

朱骏杰, 刘东峰. 用于三维点云分割和分类的高分辨率特征网络[J]. 广东工业大学学报, 2025, 42(3): 44–51. doi: 10.12052/gdutxb.240083.
Zhu Junjie, Liu Dongfeng. High-resolution feature network for 3d point cloud segmentation and classification[J]. Journal of Guangdong University of Technology, 2025, 42(3): 44–51. doi: 10.12052/gdutxb.240083.

用于三维点云分割和分类的高分辨率特征网络

朱骏杰, 刘东峰

(广东工业大学 信息工程学院, 广东 广州 510006)

摘要: 多尺度特征在点云领域的密集预测任务中至关重要。当前三维点云处理技术主要依赖编码器-解码器框架, 通过主干网络提取并融合多尺度特征。然而, 这些方法通常采用延迟融合策略, 导致特征集成不足。为解决这一问题, 本文提出了HRFN3D(High-resolution Feature Network for 3D Point Cloud)模型, 一种专为点云分类和分割任务设计的高分辨率特征网络。HRFN3D通过创新性的关系学习模块, 在早期阶段进行特征融合, 促进低分辨率高语义点与高分辨率低语义点的交互, 使高分辨率点在早期阶段就保留高语义信息, 优化后续特征学习。在后期, 结合不同池化策略生成全局特征向量, 并与原始点特征拼接, 既保留细节, 又增强全局特征的代表性。实验结果显示, HRFN3D在ShapeNetPart数据集上将类平均交并比和实例平均交并比分别提升了2.2个百分点和0.9个百分点, 并获得了最佳实例平均交并比86.3%; 在ModelNet40数据集上, 以4.3 M的参数量实现了91.5%的最高类平均精度。这些结果验证了HRFN3D在多尺度特征处理中的有效性。

关键词: 多尺度特征; 三维点云处理; 高分辨率; 特征融合; 早期阶段

中图分类号: TP391.4; TP181

文献标志码: A

文章编号: 1007-7162(2025)03-0044-08

High-resolution Feature Network for 3D Point Cloud Segmentation and Classification

Zhu Junjie, Liu Dongfeng

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

Abstract: Multi-scale features are critical in dense prediction tasks within the point cloud domain. Existing 3D point cloud processing techniques predominantly rely on encoder-decoder frameworks, which extract and integrate multiscale features via a backbone network. However, these methods often employ delayed fusion strategies, resulting in insufficient feature integration. To address this issue, this paper introduces a novel high-resolution feature network for 3D point cloud, named HRFN3D, specifically for point cloud classification and segmentation tasks. HRFN3D innovatively employs a relational learning module to perform feature fusion at an early stage, facilitating interactions between low-resolution high-semantic points and high-resolution low-semantic points. This early fusion ensures that high-resolution points retain semantic information from the outset, facilitating subsequent feature learning. In the later stage, the global feature vectors are generated by combining different pooling strategies and spliced with the original point features, preserving the details and enhancing the representation of the global features. The experimental results show that HRFN3D improves the Class mean and Instance mean Intersection over Union by 2.2 percentage point and 0.9 percentage point, respectively, and achieves the average class ratio of 86.3%. On the ModelNet40 data set, our proposed method achieves the highest class average accuracy of 91.5% with 4.3M parameters. These results validate the effectiveness of HRFN3D in multi-scale feature processing.

Key words: multi-scale features; 3D point cloud processing; high-resolution; feature fusion; early stage

收稿日期: 2024-06-21 录用日期: 2024-09-26 网络首发日期: 2025-01-03

基金项目: 广东省自然科学基金资助项目(2024A1515012058)

作者简介: 朱骏杰(2000-), 男, 硕士研究生, 主要研究方向为点云处理, E-mail: 2112203082@mail2.gdut.edu.cn

通信作者: 刘东峰(1969-), 男, 教授, 博士, 主要研究方向为虚拟现实技术、机器学习、计算机视觉, E-mail: liudf@gdut.edu.cn

近年来,三维(3D)点云因其在虚拟(增强)现实^[1-2]、自动驾驶^[3-4]、机器人^[5-6]和三维场景理解^[7-9]等各个领域的广泛应用前景而受到越来越多的关注。与二维图像不同,点云由一组三维空间中的点组成,每个点通过其笛卡尔坐标 (x, y, z) 进行特征化,提供了对真实世界三维场景的视点不变和几何精确的表示。在实际应用中,点云的稀疏性、无序性以及语义信息的缺乏构成了研究的难点,使得点云分析在三维计算机视觉中尤为重要。相较于二维图像通过有限像素展示丰富的纹理和外观,点云数据通常体积庞大,描述的场景由数万到数十万个点组成,并在三维欧几里得空间中表现出不规则、非结构化和无序的特性,这使得点云的分类和分割任务更加复杂。

从环境中获得或合成的点云通常包含不同规模的物体,从大型物品如桌子到较小物品如桌子上的笔和杯子。这就要求深度学习模型在点云中提取和区分多尺度特征。因此,有效地提取和融合多尺度特征是点云分割任务成功的关键。在传统的神经网络中,高分辨率特征被认为是包含更准确的空间位置信息的浅层特征,而低分辨率特征则是包含更丰富的语义信息的深层特征。以往的点云分割方法^[10-16]主要采用编码器-解码器框架探索多尺度信息。首先对输入点云进行编码,逐步降采样点特征,然后通过上采样将低分辨率特征逐步解码回原始规模,以生成密集预测。在此过程中,特征交互仅发生在相邻尺度之间,限制了模型对丰富多尺度语义的学习。由于最终的高分辨率表示是从低分辨率特征逐步恢复的,这会导致空间精度的降低,因此传统的编码器-解码器框架在捕获丰富多尺度上下文方面存在不足。

受二维视觉识别和早期特征融合的高分辨率结构^[17-22]的启发,本文提出了HRFN3D(High-resolution Feature Network for 3D Point Cloud)。该方法明确地、并行地维护整个网络的高低分辨率特征,并在早期执行多分辨率特征融合。与传统方法^[10-16]在每个阶段仅使用单一分辨率特征不同,HRFN3D在整个训练过程中维护多个分辨率(从高到低),并促进低分辨率和高分辨率之间的频繁通信。这种设计使得早期深层特征能够与高分辨率特征进行有效集成,从而指导后续的特征学习。实验结果证明,引入高分辨率网络结构(后文称为高分辨率架构)显著提高了网络的学习能力和模型的鲁棒性。HRFN3D在早期阶段有效集成了高级特征,提前进行了多尺度特征融合,在后期,结合不同池化策略生成全局特征向量,并与原始点特征拼接,既保留细节,又增强全局特征的代表性,在整个阶段都保持着深度特征交互,从而显著提

升了模型性能。实验结果显示,在ShapeNetPart数据集上高分辨率架构的引入使类平均交并比和实例平均交并比分别提升了2.2个百分点和0.9个百分点,并获得最佳实例平均交并比86.3%;在ModelNet40数据集上,HRFN3D的类平均精度达到了91.5%,超越了PointMLP模型的最佳水平。

1 高分辨率网络介绍

在基于图像的人类姿态估计研究领域,为了处理不同尺度的深层交互并提取详细的语义特征,文献^[17]首次提出了高分辨率网络(High-Resolution Network, HRNet)的概念,通过高分辨率网络结构在不同层次上保持特征分辨率,从而提升了姿态估计的精度。实验结果证明了高分辨率特征的优势以及对于人体姿态估计的重要性。HRNet的核心特点是其多分支结构,该结构通过并行分支进行重复的多尺度融合,允许不同分辨率的特征从其他分支接收信息,从而生成丰富的高分辨率表示。这种结构不同于传统的特征金字塔网络(Feature Pyramid Networks, FPN)结构(见图1(a)),后者通过逐层降采样和上采样来处理多尺度信息。FPN在每个阶段仅处理单一分辨率特征,并通过降采样和上采样过程逐步恢复高分辨率特征,这导致特征集成的延迟和空间精度的降低。相比之下,HRNet(见图1(b))通过其并行分支结构在整个训练过程中保持多个分辨率的特征,从而实现了高效的特征交互和融合,避免了FPN中的逐层特征丢失和恢复问题。这种方法不仅提升了特征的空间精度,还增强了对细节的捕捉能力,这对生成高质量的空间精度热图至关重要。

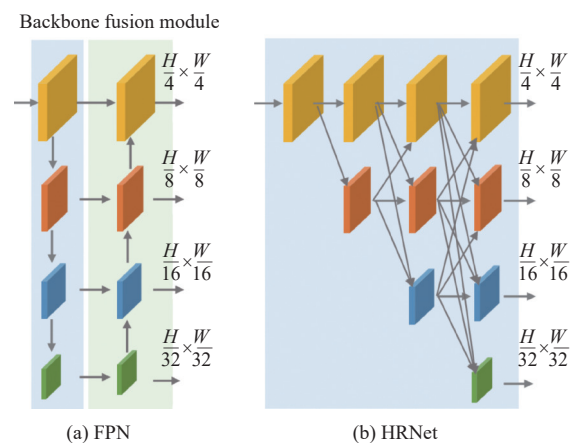


图1 特征金字塔网络和高分辨率网络结构对比图^[18]
Fig.1 Structure comparison between FPN and HRNet^[18]

在此基础上,Wang等^[19]通过维持高分辨率特征并结合多尺度信息有效提高各种视觉识别任务的性能。实验结果表明,HRNet在图像分类、目标检测和

语义分割等任务上都显著地提升了模型性能,证明了其在捕捉细粒度特征和增强特征表达能力方面的有效性。HRNet的成功展示了在处理复杂视觉数据时高分辨率特征的重要性和优势。

在后续发展中,HRFormer^[20]通过结合高分辨率特征和创新的注意力机制^[21],取代了传统高分辨率框架中的卷积层,能够在不同尺度上有效地捕捉和融合信息,从而提升了密集预测的精度。实验结果表明,HRFormer在语义分割、目标检测等任务中具有优异的性能,相比于传统方法,显著提高了预测精度和特征表达能力。HRFormer的成功验证了HRNet和注意力机制的结合在密集预测任务中的有效性和潜力。此外,Zhang等^[22]实现了HRNet用于人物再识别任务,通过引入高分辨率网络成功应对了不同分辨率输入图像所带来的挑战。最近,CEDNet^[21]对高分辨率架构进行了改进,特别是在多尺度融合方面,解决了以往方法中多尺度特征融合延迟和不足的问题。CEDNet能够在早期阶段整合高级特征,从而指导后续阶段的低级特征学习,显著提高了多尺度特征融合的有效性。

高分辨率网络已经被证明了在二维领域中发展的巨大潜力,并在提高视觉任务性能中发挥了关键作用。尽管高分辨率网络研究目前主要集中在二维图像领域,其在三维点云处理中的应用仍属前沿研究。由于三维点云数据的稀疏性、无序性及缺乏语义信息,传统特征处理方法面临挑战。通过HRFN3D模型将高分辨率架构引入三维点云领域,可以在早期阶段在多个分辨率层次上进行深度特征学习,更准确地捕捉不同尺度的细节,并有效融合多尺度特征。这不仅提高了点云分类和分割任务的准确性和鲁棒

性,还显著提升了模型对复杂场景的适应能力。这项研究推动了高分辨率特征处理技术在三维计算机视觉中的应用,为点云分析提供了新的解决方案,并为相关领域的研究和应用奠定了基础。接下来将详细介绍HRFN3D模型及其在三维点云分析中的优势。

2 本文方法

本节详细介绍用于三维点云处理的高分辨率特征网络。分别阐述网络中各组件的具体实现。

2.1 HRFN3D中的高分辨率网络

首先介绍输入点云的数据组成。给定一个点云 $P = \{[P_i, f_i] | i = 1, \dots, N\} \in \mathbf{R}^{N \times (3+D)}$, 其中 $[P_i] \in \mathbf{R}^{N \times 3}$ 表示点在笛卡尔空间中的坐标, $\{f_i\} \in \mathbf{R}^{N \times D}$ 表示其他信息,如法向量或颜色信息, D 代表信息的维度。点云处理的目标是直接训练神经网络来理解 P 的基本表示。

本文所提出的HRFN3D网络整体模型如图2所示。图的左侧展示了输入的原始点云 $N \times C_{\text{raw}}$, 首先通过多层感知器(Multi-Layer Perceptron, MLP)嵌入至 $N \times C_0$, 其中包含高分辨率分支。通过不同程度的降采样,计算得到其他低分辨率的分支。随后,局部序列算子用于学习不同分辨率的特征,并通过每个分支的向上特征融合模块将这些特征融合以得到多尺度的语义特征。每次融合操作作为一个新阶段集成到框架中,融合后的多尺度语义特征将作为下一阶段的输入,以获得更高分辨率的多尺度语义特征。在HRFN3D中,高分辨率概念对应于点的数量,而多尺度语义表示则体现了不同语义程度特征的融合。

网络中的特征交互流动框架如图3所示。图中, f_p 表示某一个点的特征,其中 $p \in \{0, 1, 2, 3, 4\}$ 表示经过

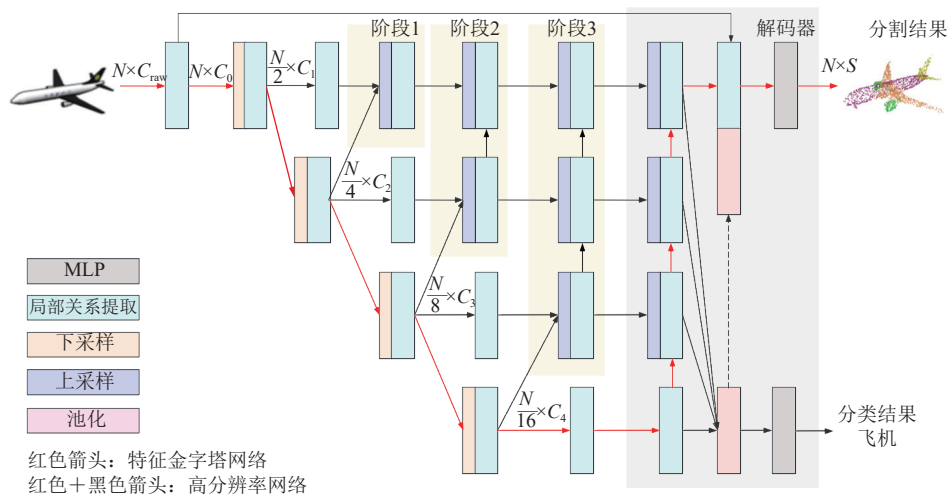


图2 整体模型

Fig.2 The overall framework of the proposed HRFN3D

合模块来实现不同分辨率分支之间的深度特征交互。对于第 j 阶段($j \in \{1, 2, 3, 4\}$)的第 p 分支 f_{pj} , 每个特征上融合模块接受两个输入: 来自同一分支的特征; 来自低一级分支的特征, 即分辨率低于当前分支的特征。低分辨率的分支特征首先经过上采样处理。随后, 这些处理后的特征进行拼接和融合, 生成第 p 分支的更新特征。例如, 在第2阶段计算出第2个分支的融合特征, 见式(3)。

$$f_{22} = \alpha (f_2 \oplus \text{upsample}(f_3)) \quad (3)$$

然后可以得到第2阶段第1分支的融合特征, 见式(4)。

$$f_{12} = \beta (f_{11} \oplus \text{upsample}(f_{22})) \quad (4)$$

其中 \oplus 为特征拼接操作, α 和 β 为用于拼接后的多尺度特征融合的MLP, 各阶段的信息融合都是逐级从低分辨率到高分辨率的融合。

HRFN3D上采样操作采用的是经典的基于距离的插值策略, 在特征传播水平上, 将点特征从 f_p 传播到 f_{p-1} , 通过插入 p_{p-1} 点坐标上的特征值 f_p 来实现特征传播。在许多插值选择中, 本文使用基于 K 个最近邻的逆距离加权平均(见式(5), 不同的 K 取决于上采样密度)。

$$f^{(j)}(x) = \frac{\sum_{p=1}^K \omega_p(x) f_p^{(j)}}{\sum_{p=1}^K \omega_p(x)},$$

$$\text{Where } \omega_p(x) = \frac{1}{d(x, x_p)^2}, j = 1, \dots, C \quad (5)$$

2.4 全局特征处理

HRFN3D在全局特征处理上有其独特之处。首先, 对最低分辨率点的深层特征进行最大池化和平均池化操作, 并将这两个结果相加, 以生成最终的全局特征向量。接着, 将该全局特征向量与原始点特征进行拼接, 并输入至分割头进行最终的分割任务。这种处理方式在保留细节的同时, 使得全局特征向量具有较强的代表性。上述过程可以用式(6)表示。

$$f^{\text{global}} = \text{MaxPool}(f_{44}) + \text{AvgPool}(f_{44}) \quad (6)$$

该设计有效地提取了点云数据中的全局特征, 对后续的点云处理任务至关重要。此过程不仅兼顾了局部信息与全局信息的平衡, 而且优化了特征的表达能力, 从而确保了模型在处理复杂点云数据时的效率和准确性。这一点将在后续消融实验中得到验证。

最终, 将局部关系输出的特征 f'_0 和全局特征 f^{global} 拼接并送入由线性层组成的分割头生成 $N \times S$ 的

预测输出, 其中 S 表示零件分割类别的数量。

3 实验部分

本文在2个广泛使用的数据集即ShapeNetPart和ModelNet40上评估所提出的点云处理模型^[24-33]。首先介绍实验的硬件设备, 然后将逐一介绍不同数据集和相对应的实验设置, 最后, 通过消融实验证明本文提出的高分辨率网络的有效性。

3.1 硬件配置

本文的实验在Windows 10操作系统中部署。编程语言是Python3.9, 深度学习框架为Pytorch2.0.1。使用GPU加速工具CUDA 12.4。硬件配置主要包括: CPU $\times 1$, 型号为Intel(R) Xeon(R) W-2245 CPU @3.90 GHz; GPU $\times 1$, 型号为NVIDIA RTX A2000 12 GB。

3.2 ShapeNetPart上的零件分割结果

在本文中, 选择ShapeNetPart数据集作为点云分割实验数据集。ShapeNetPart是一个广泛使用的3D点云分割数据集, 包含16类对象的约16 881个3D模型, 每个模型都被手动标注为多个零件。这一数据集的复杂性和多样性使其成为3D点云分割任务的标准测试集。本文训练采用了AdamW优化器, 学习率设置为0.001, 训练轮次总计250次。为了确保实验结果的公平性, 输入到点云分割网络中的点云数量设置为2 048。评价指标方面, 本文选用实例平均交并比(Instance mean Intersection over Union, Ins.mIoU)和类平均交并比(Class mean Intersection over Union, Cls.mIoU)。Ins.mIoU用于综合衡量模型在各个实例上的分割性能, 而Cls.mIoU则用于评估模型在每个类别上的分割表现, 这2个指标是评价分割算法精度和鲁棒性的关键指标。

从表1所示的实验结果中可以看出, 本文的方法在ShapeNetPart数据集上的分割结果与当前主流方法的结果相当, 尤其是在实例平均交并比上, 达到了86.3%最佳水平, 表明本文方法在3D点云分割任务中具有优异的性能。此外, 通过实验结果的可视化(见图6), 与引入高分辨率架构前的基准模型PointNet++进行比较, 发现该方法在细节分割的准确度上有显著提升, 如飞机轮胎、摩托车车身及物体部位连接等细节, 说明该模型不仅提高了预测精度, 还增强了对细节的捕捉能力, 验证了所设计方法的有效性。

3.3 ModelNet40上的分类结果

在分类任务训练方面, 选择在ModelNet40分类数据集上进行实验。ModelNet40是一个广泛使用的3D形状分类数据集, 包含来自40个不同类别的12 311

表1 ShapeNetPart上的零件分割结果
Table 1 Segmentation results of different methods on the ShapeNetPart database

方法	Cls.mIoU	Ins.mIoU	mIoU															
			飞机	书包	帽子	车	凳子	耳机	吉他	刀	灯	笔记本	摩托车	杯子	手枪	火箭	滑板	桌子
PointNet ^[10]	80.4	83.7	83.4	78.7	82.5	74.9	89.6	73.0	91.5	85.9	80.8	95.3	65.2	93.0	81.2	57.9	72.8	80.6
PointNet++ ^[11]	81.9	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
PCNN ^[29]	81.8	85.1	82.4	80.1	85.5	79.5	90.8	73.2	91.3	86.0	85.0	95.7	73.2	94.8	83.3	51.0	75.0	81.8
DGCNN ^[22]	82.3	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
PointCNN ^[12]	84.6	86.1	84.1	86.5	86.0	80.8	90.6	79.7	92.3	88.4	85.3	96.1	77.2	95.2	84.2	64.2	80.0	83.0
PointASNL ^[13]	83.4	86.1	84.1	84.7	87.9	79.7	92.2	73.7	91.0	87.2	84.2	95.8	74.4	95.2	81.0	63.0	76.3	83.2
RS-CNN ^[30]	84.0	86.2	83.5	84.8	88.8	79.6	91.2	81.1	91.6	88.4	86.0	96.0	73.7	94.1	83.4	60.5	77.7	83.6
Ours ^[31]	83.6	86.2	85.2	81.4	86.1	80.1	91.5	76.6	91.9	87.6	84.6	97.1	72.9	95.4	84.3	63.7	76.5	83.9
SPLATNet ^[31]	83.7	85.4	83.2	84.3	89.1	80.3	90.7	75.5	92.1	87.1	83.9	96.3	75.6	95.8	83.8	64.0	75.5	81.8
SpiderCNN ^[32]	82.4	85.3	83.5	81.0	87.2	77.5	90.7	76.8	91.1	87.3	83.3	95.8	70.2	93.5	82.7	59.7	75.8	82.8
KPConv ^[25]	85.1	86.2	84.6	86.3	87.2	81.1	91.1	77.8	92.6	88.4	82.7	96.2	78.1	95.8	85.4	69.0	82.0	83.6
PointMLP ^[27]	84.6	86.1	83.5	83.4	87.5	80.5	90.3	78.2	92.2	88.1	82.6	96.2	77.5	95.8	85.4	64.6	83.3	84.3
HRFN3D	84.0	86.3	83.6	85.9	88.5	80.3	91.2	79.5	92.0	87.7	84.5	95.9	73.7	95.4	83.9	63.1	75.4	83.6

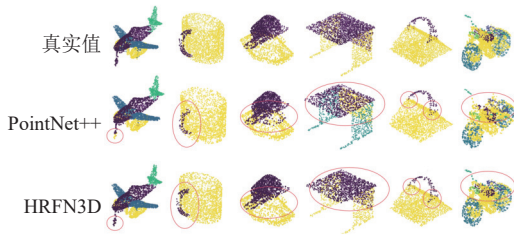


图6 ShapeNetPart数据集上的分割结果可视化
Fig.6 Some visualized segmentation results on the ShapeNetPart database

个3D CAD模型。这一数据集广泛用于评估3D形状分类算法的性能,因其多样性和复杂性而成为分类标准测试集。实验中采用AdamW优化器,初始学习率设置为0.01,并使用余弦衰减策略进行学习率优化。网络共训练250个轮次。为了评估模型的性能,本文选择了总体精度(Overall Accuracy, OA)和类平均精度(mean Accuracy, mAcc)作为评价指标。OA反映了模型在所有测试样本上的整体分类准确性,而mAcc则是对每个类别的分类准确性进行平均,能够更好地衡量模型在各个类别上的平均表现。

实验结果见表2,本文方法在分类实验中表现了优异的性能,尤其是在类平均精度上,达到了91.5%的最佳水平,同时模型仅需4.2 M的参数量。这些结果验证了该方法在3D形状分类任务中的有效性和鲁棒性。

3.4 消融实验

本节进行了全面的消融研究,旨在验证模型中提出的架构(见2.1节)与全局特征处理(见2.4节)以及其他各种参数设置的有效性。

表3所示实验结果证明了将高分辨率架构从二

表2 ModelNet40上的形状分割结果
Table 2 Classification results of different methods on the ModelNet40 dataset

方法	点数	mAcc/%	OA/%	参数量/M
PointNet ^[10]	1 024	86.0	89.2	3.47
PointNet++ ^[11]	1 024	87.7	90.7	2.10
PointNet++ ^[11]	5 000	88.7	91.9	2.10
PointCNN ^[12]	1 024	88.1	92.2	
PointConv ^[14]	1 024		92.5	18.60
Point2Sequence ^[28]	1 024	90.4	92.6	
Point Transformer ^[16]	1 024		92.8	13.50
KPConv ^[25]	7 000		92.9	15.20
DGCNN ^[26]	1 024	90.2	92.9	
PointASNL ^[13]	1 024		92.9	
PointNext ^[15]	1 024	90.6	93.2	
PCT ^[24]	1 024	90.0	93.2	
PointMLP ^[27]	1 024	91.3	94.1	12.60
HRFN3D	1 024	91.5	93.2	4.20

维图像领域引入到三维点云处理任务的有效性。在引入高分辨率架构前,实例交并比Ins.mIoU只能取得85.4%,类交并比Cls.mIoU则更低,为81.8%。引入高分辨率架构后在两个指标分别提升至86.3%和84.0%,参数量仅增加2.9 M。这种性能提升在于高分辨率架构的引入能够使高分辨率点提早进行深度特征交互从而获取丰富的多尺度点云特征,并在接下

表3 高分辨率架构的消融实验
Table 3 Ablation experimental results with a high-resolution architecture

高分辨率架构	Cls.mIoU/%	Ins.mIoU/%	参数量/M
未引入	81.8	85.4	16
引入	84.0	86.3	18.9

来指导网络后续学习,使得最终点云语义信息具备更加丰富的表征。

表4所示实验结果验证了所提出的全局特征处理在点云分割中的有效性。在注重局部信息获取的基础上补充全局信息,能够提高点云分割的精度。此外,还对领域点数量对分割准确率的影响进行了探索。表5显示,当领域点数量为24时,分割指标Cls.mIoU和Ins.mIoU均达到最佳效果,数量过多或过少都会导致这2个分割指标下降。这表明适当的领域点数量能够优化分割性能,强调了全局特征和领域点数量在分割任务中的重要性。

表4 全局特征处理的消融实验

Table 4 Ablation experimental results for global feature processing

分割中全局特征处理	Cls.mIoU/%	Ins.mIoU/%
不拼接	83.8	86.1
拼接	84.0	86.3

表5 局部关系提取中领域点数量的消融实验

Table 5 Ablation experimental results on the number of domain points in local relation extraction

K(领域点)值	Cls.mIoU/%	Ins.mIoU/%
16	83.84	86.19
20	83.95	86.27
24	84.01	86.32
28	83.92	86.18

在该模型的测试过程中,观察到下采样后通道数的设置对模型性能和效率有显著影响。如表6所示,初始使用的通道数设置效果不佳且参数量较大。当改用较少通道数的第2种设置后,模型参数量显著降低,分割效率明显提升。理论上,通道数影响模型的表达能力和计算成本。更多通道通常意味着更高的表达能力,但也会增加计算负担和内存使用。适当减少通道数可以降低计算复杂度,提高执行速度,同时避免过拟合。本文通过合理调整通道数,优化了模型的参数量和计算效率,从而提升了分割性能。这表明,适当的通道数设置能够在保持高效特征提取的同时改善模型表现。

表6 不同特征向量的通道数设置的消融实验

Table 6 Ablation experimental results with channel number settings for different feature vectors

特征向量通道数设置 [f_0, f_1, f_2, f_3, f_4]	Cls.mIoU/%	Ins.mIoU/%	参数量/M
[64, 128, 256, 512, 1024]	83.4	85.6	33.7
[48, 96, 192, 384, 512]	84.0	86.3	18.9

4 结论

本文提出了一种新的点云处理模型HRFN3D。其主要创新在于将图像处理领域中新兴的高分辨率网络思想引入点云领域,并使用了一个高效的关系学习模块来实现高分辨网络。在全局特征处理方面,通过不同池化策略结合生成全局特征向量,并与原始点特征拼接。这种方法既保留了细节,又增强了全局特征的代表性。HRFN3D通过创新的邻域关系学习模块组成的高分辨率网络和高效的全局特征处理方法,实现了对局部和全局特征的有效整合,显著提升了点云分类和分割任务的性能。实验结果表明,HRFN3D可以提早在多个分辨率层次上进行深度特征学习,有效地集成了早期阶段的高级特征,提早进行多尺度特征融合,不仅提高模型的预测精度,还增强了对细节的捕捉能力。本文提出的方法在ShapeNetPart数据集上实验得出类平均交并比(Cls.mIoU)和实例平均交并比(Ins.mIoU)分别提升了2.2个百分点和0.9个百分点,并获得最佳实例平均交并比86.3%,在ModelNet40数据集上获得的最佳类平均精度为91.5%。

尽管HRFN3D在多个指标上表现出色,但在某些实验条件下仍存在改进空间。例如,在处理极端稀疏点云时,模型的性能未能显著提升,这可能与点云数据的复杂性以及高分辨率特征的处理能力有关。未来的研究应关注进一步提高模型在不同数据集和条件下的鲁棒性,并探索处理更大规模数据集的有效策略。这将有助于增强高分辨率网络的通用性和应用范围。

参考文献:

- [1] ALEXIOU E, YANG N, EBRAHIMI T. PointXR: a toolbox for visualization and subjective evaluation of point clouds in virtual reality[C]//2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX). Athlone: IEEE, 2020: 1-6.
- [2] CHEN Y Q, WANG Q, CHEN H, *et al.* An overview of augmented reality technology[J]. Journal of Physics: Conference Series, 2019, 1237(2): 022082.
- [3] AKSOY E E, BACI S, CAVDAR S. Salsanet: fast road and vehicle segmentation in lidar point clouds for autonomous driving[C]//2020 IEEE Intelligent Vehicles Symposium (IV). Las Vegas: IEEE, 2020: 926-932.
- [4] LI Y, MA L, ZHONG Z, *et al.* Deep learning for lidar point clouds in autonomous driving: a review[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32(8): 3412-3432.
- [5] LI X, DU S, LI G, *et al.* Integrate point-cloud segmentation with 3D LiDAR scan-matching for mobile robot localiza-

- tion and mapping[J]. *Sensors*, 2019, 20(1): 237.
- [6] YANG L, LIU Y, PENG J, *et al.* A novel system for off-line 3D seam extraction and path planning based on point cloud segmentation for arc welding robot[J]. *Robotics and Computer-Integrated Manufacturing*, 2020, 64: 101929.
- [7] 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.
- [8] LUO C, YANG X, YUILLE A. Exploring simple 3D multi-object tracking for autonomous driving[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Montreal: IEEE, 2021: 10488-10497.
- [9] ZHAO L, CAI D, SHENG L, *et al.* 3DVG-Transformer: relation modeling for visual grounding on point clouds[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Montreal: IEEE, 2021: 2928-2937.
- [10] QI C, 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 (CVPR). Honolulu: IEEE, 2017: 652-660.
- [11] QI C, SU H, MO K, *et al.* PointNet++: deep hierarchical feature learning on point sets in a metric space[J]. *Advances in Neural Information Processing Systems*, 2017, 30: 5105-5114.
- [12] LI Y, BU R, SUN M, *et al.* PointCNN: convolution on x-transformed points[J]. *Advances in Neural Information Processing Systems*, 2018, 31: 828-838.
- [13] YAN X, ZHENG C, LI Z, *et al.* PointASNL: robust point clouds processing using nonlocal neural networks with adaptive sampling[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020: 5589-5598.
- [14] WU W, QI Z, FUXIN L. PointConv: deep convolutional networks on 3D point clouds[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE, 2019: 9621-9630.
- [15] 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.
- [16] ZHAO H, JIANG L, JIA J, *et al.* Point transformer[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Montreal: IEEE, 2021: 16259-16268.
- [17] SUN K, XIAO B, LIU D, *et al.* Deep high-resolution representation learning for human pose estimation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE, 2019: 5693-5703.
- [18] ZHANG G, LI Z, TANG C, *et al.* CEDNet: a cascade encoder-decoder network for dense prediction[EB/OL]. arxiv: 2302.06052(2023-02-06) [2023-10-31]. <https://arxiv.org/abs/2302.06052>.
- [19] WANG J, SUN K, CHENG T, *et al.* Deep high-resolution representation learning for visual recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 43(10): 3349-3364.
- [20] YUAN Y, FU R, HUANG L, *et al.* HRformer: high-resolution transformer for dense prediction[EB/OL]. arxiv: 2110.09408(2021-10-09) [2021-11-07]. <https://arxiv.org/abs/2110.09408>.
- [21] VASWANI A, SHAZEER N, PARMAR N, *et al.* Attention is all you need[J]. *Advances in Neural Information Processing Systems*, 2017, 30: 6000-6010.
- [22] ZHANG G, GE Y, DONG Z, *et al.* Deep high-resolution representation learning for cross-resolution person re-identification[J]. *IEEE Transactions on Image Processing*, 2021, 30: 8913-8925.
- [23] WU X, LAO Y, JIANG L, *et al.* Point transformer v2: grouped vector attention and partition-based pooling[J]. *Advances in Neural Information Processing Systems*, 2022, 35: 33330-33342.
- [24] GUO M, CAI J, LIU Z, *et al.* PCT: point cloud transformer[J]. *Computational Visual Media*, 2021, 7: 187-199.
- [25] THOMAS H, QI C, DESCHAUD J, *et al.* KPConv: flexible and deformable convolution for point clouds[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Seoul: IEEE, 2019: 6411-6420.
- [26] 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.
- [27] 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(2022-02-07) [2022-11-29]. <https://arxiv.org/abs/2202.07123v2>.
- [28] LIU X, HAN Z, LIU Y, *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 (AAAI). Hawaii: AAAI, 2019, 33(01): 8778-8785.
- [29] ATZMON M, MARON H, LIPMANY. Point convolutional neural networks by extension operators[EB/OL]. arxiv: 1803.10091(2018-03-10) [2018-03-27]. <https://arxiv.org/abs/1803.10091>.
- [30] LIU Y, FAN B, XIANG S, *et al.* Relation-shape convolutional neural network for point cloud analysis[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE, 2019: 8895-8904.
- [31] SU H, JAMPANI V, SUN D, *et al.* SplatNet: sparse lattice networks for point cloud processing[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City: IEEE, 2018: 2530-2539.
- [32] XU Y, FAN T, XU M, *et al.* SpiderCNN: deep learning on point sets with parameterized convolutional filters[C]//Proceedings of the European Conference on Computer Vision (ECCV). Munich: CVF, 2018: 87-102.
- [33] LI Z, GAO P, YUAN H, *et al.* Exploiting inductive bias in transformer for point cloud classification and segmentation [C]//2023 IEEE International Conference on Multimedia and Expo Workshops (ICMEW). Brisbane: IEEE, 2023: 140-145.

(责任编辑: 王威娜 英文审核: 费伦科)