

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-18

论文引用格式: ZHAO Fuqun, LI Jingjing. Multi-granularity lightweight point cloud classification integrating spatial domain features and frequency domain decoupling[J/OL]. Journal of Image and Graphics, XXXX:1-18. DOI: 10.11834/jig.260124. (赵夫群, 李晶晶. 融合空间域特征与频域解耦的多粒度轻量级点云分类[J/OL]. 中国图象图形学报, XXXX:1-18. DOI: 10.11834/jig.260124.) [DOI:10.11834/jig.260124]

融合空间域特征与频域解耦的多粒度轻量级点云分类

赵夫群^{1,2}, 李晶晶^{1,2}

1. 西安财经大学 信息学院, 西安 710100; 2. 智财协同可信计算陕西省高等学校重点实验室, 西安 710100

摘要: 目的 针对现有点云分类网络在局部几何结构建模、特征融合及表达能力方面的不足, 提出一种基于空间域几何特征与频域解耦的多粒度点云分类。方法 网络采用多维度多分支特征融合架构: 首先, 在空间域维度, 整合点特征分支与全局特征分支, 利用改进的边缘卷积算子及递增近邻数序列实现从局部到全局的多尺度几何特征聚合, 并引入通道注意力机制实现显著性特征的自适应增强; 在频域维度, 通过局部频谱特征提取器解耦低频、高频及频谱差异信息, 并利用选择性状态空间模型实现局部区域内的长距离特征交互; 最后, 融合正则化策略、双路池化和多重分类器, 实现点云分类。结果 该方法在 ModelNet40 数据集的总体准确率和平均准确率分别达到 93.0% 和 90.7%, 相较于现有方法分别提升约 1.0% 和 2.0%, 在 ScanobjectNN 数据集的总体准确率和平均准确率分别达到 82.4% 和 79.8%, 相较于现有方法分别提升约 3.4% 和 4.6%, 在 ShapeNet Part 数据集的实例平均交并比与类别平均交并比分别达到 85.94% 和 83.32%, 相较于现有方法分别提升约 0.8% 和 0.7%。结论 本方法能有效提升点云局部结构建模能力与尺度适应性, 在保持模型轻量化的同时, 显著提升了分类性能。

关键词: 点云分类; 空间域; 频域; 边缘卷积模型; 选择性状态空间模型

Multi-granularity lightweight point cloud classification integrating spatial domain features and frequency domain decoupling

ZHAO Fuqun^{1,2}, LI Jingjing^{1,2}

1. School of Information, Xi'an University of Finance and Economics, Xi'an, 710100, China; 2. Key Laboratory of Intelligent Finance Collaboration and Trusted Computing, Shaanxi Provincial Institutions of Higher Education, Xi'an, 710100, China

Abstract: **Objective** Point cloud classification has emerged as a fundamental pillar in 3D computer vision, primarily driven by the rapid proliferation of LiDAR sensors and depth-sensing technologies in autonomous systems, industrial inspection, and robotic perception. Unlike structured 2D images, 3D point clouds are characterized by inherent irregularity, sparsity, and a lack of canonical ordering, which poses significant challenges for high-precision recognition. While recent convolution-based and MLP-based networks have achieved considerable progress, they often struggle to capture fine-grained local geometric structures and establish comprehensive global topological dependencies simultaneously. Many existing models suffer from limited feature representation capabilities when dealing with complex object geometries or noisy environments, often failing to maintain scale adaptability and robust discriminative power. This study aims to address these critical deficiencies by proposing a novel multi-granularity classification network that synergizes spatial-domain geometric

收稿日期: 2026-03-08; 修回日期: 2026-05-28

基金项目: 国家自然科学基金(62271393); 陕西省教育厅科学研究计划项目(25JS049)

Supported by: National Natural Science Foundation of China (62271393); Scientific Research Program Funded by Education Department of Shaanxi Provincial Government (25JS049)

features with frequency-domain decoupled information. The primary objective is to enhance the model's ability to model local structures and refine feature fusion mechanisms, thereby achieving superior classification performance and structural robustness in 3D shape recognition tasks. **Method** In this study, a sophisticated multi-branch fusion architecture is developed to perform multi-granularity feature extraction and integration across different domains. The technical scheme is organized into several complementary modules designed to perceive point clouds from both spatial and spectral perspectives. The spatial domain processing is bifurcated into a point feature branch and a global geometric branch. The point feature branch utilizes Multi-Layer Perceptrons to extract robust, fine-grained geometric representations directly from the raw point coordinates, ensuring the preservation of fundamental shape information. Simultaneously, the global geometric branch employs an improved edge convolution algorithm. Unlike traditional static graph convolutions, this branch utilizes an incremental neighbor sequence, specifically a k -sequence, to realize multi-scale aggregation of the global topology. This mechanism allows the network to perceive structural evolution ranging from micro-level local neighborhoods to macro-level global contours, significantly improving the receptive field of the model. To further emphasize critical geometric characteristics, a channel attention mechanism is incorporated to adaptively enhance significant geometric responses while suppressing redundant noise. In parallel, a local spectral feature extractor is designed for frequency-domain analysis. It decouples the features into low-frequency, high-frequency, and spectral difference information, effectively capturing various levels of structural variations and geometric skeletons that are often ignored by spatial-only operations. To facilitate long-range feature interaction within local regions without excessive computational overhead, a selective state space model, specifically the Mamba architecture, is integrated into the spectral branch. Finally, the network integrates a spatial transformation regularization strategy via T-Net to ensure rotation invariance, a dual-path pooling mechanism combining maximum and average pooling to retain representative features, and multiple classifiers to achieve the final robust classification. **Result** The performance of the proposed multi-granularity network is rigorously evaluated on two mainstream benchmark datasets, namely ModelNet40 and ScanObjectNN. ModelNet40 consists of 12,311 synthetic CAD models across 40 categories, providing an ideal environment for testing geometric learning. In contrast, ScanObjectNN provides a more challenging testbed with 15,000 real-world scanned objects containing significant background noise and occlusions. The quantitative evaluation focuses on two key metrics which are Overall Accuracy and mean Accuracy. Experimental results show that the proposed method achieves an Overall Accuracy of 93.0% and a mean Accuracy of 90.7% on the ModelNet40 dataset. On the more challenging ScanObjectNN dataset, the model yields an Overall Accuracy of 82.4% and a mean Accuracy of 79.8%. On the ShapeNet Part dataset, the average instance intersection-over-union (IoU) and class intersection-over-union (IoU) ratios reached 85.94% and 83.32%. Comparative experiments demonstrate the clear superiority of this approach over several state-of-the-art models. Compared with existing mainstream methods, the classification accuracy is improved by an average of approximately 1% on ModelNet40 and 3% on ScanObjectNN. On the ShapeNet Part dataset, our approach outperforms state-of-the-art methods by about 0.8% in instance mIoU and 0.7% in class mIoU. Specifically, the integration of spectral decoupling and the selective state space model significantly enhances the model's discriminative power and scale adaptability. The experimental data confirms that the synergy of spatial-domain multi-scale features and frequency-domain decoupled information provides superior robustness against geometric deformations and real-world sensor noise, effectively reducing the performance gap between synthetic and real-world data. **Conclusion** The experimental results demonstrate that the model consistently outperforms several state-of-the-art approaches, and the proposed multi-branch fusion algorithm significantly improves the overall feature representation capability. This research offers a 3D object recognition, demonstrating that multi-granularity analysis can effectively compensate for the information loss inherent in single-domain operations. The application of this research is highly significant for fields requiring high-precision 3D perception. Future work will focus on further optimizing the model structure to maintain its lightweight characteristics while exploring the potential of this multi-granularity strategy in large-scale scene segmentation and other complex 3D vision tasks. By continuously refining the selective state space mechanism and spectral decoupling filters, we aim to develop a more generalized and efficient framework for 3D understanding that can adapt to increasingly complex sensing environments.

Key words: Point cloud classification; spatial domain; frequency domain; edge convolution model; selective state space model

0 引言

随着三维传感技术的快速发展,三维点云数据处理在自动驾驶(王正宝等,2025)、三维目标重建(Daif和Marzouk,2025)及虚拟现实(Pan等,2025)等前沿领域得到广泛应用,这对点云分类模型的判别精度与鲁棒性提出了更高要求。然而,现有点云分类网络仍面临点云的几何变换不变性处理不足,局部邻域特征的高阶表达能力有限,以及全局结构建模的计算复杂度较高的问题。因此,构建一种能够融合几何仿射变换建模、局部图卷积网络与序列化特征学习的点云分类框架,对于提升分类的判别能力与鲁棒性具有重要理论与应用价值。

早期的点云分类研究高度依赖人工提取特征与机器学习算子的结合。研究人员需预先定义点云的几何属性,并将其同既定的语义标签进行关联,进而驱动模型学习,其中,随机森林(陶帅兵等,2021)与支持向量机(释小松等,2019)等算法的应用最为广泛。近年来,深度学习在点云分类研究中展现出显著优势。相比于依赖大量先验知识的传统路径,深度学习技术能从原始海量信息中自主归纳特征分类,这种低人工干预的高效学习机制,使其成为当前主流的研究手段。其中Qi等人(2017a)提出的PointNet,为直接处理原始点数据奠定了基础,通过引入仿射变换网络来解决点云的置换不变性问题;Qi等人(2017b)提出的PointNet++,通过分层采样与局部特征汇聚的方式,更有效地建模点云中的局部结构与上下文关系。随后,点云分类研究逐步转向构建更加细致、表达能力更强的局部特征聚合算子。Qian等人(2022)提出PointNeXt方法,在PointNet++框架基础上对采样流程、邻域划分方式以及多尺度卷积结构进行优化,但在点云分布稀疏的情况下,对局部细节的刻画能力仍然受限。Wang等人(2019)提出的DGCNN(dynamic graph convolutional neural network)通过构建动态图卷积,有效地捕捉了局部邻域的几何关系,显著提升了分类性能。虽然局部特征聚合与多尺度建模显著增强了分类性能,但在面对稀疏点云或复杂几何形态时,细节表征能力仍待提升。

注意力机制通过对关键信息的自适应筛选,有效强化了特征的辨识度,已成为优化模型判别能力

的重要策略。Guo等人(2021)通过引入隐式拉普拉斯算子并结合归一化函数筛选关键特征信息,与直接使用自注意力机制的Transformer(Zhao等,2021)方法相比,该方法天然满足置换不变性要求,更契合点云学习任务。Wu等人(2024)提出的Point Transformer V3通过简化注意力模块、采用序列化邻域映射替代精确KNN搜索,显著提升点云处理的效率与可扩展性。此外,Zhang等人(2025)提出的点云Mamba(Point Cloud Mamba)虽利用最远点采样与K近邻(K-Nearest Neighbors, KNN)算法实现点云分组序列化,但忽略了局部几何特征的异质性,难以充分捕捉复杂的三维空间关联。Wang等人(2018)通过融合谱特征学习与聚类机制,利用频域将空间域的复杂关联映射为差异化频率成分,实现特征属性的精准辨识,但具有较高的计算复杂度。针对3D网格数据,Li等人(2022)提出的Laplacian Mesh Transformer利用拉普拉斯谱分解描述拓扑结构,并结合双重注意力机制捕获细粒度几何特征。

目前的深度学习分类体系以基于体素、多视图、原始点云的分类网络为主,其中基于多视图(Wang等,2019)与体素(Klokov和Lempitsky,2017)的分类通常涉及空间重采样或投影变换,易造成点云的特征信息缺失。因此,本研究聚焦于直接从原始点云中提取特征信息的技术方案,确保特征信息的完整性。此外,针对空间细节捕获不全、局部异质性建模缺失及计算复杂度较高的问题,本文提出了一种基于局部频谱特征聚合和选择性状态空间模型融合的点云分类网络。该网络包含三个核心分支:点特征分支采用简洁高效的结构提取点级细粒度几何特征;全局几何分支基于改进边缘卷积(edge convolution, EdgeConv)构建动态KNN,实现点云全局拓扑关系与空间结构的层次化建模;局部谱分支通过低频通道捕获平滑趋势,高频通道聚焦几何细节变化,谱差异通道编码基于距离的谱域差分信息,实现局部几何特征的多尺度解耦。三个分支分别从点级局部几何、全局拓扑结构和谱域特征三个维度互补建模几何表征,形成多粒度几何描述体系。

1 模型与方法

1.1 整体架构设计

本文提出一种基于空间域几何特征与频域解耦
©中国图象图形学报版权所有

的多粒度点云分类网络,其总体架构如图1所示。首先,模型输入原始三维点云数据,通过几何仿射变换网络对原始空间坐标进行对齐,增强模型对旋转、平移等空间形变的几何不变性。再通过点特征分支提取基础几何特征,为后续高阶表示学习提供初始表征。全局几何分支基于改进的边缘卷积模块构建动态KNN图,并在多层堆叠中逐步捕获点云的全局拓扑关系与空间结构,同时引入通道注意力机制,增强关键几何特征的响应。局部谱分支由低频、高频和谱差异三个子分支组成,分别建模局部平滑趋势、几何细节变化以及基于距离权重的谱域差分信息,并借助多层状态空间模块进行序列建模,从而实现局部结构的联合增强表征。三个分支输出的多尺度特征经最大池化与平均池化的双路聚合策略进行压缩,有效保留全局显著性与统计分布信息。最终,将所有聚合特征在通道维度拼接后输入全连接分类头完成预测。该架构在结构化几何建模、局部谱分析与多尺度融合之间形成互补,使模型在保持较高计算效率的同时具备更强的特征表达能力与优秀的分类性能。

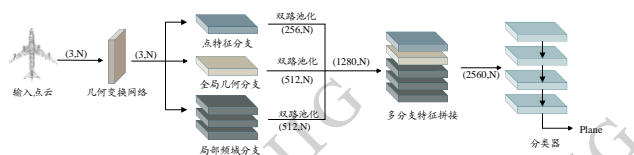


图1 网络结构

Fig. 1 Network Structure

1.2 基于多层感知机的点特征分支

点特征分支作为模型的基础特征提取模块,主要用于从原始三维点云中学习低级几何语义特征,为全局几何分支和局部谱分支提供稳定的底层表征。由于点云数据具有天然无序性,逐点处理的多层感知机(Multi-Layer Perceptron, MLP)能够有效保证特征对点的排列不敏感,同时通过非线性映射增强表达能力。

设输入点云为矩阵 \mathbf{X} , 表达式为:

$$\mathbf{X} = \{ \mathbf{x}_i \in \mathbf{R}^3 | i = 1, 2, \dots, N \} \quad (1)$$

式中, \mathbf{X} 为输入点云矩阵, \mathbf{x}_i 为第 i 个点的三维坐标向量, N 为点的数量。

首先,通过一层逐点线性映射与非线性激活函数对每个点进行编码,得到中间特征矩阵 \mathbf{U} , 计算

式为:

$$\mathbf{U} = \sigma(\mathbf{W}^{(1)}\mathbf{X} + \mathbf{b}^{(1)}) \quad (2)$$

式中, \mathbf{U} 为中间特征矩阵, $\mathbf{W}^{(1)}$ 和 $\mathbf{b}^{(1)}$ 分别为第一层的权重矩阵和偏置向量, $\sigma(\cdot)$ 为逐元素激活函数, 该层通过线性变换与非线性激活提取低级几何特征。

中间特征 \mathbf{U} 通过第二层逐点映射生成更高维的点特征 \mathbf{P} , 计算式为:

$$\mathbf{P} = \sigma(\mathbf{W}^{(2)}\mathbf{U} + \mathbf{b}^{(2)}) \quad (3)$$

式中, \mathbf{P} 为输出的点特征矩阵, $\mathbf{W}^{(2)}$ 和 $\mathbf{b}^{(2)}$ 分别为第二层的权重矩阵与偏置向量。该层通过非线性变换提取局部几何特征, 包括曲率变化、密度分布及局部结构差异等, 为多分支网络提供更精细的底层语义表示。

为将点特征 \mathbf{P} 转化为用于下游分类任务的全局特征向量, 本文采用双路聚合策略, 即全局最大池化与全局平均池化, 捕捉判别性特征, 增强特征的表达能力。先对局部特征嵌入矩阵 \mathbf{P} 进行全局最大池化, 以提取点云中最具判别性的局部特征, 强化关键通道的响应速度, 最大池化计算式为:

$$\mathbf{p}_1 = \max_{i=1}^N \{ \mathbf{p}_i \} \quad (4)$$

式中, \mathbf{p}_1 为全局最大池化后的特征向量, \mathbf{p}_i 为矩阵 \mathbf{P} 中第 i 个点的特征向量。

然后, 执行全局平均池化操作, 以反映点云特征的整体分布趋势, 提供全面且稳定的上下文特征信息, 计算式为:

$$\mathbf{p}_2 = \frac{1}{N} \sum_{i=1}^N \mathbf{p}_i \quad (5)$$

式中, \mathbf{p}_2 为全局平均池化后的特征向量。

最终, 将 \mathbf{p}_1 和 \mathbf{p}_2 在特征维度上进行拼接, 构成全局坐标特征向量 \mathbf{v}_1 , 计算式为:

$$\mathbf{v}_1 = [\mathbf{p}_1 \| \mathbf{p}_2] \quad (6)$$

式中, \mathbf{v}_1 为拼接后的全局坐标特征向量。

点特征分支具有结构简洁、计算效率高及稳定性强的显著优势。相较于全局特征提取分支, 该分支提供点级细粒度几何特征表示; 与局部谱分析分支相比, 其不依赖局部结构的拓扑关系, 可提供稳健且轻量的基础几何表征。在多分支融合架构中, 点特征分支为全局与局部特征构建统一的底层特征表示, 显著提升网络对多尺度几何信息的综合建模能力。

1.3 基于边缘卷积的全局几何特征分支

全局几何特征分支基于改进的边缘卷积模块构建动态图,通过层次化几何特征学习,捕获点云的全局拓扑关系与空间结构建模。该分支以显式编码局部邻域间的相对几何差异,为模型提供稳健的结构化表征能力。

对于特征空间中的任意点 $f_i \in \mathbf{R}^{c_n}$,先依据欧氏距离度量构建 k 个最近邻点集,再针对中心点 f_i 及其邻域点 f_j ,计算边缘特征 e_{ij} ,计算式为:

$$e_{ij} = [f_i \| (f_i - f_j)] \quad (7)$$

式中, e_{ij} 为边缘特征向量, f_i 和 f_j 分别为中心点和邻域点的特征向量。边缘特征 e_{ij} 通过融合中心点绝对特征与邻域相对位移向量,同步编码全局位置信息与局部几何结构差异。

将边缘特征 e_{ij} 通过共享参数的逐点 MLP 进行非线性映射,获得变换后的边缘特征 h_{ij} ,通过最大池化聚合关键结构特征,以提取邻域中最显著的结构信息,得到点特征 g_i ,计算式为:

$$g_i = \max_{j \in S_i} \{h_{ij}\} \quad (8)$$

式中, g_i 是边缘卷积输出特征图 G 中第 i 个点的特征向量, h_{ij} 为变换后的边缘特征向量, S_i 为中心点的邻域点索引集合。

针对固定近邻尺度难以兼顾局部细节与全局语义的问题,本文在全局几何特征分支中采用递增的近邻数序列,构建层级化的多粒度感知机制。浅层网络以较小粒度聚合细微拓扑,深层网络以较大粒度捕获全局语义。这种空间域的多粒度设计与频域的解耦特征互为补充,共同提升点云特征的判别性。在边缘卷积序列的最终输出 G 上,为多粒度特征中隐含的通道相关性,引入通道注意力机制,对不同特征通道进行自适应加权,增强高关键几何通道的响应,输出加权特征 G_1 ,计算式为:

$$G_1 = G \odot \alpha \quad (9)$$

式中, G_1 为加权后的特征矩阵, G 为输入特征矩阵, α 为通道注意力权重向量, \odot 表示通道维度的逐元素乘法。

最后,经通道注意力机制增强的特征矩阵 G_1 采用与 2.2 节相同的全局双路聚合策略,生成全局几何特征向量 v_2 。全局几何特征提取分支通过层次化动态 KNN 图构建,实现点云全局拓扑结构与空间分

布的精准建模。相较于点特征分支的局部表征,该分支专注于全局几何关系的尺度一致性描述;与局部谱分析分支的频域特征提取相比,其更强调点云整体空间结构的统一表达。在多分支融合框架中,全局几何特征提取分支为局部谱域特征提供结构化上下文支撑,显著提升模型对复杂几何语义的全局感知能力。

1.4 基于选择性状态空间模型的局部谱分支

局部谱分支通过低频、高频与谱差异三个并行子通道实现局部几何特征的多尺度解耦:低频通道捕获平滑几何趋势,高频通道聚焦噪声敏感的细节变化,谱差异通道基于距离权重编码谱域差分信息。结合 Mamba 架构实现的序列处理机制,该分支高效整合长距离依赖关系,显著提升局部结构的上下文感知能力与鲁棒性。

设输入点云数据为 X ,其中 N 为点的数量。首先,对点云数据 X 应用频率调制机制,通过学习到的频率权重参数对点云特征进行自适应缩放,生成调制后的特征表示 X_1 。这一调制步骤旨在增强点云在局部高频分量上的表示能力,计算式为:

$$X_1 = X \odot \sigma(W_1 X + B_1) \quad (10)$$

式中, X_1 为调制后的特征矩阵,为输入点云矩阵, W_1 和 B_1 为可学习的权重和偏置矩阵。

对调制后的点,在特征空间中找到其 k 个近邻点集。基于该邻域,我们并行提取三路局部特征。第一路为低频特征 l_i ,通过求邻域点特征的平均值并经过 MLP 映射,模拟局部基频信息,建模局部平滑趋势,计算式为:

$$l_i = \varphi_1 \left(\frac{1}{k} \sum_{j \in S_i} x_{1,j} \right) \quad (11)$$

式中, l_i 为低频特征向量, φ_1 为低频分支的多层感知机函数, k 为邻域点数, $x_{1,j}$ 为调制后点云 X_1 中第 j 个点的特征向量。

第二路是高频特征 q_i ,通过计算中心点特征与邻域平均值的残差,并以最大池化聚合特征,突出局部尖锐的细节信息,捕捉几何细节变化。

$$q_i = \varphi_2 \left(\max_{j \in S_i} \{x_{1,i} - x_{1,j}\} \right) \quad (12)$$

式中, q_i 为高频特征向量, φ_2 为高频分支的多层感知机函数, $x_{1,i}$ 为调制后点云中第 i 个点的特征向量。

第三路为谱差异特征 r_i ,通过可学习的距离加权函数 w_{ij} 对邻域残差进行加权求和,以模拟精细的

局部谱汇聚过程,捕获基于距离权重的谱域差分信息,计算式为:

$$\mathbf{r}_i = \varphi_3 \left(\sum_{j \in \mathcal{S}_i} \mathbf{w}_{ij} (\mathbf{x}_{1,i} - \mathbf{x}_{1,j}) \right) \quad (13)$$

式中, \mathbf{r}_i 为谱差异特征向量, φ_3 为谱差异分支的多层感知机函数, \mathbf{w}_{ij} 为基于点对距离学习的权重标量。

三路特征 \mathbf{l}_i 、 \mathbf{q}_i 和 \mathbf{r}_i 在通道维度上拼接。为利用 Mamba 实现高效序列建模,本文基于几何距离对拼接后的特征进行排序。对每个中心点的近邻点,依据其与中心点的欧氏距离升序排列。该序列化方式有效保留了局部邻域的空间连续性,为 Mamba 捕获细粒度几何上下文提供了有序的输入。考虑到点云数据固有的无序性,本文在 Mamba 序列建模层后采用全局最大池化与全局平均池化进行特征聚合。根据点云处理的对称函数理论,利用池化算子的对称函数特性,模型能够从 Mamba 提取的有序特征图中解耦出顺序无关的全局表征。经通道注意力增强后,得到最终的局部谱特征矩阵。该特征矩阵最终采用全局双路池化聚合特征策略,生成局部谱特征向量 \mathbf{v}_3 。

局部谱分支通过三支并处理,区分平滑趋势与几何细节,同时通过谱域差分机制捕获基于距离的结构差异信息。谱域分析能将不同频率的几何特征自然分离,避免空间域中特征混杂,使模型更精准地建模局部结构的频域特性。

1.5 正则化与分类模块

在三个分支完成特征提取后,本部分通过特征融合策略整合点特征、全局几何与局部谱分支的表征,生成类别预测结果。同时,协同应用空间变换网络正则化与动态随机失活(Dropout)机制,有效提升模型对点云噪声的鲁棒性与泛化能力。

首先,本模型将坐标特征向量 \mathbf{v}_1 、全局几何特征 \mathbf{v}_2 、局部谱特征向量 \mathbf{v}_3 进行拼接,构建高维融合特征 \mathbf{v} ,计算式为:

$$\mathbf{v} = [\mathbf{v}_1 \parallel \mathbf{v}_2 \parallel \mathbf{v}_3] \quad (14)$$

式中, \mathbf{v} 为融合特征向量。

再将融合特征 \mathbf{v} 送入由三层全连接层构成的分类头进行类别预测。在训练过程中,我们采用带标签平滑的交叉熵损失作为分类目标函数,以增强模型对噪声标签的鲁棒性。

为确保模型对输入点云空间变换的不变性与稳

定性,本模型采用 T-Net 空间变换网络与动态 Dropout 机制的双正则化策略。T-Net 通过预测 3×3 仿射变换矩阵 \mathbf{T} 实现点云空间对齐,生成对齐点云 $\mathbf{X}' = \mathbf{TX}$ 。为保障几何结构的度量属性,引入正交约束正则化项,计算式为:

$$L_2 = \|\mathbf{TT}^T - \mathbf{I}\|_F^2 \quad (15)$$

式中, L_2 为正交约束正则化损失, \mathbf{T} 为空间变换网络预测的 3×3 仿射变换矩阵, \mathbf{T}^T 为 \mathbf{T} 的转置, \mathbf{I} 为 3×3 单位矩阵, $\|\cdot\|_F$ 表示 Frobenius 范数。该约束确保 \mathbf{T} 近似正交,在空间对齐过程中保持几何特性。

总训练损失定义为分类损失与正则化损失的加权和,计算式为:

$$L = L_1 + \alpha L_2 \quad (16)$$

式中, L 为总训练损失, L_1 为分类损失, α 是用于平衡分类损失和正则化损失的超参数, L_2 为正则化损失。

第二种正则化策略是动态 Dropout 机制,该机制应用于分类头的全连接层,实施分阶段 Dropout 率调整:训练初期采用 0.3 以抑制过拟合,模型进入收敛阶段后降低至 0.1,在保持鲁棒性的同时提升收敛稳定性与分类精度。

本部分通过 T-Net 的正交约束保障几何不变性,结合动态 Dropout 的分阶段策略优化模型泛化能力。T-Net 的约束机制维持几何特征度量属性,动态 Dropout 在训练初期抑制过拟合、后期优化收敛,增强对噪声点云的鲁棒性。多分支特征融合的分类头有效提升点云分类的精度与稳定性。

2 实验结果与分析

2.1 数据集与评价指标

实验采用 ModelNet40、ScanObjectNN 和 ShapeNetPart 三个数据集分别验证本文算法,部分 ModelNet40、ScanObjectNN 和 ShapeNetPart 点云数据如图 2、图 3 和图 4 所示。ModelNet40 数据集包含 40 种物体,共 12311 个计算机辅助设计(computer-aided design, CAD)模型,其中 9843 个作为训练集,2468 个作为测试集。ScanObjectNN 是个在现实场景中采集的大型数据集,包含约 15000 个 CAD 模型,涵盖 15 种物体,虽然类别数量少于 ModelNet40,但 ScanObjectNN 数据集具有背景复杂、部分点缺失及变形等特点,因此更具挑战性。ShapeNetPart 数据集是点云

部件分类任务中最常用的基准之一,该数据集包含 16881 个三维模型,覆盖 16 个物体类别,每个模型被手工标注为 2 至 6 个部件,共计 50 个部件标签,其中训练集包含 12137 个模型,测试集包含 2874 个模型。模型使用 Python 和 PyTorch 实现,并在 CUDA12.8 上加速训练。训练参数设置为批次大小 16,学习率 0.0005,训练 300 轮。优化器采用 Adam,该设置能够在保证稳定收敛的同时获得较高的分类精度。

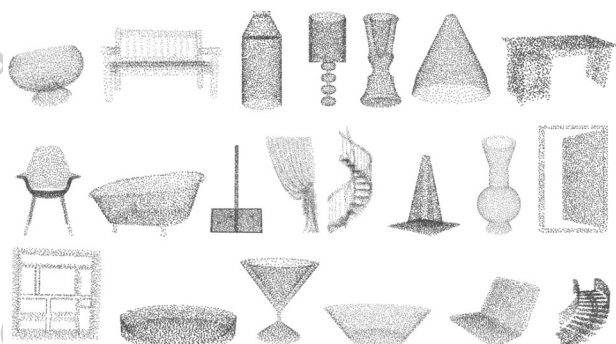


图2 部分 ModelNet40 点云数据

Fig. 2 Partial ModelNet40 point cloud data

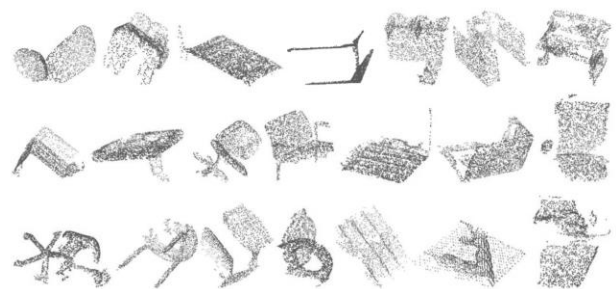


图3 部分 ScanObjectNN 点云数据

Fig. 3 Partial ScanObjectNN point cloud data

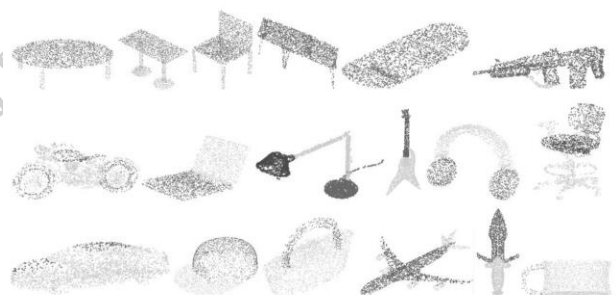


图4 部分 ShapeNetPart 点云数据

Fig. 4 Partial ShapeNetPart point cloud data

本文在 ModelNet40 和 ScanObjectNN 数据集上采用平均准确率(mean accuracy, mAcc)和总体准确率(overall accuracy, OA)作为分类评价指标。总体

准确率衡量模型对所有测试样本的整体分类正确率,平均准确率衡量模型在各个类别上的平均分类表现。在 ShapeNet Part 部件分类任务上,采用平均交并比(mean intersection over union, mIoU)作为评价指标,具体包括实例平均交并比(instance mIoU, I-mIoU)与类别平均交并比(class mIoU, C-mIoU)。I-mIoU 衡量模型在单个物体上的平均部件分类准确率,C-mIoU 衡量模型对不同部件类型的整体分类能力。

2.2 分类结果

2.2.1 本文方法与主流方法分类结果对比

本文提出的分类网络与 13 种对比模型在 ModelNet40、ScanObjectNN 和 ShapeNet Part 数据集上的实验结果如表 1、表 2 和表 3 所示。

表 1 ModelNet40 数据集分类结果对比

Table 1 Comparison of classification results on the ModelNet40 dataset

模型	mAcc/%	OA/%
PointNet(Qi 等, 2017a)	86.07	89.13
PointNet++(Qi 等, 2017b)	87.65	90.54
GVCNN(Feng 等, 2018)	80.23	91.16
PointCNN(Li 等, 2018)	88.31	91.53
KPCConv(Thomas 等, 2019)	89.62	91.37
DGCNN(Wang 等, 2019)	89.63	92.07
LDGCNN(Zhang 等, 2019)	90.25	92.34
PointASNL(Yan 等, 2020)	89.92	92.49
Point Transformer(Engel 等, 2021)	90.52	92.71
GAPointNet(Chen 等, 2021)	90.27	92.93
MLMST(Zhong 和 Han, 2022)	89.53	92.67
3DTCN(Single-scale)(Lu 等, 2022)	89.72	92.81
PointNeXt-S(Qian 等, 2022)	90.76	92.91
Point-GBLS(Zhang 等, 2025)	90.03	92.63
本文模型	90.70	92.99

注:加粗字体为每列最优结果。

在 ModelNet40 数据集上,本模型取得了 92.99% 的 OA,显著优于所对比的 13 种基线模型,包括 MLMST (multi-level multi-scale transformer)、Point Transformer、PointNeXt-S 等先进架构。在 ScanObjectNN 真实场景数据集上,本模型的 OA 和 mAcc 分

表2 ScanObjectNN数据集分类结果对比

Table 2 Comparison of classification results for the ScanObjectNN dataset

模型	mAcc/%	OA/%
PointNet(Qi等, 2017a)	63.41	68.25
PointNet++(Qi等, 2017b)	76.04	78.42
GVCNN(Feng等, 2018)	74.79	78.32
PointCNN(Li等, 2018)	75.36	78.92
KPConv(Thomas等, 2019)	77.32	78.26
DGCNN(Wang等, 2019)	73.27	79.63
LDGCNN(Zhang等, 2019)	74.54	79.34
PointASNL(Yan等, 2020)	76.27	79.92
Point Transformer(Engel等, 2021)	77.71	81.63
GAPointNet(Chen等, 2021)	76.28	80.82
MLMST(Zhong和Han, 2022)	77.12	80.29
3DTCN(Single-scale)(Lu等, 2022)	76.27	80.21
PointNeXt-S(Qian等, 2022)	79.23	81.02
Point-GBLS(Zhang等, 2025)	76.23	80.19
本文模型	79.84	82.35

注:加粗字体为每列最优结果。

别达到82.35%和79.84%,同样优于对比模型。在ShapeNet Part数据集上,本文模型的I-mIoU达到85.94%,C-mIoU为83.32%,亦优于PointNet、PointNet++等早期方法,并与主流模型持平。

与PointNet、PointNet++等早期MLP架构,及GVCNN(group-view convolutional neural network)等传统多视角流派相比,本网络通过引入自适应局部几何特征提取模块与动态图注意力机制,不仅在标准数据集ModelNet40上实现了更高精度,在复杂的真实场景ScanObjectNN上更展现了强大的泛化能力,有效弥补了传统MLP模型在背景干扰、遮挡及点云密度不均等问题下的性能短板。相较于PointCNN(point convolutional neural network)、KPConv(kernel point convolution)、PointNeXt-S及3DTCN(3d temporal convolutional network)等基于卷积操作的模型,本文在保持高效计算的同时,实现了对局部结构细节与全局语义信息的协同建模,克服了固定感受野导致的表达瓶颈,在两个数据集上均实现了稳定且显著的性能提升。相比于DGCNN、LDGCNN(linked dynamic graph convolutional neural

表3 ShapeNet Part数据集分类结果对比

Table 3 Comparison of classification results on the ShapeNet Part dataset

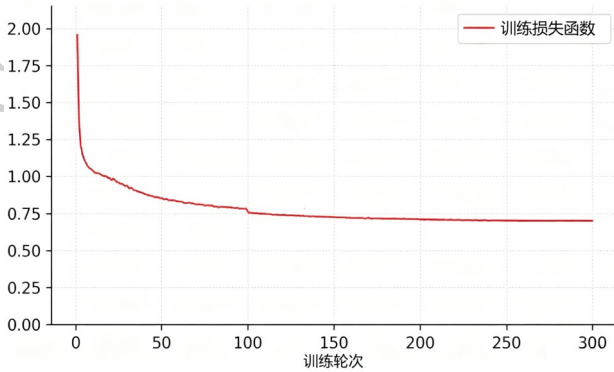
模型	I-mIoU/%	C-mIoU/%
PointNet(Qi等, 2017a)	83.71	80.44
PointNet++(Qi等, 2017b)	85.19	81.96
GVCNN(Feng等, 2018)	85.25	80.36
PointCNN(Li等, 2018)	85.16	83.27
KPConv(Thomas等, 2019)	85.42	83.73
DGCNN(Wang等, 2019)	85.67	83.92
LDGCNN(Zhang等, 2019)	85.18	83.37
PointASNL(Yan等, 2020)	85.44	83.56
Point Transformer(Engel等, 2021)	86.22	83.78
GAPointNet(Chen等, 2021)	83.79	80.43
MLMST(Zhong和Han, 2022)	85.10	82.99
3DTCN(Single-scale)(Lu等, 2022)	85.29	82.38
PointNeXt-S(Qian等, 2022)	85.70	83.23
Point-GBLS(Zhang等, 2025)	85.41	83.73
本文模型	85.94	83.32

注:加粗字体为每列最优结果。

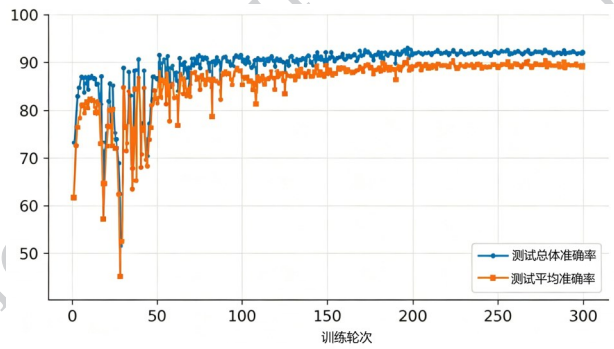
network)等图神经网络模型,本文在保留其动态拓扑感知优势的基础上,通过引入几何一致性约束与分层特征蒸馏策略,有效抑制了噪声点对特征的影响,尤其在ScanObjectNN上表现突出。从综合性能来看,在ModelNet40数据集上,本文模型的OA和mAcc相比所有对比模型的平均值分别提升了约1.05%和1.95%;在ScanObjectNN数据集上,分别提升了约3.41%和4.56%;在ShapeNet Part数据集上,本文模型的I-mIoU和C-mIoU相对对比模型平均值分别提升了0.76%和0.67%。本文模型在绝大多数指标上均优于对比模型,充分体现了该方法在点云分类与部件分类任务上的通用性。

在ModelNet40、ScanObjectNN和ShapeNetPart三个数据集验证本文算法训练过程,如图5、图6与图7所示。

两组实验的训练损失如图5(a)、图6(a)和图7(a)所示,实验结果显示,损失函数均在早期迅速下降,随后趋于稳定,表明模型能够高效收敛。训练集、Modelnet40、ScanObjectNN测试集OA和mAcc的



(a) 损失函数变化



(b) 准确率变化

((a) changes in loss function (b) changes in accuracy)

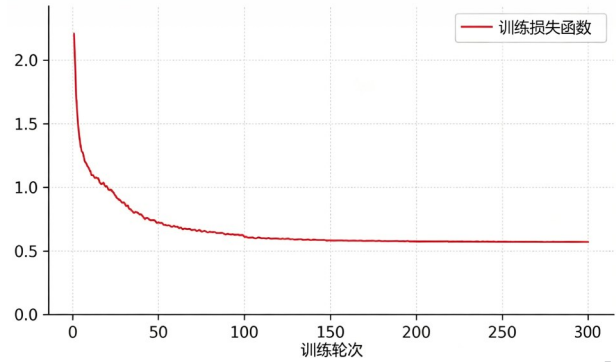
图5 ModelNet40数据集损失函数和准确率变化

Fig. 5 Loss function and accuracy variations on the ModelNet40 dataset

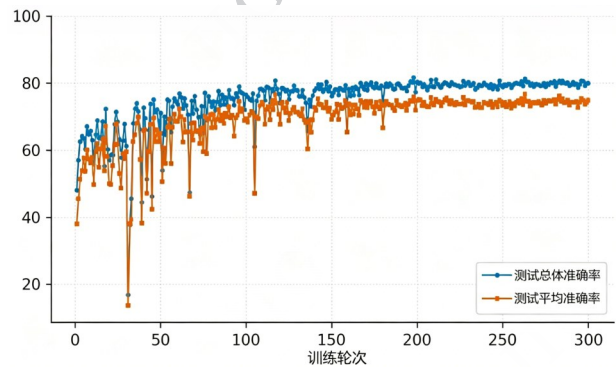
变化趋势如图5(b)、图6(b)所示,ShapeNet Part测试集 I-mIoU 和 C-mIoU 如图7(b)所示。ModelNet40实验取得最高 OA 为 92.99%, mAcc 为 90.70%, ScanObjectNN 实验取得最高 OA 为 82.35%, mAcc 为 79.84%。从两组准确率曲线可见,OA 始终略高于 mAcc,但差距较小,OA 与 mAcc 在收敛后波动幅度有限,显示模型具有良好的泛化性与稳定性。ScanObjectNN 数据集准确率波动较大,可能由于其数据集的复杂性和噪声干扰,导致模型在不同类别上表现更不稳定。ShapeNet Part 实验随着训练轮次增加,I-mIoU 与 C-mIoU 均稳步上升。经过约150轮训练后,I-mIoU 达到 84.5%,C-mIoU 达到 80.5%;最终在300轮时趋于收敛,I-mIoU 达到 85.94%左右,C-mIoU 达到 83.32%。整个训练过程中未出现明显过拟合现象,表明本文模型在部件分类任务上具有良好的收敛性与泛化能力。

2.2.2 各类别分类结果分析

为深入探究本文模型在不同数据集上的细粒度



(a) 损失函数变化



(b) 准确率变化

((a) changes in loss function (b) changes in accuracy)

图6 ScanObjectNN数据集损失函数和准确率变化

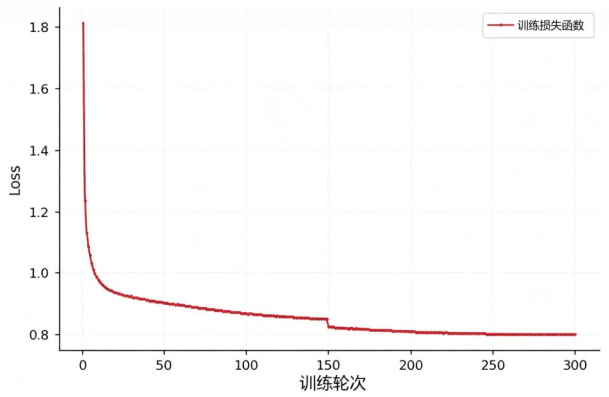
Fig. 6 Changes in loss function and accuracy for the ScanObjectNN dataset

性能差异,分别对 ModelNet40、ScanObjectNN 和 ShapeNet Part 三个数据集进行分类结果统计与可视化展示。

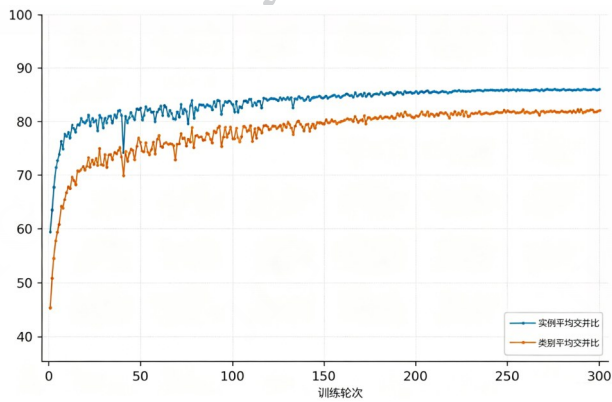
Modelnet40数据集上部分类别与分类结果的匹配情况如图8所示,各类别OA如表4所示。

实验在 Modelnet40数据集上的部分分类结果如图8所示,各类别OA结果如表4所示。大部分类别的识别效果较好,如 guitar、bookshelf、bowl 等达到了100%的OA,表明模型对这些类别有很强的识别能力。其中 cup、radio、wardrobe、xbox 等小物体或结构简单的类别因局部细节有限,容易出现误判,OA在75%。bathtub 被错误地分类为 bed 的情况较为明显,说明模型在这两个类别的区分上存在一定困难。flower_pot 的OA仅为30.0%,主要原因是该类形状多变,且与植物叶片结构易混淆。

ScanObjectNN数据集上部分类别与分类结果的匹配情况如图9所示,各类别OA如表5所示。



(a) 损失函数变化



(b) 准确率变化

((a) changes in loss function (b) changes in accuracy)

图7 ShapeNet Part数据集损失函数和平均交并比变化

Fig. 7 ShapeNet Part dataset loss function and mean cross-union ratio variation

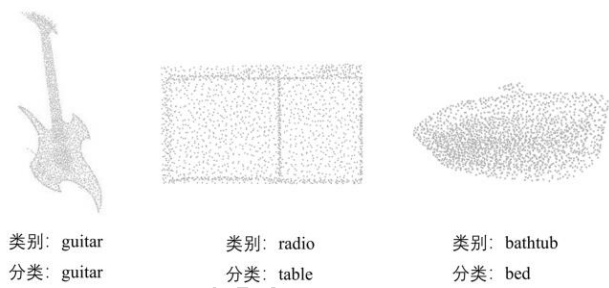


图8 ModelNet40部分数据类别与分类结果的匹配情况

Fig. 8 Matching results between ModelNet40 partial data categories and classification results

实验在 ScanObjectNN 上部分分类结果如图9所示, 各类别 OA 结果如表5所示, 实验结果显示, 各类别间差异显著。表现较好的类别为 chair(97.44%)、toilet(92.86%)、cabinet(83.0%)等, 这些类别具有相对固定的结构。模型对 box 和 cabinet 的识别更稳定准确, 而 table 和 toilet 的正确分类样本偏少, 混淆较

表4 ModelNet40数据集各类别 OA

Table 4 Per-class overall accuracy on the ModelNet40 dataset

类别	OA/%	类别	OA/%
airplane	100.00	laptop	100.00
bathtub	94.00	mantel	96.00
bed	99.00	monitor	98.00
bench	90.00	night_stand	84.00
bookshelf	100.00	person	100.00
bottle	98.00	piano	94.00
bowl	100.00	plant	86.00
car	97.00	radio	75.00
chair	93.00	range_hood	98.00
cone	95.00	sink	90.00
cup	75.00	sofa	100.00
curtain	95.00	stairs	90.00
desk	91.86	stool	85.00
door	95.00	table	77.00
dresser	87.21	tent	95.00
flower_pot	30.00	toilet	100.00
glass_box	95.00	tv_stand	87.00
guitar	100.00	vase	93.00
keyboard	100.00	wardrobe	75.00
lamp	95.00	xbox	75.00

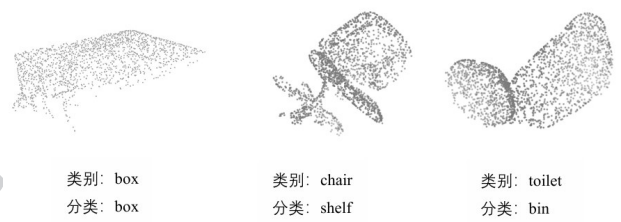


图9 ScanObjectNN部分数据类别与分类结果的匹配情况

Fig. 9 Matching results between ScanObjectNN partial data categories and classification results

明显, pillow 和 bag 则因形状可塑性高而难以区分。总体来看, 模型在 ModelNet40 上的表现更好, 而在 ScanObjectNN 这真实数据集中, 对于相似或形态复杂类别的区分仍有改进空间。

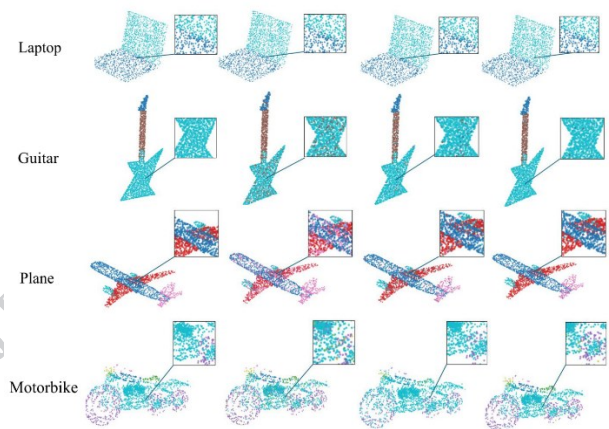
ShapeNetPart 数据集部件分类结果可视化如图10所示。ShapeNet Part 数据集上各类别的 I-mIoU 如

表5 ScanObjectNN数据集各类别OA

Table 6 Per class overall accuracy on the ModelNet40 dataset

类别	OA/%	类别	OA/%
bag	70.59%	shelf	83.67%
bin	85.29%	table	68.18%
box	90.48%	bed	85.19%
cabinet	83.00%	sofa	85.67%
chair	97.44%	sink	76.54%
desk	63.33%	toilet	67.92%
display	83.33%	pillow	64.15%
door	92.86%		

表6所示。为直观展示本文模型在部件分类上的效果,对比了本文模型与PointNet、PointNeXt-s在ShapeNet Part测试集上四个典型类别(Laptop、Guitar、Plane、Motorbike)的部件分类结果。图中每行代表一个物体类别,每列代表一种方法:(a)真实标注,(b)PointNet,(c)PointNeXt-s,(d)本文模型。四个类别分别对应“规整结构、渐变曲面、交错部件、细小细节”四种几何特征,能够较全面地反映模型的分割能力。



(a)Ground Truth (b)PointNet (c)PointNeXt-s (d)Ours
(a) ground truth (b)PointNet (c)PointNeXt-s (d)ours

图10 ShapeNetPart数据集部件分类结果可视化

Fig. 10 Visualization of part segmentation results on ShapeNet-Part dataset

实验结果可以看出,PointNet的分类结果较为粗糙,部件边界模糊,在Laptop的触控板区域、Guitar的琴头、Motorbike的车轮辐条等细小部件上存在明

显的漏分或错分。PointNeXt-s相比PointNet有了明显改善,能大致区分主要部件,但在Plane的机翼与机身交界、Motorbike的车把与油箱等几何细节复杂或部件连接处,仍存在局部粘连或误分类。而本文模型的分类结果最接近真实标注:Laptop的屏幕、键盘与触控板区域被清晰分离,边界整齐;Guitar的琴身、琴颈、琴头三个部件均被准确区分,尤其是琴头与琴颈的过渡处未出现过度分割;Plane的机翼、尾翼、机身和引擎部件划分合理,机翼与机身的几何交界面处理干净;Motorbike的车轮、油箱、车座、车把等多个细小部件均得到正确识别,车轮的圆形边界也得到了较好的保持。上述可视化结果验证了本文模型对细粒度几何结构的建模能力

表6 ShapeNet Part数据集各类别分类I-mIoU

Table 6 Per-class classification I-mIoU on the ShapeNet Part dataset

类别	I-mIoU/%	类别	I-mIoU/%
Airplane	82.05	Lamp	79.29
Bag	82.34	Laptop	94.75
Cap	68.54	Motorbike	65.22
Car	76.98	Mug	92.20
Chair	88.75	Pistol	81.51
Earphone	70.74	Rocket	57.44
Guitar	91.16	Skateboard	78.87
Knife	86.25	Table	83.27

ShapeNet Part数据集各类别分类I-mIoU如表6所示。实验结果显示,本文模型在Laptop(94.75%)、Mug(92.20%)、Guitar(91.16%)等结构规整、部件边界清晰的类别上取得了较高精度,说明模型对具有稳定几何形态的物体具有良好的部件分类能力。相比之下,Rocket(57.44%)、Motorbike(65.22%)、Cap(68.54%)、Earphone(70.74%)等类别表现偏低,主要原因在于这些类别样本数量较少、部件形态差异大或部件之间相互粘连,增加了部件分类难度。

2.3 模型复杂度分析

不同点云模型的参数量对比如表7所示。

Point-BERT、Point-MAE、Swin3D为代表的Transformer类模型虽能依赖自注意力机制构建长程依赖,特征表达能力强,但参数量通常较大;PV-

表7 各模型参数量

Table 7 Parameter Quantities for Each Model

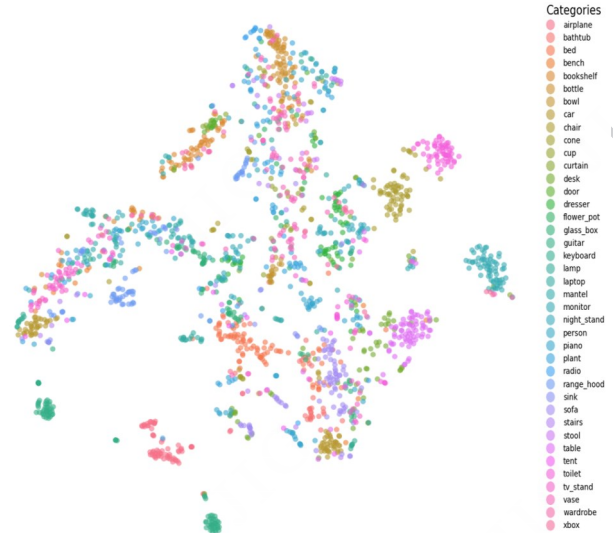
模型	参数量/MB
PointNet	3.47
PointNet++	1.48
DGCNN	1.81
PointMLP	12.95
CurveNet	4.12
Point-BERT	9.76
PointTransformer v1	12.43
Swin-3D	18.67
Transformer	21.00
PV-RCNN	25.84
本文	3.62

RCNN (point-voxel recurrent convolutional neural network) 为代表的多表征融合架构, 由于集成了点与体素双重处理分支, 参数量高达 25.84MB。这类重型普遍推理延迟与显存占用高, 难以在边缘设备上实时部署。PointNet++ (1.48MB)、DGCNN (1.81MB) 等经典轻量模型通过局部邻域聚合与动态图卷积实现分类, 但对复杂几何结构的捕捉能力仍有限。本文模型参数量仅为 3.62M, 与 CurveNet 等模型基本持平, 远低于主流 Transformer 类模型。通过构建坐标增强分支、全局几何分支和局部谱边分支三路并行结构, 并以仿射变换模型、通道注意力机制、双池化聚合策略, 增强了几何特征提取与拓扑建模能力。实验表明, 本文模型在 ModelNet40 数据集上 OA 达到 92.99%, mAcc 达到 90.70%, 超越绝大多数同参数量级方法并逼近部分重型 Transformer 模型, 实现了精度与计算效率的优良平衡。

2.4 分类特征可视化

通过 t 分布随机嵌入 (t-Distributed Stochastic Neighbor Embedding, t-SNE) 算法可将高维特征描述子降维并投影至二维空间, 实现分类结果可视化。ModelNet40 和 ScanObjectNN 数据集的可视化结果如图 11 和图 12 所示, 其中同一类别数据点为相同颜色, 并聚集形成不同的聚类簇, 不同类别的聚类簇呈现明显分离态。在此可视化结果中, 若同类数据点的类内聚集度越高, 且不同类别聚类簇之间的间隔越清晰, 则表明该特征描述子具备更强的类别区分

能力。对于 ShapeNet Part 部件分类任务, t-SNE 可视化主要面向全局特征向量的类别可分性分析, 而部件分类更关注点级特征的局部判别能力, 其评估通常采用部件着色图等直观形式 (如图 10 所示), 因此本文未在该数据集上进行 t-SNE 实验。



(a) 分类前 t-SNE 可视化结果



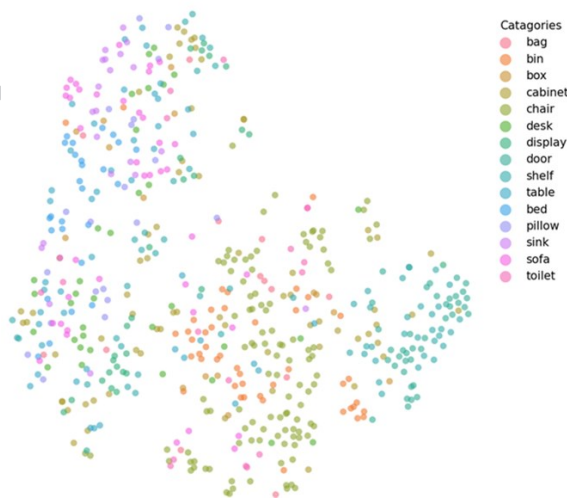
(b) 分类后 t-SNE 可视化结果

(a) visualization of t-SNE before classification (b) visualization of t-SNE after classification

图 11 在 ModelNet40 上的 t-SNE 可视化结果

Fig. 11 Visualization of t-SNE on ModelNet40

ModelNet40 和 ScanObjectNN 两个数据集分类前 t-SNE 可视化结果如图 11(a) 和图 12(a) 所示, 分类后的 t-SNE 可视化结果如图 11(b) 和图 12(b) 所示。可以看出, 分类前两个数据集的特征分布较为混杂。经过本模型提取后, 分类后的特征分布呈现出较清



(a) 分类前 t-SNE 可视化结果



(b) 分类后 t-SNE 可视化结果

(a) visualization of t-SNE before classification (b) visualization of t-SNE after classification

图 12 在 ScanObjectNN 上的 t-SNE 可视化结果

Fig. 12 Visualization results of t-SNE on ScanObjectNN

晰的类别聚类,各类别以不同颜色区分,反应模型强大的分类能力。ScanObjectNN 数据集中同类点的聚集程度更低、类间边界更模糊,反映出本文模型在真实场景下分类性能尚有不足。

2.5 鲁棒性实验

为系统评估本文模型在面对非理想点云数据的鲁棒性与泛化能力,分别在 ModelNet40 和 ScanObjectNN 数据集上开展针对噪声、遮挡及残缺三种典型退化场景的鲁棒性测试。在测试阶段对输入点云分别施加不同强度的高斯噪声、空间区域遮挡及均匀点丢弃扰动,以模拟实际三维传感中的传感器噪

声、视野遮挡及采样稀疏问题。选取 PointNet、PointNet++、DGCNN 及 PointNeXt-S 作为对比方法,通过对比各模型在不同退化程度下的 OA 下降幅度,分析本文方法在复杂环境下的性能保持能力。需要说明的是,上述鲁棒性实验未在部件分类数据集 ShapeNet Part 上实验,原因在于该数据集为干净合成数据,其部件标注基于完整拓扑结构,施加扰动会破坏部件语义完整性,导致指标无法有效评估模型能力。

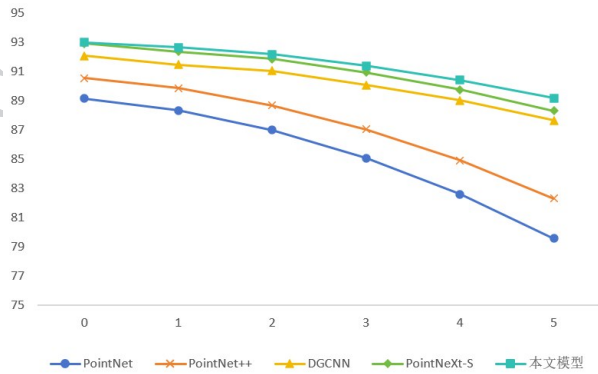
2.5.1 噪声鲁棒性实验

为评估模型抗噪性能,本文在 ModelNet40 与 ScanObjectNN 数据集上进行抗噪对比实验。在保持点云拓扑结构不变的前提下,向原始点云中每个点的三维坐标分别添加零均值的高斯随机扰动,以模拟传感器噪声。实验采用无量纲的相对强度表示法,以高斯分布的标准差 σ 衡量噪声强度,从而消除物理尺度影响并客观评价模型的鲁棒性。设置 $\sigma = 0, 0.01, 0.02, 0.03, 0.04, 0.05$ 共 6 个等级 ($\sigma = 0$ 为无噪声基准),对应图 11 中 0-5 的整数等级,覆盖从轻微扰动到较强干扰的噪声强度范围。分别在两个数据集上进行评测,实验结果如图 13 不同噪声强度下各模型的 OA 变化所示。

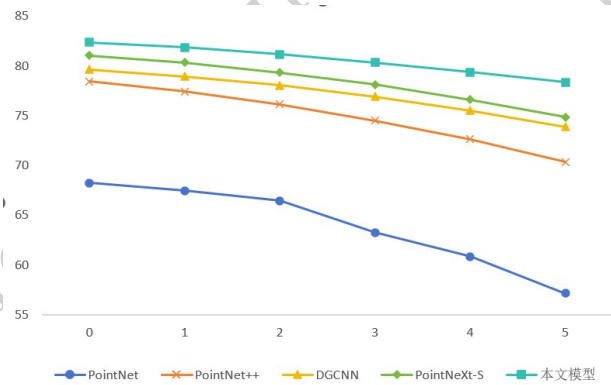
在 ModelNet40 数据集上的实验结果如图 13(a) 所示,所有模型的准确率均随 σ 增大而下降,本文模型在各噪声等级下始终保持最优。 $\sigma = 0$ 时本文模型 OA 为 92.99%,略优于 PointNeXt-S;当 $\sigma = 0.05$ 时,本文模型 OA 为 89.17%,分别高出 PointNet、PointNet++、DGCNN 和 PointNeXt-S 了 9.63%、6.88%、1.54% 和 0.88%。在更具挑战性的 ScanObjectNN 数据集上的实验结果如图 13(b) 所示,本文模型同样表现最优, $\sigma = 0.05$ 时 OA 为 78.34%,比 PointNeXt-S 高 4.45%,比 PointNet++ 高 8.12%。实验结果表明本文模型对高斯噪声具有较强的鲁棒性,即使在强噪声干扰下,本文方法仍能提取到有效的几何特征。上述抗噪优势主要通过频域解耦分支对高频噪声与几何结构的分离能力,及通道注意力机制对受污染特征通道的自适应抑制。

2.5.2 顶部遮挡鲁棒性实验

为评估模型在物体部分区域被遮挡情况下的识别能力,本文在 ModelNet40 与 ScanObjectNN 数据集上进行遮挡鲁棒性测试。模拟真实采集中的顶部遮挡场景,随机移除点云中不同比例的上半部分点。遮挡比例设为 0%、10%、20%、30%、40%、50% 共六



(a) ModelNet40噪声实验结果



(b) ScanObjectNN噪声实验结果

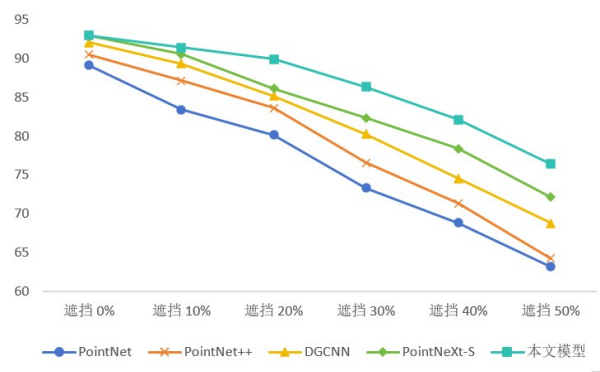
(a) noise experimental results on modelnet40 (b) noise experimental results on scanobjectnn

图 13 不同噪声强度下各模型的OA变化

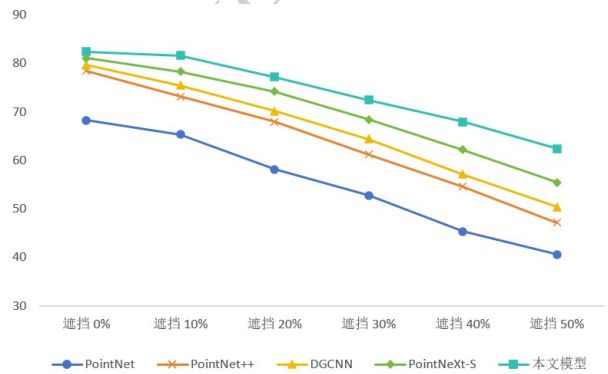
Fig. 13 Overall accuracy variation of each model under different noise intensities

个等级,其中0%表示无遮挡原始数据,50%表示上半部分点被移除。该划分覆盖了从轻微到严重遮挡的强度范围。实验结果不同遮挡强度下各模型的OA变化如图14所示。

在ModelNet40数据集上的实验结果如图14(a)所示,所有模型的准确率均随遮挡比例增加而下降,本文模型的下降幅度最为平缓。随着遮挡加剧,本文模型的领先优势持续扩大,遮挡10%时,本文模型OA为91.45%,较PointNetXt-S高0.89%;遮挡比例增至30%时,较PointNetXt-S差距扩大至3.74%;当遮挡比例达到50%时,本文模型OA仍可维持76.42%,相较PointNet、PointNet++、DGCNN和PointNetXt-S分别领先13.28%、12.19%、7.67%和4.27%。在ScanObjectNN数据集上的结果如图14(b)所示,本文模型OA始终高于对比模型。尤其当遮挡比例为50%时,本文模型OA为62.34%,比



(a) ModelNet40遮挡实验结果



(b) ScanObjectNN遮挡实验结果

(a) occlusion experimental results on modelnet40 (b) occlusion experimental results on scanobjectnn

图 14 不同遮挡强度下各模型的OA变化

Fig. 14 Overall accuracy variation of each model under different occlusion ratios

PointNetXt-S高出6.91%,比PointNet++高出15.22%。实验结果表明,本文模型对顶部遮挡具有优越的鲁棒性,即使丢失近半数的点云结构,仍能通过残留的几何信息实现分类。本文方法较强的的遮挡鲁棒性主要源于全局几何分支利用动态KNN图捕获的拓扑关系,使模型在局部区域被遮挡时能够从邻域中推断整体形状,同时多分支融合策略提供了互补的几何信息。

2.5.3 残缺鲁棒性实验

为检验本文方法在点云局部数据丢失情况下的识别能力,在ModelNet40与ScanObjectNN数据集上进行残缺鲁棒性测试。采用随机点删除方式,在点云中随机选取点并将其删除,模拟物体扫描确实。残缺强度通过删除点总数占比进行控制,设为0%、10%、20%、30%、40%、50%共六个等级,其中0%表示完整点云,50%表示删除的点覆盖约一半的物体

表面区域。该范围覆盖从轻微表面缺损到严重局部残缺的典型场景。分别在两个数据集上进行评测, 实验结果不同残缺强度下各模型的OA变化如图15所示。

在ModelNet40数据集上实验结果如图15(a)显示, 随丢弃比例上升各模型准确率均下降。丢弃20%时本文方法OA达92.29%, 分别高出PointNet++、DGCNN和PointNeXt-S了3.09%、1.15%和0.57%; 丢弃50%时仍保持88.64%, 领先幅度扩大至6.52%、2.37%和2.15%。在ScanObjectNN数据集上实验结果如图15(b)所示, 本文模型同样稳居第一: 30%残缺时OA为80.88%, 比PointNet++和PointNeXt-S高4.83%、3.04%; 50%残缺时OA为78.26%, 领先优势进一步增至4.72%和3.94%。通过实验结果可知, 即使点云严重稀疏, 本文模型仍能有效分类, 该优势主要来源于多粒度特征分支之间的信息互补, 即使点云严重稀疏, 各分支间仍可提供互补的判别信息, 共同维系分类性能。

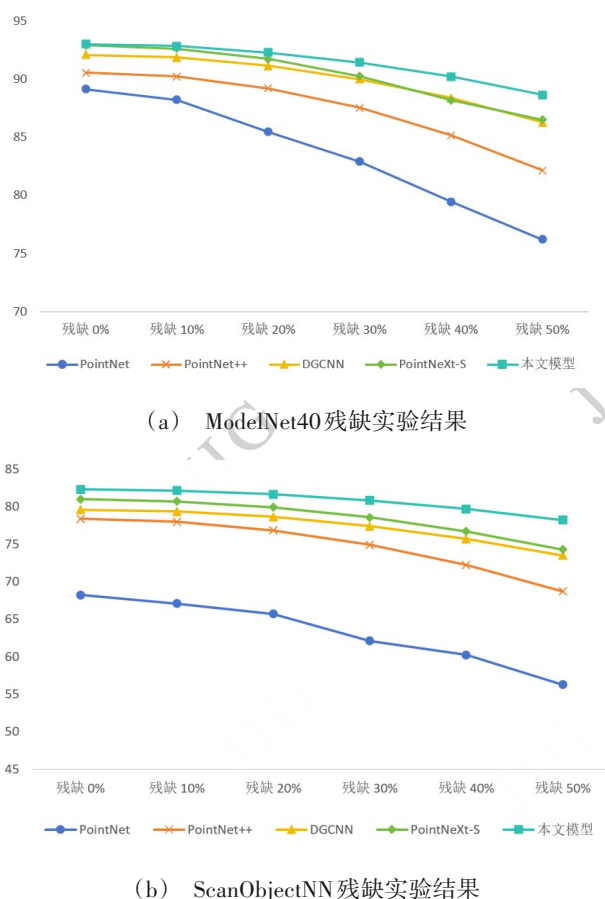
2.6 消融实验

为验证本文模型中各核心分支有效以及关键参数的影响, 设计了系列消融实验, 并在三个数据集上分别采用相应的评价指标: ModelNet40和ScanObjectNN使用OA, ShapeNet Part 部件分类任务使用IoU。实验结果如表8至表10所示。

2.6.1 特征分支消融分析

通过逐一移除三个分支进行消融实验, 三个特征分支的消融实验结果如表8所示。

由三个特征分支的消融实验结果可以看出, 完整模型在三个数据集上均显著优于各子分支。ModelNet40上OA达到92.99%, 在ScanObjectNN数据集



(a) dropout experimental results on modelnet40 (b) dropout experimental results on scanobjectnn

图15 不同残缺强度下各模型的OA变化

Fig. 15 Overall accuracy variation of each model under different dropout ratios

上OA取得82.35%, 全面领先所有子结构。单一分支OA在ModelNet40上达到仅为89.14%~89.63%, 在ScanObjectNN上达到77.25%~78.42%, 在Shap-

表8 三个特征分支的消融实验结果

Table 8 Ablation experimental results of the three characteristic branches

点特征分支	全局几何分支	局部谱分支	Modelnet40	ScanObjectNN	ShapeNet Part
×	√	×	89.14	78.25	79.23
×	×	√	89.63	77.86	80.51
√	×	×	89.47	77.42	79.82
√	√	×	91.28	79.64	83.61
√	×	√	91.74	80.21	83.92
×	√	√	91.03	80.67	82.44
√	√	√	92.99	82.35	85.94

注:加粗字体为每列最优结果。“√”表示采用该分支,“×”表示未采用该分支。

eNet Part 上 I-mIoU 达到 79.23%~85.94% 表明仅依赖任一类特征均存在明显局限。两两组合后点特征分支与局部谱分支的组合效果最佳: ModelNet40 达 91.74%, ScanObjectNN 达 80.21%, ShapeNet Part 达 83.92%, 说明基础坐标统计与频域局部细节具有互补性。完整模型融合三个分支后, 相比最优双分支, ModelNet40 提升 1.25%, ScanObjectNN 提升 2.14%, ShapeNet Part 提升 2.02%, 证实了点特征、全局几何与局部谱特征三者缺一不可。

2.6.2 边缘卷积层数的消融实验

为全局几何分支中 EdgeConv 堆叠层数对模型性能的影响, 进行相关实验, 结果如表 9 所示。

表 9 边缘卷积层数消融实验结果

Table 9 Edge convolutional layer ablation experiment results

EdgeConv 层数	参数量/MB	ModelNet40	ScanObjectNN	ShapeNet Part
1层	1.74	91.36	79.04	83.19
2层	2.19	92.19	81.25	83.51
3层	2.67	92.71	81.83	84.67
4层	3.62	92.99	82.35	85.94
5层	5.93	93.06	82.38	85.67

注: 加粗字体为每列最优结果。

实验结果表明, 随着 EdgeConv 层数从 1 层增至 4 层, 三个数据集上的性能持续提升。参数量由 1.74M 增至 3.62M, ModelNet40 的 OA 从 91.36% 升至 92.99%, 上升 1.63%。ScanObjectNN 的 OA 从 79.04% 升至 82.35%, 上升 3.31%。ShapeNet Part 的 I-mIoU 从 83.19% 升至 85.93%, 上升 2.74%。层数增至 5 层时, 参数量增加至 5.93M, 但性能几乎不再提升, ShapeNet Part 指标反而下降 0.26%, 表明出现过拟合与特征冗余。因此本文选用 4 层 EdgeConv, 在精度与效率之间取得最佳平衡。

2.6.3 边缘卷积 k 值设置的消融实验

卷积层不同 k 值策略的结果如表 10 所示。

在统一 k 值策略下, 各数据集性能随 k 值增大提升有限, ShapeNet Part 指标在 k=20 时达到峰值 84.27% 后回落。逐层递增的 k 值组合显著优于统一设置, 其中 16/20/24 的 k 值设置, 在三个数据集上均取得最佳结果: ModelNet40 的 OA 达到 92.99%、

表 10 边缘卷积不同 k 值设置实验结果对比

Table 10 Comparison of experimental results for edge convolution with different k values.

$k_1/k_2/k_3$	ModelNet40	ScanObjectNN	ShapeNet Part
16/16/16	92.44	81.52	83.72
20/20/20	92.57	81.71	84.27
24/24/24	92.61	81.96	83.96
12/16/20	92.51	82.05	85.18
16/20/24	92.99	82.35	85.94
20/24/28	92.68	82.21	85.67
24/28/32	92.63	81.75	85.08

注: 加粗字体为每列最优结果。

ScanObjectNN 的 OA 达到 82.35%、ShapeNet Part 的 I-mIoU 为 85.94%, 而 12/16/20 和 20/24/28 的 k 值设置性能居中, 24/28/32 的 k 值设置因邻域噪声增加导致性能下降。这表明浅层小 k 值利于局部细节提取、深层大 k 值利于全局上下文聚合, 由细到粗的多粒度建模策略在分类与分割任务中均有效, 因此本文采用 16/20/24 作为 EdgeConv 的 k 值配置。

2.8 输入点数敏感性实验

不同点数对模型性能的影响结果如表 11 所示。部件分类任务对点密度敏感且常采用固定点数, 故表 11 未包含 ShapeNet Part 结果。

表 11 输入点数消融实验

Table 11 Ablation Experiment with Input Points

输入点数	FLOPs/G	ModelNet40	ScanObjectNN
256	0.72	90.83	78.26
512	1.38	91.37	80.29
1024	2.71	92.99	82.35
2048	5.36	93.07	82.42

注: 加粗字体为每列最优结果。

实验显示, 随着点数从 256 增至 1024, ModelNet40 的 OA 从 90.83% 稳步提升至 92.99%, ScanObjectNN 从 79.26% 提升至 82.35%, 累计增益分别达 1.88% 和 2.09%。这主要是因为更密集的采样能更完整地保留物体表面细节与局部几何结构, 尤其在存在噪声、缺失或扰动情况下, 额外点可显著增强模型对真实形状的感知能力。继续增加至 2048 点时, 在 ModelNet40 上仅提升 0.08%、ScanObjectNN 上

提升 0.07%, 但每秒浮点运算次数 (floating-point operations per second, FLOPs) 几乎翻倍, 说明此时已出现明显的采样冗余, 边缘区域的重复点对特征表达贡献甚微。综合精度与计算开销考虑, 1024 点实现了最佳平衡, 既保证了充分的几何信息, 且避免不必要的计算浪费。

3 结论

针对现有点云分类网络在局部特征建模不足、全局拓扑建模受限及多尺度特征融合能力有限等问题, 提出一种基于空间域几何特征与频域解耦的多粒度点云分类网络。该网络包含三个核心分支: 点特征分支采用简洁高效的结构提取点级细粒度几何特征; 全局几何分支基于改进边缘卷积构建动态 KNN 图, 实现点云全局拓扑关系与空间结构的层次化建模; 局部谱分支通过低频通道捕获平滑趋势, 高频通道聚焦几何细节变化, 谱差异通道编码基于距离的谱域差分信息, 实现局部几何特征的多尺度解耦。三个分支分别从点级局部几何、全局拓扑结构和谱域特征三个维度互补建模几何表征, 形成多粒度几何描述体系。在 ModelNet40 数据集上, 本方法 mAcc 达 90.70%, OA 为 92.99%, 在 ScanObjectNN 数据集上, 本方法 mAcc 达 79.84%, OA 为 82.35%, 优于现有主流方法, 在 ShapeNet Part 部件分类任务上, 本方法取得了 85.94% 的实例平均交并比和 83.32% 的类别平均交并比, 与现有主流方法持平。尽管该方法在三类数据集上均取得了较高的分类性能, 但对于部分形态相近的类别仍存在一定程度的混淆, 在更接近真实场景的 ScanObjectNN 数据集上, 对受噪声、遮挡或结构不规则影响的样本, 其判别能力仍有不足, 表明模型在复杂环境下的鲁棒性尚待加强。在 ShapeNet Part 的 Rocket、Motorbike 等困难类别上, 部件分类精度仍有提升空间。后续工作考虑引入多头注意力机制进一步提升分类性能, 并探索自适应代表点采样策略, 以增强模型对多尺度结构变化的适应性和整体分类性能。

参考文献 (References)

Chen H J, Da F P and Gai S Y. 2021. Deep 3D point cloud classification network based on competitive attention fusion. *Journal of Zhejiang*

University (Engineering Science), 55(12): 2342-2351 (陈涵娟, 达飞鹏, 盖绍彦. 2021. 基于竞争注意力融合的深度三维点云分类网络. *浙江大学学报(工学版)*, 55(12): 2342-2351 [DOI: 10.3785/j.issn.1008-973X.2021.12.014]

Daif H and Marzouk M. 2025. Point cloud classification and part segmentation of steel structure elements. *Neural Computing & Applications*, 37(6): 10733 [DOI: 10.1007/s00521-024-10733-x]

Engel N, Belagiannis V and Dietmayer K. 2021. Point Transformer. *IEEE Access*, 9: 134826-134840 [DOI: 10.1109/ACCESS.2021.3116304]

Feng Y, Zhang Z, Zhao X, Ji R and Yue G. 2018. GVCNN: group-view convolutional neural networks for 3D shape recognition // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE: 264-272 [DOI: 10.1109/CVPR.2018.00035]

Guo M H, Cai J X, Liu Z N, Mu T J, Ralph R. Martin and Hu S M. 2021. PCT: Point cloud transformer. *Computational Visual Media*, 7(2): 187-199. [DOI: 10.1007/s41095-021-0229-5]

Klokov R and Lempitsky V. 2017. Escape from cells: deep Kd-networks for the recognition of 3D point cloud models // *Proceedings of the 16th IEEE International Conference on Computer Vision*. Venice, Italy: IEEE: 863-872 [DOI: 10.1109/ICCV.2017.99]

Li X J, Yang J and Zhang F L. 2022. Laplacian mesh transformer: Dual attention and topology aware network for 3D mesh classification and segmentation // *Proceedings of the European Conference on Computer Vision (ECCV 2022)*. Cham: Springer: 541-560. [DOI: 10.1007/978-3-031-19818-2_31]

Li Y Y, Bu R, Sun M C, Wu W, Di X H and Chen B Q. 2018. PointCNN: convolution on X-transformed points // *Proceedings of the 32nd International Conference on Neural Information Processing Systems*. Montréal, Canada: Curran Associates Inc.: 828-838

Lu D N, Xie Q, Xu L L and Li J. 2022. 3DCTN: 3D convolution Transformer network for point cloud classification [EB/OL]. [2023-02-21]. <http://arxiv.org/pdf/2203.00828.pdf>

Pan J C, Cao J Y, Xing S, Dai M F, Liu J K and Wang X Y, et al. 2025. An aerial point cloud classification using point transformer via multi-feature fusion. *Scientific Reports*, 15: 22924 [DOI: 10.1038/s41598-025-02719-z]

Qi C R, Su H, Mo K C and Guibas L J. 2017a. PointNet: deep learning on point sets for 3D classification and segmentation // *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, USA: IEEE: 77-85 [DOI: 10.1109/CVPR.2017.16]

Qi C R, Yi L, Su H and Guibas L J. 2017b. PointNet++: deep hierarchical feature learning on point sets in a metric space // *Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach, USA: Curran Associates Inc.: 5105-5114

Qian G C, Li Y, Peng H, Mai J J, Hammoud H and Elhoseiny M, et al.

2022. PointNeXt: revisiting PointNet++ with improved training and scaling strategies. [EB/OL]. [2026-03-05].
<https://doi.org/10.48550/arXiv.2206.04670>
- Shi X S, Cheng Y L, Zhao Z Y and Qin X X. 2019. Point Cloud Classification Algorithm Based on IPTD and SVM. *Laser & Optoelectronics Progress*, 56(16): 161002 (释小松, 程英蕾, 赵中阳, 秦先祥. 2019. 基于三角网滤波和支持向量机的点云分类算法. *激光与光电子学进展*, 56(16): 161002 [DOI: 10.3788/LOP56.161002])
- Tao S B, Liang C, Jiang T P, Yang Y J and Wang Y J. 2021. Sparse voxel pyramid neighborhood construction and classification of LiDAR point cloud. *Journal of Image and Graphics*, 26(11): 2703-2712 (陶帅兵, 梁冲, 蒋腾平, 杨玉娇, 王永君. 2021. 激光点云的稀疏体素金字塔邻域构建与分类. *中国图象图形学报*, 26(11): 2703-2712) [DOI: 10.11834/jig.200262]
- Thomas H, Qi C R, Deschaud J E, Marcotegui B, Goulette F and Guibas L. 2019. KPConv: flexible and deformable convolution for point clouds // *Proceedings of 2019 IEEE/CVF International Conference on Computer Vision*. Seoul, Korea (South): IEEE: 6410-6419 [DOI: 10.1109/ICCV.2019.00651]
- Wang C, Pelillo M and Siddiqi K. 2019. Dominant set clustering and pooling for multi-view 3D object recognition [EB/OL]. [2026-03-05].
<https://doi.org/10.48550/arXiv.1906.01592>
- Wang C, Samari B and Siddiqi K. 2018. Local spectral graph convolution for point set feature learning // *Proceedings of the European Conference on Computer Vision*. Munich, Germany: Springer: 52-66 [DOI: 10.1007/978-3-030-01225-0_4]
- Wang Y, Sun Y B, Liu Z W, Sarma S E, Bronstein M M and Solomon J M. 2019. Dynamic graph CNN for learning on point clouds. *ACM Transactions on Graphics*, 38(5): #146 [DOI: 10.1145/3326362]
- Wang Z B, Zeng Z X, Ouyang X, Chen H Z, Li L J and Yang J Q. 2025. Recent progress in rotation-invariant point cloud networks. *Journal of Image and Graphics*, 30(12): 3782-3803 (王正宝, 曾振轩, 欧阳轩, 陈昊哲, 李林杰, 杨佳琪. 2025. 旋转不变点云网络研究进展. *中国图象图形学报*, 30(12): 3782-3803) [DOI: 10.11834/jig.250029]
- Wu X, Jiang L, Wang P S, Liu Z and Qiao Y, et al. 2024. Point Transformer V3: simpler, faster, stronger // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE: 4840-4851 [DOI: 10.1109/CVPR.2024.00463]
- Yan X, Zheng C D, Li Z, Wang S and Cui S G. 2020. PointASNL: robust point clouds processing using nonlocal neural networks with adaptive sampling // *Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, USA: IEEE: 5588-5597 [DOI: 10.1109/CVPR42600.2020.00563]
- Zhang G Y, Zuo J X, Pan L H, Hao Z X, Guo W and Zhang X N. 2025. Point-GBLS: 3D point cloud classification network combined with deep-broad learning. *Computer Systems & Applications*, 34(03): 1-13 (张国有, 左嘉欣, 潘理虎, 等. 2025. Point-GBLS: 结合深宽度学习的三维点云分类网络. *计算机系统应用*), 34(03): 1-13 [DOI: 10.15888/j.cnki.csa.009796]
- Zhang K G, Hao M, Wang J, de Silva C W and Fu C L. 2019. Linked dynamic graph CNN: learning on point cloud via linking hierarchical features [EB/OL]. [2023-02-21].
<http://arxiv.org/pdf/1904.10014.pdf>
- Zhang T, Yuan H, Qi L, Zhang J N, Zhou Q Y and Ji S P, et al. 2025. Point cloud mamba: point cloud learning via state space model // *Proceedings of the 39th AAAI Conference on Artificial Intelligence*. Philadelphia: AAAI Press: 10121-10130 [DOI: 10.1609/aaai.v39i10.33098]
- Zhao H S, Jiang L, Jia J Y, Torr P and Koltun V. 2021. Point transformer // *Proceedings of 2021 IEEE/CVF International Conference on Computer Vision*. Montreal, Canada: IEEE: 16239-16248 [DOI: 10.1109/ICCV48922.2021.01595]
- Zhong Q and Han X F. 2022. Point cloud learning with Transformer [EB/OL]. [2023-02-21].
<http://arxiv.org/pdf/2104.13636.pdf>

作者简介

李晶晶, 2002年生, 硕士研究生, 主要研究方向为深度学习和点云分类。E-mail: lijingjing0518@126.com