

基于 SA-PointNetVLAD 的点云分类网络

李肖南¹, 王 蕾^{1,2}, 程海霞¹, 张志勇¹

(1. 东华理工大学 信息工程学院, 江西 南昌 330013;

2. 江西省核地学数据科学与系统工程技术研究中心, 江西 南昌 330013)

摘 要: 三维点云数据包含着丰富的形状和比例信息, 如何有效准确地对点云数据进行分类已经成为了目前计算机视觉领域的研究热点。由于点云在非欧氏空间中的不规则稀疏结构, 并且现有的基于深度学习的三维点云分类模型中缺乏对各个点的局部信息和全局信息的有效利用, 从而导致较低的分类精度。为了解决上述问题, 提出了一种基于 SA-PointNetVLAD 的点云分类模型框架。该网络采用逐点特征提取和汇集操作来解决无序的点云问题, 通过自注意力机制来计算每个点与其他所有点之间的关联, 充分挖掘点云的局部区域细粒度特征以及全局信息, 采用 KNN 邻近算法感知点云的局部形状结构, 通过 VLAD 层将每个点的低层几何描述符与视觉单词相关联来间接描述高层特征信息。此外, 该网络还引入一个有效的关键点描述符帮助识别整体几何图形。经过实验得出 SA-PointNetVLAD 模型在公开的 ModelNet40 数据集上仅使用 512 个点就可以达到 90.9% 的精度, 显著高于相同条件下的其他方法。

关键词: 自注意力机制; 点云分类; VLAD; 关键点; 高级特征信息

中图分类号: TP391

文献标识码: A

文章编号: 1673-629X(2022)05-0036-06

doi: 10.3969/j.issn.1673-629X.2022.05.006

Point Cloud Classification Network Based on SA-PointNetVLAD

LI Xiao-nan¹, WANG Lei^{1,2}, CHENG Hai-xia¹, ZHANG Zhi-yong¹

(1. School of Information Engineering, East China University of Technology, Nanchang 330013, China;

2. Jiangxi Engineering Technology Research Center of Nuclear Geoscience Data Science and System, Nanchang 330013, China)

Abstract: 3D point cloud data contains rich shape and scale information. How to classify the point cloud data effectively and accurately is a research hotspot in the field of computer vision. Due to the irregular and sparse structure of point cloud in non-Euclidian space, the existing 3D point cloud classification model based on deep learning lacks the effective use of the local and global information of each point, which leads to low classification accuracy. To solve the above problems, a point cloud classification model framework based on SA-PointNetVLAD is proposed. The network uses point-by-point feature extraction and aggregation operations to solve the disordered point cloud problem. By computing the association between each point and all other points through the self-attention mechanism, the local fine-grained features and global information of the point cloud are fully explored, the local shape structure of the point cloud is perceived by the KNN proximity algorithm, and the high-level feature information is indirectly described by associating the low-level geometry descriptor of each point with several visual words through the VLAD layer. In addition, the network introduces an effective key point descriptor to help identify the overall geometry. The experiments show that the SA-PointNetVLAD model can achieve 90.9% accuracy using only 512 points on a public ModelNet40 dataset, significantly higher than other methods under the same conditions.

Key words: self-attention mechanism; point cloud classification; VLAD; key points; high-level feature information

0 引言

点云作为一种表示三维模型的基本格式, 可以通过激光扫描仪很轻易地获取, 但由于其点数不确定性和排列的无序性, 不适合输入到深度神经网络中处理。近年来, 一些研究人员致力于开发 3D 点云的分类方法。2017 年, Qi 等人提出了用于直接处理不规则点云

的 PointNet^[1] 框架, 它通过对每个点独立执行特征学习, 然后应用最大池来累积特征, 从而实现点云的排列不变性。虽然 PointNet^[1] 比以前的方法更加精确和健壮, 但是它忽略了局部特征, 这限制了它对复杂场景的细粒度模式识别和感知。作为增强版, PointNet++^[2] 引入了分层特征学习来学习尺度不断增大的局部特

收稿日期: 2021-05-06

修回日期: 2021-09-08

基金项目: 国家自然科学基金资助项目(61561003, 61761003); 江西省核地学数据科学与系统工程技术研究中心基金(JETRCNGDSS201902)

作者简介: 李肖南(1997-), 女, 硕士, 研究方向为点云分类; 王 蕾, 博士, 副教授, 硕导, 研究方向为视觉信息获取与处理、计算机视觉、计算机图形图像技术等。

征,但在学习过程中仍然只对每个点进行独立操作,忽略了点与点之间的关系。为了解决这个问题,一些学者通过考虑每个点的邻域来充分提取局部特征,尽管在公共数据集上已经取得了有效的结果,但是仍然存在一些未解决的问题。首先,每个点都有自己的坐标信息和额外的属性值,即颜色、法线、反射率等。这种表示只表达点本身的物理意义,而不考虑其邻近的和上下文的意义。第二,每个点的标注不仅取决于它自己的表示,还与其他点有关,点云中复杂的全局关系还没有被明确地开发和描述。第三,不同位置的点对几何知觉的重要性是不相等的。具体来说,角点和边缘比平坦区域的视觉敏感,强调这些关键点对于提高特征的质量是至关重要的。

为了解决上述问题,该文提出了一种新的三维点云分类模型,包括低级局部几何特征和高级全局语义特征,并引入了 self-attention 机制,更灵活地捕捉局部和全局信息,可以充分挖掘三维点云的深层隐含特征。这项工作提供了以下三个主要内容:

(1)构建了一个端到端的三维点云分类模型,网络采用逐点特征提取和汇集操作来解决无序的点云问题。

(2)以 Self-Attention 机制的方式作为网络输入,而不是仅考虑每个孤立点的位置,可以灵活地捕捉全局和局部的联系,可以充分挖掘三维点云的深层隐含特征。

(3)采用 KNN 和 VLAD 模块分别提取低级几何特征和高级语义特征,并引入关键点来表示每个不同点的权重,达到对点云特征的有效提取。

1 相关工作

随着深度学习技术的发展和更多的数据集开始向公众开放,点云处理已经成为计算机视觉领域的一项重要任务。深度神经网络已经在各种视觉任务中取得了显著的成功,然而,将 CNN 应用于不规则结构的物体里(如 3D 点云)仍然具有挑战性。

PointNet^[1]是将原始点云作为深度学习输入的开创性工作,并且它是首次对点云进行直接处理,因此为很多学者对点云的研究提供了新思路。其主要思想是利用点向卷积将原始三维坐标映射到高维特征空间,然后进行 max pooling 或 average pooling 操作来消除点排列的影响。但是由于 PointNet 在提取点云特征的时候忽略了局部特征,从而限制了其分类性能。为了解决这一问题,PointNet++^[2]提出了点云的分层特征学习方法,将整个点集划分为多个子集,并对每个子集重复应用一个简化的 PointNet^[1],这些局部特征被分组组成一个全局表示。PointNet++^[2]由于分组的复杂

过程和反复的前向传播,使得调优变得非常耗时,导致结果并不能达到特别理想。这两种方法的提出使得很多学者都开始相继研究直接对点云处理的方法。

最近,已经有很多学者引入 k-d 树、图等结构来捕获非结构化点之间的局部关系^[3-8]。例如, KCNet^[8]和 FoldingNet^[7]使用基于图的 max-pooling 对图进行下采样,该图使用构建的 k-NN 图对每个节点的邻域取最大特征。但是这些方法不能保证将最重要的点(临界点)传递给下层网络,使得具有不太相关的特征的点可能被选择,而重要的点可能被删除。一些学者还提出了新的卷积策略,尝试引入各种局部特征,如到邻近点的距离和局部表面法线的夹角,并用它们来表示点云。随后的 PointCNN^[9]、PointConv^[10]等文献也聚焦于点云的局部结构,进一步提高了捕获特征的质量。根据局部特性和全局特性之间的关系,引入了 self-attention 机制、KNN 模块以及 VLAD 层来提高点云的分类精度,此外,还引入了一个有效的关键点描述符来帮助识别整体几何图形。经过大量实验,该模型在 ModelNet40 数据集上获得了 90.9% 的分类精度。

2 点云深度学习网络框架

该文提出的网络框架引入了 self-attention 层来计算每个点与其他所有点之间的关联,从而挖掘点云的局部区域细粒度特征以及全局信息,并采用 KNN 算法感知点云的局部形状结构。在点云识别中,通常可以通过一些边角上的点来判别物体形状,因此在点云分类中,一些关键点信息也不可忽略。在网络中引入了一个有效的关键点描述符,对一些点通过计算其法向量来获取不同的点权重,这是整个网络的关键之处。

文献[11]中提出了一种新的 VLAD 方法,首次提出高层几何信息是可以通过每个点的低层几何描述子和几个视觉单词之间的关系来间接描述的,这种关系使网络能够推断出三维点云的高层语义信息。因此,该文在网络框架中加入了这个模块,用于提取点云中包含的高层几何信息。总体网络框架如图 1 所示。

2.1 SA-TNet(self-attention 层)

点之间的依赖关系可以改善分类结果。在实验中,受到自然语言处理(序列对序列)任务研究的启发,该文引入了一种 self-attention 机制,可以提取点之间的上下文特征。它是一种自我注意模块,可以简单地描述为查询到键值对的映射,通常包含查询、键、值和输出。通过对键值对的查询,得到对应键下查询的权值,将权重添加到相应的查询中以得到输出。同一个查询在不同的键下有不同的输出,即不同的关注,并且 self-attention 对于输入顺序也是不变的。该文利用 self-attention 机制来获得每一个点的得到了新的考虑

了上下文信息的表征。self-attention 模型如图 2 所示。

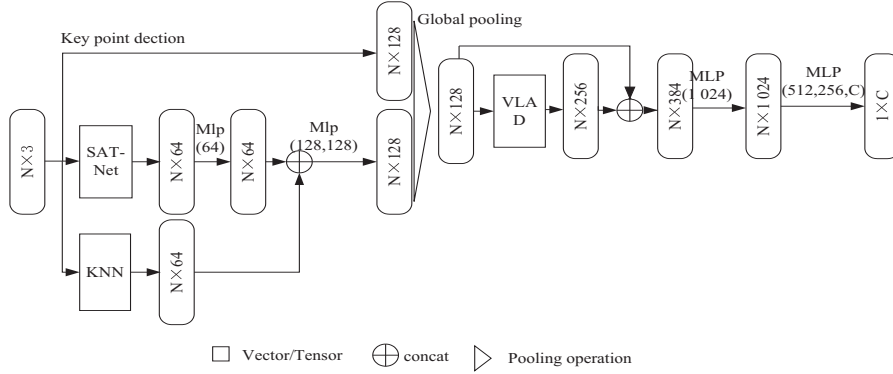


图 1 SA-PointNetVLAD 网络框架

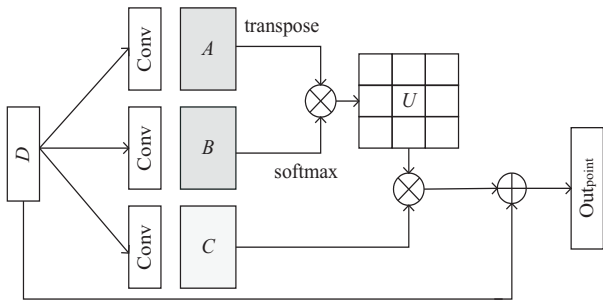


图 2 self-attention 模型

与传统的基于注意的序列对序列模型不同,在一个自注意块中,查询向量 $Q \in R^{D_q}$, key 向量 $K \in R^{D_k}$ (通常 $D_q = D_k$), Value 向量 $V \in R^{D_v}$, 它们都是从同一个输入学习的,也可以认为 Q, K 和 V 只是三个独立的 MLP 层学习的三个特征向量。注意权值通过 Q 和 K 的点积计算,然后与 V 相乘得到变换后的表示。

$$\text{Attention}(Q, K, V) = \text{softmax} \frac{QK^T}{\sqrt{d_k}} V \quad (1)$$

此时,得出的 attention value 是一个向量值,代表了某一点的编码向量。该向量中包含了该点的上下文关系,既包含了全局联系同时也拥有了局部联系,这也是 attention 的强大优势之一,它可以同时一次性获得全局联系又可以获得局部联系。

2.2 KNN 搜索

3D 数据集为 $P = \{P_1, P_2, \dots, P_n \mid P_n \in R^3\}$, 在数据集 $P - P_n$ 中寻找每个点 K ($K=8$) 个邻近点, $\hat{P}_{n,k} = \text{KNN}(P_n \mid (P - P_n))$, $\hat{P}_{n,k}$ 表示点 P_n 的 k 个邻近点,因此 KNN 的邻近点集合可以表示为 $\{\hat{P}_{n,1}, \hat{P}_{n,2}, \dots, \hat{P}_{n,k} \mid \hat{P}_{n,k} \in R^3\}$ 。

获取的特征向量包括点 P_n 的坐标和对应点标识符 P_n 的 K 个最近邻,首先,每个点 P_n 通过减去其相关坐标归一化到 K 近邻集合 $\{\hat{p}_{n,1}, \hat{p}_{n,2}, \dots, \hat{p}_{n,k} \mid \hat{p}_{n,k} \in R^3\}$ 中,表示为公式(2):

$$U_{n,k} = P_n \oplus (P_{n,k} - P_n) n \in [1, k] U_{n,k} \in R^6 \quad (2)$$

上面部分的输出形成一个 $N \times K \times 6$ 的张量,其中 N 表示 3D 点集的点数, K 表示每个点的邻域数,最后 6 维是点 p 的坐标和第 k 个最近邻点的统一坐标。为了获得足够的表达能力来将每个点特征转换成更高维度的特征,后面加了全连接层,3D 点集表示为 $N \times K \times 64$ 的张量。因此 KNN 模块可以被视为学习从输入 3D 点集提取逐点低级几何特征的组件,如图 3 所示。此外,KNN 模块以逐点方式提取低层几何特征,解决了点云无序的问题,有效提高了分类任务的准确性。

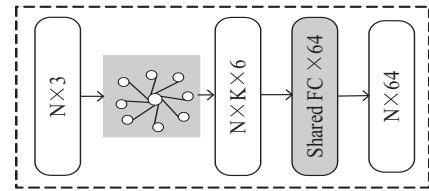


图 3 KNN 模块

2.3 关键点检测

一般情况下,关键点表示位于目标对象边缘或角上的点。现有的方法^[12]使用注意力模块来突出有利于识别的区域。在这种数据驱动的方法中,每个点的重要程度是自动学习的,然而由于缺乏真实感的关键点进行监督,使得难以区分哪个点更重要。笔者认为,利用点云的内在属性可以获得更准确的信息。众所周知,点的法线可以反映形状特征,所以通过考虑相邻区域法线的变化来为每个点分配一个响应。

$$D_r = \sum_{i \in N(r)} \sin(n_i, n_r) \quad (3)$$

n 表示点的法线。在进行全局最大池化操作之前,将计算得到的响应集成到点云的全局表示中。

2.4 VLAD 层

在本节中,主要描述了利用 VLAD 机制^[11]从 3D 点集中提取高级语义特征,如图 4 所示。VLAD 是一种流行的描述符池方法,用于实例级检索和图像分类。受 PointNetVLAD^[13]和 NetVLAD^[14]的启发,通过每个点的低级几何描述符和几个视觉单词之间的关系来间接描述高级语义特征。如图 4 所示,VLAD 模块主要

包括以下两个步骤:(1) top-K VLAD 特征选择;(2) 特征变换和融合。

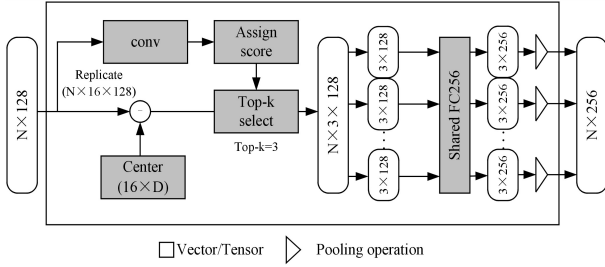


图4 VLAD 框架

2.4.1 top-K VLAD 特征选择

取 n 个低级几何特征描述符 $\{v_1, v_2, \dots, v_n \mid v_n \in R^D\}$ 作为 VLAD 模块的输入,其中 $D = 128$ 。同时,初始化 M 个视觉单词(“聚类中心”),表示为 $\{c_1, c_2, \dots, c_m \mid c_m \in R^D\}$ 。每个点的低级几何特征描述符 v_n 被分配给每个视觉单词 c_m ,由残差向量 $v_n - c_m$ 表示,该向量记录低级几何特征描述符和视觉单词之间的差异。第 n 个低级几何特征描述符 v_n 与 M 个视觉单词的关系表示为 r , r 的 (n, d) 关系计算如下:

$$r_{n,d} = \sum_{m=1}^M a_n(c_m) (v_{n,d} - c_{m,d}), d \in [1, D] \quad (4)$$

$a_n(c_m)$ 表示注意系数, $c_{m,d}$ 和 $v_{n,d}$ 分别表示在 m 的局部特征上的第 d 个维度和第 n 个低层几何特征描述子, $a_n(c_m)$ 是影响 $c_{m,d}$ 的权重。用低层几何特征描述子的 soft-assignment 来表示 $a_n(c_m)$, $a_n(c_m)$ 通过下面公式来计算:

$$a_n(c_m) = \frac{e^{W_{a_n}^T v_n + b_{a_n}}}{\sum_{m=1}^M e^{W_{a_n}^T v_n + b_{a_n}}} \quad (5)$$

三维点云中的每个点都可能与几个视觉单词有着重要的关系。因此,提供以下 top-K 定义的方法来考虑注意分数较高的视觉单词集合对高层几何特征的影响,给定低层几何描述算子 v_n 和视觉单词集合 $\{c_1, c_2, \dots, c_m \mid c_m \in R^D\}$,在经过 top-K 的选择后的子集 $\tilde{C} = \{\tilde{c}_1, \tilde{c}_2, \dots, \tilde{c}_k \mid \tilde{c}_k \in R^D\}$, $\tilde{C} \subseteq C$,对任何一个视觉单词 $\tilde{c} \in \tilde{C}$ 和 $c \in C - \tilde{C}$,有 $a_n(\tilde{c}) \geq a_n(c)$,在 top-k VLAD 模块中 r 的 (n, d) 的关系公式表示为:

$$r_{n,d} = \sum_{k=1}^{\text{top-K}} \frac{e^{W_{a_n}^T v_n + b_{a_n}}}{\sum_{k=1}^{\text{top-K}} e^{W_{a_n}^T v_n + b_{a_n}}} (v_{n,d} - c_{k,d}) \quad (6)$$

其中, $k \in [1, \text{top-K}]$ 和 $d \in [1, D]$ 。一方面, top-K 值控制剩余向量的数量,另一方面,它表示不同视觉单词之间的重叠。此外,为了改善网络的非线性变换,使用了共享的 FC 层。

2.4.2 特征转换和融合

top-kVLAD module 输出一个 $N \times 3 \times 128$ 张量,其

中 N 表示 3D 点集中的点数,3 表示 top-K 视觉单词的数量,每个点被表示为一个 3×128 矩阵。在 VLAD 模块中添加了一个 FC layer 层,经过 FC Layer 之后每个点都表示为 3×256 的矩阵,因此 3D 的数据集表示为 $N \times 3 \times 256$ 。然后将其输入到逐点的全局池化层,池化层为每个点生成高级语义特征。这里 VLAD 模块可以看作是学习从输入的 3D 点集中提取 $(n \times 256)$ 维高层几何特征的模块。VLAD 模块不仅解决了点云的无序问题,而且还有效提高了点云分类的准确性。

3 实验结果与评估

3.1 实验设置

该文提出的网络框架主要是在 ModelNet40^[15] 数据集上进行实验,数据集由来自 40 个类别的 12 311 个 CAD 模型组成,其中 9 843 个用于训练,2 468 个用于测试。由于设备限制,如果输入的点数过多将会延长运行时间,也会导致后期硬件散热问题无法长时间运行,实验中统一以 512 个点作为网络输入。设置初始学习率为 0.001,总的训练轮数为 250,在 Adam 优化器下 (momentum=0.9), Batch Size=32 参数下进行训练,实验配置如表 1 所示。

表1 实验设置

| CPU | GPU | CUDA | CUDNN | Windows | Tensor Flow | Python |
|---------|-----------|------|-------|---------|-------------|--------|
| i7-4790 | RTX-2060s | 10 | 7.4 | 10 | 1.14 | 3.6 |

3.2 K 值分析

本节主要分析了 KNN 模块中的 K 值对模型性能的影响。在 ModelNet40 上进行了大量实验,在其他条件不变的情况下,通过设置不同的 K 值以获取最佳的分类效果。 K 的大小和分类精度之间的关系如图 5 所示,由图可以看出当 epoch = 250, $K = 8$ 时,分类精度最佳。

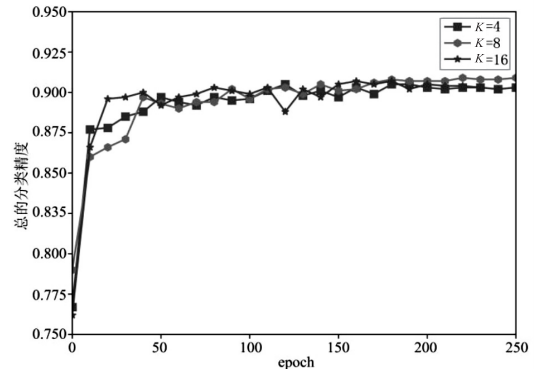


图5 不同 K 值的效果对比

由图 5 可以看出,当 K 值偏小时 (如 $K = 4$),在构建局部邻域时所搜索的邻近点数较少,“学习”的估计

误差也会随之增大,从而会降低点云分类精度。然而,随着 K 值的增大,所包含的不相关点数也会增加,由图可以看出, $K=16$ 时其分类精度比起 $K=8$ 时有所下降。因此,在构建局部邻域时, K 值的选取至关重要。将 K 值设为 8,其搜索的局部邻域包含了一定数量的点或特征,不仅避免了局部信息不足对点云分类结果产生的影响,还避免了不相关点数导致的较大学习近似误差。

3.3 模块消融分析

图 1 显示了 SA-PointNetVLAD 的网络架构,本节为了研究 SA-TNet 模块和 VLAD 模块的功效,在 ModelNet40 上进行了消融分析。

表 2 中第一行显示了只使用简单的共享层以及基础 KNN 模块,可以看出其性能很差, VLAD 模块的性能很差;第二行是基础模块 KNN 和 SA-TNet 模块的集成结果;第三行是基础模块 KNN 和 VLAD 模块的集成结果,第四行是基础模块 KNN、SA-TNet 和 VLAD 模块的集成结果。结果表明,在 KNN 模块的基础上单独引入 SA-TNet 和 VLAD 模块效果分别提升了 1.79 个百分点和 22.28 个百分点。当 SA-TNet 和 VLAD 模块全部整合在一起时,如最后一行所示, SA-PointNetVLAD 的分类精度达到 90.91%。

表 2 SA-TNet 模块、KNN 模块和 VLAD 模块对三维对象分类精度的影响 %

| KNN 模块 | SA-TNet 模块 | VLAD 模块 | 精度 | |
|-----------|---------------|------------|-------|-------|
| | | | 平均精度 | 总精度 |
| ✓ | - | - | 58.87 | 66.84 |
| ✓ | ✓ | - | 61.46 | 68.63 |
| ✓ | - | ✓ | 85.52 | 89.12 |
| ✓ | ✓ | ✓ | 88.11 | 90.91 |

实验结果分析如下:(1)虽然不同的模块对分类精度的贡献不同,但通过增加更多的模块可以进一步提高最终的分类精度;(2)在三维对象分类任务中,高层语义特征比低层几何特征更重要,充分说明 VLAD 模块有效提取了高层语义特征。

3.4 与其他方法的比较

本节主要展示了 SA-PointNetVLAD 从三维点云中提取特征时的效率。主要在 ModelNet40 数据集上设置了几组实验,分类模型的网络配置与 3.1 节中描述的相同。由于实验设备限制,主要将 numpoint 设置为 512,观察到实验结果与 pointNet 的对比效果(如图 6 所示), SA-PointNetVLAD 比 PointNet 在 ModelNet40 数据集上的分类精度明显要高。

为了进一步评估 SA-PointNetVLAD 模型的分类效果,基于 ModelNet40 数据集与几种先进的方法(包括 3DshapeNets^[16], OctNet^[17], VoxNet^[18], O-CNN^[6],

ECC^[19], So-Net^[20], A-SCN^[21], PointNet++^[2], PointNet^[1] 和 Kd-Net^[22])进行了比较。结果如表 3 所示。

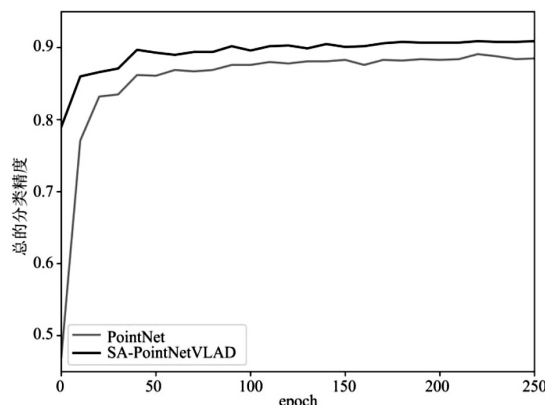


图 6 与 pointNet 的总分类精度对比

表 3 ModelNet40 数据集分类精度 %

| 网络 | 输入 | 精度 | |
|-----------------------------------|------------------|------|------|
| | | 平均精度 | 总精度 |
| 3DshapeNets ^[16] | 30 ³ | 77.3 | - |
| OctNet ^[17] | 128 ³ | 83.8 | 86.5 |
| VoxNet ^[18] | 32 ³ | 83.0 | - |
| O-CNN ^[6] | 64 ³ | - | 90.6 |
| ECC ^[19] | 1 000×3 | 83.2 | 87.4 |
| So-Net ^[20] | 2 048×3 | 87.3 | 90.9 |
| A-SCN ^[21] | 1 024×3 | 87.4 | 89.8 |
| PointNet ^[1] | 1 024×3 | 86.2 | 89.2 |
| PointNet++ ^[2] | 1 024×3 | - | 90.7 |
| Kd-Net (depth 10) ^[22] | 215×3 | 86.3 | 90.6 |
| Ours | 512×3 | 88.1 | 90.9 |

从表 3 中可以看出,由于该文在提取特征时考虑了点与点之间的几何关系,增强了描述局部信息和全局信息的能力, SA-PointNetVLAD 的性能优于其他所有只使用点作为 ModelNet40 数据集输入数据的方法。SA-PointNetVLAD 比 ECC^[19] 的分类精度高了 3.5 个百分点, SA-PointNetVLAD 比 So-Net^[20] 使用了相当小的输入数据大小,但它的效果却是差不多的。与基于体积的方法相比, SA-PointNetVLAD 获得了更好的性能,与基于点的方法相比,结果优于所有单模型结果,对于同时运用了 self-attention 机制的 A-SCN^[21], 分类精度高了 1.1 个百分点,而且 SA-PointNetVLAD 所需的输入数据大小仅为 512×3。从平衡性能和网络复杂性的角度来看, SA-PointNetVLAD 具有更多的优势。

4 结束语

针对三维点云模型中缺乏对各个点的局部信息和全局信息的有效利用,提出了一种新的点云分类框架 SA-PointNetVLAD,不仅在 PointNet 原始的框架上做

出了改进,还对整个网络框架进行了改进,最后点云分类精度达到了 90.9%。在实验过程中,发现 KNN 模块、SA-TNet 模块以及 VLAD 模块这三个模块都大大提升了网络框架的分类精度,并具有一定的实用性。在今后的工作中,将针对点云的实用性做进一步的提升,在考虑点云的法向量的同时,也可以考虑一些其他的影响因素。在检测出关键点后,没有做进一步的提取特征操作,而是简单的将结果输入到了 VLAD 模块中,后期可以考虑将这些部分进行进一步的处理以获得更好的效果。

参考文献:

- [1] QI C R, SU H, MO K, et al. PointNet: deep learning on point sets for 3D classification and segmentation[C]//Proceedings of the 2017 IEEE conference on computer vision and pattern recognition. Honolulu, HI: IEEE, 2017: 77–85.
- [2] 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 annual conference on neural information processing systems. Long Beach: Curran Associates, 2017: 5105–5114.
- [3] 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.
- [4] 党吉圣, 杨军. 深度图注意力 CNN 的三维模型识别[J]. 计算机科学与探索, 2021, 15(1): 141–149.
- [5] WANG Y, SUN Y, LIU Z, et al. Dynamic graph CNN for learning on point clouds[J]. ACM Transactions on Graphics, 2019, 38(5): 1–12.
- [6] 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): 1–11.
- [7] YANG Y, FENG C, SHEN Y, et al. FoldingNet: point cloud auto-encoder via deep grid deformation[C]//2018 IEEE/CVF conference on computer vision and pattern recognition (CVPR). Salt Lake City, UT, USA: IEEE, 2018: 206–215.
- [8] SHEN Y, FENG C, YANG Y, et al. Mining point cloud local structures by kernel correlation and graph pooling[C]//2018 IEEE/CVF conference on computer vision and pattern recognition (CVPR). Salt Lake City, UT, USA: IEEE, 2018: 4548–4557.
- [9] LI Y, BU R, SUN M, et al. PointCNN: convolution on x-transformed points[C]//Proceedings of the annual conference on neural information processing systems. Montréal: Curran Associates, 2018: 820–830.
- [10] WU W, QI Z, LI F. PointConv: deep convolutional networks on 3D point clouds[C]//2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR). Long Beach, CA, USA: IEEE, 2019: 9621–9630.
- [11] ZHANG D, HE F, TU Z, et al. Pointwise geometric and semantic learning network on 3D point clouds[J]. Integrated Computer-Aided Engineering, 2019, 27(1): 57–75.
- [12] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Advances in neural information processing systems. Long Beach, USA: IEEE, 2017: 5998–6008.
- [13] UY M A, LEE G H. PointNetVLAD: deep point cloud based retrieval for large-scale place recognition[C]//2018 IEEE/CVF conference on computer vision and pattern recognition. Salt Lake City, UT, USA: IEEE, 2018: 4470–4479.
- [14] ARANDJELOVIC R, GRONAT P, TORII A, et al. NetVLAD: CNN architecture for weakly supervised place recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 40(6): 1437–1451.
- [15] CIGNONI P, CALLIERI M, CORSINI M, et al. MeshLab: an open-source mesh processing tool[C]//Eurographics Italian chapter conference (EICC). Salerno, Italy: The Eurographics Association, 2008: 129–136.
- [16] WU N Z, SONG S, KHOSLA A, et al. 3D ShapeNets: a deep representation for volumetric shapes[C]//Proceedings of the 2015 IEEE conference on computer vision and pattern recognition. Boston, MA: IEEE, 2015: 1912–1920.
- [17] RIEGLER G, ULUSOY A O, GEIGER A. OctNet: learning deep 3D representations at high resolutions[C]//2017 IEEE conference on computer vision and pattern recognition. Honolulu, HI, USA: IEEE, 2017: 3577–3586.
- [18] 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. Los Alamitos: IEEE, 2015: 922–928.
- [19] SIMONOVSKY M, KOMODAKIS N. Dynamic edge-conditioned filters in convolutional neural networks on graphs[C]//2017 IEEE conference on computer vision and pattern recognition (CVPR). Honolulu, HI, USA: IEEE, 2017: 29–38.
- [20] LI J, CHEN B M, LEE G H. SO-Net: self-organizing network for point cloud analysis[C]//Computer vision and pattern recognition (CVPR). Salt Lake City, UT, USA: IEEE, 2018: 9397–9406.
- [21] LIU S. Attentional ShapeContextNet for point cloud recognition[C]//IEEE/CVF conference on computer vision & pattern recognition. Salt Lake City, UT, USA: IEEE, 2018.
- [22] KLOKOV R, LEMPITSKY V. Escape from cells: deep kd-networks for the recognition of 3d point cloud models[C]//Proceedings of the IEEE international conference on computer vision. Venice, Italy: IEEE, 2017: 863–872.