

基于可变形图卷积的点云表征学习

李宗民, 张玉鹏, 刘玉杰, 李华

引用本文

李宗民, 张玉鹏, 刘玉杰, 李华. [基于可变形图卷积的点云表征学习](#)[J]. 计算机科学, 2022, 49(8): 273-278.

LI Zong-min, ZHANG Yu-peng, LIU Yu-jie, LI Hua. [Deformable Graph Convolutional Networks Based Point Cloud Representation Learning](#)[J]. Computer Science, 2022, 49(8): 273-278.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[三维激光雷达点云空间多特征分割](#)

Spatial Multi-feature Segmentation of 3D Lidar Point Cloud

计算机科学, 2022, 49(8): 143-149. <https://doi.org/10.11896/jsjcx.210300275>

[基于图卷积神经网络的文本分类方法研究综述](#)

Review of Text Classification Methods Based on Graph Convolutional Network

计算机科学, 2022, 49(8): 205-216. <https://doi.org/10.11896/jsjcx.210800064>

[面向图像分类的小样本学习算法综述](#)

Survey on Few-shot Learning Algorithms for Image Classification

计算机科学, 2022, 49(5): 1-9. <https://doi.org/10.11896/jsjcx.210500128>

[一种基于遮罩的稀疏点云滤波算法](#)

Sparse Point Cloud Filtering Algorithm Based on Mask

计算机科学, 2022, 49(5): 25-32. <https://doi.org/10.11896/jsjcx.210600129>

[基于用户关联的立场检测](#)

Stance Detection Based on User Connection

计算机科学, 2022, 49(5): 221-226. <https://doi.org/10.11896/jsjcx.210400135>

基于可变形图卷积的点云表征学习

李宗民¹ 张玉鹏¹ 刘玉杰¹ 李 华²

1 中国石油大学(华东)计算机科学与技术学院 山东 青岛 266580

2 中国科学院计算技术研究所 北京 100190

(lizongmin@upc.edu.cn)

摘 要 虽然深层神经网络较为成功地解决了点云数据稀疏和不规则等问题,但是点云局部特征的学习仍然是一个非常具有挑战性的问题。现有的用于点云表征学习的网络存在点与点之间相互独立提取特征的问题,为此提出了一种全新的空域图卷积。首先,在构造图结构时提出了自适应空洞 K 近邻算法,以最大程度地捕获局部拓扑结构信息;其次,在卷积中加入了卷积核每条边与感受野图之间的角度特征,保证了更具鉴别力的特征提取;最后,为了充分利用局部特征,提出了一种全新的图金字塔池化,以更好地融合多尺度特征。在标准公开数据集 ModelNet40 和 ShapeNet 上测试该算法,分别取得了 93.2% 与 86.5% 的准确度。实验结果表明,该算法在点云表征学习中处于领先水平。

关键词: 图卷积神经网络;表征学习;点云;局部特征

中图法分类号 TP391.41

Deformable Graph Convolutional Networks Based Point Cloud Representation Learning

LI Zong-min¹, ZHANG Yu-peng¹, LIU Yu-jie¹ and LI Hua²

1 College of Computer Science and Technology, China University of Petroleum, Qingdao, Shandong 266580, China

2 Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China

Abstract Although the sparseness and irregularity of point cloud data have been successfully solved by deep neural networks. However, how to learn the local features of point clouds is still a challenging problem. Existing networks for point cloud representation learning have the problem of extracting features independently between points and points. To this end, a new spatial graph convolution is proposed. Firstly, an adaptive hole K -nearest neighbor algorithm is proposed when constructing the graph structure to maximize local topological structure information. Secondly, the angle feature between each edge of the convolution kernel and the receptive field map is added to the convolution, which ensures more discriminative feature extraction. Finally, in order to make full use of local features, a novel graph pyramid pooling is proposed. This algorithm is tested on the standard public data sets ModelNet40 and ShapeNet, and the accuracy is 93.2% and 86.5% respectively. Experimental results show that the proposed algorithm is at a leading level in point cloud representation learning.

Keywords Graph neural convolutional networks, Representation learning, Point clouds, Local feature

近年来,深度学习在点云^[1]上的应用吸引了大量研究者的注意。点云被广泛应用于各种三维应用程序中,如虚拟现实、无人机和无人驾驶车辆等。随着科技的发展,激光雷达等点云采集设备的性能越来越好且价格亲民,点云数据的获取越来越简便,直接面向点云数据的表征学习成为了研究热点。然而,三维点云数据的局部特征学习仍处于缓慢发展的阶段,这也是点云表征学习需要解决的首要问题。

目前的主流网络^[2-5]主要运用卷积神经网络来提取点云特征,忽略了目标物体本身的拓扑结构。Qi 等提出的 PointNet^[2]是第一个直接处理点云数据的网络,在点云特征学习与

形状分析中都取得了显著的成就。但是 PointNet^[2]在特征提取时点与点之间是相互独立的,忽略了局部点的几何关系,因此无法获取到点云局部细粒度特征,这就是 PointNet^[2]在复杂点云场景下表征学习能力差的原因。为了解决这个问题,Qi 等又提出了 PointNet++^[3]对点云局部划分进行多尺度的特征提取,但是对于点的特征提取依旧是相互独立的。如何更好地学习点云局部特征是三维视觉发展的关键,因为这可以使三维视觉应用有更详细和更精确的表示。

近年来,图神经网络因其强大的建模能力引起了研究者浓厚的兴趣^[6]。传统的卷积神经网络只能处理欧氏空间的

到稿日期:2021-09-03 返修日期:2022-03-24

基金项目:国家重点研发计划(2019YFF0301800);国家自然科学基金(61379106);山东省自然科学基金(ZR2013FM036, ZR2015FM011)

This work was supported by the National Key R & D Program(2019YFF0301800), National Natural Science Foundation of China(61379106) and Shandong Provincial Natural Science Foundation(ZR2013FM036, ZR2015FM011).

通信作者:张玉鹏(S19070005@s.upc.edu.cn)

数据,点云是一种典型的非欧数据,因此采用图神经网络提取点云特征是一种必然趋势。现有的基于图卷积的点云表征学习方法多采用 K 近邻算法构造图结构,然后进行特征提取。DGCNN^[7]采用动态图卷积即在每层图卷积后重新采用 K 近邻算法构造图。虽然动态图卷积可以缓解过平滑以及增大感受野,但是只能捕获到有限的边缘特征,缺少多尺度信息,最终也损失了部分局部特征。

针对上述问题,本文提出了一种新的深度学习模型用于处理和学习三维点云特征信息。该算法采用自适应空洞 K 近邻算法构造图,以更好地捕获局部细节特征;采用了一种全新的基于空域的图卷积算子,把卷积核也设置成图,这个卷积核图的形状以及权重是可以学习的。卷积算子与 3D-GCN^[8]类似,但是,3D-GCN^[8] 仅仅在数值大小上考虑了角度特征。本文对角度特征进行了筛选,在加入角度选择模块的同时融合了低层语义特征,这样可以更好地学习点云的节点特征,保持局部细节特征的提取。为了让图卷积更好地学习到目标物体的拓扑结构以及多尺度的特征,全新的图金字塔池化。实验结果表明,本文方法在点云分类分割上取得了领先的水平。

1 相关工作

1.1 用于三维点云识别的深度学习

作为人工智能主导的技术,深度学习在解决计算机视觉问题上取得了较大的成功。由于点云的稀疏性以及不规则性,深度学习在三维点云上的应用还处于刚刚起步的阶段。作为一项开创性的工作,PointNet^[2] 直接从点云中提取全局特征,但缺乏提取局部特征的能力。局部细节特征在许多三维视觉任务中起着重要的作用,因此,如何更好地提取局部特征成为了点云表征学习的主流任务。PointNet++^[3] 是 PointNet^[2] 在分层提取特征上的扩展,但是 PointNet++^[3] 与 PointNet^[2] 一样,仍然独立地考虑每个点,也就是说,他们并没有考虑点与点之间的局部关系。Zhao 等^[9] 提出了一个自适应特征调整模块,它可以在每个成对的点之间交换信息,大大增强了学习的逐点特征的表示能力。PATs^[10] 提出两个模块来捕捉点与点之间的关系以及学习层次特征,从而更好地学习局部特征。这些方法都是使用多个多层感知机 (MLPs) 对每个点建模,然后使用对称函数聚合全局特征。也有一些方法把点云转换成别的表示形式来进行特征提取,比如,之前的许多工作^[11-14] 将点特征寄存在体素网格中,或者直接使用体素化后的三维形状作为模型输入,这样就可以对这些输入数据进行标准的三维卷积,但是这类方法的计算量非常大。为了解决这个问题,Yan 等^[15] 提出了稀疏卷积,以提高训练和推理的速度,但是这种方法对于局部信息的提取仍然存在不足。

1.2 基于图的三维点云数据处理网络

图卷积神经网络是图神经网络中应用得较多的一种方法。图卷积神经网络主要分为两类方法,分别是基于谱域的方法和基于空间的方法^[16]。图卷积神经网络的工作主要分为两部分,一部分基于频域,另一部分基于空域。DGCNN (Dynamic Graph CNN for Learning on Point Clouds)^[7] 运用

边缘卷积运算动态地在每一层中构建图来提取局部几何特征。LDGCNN^[17] 去掉了 DGCNN^[7] 中的空间变换网络,它通过特征相连的方法来提升模型的精度以及降低模型的复杂度。ECC (Dynamic Edge-Conditioned Filters in Convolutional Neural Networks on Graphs)^[18] 的卷积操作和常规的二维图像卷积操作都是一种加权平均操作,它们的不同之处在于 ECC 可以作用在任何图结构上,并且其权重由节点间的边权决定。然后,DeepGCNs^[19] 在利用残差连接解决过拟合以及梯度消失问题的同时,发现利用残差连接可以对空间卷积层进行堆叠。KCNet^[20] 将点集内核定义为一组可学习的点,它们根据由核相关性测量的几何关系共同响应一组相邻数据点。3D-GCN^[8] 采用了可变形的卷积核,但是对于角度特征的选择仅仅考虑了数值上的最大值,忽略了当角度特征为负数时的最小值。Grid-GCN^[21] 通过贪心算法对体素化后的点云构造图结构,扩大了局部感受野,提高了采样效率。VA-GCN^[22] 考虑了 K 近邻算法中心点与其邻域之间向量的仰角和方位角关系,并以此构建注意力矩阵,然后采用双通道结构融合加权边缘特征和全局特征。基于图的点云处理网络虽然相比其他网络可以更好地获取局部特征,但是对于多尺度信息的学习依然存在困难。

2 图卷积方法

图卷积神经网络主要包括卷积算子和池化算子的构建。卷积算子的主要作用是学习节点的局部特征;池化算子的主要目的是学习结构化的表示,降低参数。当解决节点级别的任务时,研究者更专注于让每个节点学习到更好的特征,因此对池化算子并没有过多的关注。本文不仅对卷积算子进行了定义创新,同时对池化算子也进行了全新的设计。

本文提出了一种新的图卷积来提取点云特征,整体框架如图 1 所示。首先通过提出的自适应空洞 K 近邻算法构造图结构;然后进行点云局部语义特征的聚合和全局几何特征的传播;最终得到分类分割结果。

为了更简洁方便地介绍所提方法,先对所用的符号进行简单的定义。对于每个用点云描述的物体,我们采用一个点云的集合表示, $P = \{p_n | n = 1, 2, \dots, N\}$ 一共包含了 N 个点, p_n 表示这个集合中的第 n 个点,表示形式为一个三维坐标 (x_n, y_n, z_n) 、法向量 (v_n^x, v_n^y, v_n^z) 或者 RGB 颜色信息 (R_n, G_n, B_n) 。本文采用的模型仅仅学习点云的坐标信息特征,对于一个点云物体,我们用一个大小为 $N \times 3$ 的矩阵来表示。

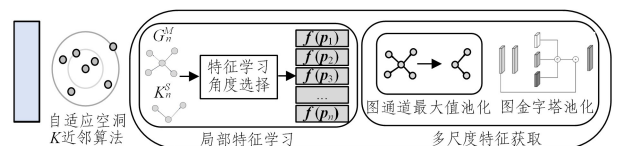


图 1 本文方法的整体框架

Fig. 1 Overall framework of our method

对于点云分类任务来说,需要处理输入的点云数据,给出模型预测的每个类别的分数 C 。对于点云分割任务来说,我们需要对输入的每一个点云进行类别预测。与分类任务不同,模型输出的大小为 $N \times C$ 。

2.1 自适应空洞 K 近邻算法

本文方法是直接对 N 个无序的点进行处理, $P = \{p_n | n = 1, 2, \dots, N\}$, 每个点 $p_n \in R^3$ 包含 3 个坐标特征。我们采用 $f(p) \in R^D$ 来表示每个点的 D 维特征向量, 对于每个点, 采用自适应空洞 K 近邻算法使得构造的图结构能获得更多的局部特征信息。通过自适应空洞 K 近邻算法扩展的作用域来提升描述符的语义级别, 避免在邻域聚集过多的点以及减少特征冗余信息, 这样有助于对局部更广泛的区域的几何理解以及节省计算成本。

自适应空洞 K 近邻算法是对 K 近邻算的改进, 首先得到它的 K 个邻居节点, 计算出中心点到第 K 个邻居节点的欧氏距离 L , 分别以 L 和 ωL 为半径进行球形查询, 如图 2 所示。 ω 是在训练过程中随机给定的, $\omega \in (1, 1.5)$ 。假定小球内点的数量为 k_L , 大球与小球之间的数量为 $k_{\omega L}$, 则采样的点分别为:

$$\begin{cases} N_L = \left\lceil \frac{k_L}{k_L + k_{\omega L}} K \right\rceil \\ N_{\omega L} = \left\lceil \frac{k_{\omega L}}{k_L + k_{\omega L}} K \right\rceil \end{cases} \quad (1)$$

然后将 $N_{\omega L} + N_L$ 个点构建图。我们把每个点构成的图定义为 G_n^M 。

$$G_n^M = \{p_n, p_m | \forall p_m \in K(p_n, M)\} \quad (2)$$

其中, $K(p_n, M)$ 表示通过自适应空洞 K 近邻算法得到的邻居节点。

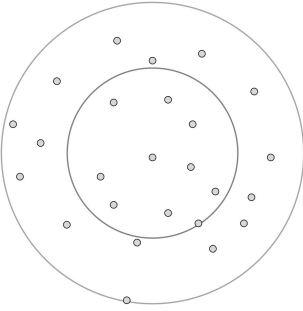


图 2 自适应空洞 K 近邻算法

Fig. 2 Adaptive hole K -nearest neighbor algorithm

在自适应空洞 K 近邻算法构建图的过程中, 记录每个邻居节点与其方向向量 $d_{m,n} = p_m - p_n$, 如图 3 所示。在定义的卷积方式中, 也增加了对方向向量的特征学习, 保证了特征学习的几何不变性, 学习到了更具鉴别力的特征。

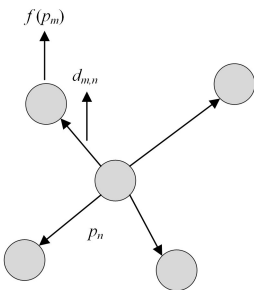


图 3 节点的特征及方向向量

Fig. 3 Node feature and direction vector

2.2 卷积算子设计

在二维卷积中, 卷积核对图像进行对应位置相乘求和, 如图 4 所示。但是点云是典型的非欧数据, 普通的二维卷积并不能在点云上适用。因此, 我们借鉴二维卷积操作, 提出了一种新的图卷积方式, 直接在点云上进行卷积。把卷积核创造性地定义为一个图的形式, 以对应 2.1 节构建出的图, 利用图形状的卷积核在图上进行相乘求和, 完美地解决了点云的无序性以及不规则性。卷积核的形状以及大小在训练过程中是可学习的。

$$\begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 5 & 4 & 3 & 2 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 5 & 4 & 3 & 2 & 1 \\ 0 & 1 & 2 & 3 & 4 \end{bmatrix} * \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 7 & 8 & 9 \\ 8 & 7 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

图 4 二维卷积

Fig. 4 Two-dimensional convolution

对卷积核进行类似的定义, 如图 5 所示。本文把卷积核定义为 K_n^S 。

$$K_n^S = \{k_n | \forall n \in (0, S)\} \quad (3)$$

其中, K_0^S 表示卷积核的中心点, $K_0^S = (0, 0, 0)$ 对应了 $K(p_n, M)$ 中的 p_0 ; S 表示卷积核的点的数量; K_n^S 包含了 $S+1$ 个点。

在卷积核中, 点的数量一开始可以由我们定义, 后续的形状是可学习的。关于卷积核的参数定义, 为每个点定义一个特征向量 $w(k) \in R^D$ 。因此, 对于特征 $f(p)$, 用相应的权重来实现卷积操作。由于卷积核节点的数量以及相应的方向可能与感受野图的节点数量以及方向不同, 因此定义卷积核节点的方向向量 $k_n = k_n - k_0$ 来描述学习到的卷积核, 同时用于角度选择模块。最后, 卷积核可以表示为 $\{w(k_0), (k, w(k_s)) | s = 1, 2, 3, \dots, S\}$, 其中每个元素数值都是在训练过程中学习得到的, 这样图卷积核的形状也会根据数值发生变换。

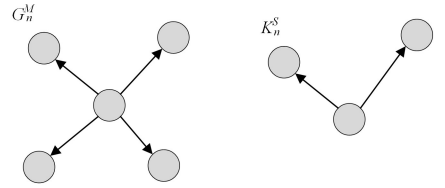


图 5 图卷积核与感受野图

Fig. 5 Image convolution kernel and receptive field map

在二维卷积中, 卷积操作可以看作卷积核与图片的相似度计算, 模型输出值较高, 表示具有高相似度。卷积算子的设计与 3D-GCN^[8] 类似, 3D-GCN^[8] 仅仅在数值上考虑了角度特征的大小, 忽略了角度值为负值的情况, 当负值绝对值大于正值时, 此时的负值记录了最大差异性, 但是此时 3D-GCN^[8] 依旧将正值作为最大差异性特征。因此, 本文加入了角度特征选择模块对角度特征进行筛选, 以确保选择的角特征可以表示卷积核与感受野图之间的差异性。对于每个点, 采用自适应 K 近邻算法得到邻居节点, 然后将其构建图, 把每个点构成的图定义为 G_n^M 。卷积核点定义的特征向量 $w(k) \in R^D$ 与特征 $f(p)$ 用相应的相似度计算来实现卷积操作。卷积核的方向向量 $k_s = k_s - k_0$ 。最后, 卷积核表示为 $\{w(k_0), (k, w(k_s)) | s = 1, 2, 3, \dots, S\}$, 把卷积算子定义为 $Conv(G_n^M, K_n^S)$ 。

来计算相应的相似度。

$$\text{Conv}(G_n^M, K_m^S) = (f(p_n), w(k_m)) + \lambda * \text{rot}(p_n, k_m) \quad (4)$$

其中, $(\cdot)$ 表示内积运算, $\forall n \in (0, M), \forall m \in (0, S); \lambda$ 是一个超参数; $\text{rot}(\cdot)$ 是特征选择模块, 它的定义如下:

$$\text{rot}(p_h, k_m) = \sum_{m=0}^S \max \begin{cases} -(f(p_n), w(k_m)) * \left(\frac{(d_{h,n}, k_m)}{\|d_{h,n}\| \|k_m\|} \right), & (d_{h,n}, k_m) < 0 \\ (f(p_n), w(k_m)) * \left(\frac{(d_{h,n}, k_m)}{\|d_{h,n}\| \|k_m\|} \right), & (d_{h,n}, k_m) > 0 \end{cases} \quad (5)$$

其中, $\forall h \in (0, M), \forall m \in (0, S), \text{rot}(\cdot)$ 记录了每个 k_m 与感受野图之间的最大差异性。因此, 角度选择模块保证模型学习到有效的特征信息。

我们记录了每个 k_m 在卷积时与 G_n^M 中每条边最大的角度差异性, 因此在卷积过程中, 可以记录卷积核的每条边与构建的图之间的角度特征信息, 增加了对几何变换的鲁棒性。

一个图卷积层包含一个提前定义好数量 L 的卷积核 K_n^S , 默认 $L=1$, 或者是上一层经过学习的卷积核, $K_n^S = \{k_n | \forall n \in (0, S)\}$ 。对于三维点云的输入 $P = \{p_n | n=1, 2, \dots, N\}$, $p_n \in \mathbb{R}^3$ 以及相应的特征输入 $f(p)^m \in \mathbb{R}^D$, 经过卷积核的操作, 最后输出 $f(p)^{\text{out}} \in \mathbb{R}^{N \times D'}$ 。因此, 点云经过卷积层可以表示为:

$$\text{GraphConv}((P, f(p)^{\text{in}}), K_n^S) = (P, f(p)^{\text{out}}) \quad (6)$$

2.3 图金字塔池化

PointNet^[2] 等方法仅仅归纳出了全局特征, 缺少对点云局部特征的描述。为了能更好地学习点云的局部特征, 我们增强了池化方式, 对输入的 G_n^M 进行通道维度上的最大值池化, 如图 1 中的图通道最大值池化部分所示, 然后采用一个超参数 γ 来进行点云数目的下采样。借鉴金字塔池化^[23] 的

方法, 提出了一种全新的用于点云的图金字塔池化层, 用多窗口对不同尺度的点云进行下采样以及提取特征, 然后进行特征融合。此过程可以定义为:

$$(P, f(p)) = \text{GraphConv}(\varphi(\text{GraphPool}_{\gamma_1}, \text{GraphPool}_{\gamma_2}, \text{GraphPool}_{\gamma_3})) \quad (7)$$

其中, P 和 $f(p)$ 表示输入的点以及对应的特征, φ 表示对不同池化层得到的特征进行聚合, $\gamma_{1,2,3}$ 表示不同的采样率。

如图 6 所示, 输入点云以及相应的特征, 通过金字塔池化对不同尺度规模的点进行汇聚, 得到多尺度的局部特征。将每个点的特征更新为金字塔池化的聚合特征, 同时与进行图金字塔池化的上一层特征进行对应元素相加, 增强局部细粒度特征, 如图 1 中的金字塔池化部分所示。

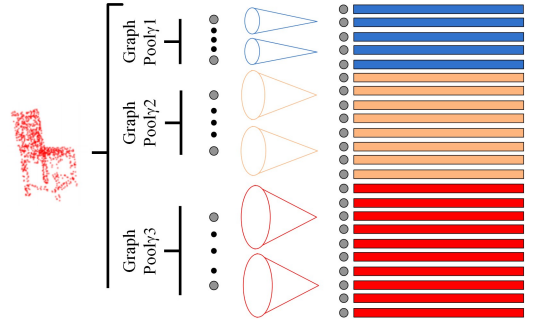


图 6 图金字塔池化

Fig. 6 Graph pyramid pooling

2.4 模型细节

将本文模型应用到点云分类分割任务中, 通过自适应空洞 K 近邻算法构建图结构。在分类任务中, 通过图卷积算子学习感知特征, 采用图金字塔池化下进行采样, 经过特征学习后, 通过多层感知机得到分类结果。在分割任务中, 通过跳层连接对感知特征的多级表示进行连续上采样以生成每个点的预测。其中, 分类分割任务都以端到端的方式进行训练, 具体细节如图 7 所示。

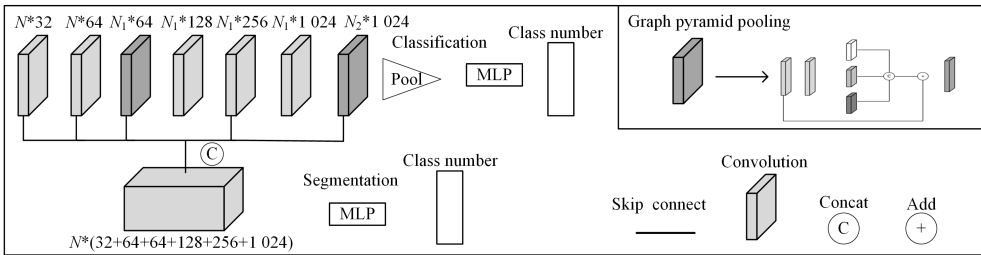


图 7 网络架构图

Fig. 7 Network architecture diagram

3 实验及结果分析

本文在 ModelNet40^[24] 和 ShapeNet^[25] 数据集上验证本文方法。ModelNet40^[24] 含有 40 个类别, 其中包括 9843 个训练样本以及 2468 个测试样本。ShapeNet^[25] 由 16 个类别的 16881 个 CAD 模型组成。特别地, ShapeNet^[25] 标定了 50 个部件标签, 每个类别包含 2~5 个部件标签。本文使用的硬件设备为 4×P100 显卡, 每个显卡有 16GB 显存。与已有方法进行了全面的比较分析, 实验结果表明, 本文方法在 Model-

Net40^[24] 上取得了领先的水平, 在 ShapeNet^[25] 上取得了最高精度。

3.1 目标分类

3.1.1 实验细节

对数据集的 CAD 模型进行随机采样生成 1024 个点。在训练过程中, 将点云沿 z 轴随机旋转, 在 $[0.7, 1.2]$ 尺度范围内进行随机缩放, 使用均值为 0、标准差为 0.02 的高斯噪声实现数据扩增。采用 mish 激活函数解决梯度下降的不平滑问题。对于分类任务, 采用图卷积进行特征提取, 卷积核

数量依次为(32,64);而后接一个图金字塔池化层,对生成的特征进行拼接;然后进行特征学习,卷积核数量依次为(128,256,1 024),后面是两层 MLP;为防止过拟合,在第一层 MLP 后添加 dropout 层,并设置 $dropout=0.4$ 。网络使用 Adam 优化器,学习率为 0.001,批处理参数为 32。若无特别说明,自适应空洞 K 近邻算法中的 K 默认为 30,卷积核点的数量 S 默认为 1,多尺度采样率 $\gamma_1=4,\gamma_2=8,\gamma_3=16$ 。

3.1.2 实验结果

表 1 列出了本文方法与已有方法在 ModelNet40 上的对比结果,可以看出,在输入类型只使用点云的情况下,本文方法比 PointNet 高 4%,比 PointNet++ 高 2.5%,比 DGCNN 高 1%,比 3D-GCN 高 1.1%,取得了最高的精度。由此证明了本文提出的几个模块在学习局部细节特征时的有效性。为了验证各模块的有效性,本文进行了消融实验,表 2 列出了实验结果。

表 1 不同方法在 ModelNet40 上的形状分类结果对比
Table 1 Comparison of shape classification results of different methods on ModelNet40

方法	输入类型	准确率/%
ECC ^[18]	点云	87.4
PointNet ^[2]	点云	89.2
PointNet++ ^[3]	点云	90.7
KCNet ^[20]	点云	91.0
Kd-Net ^[26]	点云	90.6
DGCNN ^[7]	点云	92.2
3D-GCN ^[8]	点云	92.1
本文方法	点云	93.2

表 2 消融实验
Table 2 Ablation experiment

角度选择	自适应空洞 K 近邻算法	图金字塔	准确率/%
—	—	—	85.2
✓	✓	—	92.4
✓	—	✓	92.9
✓	✓	✓	93.2

3.2 部件分割

部件分割是一项具有挑战性的三维细粒度识别任务。给定目标物体后,为每个点分配部件标签。目前已有的方法基本上是将部件分割问题转化为按点分类问题,对每个点云预测相应的类别,然后完成部件分割任务,本文的分割方法在分类网络的基础上进行,分割的特征由不同尺度的层输出连接而成,经过三层共享 MLP 进行逐点预测。其中自适应空洞 K 近邻算法中的 K 取值为 60,其余实验细节与分类一致。

表 3 列出了本文方法与 PointNet^[2],PointNet++^[3],CLCSCANet^[27],DGCNN^[7],3D-GCN^[8]方法的比较结果。本文采用平均交并比(Mean Intersection over Union,mIoU)作为评估指标,分别计算了类别和实例中的两种平均交并比。实验结果表明,本文方法比 3D-GCN^[8]的类 mIoU 和实例 mIoU 分别高出 1%和 1.4%,实现了最佳性能。这进一步说明了本文方法可以很好地学习局部细粒度特征,保证最有鉴别力特征的提取。图 8 给出了本文方法的分割实例可视化结果。

表 3 不同方法在 ShapeNet 上的部件分割结果对比
Table 3 Comparison of parts segmentation results of different methods on ShapeNet

方法	(单位:%)	
	类平均交并比	实例平均交并比
PointNet ^[2]	80.4	83.7
PointNet++ ^[3]	81.9	85.1
CLCSCANet ^[27]	82.7	85.3
DGCNN ^[7]	82.3	85.2
3D-GCN ^[8]	82.2	85.1
本文方法	83.2	86.5

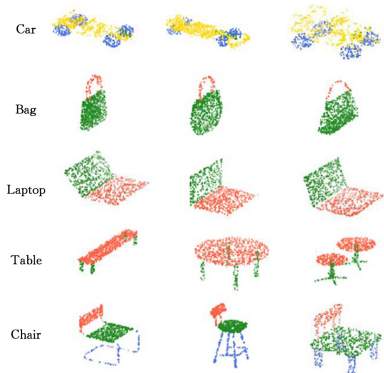


图 8 可视化结果
Fig. 8 Visualize results

3.3 补充实验

为了验证图金字塔中的采样率大小对模型结果的影响,本文设置了对比实验,结果如表 4 所列,对采样率采用不同的组合进行实验,结果发现采样率为 4,8,16 时效果最好。同时也比较了去掉图金字塔池化时的实验结果,表 4 所列的实验结果可以证明,图金字塔池化对精度有所提升。同时,对自适应空洞 K 近邻算法进行实验,验证结果表明当 K 取 30 时,效果最佳,实验结果如表 5 所列。

表 4 在 ModelNet40 上不同采样率下的结果对比
Table 4 Results comparison of different sampling rates on ModelNet40

ModelNet40	
$\gamma_1\gamma_2\gamma_3$	准确率/%
—	92.3
2,4,8	92.9
4,8,12	93.1
4,8,16	93.2
8,16,32	92.4

表 5 在 ModelNet40 上不同 K 值的结果对比
Table 5 Results comparison of different K values on ModelNet40

K	准确率/%
15	92.1
20	92.5
25	92.6
30	93.2
35	92.7

结束语 针对现有方法对点云数据进行表征学习时缺少局部细节特征以及多尺度特征这一问题,本文提出了一种新的图卷积方法,设计新的卷积算子与池化算子来学习更具鉴别力的特征。同时,针对图卷积描述符作用域小,邻域信息冗余的问题,提出了自适应空洞 K 近邻算法来构建图,对局部

几何信息进行更广泛的感知,从而增强局部细节特征。实验结果表明,本文方法在标准分类分割数据集上取得了领先水平,证明了所提方法的有效性。本文方法对角度以及位置特征进行了选择学习,下一步工作的重点是加入邻域内的重心等其他几何信息,并与语义特征信息融合,以进一步提升分类分割的准确度。

参 考 文 献

[1] GUO Y, WANG H, HU Q, et al. Deep Learning for 3D Point Clouds: A Survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 43(12): 4338-4364.

[2] QI C R, SU H, MO K, et al. Pointnet: Deep learning on point sets for 3d classification and segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 652-660.

[3] QI C R, YI L, SU H, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space[EB/OL]. <https://arxiv.org/pdf/1706.02413.pdf>.

[4] LI Y, BU R, SUN M, et al. B-PointCNN: Con-volution on X-transformed points[C]// In Conference and Work-shop on Neural Information Processing Systems(NeurIPS). 2018: 820-830.

[5] LI R, LI X, HENG P A, et al. Pointaugment: an au-to-augmentation framework for point cloud classifica-tion[C]// Proceedings of the IEEE/CVF Conference on Com-puter Vision and Pattern Recognition. 2020: 6378-6387.

[6] JIE Z A, GC A, SH A, et al. Graph neural networks: A review of methods and applications[J]. AI Open, 2020, 1: 57-81.

[7] WANG Y, SUN Y, LIU Z, et al. Dynamic graph cnn for learning on point clouds [J]. ACM Transactions On Graphics (tog), 2019, 38(5): 1-12.

[8] LIN Z H, HUANG S Y, WANG Y C F. Convolution in the cloud: Learning deformable kernels in 3D graph convolution networks for point cloud analysis[C]// Proceedings of the IEEE/ CVF Conference on Computer Vision and Pattern Recognition. 2020: 1800-1809.

[9] ZHAO H, JIANG L, FU C W, et al. PointWeb: Enhancing local neighborhood features for point cloud processing[C]// CVPR. 2019: 5565-5573.

[10] YANG J, ZHANG Q, NI 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). IEEE, 2019: 3323-3332.

[11] CHOY C B, XU D F, GWAK J Y, et al. 3D-R2N2: A Unified Approach for Single and Multi-view 3D Object Reconstruction [C]// Proceedings of the European Conference on Computer Vision(ECCV). 2016: 628-644.

[12] MATURANA D, SCHERER S. VoxNet: A 3D Convolutional Neural Network for Real-time Object Recognition[C]// Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS). IEEE, 2015: 922-928.

[13] MENG H Y, GAO L, LAI Y K, et al. VV-Net: Voxel VAE Net with Group Convolutions for Point Cloud Segmentation[J]. arXiv:1811.04337, 2018.

[14] ROYNARD X, DESCHAUD J E, GOULETTE F. Classification of Point Cloud Scenes with Multiscale Voxel Deep Network[J]. arXiv:1804.03583, 2018.

[15] YAN Y, MAO Y, LI B. Second: Sparsely embedded convolu-

tional detection[J/OL]. Sensors, 2018; 18(10): 3337. <https://doi.org/10.3390/s18103337>.

[16] TE G, HU W, ZHENG A, et al. Rgcnn: Regularized graph cnn for point cloud segmentation[C]// Proceedings of the 26th ACM international conference on Multimedia. 2018: 746-754.

[17] ZHANG K, HAO M, WANG J, et al. Linked dynamic graph cnn: Learning on point cloud via linking hierarchical features [J]. arXiv:1904.10014, 2019.

[18] SIMONOVSKY M, KOMODAKIS N. Dynamic edge-conditioned filters in convolutional neural networks on graphs[C]// CVPR. 2017: 3693-3702.

[19] LI G, MÜLLER M, THABET A, et al. DeepGCNs: Can GCNs Go as Deep as CNNs? [C]// ICCV. 2019: 9267-9276.

[20] SHEN Y, FENG C, YANG Y, et al. Mining point cloud local struc-tures by kernel correlation and graph pooling[C]// Pro-ceedings of the IEEE Conference on Computer Vision and Pat-tern Recognition. 2018: 4548-4557.

[21] XU Q, SUN X, WU C Y, et al. Grid-gcn for fast and scalable point cloud learning[C]// Proceedings of the IEEE/CVF Con-ference on Computer Vision and Pattern Recognition. 2020: 5661-5670.

[22] HU H, WANG F, LE H. VA-GCN: A Vector Attention Graph Convolution Network for learning on Point Clouds[J/OL]. <https://arxiv.org/pdf/2106.00227.pdf>.

[23] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.

[24] SHILANE P, MIN P, KAZHDAN M, et al. The princeton shape benchmark[C] // Proceedings Shape Modeling Applications, 2004. IEEE, 2004: 167-178.

[25] YI L, KIM V G, CEYLAN D, et al. A scalable active framework for region annotation in 3d shape collections[J]. ACM Transac-tions on Graphics (ToG), 2016, 35(6): 1-12.

[26] KLOKOV R, LEMPITSKY V. Escape from cells: Deep kd-net-works for the recognition of 3d point cloud models[C]// 2017 IEEE International Conference on Computer Vision (ICCV). 2017: 863-872.

[27] HAN X F, HE Z Y, CHEN J, et al. Cross-Level Cross-Scale Cross-Attention Network for Point Cloud Representation[EB/ OL]. <https://arxiv.org/pdf/2104.13053.pdf>.



LI Zong-min, born in 1965, Ph.D, pro-fessor, is a member of China Computer Federation. His main research interests include computer graphics, picture pro-cessing and scientific computing visuali-zation.



ZHANG Yu-peng, born in 1997, post-graduate. His main research interests include point cloud representation learning, graph neural network and geo-metric invariance.