

对融合特征敏感的三维点云识别与分割

朱安迪^{1,2}, 达飞鹏^{1,2}, 盖绍彦^{1,2}

(1. 东南大学自动化学院, 210096, 南京;

2. 东南大学复杂工程系统测量与控制教育部重点实验室, 210096, 南京)

摘要: 三维点云分类分割网络忽视了融合特征中的冗余信息, 缺乏放大有效特征占比能力, 不能充分挖掘特征的表达性。在 CurveNet 网络基础上, 提出了一种能够筛选和丰富融合特征的方法, 对点云的识别与分割效果达到了较先进水平。首先, 提出了对融合特征具有筛选能力的特征选择子网络, 利用结合了打分机制的 TopK 算子选出包含有效信息的融合特征, 并且能够自适应地赋予被选特征权重。其次, 在聚合曲线特征模块中增加了两个新分支, 分别学习曲线内部点距离特征和曲线之间的线距离特征, 通过快速通道相关性注意力机制提取各分支的内部相关性, 增强了网络特征的信息描述能力。实验结果表明, 分类任务在 ModelNet40 数据集上准确率达到 93.8%, 分割任务在 ShapeNet Part 数据集上平均交并比达到了 86.4%。与基准网络相比, 分类效果与分割效果均有所提高, 证明了算法的有效性。

关键词: 三维点云; 融合特征筛选; 曲线特征; 注意力机制; 分类分割

中图分类号: TP391 **文献标志码:** A

DOI: 10.7652/xjtuxb202405006 **文章编号:** 0253-987X(2024)05-0052-12

Recognition and Segmentation of 3D Point Clouds Sensitive to Fusion Features

ZHU Andi^{1,2}, DA Feipeng^{1,2}, GAI Shaoyan^{1,2}

(1. School of Automation, Southeast University, Nanjing 210096, China;

2. Key Laboratory of Measurement and Control of Complex Systems of Engineering of Ministry of Education,
Southeast University, Nanjing 210096, China)

Abstract: The 3D point cloud classification and segmentation networks ignore the redundant information in the fusion features, lack the ability to amplify the proportion of effective features, and cannot fully explore the expressiveness of features. Based on the CurveNet network, this paper proposes a method that can filter and enrich the fusion features, and the recognition and segmentation effect of point cloud reaches a relatively advanced level. Firstly, a feature selection subnetwork with filtering ability for fusion features is proposed, which combines TopK operator and a scoring mechanism to select fusion features containing valid information and adaptively assign weight to the selected features. Secondly, two new branches are added to the aggregation curve feature module, so as to learn the curve internal point distance features and the curve line distance features, respectively, and extract the internal correlation of each branch through the

quick channel affinity attention mechanism, which enhances the information description ability of the network features. The experimental results show that the accuracy of the classification task on the ModelNet40 dataset reaches 93.8%, and the average intersection over union of the segmentation task on the ShapeNet Part dataset reaches 86.4%. Compared with the benchmark network, the classification effect and segmentation effect are improved, which proves the effectiveness of the proposed algorithm.

Keywords: 3D point cloud; fusion feature screening; curve feature; attentive mechanism; classification and segmentation

物体分类和分割是计算机视觉中基本且重要的任务,其中二维图像识别和分割技术已取得了较好的发展,但实际效果易受光照和遮挡等因素影响。随着激光雷达和深度相机等三维测量工具的发展与普及,三维点云识别和分割在自动驾驶、机器人和增强现实等领域获得了广泛地应用。但是,基于三维测量工具获取的点云数据是无序且密度不均匀的,这极大地影响了点云特征的提取,对三维目标检测、场景重建等下游任务造成了一定干扰。长程全局特征含有整体空间结构信息,短程局部特征敏感于细微空间几何形状以及密度的变化,因此有效捕获局部特征和全局特征至关重要。

点云数据包含三维坐标、法向量、颜色和时间等信息。目前,针对点云数据直接提取特征的方法主要包含以下3种:①基于点的特征提取方式^[1-6],通过近邻点采样搜索中心点的相邻特征点,在PointNet网络^[1]和PointNet++网络^[2]的基础上,衍生出多种局部特征构造和提取方式来多层次细化相邻特征,利用聚合函数逐级汇聚中心点的相邻特征;②基于卷积的特征提取方式^[7-12],将点的不规则近邻点集比作二维空间中的像素,通过近邻点集内部的距离信息^[7]、潜在顺序^[8]和形状信息^[9]等特点构造特定的三维卷积核,获取近邻点集信息;③以点云动态图卷积网络(dynamic graph CNN, DGCCN)^[13]为代表的基于图的特征提取方式^[13-17]在近邻点集中设计点云的边关系,将边特征作为目标任务的组成特征,动态构建边的近邻域并聚集相应的边特征,利用图卷积探索边特征的局部结构。

边特征能在一定程度上描述点云的局部组件结构,与对距离信息更具表述能力的点云特征互为补充。PRCDRNet网络^[18]和CurveNet网络^[19]等将边特征和点特征作为基础组成特征进行融合,增加特征描述符的语义信息。CurveNet网络利用融合特征间的距离和方向关系,设计了一种长距离的特征提取方法,以相邻融合特征之间的夹角为导向,在

融合特征同构图上建立具有几何信息的曲线特征。基于卷积的特征提取方法,在利用特定卷积动态提取特征的过程中需要耗费大量的时间和计算资源。基于点的特征提取方法和基于边的特征提取方法仅利用点特征或边特征使网络缺乏描述能力,结合点特征和边特征可以增加特征的信息含量,并且避免了构造特定卷积带来的计算成本的增加。但是,边特征利用了点特征之间的距离关系,直接融合点特征和边特征会导致部分信息重叠,且没有针对性地增强融合特征中与目标任务相关的特征向量,对相关度低的特征在总体特征中的占比无正相关作用。所以,增强网络对有效融合特征的提取能力具有重要的研究意义。

为解决上述问题,本文以CurveNet网络作为基准网络,设计了一种对融合特征具有选择性和表达性的三维点云识别和分割网络。主要贡献和创新点有以下3点:①为了加强网络对融合特征的敏感性,提出了特征选择子网络模块,在选择融合特征时自适应地判断融合特征的描述性,放大被选特征的信息表达程度。②为了更好地利用曲线特征包含的语义信息,设计了两种曲线特征描述方法,增加了网络内部信道特征种类。③设计了结合快速注意力机制的自注意力机制模块,在各通道间迅速聚合特征相关性,加强了特征内部信息的表达性。

1 相关研究

1.1 基于点的特征提取方式

PointNet网络和PointNet++网络作为直接处理点云特征的开山之作,为后续点云研究工作提供了开创性启发。PointSIFT网络^[3]认为 K 近邻KNN算法对近邻点集的探寻方向单一,没有发挥出三维点云的空间优势,利用尺度不变特征变换匹配算法(SIFT)构建了具有8个方向搜索能力的方向编码单元。Mo-Net网络^[4]将刚体几何矩阵引入点云,利用多项式函数拟合点云的表面形状关系。

PointWeb 网络^[5]使用影响函数和关系函数,建立局部特征间的联系,逐元素地动态搭建特征差异图,自适应地调整局部特征。王溪波等^[6]构建双边结构,处理点云的语义信息和几何信息,设置多种池化单元跨层提取特征信息。

基于点的特征提取方法通过多种方式加强局部特征的描述能力,并关注全局特征、整体表面形状信息和颜色信息等,弥补局部描述符对整体特征的概括局限性。

1.2 基于卷积的特征提取方式

二维图像中卷积核提取特征的本质是在像素区域中进行互相关运算,点云局部特征实际是相邻点集内特征间的相关性表达。因此,利用特定的卷积核能够针对性捕获相邻点集的相关性。文献[8]基于X变换设计了将点云顺序与点云特征分离的X卷积,用于聚合中心点的近邻特征集。文献[7]为了避免点云采样过程中重叠采样,提出了可以设定任意尺寸的环形卷积核。文献[10]利用多层感知机(MLP)强大的学习能力模拟连续的三维卷积核,动态调整权重系数。文献[11]受二维图像可变性卷积的启发,设计了跳过多层感知机学习权重参数的核心点卷积。彭秀平等^[12]将点云特征划分为包含27个子立方体的大立方体,设计立方体卷积对有序子立方体内的点集进行高维特征抽取。

基于卷积的特征提取方式将二维算法迁移到点云中,对点云不规则、密度不均匀等特点针对性地设计局部区块划分方式,利用相应卷积核捕获局部区块内的特征信息。

1.3 基于图的特征提取方式

图编码将点云特征的全局坐标转换为相对坐标,利用边缘卷积,聚集边特征的语义信息,不断更新包含特征信息的特征图。DGCNN 网络^[13]在PointNet 网络的基础上设计了边缘卷积,将图概念运用在点云上,以边特征代替点特征,动态聚合局部几何特征。SPG 网络^[14]层次化地使用边缘特征,几何分割点云构建超点图,为边卷积设置条件的同时利用门控循环单元(GRU)提取目标整体上下文关系。GACNet 网络^[15]在图卷积中搭建注意力机制获得中心点与邻域点集的权重参数,使图卷积聚焦于邻域特征的内部相关性。文献[16]提出点云严格旋转不变映射,利用离差平方和约束边特征以构造多层次聚类簇,表达点云的分布特性。PointMLP 网络^[17]不以局部特征提取为重点,认为网络架构是表达目标特征的关键,通过堆叠几个简单的前馈残差

网络提取边特征的局部关系。

基于图的特征提取方法在DGCNN 网络的基础上衍变出多种更新特征图的方式,不断提升局部特征提取器及网络架构的性能。

2 对融合特征敏感的三维点云网络

本文使用 CurveNet 网络作为骨干网络。CurveNet 网络由1个局部点特征聚集模块、4个堆叠曲线干预卷积模块和1个任务指定头模块串联组成。首先通过局部点特征聚集模块学习输入点云坐标的相对局部编码,将相对点云特征和边特征投影到更高的维度以此来聚合局部特征;接着将局部特征传入4个堆叠曲线干预卷积模块中构建由曲线特征、点特征和边特征组成的融合特征;最后利用任务指定头模块对融合特征进行分类预测和分割预测。CurveNet 网络最大的贡献是利用局部特征间的几何关系构建曲线特征,提供了一种新的描述物体几何结构的局部特征聚合方法。

本文为了加强融合特征的表达性,在 CurveNet 网络原堆叠曲线干预卷积模块中的曲线聚合模块内设计了两种新的曲线特征,同时针对这两种曲线特征添加了相应的特征学习分支,并在该模块内加入了自注意力机制模块,以加强曲线特征的表达性。为了使融合特征更准确地表达出目标信息的同时减少冗余信息,在 CurveNet 网络的骨干架构上添加了4个特征选择子网络模块。特征选择子网络模块获取相关性最大的前K个高维特征向量,并将其作为被选特征与曲线特征进行特征维度的拼接,得到融合特征,再将融合特征送入简单的多层感知机中学习高维语义信息。在高维特征图中,特征选择子网络筛选有效融合特征的同时兼顾了原始网络结构和原始多级特征。整体网络结构如图1所示。本节将着重介绍特征选择子网络模块的网络结构、堆叠曲线干预卷积模块内曲线聚合模块的网络结构。

2.1 特征选择子网络模块

堆叠曲线干预卷积模块构造的曲线特征包含了几何信息,局部点特征聚集模块基于原始点云提取的高维局部边特征和点特征包含了语义信息。CurveNet 网络将点特征、边特征和曲线特征融合在一起,增加了特征的信息含量。但是,直接利用多层感知机对这些特征进行融合操作并不能最大限度地发挥融合特征的优势,且在点特征、边特征和曲线特征中出现与目标相关度低的信息重叠现象,这在

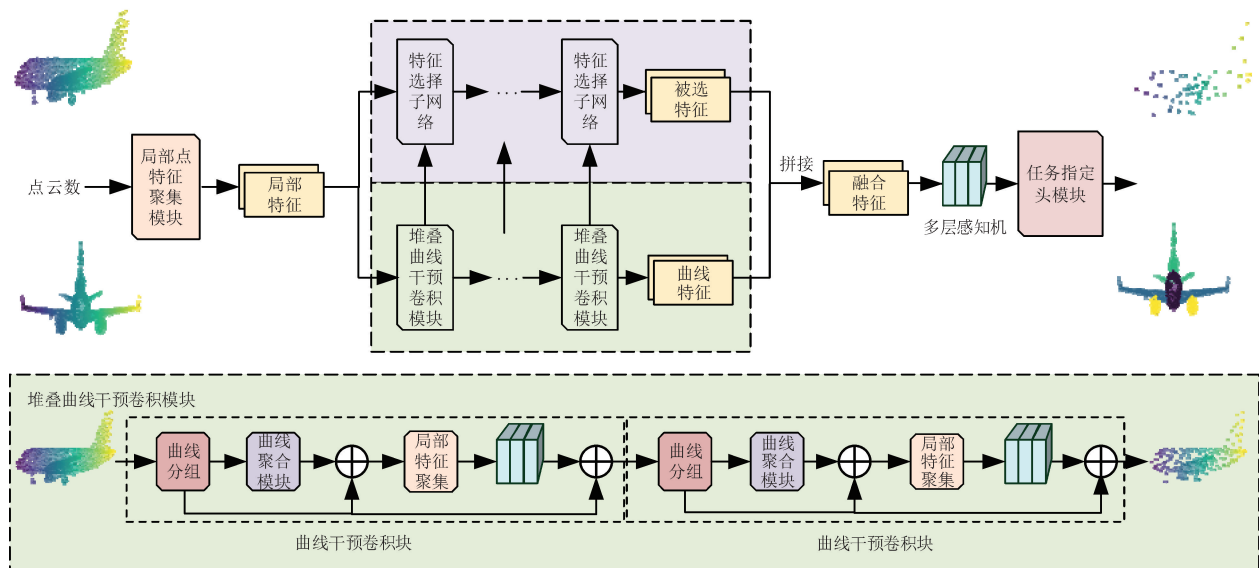


图 1 对融合特征敏感的网络结构示意图

Fig. 1 Network architecture based on sensitive to fusion features

一定程度上会对结果产生无效影响。受文献[20]将 TopK 算子用于网络输出端筛选匹配度高的源点云交互特征和目标点云交互特征的启发,本文设计了对特征有选择性的特征选择子网络。该网络旨在逐级提取不同阶段的有效融合特征,过滤相关度低的特征,留下与目标任务相关性高的特征并赋予权重,以增强网络对融合特征的敏感程度。

图 2 为特征选择子网络模块筛选融合特征的具体操作。特征选择子网络模块包含两个阶段,阶段 1 的输入特征 f_{in1} 是局部点特征聚集模块的输出特征或上一级特征选择子网络模块的输出特征,阶段 2 在接收阶段 1 输出特征的同时,将对应层级的堆叠曲线干预卷积模块输出特征作为另一个输入特征 f_{in2} 。

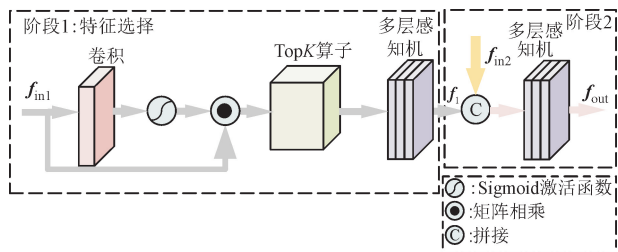


图 2 特征选择子网络结构示意图

Fig. 2 The structure of the TopK feature select block

阶段 1 中,单凭 TopK 算子虽然可以选出任意数量的融合特征,但是不能保证筛选过程中被选特征的有效性,也无法增强网络的敏感程度。所以,通过一维卷积和 Sigmoid 函数构造筛选机制,对 f_{in1} 在

网络中的作用打分,将得分作为筛选原则和权重,反馈给 f_{in1} ,加强相关度高的特征,削弱相关度低的特征,筛选出得分最高的前 K 个被赋予权重的特征,接着利用一个简单的多层感知机提取出高维特征。阶段 1 的内部过程如下

$$M = \gamma(\theta^m(\tau_{1 \times 1}^{n \rightarrow m}(x))) \quad (1)$$

$$f_1 = M(T(f_{in1} \cdot \delta(\tau_{1 \times 1}^{n \rightarrow 1}(f_{in1})))) \quad (2)$$

式中: M 为多层感知机提取特征过程; τ 为卷积操作,上标 $n \rightarrow m$ 表示卷积核的输入特征维度为 n ,输出特征维度为 m ,下标 1×1 为卷积核尺寸; θ 为批归一化处理操作,上标 m 为输入特征维度; γ 为 Relu 激活函数; δ 表示 Sigmoid 函数; $\cdot$ 表示点积操作; T 表示利用 TopK 算子设定选取点的参数,得到包含权重特征的前 K 个索引量; f_1 为阶段 1 的输出特征。

阶段 2 中,为了提高特征选择子网络模块对整体特征的选择性,将输出特征 f_{in2} 与阶段 1 的输出特征 f_1 结合,利用多层感知机进行高维特征提取。该过程表示如下

$$f_{out} = M(f_{in2} \oplus f_1) \quad (3)$$

式中: $\oplus$ 表示特征拼接操作,该操作将维度为 $B \times D \times N$ 的 f_{in2} 与 f_1 在 D 维度上进行拼接,其中 B 为批大小, D 为特征维度, N 为点云数量。

2.2 堆叠曲线干预卷积模块

堆叠曲线干预卷积模块由两个曲线干预卷积模块串联而来。在曲线干预卷积模块中输入特征先通过曲线分组子模块捕获一定数量的曲线特征,再利用曲线多特征聚合网络得到融合后的曲线特征,最

后重新设置局部点特征聚集模块提取高维融合特征。为加强曲线特征的表达性,在曲线多特征聚合网络模块中构造了两种新的曲线特征,设计对应的支路提取其与点、边特征的映射关系,同时设计了快速通道相关性注意力机制挖掘特征间的内部相关性。

2.2.1 曲线特征构造

CurveNet 对曲线特征 C_i 的定义如下:在维度为 $B \times 3 \times N$ 的点云特征 P 与维度为 $B \times D \times N$ 的边特征 F 的基础上,利用 KNN 算法得到近邻点集 $\{P_i^1, \dots, P_i^L\} \in \mathbf{R}^D$ 。在近邻点集中搜寻起始点与其余点的角度,选取角度合适的点作为该曲线特征的组成点特征,重复此操作直至选出一定数量的组成点集。曲线特征就是该组成点集构成的折线,它的尺寸为 $B \times D \times N \times L$, L 为曲线特征的组成点集数。沿用这个定义,在该曲线特征的基础上构造新的曲线特征,结合原有曲线特征一起描述目标任务。

曲线特征与点特征、边特征相比,包含了更为丰富的几何特征。在点云分类网络中,曲线特征可以更好地根据物体整体形状信息捕捉长距离特征,从物体整体结构上聚合长程特征。点云分割网络需要精准捕获物体部件的特征,利用曲线特征描绘物体各部件,可以更好地把握物体各部件的短程特征,通过部件不同的形状特点加强自身的分辨率。曲线内部点与点之间的距离可以代表曲线本身的长短和曲线的曲折程度,曲线和曲线之间的距离可以在一定程度上表达出曲线外部距离特征和结构特征。因

此,本文设计了基于曲线内部的点距离特征和基于曲线之间的线距离特征。

KNN 算法在点云中基于点到点的距离寻找中心点的近邻点集。为了使曲线特征能够利用 KNN 算法寻找近邻曲线特征集,将曲线特征看成点特征。曲线特征由 L 个维度为 D 的点特征构成,看作维度为 $L \times D$ 的点特征,此时曲线特征转换为尺寸是 $B \times (D \times L) \times N$ 的点特征。接着基于 KNN 算法找出中心曲线特征对应的 K 个维度为 $L \times D$ 的近邻点集,将其逆转为近邻曲线特征集 $\{C_i^1, \dots, C_i^K\}$ 。利用每条中心曲线特征和它的近邻曲线特征集构建曲线特征间边的关系,此时边的尺寸为 $B \times (D_i - D_i^l) \times N \times L$,将边和曲线特征进行维度拼接得到曲线间线距离特征 C_{intra} 。曲线内部点距离特征 $P_{\text{inter}} \in \mathbf{R}^{B \times (D_i^L - D_i^1)}$ 是曲线特征内部其它点到起始点的距离特征。图 3 给出了这两种曲线特征的构造过程,该过程可以如下表示

$$P_{\text{inter}} = P_i^l - P_i^1, \quad l \in C(i) \quad (4)$$

$$C_{\text{intra}} = C_i \oplus C_i^j - C_i, \quad j \in N(i) \quad (5)$$

式中: $C_i (i \in C(n))$ 代表着 N 条曲线特征中的第 i 条中心曲线特征; $C_i^j (j \in N(i))$ 表示第 i 条中心曲线特征的第 j 个近邻曲线特征, $N(i)$ 表示曲线特征 C_i 的近邻曲线特征域; $P_i^l (l \in C(i))$ 为中心曲线特征 C_i 的第 l 个组成点, $C(i)$ 是中心曲线特征 C_i 的组成点集, P_i^1 是中心曲线特征 C_i 的第一个组成点特征。

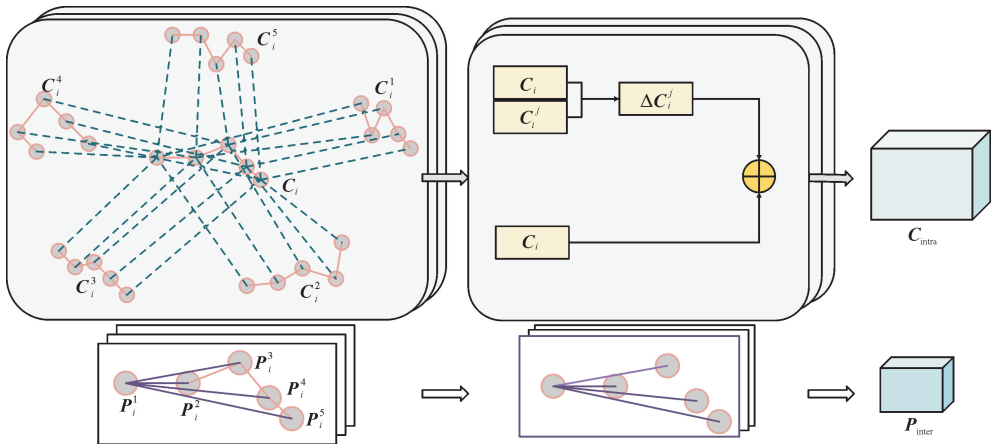


图 3 曲线特征构造过程

Fig. 3 The curve feature construction process

2.2.2 曲线多特征聚合网络模块

通过构造曲线特征,得到了由原曲线特征、曲线外部线距离特征和曲线内部点距离特征组成的多种曲线特征。CurveNet 网络通过两个分支提取原有

曲线特征的线内关系和线外关系,单独处理点、边特征,并与曲线特征相乘获得映射关系。本文在此基础上添加两个分支学习曲线内部点距离特征、曲线间线距离特征与点、边特征的映射关系。

图 4 为曲线多特征聚合网络模块结构示意图,表示了这几个分支对曲线特征和点、边特征映射关系的提取过程。曲线多特征聚合网络模块中的曲线边卷积和点边卷积为 2.2.1 节中的曲线间线距离特征和曲线内部点距离特征构造过程。这两个特征分支与其他分支结构相似,利用特征映射分支提取出各曲线特征的映射关系后,接着利用两个连续的多层感知机柔性融合曲线特征,最后模仿残差网络结构,将原始边特征与融合后的曲线特征相加。此外,曲线特征和近邻曲线特征之间在线的结构排列上存在关联,在曲线间线距离特征分支中得到曲线间线距离特征 C_{intra} 后,采用三维卷积层和三维批归一

化层,提取中心曲线特征和近邻曲线特征之间的结构排列关系。以下为两个分支的构造过程

$$P_{inter} = B_{mapping}(E_{p-conv}(C)) \quad (6)$$

$$C_{mid} = \theta^{n/2}(\tau_{1 \times 1}^{n \rightarrow n/2}(E_{c-conv}(C))) \quad (7)$$

$$C_{intra} = B_{mapping}(\beta(C_{mid})) \quad (8)$$

式中: C 为曲线特征; $\tau^{n \rightarrow n/2}$ 是卷积核为 1×1 的三维卷积层,输入特征维度为 n ,输出特征维度为 $n/2$; β 表示最大池化操作; E_{p-conv} 是点边卷积构造曲线内部点距离特征 P_{inter} 过程; E_{c-conv} 是曲线卷积构造曲线间线距离特征 C_{intra} 过程; C_{mid} 是曲线特征通过曲线边卷积、卷积层与批归一化层后得到的中间曲线特征; $B_{mapping}$ 表示各个分支提取曲线特征的和点、边特征映射过程。

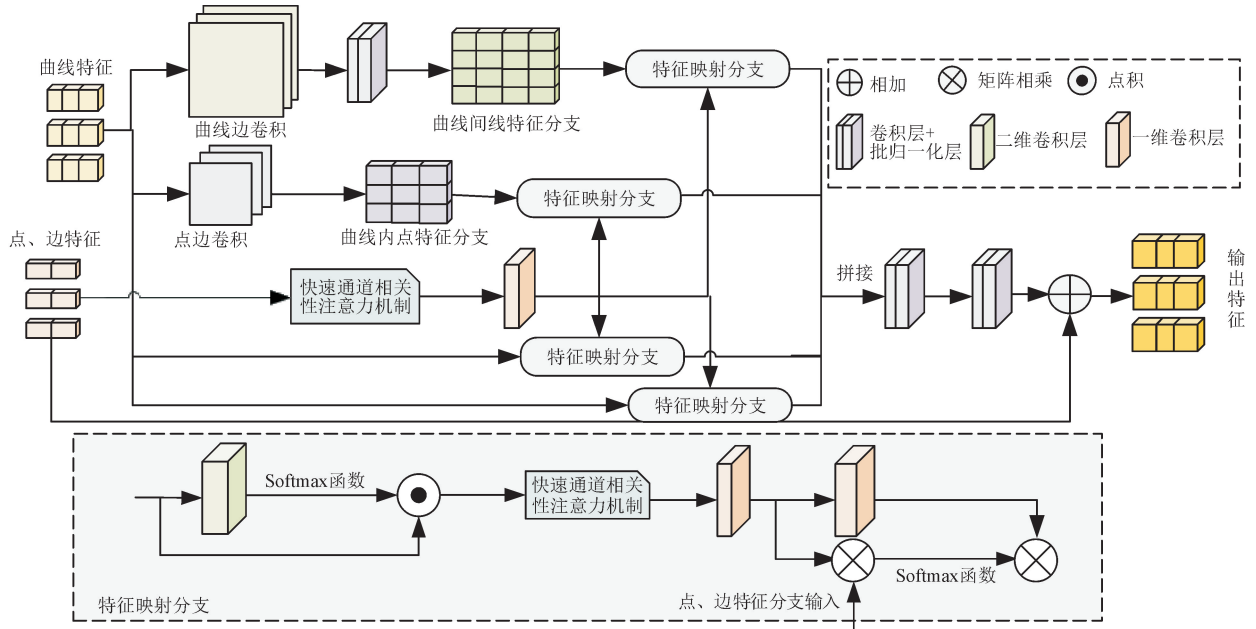


图 4 曲线多特征聚合网络结构示意图

Fig. 4 The structure of curve multi-feature aggregation network

在特征映射分支 $B_{mapping}$ 中,利用二维卷积层和 Softmax 函数为 C_{intra} 和 P_{inter} 赋予权重,再通过快速通道相关性注意力机制提取上述特征的内部相关性,利用独立的多层感知机进行特征维度对齐,最后与点、边特征矩阵相乘得到各自的映射关系。这一过程可以表示如下

$$f_{logit} = \tau_{1 \times 1}^{n/2 \rightarrow n/4}(A_{QCA}(f)) \quad (9)$$

$$f_{mid} = \tau_{1 \times 1}^{n/2 \rightarrow n/4}(A_{QCA}(\sum(C_{in} \cdot \phi(\tau_{1 \times 1}^{n/2 \rightarrow 1}(C_{in})))) \quad (10)$$

$$f_{out} = \tau_{1 \times 1}^{n/4 \rightarrow n/4}(f_{mid}) \times \phi(f_{logit} \times f_{mid}) \quad (11)$$

式中: f 为边特征向量和点特征向量; C_{in} 为相应支路转换后的曲线特征; f_{logit} 是经过维度转换的边特征,与经过维度转换的曲线中间特征 f_{mid} 矩阵相乘; ϕ 为 Softmax 函数; A_{QCA} 为快速通道相关性注意力机制

提取特征过程。在构建特征映射的时候为了使曲线特征与边、点特征维度对齐,对卷积层的输入输出特征维度进行了相应设置。

2.2.3 快速通道注意力机制

为了进一步提高曲线多特征聚合网络模块中多种曲线特征的描述能力,利用快速通道相关性注意力机制提取各支路特征的内部相关性。快速通道相关性注意力机制借鉴了通道相关性注意力机制(CAA)^[21],结构如图 5 所示。输入特征被划分为查询矩阵、键矩阵、值矩阵,利用 3 个独立的多层感知机对查询矩阵、键矩阵、值矩阵进行特征提取和特征压缩;接着利用 3 个快速注意力通道(QA)^[22] 赋予特征压缩后的查询矩阵、键矩阵和值矩阵的权重,进一步放大其内部特征关系。在提取查询矩阵与键矩

阵之间的映射矩阵 $\mathbf{Z}$ 的过程中,先将得到的查询矩阵和键矩阵相乘得到相关性矩阵 $\mathbf{S}$,利用最大池化单元得到相关性矩阵最大值,与相关性矩阵相减,再使用 Softmax 函数得到映射矩阵 $\mathbf{Z}$ 。“激励”操作中,被赋予权重的值矩阵和映射矩阵 $\mathbf{Z}$ 相乘,再次通过快速注意力通道加强内部特征,与原始输入特征向量相加。具体过程可如下表示

$$\mathbf{S} = [\mathbf{A}_{Qq}(\mathbf{M}_q(\mathbf{f}))]^T \times \mathbf{A}_{Qk}(\mathbf{M}_k(\mathbf{f})) \quad (12)$$

$$\mathbf{Z} = \phi(\beta(\mathbf{S}) - \mathbf{S}) \quad (13)$$

$$\mathbf{f}_{out} = \mathbf{f} + \mathbf{A}_Q(\mathbf{Z} \times \mathbf{A}_{Qv}(\mathbf{M}_v(\mathbf{f}))) \quad (14)$$

式中: $\mathbf{M}_q$ 、 $\mathbf{M}_k$ 、 $\mathbf{M}_v$ 分别为查询矩阵、键矩阵和值矩阵的多层感知机; $\mathbf{A}_{Qq}$ 、 $\mathbf{A}_{Qk}$ 、 $\mathbf{A}_{Qv}$ 分别为查询矩阵、键矩阵和值矩阵的快速注意力通道; $\mathbf{S}$ 、 $\mathbf{Z}$ 表示相关性矩阵和映射关系矩阵。

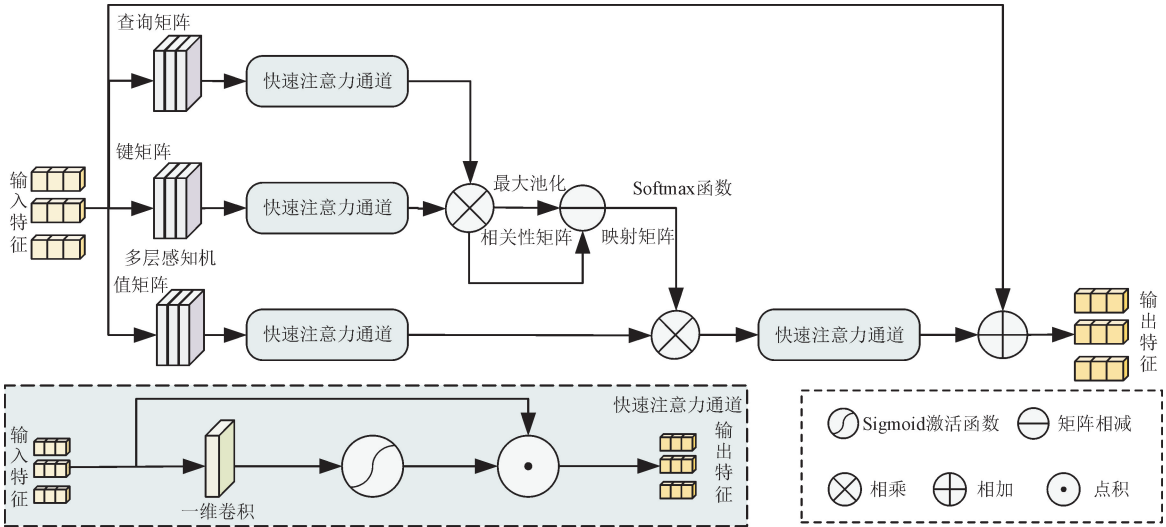


图 5 快速通道注意机制结构示意图

Fig. 5 The structure of the quick channel affinity attention

快速注意力通道利用一维卷积层和 Sigmoid 激活函数得到输入特征的权重系数,将输入特征与权重系数相乘以达到加强有效输入特征这一目的,从而得到输入特征的内部相关性。快速注意力通道结构如图 5 所示,这一过程可以用以下公式表示

$$\mathbf{f}_{out} = \delta(\tau_{1 \times 1}^{\mathbf{f}}(\mathbf{f})) \cdot \mathbf{f} \quad (15)$$

3 实验结果与分析

为了验证对融合特征敏感的曲线多特征聚合网络的有效性,在 ModelNet40 数据集上进行点云分类实验,在 ShapeNet Part 数据集上进行点云部件分割实验,点云分类和部件分割的实验环境和实验参数如表 1 所示。

表 1 实验环境与参数

Table 1 Experimental environment and parameters

环境	点云分类	部件分割
操作系统	Ubuntu18.04	Ubuntu18.04
Pytorch 版本	1.6.0	1.6.0
CUDA 版本	10.1	10.1
优化器	SGD	SGD
学习率	0.001	0.000 5
训练轮数	200	200

3.1 点云分类实验

ModelNet40 数据集常用于测试点云分类,包含 40 个类别的 CAD 模型,共 12 311 个样本。将点云数据送入网络之前,需要对点云数据进行预处理操作。预处理操作中,为了减少输入点云数量以及提高输入点云的质量,需要对点云进行采样;为了增强网络对旋转等情况的鲁棒性,需要对点云数据进行数据增强操作。与基准网络的性能进行比较时,需要排除数据预处理操作与损失函数种类的干扰,因此本文采取与基准网络相同的数据预处理操作以及损失函数种类。先对 ModelNet40 数据集所有样本中的点云坐标进行均匀采样,采样数量设置为 1 024,再对采样点云坐标进行缩放和平移处理。在网络中使用 KNN 算法捕获近邻点、边特征和近邻曲线特征时,为了排除参数 K 对识别准确率的影响,令其与基准网络保持一致($K=20$)。网络的损失函数采用三维点云分类算法中常见的交叉熵损失函数。最后,固定随机数种子以保证实验过程不受随机性的影响。

本文选取了近几年基于点、卷积和图的特征提取方法中较为经典的网络进行实验,并将其与本文

网络进行了对比分析,表 2 为各网络在 ModelNet40 数据集上的实验结果。本文方法在 ModelNet40 数据集上的识别精度达到了 93.8%,与基准网络相比,识别精度提高了 0.7%;与其他网络相比,取得了较为先进的结果。其中,本文网络仅使用 1 024 个原始点云坐标作为输入向量,输入向量种类少于 PointWeb 网络、PointConv 网络和 SO-Net 网络等将原始点云坐标和法线作为输入特征的网络,原始点云坐标数量也远少于 SO-Net 网络,且与上述网络相比,本文模型获得了更高的识别精度。这意味着在输入参数的种类以及数量更少的情况下,本文网络可以获得更准确的目标任务信息。主要原因是,特征选择子网络模块在筛选有效融合特征时,自适应地放大了有效融合特征的占比,过滤了与目标任务无关的特征,加强了网络对有效特征的提取能力。同时,在输入特征种类单一的情况下,构造多种曲线特征,采用注意力机制提取特征内部的相关性,丰富了网络特征的描述性。

表 2 不同网络在 ModelNet40 数据集中的识别效果对比

Table 2 Comparison of the recognition accuracy of different networks on the ModelNet40 dataset

网络	输入	点云数量	准确率/%
PointNet ^[1]	点云坐标	1 024	89.2
PointNet++ ^[2]	点云坐标	1 024	90.5
PointWeb ^[5]	点云坐标+法线	1 024	92.3
PointCNN ^[8]	点云坐标	1 024	92.2
RS-CNN ^[9]	点云坐标	1 024	92.9
PointConv ^[10]	点云坐标+法线	1 024	92.5
DGCNN ^[13]	点云坐标	1 024	92.9
GBNet ^[21]	点云坐标	1 024	93.8
PointASNL ^[23]	点云坐标	1 024	92.9
PAConv ^[24]	点云坐标	1 024	93.2
PCT ^[25]	点云坐标	1 024	93.2
SO-Net ^[26]	点云坐标+法线	5 000	92.5
Point-BERT ^[27]	点云坐标	1 024	93.2
Point-PN ^[28]	点云坐标	1 024	93.8
IAE ^[29]	点云坐标	1 024	93.2
I2P-MAE ^[30]	点云坐标	1 024	93.4
CurveNet ^[19]	点云坐标	1 024	93.1
本文	点云坐标	1 024	93.8

注:加粗数据为识别准确率最高值。

本文网络将特征提取过程划分为 4 个阶段,对应的特征筛选数量为 1 024→1 024→256→64。图 6 为网络后 3 个阶段关于 ModelNet40 数据集的融合

特征,筛选的特征数量分别为 1 024、256、64。在后 3 个阶段中,被选特征数量逐渐减少,但依然可以清晰地分辨出目标物体类别轮廓,这说明被筛选的特征包含了目标任务的主要信息,特征选择子网络模块有效地提取出了重要的融合特征。

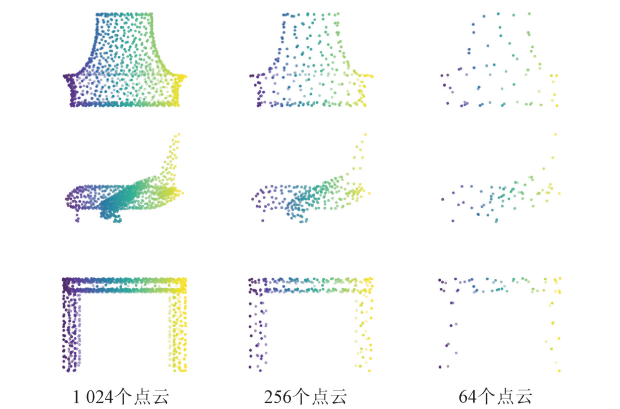


图 6 网络部分阶段点特征图

Fig. 6 Points feature map in the network part phase

3.2 点云部件分割实验

ShapeNet Part 数据集包含 16 个类别,共 16 881 个样本,每个样本中标记了不超过 6 个部位,共有 50 个部位,常用于三维点云部件分割实验。在点云部件分割实验中,与基准网络比较的时候,同样采用了相同的数据增强与损失函数。预处理操作中,在 ShapeNet Part 数据集样本中均匀采样 2 048 个点云数据,再对采样点云进行缩放与平移处理。网络中的 KNN 算法参数与基准网络保持一致 ($K \leq 32$)。网络的评价指标采用三维点云部件分割算法中常见的平均交并比损失函数。

表 3 为本文模型与近几年经典网络在 ShapeNet Part 数据集上的对比实验结果。本文模型的平均交并比达到了 86.4%,与基准网络相比,平均交并比提高了 0.1%;与其他部件分割网络相比,取得了较为先进的结果。本文模型只采用原始点云坐标作为输入特征向量,输入特征向量种类少于 PointNet++ 网络和 SO-Net 网络等以原始点云坐标和法线作为输入特征的部件分割网络。虽然与基准模型相比,本文模型的平均交并比只提高了 0.1%,但是在水杯和飞机这两个类别上实现了最近分割性能,且对其他类别的分割结果较为均衡。这是由于,本文模型在基准网络原曲线特征的基础上构建了基于曲线内部的点距离特征和基于曲线之间的线距离特征,将物体的各实例部件以多种曲线的形式划分,增强了各实例部件的部件特征,保证了模型

对各部件的信息提取程度。此外,特征选择子模型柔性化地结合了局部筛选特征和全局特征,将

部件实例的局部信息和整体信息提供给所属类别,加强了模型对部件实例的整体信息把控能力。

表 3 不同网络在 ShapeNet Part 数据集上的分割效果对比

Table 3 Comparison of the segmentation result of different networks in the ShapeNet Part dataset

网络	交并比/%																
	飞机	包	帽子	汽车	椅子	耳机	吉他	匕首	台灯	笔记本	摩托车	水杯	手枪	火箭	滑板	桌子	平均值
PointNet ^[1]	83.4	78.7	82.5	74.9	89.6	73.0	91.5	85.9	80.8	95.3	65.2	93.0	81.2	57.9	72.8	80.6	83.7
PointNet++ ^[2]	82.4	79.0	87.7	77.3	90.8	71.8	91.0	85.9	83.7	95.3	71.6	94.1	81.3	58.7	76.4	82.6	85.1
PointCNN ^[8]	84.1	86.5	86.0	80.8	90.6	79.7	92.3	88.4	85.3	96.1	77.2	95.2	84.2	64.2	80.0	83.0	86.1
RS-CNN ^[9]	83.5	84.8	88.8	79.6	91.2	81.1	91.6	88.4	86.0	96.0	73.7	94.1	83.4	60.5	77.7	83.6	86.2
DGCNN ^[13]	84.0	83.4	86.7	77.8	90.6	74.7	91.2	87.5	82.8	95.7	66.3	94.9	81.1	63.5	74.5	82.6	85.2
PointMLP ^[17]	83.5	83.4	87.5	80.5	90.3	78.2	92.2	88.1	82.6	96.2	77.5	95.8	85.4	64.6	83.3	84.3	86.1
PRCDNet ^[18]	83.8	84.8	89.1	79.9	90.0	82.4	92.0	87.8	83.0	96.1	77.1	95.2	83.2	64.8	83.5	83.2	86.0
PointASNL ^[23]	84.1	84.7	87.9	79.7	92.2	73.7	91.0	87.2	84.2	95.8	74.4	95.2	81.0	63.0	76.3	83.2	86.1
PCT ^[25]	85.0	82.4	89.0	81.2	91.9	71.5	91.3	88.1	86.3	95.8	64.6	95.8	83.6	62.2	77.6	83.7	86.4
SO-Net ^[26]	82.8	77.8	88.0	77.3	90.6	73.5	90.7	83.9	82.8	94.8	69.1	94.2	80.9	53.1	72.9	83.0	84.9
Point-BERT ^[27]	84.3	84.8	88.0	79.8	91.0	81.7	91.6	87.9	85.2	95.6	75.6	94.7	84.3	63.4	76.3	81.5	85.6
CurveNet ^[19]	85.0	85.0	88.9	80.3	91.7	77.7	91.7	88.6	85.6	96.0	68.7	95.2	82.3	60.5	76.9	83.4	86.3
本文	85.0	80.6	89.7	80.1	91.7	79.7	91.7	88.5	85.6	96.1	70.5	95.0	81.4	62.5	76.4	83.6	86.4

注:加粗数据为交并比最高值。

图 7 为 ShapeNet Part 数据集的分割结果可视化图。图中每个物体的相同颜色代表该色域中点云属于同一部件。用正方形标注同一物体不同部件在基准网络与本文网络中表现不同的地方,并在旁边放大这些局部效果。在局部放大图中,为了

更清晰地看出两个网络的表现,用圆形标注出了具有差异的点云分割结果。从图中标红部分及其相应的放大结果可以观察到,本文网络在相邻部件分界点云处的分割结果更细致,对部件的识别结果也更加准确。

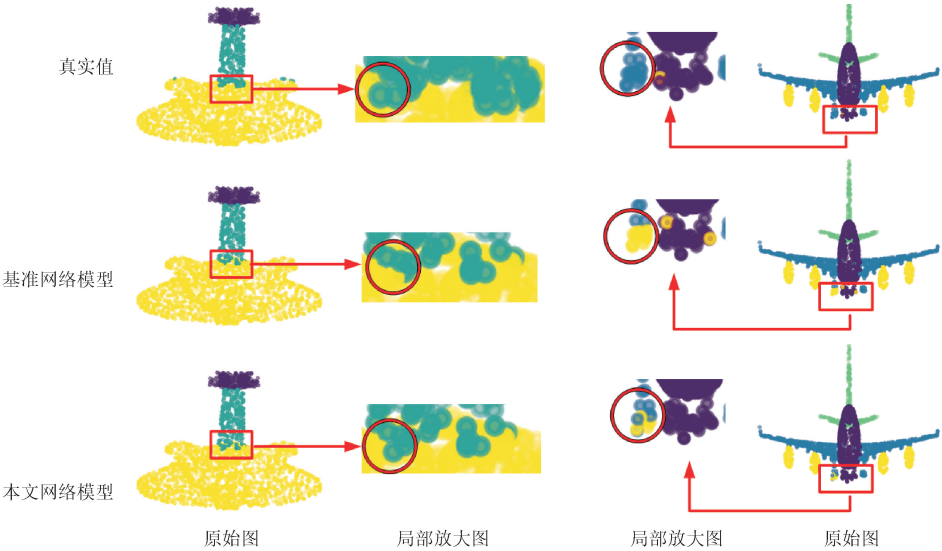


图 7 网络对 ShapenNet Part 数据集分割效果图

Fig. 7 Effect of network segmentation of the ShapenNet Part dataset

3.3 网络参数对比

本文网络与近几年经典的网络进行网络参数量对比,结果如表 4 所示。可以观察到,在分类与部件分割实验中,PointNet 网络、PointNet++ 网络、CurveNet 网络的网络参数少于本文网络,其余网络的参数量远超本文网络。本文网络在分类实验中的识别精度与部件分割实验中的平均交并比均处于先进水平。这说明了本文网络在保证输出结果高准确性的同时兼顾了网络的规模,具有一定的实际运用价值。

表 4 不同网络参数量对比

Table 4 Comparison of parameter sizes of different networks

网络	分类参数量/MB	分割参数量/MB
PointNet ^[1]	3.5	8.3
PointNet++ ^[2]	1.7	1.8
PointConv ^[10]	18.6	
KPConv ^[11]	15.2	
PointMLP ^[17]	12.6	16.8
CurveNet ^[19]	2.1	5.5
GBNet ^[21]	8.4	
本文	4.1	11.3

3.4 消融实验

本文提出了特征选择子网络模块、基于曲线内部的点距离特征与基于曲线之间的线距离特征、快速通道相关性注意力机制这 3 个模块。为了验证这 3 个模块对网络的影响,在 ModelNet40 数据集上进行了消融实验,实验结果如表 5 所示。

表 5 不同模块在 ModelNet40 数据集上的消融实验结果

Table 5 Ablation experiment results of different modules on the ModelNet40 dataset

模块	特征选择子网络	曲线多特征聚合网络			参数量/MB
		两种曲线特征	快速通道相关性注意力机制	准确率/%	
1				93.1	2.1
2	✓			93.6	3.0
3	✓	✓		93.7	3.0
4	✓	✓	✓	93.8	4.1

注:模块 1 为基准网络;✓ 表示在基准网络上加入该模块。

从表 5 中可以看出:原基准网络在 ModelNet40 数据集上的识别准确率为 93.1%;加入特征选择子

网络模块后,网络的识别准确率提升了 0.5%;在此基础上依次在曲线干预卷积模块中加入基于曲线内部的点距离特征与基于曲线之间的线距离特征、快速通道相关性注意力机制后,网络的识别精度也都提升了 0.1%。由此可得,特征选择子网络对原始网络的精度提升效果最好,证明了特征选择子网络可以提升整体网络对有效特征的敏感程度。基于曲线内部的点距离特征与基于曲线之间的线距离特征、快速通道相关性注意力机制也都对网络的识别效果起到了一定的积极影响,在一定程度上丰富了网络的信道特征种类,加深了网络对各种特征内部信息的挖掘程度,证明了丰富曲线特征种类、挖掘曲线特征信息这种思路是具备可行性的。

从参数量的角度观察网络,基准网络参数量为 2.1 MB,加入特征选择子网络模块和快速通道相关性注意力机制都会使网络参数量增加。这是因为,特征选择子网络模块为了保持原始网络优势,在不改变原始网络结构的基础上对融合特征进行筛选;快速通道相关性注意力机制为了充分提取出各曲线特征的内部信息,加深了曲线干预卷积模块的网络层数。

4 结束语

本文提出了一种用于三维点云分类和分割的对融合特征敏感的曲线多特征聚合网络。在构建基于曲线特征、点特征和边特征的融合特征的同时,对融合特征的有效性打分,筛选出描述性强的融合特征。此外,在网络中聚合曲线特征模块内部加入两种曲线特征表示方式,利用三维卷积提取出曲线特征间的上下文联系,利用快速通道相关性注意力机制加深各特征的内部联系。该方法在多点云对象信息感知中取得了较好的效果。

对融合特征敏感的曲线多特征聚合网络虽然对点云分类和分割取得了不错的效果,但是并没有考虑到加入噪声干扰点等干扰因素的影响,仍然需要不断研究网络的抗干扰能力。加强网络的鲁棒性和信息感知能力仍然是未来的重要研究方向。

参考文献:

[1] CHARLES R Q, SU Hao, KAICHUN M, et al. PointNet: deep learning on point sets for 3D classification and segmentation [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2017: 77-85.
[2] QI C R, YI Li, SU Hao, 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. Red Hook, NY, USA: Curran Associates Inc., 2017: 5105-5114.
- [3] JIANG Mingyang, WU Yiran, ZHAO Tianqi, et al. PointSIFT: a SIFT-like network module for 3D point cloud semantic segmentation [EB/OL]. (2018-11-24) [2023-08-10]. <https://arxiv.org/abs/1807.00652>.
- [4] JOSEPH-RIVLIN M, ZVIRIN A, KIMMEL R. Moment(e)t: flavor the moments in learning to classify shapes [C]//2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Piscataway, NJ, USA: IEEE, 2019: 4085-4094.
- [5] ZHAO Hengshuang, JIANG Li, FU C W, et al. PointWeb: enhancing local neighborhood features for point cloud processing [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2019: 5560-5568.
- [6] 王溪波, 曹士彭, 赵怀慈, 等. 双边特征聚合与注意力机制点云语义分割 [J]. 仪器仪表学报, 2021, 42(12): 175-183.
- WANG Xibo, CAO Shipeng, ZHAO Huaici, et al. Semantic segmentation of point cloud via bilateral feature aggregation and attention mechanism [J]. Chinese Journal of Scientific Instrument, 2021, 42(12): 175-183.
- [7] KOMARICHEV A, ZHONG Zichun, HUA Jing. A-CNN: annularly convolutional neural networks on point clouds [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2019: 7413-7422.
- [8] LI Yangyan, BU Rui, SUN Mingchao, et al. PointCNN: convolution on X-transformed points [C]//Advances in Neural Information Processing Systems. San Francisco, CA, USA: Curran Associates Inc., 2018: 820-830.
- [9] LIU Yongcheng, FAN Bin, XIANG Shiming, et al. Relation-shape convolutional neural network for point cloud analysis [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2019: 8887-8896.
- [10] WU Wenxuan, QI Zhongang, FUXIN Li. PointConv: deep convolutional networks on 3D point clouds [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2019: 9613-9622.
- [11] THOMAS H, QI C R, DESCHAUD J E, et al. KP-Conv: flexible and deformable convolution for point clouds [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2019: 6410-6419.
- [12] 彭秀平, 全其胜, 林洪彬, 等. 一种面向散乱点云语义分割的深度残差-特征金字塔网络框架 [J]. 自动化学报, 2021, 47(12): 2831-2840.
- PENG Xiuping, TONG Qisheng, LIN Hongbin, et al. A deep residual-feature pyramid network framework for scattered point cloud semantic segmentation [J]. Acta Automatica Sinica, 2021, 47(12): 2831-2840.
- [13] WANG Yue, SUN Yongbin, LIU Ziwei, et al. Dynamic graph CNN for learning on point clouds [J]. ACM Transactions on Graphics, 2019, 38(5): 146.
- [14] LANDRIEU L, SIMONOVSKY M. Large-scale point cloud semantic segmentation with superpoint graphs [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 4558-4567.
- [15] WANG Lei, HUANG Yuchun, HOU Yaolin, et al. Graph attention convolution for point cloud semantic segmentation [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2019: 10288-10297.
- [16] 李冠彬, 张锐斐, 陈超, 等. 基于旋转不变深度层次聚类网络的点云分析 [J]. 软件学报, 2022, 33(11): 4356-4378.
- LI Guanbin, ZHANG Ruifei, CHEN Chao, et al. Rotation-invariant deep hierarchical cluster network for point cloud analysis [J]. Journal of Software, 2022, 33(11): 4356-4378.
- [17] MA Xu, QIN Can, YOU Haoxuan, et al. Rethinking network design and local geometry in point cloud: a simple residual MLP framework [J/OL]. (2022-11-29) [2023-08-22]. <https://arxiv.org/abs/2202.07123>.
- [18] 杨军, 王连甲. 结合位置关系卷积与深度残差网络的三维点云识别与分割 [J]. 西安交通大学学报, 2023, 57(5): 182-193.
- YANG Jun, WANG Lianjia. Recognition and segmentation of 3D point cloud through positional relation convolution in combination with deep residual network [J]. Journal of Xi'an Jiaotong University, 2023, 57(5): 182-193.
- [19] XIANG Tiange, ZHANG Chaoyi, SONG Yang, et al. Walk in the cloud: learning curves for point clouds shape analysis [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2021: 895-904.

- [20] WU Bingli, MA Jie, CHEN Gaojie, et al. Feature interactive representation for point cloud registration [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2021: 5510-5519.
- [21] QIU Shi, ANWAR S, BARNES N. Geometric back-projection network for point cloud classification [J]. IEEE Transactions on Multimedia, 2022, 24: 1943-1955.
- [22] WAZIR S, FRAZ M M. HistoSeg: quick attention with multi-loss function for multi-structure segmentation in digital histology images [C]//2022 12th International Conference on Pattern Recognition Systems (ICPRS). Piscataway, NJ, USA: IEEE, 2022: 1-7.
- [23] YAN Xu, ZHENG Chaoda, LI Zhen, et al. Point ASNL: robust point clouds processing using nonlocal neural networks with adaptive sampling [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2020: 5588-5597.
- [24] XU Mutian, DING Runyu, ZHAO Hengshuang, et al. PAConv: position adaptive convolution with dynamic kernel assembling on point clouds [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2021: 3172-3181.
- [25] GUO Menghao, CAI Junxiong, LIU Zhengning, et al. PCT: point cloud transformer [J]. Computational Visual Media, 2021, 7(2): 187-199.
- [26] LI Jiaxin, CHEN Benm, LEE G H. SO-Net: self-organizing network for point cloud analysis [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 9397-9406.
- [27] YU Xumin, TANG Lulu, RAO Yongming, et al. Point-BERT: pre-training 3D point cloud transformers with masked point modeling [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2022: 19291-19300.
- [28] ZHANG Renrui, WANG Liuhui, GUO Ziyu, et al. Parameter is not all you need: starting from non-parametric networks for 3D point cloud analysis [EB/OL]. (2021-05-10) [2023-08-22]. <https://arxiv.org/abs/2303.08134>.
- [29] YAN S, YANG Z, LI H, et al. Implicit autoencoder for point-cloud self-supervised representation learning [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. [S. l.]: [s. n.], 2023: 14530-14542.
- [30] ZHANG Renrui, WANG Liuhui, QIAO Yu, et al. Learning 3D representations from 2D pre-trained models via image-to-point masked autoencoders [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2023: 21769-21780.

(编辑 杜秀杰)