

基于动态图卷积的混合注意力点云特征学习网络

陈奔, 谢晓尧, 刘嵩

(贵州师范大学 贵州省信息与计算科学重点实验室, 贵州 贵阳 550001)

摘要:当前针对点云模型的形状分类、部件分割等工作的深度学习网络缺乏从全局和局部两个角度综合利用上下文信息的能力,从而阻碍了点云对象在细节部分的准确度性能。因此,该文提出了一种通道注意力与局部区域注意力相结合的混合注意力机制,并基于动态图卷积构建了新的三维点云特征学习网络 EDANet。通过通道注意力优化边卷积的处理过程,充分挖掘点云的局部特征信息,同时使用局部区域注意力从全局视角提取上下文特征,并将两部分信息相结合并逐层传递,从而提高特征提取效果网络的能力。在 ModelNet40 与 ShareNet 数据集上分别进行了点云形状分类、部件分割实验,实验结果表明在 ModelNet40 数据集上,EDANet 网络的总体精度 OA 达到了 92.9%,相较于 PointNet、PointNet++、DGCNN 分别高出了 3.7%、2.3%、0.7%。在 ShareNet 数据集上平均并交比(mIoU)达到了 86%,相较于 PointNet、PointNet++、DGCNN 分别高出了 2.3%、0.9%、0.8%。相比其他深度学习网络也具有不同程度的性能提高。验证了该方法在点云处理任务中的优越性能。

关键词:点云分类;点云分割;动态图卷积;自注意力;深度学习;特征提取

中图分类号:TP391.41

文献标识码:A

文章编号:1673-629X(2023)10-0157-06

doi:10.3969/j.issn.1673-629X.2023.10.024

3D Point Cloud Feature Learning Network Based on Hybrid Attention and Dynamic Graph Convolution

CHEN Ben, XIE Xiao-yao, LIU Song

(Key Laboratory of Information and Computing Science of Guizhou Province,
Guizhou Normal University, Guiyang 550001, China)

Abstract: The current deep learning network for shape classification and part segmentation of point cloud models cannot comprehensively utilize contextual information from both global and local perspectives, thus hindering the accuracy performance of point cloud objects in detail. Therefore, we propose a hybrid attention mechanism combining channel attention and local area attention and build a new 3D point cloud feature learning network EDANet based on dynamic graph convolution. By optimizing the processing of edge convolution through channel attention, the local feature information of the point cloud is fully mined, while the context features are extracted from the global perspective by using local regional attention, and the two parts of information are combined and transmitted layer by layer, thus improving the capability of the feature extraction effect network. Point cloud shape classification and component segmentation experiments are carried out on ModelNet40 and ShareNet datasets respectively. The experimental results show that on the ModelNet40 dataset, the overall OA accuracy of the EDANet network reaches 92.9%, which is 3.7%, 2.3%, and 0.7% higher than that of PointNet, PointNet++, and DGCNN, respectively. The mean IoU ratio (mIoU) on the ShareNet dataset reached 86%, which is 2.3%, 0.9%, and 0.8% higher than that of PointNet, PointNet++, and DGCNN, respectively. Compared with other deep learning networks, it also has different degrees of performance improvement. The superior performance of the proposed method is verified in the point cloud processing task.

Key words: point cloud classification; point cloud segmentation; dynamic graph convolution; self-attention; deep learning; feature extraction

0 引言

点云模型通常由通过专业设备对物体表面采样得

到的离散点集构成,一般由各离散点的三维坐标表示,部分点云数据还包含法向量、颜色特征等信息。因为

收稿日期:2022-11-30

修回日期:2023-03-31

基金项目:四川省重点研发计划项目(2021YFS0401;2022YFS0558)

作者简介:陈奔(1997-),男,硕士研究生,研究方向为计算机视觉、点云处理;通信作者:谢晓尧(1952-),男,教授,博士,研究方向为网络通信、信息安全与人工智能。

包含信息丰富,基于点云数据的点云模型在自动驾驶、医学成像、遥感测绘、虚拟现实等领域^[1-2]均有广泛应用。传统方法手工设计几何形状^[3]或签名描述符^[4]的方法对研究人员的专业知识和针对性设计能力要求颇高,逐渐难以适应当下快速发展的需要,因此深度学习被引入点云处理领域以解决这些问题^[5]。

面对三维点云的空间分布特点,深度学习主要有两个大的研究方向:

(1)先对点云数据进行预处理,得到规则结构化的数据,再使用图像处理中成熟的卷积神经网络模型进行操作^[6-8]。

例如多视图投影^[9]和体素网络^[10]两个最具代表性的技术路线。这些方法通过将不规则、非结构化的点云数据转化为规则、结构化的图像或体素网络,随后使用二维或三维 CNN 提取点云的局部和全局特征。这些方法虽然解决了点云的无序分布问题,但也带来了许多计算上的挑战和额外的内存消耗问题。

(2)不对原始数据进行过多干预,直接送入深度学习网络中处理。通过针对原始点云数据设计特殊的卷积神经网络来学习点云的特征表示^[11]。这种方法由于内存消耗较低、结构简单越来越受到研究人员的关注。

其中 PointNet^[12]在点云分类和分割任务上取得了很好的效果,其通过多层感知机 MLP 逐点提取高维特征再通过池化操作得到全局特征。整个过程与点云数据的输入顺序无关,可以很好地适应点云数据的置换不变特性,但由于 PointNet 网络没有划分点云区域以及针对性提取区域特征,这对其处理局部细节的性能影响较大。

PointNet++^[13]在 PointNet 的基础上采用多层次结构提取特征,应用 KNN 分组和层次化结构将点云划分为多个局部区域,允许重复捕获局部信息并多次叠加得到最终效果,一定程度上弥补了 PointNet 在局部细节方面的不足。然而这种方法至少简单地划分局部区域,并没有能够进一步发掘局部点之间存在的内在关系。同时提取局部特征的过程是相互独立的,点云子区域之间没有信息交换导致部分点云结构信息的丢失。也由于它只是将局部特征简单地聚合,所以区域信息还没有被充分利用。

DGCNN^[14]为代表的图卷积技术路线^[15]则提出了一种边缘卷积的方法,在特征空间的近邻点中通过 KNN 采样构建局部图结构,得到邻边特征。通过共享权重的 MLP 对图结构的邻边特征进行卷积操作从邻边集合中提取特征信息,随后使用最大池化的方式进行聚合进一步得到局部特征,随后在网络结构的每一层中动态更新局部图结构。这种方法强化了局部结构

中点与点之间的联系,然而 DGCNN 只针对特征空间中的领域,因此具有一定的局限性。同时由于点云的稀疏性和非均匀分布特点,邻域内每个节点和中心节点的关联度都是不同的,而 DGCNN 却平等地处理每一条邻边。这种提取方法有一定局限性,导致尽管 DGCNN 能够很好地采集低级语义信息,却无法很好地描述高级语义信息和隐式高级语义特征。

综上所述,目前点云深度学习领域需求旺盛,然而现有主流点云深度学习网络^[16-17]在设计上存在对坐标空间中各区域间信息传递的不够重视、区域内近邻点聚合方式单一的问题,导致模型对区域信息、高级语义信息未能充分利用,限制了深度学习网络提取更多特征信息进行点云分类分割任务的能力。

近年来,在 NLP 等领域大火的注意力机制^[18]为点云深度学习网络的设计带来了新的思路。注意力机制的核心设计理念是在众多输入信息中提高对当前任务更为重要的信息的关注度,相对降低对其他信息的关注度,甚至直接过滤无关信息,以此解决庞大信息量带来的信息过载问题,同时提高任务处理的效率和准确性。

综上所述,为了更好地应对现有主流点云深度学习网络在设计上存在的问题,该文提出了一种新的三维点云分析网络 EDANet,尝试通过在特征空间和坐标空间中分别引入特征通道注意力和局部区域注意力机制,在坐标空间中以更广泛的视角提炼点云特征信息,同时有区别地对待同一区域内不同近邻点所提供的点云信息。以此增强点云特征的表达能力,实现点云模型的高精度形状分类和部件分割。同时由于没有点云下采样过程,让分割任务避免在上采样过程中的信息损失,进一步提高了点云分割精度。

1 方法介绍

1.1 EDANet 架构

如图 1 所示,EDANet 由多层 EdgeDualAttention 模块组成。输入原始点云集合 $P_{in} = \{P_1, P_2, \dots, P_n\}$, $P_{in} \in R^{n \times c}$,其中 n 是原始点云的点数量, c 是每个点的初始维度,一般为三维坐标参数,个别情况下还包含颜色或法向量等额外维度的信息。EDANet 首先使用一层特征嵌入通过 MLP 将输入原始点云数据的维度 c 向更高的维度 d 映射,方便后续处理的同时增强点云的信息表达能力。

其中,EdgeDualAttention 从局部和全局两个方面综合计算了注意力权重,通过对这两个模块得到的权重信息进行加权求和来深入挖掘每个点的上下文表示。逐层聚合不断提炼积累特征信息,得到不同尺寸的点云特征矩阵,最终将不同尺寸的逐点特征拼接通

过 max-pooling 与 avg-pooling 相结合的池化手段得

到全局特征用于完成形状分类任务。

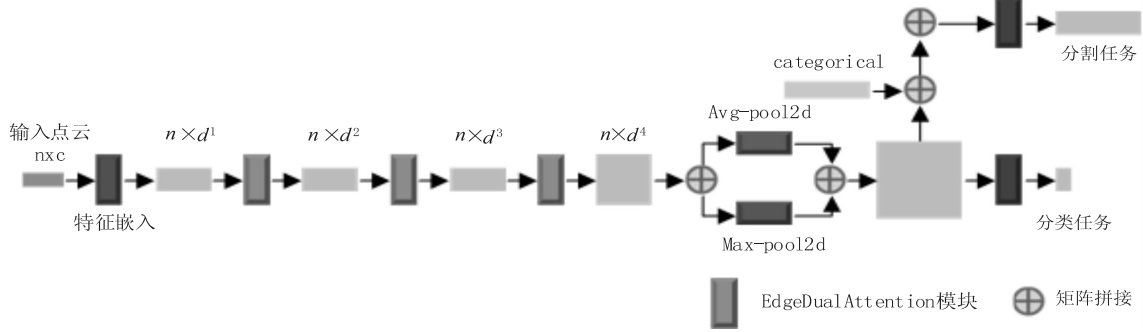


图 1 EDANet 架构

分割模型与分类模型基本一致,但由于部件分割需要获取逐点特征,因此分割网络需要在分类模型得到的全局特征基础上 repeat 并与之前各层得到的不同尺寸的逐点特征拼接,以此补全得到同时具备全局特征信息与局部特征信息的逐点特征,用于完成部件分割任务。

1.2 EdgeDualAttention 模块

如图 2 所示,EdgeDualAttention 由三个主要部分构成:Edge Create 模块负责构建一个边缘特征集合,Dual Attention 模块负责从全局和局部两个方面综合考虑点对之间的相关性,以获得更丰富的上下文信息,Feature Aggregation 模块负责聚合采集得到的特征信息并送入下一层网络。

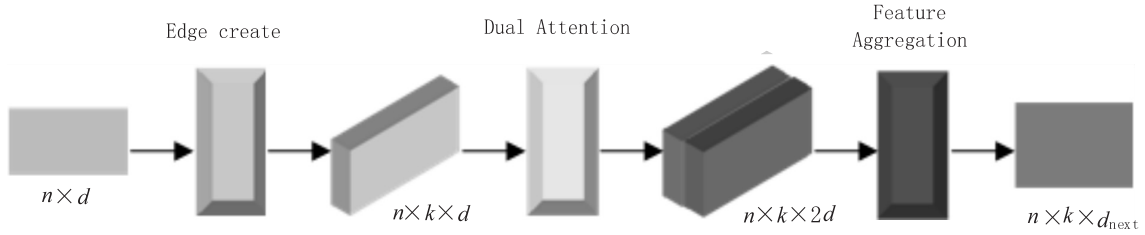


图 2 EdgeDualAttention 模块

Edge Create:输入点云数据 $P_{in} \in R^{n \times d}$,通过 KNN 采集每个点的 k 个近邻点组成近邻点集分类模型接收 n 个点作为输入,为每个点采集其在特征空间中的 k 个近邻点组成边缘特征集 $L = \{L_1, L_2, \dots, L_n\}$, $L \in R^{n \times k \times d}$,其中 $\{L_i \in R^{k \times d} \mid i = 1, 2, \dots, n\}$ 表示每个局部邻域,通过 L_i 与 K 可以唯一确定一个区域的某一个点。

Dual Attention:如图 3 所示,为了更好地表述高维

语义信息同时增加特征之间的区分度,该文构建了特征通道注意力模块 Channel Attention Module (CAM) 和局部区域注意力模块 Local Region Attention Module (LRAM) (将在后续详细描写)。这两个模块基于多头自注意力构成,主要功能分别是对任意通道之间的相互依赖进行建模以及计算任意区域之间的相互依赖权重。

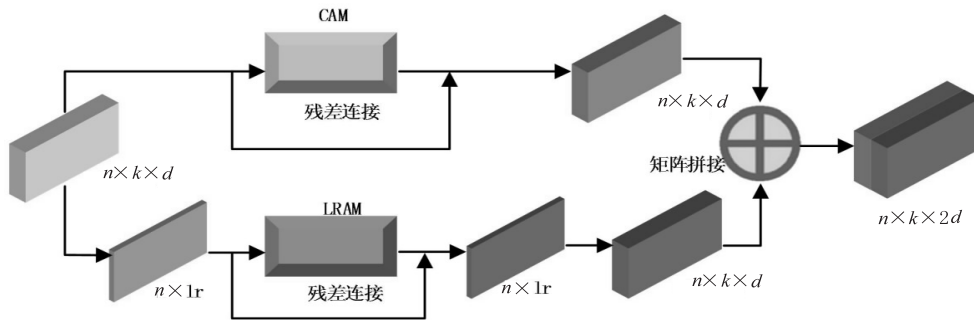


图 3 Dual Attention 子模块组成

自注意力机制作为注意力机制的改进^[19],其减少了对外部信息的依赖,更擅长捕捉数据或特征的内部相关性。常规卷积层通过卷积核和原始特征的线性组合得到输出特征,由于卷积核通常是局部的,导致许多

任务因语义信息不足影响了最终性能。而自注意力机制通过捕捉全局的信息来获得更大的感受野和上下文信息,对不同的特征和点进行了针对性的处理,为处理不同类型的信息提供了额外的灵活性,扩展了 CNN 的

表达能力。同时为了克服局部感受野的极端数值感染,为了更全面地获取不同层次的特征信息, Dual Attention 引入了多头注意力的架构。

多头注意力定义如下:

$$\text{Multihead}(E) = (\text{head}_1 \oplus \cdots \oplus \text{head}_n) W^o$$

其中, \oplus 表示矩阵拼接操作, W^o 表示一个可学习参数, 以更好地聚合多头注意力所采集的信息。 head_i 代表多头注意力架构中的第 i 个 CAM 头或者 LRAM 头。通过引入多头注意力的架构使得输入每个 CAM 或 LRAM 模块的向量集合 E 的特征维度 d 下降为 d/h 。这在一定程度上可以降低模块的计算复杂度, 增加并行运算效率。

通道注意力模块 CAM: 如图 4 所示, 首先通过卷积运算处理输入的向量集合 E , 通过不同的可学习参数 W_q, W_k, W_v 将其特征维度从 d 投影到 d' , 分别得到矩阵 $Q \in R^{n \times k \times d}, K \in R^{n \times k \times d'}, V \in R^{n \times k \times d}$ 。然后将 Q 矩阵转置, 再通过 Q^T 与 K 矩阵相乘以及归一化得到特征维度层面的相似度矩阵 $A \in R^{n \times d \times d}$, 将 A 与 V 矩阵相乘得到最终的经过权重缩放的目标矩阵 $Z \in R^{n \times k \times d}$ 。通过使用所有特征维度映射的加权和更新每个特征维度的映射来捕获任意两个特征维度映射之间的依赖关系, 以此达成对区域内信息的更高效聚合。

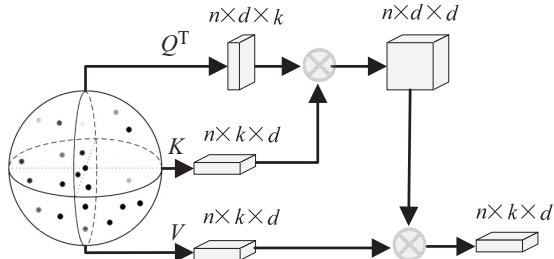


图4 通道注意力

当前大多数点云分析网络在对邻域信息采样聚合时选择了所有节点共享相同的卷积核参数, 然而近邻点对于中心点的影响是各不相同的, 这必然会限制模型的能力。通过引入特征注意力机制重新构建相邻节点的关系, 对邻域节点信息做到有区别的聚合, 实际上可以视为邻域的每个节点隐式分配不同的卷积核参数。

局部区域注意力模块 LRAM: 如图 5 所示, 首先通过与 CAM 相同的操作得到矩阵 QKV , 重构 QKV , 得到 $Q' \in R^{N \times lr}$, 其中 $lr = k * d'$, K', V 同理。然后将 Q' 矩阵转置与 K' 矩阵相乘归一化得到局部区域层面的相似度矩阵 $A' \in R^{N \times N}$, 将 A' 与 V 矩阵相乘得到最终的经过权重缩放的目标矩阵 $Z' \in R^{N \times lr}$, 经过重构得到 $Z'' \in R^{n \times k \times d}$ 。通过使用所有区域映射的加权和更新每个区域的映射对各区域之间的依赖关系加以捕获, 以此加强区域信息之间的流动。

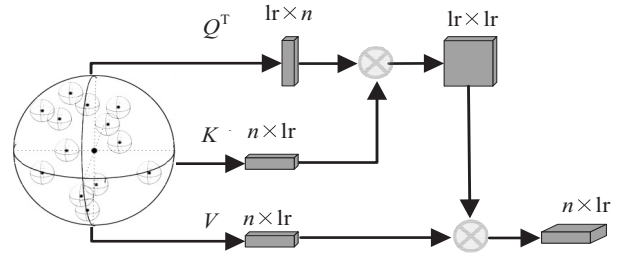


图5 局部区域注意力

当前大多数点云分析网络只是孤立地学习每个点的特征信息, 未能充分探索上下文表示, 变相地放弃了开发每个点之间的相关信息, 这直接导致具有较强上下文表示的点及区域权重过大主导整个场景。通过引入局部区域注意力机制重新评估相关区域的关系, 改进非优势目标的上下文表示, 有意识地增加非优势目标的权重。

Dual Attention 可以总结为:

$$F_{out} = \text{CAM}(F_{in}) \oplus \text{LRAM}(F_{in})$$

通过构建 CAM 和 LRAM 模块计算每个特征图通道在全局点和局部区域之间的相互依赖权重来更新特征信息, 确保所需提取的特征具有更高的特征区分度。因为特征通道与局部区域的权重需要随客观变化而变化, 所以最后通过两个可学习参数 W_{CAM} 与 W_{LRAM} , 对 CAM 得到的特征矩阵 Z 以及 LRAM 得到的特征矩阵 Z' 分别将其特征维度从 d' 映射回 d , 得到 $F_{CAM} \in R^{n \times k \times d}, F_{LRAM} \in R^{n \times k \times d}$, 再将 F_{CAM} 与 F_{LRAM} 拼接得到输出特征矩阵 $F_{out} \in R^{n \times k \times 2d}$, 以此增强局部邻域特征矩阵的灵活性。

Feature Aggregation: 特征聚合公式可以表示为:

$$G_{next} = \max - \text{avg} - \text{pool}(\text{MLP}(\text{concat}(p_i; \alpha(p_j) - p_i)))$$

其中, α 函数表示 Dual attention 模块, 输入代表局部邻域信息的边向量集合, 输出经过通道及局部区域注意力权重加成的邻域信息集合, 将其与中心点 p_i 相减得到区域内的邻域特征。最后将中心点 p_i 与其拼接送入 MLP 处理, 经过 maxpool 和 avgpooling 相结合的池化操作处理聚合得到一个融合了全局形状信息以及局部邻域信息的新的全局特征值 G_{next} 送入下一层网络。

2 实验

2.1 实验设置

实验环境为 ubuntu18.04、PyTorch 1.7.0、Python3.8、Cuda11.0, 使用两个 NVIDIA RTX 3080 Ti GPU 训练所有模型。所有任务的 batchsize 大小设置为 24, 迭代 200 次, 初始学习率 $lr = 0.002$, 权重衰减 10^{-4} , 采用 AdamW^[20] 优化器和 CrossEntropyLoss 训练。并在每个 FC 层后都加入比例为 0.5 的 dropout, 以抑制过拟合。

2.2 形状分类任务 Classification

实验采用 ModelNet40^[21] 数据集评估相关模型的形状分类任务性能,在训练分类任务时每次采样 1 024 个点送入模型。ModelNet40 由包含 40 个类别的 12 331 个 CAD 模型组成。数据集分成了 9 843 个训练模型和 2 468 个测试模型。

形状分类任务采用的评价指标为:mAcc (mean Accuracy) 和 OAcc (Overall Accuracy)。mAcc 是所有类别的平均准确率,所有 Class Accuracy 取平均值。而 OAcc 表示不分类别的全体数据的准确性。

对比算法包括经典逐点特征信息聚合方法 PointNet、PointNet++;基于树状结构重新设计了点云信息分析架构的 Kd-Net^[22];经典的体素化点云信息的网络架构 Subvolume^[23];通过多视角转化为二维图像再使用 CNN 进行分析的 MVCNN;通过 SOM 直接将高维数据降维而后处理的 SO-Net^[24];SpiderCNN^[25]则在 CNN 与点云的结合方面做出了卓越贡献、动态图卷积代表 DGCNN。

EDANet 与基于深度学习各技术路线上的诸多经典网络模型使用数据集 ModelNet40 进行了性能对比,如表 1 所示。EDANet 在与这些经典网络模型对比中,无论是 mAcc 还是 OAcc 均具有较高性能提升,由此可见引入通道特征注意力以及局部区域注意力的双重注意力机制显著地提高了分类效果,相关分析将在后续消融实验中详细介绍。

表 1 MobenNet40 上各模型性能对比 %

Model	mAcc	OAcc
PointNet	86.2	89.2
PointNet++	87.8	90.6
Kd-Net	--	88.5
Subvolume	86	89.2
MVCNN	--	90.1
SO-Net	--	90.9
SpiderCNN	--	90.5
DGCNN	90.2	92.2
EDANet	90.5	92.9

2.3 部件分割任务 Part Segmentation

点云分割是一项具有挑战性的任务,旨在将 3D 模型划分为多个有意义的部分。为了进行细粒度的 3D 分割,实验最终选取数据集 ShapeNet^[26] 进行实验评估,该数据集是一个用于零件分割的对象级数据集,注释丰富,规模较大,由 16 种不同大类组成,每个大类包含 2~6 个零件,下属合计 50 个零件标签。作者将包含 16 881 个模型的数据集分成 14 006 个训练模型和 2 875 个测试模型。在训练分割任务时每次采样

2 048 个点送入模型。

部件分割任务采用的评价指标为:IoU (Intersection over Union),表示某一类的预测结果和真实值的交集与其并集的比值。cIoU (category IoU),表示平均所有类的预测结果交并集。mIoU (mean IoU),表示取出所有类别中每一项的预测结果交并集取平均值。ShapeNet 上各模型性能对比如表 2 所示。

表 2 ShapeNet 上各模型性能对比 %

Model	cIoU	mIoU
PointNet	80.4	83.7
PointNet++	81.9	85.1
Kd-Net	77.4	82.3
SO-Net	80.8	84.6
SpiderCNN	81.7	85.3
DGCNN	82.3	85.2
EDANet	83.1	86

可见与近期相关研究相比,EDANet 方法也达到较高水平。

2.4 消融实验

为了验证 CAM 以及 LRAM 组成 Dual attention 在 EDANet 网络中的有效性,尝试在 ModelNet40 数据集上进行了消融实验。除了包含完整 CAM 和 LRAM 的 EDANet,还设计了三组额外模型进行对比,不包含 Dual attention 的模型 EDANet-Dual,包含 CAM 但不包含 LRAM 的模型 EDANet-LRAM,包含 LRAM 但不包含 CAM 的模型 EDANet-CAM。比较结果如表 3 所示。

表 3 不同模块在 EDANet 中的影响 %

Model	mAcc	OAcc
EDANet-Dual	90.1	92.0
EDANet-LRAM	90.3	92.6
EDANet-CAM	90.2	92.4
EDANet	90.5	92.9

可见当 LRAM 与 CAM 均加入 EDANet 时,分类准确率将达到最大值。

3 结束语

当前大多数点云分析网络只是从增强区域特征提取的角度来增强网络模型的表达能力,却忽略了子区域之间特征信息的交换。EDANet 尝试通过局部区域注意力机制以自适应的学习特征权重并获取子区域之间局部特征的相关性。同时特征通道注意力模块专门对每个点的特征的相互依赖关系进行建模,而局部区域注意力模块则将每个局部邻域视为一个整体,并通过捕获每个局部领域之间的特征映射的相互依赖关系来更新权重。EDANet 在多个数据集上的实验结果表

明,其在点云分类分割任务上成功实现了高性能。同时通过消融实验证明了文中网络模型的有效性。

参考文献:

- [1] HU Q, YANG B, XIE L, et al. RandLA-Net: efficient semantic segmentation of large-scale point clouds [C]//2020 IEEE/CVF conference on computer vision and pattern recognition. Seattle: IEEE, 2020: 11105–11114.
- [2] GOJCIC Z, ZHOU C, WEGNER J D, et al. Learning multiview 3D point cloud registration [C]//Proc of IEEE conference on computer vision and pattern recognition. Seattle: IEEE, 2020: 1759–1769.
- [3] OSADA R, FUNKHOUSER T, CHAZELLE B, et al. Shape distributions [J]. ACM Transactions on Graphics, 2002, 21(4): 807–832.
- [4] SUN J, OVSJANIKOV M, GUIBAS L. A concise and provably informative multi-scale signature based on heat diffusion [J]. Computer Graphics Forum, 2009, 28(5): 1383–1392.
- [5] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition [C]//Proceedings of the 2016 IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas: IEEE, 2016: 770–778.
- [6] 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.
- [7] 党吉圣, 杨军. 深度图注意力 CNN 的三维模型识别 [J]. 计算机科学与探索, 2021, 15(1): 141–149.
- [8] 景庄伟, 管海燕, 臧玉府, 等. 基于深度学习的点云语义分割研究综述 [J]. 计算机科学与探索, 2021, 15(1): 1–26.
- [9] SU H, MAJI S, KALOGERAKIS E, et al. Multi-view convolutional neural networks for 3D shape recognition [C]//Proceedings of the IEEE international conference on computer vision. Los Alamitos: IEEE, 2015: 945–953.
- [10] MATURANA D, SCHERER S. VoxNet: a 3D convolutional neural network for real-time object recognition [C]//Proceedings of the IEEE/RSJ international conference on intelligent robots and systems. Los Alamitos: IEEE, 2015: 922–928.
- [11] 李肖南, 王蕾, 程海霞, 等. 基于 SA-PointNetVLAD 的点云分类网络 [J]. 计算机技术与发展, 2022, 32(5): 36–41.
- [12] 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: IEEE, 2017: 77–85.
- [13] CHARLES R Q, SU H, YI L, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space [C]//International conference on neural information processing systems. Long Beach: NIPS, 2017: 5105–5114.
- [14] WANG Y, SUN Y, LIU Z, et al. Dynamic graph CNN for learning on point clouds [J]. ACM Transactions on Graphics, 2019, 38(5): 1–12.
- [15] 谭台哲, 黄永耀, 杨卓, 等. 基于注意力机制的点云神经网络架构搜索方法 [J]. 计算机工程与应用, 2022, 58(19): 142–151.
- [16] ENGEL N, BELAGIANNIS V, DIETMAYER K. Point transformer [EB/OL]. [2022-01-12]. <https://arxiv.org/abs/2011.00931>.
- [17] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//NIPS'17: proceedings of the 31st international conference on neural information processing systems. Long Beach: NIPS, 2017: 6000–6010.
- [18] BELLO I, ZOPH B, VASWANI A, et al. Attention augmented convolutional networks [C]//Proceedings of the IEEE/CVF international conference on computer vision. South Korea: IEEE, 2019: 3286–3295.
- [19] LOSCHILLOV I, HUTTER F. Decoupled weight decay regularization [C]//International conference on learning representations. New Orleans: ICLR, 2019: 1–19.
- [20] 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: IEEE, 2015: 1912–1920.
- [21] 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. Venice: IEEE, 2017: 863–872.
- [22] QI C R, SU H, NIEBNER M, et al. Volumetric and multi-view CNNs for object classification on 3D data [C]//2016 IEEE conference on computer vision and pattern recognition. Las Vegas: IEEE, 2016: 5648–5656.
- [23] LI Jiaxin, CHEN B M, LEE G H. So-Net: self-organizing network for point cloud analysis [C]//Proceedings of the 2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR). Salt Lake City: IEEE, 2018: 9397–9406.
- [24] XU Y F, FAN T Q, XU M Y, et al. Spidercnn: deep learning on point sets with parameterized convolutional filters [C]//Proceedings of the European conference on computer vision. Berlin: Springer, 2018: 87–102.
- [25] 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.