

文章编号 1004-924X(2024)18-2823-13

结合双注意力和加权动态图卷积的 三维点云分类与分割

肖 剑, 王晓红*, 李 炜, 杨祎斐, 罗 季
(贵州大学 矿业学院, 贵州 贵阳 550025)

摘要:针对基于深度学习的点云分类分割网络在局部上下文信息提取和近邻点特征表达上的不足,以及最大池化容易丢失次优信息的问题,提出结合双注意力和加权动态图卷积网络的点云分类分割算法。首先,加权动态图卷积利用加权K近邻算法构建鲁棒的局部结构,并引入强化边卷积模块对点特征加权以得到强化后边特征。然后,通道注意力构造通道相关性并释放各通道潜力,再利用空间注意力感知三维点云的空间结构,以增强局部语义特征的表达,并提取有效上下文信息与深层语义特征。最后,采用TopK池化添加次优特征。实验结果表明,该算法在ModelNet40分类数据集上总体分类精度达到93.36%,在ShapeNet Part部件分割数据集上平均交并比达到85.96%,能够有效提取上下文信息和增强近邻点特征表达,表明了算法的有效性。

关键词:分类与分割;三维点云;注意力机制;加权动态图卷积;K近邻算法

中图分类号:TP391 **文献标识码:**A **doi:**10.37188/OPE.20243218.2823

3D point cloud classification and segmentation based on dual attention and weighted dynamic graph convolution

XIAO Jian, WANG Xiaohong*, LI Wei, YANG Yifei, LUO Ji

(School of Mining, Guizhou University, Guiyang 550000, China)

* Corresponding author, E-mail: xhwang@gzu.edu.cn

Abstract: In response to the deficiencies in local context information extraction and neighboring point feature expression in deep learning-based point cloud classification and segmentation networks, as well as the problem of maxpooling leading to the loss of suboptimal information, a point cloud classification and segmentation algorithm that combines dual attention and weighted dynamic graph convolutional networks was proposed. Firstly, the weighted dynamic graph convolution used a weighted k-nearest neighbor algorithm to construct a robust local structure and introduced an enhanced edge convolution module to apply weights to point features, thereby obtaining enhanced edge features. Then, channel attention was used to construct channel correlations and unleash the potential of each channel, followed by spatial attention to perceive the spatial structure of 3D point clouds, enhancing the expression of local semantic features and extracting effective contextual and deep semantic information. Finally, TopK pooling was employed to add suboptimal features. Experimental results show that the algorithm achieves an overall classification accura-

收稿日期:2024-01-25;修订日期:2024-04-02.

基金项目:教育部规划基金资助项目(No. 22YJIAZH083);贵州省科技计划资助项目(No. 黔科合支撑[2022]一般204);国家青年科学基金资助项目(No. 42301440)

cy of 93.36% on the ModelNet40 classification dataset and an average intersection over union of 85.96% on the ShapeNet Part segmentation dataset, effectively extracting contextual information and enhanced neighboring point feature expression, demonstrating the effectiveness of the algorithm.

Key words: classification and segmentation; 3D point cloud; attention mechanism; weighted dynamic graph convolution; K-nearest neighbor

1 引言

随着机载激光雷达(LiDAR)和计算机技术的进步,三维点云数据在测绘地理信息、自动驾驶、文化遗产保护和计算机视觉等领域得到了广泛应用。相较于二维数据,三维点云能更精准地展现物体的形状特征及深度信息。然而,点云的不规则性、稀疏性及无序性给三维特征表达带来巨大挑战。近年来,随着深度学习技术的发展,研究者开始探索适用于三维数据的处理方法,卷积神经网络的发展更是对三维数据的分类分割准确度及效率产生了里程碑式的影响。提升三维数据的处理能力已成为当前研究工作中的重要任务,旨在开发出更多样化的三维点云处理方法。基于深度学习的分类分割方法主要分为基于多视图、基于体素和基于点云的方法。

基于多视图的方法将三维立体投影为二维图像,再输入神经网络处理。MVCNN^[1]开创将二维图像进行3D识别的先河,通常在包围物体的20面体的各顶点处获取二维图像再输入神经网络处理。但是受视角与投影方法的限制,使得大量几何语义特征丢失,且特征信息冗余。

基于体素的方法将点云转换为离散的三维网格再处理。VoxNet^[2]首次提出体素化方法,设置不同分辨率以保留原始点云的几何信息,但体素转换过程需要大量计算资源,难以处理复杂的大型数据。OctNet^[3]与O-CNN^[4]采用八叉树结构来缓解高计算量与内存占用。但是体素密度高时,仍然依赖高性能GPU,而体素稀疏时,易丢失物体表面纹理特征。

基于点云的方法主要采用端对端神经网络结构。首次对点云直接处理的PointNet^[5],提出转换网络(Transform Net, T-Net)以解决点云的无序性,利用多层感知机(Multi-layer Perceptron, MLP)提取高维空间的点云几何细节特征,

但T-Net复杂度高,忽略了局部邻域的点间几何语义关系。PointNet++^[6]采用最远点采样下采样点云,将PointNet嵌入其球形邻域提取特征,以缓解局部特征提取不足的问题,但忽略了上下文信息。DGCNN^[7]使用K近邻算法(K-nearest Neighbor, KNN)构造局部邻域,其动态图更新模块能有效提取局部特征,由于局部结构较单一,难以提取深层次语义信息,且最大池化会丢失次优信息。RandLA-Net^[8]通过随机采样规避大型数据的复杂性,并设计局部特征聚合器增加感受野来缓解随机采样导致信息丢失。KP-Conv^[9]提出一种新颖的可变逐点卷积,固定球半径进行采样,设计核函数提取局部特征。文献^[10]使用最远点采样获取中心点,再通过KNN获取局部结构,关注了构建局部后的位置与几何特征,而忽略获取局部过程的重要性。PointWeb^[11]也采用传统KNN构建邻域,并将邻域内每个点特征进行关联。王江安^[12]等提出Linked-DGCNN,对边特征进行残差连接,有效缓解了最大池化导致的细节特征丢失,但是仍采用KNN获取局部邻域。文献[13-14]设计自注意力机制用于获取点云上下文特征,丰富了全局与局部特征信息。PointFormer^[15]引入新颖的点乘注意力机制,复杂度较低的同时可区别对待点云中每个点。文献[16]则设计特定的损失函数和训练策略来提升模型表现,确保了模型训练的稳定性和效率。

目前大部分采用传统KNN构建局部邻域的网络模型,改进方向侧重于局部邻域之后的特征提取操作,忽略了邻域构建过程的重要性,而且局部上下文信息提取不够充分,对点间深层语义信息的提取与表达不足,采用最大池化不可避免地丢失部分次优信息,难以释放模型的性能。

为解决上述问题,提出结合双注意力和加权动态图卷积神经网络的点云分类分割算法,该算法具有以下创新点与贡献:(1)利用自适应池化

构建加权 KNN 模块,对 KNN 提取的欧式距离进行加权,关注构建局部邻域的过程,有助于捕捉更精细的局部几何结构,并保持局部邻域的鲁棒性;(2)构建一种双注意力模块。首先,设计通道注意力模块对各通道赋权,构建通道间相关性,抑制无用通道,强化点间局部特征。然后,构建空间注意力模块获取注意力权重矩阵,增强网络对上下文信息的理解,强化局部上下文信息与局部语义特征的表达;(3)设计 TopK 池化,选取前 N 个最优特征进行融合,向网络补全次优信息。

2 结合双注意力的加权动态图卷积神经网络

结合双注意力和加权动态图卷积的点云分类分割算法如图 1 所示。点云分类网络由加权动态图卷积模块 (Weighted Edgeconv Module,

WEM)、双注意力模块 (Dual Attention Module, DAM) 和 TopK 池化构成。点云部件分割网络主要由 WEM 和 DAM 构成。WEM 由加权 KNN 模块 (Weighted Knn Module, WKNN) 和强化边卷积模块 (Enhanced Edgeconv Module, ECC) 构成如图 2 所示。双注意力模块由通道注意力模块 (Channel Attention Module, CAM) 和空间注意力模块 (Spatial Attention Module, SAM) 构成。

点云分类网络中,首先将点云输入 WEM,其中 WKNN 获取局部邻域并使用 ECC 提取加强边特征。然后, CAM 对点云不同通道赋予权重,加强重要通道构建通道相关性,再通过 SAM 增强邻域点的空间几何关系,提取局部的深层次语义特征和上下文信息,然后使用 TopK 池化补全次优信息并映入高维空间,通过最大池化和平均池化进行特征融合。最后,输入全连接层获取分类类别。

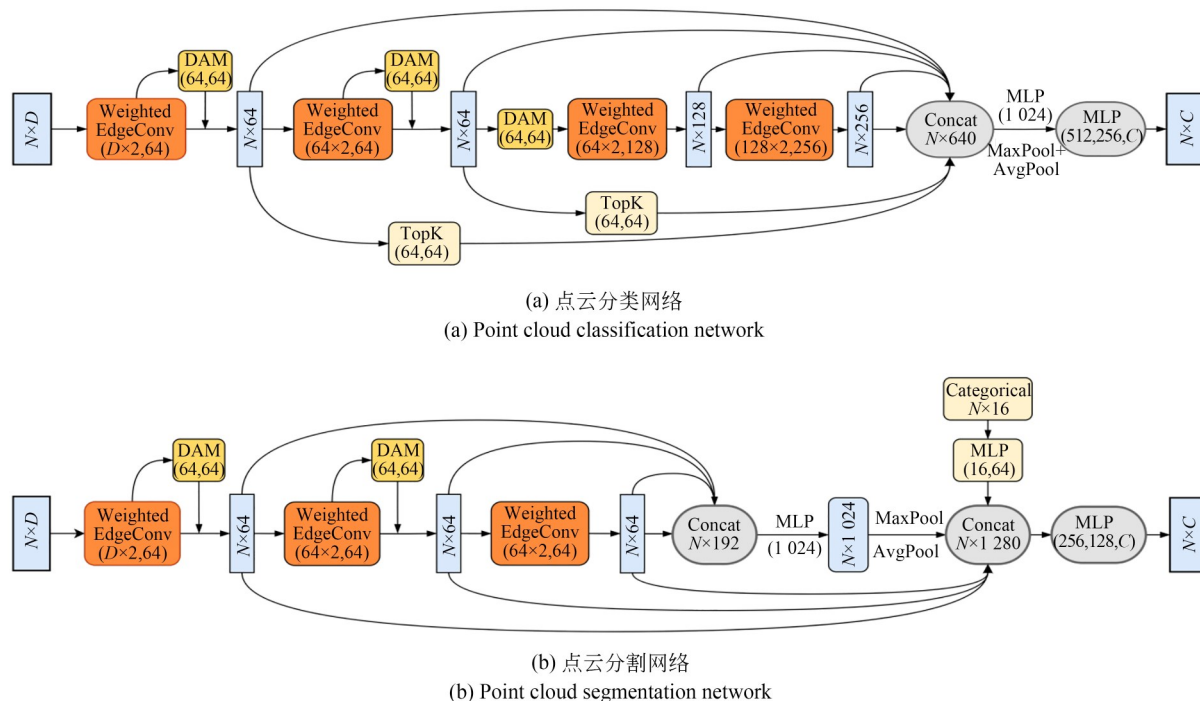


图 1 结合双注意力与加权动态图卷积的点云分类分割网络架构

Fig. 1 Point cloud classification and segmentation networks based on DAM and WEM

点云部件分割网络中,将点云输入 WEM 获取强化边特征,并经过 DAM 处理感知局部几何结构与强化邻域特征表达,不同于分类网络,分割网络仅拼接获取的三个特征向量再映入高维

空间,然后将其进行最大池化和平均池化后逐元素相加,并融合前三个特征向量与 MLP 处理后的类别信息,最后通过全连接层获取分割类别。

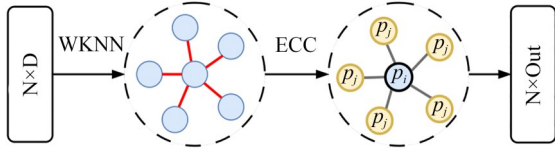


图2 加权动态图卷积模块组成结构

Fig. 2 Composition of WEM

2.1 加权 KNN 模块

受 SE-Net^[17] 启发,利用自适应池化与 MLP 设计 WKNN,对于点云 $P = \{p_1, p_2, \dots, p_n\}$,在传统 KNN 基础上将各点间欧式距离进行加权,再通过 TopK 算法选择 K 个点构成局部邻域如图 3,此过程加强了对构建局部的关注,重视了邻域的形成过程。

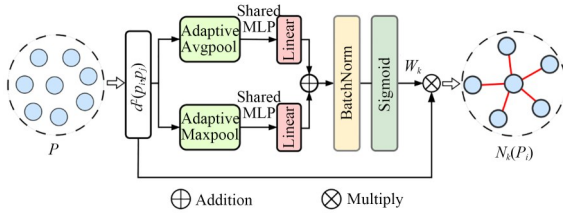


图3 加权 KNN 模块

Fig. 3 Weighted KNN module

首先,获取各点间欧氏距离:

$$d^2(p_i, p_j) = \sum_{n=1}^D (p_{in} - p_{jn})^2, \quad (1)$$

其中: D 是点云维度, p_i 和 p_j 是点云中不同点, p_{in} 和 p_{jn} 是不同维度下中心点和邻点的坐标,把 $d^2(p_i, p_j)$ 输入 WKNN,经过非线性函数处理后得到权重矩阵 W_k :

$$W_k = l_\theta(d^2(p_i, p_j)), \quad (2)$$

其中: $l_\theta \in P^{B \times C \times N}$ 是池化和 Sigmoid 等非线性函数, W_k 的大小通过 Sigmoid 控制为 $0 \sim 1$,点间距离越近权重越大。然后将 W_k 与 $d^2(p_i, p_j)$ 进行哈达玛乘积,相当于对点间相对距离赋予权重得到加权后欧式距离,增强了重要邻域点的影响:

$$D_i = W_k \cdot d^2(p_i, p_j). \quad (3)$$

最后,采用 TopK 选取 K 个邻点构造局部邻域:

$$(idx_i, dist_i) = \text{TopK}(-D_i, K) = N_k(P_i), \quad (4)$$

其中: idx_i 为点云索引, $dist_i$ 为欧式距离, $(idx_i, dist_i)$ 构成局部邻域 $N_k(P_i) = \{p_{i1}, p_{i2}, \dots,$

$p_{in}\}$ 。通过 W_k 加权欧式距离,根据点云几何结构选择合适的邻域,使得邻域更加鲁棒,有助于捕获其深层语义特征和上下文信息。

2.2 ECC 模块

DGCNN同时考虑了全局信息与局部信息:

$$e_{ij} = h_\theta(p_j - p_i, p_i). \quad (5)$$

此过程中,DGCNN 仅对特征进行简单拼接,导致信息提取不完整以及表达不足,且特征维度变换丢失大量上下文信息。在 PointWeb 等网络的启发下,对邻点 p_j 进行高斯加权如图 4 所示,在获取的邻域中建立局部邻点间联系,并增强了邻点的特征表达。

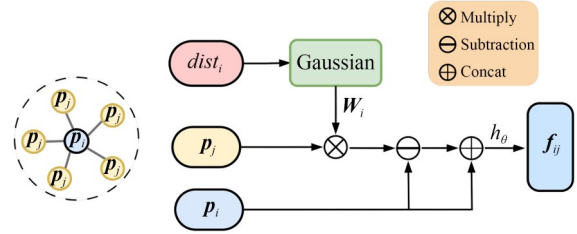


图4 ECC 模块

Fig. 4 ECC module

将 $dist_i$ 通过高斯函数处理得到权重矩阵 W_i :

$$W_i = e^{-\frac{dist_i}{2\sigma^2}}. \quad (6)$$

权重 W_i 大小与 $dist_i$ 成反比,范围控制为 $0 \sim 1$, σ 为大于 0 的标准差,用于控制权重的变化速率。 p_j 包含了自身坐标信息与中心点 p_i 的相对位置信息,将 p_j 通过 W_i 进行特征变换,不仅强化了相邻点的特征表达,还构造了更稳定的局部结构并提取到强化边特征 f_{ij} :

$$f_{ij} = h_\theta(W_i \cdot p_j - p_i, p_i) = h_\theta(n_{ij}, p_i), \quad (7)$$

其中: $f_{ij} \in P^{B \times C \times N \times K}$,与 DGCNN 获取的 e_{ij} 不同,它包含了加强的局部结构及其内部点间信息,使得网络对局部的空间感知能力增强,便于捕获三维点云的几何结构和点间语义信息。

2.3 双注意力模块

WKNN 重点关注构建点云局部邻域的过程,而 ECC 在获取的局部邻域中提取到强化边特征,但是仍缺乏上下文信息的获取。一般的注意力机制重视局部点间关系,但复杂度高,对点云通道的重视程度不够。CAM 先“Squeeze”再

“Excitation”构造了通道相关性,最大程度释放点云各通道的潜力。因此,结合CAM和SAM构造DAM如图5所示。CAM将 f_{ij} 作为输入,对其最后两个维度进行自适应最大和平均池化得到 $F_{avg}, F_{max} \in P^{B \times C \times 1 \times 1}$,降低数据复杂度的同时对通道信息进行感知,经过卷积变换后通过激活函数获得通道权重:

$$W_{avg} = \text{Sigmoid} [\text{Conv}(\text{Re lu}(\text{Conv}(F_{avg})))], \quad (8)$$

$$W_{max} = \text{Sigmoid} [\text{Conv}(\text{Re lu}(\text{Conv}(F_{max})))], \quad (9)$$

其中: $W_{avg}, W_{max} \in P^{B \times \frac{C}{2} \times 1 \times 1}$,变化 f_{ij} 的通道维度以方便进行矩阵相乘,得到 $F_{ij} \in P^{B \times \frac{C}{2} \times 1 \times 1}$ 。然后,将 F_{ij} 分别与 W_{avg} 和 W_{max} 逐元素相乘再拼接得到特征图 O :

$$O = \text{Concat}(W_{avg} \cdot F_{ij}, W_{max} \cdot F_{ij}). \quad (10)$$

O 融合了 W_{avg} 与 W_{max} 的权重信息,能够有效强化有用通道,加强通道间相关性。

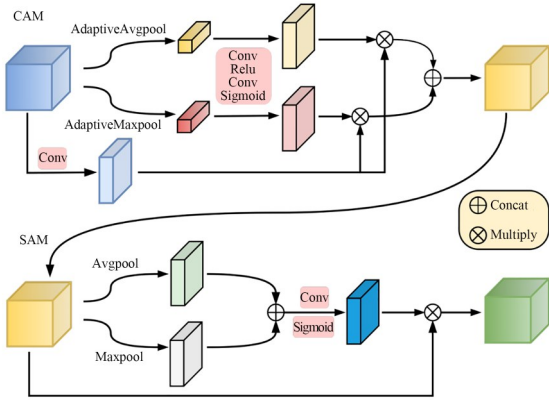


图5 双注意力模块

Fig. 5 Dual attention module

SAM将 O 作为输入,使用普通池化将通道维度降为1,其他维度保持不变得到:

$$O_{avg} = \text{Mean}(O, \text{dim} = 1) \in P^{B \times 1 \times N \times K}, \quad (11)$$

$$O_{max} = \text{Max}(O, \text{dim} = 1) \in P^{B \times 1 \times N \times K}. \quad (12)$$

然后拼接 O_{avg} 和 O_{max} ,输入 1×1 卷积层并经过Sigmoid得到空间注意力权重 O_{att} :

$$O_{att} = \text{Sigmoid}(\text{Conv}(\text{Concat}(O_{avg}, O_{max}))). \quad (13)$$

$O_{att} \in P^{B \times 1 \times N \times K}$ 经过一系列非线性函数处理,感知了点云的空间结构,有效提取到局部空间的上下文信息,且经过CAM处理后通道具有

强相关性,所以将 O_{att} 与 O 逐元素相乘得到 O_{final} :

$$O_{final} = (O_{att} \cdot O) \in P^{B \times C \times N \times K}. \quad (14)$$

使得局部邻域中不仅通道间相关性增强,也具备了有效的上下文信息。

2.4 TopK池化

目前基于点云的模型大都采用最大池化聚合特征,容易丢失次优特征。为缓解该问题,在KNN中TopK算法的启发下,设计TopK池化如公式(15):

$$\text{TopKpooling} = \text{Mean}(\text{Max}(\sum_{k=0}^N O_{final; k}, \text{dim} = -1)), \quad (15)$$

其中: K 为局部邻域的 K 个点,在 O_{final} 特征向量最后一个维度选择前 N 个最大特征向量,再进行平均操作以获得新的特征向量,将其与提取的特征融合来补全丢失的次优信息。TopK池化添加在分类网络的第二和第三层,一般最大池化只选择特征向量中最大的一个作为下一层的输入,而TopK池化对所有向量进行排序,选择前 N 个最大向量,该 N 个向量就包含最大池化可能忽略的次优向量,最后将 N 个向量进行平均操作以满足下一层的输入维度。

3 实验结果与分析

3.1 数据集与实验环境参数设置

采用点云分类数据集ModelNet40^[18],点云部件分割数据集ShapeNet Part^[19],对本文分类分割网络的性能进行评估。ModelNet40包含40个三维模型类别,含有12 311个已标注的CAD模型,9 842个模型用于训练,2468个模型用于测试,评估指标采用总体分类精度(Overall Accuracy, OA)和平均分类精度(Mean Accuracy, mAcc)。ShapeNet Part包含16种三维模型,目前约有16 800个已标注的CAD模型,评估指标采用平均交并比(Mean Intersection-over-union, mIoU)。

实验硬件环境由Intel i5 13490F CPU和GeForce RTX 3080ti(12GB)显卡组成。软件环境为Linux Ubuntu 20.04系统,深度学习框架Pytorch 1.7.0,语言Python 3.7, CUDA 11.0和cuDNN 8.0.4。实验主要参数设置见表1,参考DGCNN网络的参数设置进行设置,表中参数为

获得网络最佳结果时的参数组合。采用 AdamW 优化器训练,因其相比 Adam 更加鲁棒,能更快收敛。由于批大小和验证集大小较小且 AdamW 收敛较快,所以将学习率设置得较小,同时增加训

练次数让网络能够充分学习数据集。余弦退火设置根据训练末期网络的分类精度进行调整,在训练末期调整为较小的学习率可让网络准确率平缓收敛到较好的数值。

表 1 实验主要参数设置

Tab. 1 Main parameter settings of the experiment

Dataset	Points	Optimizer	Learning rate	Cosine Annealing	Batch size	Test batch size	Epochs
ModelNet40	1 024	AdamW	0.001	1×10^{-5}	16	8	250
ShapeNet Part	2 048	AdamW	0.001	1.5×10^{-5}	8	4	200

3.2 点云分类结果与复杂度分析

3.2.1 点云分类结果与分析

表 2 是本文分类网络和其他主流网络模型在 ModelNet40 数据集的实验结果。本文分类网络的 OA 和 mAcc 分别达到 93.36% 和 90.74%。整体来看,基于点云的方法比基于多视图和基于体素的方法好,本文分类网络无需转换为二维图像和体素,减少了信息损失,mAcc 几乎达到表中最优。从 OA 来看,比 MVCNN^[1]高出 3.26%,优于 VoxNet^[2]7.46%,相较于 DGCNN^[7]和 KP-Conv^[9]均提高 0.46%。比近期提出的 GCN3D^[20]

高出 0.36%,主要因为 WKNN 构建了有效且鲁棒的局部邻域,ECC 模块突出了点的特征表达,以及 TopK 补全了次优信息。3k 点输入的 Linked-DGCNN^[12]网络达到 92.6% 的 OA,而本算法在 1k 点输入情况下 OA 较其提高 0.76%,mAcc 提高 1.04%。与当下热门网络 Point-Transformer^[21]相比仍有提升空间,Transformer 网络更深,多头注意力能更好地感知三维空间,但是复杂度高依赖高性能 GPU。与目前较新网络 RF-DGCNN^[22]相比稍有逊色,因为其残差结构构建了更深层次的网络。

本文分类网络由于硬件限制,难以设计深层次和高复杂度的网络,在算法复杂度和网络结构方面仍存在优化空间。与主流模型 PointWeb^[11],PointConv^[23],PointASNL^[24],SpiderCNN^[25]相比提高了 1.06%,0.96%,0.16%,0.96%。由此可见本网络在点云分类任务中可达到较高性能,具有较好的点云分类能力。

图 6 是训练次数与总体精度曲线图,因为采

表 2 ModelNet40 数据集三维点云分类结果对比
Tab. 2 Comparison of 3d point cloud classification results on modelnet40 dataset

Methods	Input Type	OA/%	mAcc/%
MVCNN ^[1]	Image	90.1	—
VoxNet ^[2]	Voxel	85.9	—
PointNet ^[5]	Points(1k)	89.2	86.2
PointNet++ ^[6]	Points(5k)+Normal	91.7	89.5
PointWeb ^[11]	Points(1k)+Normal	92.3	89.4
Point-Transformer ^[21]	Points(1k)	93.7	—
PointConv ^[23]	Points(1k)+Normal	92.4	—
PointASNL ^[24]	Points(1k)+Normal	93.2	—
SpiderCNN ^[25]	Points(1k)+Normal	92.4	—
GCN3D ^[20]	Points(0.768k)	93.0	90.3
RF-DGCNN ^[22]	Points(1k)	93.62	90.86
KPConv ^[9]	Points(1k)	92.9	—
DGCNN ^[7]	Points(1k)	92.9	90.2
Linked-DGCNN ^[12]	Points(3k)	92.6	89.7
Ours	Points(1k)	93.36	90.74

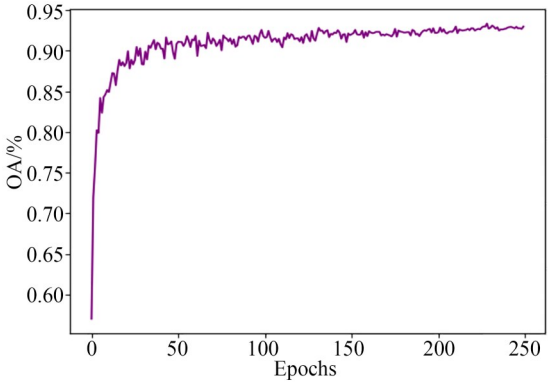


图 6 训练次数与总体精度曲线图

Fig. 6 Training epochs and overall accuracy curve graph

用 AdamW 优化器,网络在前 50 次训练中快速收敛,在 150 次时学习率余弦退火到较小数值,训练逐渐平缓,说明训练过程各超参数已达最优状态。

3.2.2 点云分类网络时间及空间复杂度分析

表 3 为本文分类网络与其他经典网络的参数量和前向传播时间结果对比。通过对比参数量和前向传播时间来分析网络的时间和空间复杂性。参数量包括网络的权重和偏置,前向传播时间是在单个 GeForce RTX 3080ti GPU 上计算得到。

表 3 不同分类网络参数量和前向传播时间的结果对比

Tab.3 Comparison of the results for the parameters and forward time of different classification networks

Methods	Parameters/M	Forward Time/ms	OA/%
PointNet ^[5]	3.48	3.2	89.2
PointNet++ ^[6]	1.48	4.8	91.7
DGCNN ^[7]	1.84	6.3	92.9
Point-Transformer ^[21]	1.73	136.4	93.7
Ours	2.08	8.3	93.36

本文网络在 DGCNN 基础上改进,构建了 WKNN,ECC 和 DAM 模块。而 DGCNN 参数量低于 PointNet 略高于 PointNet++,所以本文网络参数量虽然低于 PointNet,但略高于 PointNet++,DGCNN 和 Point-Transformer 等网络,说明增加的模块具有一定参数量,未使得网络参数冗余。上述三模块的设计使得前向传播时间略高于 PointNet,PointNet++ 和 DGCNN 网络,

但是分类总体精度均高于该三个网络。与 Point-Transformer 相比效率较高,消耗更少的时间达到近似 Point-Transformer 的分类效果。本文设计的各模块导致网络参数量增大,前向传播时间有所下降,但也提升了分类精度,综合性能较好,展现了该网络对于分类任务的有效性。

3.3 点云部件分割结果与分析

表 4 为本文部件分割网络与其他主流网络模型在 ShapeNet Part 数据集上的实验结果,表中包含 mIoU 和 16 个类别的 IoU。由表 4 可知,本文分割网络 mIoU 达到 85.96%,比 DGCNN^[7]高出 0.76%,相较于 Linked-DGCNN^[12]高出 1.26%,比近三年模型 GCN3D^[20],DTNet^[26],3D-GCN^[27]分别提高 0.76%,0.36%,0.86%,与主流模型 PointNet^[5],PointNet++^[6],PointConv^[23]相比提高 2.26%,0.86%,0.26%。虽然低于 PointASNL^[24]的 86.1%,但是在 16 个类别中 11 类达到最优,分别是:飞机、包、帽子、汽车、耳机、吉他、笔记本电脑、马克杯、手枪、火箭和滑板。

图 7 是部分部件分割可视化结果,从图中可看出本文分割网络较 DGCNN 更接近真实值,本文网络正确地对飞机的细节处进行了分割,这是由于 WKNN 增强了对局部邻域选择的关注,DAM 对点云整体的空间观察得更加细致,反观 DGCNN 采用传统 KNN 未能注意到该细节部分,对于摩托车本文网络在车灯处仅有少许错分点,而 DGCNN 直接未分割出车灯部分。综合表 3 与图 6,本文网络对物体部件分割效果较好,相比大部分主流网络模型具有更好的分割性能。

表 4 ShapeNet Part 数据集三维点云部件分割结果对比

Tab.4 Comparison of 3d point cloud segmentation results on shapenet part dataset

Methods	mIoU/%	airplane	Bag	Cap	Car	Chair	Earphone	Guitar	Knife
PointNet ^[5]	83.7	83.4	78.7	82.5	74.9	89.6	73.0	91.5	85.9
PointNet++ ^[6]	85.1	82.4	79.0	87.7	77.3	90.8	71.8	91.0	85.9
PointConv ^[21]	85.7	—	—	—	—	—	—	—	—
PointASNL ^[22]	86.1	84.1	84.7	87.9	79.7	92.2	73.7	91.0	87.2
GCN3D ^[24]	85.2	83.3	78.8	83.2	77.5	90.8	75.8	90.8	86.0
DTNet ^[26]	85.6	83.0	81.4	84.3	78.4	90.9	74.3	91.0	87.3
3D-GCN ^[27]	85.1	83.1	84.0	86.6	77.5	90.3	74.1	90.9	86.4
DGCNN ^[7]	85.2	84.0	83.4	86.7	77.8	90.6	74.7	91.2	87.5
Linked-DGCNN ^[12]	84.7	—	85.4	88.6	—	90.8	—	90.9	88.6
Ours	85.96	86.5	86.6	89.4	80.4	91.7	81.3	91.9	88.4

续表 4 ShapeNet Part数据集三维点云部件分割结果对比

Tab. 4 Comparison of 3d point cloud segmentation results on shapenet part dataset

Methods	Lamp	Laptop	Motorbike	Mug	Pistol	Rocket	Skateboard	Table
PointNet ^[5]	80.8	95.3	65.2	93.0	81.2	57.9	72.8	80.6
PointNet++ ^[6]	83.7	95.3	71.6	94.1	81.3	58.7	76.4	82.6
PointConv ^[21]	—	—	—	—	—	—	—	—
PointASNL ^[22]	84.2	95.8	74.4	95.2	81.0	63.0	76.3	83.2
GCN3D ^[24]	83.8	95.2	66.4	93.4	81.3	51.2	67.1	83.0
DTNet ^[26]	84.7	95.6	69.0	94.4	82.5	59.0	76.4	83.5
3D-GCN ^[27]	83.8	95.6	66.8	94.8	81.3	59.6	75.7	82.8
DGCNN ^[7]	82.8	95.7	66.3	94.9	81.1	63.5	74.5	82.6
Linked-DGCNN ^[12]	82.4	95.4	—	—	—	—	74.7	81.9
Ours	84.5	96.4	72.9	96.0	85.4	63.6	78.8	82.9

Bold numbers indicates the best value in the category.

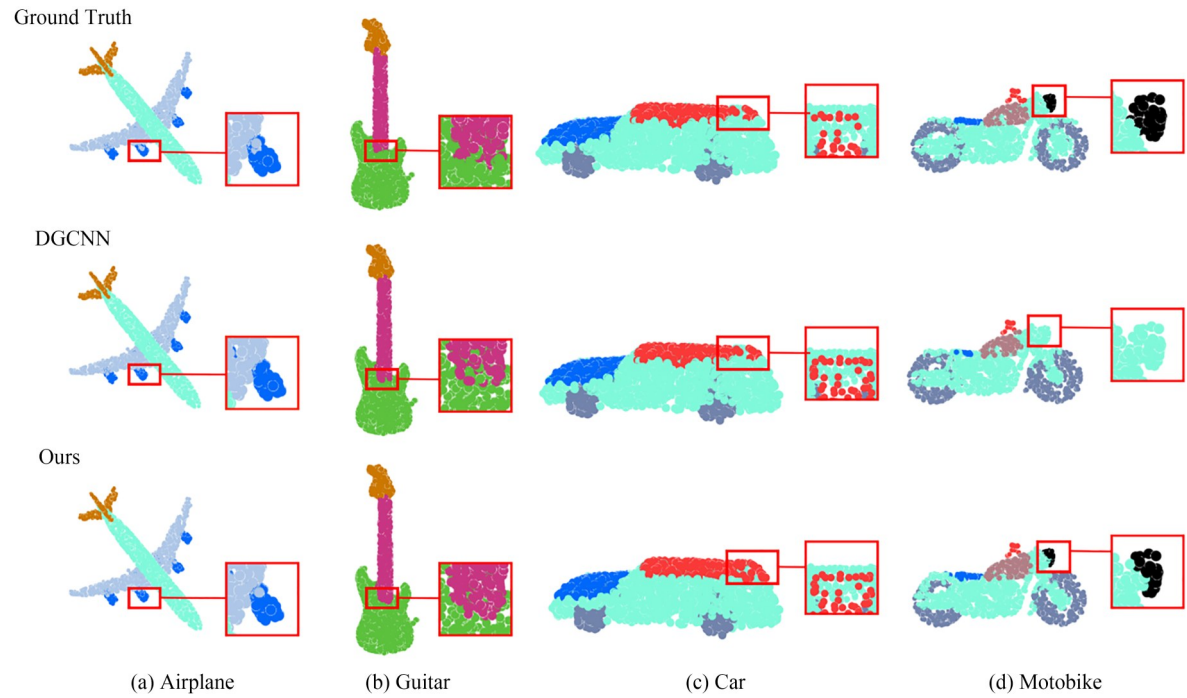


图7 部件分割可视化结果

Fig. 7 Visualization results of part segmentation

3. 4 消融实验与鲁棒性分析

3. 4. 1 点云分类任务消融实验

为验证各模块对点云分类任务的影响,实验设置及结果如表5所示,表中粗体为最佳性能。

由表5可得,网络B的OA和mAcc比原始网络A提升了0.65%和1.14%,证明DAM能够充分感知局部邻域的点云空间结构,并增强通道相

关性;网络C和E,均比网络B精度高,网络G与网络C,E相比精度也有提升,表明在DAM基础上添加ECC+KNN或单独添加WKNN,可有效提升分类精度,且在ECC+WKNN的组合下网络的分类性能提升最大,综合反映了ECC和WKNN较原始边缘卷积和KNN的优异性。而添加TopK池化的网络D,F,H与未添加TopK

表 5 点云分类任务中各模块有效性验证

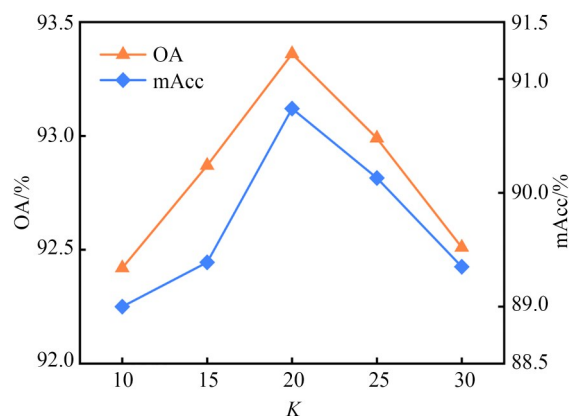
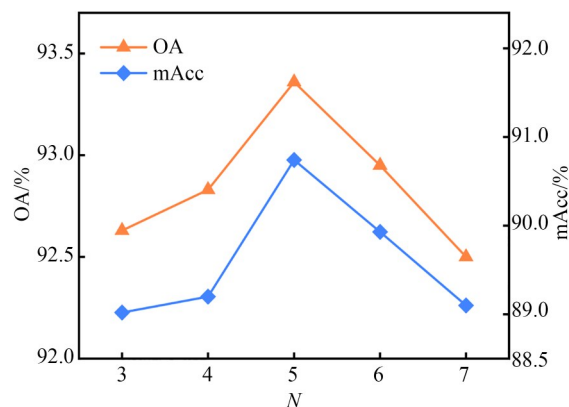
Tab. 5 Validation of the effectiveness of each module in point cloud classification task

Nets	DAM	ECC/ KNN/WKNN	TopK	OA/ %	mAcc/ %
A	×	KNN	×	91.45	87.68
B	✓	KNN	×	92.10	88.82
C	✓	ECC+KNN	×	92.18	88.83
D	✓	ECC+KNN	✓	92.30	89.31
E	✓	WKNN	×	92.42	89.25
F	✓	WKNN	✓	92.63	88.67
G	✓	ECC+WKNN	×	92.91	89.89
H	✓	ECC+WKNN	✓	93.36	90.74

池化的网络 C, E, G 对应相比, 添加 TopK 池化的网络精度更高, 说明 TopK 池化优于单一的最大池化, 显著缓解了次优信息的丢失, 在网络映射入高维空间时补全了次优信息。以上实验结果表明, WKNN 较 KNN 能构建更优的局部邻域, ECC 在局部邻域提取到更具代表性的边特征, DAM 则在之后的操作中强化了对点云空间和通道的感知, TopK 池化在网络末端补全次优信息, 综合证明各模块的有效性, 所有模块均使用时达到网络的最优性能。

构建局部邻域过程中, K 值对局部特征的提取十分重要, 合理的 K 值决定了网络的性能, 在一定范围内设定不同 K 值, 研究其对分类网络性能的影响, 具体实验结果如图 8 所示。实验结果表明, 一定范围内随着 K 的增大网络分类精度也在提升, 在 $K=20$ 时达到网络最佳分类性能。当 $K<20$ 时, 由于局部点较少点密度低, 即使 WKNN 构建了鲁棒的局部结构, 但 ECC 也难以提取到局部中有效的边特征, 导致分类精度较低。而 $K>20$ 时, 局部邻域点过多, 不论 KNN 或 WKNN 都会因点密度高而造成计算量大与网络复杂度高等问题, 使得特征产生冗余, 算法性能下降。

设置 TopK 消融实验如图 9 所示, 研究 N 值对分类网络中次优信息补全能力的影响。通过实验证明, 一定范围内, 设定合适的 N 值可获取有效次优信息。当 $N<5$ 时, 特征个数较少可能难

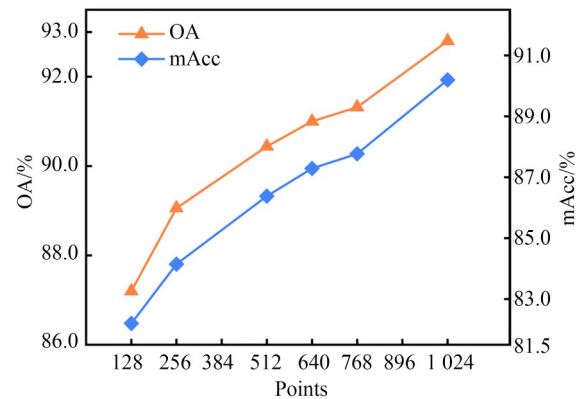
图 8 K 值对分类网络性能影响对比Fig. 8 Comparison of the impact of K on the performance of classification network图 9 N 值对分类网络性能的影响对比Fig. 9 Comparison of the impact of N on the performance of classification network

以捕获到次优信息, 无法弥补信息丢失的问题。在 $N=5$ 时, 特征个数适宜, 能够获得适量的特征向量来补充次优信息, 从而达到算法最优性能。而 $N>5$ 后, 由于特征向量过多, 特征产生冗余, 可能忽略其中的次优特征, 导致网络性能逐渐下降。

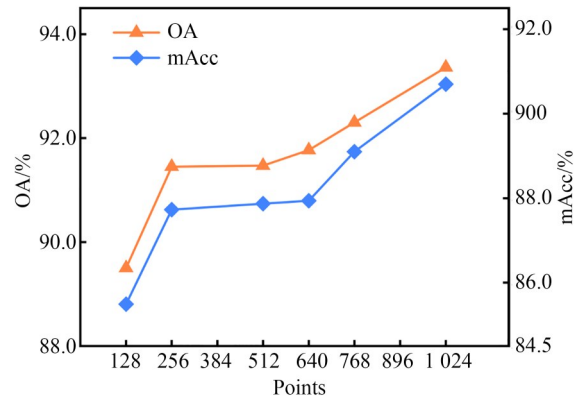
3.4.2 点云分类任务鲁棒性分析

设置鲁棒性实验分析不同点云密度下本文网络的鲁棒性, 与 DGCNN 网络进行对比。如图 10 所示, 在采样点为 128, 256, 512, 640, 768 和 1 024 时训练网络并获得其 OA 与 mAcc。图 10(a) 是 DGCNN 实验结果, 图 10(b) 是本文分类网络实验结果。随着点数增加, 本文网络的 OA 与 mAcc 逐渐升高并始终高于 DGCNN。当点数

低于 512 时,DGCNN 分类精度发生大幅下降,OA 在 128 点下降至 87.2%,mAcc 下降至 82.2%,而本文网络在 128 点时,OA 与 mAcc 仍能达到 89.5% 和 85.5%。通过以上实验表明,在处理不同点云密度时,通过 DAM 感知点云空间,以及 WKNN 对构建局部过程的关注,使得本文网络具有优异的鲁棒性。



(a) DGCNN 鲁棒性实验结果
(a) Robustness experiment results of DGCNN



(b) 本文网络鲁棒性实验结果
(b) Robustness experiment results of our network

图 10 鲁棒性分析实验结果

Fig. 10 Robustness analysis experiment results

3. 4. 3 点云部件分割任务消融实验

为探究各模块对部件分割网络性能的影响,设置以下消融实验,具体设置及结果如表 6 所示,其中粗体为最佳性能。

由表 6 可得,采用 KNN 的网络添加 DAM 后 mIoU 提高了 0.29%;在具备 DAM 后采用 ECC+KNN 或 WKNN 模块均有效提升了网络

表 6 点云分割任务中各模块有效性验证

Tab. 6 Validation of the effectiveness of each module in point cloud segmentation task

DAM	ECC+WKNN/KNN	OA/%	mIoU/%
×	KNN	94.03	84.65
✓	KNN	94.13	84.94
✓	ECC+KNN	94.28	85.29
✓	WKNN	94.25	85.32
✓	ECC+WKNN	94.56	85.96

性能;而同时包含 DAM、ECC 和 WKNN 模块的算法获得最佳分割效果。说明相较于普通 KNN,WKNN 更注重邻域的构建过程,得到更有效的局部邻域,ECC 在更有效邻域基础上增强了邻域点的特征表达,之后便通过 DAM 感知点云结构并提取有效的上下文信息和深层语义特征,部件分割任务采用 2 048 点输入,为避免特征冗余,未加入 TopK 池化补全次优信息。

为探讨 K 值对点云部件分割算法性能的影响,将 K 设置为 30,35,40,45 等,具体设置和实验结果如图 11。从图 11 看出,K=40 时 OA 和 mIoU 均达到最优,达到网络最佳性能。K<40 时,局部点密度低,提取特征不充分,限制了网络的分割性能,而 K>40 后,点密度过大,破坏了局部结构,大幅降低了网络性能。因此,在合适的区间选择适当的 K 对网络分割性能影响重大。

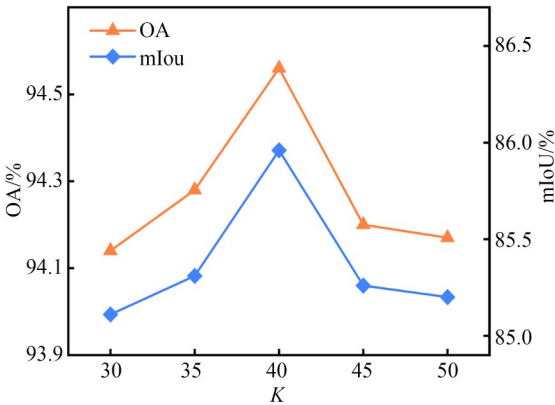


图 11 K 值对部件分割网络性能的影响对比

Fig. 11 Comparison of the impact of K on the performance of segmentation network

4 结 论

本文提出一种结合双注意力和加权动态图卷积的三维点云分类分割算法。首先,通过加权动态图卷积以重视局部邻域的构建过程,并在邻域中强化边缘特征表达;然后,通过双注意力模块强化通道间关联以及增强网络对点云的空间感知能力;最后,设计 TopK 池化补全未提取到的次优特征。通过消融实验和鲁棒性实验分析网络的参数设置,并进一步提升网络性能,在 ModelNet40 和 ShapeNet Part 数据集上实现了对点云

的高性能处理。

本文网络在点云分类与部件分割数据集上展示了优秀的性能。然而,当应用于大型室内外语义分割数据集时,采用 WKNN 获取邻域可能导致计算成本较高,这表明本方法在语义分割性能上仍有进步空间。此外,当前的 KNN 方法仍需手工设置 K 值,网络的复杂度也存在优化空间。因此,研究网络在大型室内外语义分割数据集上的分割效率及性能,以及通过训练神经网络实现自动更新 K 值是未来进一步的研究工作。

参考文献:

- [1] SU H, MAJI S, KALOGERAKIS E, *et al.* Multi-view convolutional neural networks for 3D shape recognition[C]. 2015 *IEEE International Conference on Computer Vision (ICCV)*. Santiago, Chile. IEEE, 2015: 945-953.
- [2] MATURANA D, SCHERER S. Voxnet: a 3D convolutional neural network for real-time object recognition[C]. 2015 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Hamburg, Germany. IEEE, 2015: 922-928.
- [3] RIEGLER G, ULUSOY A O, GEIGER A. OctNet: learning deep 3D representations at high resolutions[C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI, USA. IEEE, 2017: 6620-6629.
- [4] WANG P S, LIU Y, GUO Y X, *et al.* O-CNN[J]. *ACM Transactions on Graphics*, 2017, 36(4): 1-11.
- [5] CHARLES R Q, HAO S, MO K C, *et al.* PointNet: deep learning on point sets for 3D classification and segmentation[C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI, USA. IEEE, 2017: 77-85.
- [6] QI C R, YI L, SU H, *et al.* Pointnet++: deep hierarchical feature learning on point sets in a metric space[J]. *Advances in Neural Information Processing Systems*, 2017, 30: 5099-5108.
- [7] 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.
- [8] HU Q Y, YANG B, XIE L H, *et al.* RandLA-net: efficient semantic segmentation of large-scale point clouds[C]. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, WA, USA. IEEE, 2020: 11105-11114.
- [9] THOMAS H, QI C R, DESCHAUD J E, *et al.* KPConv: flexible and deformable convolution for point clouds[C]. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*. Seoul, Korea (South). IEEE, 2019: 6410-6419.
- [10] 杨军,王连甲. 结合位置关系卷积与深度残差网络的三维点云识别与分割[J]. *西安交通大学学报*, 2023, 57(5): 182-193.
YANG J, WANG L J. Recognition and segmentation of 3D point cloud through positional relation convolution in combination with deep residual network[J]. *Journal of Xi'an Jiaotong University*, 2023, 57(5): 182-193. (in Chinese)
- [11] ZHAO H S, JIANG L, FU C W, *et al.* PointWeb: enhancing local neighborhood features for point cloud processing[C]. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, CA, USA. IEEE, 2019: 5560-5568.
- [12] 王江安,何娇,庞大为. 基于动态图卷积网络

- 点云分类和分割网络[J]. 激光与光电子学进展, 2021, 58(12): 1215008.
- WANG J A, HE J, PANG D W. Point cloud classification and segmentation network based on dynamic graph convolutional network[J]. *Laser & Optoelectronics Progress*, 2021, 58 (12) : 1215008. (in Chinese)
- [13] 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. (in Chinese)
- [14] 杨军, 李博赞. 基于自注意力特征融合组卷积神经网络的三维点云语义分割[J]. 光学精密工程, 2022, 30(7): 840-853.
- YANG J, LI B Z. Semantic segmentation of 3D point cloud based on self-attention feature fusion group convolutional neural network[J]. *Opt. Precision Eng.*, 2022, 30(7): 840-853. (in Chinese)
- [15] CHEN Y J, YANG Z L, ZHENG X W, *et al.* Pointformer: a dual perception attention-based network for point cloud classification[C]. WANG L, GALL J, CHIN TJ, *et al.* *Asian Conference on Computer Vision*. Cham: Springer, 2023: 432-449.
- [16] YANG Z L, CHEN Y J, ZHENG X W, *et al.* Conditional GAN for point cloud generation[C]. WANG L, GALL J, CHIN TJ, *et al.* *Asian Conference on Computer Vision*. Cham: Springer, 2023: 117-133.
- [17] HU J, SHEN L, SUN G. Squeeze-and-excitation Networks [C]. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT, USA. IEEE, 2018: 7132-7141.
- [18] 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)*. Boston, MA, USA. IEEE, 2015: 1912-1920.
- [19] 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.
- [20] 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. (in Chinese)
- [21] ZHAO H S, JIANG L, JIA J Y, *et al.* Point transformer [C]. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*. Montreal, QC, Canada. IEEE, 2021: 16239-16248.
- [22] 梁正友, 蔡俊民, 孙宇, 等. 结合残差动态图卷积与特征强化的点云分类[J]. 广西师范大学学报(自然科学版), 2023, 41(5): 37-48.
- LIANG Z Y, CAI J M, SUN Y, *et al.* Point cloud classification based on residual dynamic graph convolution and feature enhancement [J]. *Journal of Guangxi Normal University (Natural Science Edition)*, 2023, 41(5): 37-48. (in Chinese)
- [23] WU W X, QI Z A, LI F X. PointConv: deep convolutional networks on 3D point clouds [C]. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, CA, USA. IEEE, 2019: 9613-9622.
- [24] 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)*. Seattle, WA, USA. IEEE, 2020: 5588-5597.
- [25] XU Y F, FAN T Q, XU M Y, *et al.* SpiderCNN: deep learning on point sets with parameterized convolutional filters [C]. *European Conference on Computer Vision*. Cham: Springer, 2018: 90-105.
- [26] HAN X F, JIN Y F, CHENG H X, *et al.* Dual transformer for point cloud analysis [J]. *IEEE*

Transactions on Multimedia, 2022, 25: 5638-5648.

- [27] 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]. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, WA, USA. IEEE, 2020: 1797-1806.

作者简介:



肖 剑(2000—),男,贵州安顺人,硕士研究生,2022年于东华理工大学获得学士学位,主要从事为三维点云处理与深度学习方面的研究。E-mail: 1332553410@qq.com



王晓红(1970—),男,贵州贵阳人,博士,副教授,硕士生导师,2003年于武汉大学获得硕士学位,2013年于武汉大学获得博士学位,现为贵州大学矿业学院副教授,主要从事遥感数据处理与分析,3s技术应用等方面的研究。E-mail: xhwang@gzu.edu.cn