



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2021

04

VOL.26

ISSN1006-8961
CN11-3758/TB



细粒度图像分类 P847



第26卷第4期 (总第300期)

2021年4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 中国科学院空天信息创新研究院

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 吴一戎

编辑出版 《中国图象图形学报》编辑部

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 Aerospace Information Research Institute, CAS

China Society of Image and Graphics

Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong

Editor, Publisher Editorial Office 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

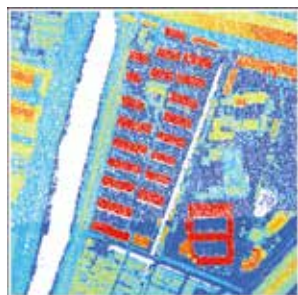
国内定价 60.00元



YOLOv3剪枝模型的多人姿态估计(第0837页)



融入时序和速度信息的自适应更新目标跟踪(第0883页)



可变半径Alpha Shapes提取机载LiDAR点云建筑物轮廓(第0910页)

综述

从图像到语言: 图像标题生成与描述

谭云兰, 汤鹏杰, 张丽, 罗玉盘 0727

CT影像肺结节分割研究进展

董婷, 魏珑, 聂生东 0751

图像处理和编码

全局注意力门控残差记忆网络的图像超分重建

王静, 宋慧慧, 张开华, 刘青山 0766

多层次感知残差卷积网络的单幅图像超分重建

何蕾, 程佳豪, 占志钰, 杨雯博, 刘沛然 0776

采用双尺度图像分解的水下彩色图像增强

李健, 张显斗, 李熹飞, 吴子朝 0787

图像分析和识别

姿态特征结合2维傅里叶变换的步态识别

王新年, 胡丹丹, 张涛, 白桂欣 0796

复杂背景下的手势识别

王银, 陈云龙, 孙前来 0815

结合形变模型与图像修复的人脸姿态矫正

吴从中, 郑荣生, 臧怀娟, 刘明威, 徐甲甲, 詹曙 0828

YOLOv3剪枝模型的多人姿态估计

蔡哲栋, 应娜, 郭春生, 郭锐, 杨鹏 0837

YOLOv3和双线性特征融合的细粒度图像分类

闫子旭, 侯志强, 熊磊, 刘晓义, 余旺盛, 马素刚 0847

深度多模态融合服装风格检索

苏卓, 柯司博, 王若梅, 周凡 0857

超像素条件随机场下的RGB-D视频显著性检测

李贝, 杨铀, 刘琼 0872

融入时序和速度信息的自适应更新目标跟踪

尹宽, 李均利, 胡凯, 李丽 0883

融合逆密度函数与关系形状卷积神经网络的点云分析

邱云飞, 朱梦影 0898

图像理解和计算机视觉

可变半径Alpha Shapes提取机载LiDAR点云建筑物轮廓

伍阳, 王丽妍, 胡春霞, 程亮 0910

3D遮挡模型引导的光场图像深度获取

吴迪, 张旭东, 张骏, 范之国, 孙锐 0924

计算机图形学

基于全正基的非均匀三次加权B样条

姬佩佩, 张贵仓, 汪凯, 孟建军 0939

遥感图像处理

LBP特征分类的极化SAR图像机场跑道检测

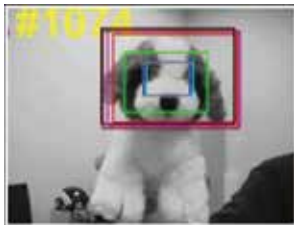
韩萍, 万义爽, 刘亚芳, 韩宾宾 0952

CONTENTS

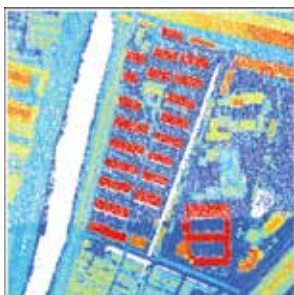
JOURNAL OF IMAGE AND GRAPHICS



Research on multiperson pose estimation combined with YOLOv3 pruning model(P0837)



Adaptive update object tracking algorithm incorporating timing and speed information(P0883)



Extraction of building contours from airborne LiDAR point cloud using variable radius Alpha Shapes method(P0910)

Review

From image to language: image captioning and description

Tan Yunlan, Tang Pengjie, Zhang Li, Luo Yupan 0727

Research progress of lung nodule segmentation based on CT images

Dong Ting, Wei Long, Nie Shengdong 0751

Image Processing and Coding

Learning global attention-gated multi-scale memory residual networks for single-image super-resolution

Wang Jing, Song Huihui, Zhang Kaihua, Liu Qingshan 0766

Single image super-resolution reconstruction based on multi-level perceptual residual convolutional network

He Lei, Cheng Jiahao, Zhan Zhiyu, Yang Wenbo, Liu Peiran 0776

Underwater color image enhancement based on two-scale image decomposition

Li Jian, Zhang Xiandou, Li Shangfei, Wu Zizhao 0787

Image Analysis and Recognition

Gait recognition using pose features and 2D Fourier transform

Wang Xinnian, Hu Dandan, Zhang Tao, Bai Guixin 0796

Hand gesture recognition in complex background

Wang Yin, Chen Yunlong, Sun Qianlai 0815

Face pose correction based on morphable model and image inpainting

Wu Congzhong, Zheng Rongsheng, Zang Huaijuan, Liu Mingwei, Xu Jiajia, Zhan Shu 0828

Research on multiperson pose estimation combined with YOLOv3 pruning model

Cai Zhedong, Ying Na, Guo Chunsheng, Guo Rui, Yang Peng 0837

Fine-grained classification based on bilinear feature fusion and YOLOv3

Yan Zixu, Hou Zhiqiang, Xiong Lei, Liu Xiaoyi, Yu Wangsheng, Ma Sugang 0847

Fashion style retrieval based on deep multimodal fusion

Su Zhuo, Ke Sibao, Wang Ruomei, Zhou Fan 0857

RGB-D video saliency detection via superpixel-level conditional random field

Li Bei, Yang You, Liu Qiong 0872

Adaptive update object tracking algorithm incorporating timing and speed information

Yin Kuan, Li Junli, Hu Kai, Li Li 0883

Point cloud analysis model combining inverse density function and relation-shape convolution neural network

Qiu Yunfei, Zhu Mengying 0898

Image Understanding and Computer Vision

Extraction of building contours from airborne LiDAR point cloud using variable radius Alpha Shapes method

Wu Yang, Wang Liyan, Hu Chunxia, Cheng Liang 0910

Light field depth estimation guided by 3D occlusion model

Wu Di, Zhang Xudong, Zhang Jun, Fan Zhiguo, Sun Rui 0924

Computer Graphics

Non uniform weighted cubic B-spline based on all positive basis

Ji Peipei, Zhang Guicang, Wang Kai, Meng Jianjun 0939

Remote Sensing Image Processing

Airport runway detection based on LBP feature classification in PolSAR images

Han Ping, Wan Yishuang, Liu Yafang, Han Binbin 0952

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2021)04-0898-12

论文引用格式: Qiu Y F and Zhu M Y. 2021. Point cloud analysis model combining inverse density function and relation-shape convolution neural network. Journal of Image and Graphics, 26(04): 0898-0909 (邱云飞, 朱梦影. 2021. 融合逆密度函数与关系形状卷积神经网络的点云分析. 中国图象图形学报, 26(04): 0898-0909) [DOI: 10.11834/jig.200159]

融合逆密度函数与关系形状 卷积神经网络的点云分析

邱云飞¹, 朱梦影^{1,2*}

1. 辽宁工程技术大学软件学院, 葫芦岛 125100; 2. 中国科学院海西研究院泉州装备制造研究所, 泉州 362216

摘要: 目的 3D点云与以规则的密集网格表示的图像不同, 不仅不规则且无序, 而且由于输入输出大小和顺序差异, 具有密度不均匀以及形状和缩放比例存在差异的特性。为此, 提出一种对3D点云进行卷积的方法, 将关系形状卷积神经网络(relation-shape convolution neural network, RSCNN)与逆密度函数相结合, 并在卷积网络中增添反卷积层, 实现了点云更精确的分类分割效果。**方法** 在关系形状卷积神经网络中, 将卷积核视为由权重函数和逆密度函数组成的3D点局部坐标的非线性函数。对给定的点, 权重函数通过多层感知器网络学习, 逆密度函数通过核密度估计(kernel density estimation, KDE)学习, 逆密度函数的引入对点云采样率不均匀的情况进行弥补。在点云分割任务中, 引入由插值和关系形状卷积层两部分组成的反卷积层, 将特征从子采样点云传播回原始分辨率。**结果** 在ModelNet40、ShapeNet、ScanNet数据集上进行分类、部分分割和语义场景分割实验, 验证模型的分类分割性能。在分类实验中, 与PointNet++相比, 整体精度提升3.1%, 在PointNet++将法线也作为输入的情况下, 精度依然提升了1.9%; 在部分分割实验中, 类平均交并比(mean intersection over union, mIoU)比PointNet++在法线作为输入情况下高6.0%, 实例mIoU比PointNet++高1.4%; 在语义场景分割实验中, mIoU比PointNet++高13.7%。在ScanNet数据集上进行不同步长有无逆密度函数的对比实验, 实验证明逆密度函数将分割精度提升0.8%左右, 有效提升了模型性能。**结论** 融合逆密度函数的关系形状卷积神经网络可以有效获取点云数据中的局部和全局特征, 并对点云采样不均匀的情况实现一定程度的补偿, 实现更优的分类和分割效果。

关键词: 关系形状卷积神经网络(RSCNN); 逆密度函数; 非均匀采样; 反卷积层; 点云的分类与分割

Point cloud analysis model combining inverse density function and relation-shape convolution neural network

Qiu Yunfei¹, Zhu Mengying^{1,2*}

1. College of Software, Liaoning Technical University, Huludao 125100, China;

2. Quanzhou Institute of Equipment Manufacturing Haixi Institutes, Chinese Academy of Sciences, Quanzhou 362216, China

Abstract: **Objective** 3D point clouds have received widespread attention for their wide range of applications, such as robotics, autonomous driving, and virtual reality applications. However, due to the characteristics of input and output size and order differences, uneven density, and differences in shape and scaling, point cloud processing is very challenging, and because of this, various shapes formed by irregular points are often difficult to distinguish. For this problem, sufficient contextual semantic information must be captured to thoroughly grasp the elusive shapes. In 2D images, convolutional neural

收稿日期: 2020-06-16; 修回日期: 2020-08-13; 预印本日期: 2020-08-20

* 通信作者: 朱梦影 3066964438@qq.com

networks (CNN) have fundamentally changed the landscape of computer vision by greatly improving the results of almost all vision tasks. CNN succeeds by using translation invariance, so the same set of convolution filters can be applied to all positions in the image, thereby reducing the number of parameters and improving the generalization ability. We hope to transfer these successes to 3D point cloud analysis. However, a 3D point cloud is an unordered set of 3D points, each with or without additional features (such as RGB), which does not conform to the regular lattice grid in a 2D image. Applying conventional CNNs to such unordered inputs is difficult. Some works convert point clouds into regular voxels or multi-view images. However, these conversions usually cause a large amount of inherent geometric information to be lost, and the amount of data in the 3D point cloud is huge, increasing the complexity of the conversion. Another solution is to learn directly from an irregular point cloud. PointNet learns each point independently, and then applies a symmetric function to accumulate features. This direction can realize the invariance of point replacement. Although impressive, it ignores local patterns that have proven to be important for advanced visual semantics in abstract image CNNs. To correct this, KCNet mines local patterns by creating KNN (K-nearest neighbor) graphs at each point of PointNet. However, it does not have a pooling layer that can explicitly increase the semantic level. PointNet++ hierarchical groups point clouds into local subsets and learn them through PointNet. This design does work similar to CNN, but the basic operation of PointNet requires high complexity to achieve sufficient effectiveness, which usually results in huge computational costs. **Method** To solve this problem, some works divide the point cloud into several subsets by sampling. Then, a hierarchical structure is constructed to learn the context representation from local to global. However, this greatly depends on effective inductive learning of local subsets, which is difficult to achieve for point clouds with uneven density and irregular shape. Inspired by inverse density function in point cloud networks, we propose a relational shape CNN (RSCNN) fused with inverse density function. The key of the network is to learn from the relationship, that is, to learn from the geometric topology constraints between the points, which encode meaningful shape information in the 3D point cloud, and the convolution kernel is regarded as a nonlinear function of the local coordinates of 3D points composed of 3D points. For a given point, predefined geometric prior convolution weights are used in advanced relational expressions, and kernel density estimation is used to study inverse density functions. The introduction of the inverse density function is further applied to non-uniformly sampled point clouds. In addition, for the segmentation task, we propose a deconvolution operator. Relation-shape deconvolution layer (RSDeconv) consists of two parts: interpolation and Relation-shape convolution layer (RSConv). The feature is propagated from the sampling point cloud to the original resolution through convolution. The spatial layout of the points can be used to calculate the induction through convolution, and it can reflect the potential shape formed by irregular points differently, so as to realize the context shape awareness learning of point cloud analysis. Benefit from the geometric priors, the invariance to point displacement and the robustness to rigid transformations (such as translation and rotation) are realized. **Result** Classification, partial segmentation, and semantic scene segmentation experiments were conducted on ModelNet40, ShapeNet, and ScanNet datasets to verify the classification and segmentation performance of the model. In the classification experiment of ModelNet40, its overall accuracy is improved by 3.1% compared with PointNet++. Even when PointNet++ takes normal as input, its accuracy is improved by 1.9% compared with PointNet++. In the ShapeNet partially segmented dataset, the mean intersection-over-union (mIoU) is 6.0% higher than that of PointNet++, and the instance mIoU is 1.4% higher than that of PointNet++. In the ScanNet indoor scene dataset, the mIoU is 13.7% higher than that of PointNet++. A comparison experiment is conducted on the ScanNet dataset to compare the unsynchronized length with and without density function. The experiment proves that the inverse density function improves the segmentation accuracy by about 0.8% and effectively improves the model performance. **Conclusion** The experimental results show that the heterogeneous sampling of point clouds can be realized by introducing the inverse density function into the relational CNN. In addition, the deconvolution layer further propagates the feature from the subsampling point cloud back to the original resolution, which improves the segmentation accuracy. In general, the proposed network model can effectively obtain the global and local characteristics of point cloud data, achieving better classification and segmentation effect.

Key words: relation-shape convolutional neural network (RSCNN); inverse density function; non-uniform sampling; deconvolution layer; classification and segmentation of point cloud

0 引言

随着 3D 点云在自动驾驶 (Qi 等, 2018) 和导航定位 (Zhu 等, 2017; 文威威 等, 2019) 等诸多领域的应用, 3D 点云引起了越来越多的关注。由于很难推断不规则点云形成的基本形状, 此项工作非常具有挑战性, 为此提出了很多方法, 主要有基于体素和基于视图的方法。

鉴于卷积神经网络 (convolution neural network, CNN) 在图像识别领域强大的语义信息抽象能力 (Chen 等, 2018; 黄凯奇 等, 2014), 很多方法将其在图像分析即规则网格数据上的处理用于不规则点云处理。为使点数据适用于卷积, 一种简单的方法是将其体素化为 3D 网格结构 (Maturana 和 Scherer, 2015; Wu 等, 2015), 但是由于大多数体素通常未能占用, 因此体素的表示方式效率不高。为此, Riegler 等人 (2017) 提出 OctNet, 探索了体素数据的稀疏性并缓解了这个问题, 但是当涉及更深的神经网络时, 内存占用仍然很高, 而且由于体素是空间的离散表示, 因此该方法需要高分辨率的网格, 对存储器容量提出了较高要求。多视图 (Qi 等, 2016; Su 等, 2015) 是另一种常见的 3D 表示形式, 多视图将点数据投影到 3D 空间中的各种特定图像平面上形成 2D 图像, 通过这种方式可以在 2D 图像上使用常规卷积处理点数据, 但是这种方法忽略了 3D 点的内在几何关系, 投影导致的 3D 数据中的遮挡部分未得到处理, 并且由于点云数据密度分布的不均匀性, 这些方法无法充分识别点云细粒度的局部结构。Qi 等人 (2017a) 提出的 PointNet 通过在每个点上独立学习并通过最大池收集最终特征, 然而这种设计同样忽略了局部结构信息。事实证明, 局部结构信息对于 CNN 的成功至关重要。为了解决这个问题, Qi 等人 (2017b) 提出 PointNet++ 将 PointNet 分层应用到点云的多个子集, 使其具有可扩展性, 然而该方法是在点周围半径内选择固定数量的随机样本, 在计算中没有考虑点密度, 因此导致误差。Shen 等人 (2018) 提出使用内核相关层来开发局部几何结构, 并提出了一个基于图的池化层, 以利用局部特征结构来增强网络的鲁棒性。在小规模点集上表现出良好的性能, 但对底层的图结构仍然特别敏感。大多数卷积方法无论是否使用层次结构都是基于 K 个

最近邻的卷积 (Klokov 和 Lempitsky, 2017; Li 等, 2018; Wang 等, 2019; Xu 等, 2018)。其中, Li 等人 (2018) 提出的 PointCNN 在整个邻域上使用多层感知器 (multilayer perceptron, MLP) 来学习一个变换矩阵, 然后用其加权和置换相邻点的输入特征, 并对变换后的特征进行标准图像卷积; Wang 等人 (2019) 提出的动态图卷积 (dynamic graph CNN, DGCNN) 在相邻点图的边上使用了可学习操作的一般概念, 与其他方法相比, 在学习过程中改变了邻域, 使得该方法的实现稍微复杂一些, 在大型点云上效率较低。这些基于 K 近邻的方法在非均匀采样环境下并不健壮, 因为在空间中的一个区域中添加更多的点会使捕捉到的特征与内核训练时不同, 产生点云稠密区域感受野收缩与点云稀疏区域感受野扩大的视觉差。有些方法将点云映射到高维空间, 以促进经典 CNN 的应用。SPLATNet (sparse lattice networks for point cloud processing) (Su 等, 2018) 将输入点映射到稀疏网格上, 然后使用双边卷积进行处理 (Jampani 等, 2016)。PCNN (point convolutional neural networks by extension operators) (Atzmon 等, 2018) 将点云上的函数扩展为环境空间上的连续体积函数, 虽然没有考虑非均匀性, 但将点云转换为对非均匀采样点云具有鲁棒性的体积函数。然而这些方法的部分计算在点数上是二次的, 因而不可扩展。

以上方法很大程度上解决的只是点云积分的不规则性问题。对此, 有些研究在解决点云积分不规则性的同时考虑了点云采样密度不均匀问题。Hermosilla 等人 (2018) 利用蒙特卡罗卷积解决点云上的采样不均匀问题, 但包含多个点的接收域计算速度较慢, 效率低。Wu 等人 (2019) 提出的 PointConv 将卷积核视为由权函数和密度函数组成的 3 维点的局部坐标的非线性函数。Li 等人 (2019) 设计了一个密度感知卷积模块, 利用每个点的点态密度加权卷积核的可学习权重。

关系卷积神经网络是从关系中学习, 关系即点之间的几何拓扑约束。受密度函数的启发, 本文提出将关系形状卷积神经网络 (relation-shape convolution neural network, RSCNN) (Liu 等, 2019) 与逆密度函数相融合, 得到点之间的高级关系表达式, 以此捕获点的空间布局, 并逐步执行上下文局部到全局形状的特征提取。通过学习点之间几何关系的方法, 对刚性变换实现了很好的鲁棒性。逆密度函数通过

核密度估计(kernel density estimation, KDE)来学习,目的是对点云采样率不均匀情况进行补偿,实现对点云进行非均匀采样。同时在点云分割任务中,引入关系卷积神经网络的反卷积层(Deconv),该层更好地将特征从子采样点云传播回原始分辨率。实验证明,本文方法对点云分割精确度的提升起到了很好的作用。

1 算法流程

本文提出一种对3D点云进行卷积的方法,能够对点云采样率不均匀的情况进行一定程度的补偿,并且使用逆密度函数对多层感知器学习的连续函数进行加权。在分割任务中,添加一个基于关系形状卷积神经网络的反卷积层,使用反卷积层将特征信息从较粗层传播到较细的层,实现了改进语义分割任务的性能。本文算法流程如图1所示,图中, $N(x_i)$ 表示采样点 x_i 的邻居点, f_{x_j} 表示 x_j 的特征向量, h_{ij} 表示 x_j 与 x_i 之间的低级关系向量,函数 A 用来合并所有特征, σ 表示非线性激活函数,映射 M 抽象出两点之间高级关系表达式。

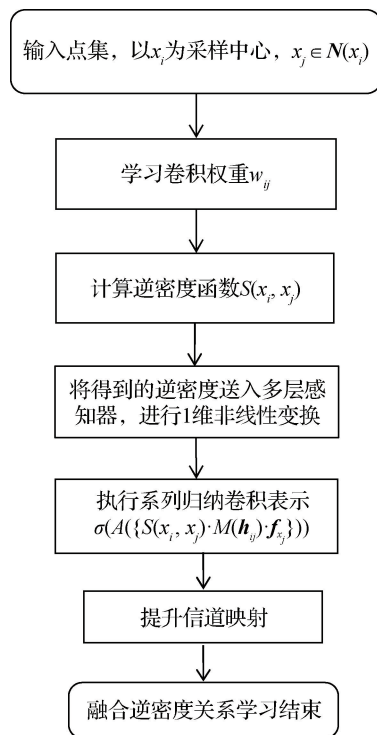


图1 融合逆密度函数的关系形状卷积神经网络流程图

Fig. 1 Flow chart of relational shape convolutional neural network fused with inverse density function

2 原理与方法

融合逆密度函数的关系形状卷积神经网络从3D空间中的几何先验学习高级关系表达式,并执行上下文局部到全局的形状学习。

2.1 关系学习

从关系中获取与数据相关权重的方法在图像和视频分析领域已经有所应用。空间变换器(Jaderberg等,2015)通过学习过渡矩阵对齐2D图像,非局域网(Wang等,2018)学习跨视频帧的长期关系,关系网络(Hu等,2018)学习跨对象的位置关系。还有一些研究专注于对3D点云中的关系进行学习,DGCNN(Wang等,2019)通过学习高维特征空间中的点关系来捕获相似的局部形状,但这种关系在某些情况下可能不可靠。

2.2 关系形状卷积

卷积运算可以看做是连续卷积算子的离散逼近。在3D空间中,可以将此卷积运算符的权重视为相对于参考3D点的局部点坐标的连续函数,连续功能可以通过多层感知器进行近似。在关系形状卷积神经网络中,通过将采样点 x 作为质心并将周围点作为其邻居 $N(x_i)$ 来构造卷积权重,从预定义的几何先验(即 x 和 $N(x_i)$ 之间的直观低级关系)中学习高级关系表达式。通过这种方式进行卷积可以获得具有关于点空间布局的归纳表示,可辨别地反映不规则点形成的潜在形状。受益于几何先验,实现了点的排列不变性和对刚性变换(例如平移和旋转)的鲁棒性,但关系形状卷积神经网络更加适用于均匀采样的点集,而忽略了采样率不均匀的情况。对此,本文提出融合逆密度函数的关系形状卷积网络,并使用反卷积层将特征信息传播回原始的分辨率,从而改进语义分割任务的性能。

关系形状卷积(RSConv)将规则网格CNN扩展到不规则配置,并最终学习点云的上下文形状感知表示,但是非常依赖从不规则点子集中进行形状感知的归纳学习。对此,关系形状卷积网络将局部点子集 $P_{\text{sub}} \subset \mathbf{R}^3$ 建模成为球面邻域(如图2所示),以采样点 x_i 为质心,周围的点为其邻居 $x_j \in N(x_i)$,然后学习该邻域的归纳表示 $f_{p_{\text{sub}}}$,关系形状卷积网络的卷积运算为

$$f_{p_{\text{sub}}} = \sigma(A(\{T(f_{x_j}), \forall x_j\}))$$

$$d_{ij} < r, \forall x_j \in N(x_i) \quad (1)$$

式中, x 是 3D 点, f 是特征向量, d_{ij} 是 x_i 与 x_j 之间的欧氏距离, r 是球域半径。首先用函数 T 变换 $N(x_i)$ 中所有点的特征, 然后用函数 A 合并所有特征, 最后通过非线性激活因子 σ 来获得。式(1)中, 两个函数 A 和 T 是关键, 只有当 A 是对称的, 并且在 $N(x_i)$ 中的每个点上共享 T 时, 才能实现点集的排列不变性。

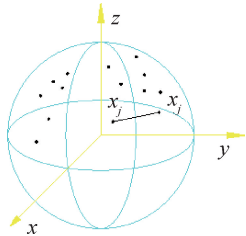
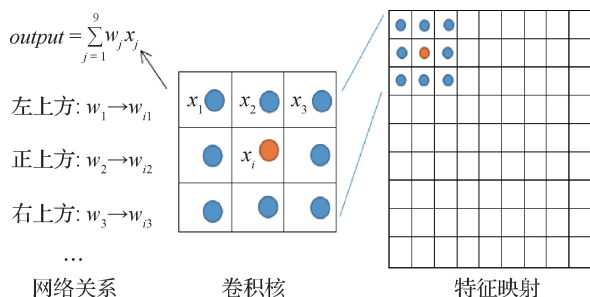


图2 建模球邻域图

Fig. 2 Modeling spherical neighborhood graph

2.3 权重函数

在经典 CNN 中, T 实现为 $T(f_{x_j}) = w_j \cdot f_{x_j}$, 其中 w_j 是可学习的权重, “ $\cdot$ ” 表示逐元素相乘。如图 3 所示, 在 2D 点云上应用时, 此卷积的主要限制有: 1) w_j 未在 $N(x_i)$ 中每个点上共享, 导致点置换的变化和无法处理不规则局部点子集 P_{sub} ; 2) 反向传播中, w_j 的梯度仅与孤立点 x_j 有关, 这会给 $f_{p_{\text{sub}}}$ 带来较少的形状感知和鲁棒性。

图3 内核为 3×3 的 2 维网格卷积图Fig. 3 Illustration of 2D grid convolution with a kernel of 3×3

在 3D 空间的邻域中, x_i 及其所有邻域 $N(x_i)$ 之间的几何关系是点空间布局的明确表达, 能够反映其基本形状。在关系形状卷积神经网络中用 w_{ij} 替换经典 CNN 中的 w_j , 用权重 w_{ij} 反映点 x_i 与点 x_j 之间存在的位置关系 (如图 3 所示), 通过对关系向量 h_{ij} 的映射 M 学习得到, h_{ij} 即 x_i 与 x_j 之间的预定义几何先验, 称为低级关系。该过程可以描述为

$$T(f_{x_j}) = w_{ij} \cdot f_{x_j} = M(h_{ij}) \cdot f_{x_j} \quad (2)$$

映射 M 的目的是抽象两点之间的高级关系表达式, 可以对其空间布局进行编码。鉴于多层感知器强大的映射功能, 使用共享的多层感知器来实现 M , 则式(1)中的 $f_{p_{\text{sub}}}$ 变为

$$f_{p_{\text{sub}}} = \sigma(A(\{M(h_{ij}) \cdot f_{x_j}, \forall x_j\})) \quad (3)$$

这种卷积表示包含 x_i 与 $N(x_i)$ 之间所有关系的总和, 可以实现相关点空间布局的显式推理, 从而获取形状感知。对于几何先验, 可以使用 3D 欧氏距离作为对低级关系 h_{ij} 的直观描述。此外, h_{ij} 还可以灵活定义, 因为 M 可以将其映射到高维关系向量使其与 f_{x_j} 通道对齐, 从而易于点乘。在式(3)中, $f_{p_{\text{sub}}}$ 的通道数与输入特征 f_{x_j} 相同, 这与经典图像 CNN 不一致。经典图像 CNN 增加了通道数, 同时降低了图像分辨率, 以获得更抽象的表示。例如, 在 VGG (Visual Geometry Group) 网络中依次设置了通道数 64—128—256—512。因此, 在 $f_{p_{\text{sub}}}$ 上添加了共享的多层感知器, 用于进一步的信道提升映射。

3 逆密度函数

从点云空间分布的角度出发, 3 维点的不规则结构会导致某些区域点集的过采样或欠采样, 从而导致卷积核上的密度不同。在 2D 图像中, 这种视差类比为 1 个像素接收多个值, 当更新内核时就会产生偏差, 这是因为在过采样区域设置了过多的权重, 在欠采样区域设置了较少的权重。在数字化图像中, 像素均匀地放置在 2 维网格上, 相邻像素之间的距离是恒定的, 网格中的每个单元格都只有 1 个值, 并且没有空的单元格, 整个栅格密度均匀, 但这并不适用于不规则的 3 维点云。为了克服这种由于点云不规则带来的不平衡, 同时考虑到本文的基础网络——关系形状卷积神经网络更加适用均匀采样的点云而没有考虑采样率不均匀的情况, 在关系形状卷积神经网络中引入逆密度函数, 利用局部区域的密度估计来重新缩放核函数, 对非均匀采样的点进行补偿。由此, 卷积公式变换为

$$f_{p_{\text{sub}}} = \sigma(A(\{S(x_i, x_j) \cdot M(h_{ij}) \cdot f_{x_j}, \forall x_j\})) \quad (4)$$

式中, $S(x_i, x_j)$ 是点的逆密度函数, 因为点云可能非均匀采样。直观地, 局部点的数量在整个点云中

变化,如图4所示,点 $p_3, p_4, p_6, p_7, p_8, p_9$ 非常接近另一个,因此每个点的权重应该较小。

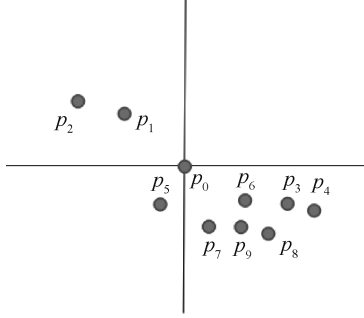


图4 点非均匀分布图

Fig. 4 Point non-uniform distribution map

本文的主要思想是通过对3D点的坐标进行多层感知器的高层映射和通过核密度估计的逆密度函数的点乘来近似权重函数 w_{ij} ,然后使用多层感知器实现非线性变换,这样的权重函数能够进一步反映输入点云的分布,网络中多层感知器的权重在所有点之间共享,以保持排列的不变性。

为了计算逆密度函数 $S(x_i, x_j)$,首先使用核密度估计计算点云中每个点的密度,将 $x_j \in N(x_i)$ 周围的局部区域视为半径为 r 的球,并应用Parzen-Rosenblatt窗口方法(Parzen, 1962)沿内部每个轴拟合密度分布球,具体为

$$S(x_i, x_j)_{\text{KDE}} = \frac{1}{Nh} \sum_{x_j \in N(x_i)} G\left(\frac{x_i - x_j}{h}\right) \quad (5)$$

式中, G 是一个具有平滑因子 h 和归一化常数 N 的多元高斯核,归一化常数 N 是 x_i 局部区域中的点数,KDE表示核密度估计(kernel density estimation)。然后将密度输入多层感知器进行1维非线性变换(如图5所示,图中 K 表示每个局部空间的点数量), C_{in} 为输入的通道数,使网络可以自适应地决定是否使用核密度估计,对其取倒数即为此时的逆密

度函数,即

$$S(x_i, x_j) = \frac{1}{MLP(S(x_i, x_j)_{\text{KDE}})} \quad (6)$$

式中, MLP 函数表示多层感知器(multi-layer perception)。

本文没有使用绝对距离,而是使用从相邻点 $x_j \in N(x_i)$ 到质心 x_i 的相对距离(即 $x_i - x_j$)来估计 $S(x_i, x_j)$,将逆密度函数限制为半径为 r 的球域内。通过固定域,模型可以将其核密度估计值推广到具有相同半径局部区域的任意点。

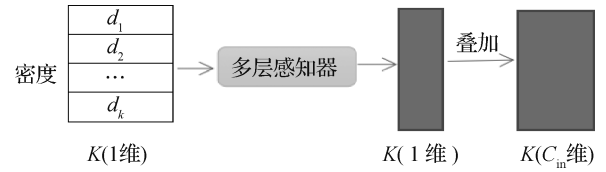


图5 计算逆密度函数流程示意图

Fig. 5 Schematic diagram of calculating the inverse density function

4 反卷积的特征传播

对于分割任务需要进行逐点预测,为了获得所有输入点的特征,需要一种将特征从子采样点云传播到更密集点云的方法。PointNet++使用基于距离的插值来传播特征,主要用于从粗糙级别捕获传播信息的局部相关性,但是并没有充分利用反卷积的操作。对此,本文在关系形状卷积神经网络的最后一层添加一个由插值和关系形状卷积两部分组成的反卷积层,如图6所示。 c 代表经过逆密度关系卷积之后的点输出层数。首先采用插值法传播来自上一层的粗略特征,使用PointNet++的插值方法,通过对来自选定点的3个最近点的特征进行线性插

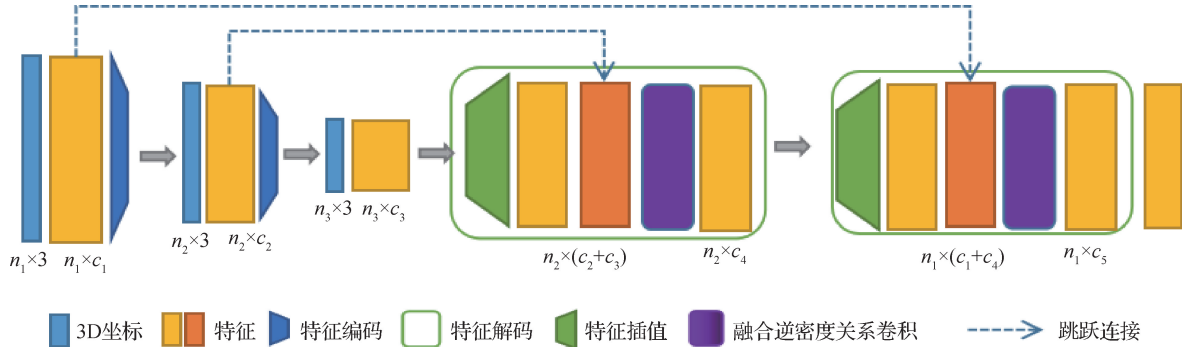


图6 反卷积特征传播图

Fig. 6 Deconvolution feature propagation diagram

值。然后使用跳跃连接将内插特征与具有相同分辨率的卷积层中的特征连接在一起。连接后,将关系形状卷积应用于连接后的特征值以获得最终的反卷积输出,类似于图像反卷积层(Noh 等,2015),重复应用此过程,直至所有输入点的特征都传播到原始分辨率为止。

5 点云模型分析

融合逆密度函数的关系形状卷积神经网络可以开发出类似经典 CNN 的分层形状感知学习架构用于点云分析,表示为

$$\mathbf{F}_{p_{N_l}}^l = \text{RSCONV}(\mathbf{F}_{p_{N_{l-1}}}^{l-1}) \quad (7)$$

式中, p 代表点集, $\mathbf{F}$ 是采样点集层的特征, N 是采样点集的数量,该层特征通过对前一层中的特征经融合逆密度函数的关系形状卷积获得。将本文模型用于点云分类和分割任务。在点云分类中,首先学习一组层次的形状感知表示形成最终的全局表示形式,然后通过完全连接层(fully connected layers, FC)进行分类(如图 7 所示)。在分割任务中,通过特征传播对学习的多级表示进行连续上采样生成每点预测(如图 8 所示)。两者都能以端到端的方式进行训练。

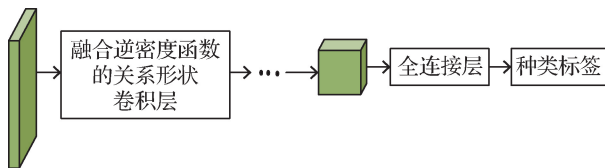


图 7 分类的网络架构

Fig. 7 Classified network architecture

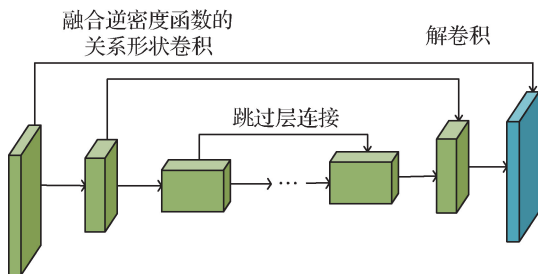


图 8 分割的网络架构

Fig. 8 Segmented network architecture

6 实验结果与分析

6.1 实验配置

在实验中,本文提出的融合逆密度函数的关系

形状卷积神经网络的聚合函数使用具有对称性质的最大池化层,ReLU(rectified linear unit)用做非线性激活因子。对映射函数 M ,使用 3 层共享多层感知器,因为理论上多层感知器适合任意连续映射。底层关系定义为 3D 欧氏距离。信道提升映射通过单层共享多层感知器实现,每个多层感知器都进行批量归一化。

为验证本文模型的有效性,在 ModelNet40(Yi 等,2016)、ShapeNet(Wu 等,2015)和 ScanNet(Dai 等,2017)数据集上进行实验。实验环境为 Ubuntu14.04 操作系统、GTX1080Ti 显卡、GPU 显卡、Python3.0,基于 Pytorch 框架。

6.2 点云分类

6.2.1 数据集

ModelNet40 数据集由 40 个种类的 12 311 个 CAD 模型组成,包含 9 843 个训练模型和 2 468 个测试模型。在采样率不均匀的情况下随机采样 1 024 个点并将其标准化为单位球体。在训练过程中,使用 $[-0.66, 1.5]$ 范围内的随机缩放和 $[-0.2, 0.2]$ 范围内的平移来扩充输入数据。为防止过拟合,在全连接层中使用丢弃技术,dropout = 0.5。在测试过程中,执行 10 次具有随机缩放比例的投票测试,并对预测结果取平均。

6.2.2 分类结果

表 1 是本文方法与 PointNet(Qi 等,2017)、PointNet++(Qi 等,2017)、PointCNN(Li 等,2018)、DGCNN(Wang 等,2019)、PCNN(Atzmon 等,2018)、RSCNN(Liu 等,2019)和 SpiderCNN(Xu 等,2018)方法在 ModelNet40 数据集上点云分类性能的定量比较。可以看出,本文提出的融合逆密度函数的关系形状卷积神经网络优于所有 xyz 点坐标输入的对比方法。与 PointNet++ 方法相比,本文方法的分类精度提升了 3.1%,并超过其高级版本,该版本输入更加密集的点(5 120 个)且将点云的法线也作为输入;与 RSCNN 相比,本文方法将精度提升了 0.6%。实验结果表明,本文模型能够通过将逆密度函数与几何拓扑的关系学习结合,有效捕获点云的形状信息。

6.3 点云分割

3D 对象零件分割是一项具有挑战性的细粒度 3D 识别任务。给定 3D 网格模型后,任务是为每个点或每个面分配零件类别(例如椅子腿、杯柄等)标签。

表1 不同方法在 ModelNet40 数据集上的分类精确度对比

Table 1 Comparison of classification accuracy on ModelNet40 dataset among different methods

方法	输入	采样点数	精确度/%
PointNet (Qi 等, 2017)	点坐标	1 024	86.3
PointNet ++ (Qi 等, 2017)	点坐标	1 024	90.1
PointCNN (Li 等, 2018)	点坐标	1 024	91.1
DGCNN (Wang 等, 2019)	点坐标	1 024	91.8
PCNN (Atzmon 等, 2018)	点坐标	1 024	91.5
RSCNN (Liu 等, 2019)	点坐标	1 024	92.6
PointNet ++ (Qi 等, 2017)	点坐标 + 法线	5 120	91.3
SpiderCNN (Xu 等, 2018)	点坐标 + 法线	5 120	91.7
本文	点坐标	1 024	93.2

注:加粗字体表示最优结果。

6.3.1 数据集

ShapeNet 数据集包含来自 16 个类别的 50 个零件的 16 881 个形状。任务的输入是由点云表示的形状,本文将零件细分问题转化为按点分类问题,评估指标按点计算。将网络模型应用在 ShapeNet 数据集的部件基准上,并进行评估零件分割任务,不均匀地随机选择 2 048 个点作为输入,并将对象标签的一键编码连接到最后一个特征层。在测试期间,依然使用随机缩放比例进行 10 次投票测试,最后对预测结果取平均。

6.3.2 分割结果

表2是本文方法与 PointNet、PCNN、DGCNN、

表2 不同方法在 ShapeNet 数据集上的分割结果对比

Table 2 Comparison of segmentation results on ShapeNet dataset among different methods

方法	输入(采样点数)	类平均交并比/%	实例平均交并比/%
PointNet	点坐标(2 048)	80.4	83.7
PCNN	点坐标(2 048)	81.8	85.1
DGCNN	点坐标(2 048)	82.3	85.1
RSCNN	点坐标(2 048)	84.0	86.2
PointNet ++	点坐标 + 法线(2 048)	81.9	85.1
本文	点坐标(2 048)	87.9	86.5

注:加粗字体表示各列最优结果。

RSCNN 和 PointNet ++ 方法在 ShapeNet 数据集上点云分割性能的定量比较。实验使用平均交并比(mean intersection over union, mIoU)作为评估指标,分别计算了在所有类别和所有实例中两种类型的平均交并比。可以看出,本文方法的类 mIoU 和实例 mIoU 分别为 87.9% 和 86.5%,实现了最佳性能,不仅超过了 RSCNN 的验证结果,而且在 PointNet ++ 高级版本输入情况下,结果也得到了较高提升。实验结果表明,逆密度函数能够较好地实现对不均匀采样的点云密度补偿,反卷积层的引入也进一步提升了点云分割对各种形状结构的鲁棒性。

图9显示了本文方法在 ShapeNet 数据集上的一些分割示例。可以看到,尽管隐含在不规则点中的零件形状多种多样,并且可能很难识别,但是本文

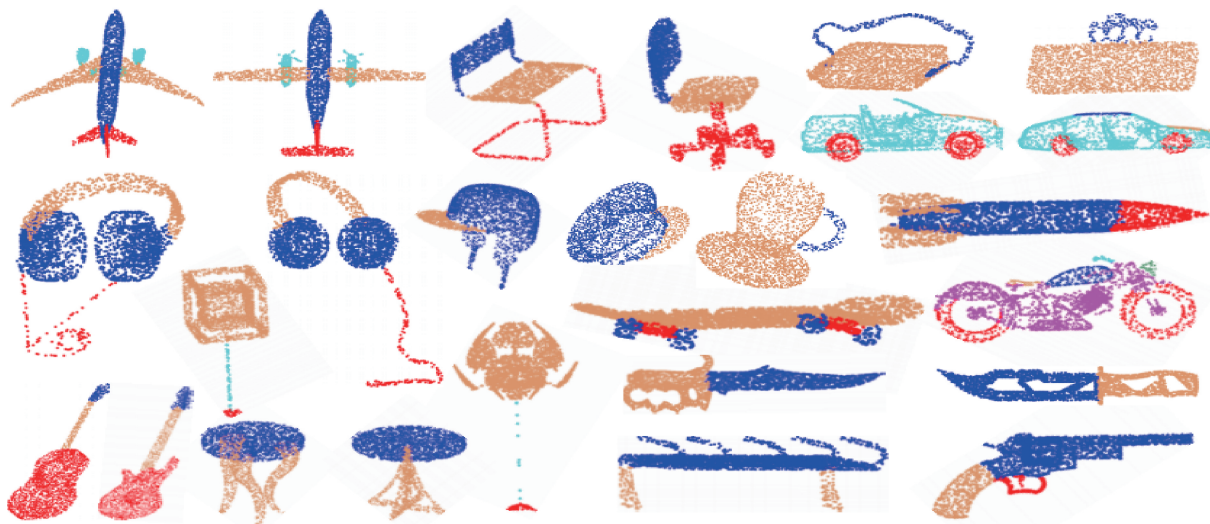


图9 本文方法在 ShapeNet 数据集上的零件分割示例

Fig. 9 Examples of part segmentation on ShapeNet dataset by our method

方法依然可以对其以相当不错的准确率进行点云分割。

6.4 法线估计

点云中的法线估计对于许多应用(例如曲面重建和渲染)是至关重要的一步,由于需要更高层次的推理,超越了潜在的形状识别,此任务非常具有挑战性。本文将法线估计(袁小翠等,2017)作为有监督的回归任务并使用细分网络实现。将归一化输出与真值法线之间的余弦损失用于回归训练,使用ModelNet40数据集进行评估,以统一采样的1 024个点作为输入。

表3是本文方法与PointNet、PointNet++和PCNN方法在ModelNet40数据集上法线估算的定量比较。可以看出,本文方法的误差为0.14,表现优于其他方法,将PointNet++的误差减少了0.15。实验结果表明,本文方法能从点云的几何拓扑关系学习中以更低的误差实现对法线的估计,更好地对点云的法线进行预测。

表3 不同方法在ModelNet40数据集上的法线估计误差对比
Table 3 Comparison of normal estimation errors on ModelNet40 dataset among different methods

方法	采样点数	误差
PointNet	1 024	0.47
PointNet++	1 024	0.29
PCNN	1 024	0.19
本文	1 024	0.14

注:加粗字体表示最优结果。

6.5 语义场景标签

为了评估本文方法处理包含大量嘈杂数据的现实点云的能力,在ScanNet数据集上进行语义场景分割。任务是在给定由点云表示的室内场景的情况下,预测每个3D点上的语义对象标签。

ScanNet数据集的最新版本包含所有1 513个ScanNet扫描和100个新测试扫描的更新注释,其中所有语义标签均不公开,算法的输入仅使用3D坐标数据和RGB信息。在实验中,通过从室内房间随机采样3 m × 1.5 m × 1.5 m立方体来生成训练样本,并在整个扫描过程中使用滑动窗口进行评估,使用平均交并比(mIoU)作为主要衡量标准。

表4是本文方法与SPLATNet、PointNet++和ScanNet方法在ScanNet数据集上所有类别IoU的平均值(mIoU)的定量比较。可以看出,本文模型的语义分割精度比PointNet++方法提升了13.7%,相较其他算法均有较大幅度的提升。实验结果表明,本文方法能够较好地实现具有大量噪声的真实点云场景数据的语义分割,体现了处理真实点云的鲁棒性和良好的泛化能力。

表4 不同方法在ScanNet数据集上的语义场景分割结果对比
Table 4 Comparison of semantic scene segmentation results on ScanNet dataset among different methods

方法	类平均交并比/%
ScanNet	30.6
PointNet++	33.9
SPLATNet	39.3
本文	47.6

注:加粗字体表示最优结果。

6.6 逆密度函数评估

为验证逆密度函数S的有效性,在ScanNet数据集上进行对比实验,因为ScanNet数据集中的点云是从真实的室内场景生成的。遵循ScanNet数据集提供的标准训练/验证集拆分,分别在具有和不具有RGB信息、有和没有逆密度函数和不同步长大小滑动窗口等情况下训练网络,对比实验结果如表5所示。可以看出,1)具有逆密度函数的关系卷积神经网络的性能比不具有逆密度函数的关系卷积神经网络的性能提高约1%,证明了逆密度函数在关系卷积神经网络进行点云分割中的有效性。在实验中,观察到逆密度函数在靠近输入的层中更有效,主要是因为更深层的卷积网络中,多层感知器倾向于逐渐减少逆密度函数的影响。一个可能的原因是,本文算法使用最远的点采样算法作为子采样算法,更深层的点云倾向于更均匀地分布,这时逆密度函数的作用将减小。2)与ScanNet数据集上没有逆密度函数的方法相比,不使用非线性变换而直接应用逆密度函数的结果相对于使用非线性变换的结果有一定程度的降低,表明非线性变换能够对是否使用逆密度函数进行评估,从而更好地提升点云的分割精度。

表5 在ScanNet数据集上有无逆密度函数的实验结果对比

Table 5 Comparison of experimental results with and without inverse density function on ScanNet dataset

输入	步长	/%		
		mIoU	mIoU1	mIoU2
点坐标	0.5	61.0	60.2	60.1
点坐标	1.0	59.0	58.2	57.7
点坐标	1.5	58.2	56.9	57.3
点坐标 + RGB	0.5	60.8	58.9	-
点坐标 + RGB	1.0	58.6	56.7	-
点坐标 + RGB	1.5	57.5	56.1	-

注:加粗字体表示各列最优结果,“-”表示未知结果。mIoU1表示无逆密度情况,mIoU2表示有逆密度、无线性层情况。

6.7 鲁棒性及复杂性测试

为验证点置换和刚性变换的鲁棒性,将本文方法与 PointNet 和 PointNet ++ 方法进行点云的鲁棒性比较。在实验过程中,所有模型都经过训练,没有相关的数据扩充(例如平移或旋转),以避免在测试中造成混淆。此外,尽管融合逆密度函数的关系卷积神经网络中的关系学习对旋转具有鲁棒性,但3D坐标的初始输入特征对旋转的鲁棒性仍有一定影响,对此,通过将每个采样点的子集归一化为相应的局部坐标系来解决,该局部坐标系由每个采样点及其法线确定。为了公平比较,对 PointNet ++ 执行此归一化,因为这样可以局部子集中学习。在测试中,将3D欧氏距离作为本文方法中的几何关系,测试结果如表6所示。可以看出,所有方法对于排列都是不变的。但是,PointNet 易受平移和旋转的影响,而 PointNet ++ 对旋转敏感。相比之下,本文方法对这些扰动是不变的,表明本文方法对点云平移或旋转变换之后的形状识别仍然具有较大

表6 不同模型在点云平移或旋转后的精确度对比

Table 6 Comparison of accuracy after point cloud translation or rotation among different models

方法	/%			
	平移	平移	旋转	旋转
	0.2	-0.2	90°	180°
PointNet	70.8	70.6	42.5	38.6
PointNet ++	88.2	88.2	47.9	39.7
本文	90.3	90.3	90.3	90.3

注:加粗字体表示各列最优结果,旋转为沿y轴旋转。

潜力。

6.8 复杂度分析

表7是在输入为1024个点时本文方法与 PointNet 和 PointNet ++ 方法分类的空间(参数数量)和运行时间(浮点运算/采样)复杂度对比。可以看出,与 PointNet 相比,本文模型的参数减少了59.4%,采样复杂度减少了21.1%。实验结果表明,本文模型在实时应用(例如自动驾驶中的场景解析)中具有巨大的潜力。

表7 不同模型在点云分类中的复杂度对比

Table 7 Comparison of complexity of point cloud classification among different models

方法	参数	/MB	
		浮点计算(采样)	
PointNet	3.5	440	
PointNet ++	1.48	1684	
本文	1.42	347	

注:加粗字体表示各列最优结果。

7 结论

本文通过引入逆密度函数对关系形状卷积神经网络(RSCNN)进行改进,以使关系形状卷积神经网络对采样率不均匀的点云更加鲁棒,并在原网络架构的基础上添加了关系形状反卷积层,用于从粗糙级别捕获传播信息的局部相关性,同时保持关系形状卷积神经网络本身具备的排列不变性、刚性转换的鲁棒性和权重共享等关键特性,将规则网格CNN扩展到不规则配置以进行点云分析,可以从关系(即点之间的几何拓扑约束)中学习。通过对点的空间布局进行明确推理,获得有区别的形状意识。通过点云分类、部分分割和语义分割3项任务的具有挑战性的基准测试的广泛实验,以及详尽的经验和理论分析,证明融合逆密度函数的关系形状卷积神经网络相对于其他网络模型都有一定程度的提升。

但是,每个局部点子集的强制归一化一定程度上会给形状识别带来困难,另一方面融合逆密度函数的关系形状卷积网络在采样率相对均匀情况下的分类精度与原始关系卷积神经网络呈现的精度结果相当,这主要是因为 ModelNet40 数据集在整体上呈现均匀状态,尽管局部存在非均匀情况,但实际使用

的数据集出现不均匀情况是比较常见的,这也是本文网络今后的改进点。下一步工作是将原始坐标和几何特征融合到同一特征空间,进一步提升点云识别精度。

参考文献 (References)

- Atzmon M, Maron H and Lipman Y. 2018. Point convolutional neural networks by extension operators. *ACM Transactions on Graphics*, 37(4): #71
- 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]
- Dai A, Chang A X, Savva M, Halber M, Funkhouser T and Nießner M. 2017. ScanNet: richly-annotated 3D reconstructions of indoor scenes//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA; IEEE: 2432-2443 [DOI: 10.1109/CVPR.2017.261]
- Hermosilla P, Ritschel T, Vázquez P P, Vinacua À and Ropinski T. 2018. Monte Carlo convolution for learning on non-uniformly sampled point clouds. *ACM Transactions on Graphics*, 37(6): #235 [DOI: 10.1145/3272127.3275110]
- Hu H, Gu J Y, Zhang Z, Dai J F and Wei Y C. 2018. Relation networks for object detection//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA; IEEE: 3588-3597 [DOI: 10.1109/CVPR.2018.00378]
- Huang K Q, Ren W Q and Tan T N. 2014. A review on image object classification and detection. *Chinese Journal of Computers*, 37(6): 1225-1240 (黄凯奇, 任伟强, 谭铁牛. 2014. 图像物体分类与检测算法综述. *计算机学报*, 37(6): 1225-1240) [DOI: 10.3724/SP.J.1016.2014.01225]
- 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
- Jampani V, Kiefel M and Gehler P V. 2016. Learning sparse high dimensional filters: image filtering, dense CRFs and bilateral neural networks//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA; IEEE: 4452-4461 [DOI: 10.1109/CVPR.2016.482]
- 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 (ICCV). Venice, Italy; IEEE: 863-872 [DOI: 10.1109/ICCV.2017.99]
- Li X, Wang M Y, Wen C C, Wang L J, Zhou N and Fang Y. 2019. Density-aware convolutional networks with context encoding for airborne LiDAR point cloud classification [EB/OL]. [2020-08-10]. <https://arxiv.org/pdf/1910.05909.pdf>
- 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]. [2020-08-13]. <https://arxiv.org/pdf/1801.07791.pdf>
- Liu Y C, Fan B, Xiang S M and Pan C H. 2019. Relation-shape convolutional neural network for point cloud analysis//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA; IEEE: 8887-8896 [DOI: 10.1109/CVPR.2019.00910]
- 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 (IROS). Hamburg, Germany; IEEE: 922-928 [DOI: 10.1109/IROS.2015.7353481]
- Noh H, Hong S and Han B. 2015. Learning deconvolution network for semantic segmentation//Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile; IEEE: 1520-1528 [DOI: 10.1109/ICCV.2015.178]
- Parzen E. 1962. On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, 33(3): 1065-1076 [DOI: 10.1214/aoms/1177704472]
- Qi C R, Liu W, Wu C X, Su H and Guibas L J. 2018. Frustum PointNets for 3D object detection from RGB-D data//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA; IEEE: 918-927 [DOI: 10.1109/CVPR.2018.00102]
- 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 (CVPR). Las Vegas, USA; IEEE: 5648-5656 [DOI: 10.1109/CVPR.2016.609]
- Qi C R, Su H, Kaichun M and Guibas L J. 2017a. PointNet: deep learning on point sets for 3D classification and segmentation//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA; IEEE: 77-85 [DOI: 10.1109/CVPR.2017.16]
- 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]. [2020-07-25]. <https://arxiv.org/pdf/1706.02413.pdf>
- Riegler G, Ulusoy A O and Geiger A. 2017. OctNet: learning deep 3D representations at high resolutions//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA; IEEE: 6620-6629 [DOI: 10.1109/CVPR.2017.701]
- Shen Y R, Feng C, Yang Y Q and Tian D. 2018. Mining point cloud local structures by kernel correlation and graph pooling//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA; IEEE: 4548-4557 [DOI: 10.

- 1109/CVPR. 2018. 00478]
- Su H, Jampani V, Sun D Q, Maji S, Kalogerakis E, Yang M H and Kautz J. 2018. SplatNet: sparse lattice networks for point cloud processing//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 2530-2539 [DOI: 10.1109/CVPR.2018.00268]
- 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 (ICCV). Santiago, Chile: IEEE: 945-953 [DOI: 10.1109/ICCV.2015.114]
- Wang X L, Girshick R, Gupta A and He K M. 2018. Non-local neural networks//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 7794-7803 [DOI: 10.1109/CVPR.2018.00813]
- Wang Y, Sun Y B, Liu Z W, Sarma S E, Bronstein M M and Solomon J M. 2019. Dynamic graph CNN for learning on point clouds. ACM Transactions on Graphics, 38(5): #146
- 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 W X, Qi Z G and Li F X. 2019. PointConv: deep convolutional networks on 3D point clouds//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE: 9613-9622 [DOI: 10.1109/CVPR.2019.00985]
- 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 (CVPR). Boston, USA: IEEE: 1912-1920 [DOI: 10.1109/CVPR.2015.7298801]
- Xu Y F, Fan T Q, Xu M Y, Zeng L and Qiao Y. 2018. SpiderCNN: deep learning on point sets with parameterized convolutional filters [EB/OL]. [2020-08-06]. <https://arxiv.org/pdf/1803.11527.pdf>
- 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]
- Yuan X C, Chen H W and Li Y W. 2017. Normal estimation method for regular point cloud surface with sharp feature. Journal of Image and Graphics, 22(3): 334-341 (袁小翠, 陈华伟, 李彧雯. 2017. 规则特征曲面点云法向估计. 中国图象图形学报, 22(3): 334-341) [DOI: 10.11834/jig.20170307]
- Zhu Y K, Mottaghi R, Kolve E, Lim J J, Gupta A, Li F F and Farhadi A. 2017. Target-driven visual navigation in indoor scenes using deep reinforcement learning//Proceedings of 2017 IEEE International Conference on Robotics and Automation (ICRA). Singapore, Singapore: IEEE: 3357-3364 [DOI: 10.1109/ICRA.2017.7989381]

作者简介



邱云飞, 1976年生, 男, 教授, 主要研究方向为数据挖掘、机器学习、智能数据处理。
E-mail: 7415575@qq.com



朱梦影, 通信作者, 女, 硕士研究生, 主要研究方向为点云数据处理。
E-mail: 3066964438@qq.com