

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

融合多分辨率特征的点云分类与分割网络

陶志勇, 李衡, 豆森森, 林森

引用本文:

陶志勇, 李衡, 豆森森, 等. 融合多分辨率特征的点云分类与分割网络[J]. *光电工程*, 2023, **50**(10): 230166.

Tao Z Y, Li H, Dou M S, et al. Multi-resolution feature fusion for point cloud classification and segmentation network[J]. *Opto-Electron Eng*, 2023, **50**(10): 230166.

<https://doi.org/10.12086/oe.2023.230166>

收稿日期: 2023-07-07; 修改日期: 2023-09-13; 录用日期: 2023-09-20

相关论文

基于图像信息约束的三维激光点云聚类方法

夏金泽, 孙浩铭, 胡盛辉, 梁冬泰

光电工程 2023, **50**(2): 220148 doi: [10.12086/oe.2023.220148](https://doi.org/10.12086/oe.2023.220148)

融合空间掩膜预测与点云投影的多目标跟踪

陆康亮, 薛俊, 陶重彝

光电工程 2022, **49**(9): 220024 doi: [10.12086/oe.2022.220024](https://doi.org/10.12086/oe.2022.220024)

模拟水体湍流环境下目标激光点云数据的三维重建与分析

王明军, 彭月, 刘燕荣, 李勇俊, 李乐

光电工程 2023, **50**(6): 230004 doi: [10.12086/oe.2023.230004](https://doi.org/10.12086/oe.2023.230004)

更多相关论文见光电期刊集群网站 

 **光电工程**
Opto-Electronic Engineering

<http://cn.ojournal.org/oe>



 OE_Journal



Website

DOI: 10.12086/oe.2023.230166

融合多分辨率特征的点云分类与分割网络

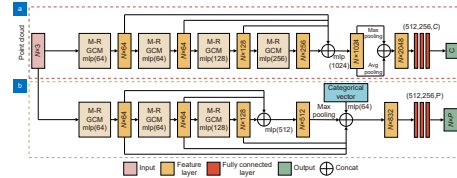
陶志勇¹, 李衡^{1*}, 豆淼森¹, 林森²¹ 辽宁工程技术大学电子与信息工程学院, 辽宁 葫芦岛 125100;² 沈阳理工大学自动化与电气工程学院, 辽宁 沈阳 110159

摘要: 针对现有网络难以有效学习点云局部几何信息的问题, 提出一种融合点云多分辨率特征的图卷积网络。首先, 通过k-最近邻算法对点云构建局部图结构, 以更好地表示点云的局部几何结构。其次, 基于最远点采样算法提出一个并行通道分支, 该分支通过对点云进行下采样来获得不同分辨率的点云, 然后对其进行分组处理; 为克服点云的稀疏特性, 提出一种几何映射模块对分组点云执行正态化操作。最后, 提出一种特征融合模块对图特征和多分辨率特征进行聚合, 以更有效地获得全局特征。实验使用 ModelNet40、ScanObjectNN 和 ShapeNet Part 数据集进行评估, 结果表明, 提出的网络具有良好的分类与分割性能。

关键词: 点云; 图卷积网络; 多分辨率点云; 特征融合模块

中图分类号: TP391.41

文献标志码: A



陶志勇, 李衡, 豆淼森, 等. 融合多分辨率特征的点云分类与分割网络[J]. 光电工程, 2023, 50(10): 230166

Tao Z Y, Li H, Dou M S, et al. Multi-resolution feature fusion for point cloud classification and segmentation network[J]. *Opto-Electron Eng*, 2023, 50(10): 230166

Multi-resolution feature fusion for point cloud classification and segmentation network

Tao Zhiyong¹, Li Heng^{1*}, Dou Miaosen¹, Lin Sen²¹ School of Electronic and Information Engineering, Liaoning Technical University, Huludao, Liaoning 125100, China;² School of Automation and Electrical Engineering, Shenyang Ligong University, Shenyang, Liaoning 110159, China

Abstract: To address the problem that existing networks find it difficult to learn local geometric information of point cloud effectively, a graph convolutional network that fuses multi-resolution features of point cloud is proposed. First, the local graph structure of the point cloud is constructed by the k-nearest neighbor algorithm to better represent the local geometric structure of the point cloud. Second, a parallel channel branch is proposed based on the farthest point sampling algorithm, which obtains point clouds with different resolutions by downsampling them and then groups them. To overcome the sparse characteristics of the point cloud, a geometric mapping module is proposed to perform normalization operations on the grouped point cloud. Finally, a feature fusion module is proposed to aggregate graph features and multi-resolution features to obtain global features more effectively. Experiments are evaluated using ModelNet40, ScanObjectNN, and ShapeNet Part datasets. The experimental results show that the proposed network has state-of-the-art classification and segmentation performance.

Keywords: point cloud; graph convolutional network; multi-resolution point cloud; feature fusion module

收稿日期: 2023-07-07; 修回日期: 2023-09-13; 录用日期: 2023-09-20

基金项目: 辽宁省科技厅应用基础研究项目 (2022JH2/101300274)

*通信作者: 李衡, PaperLH@163.com。

版权所有©2023 中国科学院光电技术研究所

1 引言

随着激光雷达等 3D 扫描技术的快速发展^[1-2], 点云的获取变得容易, 近年来, 点云分析已经成为 3D 视觉任务的热门话题, 在自动驾驶^[3-4]、医学影像^[5]、机器人^[6]、三维重建^[7]等领域得到了广泛应用。尽管深度学习在点云处理方面取得了不错的成果, 但是由于点云无序且不规则的特性, 有效处理点云的数据仍然具有挑战性。

将神经网络适用于点云数据并非易事。早期的点云处理主要分为基于体素的方法和基于多视图的方法。前者将点云转换为体素, 通过规则的 3D 网格结构表征点云。如: Maturana 等人^[8]提出了 VoxNet, 通过体素化点云以实现 3D 形状的识别任务; Riegler 等人^[9]提出了 OctNet, 对八叉树结构进行有效编码, 并对每个体素的特征向量进行简单的算法索引; 文献^[10]针对传统体素网格降采样存在采样点不均匀的问题, 提出了一种新的体素网格降采样方式, 提高了计算效率。但随着分辨率的增加, 这类方法的计算量和内存占用量会以立方形式增长。基于多视图的方法是将点云从不同角度投影至多个平面, 使其适用于 2D 卷积。如: Su 等人^[11]提出了 MVCNN, 将多视图特征最大池化为一个全局描述符, 但是最大池化操作只保留了特定视图中的最大元素, 不可避免地导致信息缺失。为了不引入中间表示(体素和多视图)来直接处理点云数据, 基于点的方法得到了发展。Qi 等人^[12]提出的 PointNet 是该类方法的先驱, 通过几个多层感知机 (MLP) 独立地学习逐点特征, 克服了点云的无序性, 但同时却忽略了点对之间的局部关系。为此, Qi 等人^[13]扩展了 PointNet, 提出新的网络 PointNet++, 通过在不同的层次上使用 PointNet 提取特征, 但是本质上还是在独立地处理局部区域的每一个点, 点对之间的几何联系仍然被忽略。点云作为一种不规则的非欧式数据, 类图结构能对其进行有效地表征。Simonovsky 等人^[14]是点云图论的先驱, 将每一个点视为图的顶点, 通过有向边将顶点和其邻域点进行连接, 然后基于滤波器和 MLP 提出了边缘条件卷积 (edge-conditioned convolution, ECC)。Wang 等人^[15]在 DGCNN 中提出一种边缘卷积算子 (edge convolution, EdgeConv), 通过在特征空间中构造局部图结构表示点云的局部联系, 并在网络的每一层之后动态更新图特征。Zhang 等人^[16]基于 DGCNN 提出了 LDGCNN, 删去 DGCNN 中的变换网络并将不同层次的特征进行

连接, 以提高网络的性能。文献^[17]通过引入位置关系改进 DGCNN, 强化局部特征的提取能力。此外, 一些方法通过引入注意力机制来提升网络性能。如: Sun 等人^[18]提出了基于双向注意力机制的残差图卷积网络, 以更好地区分任务相关特征; Gou 等人^[19]基于 transformer 提出了名为 PCT 的网络执行点云分析任务; Liu 等人^[20]将 SENet 和注意力机制引入 PointNet++, 优化了网络对于重要特征的学习能力。

为了能更有效地提取点云局部特征, 提高网络分类与分割性能, 提出一种融合多分辨率特征的图卷积网络。首先考虑到点云的无序不规则特性, 通过 k-最近邻算法 (k-nearest neighbor, kNN) 在特征空间中构造局部图结构, 通过图结构表征点云局部区域中点对之间的几何联系。其次, 通过最远点采样算法 (farthest point sampling, FPS) 对点云进行多级下采样操作, 以获得不同分辨率的点云, 考虑其局部区域的几何关系, 采用 kNN 算法进行分组; 为克服点云的稀疏性质, 引入几何映射模块将分组点集正态化。最后, 为了实现良好的分类与分割性能, 提出一种特征融合方式, 用点云多分辨率特征对局部图特征进行补偿。通过在 ModelNet40^[21]、ScanObjectNN^[22] 和 ShapeNet Part^[23] 数据集上对网络的分类、分割性能进行评估, 验证本文方法的有效性。所提出的多分辨率图卷积模块 (multi-resolution graph convolution module, M-R GCM) 算法流程如图 1 所示。

2 原理与方法

2.1 网络整体框架

网络框架如图 2 所示, 输入是点数为 N , 维度为 3 的点云数据。受启发于 DGCNN 模型^[15], 并以其为基线网络, 去除变换网络, 通过融合多分辨率特征对 EdgeConv 进行改进。分类网络用四个多分辨率图卷积模块学习点云上下文局部信息, 为克服点云的无序特性, 从而引入池化策略 (pooling strategy), 其中最大池化 (max pooling) 和平均池化 (avg pooling) 分别保留特征图中的最显著特征以及整体特征, 然后对其进行拼接后获得全局特征作为分类器的输入。最后由全连接层 (fully connected layer) 作为分类器输出分类得分 C ; 分割网络主干包含三个多分辨率图卷积模块, 池化策略选择 max pooling, 最终输出 P 个语义标签作为每个点的得分。

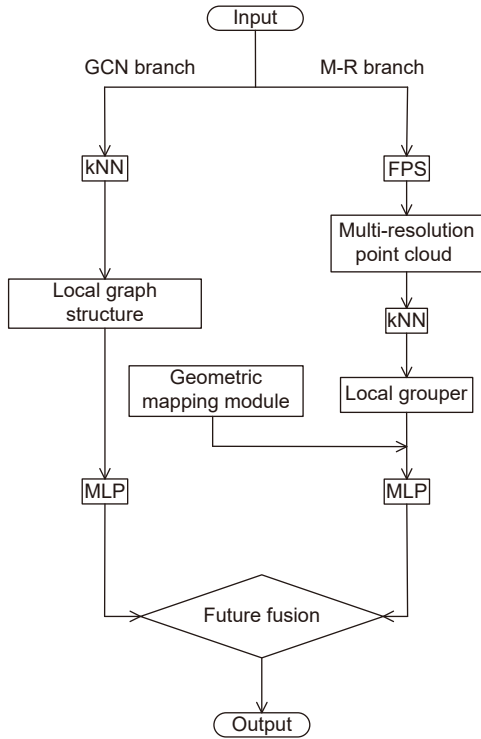


图 1 多分辨率图卷积模块算法流程图
Fig. 1 Multi-resolution graph convolution module algorithm flow chart

2.2 图卷积分支

与基于图方法的 DGCNN^[15] 类似, 将类图结构应用于点云特征学习可以更为有效地处理点云这类非结构化数据, 其核心在于选取点云表面的点作为节点, 并与邻域点之间建立边, 从而构建局部点对之间的几何联系。图卷积即为对该图结构进行卷积操作, 具体过程如图 3 所示。

首先, 定义输入的点云为包含 n 个点的点集 $P = \{p_1, p_2, \dots, p_n\}_{i=1,2,\dots,n} \subseteq \mathbf{R}^F$, 其中 \mathbf{R} 为实数集合, F 为点云的维度, 一般取 $F=3$, 即每一个点 p_i 用其空间坐标 (x_i, y_i, z_i) 表示, 此外还可以包括 RGB 信息、法线等额外的维度信息。由于在神经网络中, 下一层的输入一般为上一层的输出, 因此一般而言 F 表示给定的特征维度。

然后, 在点云的表面构建有向图 G , 表示为

$$G = (V, E), \quad (1)$$

其中: $V = \{p_1, p_2, \dots, p_n\}$ 和 $E \subseteq V \times V$ 分别为有向图的顶点和边。在图卷积模块中, 采用 kNN 算法对每一个顶点 p_i 进行检索, 找到其周围的 k 个邻域点构成邻域点集, 表示为

$$N_i = \{p_{i,1}, p_{i,2}, \dots, p_{i,j}\}_{j=1,2,\dots,k}, \quad (2)$$

其中, $p_{i,j}$ 为节点 p_i 的第 j 个邻域点, 通过计算 p_i 与 $\{p_{i,j}\}_{j=1,2,\dots,k}$ 之间的欧几里得距离作为图 G 的边, 表示为

$$e_{i,j} = \{p_i - p_{i,j}\}_{j=1,2,\dots,k}. \quad (3)$$

由式 (3) 计算边特征为

$$e_i = h_{\theta}\{p_i, p_i - p_{i,j}\}, \quad (4)$$

其中, h_{θ} 是一个包含一组可学习参数 θ 的非线性函数, 可实现特征维度 $\mathbf{R}^F \times \mathbf{R}^F \rightarrow \mathbf{R}^F$ 。此外, 由于对 h 的选择方式不同, 边特征的计算方式也不同, 其余计算方式如

$$e_i = h_{\theta}\{p_i\}, \quad (5)$$

$$e_i = h_{\theta}\{p_{i,j}\}. \quad (6)$$

式 (5) 为 PointNet^[12] 的方式, 只考虑每一个点,

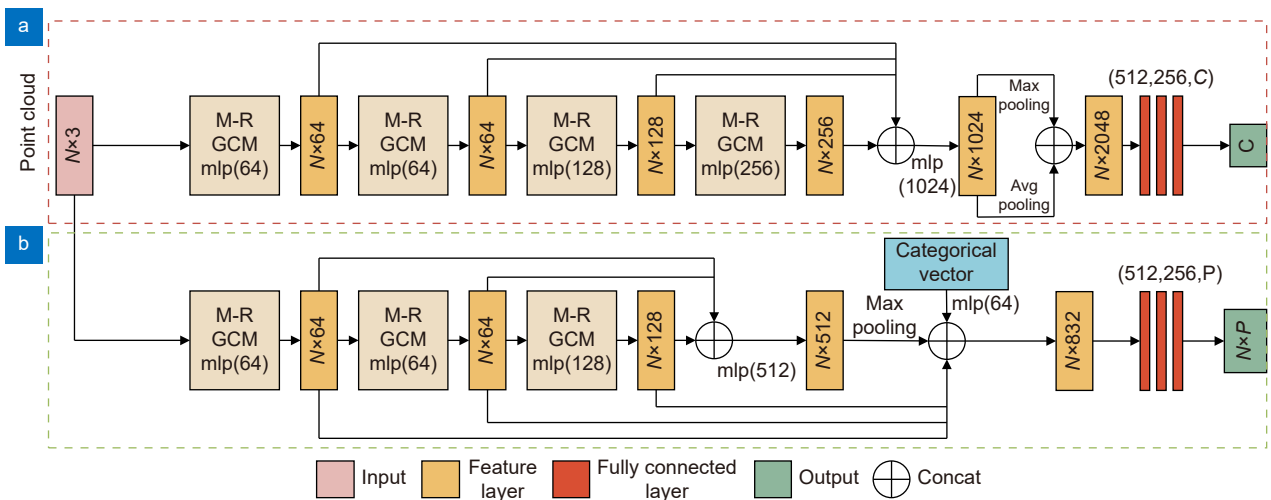


图 2 网络框架。(a) 分类网络; (b) 分割网络
Fig. 2 Network framework. (a) Classification network; (b) Segmentation network

忽视点云的局部结构, 式 (6) 为只考虑了点云的局部关系却忽视全局结构。因此本文采用式 (4) 所示的方法, 由 p_i 和 $\{p_i - p_{i,j}\}$ 同时兼顾全局和局部结构。

最后, 通过聚合函数对特征 $\{e_i\}_{i=1,2,\dots,n}$ 进行操作, 将更新后的图特征 f_G 表示为

$$f_G = \square_{j:(i,j) \in E} \{h_\theta(p_i, e_{i,j})\}, \quad (7)$$

其中: \square 表示 Max pooling 策略, h_θ 为 MLP 操作, 表示为 $h_\theta(p_i, e_{i,j}) = \text{LeakyReLU}(\theta_i \cdot p_i + \phi_i \cdot e_{i,j})$, 可学习参数 $\theta = (\theta_1, \dots, \theta_n, \phi_1, \dots, \phi_n)$ 。

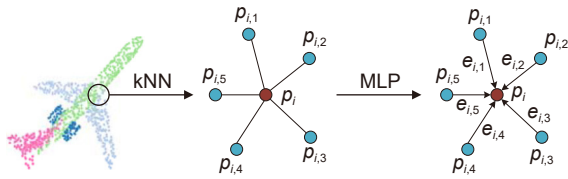


图 3 图卷积操作过程

Fig. 3 The operation procedure of graph convolution

考虑到图卷积网络的性能受限于预定义邻域的大小 (kNN 中 k 的取值, 实验部分将进行验证), 本文考虑用不同分辨率的点云特征对其进行融合补偿, 从而提高网络的性能。

2.3 多分辨率分支

在图像处理任务中, 神经网络层数的增加能更为有效地提取复杂特征, 但往往会以牺牲空间分辨率的方式来换取特征通道数的增加。本文从弥补空间分辨率的角度, 设计多分辨率分支对基线网络 DGCNN 的图卷积策略进行改进, 从而提高模型性能。多分辨率点云特征的学习过程如图 4 所示。

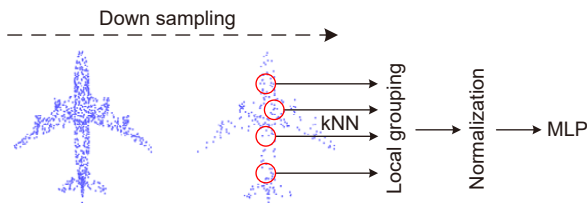


图 4 多分辨率点云特征学习过程

Fig. 4 The process of learning multi-resolution point cloud features

该分支对点云的处理分为三个阶段, 第一个阶段为基于 FPS 算法的下采样阶段 (down sampling)。当输入为包含 n 个点的点云 $X \subseteq \mathbf{R}^d$, 其中 d 为点云的特征维度, 每一个点表示为 $\{x_i\}_{i=1,2,\dots,n} \in X$, 选择其中的一个点 x_1 作为起始点得到第一个采样点集合, 表示为

$S = \{x_1\}$; 然后计算所有点与 x_1 的距离, 用数组 $L = (l_1, l_2, l_3, \dots)$ 进行储存, 选择其中距离最远的点 x_1 , 更新采样点集合为 $S = \{x_1, x_2\}$; 再计算所有点与 x_1 的距离, 若其中的一个点 x_i 与 x_1 的距离小于 l_i , 则将 l_i 进行更新, 数组 L 始终记录最小距离; 选择其中距离最远的点 x_3 更新采样点集合为 $S = \{x_1, x_2, x_3\}$, 重复采样 n' 次后, 得到的采样点集表示为

$$S = \{x_1, x_2, x_3, \dots, x_{n'}\}. \quad (8)$$

上述操作在改变点云分辨率的同时, 新的采样点集 S 仍然能够较好地表征点云的表面形状。第二个阶段为局部分组阶段 (local grouping), 在对不同分辨率点云进行处理时, 考虑其局部几何关系, 采用 kNN 算法对式 (8) 的点集进行局部组划分, 每一个局部组点集表示为

$$S_i = \{x_{i,1}, x_{i,2}, \dots, x_{i,k}\}. \quad (9)$$

考虑点云的稀疏特性, 为提高网络的性能, 在第三个阶段 (normalization) 中, 对每一个局部组 S_i 正态化, 具体操作如下

$$S'_i = \frac{S_i - \text{mean}(S_i)}{\sqrt{\text{bias}}}, \quad (10)$$

$$\text{bias} = \frac{\sum_{i=1}^{n'} \sum_{j=1}^k (x_{i,j} - x_i)^2}{k \times n' \times d}, \quad (11)$$

其中: mean 为求均值操作, bias 表示局部组中元素的方差, k 为局部组中邻域点数量。因此, 多分辨率特征表示为

$$f_{MR} = \text{LeakyReLU}(a \odot S'_i + b), \quad (12)$$

式中: $a \in \mathbf{R}^d$ 和 $b \in \mathbf{R}^d$ 是两个可学习参数, \odot 表示 Hadamard 乘积。

并且, 多分辨率分支作为并行分支而不是串行分支连接进网络。在每个多分辨率图卷积模块输出特征之前, 特征融合模块基于图卷积分支和多分辨率分支的特征, 通过网络计算获得一个自适应权重系数对两路特征进行加权输出, 并作为下一个多分辨率图卷积模块的输入, 在之后的每一层网络中动态更新, 直至进入全连接层进行分类。在整个过程两个分支不断进行特征交互, 以此实现特征图更新。

2.4 特征融合模块

为了使多分辨率特征更有效地对图特征进行补偿, 提出了一种特征融合模块, 相比较于线性特征聚合方式 (如加法运算), 该融合模块更具特征的自适应性。

特征融合原理如图 5 所示。

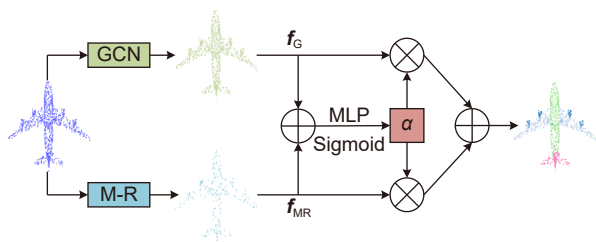


图 5 特征融合

Fig. 5 The operation of feature fusion

在每一个多分辨率图卷积模块中, 通过两个分支的网络分别学习到图特征 f_G 与多分辨率特征 f_{MR} , 为更有效地集成特征, 首先将两路特征进行逐元素求和, 以紧凑地表示点云特征 f :

$$f = f_G + f_{MR}. \quad (13)$$

接着基于该特征矩阵生成一个权重向量, 表示为

$$\alpha = \sigma(\Phi(f)), \quad (14)$$

式中: σ 为 Sigmoid 激活函数, Φ 为 MLP 操作。最终由该权重向量对两个分支的特征进行加权输出, 特征图表示为

$$f' = \alpha \odot f_G + (1 - \alpha) \odot f_{MR}, \quad (15)$$

其中: $\mathbf{1}$ 为全元素为 1 的向量, \odot 表示 Hadamard 乘积。

3 实验结果与分析

3.1 实验环境配置

实验参数设置如表 1 所示。为验证本文方法的有效性, 在 ModelNet40^[21]、ScanObjectNN^[22] 数据集上进行分类实验, 在 ShapeNet Part^[23] 数据集上进行部分分割实验。实验参数设置如表 1 所示, 所有实验均基于 Linux Ubuntu 系统, 训练所用 GPU 为 GeForce

表 1 实验参数设置

Table 1 Experimental parameter setting

参数项	分类网络	分割网络
输入点数	1024	2048
多分辨率点云点数	[896,768,640,512]	[896,768,640]
图卷积分支 k 取值	20	20
训练周期	250	300
优化器	SGD	SGD
训练批次	32	32
测试批次	16	16
初始学习速率	0.1	0.003

RTX 3090, 学习框架为 Python3.7+Pytorch-1.7.1。

3.2 基于 ModelNet40 数据集的点云分类

3.2.1 数据集描述

ModelNet40 数据集包含 40 个类别的 12311 个 CAD 模型, 其中 9843 个用于训练, 2468 个用于测试。在训练和测试期间没有执行任何数据预处理操作, 保证实验的有效性。

3.2.2 分类结果与分析

在 ModelNet40 数据集上的分类结果如表 2 所示, 评价指标为总体准确率 (overall accuracy, OA) 和类平均准确率 (mean class accuracy, mAcc)。

$$OA = \frac{TP}{TP + FN}, \quad (16)$$

$$mAcc = \frac{1}{c} \sum_{i=1}^c \frac{TP_i}{TP_i + FN_i}, \quad (17)$$

其中: TP 表示正确预测正样本的数目, FN 为错误预测正样本的数目, 而 TP_i 和 FN_i 分别表示每一个类别中正样本被预测正确和错误的样本数。

表 2 不同方法在 ModelNet40 数据集上的分类精度对比

Table 2 Comparison of classification accuracy with different methods on ModelNet40 dataset

方法	输入	点数/ 10^3	mAcc/%	OA/%
VoxNet ^[9]	体素	-	83.0	85.9
MVCNN ^[11]	多视图	-	-	90.1
PointNet ^[12]	坐标	1	86.0	89.2
PointNet++ ^[13]	坐标+法线	5	-	91.9
文献[24]	坐标+法线	1	89.8	91.6
文献[25]	坐标+法线	1	-	93.0
3D-GCN ^[26]	坐标	1	-	92.1
DGCNN ^[15]	坐标	1	90.2	92.9
LDGCNN ^[16]	坐标	1	90.3	92.9
DDGCN ^[27]	坐标	1	90.4	92.7
DRNet ^[28]	坐标	1	-	93.1
DGANet ^[29]	坐标	1	89.4	92.3
PCT ^[19]	坐标	1	-	93.2
AFM-Net ^[30]	坐标	1	89.4	92.85
文献[31]	坐标	1	89.02	92.5
Our	坐标	1	91.2	93.4

从表 2 中可以看出, 对比基于体素的经典方法 VoxNet 和基于多视图的 MVCNN, 本文的方法在 OA 这个指标上提升了 7.5% 和 2.7%; 与 PointNet 和 PointNet++ 相比也有 4.2% 和 1.5% 的提升, 并和基于

图的网络 DGCNN、LDGCNN、DGANet 相比分别提升了 0.5%、0.5% 和 1.1%。与新颖的方法如 PCT、AFM-Net 以及文献 [31] 相比, 也有一定的精度提升。实验结果表明, 改进后的多分辨率图卷积网络在分类任务中具有一定的优越性。

3.3 基于 ScanObjectNN 数据集的点云分类

3.3.1 数据集描述

ScanObjectNN 数据集是一个以真实世界对象为模型的数据集, 包含 15 个类别的 15000 个对象, 其中有 2902 个唯一对象实例。由于模型中存在复杂背景、噪声以及遮挡, 该数据集对于现有点云分类任务更具有挑战性。

3.3.2 分类结果与分析

基于 ScanObjectNN 数据集的分类结果如表 3 所示, 所有模型的输入均为坐标形式, 本文模型的 OA 和 mAcc 分别达到了 83.3%、81.7%。从表 3 中可以看出, 本文的方法在评估指标上比 PointNet 提高了 15.1%(OA)、18.3%(mAcc); 与基线网络 DGCNN 相比提高了 5.2%(OA)、8.1%(mAcc); 与新颖的方法 GBNet 相比提高 2.8%(OA)、3.9%(mAcc)。实验结果表明, 当数据集处于噪声干扰的情况下, 本文提出的方法通过弥补一定的空间分辨率, 能够更好地识别形状特征, 比起其它方法具有更强的形状分类能力。

表 3 不同方法在 ScanObjectNN 数据集上的分类精度对比

Table 3 Comparison of classification accuracy with different methods on ScanObjectNN dataset

方法	输入	mAcc/%	OA/%
PointNet ^[12]	坐标	63.4	68.2
PointNet++ ^[13]	坐标	75.4	77.9
DGCNN ^[15]	坐标	73.6	78.1
DRNet ^[28]	坐标	78.0	80.3
GBNet ^[32]	坐标	77.8	80.5
PRANet ^[33]	坐标	79.1	82.1
Ours	坐标	81.7	83.3

3.4 基于 ShapeNet Part 数据集的点云分割

3D 点云的部分分割是一项具有挑战性的细粒度识别任务。简单而言, 分割任务需要为每个点分配类别标签, 如机翼、机身等。为验证模型的分割性能, 在 ShapeNet Part 数据集上进行部分分割实验。

3.4.1 数据集描述

ShapeNet Part 数据集由 16 个类别的 16881 个模型组成, 共标记了 50 个零件标签, 其中 13998 个模型用于训练, 2874 个用于测试, 每一个模型分为 2 到 6 个不同的部分进行标注。训练时从每个形状中采样 2048 个点。

3.4.2 分割结果与分析

以 ShapeNet Part 为基准数据集, 部分分割的结果如表 4 所示, 分割实验的评价指标为平均交并比

表 4 ShapeNet Part 数据集上的部分分割结果

Table 4 Part segmentation results on the ShapeNet Part dataset

方法	PointNet ^[12]	PointNet++ ^[13]	文献[25]	3D-GCN ^[26]	LDGCNN ^[16]	DGANet ^[29]	DGCSA ^[34]	DGCNN ^[15]	本文
飞机	83.4	82.4	83.8	83.1	84.0	84.6	84.2	84.0	83.6
包	78.7	79.0	77.5	84.0	83.0	85.7	73.3	83.4	83.4
帐篷	82.5	87.7	87.9	86.6	84.9	87.8	82.3	86.7	88.4 ↑
车	74.9	77.3	78.7	77.5	78.4	78.5	77.7	77.8	78.4↑
椅子	89.6	90.8	90.8	90.3	90.6	91.0	91.0	90.6	89.7
耳机	73.0	71.8	77.3	74.1	74.4	77.3	75.3	74.7	80.5 ↑
吉他	91.5	91.0	91.8	90.9	91.0	91.2	91.2	91.2	91.8 ↑
刀	85.9	85.9	87.9	86.4	88.1	87.9	88.6	87.5	88.6 ↑
台灯	80.8	83.7	84.2	83.8	83.4	82.4	85.3	82.8	81.6
手提电脑	95.3	95.3	95.9	95.6	95.8	95.8	95.9	95.7	95.8↑
摩托	65.2	71.6	71.8	66.8	67.4	67.8	58.9	66.3	69.6↑
马克杯	93.0	94.1	95.1	94.8	94.9	94.2	94.3	94.9	94.4
手枪	81.2	81.3	80.9	81.3	82.3	81.1	81.8	81.1	83.7 ↑
火箭	57.9	58.7	59.6	59.6	59.2	59.7	56.9	63.5	62.5
滑板	72.8	76.4	76.6	75.7	76.0	75.7	75.4	74.5	82.0 ↑
桌子	80.6	82.6	82.4	82.8	81.9	82.0	82.7	82.6	83.0 ↑
mIoU	83.7	85.1	85.4	85.1	85.1	85.2	85.3	85.2	85.4 ↑

(mean intersection over union, mIoU)。

$$mIoU = \frac{TP}{TP+FP+FN}, \quad (18)$$

其中: TP 表示对正样本预测正确的数目, FP 为对负样本预测错误的数目, FN 为对正样本预测正确的数目。本文模型在部分分割任务中, 达到了 85.4% 的精度。对比经典算法 PointNet、PointNet++ 分别提高了 1.6% 和 0.2%。并且与基于图结构的算法 DGCNN、LDGCNN 和 DGANet 相比, 多分辨率图卷积网络的精度分别提高了 0.1%、0.2% 和 0.1%。在该实验中, 有 10 个类别的物体分割精度较于基线网络 DGCNN 均有所改善 (“↑” 代表有改善的物体类别), 实验结果表明, 多分辨率图卷积网络在分割任务中也实现了一定的优化。

为了能更直观地展示本文模型的分割效果, 将 16 个类别的物体进行分割结果可视化如图 6 所示。

同时, 将 DGCNN 作为基线网络, 展示了经本文方法改进后部分分割的细节对比如图 7 所示, 红色方框内效果差异。

通过上述定量、定性分析比较, 融合多分辨率特征的图卷积网络能正确分割大部分点, 并且对比基线网络, 本文方法在目标边界处的分割能力也有所提升, 相较于改进之前的网络有一定的优越性。

4 消融实验

为验证网络的有效性和鲁棒性, 以 ModelNet40 为基准数据集, 对模型各模块及超参数的选择进行分析验证。

4.1 图卷积分支中 k 取值对性能的影响

回顾 2.2 小节, 局部图结构基于 kNN 算法构造, 该预定义邻域的大小影响模型的性能, 通过改变 k 的

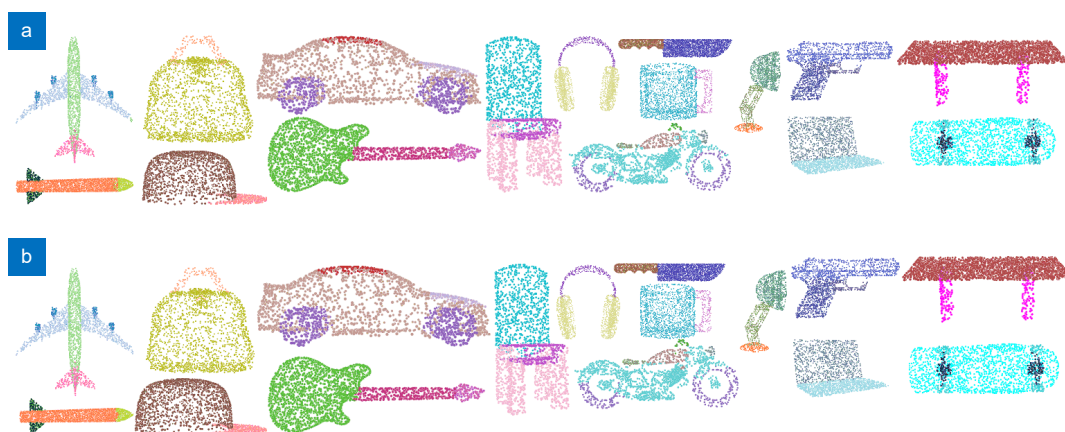


图 6 部分分割可视化结果。(a) 真实值; (b) 本文的分割结果

Fig. 6 The results of the part segmentation visualization. (a) Groud truth; (b) Ours

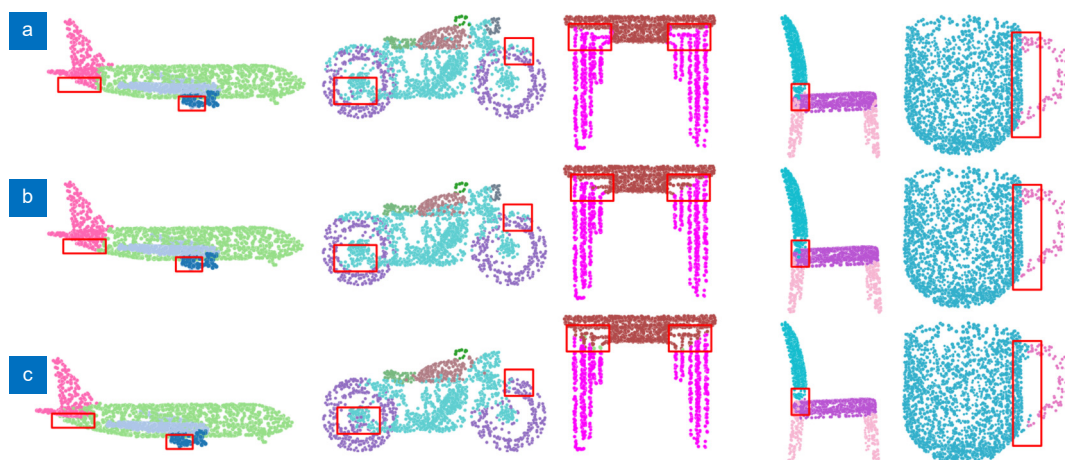


图 7 分割细节对比。(a) 真实值; (b) 本文方法; (c) 基线

Fig. 7 Comparison of segmentation details. (a) Groud truth; (b) Ours; (c) Baseline

取值进行实验以寻找合适的邻域大小。实验结果如表 5 所示。

表 5 不同 k 值对模型性能的影响Table 5 Effect of different k values on model performance

k	OA(%)	用多分辨率分支补偿后OA/%	提升/%
5	20.7	35.1	+14.4
10	85.4	88.3	+2.9
15	91.9	92.1	+0.2
20	92.5	93.4	+0.9
25	92.1	92.3	+0.2

实验表明, 当 k 值过小或过大都会导致网络性能下降, 取值为 20 的时候性能最佳。并且由表 5 可以看出, 当引入多分辨率分支对图特征进行补偿后, 网络的性能都得到了提升, 证明了多分辨率分支的有效性。

4.2 多分辨率图卷积模块的消融实验

在本节中设置四组实验以验证多分辨率图卷积模块各部分对网络性能的影响, 实验结果如表 6 所示。

表 6 多分辨率图卷积模块消融实验

Table 6 Ablation experiments of multi-resolution GCN module

实验	GCN分支	M-R分支	融合	mAcc/%	OA/%
1	√	×	×	89.9	92.5
2	×	√	×	84.0	89.1
3	√	√	×	89.9	92.6
4	√	√	√	91.2	93.4

实验 1: 基线网络, 仅使用图卷积分支进行特征学习;

实验 2: 仅使用多分辨率分支进行特征学习;

实验 3: 使用两个分支进行特征学习, 但特征集成方式采用线性组合方式;

实验 4: 在实验 3 的基础上将特征集成方式替换为所提出的特征融合模块。

实验结果表明, 同时使用两个分支的网络性能均优于使用单一分支, 并且当采用所提出的特征融合模块来聚合两个分支的特征后, 网络的精度进一步获得提升。

4.3 多分辨率点云规模的选取

在本文方法中, 点云的空间分辨率影响网络的性能, 每一个多分辨率图卷积模块中都存在图特征与不

同分辨率特征的交互, 表 7 展示了不同分辨率的点云对网络性能的影响。

表 7 不同分辨率点云对网络性能的影响

Table 7 The effect of different resolution point cloud on network performance

不同分辨率的点云	mAcc/%	OA/%
[512,384,256,128]	90.4	92.7
[640,512,384,256]	90.6	92.8
[768,640,512,384]	90.9	93.0
[896,768,640,512]	91.2	93.4

实验结果表明, 当四个多分辨率图卷积模块中不同分辨率点云的点数为 [896,768,640,512] 时, 模型的性能达到最优。

4.4 噪声鲁棒性测试

现实生产应用中, 通过扫描设备所获取的点云数据往往存在噪声干扰, 为测试网络对于噪声的鲁棒性, 对初始点云添加不同水平的随机噪声模拟实际生产应用的情况, 实验结果如图 8 所示。

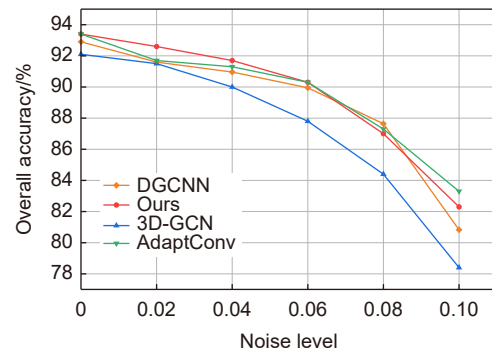


图 8 噪声鲁棒性测试

Fig. 8 Noise robustness testing

图 8 展示了在不同水平噪声干扰下本文方法与基线网络 DGCNN 以及基于图方法的 3D-GCN、AdaptConv^[35] 的精度对比, 可以看出在噪声干扰下, 本文方法的性能总体上优于 DGCNN, 与其它两种模

表 8 多种模型的噪声鲁棒性比较

Table 8 Comparison of the noise robustness of the several methods

噪声水平	下降程度			
	3D-GCN	AdaptConv	DGCNN	Ours
0.02	0.7 ↓	1.8↓	1.4↓	0.9↓
0.04	2.2↓	2.2↓	2.2↓	1.8 ↓
0.06	4.6↓	3.3↓	3.2↓	3.1 ↓
0.08	8.4↓	6.5↓	5.7 ↓	6.4↓
0.1	14.9↓	10.8 ↓	13.1↓	11.7↓

型对比也有一定优势。表 8 展示了两种方法在噪声干扰下精度下降的程度, 下降程度的计算公式如式 (19):

$$\downarrow \% = \frac{OA(\text{噪声下}) - OA(\text{无噪声})}{OA(\text{无噪声})} \quad (19)$$

结合图表分析, 多分辨率图卷积网络的噪声鲁棒性对比基线网络得到了提高, 在抗噪能力上具有一定的优越性。

4.5 特征提取模块数量消融实验

在图 2 的分类网络框架中, 由四个局部特征提取模块(多分辨率图卷积模块)组成, 在本节实验中, 将使用不同数量的模块以寻找最优的模型框架, 实验结果如表 9 所示。

表 9 不同数量特征提取模块对网络性能的影响
Table 9 The impact of different number of feature extraction modules on network performance

模块数量	mAcc/%	OA/%	每轮训练时间/s	模型参数量/M
3	89.7	92.4	63	2.8
4	91.2	93.4	139	3.6
5	90.6	93.1	323	4.8

在上述实验中, 使用三个模块时的点云分辨率为 [896,768,640], 四个模块为 [896,768,640,512], 五个模块为 [896,768,640,512,384], 分析实验结果可知特征提取模块的数量越多, 网络的性能并不会一直上升, 反而导致模块复杂度与所需的内存占用增加。因此, 为兼顾精度与效率, 选择四个特征提取模块最为合适。

5 结 论

本文在图卷积网络 DGCNN 的基础上提出了一种融合多分辨率特征的图卷积模块, 通过引入不同分辨率的点云特征对传统的图结构信息进行补偿, 克服传统图结构信息受限于预定义邻域规模的问题, 用空间分辨率丰富了点云的局部特征。并且相较于一般的特征线性聚合方式, 所提出的特征聚合模块更具有自适应性, 能更加有效地进行点云特征学习。通过在 ModelNet40、ScanObjectNN 以及 ShapeNet Part 三个具有挑战性的基准数据集上进行分类分割实验, 证明融合多分辨率特征的图卷积网络性能相比于其它模型均有一定提升。此外, 大量消融实验也验证了多分辨率图卷积网络在点云分析任务中的有效性和鲁棒性。

虽然, 融合多分辨率特征的图卷积网络在基准数据集上取得了令人满意的结果, 但在实际应用中, 样本的分布并不总是均匀的, 并且该网络在分割实验中, 样本边界的识别能力也有所欠缺, 这些都是本文网络将来的优化方向。今后的工作将聚焦于捕获点云细粒度特征的研究前沿, 进一步提高网络在复杂场景下的识别能力, 并与实际生产应用相结合。

参考文献

- [1] Zhang X Y, Chen M L, Liu X J, et al. Classification of terrestrial point cloud considering point density and unknown angular resolution[J]. *Laser Technol*, 2023, **47**(1): 59–66.
张昕怡, 陈茂霖, 刘祥江, 等. 顾及点密度与未知角分辨率的地面点云分类[J]. *激光技术*, 2023, **47**(1): 59–66.
- [2] Li J N, Wang Z, Xu T F. Three-dimensional object detection technology based on point cloud data[J]. *Acta Opt Sin*, 2023, **43**(15): 1515001.
李佳男, 王泽, 许廷发. 基于点云数据的三维目标检测技术研究进展[J]. *光学学报*, 2023, **43**(15): 1515001.
- [3] Lu K L, Xue J, Tao C B. Multi target tracking based on spatial mask prediction and point cloud projection[J]. *Opto-Electron Eng*, 2022, **49**(9): 220024.
陆康亮, 薛俊, 陶重森. 融合空间掩膜预测与点云投影的多目标跟踪[J]. *光电工程*, 2022, **49**(9): 220024.
- [4] Chen Z S, Li C L, Zuo W, et al. Study on improved point cloud registration algorithm enhanced by double down-sampling[J]. *Automot Eng*, 2023, **45**(4): 572–578.
陈仲生, 李潮林, 左旺, 等. 双重下采样增强的点云改进配准算法研究[J]. *汽车工程*, 2023, **45**(4): 572–578.
- [5] Li M J, Yu Z K, Liu X, et al. Progress of point cloud algorithm in medical field[J]. *J Image Graphics*, 2020, **25**(10): 2013–2023.
李美佳, 于泽宽, 刘晓, 等. 点云算法在医学领域的研究进展[J]. *中国图象图形学报*, 2020, **25**(10): 2013–2023.
- [6] Xia J Z, Sun H M, Hu S H, et al. 3D laser point cloud clustering method based on image information constraints[J]. *Opto-Electron Eng*, 2023, **50**(2): 220148.
夏金泽, 孙浩铭, 胡盛辉, 等. 基于图像信息约束的三维激光点云聚类方法[J]. *光电工程*, 2023, **50**(2): 220148.
- [7] Bai H Q, Xia Y H, Yang M L, et al. Multi-resolution 3D reconstruction of karst caves based on the feature line extraction of 3D laser point cloud[J]. *Laser Optoelectron Prog*, 2020, **57**(20): 202802.
柏宏强, 夏永华, 杨明龙, 等. 基于三维激光点云特征线提取的溶洞多分辨率三维重建方法研究[J]. *激光与光电子学进展*, 2020, **57**(20): 202802.
- [8] Maturana D, Scherer S. Voxnet: a 3D convolutional neural network for real-time object recognition[C]//*Proceedings of 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2015: 922–928. <https://doi.org/10.1109/IROS.2015.7353481>.
- [9] Riegler G, Osman Ulusoy A, Geiger A. OctNet: learning deep 3D representations at high resolutions[C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 6620–6629. <https://doi.org/10.1109/CVPR.2017.701>.

- [10] Xiao Z T, Gao J, Wu D Q, et al. A uniform downsampling method for three-dimensional point clouds based on voxel grids[J]. *Mach Des Manuf*, 2023(8): 180–184.
肖正涛, 高健, 吴东庆, 等. 一种基于体素网格的三维点云均匀降采样方法[J]. *机械设计与制造*, 2023(8): 180–184.
- [11] Su H, Maji S, Kalogerakis E, et al. Multi-view convolutional neural networks for 3D shape recognition[C]//*Proceedings of 2015 IEEE International Conference on Computer Vision*, 2015: 945–953. <https://doi.org/10.1109/ICCV.2015.114>.
- [12] Charles R Q, Su H, Mo K C, et al. PointNet: deep learning on point sets for 3D classification and segmentation[C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 77–85. <https://doi.org/10.1109/CVPR.2017.16>.
- [13] 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 31st International Conference on Neural Information Processing Systems*, 2017: 5105–5114. <https://doi.org/10.5555/3295222.3295263>.
- [14] Simonovsky M, Komodakis N. Dynamic edge-conditioned filters in convolutional neural networks on graphs[C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 29–38. <https://doi.org/10.1109/CVPR.2017.11>.
- [15] Wang Y, Sun Y B, Liu Z W, et al. Dynamic graph CNN for learning on point clouds[J]. *ACM Trans Graphics*, 2019, 38(5): 146.
- [16] Zhang K G, Hao M, Wang J, et al. Linked dynamic graph CNN: learning on point cloud via linking hierarchical features[Z]. arXiv: 1904.10014, 2019. <https://doi.org/10.48550/arXiv.1904.10014>.
- [17] Liu B, Fan Y C. A point cloud classification model based on improved dynamic graph convolution[J]. *China Sci*, 2022, 17(11): 1230–1235, 1266.
刘斌, 樊云超. 基于改进动态图卷积的点云分类模型[J]. *中国科技论文*, 2022, 17(11): 1230–1235, 1266.
- [18] Sun Q, Liu H Y, He J, et al. DAGC: employing dual attention and graph convolution for point cloud based place recognition[C]//*Proceedings of 2020 International Conference on Multimedia Retrieval*, 2020: 224–232. <https://doi.org/10.1145/3372278.3390693>.
- [19] Guo M H, Cai J X, Liu Z N, et al. PCT: point cloud transformer[J]. *Comput Vis Med*, 2021, 7(2): 187–199.
- [20] Liu H, Tian S H. Deep 3D point cloud classification and segmentation network based on GateNet[J]. *Vis Comput*, 2023. <https://doi.org/10.1007/s00371-023-02826-w>.
- [21] Wu Z R, Song S R, Khosla A, et al. 3D shapeNets: a deep representation for volumetric shapes[C]//*Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*, 2015: 1912–1920. <https://doi.org/10.1109/CVPR.2015.7298801>.
- [22] 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 2019 IEEE/CVF International Conference on Computer Vision*, 2019: 1588–1597. <https://doi.org/10.1109/ICCV.2019.00167>.
- [23] Yi L, Kim V G, Ceylan D, et al. A scalable active framework for region annotation in 3D shape collections[J]. *ACM Trans Graphics*, 2016, 35(6): 210.
- [24] Wang Z X, Ren M W. DST-Pointnet++: A novel point cloud classification network based on pointnet++[J]. *Comput Digit Eng*, 2022, 50(11): 2497–2501.
王子璇, 任明武. DST-Pointnet++: 基于Pointnet++改进的点云分类网络[J]. *计算机与数字工程*, 2022, 50(11): 2497–2501.
- [25] Wang B J, Nong L P, Zhang W H, et al. 3D point cloud classification and segmentation network based on Spider convolution[J]. *J Comput Appl*, 2020, 40(6): 1607–1612.
王本杰, 农丽萍, 张文辉, 等. 基于Spider卷积的三维点云分类与分割网络[J]. *计算机应用*, 2020, 40(6): 1607–1612.
- [26] Lin Z H, Huang S Y, Wang Y C F. Convolution in the cloud: learning deformable kernels in 3D graph convolution networks for point cloud analysis[C]//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020: 1797–1806. <https://doi.org/10.1109/CVPR42600.2020.00187>.
- [27] Chen L F, Zhang Q. DDGCN: graph convolution network based on direction and distance for point cloud learning[J]. *Vis Comput*, 2023, 39(3): 863–873.
- [28] Qiu S, Anwar S, Barnes N. Dense-resolution network for point cloud classification and segmentation[C]//*Proceedings of 2021 IEEE Winter Conference on Applications of Computer Vision*, 2021: 3812–3821. <https://doi.org/10.1109/WACV48630.2021.00386>.
- [29] Wan J, Xie Z, Xu Y Y, et al. DGANet: a dilated graph attention-based network for local feature extraction on 3D point clouds[J]. *Remote Sens*, 2021, 13(17): 3484.
- [30] Zhang R M, Cheng T, Yin L, et al. Attention fusion and multi-scale point cloud classification network[J]. *J Huaibei Normal Univ (Nat Sci)*, 2023, 44(1): 70–75.
张润梅, 程婷, 尹蕾, 等. 一种注意力融合的多尺度点云分类网络[J]. *淮北师范大学学报(自然科学版)*, 2023, 44(1): 70–75.
- [31] Guo Y E, Ren M W. Improved point cloud classification network based on PointConv[J]. *Comput Digit Eng*, 2022, 50(12): 2737–2740, 2764.
国玉恩, 任明武. 基于PointConv改进的点云分类网络[J]. *计算机与数字工程*, 2022, 50(12): 2737–2740, 2764.
- [32] Qiu S, Anwar S, Barnes N. Geometric back-projection network for point cloud classification[J]. *IEEE Trans Multimedia*, 2021, 24: 1943–1955.
- [33] Cheng S L, Chen X W, He X W, et al. PRA-Net: point relation-aware network for 3D point cloud analysis[J]. *IEEE Trans Image Process*, 2021, 30: 4436–4448.
- [34] Song W, Cai W Y, He S Q, et al. Dynamic graph convolution with spatial attention for point cloud classification and segmentation[J]. *J Image Graphics*, 2021, 26(11): 2691–2702.
宋巍, 蔡万源, 何盛琪, 等. 结合动态图卷积和空间注意力的点云分类与分割[J]. *中国图象图形学报*, 2021, 26(11): 2691–2702.
- [35] Zhou H R, Feng Y D, Fang M S, et al. Adaptive graph convolution for point cloud analysis[C]//*Proceedings of 2021 IEEE/CVF International Conference on Computer Vision*, 2021: 4945–4954. <https://doi.org/10.1109/ICCV48922.2021.00492>.

作者简介



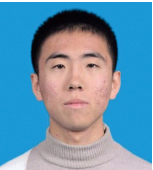
陶志勇 (1978-), 男, 博士, 教授, 主要研究方向为机器学习、图像处理、智能信息处理。

E-mail: xyzmail@126.com



【通信作者】李衡 (1998-), 男, 硕士研究生, 主要研究方向为机器学习、三维点云数据处理。

E-mail: PaperLH@163.com



豆淼森 (2001-), 男, 硕士研究生, 主要研究方向为三维点云数据处理。

E-mail: doumiaosen@163.com



林森 (1980-), 男, 博士, 副教授, 主要研究方向为图像处理与模式识别。

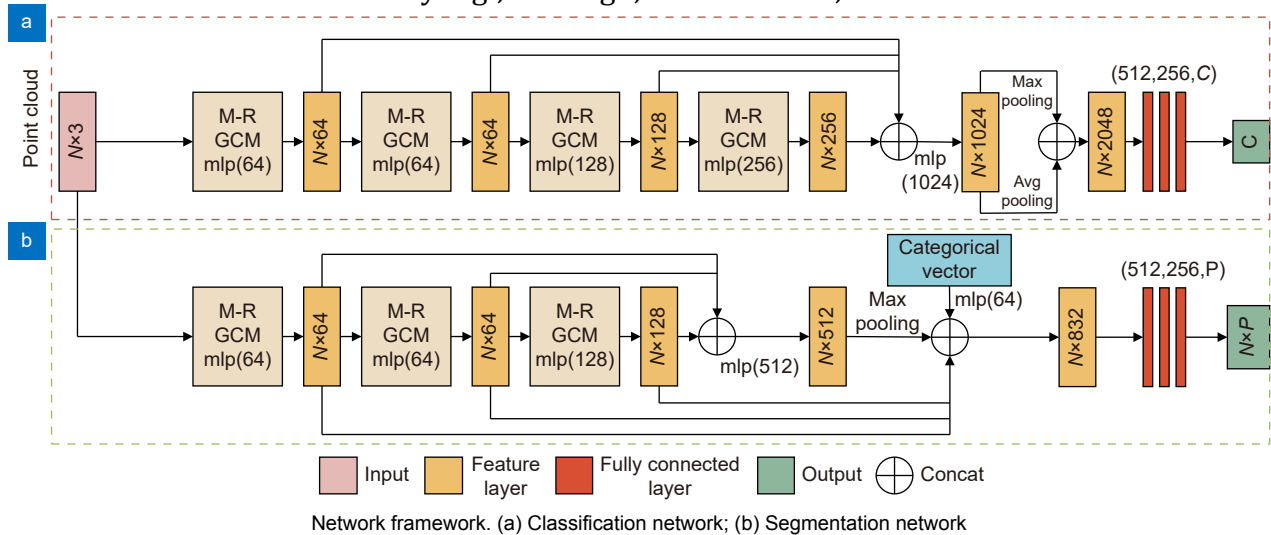
E-mail: lin_sen6@126.com



扫描二维码, 获取PDF全文

Multi-resolution feature fusion for point cloud classification and segmentation network

Tao Zhiyong¹, Li Heng^{1*}, Dou Miaosen¹, Lin Sen²



Overview: In recent years, 3D point cloud analysis has become a hot topic in computer vision and been widely used in mapping, medical imaging, and autonomous driving. As a 3D image representation, point cloud contains rich geometric information. With the development of 3D scanning technologies such as LiDAR, the acquisition of point clouds is becoming more accessible. Since convolutional neural networks (CNN) have greatly improved the results of computer vision tasks, neural networks are becoming a mainstream approach in image processing. Traditional 2D images comprise regular and dense pixels, and CNNs apply to 2D image processing. However, point cloud data are sparse and disordered; each point does not contain additional information (e.g., RGB). Using traditional CNNs for point cloud learning tasks is a challenging task. The graph-like structure can effectively represent non-Euclidean data like point clouds, and this method largely solves the problem of difficulty in learning the local features of point clouds. Since the graph structure construction process is generally based on the k -nearest neighbor algorithm (kNN), the size of the predefined neighborhood limits the effectiveness of the local graph structure. If the value of k is too small, it will lead to an incomplete representation of local information. At the same time, too large a value of k will introduce information redundancy, leading to performance degradation. To this end, we propose a multi-resolution graph convolutional network to perform the point cloud analysis task. The network learns the local features of point clouds by constructing graph structures and then downsamples the point clouds using the farthest point sampling method (FPS) to obtain multi-resolution point cloud data, followed by feature learning for point clouds at different resolutions. To overcome the effect of predefined neighborhoods, we compensate local features with multi-resolution features and efficiently aggregate point cloud features by the feature fusion module. To verify the classification and segmentation performance of the model, we perform classification experiments on ModelNet40 and ScanObjectNN datasets and part segmentation experiments on ShapeNet Part dataset. It is experimentally verified that the compensation of point cloud local graph structure information with multi-resolution features can enhance point clouds' local feature learning ability. The multi-resolution graph convolutional network proposed in this paper can effectively capture the local features of point clouds and achieve good results in classification and segmentation tasks. A large number of ablation experiments verify the effectiveness and robustness of the model.

Tao Z Y, Li H, Dou M S, et al. Multi-resolution feature fusion for point cloud classification and segmentation network[J]. *Opto-Electron Eng*, 2023, 50(10): 230166; DOI: [10.12086/oe.2023.230166](https://doi.org/10.12086/oe.2023.230166)

Foundation item: Project supported by Department of Science & Technology of Liaoning Province Application Fundamental Research Project(2022JH2/101300274)

¹School of Electronic and Information Engineering, Liaoning Technical University, Huludao, Liaoning 125100, China; ²School of Automation and Electrical Engineering, Shenyang Ligong University, Shenyang, Liaoning 110159, China

* E-mail: PaperLH@163.com