

# DGPoint: 用于三维点云语义分割的动态图卷积网络

刘友群, 敖建锋\*, 潘仲泰

江西理工大学土木与测绘工程学院, 江西 赣州 341000

**摘要** 三维点云语义分割在三维场景理解和重建、自动驾驶和机器人导航等领域发挥着重要作用。针对现有方法对点云的局部特征提取不足导致分割精度不高这一问题, 基于 PointNet++ 架构提出了一个动态图卷积网络 DGPoint。首先, 边缘卷积的特征聚合函数使用双通道池化操作来补偿信息丢失, 可以更好地保留点云的细粒度局部特征; 其次, 在进行边缘卷积之前使用 K 近邻算法确定新的局部区域以达到动态图更新的效果; 此外, 为了保证边缘提取准确, 将设计的编码器重复多次, 并将提取的特征以跳跃连接的方式连接起来输入到解码器中。在 S3DIS 数据集上的实验结果表明, 与其他方法相比, DGPoint 有效解决了局部特征提取不足的缺点, 并提高了语义分割的准确性: 在 S3DIS 数据集的平均交并比为 68.3%, 总体准确率为 86.2%。

**关键词** 图像处理; 图卷积网络; 点云; 语义分割; 边缘卷积

中图分类号 TP391 文献标志码 A

DOI: 10.3788/LOP202259.1610014

## DGPoint: A Dynamic Graph Convolution Network for Point Cloud Semantic Segmentation

Liu Youqun, Ao Jianfeng\*, Pan Zhongtai

School of Architectural and Surveying Engineering, Jiangxi University of Science and Technology,  
Ganzhou 341000, Jiangxi, China

**Abstract** Semantic segmentation of point cloud data plays an important role for 3D scene understanding and reconstruction, autonomous driving, and robot navigation. In this study, DGPoint, a dynamic graph convolution network based on the PointNet++ architecture, is proposed to address the insufficient segmentation accuracy due to insufficient local feature extraction of point clouds by existing methods. First, the feature aggregation function in edge convolution compensates for loss using a dual-channel pooling operation, which can better retain the fine-grained local characteristics of the point cloud. Then, to accomplish the impact of dynamic graph updates, K-nearest neighbors algorithm is used to determine new local regions prior to edge convolution. Additionally, to ensure the accuracy of edge feature extraction, the designed encoder is repeated multiple times, and the extracted features are concatenated in a jump-connection style before being input to the decoder. Experimental results of the S3DIS data set show that DGPoint effectively solves the shortcomings of the insufficient local feature extraction and improves semantic segmentation accuracy with the mean intersection over union of 68.3% and overall accuracy of 86.2% compared with other methods.

**Key words** image processing; graph convolutional network; point cloud; semantic segmentation; edge convolution

## 1 引言

点云语义分割作为三维场景分析和理解的关键技术之一, 在三维重建、自动驾驶和机器人技术领域都发挥着重要作用<sup>[1-5]</sup>。自动驾驶车辆需实时精确地感知其周围环境中的物体及其运动状态, 以使车辆能够及时制定有效的驾驶策略以避免事故发生。对于机器人

导航来说, 对周围场景的重建和解析是导航的前提。与 2D 图像<sup>[6-7]</sup>不同, 点云具有非结构性、无序性、分布不均和数据量大等特点, 这使得点云难以使用常规方法进行处理分析, 因此用深度学习的方式实现点云语义分割受到了极大的关注, 然而, 如何有效地从无序点云中学习特征仍然是一个极具挑战性的问题。

早期, 为了解决点云无序性导致的特征学习困难,

收稿日期: 2021-09-30; 修回日期: 2021-11-05; 录用日期: 2021-11-29

基金项目: 国家自然科学基金青年科学基金(42004158)

通信作者: \*jiao008@163.com

许多研究者都致力于将 3D 点云数据转化为规则的结构表示,例如多视图投影和体素化。然而,投影不可避免导致信息丢失,体素化则会带来更大的计算成本。所以近年来研究者们将目光聚焦在以原始点云作为输入进行端到端的学习方法上,PointNet<sup>[8]</sup>是这种方法的开创性成果,但其在特征提取过程中并未考虑局部邻域信息。PointNet++<sup>[9]</sup>对其进行改进,采用分层抽取特征的思想,递归地将 PointNet 循环使用以提取点云局部特征,但在采样的区域内仍然使用了 PointNet,使得建立的局部几何特征对应关系效果不佳,从而影响了分割精度。

针对上述问题,本文工作主要考虑如何解决 PointNet++ 中局部特征提取不当的问题。提出了一个动态图卷积网络 DGPoint,该网络集成了 PointNet++ 和边缘卷积模块,通过使用 K 近邻(KNN)确定新的局部区域以达到动态图更新的效果。DGPoint 构造图的边缘上执行卷积以通过多个多层感知机(MLP)提取具有代表性的局部特征,然后通过两个不同的池化操作来补偿信息丢失。此外,DGPoint 将设计的特征编码块重复多次,把提取的特征连接起来输入到上采样层进行复原,从而提高语义分割的准确性。

## 2 基于深度学习的语义分割方法

点云语义分割的任务是为输入的每一个点云分配一个语义类别。与传统语义分割方法相比,基于深度学习的分割方法可以达到更高的准确度,且泛化性更好<sup>[10]</sup>。目前,基于深度学习的语义分割方法可以分为间接语义分割和直接语义分割两大类。

### 2.1 间接语义分割

多视图卷积神经网络(MVCNN)<sup>[11]</sup>被用于 3D 点云分割,该方法在投影上会丢失大量空间信息且受视图选择的影响,导致分割精度并不高。因此,Feng 等<sup>[12]</sup>在 MVCNN 的基础上提出了群视图卷积神经网络(GVCNN),可通过对不同视图下提取的视觉描述子进行分组来有效利用图像特征之间的关系。Ibrahim 等<sup>[13]</sup>提出了一种构建伪图像的方法,先将点云按角度拆分为多个 3D 切片,再将切片投影到 2D 网格上构建伪图像。RGB-D 数据因其特有的深度信息受到了语义分割领域研究者的青睐。SnapNet<sup>[14]</sup>对三维场景生成的二维快照进行语义标记后,再投影回原始点云上。SnapNet-R<sup>[15]</sup>则可以对多个视图直接处理以实现密集的三维点标记,但是生成二维快照的过程仍破坏了三维数据的内在几何关系。网络的轻量化一直以来都是研究的重点,SqueezeNet 作为轻量级网络架构,可以在提高效率的同时保持精度。在 SqueezeNet 的基础上,Wu 等<sup>[16]</sup>提出了 SqueezeSeg,该网络利用条件随机场作为递归层优化了分割结果,但受失调噪声影响使得模型泛化性较差。因此,SqueezeSegV2<sup>[17]</sup>通过添加上下文聚合模块(CAM)来

提取上下文信息,从而增强网络对失调噪声的鲁棒性。Kang 等<sup>[18]</sup>则通过使用自适应断点检测器(ABD)来减小将 2D 分割结果投影回 3D 点云时的精度损耗。

点云体素化是将不规则的点云转化为规则的体素。VoxNet 模型<sup>[19]</sup>最早将点云采样成体积网格后输入 Volumetric CNN 网络中实现点云分割,相较于二维图像数据,其计算开销更大,效率也更低。针对体素网格低分辨率的限制,SegCloud<sup>[20]</sup>网络引入了三线性插值(TI)和全连接条件随机场(FC-CRF)来实现细粒度和全局一致性。针对体素化带来的复杂计算问题,OctNet<sup>[21]</sup>和 VGS&SVGS<sup>[22]</sup>实现了在不影响网络性能条件下实现网络轻量化。MSNet<sup>[23]</sup>网络将点云划分为多尺度的体素,可以自适应地学习局部特征。针对点云数据稀疏性这一特征,Meng 等<sup>[24]</sup>利用基于径向基函数(RBF)的变分自动编码器(VAE)网络编码每个体素内的局部几何结构从而提高分割精度。

间接语义分割的方法虽然解决了点云数据的非结构化的问题,但投影和体素化步骤不可避免地会导致信息丢失。虽然使用高分辨率网格可以弥补这一缺陷,但高耗存和高计算成本使网络难以训练,因此实用性更低。

### 2.2 直接语义分割

PointNet<sup>[8]</sup>作为直接从点云中学习特征的开创性网络,解决了点云无序性、置换不变性和旋转不变性这 3 个核心问题,但特征提取时没有考虑局部邻域信息。受 2D-CNN 启发,PointNet++<sup>[9]</sup>采用递归的方式循环使用 PointNet 以提取局部特征,但仍然忽略了每个区域的局部特征。PointSIFT 模块<sup>[25]</sup>的方向编码单元在 8 个方向上对最近点的特征进行卷积,从而能够提取更可靠和稳定的表征点,但计算量的增加使得效率不高,且采样不均衡也一定程度上限制了分割的效果。ShellNet<sup>[26]</sup>使用同心球壳的统计信息来定义有代表性的特征并解决了点云无序性这一问题。RandLANet<sup>[27]</sup>是一种大规模点云处理的轻量级网络,在存储和计算方面都具有优势。Wang 等<sup>[28]</sup>提出了动态图边缘卷积神经网络(DGCNN),其使用边缘卷积(edgeconv)可以很好地提取点云的局部特征,但特征提取时忽视了相邻点之间的方向信息,所以损失了部分几何特征。Landrieu 等<sup>[29]</sup>提出了超点图(SPG)能够详细描述相邻目标之间的关系,解决了点与点之间缺乏联系这一问题。采用深度学习技术对点云进行分割需要大量精细标记的点云用来训练网络,这使得深度学习在语义分割中的应用受到限制。为了减少所需注释标签的数量,3DLEB-Net<sup>[30]</sup>通过将 DGCNN 编码器和基于折叠的编码器组合使用,实现在只有少量标签的情况下对其进行重建并提取特征信息,但该网络受块大小的限制,在小区域上训练时,性能下降。Engelmann 等<sup>[31]</sup>提出了一种扩张点卷积操作来积累扩张的相邻特征,缓解了卷积网络感受野受限的问题。MSSCN<sup>[32]</sup>通过聚合不同密度的点云特征来提高语义

分割性能,但在特征融合过程中仍会丢失部分信息。KVGCN<sup>[33]</sup>通过在特征编码块中嵌入 VLAD 层以有效聚合局部特征和全局上下文信息,提高了语义分割精度,但网络参数较多,训练效率不高。

综上所述,现有网络基本解决了直接从点云中学习特征的困难。然而,如何从点云数据中有效提取局部特征这一问题仍需要思考。所以本文围绕这一问题做进一步探索,提出了一个动态图卷积网络 DGPoint,可以有效地聚合局部特征,提高语义分割的准确率。

### 3 本文方法

受 PointNet++<sup>[9]</sup>和 DGCNN<sup>[28]</sup>的启发,本文将图

卷积神经网络的思想融入到 PointNet++ 网络上,通过构建局部区域图并应用边缘卷积来有效提取局部特征。本文网络还引入了动态图更新的思想,网络的每一层都会根据 KNN 确定新的局部区域,网络架构如图 1 所示。相较于 DGCNN<sup>[28]</sup>,DGPoint 是在 PointNet++ 的基础上引入边缘卷积来提取局部特征,这保留了 PointNet++ 在上采样层过程中使用线性插值方法的优势。此外,对于特征聚合函数  $\square$ ,DGCNN 仅使用了最大池化函数(max pooling)来聚合特征,DGPoint 则使用双通道池化(max pooling 和 avg pooling)操作来补偿信息丢失,这能更好地保留点云的细粒度局部特征。

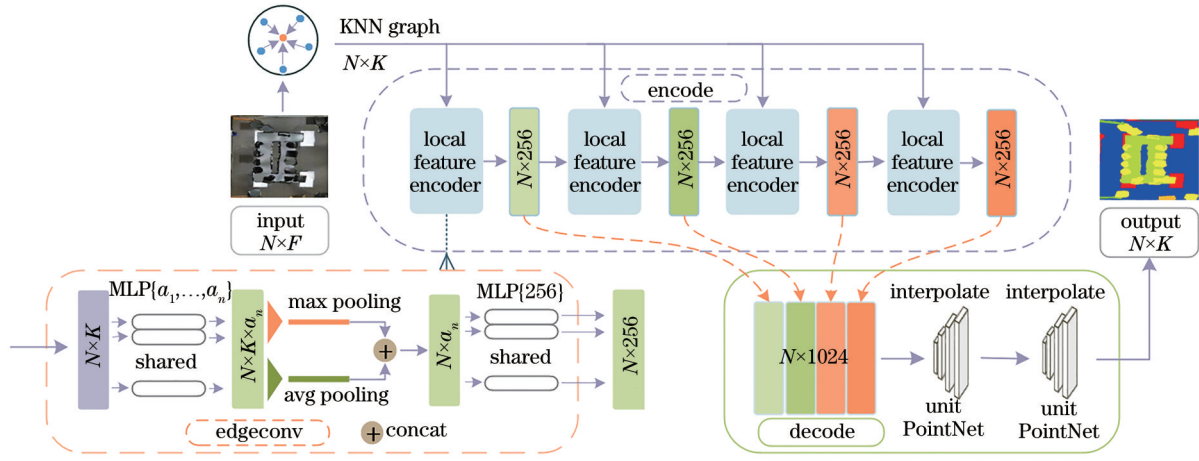


图 1 网络架构:4个局部特征编码块(上部虚线框和左下虚线框内显示单个局部特征编码块的详细结构)被递归组装以编码局部特征,然后将4个输出(右下实线框内)连接到解码块的输入

Fig.1 Network architecture: four local feature encoding blocks (dotted box above, with detail structure of a single local feature encoding block shown in lower left dotted box) are assembled recursively to encode local features, then the four outputs (lower right solid box) are concatenated to be input to decoding blocks

#### 3.1 边缘卷积

假设具有  $n$  个点的  $F$  维点云,表示为  $X = \{x_1, \dots, x_n\} \subseteq \mathbb{R}^F$ , 每个点包含坐标信息  $x_i = (x_i, y_i, z_i)$ , 颜色信息红绿蓝(RGB)以及其他信息。用一个有向图  $\vartheta = (V, \epsilon)$  表示局部点云结构,其中  $V = \{1, \dots, n\}$  和  $\epsilon \subset V \times V$  分别是顶点和边。最简单的情况下,将  $\vartheta$  构造为  $\mathbb{R}^F$  中  $X$  的 KNN 图,边缘特征定义为  $e_{ij} = h_\theta(x_i, x_j)$ , 其中,  $h_\theta: \mathbb{R}^F \times \mathbb{R}^F \rightarrow \mathbb{R}^F$  (将两个  $F$  维的点笛卡儿积的维度映射到  $\mathbb{R}^F$  上)是具有一组可学习参数  $\theta$  的非线性函数。

在边缘特征上,使用一个通道聚合函数操作  $\square$  (比如 sum, max)来定义边缘卷积操作,则第  $i$  个顶点的边缘卷积输出为

$$x'_i = \square_{j:(i,j) \in \epsilon} h_\theta(x_i, x_j), \quad (1)$$

类比图像的卷积,可将  $x_i$  视为中心像素,将  $\{x_j: (i, j) \in \epsilon\}$  视为其领域结构。边函数( $h$ )的选择和聚合操作  $\square$  对边缘卷积的属性具有至关重要的影响,在先前的工作中,研究者们对于边函数( $h$ )的选择主要有以下4种:

1)当  $x_1, \dots, x_n$  表示规则格网上的图像像素,且图  $\vartheta$  是表示每个像素周围固定大小的联通区域时,选择  $\theta_m \cdot x_j$  作为边缘函数, sum 作为聚合操作,则有

$$h_\theta(x_i, x_j) = \theta_m \cdot x_j, \quad (2)$$

$$x'_i = \sum_{j:(i,j) \in \epsilon} \theta_m \cdot x_j, \quad (3)$$

式中:每个  $\theta_m$  与  $x$  的维度相同;  $\cdot$  表示欧几里得内积;

2)PointNet使用的边函数( $h$ ),只编码全局形状信息而忽略局部邻域结构,则有

$$h_\theta(x_i, x_j) = \bar{h}_\theta(x_i); \quad (4)$$

3)PointNet++使用的边函数( $h$ )为

$$h_\theta(x_i, x_j) = \bar{h}_\theta(x_j); \quad (5)$$

$$x'_i = \sum_{j \in V} \{h_\theta(x_j) g[u(x_i, x_j)]\}, \quad (6)$$

式中:  $g$  是高斯核;  $u$  是计算欧几里得空间中的成对距离;

4)仅编码局部信息,而丢失了全局信息,则有

$$h_\theta(x_i, x_j) = \bar{h}_\theta(x_j - x_i). \quad (7)$$

由于以上4种边函数( $h$ )的选择均存在局限性,考

考虑到编码全局信息的同时兼顾局部信息,本文采用非对称度量边函数,由区域中心  $x_i$  坐标来捕获全局形状信息,  $x_j - x_i$  捕获局部邻域信息,则有

$$h_{\theta}(x_i, x_j) = \bar{h}_{\theta}(x_i, x_j - x_i). \quad (8)$$

图 2 展示了边缘卷积操作,将中心点的特征与两点的特征差串联后输入 MLP,这样边缘特征就融合了点的局部信息和全局信息。假设一个中心点  $x_i$  周围有

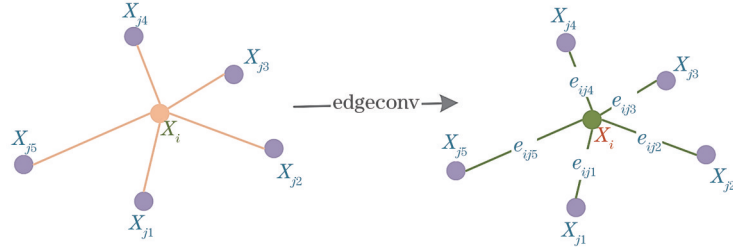


图 2 边缘卷积操作示例

Fig. 2 Example of edge convolution operation

### 3.2 动态图更新

在逐层进行边缘卷积时,每一层都会输入新的点云图结构产生新的特征空间,考虑到不同的特征空间之间的差异,每一层都应在新结构上定义新的局部区域,所以在进行卷积之前,首先计算特征空间中的成对距离矩阵,再根据 KNN 确定新的局部区域,这样在每一层就会有一个不同的图结构  $\vartheta^{(l)} = (V^{(l)}, \epsilon^{(l)})$ ,其中,第  $l$  层的边缘形式为  $(i, j_{i1}), \dots, (i, j_{ik_i})$ 。简而言之,本文的网络架构学习如何构建每一层中使用的图  $\vartheta$ ,而不是将其视为固定常数。

$m$  个近邻点(图 2 中  $m=5$ ),  $\max$  和  $\text{avg}$  作为聚合操作,则边缘特征计算和边缘卷积输出为

$$e_{ijm} = \text{ReLU}[\theta_m \cdot (x_j - x_i) + \phi_m \cdot x_i], \quad (9)$$

$$x'_{im} = \max_{j:(i,j) \in \epsilon} e_{ijm} + \text{avg}_{j:(i,j) \in \epsilon} e_{ijm}, \quad (10)$$

式中,  $\Theta = (\theta_1, \dots, \theta_m, \phi_1, \dots, \phi_m)$ ,  $\text{ReLU}$  为神经网络中常用的激活函数。

## 4 实验

为了评估 DGPoint 的网络性能,本文在目前较为主流的数据集 S3DIS<sup>[34]</sup> 上进行了实验, S3DIS 数据集的 6 个区域如图 3 所示。S3DIS 数据集是斯坦福大学研究团队提出的多模态、大规模的室内空间数据集,具有实例级语义和几何注释。其包含 6 个不同的大型室内区域 272 个 3D 房间场景,占地面积超过 6000 m<sup>2</sup>,分为 13 个语义类别(墙壁、桌子、椅子等),每个点包含 3D 坐标、颜色、归一化坐标和其他信息。

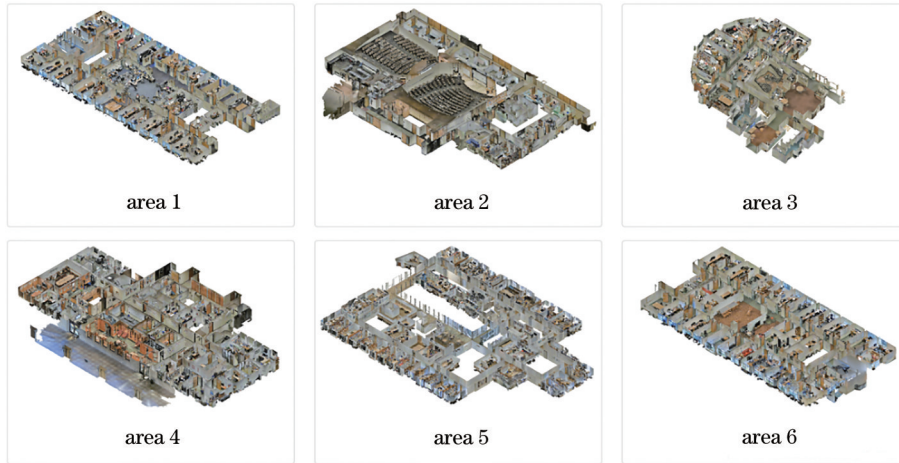


图 3 S3DIS<sup>[34]</sup>数据集

Fig. 3 S3DIS data set

### 4.1 训练

S3DIS 数据集分为 6 个不同的区域,为了方便对提出的网络进行评估,本文使用 6 折交叉验证的方法进行模型评估,每次选择 5 个区域作为训练集,剩下的 1 个区域作为测试集,总共建立了 6 个模型来覆盖整个数据集。还使用了 3 个评估指标:总体准确度(oAcc)、交并比(IoU)和平均交并比(mIoU)来评估模型的分割性能。

使用 Tensorflow 在 Adam 优化器 GTX 1080Ti GPU 上训练模型,分别训练了 PointNet++ 和 DGPoint 两个模型,训练时超参数设置为:初始学习率为 0.001、初始衰减率为 0.5、batch\_size 为 24、优化方法为 Adam、Momentum 为 0.9、max\_epoch 为 201。实验时根据房间号区分 S3DIS 数据集,将每个房间采样为边长为 1 m 的立方体区域,然后以 0.5 m 的步长遍历整个房间,每个

点用一个 9 维向量表示, 向量分量为:  $X, Y, Z, R, G, B$ , 以及所属房间的归一化坐标  $X_n, Y_n, Z_n$ 。图 4 显示了训

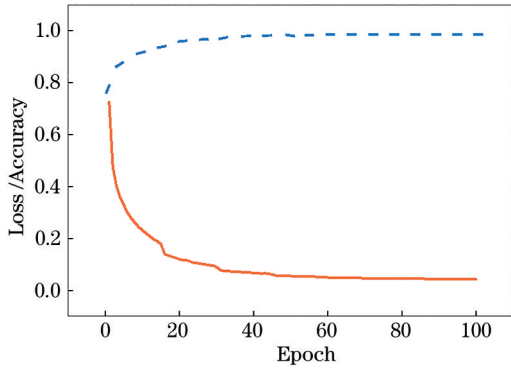


图 4 精度(虚线)和损失函数(实线)曲线

Fig. 4 Accuracy (dotted line) and loss function (solid line) curves

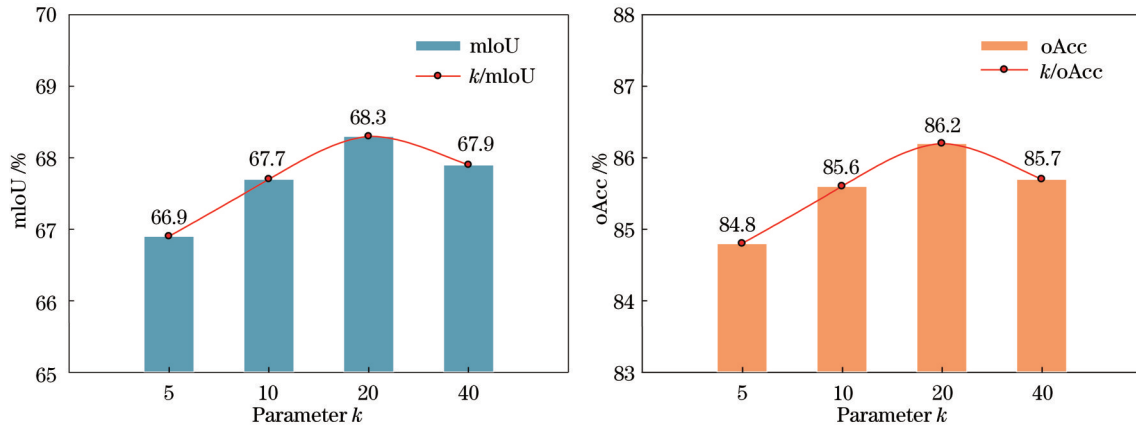


图 5 S3DIS 数据集对不同 k 的分割精度

Fig. 5 Segmentation accuracy of S3DIS data set for different k

### 4.3 语义分割结果比较

本文通过 6 折交叉验证在 S3DIS 数据集上得到所提出的 DGPoint 模型的分割结果, 并与最先进的方法进行对比。模型的 mIoU 和 oAcc 对比如表 1 所示, 每个语义类的 IoU 对比如表 2 所示。DGPoint 的 mIoU、oAcc 分别为 68.3%、86.2%, 优于 PointNet<sup>[8]</sup>、PointNet++<sup>[9]</sup>、DGCNN<sup>[28]</sup> 和 ShellNet<sup>[26]</sup>。与

表 1 整个 S3DIS 数据集上的语义分割准确度比较

Table 1 Comparison of semantic segmentation accuracy on entire S3DIS data set

Method	mIoU / %	oAcc / %
PointNet <sup>[8]</sup>	47.6	78.6
PointNet++ <sup>[9]</sup>	50.8	82.2
DGCNN <sup>[28]</sup>	56.1	84.1
ShellNet <sup>[26]</sup>	66.8	—
KVGCN <sup>[33]</sup>	60.9	<b>87.4</b>
<b>DGPoint</b>	<b>68.3</b>	86.2

PointNet++<sup>[9]</sup>相比, DGPoint 的 mIoU 提高 17.5 个百分点, oAcc 提高 4.0 个百分点, 说明边缘卷积的引入使网络可以有效聚合局部特征, 提高分割精度。但整体

训练期间的准确率和损失曲线, 可以看出随着网络训练过程中迭代次数的增加, 损失函数值不断减小, 最终稳定在 0.05 以下, 此时模型的精度接近 1。

### 4.2 参数 k 评估

最近邻 K 值控制邻居搜索的范围和用于训练的本地信息量, K 值太小会导致捕获的局部信息不足影响分割精度。但这并不意味着 K 值越大越好, 过大的 K 值不但会使计算成本过高, 而且欧氏距离无法近似测地距离, 从而会破坏每个 patch 的几何形状信息。本文对比了 4 个不同数量的最近邻 K 值, 如图 5 所示, 虽然只进行了 4 个对比实验, 但可以观察到随着 K 值的增加, mIoU 和 oAcc 先增加然后减少。因此可以得出结论, 参数 k 的最佳设置是 k=20。鉴于此, 在下面的评估中, 参数 k 固定为 20。

精度不如 KVGCN<sup>[33]</sup>, KVGCN 可以通过在特征编码块中嵌入 VLAD 层来有效地聚合局部特征和全局上下文信息。从每个语义类别的 IoU 来看, DGPoint 模型改进了某些类别的 IoU (表 2)。这些定量结果表明, 边缘卷积可以有效地聚合点云的局部特征, 提高点云分割的准确性。

为了呈现直观的观察结果, 本文选择了 3 个样本来可视化分割结果, 如图 6 所示。第一列表示真实分割结果; 第二列表示使用 PointNet++ 的分割结果; 第三列表示使用 DGPoint 的分割结果。与 PointNet++ 相比 (由图 6 中实线框可看出), DGPoint 可以更好地提取局部信息, 优化语义分割的边界, 取得与地面实况更接近的分割结果。点云数据集由于数据量庞大, 场景中大量噪声, 因此两个语义类别之间往往存在大量交集点云, DGPoint 可以更清晰地分离这些交集点云, 获得更清晰的语义分割边界。

## 5 结 论

设计了一种新的图卷积网络架构 DGPoint, 通过建立查询点的动态图结构, 使用边缘卷积以获得强大

表 2 S3DIS 数据集中部分类别的 IoU

Table 2 IoU of some semantic class in S3DIS data set

unit: %

Method	Celling	Floor	Wall	Door	Table	Chair	Bookcase	Clutter
PointNet <sup>[8]</sup>	88.8	97.3	69.8	10.8	58.9	52.6	40.3	33.2
PointNet++ <sup>[9]</sup>	93.1	97.6	77.2	58.1	62.3	65.7	45.2	50.4
SegCloud <sup>[20]</sup>	90.1	96.1	69.9	23.1	70.4	<b>78.6</b>	<b>58.4</b>	41.6
SPG <sup>[29]</sup>	89.9	95.1	72.0	60.0	69.2	73.5	3.2	52.9
KVGCN <sup>[33]</sup>	94.5	94.1	<b>79.5</b>	<b>63.2</b>	64.3	67.5	23.6	53.2
<b>DGPoint</b>	<b>97.8</b>	<b>98.6</b>	78.7	62.6	<b>76.9</b>	73.3	30.3	<b>55.9</b>

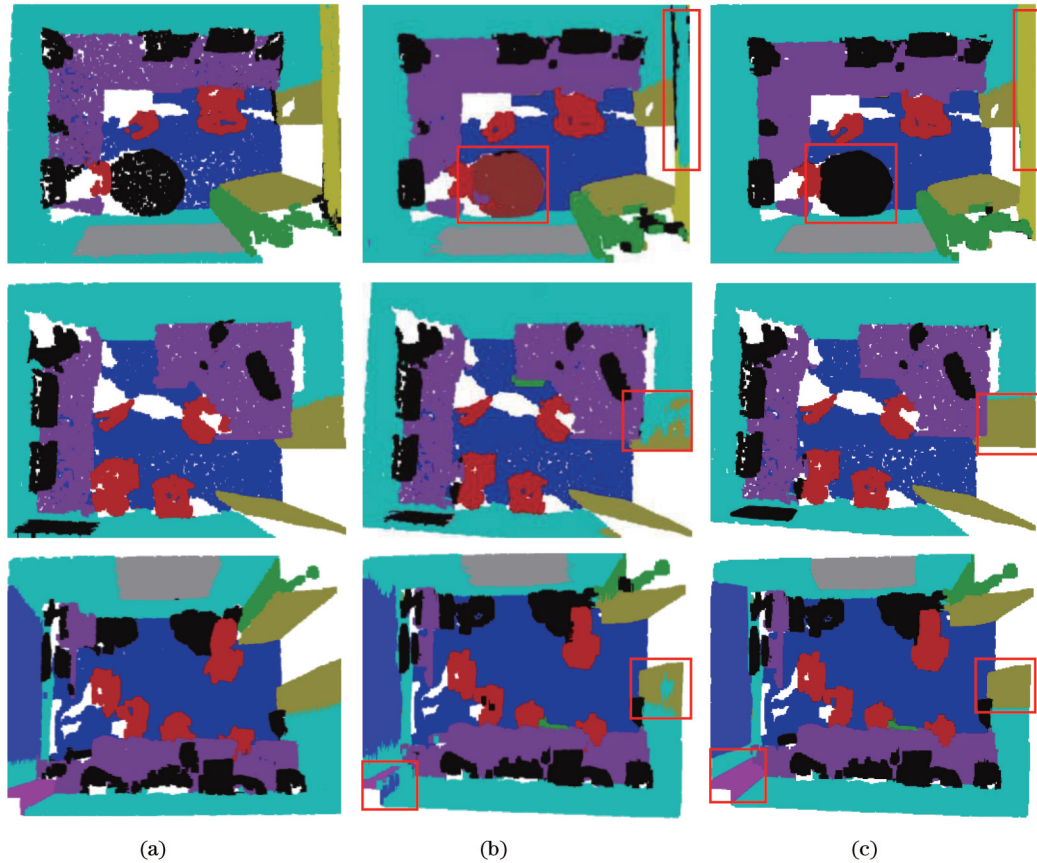


图 6 S3DIS 数据集的样本语义分割结果。(a) 真实分割结果; (b) PointNet++ 的分割结果; (c) DGPoint 的分割结果  
Fig. 6 Sample semantic segmentation results of S3DIS data set. (a) Ground truth; (b) segmentation result of PointNet++; (c) segmentation result of DGPoint

的局部特征,在特征表示方面具有很大优势。在 S3DIS 数据集上评估了所提出的网络,与 PointNet++<sup>[9]</sup>相比,DGPoint 可以从图结构中提取复杂的特征,体现出其分割性能的优越性。虽然提出网络整体上提高了分割效果,但仍然需要进一步改进,例如,提出网络使用 max pooling 来提取全局特征,max pooling 只能保留最明显的全局特征信息,而丢失了一些不太明显但有用的细节特征,所以如何解决全局特征提取不当的问题还需要进一步研究。此外,所提出的网络在提高模型表达能力的同时增加了计算负荷,违背了网络轻量化的追求,因此,如何降低计算负荷、提高存储效率也将成为下一步研究的重点。

## 参 考 文 献

- [1] Zhang J Y, Zhao X L, Chen Z, et al. A review of deep learning-based semantic segmentation for point cloud[J]. IEEE Access, 2019, 7: 179118-179133.
- [2] Zhou H, Fang Z J, Gao Y B, et al. Feature fusion network based on attention mechanism for 3D semantic segmentation of point clouds[J]. Pattern Recognition Letters, 2020, 133: 327-333.
- [3] Shi B G, Bai S, Zhou Z C, et al. DeepPano: deep panoramic representation for 3-D shape recognition[J]. IEEE Signal Processing Letters, 2015, 22(12): 2339-2343.
- [4] 赵亮, 胡杰, 刘汉, 等. 基于语义分割的深度学习激光点云三维目标检测[J]. 中国激光, 2021, 48(17):

1710004.  
Zhao L, Hu J, Liu H, et al. Deep learning based on semantic segmentation for three-dimensional object detection from point clouds[J]. Chinese Journal of Lasers, 2021, 48(17): 1710004.
- [5] 柴玉晶, 马杰, 刘红. 用于点云语义分割的深度图注意力卷积网络[J]. 激光与光电子学进展, 2021, 58(12): 1210016.  
Chai Y J, Ma J, Liu H. Deep graph attention convolution network for point cloud semantic segmentation[J]. Laser & Optoelectronics Progress, 2021, 58(12): 1210016.
- [6] 鲍海龙, 万敏, 刘忠祥, 等. 基于区域自我注意力的实时语义分割网络[J]. 激光与光电子学进展, 2021, 58(8): 0810018.  
Bao H L, Wan M, Liu Z X, et al. Real-time semantic segmentation network based on regional self-attention[J]. Laser & Optoelectronics Progress, 2021, 58(8): 0810018.
- [7] 王宇彤, 李琦. 进化算法优化区域生长的太赫兹全息再现图像分割[J]. 中国激光, 2020, 47(8): 0814003.  
Wang Y T, Li Q. Terahertz holographic reconstructed image segmentation based on optimized region growth by evolutionary algorithm[J]. Chinese Journal of Lasers, 2020, 47(8): 0814003.
- [8] 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, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 77-85.
- [9] 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 Advances in Neural Information Processing Systems, December 4-9, 2017, Long Beach, CA, USA. [S.l.: s.n.], 2017: 5099-5108.
- [10] Liu W B, Wang Z D, Liu X H, et al. A survey of deep neural network architectures and their applications[J]. Neurocomputing, 2017, 234: 11-26.
- [11] 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, December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 945-953.
- [12] Feng Y F, Zhang Z Z, Zhao X B, et al. GVCNN: group-view convolutional neural networks for 3D shape recognition[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 264-272.
- [13] Ibrahim M, Akhtar N, Ullah K, et al. Exploiting structured CNNs for semantic segmentation of unstructured point clouds from LiDAR sensor[J]. Remote Sensing, 2021, 13(18): 3621.
- [14] Boulch A, Guerry J, le Saux B, et al. SnapNet: 3D point cloud semantic labeling with 2D deep segmentation networks[J]. Computers & Graphics, 2018, 71: 189-198.
- [15] Guerry J, Boulch A, le Saux B, et al. SnapNet-R: consistent 3D multi-view semantic labeling for robotics [C]//2017 IEEE International Conference on Computer Vision Workshops, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 669-678.
- [16] 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, May 21-25, 2018, Brisbane, QLD, Australia. New York: IEEE Press, 2018: 1887-1893.
- [17] 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), May 20-24, 2019, Montreal, QC, Canada. New York: IEEE Press, 2019: 4376-4382.
- [18] Kang D W, Wong A, Lee B, et al. Real-time semantic segmentation of 3D point cloud for autonomous driving [J]. Electronics, 2021, 10(16): 1960.
- [19] 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), September 28-October 2, 2015, Hamburg, Germany. New York: IEEE Press, 2015: 922-928.
- [20] Tchapmi L, Choy C, Armeni I, et al. SEGCloud: semantic segmentation of 3D point clouds[C]//2017 International Conference on 3D Vision (3DV), October 10-12, 2017, Qingdao, China. New York: IEEE Press, 2017: 537-547.
- [21] 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, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6620-6629.
- [22] Xu Y, Hoegner L, Tattas S, et al. Voxel- and graph-based point cloud segmentation of 3D scenes using perceptual grouping laws[J]. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2017, IV-1/W1: 43-50.
- [23] Wang L, Huang Y C, Shan J, et al. MSNet: multi-scale convolutional network for point cloud classification[J]. Remote Sensing, 2018, 10(4): 612.
- [24] Meng H Y, Gao L, Lai Y K, et al. VV-net: voxel VAE net with group convolutions for point cloud segmentation [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 8499-8507.
- [25] Jiang M Y, Wu Y R, Zhao T Q, et al. PointSIFT: a SIFT-like network module for 3D point cloud semantic segmentation[EB/OL]. (2018-07-02)[2021-02-06]. <https://arxiv.org/abs/1807.00652>.
- [26] Zhang Z Y, Hua B S, Yeung S K. ShellNet: efficient point cloud convolutional neural networks using concentric shells statistics[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2,

- 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 1607-1616.
- [27] 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), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 11105-11114.
- [28] 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.
- [29] Landrieu L, Simonovsky M. Large-scale point cloud semantic segmentation with superpoint graphs[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 4558-4567.
- [30] Cao Y W, Scaioni M. 3DLEB-net: label-efficient deep learning-based semantic segmentation of building point clouds at LoD3 level[J]. Applied Sciences, 2021, 11(19): 8996.
- [31] Engelmann F, Kontogianni T, Leibe B. Dilated point convolutions: on the receptive field size of point convolutions on 3D point clouds[C]//2020 IEEE International Conference on Robotics and Automation, May 31-August 31, 2020, Paris, France. New York: IEEE Press, 2020: 9463-9469.
- [32] Du J, Jiang Z N, Huang S F, et al. Point cloud semantic segmentation network based on multi-scale feature fusion [J]. Sensors, 2021, 21(5): 1625.
- [33] Luo N, Yu H Q, Huo Z F, et al. KVGCN: a KNN searching and VLAD combined graph convolutional network for point cloud segmentation[J]. Remote Sensing, 2021, 13(5): 1003.
- [34] Armeni I, Sener O, Zamir A R, et al. 3D semantic parsing of large-scale indoor spaces[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1534-1543.