



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所

中国图象图形学学会

北京应用物理与计算数学研究所

中国图象学报 图形学报

2020

06

VOL.25

ISSN1006-8961

CN11-3758/TB



三维点云分割 P1201



第25卷第6期 (总第290期)

2020年6月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿, 均视为同意在本刊网站及CNKI等全文数据库出版, 所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用, 所有内容, 未经许可, 不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院

主办单位 中国科学院遥感与数字地球研究所

中国图象图形学学会

北京应用物理与计算数学研究所

主 编 顾行发

编辑出版 《中国图象图形学报》编辑出版委员会

通信地址 北京市海淀区北四环西路19号

邮 编 100190

电子信箱 jig@radi.ac.cn

电 话 010-58887035

网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号

总 发 行 北京报刊发行局

订 购 全国各地邮局

海外发行 中国国际图书贸易集团有限公司

(邮政信箱: 北京399信箱 邮编: 100048)

印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian

Monthly, Started in 1996

Superintended by Chinese Academy of Sciences

Sponsored by Institute of Remote Sensing and Digital Earth, CAS

China Society of Image and Graphics

Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Gu Xingfa

Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics

Address No. 19, North 4th Ring Road West, Haidian District, Beijing, P. R. China

Zip code 100190

E-mail jig@radi.ac.cn

Telephone 010-58887035

Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals

Domestic All Local Post Offices in China

Overseas China International Book Trading Corporation

(P.O.Box 399, Beijing 100048, P.R.China)

Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB

ISSN 1006-8961

CODEN ZTTXFZ

国外发行代号 M1406

国内邮发代号 82-831

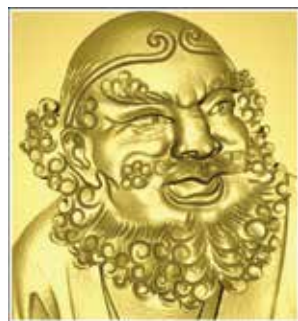
国内定价 60.00元



深度卷积神经网络图像语义分割研究进展(第1069页)



变分调整约束下的反向低秩稀疏学习目标跟踪(第1142页)



二次联合稀疏表示和低秩近似的浅浮雕优化(第1245页)

综述

深度卷积神经网络图像语义分割研究进展

青晨, 禹晶, 肖创柏, 段娟 1069

图像处理和编码

稳健局部特征非下采样小波域数字水印

牛盼盼, 杨思宇, 沈鑫, 杨红颖, 石齐良, 王向阳 1091

图像分析和识别

融合背景块再选取过程的显著性检测

陈星宇, 叶锋, 黄添强, 翁彬, 林文忠 1104

多尺度卷积神经网络显著物体检测

张晴, 左保川, 石艳娇, 戴蒙 1116

结合空间注意力多层特征融合显著性检测

陈凯, 王永雄 1130

变分调整约束下的反向低秩稀疏学习目标跟踪

田丹, 张国山, 张娟瑾 1142

自适应尺度突变目标跟踪

任俊丽, 郭浩, 董亚飞, 刘茹, 安居白, 王妍 1150

自适应多特征融合相关滤波目标跟踪

张艳琳, 钱小燕, 张淼, 葛红娟 1160

新型雷达辐射源识别

高欣宇, 张文博, 姬红兵, 欧阳成 1171

特征点聚类高精度视差图像拼接

谢从华, 张冰, 高蕴梅 1180

动态生成掩膜弱监督语义分割

陈辰, 唐胜, 李锦涛 1190

融合图卷积和差异性池化函数的点云数据分类分割模型

张新良, 付鹏飞, 赵运基, 谢恒, 王琬如 1201

图像理解和计算机视觉

外观表征分析下动态更新相关滤波跟踪

强壮, 石繁槐 1209

结合混合域注意力与空洞卷积的3维目标检测

严娟, 方志军, 高永彬 1221

计算机图形学

字典神经网络方法快速湍流烟雾合成

柏凯, 李伟, 刘晓培 1235

二次联合稀疏表示和低秩近似的浅浮雕优化

张倩倩, 计忠平, 付晓峰 1245

医学图像处理

双层水平集描述眼底图像视杯视盘分割

王莹, 于晓升, 迟剑宁, 雷晓亮, 吴成东 1260

遥感图像处理

结合深度学习和超像元的高分遥感影像变化检测

王艳恒, 高连如, 陈正超, 张兵 1271

CONTENTS

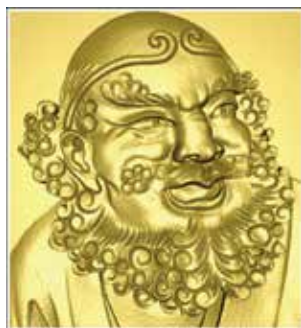
JOURNAL OF IMAGE AND GRAPHICS



Deep convolutional neural network for semantic image segmentation (P1069)



Object tracking via reverse low-rank sparse learning with variation regularization(P1142)



Bas-relief optimization based on twice-joint sparse representation and low-rank approximation (P1245)

Review

- Deep convolutional neural network for semantic image segmentation
Qing Chen, Yu Jing, Xiao Chuangbai, Duan Juan 1069

Image Processing and Coding

- Digital watermarking algorithm in the UDWT domain based on robust local features
Niu Panpan, Yang Siyu, Shen Xin, Yang Hongying, Shi Qiliang, Wang Xiangyang 1091

Image Analysis and Recognition

- Saliency detection based on the background block reselection method
Chen Xingyu, Ye Feng, Huang Tianqiang, Weng Bin, Lin Wenzhong 1104
- A multi-scale convolutional neural network for salient object detection
Zhang Qing, Zuo Baochuan, Shi Yanjiao, Dai Meng 1116
- Saliency detection based on multi-level features and spatial attention
Chen Kai, Wang Yongxiong 1130
- Object tracking via reverse low-rank sparse learning with variation regularization
Tian Dan, Zhang Guoshan, Zhang Juanjin 1142
- Adaptive scale sudden change object tracking
Ren Junli, Guo Hao, Dong Yafei, Liu Ru, An Jubai, Wang Yan 1150
- Correlation filter target tracking algorithm based on adaptive multifeature fusion
Zhang Yanlin, Qian Xiaoyan, Zhang Miao, Ge Hongjuan 1160
- New radar emitter identification method
Gao Xinyu, Zhang Wenbo, Ji Hongbing, Ouyang Cheng 1171
- High-precision parallax image stitching method using feature points clustering
Xie Conghua, Zhang Bing, Gao Yunmei 1180
- Weakly supervised semantic segmentation based on dynamic mask generation
Chen Chen, Tang Sheng, Li Jintao 1190
- Point cloud data classification and segmentation model using graph CNN and different pooling functions
Zhang Xinliang, Fu Pengfei, Zhao Yunji, Xie Heng, Wang Wanru 1201

Image Understanding and Computer Vision

- Dynamic update correlation filter tracking based on appearance representation analysis
Qiang Zhuang, Shi Fanhuai 1209
- 3D object detection based on domain attention and dilated convolution
Yan Juan, Fang Zhijun, Gao Yongbin 1221

Computer Graphics

- Fast turbulent smoke synthesis via a dictionary-based neural network
Bai Kai, Li Wei, Liu Xiaopei 1235
- Bas-relief optimization based on twice-joint sparse representation and low-rank approximation
Zhang Qiankan, Ji Zhongping, Fu Xiaofeng 1245

Medical Image Processing

- Segmentation of optic cup and disc based on two-layer level set describer in retinal fundus images
Wang Ying , Yu Xiaosheng, Chi Jianning, Lei Xiaoliang, Wu Chengdong 1260

Remote Sensing Image Processing

- Deep learning and superpixel-based method for high-resolution remote sensing image change detection
Wang Yanheng, Gao Lianru, Chen Zhengchao, Zhang Bing 1271

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2020)06-1201-08

论文引用格式: Zhang X L, Fu P F, Zhao Y J, Xie H and Wang W R. 2020. Point cloud data classification and segmentation model using graph CNN and different pooling functions. Journal of Image and Graphics, 25(06): 1201-1208 (张新良, 付鹏飞, 赵运基, 谢恒, 王琬如. 2020. 融合图卷积和差异性池化函数的点云数据分类分割模型. 中国图象图形学报, 25(06): 1201-1208) [DOI: 10.11834/jig.190367]

融合图卷积和差异性池化函数的 点云数据分类分割模型

张新良, 付鹏飞, 赵运基, 谢恒, 王琬如

河南理工大学电气工程与自动化学院, 焦作 454000

摘要: 目的 深度网络用于3维点云数据的分类分割任务时,精度与模型在全局和局部特征上的描述能力密切相关。现有的特征提取网络,往往将全局特征和不同尺度下的局部特征相结合,忽略了点与点之间的结构信息和位置关系。为此,通过在分类分割模型中引入图卷积神经网络(graph convolution neural network, GCN)和改进池化层函数,增强局部特征表征能力和获取更丰富的全局特征,改善模型对点云数据的分类分割性能。**方法** GCN模块通过K近邻算法构造图结构,利用相邻点对的边缘卷积获取局部特征,在深度网络模型中动态扩展GCN使模型获得完备的局部特征。在池化层,通过选择差异性的池化函数,联合提取多个全局特征并进行综合,保证模型在数据抖动时的鲁棒性。**结果** 在ModelNet40、ShapeNet和S3DIS(stanford large-scale 3D indoor semantics)数据集上进行分类、部分分割以及语义场景分割实验,验证模型的分类分割性能。与PointNet相比,在ModelNet40分类实验中,整体精度和平均分类精度分别提升4%和3.7%;在ShapeNet部分分割数据集和S3DIS室内场景数据集中,平均交并比(mean intersection-over-union, mIoU)分别高1.4%和9.8%。采用不同的池化函数测试结果表明,本文提出的差异性池化函数与PointNet提出的池化函数相比,平均分类精度提升了0.9%,有效改善了模型性能。**结论** 本文改进的网络模型可以有效获取点云数据中的全局和局部特征,实现更优的分类和分割效果。

关键词: 点云;深度学习;图卷积神经网络(GCN);差异性池化函数;分类分割;联合特征

Point cloud data classification and segmentation model using graph CNN and different pooling functions

Zhang Xinliang, Fu Pengfei, Zhao Yunji, Xie Heng, Wang Wanru

College of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454000, China

Abstract: Objective The depth feature representation of the 3D model is the key and premise of 3D target recognition and 3D model semantic segmentation. It has broad application prospects in the fields of robot, automatic driving, virtual reality, and remote sensing mapping. Semantic segmentation has achieved great progress with the help of deep learning, but most of the methods are used to process 2D images. Given the large amount of data, uneven density, and irregular shape of unstructured 3D point clouds, their classification and segmentation still have enormous challenges. Traditional convolutional neural networks (CNNs) require regularized data as input. The point cloud needs to be converted into multi-view or a voxel

收稿日期: 2019-07-19;修回日期: 2019-10-24;预印本日期: 2019-10-31

基金项目: 国家自然科学基金项目(U1404612);河南省教育厅项目(13B413037, 16A413009, 16A470001);河南理工大学教改项目(2018YJ04)

Supported by: National Natural Science Foundation of China (U1404612)

mesh to process. The existing deep learning network used for directly processing point cloud data solves the disorder problem of point cloud through the pooling layer of CNN. Thus, the network model can directly classify and segment the point cloud data. As for the classification and segmentation model dealing with point cloud data, its accuracy is closely related to the ability of the network to describe global and local features. Existing feature extraction networks often combine global features with local features at different scales, ignoring the structural information and position relationship among points. Thus, the global feature vectors with more significant features cannot be generated in the pooling layer, resulting in low classification and segmentation accuracy. **Method** To improve the performance of the network model, the graph CNN (GCN) and the improved pooling layer function are introduced in the classification and segmentation model. The method can enhance the ability of local feature representation and obtain more abundant global features. The processing ability of the network model to point cloud data can be improved. In the GCN, a graph structure is constructed by connecting the vertex with the nearest K points through the K -nearest neighbor algorithm. The convolution operation is then carried out on the edge and relative position relationship of the adjacent point pairs in the graph structure. Consequently, the more detailed local features implicit in the point cloud data are extracted. The graph structure in the GCN model is not fixed. It is dynamically updated and the graph convolution module can be stacked numerous times in the network to further perceive the local characteristics of point cloud data. In the network pooling layer, a hybrid pooling structure is adopted composed of two parallel pooling channels to obtain the global feature vectors. The maximum pooling channel is used to obtain the maximal feature vector, while another maximum-average pooling channel is used to obtain a synthetic feature concerning the maximal and mean feature vectors. The acquired characteristic vectors are concatenated to obtain the final global feature vector of the network. Consequently, the network provides good robustness for the jittered data. **Result** The datasets ModelNet40, ShapeNet, and Stanford 3D indoor semantics (S3DIS) are mostly used for testing the performance of classification, partial segmentation, and semantic scene segmentation. Several experiments are carried out on the above three datasets to validate the performance of the model. In the classification experiment of ModelNet40, the proposed model achieves a better classification effect compared with the other competitive models. The overall accuracy and average classification accuracy are improved by 4% and 3.7%, respectively, compared with PointNet. In the partial segmentation ShapeNet dataset, the mean intersection-over-union (mIoU) is used as the index for evaluating model segmentation performance. In the comparison test, the proposed model in this study also obtains a satisfactory segmentation result. Specifically, our model's mIoU is 1.4% higher than that of PointNet. In S3DIS indoor scene dataset, our model's mIoU is 9.8% higher than PointNet. Furthermore, different pooling functions are tested and investigated to verify the effectiveness of the proposed hybrid pooling function in this study. Results show that the proposed hybrid pooling function in this study improves the average classification accuracy, exhibiting a 0.9% increase compared with the pooling function by PointNet. **Conclusion** Experimental results show that the local features of point cloud data can be effectively extracted by introducing GCN into the network model. The hybrid pooling function also yields great improvement in generating global characteristics with additional information. In general, the proposed network model can effectively obtain the global and local features of point cloud data and achieve better classification and segmentation effects.

Key words: point cloud; deep learning; graph convolution neural network (GCN); hybrid pooling function; classification and segmentation; joint feature

0 引言

深度网络已经广泛用于2D图像物体的分类分割并取得了丰富的研究成果(Chen等,2018;汪然等,2018),而非结构化的3D点云由于数据量大、密度不均匀以及形状不规则等特点,分类分割仍然面临巨大挑战(Yu等,2017;白静等,2019)。

现有的深度学习网络用于3D点云处理的研究成果分为两类。一类是对点云数据进行预处理,使其具有结构性特征,然后使用成熟的2维神经网络模型对数据进行处理(Qi等,2016;Ren等,2017)。如Su等人(2015)将点云数据投影到20个平面中,使用2D图像卷积神经网络对投影平面训练,实现分类分割任务。Maturana和Scherer(2015)提出一种3D ShapeNets网络,将点云数据体素化,通过卷积

深度置信网络学习3维数据和标签之间的联合概率分布,完成分类分割。另一类是改进网络模型结构,使其具有直接处理点云数据的能力。如 Qi 等人(2017a)构造了一种 PointNet 模型,通过多层感知器(multi-layer perceptron, MLP)提取点云的特征信息,使用最大池化层提取全局特征向量,解决了点云输入的无序性问题,在3D模型的分类以及语义分割上都取得了较好的效果,但是缺乏提取局部特征的能力。随后, Qi 等人(2017b)在 PointNet 的基础上加入一个类似图像神经网络的分层结构,提取不同尺度下的局部特征,提升了分类分割性能。Jiang 等人(2018)提出了一种方向编码单元,可以捕获更多的局部特征,并且能够在 PointNet++ 的结构上多次堆叠,使网络模型具有多尺度感知的能力。但是,上述网络模型大多仅考虑了点云数据中最直接的原始信息,忽略了点间的几何结构以及相对位置关系(Li 等, 2018),没有将深层次的几何约束关系由网络模型学习完成,导致识别细粒度模式的能力以及对复杂场景的泛化能力较差(文威威 等, 2019)。

基于此,本文通过引入图卷积神经网络(graph convolution neural networks, GCN)构造一种图结构,建立点之间的几何结构和相对位置关系,改善模型性能。同时,通过构造差异性池化函数,获取不同的全局特征向量,保证模型的鲁棒性。

1 动态 GCN 模型

受 PointNet 单点卷积操作提取全局特征方法启发,引入 GCN,通过 K 近邻(K nearest neighborhood, KNN)算法构造一种图结构,并对图结构中点对之间的位置关系卷积运算,使得网络可以捕捉点与点之间的相关性特征。

在网络模型中,输入为 $X = \{x_1, \dots, x_i, \dots, x_n\} \subseteq \mathbf{R}^F$, GCN 通过 KNN 算法查找距离顶点 x_i 最近的 k 个点并连接构成拓扑图 G , 如图 1 所示。其中, $G = (\nu, \varepsilon)$, ν 是由 k 个最近邻点构成的点的集合, $\nu = \{x_{i1}, x_{i2}, \dots, x_{ik}\}$, $\varepsilon = \nu \times \nu$ 表示图结构边缘集合。相邻两点的边缘特征可描述(Wang 等, 2018)为

$$e_{ij} = h_{\phi}(x_i, x_j) \quad (1)$$

式中, h_{ϕ} 是具有确定结构的非线性函数,参数 ϕ 表示模型中权重等参数构成的集合,当模型结构确定

时, ϕ 为待训练参数。

对于 GCN 模块,定义输出为对应顶点周围 k 个点的边缘特征的聚合,即

$$x'_i = \sum_{j: (x_i, x_j) \in \varepsilon} h_{\phi}(x_i, x_j) \quad (2)$$

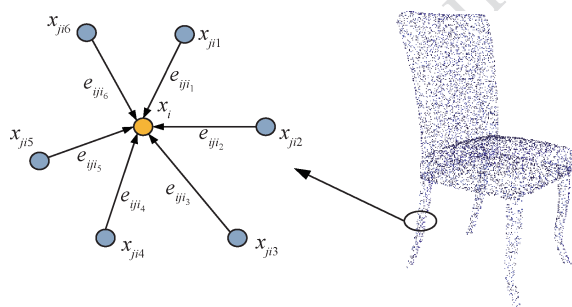


图1 图卷积示意图($k=6$)

Fig. 1 Schematic diagram of GCN($k=6$)

在图卷积过程中,边缘函数的选择对局部特征的提取具有重要作用,对于式(1),如果选择边缘函数为

$$h_{\phi}(x_i, x_j) = h_{\phi}(x_i) \quad (3)$$

则此时只对全局点云进行编码,未考虑点对之间的结构信息,即为 Qi 等人(2017a)提出的 PointNet 网络结构中的点特征提取方法。引入局部结构编码如下

$$h_{\phi}(x_i, x_j) = h_{\phi}(x_i, x_j - x_i) \quad (4)$$

可以看出,改进后的边缘函数同时包含由 x_i 编码的点云全局特征和由 $x_j - x_i$ 编码的局部邻域特征。

通过设置 GCN 模块在网络中多次堆叠实现图结构的动态更新,以获取对图结构的感知能力,捕捉更复杂的特征,即第 l 层输入和拓扑图分别为

$$X^{(l)} = \{x_1^{(l)}, x_2^{(l)}, \dots, x_n^{(l)}\} \subseteq \mathbf{R}^{F_l} \quad (5)$$

$$G^{(l)} = (\nu^{(l)}, \varepsilon^{(l)}) \quad (6)$$

第 $l+1$ 次图卷积操作输出为

$$x_i^{(l+1)} = \sum_{j: (x_i, x_j) \in \varepsilon^{(l)}} h_{\phi}^{(l)}(x_i^{(l)}, x_j^{(l)} - x_i^{(l)}) \quad (7)$$

引入 GCN 模块后构造的点云数据深度网络模型架构如图 2 所示。对原始点云数据使用空间变换网络,保证点云数据的旋转不变性(Jaderberg 等, 2015)。

在分类任务中,网络模型以 c 类 $n \times 3$ 的点云坐标信息作为输入,由 GCN 模块、MLP 和并联的池化层组成特征提取网络。GCN 模块提取局部特征,并联的两个池化层通道生成不同的池化特征,得到全

局特征向量,两个全连接层(512, 256)转换全局特征,输出目标所属类别概率的 c 维向量。

分割网络从包含 p 个语义标签的点云模型中采样 n 个点作为输入,网络结构与分类模型类似,在分

割网络中,将所有 GCN 模块输出的局部特征描述符与池化后的全局特征串联。使用 3 个全连接层(512, 256, 128)转换点态特征,输出对应每个点语义标签的 $n \times p$ 维的特征矩阵。

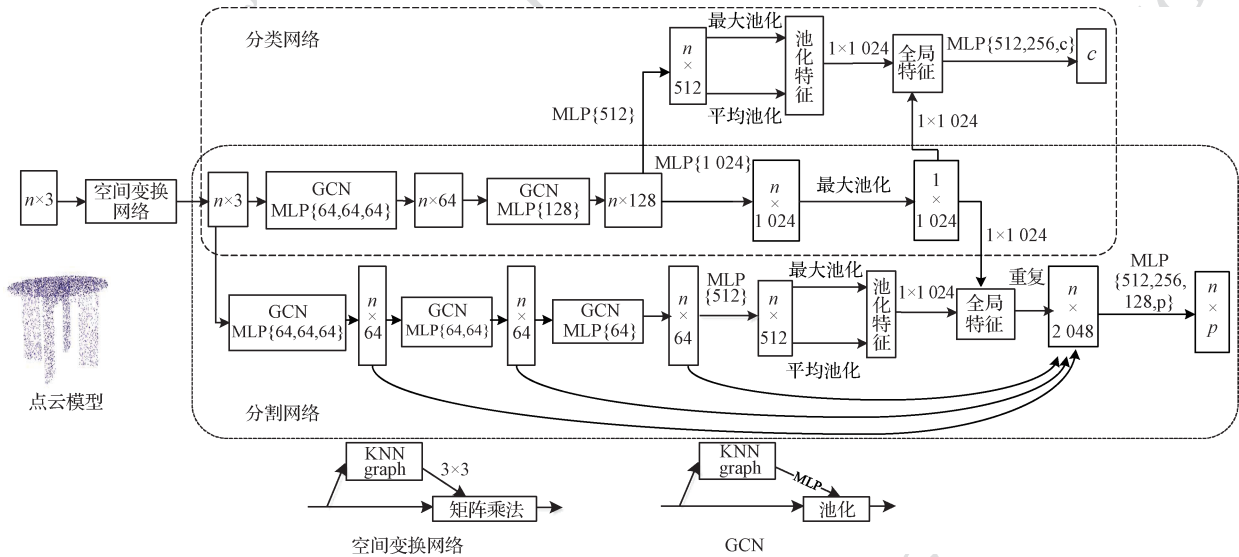


图2 网络整体结构示意图

Fig. 2 Schematic diagram of the overall network structure

2 差异性池化函数

在网络模型中,池化层对提取的点云特征进行池化操作生成全局特征向量。为了解决输出特征点序不变性问题,通常采用具有对称特性的最大池化、均值池化函数。

输入为 $X = \{x_1, \dots, x_2, \dots, x_n\} \subseteq \mathbf{R}^F$ 的点云模型,其中一点 x_j 经网络卷积后映射到中层特征为

$$\Gamma_N = K(x_j) \quad (8)$$

式中, Γ_N 为映射后的特征向量, N 为特征向量长度, K 为特征映射函数。

对所有的输入数据 X ,最大池化的全局特征向量为

$$\mathbf{M}\{\rho_1, \rho_2, \dots, \rho_N\} = \max(K(x_1), K(x_2), \dots, K(x_n)) \quad (9)$$

式中, $(K(x_1), \dots, K(x_j), \dots, K(x_n))$ 为经过深度网络映射后的中层特征矩阵,典型的点云处理网络模型中 $N = 1\,024$, $\max$ 为最大值函数。

均值池化是将某一维数据的均值作为该维度的特征值。其特征向量为

$$\mathbf{A}\{\rho_1, \rho_2, \dots, \rho_N\} =$$

$$\text{avg}(K(x_1), K(x_2), \dots, K(x_n)) \quad (10)$$

式中, avg 为均值函数。

最大池化和均值池化可获得点云模型的不同显著性特征。设计网络模型采用并联双通道池化层提取全局特征。第1通道由最大池化函数提取全局特征;另一通道采用差异性池化函数,其全局特征由最大池化和均值池化两种函数提取,描述为

$$\mathbf{T}\{\rho_1, \rho_2, \dots, \rho_N\} = \boldsymbol{\theta}\left(\frac{N}{2}\right) \oplus \boldsymbol{\psi}\left(\frac{N}{2}\right) \quad (11)$$

式中, $\boldsymbol{\theta}\left(\frac{N}{2}\right) = \mathbf{M}\{\rho_1, \rho_2, \dots, \rho_{N/2}\}$, $\boldsymbol{\psi}\left(\frac{N}{2}\right) = \mathbf{A}\{\rho_1, \rho_2, \dots, \rho_{N/2}\}$, $\mathbf{T}$ 为合并后的高层全局特征, $\oplus$ 为串联操作符。当特征向量长度 $N = 1\,024$ 时,此时 $N/2 = 512$ 。

最终的全局特征向量为

$$\mathbf{T}\{\rho_1, \rho_2, \dots, \rho_{2N}\} = \boldsymbol{\theta}(N) \oplus \mathbf{T}\{\rho_1, \rho_2, \dots, \rho_N\} \quad (12)$$

式中, $\boldsymbol{\theta}(N)$ 为单通道最大池化特征向量, $\boldsymbol{\theta}(N) = \mathbf{M}\{\rho_1, \rho_2, \dots, \rho_N\}$, $\mathbf{T}\{\rho_1, \rho_2, \dots, \rho_{2N}\}$ 为合并后的全局特征。即若模型中单通道特征向量长度 $N = 1\,024$,则全局向量长度为 $2N = 2\,048$ 。

3 实验及结果分析

为了验证提出模型的有效性,在 ModelNet40 (Wu 等,2015)、ShapeNet (Yi 等,2016)、S3DIS (Armeni 等,2016)等广泛使用的数据集上进行测试。实验环境为2个 GTX 1080Ti GPU 显卡,12 GB 显存,Ubuntu16.04 操作系统, tensorflow 1.9 版本。

3.1 3 维目标的分类实验

在 ModelNet40 模型分类数据集上评估本文模型对3维目标分类的性能。ModelNet40 数据集包含40个类别的12 311个CAD(computer aided design)模型,其中,9 843个模型用于训练,2 468个模型用于测试。在模型表面均匀采样1 024个点,并归一化到单位球面。为保证实验条件一致性,采用与 PointNet 相同的数据增强策略,即随机将点云沿 z 轴旋转,并使用均值为零、标准差为0.02的高斯噪声对每个点进行抖动。

3维点云数据分类网络架构见图2,使用Adam 优化器对网络进行优化, $K=20$, $\text{dropout}=0.5$,实验结果见表1。可以看出,与 PointNet 相比,模型平均分类精度和整体精度分别提高3.7%和4%;与分类效果较好的 MGCNN 网络相比,平均分类精度和整体精度分别提高0.9%和1%。实验结果表明,本文的网络模型在平均分类精度和整体精度上均实现了最优的分类效果。

3.2 点云部分分割实验

在部分分割实验中,分割任务是对3维模型中

的每个点分配部件类别标签(如椅子腿,杯柄等)。在 ShapeNet 数据集上进行实验验证,选取16类共16 881个3维形状的物体作为分割对象。3维对象中每个点对应1个部件注释标签,共50个部件,每个对象由少于6个部件组成。同时,分割网络从每个CAD模型中采样2 048个点作为输入。考虑输入采样点密度增加,分割网络中选择 $K=30$ 。选用平均交并比(mean intersection-over-union, mIoU)作为模型优劣的评价指标,典型网络的 mIoU 和每个类别 IoU 的测试结果如表2所示。可以看出,本文模型 mIoU 与 PointNet++ 相当,均为最优的85.1%,比 PointNet、Kd-Net (Klokov 和 Lempitsky, 2017)以及 DGCNN (Wang 等,2018)分别提高1.4%、2.8%和0.3%。充分验证本文网络模型对于细粒度模式和复杂场景具有很好的分割性能。

表1 ModelNet40 数据集分类性能比较

Table 1 Comparison of classification performance on ModelNet40 dataset

方法	平均分类精度	整体精度
3DShapeNets	77.3	84.7
VoxNet	83.0	85.9
Subvolume	86.0	89.2
PointNet	86.0	88.1
PointNet++	-	90.7
DGCNN	88.8	91.1
本文	89.7	92.1

注:加粗字体表示最优结果,“-”表示未知结果。

表2 ShapeNet 部分数据集的分割结果(IoU)

Table 2 Segmentation results on the ShapeNet partial data set(IoU)

方法	mIoU	飞机	书包	帽子	汽车	椅子	耳机	吉他	刀	灯	手提电脑	摩托	大杯子	手枪	火箭	滑板	桌子
PointNet	83.7	83.4	78.7	82.5	74.9	89.6	73.0	91.5	85.9	80.8	95.3	65.2	93.0	81.2	57.9	72.8	80.6
PointNet++	85.1	82.4	79.0	87.7	77.3	90.8	71.8	91.0	85.9	83.7	95.3	71.6	94.1	81.3	58.7	76.4	82.6
Kd-Net	82.3	80.1	74.6	74.3	70.3	88.6	73.5	90.2	87.2	81.0	94.9	57.4	86.7	78.1	51.8	69.9	80.3
DGCNN	84.8	83.9	82.4	85.7	78.6	90.9	71.1	90.6	87.2	82.8	96.1	64.8	94.1	81.8	63.3	76.2	81.7
本文	85.1	84.7	84.5	83.0	77.9	91.1	69.8	91.3	87.2	83.3	95.9	67.7	94.7	80.2	63.1	74.9	81.7

注:加粗字体为各列最优结果。

图3为可视化的部分分割结果。可以看出,本文模型对刀身和刀柄,火箭主体和喷射器,耳机听筒和耳机线有着较好的分割效果,由误分割图可见,错

误分割大多存在于目标粘连部分。

3.3 场景语义分割实验

室内场景分割和部分分割使用同一种网络模

型。室内场景点云来自实际环境,每个场景中包含多类目标以及丰富的噪声,因此相比于点云部分分割,室内场景的点云分割更能体现网络模型对复杂场景的泛化能力,更不失一般性。在斯坦福大学大规模 3D 室内场景数据集下进行语义场景分割评估,该数据集包含 6 个区域的 271 个房间的 3D 扫描数据。扫描中的每个点都对应椅子、桌子、地板、墙壁等 13 个类别语义标签。实验过程中,遵从与 PointNet 相同的设置,将所有点按照房间编号分开,且将每个房间划分为 $1\text{ m} \times 1\text{ m}$ 的小区域,每个点用 9D 向量(3 维坐标、法向量以及归一化后的 3 维坐标)表示。在每个小区域,随机采样 4 096 个点作为训练输入。本文模型分割结果使用 mIoU 和整体精度作为评价指标,与 PointNet 和 PointNet++ 的分割结果比较见表 3。

在平均交并比 mIoU 上,本文方法比 PointNet 高 9.8%,在整体精度上,本文方法比 PointNet 和 PointNet++ 分别高 5.6% 和 0.8%。实验结果充分表明本文 GCN 模块能很好地提取局部特征,部分室内场景分割可视化结果如图 4,依次为走廊、开阔空间和前厅。

表 3 场景语义分割结果比较

Table 3 Comparison of scene semantic segmentation results

方法	平均交并比	整体精度 /%
PointNet	47.6	78.5
PointNet++	—	83.3
本文	57.4	84.1

注:加粗字体表示最优结果,“—”表示未知结果。

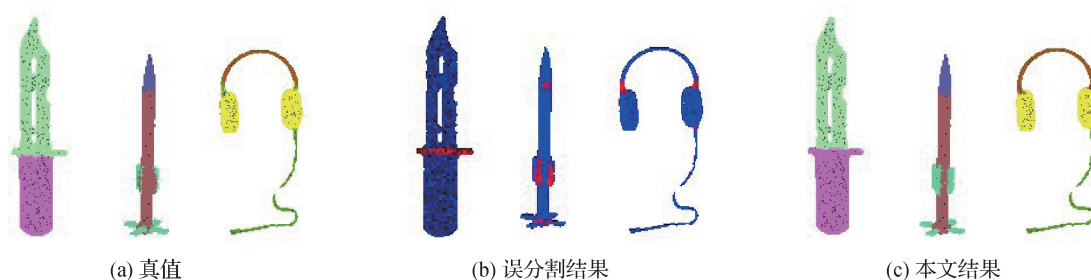


图 3 部分分割结果

Fig. 3 Partial segmentation results ((a) ground truth; (b) incorrect; (c) ours)



图 4 场景语义分割结果

Fig. 4 Results of scene semantic segmentation ((a) ground truth; (b) ours)

3.4 差异性池化对比实验及鲁棒性测试

在 3 维数据的分类任务中,不同的池化函数会生成不同的全局特征,从而影响识别精度。为了研究不同池化函数对深度网络的分类性能的影响,使用不同池化函数进行对比测试,结果如表 4 所示。

在实验中单独使用均值池化和最大值池化得到 86.2% 和 88.8% 的平均分类精度,而使用差异性池化实现了 89.4% 的平均分类精度,充分验证不同信

表 4 不同池化函数组合对比

Table 4 Comparison of different pooling functions

池化函数	平均精度	整体精度 /%
avg	86.2	89.8
max	88.8	91.1
max + avg	89.4	91.7
max + (max + avg)	89.7	92.1

注:加粗字体表示各列最优结果。

息的全局特征串联可以提升网络模型的分类精度。在本文模型中,通过最大池化与差异性池化并联,实现了最优的平均分类精度,为 89.7%。验证了在最大池化加均值池化基础上加入更高维度的最大池化,可以得到包含更多特征信息的全局特征向量。

为了验证网络对于输入点云稀疏性表达的鲁棒性,在对 ModelNet40 数据集原始采样 1 024 个点的基础上,随机丢弃 25%、50%、62.5%、75% 以及 87.5% 的采样点,测试网络模型的平均分类精度,并与 PointNet 模型进行对比。实验中设置测试样本与训练样本具有相同采样点数,每个采样点数进行 3 次训练与测试并求取平均值,测试结果见图 5。实验结果表明,与 PointNet 相比,本文模型在不同采样点数下均实现了更高的平均分类精度,当随机丢弃 25% 的采样点时,本文模型平均分类精度为 89.3%,比原始采样点数分类精度低 0.4%。当随机丢弃 87.5% 的采样点时,模型依然得到 86.0% 的平均分类精度,比原始采样点数低 3.7%,表明模型对点云的稀疏性有较好的鲁棒性。

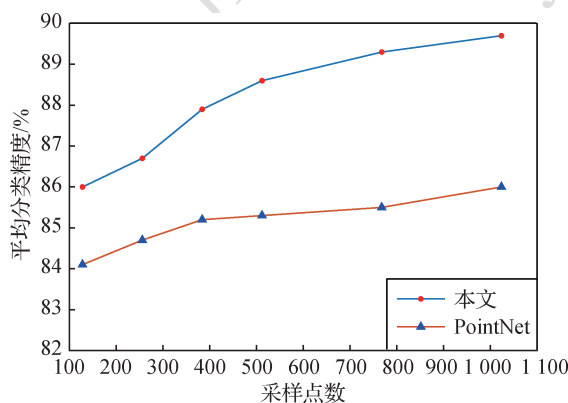


图5 网络鲁棒性测试

Fig. 5 Robustness test for network

4 结 论

针对点云数据网络模型的分类分割性能与局部特征和全局特征的提取能力密切相关,本文提出了一种由局部特征提取 GCN 模块和全局特征差异性池化函数整合的网络模型,使得网络具有更强的鲁棒性和分类分割性能。动态图卷积网络 GCN 工作机制的引入使得网络具有对图结构的感知能力,捕捉更复杂的特征。在分类、部分分割和场景语义分割数据集上的分类分割实验表明,本文网络模型可

以更好地提取局部特征和全局特征,实现了较好的分类分割效果。在扩展网络模型结构处理 3D 点云数据方面具有重要的参考价值。

未来的主要工作将通过图结构中点对之间的方向信息进行编码,进一步提取点云数据中局部特征,增强网络模型对点云的处理能力,并且通过改进网络模型使网络适用于现实点云场景的语义分割。

参考文献 (References)

- Armeni I, Sener O, Zamir A R, Jiang H, Brilakis I, Fischer M and Savarese S. 2016. 3D semantic parsing of large-scale indoor spaces// Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE: 1534-1543 [DOI: 10.1109/CVPR.2016.170]
- Bai J, Si Q L and Qin F W. 2019. Lightweight real-time point cloud classification network LightPointNet. Journal of Computer-aided Design and Graphics, 31(4): 612-621 (白静, 司庆龙, 秦飞巍. 2019. 轻量级实时点云分类网络 LightPointNet. 计算机辅助设计与图形学学报, 31(4): 612-621) [DOI: 10.3724/SP.J.1089.2019.17328]
- Chen L C, Papandreou G, Kokkinos I, Murphy K and Yuille A L. 2018. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4): 834-848 [DOI: 10.1109/TPAMI.2017.2699184]
- Jaderberg M, Simonyan K, Zisserman A and Kavukcuoglu K. 2015. Spatial transformer networks//Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, Canada: MIT Press: 2017-2025
- Jiang M Y, Wu Y R, Zhao T Q, Zhao Z L and Lu C W. 2018. PointSIFT: a SIFT-like network module for 3D point cloud semantic segmentation [EB/OL]. <https://arxiv.org/pdf/1807.00652.pdf>
- Klokov R and Lempitsky V. 2017. Escape from cells: deep Kd-networks for the recognition of 3D point cloud models//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: 863-872 [DOI: 10.1109/ICCV.2017.99]
- Li Y Y, Bu R, Sun M C, Wu W, Di X H and Chen B Q. 2018. PointCNN: convolution On X-transformed points [EB/OL]. <https://arxiv.org/pdf/1801.07791.pdf>
- Maturana D and Scherer S. 2015. VoxNet: a 3D convolutional neural network for real-time object recognition//Proceedings of 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems. Hamburg, Germany: IEEE: 922-928 [DOI: 10.1109/IROS.2015.7353481]
- Qi C R, Su H, Kaichun M and Guibas L J. 2017a. PointNet: deep learning on point sets for 3D classification and segmentation//Pro-

- ceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA; IEEE; 77-85 [DOI: 10.1109/CVPR.2017.16]
- Qi C R, Su H, Nießner M, Dai A, Yan M Y and Guibas L J. 2016. Volumetric and multi-view CNNs for object classification on 3D data//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA; IEEE; 5648-5656 [DOI: 10.1109/CVPR.2016.609]
- Qi C R, Yi L, Su H and Guibas L J. 2017b. PointNet++: deep hierarchical feature learning on point sets in a metric space [EB/OL]. <https://arxiv.org/pdf/1706.02413.pdf>
- Ren M W, Niu L and Fang Y. 2017. 3D-A-Nets: 3D deep dense descriptor for volumetric shapes with adversarial networks [EB/OL]. <https://arxiv.org/pdf/1711.10108.pdf>
- Su H, Maji S, Kalogerakis E and Learned-Miller E. 2015. Multi-view convolutional neural networks for 3D shape recognition//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile; IEEE; 945-953 [DOI: 10.1109/ICCV.2015.114]
- Wang R, Xue X Y, Ping X J, Niu S Z and Zhang T. 2018. Steganalysis of JPEG images based on image classification and segmentation. *Journal of Image and Graphics*, 23(10): 1472-1482 (汪然, 薛小燕, 平西建, 牛少彰, 张涛. 2018. 分类与分割相结合的 JPEG 图像隐写分析. *中国图象图形学报*, 23(10): 1472-1482) [DOI: 10.11834/jig.180037]
- Wang Y, Sun Y B, Liu Z W, Sarma S E, Bronstein M M and Solomon J M. 2018. Dynamic graph CNN for learning on point clouds [EB/OL]. <https://arxiv.org/pdf/1801.07829.pdf>
- Wen W W, Wen G J, Hui B W and Chen D X. 2019. Model library construction by combining global and local surfaces for 3D object recognition. *Journal of Image and Graphics*, 24(2): 248-257 (文威威, 文贡坚, 回丙伟, 陈鼎新. 2019. 结合全局与局部信息的点云目标识别模型库构建. *中国图象图形学报*, 24(2): 248-257) [DOI: 10.11834/jig.180270]
- Wu Z R, Song S R, Khosla A, Yu F, Zhang L G, Tang X O and Xiao J X. 2015. 3D ShapeNets: a deep representation for volumetric shapes//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA; IEEE; 1912-1920 [DOI: 10.1109/CVPR.2015.7298801]
- Yi L, Kim V G, Ceylan D, Shen I C, Yan M Y, Su H, Lu C W, Huang Q X, Sheffer A and Guibas L. 2016. A scalable active framework for region annotation in 3D shape collections. *ACM Transactions on Graphics*, 35(6): 210 [DOI: 10.1145/2980179.2980238]
- Yu B, Yin H T and Zhu Z X. 2017. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting//Proceedings of the 27th International Joint Conference on Artificial Intelligence. [s. l.]; IJCAI; 3634-3640 [DOI: 10.24963/ijcai.2018/505]

作者简介



张新良, 1978 年生, 男, 副教授, 硕士生导师, 主要研究方向为智能控制、检测技术与自动化装置。
E-mail: zxldq@hpu.edu.cn

付鹏飞, 男, 硕士研究生, 主要研究方向为点云处理。

E-mail: 936509288@qq.com

赵运基, 男, 博士, 副教授, 硕士生导师, 主要研究方向为模式识别与智能控制。E-mail: auyjz@hpu.edu.cn

谢恒, 男, 硕士研究生, 主要研究方向为模式识别与人工智能。E-mail: 708998966@qq.com

王琬如, 女, 硕士研究生, 主要研究方向为模式识别与人工智能。E-mail: 870925329@qq.com