

面向局部学习的点云分割分类网络

冯锦梁,王 蕾,温智成,叶森辉,马 晗
(东华理工大学,江西 南昌 330013)

摘要:在各种3D对象表示中,三维点云越来越受欢迎,其中基于点的方法在各种数据集上都表现出了良好的性能。针对PointNet++只关注了点本身的信息,并未关注相邻点的信息,同时它采用最大池化聚合局部信息,导致丢失次最大值信息。由此,文中提出Con-PointNet++网络,该网络利用增强局部信息模块,以关注相邻点的信息,从而增强局部信息特征提取;采用局部注意力机制下的融合池化模块,将最大池化与注意力池化特征信息融合,得到更为丰富的局部特征信息。本文方法在室内数据集S3DIS的Area_5区域上评估模型语义分割能力,mIoU达55.2%;在数据集ModelNet40上评估模型分类效果,OA达91.2%。与其他方法相比,所提模型性能均有提升,进一步证明了本文方法的有效性。

关键词:三维点云;语义分割;分类;局部注意力机制;局部增强模块

中图分类号:TP391.41;TN249 **文献标识码:**A **DOI:**10.3969/j.issn.1001-5078.2024.01.004

A point cloud segmentation classification network for local learning

FENG Jin-liang, WANG Lei, WEN Zhi-cheng, YE Sen-hui, MA Han
(East China University of Technology, Nanchang 330013, China)

Abstract:3D point clouds are becoming increasingly popular in various 3D object representations, where point-based methods have shown good performance on a variety of datasets. In response to PointNet++ only focuses on the information of the points themselves and not the information of the neighboring points, while it uses max-pooling to aggregate local information, resulting in the loss of sub-maximal information. As a result, the Con-PointNet++ network is proposed to make full use of the enhanced Local Information Module to focus on the information of neighboring points and thus enhance the local information feature extraction. Then, the fusion pooling module under the local attention mechanism is used to fuse the max-pooling and attention pooling feature information to obtain richer local feature information. The proposed method evaluates the model semantic segmentation ability on Area_5 of indoor dataset S3DIS, with mIoU of 55.2%; and the model classification effect on dataset ModelNet40, with OA of 91.2%. Compared with other methods, the performance of the proposed model is improved, which further demonstrates the effectiveness of the proposed method.

Keywords:3D point cloud; semantic segmentation; classification; local attention mechanism; local enhancement module

基金项目:江西省核地学数据科学与系统工程技术研究中心基金项目(No. JELRGBDT202202);江西省放射性地学大数据技术工程实验室开放基金项目(No. JELRGBDT202103)资助。

作者简介:冯锦梁(1998-),男,硕士研究生,主要从事三维点云语义分割与分类技术的研究。E-mail:179815700@qq.com

通讯作者:王 蕾(1979-),女,博士,教授,硕士生导师,主要研究方向为计算机视觉,机器学习,图像处理,三维点云数据处理,虚拟现实可视化等。E-mail:wlei598@163.com

收稿日期:2023-03-29; **修订日期:**2023-05-11

1 引言

在各种 3D 对象表示中,点云越来越受欢迎,成为最基本的 3D 数据表示方法之一。这种流行源于立体匹配算法的发展和三维传感器的日益普及,比如激光雷达,它可以产生点云作为原始输出。高质量的点云是连接虚拟世界和现实世界的桥梁。例如,语义信息可以丰富周围的信息。点云的语义分割、分类任务是计算机视觉^[1-3]、智能驾驶^[4]、遥感测绘^[5-6]、文物古迹^[7-8]等研究的基础。与图像不同的是,点云是无序的、无结构的,在不同距离上密度分布不均匀,这些特征使得现有深度学习在点云语义分割与分类方法上比图像更加困难。现有深度学习方法在点云分类分割上具有效率低、局部细节识别易混淆等缺点^[9]。如何有效地对三维点云模型进行高精度识别成为一个亟待解决的问题^[10]。

2 相关工作

针对深度学习下的点云分割或分类,目前存在有几种方法。它们可以大致分为基于多视图的方法、基于体素的方法和基于点的方法。

基于多视图的方法首先使用虚拟相机捕捉不同位置下多个视角的图,然后通过卷积层与池化层聚合输出,如多视图卷积神经网络(Multi-View CNN, MVCNN)^[11],不过该方法不能动态选择视图,导致丢失大量关键的几何空间信息。后续针对 MVCNN 的不足,Feng 等提出组视图卷积神经网络(Group-View CNN, GVCNN)^[12],该方法将多视图先按内容分组,再投影到指定的 2D 平面上,然后应用 2D FCNs(Fully Convolutional Networks)来识别 3D 对象。不过这类方法没有充分利用数据,因为在二维投影中丢失了第 z 维的信息,限制了他们在 3D 场景中分析复杂 3D 物体的能力。

基于体素的方法将点云离散在体积 3D 网格上,使用标准 3D FCNs 进行形状语义分割。如 VoxNet^[13]、VoxelNet^[14]与将体素简化为 pillars(即仅在平面中进行体素化)的 PointPillars^[15]等网络,这类方法需要进行数据转换,由于网格结构的分辨率有限,导致了额外的计算成本和信息损失。此外,基于体素的方法对 GPU 内存的要求也很高,因此限制其应用于单个对象和小场景领域中的发展。

与上述两种方法相比,基于点的方法直接将点云作为网络的输入,在各种数据集上都表现出了良

好的性能。2017 年,Charles 团队推出的 PointNet^[16]网络是该方向的第一个成功尝试,它采用了一系列多层感知器(mlp)和 max-pooling 算子解决点云的无序问题以及学习全局特征,在 3D 形状分类和场景分割的场景中取得了巨大成功。但是,PointNet 针对的是孤立点的学习,忽略了点与点之间的联系关系,因此该网络在不同尺度上提取局部信息的能力较弱。针对 PointNet 网络出现的不足之处,Charles 团队后续推出了 PointNet++^[17]网络,它提出了一个新颖的 Set Abstraction 模块,该模块对点云进行子采样,对邻域进行分组,通过一组 mlp 提取局部信息,然后通过约简层(即池化)聚合局部信息,并且 PointNet++ 通过度量点与点之间空间的距离,能够利用上下文尺度的增长学习局部特征,使得精度进一步提升。Yang 等人^[18]提出 Point Attention Transformers(PATs),该方法通过其自身的绝对位置和相对于其相邻的相对位置来表示每个点,并通过 MLPs 学习高维特征。然后由集体重新排序注意力模块(Group Shuffle Attention, GSA)来捕获点之间的关系,使用 Gumbel 子集采样层(Gumbel Subset Sampling, GSS)来学习分层特征。Zhao 等人^[19]基于 PointNet++ 提出的 PointWeb 网络,通过利用学习局部邻域上下文信息与 Adaptive Feature Adjustment(AFA)模块来自适应的调整特征。Yan 等人^[20]提出的 PointASNL 中,利用自适应采样模块(Adaptive Sampling, AS)来自适应调整最远点采样算法(Furthest Point Sampling, FPS)所采样的点的坐标和特征,并提出一个 local-non local(L-NL)模块来捕捉这些采样点的局部和远距离依赖性,进一步减少噪声对模型的影响。

本文受到点云语义分割网络 RandLA-Net^[21]的启发,将局部注意力机制加入基础框架 PointNet++ ,使得网络可以自适应地筛选局部信息,关注更为重要的特征信息。同时加入增强局部信息模块,通过添加局部信息模块获取更为丰富的局部信息。两者通过增加点云的细粒度局部特征的学习能力,使得网络分割与分类精度进一步提升,本文将这种新型的网络命名为 Con-PointNet++ 。

3 本文方法

3.1 基准模型 PointNet++

PointNet++ 核心思想:在局部的区域区域内,重复的,迭代性使用 PointNet。由于局部区域具有置换

不变性,所以整体也具有局部不变性。PointNet++ 这个核心思想主要借鉴了 CNN 的多层感受野的思想。首先从小的局部区域捕获细粒度的局部结构来提取局部特征,之后局部特征被分组到更大的单元,通过迭代得到更为丰富的局部特征,最后获取到整个输入点集的特征。PointNet++ 包含了分割与分类网络,具体步骤如下:

1) 其先通过 sampling 与 grouping 方法在整个点云的局部采样并划一个范围,将里面的点作为局部的特征;

2) 用 PointNet 进行一次特征的提取;

3) 重复完成上述两个步骤,即降采样操作,使得原本的点集个数变得越来越少,而每个点都带有上一层通过 PointNet 提取出来的局部特征,也就是降采样操作后每个点包含的信息变多了;

4) 分类网络只需要将 PointNet 提取出来的局部特征进行最大池化,然后经过全连接层进行分类即可;而分割网络相对复杂,需要先使用编码器主簿提取局部特征,通过解码器逐步对特征进行上采样,并与对于编码器输出的特征进行连接,最后使用全连接层对输出结果进行预测。

3.2 增强模式下的局部特征学习

由于 3D 点云数据规模庞大,PointNet++ 等系列网络模型均将原始输入点云数据划分成不同子区域,在对子区域中的原始点云采样后分批次输入卷积神经网络 CNN 中学习特征。PointNet++ 采用最远点采样法 (Farthest Point Sampling, FPS) 以维持子区域中点云数据的结构。但由于 PointNet++ 基础框架只考虑点本身的信息,其点领域的信息并没有考虑,采样后不可避免丢失了中心点与其邻域的邻接关系,从而限制了局部几何特征的充分表达与网络对其的学习能力,从而导致模型的性能下降。

针对上述不足,本文使用增强局部信息的模块,在原有 PointNet++ 基础框架中引入邻域信息添加模块,增加局部特征信息量,如图 1 所示。

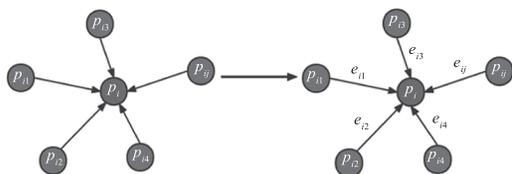


图 1 增强局部信息模块

Fig. 1 Enhanced local information module

首先使用 K 近邻搜索算法 (KNN) 获取每个中心点 (p_i, f_i) 的相邻点 (p_{ij}, f_{ij}) , 公式如下:

$$p_{ij}, f_{ij} = KNN(p_i, f_i) \quad (1)$$

然后分别计算相邻点坐标和对应特征与中心点坐标和对应特征的差值,并将两者结果在特征维度上进行拼接得到增强后的局部特征 (f_{aug}) , 公式如下:

$$f_{aug} = \text{concat}(p_{ij} - p_i, f_{ij} - f_i) \quad (2)$$

最后使用 MLP 学习局部特征 (F_{aug}) , 公式如下:

$$F_{aug} = MLP(f_{aug}) \quad (3)$$

如上述表达式所示,MLP 为多层感知机,concat 为拼接特征方法; p_i 为中心点坐标; f_i 为中心点特征; $p_{i1}, p_{i2} \dots p_{ij}$ 为邻域点坐标; $f_{i1}, f_{i2} \dots f_{ij}$ 为邻域点特征。

3.3 局部注意力机制下的融合池化

现有的卷积神经网络通常使用最大/平均池化操作,可以通过集成邻域特征以生成全局特征向量,但是该操作很难整合到相邻点的特征,导致丢失大量特征信息。本文将注意力机制引入局部特征提取机制中,同时为了保持网络特征输出的平移不变性,将注意力机制下的细节特征与最大池化获取的显著特征进行融合操作,获取到特征信息更为丰富的局部特征,其结构如图 2 所示。

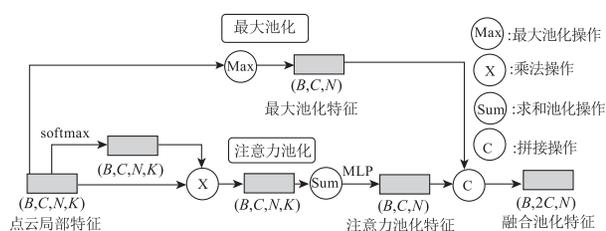


图 2 局部注意力机制下的融合池化

Fig. 2 Fusion pooling under local attention mechanism

局部注意力机制下的融合池化模块分为 2 部分,首先对该局部特征使用注意力机制自适应地学习到局部点的权重 (score), 表达式为:

$$score = \text{Softmax}(MLP(f_{ij})) \quad (4)$$

然后利用公式(3)得到的权重信息对局部特征进行加权求和获得局部注意力池化特征 (f_{atten}) , 表达式为:

$$f_{atten} = MLP(\text{Sum}(score \cdot f_{ij})) \quad (5)$$

同时将点云局部特征进行最大池化得到最大池化特征 (f_{max}) , 表达式为:

$$f_{max} = \text{Max}(f_{ij}) \quad (6)$$

最后将公式(4)得到的局部注意力池化特征与公式(5)得到的最大池化特征进行拼接得到融合池化特征($result$),既包含了更多有用信息,同时保证网络特征输出的平移不变性,表达式为:

$$result = f_{atten} + f_{max} \quad (7)$$

3.4 Con-pointnet++ 网络结构设计

Con-pointnet++ 网络结构如图3所示,分割网络由 Encoder 层、Decoder 层和分类器3部分组成;分类网络由 Encoder 层和分类器组成。

(1)Encoder 层包括点云数据采样、增强局部信息模块和融合池化层。数据采样层使用最远点采样法和邻域点采样法获取局部信息,同时增强局部信息模块将中心点的邻域信息考虑进来,组成更为丰富的局部信息。最后融合池化模块分别通过对局部特征进行最大池化与注意力池化,将二者拼接得到更为丰富的局部特征信息;(2)分割网络中 Decoder 层主要使用插值法来实现,首先计算出逆向距离的权重,得到插值后的点特征,再将该特征与 Encoder 层输出的特征进行拼接,最后经由分类器输出语义分割结果;在分类网络中直接使用分类器将 Encoder 层输出的结果进行分类。

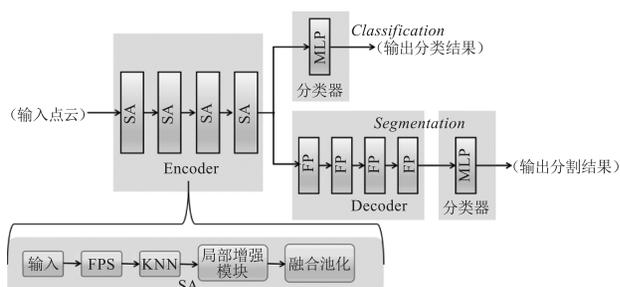


图3 Con-pointnet++ 网络结构示意图

Fig. 3 Con-pointnet++ network structure schematic

4 实验与结果分析

4.1 点云分割实验配置

本文使用室内场景数据集 S3DIS 与 ModelNet40 数据集分别评估 Con-PointNet++ 的 3D 点云分割性能和分类性能,并在最后一小节进行消融实验,进一步验证增强局部信息模块与注意力机制下的融合池化模块对于 Con-PointNet++ 模型的有效性。实验环境配置如下:硬件为显卡 NVIDIA GeForce RTX 3080 10 GB GPU、处理器 Intel(R) Core(TM) i9-9900kCPU @3.60 GHz、内存 32 GB;软件为 UBUNTU20.04 LTS + CUDA11.2 + Python 3.6,深度学习

框架为 Pytorch 1.12.1。

网络训练参数设置如下:采用余弦退火算法使学习率逐渐衰减;采用 Adam 优化器,初始学习率为 0.01;选择使用 ReLu 激活函数。训练时, batch size 设置为 32;测试时, batch size 也同样设置为 32。分割网络训练 100 个 epoch,分类网络训练 250 个 epoch。

4.2 评价指标说明

网络语义分割精度评估指标采用平均交并比(mIoU)与(mAcc);网络分类精度评估指标采用平均类精度(mAcc)与总体分类精度(OA)。在三维点云分割与分类实验中,mAcc、OA 和 mIoU 是最常用的衡量方法精度的指标^[22]。mIoU 指计算两个集合的交并比,在语义分割的问题中,这两个集合分别表示为真实值(ground truth)和预测值(predicted segmentation);mAcc 用于计算每个类别的分类精度再除以类别数获取到平均类别精度;OA 用于计算。用计算公式分别如下:

$$mIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{P_{ii}}{\sum_{j=0}^k P_{ij} + \sum_{j=0}^k P_{ji} - P_{ii}} \quad (8)$$

$$mAcc = \frac{1}{k+1} \sum_{i=0}^k \frac{P_{ii}}{\sum_{j=0}^k P_{ij}} \quad (9)$$

$$OA = \frac{\sum_{i=0}^k P_{ii}}{\sum_{j=0}^k P_{ij}} \quad (10)$$

其中, k 表示数据集中点云的类别个数(包含空类);有 $K+1$ 个类, p_{ij} 是类 i 隐含属于类 j 的最小单位(如像素、体素、网格、点),即 p_{ii} 代表真阳性, p_{ij} 与 p_{ji} 分别代表假阳性与假阴性^[22]。mIoU、acc 和 mAcc 值越大,代表分割、分类效果越好。

4.3 实验

4.3.1 室内场景 S3DIS 语义分割实验

本文语义分割实验以 S3DIS 数据集为研究对象,该数据集是斯坦福大学提供的大场景室内 3D 点云数据集,包含 6 个教学和办公 Area,共有 271 个独立房间,特征点超过了 2.15 亿。

训练时将房间分为 1×1 (m) 的块,随机对每块区域选取 4096 个点输入,其中每个点包含 XYZ 坐标信息、归一化坐标和 RGB 颜色信息。以数据集中的 Area_1、Area_2、Area_3、Area_4、Area_6 为训练

样本, Area_5 为测试区域评估模型分割精度。结果如表 1 所示, 可以看出, 与基准网络 PointNet++ 相比较, 本文 Con-PointNet++ 在 mIoU 上可提升 5.2%; 与其他网络 (PointNet、DGCNN、SegCloud、ASIS (Associatively Segmenting Instances and Semantics) 和 DeepGCN) 相比, mAcc 与 mIoU 都有着不同程度的提升, 因此可以验证加入增强局部信息、与注意力机制下的融合池化模块下的 Con-PointNet++ 在整体性能上明显优于其他 3D 点云场景语义分割网络。

表 1 不同方法在 S3DIS Area_5 中的结果对比

Tab. 1 Comparison of the results of different methods in S3DIS Area_5

方法	mAcc/%	mIoU/%
PointNet ^[16]	49.0	41.1
PointNet++ ^[10]	/	50.0
DGCNN ^[23]	/	47.0
SEGCloud ^[24]	57.4	48.9
ASIS ^[25]	/	51.1
DeepGCN ^[26]	/	52.5
Ours	64.1	55.2

选取 S3DIS 数据集中的 Area5_conferenceRoom_1、Area5_hallway_12 和 Area5_office_15 三个房间作为可视化分割场景样本, 将所提算法的可视化分割结果与基准模型 PointNet++ 进行对比。图 4(b)~(c) 分别为 Area5_conferenceRoom_1、Area5_hallway_12 和 Area5_office_15 的原始点云、真实分割结果、PointNet++ 分割结果和所提算法的分割结果可视化图, 其中每一列表示一个场景, 每一行表示对应方法的结果。

由图 4 可看出, 加入了增强局部特征模块与注意力机制下的融合池化模块后的网络, 分割能力明显增强。本文所提方法与 PointNet++ 相比较, 在场景(a)中, 椅子细节处、墙体与墙上的图画处分割明显更优, PointNet++ 几乎丢失了墙上的图画全部信息; 在场景(b)中, PointNet++ 对于走廊两边的门框与墙上图画细节丢失较多, 本文所提算法则相对精确的还原了真实分割场景; 场景(c)是较为杂乱的办公室场景, PointNet++ 分割效果明显变差, 左墙投影幕布部分识别成墙体, 中间的办公桌错误识别成椅子, 而本文所提算法识别精准, 且相对 PointNet++ 分

割细节保留更多, 在办公桌与椅子连接处的特征点识别精度明显更高。

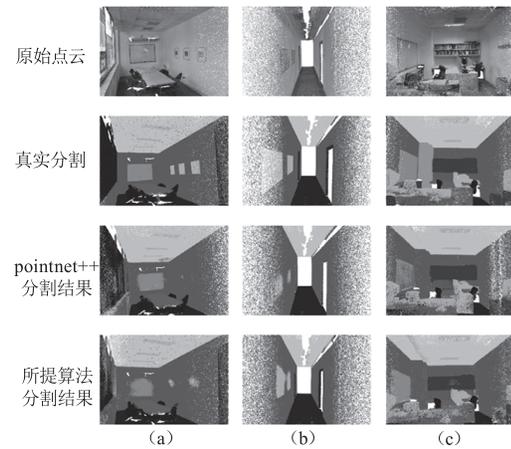


图 4 Area5_conferenceRoom_1

Fig. 4 Comparison of semantic segmentation visualization results of different algorithms for Area5_conferenceRoom_1

4.3.2 ModelNet40 上的分类实验

本文分类实验在标准公开数据集 ModelNet40 上进行, 该数据集由普林顿大学提供, 共包含 12311 个三维 CAD 模型, 9843 个模型用以训练, 2468 个模型用以测试, 总共包含 40 个类别。同时为了让网络学习到旋转不变性、放缩不变性和对噪声具有一定的抗干扰能力, 通过对数据进行随机旋转、放缩以及添加噪声进行数据增强, 增强网络的鲁棒性。

结果如表 2 所示, 可以看出, 对于基准网络 PointNet++, 本文网络在整体分类精度达到 91.2%, 提升约 0.5%, 同时对于所提其他网络 (PointNet、MVCNN 与 VoxNet) 相比, 因加入了增强局部特征模块与注意力机制下的融合池化模块, 在 OA 与 mAcc 性能均有提升, 通过实验结果可看出, 本文所提网络在分类任务上同样具有较强能力。

表 2 不同方法在 ModelNet40 中的结果对比

Tab. 2 Comparison of the results of different methods in ModelNet40

方法	mAcc/%	OA/%
PointNet ^[16]	86.0	89.2
MVCNN ^[9]	/	90.1
VoxNet ^[11]	83.0	85.9
PointNet++ ^[17]	/	90.7
Ours	88.4	91.2

4.4 消融实验

为了更好地验证增强局部信息(+ Loc)、与注意力机制下的融合池化(+ Mer)模块对 Con-PointNet++ 的有效性,将基准模型 PointNet++ 与不同模块组合,用以分析各模块对网络精度的影响。

分割网络保持在 S3DIS Area_5 上进行消融实验。如表 3 所示,相比于基准模型 PointNet++ 的 mIoU(约 50.0%),各个模块的加入均使得基准模型的 3D 点云场景语义分割能力获得明显提升。

表 3 针对 S3DIS Area_5 数据的消融实验结果

Tab.3 Results of ablation experiments for S3DIS Area_5 data

方法	mAcc/%	mIoU/%
PointNet++ ^[10]	/	50.0
+ Loc	62.8	53.6
+ Mer	64.2	54.8
Con-PointNet++	64.1	55.2

分类网络保持在 ModelNet40 数据集上进行消融实验。如表 4 所示,相比于基准模型 PointNet++ 的 OA(约 90.7%),各个模块的加入也均使得基准模型的 3D 点云分类能力获得明显提升。

表 4 针对 ModelNet40 的消融实验结果

Tab.4 Results of ablation experiments for ModelNet40

方法	mAcc/%	OA/%
PointNet++ ^[17]	/	90.7
+ Loc	87.7	90.8
+ Mer	87.9	91.0
Con-PointNet++	88.4	91.2

5 结 语

针对 PointNet++ 下采样时只关注点本身的信息以及使用最大池化导致丢失次最大值信息等问题,本文提出 Con-PointNet++ 网络。通过加入增强局部信息模块将相邻点的信息添加关注;使用局部注意力机制下的融合池化,将最大池化获取的最大值信息与注意力池化自适应地获取细节信息进行融合,获取到更为丰富的局部特征信息。数据集 S3DIS、ModelNet40 的实验以及消融实验证明所提方法的有效性。最后,本文期望 Con-PointNet++ 网络随着点云的发展能够释放更深层次的潜力。

参考文献:

- [1] Wang C, Samari B, Siddiqi K. Local spectral graph convolution for point set feature learning [C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018:52-66.
- [2] Hua B S, Tran M K, Yeung S K. Pointwise convolutional neural networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018:984-993.
- [3] GAO Jiayue, XU Hongli, SHAO Kailang, et al. An adaptive edge detection method based on local edge feature descriptor [J]. Chinese Journal of Lasers, 2020, 47(6):193-201. (in Chinese)
高佳月, 许宏丽, 邵凯亮, 等. 基于局部边缘特征描述子的自适应边缘检测算法 [J]. 中国激光, 2020, 47(6):193-201.
- [4] Jia Z, Yuan H, Zhao X, et al. Single-cell genetic analysis of lung tumor cells based on self-driving micro-cavity array chip [J]. Talanta, 2021, 226:122172.
- [5] Lu Junyan, Jia Hongguang, Gao Fang, et al. Reconstruction of digital surface model of single-view remote sensing image by semantic segmentation network [J]. Journal of Electronics and Information, 2021, 43(4):974-981. (in Chinese)
卢俊言, 贾宏光, 高放, 等. 语义分割网络重建单视图遥感影像数字表面模型 [J]. 电子与信息学报, 2021, 43(4):974-981.
- [6] Cheng Xianghong, Fan Shimiao. Underwater terrain-aided navigation based on improved Gauss and particle filtering [J]. Journal of Chinese Inertial Technology, 2019, 27(2):199-204. (in Chinese)
程向红, 范时秒. 基于改进高斯和粒子滤波的海底地形辅助导航 [J]. 中国惯性技术学报, 2019, 27(2):199-204.
- [7] Yu Peiyong, Yang Songyong. Research on restoration of ancient cultural relics based on 3D laser scanning technology [J]. Digital Space, 2017, 141(7):266. (in Chinese)
余培永, 杨松勇. 基于三维激光扫描技术的古文物修复研究 [J]. 数码世界, 2017(7):266-266.
- [8] 叶珉吕. 基于点云的历史建筑形变信息提取方法研究 [J]. 测绘与空间地理信息, 2019, 42(11):111-113.
- [9] 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.
- [10] 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.
- [11] 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. 2015: 945 – 953.
- [12] Feng Y, Zhang Z, Zhao X, et al. Gvcnn: Group-view convolutional neural networks for 3d shape recognition [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 264 – 272.
- [13] Maturana D, Scherer S. Voxnet: A 3d convolutional neural network for real-time object recognition [C]//2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2015: 922 – 928.
- [14] Zhou Y, Tuzel O. Voxelnet: End-to-end learning for point cloud based 3d object detection [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 4490 – 4499.
- [15] Lang A H, Vora S, Caesar H, et al. Pointpillars: Fast encoders for object detection from point clouds [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 12697 – 12705.
- [16] 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. 2017: 652 – 660.
- [17] Qi C R, Yi L, Su H, et al. Pointnet ++: Deep hierarchical feature learning on point sets in a metric space [J]. Advances in Neural Information Processing Systems, 2017, 30.
- [18] Yang J, Zhang Q, Ni B, et al. Modeling point clouds with self-attention and gumbel subset sampling [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 3323 – 3332.
- [19] 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, 2019: 5565 – 5573.
- [20] 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, 2020: 5589 – 5598.
- [21] Hu Q, Yang B, Xie L, et al. Randla-net: Efficient semantic segmentation of large-scale point clouds [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11108 – 11117.
- [22] He Y, Yu H, Liu X, et al. Deep learning based 3D segmentation: A survey [J]. ArXiv Preprint ArXiv: 2103.05423, 2021.
- [23] 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.
- [24] Tchapmi L, Choy C, Armeni I, et al. Segcloud: Semantic segmentation of 3d point clouds [C]//2017 International Conference on 3D Vision (3DV). IEEE, 2017: 537 – 547.
- [25] Wang X, Liu S, Shen X, et al. Associatively segmenting instances and semantics in point clouds [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 4096 – 4105.
- [26] Li G, Müller M, Qian G, et al. Deepgens: making gens go as deep as cnns [J]. arXiv: 1910.06849.