分类	超体素 LSTM 优化	基于原始点云 RF 分类	PointCNN	PVCNN++	PointNet++	RandLA-Net
mIoU	39.2	46.3	8.7	65.4	59.0	54.5	70
mAcc	72.4	83.2	24.3	75.6	87.1		82

表2 不同方法在不同类型物体中分类精度对比/%

Tab.2 Comparison of Classification Accuracy of Different Methods in Different Types of Objects/%

方法	天花板	地板	墙	梁	柱	窗户	门	桌子	椅子	沙发	书柜	板	杂物
超体素 LSTM 优化	97	99	80	56	20	38	35	59	20	5	34	26	33
RandLA-Net	93.1	96.1	80.6	62.4	48	64.4	69.4	69.4	76.4	60	64.2	65.9	60.1

表3 不同方法的训练时间和预测时间对比/min
 Tab.3 Comparison of Training Time and Prediction
 Time of Different Methods/min

方法	训练时间	预测时间
PointCNN	288	19
PVCNN++	305	14
PointNet++	269	13
RandLA-Net	241	8
超体素 LSTM 优化	10	2

4 结语

本文提出了一种超体素随机森林与LSTM网络联合优化的室内点云高精度分类方法,该方法充分利用超体素具备内部特征一致性的特点,将原始点云进行超体素划分,并以超体素为基本单元对超体素几何、颜色和形状特征进行计算,搭建了室内点云超体素随机森林分类模型,实现点云数据的粗分类。基于由粗到精的分类思路,本文考虑超体素之间连接关系中蕴含的不同类型要素之间的关联特性,引入LSTM对粗分类的超体素邻域连接关系进行模型训练与预测,实现超体素粗分类结果的优化。并基于开放数据集对本文方法进行有效性和精度验证,结果表明,本文方法在公开数据集中可达到83.2%的分类精度。由于当前的LSTM网络结构更偏向于具有逻辑顺序的序列输入,而由超体素提取空间序列数据则存在两两之间的联系,因此如何设计一个非顺序超体素排列的序列神经网络结构将会是一个重要的研究方向。

参考文献

- [1] Sequeira V, Goncalves J G M, Ribeiro M I. 3D Reconstruction of Indoor Environments [C]//The 3rd IEEE International Conference on Image Processing. Lausanne, Switzerland, 2002.
- [2] Choi J. Development of BIM-Based Evacuation Regulation Checking System for High-Rise and Complex Buildings [J]. *Automation in Construction*, 2014, 46: 38-49.
- [3] Taira H, Okutomi M, Sattler T, et al. InLoc: Indoor Visual Localization with Dense Matching and View Synthesis [C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018.
- [4] Yang Bisheng, Dong Zhen. Progress and Perspective of Point Cloud Intelligence[J]. *Acta Geodaetica et Cartographica Sinica*, 2019, 48(12): 1575-1585. (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.)
- [5] Yang Bisheng, Liang Fuxun, Huang Ronggang. Progress, Challenges and Perspectives of 3D LiDAR Point Cloud Processing [J]. *Acta Geodaetica et Cartographica Sinica*, 2017, 46 (10) : 1509-1516. (杨必胜, 梁福逊, 黄荣刚. 三维激光扫描点云数据处理研究进展、挑战与趋势[J]. 测绘学报, 2017, 46(10): 1509-1516.)
- [6] Hu Q Y, Yang B, Xie L H, et al. RandLA-Net: Efficient Semantic Segmentation of Large-scale Point Clouds[C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, USA, 2020.
- [7] Xiong Hanjiang, Zheng Xianwei, Ding Youli, et al. Semantic Segmentation of Indoor 3D Point Cloud Model Based on 2D-3D Semantic Transfer[J]. *Geomatics and Information Science of Wuhan University*, 2018, 43(12): 2303-2309. (熊汉江, 郑先伟, 丁友丽, 等. 基于2D-3D语义传递的室内三维点云模型语义分割[J]. 武汉大学学报(信息科学版), 2018, 43(12): 2303-2309.)
- [8] Zhang Ruiju, Zhou Xin, Zhao Jianghong, et al. A Semantic Segmentation Algorithm of Ancient Building's Point Cloud Data [J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(5) : 753-759. (张瑞菊, 周欣, 赵江洪, 等. 一种古建筑点云数据的语义分割算法[J]. 武汉大学学报(信息科学版), 2020, 45(5): 753-759.)
- [9] Tran H. Geometric Comparison and Quality Evaluation of 3D Models of Indoor Environments [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2019, 149: 29-39.
- [10] Wang Peng, Liu Rufei, Ma Xinjiang, et al. An Automatic Extraction Method for Pole-like Objects from Vehicle-Borne Laser Point Cloud[J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(7) : 1035-1042. (王鹏, 刘如飞, 马新江, 等. 一种车载激光点云中杆目标自动提取方法[J]. 武汉大学学报(信息科学版), 2020, 45(7): 1035-1042.)
- [11] Charles R Q, Hao S, Mo K C, et al. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation [C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, USA, 2017.
- [12] Su H, Maji S, Kalogerakis E, et al. Multi-view Convolutional Neural Networks for 3D Shape Recognition[C]//IEEE International Conference on Com-

- puter Vision (ICCV), Santiago, Chile, 2016.
- [13] Feng Y F, Zhang Z Z, Zhao X B, et al. GVCNN: Group-view Convolutional Neural Networks for 3D Shape Recognition [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018.
- [14] Wang C, Pelillo M, Siddiqi K. Dominant Set Clustering and Pooling for Multi-view 3D Object Recognition [C]//The British Machine Vision Conference, London, UK, 2017.
- [15] Kanezaki A, Matsushita Y, Nishida Y. Rotation-Net: Joint Object Categorization and Pose Estimation Using Multiviews from Unsupervised Viewpoints [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018.
- [16] Wu Z R, Song S R, Khosla A, et al. 3D ShapeNets: A Deep Representation for Volumetric Shapes[C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, USA, 2015.
- [17] Maturana D, Scherer S. VoxNet: A 3D Convolutional Neural Network for Real-time Object Recognition [C]// IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg, Germany, 2015.
- [18] Sedaghat N, Zolfaghari M, Amiri E, et al. Orientation-Boosted Voxel Nets for 3D Object Recognition [C]//The British Machine Vision Conference, London, UK, 2017.
- [19] Riegler G, Ulusoy A O, Geiger A. OctNet: Learning Deep 3D Representations at High Resolutions[C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, USA, 2017.
- [20] Qi C R, Yi L, Su H, et al. PointNet++ : Deep Hierarchical Feature Learning on Point Sets in a Metric Space [J]. *arXiv*, 2017, DOI: 1706.02413.
- [21] Shen Y R, Feng C, Yang Y Q, et al. Mining Point Cloud Local Structures by Kernel Correlation and Graph Pooling [C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018.
- [22] Klokov R, Lempitsky V. Escape from Cells: Deep Kd-Networks for the Recognition of 3D Point Cloud Models [C]// IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017.
- [23] Hochreiter S, Schmidhuber J. Long Short-Term Memory [J]. *Neural Computation*, 1997, 9 (8) : 1735–1780.