

基于动态图卷积网络的点云分类和分割网络

王江安*, 何娇, 庞大为

长安大学信息工程学院, 陕西 西安 710064

摘要 点云分类和分割是三维场景理解中的关键步骤。针对稀疏点云输入和遮挡不能有效识别点云的问题, 提出一种改进型分类和分割网络 Linked-DGCNN。在动态图卷积网络(DGCNN)的基础上增加 EdgeConv 卷积层数以提取深层次点云特征; 去除 DGCNN 的转换网络以简化网络结构; 引入深度残差网络的思想连接不同网络层的输出特征, 形成点云特征, 同时使网络训练更加稳定。基于 ModelNet40 和 ShapeNet Parts 数据集将该网络与其他点云网络进行对比实验, 实验结果表明, 该网络在稀疏点云输入和遮挡情况下, 相比其他方法有较高的点云分类和分割精度, 由此说明该网络具有较强的鲁棒性。

关键词 机器视觉; 深度学习; 点云分类与分割; 图卷积神经网络; 深度残差网络

中图分类号 TP391.41

文献标志码 A

doi: 10.3788/LOP202158.1215008

Point Cloud Classification and Segmentation Network Based on Dynamic Graph Convolutional Network

Wang Jiang'an*, He Jiao, Pang Dawei

School of Information Engineering, Chang'an University, Xi'an, Shaanxi 710064, China

Abstract Point cloud classification and segmentation are key steps in understanding three-dimensional (3D) scenes. Aiming at the problem that sparse point cloud input and occlusion cannot effectively identify point clouds, an improved classification and segmentation network linked-dynamic graph convolutional neural network (DGCNN) is proposed. First, the deep-level point cloud features were extracted by increasing the number of EdgeConv convolutional layers based on DGCNN. Next, the transformation networks of DGCNN were removed to simplify the network structure. Finally, the idea of introducing a deep residual network was used to link the output features of different network layers to form point cloud features, making the network training more stable. The proposed network was compared with other point cloud networks on ModelNet40 and ShapeNet Parts datasets. The experimental results show that the network has higher accuracy of point cloud classification and segmentation than other methods under the sparse point cloud input and occlusion. Besides, it has stronger robustness.

Key words machine vision; deep learning; point cloud classification and segmentation; graph convolutional neural network; deep residual network

OCIS codes 150.1135; 280.3640; 150.6910

1 引言

在深度学习时代, 深度神经网络几乎已经被应用到所有计算机视觉任务中^[1]。其中利用深度神经

网络处理无序排列的数据点来产生语义上有意义的结果, 是一种新兴的深度学习领域, 称为几何深度学习^[2]。这类无规则排列的数据点即三维(3D)点云, 是通过传感器对三维物体表面采样形成的。点云是

收稿日期: 2020-09-14; 修回日期: 2020-10-19; 录用日期: 2020-12-02

基金项目: 国家自然科学基金面上项目(61771075)、陕西省自然科学基金(2017JQ6048)、中央高校复杂城市环境下 GPS 非视距多径智能实时抑制方法研究(310824161009)

*E-mail: wangjiangan@126.com

最常见的三维几何数据表示形式。因此,有效识别点云对三维场景理解、自动驾驶等领域的快速发展具有重要意义。

为了有效识别点云,早期的深度学习技术将 3D 点云投影到多个视图或体素等网格结构上,然后在投影上使用卷积神经网络(CNN)。Su 等^[3]提出的多视图卷积神经网络(MVCNN)是从点云中抽取 80 个视图,并将其投影到二维(2D)网格,再由 2D CNN 处理。通过池化层将从不同图像学习到的特征聚合为全局特征,并利用该全局特征对物体进行分类。Maturana 等^[4]提出的 VoxNet 将 3D 点投影到 $32 \times 32 \times 32$ 的 3D 网格上,该网格由 3D CNN 处理。但基于多个视图或体素的数据存储和处理,计算量较大,且在平面或 3D 网格上投影点云数据会导致重要信息丢失,因此用深度学习技术直接处理点云数据已成为当前的研究热点。

近些年,斯坦福大学的 Charles 等^[5]开创了“端到端”处理点云的方法,提出了 PointNet 模型,并假定学习 N 个数据点的对称函数可以对 $N!$ 种点云输入点的排列具有置换不变性。PointNet 具有独创性,可以直接对点云进行分类和分割。但是它只能单独处理每个点,并未考虑到点云中每个点与其邻域之间的关系。由于点云可以旋转或平移,因此相邻点之间的局部特征可能会比独立的点坐标更具有鲁棒性。为了解决上述问题。PointNet 的研究者又在 PointNet++^[6]中引入了局部信息,采用最远点采样(FPS)算法统一采样点云,构建局部区域,并利用基于半径的球查询方法对它们进行分组,确定局部区域。然而,PointNet++ 仍然只单独处理局部点集中的每个点,并未考虑某点及其邻域之间如距离和边缘向量等的依赖关系。

自从 PointNet 家族体系结构问世以来,研究者们不断利用现有技术构建网络结构,以提高点云网络的性能。Li 等^[7]提出的 SO-Net 使用自组织映射(SOM)网络进行分层特征提取。Fan 等^[8]提出的 SpiderCNN 和山东大学的 Li 等^[9]提出的 PointCNN 都是基于局部几何信息定义点云上的卷积运算,而不使用对称函数处理点云。上述网络都使用局部点捕获局部信息,但并未获取到点与点之间的关系。Wang 等^[10]提出的动态图卷积神经网络(DGCNN)对输入点云上的每一个点计算相邻点之间的边缘特征,从而得到某点与其邻域之间的依赖关系。该网络根据每层的输入动态计算 k 最近邻图,因此称为动态网络。该模型较好地捕获了点的

局部特征,当输入点云仅有 1024 个点时,DGCNN 在 ModelNet40 数据集上实现了较高的分类精度。因此,DGCNN 似乎更实用,但还有一个明显的缺陷,即在稀疏点云输入和遮挡情况下无法有效识别点云。

为了解决上述问题,本文结合 PointNet 和 DGCNN 的思想,对 DGCNN 的网络架构进行优化设计,提出一种新的点云分类和分割网络——Linked-DGCNN。首先在 DGCNN 的基础上增加 EdgeConv 层数,然后引入残差网络的思想连接不同网络层的输出特征,形成最终的点云特征,最后使用该特征进行点云分类与分割。该网络解决了诸如稀疏点输入和遮挡等问题,可实现点云的准确分类和分割,为三维场景理解提供了更加可靠的信息。

2 基本原理

2.1 问题描述

本文设计了一种以无序点集为输入的点云分类分割网络,将点云表示为一组点集 $\mathbf{P} = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n\} \subseteq \mathbb{R}^D$,其中每个点 \mathbf{p}_i 包含其坐标 (x_i, y_i, z_i) 以及其他一些特征,如颜色、法向量等。为简洁起见,本文仅将点 \mathbf{p}_i 的坐标信息 (x_i, y_i, z_i) 作为网络的输入。在深度神经网络结构中,每一网络层都是在前一层的输出特征上进行操作,因此, D 通常表示某一网络层的特征维度。本文所提网络将进行两个点云处理任务:点云分类和分割。对于点云分类,需要对整个点云进行分类。因此,需要找到一个分类函数 f_c ,并将输入点云转换为每个类别上的概率分布 $\mathbf{P}_c; \mathbf{P}_c = f_c(\mathbf{P})$ 。对于点云分割任务,点云中的每点 \mathbf{p}_i 都被指定为特定类别。因此,需要找到一个分割函数 f_s 来计算每个点 \mathbf{p}_i 在每个类别上的概率分布: $\mathbf{P}_{c_i} = f_s(\mathbf{p}_i)$ 。

2.2 图卷积神经网络

由于图卷积神经网络不需考虑节点的输入次序,而是在每个节点上独立传播,通过提取节点间的边特征来学习两个节点间的依赖关系^[11],因此是最适用于点云处理的方法。若要将图卷积神经网络应用于点云处理,需先将其转换为有向图。图 \mathbf{G} 由顶点 V 和边 E 组成: $\mathbf{G} = (V, E), V \subseteq \mathbb{R}^D, E \subseteq V \times V$ 。但由于计算机存储容量有限,本文算法不为点云构造全连接的边。一种最简单的方法是利用 K 近邻分类(KNN)算法^[12]构造一个局部有向图 \mathbf{G} ,如图 1 所示。在局部有向图 \mathbf{G} 中,每个点 \mathbf{p}_i 都是一个中心节点,并计算中心节点与其 K 个相邻点之间的边 \mathbf{e}_i :

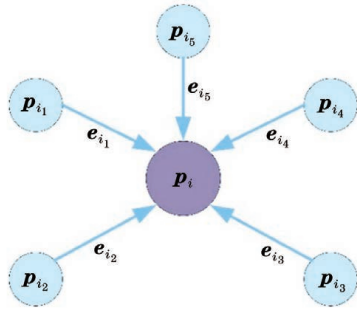


图 1 点云局部有向图

Fig. 1 Local digraph of point cloud

$$\begin{cases} \mathbf{G} = (\mathbf{V}, \mathbf{E}) \\ \mathbf{V} = \{p_i | i = 1, 2, \dots, n\} \\ \mathbf{E} = \{e_i = (e_{i_1}, e_{i_2}, \dots, e_{i_K}) | i = 1, 2, \dots, n\} \\ e_{i_j} = p_{i_j} - p_i \end{cases} \quad (1)$$

其中 p_{i_j} 和 e_{i_j} 分别为中心点 p_i 的一个相邻节点以及从 p_{i_j} 到中心点 p_i 的有向边。

本文算法与 DGCNN 结构相似, 在每个卷积层前应用 KNN 算法, 在欧氏空间和特征空间中分别构造局部图。因此, 点 p_i 也表示特征空间中的中心点, 且点 p_{i_j} 为特征空间中心点 p_i 的相邻节点。

2.3 动态图特征提取

构造局部有向图后, 基于 EdgeConv 层提取局部图特征。特征提取函数 f_c 对于所有点都是相同的, 因此以其中一个中心点 p_i 及其 K 个相邻节点为例来表示 EdgeConv 卷积运算。EdgeConv 网络层以中心点 p_i 的局部图作为输入, 输出是点云局部特征 l_i , 其表达式为:

$$l_i = f_c[G(p_i, e_i)] = \max\{h(p_i, e_{i_1}), h(p_i, e_{i_2}), \dots, h(p_i, e_{i_K})\}, \quad (2)$$

其中 $h(p_i, e_{i_j})$ 是中心点 p_i 和一个边缘向量 e_{i_j} 的隐含特征向量。然后将对称聚合函数, 例如求和或最大池化, 应用于与每个顶点发出的所有相互关联的边缘特征。由于不受相邻节点次序的影响, 并且需要提取所有边特征向量中最主要的特征, 本文使用最大池化操作。隐含特征向量 $h(p_i, e_{i_j})$ 通过多层感知器 (MLP) 提取^[13], 即

$$h(p_i, e_{i_j}) = [h_1(p_i, e_{i_j}), h_2(p_i, e_{i_j}), \dots, h_{c'}(p_i, e_{i_j})], \quad (3)$$

$$h_{c'}(p_i, e_{i_j}) = \sum_{c=1}^c \omega_{c'} p_{ic} + \omega_{c'(c+C)}(p_{i_j c} - p_{ic}) + b_{c'}, \quad (4)$$

式中: p_{ic} 和 $p_{i_j c}$ 为通道 c 中的中心点 p_i 及其相邻

节点 p_{i_j} 的值; c 和 c' 为输入点 p_i 和输出隐含特征向量 $h(p_i, e_{i_j})$ 的通道数; $h_{c'}(p_i, e_{i_j})$ 表示隐含特征向量的通道 c' 中的值; 且 $\omega_{c'c}$, $\omega_{c'(c+C)}$ 和 $b_{c'}$ 是 MLP 的可训练参数。

本文构造的局部有向图不是固定的, 而是在网络的每一层之后动态更新, 即中心点 p_i 的 K 最近邻点集在网络中逐层变化。动态更新局部有向图, 使点云具有介于平移不变性和非局部性之间的性质^[14]。在稀疏点云输入和遮挡情况下, 局部感受野与点云的直径相同, 因此不会影响点云的分类和分割结果。

2.4 Linked-DGCNN 网络结构设计

为有效提高点云分类和分割精度, 本文在 DGCNN 的基础上增加 EdgeConv 卷积层数以提取深层特征, 并引入残差网络的思想连接不同网络层的输出特征以形成点云特征。因此本文将该网络命名为 Linked-DGCNN, 其网络结构如图 2 所示。

分类网络将 n 个点坐标作为输入, 计算 EdgeConv 层上每个点的 k 个大小的边缘特征集, 并聚合每个集合中的特征以计算相应点的 EdgeConv 响应。最后一个 EdgeConv 层的输出特性被全局池化, 形成一个 1 维全局特征, 用于生成 m 类分类分数。分割网络是分类网络的一种扩展, 包含 s 个标签的点云模型, 且每个模型均以 n 个点作为输入, 通过连接 1 维全局特征和每个点所有 EdgeConv 层输出的局部特征形成点云特征, 并用 3 个全连接层 (256, 256, 128) 生成每点语义标签的分类分数。

本文网络与 DGCNN 结构相似, 但本文提出的 Linked-DGCNN 和 DGCNN 存在一些差异:

1) 引入残差网络思想连接不同网络层的输出特征以增强网络

通过加深网络层数可以强化深度神经网络的特征。因此, 为了获得更高层次的特征, 本文使用更深层次的网络结构。但是在使用深度网络结构时, 常会遇到两个问题: 网络的退化和网络收敛的延迟。为了解决简单增加网络层数引起的网络性能下降的问题, 本文引入残差网络的思想。深度残差网络是由微软亚洲研究院的 He 等^[15]提出的一种跳跃连接的卷积神经网络结构, 并在 2015 年 ImageNet 大规模视觉识别挑战赛上获得了图像分类和目标检测组冠军。残差网络使得网络具有较强的恒等映射能力, 不但使网络达到了较深的层次, 同时提升了网络性能。

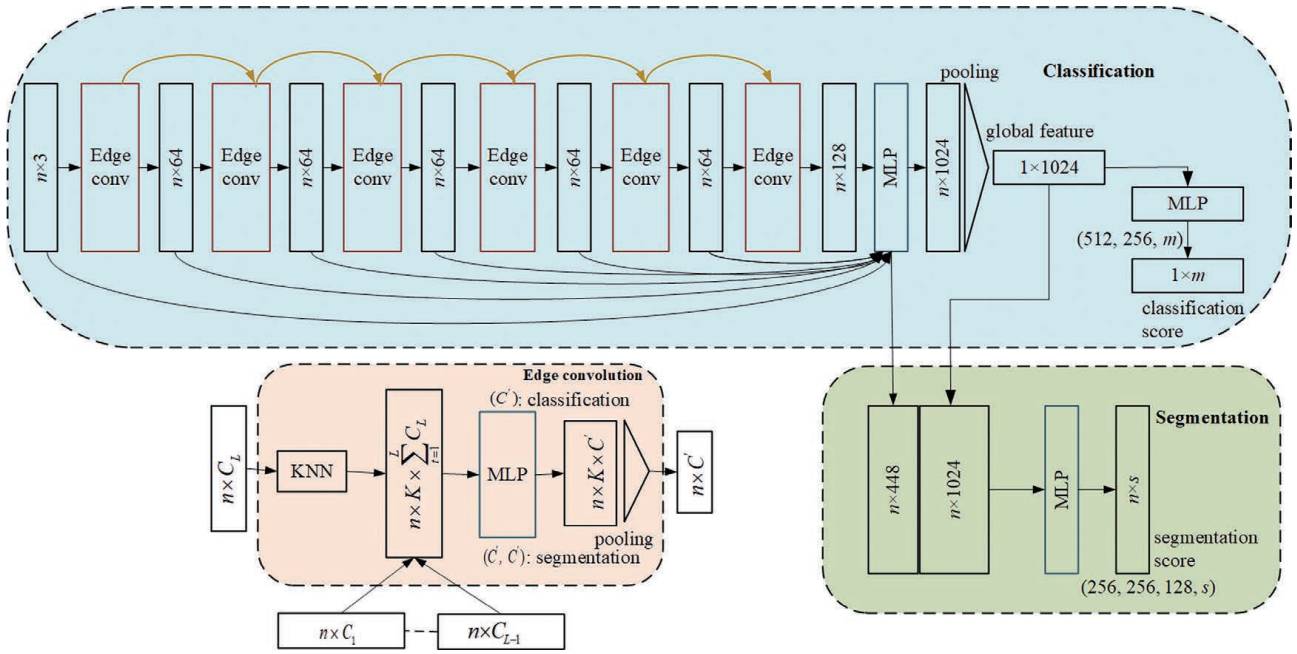


图 2 Linked-DGCNN 网络结构

Fig. 2 Linked-DGCNN network structure

2) 去除输入转换网络以简化网络结构

PointNet^[5] 和 DGCNN^[10] 都依赖输入转换网络来估计点云的旋转矩阵, 最终实现点云偏差补偿。但是, 转换网络使网络的大小增加了一倍。但实验结果表明, 该网络在去除输入转换网络后仍有较好的性能。

转换网络的输出是一个 3×3 矩阵 R , 它可以弥补点云旋转或平移带来的偏差:

$$P_{\text{offset}} = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots \\ x_n & y_n & z_n \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}, \quad (5)$$

其中 r_{ij} 是第 i 行和第 j 列上旋转矩阵 R 的值。

若将 MLP 转换成矩阵形式, 可以得到隐含特征向量:

$$h(P) = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots \\ x_n & y_n & z_n \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1c'} \\ w_{21} & w_{22} & \cdots & w_{2c'} \\ w_{31} & w_{32} & \cdots & w_{3c'} \end{bmatrix} + [b_1 \ b_2 \ \cdots \ b_{c'}]. \quad (6)$$

对比(4)式与(5)式, 可以发现它们是相似的, 区别在于: 变换网络将每个点云转换为一个特定的矩阵, 而 MLP 对于所有点云都是固定的。不同通道的 MLP 参数能以不同的角度旋转点云, 也可以通过不同的平移向量来移动点云。考虑到本文网络至少有

64 个输出通道, MLP 可以沿至少 64 个方向计算点云特征, 故网络可近似为旋转不变。因此, 即使去掉了输入转换网络, 点云分类与分割的精度同样不会受到影响。

3 实验与结果分析

3.1 实验数据集

对于点云分类任务, 使用 ModelNet40 数据集^[16] 进行实验。ModelNet40 包含 40 个对象类别的 12311 个 CAD 模型。该数据集被分为 9843 个模型的训练集和 2468 个模型的测试集。本文以相同的数据集划分比例进行实验, 并在测试集上显示分类结果。对于点云分割任务, 使用 ShapeNet Parts 数据集^[17] 进行点云分割实验。该数据集包含 16 个对象类别的 16881 个 3D 模型和 50 个部件分割类别标签。给定 3D 点云, 对点云部件进行语义分割, 如对飞机而言, 需要分割机翼、机身等部件。

3.2 网络参数设置和预处理

本文所提网络 Linked-DGCNN 的训练和测试实验参数设置如下: 学习率为 0.001, 迭代次数为 250, batch size 为 32, 且使用 Adam 优化器^[18]。配置显卡为 NVIDIA GTX 1060, 实验软件环境为 Ubuntu 16.04.5 系统、python 3.7、Tensorflow-GPU 1.13.1、CUDA 10.0。

本文采用与其 DGCNN 实验相同的预处理方法。在输入训练网络之前, 首先, 对每个 3D 模型,

从网格面均匀采样 1024 个点,并将这些点归一化到一个单位球体。在训练期间,使用数据增强策略,对输入随机旋转、缩放以及逐点抖动,以扰乱点的位置,增加数据量。

3.3 点云分类实验结果分析

为了验证本文算法的优越性,在 ModelNet40 数据集上与最新典型点云分类网络进行了对比实验,结果如表 1 所示。本文算法的总体分类精度和平均分类精度两个指标都达到了最好的效果。

表 1 ModelNet40 数据集点云分类结果

Table 1 Results of dataset point cloud classification of ModelNet40

Method	Input(size)	Model size /MB	Forward time /ms	Mean class accuracy /%	Accuracy overall /%
3D ShapeNets	Voxels(1)	—	180.2	77.5	84.5
VoxNet	Voxels(12)	—	87.0	82.8	85.5
Subvolumes ^[17]	Voxels(20)	190.8	120.3	86.1	89.3
MVCNN	Views(80)	690	163.2	88.8	—
PointNet	Points(1024×3)	40	25.3	84.3	87.8
PointNet++	Points+normal(5000×6)	24	44.3	—	90.5
DGCNN	Points(1024×3)	21.2	98.03	89.0	91.0
Ours	Points(1024×3)	12.4	88.5	89.7	92.6

图 3 和图 4 分别给出了 Linked-DGCNN 在 ModelNet40 数据集上的分类训练精度曲线和训练损失曲线。从图中可以看出,在训练开始时,随着迭代次数的增加,识别精度不断提高,训练损失逐渐下降。由于在网络训练过程中,参数是不断优化的,网络学习到的点云特征描述不断增多,准确率逐渐增高。当迭代次数达到 150 次时,识别精度和训练误差处于动态平衡状态,网络趋于收敛趋势,这说明网络中的参数已经达到最优。

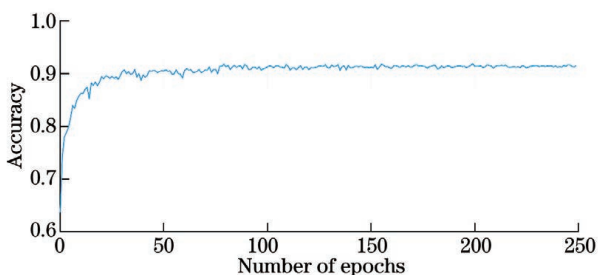


图 3 分类测试精度曲线

Fig. 3 Classification test accuracy curve

3.4 点云分割实验结果分析

为了评估本文网络架构在语义分割任务方面的性能,在 ShapeNet Parts 数据集上与最新点云分割

虽然本文仅用 1024 个点作为输入,但其实验效果明显优于使用 5000 个点作为输入的模型,这在很大程度上证明 Linked-DGCNN 具有提取点云深层次特征的能力。此外,利用模型大小和前向传播时间评估网络的复杂性,如表 1 所示,记录了在单个 GTX 1060GPU 上, batch size 为 32 进行网络训练的前向传播时间。实验结果表明,本文模型以极高的参数效率和良好的运行速度达到了较好的性能。

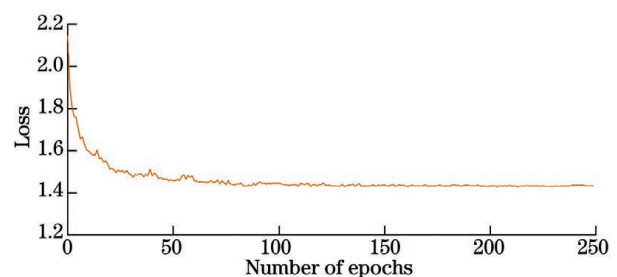


图 4 训练损失曲线

Fig. 4 Training loss curve

算法进行对比,部分实验结果如表 2 所示。分割任务的评估指标是特定类别中所有形状的并集的平均交集即平均交并比(mIoU)^[19]。

由表 2 可以看出,本文网络结构以 84.7% 的 mIoU 值,达到了比较好的点云分割性能。虽然这仅比此类网络的最高水平低 0.3%,本文网络结构没有获得最先进的结果,但在大部分物体上获得了接近该对象的最佳分割精度值。图 5 为本文算法在部分物体上 Linked-DGCNN 与 DGCNN 的点云分割可视化结果对比,其中 DGCNN_diff 和 Linked-DGCNN_diff 分别表示了预测结果与真实标签的错分结果。与 DGCNN 相比,本文算法的点云分割结

表 2 ShapeNet 数据集物体部件分割精度

Table 2 Segmentation accuracy of object components in ShapeNet dataset

unit: %

Method	mIoU	Bag	Cat	Chair	Guitar	Knife	Laptop	Lamp	Board	Table
PointNet	83.6	78.6	82.4	89.5	91.5	85.8	80.7	95.2	72.7	80.5
PointNet++	85.0	78.9	87.6	90.7	90.9	85.8	83.6	95.2	76.3	82.5
SO-Net	84.6	83.5	84.8	90.8	90.1	83.6	82.3	95.1	72.0	82.5
DGCNN	85.0	83.6	84.3	90.6	91.4	87.2	82.8	96.0	75.4	81.8
Ours	84.7	85.4	88.6	90.8	90.9	88.6	82.4	95.4	74.7	81.9

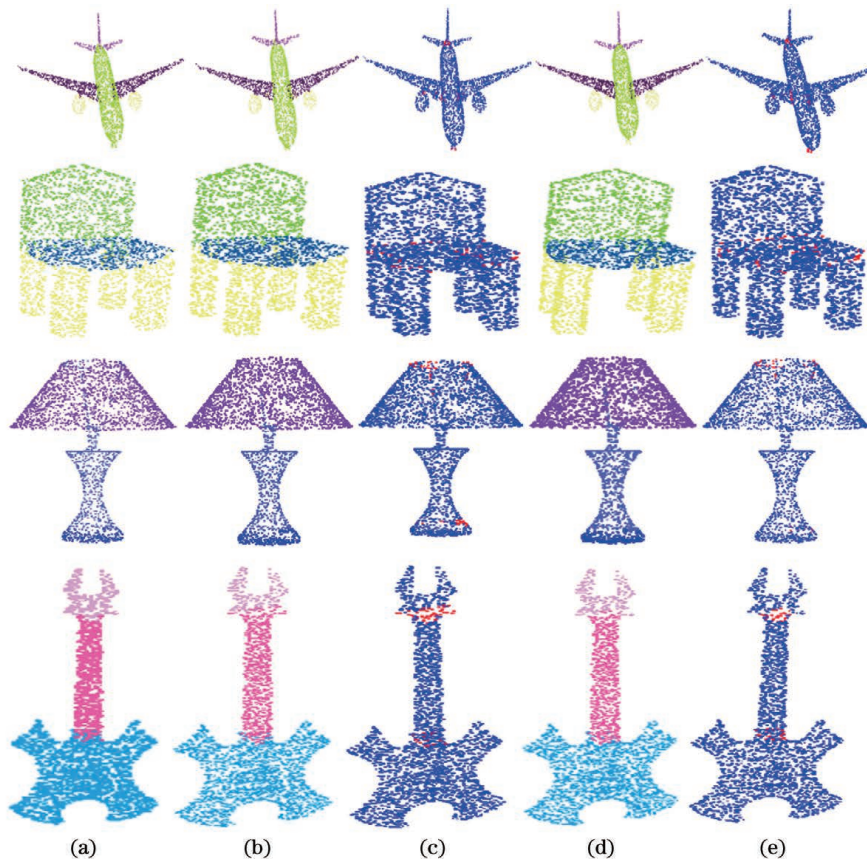


图 5 本文算法与 DGCNN 在 ShapeNet 数据集中对于物体部件分割结果对比。(a)真值图;(b)DGCNN;(c)DGCNN_diff;(d)Linked-DGCNN;(e)Linked-DGCNN_diff

Fig. 5 Comparison of the segmentation results of the algorithm in this paper and DGCNN for object components in the ShapeNet dataset. (a)Ground truth; (b)DGCNN; (c)DGCNN_diff; (d)Linked-DGCNN; (e)Linked-DGCNN_diff

果与真实标签结果相近,尤其是在模型边界处的分割准确率明显提升,进一步验证了本文算法的优越性。

3.5 网络结构设计分析

3.5.1 Linked-DGCNN 的 EdgeConv 卷积层数与 K 近邻数的选择分析

Linked-DGCNN 主要是通过增加 EdgeConv 卷积层数来提取点云深层次特征,因此 EdgeConv 卷积层数对点云分类和分割精度的提高有决定性作用。为了研究 EdgeConv 卷积层数对网络性能的影响,本文在 ModelNet40 数据集上进行对比实验。

在实验中,卷积层数选择 3~7,结果如图 6 所示。EdgeConv 的卷积层数为 6 时,分类精度最高,因此最终选择 6 层网络。

选取不同的最近邻数(K 为 5、10、20、40),以 ModelNet40 为测试对象进行对比实验,实验结果如图 7 所示。尽管没有对所有可能的 K 进行详尽的实验,但从实验结果可以发现,当 K 过大时,网络性能会急剧下降。由此可以得出结论:在一定的密度条件下,当 K 较大时,欧氏距离不能近似为测地距离^[20],故会破坏每个局部区域的几何结构,使分类精度大大下降。当 K=20 时,分类精度最高,因此

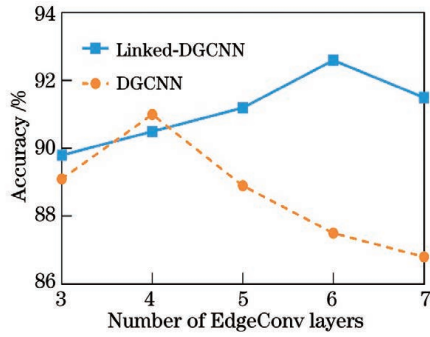


图 6 EdgeConv 卷积层数对精度的影响

Fig. 6 Influence of number of EdgeConv convolution layers on accuracy

选择合适的 K 值对实验结果有着至关重要的影响。

3.5.2 鲁棒性分析

为了验证该网络对于稀疏输入和遮挡问题的鲁棒性,选取不同的采样点个数(128、256、384、512、768、1024),使用与 3.3 节中讨论的实验参数运行了 75 个 epoch 的网络,与 DGCNN 进行对比实验,比

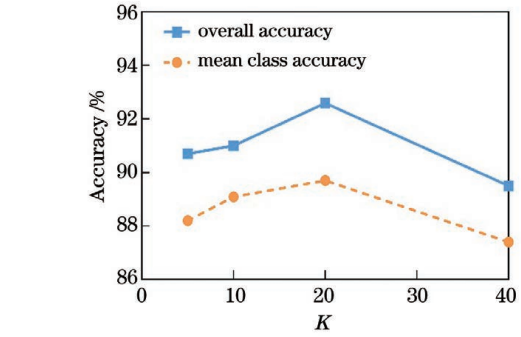
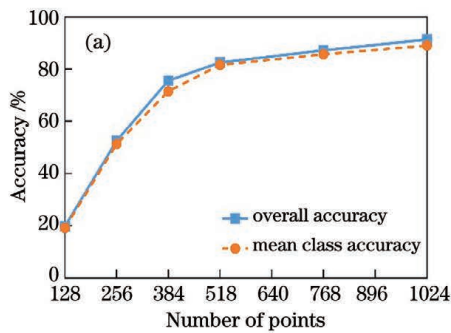


图 7 K 值对分类精度的影响

Fig. 7 Influence of K value on accuracy

较结果如图 8 所示。可以看出,随着采样点数目的增加,Linked-DGCNN 的分类精度逐渐上升且始终高于 DGCNN 的分类精度。此外,当点数低于 384 时,DGCNN 的性能明显下降。然而,即使输入到本文网络的点数下降到 128 点,网络的分类精度仍能达到 86.2%。这表明在解决稀疏点输入、遮挡等问题时,通过增加网络深度并引入残差网络可以使网络具有较强的鲁棒性。

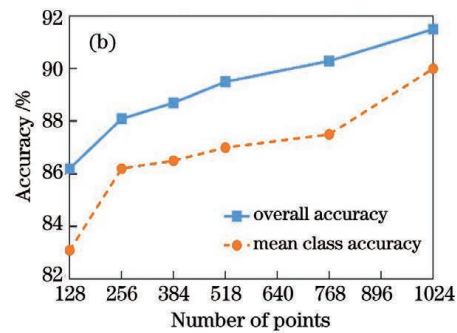


图 8 不同采样点个数对分类精度的影响。(a) DGCNN; (b) Linked-DGCNN

Fig. 8 Effect of number of different sampling points on classification accuracy. (a) DGCNN; (b) Linked-DGCNN

4 结 论

为有效提高点云分类和分割的精度,本文提出一种基于动态图卷积神经网络(DGCNN)的点云分类和分割网络: Linked-DGCNN。本文网络在 ModelNet40 分类数据集上达到了较高的分类与分割精度,在分割数据集上也具有极强的竞争力。通过一系列实验评估本文网络,实验结果表明,增加卷积层数并引入残差网络思想连接不同层次的特征,可以有效提高点云分类与分割精度,是一种行之有效的方法。与 DGCNN 相比,本文网络针对稀疏点输入和遮挡的问题表现出较强的鲁棒性。但本文网络的结构依然较为复杂,下一步的工作重心将是在保证精度的同时降低网络复杂度,设计出一个轻量级网络。

参 考 文 献

- [1] Zhao Z Y, Cheng Y L, Shi X S, et al. Terrain classification of LiDAR point cloud based on multi-scale features and PointNet[J]. Laser & Optoelectronics Progress, 2019, 56(5): 052804.
赵中阳, 程英蕾, 释小松, 等. 基于多尺度特征和 PointNet 的 LiDAR 点云地物分类方法[J]. 激光与光电子学进展, 2019, 56(5): 052804
- [2] Wang X J, Ma J, Wang N N, et al. Deep learning model for point cloud classification based on graph convolutional network [J]. Laser & Optoelectronics Progress, 2019, 56(21): 211004.
王旭娇, 马杰, 王楠楠, 等. 基于图卷积网络的深度学习点云分类模型[J]. 激光与光电子学进展, 2019, 56(21): 211004.
- [3] Su H, Maji S, Kalogerakis E, et al. Multi-view

- convolutional neural networks for 3D shape recognition [C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 945-953.
- [4] 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), September 28-October 2, 2015, Hamburg, Germany. New York: IEEE Press, 2015: 922-928.
- [5] Charles R Q, Hao S, Mo K C, et al. PointNet: deep learning on point sets for 3D classification and segmentation [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 77-85.
- [6] 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, December 4, 2017, Red Hook, NY, USA. Red Hook: Curran Associates Inc., 2017: 5105-5114.
- [7] Li J X, Chen B M, Lee G H, et al. SO-Net: self-organizing network for point cloud analysis [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 9397-9406.
- [8] Xu Y F, Fan T Q, Xu M Y, et al. SpiderCNN: deep learning on point sets with parameterized convolutional filters [M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11212: 90-105.
- [9] Li Y Y, Bu R, Sun M C, et al. PointCNN: convolution on X-transformed points [C]//Advances in Neural Information Processing Systems, December 3-8, 2018, Montréal, Canada: NIPS, 2018: 820-830.
- [10] Wang Y, Sun Y B, Liu Z W, et al. Dynamic graph CNN for learning on point clouds [J]. ACM Transactions on Graphics, 2019, 38(5): 1-12.
- [11] Wang B J, Nong L P, Zhang W H, et al. 3D point cloud classification and segmentation network based on Spider convolution [J]. Journal of Computer Applications, 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.
- [12] Shi X S, Cheng Y L, Zhao Z Y, et al. Point cloud classification algorithm based on IPTD and SVM [J]. Laser & Optoelectronics Progress, 2019, 56(16): 161002.
释小松, 程英蕾, 赵中阳, 等. 基于三角网滤波和支持向量机的点云分类算法 [J]. 激光与光电子学进展, 2019, 56(16): 161002.
- [13] Zhang J Y, Zhao X L, Chen Z, et al. Review of semantic segmentation of point cloud based on deep learning [J]. Laser & Optoelectronics Progress, 2020, 57(4): 040002.
张佳颖, 赵晓丽, 陈正, 等. 基于深度学习的点云语义分割综述 [J]. 激光与光电子学进展, 2020, 57(4): 040002.
- [14] Bai J, Si Q L, Qin F W, et al. Lightweight real-time point cloud classification network LightPointNet [J]. Journal of Computer-Aided Design & Computer Graphics, 2019, 31(4): 612-621.
白静, 司庆龙, 秦飞巍, 等. 轻量级实时点云分类网络 LightPointNet [J]. 计算机辅助设计与图形学学报, 2019, 31(4): 612-621.
- [15] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [16] Wu Z R, Song S R, Khosla A, et al. 3D ShapeNets: a deep representation for volumetric shapes [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 1912-1920.
- [17] Qi C R, Su H, Nießner M, et al. Volumetric and multi-view CNNs for object classification on 3D data [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 5648-5656.
- [18] Kingma D P, Ba J. Adam: a method for stochastic optimization [C]//3rd International Conference on Learning Representations, May 7-9, 2015, San Diego. Ithaca: arXiv.org, 2015: 13.
- [19] Yi L, Kim V G, Ceylan D, et al. A scalable active framework for region annotation in 3D shape collections [J]. ACM Transactions on Graphics, 2016, 35(6): 210.
- [20] Klokov R, Lempitsky V. Escape from cells: deep Kd-networks for the recognition of 3D point cloud models [C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 863-872.