

邹舒航, 李光平, 高祥. 基于隐藏特征的点云模块[J]. 广东工业大学学报, 2025, 42(6): 62–69. doi: 10.12052/gdutxb.240153.
Zou Shuhang, Li Guangping, Gao Xiang. Point cloud module based on hidden features[J]. Journal of Guangdong University of Technology, 2025, 42(6): 62–69. doi: 10.12052/gdutxb.240153.

基于隐藏特征的点云模块

邹舒航, 李光平, 高祥

(广东工业大学 信息工程学院, 广东 广州 510006)

摘要: 点云全局特征和局部特征之间普遍存在隐藏特征, 挖掘并利用隐藏特征可增强点云表示能力。但现有理论中, 对隐藏特征的分析利用并不深入。为进一步挖掘利用隐藏特征, 提出一种基于注意力机制的隐藏特征利用 (Attention-based Hidden Feature Utilization, AHU) 模块。模块由两部分子模块组成, 一方面, 基于通道注意力机制的子模块加强特征通道间的依赖关系, 提升隐藏特征的显著性; 另一方面, 基于交叉注意力机制的子模块将学习到的隐藏特征反投影到原始局部特征上, 并建立二者的长距离依赖关系, 促进信息融合并提高模块泛化能力。本文拓展了隐藏特征的相关理论, 并以实验证明, AHU模块可以集成到现有的先进网络中, 显著提高其性能。

关键词: 点云; 注意力机制; 分类; 分割; 即插即用

中图分类号: TP394.1

文献标志码: A

文章编号: 1007-7162(2025)06-0062-08

Point Cloud Module Based on Hidden Features

Zou Shuhang, Li Guangping, Gao Xiang

(School of Information Engineering, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: Existing research has shown that there are hidden features between the global and local features of point cloud, and the representation capability of point cloud can be enhanced by mining and utilizing the hidden features. However, existing theories have not delved deeply into the analysis and utilization of hidden features. To further exploit the hidden features, we propose an Attention-based Hidden Feature Utilization (AHU) module, which consists of two sub-modules. On one hand, a sub-module based on channel attention mechanism enhances the inter-channel dependencies of features, improving the significance of hidden features; on the other hand, another sub-module based on cross-attention mechanism projects the learned hidden features back to the original local features, establishing long-distance dependencies between them and promoting information fusion, such that the generalization ability of the module can be improved. This paper extends the theory about hidden features and the experimental results demonstrate that the AHU module can be integrated into existing state-of-the-art networks to significantly improve the performance.

Key words: point cloud; attention mechanism; classification; segmentation; plug and play

随着激光扫描技术和多样化的3D传感器的发展, 人们得以更加高效地获取点云, 并将其广泛应用于自动驾驶^[1]、虚拟现实、增强现实以及机器人^[2]等领域, 客观上推动了点云的发展。在早期的方法中, 常利用手工制作的特征^[3]来捕获点云的几何属性, 用以解决点云在分类和分割任务上的问题。近年来出现了基于点的方法, 为了处理点云的不规则属性, 最先进的网络往往是直接基于单个点进行设计, 而不是

引入中间表示。点网络(PointNet)^[4]使用数个多层感知机和对称网络提取点云特征, 使点云对旋转和排列具有不变性, 但是仍存在特征表示不足等缺陷。对此, Qi等^[5]提出PointNet++作为改进模型, 该模型通过在不同分层架构上使用PointNet层捕捉不同尺度区域信息。现有许多方法表明^[6-9], 使用点云的局部特征能够帮助网络模型学习点云。Chen等^[10]没有聚焦于卷积算子, 而是创新性提出局部和全局之间存在

收稿日期: 2024-12-06 录用日期: 2025-02-28 网络首发日期: 2025-06-03

基金项目: 国家自然科学基金资助项目(61601130); 大亚湾科技计划项目(2020010203)

作者简介: 邹舒航(1999-), 男, 硕士研究生, 主要研究方向为点云处理、特征提取、点云分类, E-mail: 875936772@qq.com

通信作者: 李光平(1981-), 男, 副教授, 博士, 主要研究方向为机器学习与信号处理, E-mail: gpli@gdut.edu.cn

的隐藏特征理论,表明隐藏特征能够有效提高点云网络对点云的学习能力。隐藏特征的灵感来自人脑的思考过程,为点云相关研究提供了新的角度,但有些问题仍待解决。具体来说,在有关于隐藏特征的理论中,对特征通道之间的相关性关注不足,导致现有隐藏特征研究普遍缺乏重新校准特征通道的能力。其次,目前关于隐藏特征的已有方法无法有效地整合隐藏特征和原始局部特征之间的几何和语义信息,限制了信息传递,对特征表示能力仍存在优化空间。

本文优化和改进了Chen等^[10]的工作,并创新性提出基于注意力机制的隐藏特征利用模块,提出用通道注意力捕获点云中的隐藏特征,有效地提升特征通道之间的相互依赖性,提高隐藏特征表达能力。此外,利用交叉注意力机制作为建立隐藏特征与原始局部特征间长距离依赖关系的手段,将学习到的隐藏特征映射至原始局部特征。

综上所述,本文模块贡献如下:

(1) 对特征通道之间的相互依赖关系进行建模,自适应地重新校准通道级特征响应以捕获隐藏特征。

(2) 利用交叉注意力机制实现隐藏特征与原始局部特征的融合,建立二者间的长距离依赖关系,以此增强模块表达能力。

(3) 实验结果表明,AHU模块能够集成到现有有点云网络,并在点云分类和分割任务中实现显著的性能提升。

1 相关工作

1.1 局部特征学习

点网络PointNet模型的出现开创了点云深度学习的新趋势。PointNet模型以原始点云为输入,并从点云中提取特征进行分类,是基于点的方法的开创性工作,但其几何信息获取能力仍存在不足。为解决上述问题,PointNet++引入K近邻采样算法,聚合局部特征信息,提取特征点间的几何联系。

图卷积神经网络(Dynamic Graph CNN, DGCNN^[11])模型将图卷积思想与点云特征学习相结合,探索中心点与邻近点间的关系并通过边缘卷积算子提取特征。自适应图卷积模型(Adaptive Graph Convolution, AdaptConv^[12])动态学习中心点特征,并生成自适应卷积内核,从而提取更灵活的局部几何特征。

1.2 点云的注意力机制

自注意力模型Transformer^[13]在自然语言处理中的优异表现开创了注意力机制在点云相关领域的应

用热潮。基于注意力的序列到序列网络学习3D点云表示模型Point2Sequence^[14]通过将点云表示为序列模型作为一种新颖的方法,通过注意力机制捕获局部区域特征之间的相关性,进而突出区域在多尺度上的重要性。点云自注意力模型PCT^[15] (Point Cloud Transformer)引入偏移注意力模块,计算并利用自注意力模块的输出特征与输入特征间的偏移量,强调每个点的相对坐标,提高了网络的鲁棒性。点自注意力模型Point Transformer^[16]设计出具有表现力的点变换层和点变换器块,应用于每个中心点周围的邻近点,从而生成用于分类或语义分割的特征表示。

1.3 隐藏特征

由Chen等^[10]提出的基于局部全局反投影的误差特征学习模块(Error feature Back-projection based Local-Global feature learning module, EB-LG)开创性地提出隐藏特征概念,为点云的深度学习引入了全新角度。EB-LG模块捕获局部和全局特征之间的隐藏特征并将其反向投影到原始局部特征,达到增强原始局部特征的目的,在分类和分割任务上表现优异。通过实验证明捕获和利用隐藏特征对现有有点云网络能实现有效提升。

2 方法介绍

基于注意力机制的隐藏特征利用(Attention-based Hidden Feature Utilization, AHU)模块总体上(见图1)由两个子模块组合而成。第一子模块为隐藏特征提取模块(见图2),第二子模块为偏移量反向投影模块(见图3)。图2中包含全局特征提取器(见图4)。

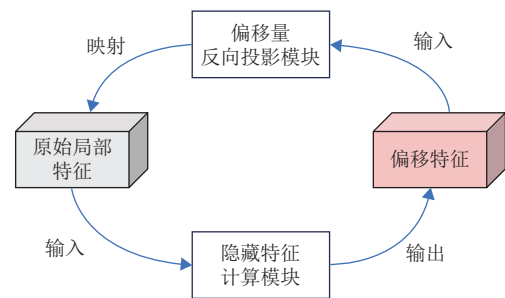


图1 AHU模块总览

Fig.1 The overview of the AHU module

2.1 隐藏特征计算模块

首先如图2所示,无序点云集合 $\mathcal{X} = \{x_i | i = 1, 2, \dots, N\} \in \mathbb{R}^{N \times 3}$,其中, x_i 代表第*i*个点的(*x*,*y*,*z*)坐标,经过现有有点云网络模型中的卷积算子编码计算得到点云局部特征表示 $\mathbf{F} = \{f_i | i = 1, 2, \dots, N\} \in \mathbb{R}^{N \times D}$,其

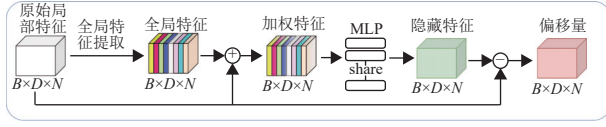


图2 隐藏特征计算模块

Fig.2 The illustration of the hidden feature calculation module

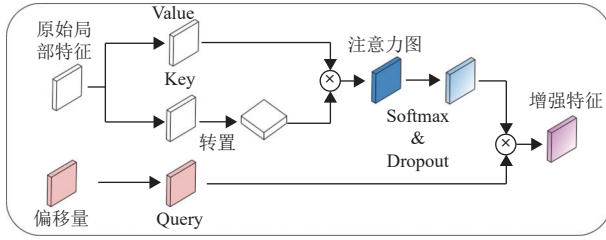


图3 基于交叉注意力机制的偏移量反向投影模块

Fig.3 Cross-attention based offset feature back-projection module

中, B 代表Batch size, 即单一批次训练使用的样本数量, N 代表点的数量, D 代表第 i 个点的特征维度, f_i 表示第 i 个点的特征。将其作为模块的原始局部特征输入, 并从中聚合全局特征信息 G (由式(1)计算得

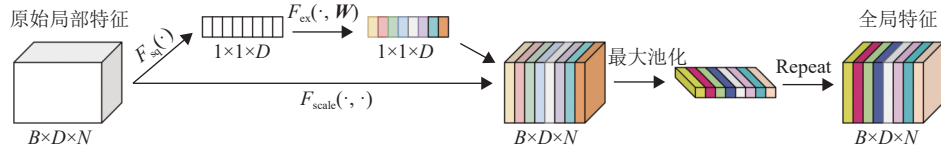


图4 全局特征提取

Fig.4 Global feature extractor

为将全局信息压缩到一个通道描述符中, 先使用全局平均池化方法整合通道信息。从形式上讲, 首先收缩原始局部特征 F 的 $B \times N$ 上的空间维度, 生成统计量 $z \in \mathbb{R}^D$ 。通道 c 的全局描述由式(4)计算得到

$$z_c = F_{sq}(f_c) = \frac{1}{B \times N} \sum_{i=1}^B \sum_{j=1}^N f_c(i, j) \quad (4)$$

式中 $F_{sq}(\cdot)$ 表示对输入的原始局部特征相应特征通道的挤压操作, $f_c(i, j)$ 表示第 c 个通道在位置 (i, j) 处的值。图4中 $F_{ex}(\cdot)$ 旨在建模特征通道间的依赖关系, 以式(4)计算得到的 $z = [z_1, z_2, \dots, z_D]$ 作为输入, 计算得到通道注意力描述符 $s \in \mathbb{R}^D$ 作为加权函数, 用于重新校准输入特征, 从而突出信息通道的贡献。 s 的定义为

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \quad (5)$$

式中 δ 为RELU^[18]激活函数, σ 为sigmoid激活函数, $W_1 \in \mathbb{R}^{\frac{D}{r} \times D}$, $W_2 \in \mathbb{R}^{D \times \frac{D}{r}}$ 分别为两层具有可学习参数的全连接层, r 为降维比率, 旨在限制模块复杂度并

到)。应用多层感知机对结果进行进一步计算得到隐藏特征 F_h , 并进一步计算隐藏特征与输入的原始局部特征的偏移特征 F_e 。相关计算公式如式(1)~(3)所示。

$$G = h(f_1, f_2, \dots, f_N) \in \mathbb{R}^{N \times D} \quad (1)$$

$$F_h = \text{MLP}(F \oplus G) \in \mathbb{R}^{N \times D} \quad (2)$$

$$F_e = F - F_h \quad (3)$$

式(1)~(3)中 $h(\cdot)$ 代表全局特征提取过程(由式(4)~式(7)组合而成), 经上述过程得到全局特征 G , MLP代表用于学习隐藏特征高级表示的多层感知机网络, F 、 F_h 和 F_e 分别代表局部特征、隐藏特征和二者偏移量, $\oplus$ 代表逐点相加运算。

2.2 全局特征提取过程

不同于EB-LG, 本文选择基于通道注意力机制的描述符^[17], 意在获取更加精确的全局特征表示。通过利用通道注意力机制, 模块能够选择性地强调有效特征, 同时忽略不太相关的特征(如图4所示)。

提高泛化能力, 本文遵循文献^[17]的实验设置, 将其设置为16, 在该条件下, 通道注意力机制能够实现性能与计算复杂度的最佳平衡。通过应用通道描述符 s 达到重新缩放原始输入特征的目的。

$$\tilde{f}_c = F_{scale}(f_c, s_c) = s_c \cdot f_c \quad (6)$$

$$G = \text{Repeat}(\text{Maxpooling}(\tilde{F})) \quad (7)$$

式(6)~(7)中 $\tilde{F} = [\tilde{f}_1, \tilde{f}_2, \dots, \tilde{f}_D]$, $F_{scale}(f_c, s_c)$ 代表第 c 个输入原始局部特征通道与标量 s_c 进行通道级相乘。 G 为最终得到的具有较高可分辨性的全局特征。

2.3 偏移量反投影模块

交叉注意力机制^[19]作用于两个不同特征序列时具有选择特征高效、计算成本低的优势, 本文决定将其运用于偏移量的反投影过程中。与EB-LG相比, 本文提出的偏移量反向投影模块在利用偏移量对原始局部特征进行增强的过程中, 更加关注偏移量中与原始局部特征有关的信息, 从而能够更好地实现偏

移量与原始局部特征之间的信息交互。与简单的加法运算相比,交叉注意力机制能够理解原始局部特征和隐藏特征间的偏移量并提高对其的表达能力。具体实现过程如图3所示。算式如下:

$$\mathbf{Q} = \mathbf{F}_e \mathbf{W}_q, \mathbf{K} = \mathbf{F} \mathbf{W}_k, \mathbf{V} = \mathbf{F} \mathbf{W}_v \quad (8)$$

$$\mathbf{A} = \tilde{h}_i \left(\text{softmax} \left(\frac{\mathbf{Q} \mathbf{K}^T}{\sqrt{d_k}} \right) \right) \quad (9)$$

$$\mathbf{Z} = \mathbf{A} \mathbf{V} \quad (10)$$

式(8)~(10)中 $\mathbf{W}_q, \mathbf{W}_k, \mathbf{W}_v \in \mathbb{R}^{N \times d_k}$ 为可学习参数矩阵,分别用于生成代表输入特征的查询 $\mathbf{Q}$,作为参考信息的键 $\mathbf{K}$ 和包含实际聚合信息的值 $\mathbf{V}$, $\mathbf{A}$ 是由 $\mathbf{Q}$ 和 $\mathbf{K}$ 计算得到的相关性权重, $\tilde{h}_i$ 代表dropout操作^[20], $\mathbf{Z}$ 代表经过增强的原始局部特征。

2.4 嵌入形式

2.4.1 并行嵌入

如图5(a)所示,将4个AHU模块依次放置于4个局部特征提取器后。AHU模块的并行嵌入方式并未改变特征网络在原网络中的传播路径,而是针对每个局部特征提取器的计算结果进行特征增强。并行嵌入的方式更适合缺少上采样过程的网络,如图卷积神经网络DGCNN。

2.4.2 级联嵌入

对于包含上采样过程的网络,本文引入另一种嵌入方式,即级联嵌入,如图5(b)所示。具体来说,由AHU模块学习到的经过增强的局部特征被直接作为下一个卷积层的输入,以避免由上采样和插值导致的不连续特征问题。

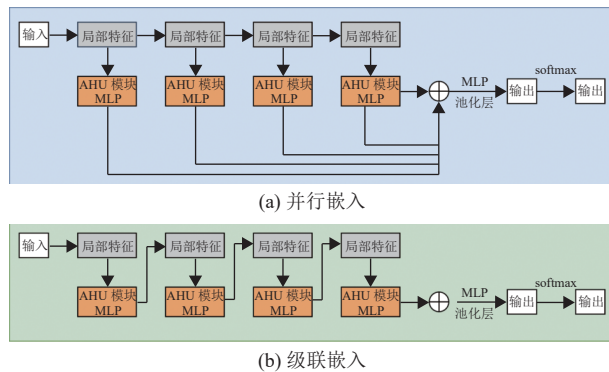


图5 AHU模块的嵌入方式

Fig.5 Embedding modes of the AHU module

2.4.3 分类任务设置

对于图卷积神经网络DGCNN^[11]、自适应图卷积网络AdaptConv^[12]、点自注意力网络PointTransformer^[16]以及点云多层感知机网络PointMLP^[21],选择并

行嵌入模式。在相应的4个边缘卷积层、自适应图卷积层、点网络块以及多层感知机层后分别放置4个AHU模块。对于PointNet++,通过级联嵌入的方式,在两个点集抽取层后嵌入两个AHU模块。

2.4.4 分割任务设置

将3个AHU模块按照并行嵌入模式,放置在DGCNN分割网络的3个边缘卷积层后,其余框架结构与DGCNN保持一致。下一步,在AdaptConv和PointNet++分割网络中应用级联嵌入模式。具体来说,对于AdaptConv,分别在4个自适应图卷积层和1个图卷积层后使用共5个AHU模块,对于PointNet++,在2个点集抽取层后使用2个AHU模块。

3 实验评估

本文将AHU模块插入到现有的3个经典点云网络架构(PointNet++, DGCNN和AdaptConv)中,并进行点云分类和分割实验,以此展示模块的性能。

3.1 实验设置

数据集:选择在3个数据集上完成相关实验,其中ModelNet40^[22]和ScanObjectNN^[23]用于点云分类实验,ShapeNet^[24]用于部件分割实验。

ModelNet40是三维图像分类数据集,其中包含了来自40个类别的12 332个手工绘制点云图像(9 843个用于训练,2 468个用于测试)。

ScanObjectNN是真实3D点云分类数据集,由于受背景、部分缺失以及变形等因素影响,该数据集极具挑战性。其中包含来自15个类别的2 902个三维物体,均为从真实世界室内扫描提取的对象。

ShapeNet是由三维CAD物体模型表示的、具有丰富注释的大型形状库。它总共包含16个父类别,其下还包含数量庞大的子类别,能够完成部件分割任务。

相关设置细节:本文参考文献[5, 10-12, 21]的实验设置,将所有层的近邻点数量 k 设置为20。其次,多尺度的特征通道共同设置为64, 64, 128, 256,并用一个共享的全连接层进行拼接,用于最后的分类任务。使用随机梯度下降优化器,初始学习率为0.1,初始动量设置为0.9,并利用余弦退火^[25]逐渐降至0.001。训练和测试批大小分别设为32和16。以上实验均在PyTorch以及两个NVIDIA RTX 3090 GPU进行。

3.2 点云分类任务

表1展示在ModelNet40数据集上的点云分类性能。在这项实验中,用于评估分类结果的指标是OA(整体准确率)和mAcc(类平均准确率)。如表1所

示,本文提出的模块使PointNet++得到2.8个百分点的整体准确率提升(优于EB-LG的1.9个百分点),DGCNN以及AdaptConv则分别得到1.1个百分点以及0.5个百分点的整体准确率优化,对于PointTransformer,整体准确率能优化至94.1%,意味着本文提出的模块在4个基线网络上效果均优于EB-LG及其基线网络。

表 1 ModelNet40数据集分类结果
Table 1 Classification results on the ModelNet40 benchmark %

模型	mAcc	OA
PointNet ^[4]	86.0	89.2
PointNeXt ^[26]	91.1	90.0
PointConT ^[27]		93.5
PCP-MAE ^[28]		94.2
PointNet++ ^[5]		90.7
PointNet++ + EB-LG ^[10]	90.7	92.6
PointNet++ + AHU	91.5	93.5
DGCNN ^[11]	90.2	92.9
DGCNN + EB-LG ^[10]	90.8	93.2
DGCNN+ AHU	91.4	94.0
AdaptConv ^[12]	90.7	93.4
AdaptConv + EB-LG ^[10]	91.2	93.5
AdaptConv + AHU	91.7	93.9
PointTransformer ^[16]	90.6	93.7
PointTransformer + AHU	90.9	94.1

为进一步验证模块的泛用性,在ScanObjectNN数据集上进行检验.选择与基准点云网络模型相同的输入条件(1 024个点及其相应的3D坐标),使用统一的训练策略.结果如表2所示,集成AHU模块之后,

表 2 ScanObjectNN数据集分类结果
Table 2 Classification results on the ScanObjectNN benchmark %

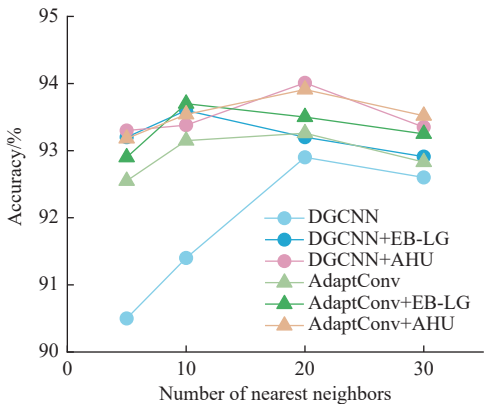
模型	mAcc	OA
APP-Net ^[29]		84.1
Point-LGMask ^[30]		85.3
PointConT ^[27]	86.0	86.0
PointNet++ ^[5]	75.4	77.9
PointNet++ + EB-LG ^[10]	77.1	80.4
PointNet++ + AHU	77.9	81.6
DGCNN ^[11]	73.6	78.1
DGCNN + EB-LG ^[10]	78.6	82.0
DGCNN + AHU	80.5	83.4
AdaptConv ^[12]	76.0	79.3
AdaptConv + EB-LG ^[10]	78.0	81.0
AdaptConv + AHU	78.9	82.0
PointMLP ^[21]	83.9	85.4
PointMLP + AHU	84.7	86.3

基准网络模型得到不同程度的改进.具体来说:在类平均准确率方面,PointNet++取得了2.5个百分点的提升,DGCNN和AdaptConv分别提升6.9和2.9个百分点;在整体准确率方面,PointNet++,DGCNN,AdaptConv以及PointMLP的优化程度分别为3.7,5.3,2.7以及0.9个百分点.以上实验结果表明,AHU模块在点云分类方面对基于局部卷积核的网络模型有着显著的提升,相较于首先提出隐藏特征理论的EB-LG模块,AHU模块在全局信息获取以及偏移影射方面表现更加优秀,同时能够将原有基线网络性能提升至现有先进网络水平,更能充分体现隐藏特征在点云分类性能方面的应用潜力。

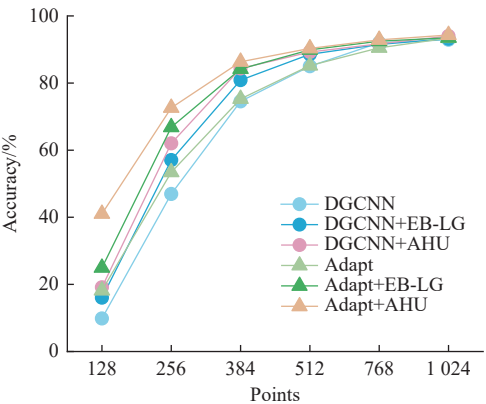
3.3 鲁棒性研究

3.3.1 不同近邻点k下的表现

如图6(a)和图6(b)所示,实验表明较小取值的k对于基准网络模型DGCNN和AdaptConv有较大的干扰.相较于EB-LG,集成AHU模块后,基准网络能够在一定程度上表现得更加平滑和稳定。



(a) 不同近邻点时的分类准确度



(b) 不同输入点时的分类准确度

图 6 在ModelNet40分类任务的鲁棒性实验
Fig.6 Robustness test on the ModelNet40

需要注意的是,当AHU模型使用20个而不是10个近邻点进行训练时,网络模型达到最高的精度,

这是因为在一定近邻点取值范围内,随着近邻点数量的增多,通道注意力描述符强调显著特征通道,同时抑制非重要特征通道的优势更加明显。而当近邻点超过一定范围后,模块会将更多与中心点关联性较弱的近邻点模块考虑在内,意味着噪声增加,致使性能下降。

3.3.2 输入点云数量的鲁棒性研究

在两个基准网络(DGCNN及AdaptConv)上进行点云密度的鲁棒性测试:给定1 024个点进行训练,选取不同数量点云进行验证。如图6(b)所示,集成AHU模块后,基准点云网络均表现出更强的抗干扰能力。即使丢弃87.5%的点(即使用128个点进行测试),AHU在AdaptConv网络模型上的表现仍比EB-LG模块多16个百分点的整体精度值提升。

3.3.3 模型复杂度比较

为了评估AHU模块效率,本文在表3和表4中分别展示单样本情况下的模型参数量及每秒浮点运算量。经分析认为模块的参数以及浮点运算数主要集中在基于交叉注意力机制的反向投影子模块。

表 3 单样本下参数量比较

Table 3 Comparison of the number of parameters for one sample

M

模型	基准模型	+ EB-LG	+ AHU
PointNet++ ^[5]	1.48	1.72	1.74
DGCNN ^[11]	1.81	2.08	2.12
AdaptConv ^[12]	1.85	2.21	2.25

通过以级联方式嵌入2个AHU模块,PointNet++网络的参数量和浮点运算数分别增长0.26 M和73 M,相较于PointNet++,二者指标分别增加了17.5%和8.9%;以并行嵌入方式集成4个AHU模块,DGCNN网络的参数量增加了17%,即0.31 M,浮点运算数增加

14.3%,即359 M;而AdaptConv网络在以上2个指标的增量分别为0.4 M以及347 M,分别增长21.6%和9.9%。总的来说,AHU模块相较于EB-LG基准模块,在以少量运算增加为代价的情况下,对现有经典点云网络实现了明显的性能提升。

表 4 单样本下浮点运算数比较

Table 4 Comparison of the number of FLOPs

M

模型	基准模型	+ EB-LG	+ AHU
PointNet++ ^[5]	863	915	940
DGCNN ^[11]	2 495	2 781	2 854
AdaptConv ^[12]	3 493	3 780	3 840

3.4 点云分割任务

为进一步验证模块在经典点云网络架构中的优势,本文在ShapeNet部件分割数据集进行实验。在训练设置上,遵循DGCNN及ShapeNet的配置。

在网络架构方面,选择与EB-LG相同的集成方式,在不同的尺度上嵌入AHU模块,并在最后的全连接层中进行拼接。仅将2 048个点作为输入,并用mIoU(平均交并比)衡量分割效果,是所有测试形状的平均值。

表5和图7展示两种模块分别集成于DGCNN网络模型时,在点云部件分割任务上的性能表现。图7所展示的结果中,第1行是真实值,第2和第3行分别为DGCNN+EB-LG和DGCNN + AHU的分割效果,其中蓝色点为正确分类,红色点为错误分类。在AHU模块作用下,PointNet++以及DGCNN网络的mIoU指标分别提升0.8以及0.5个百分点。从结果来看,在PointNet++基准模型上,AHU模块相较于EB-LG模块在11个类别上表现更优秀,(在DGCNN和AdaptConv模型上为10个)。

表 5 ShapeNet数据集部件分割结果

Table 5 Part segmentation results on the ShapeNet part dataset

%

模型	mIoU	Air	Bag	Cap	Car	Chair	Ear	Gui	Kni	Lam	Lap	Mot	Mug	Pis	Roc	Ska	Tab
PointNet++ ^[5]	85.1	82.4	79.0	87.7	77.3	90.8	71.8	91.0	85.9	83.7	95.3	71.6	94.1	81.3	58.7	76.4	82.6
PointNet++ + EB-LG ^[10]	85.6	84.4	82.2	87.0	79.0	91.1	74.4	91.6	88.2	84.3	95.8	67.6	94.4	81.9	59.1	74.1	82.6
PointNet++ + AHU	85.9	84.7	82.5	87.4	79.4	90.9	76.3	91.4	88.6	84.1	95.5	73.2	94.3	83.7	61.2	76.8	83.1
DGCNN ^[11]	85.2	84.0	83.4	86.7	77.8	90.6	74.7	91.2	87.5	82.8	95.7	66.3	94.9	81.1	63.5	74.5	82.6
DGCNN + EB-LG ^[10]	85.4	83.8	82.5	87.1	78.3	90.8	77.6	91.9	87.9	84.0	95.9	68.0	94.3	81.8	62.4	74.7	82.5
DGCNN + AHU	85.7	84.6	83.7	87.3	78.0	91.2	79.2	91.5	87.6	84.5	93.3	70.4	94.7	82.5	62.1	73.9	82.9
AdaptConv ^[12]	86.4	84.8	81.2	85.7	79.7	91.2	80.9	91.9	88.6	84.8	96.2	70.7	94.9	82.3	61.0	75.9	84.2
AdaptConv + EB-LG ^[10]	86.1	84.8	81.7	88.2	80.1	91.4	81.2	91.7	88.6	84.8	96.0	74.7	94.9	82.7	61.7	75.9	83.0
AdaptConv + AHU	86.3	85.1	82.2	88.4	80.4	91.4	81.0	91.6	89.4	85.1	89.8	76.8	94.7	83.2	63.4	76.2	83.9

AHU模块加强了基线网络对点云的全局特征关注,鼓励网络从整个点云的角度查找出所有相似特

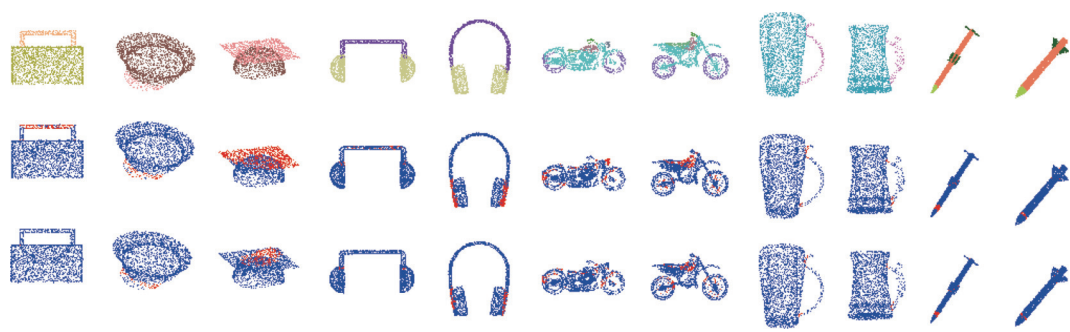


图 7 分割任务结果比较
Fig.7 Comparison of the segmentation results

征的点,然后根据所有查询到的点,标记相应的点。通过这种方式,AHU模块通常能够在3种基线网络的基础上实现更合理的分割。

值得一提的是,由于Tab类别的几何结构以及标注的特殊性,EB-LG模块和AHU模块在该类别表现欠佳,而得益于通道注意力机制,AHU模块能够基于特征通道亲和度强调重要特征通道,弥补了EB-LG模块在该项分割任务上的不足。但是对于Lap类别,不同特征部位的标记存在部分相同的情况(例如键盘与外壳),在AHU模块强调全局特征的情况下,可能造成错误标记,导致性能下降。

3.5 消融研究

本节将通过三维点云进行消融研究,以证明模块方法的有效性。以DGCNN网络为基准模型,在取 $k=20$ 的条件下进行点云分类任务,相关结果记录在表6中。“DGCNN+隐藏特征模块”表示本文只保留隐藏特征计算模块,并以逐点相加的方式替代偏移量反投影模块。“DGCNN+反投影模块”表示仅保留偏移量反投影模块,使用一组多层感知机代替隐藏特征计算模块。

表 6 消融研究结果 Table 6 The results of ablation experiments			%
方法	mAcc	OA	
DGCNN	90.2	92.9	
DGCNN +隐藏特征模块	90.8	93.2	
DGCNN+反投影模块	90.0	92.8	
DGCNN+AHU模块	91.4	94.0	

从表6可看出,DGCNN基准网络的整体准确率为92.9%,平均准确率为90.2%。仅使用隐藏特征模块,整体准确率和平均准确率分别为93.2%和90.8%,相较基准网络有略微提高。仅使用偏移量反投影模块,整体准确率和平均准确率为92.8%和90.0%,与基

准网络性能基本持平。只有联合使用隐藏特征模块和偏移量反投影模块时,DGCNN才能分别提高1.1个百分点的整体准确率和1.2个百分点的平均准确率。

4 结论

本文提出了基于注意力机制的隐藏特征利用(AHU)模块。该模块首先在建模通道之间相互依赖性的基础上获得隐藏特征,然后计算隐藏特征与原始局部特征间的偏移量,通过交叉注意力机制映射至原始局部特征,达到增强局部特征的目的。实验表明,AHU模块能够以足够便捷的方式嵌入现有点云经典网络中,并带来可观的性能提升,在多种现有点云先进网络的分类和分割任务中均取得较大提升。在未来,将继续挖掘隐藏特征的利用价值,并探索AHU模块对其他深度学习架构和应用程序的影响,并探索隐藏特征在点云语义分割以及目标检测任务上的作用。

参考文献:

[1] WANG S, SUO S, MA W C, *et al.* Deep parametric continuous convolutional neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 2589-2597.

[2] DUAN H, WANG P, HUANG Y, *et al.* Robotics dexterous grasping: the methods based on point cloud and deep learning[J]. *Frontiers in Neurorobotics*, 2021, 15: 658280.

[3] TUZEL O, LIU M Y, TAGUCHI Y, *et al.* Learning to rank 3D features[C]//Computer Vision—ECCV 2014: 13th European Conference. Zurich, Switzerland: Springer, 2014: 520-535.

[4] QI C R, SU H, MO K, *et al.* Pointnet: deep learning on point sets for 3D classification and segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017: 652-660.

[5] QI C R, YI L, SU H, *et al.* Pointnet++: deep hierarchical fea-

- ture learning on point sets in a metric space[J]. *Advances in Neural Information Processing Systems*, 2017, 30: 5099-5108.
- [6] GUO Y, WANG H, HU Q, *et al.* Deep learning for 3D point clouds: a survey[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 43(12): 4338-4364.
- [7] ZHAO H, JIANG L, FU C W, *et al.* Pointweb: enhancing local neighborhood features for point cloud processing[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach, USA: IEEE, 2019: 5565-5573.
- [8] ZHANG Y X, SUN Z L, ZENG Z, *et al.* Partial point cloud registration with deep local feature[J]. *IEEE Transactions on Artificial Intelligence*, 2022, 4(5): 1317-1327.
- [9] HUI L, CHENG M, XIE J, *et al.* Efficient 3D point cloud feature learning for large-scale place recognition[J]. *IEEE Transactions on Image Processing*, 2022, 31: 1258-1270.
- [10] CHEN J, ZHANG Y, MA F, *et al.* EB-LG module for 3D point cloud classification and segmentation[J]. *IEEE Robotics and Automation Letters*, 2022, 8(1): 160-167.
- [11] WANG Y, SUN Y, LIU Z, *et al.* Dynamic graph CNN for learning on point clouds[J]. *ACM Transactions on Graphics (tog)*, 2019, 38(5): 1-12.
- [12] NEZHADARYA E, TAGHAVI E, RAZANI R, *et al.* Adaptive hierarchical down-sampling for point cloud classification[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020: 12956-12964.
- [13] VASWANI A, SHAZEER N, PARMAR N, *et al.* Attention is all you need[EB/OL]. arXiv: 1706.03762(2017-06-12) [2024-12-05]. <https://doi.org/10.48550/arXiv.1706.03762>.
- [14] LIU X, HAN Z, LIU Y S, *et al.* Point2sequence: learning the shape representation of 3D point clouds with an attention-based sequence to sequence network[C]// *Proceedings of the AAAI Conference on Artificial Intelligence*. Honolulu, USA: IEEE, 2019: 8778-8785.
- [15] GUO M H, CAI J X, LIU Z N, *et al.* PCT: point cloud transformer[J]. *Computational Visual Media*, 2021, 7: 187-199.
- [16] ZHAO H, JIANG L, JIA J, *et al.* Point transformer[C]// *Proceedings of the IEEE/CVF International Conference on Computer Vision*. Nashville, USA: IEEE, 2021: 16259-16268.
- [17] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018: 7132-7141.
- [18] NAIR V, HINTON G E. Rectified linear units improve restricted boltzmann machines[C]// *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*. Haifa, Israel: IMLS, 2010: 807-814.
- [19] LIN H, CHENG X, WU X, *et al.* CAT: cross attention in vision transformer[C]// *2022 IEEE International Conference on Multimedia and Expo (ICME)*. Taiwan, China: IEEE, 2022: 1-6.
- [20] ZEHUI L, LIU P, HUANG L, *et al.* Dropattention: a regularization method for fully-connected self-attention networks [EB/OL]. arXiv:1907.11065(2019-07-25) [2024-12-05]. <https://doi.org/10.48550/arXiv.1907.11065>
- [21] MA X, QIN C, YOU H, *et al.* Rethinking network design and local geometry in point cloud: a simple residual MLP framework[EB/OL]. arXiv: 2202.07123(2022-02-15) [2024-12-05] <https://doi.org/10.48550/arXiv.2202.07123>.
- [22] WU Z, SONG S, KHOSLA A, *et al.* 3D shapenets: a deep representation for volumetric shapes[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Boston: IEEE, 2015: 1912-1920.
- [23] UY M A, PHAM Q H, HUA B S, *et al.* Revisiting point cloud classification: a new benchmark dataset and classification model on real-world data[C]// *Proceedings of the IEEE/CVF International Conference on Computer Vision*. Long Beach, USA: IEEE, 2019: 1588-1597.
- [24] CHANG A X, FUNKHOUSER T, GUIBAS L, *et al.* Shapenet: an information-rich 3D model repository[EB/OL]. arXiv: 1512.03012(2015-10-09) [2024-12-05]. <https://doi.org/10.48550/arXiv.1512.03012>.
- [25] LOSHCHILOV I, HUTTER F. SGDR: stochastic gradient descent with warm restarts[EB/OL]. arXiv: 1608.03983 (2016-08-13) [2024-12-05] <https://doi.org/10.48550/arXiv.1608.03983>.
- [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] LIU Y, TIAN B, LYU Y, *et al.* Point cloud classification using content-based transformer via clustering in feature space[J]. *IEEE/CAA Journal of Automatica Sinica*, 2023, 11(1): 231-239.
- [28] ZHANG X, ZHANG S, YAN J. PCP-MAE: learning to predict centers for point masked autoencoders[EB/OL]. arXiv: 2408.08753(2024-08-16) [2024-12-05] <https://doi.org/10.48550/arXiv.2408.08753>.
- [29] WANG X, CHEN S, SU J. App-Net: a hybrid neural network for encrypted mobile traffic classification[C]// *IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. Toronto: IEEE, 2020: 424-429.
- [30] TANG Y, LI X, XU J, *et al.* Point-LGMask: local and global contexts embedding for point cloud pre-training with multi-ratio masking[J]. *IEEE Transactions on Multimedia*, 2023, 26: 8360-8370.

(责任编辑: 王威娜 英文审核: 费伦科)