

# 融合栅格与表面特征编码的轻量级点云分类网络

杨官学, 周昊, 刘慧, 沈跃, 徐婕

(江苏大学电气信息工程学院, 江苏镇江 212013)

**摘要:** 点云携带着丰富的三维特征, 其分类问题一直是深度学习领域的热点。现有点云分类网络的精度已经较为理想, 但参数量与计算量过于庞大, 不利于实际场景下的部署。针对该问题, 提出一种轻量级点云分类网络 GridPoint。首先设计点云栅格化模块 Gridding, 根据点的坐标位置进行栅格区域划分; 然后扩展坐标的高次项函数, 对原始点云进行表面特征编码, 增强对轮廓特征的表达; 最后使用两次全局池化完成局部特征的提取与全局特征的聚合。使用经典点云数据集 ModelNet40、ShapeNetCore 与真实数据集 ScanObjectNN 进行分类与消融实验。实验结果表明, GridPoint 的分类精度接近 PointNet++ 等主流网络, 差距在 0.3%~2.3% 之间; 网络参数量与计算量分别为 0.11 M 与 0.05 G, 相较主流网络分别减少了 81.7% 和 88.9% 以上, 在轻量化方面优势显著, 具有良好的实用价值。

**关键词:** 深度学习; 点云分类; 轻量级网络; 点云栅格化; 表面特征编码

DOI: 10.11907/rjdk.241137

开放科学(资源服务)标识码(OSID):

中图分类号: TP391

文献标识码: A

文章编号: 1672-7800(2024)005-0009-08



## Lightweight Network for Point Cloud Classification Based on Gridding and Surface Features Encoder

YANG Guanxue, ZHOU Hao, LIU Hui, SHEN Yue, XU Jie

(School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, China)

**Abstract:** Point clouds carry rich three-dimensional features, and their classification problem has always been a hot topic in the field of deep learning. The accuracy of existing point cloud classification networks is already relatively ideal, but the parameter and computational complexity are too large, which is not conducive to deployment in practical scenarios. A lightweight point cloud classification network, GridPoint, is proposed to address this issue. Firstly, design a point cloud gridding module, which divides the grid area based on the coordinate position of the points; Then expand the higher-order term function of the coordinates, encode the surface features of the original point cloud, and enhance the expression of contour features; Finally, two rounds of global pooling are used to extract local features and aggregate global features. Perform classification and ablation experiments using the classic point cloud dataset ModelNet40, ShapeNetCore, and the real dataset ScanObject NN. The experimental results show that the classification accuracy of GridPoint is close to mainstream networks such as PointNet++, with a difference between 0.3% and 2.3%; The network parameters and computational complexity are 0.11 M and 0.05 G, respectively, which are reduced by more than 81.7% and 88.9% compared to mainstream networks. They have significant advantages in lightweight and have good practical value.

**Key Words:** deep learning; point cloud classification; lightweight network; point cloud gridding; surface feature encoder

### 0 引言

随着激光雷达扫描速度与精度的不断提升, 越来越多

的视觉感知场景开始引入三维点云作为替代或补充, 如机器人定位<sup>[1]</sup>、自动驾驶<sup>[2]</sup>、遥感<sup>[3]</sup>等。相较于图像、体素等传统三维数据格式, 点云包含更丰富的位置、深度与轮廓信息, 逐渐成为深度学习领域一种流行的数据形式<sup>[4]</sup>。如

收稿日期: 2024-02-20

基金项目: 国家自然科学基金项目(32171908)

作者简介: 杨官学(1984-), 男, 博士, 江苏大学电气信息工程学院讲师、硕士生导师, 研究方向为复杂网络重构与分析、机器学习等; 周昊(1999-), 男, 江苏大学电气信息工程学院硕士研究生, 研究方向为深度学习; 刘慧(1980-), 女, 博士, 江苏大学电气信息工程学院教授、博士生导师, 研究方向为智能感知与控制。本文通讯作者: 刘慧。

何利用深度学习方法高效处理点云并完成分类识别等任务成为当前的研究热点。

在传统图像识别领域,由卷积层、池化层与全连接层构建的卷积神经网络(Convolutional Neural Network, CNN)有着广泛且成熟的应用。然而,与图片像素那样规则的排列不同,三维点云具有无序性、离散性和旋转不变性,无法适用常规的卷积计算,因此有部分学者将点云先转换为其他类型的数据,再使用CNN方法进行分类。例如, Su等<sup>[5]</sup>提出基于多视图投影的MVCNN模型,将点云转换为各个角度的投影后再进行图像识别; Maturana等<sup>[6]</sup>提出基于体素的VoxNet模型,先将点云转换为三维体素,再利用三维卷积进行处理。然而以上方法在进行点云转换时会不可避免地丢失大量几何信息,并且会增加大量繁杂的计算过程。

与数据转换相比,基于原始点云的方法更为直观简洁。首个基于原始点云数据的分类网络是由Qi等<sup>[7]</sup>提出的PointNet。该方法使用一维卷积构建多层感知机(Multi-layer Perceptron, MLP),以映射出点的高维特征,同时使用全局对称的池化函数MaxPool进行特征聚合。Qi等<sup>[8]</sup>随后提出基于PointNet改进的PointNet++,通过多层级的球查询与局部特征提取进一步提升了分类精度。此后,更多基于原始点云输入的分类网络被提出。例如, Li等<sup>[9]</sup>提出使用X-Transform对点进行重排序的PointCNN模型; Wang等<sup>[10]</sup>提出使用边缘卷积构建局部图结构的DGCNN模型; Thomas等<sup>[11]</sup>提出使用内核点卷积的KPCConv模型; Guo等<sup>[12]</sup>提出的PCT(Point Cloud Transformer)将自注意力机制引入点云分类领域; Yan等<sup>[13]</sup>提出通过自注意力调整采样点坐标,从而抑制点云噪声的影响; 兰红等<sup>[14]</sup>提出一种三维方向卷积来捕捉节点内的方向信息。然而,以上网络模型在不断提高分类精度的同时,大小、参数量与计算复杂度也在不断增加。

为满足实际应用需求,部分学者提出轻量级的点云分类网络。例如, Li等<sup>[15]</sup>提出的LPCCNet模型使用密集连接的卷积层与分组卷积减少参数量; Cui等<sup>[16]</sup>在PointNet++中引入一种轻量级注意力模块(Lightweight Attention Module, LAM)以降低网络复杂度; Wang等<sup>[17]</sup>通过减少DGCNN中的全连接层数来减少参数量,并引入通道注意力来改善性能损失。然而,以上轻量化方法效果不够理想,没有解决遍历临近点所带来的较大计算量问题,难以满足实际应用场景中需兼顾运算速度、内存占用以及分类精度的要求。为此,本文提出一种轻量级点云分类网络GridPoint。该网络基于经典方法PointNet进行重构,并引入点云的栅格化模块Gridding与表面特征编码(Surface Feature Encoder, SFE)分别提取点云的局部特征与表面轮廓特征,在轻量化模型的同时保证了足够的分类性能。Gridding模块在三维方向上根据坐标位置对点云进行局部栅格划分,以较少的计算量完成局部特征提取; SFE模块通过扩展不同的

高次项函数并进行编码以表达出点云的轮廓特征,使网络在层数较少的情况下获得更好的表达能力与分类性能。在公共的合成点云与真实点云数据集上进行的分类与消融实验结果验证了GridPoint网络的鲁棒性、轻量性以及模块设置的合理性。

## 1 GridPoint网络

点云一般通过激光雷达扫描获得,其数据格式定义为空间中 $N$ 个 $F$ 维向量所构成的集合 $S_N$ 。表示为:

$$S_N = \{X_0, X_1, \dots, X_N\}, X_i \subseteq R^F \quad (1)$$

一般情况下 $X_i \in R^3$ ,即三维向量 $(x_i, y_i, z_i)$ 代表了点的空间坐标。有时 $X_i$ 会包含颜色、反射率等其他信息,维度也会相应增加。作为分类网络的输入时,点云数据表示为 $N \times 3$ 的矩阵。为统一点数以便进行训练,需要对点云进行降采样,通常取点数 $N = 1024$ 。此外,为使分类网络具有更好的收敛速度,避免梯度爆炸,需要对采集到的点云数据进行归一化。表示为:

$$\begin{cases} r = \max(x_{\max} - x_{\min}, y_{\max} - y_{\min}, z_{\max} - z_{\min}), \\ (x'_i, y'_i, z'_i) = \frac{(x_i - x_{\min}, y_i - y_{\min}, z_i - z_{\min})}{r} \end{cases} \quad (2)$$

式中: $r$ 表示点云的最大尺度半径; $(x'_i, y'_i, z'_i)$ 表示归一化后的点。通过同一尺度下的等比例缩小,点云能在保持形状不变的同时坐标均能够处于区间 $[-1, 1]$ 内。

### 1.1 网络整体结构

GridPoint以经典方法PointNet为基础,借助主流方法局部特征提取的思想,对原始网络进行了简化与重构。整体结构如图1所示。图中, $N$ 表示输入点云的总数; $3$ 表示输入空间坐标 $(x, y, z)$ 的维数; $n$ 表示栅格化后每个栅格内的点数; $C$ 表示拼接操作(Concat); $k$ 表示分类任务的类数。网络在点云数据的输入端首先使用Gridding模块将点云依据坐标在三维方向上划分成8个局部栅格区域;然后使用SFE模块扩展坐标的高次项函数,并生成特征编码。随后使用一维卷积构建MLP,将点云特征向高维映射,由原本的15维升至256维。在升维过程中,将原始点云特征与高维特征进行拼接,以保证低维特征不会丢失。在特征维度上使用最大池化,提取局部特征为 $8 \times 256$ 维;然后使用最大池化与平均池化相加进行全局特征融合;最后通过全连接层完成分类。

GridPoint网络包含4个卷积层、2个池化层和1个全连接层,所使用的卷积核大小均为1。由于网络层数较浅,并未在层与层之间设置任何激活函数、批量归一化或dropout层。

### 1.2 点云栅格化模块

以PointNet++为代表的主流网络均证实提取点云的局部特征是分类任务的重要步骤。然而,该过程往往需要使用复杂算法对区域进行划分以及查找区域内的邻近点,例

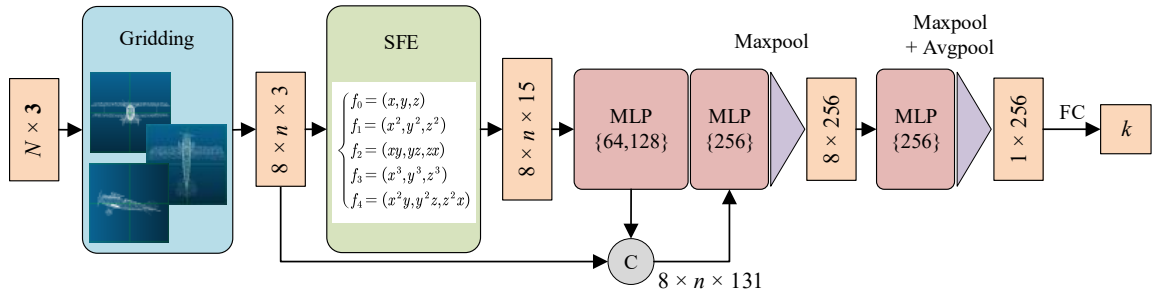


Fig. 1 GridPoint network structure

图1 GridPoint网络结构

如球查询法与 K-最邻近算法 (K-Nearest Neighbors, KNN), 这些算法会带来额外的计算量, 进而影响网络模型的运行速度。早期的体素化方法 VoxNet 先将点云转化为网格, 以便使用三维卷积; Jiang 等<sup>[18]</sup>将尺度不变特征变换 (Scale-invariant Feature Transform, SIFT) 模块应用于点云分类, 从 8 个方向尺度上获取中心点的临近点, 用于代替 KNN。Gridding 模块同样基于 8 个方向的空间尺度, 提出一种基于栅格的、低复杂度的局部特征提取方法, 采用最小分辨率的栅格对点云进行局部区域划分, 在运算中只进行数值大小比较, 不需要进行距离计算。具体来说, 根据坐标位置将点云中的  $N$  个点划分到 8 个同等大小的栅格区域中。点云栅格化示意如图 2 所示。

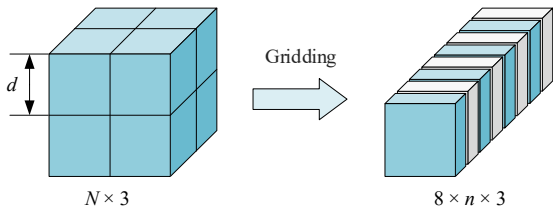


Fig. 2 Schematic of point cloud gridding

图2 点云栅格化示意

在每个栅格内随机采样  $n$  个点云,  $n$  根据栅格边长  $d$  和点总数  $N$  来决定, 并向下取整。表示为:

$$n = N \cdot \frac{d^3}{2^3} \quad (3)$$

若栅格内原始点数不足则进行复制; 若栅格内点数为 0 则从原始点云中随机采样作为替代。由于点坐标被归一化在  $[-1, 1]$  内, 为避免栅格划分时出现点的遗漏, 栅格边长  $d$  应不小于 1。通常总点数  $N = 1024$ , 因此边长  $d$  设置为 1 时, 栅格内的采样点数  $n = 128$ ; 而当  $d$  大于 1 时, 栅格的空间感受野将出现重叠, 这种重叠是可以接受的。例如当  $d = 1.1$  时, 采样点数  $n = 170$ 。

然而, Gridding 模块并未立刻完成对局部特征的提取, 而是仅将点云划分为离散的局部区域。这些局部区域在网络的后续运算中会共享 MLP, 对栅格内的点云特征进行升维, 最后通过两次全局池化完成对局部特征的提取与融合。

### 1.3 点云 SFE

激光雷达扫描并不能穿透物体, 因此得到的点云数据表达的是物体的表面轮廓。而物体表面往往由连续的曲面构成, 可由多次多项式进行拟合<sup>[19]</sup>。例如二次曲面方程的一般表达式为:

$$a_1x^2 + b_1y^2 + c_1z^2 + a_2xy + b_2yz + c_2zx + a_3x + b_3y + c_3z + d = 0 \quad (4)$$

另一种经典约束的三次曲面方程为:

$$z = a_1x^3 + a_2y^3 + b_1x^2y + b_2xy^2 + c_1x^2 + c_2y^2 + c_3xy + d_1x + d_2y + e \quad (5)$$

可以看出, 曲面方程主要包含由  $x, y, z$  构成的多次项。关于曲面相似性的研究指出: 一组有限的几何矩 (Shape Momoent) 通常能够定义曲面的紧凑特征<sup>[20]</sup>, 该结论可以推广至离散采样的点云表面。而在神经网络中, Joseph-Rivlin 等<sup>[21]</sup>指出: 近似出几何矩往往需要大量权重与层数, 但对原始点云的三维向量扩展二次函数后能够更容易地完成拟合。换言之, 对坐标向量扩展高次项函数有利于增强对轮廓特征的表达。

本文参考曲面方程中的几种高次项, 扩展多种不同形式的二次、三次项函数, 包括平方项、乘积项等。通过对这些高次项函数进行拼接构成表达点云轮廓的 SFE。在卷积过程中, 这些多次项会进行线性加权, 从而自然地构成曲面方程的形式, 拟合出点云轮廓。

SFE 示意如图 3 所示, 其可将点云中每个点原本的 3 维向量扩展至 15 维向量。而在实际网络中, 经过 Gridding 模块后, 点云的特征维度会先由  $N \times 3$  变为  $8 \times n \times 3$ , 再经过 SFE 模块扩展至  $8 \times n \times 15$ 。

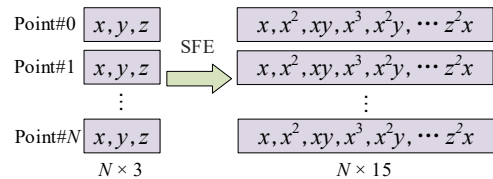


Fig. 3 Schematic of surface feature encoder

图3 SPE示意

SFE 模块对点云扩展函数的具体形式如式 (6) 所示, 其中  $f_0$  表示原始点云向量, 其余  $f_i (i = 1, 2, 3, 4)$  则表示不同的扩展形式。最终编码由所有  $f$  按顺序拼接在一起

构成。

$$\begin{cases} f_0 = (x, y, z) \\ f_1 = (x^2, y^2, z^2) \\ f_2 = (xy, yz, zx) \\ f_3 = (x^3, y^3, z^3) \\ f_4 = (x^2y, y^2z, z^2x) \end{cases} \quad (6)$$

## 2 实验方法与结果分析

### 2.1 实验环境

使用 Dell Precision 7820 塔式工作站,具体配置如表 1 所示。优化器选择自适应矩估计算法 Adam,学习率设置为 0.001,权重衰减设置为 0.000 1。训练与测试过程中的批处理大小均设置为 18,进行 200 轮次的训练。

Table 1 Workstation configuration

表 1 工作站配置

项目	参数
CPU	Intel(R) Xeon(R) Gold 6226R
GPU	NVIDIA GeForce RTX 3090
内存/G	64
操作系统	Ubuntu 20.22
Python	3.10.1
CUDA	11.3
CUDNN	8.2.1
Pytorch	1.12.1

### 2.2 数据集

为评估 GridPoint 的点云分类性能,选择合成点云数据集 ModelNet40<sup>[22]</sup>与 ShapeNetCore<sup>[23]</sup>进行测试。ModelNet40 为点云分析中应用最广泛的经典数据集,其点云形状简洁、构造良好。该数据集中包含 12 311 个物体,分为 40 个类别,其中训练集点云个数为 9 843,测试集点云个数为 2 468。ShapeNetCore 为大规模点云数据集,相较 ModelNet40 有着更丰富的模型类别与数据量,共包含 55 种常见的物品类别,共计 513 000 个三维模型。此外,为进一步验证该网络的准确性与鲁棒性,使用扫描自真实世界物体的点云数据集 ScanObjectNN<sup>[24]</sup>进行实验。该数据集包含 15 个类别,约 15 000 个对象,虽然种类较少,但包含复杂的背景、残缺的部件以及无规律的形变,对于分类网络性能有着更高的要求。

### 2.3 评价指标

本文选择整体精度(Overall Accuracy, OA)与平均精度(Mean Accuracy, mAcc)作为分类结果的评价指标。具体定义为:

$$OA = \frac{TP}{TP + FN} \quad (7)$$

$$mAcc = \frac{1}{k} \sum_{i=1}^k \frac{TP_i}{TP_i + FN_i} \quad (8)$$

式中:TP 和 FN 分别表示对正样本的正确预测和错误

预测;TP+FN 构成整个样本集;k 表示类别数。因此,OA 代表测试实例全体的准确率,mAcc 代表各类别的平均准确率。

在网络复杂度实验中,采用参数量(Params)、浮点运算次数(Floating-point operations, FLOPs)、200 轮训练时间及单轮测试时间 4 项指标评价网络的空间复杂度与时间复杂度。以上指标数值越低,表明网络轻量化程度越高。

### 2.4 分类实验

分类精度为分类网络的基本性能。GridPoint 作为轻量级网络,其层级较浅,因此在精度上存在一定的妥协,与主流网络相比会有些许落后。选择 VoxNet<sup>[6]</sup>、PointNet<sup>[7]</sup>、PointNet++<sup>[8]</sup>、PointCNN<sup>[9]</sup>、DGCNN<sup>[10]</sup>、PCT<sup>[12]</sup>、GCN3D<sup>[14]</sup> 网络作为对照,与 GridPoint 在 ModelNet40 数据集上进行分类性能比较,结果见表 2。可以看出,GridPoint 在 ModelNet40 数据集上的整体分类精度为 91.2%,平均精度为 88.4%,相较原始网络 PointNet 分别提升了 2.0% 与 2.4%,同时远高于基于体素的方法 VoxNet。而与 PointNet++、DGCNN 等主流网络相比,GridPoint 的分类精度略低,差距在 1% 左右;而与性能最好的 PCT 和 GCN3D 相比,GridPoint 的整体精度与平均精度分别只落后 2.0% 与 1.9%。

Table 2 Comparison of classification performance of different models on the ModelNet40 dataset

表 2 不同模型在 ModelNet40 数据集上的分类性能比较 (%)

网络模型	mAcc	OA
VoxNet <sup>[6]</sup>	83.0	85.9
PointNet <sup>[7]</sup>	86.0	89.2
PointNet++ <sup>[8]</sup>	89.5	91.7
PointCNN <sup>[9]</sup>	88.1	92.2
DGCNN <sup>[10]</sup>	89.1	92.5
GCN3D <sup>[14]</sup>	90.3	93.0
PCT <sup>[12]</sup>	90.2	93.2
GridPoint	88.4	91.2

各网络在 ShapeNetCore 与 ScanObjectNN 数据集上的分类实验结果分别如表 3 和表 4 所示。其中,ShapeNetCore 数据集常用于分割实验,因此该数据集上进行的分类实验仅有 OA 可作为评价指标。表 3 实验结果表明,GridPoint 在 ShapeNetCore 数据集上的分类整体精度为 84.4%,相较原始网络 PointNet 提升了 0.7%;与主流网络 DGCNN 相比,GridPoint 仅落后 0.3%;而与精度最高的 KPConv 和 PCT 相比,GridPoint 则落后 1.8%。

由表 4 可知,鉴于 ScanObjectNN 数据集中点云的残缺性和复杂性,大部分网络的分类效果均不理想。其中,GridPoint 的分类整体精度为 76.8%,平均精度为 73.1%,相较 PointNet 分别提高了 8.6% 与 9.7%,性能提升较为明显。与主流网络相比,GridPoint 的分类整体精度比最高的 PCT 落后 2%,平均精度则比最高的 PointNet++ 落后约 2.3%。

综上所述,GridPoint 通过对点云局部与表面轮廓进行特征融合,在结构精简的同时相较原始网络 PointNet 有着最高 9.7% 的提升。但相比使用了 KNN、注意力机制等结

**Table 3 Comparison of classification performance of different models on the ShapeNetCore dataset**

表 3 不同模型在 ShapeNetCore 数据集上的分类性能比较 (%)

网络模型	OA
PointNet <sup>[7]</sup>	83.7
DGCNN <sup>[10]</sup>	84.7
PointNet++ <sup>[8]</sup>	85.1
GCN3D <sup>[14]</sup>	85.2
KPConv <sup>[11]</sup>	86.2
PCT <sup>[12]</sup>	86.2
GridPoint	84.4

**Table 4 Comparison of classification performance of different models on the ScanObjectNN dataset**

表 4 不同模型在 ScanObjectNN 数据集上的分类性能比较 (%)

网络模型	mAcc	OA
PointNet <sup>[7]</sup>	63.4	68.2
PointNet++ <sup>[8]</sup>	75.4	77.9
GCN3D <sup>[14]</sup>	74.7	77.4
DGCNN <sup>[10]</sup>	73.6	78.1
PointCNN <sup>[9]</sup>	75.1	78.5
PCT <sup>[12]</sup>	75.2	78.8
GridPoint	73.1	76.8

构的主流网络,GridPoint对局部特征的利用稍有不足,在精度上落后0.3%~2.3%。文献[16][17]分别对PointNet++和DGCNN进行了轻量化处理,其精度分别下降了1.5%和3%,因此可以认为GridPoint在轻量化的同时依然保持了较好的分类性能。

**2.5 网络复杂度实验**

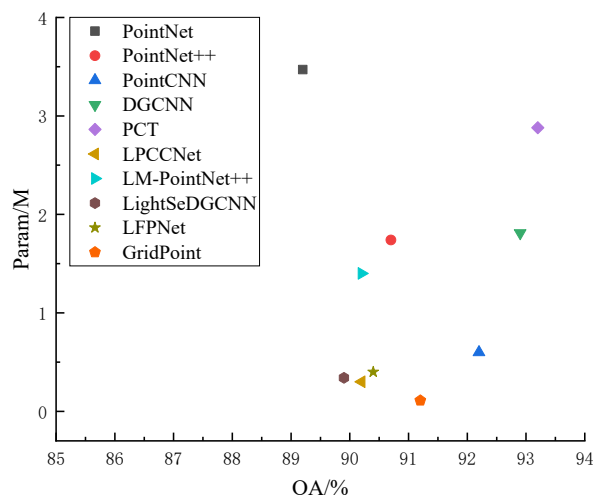
GridPoint作为轻量级网络,虽然在分类精度上稍低,但是其主要优势体现在更精简的网络结构、更少的运算量与更快的速度方面。因此设计网络复杂度实验,选取ModelNet40作为测试数据集,分析GridPoint与其他网络在分类任务中的计算复杂度。除以上对照网络外,还将LPCCNet<sup>[15]</sup>、LM-PointNet++<sup>[16]</sup>、LightSeDGCNN<sup>[17]</sup>以及LFPNet<sup>[25]</sup>4种轻量级网络加入比较。实验结果如表5与图4所示。

**Table 5 Comparison of network complexity**

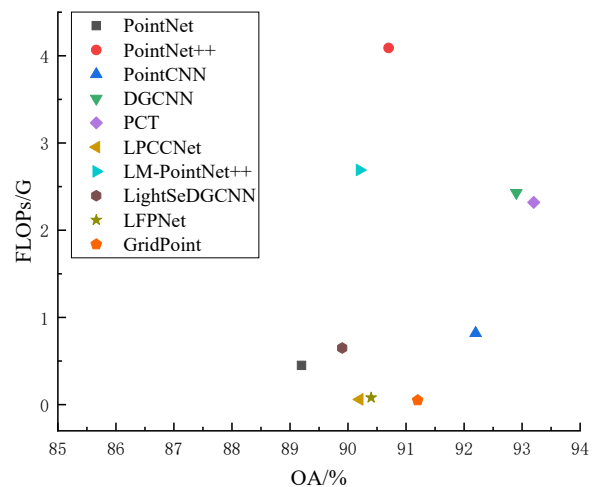
表 5 网络复杂度比较

网络模型	Params/M	FLOPs/G	训练时间/h	测试时间/s	OA/%
PointNet <sup>[7]</sup>	3.47	0.45	4.5	23	89.2
PointNet++ <sup>[8]</sup>	1.74	4.09	6.8	38	91.7
PointCNN <sup>[9]</sup>	0.60	0.82	3.2	28	92.2
DGCNN <sup>[10]</sup>	1.81	2.43	9.8	85	92.9
PCT <sup>[12]</sup>	2.88	2.32	7.4	39	93.2
LPCCNet <sup>[15]</sup>	0.30	0.06	2.8	22	90.2
LM-PointNet++ <sup>[16]</sup>	1.40	2.69	3.4	25	90.2
LightSeDGCNN <sup>[17]</sup>	0.34	0.65	3.2	24	89.9
LFPNet <sup>[25]</sup>	0.40	0.08	2.5	20	90.4
GridPoint	0.11	0.05	2.1	16	91.2

由表5可知,GridPoint的网络参数量仅为0.11 M,浮点运算次数仅为0.05 G,为所有网络中的最小值。主流网络中精度最高的PCT参数量为2.88 M,浮点运算次数2.32 G,远高于GridPoint;参数量较小的PointCNN也达到了0.6 M,浮点运算次数较低的PointNet达到0.45 G。相比之下,GridPoint在网络参数量与浮点运算量方面分别达到81.7%与88.9%以上的降低,轻量化效果显著。在运算速度方面,GridPoint完成200轮次训练仅需2.1 h,一轮测试时间为16 s,均为各网络中的最优值,在实时性方面亦有明显优越性。



(a) Parameter quantity-precision scatter plot  
(a) 参数量-精度散点图



(b) Calculation quantity-precision scatter plot  
(b) 计算量-精度散点图

**Fig. 4 Network complexity-accuracy scatter plot**

图 4 网络复杂度-精度散点图

与其他轻量级网络相比,GridPoint的分类精度最高,领先约1%。其中LM-PointNet++基于PointNet++进行优化,但其参数量与计算量依然较高;LPCCNet和LFPNet的浮点运算次数与GridPoint相当,但参数量较高;LightSeDGCNN未能较好地降低复杂度,精度损失也过多。在

运算速度方面,其他轻量级网络的测试时间均超过 20 s,综合来看,GridPoint 在轻量级网络中同样有着较为明显的优势。

图 4 为各模型网络的复杂度—精度散点图,是对表 5 中参数量与计算量的可视化。图中越靠右的网络分类精度越高,越靠下的网络轻量化程度越高。主流网络大多处于图的右上角,这是由于 PointNet++ 使用多层级的球查询, DGCNN 的边缘卷积使用大量 KNN, PointCNN 构建了额外的 X 变换矩阵, PCT 则引入自注意力机制,它们都能更好地提取局部特征,获得更高精度,但也使得网络过于复杂、参数量大且运行速度缓慢。轻量级网络则大多处于图的左下角,均妥协了一定的分类精度以换取更快的运算速度。GridPoint 始终处于右下角的位置,其中的 Gridding 模块仅进行数值大小的比较,而无需计算欧氏距离,以较低的复杂度完成局部特征提取; SFE 模块则通过表达点云轮廓丰富其原始特征。两个模块的结合使得 GridPoint 获得最佳轻量化表现的同时也有着较为理想的分类性能。

综上所述,GridPoint 虽然妥协了一定的分类精度,但与主流方法差距较小,且在轻量化方面优势显著,有着更少的资源占用、更快的推理速度,满足实际场景中实时性的要求。

## 2.6 消融实验

为验证本文所提 Gridding 与 SFE 两个模块的有效性,设计不同的网络结构和模块内部参数在 ModelNet40 数据集上进行消融实验。表 6 为 GridPoint 中两个主要模块的消融实验结果,其中“√”表示添加了对应模块,“×”表示没有添加模块。实验结果表明,Gridding 模块和 SFE 模块对网络的分类精度均有着显著的提升作用。由于 GridPoint 本身层数较浅,两种模块均未添加时网络的分类性能明显不足;而当两个模块同时使用时,网络的分类精度可以从 71.8% 大幅提升至 91.2%,该实验验证了所提模块的有效性与合理性。

Table 6 Results of ablation experiments on two main modules

表 6 两个主要模块的消融实验结果

Method	Gridding	SFE	OA/%
0	×	×	71.8
1	√	×	87.4
2	×	√	88.3
3	√	√	91.2

在 Gridding 模块内部,栅格边长  $d$  和栅格数量均为重要参数。栅格边长  $d$  越大,空间感受野就越大,相互重叠的部分也越多。然而,体积过大的栅格会导致局部特征减弱,趋向全局特征。此外,栅格数量表示栅格化的分辨率,更高的分辨率意味着更大的计算量,也意味着更小的  $d$  和  $n$ 。例如,将点云划分为 27 个栅格时, $d$  的取值只需大于 0.66 即可。表 7 为 Gridding 模块中栅格边长  $d$  以及栅格数量对分类结果的影响,其中分别设置了 8 分栅格与 27 分栅格,以及相应的栅格边长  $d$  和栅格内采样点数  $n$ 。实验结

果表明,在 27 分栅格情况下,取栅格边长  $d = 1.0$ ,采样点数  $n = 128$  时,网络的分类精度达到最高的 89.7%;在 8 分栅格情况下,取  $d = 1.1, n = 170$  时,网络的分类精度达到最高的 91.2%。通过比较可以看出,8 分栅格在结构更简单的同时分类性能也优于 27 分栅格。同时,不断增大栅格边长  $d$  也并不能持续提升分类精度,在达到最大精度后继续增大  $d$  会导致精度开始下降。

Table 7 Ablation studies about different gridding setting

表 7 不同栅格设置消融实验

栅格数量	$d$	$n$	OA/%
27	0.7	43	88.2
27	1.0	128	89.7
27	1.1	170	89.2
8	1.0	128	89.9
8	1.1	170	91.2
8	1.2	221	90.8
8	1.3	281	88.8

SFE 模块内部使用高次多项式对点云的表面轮廓进行拟合,但不同的多项式编码形式对分类结果产生的影响值得研究。表 8 为 SFE 模块对点云进行不同编码的消融实验结果,其中“√”表示扩展了对应函数,“×”则表示没有扩展; $f_1 \sim f_4$  分别对应式(6)中的平方项、立方项等不同拓展编码。实验结果表明,增加编码中的高次项函数会对分类性能产生正面影响。在只增加  $f_1$ ,即平方项的情况下,SFE 模块能将网络的分类精度从 87.4% 提升至 89.4%;而将多种不同的高次项函数全部加入编码时,网络的分类精度逐步提升,最终达到最大值 91.2%。

Table 8 Different encoding point cloud ablation experiments

表 8 不同编码点云消融实验

方法	$f_1$	$f_2$	$f_3$	$f_4$	总维数	OA/%
0	×	×	×	×	3	87.4
1	√	×	×	×	6	89.4
2	√	√	×	×	9	89.6
3	√	√	√	×	12	90.3
4	√	√	√	√	15	91.2

以上 3 个消融实验验证了 Gridding 与 SFE 模块在网络中的有效性以及内部构造的合理性。

## 2.7 稀疏点云鲁棒性实验

在实际应用场景中采集到的点云往往不够理想,会伴随一定干扰,因此需要对网络的鲁棒性进行验证。对点云进行下采样形成稀疏点云,以检验网络的抗干扰能力。从 ModelNet40 数据集中随机选择 3 种物品,分别为杯子(cup)、台灯(lamp)和椅子(chair)。图 5 展示了它们在不同采样点数下的稀疏点云,其中采样点数分别为 1 024、768、512 和 256。可以看出,随着采样点的减少,点云的分类难度随之上升,物品部分结构也出现残缺与断裂,如台灯与椅子中的连接杆。

鲁棒性实验同样选取 ModelNet40 数据集,在不同采样点数下进行分类实验,同时加入 PointNet、PointNet++ 与 DGCNN 进行比较,实验结果如图 6 所示。可以看出,当点

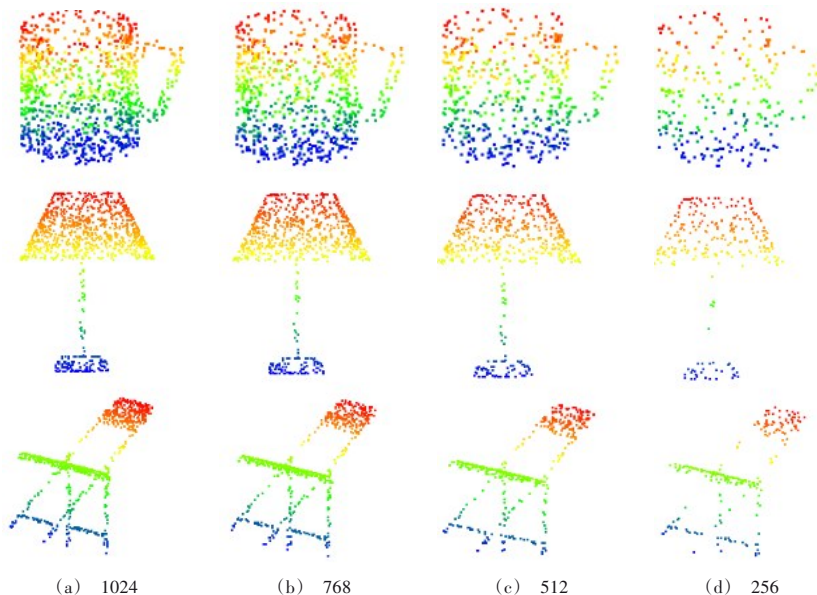


Fig. 5 Sparse point cloud of different point number

图 5 不同点数的稀疏点云

云逐渐稀疏时,所有网络的分类精度均有所下降,其中 DGCNN 由于需要点云中的点构成图结构,在采样点降至 512 后分类精度大幅下降;而 GridPoint 则始终保持着较好的性能,在采样点分别为 768、512、256 时,整体分类精度分别下降至 90.9%、90.4% 与 90.2%,始终高于 PointNet,且与 PointNet++ 的性能差距也一直保持在 0.5% 内。综上所述,GridPoint 在应对稀疏点云时有着较好的鲁棒性。

Point 网络结构的合理性、各模块的有效性以及应对稀疏点云时的鲁棒性。与其他主流算法相比,GridPoint 在较少妥协分类精度的同时显著降低了参数量与计算复杂度,极大减少了对内存和算力的占用,使其能够更加适配移动终端的部署。当然 GridPoint 在分类精度方面仍有提升空间,后续计划在网络中融合注意力机制进一步提高分类性能<sup>[26-27]</sup>,并基于 GridPoint 构建分割网络。

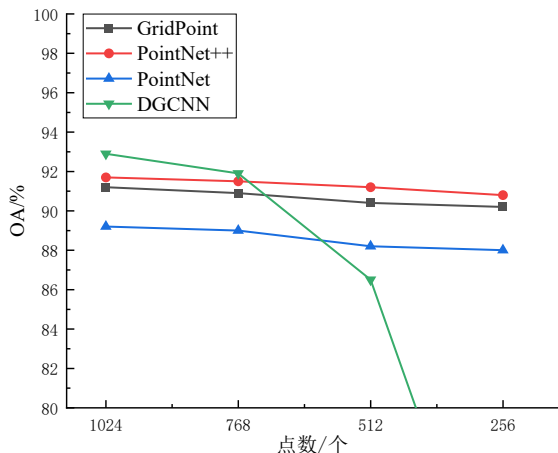


Fig. 6 Robustness result of sparse point cloud

图 6 稀疏点云鲁棒性实验结果

### 3 结语

针对传统点云分类网络复杂度过高的问题,本文提出一种轻量级网络 GridPoint。其利用栅格化 Gridding 与 SFE 两个模块提取并聚合了点云的局部与轮廓特征,在轻量化重构的同时保证了网络的分类精度,相较原始网络 PointNet 性能有一定提升。此外,还通过相关实验验证了 Grid-

### 参考文献:

- [1] JIANG S, WANG S, YI Z, et al. Autonomous navigation system of greenhouse mobile robot based on 3D lidar and 2D lidar SLAM[J]. *Frontiers in Plant Science*, 2022, 13: 815218.
- [2] YURTSEVER E, LAMBERT J, CARBALLO A, et al. A survey of autonomous driving: common practices and emerging technologies[J]. *IEEE Access*, 2020, 8: 58443-58469.
- [3] LI Y, MA J, ZHANG Y. Image retrieval from remote sensing big data: a survey[J]. *Information Fusion*, 2021, 67: 94-115.
- [4] GONG J Y, LOU Y J, LIU F Q, et al. Scene point cloud understanding and reconstruction technologies in 3D space[J]. *Journal of Image and Graphics*, 2023, 28(6): 1741-1766.  
龚靖瑜,楼雨京,柳奉奇,等. 三维场景点云理解与重建技术[J]. *中国图象图形学报*, 2023, 28(6): 1741-1766.
- [5] SU H, MAJI S, KALOGERAKIS E, et al. Multi-view convolutional neural networks for 3d shape recognition[C]// *Proceedings of the IEEE International Conference on Computer Vision*, 2015: 945-953.
- [6] 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*, 2015: 922-928.
- [7] QI C R, SU H, MO K, et al. Pointnet: deep learning on point sets for 3d classification and segmentation[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 652-660.
- [8] QI C R, YI L, SU H, et al. Pointnet++: deep hierarchical feature learning

- on point sets in a metric space [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017: 5105–5114.
- [9] LI Y, BU R, SUN M, et al. PointCNN: convolution on x-transformed points [DB/OL]. <https://arxiv.org/abs/1801.07791>.
- [10] WANG C, SAMARI B, SIDDIQI K. Local spectral graph convolution for point set feature learning [C]//Proceedings of the European Conference on Computer Vision, 2018: 52–66.
- [11] THOMAS H, QI C R, DESCHAUD J E, et al. Kpconv: flexible and deformable convolution for point clouds [C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 6411–6420.
- [12] GUO M H, CAI J X, LIU Z N, et al. PCT: point cloud transformer [J]. *Computational Visual Media*, 2021, 7: 187–199.
- [13] YAN X, ZHENG C, LI Z, et al. PointASNL: robust point clouds processing using nonlocal neural networks with adaptive sampling [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 5589–5598.
- [14] LAN H, CHEN H, ZHANG P F. Point cloud classification and segmentation model based on graph convolution and 3D direction convolution [J]. *Computer Engineering and Applications*, 2023, 59(8): 182–191.  
兰红, 陈浩, 张蒲芬. 集图卷积和三维方向卷积的点云分类分割模型 [J]. *计算机工程与应用*, 2023, 59(8): 182–191.
- [15] LI M, HU Y, ZHAO N, et al. LPCCNet: a lightweight network for point cloud classification [J]. *IEEE Geoscience and Remote Sensing Letters*, 2019, 16(6): 962–966.
- [16] CUI Y, AN Y, SUN W, et al. Lightweight attention module for deep learning on classification and segmentation of 3-D point clouds [J]. *IEEE Transactions on Instrumentation and Measurement*, 2020, 70: 1–12.
- [17] WANG C X, SHI T B, TANG L L, et al. Design of neural network model for lightweight 3D point cloud classification [J]. *JNI*, 2020, 5: 122–128.
- [18] JIANG M, WU Y, ZHAO T, et al. Pointsift: a sift-like network module for 3d point cloud semantic segmentation [DB/OL]. <https://arxiv.org/abs/1807.00652>.
- [19] MINEO C, PIERCE S G, SUMMAN R. Novel algorithms for 3D surface point cloud boundary detection and edge reconstruction [J]. *Journal of Computational Design and Engineering*, 2019, 6(1): 81–91.
- [20] BRONSTEIN A M, BRONSTEIN M M, KIMMEL R. Numerical geometry of non-rigid shapes [M]. Berlin: Springer Science & Business Media, 2008.
- [21] JOSEPH-RIVLIN M, ZVIRIN A, KIMMEL R. Momen (e) t: flavor the moments in learning to classify shapes [C]// Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision Workshop, 2019: 4085–4094.
- [22] WU Z, SONG S, KHOSLA A, et al. 3d shapenets: a deep representation for volumetric shapes [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1912–1920.
- [23] 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 (ToG)*, 2016, 35(6): 1–12.
- [24] UY M A, PHAM Q H, HUA B S, et al. Revisiting point cloud classification: a new benchmark dataset and classification model on real-world data [C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 1588–1597.
- [25] YU Q, YANG H, GAO Y, et al. LFPNet: lightweight network on real point sets for fruit classification and segmentation [J]. *Computers and Electronics in Agriculture*, 2022, 194: 106691.
- [26] ZHAO J. Multi scale feature point cloud completion method based on fusion attention mechanism [D]. Fuxin: Liaoning University of Engineering and Technology, 2023.  
赵静. 融合注意力机制的多尺度特征点云补全方法 [D]. 阜新: 辽宁工程技术大学, 2023.
- [27] LU B, SUN Y, YANG Z Y. A three-dimensional object detection method based on the self attention mechanism of the original point cloud grid [J]. *Journal of Communications*, 2023, 44(10): 72–84.  
鲁斌, 孙洋, 杨振宇. 基于原始点云网格自注意力机制的三维目标检测方法 [J]. *通信学报*, 2023, 44(10): 72–84.

(责任编辑: 尹晨茹)