

doi:10.16055/j.issn.1672-058X.2022.0005.001

融合边缘检测的 3D 点云语义分割方法研究

陈玲, 许钢, 伏娜娜, 胡志锋, 郑书展

(安徽工程大学 检测技术与节能装置安徽省重点实验室, 安徽 芜湖 241000)

摘要:针对点云分割中分割目标不明确,边缘不清晰,全局特征与边缘特征未能有效融合等问题,提出了一种融合边缘检测的 3D 点云语义分割算法。首先,通过 3D 点云语义分割网络对点云数据进行初步提取区域内的全局语义特征;然后,采用引入了注意力机制的语义边缘检测网络,能够更好地对点云数据中的物体进行特征提取增强,抑制非边缘信息的产生,得到了具有丰富的语义信息的边缘特征;最后,通过融合模块将属于同一物体的语义特征融合起来进行分割细化处理,使得分割目标更精确;此外,使用了双重语义损失函数,使网络产生具有更好边界的语义分割结果。通过搭建实验平台和使用 S3DIS 标准数据集进行测试,改进后的算法在数据集上的平均交互比为 70.21%,在精度上较 KPConv 语义分割算法有所提高。实验结果表明:该算法能够有效改善物体边界分割不清晰、边缘信息模糊等问题,总体分割性能良好。

关键词:3D 点云;语义分割;语义边缘检测;特征融合

中图分类号:TP391

文献标志码:A

文章编号:1672-058X(2022)05-0001-09

0 引言

随着 3D 采集技术的不断发展与完善,3D 数据运用于各个不同的领域中。点云作为一种常用的 3D 数据表达形式,能够在空间里将原始的几何信息完整地保留下来,不会产生离散化。作为理解和分析 3D 场景基础的点云语义分割^[1],更是未来研究的热点。点云语义分割是区分点云中的每个点,将点云划分为一块块具有特定语义信息的区域,在计算机视觉中具有重要研究意义。

随着卷积神经网络、深度学习的快速发展,也推动了点云语义分割技术不断前进,在不同领域中应用范围广泛,比如在无人驾驶、室外建筑物识别^[2]

和室内机器人定位等领域。近年来,为了解决 3D 点云不规则且无序性等问题,许多学者利用深度学习^[3]的方法将其转换为可用于卷积神经网络处理的规则结构。在此基础上 Qi 等^[4]提出了 PointNet 利用多层感知机学习逐点数据提取全局特征,再使用 MLP 对每个点进行分割,但该方法对局部特征的感知较差。为此 Lian 等^[5]提出了一种高密度的 PointNet++,它能够在不同尺度上提取局部特征;柴玉晶等^[6]引入了图注意力机制增强局部特征的感知能力来更好地学习点云的空间分布;Huang 等^[7]提出了条件随机场的方法,认为可在语义分割后期改进物体边缘的分割效果;Ye 等^[8]提出了递归神经网络(RNN)进行端到端学习,更好地捕获点云中的上下文特征并将其合并。但这些方法在融合全局特

收稿日期:2021-07-13;修回日期:2021-09-30.

基金项目:安徽省重点实验室开放课题(2017070503B026-A01).

作者简介:陈玲(1997—),女,安徽池州人,硕士研究生,从事智能控制技术研究.

通讯作者:许钢(1972—),男,安徽芜湖人,博士,教授,从事检测技术与自动化装置、语音与图像信号处理、机器视觉等研究. Email: xugang@ahpu.edu.cn.

征和局部特征时会丢失点云丰富的几何特征和空间密度分布。

为了更好地融合全局特征与局部特征, TAKIKAWA 等^[9]设计了一个双流 CNN 结构,将形状流与经典流并行处理,这种结构可以在物体边缘产生更清晰的预测,并且对小物体和瘦物体有更好的性能。但这种方法现用于 2D 图像中且未对分割出的物体边界进行语义信息的提取。综上所述,在现有的方法中,物体边缘的分割精度仍有待加强且未能将全局信息及局部信息进行有效融合。为此,本文提出了一种融合语义边缘检测的三维点云语义分割算法。在文献[10]的基础上引入了注意力机制对点云的边缘信息能够有效提取,减少感知歧义,生成精确的语义边缘。同时设计了一个融合模块,将语义边缘检测网络得到的边缘特征和语义分割网络得到的区域特征进行有效融合并对其进行优化。融合后的网络改善了语义分割结果中分割目标不明确、边缘不清晰等缺点,同时对小物体也有更好的分割性能。

1 相关工作

1.1 2D 语义边缘检测

目前,在二维图像中,一些工作利用基于 CNN 的方法进行边缘提取与分类。CASENet^[11]利用全局嵌套边缘检测(HED)网络^[12]增加类别感知的语义信息,将低层和高层特征与多标签损失函数相结合进行监督学习。在后续工作中,SEAL 网络^[13]通过减少标注过程中不可避免的不对齐导致的训练标签噪声来优化 CASENet。为解决固定权重融合的限制性,提出一种动态特征融合(DFE)模型^[14],根据特定的输入为每个位置推断出合适的多层次特征融合权重,从而有助于产生更准确的边缘预测。

与现有工作不同,本文为了对 3D 点云进行有效的语义边缘检测,从 2D 语义边缘检测中借鉴了提取增强特征的思想,对 3D 点云中的对象进行物体边界的提取,从而更好地识别三维物体。

1.2 融合边缘检测的语义分割

对于二维图像,将边缘检测任务和语义分割任务进行联合学习,对提升语义分割学习速率、预测精度是

一个可取的做法。在 Gated-SCNN 中,利用了分割和边缘预测任务之间的对偶性,将边界信息融入分割 CNN 中来辅助分割,并引入了双重损失来细化语义分割和边缘预测。DecoupleSegNets^[15]则是通过对不同监督下的主体和边缘进行解耦,再与相应的两个正交部分损失函数将其合并完成最终分割任务。以上两个网络都是利用二值边缘检测联合语义分割来提升性能。而 RPCNet^[16]是将语义边缘检测与语义分割任务进行联合学习,识别属于目标边界的像素类别,以提高边缘分割的正确性。

利用类别感知的 3D 边缘检测方法与 3D 点云语义分割方法进行联合学习,两者的耦合性会更加一致。此外,与以前的工作不同,它将分割和边缘检测之间的交互限制在特征和网络结构的共享上,同时利用了语义分割的全局特征和语义边缘检测的边缘特征之间的密切关系。

2 网络结构设计

2.1 网络整体结构

如图 1 所示,该架构是融合语义边缘检测的语义分割网络整体结构,由两个主流网络构成。首先通过一个共同的特征编码器对点云进行特征提取与编码,保留完整且相同的点云空间信息,然后经过不同的解码结构分别得到不同的特征信息。语义分割解码模块得到的是语义全局特征,语义边缘检测解码模块得到的具有语义信息的边缘特征,为了更好地融合全局特征与边缘特征,采用融合算法来处理特征合并,以达到最终理想的语义分割结果。

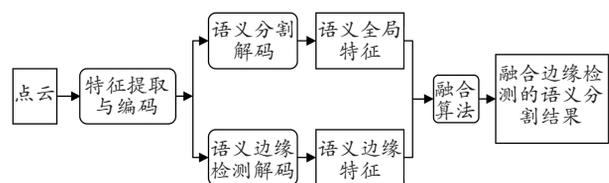


图 1 语义分割网络整体框架

Fig. 1 The overall framework of semantic segmentation network

2.2 语义分割网络

2.2.1 特征提取

为了减少信息的损失,采用直接对点云进行卷积处理的方法,不使用点云的其他中间表示形式。该网

络采用基于核的方法 KPConv (Kernel Point Convolution)^[17] 建立骨干网络,克服了之前卷积方法的局限性。KPConv 分为可变卷积核和不变卷积核,文中采用的是不变卷积核,它具有良好的收敛性能。

将点云中的点定义为 $p \in R^{N \times 3}$, 点特征定义为 $F \in R^{N \times D}$, 并确定一个以点 x 为球心, r 为半径的一个球体 B_r^3 。在球体内确定 Q 个 Kernel Points 定义为 $\{\tilde{x}_q | q < Q\} \subset B_r^3$, 每个核心点都有一个权重矩阵 W_q 定义为 $\{W_q | q < Q\} \subset R^{D_{in} \times D_{out}}$ 。对于落在球体范围内的点 x_i , 核函数被用来计算该点的权重矩阵, 该点特征由这个矩阵进行变换, 即:

$$g(x_i - x) = \sum_{q < Q} h(x_i - x, \tilde{x}_q) W_q$$

其中, $x_i - x$ 为点 x_i 相对于 x 的距离, $g(x_i - x)$ 等价于 q 个权重矩阵加权和, 权重系数 $h(x_i - x, \tilde{x}_q)$ 定义为

$$h(x_i - x, \tilde{x}_q) = \max(0, 1 - \frac{\|(x_i - x) - \tilde{x}_q\|}{\sigma})$$

权重随着相对距离的增大而减小, 取值范围为 $[0, 1]$ 。

对于落在球内的点 x_i , 对应于它们特征 f_i , 通过上一步的方法获得新的特征, 最后将特征累加作为点 x 的特征, 即:

$$(F * g)(x) = \sum_{x_i \in N_x} g(x_i - x) f_i$$

以上就是对点云进行特征提取的过程。

2.2.2 编码器与解码器

语义分割网络由编码器—解码器构成, 如图 2 所示, 编码器内表示的是特征提取与编码部分, 语义分割解码器内表示的是解码部分。

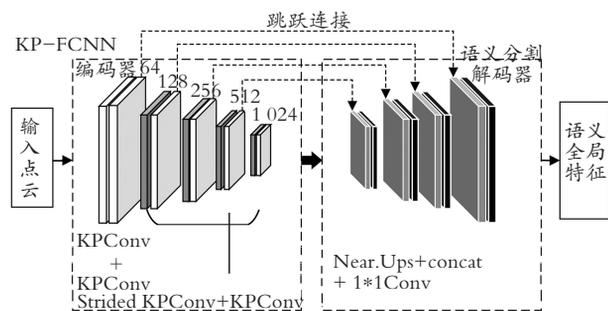


图 2 语义分割网络模块

Fig. 2 Network module of semantic segmentation

如图 2 所示的 KP-FCNN 是一个全连接的卷积神经网络, 其中编码部分被设计为基于 ResNet^[18] 网络结构, 其中包含五层卷积网络, 每一层卷积网络中都由两个卷积模块构成, 除了第一层由两个相同的

KPConv 卷积块构成, 其他每一层都由 KPConv 和 Strided KPConv 卷积块构成。

如图 3 所示, KPConv 卷积块代替了每一层中 BN (Batch Normalization) 和 Leaky ReLU Activation 的作用, 其中 BN 层主要对点云的分布进行归一化, 使用更大的学习速率使得梯度传播能够更稳定地进行, 增加网络的泛化能力。ReLU 激活函数的作用是在使用梯度下降法时, 会使收敛速度更快。与一般的网络结构不同, 该网络已经有了网格采样, 因此在每一层新的位置上不需要池化层进行下采样^[19] 操作, 在网络结构中使用 KPConv 卷积提取小区域的局部特征向量并且称为“跨步 KPConv”, 类比图像中的跨步卷积。编码器阶段主要对点云进行不同尺度的特征提取。

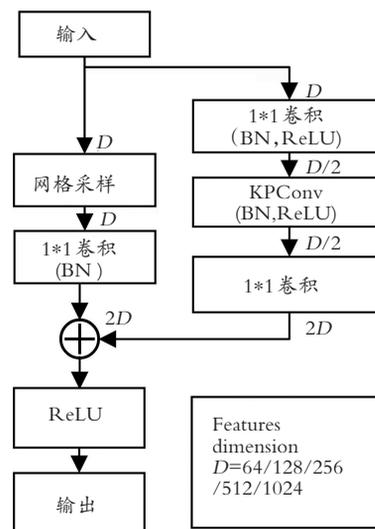


图 3 核点网络层

Fig. 3 KPConv network layer

解码部分使用了最邻近上采样来获得最终的逐点特征。编码层和解码层之间的特征使用跳跃连接来传递, 以上特征被并联到上采样部分通过一维卷积来处理。由以上所述的网络结构得到了具有语义信息的全局特征, 也就是初步语义分割结果。

2.3 语义边缘检测网络

语义边缘检测网络如图 4 所示, 编码器内是与语义分割网络相同的编码结构对点云的特征进行提取, 语义边缘检测解码器内表示的是对提取的特征增强, 捕获目标的边缘信息并对其添加语义信息。

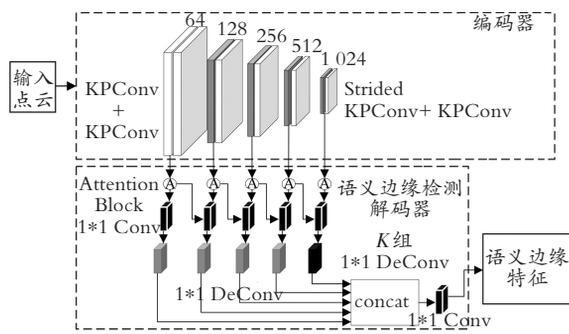


图 4 语义边缘检测网络模块

Fig. 4 Network module of semantic edge detection

语义边缘检测解码器的输入与语义分割解码器的输入均来自前一级编码器所使用的五层卷积网络,该网络去除了所有的连接层得到了全卷积网络,去除了第五层的池化操作,避免了高层信息的丢失,增加了感受野,对特征边缘及小物体进行更好地定位。

为了对提取的不同尺度特征进行增强处理,语义边缘检测解码加入了注意力机制^[20]模块(Attention Block)。如图 4 所示,在每一层后加入了注意力机制模块,且将每一层经过注意力机制模块提取后的特征加入后一层卷积块中,多种不同尺度的特征进行提取融合。如图 5 所示的注意力机制模块,将输入的点云数量设为 N ,特征维度为 C ,通道数为 D 。在该模块内,通过卷积层、全连接层和 softmax 得到各点之间的权重概率分布,然后对各点特征进行加权处理,得到带不同点的注意力新特征。

$$H(x) = W_i(x) * T_i(x) + x$$

其中, $W_i(x)$ 是第 i 个点经过 softmax 归一化后的权重, $W_i \in [0, 1]$ 且有 $\sum_{i=1}^N W_i = 1$, $T_i(x)$ 为经过卷积变换后的第 i 个点的特征, x 为输入点特征。

在增强特征提取操作后加上一个 Deconv layer 用于放大特征信息。为了使边缘轮廓加入语义信息,在以上基本网络的第五层加入了一个分类模块,该分类模块采用一个双线性上采样(由 K 组反卷积层去完成)以产生 K 个激活边缘结果(其中 K 表示语义类别),每个结果的大小都与输入点云相同。每一层经过反卷积操作得到的输出结果进行 concat 融合,随后使用一个 1×1 Conv layer 进行特征融合,得出输出点云大小与输入点云大小相同的结果。由以上网络结构得到了具有语义信息的边缘轮廓特征结果。

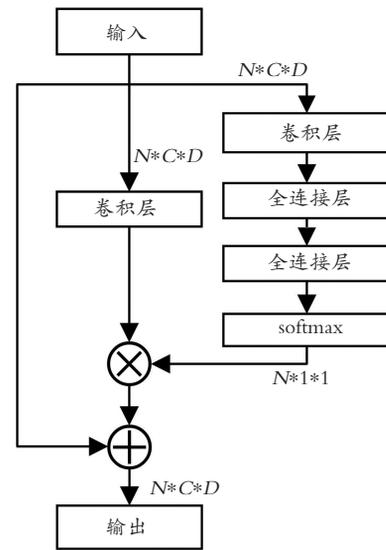


图 5 注意力机制模块

Fig. 5 Attention block

2.4 融合结构

融合结构如图 6 所示,将语义分割网络得到的全局特征结果与语义边缘检测网络得到的边缘特征结果进行融合,将融合后的结果进行分割细化得到最终语义分割结果。

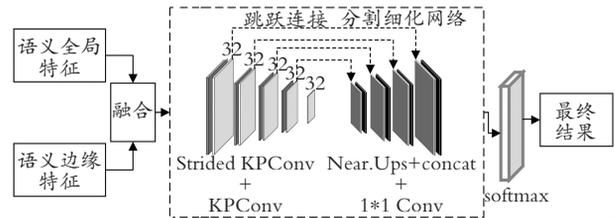


图 6 融合模块

Fig. 6 Fusion module

在进行两路特征融合时,采用加权融合^[21]的方式利用两路信息,其中语义全局特征为 A ,语义边缘特征为 B ,融合后的特征 M 可表示为

$$M = W(A+B) \otimes A + (1-W(A+B)) \otimes B$$

其中, W 表示为全局特征映射的权值,两路特征映射的权值之和为 1。

为了更好地融合全局和边缘特征,构建了一个分割细化网络。该网络如图 6 中分割细化网络所示,在编码结构中有五层卷积网络,除第五层以外,每一层都由 KPConv 卷积块和 Strided KPConv 卷积块构成,其中各层的通道数的大小为 32。解码部分同样使用了最邻近上采样来获得最终的全局特征。使用跳跃连接在编码和解码中间层之间传递特征并用 concat 函数进行融合,将局部特征与全局特征同

时考虑,所有特征被并联到上采样部分通过一个 1×1 卷积进行处理。

Softmax用于多分类过程中,首先将所有预测的结果转化为非负数,然后对转换后的结果进行归一化处理,最后预测边界的点将具有较大的激活值。采用softmax函数进行分类处理,得到最终的语义分割结果。

3 损失函数设计

本文模型是将语义边缘检测及语义分割进行联合学习,故使用了双重语义损失函数,使网络具有更好的语义边界结果。

针对语义边缘检测网络中损失函数的选取,采用了CASENet^[15]中通过多标签学习框架来解决点云中语义边缘检测问题的思想。因此,将多标签损失函数应用于语义边缘检测网络中。假设输入的一组点云 P ,根据语义类别 K ,可以获得 K 个标签边缘图映射 $\{Y_1, Y_2, \dots, Y_k\}$ 及 K 个标签边缘点映射 $\{\bar{Y}_1, \bar{Y}_2, \dots, \bar{Y}_k\}$,其中 \bar{Y}_k 表示第 K 类语义边缘的真值,其损失函数定义如下:

$$loss_{edge} = \sum_{k \in K} L_k = \sum_{k \in K} \sum_{p \in P} \left\{ -\beta \bar{Y}_k(p) \log Y_k(p) - (1 - \beta) (1 - \bar{Y}_k(p)) \log (1 - Y_k(p)) \right\}$$

其中, β 是 K 类点云中非边缘点的百分比,以解释样本数的偏度。该损失函数用来监督边缘信息的提取,强迫边缘检测部分只学习边缘信息。

针对语义分割网络中损失函数的选取,采用了一个标准的多类别交叉熵损失函数。假设输入的一组点云 P ,具有语义类别 K ,其损失函数定义如下:

$$loss_{seg} = - \sum_{k \in K} \sum_{p \in P} \hat{S}_k(p) \log (S_k(p))$$

其中, $\hat{S}_k(p)$ 表示真实值, $S_k(p)$ 表示预测值。

4 实验设计与结果分析

4.1 实验数据集及评价指标

S3DIS数据集包括三个不同建筑中的5个大型室内区域,每个区域的面积约为1 900、450、1 700、870和1 100 m²(总共6 020 m²)。这些区域在以下

方面表现出不同的建筑风格和外观属性,主要包括办公区域、教育和展览空间,以及会议室、个人办公室、卫生间、开放空间、大堂、楼梯和走廊,同时包含了13类室内场景对象。

采用的评价指标平均交并比(Mean Intersection Over Union, MIOU)为语义分割的标准度量其值为 f_{MIOU} 表示。计算两个集合的交集和并集之比,在语义分割的问题中,这两个集合为真实值(Ground Truth)和预测值(Predicted Segmentation)。这个比例可以变形为正真数(Intersection)比真正、假负、假正(并集)之和。在每个类上计算IOU(Intersection Over Union),之后平均。

$$f_{MIOU} = \frac{1}{k+1} \sum_{i=0}^k \frac{P_{ii}}{\sum_{j=0}^k P_{ij} + \sum_{j=0}^k P_{ji} - P_{ii}}$$

4.2 网络参数设置及实验方法

本文的算法实验系统为Linux 16.04,使用的GPU是NVIDIA GTX 1080Ti,运算平台为CUDA-Toolkit 9.0,深度学习框架为Pytorch,版本号为1.1.0。在模型训练过程中,使用一个具有动量的梯度下降优化器,以0.95的动量和0.01的初始学习速率,去最小化逐点交叉熵损失。学习速率呈指数下降。将weight_decay设置为0.000 1, batch_size设置为4。

该网络采用的骨干网络是基于KPCConv的算法框架,除了对基本网络参数的设置,还要对其他参数进行一些设置。由于三维场景点云数量巨大,所选取的数据集是室内场景,在随机采样时采样半径为2 m的球体,并用高斯噪声^[22-23]、随机缩放和随机旋转来增强。输入的点云被下采样,设置网格大小为4 cm。在训练过程中,设置网络最多需要120个epoch才能收敛,对于分割任务可以生成输入球形领域的任意个数,若一个epoch为500次优化,相当于网络可以看到2 000个领域球。

4.3 实验结果与分析

4.3.1 S3DIS数据集区域5测试结果

图7-图10是S3DIS数据集区域5的语义分割可视化结果,选取了3个场景对语义分割结果进行了对比,并与原始场景和Ground Truth语义分割结果进行了对照。

将Rigid KPCConv算法实验结果与本文算法实验结果对比,能够清晰看到门、墙以及杂物的分割

更为精准,同时也减少了误分割、分割不准确等问题。例如图 7(a)中,算法对门的分割更为清楚,有明显的轮廓;图 7(b)中,算法对墙上的壁画和周

围物体的边界有着较为良好的分割性能;图 7(c)中,算法对墙、门以及书架有了更为清晰的认知能力,同时对小物体也有较好的分割性能。



图 7 原始点云数据

Fig. 7 Cloud data of original points

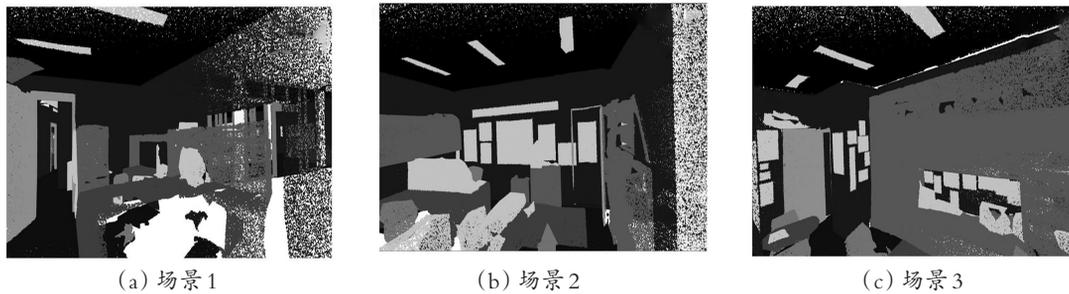


图 8 真实值

Fig. 8 True values

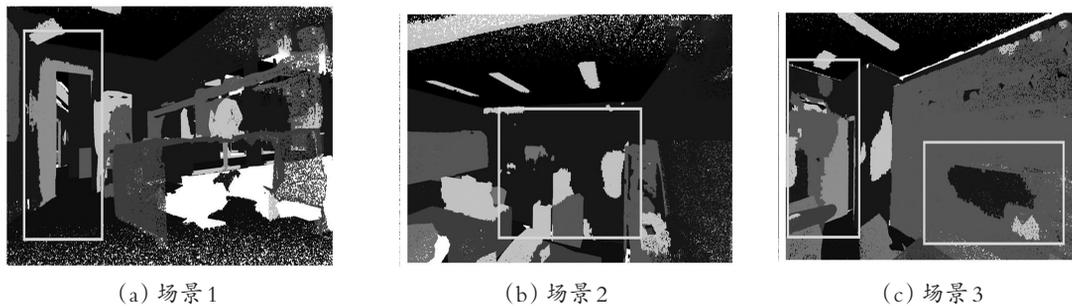


图 9 Rigid KPConv

Fig. 9 Rigid KPConv

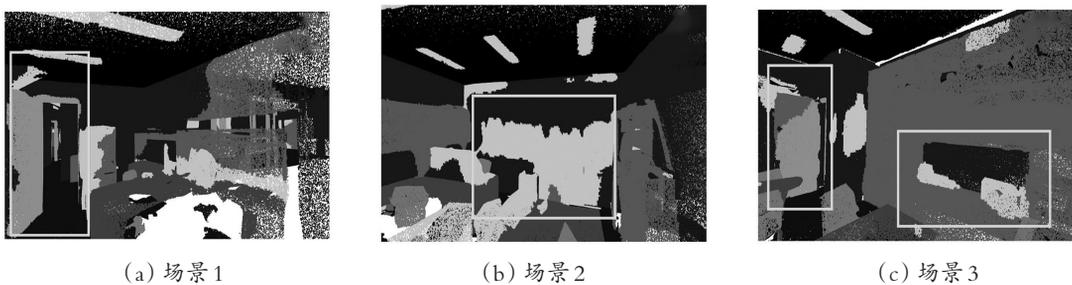


图 10 本文算法

Fig. 10 Algorithms used in this paper

表 1 是算法在 S3DIS 数据集区域 5 内测试的结果。实验结果表明与 Rigid KPConv 算法相比,本文

算法在对柱状物体、门、窗户等物体有较为良好的分割结果。例如在表 1 中柱状物体的 f_{MIou} 提高了

6%,门的 f_{MIOU} 提高了 11%,对窗户的 f_{MIOU} 提高了 23%,有着较为良好的分割性能。但仍对椅子、书架这类物体分割性能有稍许欠缺,可能与物体的复杂性有关。

由图 7、表 1 可知融合了语义边缘检测的语义分割网络具有更好的分割性能,对物体的边缘信息能够更好地完善。在 S3DIS 数据集区域 5 上的实验结果表明,本文算法较卷积核不变的 KPConv 算法在语义分割结果精度上提高了 4.81%。

4.3.2 消融实验

为了进一步分析网络中各因素带来的影响,通过消融实验研究加入不同模块对测试结果的影响。实验在数据集 S3DIS 区域 5 内进行。

表 1 S3DIS 数据集区域 5 上语义分割的详细 MIOU 分数

Table 1 Detailed MIOU scores of semantic segmentation in region 5 of S3DIS dataset

Method	Rigid KPConv	本文算法
Ceiling	92.60	92.64
Floor	97.30	97.77
Wall	81.40	82.39
Beam	0.00	0.00
Column	16.50	22.54
Window	54.50	77.76
Door	69.50	80.24
Chair	90.10	89.73
Table	80.20	81.97
BookCase	74.60	73.58
Sofa	66.40	68.68
Board	63.70	70.93
Clutter	58.10	74.50
$f_{\text{MIOU}}/\%$	65.40	70.21

如表 2 所示,测试结果中只考虑语义分割模块得到的初步语义分割结果分割精度为 64.43%,在本文中语义分割网络采用的是与基于 KPConv 算法相同的网络架构,得到的分割精度与 Rigid KPConv 算法得出的分割精度仍有差距。对于加入语义边缘检测任务中来看,在没有引入注意力机制网络的情况下,产生了具有语义信息的边缘轮廓与语义分割模块得到的结果融合细化,输出的最终语义分割结果的分割精度为 68.90%, f_{MIOU} 提高了 4.47%;在引入注意力机制网络后增强了特征信息的同时抑制了非边缘的语义信息的产生,改进了语义边缘检测的性能,得到最终语义分割实验结果的分割精度为

70.21%, f_{MIOU} 提高了 1.31%,总体提高了 5.78%。这表明了语义边缘检测模块及注意力机制网络对语义分割精度的提高是有效的。

表 2 不同模块的有效性对比

Table 2 Comparison of effectiveness of different modules

Method	未加语义	加入语义	加入语义
	边缘检测	边缘检测	边缘检测
	模块	模块(未引入	模块(引入
		注意力机制)	注意力机制)
Ceiling	94.23	92.54	92.64
Floor	96.61	97.76	97.77
Wall	80.67	81.69	82.39
Beam	0.00	0.00	0.00
Column	12.86	22.10	22.54
Window	58.39	76.68	77.76
Door	61.34	77.45	80.24
Chair	88.79	88.75	89.73
Table	82.15	81.35	81.97
BookCase	74.75	72.84	73.58
Sofa	67.31	66.05	68.68
Board	60.59	76.11	70.93
Clutter	59.54	62.38	74.50
$f_{\text{MIOU}}/\%$	64.43	68.90	70.21

5 结束语

提出将语义分割网络及语义边缘检测网络进行联合学习,是一种双流 FCN 结构。语义分割网络是对点云数据进行初步提取区域内的全局语义特征信息。在语义边缘检测网络中引入了注意力机制,对点云数据中物体的边缘进行特征提取增强,抑制了非边缘信息的产生,给边缘信息赋予了丰富的语义特征信息。然后将两路网络得到的语义特征信息进行融合处理,同时对融合结果进行输出细化。此外,使用了双重语义损失函数,使网络产生更好的语义边界,以此得出最终的语义分割结果。实验结果表明,融合了语义边缘检测的语义分割网络对物体的边界分割效果更加清晰明确,改善了边缘信息模糊的问题,提高了总体分割的精度。在数据集 S3DIS 上验证了该网络的有效性,能够得到较为理想的分割效果。

参考文献(References):

- [1] 李勇,佟国峰,杨景超,等. 三维点云场景数据获取及其场景理解关键技术综述[J]. 激光与光电子学进展, 2019,56(4):21—34.
LI Yong, TONG Guo-feng, YANG Jing-chao, et al. Overview of key technologies of 3D point cloud scene data acquisition understanding [J]. Laser & Optoelectronics Progress, 2019, 56(4): 21—34.
- [2] 邓瑞,林金朝,杨宏志. 基于深度学习的建筑物识别[J]. 重庆工商大学学报(自然科学版), 2019, 36(4):17—22.
DENG Rui, LIN Jin-chao, YANG Hong-zhi. Building recognition based on deep learning[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2019, 36(4): 17—22.
- [3] 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.
JING Zhuang-wei, GUAN Hai-yan, ZANG Yu-fu, et al. A review on semantic segmentation of point cloud based on deep learning [J]. Journal of Frontiers of Computer Science and Technology, 2021, 15(1): 1—26.
- [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, 2017: 652—660.
- [5] LIAN Y, FENG T, ZHOU J. A dense pointnet++ architecture for 3D point cloud semantic segmentation [C]//IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium, 2019: 5061—5064.
- [6] 柴玉晶,马杰,刘红. 用于点云语义分割的深度图注意力卷积网络[J]. 激光与光电子学进展, 2021, 58(12):1210016.
CHAI Yu-jin, MA Jie, LIU Hong. A depth map convolutional network for point cloud semantic segmentation [J]. Laser & Optoelectronics Progress, 2021, 58(12): 1210016.
- [7] HUANG Q, WANG W, NEUMANN U. Recurrent slice networks for 3D segmentation of point clouds [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 2626—2635.
- [8] YE X, LI J, HUANG H, et al. 3D recurrent neural networks with context fusion for point cloud semantic segmentation [C]//Proceedings of the European Conference on Computer Vision (ECCV), 2018: 403—417.
- [9] TAKIKAWA T, ACUNA D, JAMPANI V, et al. Gated-SCNN: gated shape CNNs for semantic segmentation [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 5229—5238.
- [10] HU Z, ZHEN M, BAI X, et al. JSENet: joint semantic segmentation and edge detection net-work for 3D point clouds [C]//European Conference on Computer Vision. Cham: Springer, 2020: 222—239.
- [11] YU Z, FENG C, LIU M Y, et al. Casenet: deep category-aware semantic edge detection [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 5964—5973.
- [12] XIE S, TU Z. Holistically-nested edge detection [C]//Proceedings of the IEEE international conference on computer vision. 2015: 1395—1403.
- [13] YU Z, LIU W, ZOU Y, et al. Simultaneous edge alignment and learning [C]//Proceedings of the European Conference on Computer Vision (ECCV), 2018: 388—404.
- [14] MA W, GONG C, XU S, et al. Multi-scale spatial context-based semantic edge detection [J]. Information Fusion, 2020, 64: 238—251.
- [15] LI X, LI X, ZHANG L, et al. Improving semantic segmentation via decoupled body and edge supervision [C]//European Conference on Computer Vision. Springer, Cham, 2020: 435—452.
- [16] ZHEN M, WANG J, ZHOU L, et al. Joint semantic segmentation and boundary detection using iterative pyramid contexts [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 13666—13675.
- [17] THOMAS H, QI C R, DESCHAUD J E, et al. Kpconv: flexible and deformable convolution for point clouds [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 6411—6420.
- [18] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770—778.
- [19] 罗华美. 基于最大池化层参数的优化模型在引力波天文学中的应用[J]. 重庆工商大学学报(自然科学版), 2020, 37(1): 59—64.
LUO Hua-mei. Application of optimal model based on maximum pooling layer parameters in gravitational wave

- astronomy [J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2020, 37(1):59—64.
- [20] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018:7132—7141.
- [21] LU C, CHEN F, SU X, et al. MS-AFF: a novel semantic segmentation approach for buried object based on multi-scale attentional feature fusion[J]. Optical and Quantum Electronics, 2021, 53(6):1—17.
- [22] LUO S, HU W. Score-based point cloud denoising[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021:4583—4592.
- [23] FU K, LIU S, LUO X, et al. Robust point cloud registration framework based on deep graph matching [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021:8893—8902.

Research on 3D Point Cloud Semantic Segmentation Method Fused with Edge Detection

CHEN Ling, XU Gang, FU Na-na, HU Zhi-feng, ZHENG Shu-zhan

(Key Laboratory of Detection Technology and Energy Saving Devices of Anhui Province, Anhui University of Technology, Anhui Wuhu 241000, China)

Abstract: Aiming at the problems of unclear segmentation targets, unclear edges, and ineffective fusion of global features and edge features in point cloud segmentation, a 3D point cloud semantic segmentation algorithm fused with edge detection was proposed. First, the global semantic features in the region are initially extracted from the point cloud data through the 3D point cloud semantic segmentation network. Then, the semantic edge detection network with the introduction of an attention mechanism is adopted, which can better extract and enhance the features of objects in the point cloud data, suppress the generation of non-edge information, and obtain edge features with rich semantic information. Finally, the semantic features belonging to the same object are fused by the fusion module for segmentation and refinement processing, which makes the segmentation target more accurate. In addition, dual semantic loss functions are used, which enables the network to produce semantic segmentation results with better boundaries. By building an experimental platform and using the S3DIS standard data set to test, the average intersection ratio of the improved algorithm on the data set is 70.21%, which is better than the KPConv semantic segmentation algorithm in precision. The experimental results show that the algorithm can effectively solve the problems of unclear boundary segmentation and fuzzy edge information, and the overall segmentation performance is good.

Key words: 3D point cloud; semantic segmentation; semantic edge detection; feature fusion

责任编辑:罗姗姗

引用本文/Cite this paper:

陈玲,许钢,伏娜娜,等.融合边缘检测的3D点云语义分割方法研究[J].重庆工商大学学报(自然科学版),2022,39(5):1—9.

CHEN Ling, XU Gang, FU Na-na, et al. Research on 3D point cloud semantic segmentation method fused with edge detection [J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2022, 39(5):1—9.