

# 基于特征负反馈卷积的点云分析方法

邓林涛, 方志军\*

上海工程技术大学电子电气工程学院, 上海 201620

**摘要** 针对不规则性、无序性和稀疏性给点云分析带来的困难与挑战,提出了融合局部信息提取与全局特征推理的点云分析方法。首先,为了更加有效地进行局部点分组,使用结构感知K近邻(KNN)搜索局部邻域点。其次,基于边卷积改进提出一种特征负反馈卷积模块,在映射的高维空间中提取更为准确的局部特征。此外,设计了基于注意力机制的全局语义推理模块,通过强调不同区域的分组点来避免潜在的信息冗余,从而全面地获取点云特征。通过在公开的点云数据集 ModelNet40 和 ShapeNet 上进行测试,该方法总体分类精度和总体平均交并比分别达到 93.8% 和 86.4%,定量的评估指标以及定性的可视化实验证明了该方法的准确性和鲁棒性。

**关键词** 图像处理; 点云分析; 特征负反馈卷积; 注意力机制; 全局语义推理模块

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP202259.1210006

## Point Cloud Analysis Method Based on Feature Negative Feedback Convolution

Deng Lintao, Fang Zhijun\*

*School of Electronic and Electrical Engineering, Shanghai University of Engineering Science,  
Shanghai 201620, China*

**Abstract** Aiming at the difficulties and challenges caused by the irregularity, disorder and sparsity to point cloud analysis, a point cloud analysis method that combines local information extraction and global feature reasoning is proposed. First, in order to group local points more effectively, the structure-aware K nearest neighbor (KNN) is used to search for local neighborhood points. Secondly, a feature negative feedback convolution module is improved based on edge convolution to extract more accurate local features in the mapped high-dimensional space. In addition, a global semantic reasoning module based on the attention mechanism is designed to avoid potential information redundancy by emphasizing the grouping point of different regions, so as to obtain point cloud features more comprehensively. Through tested on the public point cloud data sets ModelNet40 and ShapeNet, the overall classification accuracy and overall mean intersection over union (mIoU) of the proposed method reach 93.8% and 86.4%, respectively. Quantitative evaluation indicators and qualitative visualization experiments prove the accuracy and robustness of the proposed method.

**Key words** image processing; point cloud analysis; feature negative feedback convolution; attention mechanism; global semantic reasoning module

## 1 引言

在自动驾驶<sup>[1-4]</sup>和机器人导航<sup>[5-6]</sup>等领域的推动

下,3D点云分析引起了广泛关注。然而点云的稀疏性、无序性、不规则性,给点云分析带来了挑战。对此,研究人员将卷积神经网络(CNN)在图像分析领

收稿日期: 2021-04-26; 修回日期: 2021-05-21; 录用日期: 2021-06-10

通信作者: \*zjfang@foxmail.com

域<sup>[7-8]</sup>的成功应用搬到点云处理中。文献[9-11]将点云转换为常规体素,文献[12-14]采用点云的多视图图像。但这些转换导致3D点云中几何信息的大量丢失并带来较高复杂性。为了直接处理点云,PointNet<sup>[15]</sup>独立地学习每个点并最终聚合所有特征,尽管这是一项卓有成效的工作,但它忽略了点云数据的局部结构。改进后的PointNet++<sup>[16]</sup>通过构建分层的神经网络以提取每个局部区域中的特征,但在局部区域中点之间的关系仍未引起足够重视。随后,动态图形卷积神经网络(DGCNN)<sup>[17]</sup>中提出了边卷积(EdgeConv)思想,它捕捉点云的局部几何特征以及排列不变性。在边卷积思想基础上,关系形状卷积神经网络(RS-CNN)<sup>[18]</sup>中通过使用球搜索(BQ)的邻近点来执行局部区域关系推断。这被证明是获取每个局部区域几何关系的有效方法。

与之前的工作不同,本文尝试在感知局部点云结构的基础之上,为每个中心点自适应地定义局部区域的密度分布。尽管需要少量人工干预和参数调整,但点邻域将更加合理。此外为了优化点云空间中基于CNN的特征学习,提出了一种特征负反馈模块,该模块在结合局部语义信息的基础上设计了误差反馈机制。这种反馈机制旨在通过输出的响应修改原始输入,从而引导视觉感知任务获得相关结果。早期的工作尝试在CNN中使用反馈连接来纠正误差:例如2D人体姿势估计<sup>[19]</sup>、图像超分辨率(SR)<sup>[20-21]</sup>。在3D中,点云上采样网络(PU-GAN)<sup>[22]</sup>中提出了误差反馈单元,该方法实现了针对点云生成的特征扩展。但是这种反馈结构尚未应用于点云特征学习领域。本文通过比较输入与经过网络学习后的特征,反馈两者之间的差异来合理补充输出特征图,利用这种结构强化网络表征能力。

关于3D点云的高级表达,注意力机制是广泛应用的策略。这种机制可以帮助网络将更多的权重放在有效的信息上<sup>[23]</sup>。在众多2D视觉问题例如图像分割<sup>[24-26]</sup>、图像去噪<sup>[27]</sup>等,注意力模块已得到应用。对于3D点云,文献[28-30]利用自注意的概念找到重要的点特征,而最近的工作<sup>[31]</sup>则在邻域点上分配权重以进行局部信息聚合。在分类任务方面,给定数量有限的点数,每个点特征都将提供信息。为了在点云分割中捕获远距离特征,以及合理分析推断不同距离特征的影响,本文提出了更合理且更紧凑的推理模块以捕获分布在不同区域的语义信

息,无需堆叠卷积层同时扩大网络的感知领域。实验表明,本文提出算法在3D点云分类分割的基准上是具有竞争力的。

## 2 基本方法

对比利用体素以及多视图的方法优缺点,本文采用直接基于点的学习方法,以期减少预处理的繁琐流程,获得更优异的表现。如引言中所述,当前方法尚存改进空间。为了提高网络分析点云的准确性和对噪点的鲁棒性,从DGCNN出发,修改了中心点局部区域的选择方式和特征聚合分析模块。但是经典的多层感知器(MLP)映射<sup>[15]</sup>和边特征<sup>[17]</sup>依然不可或缺。

MLP在点云特征图上呈现为 $1 \times 1$ 的卷积,通常后跟批处理归一化层和激活函数:

$$\mathcal{M}\{\cdot\} = \partial\{\text{BN}[c_{1 \times 1}(\cdot)]\}, \quad (1)$$

式中: $\mathcal{M}$ 是MLP; $\partial$ 是激活函数;BN是批处理归一化; $c$ 是卷积,其下标表示滤波器的大小。

边特征 $f_{\psi}$ 被认为是空间中的隐藏几何特征,它是在高维特征空间基于点间几何关系设计的:

$$f_{\psi}(x_i) = \bar{f}_{\psi}(x_i, x_j - x_i), \quad x_i \in \mathbb{R}^d, \quad (2)$$

式中, $\bar{f}$ 为具有一组可学习参数 $\Psi$ 的非线性函数; $x_i$ 是 $d$ 维嵌入空间中的点特征,其特征图可表达成 $\chi_{N \times d} = [x_1, x_2, \dots, x_N]$ , $\mathbb{R}^d$ 表示 $d$ 维的特征空间; $x_j$ 是 $x_i$ 的局部近邻点,一般通过K近邻(KNN)算法搜索。

此外,MLP可以对边特征给定的局部语义信息编码,表示为边卷积 $\mathcal{E}$ :

$$\mathcal{E}(\cdot) = \mathcal{M}[f_{\psi}(\cdot)]. \quad (3)$$

通过边卷积构建径向局部图网络,该图由从邻点指向中心点的边组成,由共享权重的MLP映射到特征空间中。应用对称函数后,将每个点特征信息编码到相应的局部区域。

基于以上图卷积理论,本节介绍在此基础上的改进与创新。设计的网络结构与流程如图1所示,共完成形状分类和部分分割两个任务。在图1上半分支实线框的部分分割任务中,将两层负反馈卷积模块层加一层全局语义推理模块共3层作为一个模块组,输入 $3 \times N$ 的点云数据后,按顺序通过4个模块组的学习。每一批输入的数据在顺序通过4个模块组后,取各组末层全局语义推理层学习出的特征拼接在一起,经过共享参数的MLP层后,在通道方

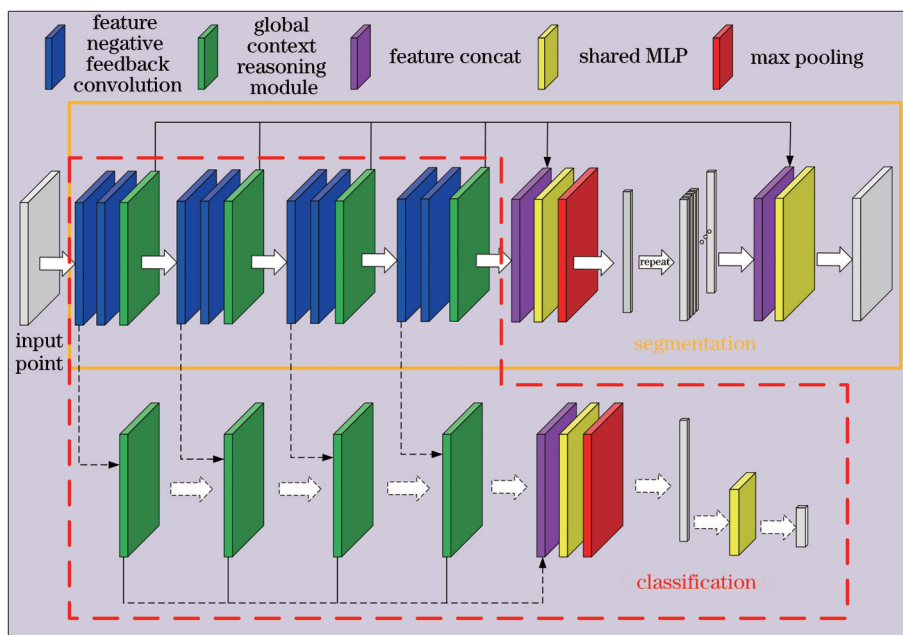


图 1 网络整体结构示意图

Fig. 1 Schematic diagram of overall network structure

向池化。池化的特征图在通道方向重复 1024 次,与前述拼接后的特征图拼接,通过另一个共享参数的 MLP 层学习,最终为点云的每一个像素点分配一个标签。

在图 1 下半分支虚线框的形状分类任务中,考虑到计算成本和分类准确性的平衡,只将一层负反馈卷积模块层与一层全局语义推理层组合。输入的点云数据同样按顺序通过 4 个模块层学习,取 4 组中全局语义推理层输出特征拼接,经过共享参数的 MLP 层后,在通道方向池化,再次通过 MLP 学

习后,为每一批点云数据给出  $1 \times 40$  的分类结果。

### 2.1 结构感知 KNN

当前主流分组方法 BG 和 KNN 在结构特征利用方面仍有不足。例如选择区域较为僵化,无法自适应地搜寻并将局部结构内较为相关的点划分到同一邻域。为了降低这些问题带来的影响,本文提出了结构感知 KNN 点分组方法(如图 2 所示),它的输入是点云各点特征,输出是邻域点的索引。与经典 KNN 相比,结构感知的 KNN 点分组方法可在最近邻搜索中收集更多信息以选择合适的分组

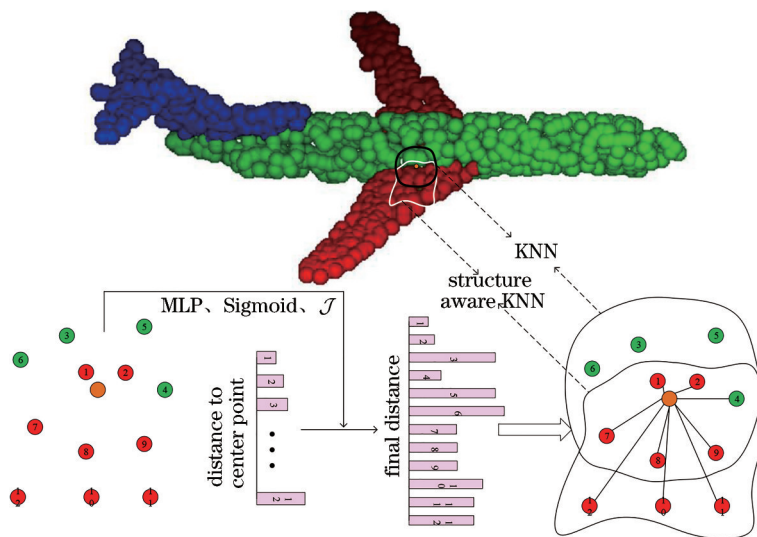


图 2 结构感知 KNN 和经典 KNN 示意图

Fig. 2 Diagram of structure aware KNN and classic KNN



并减少噪点干扰,这种方法可以根据某个数据集动态控制候选邻域范围。在此过程中边界上的点得到更精确的划分,从而提升网络整体的分类精度。

首先将各点分别作为中心点,搜索中心点候选邻点。将特征空间中的各成对点的欧氏距离矩阵  $E_{N \times N}$  定义为度量,  $N$  为总的点数,矩阵表达了点的密度分布。根据计算出的距离度量,可以确定每个点的  $k \cdot d$  个候选最近邻。这里  $k$  是邻点数,  $d$  是设置的候选参数,通过它调节进入候选的点数。在实践中,按升序对  $E_{N \times N}$  矩阵的行进行排序,保留前  $k \cdot d$  个最大元素的度量值和索引。因此,在  $E_{N \times N}$  的每行中前  $k \cdot d$  个度量值的元素被标识为每个点的候选邻点。记录下所搜索的候选邻居的度量值和索引。

为了在感知点云结构的基础上为每个点构造邻域,基于候选邻点的已知信息为每个点学习出一个扩张因子。在实际操作中,通过  $E_{N \times (k \cdot d)}$  和 CNN 运算来学习扩张因子。具体来说,将 MLP ( $\mathcal{M}$ ) 和 Sigmoid 函数 ( $\sigma$ ) 应用于候选  $E_{N \times (k \cdot d)}$  的度量值,以学习局部区域的点分布信息。投影函数  $\mathcal{J}$  将这些值映射到期望的扩张因子,通过扩张因子指导所选点的索引:

$$\mathcal{D}_N = \mathcal{J} \left\{ \sigma \left[ \mathcal{M} \left( E_{N \times (k \cdot d)} \right) \right] \right\}. \quad (4)$$

最后,根据逐点学习出的扩张因子,在每个中心点的候选点索引中选取  $d_i$ ,以形成每个中心点的  $k$  个邻点,  $d_i$  是矩阵  $\mathcal{D}_N$  中的元素。

## 2.2 特征负反馈卷积

在确定各中心点的邻域分组后,需要通过卷积网络学习各邻域的特征。通常,在经过边卷积后将池化操作应用于各局部图的  $k$  个邻点,聚合局部语义信息为中心点的特征表示。但是这一过程中可能存在偏差:一方面,局部图从初始 3D 空间起缺少几何正则化;另一方面,最大池化操作在合并特征时仅保留最突出的特征,而在特征空间中丢弃局部细节。在这种情况下,特征负反馈机制有助于避免在高维特征学习过程中出现偏差累积。

控制理论中的负反馈机制是指:通过监视过程的实际输出,找到与输入信号之间的差异,以此获得所要进行的调整激励,从而精确地控制过程向理想方向进行。受此启发,本文设计一种结构,通过向前投射经最大池化后具有代表性的邻域特征恢复信号来调节学习过程。在神经网络训练机制下,随着训练的持续,预期得到一个特定含义的误差反馈激励信号,该信号将约束特征学习并减轻对期望

输出的可能偏差。

类似于先前的注释,  $x_i \in \mathbb{R}^d$  是特征图  $\chi_{N \times d}$  的点特征。具体来说,将不同下标表达的点特征添加至式(1)和式(3)中定义的运算中。使用 MLP 编码局部信息的边卷积表示为  $\mathcal{E}_\phi$  作为输入特征图,此操作可以表达特征空间中局部区域的关系。应用式(3),  $\mathcal{E}_\phi$  的输出为

$$f_\phi = \mathcal{E}_\phi(x_i) = \mathcal{M}_\phi[f_\psi(x_i)]; \quad f_\phi \in \mathbb{R}^{d' \times k}, \quad (5)$$

式中  $k$  是点的邻居数。

通过边卷积  $\mathcal{E}_\phi$  学习后,理想的输出特征图应完全编码局部几何细节。如果输出  $f_\phi$  包含局部信息,可以通过反向投影从  $f_\phi$  恢复原始输入  $x_i$ ,这种反向投影应能模拟边卷积的反向过程。将经过最大池化后聚合表达的特征反投影重建到原始输入  $x_i$ ,以此观察这些具有代表性的特征与边卷积在特征空间中局部邻域点特征投影到中心点之间的差异,这里直接使用全连接层(FC),这一操作的物理含义是通过可学习的权重将筛选出的  $k$  组中具有代表性的特征学习到先前的特征空间  $d' \rightarrow d$ :

$$Y(\bullet) = \partial \left\{ \text{BN} \left[ c_{1 \times d'}(\bullet) \right] \right\}, \quad (6)$$

式中  $Y$  是 FC。在数学上,FC 是具有内核大小  $1 \times d'$  的共享参数 MLP,可以完全连接特征空间中的  $d'$  个特征。将重建的特征定义为负反馈信号  $x'_i$ 。根据式(5)和式(6)可得

$$x'_i = Y \left\{ P \left\{ f_\phi \right\} \right\} = Y \left\{ P \left\{ \mathcal{M}_\phi \left[ f_\psi(x_i) \right] \right\} \right\}; \quad x'_i \in \mathbb{R}^d, \quad (7)$$

式中  $P$  为最大池化操作。此外,原始输入与重建的输入(即特征的负反馈信号)之差  $\Delta x_i$  可以表示为相应的差值激励信号:

$$\Delta x_i = x_i' - x_i; \quad \Delta x_i \in \mathbb{R}^d. \quad (8)$$

通过另一个边卷积投影可以得到误差信号对  $f_\phi$  的影响:

$$f_{\theta_i} = \mathcal{E}_\theta(\Delta x_i) = \mathcal{M}_\theta[f_\psi(\Delta x_i)]; \quad f_{\theta_i} \in \mathbb{R}^{d' \times k}. \quad (9)$$

最后添加误差信号影响  $f_\phi$ , 调整输出  $f_{\phi_i}$ 。

网络结构如图 3 左半分支即 attentive pooling 框外部分所示,应用 2.1 节所述结构感知 KNN 将点云划分为局部区域,每个局部区域包括中心点及其  $k$  个邻点。再将  $k$  组特征图输入 MLP 层学习,每组学习到  $d'$  维数特征,  $N \times k \times d'$  维矩阵特征图经最大池化与 FC 反馈学习到初始输入维度  $N \times d$ 。与初始输入逐像素比较做差得到反馈激励信号,该信

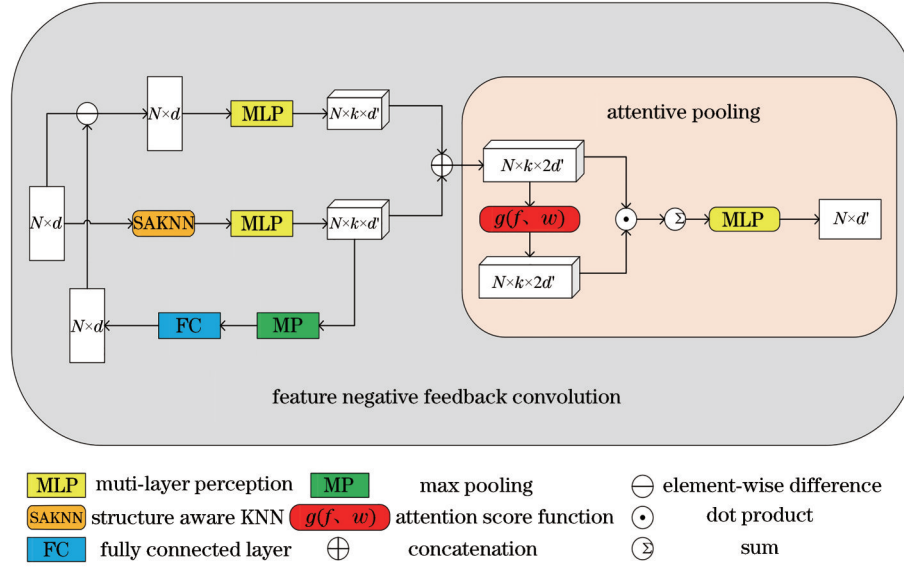


图 3 特征负反馈卷积模块

Fig. 3 Feature negative feedback convolution module

号通过 MLP 学习出补充特征图。补充特征图与第一次输出的特征图拼接,由注意力池化筛选得到卷积模块的输出。

### 2.3 注意力池化

尽管在特征负反馈卷积中将池化操作的结果反投影到输入维度,但是最终目的依然是聚合这  $k$  组特征筛选最具代表性的特征。这里不再使用最大池化,而是使用基于注意力的池化操作。根据网络自适应学习出的特征权重进行筛选将有助于得到最优的池化结果。

注意力池化的神经网络单元用于汇总邻近点特征  $F_i$  的集合。现有方法通常使用最大池化或者平均池化来集成邻域特征,从而导致部分信息丢失。相比之下,本文使用注意力机制来自动聚合重要的局部特征。本节的池化单元包括以下步骤,以  $k$  组其中之一中心点及其邻域来说明:计算注意力分数,在得到卷积网络输出的局部特征  $F_i = \{f_i^1, \dots, f_i^K\}$  后,设计函数  $g$  为每个特征学习唯一的注意力得分。函数  $g$  由一个共享参数的 MLP 组成,并使用软极大值使注意力权重更集中。它的形式定义如下:  $s_i^k = g(f_i^k, W)$ 。这里  $W$  是共享参数的 MLP 中可学习权重。学到的注意力分数可以看作是选择重要特征的掩膜。这些函数的加权总和如下:

$$\tilde{f}_i = \sum_{k=1}^K (f_i^k \cdot s_i^k). \quad (10)$$

总结本节所述,如图 3 特征负反馈卷积模块层中 attentive pooling 框中结构所示。对于点云的中

心点  $i$ ,通过  $g$  函数计算其  $k$  个最近邻点特征的注意力得分,并将注意力得分矩阵点乘特征图,点乘后的特征图在  $k$  维度上求和,通过共享参数的 MLP 层学习,生成信息丰富的特征向量  $\tilde{f}_i$ 。对每一个中心点及其邻域都执行如上注意力池化操作,可以聚合所有  $\tilde{f}_i$  得到最终输出特征图。图 3 中记述了特征图维度的变化。

### 2.4 全局语义推理模块

鉴于特征负反馈卷积只能分散地捕获局部语义信息,如果能更好地获取全局语义信息作为补充将有效增强网络整体表达力。基于传统卷积的方法需要套用多层卷积层以扩大接收场。这种方式可能消耗大量的内存并使用更长的推理时间。本文采用全局结构推理来避免这些问题,目标是通过特征负反馈卷积模块输出的特征图  $F^{out}$  中任何一对元素的相似性,从而提取整个特征图  $F^{out}$  中非局部的语义信息。

如图 4 所示,在特征负反馈卷积模块输出特征图  $F^{out} \in \mathbb{R}^{N \times d}$  后,全局推理模块在  $F^{out}$  上应用 3 个不同的  $1 \times 1$  卷积滤波器以获得 3 个特征图:查询特征图  $Q \in \mathbb{R}^{N \times d}$ 、键特征图  $K \in \mathbb{R}^{N \times d}$ 、值特征图  $V \in \mathbb{R}^{N \times d}$ 。查询特征图  $Q$  和键特征图  $K$  在测量  $F^{out}$  中每对元素的成对相似性中起作用。对于空间位置  $u$ ,选择特征图  $Q$  的特征向量  $Q_u \in \mathbb{R}^d$  进行相似性比较,其中  $u = [1, \dots, N]$ 。另外,  $K_v \in \mathbb{R}^d$  是每个查找位置  $v = [1, \dots, N]$  的键特征图  $K$  的特征向量。然后,在  $Q_u$  和  $K_v$  之间应用相似性运算以获得相似度分数:

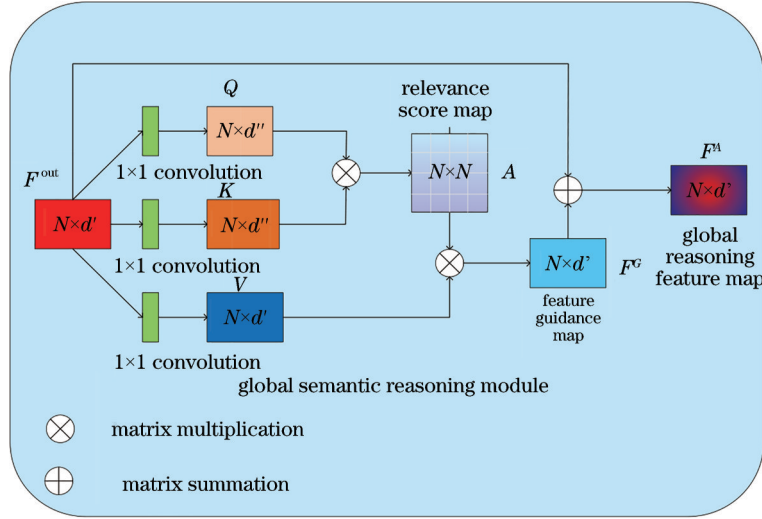


图 4 全局语义推理模块

Fig. 4 Global semantic reasoning module

$$r_{u,v} = Q_u^T K_v, \quad (11)$$

式中  $r_{u,v} \in \mathbf{R}$  表示特征向量  $Q_u$ 、 $K_v$  和矩阵  $\mathbf{R} \in \mathbb{R}^{N \times N}$  是相似度分数图。接着对每个元素  $r_{u,v} \in \mathbf{R}$  进行 softmax 运算以获得全局推理权重  $a_{u,v}$ ：

$$a_{u,v} = \frac{\exp(r_{u,v})}{\sum_{v=1}^N \exp(r_{u,v})}, \quad (12)$$

构成相似度分数图  $\mathbf{A} \in \mathbb{R}^{N \times N}$  的元素携带了整个特征图  $\mathbf{F}^{\text{out}}$  中的归一化相似度。

值特征图  $\mathbf{V}$  是特征图  $\mathbf{F}^{\text{out}}$  的变换, 提取其中包含的全局语义信息。全局推理权重  $a_{u,v}$  反映所有输入元素之间的相似性, 具有更高相似性的特征包括位于非局部的特征, 也能够被分配足够的权重, 从而进行后续的特征增强。因此, 将全局推理权重  $a_{u,v}$  应用于值特征图  $\mathbf{V}$  中的所有特征向量, 以获得全局特征向量  $\mathbf{F}_u^G$ ：

$$\mathbf{F}_u^G = \sum_{v=1}^N a_{u,v} \mathbf{V}_v, \quad (13)$$

式中  $\mathbf{F}_u^G \in \mathbb{R}^d$  表示特征引导图  $\mathbf{F}^G \in \mathbb{R}^{N \times d'}$  在空间位置  $u$  的特征向量, 该特征向量捕获了包含非局部区域的全局语义信息。然后, 将其添加到原始局部特征图  $\mathbf{F}^{\text{out}}$  以获得全局推理特征图  $\mathbf{F}^A = \mathbf{F}^G + \mathbf{F}^{\text{out}}$ 。最终, 推理聚合了局部与非局部各邻域语义信息, 增强了各中心点特征的表达, 从而更好地执行语义分类和分割任务。

### 3 实验与讨论

为了验证提出算法在点云分析方面的有效性,

本文在 ModelNet40<sup>[9]</sup> 和 ShapeNet 部件<sup>[32]</sup> 上对该算法性能进行了评估, 并与最先进的方法进行了比较。此外, 还进行了消融实验和可视化, 以说明提出方法的有效性。所有实验均在配备单个 NVIDIA GTX 2080 GPU 的台式计算机上进行, 使用 Pytorch 进行训练和测试。

#### 3.1 形状分类

##### 3.1.1 数据集

在 ModelNet40 数据集上证明了分类网络对形状分类的有效性<sup>[9]</sup>。ModelNet40 是当前点云分析领域流行且权威的基准, 因为它的形状清晰、建模良好, 但是由于各种类别模型数量的有限性和不平衡性也充满了挑战。ModelNet40 包含 12311 个计算机辅助设计 (CAD) 模型, 这些模型分为 40 类。为了获得更好的训练效果, 分别使用官方划分的 9843 个模型来训练, 以及 2468 个模型来测试。从每个模型的网状表面均匀采样 1024 个点, 并将它们移动至原点并统一到单位球体。

##### 3.1.2 网络结构

形状分类任务的网络架构如图 1 所示。为了更好地对点云的局部区域信息进行编码, 使用包含注意力池化的特征负反馈卷积模块学习局部区域特征, 并通过全局语义推理模块捕获全局特征。通过这种将局部特征和全局特征组合的方式, 网络学习到了具有较强描述能力的全局特征。

将特征负反馈卷积模块与全局语义推理模块叠加作为一组, 共经过 4 组。这 4 组中特征负反馈卷积模块中 MLP 层卷积核个数分别为 64、64、128、



256。其中每个特征负反馈卷积模块包含结构感知 KNN, MLP 等主要结构, 如图 2 所示。

4 组卷积模块与全局语义推理模块的结果拼接后经 MLP 和最大池化得到  $1 \times 1024$  维全局特征, 应用 FC(512, 256, 40) 将全局特征转换为 40 个类别。在训练中, 使用经典的交叉熵损失来监督点云的分类。在测试过程中, 执行了与许多先前研究相似的投票测试<sup>[33-34]</sup>, 即对测试数据进行 10 次随机投票测试, 并对预测结果取平均值。此外, 除最后一层外, 每层都使用了批处理归一化(BN)和修正线性单元(ReLU)。

网络训练了 300 个批次, 并且批次大小设置为 32。使用 Adam<sup>[35]</sup> 优化器优化。初始学习速率为 0.001, 然后每 24 个批次降低 0.7 的学习速率。BN 的动量初始设置为 0.9, 每训练 24 个周期以 0.5 的速率衰减。使用文献<sup>[36]</sup>介绍的初始化技术来初始化权重, 并将偏差初始化为 0。根据文献<sup>[17, 37]</sup>的经验, 在结构感知 KNN 中的参数设置采用  $k=20$

和  $d=5$ 。此外, 参考以往研究<sup>[38]</sup>中数据增强的方法, 在  $[-0.8, 1.25]$  的尺度内非均匀缩放输入数据, 并在  $[-0.1, 0.1]$  的范围内平移。

### 3.1.3 结果分析

表 1 显示了本文以及其他先进方法在形状分类任务中的定量评估结果。结果表明, 提出方法的平均分类精度为 91.0%, 总体精度为 93.8%, 这两项指标均超过了近年来最先进的方法 RS-CNN<sup>[18]</sup>。相比于改进的基础 DGCNN<sup>[17]</sup> 网络, 提出方法的总体精度提高了 1.6 个百分点。而当 DGCNN<sup>[17]</sup> 使用 2k 个点坐标时, 提出方法的总体精度仍然高出 DGCNN 0.3 个百分点。同样, 提出方法也优于 SO-NET<sup>[39]</sup>, SO-NET 使用更为丰富的信息, 包含 5k 个点坐标及法向量, 实现了 93.4% 的总体精度和 90.8% 平均分类精度。提出方法不仅提高了准确性, 而且增强了系统的鲁棒性, 这进一步证明了该方法的有效性。

表 1 ModelNet40 数据集上的分类精度

Table 1 Classification accuracy on ModelNet40 data set

Method	Input type	Points	Mean class accuracy / %	Overall accuracy / %
Point-CNN <sup>[40]</sup>	coords	1k	88.1	91.7
PointNet <sup>[15]</sup>	coords	1k	86.0	89.2
A-SCN <sup>[28]</sup>	coords	1k		90.0
Kd-Net <sup>[38]</sup>	coords	1k		90.6
PointNet++ <sup>[16]</sup>	coords	1k		90.7
KCNet <sup>[41]</sup>	coords	1k		91.0
Spec-GCN <sup>[42]</sup>	coords	1k		91.5
DGCNN <sup>[17]</sup>	coords	1k	90.2	92.2
RS-CNN <sup>[18]</sup>	coords	1k		93.6
KP-Conv <sup>[43]</sup>	coords	1k		92.9
PointASNL <sup>[44]</sup>	coords	1k		93.2
Proposed	coords	1k	<b>91.0</b>	<b>93.8</b>
SpiderCNN <sup>[45]</sup>	coords + norm	5k		92.4
DensePoint <sup>[46]</sup>	coords + norm	1k		93.2
SO-NET <sup>[39]</sup>	coords + norm	5k	90.8	93.4
PointNet++ <sup>[16]</sup>	coords + norm	5k		91.9
DGCNN <sup>[17]</sup>	coords	2k	90.7	93.5

## 3.2 部分分割

### 3.2.1 数据集

在 ShapeNet 部件分割任务上评估了分割模型。部件分割是指对对象各个部分(例如飞机机翼、笔记本电脑的屏幕、桌腿等)的预定义类别标签的分配。ShapeNet Part<sup>[32]</sup> 数据集是 ShapeNet<sup>[47]</sup> 数据库的子集。它由 16881 个 CAD 形状组成, 其中包括

16 个类别和 50 个标记部件。此外, 在此数据集中, 每个形状由 2~6 个部分组成, 这导致相同类型形状中的部分数量有所不同。例如, 有线耳机包含 3 个部分, 而无线耳机包含 2 个部分。不平衡的形状类别和形状的各部分对细粒度形状分析的分割任务提出了很大的挑战。使用第 3.1.1 节中提到的相同策略, 为每个形状统一采样 2048 个点。采用与

文献[47]一致的数据划分方法将数据划分为训练集、验证集和测试集。

### 3.2.2 网络结构

图1上半部分分支显示,分割模型对于点云中每个点的部件类别都预测了标签。使用第3.1节形状分类相同的特征负反馈卷积模块和全局语义推理模块的组合进行特征学习,区别在于分割网络的特征负反馈卷积模块MLP卷积核个数分别为(64, 64)、(64, 64)、(64, 128)、(128, 256)。经过全局语义推理模块后的特征拼接后通过MLP和最大池化得到全局特征。将全局特征堆叠2048次并与对象的标签(例如飞机、笔记本电脑,书桌等)相连后再次和前述特征拼接,最后通过MLP(512, 256, 128, 50)逐点变换特征并为每个点生成最终预测标签。在测试过程中,使用了具有随机缩放比例的投票测试方法,并使用验证集保存了具有更好推理能力的模型。投票测试进行了10次,且对测试数据进行了随机缩放。

部件的细分任务网络配置设置与分类任务中的配置类似,结构感知KNN中的参数设置与分类网络一样。利用Adam<sup>[35]</sup>优化进行200个时期的训练。学习率从0.001开始,每20个周期以0.5的速率衰减。这两个任务的批处理大小均为32。

### 3.2.3 结果分析

在每个类别上使用标准交并比(IoU),将所有类别的平均IoU(overall mIoU)用作评估指标。每个类别的标准IoU通过对属于同一类别的所有部件的IoU平均计算得出。表2显示了网络在ShapeNet Part<sup>[32]</sup>数据集上的性能以及与其他方法的比较结果。与改进的基准DGCNN<sup>[17]</sup>相比,提出方法的overall mIoU比DGCNN高1.2个百分点。可以观察到,提出方法在多个类别中效果最佳。对于相对复杂的局部结构类别,例如飞机和汽车,该方法表现出色,这是因为网络在点云复杂结构上分组学习时就能更加精细地划分各中心点邻域。通过特征负反馈卷积模块,各邻域中最具代表性的特征被最大限度地选取出来。此外全局语义推理模块也加强了每个点与点之间的联系,更好地利用局部结构信息。然而提出方法在滑板、摩托车和火箭等类别的分割结果并不如人意。推测是这些模型中许多小而相似的部件结构影响了该方法的判断。

分割结果被保存为obj格式,通过Meshlab可视化了部分分割结果。图5的示例包括飞机、手提包、耳机、小刀、杯子、椅子、滑板、手枪、汽车、摩托、灯、火箭、帽子和桌子等。在这些类别中随机选择一个对象进行可视化。

表2 ShapeNet Part数据集上部件分割平均交并比

Table 2 Part segmentation results [mIoU (%)] on ShapeNet Part data set

Method	PointNet <sup>[15]</sup>	PointNet++ <sup>[16]</sup>	Spider CNN <sup>[45]</sup>	SO-NET <sup>[39]</sup>	A-SCN <sup>[28]</sup>	P2Sequence <sup>[48]</sup>	PCNN <sup>[49]</sup>	DGCNN <sup>[17]</sup>	RS-CNN <sup>[18]</sup>	Point ASNL <sup>[44]</sup>	Proposed
Overall mIoU	83.7	85.1	85.3	84.6	84.6	85.2	85.1	85.2	86.2	86.1	<b>86.4</b>
Air plane	83.4	82.4	83.5	81.9	83.8	82.6	82.4	84.0	83.5	84.1	<b>84.3</b>
Bag	78.7	79.0	81.0	83.5	80.8	81.8	80.1	83.4	84.8	84.7	<b>85.1</b>
Cap	82.5	87.7	87.2	84.8	83.5	87.5	85.5	86.7	<b>88.8</b>	87.9	88.6
Car	74.9	77.3	77.5	78.1	79.3	77.3	79.5	77.8	79.6	79.7	<b>79.9</b>
Chair	89.6	90.8	90.7	90.8	90.5	90.8	90.8	90.6	91.2	<b>92.2</b>	91.3
Ear phone	73.0	71.8	76.8	72.2	69.8	77.1	73.2	74.7	<b>81.1</b>	73.7	79.2
Guitar	91.5	91.0	91.1	90.1	91.7	91.1	91.3	91.2	91.6	91.0	<b>91.8</b>
Knife	85.9	85.9	87.3	83.6	86.5	86.9	86.0	87.5	88.4	87.2	<b>89.0</b>
Lamp	80.8	83.7	83.3	82.3	82.9	83.9	85.0	82.8	<b>86.0</b>	84.2	85.2
Laptop	95.3	95.3	95.8	95.2	<b>96.0</b>	95.7	95.7	95.7	<b>96.0</b>	95.8	95.7
Motobike	65.2	71.6	70.2	69.3	69.2	70.8	73.2	66.3	73.7	<b>74.4</b>	72.3
Mug	93.0	94.1	93.5	94.2	93.8	94.6	94.8	94.9	94.1	<b>95.2</b>	94.5
Pistol	81.2	81.3	82.7	80.0	82.5	79.3	83.3	81.1	<b>83.4</b>	81.0	82.0
Rocket	57.9	58.7	59.7	51.6	62.9	58.1	51.0	<b>63.5</b>	60.5	63.0	60.3
Skate board	72.8	76.4	75.8	72.1	74.4	75.2	75.0	74.5	<b>77.7</b>	76.3	76.4
Table	80.6	82.6	82.8	82.6	80.8	82.8	81.8	82.6	83.6	83.2	<b>84.4</b>



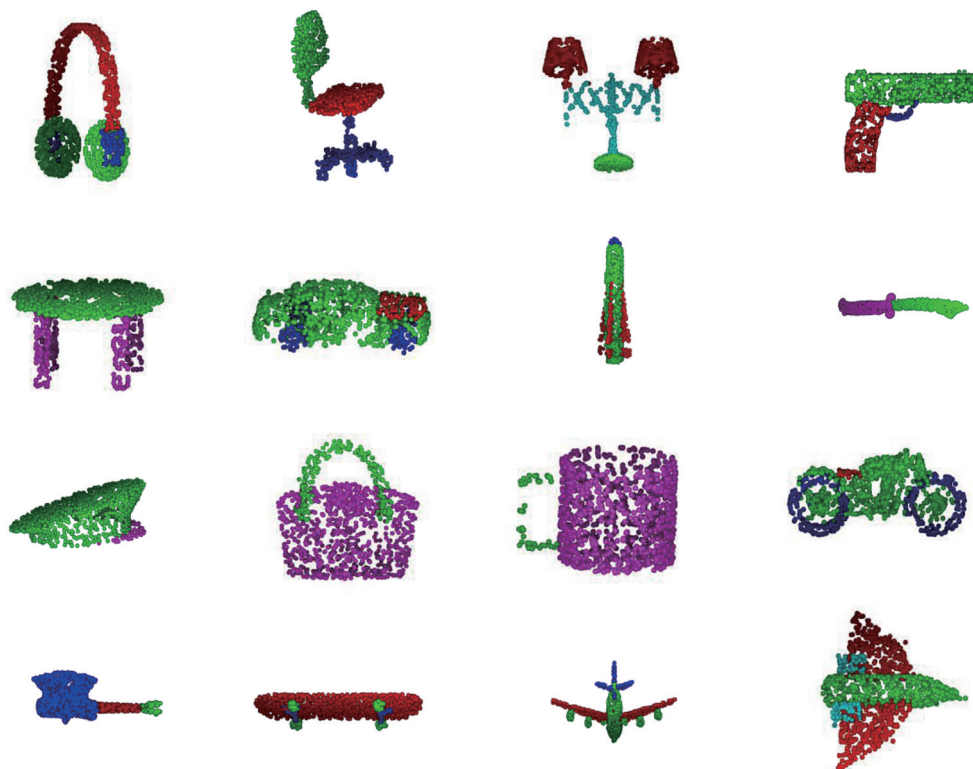


图 5 部件分割结果可视化

Fig. 5 Visualization of part segmentation results

### 3.3 消融实验

图 1 所示的网络架构由串联的组合模块构成, 主要包含结构感知 KNN、特征负反馈卷积、全局语义推理等组件。为研究各组件作用, 在 ShapeNet Part 上使用不同的设置来定量评估其对点云形状分类结果的影响。基线设置为堆叠由 MLP 组合的边卷积的 DGCNN 网络, 分组方式选择经典 KNN, 池化操作选择最大池化。

实验表明(结果见表 3)基线准确性为 84.9%。将结构感知 KNN 添加到基线后, overall mIou 提高了 0.3 个百分点。对此可以合理推测相比于经典 KNN, 本文提出的结构感知 KNN 在细致划分邻域时更加合理, 能够提前将可能相关的点划分到同一

组, 为后续卷积学习打下基础。将基线的最大池化操作改为注意力池化后, overall mIou 提高了 0.2 个百分点。这表明提出方法在挑选各邻域最具代表性的特征时, 机械地选择各通道极值作为代表可能并不是最优选项。当使用包含结构感知 KNN 以及注意力池化的特征负反馈卷积模块替代基线的边卷积操作时, overall mIou 进一步增加 0.4 个百分点。这是因为设计的负反馈机制能够改善特征在前向学习传播中受误差干扰的情况, 另外这一结构也提供了类似正则化的帮助, 从而提升了网络的准确率。添加了全局语义推理的模块的网络达到了 86.1% 的 overall mIou, 这充分说明在局部点或区域之间构筑相互联系的必要性。最后经过在

表 3 ShapeNet 部件数据集上关于不同网络组件作用的消融实验

Table 3 Ablation experiments about effects of different network components on ShapeNet part data set

Model	SAKNN	Attentive pooling	Feature negative feedback convolution	Global context reasoning module	Overall mIou / %
0					84.9
1	✓				85.2
2	✓	✓			85.4
3	✓	✓	✓		85.8
4	✓	✓	✓	✓	86.1

ShapeNet Part 上的 10 次投票测试, 网络得到了 86.4% 的 overall mIoU。

为了进一步验证模型的鲁棒性, 训练时也采取随机丢失的方法将稀疏后的点云数据集(1024、768、512、256、128)输入模型。将提出方法与 PointNet<sup>[16]</sup>, PointNet++<sup>[17]</sup>, DGCNN<sup>[18]</sup> 和最新的具有优秀鲁棒性的模型 PointASNL<sup>[44]</sup> 进行比较。从图 6 中可以看出, 相比改进前的 DGCNN, 提出方法的鲁棒性得到了很大提升, 特别是在输入 384 个点以下时。实验结果充分证明了特征负反馈和全局推理学习方法的使用可以大大提高模型的鲁棒性。这种鲁棒性使提出方法可以适用于小样本带有噪点的点云, 为将来在户外等存在干扰的场景下的应用提供了思路。

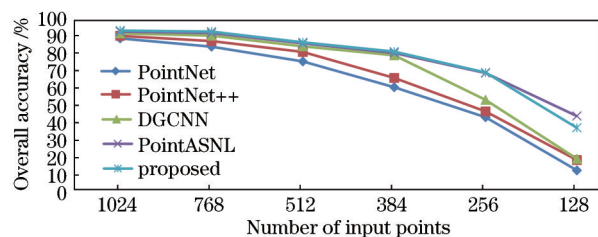


图 6 稀疏点的测试结果

Fig. 6 Test results of sparser points

## 4 结 论

提出了基于特征负反馈卷积的网络用于点云分析, 也即通过引入结构感知分组方法以实现自适应的柔性点分组, 提出了特征负反馈的卷积方法和注意力池化, 以筛选捕获更具代表性的局部特征信息。此外, 引入全局推理模块通过聚焦不同区域之间的联系来优化特征。实验结果表明, 在 ShapeNet 部件分割和 ModelNet40 形状分类数据集上, 同最先进的的方法相比, 提出方法也是具有竞争力的。然而在小而相似的精细结构划分上, 该方法尚存改进空间。在未来的工作中, 将进一步优化网络拓扑结构, 完善算法在构筑细微部件与整体之间联系方面的不足, 并尝试推广到大规模室外场景的点云语义分析中。

## 参 考 文 献

[1] Zhou Y, Tuzel O. VoxelNet: end-to-end learning for point cloud based 3D object detection[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: 4490-4499.

- [2] Qi C R, Liu W, Wu C X, et al. Frustum PointNets for 3D object detection from RGB-D data[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: 918-927.
- [3] Ku J, Mozifian M, Lee J, et al. Joint 3D proposal generation and object detection from view aggregation [C]//2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), October 1-5, 2018, Madrid, Spain. New York: IEEE Press, 2018: 1839-1845.
- [4] Liu Z Z, Chen H Y, Di H J, et al. Real-time 6D lidar SLAM in large scale natural terrains for UGV [C]//2018 IEEE Intelligent Vehicles Symposium, June 26-30, 2018, Changshu, China. New York: IEEE Press, 2018: 662-667.
- [5] Zhu Y K, Mottaghi R, Kolve E, et al. Target-driven visual navigation in indoor scenes using deep reinforcement learning[C]//2017 IEEE International Conference on Robotics and Automation, May 29-June 3, 2017, Singapore. New York: IEEE Press, 2017: 3357-3364.
- [6] 戴仁月, 方志军, 高永彬. 融合扩张卷积网络与 SLAM 的无监督单目深度估计[J]. 激光与光电子学进展, 2020, 57(6): 061007.
- Dai R Y, Fang Z J, Gao Y B. Unsupervised monocular depth estimation by fusing dilated convolutional network and SLAM[J]. Laser & Optoelectronics Progress, 2020, 57(6): 061007.
- [7] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84-90.
- [8] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2014-09-04) [2021-05-04]. <https://arxiv.org/abs/1409.1556>.
- [9] 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, June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 1912-1920.
- [10] 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.
- [11] Gadelha M, Wang R, Maji S. Multiresolution tree networks for 3D point cloud processing[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 105-122.
- [12] 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, December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 945-953.
- [13] Feng Y F, Zhang Z Z, Zhao X B, et al. GVCNN: group-view convolutional neural networks for 3D shape recognition[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: 264-272.
- [14] Guo H Y, Wang J Q, Gao Y, et al. Multi-view 3D object retrieval with deep embedding network[J]. IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society, 2016, 25(12): 5526-5537.
- [15] 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, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 77-85.
- [16] Qi C R, Yi L, Su H, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space[EB/OL]. (2017-06-07) [2021-05-07]. <https://arxiv.org/abs/1706.02413>.
- [17] 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): 146.
- [18] Liu Y C, Fan B, Xiang S M, et al. Relation-shape convolutional neural network for point cloud analysis [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 8887-8896.
- [19] Carreira J, Agrawal P, Fragkiadaki K, et al. Human pose estimation with iterative error feedback[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 4733-4742.
- [20] Haris M, Shakhnarovich G, Ukita N. Deep back-projection networks for super-resolution[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: 1664-1673.
- [21] Liu Z S, Wang L W, Li C T, et al. Hierarchical back projection network for image super-resolution [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 16-17, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 2041-2050.
- [22] Li R H, Li X Z, Fu C W, et al. PU-GAN: a point cloud upsampling adversarial network[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 7202-7211.
- [23] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA. [S.l.: s.n.], 2017: 5998-6008.
- [24] Wang X L, Girshick R, Gupta A, et al. Non-local neural networks[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: 7794-7803.
- [25] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[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: 7132-7141.
- [26] Fu J, Liu J, Tian H J, et al. Dual attention network for scene segmentation[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 3141-3149.
- [27] Anwar S, Barnes N. Real image denoising with feature attention[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 3155-3164.
- [28] Xie S N, Liu S N, Chen Z Y, et al. Attentional ShapeContextNet for point cloud recognition[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: 4606-4615.
- [29] Feng M T, Zhang L, Lin X F, et al. Point attention network for semantic segmentation of 3D point clouds [EB/OL]. (2019-09-27) [2021-06-04]. <https://arxiv.org/abs/1909.12663>.
- [30] Liu X H, Han Z Z, Wen X, et al. L2G auto-encoder: understanding point clouds by local-to-global reconstruction with hierarchical self-attention [C]//MM '19: Proceedings of the 27th ACM International Conference on Multimedia, October 21-25, 2019, Nice, France. New York: ACM Press, 2019: 989-997.
- [31] Hu Q Y, Yang B, Xie L H, et al. RandLA-net: efficient semantic segmentation of large-scale point clouds[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 11105-11114.
- [32] 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.
- [33] Wang P S, Liu Y, Guo Y X, et al. O-CNN: octree-based convolutional neural networks for 3D shape analysis[J]. *ACM Transactions on Graphics*, 2017, 36(4): 72.
- [34] Wang P S, Sun C Y, Liu Y, et al. Adaptive O-CNN: a patch-based deep representation of 3D shapes [J]. *ACM Transactions on Graphics*, 2018, 37(6): 217.
- [35] Kingma D P, Ba J. Adam: a method for stochastic optimization[EB/OL]. (2014-12-22) [2021-05-06]. <https://arxiv.org/abs/1412.6980>.
- [36] He K M, Zhang X Y, Ren S Q, et al. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification[C]//2015 IEEE International Conference on Computer Vision, December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 1026-1034.
- [37] Engelmann F, Kontogianni T, Leibe B. Dilated point convolutions: on the receptive field size of point convolutions on 3D point clouds[EB/OL]. (2019-07-28)[2021-05-08]. <https://arxiv.org/abs/1907.12046>.
- [38] Klovov 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, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 863-872.
- [39] Li J X, Chen B M, Lee G H. 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.
- [40] Li Y Y, Bu R, Sun M C, et al. PointCNN: convolution on X-transformed points[C]//Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, December 3-8, 2018, Montréal, Canada. [S.l.: s.n.], 2018.
- [41] Shen Y R, Feng C, Yang Y Q, et al. Mining point cloud local structures by kernel correlation and graph pooling[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: 4548-4557.
- [42] Wang C, Samari B, Siddiqi K. Local spectral graph convolution for point set feature learning[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11208: 56-71.
- [43] Thomas H, Qi C R, Deschard J E, et al. KPConv: flexible and deformable convolution for point clouds [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 6410-6419.
- [44] Yan X, Zheng C D, Li Z, et al. PointASNL: robust point clouds processing using nonlocal neural networks with adaptive sampling[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 5588-5597.
- [45] 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.
- [46] Liu Y C, Fan B, Meng G F, et al. DensePoint: learning densely contextual representation for efficient point cloud processing[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV),



- October 27-November 2, 2019, Seoul, Korea (South).  
New York: IEEE Press, 2019: 5238-5247.
- [47] Chang A X, Funkhouser T, Guibas L, et al. ShapeNet: an information-rich 3D model repository [EB/OL]. (2015-12-09) [2021-03-07]. <https://arxiv.org/abs/1512.03012>.
- [48] Liu X H, Han Z Z, Liu Y S, et al. Point2Sequence: learning the shape representation of 3D point clouds with an attention-based sequence to sequence network [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33: 8778-8785.
- [49] Atzmon M, Maron H, Lipman Y. Point convolutional neural networks by extension operators [J]. ACM Transactions on Graphics, 2018, 37(4): 71.