

文章编号:1671-6833(2024)02-0033-09

基于图游走和图注意力的点云分类与分割

李文举¹, 姬倩倩¹, 沙利业², 储王慧¹, 崔柳¹

(1. 上海应用技术大学 计算机科学与信息工程学院, 上海 201418; 2. 上海普利森配料系统有限公司, 上海 201108)

摘要: 针对点云特征提取中远距离特征和局部几何结构信息欠缺的问题, 提出了一种基于图游走和图注意力的点云分类与分割网络。首先, 利用带有导向性的图游走算法, 对点云全局特征补充额外的几何信息和远距离特征信息; 其次, 嵌入图注意力机制, 使模型聚焦于点云的关键区域, 提升网络的特征提取能力; 最后, 在初始点云中提取距离特征作为初始残差嵌入到网络中, 避免网络过平滑。在 ModelNet40 数据集、ScanObjectNN 数据集进行了点云分类实验, 在 ShapeNetPart 数据集与 Toronto-3D 数据集上分别进行了点云部件分割与点云语义分割实验, 实验结果表明: 相较于基准网络 DGCNN, 分类精度分别提升了 1.3 百分点、5.6 百分点; 分割精度分别提升了 1.2 百分点、33.1 百分点。通过在 ModelNet40-C 数据集上进行稳健性分析, 验证了所提网络具有较强的鲁棒性。

关键词: 点云分类; 点云分割; 图神经网络; 图游走; 图注意力机制

中图分类号: TP391.41; TP183

文献标志码: A

doi: 10.13705/j.issn.1671-6833.2024.02.006

随着三维采集技术的飞速发展, 点云分类与分割任务作为三维场景理解中重要的基础性问题, 被广泛地应用在自动驾驶、机器人导航、医学成像、室内场景分析等实际场景中^[1-3]。

由于点云具有无序性、稀疏性及非结构化特性, 使用传统的神经网络模型处理点云时需要将不规则的点云形式转换为高度规则的输入格式, 例如使用投影、体素格划分等方法, 这种间接的转换过程不仅会造成数据的几何信息丢失, 还会增加额外的计算量, 因此, 后续的模型主要使用直接基于点云的方法^[4-5]。

PointNet^[6] 网络直接基于原始点云进行特征处理, 该网络通过多层感知机 (multi-layer perceptron, MLP) 学习单个点的特征并用对称函数编码全局信息, 解决了点云的无序性与置换不变性等问题, 缺点在于只捕捉到单个点与全局点的特征信息, 忽略了局部特征信息。针对其缺点改进的 PointNet++^[7] 模型利用最远点采样和球查询的方法构建局部邻域以提取局部特征, 在一定程度上缓解了上述问题, 但忽略了点云邻域关系和空间特征信息, 不利于获取语

义信息。

为了更有效地分析点云这类不规则数据, 越来越多的研究开始使用图神经网络 (graph neural network, GNN) 将图学习应用到点云数据上。在图结构中, 每个节点都含有丰富的特征信息, 边代表节点之间的依赖关系。基于图神经网络的点云处理方法可以将点云中的每个点及其邻接点构建成图结构, 进而直接对点云进行卷积, 可以充分利用点云的结构特点, 有效减少空间几何信息的损失。DGCNN^[8] 将图神经网络应用到点云处理的过程中, 使用 K 近邻 (K -NN) 算法构建每个点的局部邻域, 并使用边缘卷积算子替换了 PointNet 网络中叠加的多层感知机, 实现了在保持置换不变性的基础上捕捉局部几何信息, 但由于忽略了点云的几何形状信息导致丢失了部分有效特征。DeepGCNs^[9] 集成残差连接、密集连接和空洞卷积构建一个深层次的 GCN (graph convolution network) 模型。该网络在不损失分辨率的情况下捕获到了较大的感受野, 实现了较高性能的点云语义分割, 但由于网络层次较深, 会造成过多的内存开销。SNEDA^[10] 模型基于图注意力方法学

收稿日期: 2023-09-13; 修订日期: 2023-11-02

基金项目: 国家自然科学基金资助项目 (61903256, 61973307); 上海市生物医药科技支撑专项项目 (22S31903900)

作者简介: 李文举 (1964—), 男, 辽宁营口人, 上海应用技术大学教授, 博士, 主要从事计算机视觉、模式识别、智能检测方面的研究, E-mail: wjli@sit.edu.cn。

通信作者: 崔柳 (1984—), 女, 辽宁锦州人, 上海应用技术大学讲师, 博士, 主要从事制导导航与控制、微传感器和生物医学信号处理的研究, E-mail: cui.liu8475@msn.com。

引用本文: 李文举, 姬倩倩, 沙利业, 等. 基于图游走和图注意力的点云分类与分割[J]. 郑州大学学报(工学版), 2024, 45(2): 33-41. (LI W J, JI Q Q, SHA L Y, et al. Point cloud classification and segmentation based on graph walk and graph attention [J]. Journal of Zhengzhou University (Engineering Science), 2024, 45(2): 33-41.)

习符号网络节点嵌入,有效捕获了不同邻居节点对中心节点的贡献和影响。GACNet^[11]根据不同相邻点和特征通道的空间位置以及特征差异为每个邻接点分配注意力权值,充分学习和聚合了局部结构信息,但忽略了远程点信息和空间几何信息。GAPointNet^[12]将自注意力机制与 GCN 结合,提供了一种基于多头注意力机制的点云分割的网络,但其对点云的泛化能力较差。PointMLP^[13]是一种仅使用残差 MLP 模块进行点云分析的模型,避免了由复杂的局部特征提取引起的高额计算开销,但分割精度有待提升。

针对上述问题,本文提出了一种基于图游走与图注意力的点云分类与分割网络。首先,通过构造点云同构图,进而在图上进行有导向的游走,聚集出一系列固定数量和长度的游走曲线,以此捕获远距离的点云特征和额外的几何信息。其次,结合图注意力算法自适应地学习图拓扑结构来聚合邻域特征,从而实现不同节点间有区分度的特征表示。最后,对初始点云提取距离特征作为初始残差进行残差连接,利用连接后的特征进行后续的分类与分割任务。本文的主要贡献如下。

(1)通过引入图游走技术,解决了特征提取过程中 3D 点云的几何特征和长距离特征匮乏的问题;对游走得到的节点使用图注意力机制,增强了对重要节点的特征提取能力,实现了更高效地聚合邻域信息来更新节点特征。

(2)设计了一种嵌入初始残差的三维点云分类与分割网络,通过在初始点云中提取距离特征计算初始残差,解决了网络图卷积层数过多时出现的过平滑的问题。

(3)在合成数据集和真实世界数据集上对点云分类、分割,进行了多次实验,结果验证了所提网络模型的有效性。

1 模型介绍

基于图游走和图注意力的点云分类与分割网络模型主要由多个 GWA(graph walk and attention)层构成,GWA 层由最远点下采样模块(farthest point sampling, FPS)、初始残差模块、图游走模块和图注意力模块组成。其中,最远点下采样模块获取非局部特征和缩放图模型;初始残差模块主要用于获取特征表达更丰富的初始残差以缓解网络的过平滑问题;图游走模块主要用于提取点的远距离特征和几何形状特征来得到丰富的上下文信息;图注意力模块主要用来获取节点间的局部邻域特征;最后通过

最大池化层和平均池化层来获取点云的全局特征,并使用全连接层获取用于预测的特征向量。

1.1 初始残差模块

当图神经网络堆叠多层时,网络会出现过平滑问题使得性能急剧下降。过平滑问题是指经过多次图卷积操作后节点的特征表示变得无法区分现象。因此,大多数图神经网络只有 2~3 层的浅层架构,这限制了它们的表达能力和对深层邻域的利用。为了解决这一问题,受 GCNII^[14]的启发,在网络的特征融合过程中嵌入了初始残差模块,如图 1 所示。

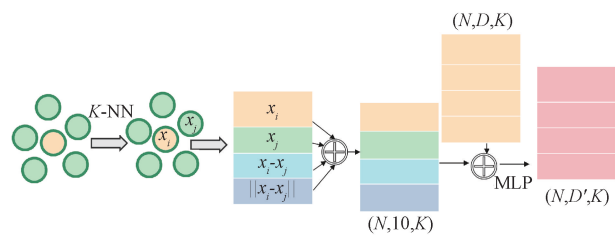


图 1 初始残差模块

Figure 1 Initial residual module

初始残差的连接可以确保在网络中堆叠了多个图卷积层的情况下,每个节点的最终表示仍保留输入层的部分特征信息,可缓解节点特征趋于一致的过平滑问题。

计算初始残差的步骤如下。

步骤 1 将原始点云 P 中的每个点 p_i 看作中心节点,利用 K -NN 算法寻找离中心节点最近的 K 个邻居节点 p_j 来构建局部邻域。

步骤 2 计算邻居节点与中心节点的距离特征,融合邻居节点、中心节点的空间信息及距离信息得到初始残差 p ,数学描述为

$$p = p_i \oplus p_j \oplus (p_j - p_i) \oplus \|p_j - p_i\|。 \quad (1)$$

式中: p_i 为中心节点; p_j 为邻居节点; $\| \cdot \|$ 表示中心节点和邻居节点间的欧式距离; $(p_j - p_i)$ 表示两者的特征差异。

步骤 3 将初始残差 p 和当前特征 f 融合以更好地聚集上下文信息,计算公式为

$$f' = \max(\text{MLP}(p \oplus f))。 \quad (2)$$

式中: \oplus 表示向量的拼接操作;MLP 表示多层感知机。

1.2 图游走模块

由于离散点云对象缺乏足够的三维几何图形的形状描述符,受 CurveNet^[15]的启发,引入一种图游走方法来捕获点云中的几何信息和远距离特征,即对原始点云建立同构图,通过在图上进行有导向的游走采样出不同的游走曲线,然后对曲线进行聚合以增强其点云中逐个点的特征。

1.2.1 图构建

图游走模块的输入为一个具有 N 个点的三维点云集,即每个点的初始化特征为三维的空间坐标信息,表示为 $P = \{p_1, p_2, \dots, p_N\} \in \mathbf{R}^3$ 。构建点云同构图 $G = (V, E)$, 其中 $V = \{1, 2, \dots, N\}$ 为图中的顶点集, $E = (e_{ij})_{i,j=1}^N \subseteq |V| \times |V|$ 为边集, $e_{ij} \in E$ 。 $F = \{f_1, f_2, \dots, f_N\} \in \mathbf{R}^N$ 为输入顶点的特征。顶点 i 的邻接点集为 $N(i) = \{j\} \cup \{i\}$ 。

1.2.2 分组

(1) 选择起始点。采用 TOP- K 选择策略,通过 MLP 学习每个点的特征,并使用 sigmoid 门控得到每个点的得分,得分最高的 M 个点作为游走的起始点 $S = \{s_1, s_2, \dots, s_M\}$, 由此引出 M 条游走曲线。

(2) 构建游走策略。在得到起始点集后,对每个节点 s_i 使用式(3)中的可学习的 $\pi(\cdot)$ 策略找到下一个要游走的点 s_{i+1} 。即对 M 个起始点计算 K 近邻,并使用 MLP 学习 K 个邻接点的特征分布,为了便于梯度回传,使用 gumbel-softmax 进行归一化并计算出每个点的概率得分,选出概率最大的点作为下一个要游走的点,数学表达如式(4)所示。对预定义的 $\pi(\cdot)$, 通过反复执行等式(3) L 次,对曲线进行分组。游走曲线如图2所示。

$$s_{i+1} = \pi(s_i); \quad (3)$$

$$\pi(s) = \sum_{j=1}^k \text{gumbel-softmax}(\text{MLP}(h_s^j)) \cdot N_s. \quad (4)$$

式中: h_s^j 为节点的 K 近邻 s^j 的状态描述符,由特征 s_i 和曲线描述参数 r_i 拼接后得到。为了避免曲线出现交叉和圆圈的情况,采用动态编码和交叉抑制对 h_s^j 进行处理。

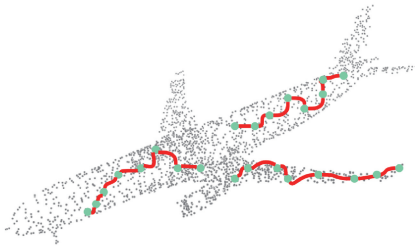


图2 游走曲线示意图

Figure 2 Diagram of the walking curve

1.2.3 聚合

为了考虑游走曲线之间和曲线内部的联系,分别沿曲线的个数维度和曲线内部点的数量维度聚合出曲线间特征向量 f_{inter} 和曲线内特征向量 f_{intra} , 并利用不同的多层感知机进行特征提取,通过矩阵乘法的形式,建立一个映射关系,学习曲线上点与点之间的联系,然后与 softmax 计算出的分数相融合,得到2个更具有细粒度表示的特征向量,并沿着特征

轴将 f'_{inter} 和 f'_{intra} 拼接起来再进行特征融合,最终聚合得到各个点更新后的特征。具体细节如图3所示。

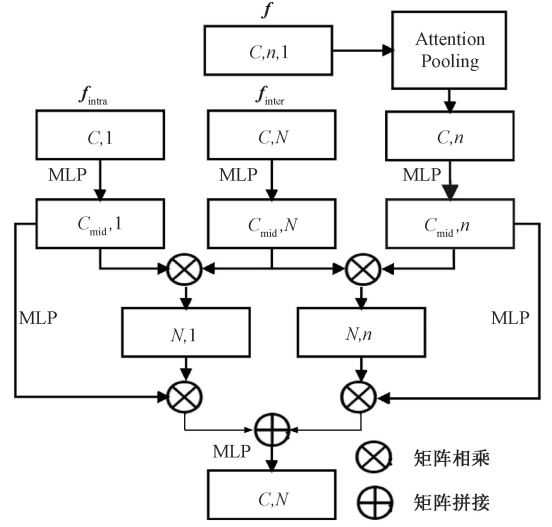


图3 游走特征聚合模块

Figure 3 Walk feature aggregation module

图游走机制的伪代码如下。

输入: 节点特征 F , 节点集合 V ;

输出: 游走后各节点更新后的特征向量 f 。

- ① for each $V_i \in V$:
 $x = \text{torch.sigmoid}(\text{self.MLP}(V_i))$;
 $\text{start} = \text{torch.topk}(x, \text{curve_num})$;
 end for
- ② 由 start 找到游走起始点 $S = \{s_1, s_2, \dots, s_M\}$;
- ③ $s^j = \text{KNN}(s, k)$; /* 找到 s 节点的 K 近邻 */
- ④ $h_s^j = \text{torch.matmul}(s^j, \text{curve_num})$ /* 得到 s 节点的 K 近邻状态描述符 */
- ⑤ $x = \text{gumbel-softmax}(\text{self.MLP}(h_s^j))$; /* 计算每个点的得分 */
- ⑥ 与 $N(s)$ 相乘后求和加得到下一个游走点 s_{i+1} ;
- ⑦ $\text{curves.append}(s_i)$; /* 得到曲线特征 */
- ⑧ for each $x_i \in \text{curves}$:
 $f = \text{self.attpool}(x_i)$;
 $f'_{\text{inter}}, f'_{\text{intra}} \leftarrow \text{Conv}(f_{\text{inter}}, f_{\text{intra}}, f)$;
 end for
- ⑨ $f = \text{torch.concat}(f'_{\text{inter}}, f'_{\text{intra}})$;
- ⑩ 结束。

1.3 图注意力模块

由于点云属于非规则化数据,之前大多数方法往往使用固定的核函数来更新节点的特征信息,忽略了属于同一对象的点之间的结构联系和不同邻居节点对中心节点的重要程度不同等问题。图注意力可以动态学习不同节点的特征,选择性地关注其中

2 实验

2.1 实验环境及数据集

本文在 ModelNet40 数据集^[16]和 ScanObjectNN 数据集^[17]上进行点云分类实验,在 ShapeNetPart 数据集^[18]上进行点云部件分割实验,在 Toronto-3D 数据集^[19]上进行点云语义分割实验,以此验证所提出方法的有效性。实验在如下环境下进行:显卡型号 NVIDIA RTX2080Ti,Pytorch 版本 1.10.0,python 版本 3.8,CUDA 版本 11.3。

2.1.1 点云分类数据集

ModelNet40 是一个含有 40 个类别的 12 311 个网格 CAD 模型的数据集,其中 9 843 个模型用于训练,2 468 个模型用于测试。对于每个模型,按照与 PointNet 中相同的实验设置,从网格面上均匀采样 1 024 个点作为网络的输入,并使用随机缩放和扰动来进行数据增强。*batchsize* 设置为 48,其中,32 个用于训练,16 个用于测试,*epoch* 为 300,使用 SGD 优化器优化模型,动量系数设置为 0.9。初始学习率设置为 0.1,并使用余弦退火衰减策略来动态调整学习率。

ModelNet40 数据集虽然作为点云分析的规范基准,但在训练过程中有着极大的不稳定性和随机性,故本文还在 ScanObjectNN 数据集上进行实验,该数据集包含了背景、噪声和遮挡的干扰,故在该基准上测试可以得到一个更稳健且公平的实验结果。该数据集包含 15 000 个对象,分为 15 个类别和 2 902 个独特的实例对象。除 *epoch* 数量为 200 外,其他实验设置与 ModelNet40 上的设置相同。

2.1.2 点云部件分割数据集

ShapeNetPart 数据集是一个用于细粒度形状识别的点云部件分割数据集,包含 16 881 个预先对齐的点云形状,可以分为 16 个形状类和总共 50 个分割类。形状类包含飞机、包、车、椅子、耳机等,每个形状类被划分为 5 个以内的分割类,如飞机形状的点云会被划分为机身、机翼、尾翼、发动机这 4 个零件分割类。在训练和测试时遵循和 PointNet 一样的分割方案,将 12 137 个模型用作训练样本,其余的用作验证和测试,从每个模型中统一采样 2 048 个点作为网络的输入。实验设置与 ScanObjectNN 数据集上的参数设置一致。

2.1.3 点云语义分割数据集

Toronto-3D 数据集是由车载 MLS 系统采集的大型城市室外点云数据集,由大约 7.83×10^7 个点组成,该数据集被手动标记为 8 个类别,包括道路、路

标、树木、建筑、电线、电线杆、汽车和围栏。参数设置与部件分割时的设置一致。

2.2 实验结果与分析

2.2.1 分类结果分析

本文在 ModelNet40 数据集上的测试结果如表 1 所示。由表 1 可以看出,分类的总体精度 *OA* 达到了 94.2%,平均分类精度 *mAcc* 达到了 91.6%,与基准网络 DGCNN 相比,本文算法的 GWA 层融合了远距离特征和几何关系特征,得到了更为丰富的语义信息,使得总体分类精度提高了 1.3 个百分点,平均分类精度提高了 1.4 个百分点。相比于近期提出的 SimpleView 和 Diffconv 网络,本文网络的分类性能也具有显著优势。

表 1 不同模型在 ModelNet40 数据集上的实验结果
Table 1 Experimental results of different models on ModelNet40 dataset

模型	<i>mAcc</i>	<i>OA</i>
PointNet ^[6]	86.2	89.2
PointNet++ ^[7]	—	90.7
DGCNN ^[8]	90.2	92.9
PointASNL ^[20]	—	93.2
SimpleView ^[21]	—	93.6
DiffConv ^[22]	90.6	93.6
本文模型	91.2	94.0
本文模型*	91.6	94.2

注:* 代表使用了投票机制。

本文在真实数据集 ScanObjectNN 上的实验结果如表 2 所示,将数据划分为 3 组进行测试,分别为有背景干扰(OBJ_BG)、无背景干扰(OBJ_ONLY)及带有数据扰动(T50RS)的对象。其中,T50RS 代表边界框被平移(translate)、旋转(rotate)、缩放(scale)了 50%。采用基于 DGCNN 架构的基准模型,共有 4 个 GWA 卷积块,输出维度为 64、128、256、512,邻居个数为 24,*epoch* 为 250。

表 2 不同模型在 ScanObjectNN 数据集上的实验结果
Table 2 Experimental results of different models on ScanObjectNN dataset

模型	<i>OA</i> /%		
	OBJ_ONLY	OBJ_BG	T50RS
PointNet ^[6]	79.2	73.3	68.2
PointNet++ ^[7]	84.3	82.3	77.9
DGCNN ^[8]	86.2	82.8	78.1
CurveNet ^[15]	84.3	84.4	—
DiffConv ^[22]	86.6	84.9	—
本文模型	86.8	87.7	83.7

由表 2 可以看出,与现有方法相比,本文在无背景干扰、有背景干扰及数据扰动情况下的分类效果都优于其他网络,总体分类精度分别达到了 86.8%、87.7%、83.7%,在数据扰动的情况下,较基准网络 DGCNN 提升了 5.6 百分点,取得了良好的分类效果。

2.2.2 部件分割结果分析

在 ShapeNetPart 数据集上进行点云部件分割实验,本文实验结果和近年来其他模型的实验结果如表 3 所示。由表 3 中数据可以计算出本文实验、基准模型 DGCNN 及经典模型 PointNet++ 的平均交并比 $mIoU$ 分别为 86.4%、85.2% 和 85.1%,相比之下,本文方法的 $mIoU$ 较 DGCNN 和 PointNet++ 分别提高了 1.2 百分点、1.3 百分点。如表 3 所示,本文模型在飞机、车和吉他等 8 个类别上的分割性能最好,

表 3 不同模型在 ShapeNetPart 数据集上的实验结果
Table 3 Experimental results of different models on ShapeNetPart dataset

模型	交并比 $IoU/\%$				本文模型
	PointNet++ ^[7]	DGCNN ^[8]	GCN3D ^[23]	SADGCNN ^[24]	
飞机	82.4	84.0	83.8	83.3	85.8
包	79.0	83.4	78.8	81.1	81.9
帽子	87.7	86.7	83.2	81.5	82.2
车	77.3	77.8	77.5	76.7	80.8
椅子	90.8	90.6	90.8	91.0	91.5
耳机	71.8	74.7	75.8	77.8	76.7
吉他	91.0	91.2	90.8	89.3	92.3
刀	85.9	87.5	86.0	88.8	87.4
灯	83.7	82.8	83.8	85.0	84.9
电脑	95.3	95.7	95.2	95.4	96.0
摩托	71.6	66.3	66.4	62.6	75.9
杯子	94.1	94.9	93.4	93.6	94.7
手枪	81.3	81.1	81.3	80.8	83.4
火箭	58.7	63.5	51.2	65.3	61.1
滑板	76.4	74.5	67.1	74.3	78.2
桌子	82.6	82.6	83.0	85.7	83.1

可见本文方法对细粒度较高的语义类别具有更好的分割性能,充分验证了本文模型在点云部件分割任务中具有较强的竞争力。

为了更好地展示本文方法在点云部件分割上的效果,如图 7 所示,对物体部件分割进行可视化,可以看出本文网络能够非常清晰、准确地分割出物体的各个部件。将本文网络与基准网络 DGCNN 及真实标签值(ground truth,GT)进行可视化对比分析,如图 8 所示,从左到右分别是对同一个包、帽子、刀和杯子的部件分割效果图,标记的黑框处代表分割时存在点的语义标签分割混乱及错误的情况。可以看出,与基准网络 DGCNN 相比,本文网络语义标签边界处的分割效果有了大幅度提升,基本接近真实标签的分割效果。

2.2.3 语义分割结果分析

表 4 为本文方法在 Toronto-3D 数据集上与其他经典方法的语义分割定量结果对比。由表 4 中数据可以计算出本文算法分割的 $mIoU$ 为 82.7%,DGCNN 的 $mIoU$ 为 49.6%,较 DGCNN 提升了 33.1 百分点。分割效果较基准网络 DGCNN 有了显著提升,并且在路标、建筑、电线、电线杆、汽车等 5 个类别上也取得了最好的分割结果。由于道路、围栏和树木的细节信息较少,本文方法对其分割效果没有显著提高,但仍取得了与近年来其他网络相当的分割效果。

2.3 消融实验

为了进一步说明本文引入的图游走与图注意力模块的有效性,在 ModelNet40 数据集与 ShapeNet-Part 数据集上进行了消融实验,结果如表 5 所示。其中,为了避免图卷积层过多导致的网络过平滑现象,所有网络模型均在融入初始残差特征的基础上进行。

由表 5 可以看出,以 DGCNN 网络为基准网络模型添加图游走模块后,即对点云特征补充几何语

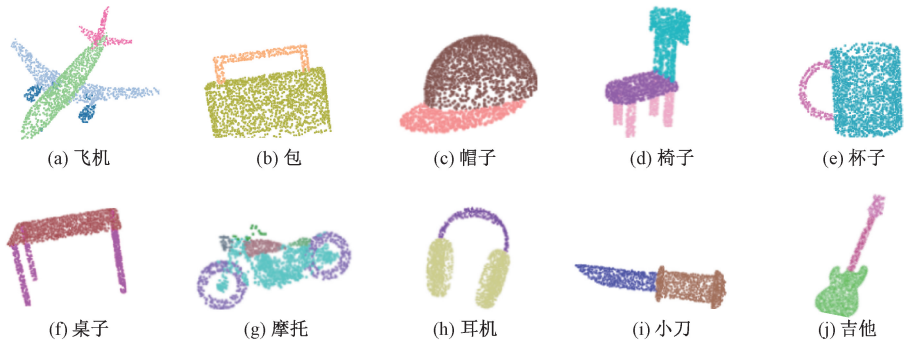


图 7 分割效果可视化

Figure 7 Visualization of segmentation effect

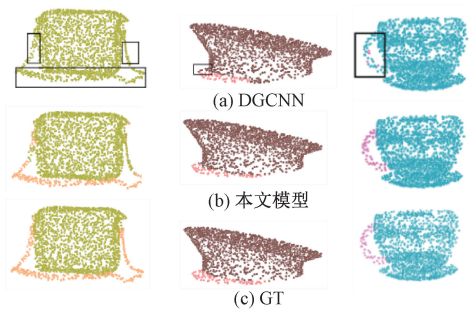


图 8 分割结果对比图

Figure 8 Visualization of segmentation effect

表 4 不同模型在 Toronto-3D 数据集上的实验结果

Table 4 Experimental results of different models on the Toronto-3D dataset

模型	IoU/%				
	DGCNN ^[8]	MS-TGNet ^[19]	PCT ^[25]	DiffConv ^[22]	本文
道路	90.6	90.9	79.8	83.3	83.1
路标	0.4	18.8	59.5	51.1	59.8
树木	81.3	92.2	75.8	69.0	85.3
建筑	64.0	80.6	84.3	79.6	86.9
电线	47.1	69.4	77.8	80.5	81.9
电线杆	56.9	71.2	82.0	84.4	88.2
汽车	49.3	51.1	79.5	76.2	80.6
围栏	7.3	13.6	95.9	89.8	95.7

表 5 不同模块的消融实验

Table 5 Ablation experiments for different modules %

模型	模块	<i>mIoU</i>	<i>mAcc</i>	<i>OA</i>
		PartSeg	M40	M40
DGCNN	基准网络	85.2	90.2	92.9
Model ₁	基准网络+	86.0	91.4	93.4
	初始残差+			
Model ₂	图游走	85.9	91.2	93.8
	基准网络+			
Model ₃	初始残差+	86.4	91.6	94.2
	图注意力			

义特征和长距离特征,使点云分类的平均准确率、总体准确率和分割的平均交并比都得到了不同程度的提升,分别提升了 1.2 百分点、0.5 百分点和 0.8 百分点;添加图注意力模块后,即仅选择性地强调重要特征信息,使得网络在 *OA* 指标上有 0.9 百分点的提升,在 *mIoU* 指标上有 0.7 百分点的提升;在同时加入图游走模块和图注意力模块后,网络模型的表现更加突出,与基准网络模型相比,其 *mIoU* 提升 1.2 百分点, *OA* 提升 1.3 百分点, *mAcc* 提升 1.4 百分点,优于消融实验中的其他各个网络。

由上述模块间的消融实验可知,本文网络对点云数据具有较好的学习能力,对点云几何特征和长距离特征与重要点特征都实现了有效补充与增强,对点云分类、分割任务起到了关键作用,有效地提升了网络的性能。

2.4 不同最近邻数对比实验

本文提出的图注意力模块在提取点云局部特征的过程中采用 *K*-NN 算法搜集 *K* 个近邻点构建局部有向图,近邻点的个数 *K* 在一定程度上会影响网络提取到的局部特征。现有的主流网络模型通常将 *K* 值设置为 16、24、32、40,本文分别在 ModelNet40 数据集与 ShapeNetPart 数据集上进行了对比实验,测试了不同 *K* 值对模型性能的影响,如表 6 所示。

表 6 不同 *K* 值的对比实验

Table 6 Comparative experiments with different *K* %

<i>K</i>	<i>mIoU</i>	<i>mAcc</i>	<i>OA</i>
	ShapeNetPart	ModelNet40	ModelNet40
16	85.9	90.6	93.5
24	86.1	91.6	94.2
32	86.4	91.1	93.9
40	85.8	90.7	93.5

从表 6 可以看出,网络的分类性能在 *K* 为 24 时表现最为优秀,分割精度在 *K* 为 32 时最高。当 *K* 值较小时,构建的邻域范围过小会使模型的感受野有限,导致模型不能全面地学习到点云的几何特征,分类精度与分割精度都偏低;当 *K* 值过大时,近邻点过多会使不同邻域内的点特征信息重叠冗余,分类、分割精度均有所下降,因此,选择合适的 *K* 值会使模型性能表现更为优异。

2.5 鲁棒性实验

鉴于 3D 点云数据的收集在现实生活中存在不确定性,如遮挡情况下只能获取部分可见的点云。故本文采用带有场景扰动的 ModelNet40-C^[26] 失真点云数据集,对所提出的点云识别模型进行全面的鲁棒性测试,测试结果如表 7 所示,其中损坏错误率 *CER* 越低代表模型的鲁棒性越好,可以看出本文模型的 *CER* 较低,取得了较为稳健的结果,模型鲁棒性较好。

表 7 不同模型在 ModelNet40-C 数据集上的实验结果

Table 7 Experimental results of different models on ModelNet40-C dataset

模型	<i>CER</i> /%
PointNet ^[6]	28.3
PointNet++ ^[7]	23.6
DGCNN ^[8]	25.9
DiffConv ^[22]	21.4
本文模型	22.5

3 结论

为了提高深度学习网络直接对点云数据分类与分割的性能,本文提出了一种基于图游走与图注意力的点云分类与分割网络。该网络模型采用图游走机制来提升对点云的远距离特征和几何结构特征的表达能力,并在特征提取过程中使用图注意力机制,使模型聚焦于重要的特征通道和点云区域,增强了关键点的特征表示。通过在公共数据集上进行点云分类与分割对比实验,取得了较好的精度效果。同时研究了在失真点云数据上的点云分类效果,实验结果表明,网络具有较好的鲁棒性。由于图特征收集使得模型的计算量及参数量有所上升,今后的工作重心将放在如何使模型的参数量减少的同时不损失或提升精度,进而使网络可以更高效地处理更复杂的点云任务。

参考文献:

[1] 崔建明, 蔺繁荣, 张迪, 等. 基于有向图的强化学习自动驾驶轨迹预测[J]. 郑州大学学报(工学版), 2023, 44(5): 53-61.
CUI J M, LIN F R, ZHANG D, et al. Reinforcement learning autonomous driving trajectory prediction based on directed graph[J]. Journal of Zhengzhou University (Engineering Science), 2023, 44(5): 53-61.

[2] BELLO S A, YU S S, WANG C, et al. Review: deep learning on 3D point clouds[J]. Remote Sensing, 2020, 12(11): 1729.

[3] 何雄辉, 谭杰夫, 刘哲, 等. 基于稀疏点云分割的适应视角变化的场景识别方法[J]. 计算机科学, 2023, 50(1): 87-97.
HE X H, TAN J F, LIU Z, et al. Viewpoint-tolerant scene recognition based on segmentation of sparse point cloud[J]. Computer Science, 2023, 50(1): 87-97.

[4] GUO Y L, WANG H Y, HU Q Y, et al. Deep learning for 3D point clouds: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(12): 4338-4364.

[5] 王文曦, 李乐林. 深度学习在点云分类中的研究综述[J]. 计算机工程与应用, 2022, 58(1): 26-40.
WANG W X, LI L L. Review of deep learning in point cloud classification[J]. Computer Engineering and Applications, 2022, 58(1): 26-40.

[6] QI C R, SU H, M 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. Piscataway: IEEE, 2017: 652-660.

[7] QI C R, YI L, SU H, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space [C] // Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM, 2017: 5105-5114.

[8] WANG Y, SUN Y B, LIU Z W, et al. Dynamic graph CNN for learning on point clouds[J]. ACM Transactions on Graphics, 2019, 38(5): 1-12.

[9] LI G H, MÜLLER M, THABET A, et al. DeepGCNs: can GCNs go as deep as CNNs? [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2019: 9266-9275.

[10] 逯泽锟, 于千城, 王晓峰, 等. 基于双重注意力机制的符号网络节点嵌入[J]. 郑州大学学报(工学版), 2023, 44(2): 68-74.
LU Z K, YU Q C, WANG X F, et al. Learning signed network node embedding via dual attention mechanism [J]. Journal of Zhengzhou University (Engineering Science), 2023, 44(2): 68-74.

[11] WANG L, HUANG Y C, HOU Y L, et al. Graph attention convolution for point cloud semantic segmentation [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2019: 10288-10297.

[12] CHEN C, FRAGONARA L Z, TSOURDOS A. GAPointNet: graph attention based point neural network for exploiting local feature of point cloud[J]. Neurocomputing, 2021, 438: 122-132.

[13] MA X, QIN C, YOU H X, et al. Rethinking network design and local geometry in point cloud: a simple residual MLP framework[EB/OL]. (2022-02-15) [2023-07-12]. <https://arxiv.org/abs/2202.07123>. pdf.

[14] CHEN M, WEI Z W, HUANG Z F, et al. Simple and deep graph convolutional networks [C] // Proceedings of the 37th International Conference on Machine Learning. New York: ACM, 2020: 1725-1735.

[15] XIANG T G, ZHANG C Y, SONG Y, et al. Walk in the cloud: learning curves for point clouds shape analysis [C] // 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2021: 895-904.

[16] WU Z R, SONG S R, KHOSLA A, et al. 3D ShapeNets: a deep representation for volumetric shapes [C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2015: 1912-1920.

[17] UY M A, PHAM Q H, HUA B S, et al. Revisiting point cloud classification: a new benchmark dataset and classification model on real-world data [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2019: 1588-1597.

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

[19] TAN W K, QIN N N, MA L F, et al. Toronto-3D: a large-scale mobile LiDAR dataset for semantic segmentation of urban roadways[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE, 2020: 797-806.

[20] YAN X, ZHENG C D, LI Z, et al. PointASNL: robust point clouds processing using nonlocal neural networks with adaptive sampling[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020: 5588-5597.

[21] GOYAL A, LAW H, LIU B W, et al. Revisiting point cloud shape classification with a simple and effective baseline[EB/OL]. (2021-01-09) [2023-07-12]. <https://arxiv.org/abs/2106.05304>. pdf.

[22] LIN M X, FERAGEN A. DiffConv: analyzing irregular point clouds with an irregular view[C]//European Conference on Computer Vision. Cham: Springer, 2022: 380-397.

[23] 兰红, 陈浩, 张蒲芬. 集图卷积和三维方向卷积的点云分类分割模型[J]. 计算机工程与应用, 2023, 59 (8): 182-191.

LAN H, CHEN H, ZHANG P F. Point cloud classification and segmentation model based on graph convolution and 3D direction convolution[J]. Computer Engineering and Applications, 2023, 59(8): 182-191.

[24] 沈露, 杨家志, 周国清, 等. 集自注意力与边卷积的点云分类分割模型[J]. 计算机工程与应用, 2023, 59 (19): 106-113.

SHEN L, YANG J Z, ZHOU G Q, et al. Point cloud classification segmentation model based on self-attention and edge convolution[J]. Computer Engineering and Applications, 2023, 59(19): 106-113.

[25] GUO M H, CAI J X, LIU Z N, et al. PCT: point cloud transformer[J]. Computational Visual Media, 2021, 7 (2): 187-199.

[26] SUN J C, ZHANG Q Z, KAILKHURA B, et al. Benchmarking robustness of 3D point cloud recognition against common corruptions[EB/OL]. (2022-01-28) [2023-07-12]. <https://arxiv.org/abs/2201.12296>. pdf.

Point Cloud Classification and Segmentation Based on Graph Walk and Graph Attention

LI Wenju¹, JI Qianqian¹, SHA Liye², CHU Wanghui¹, CUI Liu¹

(1. School of Computer Science and Information Engineering, Shanghai Institute of Technology, Shanghai 201418, China; 2. Shanghai Precision Dosing & Weighing System Co., Ltd., Shanghai 201108, China)

Abstract: Aiming at the shortage of distance feature and local geometric structure information in feature extraction, a point cloud classification and segmentation network based on graph walk and graph attention was proposed. Firstly, a guided graph walk algorithm was used to supplement additional geometric information and remote feature information to the whole feature of point cloud. Secondly, the graph attention mechanism was embedded to make the model on the key areas of the point cloud and improve the feature extraction ability of the network. Finally, distance features were extracted from the initial point cloud and embedded into the network as initial residuals to avoid oversmoothing. Point cloud classification experiments were carried out on ModelNet40 dataset and ScanObjectNN dataset, and point cloud component segmentation and point cloud semantic segmentation experiments were carried out on ShapeNetPart dataset and Toronto-3D dataset, respectively. The experiment results showed that, compared with the benchmark network DGCNN, classification accuracy increased by 1.3 percentages and 5.6 percentages, respectively; The segmentation accuracy was improved by 1.2 percentages and 33.1 percentages respectively. Through the robust analysis on ModelNet40-C dataset, it was proved that the proposed network had strong robustness.

Keywords: point cloud classification; point cloud segmentation; graph neural network; graph walk; graph attention mechanism