

用于点云语义分割的深度图注意力卷积网络

柴玉晶, 马杰*, 刘红

河北工业大学电子信息工程学院, 天津 300401

摘要 相比于卷积神经网络,图卷积网络更适合处理不规则的点云数据,但其存在网络层数受限以及固定的标准化聚集方式影响点云语义分割结果的问题。针对这些问题,提出一种用于点云语义分割的深度图注意力卷积网络。该网络通过残差连接来加深图卷积网络层数,可以有效解决网络过深所导致的梯度消失和网络退化的问题,采用注意力机制可以使网络有选择地关注最相关的邻域点并为其分配不同的注意力权重。同时在每层图卷积之后重新构建图形,从而更好地表征图结构。实验结果表明,该网络在斯坦福大规模的三维室内空间数据集上的平均交并比达到 64.5%。

关键词 图像处理; 点云语义分割; 图卷积; 注意力机制; 残差连接; 动态图卷积

中图分类号 TP391

文献标志码 A

doi: 10.3788/LOP202158.1210016

Deep Graph Attention Convolution Network for Point Cloud Semantic Segmentation

Chai Yujing, Ma Jie*, Liu Hong

School of Electronic Information Engineering, Hebei University of Technology, Tianjin 300401, China

Abstract Compared with convolutional neural networks, graph convolution network is more suitable for processing irregular point cloud data. However, it has the problem that the number of network layers is limited and the fixed and standardized aggregation method affects the result of point cloud semantic segmentation. To solve these problems, a depth graph attention convolutional network for point cloud semantic segmentation is proposed herein. The network uses residual connections to deepen the number of layers of the graph convolutional network, which can effectively solve the problems of gradient disappearance and network degradation caused by the network being too deep. The attention mechanism is used to make the network selectively focus on the most relevant neighborhood points, and it assigns different attention weights to it. Simultaneously, the graph is reconstructed after each layer of graph convolution to better characterize the graph structure. Experimental results show that the average intersection ratio of the network on Stanford large-scale three-dimensional indoor spatial dataset reaches 64.5%.

Key words image processing; point cloud semantic segmentation; graph convolution; attention mechanism; residual connection; dynamic graph convolution

OCIS codes 100.6890; 100.4996; 150.1135

1 引言

随着三维扫描技术的发展,人们的研究重心逐渐从二维图像数据转移到三维数据,而点云是三维

数据中的一种重要数据类型,广泛应用于导航定位和文化遗产保护等领域。最初的点云处理方法是手动来提取特征的,方法包括点特征直方图(Point Feature Histogram, PFH)、快速点特征直

收稿日期: 2020-09-07; 修回日期: 2020-10-02; 录用日期: 2020-10-16

基金项目: 河北省自然科学基金(F2020202045)、河北省研究生创新资助项目(CXZZBS2020026)

* E-mail: jma@hebut.edu.cn

方图(Fast Point Feature Histogram, FPFH)和视点特征直方图(Viewpoint Feature Histogram, VFH)等^[1-5]。然而,这些传统方法均依赖于专业领域的知识,当处理种类和数量越来越多的三维数据时效果不佳。随着深度学习的兴起,越来越多的研究者利用深度神经网络来处理点云^[6-7]。Charles 等^[8]提出了以原始点云为输入的深度学习网络 PointNet,该网络可以直接对点云数据进行处理,但是 PointNet 是对点云的全局特征进行分类分割,忽略了局部特征。为此, Qi 等^[9]进一步提出了 PointNet++,用来解决对点云的局部划分以及提取局部特征的问题。

点云是一种不规则的数据类型。相比于用于处理规则数据的卷积神经网络(Convolutional Neural Network, CNN),用于处理非欧几里得数据的图卷积神经网络(Graph Convolutional Network, GCN)^[10-11]更适合处理点云。动态图卷积网络(Dynamic Graph Convolutional Network, DGCNN)^[12]将图卷积网络应用在点云处理的过程中,在保证置换不变性的同时捕获局部几何信息。值得一提的是, DGCNN 使用了动态的图卷积,在每层图卷积之后重构图结构。相比于固定结构的图卷积,动态的图卷积能够有效解决过度平滑的问题,增大了感受野。DGCNN 虽然改善了聚集方式,但仍是一种固定的标准化聚集方式,而注意力机制^[13]可以从众多信息中找到对当前任务更重要的信息。图注意力网络(Graph Attention Network, GAT)通过注意力机制为不同的邻域点分配不同的权重,从而能够更好地聚集邻域信息。图注意力卷积(Graph Attention Convolution, GAC)^[14]结合了 DGCNN 与 GAT,可以将图注意力网络应用在点云分割方面。由于网络存在梯度消失的问题,深层的 GCN 在聚集过程中容易过度平滑,而上述 GCN 算法的图卷积层数不超过 3 层和 4 层。借鉴 CNN 加深网络层数的方法, Li 等^[15]通过残差连接^[16](或密集连接^[17])和膨胀卷积来解决梯度消失的问题,成功训练了用于点云分割且具有 28 层的 DeepGCNs,然而 DeepGCNs 仍然采用固定的标准化聚集方式。

为此,本文设计一种用于点云语义分割的深度图注意力卷积网络(Deep Graph Attention Convolution Network, DeepGAC)。该网络将 GAC 与 DeepGCNs 结合,利用残差连接来解决梯度消失的问题,采用注意力机制为不同的邻域点分配不同的权重,从而更好地聚集邻域信息,并且通过动态图

卷积来生成更好的图结构。实验结果表明,该网络在窗户、沙发和桌子这类形状相似以及难以区分的目标物上效果良好。

2 深度图注意力卷积神经网络

2.1 图注意力卷积

在点云上构造图像 $G(V, E)$, 其中 $V \in \{1, 2, \dots, N\}$ 表示图的顶点, $E \in |V| \times |V|$ 表示图的边缘。设中心点为 i , 邻域点为 j , 则 $j \in N(i)$, 其中 $N(i)$ 表示中心点 i 的所有邻域点。一组点云为 $\{P_i \in \mathbf{R}^3 | i=1, 2, \dots, n\}$, 其中 n 为点的数目。每个点的特征为 $\{h_i \in \mathbf{R}^D | i=1, 2, \dots, n\}$, 其中 D 为特征的维度。

图注意力卷积步骤如下。

1) 计算图注意力系数 e_{ij} , 表达式为

$$e_{ij} = M_{MLP}(P_j - P_i \parallel h_j - h_i), \quad (1)$$

式中: \parallel 为拼接操作; M_{MLP} 为多层感知器。(1)式同时考虑了空间位置信息和特征属性信息, 所求得的图注意力系数 e_{ij} 表示邻域点 j 对于中心点 i 的重要程度。 e_{ij} 值越大, 邻域点 j 对中心点 i 越重要。

2) 使用 Softmax 函数归一化图注意力系数, 使得所有的权重相加值为 1, 以便简化运算, 表达式为

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{j \in N(i)} \exp(e_{ij})}. \quad (2)$$

3) 采用一层 MLP 将特征维度从 D 映射到 D' , 然后加权求和聚集邻域点的特征到下一层的中心点特征, 表达式为

$$h'_i = \sum_{j \in N(i)} \alpha_{ij} \cdot M_{MLP}(h_j) + b_i, \quad (3)$$

式中: $h'_i \in \mathbf{R}^{D'}$ 为下一层的中心点特征; $b_i \in \mathbf{R}^{D'}$ 为偏置。

图 1 为图注意力卷积的过程。

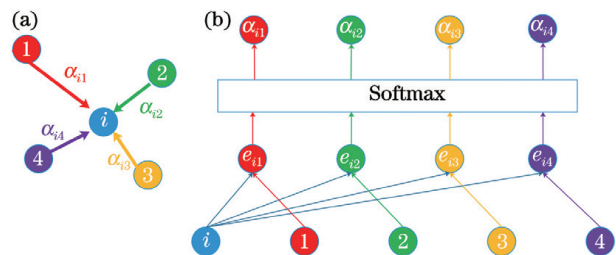


图 1 GAC 的过程。(a)邻域点向中心点聚集;

(b)图注意力系数的计算过程

Fig. 1 Process of GAC. (a) Neighborhood points cluster toward center point; (b) calculation process of graph attention coefficient

2.2 残差连接在图卷积网络的运用

2.2.1 图卷积网络

不同于图像,点云是一种不规则的数据类型。相比于用于处理规则数据的卷积神经网络,用于处理非欧几里得数据的图卷积神经网络更适合处理点云。图卷积主要有聚集和更新两个步骤^[18-19],首先聚集邻域点的特征到中心点,然后通过多层感知器和非线性激活函数来更新顶点特征。图像 \mathbf{G} 在第 $L+1$ 层的输出为

$$\mathbf{G}_{L+1} = F(\mathbf{G}_L, \mathbf{W}) =$$

$$F_{\text{update}} [F_{\text{agg}}(\mathbf{G}_L, \mathbf{W}_{\text{agg},L}), \mathbf{W}_{\text{update},L}], \quad (4)$$

式中: $\mathbf{G}_L = (\mathbf{V}_L, \mathbf{E}_L)$; F 为一层图卷积函数; F_{agg} 和 F_{update} 分别为聚集和更新函数; $\mathbf{W}_{\text{agg},L}$ 和 $\mathbf{W}_{\text{update},L}$ 为聚集和更新函数的权重。聚集函数有最大池化和平均等操作,实验采用图注意力聚集方式将 2.1 节求得的图注意力系数作为聚集权重,则 $[\alpha_{i1} \alpha_{i2} \dots \alpha_{ij}] = \mathbf{W}_{\text{agg},L}$, 所以点 i 在第 $L+1$ 层的输出为

$$h_{i_{L+1}} = \varphi \left[\sum_{j \in N(i)} \alpha_{ij} \cdot M_{\text{MLP}}(h_{j_L}) + b_{i_L} \right], \quad (5)$$

式中: φ 为非线性激活函数。由 φ 和一层 MLP 构成顶点特征更新函数,由邻域点特征与图注意力系数加权求和构成聚集函数。

2.2.2 残差连接

在神经网络中,后一层的输入是对前一层输出的加权求和处理后的结果,即前一层的特征在后一层被抽象化。随着神经网络的加深,网络所学的特征逐渐从颜色和边缘等简单的特征抽象化为纹理和物体部件等复杂的特征。在深度学习中,越复杂的特征越能表征目标特征,通过加深网络层数能够提高特征的抽象化程度,使网络取得更好的效果。然而,网络过深会导致梯度消失和网络退化等问题,对此采用残差连接来搭建整体网络结构,这可以有效解决这一问题,从而实现了网络的加深。残差网络学习的是输入与输出之差,并使用跳跃连接将前一层的输出直接传递到下一层的输出,从而实现恒等映射。残差连接的基本公式为

$$H(x) = C(x) + x, \quad (6)$$

式中: x 和 $H(x)$ 分别为输入与输出; $C(x)$ 为二者之差。在前向传播的过程中,当输入 x 代表的特征已经达到最优时,任何对于特征 x 的改变都会增大损失值,那么 $C(x)$ 会趋向于 0,即 $H(x) = x$,从而实现恒等映射,达到加深网络不会使结果变差的效果。在反向传播的过程中,因为残差模块学习的是输入与输出之差,这可以显著减小网络中的参数值,

相对增大了反向传播过程中的损失值,从而使网络中的参数对损失值响应更敏感。

借鉴残差网络在卷积神经网络上的应用,将残差连接的思想应用在图卷积中,则图像 \mathbf{G} 在第 $L+1$ 层的输出变为

$$\mathbf{G}_{L+1} = O(\mathbf{G}_L, \mathbf{W}_L) =$$

$$F(\mathbf{G}_L, \mathbf{W}_L) + \mathbf{G}_L = \mathbf{G}_{L+1,\text{res}} + \mathbf{G}_L, \quad (7)$$

式中: O 为一层映射函数; $\mathbf{G}_{L+1,\text{res}}$ 为残差模块。由 (7) 式可知,第 $L+1$ 层的输出等于第 L 层的输出加上残差模块的输出。点 i 在第 $L+1$ 层的输出变为

$$h_{i_{L+1},\text{res}} = \varphi \left[\sum_{j \in N(i)} \alpha_{ij} \cdot M_{\text{MLP}}(h_{j_L}) + b_{i_L} \right], \quad (8)$$

$$h_{i_{L+1}} = h_{i_{L+1},\text{res}} + h_{i_L}. \quad (9)$$

2.3 动态图卷积

传统的图卷积只在每次迭代的过程中更新顶点特征,不改变图结构。实验采用动态图卷积在每层图卷积之后重新构建图形,从而更好地表征图结构,这可以有效解决过拟合的问题。为了更新图结构,在每层图卷积之后采用 k 最近邻 (k -NN) 算法来找到中心点在特征空间上距离最近的 k 个点,并将其作为邻域点,从而获得新的边缘以更新图结构。因此第 L 层的图结构为 $\mathbf{G}_L(\mathbf{V}_L, \mathbf{E}_L)$, 边缘部分为 $(i, j_{i_1}), \dots, (i, j_{i_k})$, 其中 j_{i_k} 代表中心点 i 在特征空间上第 k 近邻的邻域点。

图 2 为动态图卷积的过程。

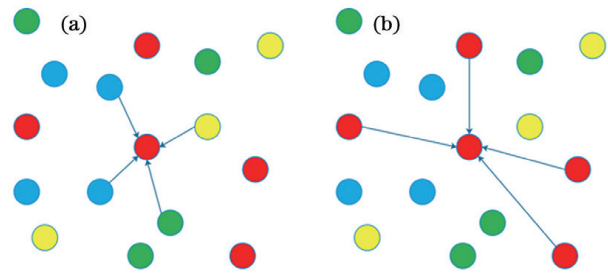


图 2 动态图卷积的过程。(a)中心点在三维空间中寻找邻域点;(b)中心点在特征空间中寻找邻域点

Fig. 2 Process of dynamic graph convolution.

(a) Center point looks for neighborhood point in three-dimensional space; (b) center point looks for neighborhood point in feature space

2.4 深度图注意力卷积网络结构

DeepGAC 由图卷积、特征融合和预测输出三个模块构成。图卷积模块的整体网络采用残差连接结构,将图注意力卷积所得的残差模块的输出与上一层的输出相加后得到下一层的输出。特征融合模块将各层图的特征进行连接得到局部特征,局部特征经过 1×1 的卷积后,固定通道数为 1024,然后通过

全局最大池化得到全局特征。将局部特征与全局特征连接在一起可以同时考虑全局特征和局部特征,从而获得更好的特征信息。预测输出模块通过三层

1×1 的卷积进行降维,将通道数降为所要分割数据集的语义类数(此处为 13),最终输出预测值。整体网络结构如图 3 所示。

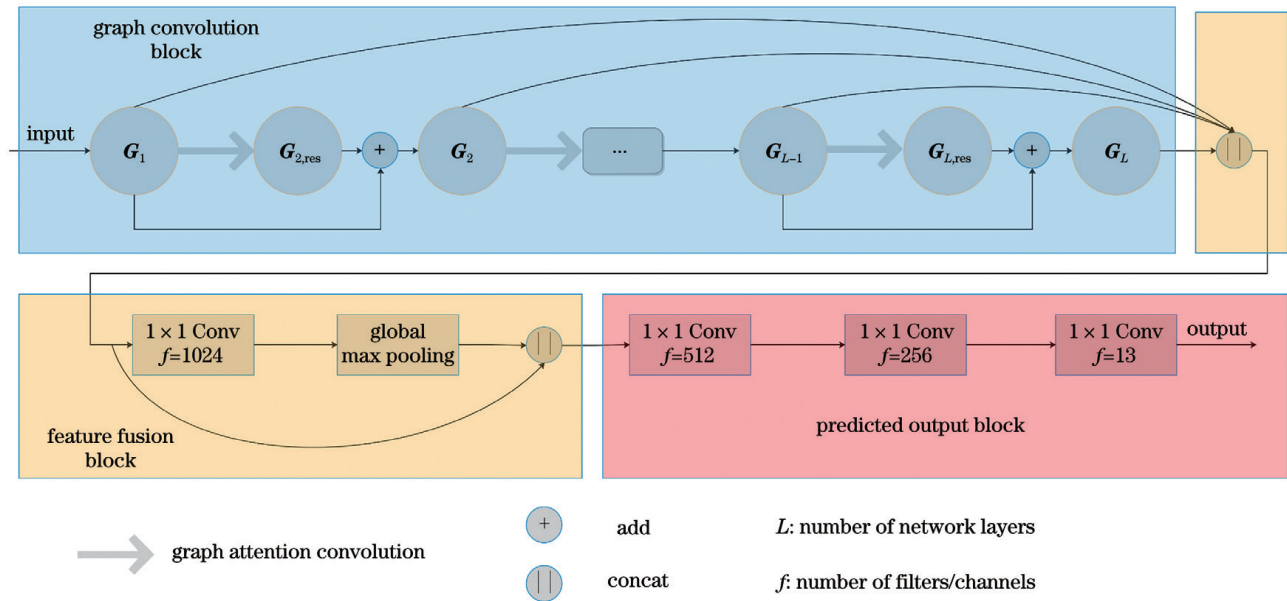


图 3 DeepGAC 的结构

Fig. 3 Structure of DeepGAC

3 实验与分析

为了评估所设计的 DeepGAC 的性能,在斯坦福大规模的三维室内空间数据集(Stanford Large-Scale 3D Indoor Spaces dataset, S3DIS)^[20]上进行训练和测试,将实验结果与当前相关算法进行对比。实验使用 GTX 1080Ti 显卡,Ubuntu16.04 LTS 操作系统,PyThon(3.6 版),TensorFlow(1.12.0 版),Vtk(9.0.1 版)和 Jupyter Notebook(6.0.3 版),其中 Vtk 和 Jupyter Notebook 用于可视化点云语义分割结果。实验采用深度为 28 层残差连接的图注意力卷积网络(ResGAC-28)。通过在各类语义标签上的交并比(intersection over union, IoU)、平均交并比(mean intersection over union, mIoU)和整体准确率(overall accuracy, OA)来评估测试结果,其中 OA 是模型在测试集上预测正确的数量与测试集总体数量的比值,IoU 是某一类别的预测值和真实值两个集合的交集与并集之比,mIoU 是对各类别的 IoU 值取平均值的结果。

3.1 k-折交叉验证法

S3DIS 包括 6 个室内区域的 271 个房间,有 13 类语义标签。首先对数据集进行预处理,将房间采样为 $1 \text{ m} \times 1 \text{ m}$ 大小的块,每个块中随机采样 4096 个点并作为一组点云。将数据集随机分成

6 份,采用 k -折交叉验证法进行 6 次训练和测试,每次都将其其中一份作为测试集,其余 5 份作为训练集,可以充分利用数据集评估模型的泛化能力。对当下主流的点云语义分割网络的测试结果与所提的深度图注意力卷积网络进行对比,结果如表 1 所示,其中 A-CNN 为环状卷积神经网络。从表 1 可以看到,ResGAC-28 的 mIoU 值达到了 64.5%,比 ResGCN-28 提高 4.5 个百分点,证明了图注意力聚集方式的有效性;得益于图注意力卷积和动态图卷积,ResGAC-28 充分利用了特征信息,在沙发和窗户等物体上均取得了较好的分割结果;同时对于柱体和杂物等物体,由于其特征信息不够显著,利用价值较低,则 ResGAC-28 的分割结果略差,通过相对位置的信息来增强特征属性信息或者充分利用空间几何信息也许能够提升此类物体的分割结果。

3.2 区域 5 测试

S3DIS 由三栋不同的建筑扫描得到,将其分为 6 个区域,其中区域 5 的建筑物与其他区域不同,没有重叠,且与其他区域所包含的物体存在一些差异。因此将区域 5 用于测试,其他区域用于训练,可以评估框架的可推广性。图 4 为 S3DIS 中区域 5 的语义分割可视化结果,对 ResGCN-28、GAC 和 ResGAC-28 的语义分割结果在选取的 4 个场景中进行对比,并与原始场景和真值图的语义分割结果进行对照,

表 1 6-折交叉验证法在 S3DIS 上的语义分割结果

Table 1 Semantic segmentation results of 6-fold cross validation method on S3DIS

unit: %

Type	PointNet	PointNet++	DGCNN	ResGCN-28	A-CNN	ResGAC-28 (ours)
OA	78.5	—	84.1	85.9	87.3	88.2
mIoU	47.6	53.2	56.1	60.0	62.9	64.5
Ceiling	88.0	90.2	—	93.1	92.4	92.6
Floor	88.7	91.7	—	95.3	96.4	96.5
Wall	69.3	73.1	—	78.2	79.2	81.1
Beam	42.4	42.7	—	33.9	59.5	36.8
Column	23.1	21.2	—	37.4	34.2	37.5
Window	47.5	49.7	—	56.1	56.3	59.3
Door	51.6	42.3	—	68.2	65.0	68.1
Chair	42.0	59.0	—	61.0	78.0	62.5
Table	54.1	62.7	—	64.9	66.5	69.1
Bookcase	38.2	45.8	—	51.5	56.9	54.7
Sofa	9.6	19.6	—	34.6	28.5	60.9
Board	29.4	48.2	—	51.1	48.0	66.7
Clutter	35.2	45.6	—	54.4	56.8	52.7

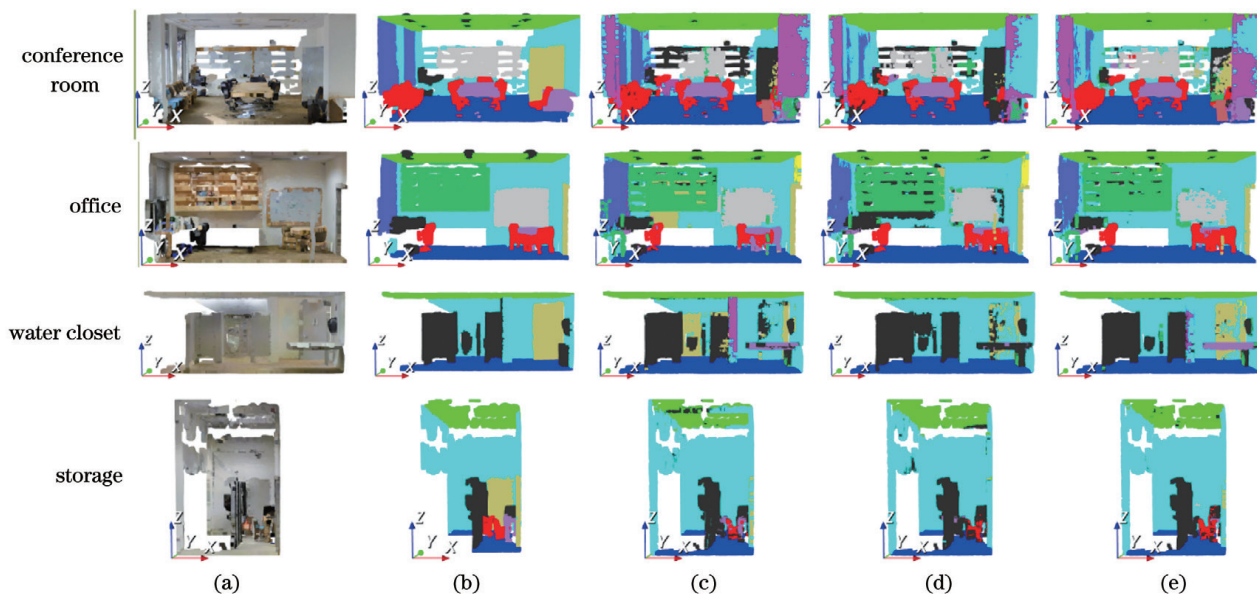


图 4 S3DIS 在不同方法下的语义分割定性结果。(a)原始图像;(b)真值图;(c) ResGCN-28;(d) GAC;(e) ResGAC-28
Fig. 4 Qualitative results of semantic segmentation of S3DIS under different methods. (a) Original images; (b) ground truth; (c) ResGCN-28; (d) GAC; (e) ResGAC-28

其中 4 个场景分别为会议室、办公室、厕所和仓库。

表 2 为不同方法在 S3DIS 中区域 5 的测试结果,其中 SPG 为超点图。从表 2 可以看到,ResGAC-28 在窗户、桌子和沙发这类几何形状上十分相似且难以区分的物体上均取得了很好的效果,mIoU 达到了 63.3%。这是因为所提的 DeepGAC

在计算分配邻域点权重的图注意力系数的过程中,不仅考虑了空间位置信息,还考虑了特征属性信息,当面对空间几何上难以区分的点云时,仍然可以依据特征属性来鉴别点云。对于柱体和杂物等物体,ResGAC-28 网络的语义分割结果略差,这或许是因为这类物体的特征属性信息不显著。

表 2 不同方法在 S3DIS 中区域 5 的测试结果

Table 2 Test results of different methods in region 5 of S3DIS

unit: %

Type	PointNet	ResGCN-28	SPG ^[21]	PointCNN ^[22]	GAC	ResGAC-28 (ours)
OA	—	—	86.4	85.9	87.8	88.0
mIoU	41.1	52.5	58.0	57.3	62.9	63.3
Ceiling	88.8	—	89.4	92.3	92.3	92.4
Floor	97.3	—	96.9	98.2	98.3	97.5
Wall	69.8	—	78.1	79.4	81.9	80.1
Beam	0.1	—	0	0	0	0
Column	3.9	—	42.8	17.6	20.4	21.6
Window	46.3	—	48.9	22.8	59.1	60.4
Door	10.8	—	61.6	62.1	40.9	51.5
Chair	52.6	—	75.4	80.1	78.5	78.1
Table	58.9	—	84.7	74.4	85.8	85.9
Bookcase	40.3	—	69.8	66.7	61.7	64.6
Sofa	5.9	—	52.6	31.7	70.8	71.0
Board	26.4	—	2.1	62.1	74.7	69.8
Clutter	33.2	—	52.2	56.7	52.8	50.0

3.3 消融实验与分析

为了进一步分析网络中各因素带来的影响,通过多项消融实验来研究不同的聚集方式、空间位置信息、特征属性信息、残差连接、动态图卷积和模型参数对测试结果的影响。

3.3.1 聚集方式的影响

为了评估注意力聚集方式的有效性,将其与几种常见的图卷积聚集方式进行比较。表 3 对比几种聚集方式的测试结果。从表 3 可以看到,注意力聚集的 mIoU 值比最大池化方式和平均方式分别提高了 3.7 个百分点和 9.4 个百分点。这是由于最大池化方式倾向于学习边缘部分,平均方式偏向学习分布信息,而注意力聚集方式可以根据邻域点的重要程度来选择性地聚集邻域点特征。

表 3 不同聚集方式的测试结果

Table 3 Test results for different aggregation modes

Aggregation mode	OA / %	mIoU / %
Max-pooling	86.0	60.8
Mean	83.3	55.1
ResGAC-28(ours)	88.2	64.5

3.3.2 空间位置信息和特征属性信息的影响

在计算图注意力系数 e_{ij} 的同时考虑了空间位

置信息和特征属性信息,为了评估二者对实验结果的影响,分别只考虑空间位置和特征属性,测试结果如表 4 所示。从表 4 可以看到,不考虑特征属性的 mIoU 值降低了 1.7 个百分点,不考虑空间位置的 mIoU 值降低了 1.9 个百分点。这表明空间位置信息与特征属性信息同样重要,空间位置信息使点连接成有意义的物体,特征属性信息通过给邻域点赋予不同的权重来区分目标边界。

表 4 空间位置信息和特征属性信息的测试结果

Table 4 Test results of spatial location information and characteristic attribute information

Formula mode	OA / %	mIoU / %
Spatial location information	87.9	62.8
Characteristic attribute information	87.7	62.6
ResGAC-28(ours)	88.2	64.5

3.3.3 残差连接和动态图卷积的有效性

为了评估残差连接和动态图卷积的有效性,分别设计去除残差连接的网络(No Res-28)和使用固定图结构的网络(Fixed-28),残差连接和动态图卷积的有效性测试结果如表 5 所示。从表 5 可以看到,使用残差连接和动态图卷积的 mIoU 值分别提高了 11.8 个百分点和 3.7 个百分点,证明了二者的有效性。

表 5 残差连接和动态图卷积的有效性测试结果

Table 5 Validity test results of residual join and dynamic graph convolution

Network structure	OA / %	mIoU / %
No Res-28	82.6	52.7
Fixed-28	86.1	60.8
ResGAC-28(ours)	88.2	64.5

3.3.4 模型参数的调整

为了评估网络层数 L 、滤波器的个数/通道数 n 和邻域点个数 k 对测试结果的影响,通过消融实验来调整模型的参数,模型参数调整的测试结果如表 6 所示。从表 6 可以看到,当网络层数 L 为 28、滤波器的个数/通道数 n 为 64 和邻域点个数 k 为 16 时,结果最优。

表 6 模型参数调整的测试结果

Table 6 Test results of model parameter adjustment

Parameter	mIoU / %	Δ mIoU / %	L	n	k
Reference	64.5	0	28	64	16
Depth	61.3	-3.2	7	64	16
	62.2	-2.3	14	64	16
Width	58.6	-5.9	28	16	16
	61.5	-3.0	28	32	16
Neighbors	61.6	-2.9	28	64	4
	62.4	-2.1	28	64	8

4 结 论

提出一种用于点云语义分割的深度图注意力卷积神经网络。先优化图卷积聚集函数,然后应用图注意力卷积将图卷积从固定标准化的聚集方式变为依据图注意力机制分配权重的聚集方式,这可以更好地学习点云的结构特征。通过残差连接来解决图卷积网络过深而导致的梯度消失和网络退化的问题,从而加深网络层数,进而获得更复杂的特征。使用动态图卷积在每层图卷积之后重新构建图形,可以更好地表征图结构。实验在 S3DIS 上对网络进行训练和测试,使用 k -折交叉验证法和区域 5 进行测试,发现 ResGAC-28 的 mIoU 值分别达到了 64.5% 和 63.3%。实验结果表明,该网络在沙发和桌子这类几何形状上相似且难以区分的物体上表现效果较好,但对于柱体和杂物这类特征信息不够明显的物体,测试结果略差。如何更好地处理这类物

体,将是下一步研究的重点。

参 考 文 献

- [1] Rusu R B, Marton Z C, Blodow N, et al. Persistent point feature histograms for 3D point clouds [C] // International Conference on Intelligent Autonomous Systems (IAS-10), November 3-7, Baden, Germany. [S.l. : s.n.], 2009: 119-128.
- [2] Rusu R B, Blodow N, Beetz M. Fast point feature histograms (FPFH) for 3D registration [C] // 2009 IEEE International Conference on Robotics and Automation, May 12-17, 2009, Kobe, Japan. New York: IEEE Press, 2009: 3212-3217.
- [3] Rusu R B, Bradski G, Thibaux R, et al. Fast 3D recognition and pose using the Viewpoint Feature Histogram [C] // 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, October 18-22, 2010, Taipei, Taiwan, China. New York: IEEE Press, 2010: 2155-2162.
- [4] Zhang X X, Ji X G, Hu H T, et al. Point cloud segmentation method for complex micro-surface based on feature line fitting [J]. Laser & Optoelectronics Progress, 2020, 57(6): 061502.
张溪溪, 纪小刚, 胡海涛, 等. 基于特征线拟合的微型复杂曲面点云分割方法 [J]. 激光与光电子学进展, 2020, 57(6): 061502.
- [5] Hou B, Jin S Z, Wang Y, et al. Comparison of point cloud registration methods in coarse registration [J]. Laser & Optoelectronics Progress, 2020, 57(8): 081502.
侯彬, 金尚忠, 王赞, 等. 点云配准方法在粗配准中的比较 [J]. 激光与光电子学进展, 2020, 57(8): 081502.
- [6] Zhang J Y, Zhao X L, Chen Z. Review of semantic segmentation of point cloud based on deep learning [J]. Laser & Optoelectronics Progress, 2020, 57(4): 040002.
张佳颖, 赵晓丽, 陈正. 基于深度学习的点云语义分割综述 [J]. 激光与光电子学进展, 2020, 57(4): 040002.
- [7] Hu W, Huang Y Y, Wei L, et al. Deep convolutional neural networks for hyperspectral image classification [J]. Journal of Sensors, 2015, 2015: 258619.
- [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 (CVPR), 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]//Advances in Neural Information Processing Systems, December 4-9, 2017, Long Beach, California, USA. New York: Curran Associates, 2017: 5105-5114.
- [10] Wang X J, Ma J, Wang N N, et al. Deep learning model for point cloud classification based on graph convolutional network[J]. *Laser & Optoelectronics Progress*, 2019, 56(21): 211004.
王旭娇, 马杰, 王楠楠, 等. 基于图卷积网络的深度学习点云分类模型[J]. *激光与光电子学进展*, 2019, 56(21): 211004.
- [11] Hou X D, Yu X X, Liu H P. 3D point cloud classification and segmentation model based on graph convolutional network[J]. *Laser & Optoelectronics Progress*, 2020, 57(18): 181019.
侯向丹, 于习欣, 刘洪普. 基于图卷积网络的三维点云分类分割模型[J]. *激光与光电子学进展*, 2020, 57(18): 181019.
- [12] Wang Y, Sun Y, Liu Z W, et al. Dynamic graph CNN for learning on point clouds[EB/OL]. (2019-06-11) [2020-09-06]. <https://arxiv.org/abs/1801.07829>.
- [13] Xu K, Ba J, Kiros R, et al. Show, attend and tell: neural image caption generation with visual attention [C]//2015 IEEE International Conference on Machine Learning, July 6-11, 2015, Lille, France. New York: IEEE Press, 2015: 2048-2057.
- [14] 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. New York: IEEE Press, 2019: 10288-10297.
- [15] 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), October 27-November 2, 2019, Seoul, Korea. New York: IEEE Press, 2019: 9266-9275.
- [16] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [17] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2261-2269.
- [18] Hamilton W L, Ying Z, Leskovec J, et al. Inductive representation learning on large graphs[C]//Advances in Neural Information Processing Systems, December 4-9, 2017, Long Beach, California, USA. New York: Curran Associates, 2017: 1024-1034.
- [19] Yan S J, Xiong Y J, Lin D H. Spatial temporal graph convolutional networks for skeleton-based action recognition [C] // The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18), February 2-7, 2018, Hilton New Orleans Riverside, New Orleans, Louisiana, USA. New York: AAAI, 2018: 7444-7452.
- [20] 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 (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1534-1543.
- [21] 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.
- [22] Li Y, Bu R, Sun M C, et al. PointCNN: convolution on X-transformed points [C] // Advances in Neural Information Processing Systems, December 3-8, 2018, Montreal, Candad. New York: Curran Associates, 2018: 820-830.