文章编号:1674-2974(2024)12-0139-14 DOI:10.16339/j.cnki.hdxbzkb.2024291

自适应生成卷积核的动态图注意力三维点云识别及分割

杨军[1,2†],郭佳晨[2]
(1. 兰州交通大学 电子与信息工程学院,甘肃 兰州 730070;
2. 兰州交通大学 自动化与电气工程学院,甘肃 兰州 730070)

摘 要:针对现有算法在对点云数据进行平移、缩放以及旋转等几何变换时网络不能充分提取局部特征,导致网络精度显著下降的问题,提出基于自适应生成卷积核的动态图注意力三维点云识别及分割算法.首先,利用感受野中心点位置信息增强邻点感知上下文信息能力,通过改进的自注意力机制重构感受野,使感受野内特征信息充分交互,增强感受野的上下文信息.其次,构造自适应生成卷积核,通过捕获变化的点云拓扑信息,自适应生成卷积核权重,进而提升网络性能.最后,构建动态图注意力卷积算子,并设计点云识别的动态网络与分割的U形网络.实验结果表明,本文算法在ModelNet40点云识别数据集的识别精度达到了94.0%,在ShapeNet Part点云部件语义分割数据集的平均交并比达到了86.2%.本文算法能够提取三维点云的关键特征信息,具有较好的三维点云识别与分割能力.

关键词:三维点云;动态图注意力卷积;自适应算法;模型识别;语义分割
中图分类号:TP391 **文献标志码**:A

Recognition and Segmentation of 3D Point Cloud by Dynamic Graph Attention with Adaptive Generated Convolutional Kernel

YANG Jun[1,2†], GUO Jiachen[2]
(1. School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China;
2. School of Automation and Electrical Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China)

Abstract: As the current algorithms fail to fully extract local features and result in significant degradation of network accuracy when performing geometric transformations such as translation, scaling, and rotation on point cloud data, this paper proposes a dynamic graph attention-based 3D point cloud recognition and segmentation algorithm based on adaptive generated convolutional kernels. Firstly, the positional information of the center point in the receptive field is used to enhance the contextual information perception of neighboring points. The receptive field is reconstructed to enable sufficient interaction of feature information within the receptive field and enhance the contextual information by improving the self-attention mechanism. Then, an adaptive generated convolutional kernel is constructed to capture changing point cloud topology information and adaptively generate convolutional kernel

* **收稿日期**:2023-09-14
基金项目:国家自然科学基金资助项目(42261067), National Natural Science Foundation of China(42261067)
作者简介:杨军(1973—),男,宁夏吴忠人,兰州交通大学教授,博士
† 通信联系人,E-mail: yangj@mail.lzjtu.cn

weights to enhance network performance. Finally, a dynamic graph attention convolutional operator is built, and a dynamic network for point cloud recognition and a U-shaped network for segmentation are designed. The experimental results show that the proposed algorithm achieves a recognition accuracy of 94.0% in the ModelNet40 point cloud recognition dataset, and the instance mean intersection over union reaches 86.2% in the ShapeNet Part point cloud semantic segmentation dataset. The algorithm proposed can extract critical feature information from 3D point clouds and is capable of 3D point cloud recognition and segmentation.

Key words: 3D point cloud; dynamic attention graph convolution; adaptive algorithms; model recognition; semantic segmentation

近年来,三维视觉技术的发展促进了增强现实、桥隧工程以及自动驾驶[1-2]等技术的应用.三维点云识别作为计算机视觉的重要研究课题之一,旨在分析三维点云模型并输出所属类别,而三维点云分割旨在预测点云场景中每个点所属的物体,或所属的部件类别[3].由于点云数据具有不规则、无序性和稀疏性等特点,自动、精确且在数据变换时具有鲁棒性的点云识别与分割的研究成为一项具有挑战性的任务.

随着深度学习技术的显著进步,研究者尝试将卷积神经网络(convolution neural networks, CNNs)应用到点云数据的识别和分割任务中,取得了较好的结果.起初学者们将三维点云转化为二维图像,输入CNN中提取数据特征,于是,基于投影的方法应运而生.该方法首先获取三维点云数据在不同视角下的二维视图,然后对每个视图进行卷积计算,并通过池化层和全连接层对每个视图的特征进行聚合.然而,投影过程会造成部分信息缺失,其忽略了三维点云的空间内在几何关系,导致识别及分割结果并不理想.基于体素化的方法[4-5]将三维点云描述为体素网格,生成三维点和其对应的体素之间多对一映射,并利用标准的三维卷积进行分析处理.然而,该类方法通常存在体素网格分辨率受限、内存成本较高以及难以捕获细粒度几何特征等问题.

与图像不同,点云是非结构化的,离散分布在采样对象的表面.基于投影和基于体素化的方法均是对其进行规范化处理后输入神经网络,而PointNet[6]开创性地提出直接对点云进行处理,应用多个全连接层对三维点云编码,在三维点云识别和语义分割领域取得了较好的结果,然而该方法聚焦于提取点云全局特征,忽略了局部点之间的几何联系,存在局部特征丢失的问题.

为了更好地描述三维点云的局部信息,一些研究工作将点云表示为图数据,并将二维卷积推广到三维,以提取丰富的局部信息.文献[7]提出了边缘条件卷积,由边缘标签动态生成滤波器权重,将二维卷积算子推广至三维数据的同时避免了频域运算.文献[8]和文献[9]则定义具有固定模式的内核提取点云特征,文献[10-13]为点云构建类图结构并改进局部特征提取器,深入挖掘局部特征.为了处理不同邻域大小的点集,标准图卷积使用共享权值函数来提取边缘特征,致使出现各向同性核,忽略了感受野内不同边具有不同权重的问题.此外,由于现有方法大多考虑点云的全局坐标,如果没有进行归一化处理,当点云发生平移和缩放等刚性变换时网络模型的性能会大幅下降.

括而言之,目前大多数基于深度学习的点云识别与分割的方法缺乏生成卷积核权重的自适应性,并且网络模型过度关注提升精度,忽略了网络在点云数据变化时鲁棒性较差的问题,致使模型性能下降.为此,本文提出了自适应生成卷积核的动态图注意力神经网络架构.主要创新点和贡献如下:

1)提出将中心点位置信息嵌入邻点,并利用改进的自注意力机制重构感受野,使得感受野内特征信息充分交流,提升感受野模拟点云局部结构的能力.

2)构建动态图注意力卷积,利用点云特征信息增强卷积核,并自适应生成卷积核权重,提高网络在点云数据变换时的鲁棒性.

1 相关研究工作

目前,三维点云的识别及分割算法主要分为三类:基于多视图、基于体素与基于点云的方法.

1.1 基于多视图的方法

由于标准的卷积神经网络不能直接应用于类似三维点云这样的非结构化数据,基于投影的方法将3D点云从多个视角投影到2D平面上,融合二维卷积提取的多个视图特征,得到预测结果.MVCNN[14]通过12个不同的视图呈现三维模型,并分别利用卷积操作提取视图的特征,再将其跨视图融合后获得全局特征描述符.但该方法会造成特征冗余,且忽略了各个视角间的特征关联.Lawin等[15]从多个虚拟视角将点云投影到平面上,使用多流全卷积网络进行像素级语义分割,再将分割结果反投影到点云进行融合得到点的语义标签.基于SqueezeNet[16]提出了SqueezeSeg[17]和SqueezeSegV2[18],其利用了球面投影技术捕获点云信息.SqueezeSeg利用SqueezeNet提取投影得到前视图特征,条件随机场(conditional random field, CRF)作为递归层优化判别模型.SqueezeSegV2设计上下文聚合模块,并使用无监督域自适应训练管道,减少了合成数据和真实数据之间的分布差距.基于多视图的方法解决了将神经网络从二维数据推广至三维数据的问题,但由于该方法是将获取的多个视点的图像进行特征融合,因此会受到不同视点和投影角度选择的影响,不可避免地会丢失点云的几何和结构信息.

1.2 基于体素的方法

基于体素的方法将不规则的点云数据转化为一系列占用一定空间的体素,利用标准的3D卷积分析和处理.VoxelNet[19]将点云划分为等间距的体素,并堆叠体素特征编码(voxel feature encoding, VFE)层对每个体素进行编码,再利用3D卷积进一步聚集局部体素特征,将点云转换为高维体积表示,然后连接到区域候选网络(region proposal network,RPN)以生成检测结果.体素化数据与二维图像数据相比,其计算开销更大,效率更低.为有效克服VoxNet[20]等方法在体素规模上的限制,PointGrid[21]采用补"0"与随机采样的方式,为每个体素单元采样具有结构化分布的恒定数量的点.为了降低计算量和内存占用,Riegler等[22]采用八叉树的方法构建OctNet,集中计算和分配存储器,保持分辨率不变的同时构建更深层网络;Zeng等[23]基于Kd-tree构建了提高点云计算和存储效率的Kd-Net.但树结构方法过于依赖体素边界,未充分利用其局部几何结构. point-voxel transformer(PVT)[24]采用稀疏窗口注意(SWA)模块,从非空体素中收集粗粒度的局部特征,避免了复杂的不规则数据结构和无效的空体素计算,同时利用两种不同的自注意力变体,根据不同的点云尺度,挖掘全局细粒度特征.

1.3 基于点云的方法

以上两种方法对点云加以转换再进行处理,无疑会导致局部特征提取不充分以及增加计算量.PointNet是直接对点云数据进行处理的开山之作,使用多个共享的全连接层来处理输入的无序三维点云,通过通道级最大池化提取全局特征来表示三维点云数据.PointNet的本质是学习输入对象的关键点特征,其局部几何信息不是直接编码的;此外,其对输入数据的平移和缩放变化很敏感.由于不能很好地提取局部精细特征,PointNet++[25]通过采样分组,对每个局部空间采用PointNet的基础结构提取局部特征.尽管PointNet++设计了局部特征提取模块,但其仍然忽略了单个点与其邻域点的几何关系.

Qian等[26]提出了自适应采样和动态权重调整方法,解决了不同类别之间的样本不平衡问题,提高了PointNet++模型在点云数据上的训练效果.Zhou等[27]提出的自适应图卷积网络通过动态学习点的特征生成自适应核,改进了先前算法较简单的为邻域点分配不同权重的方法,提高了图卷积的灵活性和精确性.受注意力机制思想的启发,Wu等[28]引入了注意力机制对点云边缘采样,首先通过注意力网络对点云中的每个点进行编码,然后利用特征表示计算每个点的注意力权重,最后对点云进行采样,选择具有较高权重的点作为边缘点,从而实现了对点云边缘的有效提取.ECC[7]在空间域中构建了图卷积神经网络,卷积核权值依赖图结构中边的值,并对每个特定的输入进行动态更新,且在每个边值上都加入卷积核权值的条件.KCNet[10]将卷积核定义成一组可学习三维点,通过点云位置计算得到最近邻域,然后递归地进行特征聚合,提取特征信息.KPConv[12]引入了可变形卷积核的概念,每个卷积核由一组预定义的关键点组成,其位置和形状可以自适应地进行调整,通过计算输入点云中的每个点与卷积核关键点之间的关系,实现点卷积操作.DGCNN[13]采用动态图卷积层,该层通过K-最近邻(K-nearest neighbor, KNN)算法找出每个节点的邻点,并根据邻点的位置信息动态生成邻接矩阵,此动态性使得模型能够适应图的局部结构变化,提高了模型对异构图的处理

能力.MRTNet[29]通过在多个分辨率级别上表示三维形状,并使用空间分割树对点云进行空间排序,维持三个不同分辨率的点云表示,并进行上采样和池化操作,使之能够结合全局和局部特征信息,但其对于稀疏或不均匀分布的点云需要额外的预处理步骤,同时,多分辨率架构增加了网络的复杂性和训练时间.SO-Net[30]用自组织网络SOM建模输入点云,然后聚合成一个全局向量.SOM是一种无监督的聚类方法,第一步选择激活节点,第二步更新激活节点及临近节点权值.该方法具有置换不变性的优点,但其训练较为复杂.SPH3D-GCN[31]提出了一种基于球面的离散卷积核,用于在3D空间中对点云进行图卷积操作.球面核的性能对超参数选择敏感,需要额外的调整和实验.文献[32]提出了自监督学习点云特征的方法,该方法首先构建图结构,再利用图卷积神经网络对图进行特征学习,并设计自监督任务引导网络,采用自编码器的思想,对部分节点的特征进行掩码处理,重构被掩码的节点特征,从而使网络学习到语义信息丰富的特征表示,用于后续的点云识别及分割等任务.文献[33]通过动态图卷积操作构造边缘特征,并引入空间注意力模块建立邻域点之间依赖关系,最后加入通道注意力模块,为不同通道赋予不同权值,使网络聚焦于有用通道.但其使用KNN算法构造局部邻域,对于稀疏点云数据网络鲁棒性较差.

综上所述,基于多视图与基于体素的方法均是将点云转换后输入神经网络,忽略了点云本身的空间位置信息;基于点云的方法直接利用原始点云训练网络模型,缺乏对局部特征的深入挖掘;图卷积所使用的图数据在结构上类似于点云,可以很好地模拟点云的局部结构.当数据发生变换时,这些网络模型的识别精度大幅降低,即网络鲁棒性较差,且不能自适应地学习卷积核.因此,如何设计一个在数据变化时依旧健壮,并且可以自适应地学习卷积核的图卷积神经网络是三维点云识别与分割研究的一个亟待解决的问题.

2 动态图注意力卷积神经网络

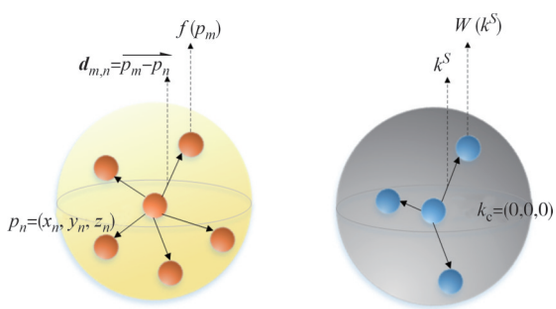
本文构建了自适应生成卷积核的动态图注意力卷积神经网络,主要利用改进的自注意力机制对感受野邻点特征信息进行增强,并提出动态图注意力卷积(dynamic attention graph convolution, DA-GraphConv)算子.首先通过特征增强模块(feature enhancement, F-E)构造增强特征信息的感受野,并构造自适应卷积核,然后利用动态图注意力卷积算子提取点云几何特征.

2.1 感受野构建

设N个点的三维点云表示为$P=\{p_n|n=1,2,\cdots,N,p_n\in\mathbb{R}^3\}$,$f(p)\in\mathbb{R}^d$表示每个点$p$的$d$维输出特征向量.图1为三维感受野和三维卷积核示意图.为了提取每个点p_n的局部几何信息,如图1(a)所示,定义M个相邻点来确定点p_n的三维感受野R_n^M.

$$R_n^M=\left\{p_n,p_m\middle|\forall p_m\in H(p_n,M)\right\} \tag{1}$$

式中,$H(p_n,M)$表示点p_n的M个基于欧氏距离$\|p_m-p_n\|_2$的最近邻点,邻点p_m和中心点p_n的相对位置用方向向量$\boldsymbol{d}_{m,n}=\overrightarrow{p_m-p_n}$表示.点$p_n$的大小为$M$的感受野内的特征表示为$\left\{f(p_n),f(p_m)\middle|\forall p_m\in H(p_n,M)\right\}$.文中使用KNN算法构建感受野.文献[12]预定义半径为r的球体,并为每个点选择相邻的点.该方法虽然可以通过调整参数r来正确地描述三维点云的局部结构,但模型很容易受到不同密度点云的影响,在点云发生尺度变化时网络鲁棒性较差.



(a)三维感受野 (b)三维卷积核

图1 三维感受野和三维卷积核示意图

Fig.1 Illustration of the 3D receptive field and 3D convolutional kernel

为了使感受野的中心点与邻点可以交互空间位置信息,本文提出将中心点空间位置信息传入邻点,融合二者空间位置信息后,通过双线性插值(bilinear interpolation)操作对齐特征维度,将原感受野信息与加强后的感受野信息利用自注意力机制重新分配权重,使网络聚焦于需要关注的部分,进一步提升网络性能.原始的自注意力机制旨在利用自身聚焦来提高网络识别和分割精度,使用自身生成自注意力机制的查询向量$\boldsymbol{q}$,值向量$\boldsymbol{v}$与键向量$\boldsymbol{k}$值进行计算.本

文对其进行改进，如图2所示，利用自身生成注意力机制的查询向量$\boldsymbol{q}$，并利用加强特征信息后的感受野生成值向量$\boldsymbol{v}$与键向量$\boldsymbol{k}$，进一步构建特征信息丰富的感受野.

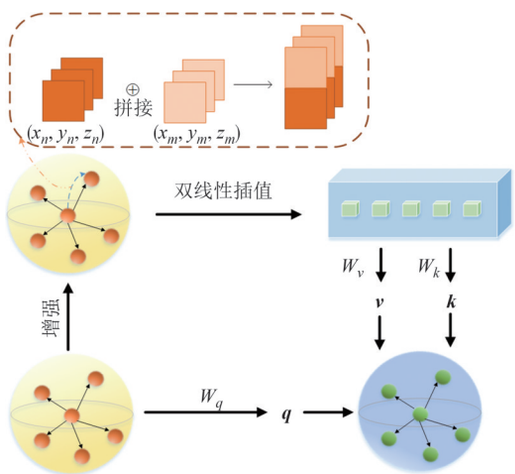


图2　增强感受野

Fig.2　Strengthened receptive field

2.2　自适应可变形卷积核

对二维图像的特征提取大多使用固定大小的卷积核，对于三维点云，使用固定大小的卷积核会限制网络提取特征的能力.因此，本文提出一种自适应生成的可变形卷积核K^S，S表示卷积核K中邻域点的数量.

$$K^S=\{k_C,k_1,k_2,\cdots,k_S\} \tag{2}$$

式中：k_C=(0, 0, 0)是卷积核的中心点；k_1到k_S表示卷积核中心点相关的邻点.为每个卷积核中的点定义权值向量$W(k)\in\mathbb{R}^D$；利用相应的权值向量对特征$f(p)$加权即可实现卷积运算.如图1(b)所示，三维卷积核定义为$\{W(k_C),W(k_s)|s=1,2,\cdots,S\}$，其中每一项权重都是通过训练来学习的，因此内核是可变形的，使得三维图卷积能够充分提取点云特征.

点云所包含的信息分为两类，一类是位置信息，另一类是特征信息.目前已有的大量研究工作都只使用了点云的位置信息，而忽略了点云的特征信息，例如颜色、强度信息等，这就导致了局部特征提取不充分和对点云信息未充分利用的问题.因此，在感受野的构建中，已利用了点云的位置信息，增强了感受野中心点与邻点的相互感知能力.在卷积核的构造过程中，将点云的特征信息加以利用，如公式(3)所示，通过通道传输点云特征信息，增强卷积核，使其更好模拟感受野，以提升网络精度.

$$\begin{aligned}&\mathrm{ConvLayer}\left(\left(p,R,G,B,\boldsymbol{F}^{\mathrm{in}}\right),K_i^S\right)=\\&\quad\left(p,R,G,B,\boldsymbol{F}_i^{\mathrm{out}}\right)\end{aligned} \tag{3}$$

式中：输入特征为$\boldsymbol{F}^{\mathrm{in}}\in\mathbb{R}^{N\times D}$；输出特征为$\boldsymbol{F}^{\mathrm{out}}\in\mathbb{R}^{N\times L}$；通道$i$=1，2，…，$L$，($R$，$G$，$B$)表示点云的颜色信息.由式(3)可以看出，将点云颜色信息与点云位置信息一同传入网络进行学习，使网络能够充分提取更加全面的特征.

2.3　动态图注意力卷积

通过计算R_n^M和K^S之间的相似性来定义动态图注意力卷积$\mathrm{Conv}(R_n^M,K^S)$，如图3所示，由于图与图之间没有明显的一一对应关系，因此考虑所有可能的(p_m,K^S)一一对应关系.由此，动态图注意力卷积定义为：

$$\mathrm{Conv}\left(R_n^M,K^S\right)=\left\langle f\left(p_n\right),W\left(k_C\right)\right\rangle+g(A) \tag{4}$$

式中：<，>表示内积；函数g表示卷积核中每个中心点k_s与所有相邻点之间的最大相似度值，即sim值，具体如下：

$$\begin{aligned}&g(A)=\sum_{s=1}^{S}\max\nolimits_{m\in(1,M)}\left\{\mathrm{sim}\left(p_m,k_s\right)\right\},\\&A=\left\{\mathrm{sim}\left(p_m,k_s\right)\middle|\,\forall m\in(1,M),\forall s\in(1,S)\right\},\\&\mathrm{sim}\left(p_m,k_s\right)=\left\langle f\left(p_m\right),W\left(k_s\right)\right\rangle\frac{\left\langle\boldsymbol{d}_{m,n},k_s\right\rangle}{\left\|\boldsymbol{d}_{m,n}\right\|\left\|k_s\right\|}\end{aligned} \tag{5}$$

式中：利用余弦相似度$\cos\theta=\dfrac{\left\langle\boldsymbol{d}_{m,n},k_s\right\rangle}{\left\|\boldsymbol{d}_{m,n}\right\|\left\|k_s\right\|}$得出$p_m$与$k_s$的相似性大小.可以看出，较大的向量间相似度sim值是由特征空间$\mathbb{R}^D$的较大内积和欧几里得空间$\mathbb{R}^3$的更为相似的方向向量产生的.

式(4)由两项组成，第一项为感受野中心点p_n与卷积核中心点k_C的卷积计算，第二项为S组最大相似度累加求和，构成感受野邻域点p_m与卷积核邻点k_s的卷积计算.函数A表示各组相似度函数计算的具体过程，由感受野邻点特征与卷积核邻点权重求内积后与余弦相似度函数相乘得到一组相似度结果.式(4)中第一项为常规卷积，如图3中a部分所示，第二项卷积过程如图3中b、c两部分所示.图3以两个邻域点构成的感受野说明整个动态图卷积过程，相似值为100表示感受野与学习到的卷积核完全相同.以卷积核中红色向量为基准，计算卷积核中红色向量与感受野中蓝色向量之间的余弦相似度，进而得到卷积结果.当卷积核中红色向量在图中位置1时(见图3中b部分)，感受野中蓝色向量在位置2或位置3，分别计算两组余弦相似度值，由式(5)可以看出，取两组中较大的余弦相似度值作为加和值.同

理,当卷积核中红色向量在图中位置4时(见图3中c部分),感受野中蓝色向量在位置5或位置6,分别计算两组余弦相似度值,取两组中较大的余弦相似度值作为加和值.

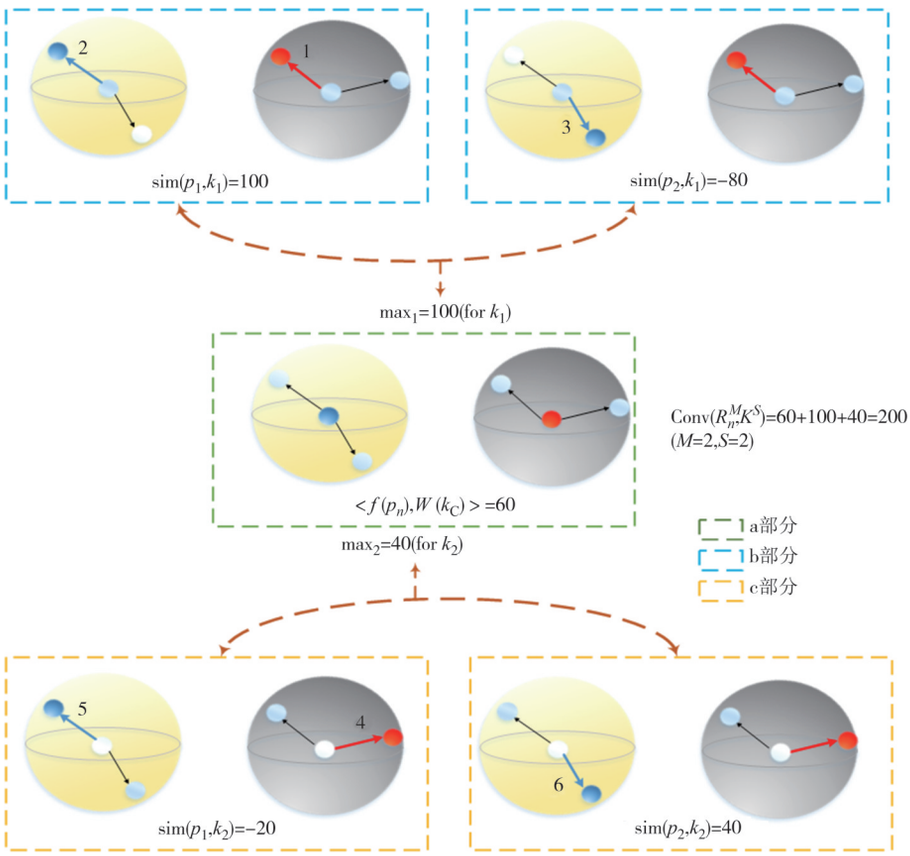


图3 动态图注意力卷积

Fig.3 Dynamic graph attention convolution

为了评估所构建的自适应生成可变形三维卷积核和动态图卷积神经网络架构对点云识别及分割的有效性,选用ModelNet40和ShapeNet Part数据集进行实验.本文实验硬件条件为Intel Core i9−10900K CPU,NVIDIA RTX 3090 GPU(显存24 G),软件为Linux Ubuntu 18.04和深度学习框架PyTorch1.11.0.实验中采用随机梯度下降SGD优化算法(动量因子为0.99),并使用Adam优化器更新SGD步长.网络识别性能评价指标采用总体准确率(overall accuracy,OA),网络分割性能的评价指标采用每一个部件的评价指标为交并比(mIoU)计算实例平均交并比(instance mean intersection over union,Inst.mIoU)和类别平均交并比(class mean intersection over union,Cls.mIoU),公式如下:

$$\mathrm{OA}=\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}} \tag{6}$$

$$\mathrm{Inst.mIoU}=\frac{\sum_{i=1}^{s}\sum_{i=1}^{c}\mathrm{TP}}{\sum_{i=1}^{s}\sum_{i=1}^{c}\mathrm{TP}+\sum_{i=1}^{s}\sum_{i=1}^{c}\mathrm{FP}+\sum_{i=1}^{s}\sum_{i=1}^{c}\mathrm{TN}} \tag{7}$$

$$\mathrm{Cls.mIoU}=\frac{1}{s}\sum_{i=1}^{s}\frac{\sum_{i=1}^{c}\mathrm{TP}}{\sum_{i=1}^{c}\mathrm{TP}+\sum_{i=1}^{c}\mathrm{FP}+\sum_{i=1}^{c}\mathrm{TN}} \tag{8}$$

式中:s表示类别数目;c表示类别中点的个数;TP为被模型预测为正类的正样本;FP为被模型预测为正类的负样本;TN为被模型预测为负类的负样本;FN为被模型预测为负类的正样本.

2.4 网络架构

识别任务的网络结构如图4所示.图中字符(3,D1)中,第一项"3"表示将点云位置信息(x, y, z)输入网络,共三个通道,第二项中"D"表示特征维度.首先利用动态图注意力卷积提取点云特征,将点云输入网络后利用F-E模块增强感受野,再利用本文构建的动态图注意力卷积提取特征,同时将点云特征

信息通过通道传输到网络，进一步提升卷积核模拟感受野的能力．经过两层卷积后，通过图池化操作聚合特征，再使用采样率为N/r的下采样，其中N为点云点数总和，实验中取r=4，指将点云点数减少至1/4，达到提取深层语义信息的目的．而后堆叠两层动态图注意力卷积，再利用采样率为N/r^2的下采样进一步减少点数，经过一层卷积后，通过全局最大池化获得代表性特征，然后添加多层感知机（multi-layer perceptron，MLP）来预测理想的输出．识别模型采用动态结构，将点云颜色信息通过通道传输到每层卷积，动态匹配特征维度，同时学习更新卷积核权重，挖掘深层语义信息．

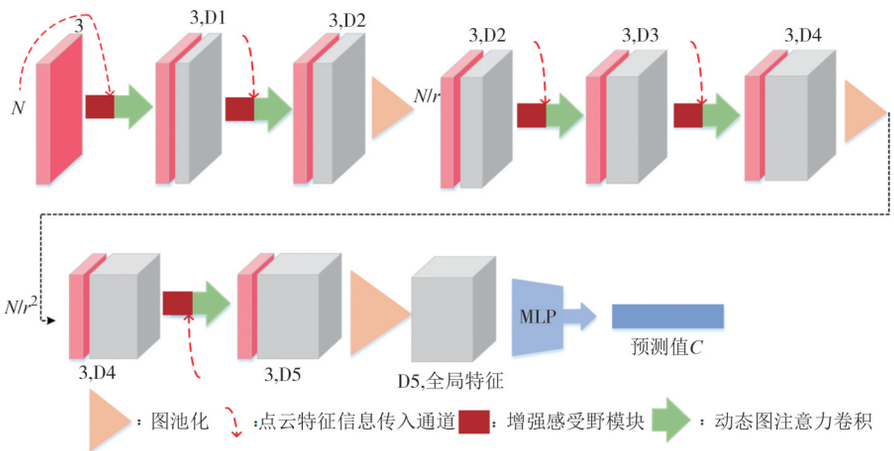


图4　识别任务的网络结构

Fig.4　Network architecture for recognition tasks

分割任务的网络结构如图5所示．首先利用两层动态图注意力卷积提取特征，然后经过图池化操作，目的是保留主要特征的同时减少参数量，再使用采样率为N/r的下采样进行特征提取，增大感受野，使后续网络能够学习到全局特征，重复堆叠两层动态注意图卷积与图池化操作，提取深层次语义信息，再使用采样率为N/r^2的下采样提取点云特征．下采样学习不同尺度特征，提取丰富点云特征．对前五层神经网络提取的特征，采用插值（Interpolation）的方法提高特征维度，经过上采样（Up-sample）还原点云个数后，与网络已得到的特征合并，再经过动态图注意力卷积提取特征，获得最终分割结果．分割模型设计为U形结构，并用图池化和插值操作来构建层次结构，图池化操作如图6所示．

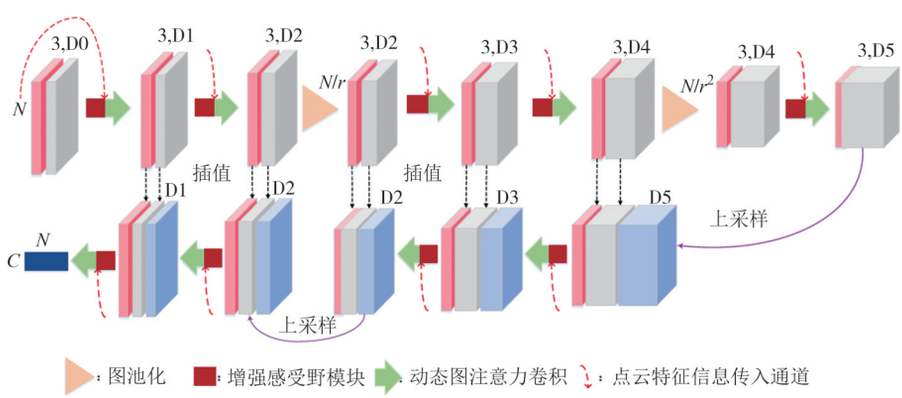


图5　分割任务的网络结构

Fig.5　Network architecture for segmentation tasks

从点云识别任务的网络架构可以看出，网络采用本文构造的动态图注意力卷积进行特征提取，利用相对坐标构造动态图注意力卷积算子，在数据变换时相对坐标不发生改变，因而网络具有良好的鲁棒性，并在每次特征提取前利用F-E模块增强感受野，通过5次动态图注意力卷积逐步扩大感受野，逐层提取丰富的语义信息．分割任务的网络框架同样采用动态图注意力卷积捕获点云特征信息，通过下

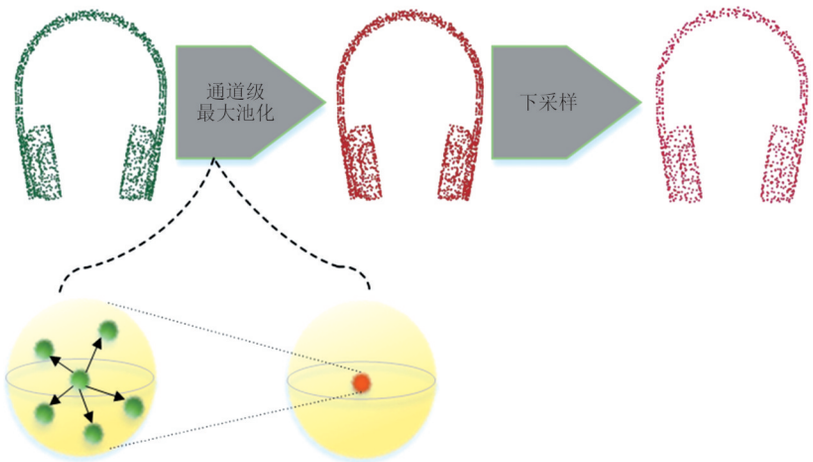


图6 图池化示意图

Fig.6 Illustration of the graph pooling

采样与上采样获取更为精确的点云特征,从而进行对模型的分割任务.

识别任务需要识别出点云模型,并输出所属类别,分割任务需在识别点云模型后,将组成模型的各个部件进行再次分割.因此分割网络在识别网络的基础上增加了上采样的过程,意在通过增加点的数量来提高点云数据的密度,从而更好地保留细节信息,使得分割网络可以更好地感知理解整个点云,从而提高分割结果的准确性,同时,上采样操作也可以帮助网络更好地处理不同尺度的特征,使网络对不同物体都能有较好的分割效果.

本文与现有工作之间技术要点的区别如表1所示.首先,三维点云的感受野$H(p_n)$的定义在不同的工作中有所不同.PointNet++和KPConv将它们定义为$M(p_n)=\left\{p_m \middle| p_m \in P, \left\| p_m - p_n \right\| \leqslant z\right\}$,其中$z$是为不同数据集设计的超参数.DGCNN在派生的特征空间中选择KNN算法.本网络选择在三维坐标空间中以KNN的方式定义$H(p_n)$.因此,其在尺度变化的同时保持局部几何信息不变.另外,文献[12]、[13]、[24]的工作将坐标差(p_j-p_i)作为函数s的输入,因此得到的输出对尺度变化很敏感.由表1可以看出,本文算法聚焦于局部三维点云之间的方向信息进行卷积,这使网络能够捕捉几何特征,同时具有位移、尺度和旋转不变性的特性.

表1 不同方法的比较

Tab.1 Comparison of different methods

方法	池化机制	邻点$H(p_n)$	聚合函数	卷积层特征
PointNet++[25]	√	$\left\{p_m \middle| p_m \in P, \left\| p_m - p_n \right\| \leqslant z\right\}$	max()	$\left\langle p_j - p_i, k^\theta \right\rangle + \left\langle f_j, w^\theta \right\rangle$
DGCNN[13]	×	用特征空间KNN定义	max()	$\left\langle f_j - f_i, w_1^\theta \right\rangle + \left\langle f_i, w_2^\theta \right\rangle$
KPConv[12]	√	$\left\{p_m \middle| p_m \in P, \left\| p_m - p_n \right\| \leqslant z\right\}$	sum()	$\max\left(0, 1 - \frac{\left\| k^\theta - p_j + p_i \right\|}{\sigma}\right) \cdot \left\langle f_j, w^\theta \right\rangle$
本文	√	用三维空间KNN定义	max()	$\frac{\left\langle p_j - p_i, k^\theta \right\rangle}{\left\| p_j - p_i \right\| \left\| k^\theta \right\|} \cdot \left\langle f_j, w^\theta \right\rangle$

3 实验结果与分析

3.1 识别任务

ModelNet40[5]数据集包括40个类别的12 311个CAD模型,其中,用于训练的模型为9 843个,用于测试的模型为2 468个.使用总体精度(OA)作为识别任务的性能评估指标.

将本文算法与其他代表性算法进行识别实验比较,结果如表2所示.可以看出,当测试数据没有任何变化时,本文算法具有较好的点云识别能力.本文算法为基于点云的方法,无须转换点云信息,并且无预处理过程,减少了点云处理过程中信息丢失,网络精度大幅领先于Kd-Net,这是因为其参考Kd-tree的

结构,每次网络需要前向操作时,需将点云转为Kd-tree结构.与基于点云的方法对比,比PointNet分割精度提升4.8个百分点,主要由于其未考虑局部特征提取,无法应用于需要识别精细特征的任务.本文对比了PointNet++模型的两个版本,一个输入原始点云,即表2中输入为*xyz*的模型,另一个输入归一化后的点云,即表2中输入为*xyz*, normal的模型.PointNet++在PointNet基础上进行改进,通过分层的特征提取逐步扩大感受野,提取不同层次的特征,其选用采样部分点的方式训练网络,并未覆盖全部点云,因此本文精度比原始点云高出3.3个百分点.

同基于体素的方法PVT网络相比,将点云转换为体素需要额外的操作,导致局部特征提取不充分以及计算量增加的问题,因此本网络精度高于该网络.与同样采用图卷积方法的DGCNN相比,DGCNN共享卷积核参数,而本文构建自适应可变形卷积核,网络精度更佳;与AGConv[35]相比,本文增添了F-E模块增强感受野,加强了感受野模拟点云的能力,进一步提升网络精度.本网络的识别可视化结果如图7所示.

表2　ModelNet40数据集上识别精度对比

Tab.2　Comparison of recognition accuracy on the ModelNet40 dataset

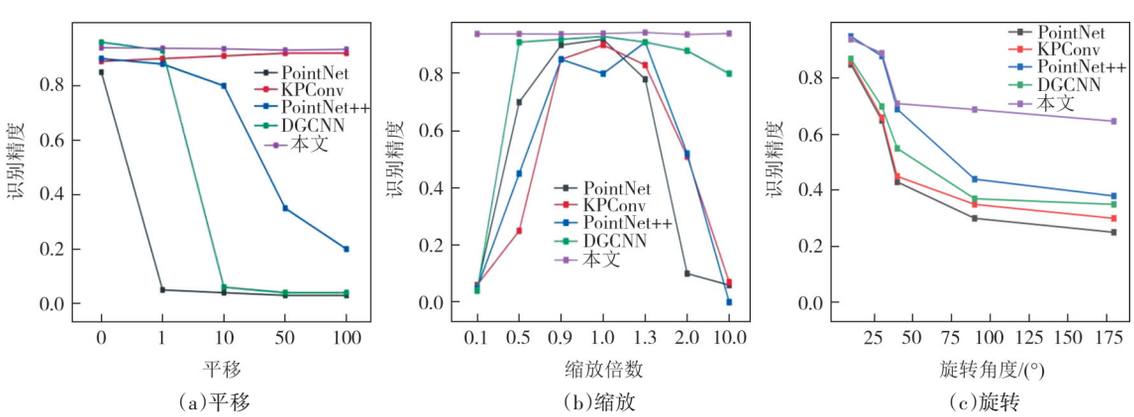
算法	输入	输入点数	OA/%
PointNet[6]	*xyz*	1 k	89.2
Kd-Net[34]	voxel	—	90.6
PVT[24]	voxel	—	93.7
ECC[7]	*xyz*	1 k	87.4
KCNet[10]	*xyz*	1 k	91.0
MRTNet[29]	*xyz*	1 k	91.2
PointNet++[25]	*xyz*	1 k	90.7
DGCNN[13]	*xyz*	1 k	92.9
SO-Net[30]	*xyz*	2 k	90.9
KPConv[12]	*xyz*	6.8 k	92.9
SPH3D-GCN[31]	*xyz*	10 k	92.1
PointNet++[25]	*xyz*, normal	5 k	91.9
SO-Net[30]	*xyz*, normal	5 k	93.4
AGConv[35]	*xyz*	1 k	93.4
APES[28]	—	—	93.8
本文	*xyz*	1 k	**94.0**



图7　ModelNet40数据集上识别可视化结果

Fig.7　Visualization of recognition on the ModelNet40 dataset

为了进一步评估本文算法在数据变化时的鲁棒性,分别通过平移、缩放和旋转对模型进行变换,并测试不同算法的识别精度,对比实验结果分别如图8(a)、8(b)、8(c)所示.如图8(a)所示,PointNet和DGCNN的性能随着坐标平移而显著下降,这是由于其从全局坐标中提取特征引起的性能下降.PointNet++架构中联合使用了三维点云的局部相对坐标和全局坐标,点云平移时,全局坐标发生改变,致使网络精度下降幅度较大.可以看出,随着坐标平移,本文模型精度基本保持不变.当对模型进行缩放变换时,只有本文模型能够以令人满意的性能进行识别,如图8(b)所



(a)平移　(b)缩放　(c)旋转

图8　ModelNet40上的数据不变性评估

Fig.8　Evaluation of invariance properties on the ModelNet40

示,由于动态图注意力卷积采用局部坐标设计卷积算子,当点云按比例缩放一定倍数时,局部坐标按比例更改,因此网络相对健壮.当对模型旋转一定角度时,本文算法表现出更好的鲁棒性,如图8(c)所示,这是因为本文算法利用余弦相似度函数构造卷积算子,当点云数据整体旋转一定角度时,两向量夹角不变,向量模不变,且具有自适应学习卷积核权重的能力.图9展示了本文模型在ModelNet40数据集上的训练轮数与总体精度OA之间的关系,在60轮内收敛迅速,第265轮时识别精度最高.

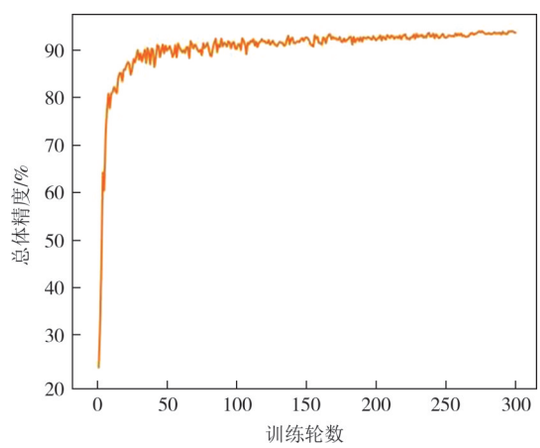


图9 训练轮数与总体精度的关系

Fig.9 The relationship of epoch and overall accuracy

3.2 分割任务

ShapeNet Part[36]数据集包含16种共15 881个CAD模型,每类包含2~6个不同部件,共50个部件语义标签.不同算法的部件分割结果如表3所示,其中每一个部件的评价指标为交并比(mIoU),本文网络取得了与现有方法基本相当或更好的分割结果.

SPH3D-GCN[31]通过将点云数据映射到球面上进行卷积操作,实现了高效的图卷积计算,而本文构造了精细的局部特征提取算子,随着网络的加深,出现特征冗余的情况,因此在摩托车等较为复杂的物体分割结果上表现欠佳.与DGCNN模型相比,本文算法融合了点云特征信息,提取了丰富的语义特征,自适应生成卷积核,减少了信息冗余,使得Cls.mIoU与Inst.mIoU分别提升了2.4与1个百分点.虽然精度略逊于KPConv和AGConv,但在车辆、耳机、吉他等类别实现了最佳分割性能.与需要预处理的网络模型Kd-Net和RS-Net相比,Kd-Net前向操作时,需将点云转为Kd-tree结构,RS-Net通过x、y、z三个方向的切片池化层将无序点云转换为有序序列,而本文模型输入原始点云,保留了点云原始信息,充分学习点云局部特征,各项评价指标提升幅度较大.实验结果表明本文方法在点云分割任务中取得了较好的分割效果.

表3 不同算法的分割实验结果对比

Tab.3 Comparison of segmentation results of different algorithms %

算法	类别平均交并比	实例平均交并比	飞机	杯子	包	车辆	椅子	耳机	小刀	吉他	灯	马克杯	摩托车	滑板	桌子	手枪
PointNet[6]	80.4	83.7	83.4	82.5	78.7	74.9	89.6	73.0	85.9	91.5	80.8	93.0	65.2	72.8	80.6	81.2
Kd-Net[34]	77.4	82.3	80.1	74.3	74.6	70.3	88.6	73.5	87.2	90.2	81.0	86.7	87.4	69.9	80.3	78.1
PVT[24]	84.0	—	84.3	82.1	**88.7**	**82.1**	**92.4**	75.5	88.5	91.0	85.6	94.7	76.2	75.3	81.7	84.2
MRTNet[29]	79.3	83.0	81.0	87.0	76.7	87.0	89.1	67.6	85.4	90.6	80.6	91.8	64.4	69.1	80.6	79.7
KCNet[10]	82.2	84.7	82.8	86.4	81.5	77.6	90.3	76.8	87.2	91.0	84.5	94.4	69.2	75.2	81.3	81.6
SO-Net[30]	81.0	84.9	82.8	88.0	77.8	77.3	90.6	73.5	83.9	90.7	82.8	94.2	69.1	72.9	83.0	80.9
RS-Net[37]	81.4	84.9	82.7	84.1	86.4	78.2	90.4	69.3	87.0	91.4	83.5	92.6	66.0	75.8	82.2	81.8
PointNet++[25]	81.9	85.1	82.4	87.7	79.0	77.3	90.8	71.8	85.9	91.0	83.7	94.1	71.6	76.4	82.6	81.3
DGCNN[13]	82.3	85.2	84.0	86.7	83.4	77.8	90.6	74.7	87.5	91.2	82.8	94.9	66.3	74.5	82.6	81.1
KPConv[12]	**85.1**	86.4	84.6	87.2	86.3	81.1	91.1	77.8	88.4	**92.6**	82.7	**95.8**	78.1	**82.0**	83.6	**85.4**
SPH3D-GCN[31]	84.9	**86.8**	84.4	**89.2**	86.2	81.4	91.5	77.4	88.2	92.5	**85.7**	95.6	**78.6**	78.5	84.0	84.7
AGConv[35]	83.4	86.4	84.8	85.7	81.2	79.7	91.2	**80.9**	88.6	91.9	84.8	94.9	70.7	75.9	**84.2**	82.3
本文	84.7	86.2	**84.4**	87.1	85.9	81.4	90.8	77.9	**88.6**	**92.6**	82.8	94.9	74.5	81.3	82.8	83.7

为了进一步验证本文算法的鲁棒性,对比分析了不同模型在平移、缩放和旋转变换下的分割可视化结果与定量结果,如表4与表5所示.从表4可以看出,PointNet++与DGCNN在平移、缩放和旋转变化时,不能产生较好的分割结果,而本文模型设计的卷积算子由全局坐标与余弦相似度函数构造,因此在数据变化时具有较好鲁棒性.图10展示了本文算法在ShapeNet Part数据集上的训练轮数与分割精度之间的变化关系,其于100轮时收敛,220轮时取得最佳的分割结果.

表 4　不同算法的分割实验可视化结果

Tab.4　Visualization of segmentation results of different algorithms

		PointNet++			DGCNN			本文		
类别	真实标签	平移	缩放	旋转	平移	缩放	旋转	平移	缩放	旋转
小刀										
手枪										
摩托车										
吉他										

表 5　ShapeNet Part 数据集上模型平移、缩放和旋转变化后的分割 cls. mIoU 对比

Tab.5　Part segmentation in terms of cls. mIoU with translate, scale and rotation variations on ShapeNet Part dataset　%

算法	平移				缩放				旋转/(°)			
	1	10	50	100	0.1	0.5	5	10	30	60	90	120
PointNet[6]	23.7	16.1	16.0	15.8	30.8	66.6	40.8	35.1	68.0	58.5	54.5	50.8
PointNet++[25]	43.3	23.9	17.6	15.1	29.1	54.7	38.8	36.2	74.5	66.6	62.0	58.8
DGCNN[13]	45.5	19.4	16.8	16.1	37.9	69.2	50.5	27.4	72.5	67.0	64.2	61.1
KPConv[12]	36.8	22.2	21.1	20.9	30.5	46.0	51.2	48.9	72.2	61.1	51.4	47.2
本文	84.1	84.4	83.9	84.2	84.4	84.1	84.2	84.2	82.8	80.6	69.3	66.4

图 10　训练轮数与分割精度

Fig. 10　Epoch and segmentation accuracy

3.3　消融实验

3.3.1　邻点数目

本文在构建局部邻域时使用 KNN 算法,邻点数目 M 越大,则感受野越大.其中,邻点数目 M 的值在网络进行局部特征提取时能够影响网络的性能,在构建局部邻域时若设置较小的邻点数目 M,局部邻域的构建速度较快,网络模型的计算量偏低,但邻域范围过小会使网络无法充分学习到有效的局部特征,导致网络模型精度较差;而当设置较多的邻域点数目时,局部邻域范围重叠,难免会学习到冗余的特征,影响网络模型学习几何特征的能力.此外,每个数据集的点云规模不同,不同数据集选择的邻点数 M 也存在差异性.本文消融实验选取 ModelNet40 数据集与 ShapeNet Part 数据集分别测试邻点数目 M 对识别精度与分割精度的影响,结果如表 6 所示.从表 6 中可以看出,识别任务在 M 为 25 时网络精度达到最佳,而分割任务在 M 为 50 时网络精度最高,这是由于 ShapeNet Part 点云分割数据集的输入点个数是 ModelNet40 点云识别数据集的 2 倍,因此当 M 较大时,分割网络能提取更丰富的几何特征.

3.3.2　增强感受野模块及动态图注意力卷积算子

为了进一步说明本文模型增强感受野模块以及

表 6　感受野大小对点云识别与分割的影响

Tab.6　Influence of the receptive field size on recognition and segmentation results　%

邻点数目 M	识别总体精度	分割平均交并比
10	93.1	85.4
25	**94.0**	85.9
50	93.6	**86.2**
60	93.2	85.7

动态图注意力卷积的有效性,分别在ModelNet40数据集与ShapeNet Part数据集上进行消融实验,结果分别如表7和表8所示.Model-1指在baseline上将标准图卷积替换为动态图注意力卷积后的网络;Model-2指在baseline上增添增强感受野模块后的网络.可以看出,由于动态图注意力卷积算子构造了自适应生成的卷积核,其可以随着点云局部结构的改变,自适应学习不同的卷积核权重,从而构建不同的卷积核,并将点云特征信息通过通道传入网络,使网络获得更加全面的点云信息,而标准图卷积共享卷积核权重,无法应对点云局部结构之间的变化,因此,Model-1相较于baseline网络精度增幅较大.增强感受野模块提升了感受野模拟点云局部结构的能力,使中心点与邻点可以交互空间位置信息,并用改进的自注意力机制进一步重构感受野,从而网络可以更充分地提取局部特征,因此Model-2相较于baseline精度提升.

表7 F-E模块和动态图注意力卷积在识别任务上的有效性

Tab.7 Validation of F-E module and dynamic attention graph convolution on recognition

网络模块	识别总体精度/%
Model-1	92.7
Model-2	91.6
baseline	90.8
本文	**94.0**

表8 F-E模块和动态图注意力卷积在分割任务上的有效性

Tab.8 Validation of F-E module and dynamic attention graph convolution on segmentation

网络模块	分割平均交并比/%
Model-1	84.1
Model-2	82.5
baseline	81.7
本文	**86.2**

在增强感受野模块中,利用改进的自注意力机制替代原始注意力层的有效性需进一步验证.消融实验在ModelNet40数据集与ShapeNet Part数据集上进行,定量结果如表9所示.原始注意力层从自身生成查询向量$\boldsymbol{q}$、值向量$\boldsymbol{v}$与键向量$\boldsymbol{k}$,本文利用原感受野生成查询向量$\boldsymbol{q}$与增强特征信息后的感受野生成值向量$\boldsymbol{v}$与键向量$\boldsymbol{k}$,关注了增强特征信息后的感受野的同时,将原感受野特征信息加以利用,从而获得具有更全面特征信息的感受野,进而提升网络精度.表10为消融实验在ShapeNet Part数据集上的可视化结果.

表9 F-E模块中改进的自注意力机制有效性验证

Tab.9 Validation of effectiveness of the improved self-attention mechanism in module F-E %

注意力机制	识别总体精度	分割平均交并比
原始自注意力机制	93.2	85.1
本文算法	**94.0**	**86.2**

表10 改进的自注意力机制有效性验证的可视化结果

Tab.10 Visualization results of validation of effectiveness of the improved self-attention mechanism

类别	小刀	手枪	摩托车	吉他
原始自注意力机制				
本文算法				
真实标签				

4 结 论

本文提出了自适应生成卷积核的动态图注意力三维点云识别及分割算法.首先,将感受野中心点位置信息传入邻点,加强感受野中心点与邻点的空间位置信息交流,再利用改进的自注意力机制重构感受野,进一步加强感受野模拟点云局部结构的能力;其次,将颜色、强度等点云特征信息通过通道传入网络,使网络获得更加全面的点云信息;然后,利用余弦相似度函数构建动态图注意力卷积算子,旨在使网络在数据变化时可以表现出较强的鲁棒性;最后,设计点云识别的动态网络架构及语义分割的U形网络架构.本网络模型在点云识别数据集ModelNet40上总体精度达到了94.0%,在点云分割数据集ShapeNet Part上平均交并比达到了86.2%,具有较好的点云识别与分割能力,并且在数据变化时网络具有较强鲁棒性.本文算法在小规模点云数据集上取得了较好的成绩,但在大规模场景语义分割任务上还存在较大的提升空间.此外,本文采用KNN算法构建感受野,由于稀疏点云数据具有分布不均匀的特性,从而影响点云识别与分割的结果.因此,如何提高本方法对场景语义分割的能力以及如何处理稀

疏点云数据,是将来要着重研究的内容.

参考文献

[1] 程云建,仇文革,雷劲.基于三维点云的隧道全局中线提取方法及应用[J].湖南大学学报(自然科学版),2017,44(9):146-150.
CHENG Y J, QIU W G, LEI J. Application of method for global extraction of tunnel centerlines based on 3D point clouds [J]. Journal of Hunan University (Natural Sciences), 2017, 44(9): 146-150. (in Chinese)

[2] LUO W J, YANG B, URTASUN R. Fast and furious: real time end-to-end 3D detection, tracking and motion forecasting with a single convolutional net [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018, Salt Lake City, UT, USA: IEEE, 2018: 3569-3577.

[3] 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)

[4] 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 2016 European Conference on Computer Vision (ECCV). Cham, Switzerland: Springer, 2016: 628-644.

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

[6] 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). July 21-26, 2017, Honolulu, HI, USA: IEEE, 2017: 77-85.

[7] SIMONOVSKY M, KOMODAKIS N. Dynamic edge-conditioned filters in convolutional neural networks on graphs [C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE, 2017: 29-38.

[8] MASCI J, BOSCAINI D, BRONSTEIN M M, et al. Geodesic convolutional neural networks on Riemannian manifolds [C]//2015 IEEE International Conference on Computer Vision Workshop (ICCVW). December 7-13, 2015, Santiago, Chile: IEEE, 2015: 832-840.

[9] BOSCAINI D, MASCI J, RODOLÀ E, et al. Learning shape correspondence with anisotropic convolutional neural networks [EB/OL]. (2016-05-20) [2023-09-14]: 1605.06437. https://arxiv.org/abs/1605.06437v1.

[10] SHEN Y R, FENG C, YANG Y Q, et al. Mining point cloud local structures by kernel correlation and graph pooling [C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 4548-4557.

[11] LIU Y C, FAN B, XIANG S M, et al. Relation-shape convolutional neural network for point cloud analysis [C]// Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 15-20, 2019, Long Beach, CA, USA: IEEE, 2019: 8887-8896.

[12] THOMAS H, QI C R, DESCHAUD J E, et al. Kpconv: flexible and deformable convolution for point clouds [C] // Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South): IEEE, 2019: 6410-6419.

[13] 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.

[14] 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). December 7-13, 2015, Santiago, Chile: IEEE, 2015: 945-953.

[15] LAWIN F J, DANELLJAN M, TOSTEBERG P, et al. Deep projective 3D semantic segmentation [M]//FELSBERG M, HEYDEN A, KRÜGER N, eds. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2017: 95-107.

[16] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size [EB/OL]. (2016-02-24) [2023-09-14]: 1602.07360. https://arxiv.org/abs/1602.07360v4.

[17] WU B C, WAN A, YUE X Y, et al. SqueezeSeg: convolutional neural nets with recurrent CRF for real-time road-object segmentation from 3D LiDAR point cloud [C]// 2018 IEEE International Conference on Robotics and Automation (ICRA). Brisbane, Australia: IEEE, 2018: 1887-1893.

[18] WU B C, ZHOU X Y, ZHAO S C, et al. SqueezeSegV2: improved model structure and unsupervised domain adaptation for road-object segmentation from a LiDAR point cloud [C]// 2019 International Conference on Robotics and Automation (ICRA). Montreal, Canada: IEEE, 2019: 4376-4382.

[19] ZHOU Y, TUZEL O. VoxelNet: end-to-end learning for point cloud based 3D object detection [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018, Salt Lake City, UT, USA: IEEE, 2018: 4490-4499.

[20] 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.

[21] LE T, DUAN Y. PointGrid: a deep network for 3D shape understanding [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018, Salt Lake City, UT, USA: IEEE, 2018: 9204-9214.

[22] 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). July 21-26, 2017, Honolulu, HI, USA: IEEE, 2017: 6620-6629.

[23] ZENG W, GEVERS T. 3DContextNet: K-d tree guided hierarchical learning of point clouds using local and global contextual cues[M]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2019: 314-330.

[24] ZHANG C, WAN H C, SHEN X Y, et al. PVT: point-voxel transformer for point cloud learning[J]. International Journal of Intelligent Systems, 2022, 37(12): 11985-12008.

[25] QI C, YI L, SU H, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space[EB/OL]. (2017-06-07)[2023-09-14]: 1605.06437. https://arxiv.org/pdf/1706.02413.

[26] QIAN G, LI Y, PENG H, et al. Pointnext: revisiting pointnet++ with improved training and scaling strategies[J]. Advances in Neural Information Processing Systems, 2022, 35: 23192-23204.

[27] ZHOU H R, FENG Y D, FANG M S, et al. Adaptive graph convolution for point cloud analysis [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). October 10-17, 2021, Montreal, QC, Canada: IEEE, 2021: 4945-4954.

[28] WU C Z, ZHENG J W, PFROMMER J, et al. Attention-based point cloud edge sampling[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 17-24, 2023, Vancouver, BC, Canada: IEEE, 2023: 5333-5343.

[29] GADELHA M, WANG R, MAJI S. Multiresolution tree networks for 3D point cloud processing[M]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2018: 105-122.

[30] LI J X, CHEN B M, LEE G H. SO-net: self-organizing network for point cloud analysis [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018, Salt Lake City, UT, USA: IEEE, 2018: 9397-9406.

[31] LEI H, AKHTAR N, MIAN A. Spherical kernel for efficient graph convolution on 3D point clouds[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(10): 3664-3680.

[32] 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). June 15-20, 2019, Long Beach, CA, USA: IEEE, 2019: 10288-10297.

[33] 杨军, 张琛. 融合双注意力机制和动态图卷积神经网络的三维点云语义分割[J/OL]. 北京航空航天大学学报: 1-12[2023-09-09]. DOI: 10.13700/j.bh.1001-5965.2022.0775.
YANG J, ZHANG C. Semantic segmentation of 3D point cloud by fusing dual attention mechanism and dynamic graph convolution neural network [J/OL]. Journal of Beijing University of Aeronautics and Astronautics: 1-12 [2023-09-09]. DOI: 10.13700/j.bh.1001-5965.2022.0775(In Chinese)

[34] KLOKOV R, LEMPITSKY V. Escape from cells: deep Kd-networks for the recognition of 3D point cloud models[C]//2017 IEEE International Conference on Computer Vision (ICCV). October 22-29, 2017, Venice, Italy: IEEE, 2017: 863-872.

[35] WEI M Q, WEI Z Y, ZHOU H R, et al. AGConv: adaptive graph convolution on 3D point clouds[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(8): 9374-9392.

[36] 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.

[37] HUANG Q G, WANG W Y, NEUMANN U. Recurrent slice networks for 3D segmentation of point clouds[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018, Salt Lake City, UT, USA: IEEE, 2018: 2626-2635.