

中国激光
Chinese Journal of Lasers
ISSN 0258-7025, CN 31-1339/TN

《中国激光》网络首发论文

题目：基于深度学习的三维点云处理方法研究进展
作者：吴一全，陈慧娴，张耀
收稿日期：2023-06-19
网络首发日期：2023-08-18
引用格式：吴一全，陈慧娴，张耀. 基于深度学习的三维点云处理方法研究进展[J/OL]. 中国激光. <https://link.cnki.net/urlid/31.1339.TN.20230817.1813.016>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于深度学习的三维点云处理方法研究进展

吴一全*, 陈慧娴, 张耀

南京航空航天大学电子信息工程学院, 江苏 南京 211106

摘要 随着传感器技术不断发展, 三维点云被广泛应用于自动驾驶、机器人、遥感、文物修复、增强现实、虚拟现实等领域的视觉任务中。然而直接应用收集到的海量原始点云数据的效果不佳, 因此基于深度学习的点云处理方法受到越来越多的关注和研究。该文综述了近六年来基于深度学习的三维点云处理方法的研究进展。首先给出了三维点云的基本概念和获取方式, 简述了4种点云处理任务; 然后针对点云去噪和滤波、点云压缩、点云超分辨率以及点云修复、补全和重建任务, 重点阐述了相应的深度学习方法的原理并分析了优缺点; 随后介绍了22种点云数据集和4类评价指标, 给出了性能对比结果; 最后探讨了点云处理方法目前存在的问题, 对未来的研究趋势进行了展望。

关键词 三维点云处理; 深度学习; 点云去噪; 点云压缩; 点云修复

中图分类号 TP391.41 **文献标志码** A

OCIS codes 280.3640; 100.6890; 100.4996

1 引言

随着三维激光雷达等三维采集设备种类的扩展以及数量的增加, 三维点云数据更容易被获取, 精度也得到提升, 逐渐成为三维数据最重要的一种表达方式。三维点云“处理”任务主要包括点云去噪和滤波、点云压缩、点云超分辨率以及点云修复、补全和重建(点云目标检测与跟踪、分类分割、配准和匹配以及拼接归为三维点云“分析”任务, 因篇幅限制, 已另文综述), 在自动驾驶^[1]、机器人^[2]、遥感测绘^[3]、文物修复^[4]、农业生产^[5]、医疗^[6]、电力^[7]、建筑^[8]、铁路^[9]、增强现实(Augmented Reality, AR)^[10]、虚拟现实(Virtual Reality, VR)^[11]等领域运用较多。因此, 研究三维点云处理方法在实际应用中具有重要意义。

目前三维点云处理方法分为传统方法和学习方法两大类。传统方法包括几何分析处理、特征描述子提取等, 前者主要通过估计点云的法向量、曲率、密度等几何信息来分析点云数据, 后者通过传统特征描述子获取对应区域的点云几何与属性信息。采用传统方法处理点云的准确度较高, 但是针对大幅度旋转等复杂点云场景的效果不够理想, 且涉及的函数计算极为繁琐。学习方法可分为经典的机器学习方法和深度学习方法。机器学习方法能够处理三维数据, 学习有效特征信息, 但是由于三维数据量巨大, 训练样本的数量也极大, 这会增加经典机器学习方法训练的难度。同时, 机器学习需要专家识别特征, 特征提取的准确度直接影响机器学习方法的性能, 而三维数据的标记不仅多, 标注难度也远大于二维图像^[12]。深度学习是机器学习的一种, 能够训练和计算大型数据, 自主学习输入信息中的潜在空间特征和高

基金项目: 国家自然科学基金(61573183)

通信作者: *E-mail: nuaaimage@163.com

级规律,适用于处理海量的点云数据。深度学习方法采用端到端架构,输入数据信息,经所设计的网络结构的处理分析,输出需要的结果。虽然深度学习方法耗费了大量的时间训练样本以学习参数信息,但是相较于机器学习方法的分块解决方式,深度学习方法的测试时间更少,可以一次性得到更精准的预测结果^[13]。因此,本文讨论基于深度学习的三维点云处理方法。深度学习的基础是神经网络,目前,常用的神经网络模型有以下 4 种:卷积神经网络(Convolutional Neural Network, CNN)、图卷积神经网络(Graph Convolutional Network, GCN)、生成对抗网络(Generative Adversarial Network, GAN)和自编码器(Auto-Encoder, AE)。鉴于三维点云的内部结构不规则,稀疏且不均匀,直接采用二维深度学习方法处理点云数据十分困难,如何高效地实现三维点云处理是近年来研究人员的工作重点。

当前相关综述文献情况如下:文献[14]回顾了近五年提出的点云去噪算法并汇总了去噪后点云的质量评估方法,但点云去噪的深度学习方法不够全面;文献[15]综述了国内外传统及深度学习的点云压缩方法,总结了 3 种开源数据集并进行了实验性能分析,但是深度学习方法的篇幅较少;文献[16]汇总了基于传统和基于机器学习两类超分辨率算法,但是仅提及了 4 种基于深度学习的点云超分辨率方法,且缺乏系统性对比;文献[17]详细分析了三维点云修复的相关模型和关键技术,但缺少各种方法之间的对比分析;上述所有文献仅对一种点云处理任务进行了探讨,但比较片面,本文在前人的基础上,系统性地整理并总结了目前国内外常用的三维点云处理的深度学习方法,对上述综述文献进行了扩展、更新和深化。

本文回顾了近六年来国内外基于深度学习的三维点云处理方法。论文结构如下:第一部分给出了三维点云的概念和获取方式,概述了 4 种点云处理任务。第二部分归纳了基于深度学习的三维点云处理方法,对点云去噪和滤波、点云压缩、点云超分辨率以及点云修复、补全和重建这 4 种任务的深度学习方法分类,并对优缺点进行了对比分析。第三部分整理了目前常用的 22 种点云数据集和性能评价指标,并给出了性能对比结果。最后针对三维点云处理方法存在的问题进行了展望,期望能为点云处理的研究人员提供灵感和思路。

本文的贡献有以下 6 点:(1) 整理归纳了目前最常用的 4 种点云处理任务,简述了 4 种任务的主要目标;(2) 从卷积网络、上采样、滤波、梯度和其他方法 5 个方面对基于深度学习的点云去噪方法进行了汇总和分析,对比了各种方法的优劣;(3) 总结了点云有损压缩和无损压缩的深度学习方法,并进行了对比分析;(4) 对点云超分辨率任务进行了全面的总结,归纳了基于卷积神经网络、图卷积神经网络、生成对抗网络以及其他结构的深度学习方法;(5) 整理了点云修复、补全和重建的最新方法;(6) 梳理了 22 种点云处理任务常用的数据集和网址,并给出了评价指标和性能对比结果。

2 三维点云获取及 4 种处理任务

2.1 三维点云及获取

点云是在同一坐标系下的目标表面特性空间采样点的集合，采样点包含三维坐标、大小等几何信息以及物体色彩、纹理特征等属性信息。根据组成特点可以将点云划分为有序点云和无序点云两种，有序点云一般是由深度图还原的按顺序排列的点云，而无序点云的排列没有任何顺序。点云数据是一组具有位置特征属性的向量集合，是由物体表面一系列空间采样点构成的、描述物体形体的一种边界表示形式。点云数据结构简单，不仅可以表示物体的几何位置和表面细节信息，还能使用传感器融合技术丰富点云信息，例如相机可以获取彩色图像来表示颜色信息(RGB)，激光雷达通过收集的回波强度表示反射强度(Intensity)。

三维点云获取是指采用现代化测量手段采集物体表面的实体信息并转化为点的坐标数据的数字化过程。目前，常用的点云获取方式有激光扫描仪(Laser Scanner)、深度相机(Depth Camera)和多传感器融合 3 种。激光扫描仪根据搭载平台的不同可分为地面激光扫描仪(Terrestrial Laser Scanner, TLS)、车载激光扫描仪(Mobile Laser Scanner, MLS)、机载激光扫描仪(Airborne Laser Scanner, ALS)、星载激光扫描仪(Satellite Laser Scanner, SLS)等。激光扫描仪通过记录接收到的被测物体表面的点的信息，如三维坐标、反射率等，来确定被测物体的空间位置，得到点云数据并重建其三维模型。深度相机利用近红外激光器将具有结构特征的光线投影到物体上，通过红外摄像头采集得到被测目标的深度信息。根据深度相机的工作原理可将其分为 3 种：结构光、RGB 双目和飞行时间(Time of Flight, TOF)，3 种深度相机的优劣对比如表 1 所示。多传感器融合技术是将摄影测量(深度相机)和激光测量(激光雷达)原理结合得到的新型技术，采集到的点云数据包括几何坐标、色彩和反射强度等信息。

表 1 3 种深度相机性能参数对比

Table 1 Comparison of performance parameters of three depth cameras

Camera type	Structured light		Binocular vision		Time of flight
Principle	Project	special structural patterns onto the object	Calculate	depth information from two RGB images	Direct measurement based on the time of flight of light
Accuracy	Achieve high precision of 0.01 mm-1 mm in short distance		Up to millimeter precision in short distance		Up to centimeter-level accuracy
Range	Within 10 m		Within 2 m (baseline 10 mm)		Within 100 m
Resolution	Up to 1080×720		Up to 2K resolution		Less than 640×480
Frame rate	30 fps		From high to low		Higher, up to hundreds of fps
Influencing factors	Reflections		Illumination changes and object textures, unavailable at night		Illumination changes and object textures, multiple reflections

Software complexity	Medium	High	Low
Representative	Kinect v1, Pickit, PrimeSense	PointGrey Bumblebee, ZED	Kinect v2, Terabee, Basler

2.2 点云处理任务

目前，常用的点云处理任务有 4 种：点云去噪和滤波、点云压缩、点云超分辨率、点云修复、补全和重建。

(1) 点云去噪和滤波：点云在采集过程中难免会受到传感器及环境的干扰，导致噪声点和一些远离目标点云的离群点的出现，降低了点云的质量，进而影响点云处理任务的性能，因此需要对点云进行去噪处理。同时，滤波操作不仅可以去除噪声点和离群点，还能根据具体需求平滑不规则数据并获取有效的点。

(2) 点云压缩：本质是利用点与点之间的空间与时间相关性来减少冗余。点云压缩主要是通过编码解码来实现的，根据压缩后的点云是否存在信息丢失可划分为点云有损压缩和点云无损压缩两类。

(3) 点云超分辨率：以低分辨率(Low-resolution, LR)点云为输入，结合给定的先验信息生成具有丰富细节信息的高分辨率(High-resolution, HR)点云。

(4) 点云修复、补全和重建：根据采集到的点云数据填补缺失的点云区域，以生成完整的点云形状，提高点云质量从而实现点云的复原和重构。

3 基于深度学习的三维点云处理方法

硬件配置图形处理器(Graphics Processing Unit, GPU)的迅速发展大大提高了计算机处理复杂计算的速度，解决了深度学习计算量大且复杂的难题，使其在计算机视觉领域中的应用得到进一步推广，处理大规模的三维点云数据也有较好的效果。下面将从点云去噪和滤波、点云压缩、点云超分辨率以及点云修复、补全和重建 4 个点云处理任务总结其深度学习方法。图 1 展示了近六年来 4 种点云处理任务最常用的深度学习的发展路线。

3.1 点云去噪和滤波

传统的点云去噪方法基于信号处理或者几何学理论，结合特征描述子估计点云的边界点和法线位置，调整点的坐标信息来去除噪声，但对噪声、点云采样和曲率变化等不够鲁棒，且过于依赖几何先验。基于点云的神经网络的出现与发展，使得基于深度学习的点云去噪方

法得到了广泛关注。点云去噪的目的是从噪声扰动输入中恢复干净的点云。根据选取的去噪技术的不同，点云去噪的深度学习方法可以分为以下 5 种：基于卷积网络、基于上采样、基于滤波、基于梯度和其他方法。

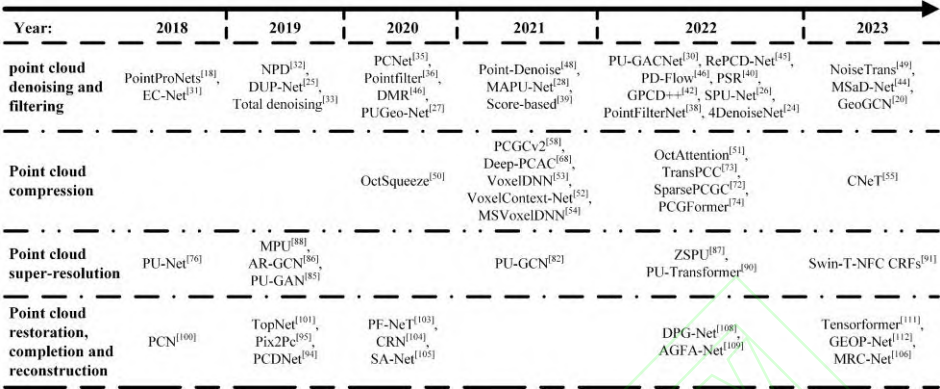


图1 4种点云处理任务的深度学习方法的发展路线

Fig. 1 Development route of deep learning methods for four point cloud processing tasks

3.1.1 基于卷积网络的方法

基于卷积网络的去噪方法是指利用卷积神经网络的结构或者对网络加以改进达到去除噪声点云的目的。由于几何先验会导致点云特征和细节的丢失，PointProNet^[18]采用一个完全可微的 CNN 架构，通过可微分点投影层，将无序的三维点云转换为定期采样的二维高度图，再利用高度图去噪网络(HDN)对高度图卷积去噪。GPDNet^[19]和 GeoGCN^[20]借助图卷积网络去噪。为了在去除异常值的同时对剩余点进行去噪，GPDNet^[19]引入了图卷积神经网络处理点云排列不变性问题，利用残差网络估计输入的噪声成分，然后将噪声投影到三维空间并从噪声点云中去除估计的噪声。GeoGCN^[20]融合了全局几何特征信息，设置两种表面法线：真实法线和虚拟法线。前者更新顶点坐标以保留噪声点云的局部细节，后者作为损失函数来约束噪声点云的全局形状，通过几何双域图卷积网络充分学习空间域和普通域的特征细节，提高了网络的噪声鲁棒性。Lu 等人^[21]提出了一种基于 CNN 的特征保留法线估计方法，可以自动估计输入的嘈杂点云的法线，然后根据估计的法线更新点位置，通过多次迭代法线估计和更新位置得到更好的去噪结果。点云去噪在医学、雷达处理等领域也有重要应用。文献[22]构建了类似 U-Net 架构的神经网络，将人的单个前视图深度图像作为输入，为前视图和后视图训练神经网络，并输出去噪且完整的前视图深度图像或预测的后视图深度图像。由于文献[23]仅考虑了雨雾两种恶劣天气，4DenoiseNet^[24]通过对连续点云进行 k 最近邻搜索卷积来捕获空间和时间信息，采用运动引导注意力机制融合时间特征和空间特征，在雨、雾、轻雪、中雪和大雪场景都实现了良好的去噪性能，计算成本更低。

点云的无序性和不规则性使得在三维点云领域使用 CNN 具有挑战性。目前基于卷积网络的去噪方法还是会在数据预处理阶段将点云转化为体素或者网格,但是会损失部分点云信息。大多数方法直接利用卷积神经网络的基本结构进行训练,通过捕获特征差异剔除噪声点。

3.1.2 基于上采样的方法

实际扫描得到的点云通常是稀疏、嘈杂且不均匀的,处理起来比较麻烦。为此,研究人员提出了点云上采样技术进行修正,通过各种上采样网络生成一个密集、完整且均匀分布的点云,在上采样的过程中去噪。DUP-Net^[25]通过统计异常值去除层和上采样网络缓解点云的对抗效应,利用去噪层的不可微分性统计并去除离群点,以增强对对抗样本的防御,避免网络做出错误判断。考虑到在实际扫描时,原始点的表面分布会被破坏,为了保持原点分布,SPU-Net^[26]通过点特征提取和点特征扩展的重构框架促进上采样,采用自投影优化方法将含噪声的点沿投影方向交互式地投影到底层物体表面,优化生成的密集点云分布,进一步提高点云上采样质量。鉴于先前的深度学习方法只生成点,难以估计原始点和生成点的法线,PUGeo-Net^[27]构建了基于离散微分几何的神经网络,以纯几何的方式进行采样,通过学习每个点及其法线方向的局部参数来生成具有准确法线的密集点云,网络可以处理嘈杂和非均匀分布的输入。上述去噪方法适用于特定的上采样因子,模型复杂且训练时间长,MAPU-Net^[28]可以处理多种类型的上采样因子,能够在一次性训练后对具有灵活因子的稀疏点云进行上采样生成密集点云,通过投影到切平面上的相邻点的仿射组合来生成新样本,产生更丰富的几何细节。考虑到现有的上采样方法忽略了点云块之间的关联性,PC²-PU^[29]引入点云块相关单元和位置校正单元,利用相邻点云块的信息和点之间的相对位置关系,通过距离感知编码器动态调整生成的点以减轻噪声点和离群值的影响。为了获取点云全局和局部特征,PU-GACNet^[30]利用图注意力卷积模块分配不同的注意力权重,动态组合全局和局部特征,通过边缘感知节点缓存模块对点特征进行平滑地上采样,以保留局部几何细节和边缘。

点云上采样技术的效率和有效性取决于所用的特征提取器和特征扩展器,因此,基于上采样的去噪方法基本是对这两个部分加以修改。在上采样技术中融入网络的方法可以处理多特征等复杂结构,但主要集中在监督学习上,使用匹配的稀疏密集点云对进行学习,容易忽略一些局部特征,在合成数据集的去噪效果较好,但在真实数据集的效果不佳。

3.1.3 基于滤波的方法

点云滤波需要在保留几何特征的同时去除噪声或异常值,这也是点云滤波的技术难点。

为了整合稀疏嘈杂的点云, EC-Net^[31]通过边缘感知联合损失函数计算从输出点到 3D 网格和边缘的距离, 使输出点均匀分布在表面。针对三维点云体素化导致的量化伪影问题, 神经投影去噪方法(NPD)^[32]直接使用原始点云, 借助网络预估参考平面, 将噪声点投影到参考平面上来平滑 3D 噪声点云。文献[33]提出了一种无监督方法来过滤噪声点云, 引入空间先验项, 通过强加先验可以直接在噪声数据上进行训练, 不需要获取噪声对辅助训练。文献[35]基于 PCPNet^[34]网络架构设计了两阶段的去噪网络 PointCleanNet(PCNet), 用于去除无序点云的噪声和异常值, 首先利用 PCPNet 网络检测并去除异常值, 然后估计剩余点的局部属性并进行去噪, 具体的网络结构如图 2 所示。为了在保留点云特征的同时保持较好的泛化能力, Pointfilter^[36]利用端到端神经网络去除噪声, 并保留其清晰特征, 实现了自动稳健地滤波。鉴于目前的点云滤波方法不能在保持清晰特征的同时输出均匀分布的点, 文献[37]设计了一种综合考虑点分布和特征保留的点云滤波方法, 主要思想是在目标函数中加入斥力项和数据项, 斥力项保证点均匀分布, 而数据项在保持几何特征的情况下逼近噪声表面, 在不同噪声和密度条件下的滤波效果很好。考虑到先前的方法需要手动定义滤波系数, PointFilterNet(PFN)^[38]将深度学习与滤波结合, 通过异常值识别器辨识异常值并分配小系数来减轻异常值干扰, 再经降噪器逐步降噪, PFN 利用网络学习的系数生成滤波去噪点, 点云去噪的鲁棒性和有效性得到改善。

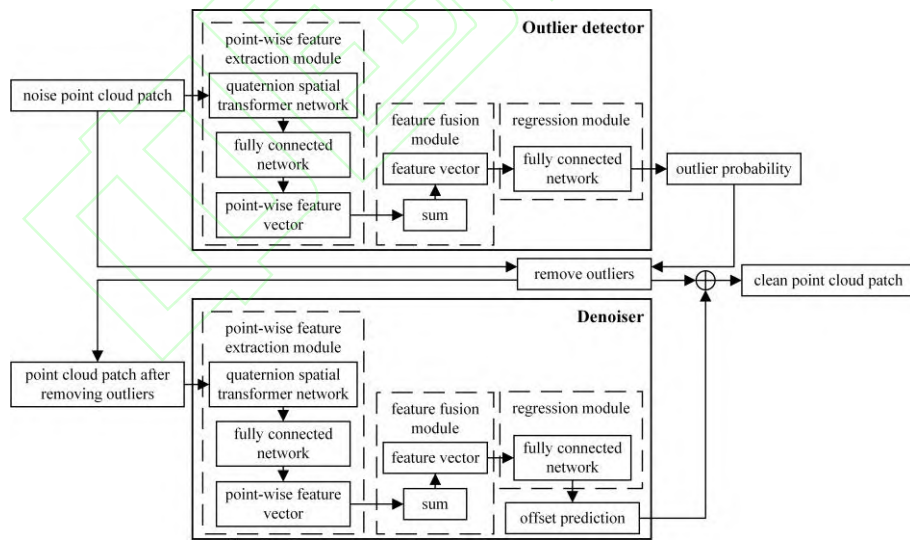


图2 PointCleanNet网络框架

Fig. 2 PointCleanNet Framework

基于滤波的去噪方法主要通过构建神经网络结构保留特征并去噪, 但神经网络方法通常基于局部信息, 采用法线估计进行滤波是一种可以尝试并优化的深度学习方法。

3.1.4 基于梯度的方法

基于梯度的去噪方法的核心是利用神经网络估计点云梯度场,用对数概率密度函数的梯度来表示点云,通过梯度上升迭代更新每个点的位置进行去噪。基于分数的点云去噪方法^[39]设计了一种保留细节的神经网络结构来估计输入噪声点云在对数概率下的分布分数,即对数概率函数的梯度,通过估计梯度迭代更新每个点的位置,引导梯度上升来去除异常值。不同于分数方法^[39]估计每个输入点周围的局部梯度场,文献[40]学习支持整个点云分布的全局梯度场,提出了一种新的恢复点集重采样方法 PSR,通过梯度场模拟退化点云的分布使梯度连续,对输入点云执行基于梯度的马尔可夫链蒙特卡罗(Markov Chain Monte Carlo, MCMC),在保留信息细节的同时能降低噪声和离群值的影响。为了避免文献[39]中分数梯度的波动,文献[41]在其基础上开发了动量梯度上升方法,利用分数方法中的迭代信息确定点的轨迹,提高了解的稳定性,减少了推理时间,进而提高了梯度去噪方法的性能和效率。上述 3 种基于梯度的方法未考虑点之间的关联性,文献[42]提出了 GPCD++框架,骨干网沿用分数方法^[39]或点集重采样方法^[40]来模拟三维梯度场,然后采用迭代梯度上升法进行去噪,在此基础上使用 UniNet 超轻量网络模拟点之间的相互作用,以实现局部均匀分布。

基于梯度的去噪方法取得了一定的成果,但目前研究还比较少,根据噪声点云的分布特性构建梯度场的思路是可行的,梯度去噪方法的研究空间很宽广。

3.1.5 其他方法

考虑到三维点云中点的特征通道受噪声影响的程度不同,文献[43]和 MSaD-Net^[44]加入了通道注意力模块。文献[43]设计了一种特征提取-压缩网络 ESANet,引入了注意力模块,根据各个通道受噪声扰动程度的不同来分配权值,通过提取并学习关键的特征通道进一步提升网络的去噪性能。MSaD-Net^[44]则引入通道注意力模块来获取不同通道之间的特征关系,通过局部信息嵌入模块快速选取重要信息点并建立局部邻域信息,再基于 Transformer 结构建立全局信息的依赖关系来优化特征信息,运行速度大幅提升。点云进行去噪处理时,噪声去除和几何细节保留之间的权衡也为点云去噪方法提供了新角度。Chen 等人^[45]提出了一种特征感知循环点云去噪网络 RePCD-Net,通过基于 BRNN 的多尺度特征聚合模块在不同的去噪阶段提取自适应特征,引入循环传播层使用深层特征,以恢复平滑的精细几何特征。为了保留点云的细粒度信息,DMR^[46]和文献[47]借助流形框架去除噪声点云。DMR^[46]采用自适应可微池化算子缩小重建底层流形的潜在空间,并对每个流形块采样生成干净点云。Mao 等人^[47]则设计了简单的 PD 流框架,将噪声点云视为三维形状和噪声联合分布的样本,利用规范化流学习噪声点集的分布,将噪声扰动和干净点转换为不纠缠的代码空间,再通过噪声

解缠技术去噪。Point-Denoise^[48]基于无监督思想，采用孤立森林和椭圆包络两种单类别分类技术来检测异常值并增强原始 3D 点云，但是该方法的时间复杂度较高。NoiseTrans^[49]尝试使用 Transformer 解决点云去噪任务，结合具有自注意机制的点嵌入模块学习多尺度点云的局部特征，引入可学习的稀疏编码感知点云结构关系，对原始点云的去噪性能大幅提升。

ESANet^[43]和 MSaD-Net^[44]引入了通道注意力模块，RePCD-Net^[45]借助了循环神经网络，DMR^[46]和 PD 流^[47]结合了规范化流和解缠技术，Point-Denoise^[48]引入了无监督技术，NoiseTrans^[49]借鉴了 Transformer 思想，5 类方法都实现了点云精细特征的有效提取和保留，同时，也取得了出色的点云去噪结果。

3.1.6 点云去噪和滤波方法小结

点云固有的无序性给点云去噪和滤波任务带来了极大的挑战。基于卷积神经网络的方法自主学习能力很强；基于上采样的方法能将杂乱点云转变为结构化点云，但是目前主要集中在监督学习上，即使用匹配的稀疏点云对进行学习；基于滤波的方法可以在去噪的同时保留特征信息；基于梯度的方法新颖且对异常值的鲁棒性很强；应用注意力模块、循环神经网络、流形框架和 Transformer 结构去噪能够有效学习点云细节信息。因此使用深度学习方法去除噪声点云有很大的优越性。基于深度学习的点云去噪和滤波方法的对比如表 2 所示。

表 2 基于深度学习的点云去噪和滤波任务主要方法对比

Table 2 Comparison of main methods for point cloud denoising and filtering tasks based on deep learning

Type	Method	Specific structure	Contribution	Limitation
CNN-based	[18]	Fully differentiable CNN	Height map denoising network	Poor denoising effect on larger holes
	[19]	GCN	Robust to high levels of noise	Neighborhood size affects performance
	[20]	Geometric Dual Domain Graph Convolutional Networks	Real and virtual normals are defined	Longer training time
	[21]	Feature preserving normal estimation	Automatically estimate normals and update point locations	Unsuitable for severe noise and large outliers
Upsampling-based	[25]	Denoiser and upsampler combined	Effectively resist attacks from other point cloud datasets	Unsuitable for defending against black box attacks
	[27]	Networks based on discrete differential geometry	Preserve features and geometric details	Incomplete datasets are not considered
	[29]	Patch correlation unit and position correction unit	Considered noise and outliers in practical applications	The patch selection strategy will affect the stability of the algorithm
	[30]	Graph attention convolution and edge-aware node caching	Fine-grained edge detail is preserved with high quality	GAC modules increase computational complexity
Filter-based	[31]	Edge-aware integrated network	Suitable for dense point clouds with structure-invariant scale	Training time is long
	[32]	Projection denoising method	Direct point cloud denoising	Need enough training samples

		based on neural network	using deep learning techniques	
	[37]	Add repulsion term and data term to the objective function	Capable of handling fine-scale features and sharp features	Depend on the quality of the input normals
	[38]	Outlier recognizer and denoiser	Identify and remove points that are far from the surface	Runtime can also be optimized
Gradient-based	[39]	Score estimation network	More robust to outliers	The gradient is discontinuous
	[41]	Momentum gradient ascent	The gradient field is continuous	Need to construct an effective global gradient field
	[42]	GPCD++ network framework	Lightweight network UniNet	Cannot handle large pores
Other methods	[43]	Channel attention module	Stitching local features of point clouds at multiple scales	The capture of neighborhood feature information is biased
	[44]	Hybrid self-attention network	Enhance local information through Transformer	Longer training time
	[48]	Unsupervised machine learning	Detect outliers by isolation forests and elliptical envelopes	Time complexity needs to be improved
	[49]	Transformer-based	Extract multi-scale local features	High computational complexity

3.2 点云压缩

考虑到直接存储和传输海量的点云数据会占用大量的内存和带宽资源，因此需要进行点云压缩以实现低失真、高效率的传输和存储。目前点云压缩的深度学习方法的研究主要集中在点云信息的还原质量，根据信息还原质量的高低可以分为无损压缩和有损压缩两大类。

3.2.1 点云无损压缩的方法

点云无损压缩方法利用深度学习的网络结构或点云的属性信息估计每个点的概率分布，辅助熵模型进行编码压缩，可分为八叉树、混合表示和颜色属性 3 类。

(1) 基于八叉树的方法

考虑到稀疏的雷达点云数据缺少内部结构，OctSqueeze 模型^[50]将点云编码为八叉树，创造性地利用神经网络训练八叉树结构的熵模型，结合上下文信息预估节点符号并进行压缩编码。文献[51]在文献[50]的基础上，引入兄弟节点的高分辨率特征，借助注意力机制强调上下文中的重要节点，使用掩码操作减少编码时间。

基于八叉树的方法减小了点云传输消耗的带宽，但是文献[50]仅利用父节点信息构建上下文，忽略了相邻节点间的信息。而文献[51]通过条件熵模型和注意力机制实现了点云压缩性能和编码时间的平衡。

(2) 基于混合表示的方法

虽然基于八叉树的方法实现了点云结构化处理,但是构建八叉树结构会引入失真,一些研究人员在八叉树结构中引入体素,尽可能地减少点云空间信息的丢失。**VoxelContext-Net**^[52]首先在三维点云中构建八叉树结构,通过体素化方法获取空间邻域信息,再利用体素上下文信息训练深度熵模型,静态和动态点云压缩的性能都得到了改善。为了缓解点云的稀疏性带来的干扰, **VoxelDNN**^[53]自适应地将点云划分为多个体素块,利用三维卷积网络按顺序预测体素块占用的概率,该方法融合了自回归模型的思想,点云无损压缩率得到提升,但是顺序概率预测会导致计算成本很高。考虑到自回归模型的计算成本较高, Nguyen 等人^[54]提出了 **MSVoxelDNN** 方法,采用多尺度上下文模型代替 **VoxelDNN**^[53]中的自回归模型,以低分辨率尺度辅助预估高分辨率尺度的体素占用概率,同时采用并行预测加快了编解码时间。

基于混合表示的点云压缩方法极大地保留了原始点云的细节信息,点云无损压缩率有效降低,但是体素化数据增加的计算成本还有待改进。

(3) 基于颜色属性的方法

颜色是点云数据的重要属性之一,点云在经过色彩空间变换后仍旧存在色彩相关性,部分研究人员尝试利用颜色属性辅助进行点云无损压缩。**CNeT**^[55]是一种基于学习条件概率模型的静态点云无损压缩方法,将 RGB 转换为亮度-色度色彩空间,并对变换后的颜色特征进行顺序编码并预估概率分布,再通过基于稀疏张量的网络学习预估的概率分布,从而生成精确的自回归上下文模型。文献[56]提出了一种轻量化的点云无损压缩方法,通过跨尺度、跨组、跨颜色相关性预测概率分布,使用稀疏卷积网络学习邻域相关性,对 YCoCg/RGB 等多通道属性进行交叉颜色预测,消除了跨色冗余,平衡了编码效率和模型复杂度。

上述两种基于颜色属性的点云无损压缩方法充分利用了点云的颜色信息,实现了高效压缩,提供了点云压缩研究的可尝试的新方向。

小结: 基于深度学习的点云无损压缩方法提高了压缩率,利用神经网络提升了点云概率预测的准确性,降低了编解码的时间成本,但是压缩性能有限。考虑到在编码和解码过程中存在信息丢失,该方法并不能做到完全的无损压缩,但可以使解码后的点云具有和原始点云相同的点数量以及点之间的关联信息,基本保留了原始数据的特征,实现点云“无损压缩”。

3.2.2 点云有损压缩的方法

点云有损压缩方法利用神经网络学习点之间的关联特征,通过特征重建点云,弥补了量化导致的点云信息丢失。有损压缩方法可以细分为八叉树、体素、自编码器和其他方法 4 种。

(1) 基于八叉树的方法

基于八叉树的有损压缩方法和无损压缩方法的原理相似,都是通过八叉树结构划分三维空间再压缩。由于仅使用八叉树结构在修剪八叉树级别时会丢失大量点,失真严重且点云有损压缩性能差,文献[57]提出了一种基于神经网络的学习逼近模型 GeoCNNv2,利用八叉树分区将点云划分为大小相同的块,通过尺度超先验熵模型编码,再利用多层卷积变换学习分区点云的特征。为了减少 3D 卷积占用的计算资源,文献[58]提出了多尺度网络 PCGCv2,在多尺度重采样模块采用稀疏卷积处理张量,使计算的空间和时间复杂度有效降低。

基于八叉树的方法通过八叉树编码器对非结构化点云进行编码,但需要解决八叉树结构带来的一系列问题。文献[57]和文献[58]分别针对信息丢失和计算复杂的问题进行了优化,有效改善了基于深度学习的点云有损压缩性能。

(2) 基于体素的方法

除了八叉树结构外,一些研究人员通过体素化操作获取结构化点云。Learned-PCGC 方法^[59]首先对点云进行体素化处理,经过缩放和分区得到相同尺寸的不重叠方块,利用基于 CNN 的变分自编码器(Variational Autoencoder, VAE)来学习潜在特征,完成了平滑的表面重建。

基于体素的点云有损压缩方法在编码前将点云坐标量化为整数精度,可以直接进行三维卷积,但是体素化会导致计算复杂度的指数式升高。

(3) 基于自编码器的方法

自编码器可以灵活地学习输入点云的隐含特征并进行编码,再经过解码操作重构三维点云。受二维图像压缩方法的启发,文献[60]提出了一种基于深度自编码器的点云有损压缩框架,对原始点云进行下采样并输入编码器得到压缩表示,经量化操作合并重复点以减少点数量,最后使用全连接解码器输出重建的点云。文献[61]将三维卷积与自编码器结合,使用权衡参数对速率和失真进行联合优化,借助了体素表示方法,将解码过程转换为体素占用的二分类,在低比特率下也能得到良好的压缩性能。为了充分使用点云自身的信息和属性,文献[62]基于自编码器网络设计了一种深度学习点云编码方法,将点云划分为多个大小相同的 3D 块,通过三维卷积提取潜在特征,实现了编解码的有效变换,但时间较长。考虑到现有的深度学习方法在局部细节重构方面表现不佳,文献[63]和文献[64]进行了改进。文献[63]设计了一种深度自编码器,采用多尺度分层编码器提取局部特征,再通过池化层生成综合特征,在优化重建质量的同时也提升了点云压缩比。文献[64]则重点关注局部密度,提出了一种密度保持深点云压缩方法,使用密度嵌入、局部位置嵌入和祖先嵌入来获取局部几何分布和密度,设计子点卷积学习丰富的上采样特征,缓解了编码过程中的聚类点问题,同时实现了速率和

失真的权衡。文献[65]提出了一种自适应八叉树引导的点云自编码器网络，将密集点云分解为不重叠的点云块，自适应地输入到网络中获取丰富特征信息，剔除了异常值，重建误差降低，可以高效处理任意大小的复杂点云。文献[66]引入了分布式源编码(Distributed Source Coding, DSC)的思想，使用基于 UNet 架构的卷积自编码器，将隐变量描述为综合征，从不同的参考数据重构点云，点云压缩的鲁棒性和灵活性得到了有效提升。鉴于先前的方法大多仅压缩单个点云帧，文献[67]利用时间和空间冗余进行编码，压缩一系列点云，在提升压缩率、压缩速度和精度的同时也减少了冗余计算。由于先前的压缩方法存在点卷积效率低、解码器结构不适合属性压缩的问题，Deep-PCAC^[68]借助几何学直接对点云属性进行编解码，设计了二阶点卷积和密集点-起始块来充分学习离散点之间的空间相关性，引入多尺度损失以引导自动编码器优化所有点的质量，网络的特征提取和传播能力得到提升，训练稳定性增强。

基于自编码器的方法在网络中使用三维卷积学习特征，虽然点云重建质量和压缩性能都有提升，但在实际应用中还需要考虑点云有损压缩耗费的计算资源和功率。

(4) 其他方法

除上述方法外，部分研究人员采用折叠网络、二值网络、稀疏卷积网络、Transformer 等其他网络结构进行点云有损压缩。文献[69]将点云视为二维空间的三维离散流形，提出了一种折叠网络，将原始点云属性映射到二维网格上，使用传统的图像压缩方法进行处理。由于先前的点云压缩方法并未有效消除点云帧之间存在的时间冗余，BPNet^[70]采用具有三维时空卷积的非对称残差模块获取时间和空间特征，根据预测的残差自适应地选择编码方式，实现了实时、高效的双向帧预测。文献[71]提出了一种基于遗传算法的二值量化网络，逐层搜索确定初始最佳尺度恢复因子，借助网络学习并更新，通过统计自适应池化损失最小化方法有效缓解了特征信息损失的问题。为了降低计算复杂性，SparsePCGC^[72]融合了稀疏卷积来分析局部邻域变化，使用多尺度表示有效聚合局部邻域信息，同时降低了点云有损压缩模型的空间复杂度。由于先前的方法在大规模密集点云场景的压缩效果不佳，TransPCC^[73]引入 Transformer 架构对输入的原始点云进行编码，通过自注意模块充分捕获每个点的最近邻特征，并采用逐点操作学习点之间的复杂关联，进而实现大规模点云压缩，有效降低重建误差。PCGFormer^[74]也以 Transformer 框架为基础，使用 K 最近邻搜索构建每个点的动态局部邻域，在多个尺度引入局部自注意模块并堆叠稀疏卷积以学习邻域中的典型特征表示，进一步提升了重构点云的质量。为了获取并增强有用的点云特征，文献[75]还引入了注意力机制，采用稀疏卷积对体素化点云进行多尺度下采样和特征信息提取，使用两个注意力模块增强下采样后的特征，再通过 Transformer 模块提取并恢复有用特征，重建点云的结构信息更加丰富。

上述 7 种方法构建了各自的编解码网络，将二维图像压缩领域应用较好的网络扩展到三维点云压缩中，取得了良好的实验效果。

小结：基于深度学习的点云有损压缩方法提高了压缩率和点云的重建质量，八叉树、体素、自编码器以及其他网络结构都能够获取点云的有效特征，但是三维计算复杂度也随之增加，如何减少编解码时间是点云有损压缩方法的一大难点。

3.2.3 点云压缩方法小结

基于深度学习的点云压缩方法通过神经网络对点云数据进行编码，量化潜在特征，再通过熵模型压缩得到比特流。深度学习方法可以适应复杂的特征结构，从大量点云数据中学习编码器与解码器的变换，是今后编码压缩的一个极具潜力的研究方向。但无论是无损压缩还是有损压缩，都需要解决三维计算的内存占用问题。表 3 和表 4 分别给出了基于深度学习的点云无损压缩和有损压缩方法的对比情况。

表 3 基于深度学习的点云无损压缩任务主要方法对比

Table 3 Comparison of main methods for point cloud lossless compression tasks based on deep learning

Type	Method	Specific structure	Contribution	Limitation
Octree-based	[50]	Octree encoding	Using network training entropy model	Neighborhood information not used
	[51]	Multi-context deep learning	Using the feature of sibling nodes	Decoding speed can be accelerated
Hybrid representation	[52]	Voxel context compression octree structured	Suitable for static and dynamic point cloud compression	Higher resolution features are ignored
	[53]	Deep autoregressive generative models	Apply autoregressive generative models to 3D	Long encoding and decoding time
	[54]	Multiscale deep context model	Parallel voxel prediction	Sparse point cloud effect is poor
Other methods	[55]	Based on learning conditional probability model	Capture features and relationships of point clouds by sparse tensors	Runtime is highly dependent on the number of occupied blocks
	[56]	Combination of multi-scale and sparse convolutional network	Use cross-scale, cross-group and cross-color correlations to approximate attribute probabilities	When the prediction module increases, the algorithm complexity will also increase

表 4 基于深度学习的点云有损压缩任务主要方法对比

Table 4 Comparison of main methods for point cloud lossy compression tasks based on deep learning

Type	Method	Specific structure	Contribution	Limitation
Octree-based	[57]	Learning approximation model based on neural network	Use octree partition to divide point cloud patches of the same size	Long training time
	[58]	Multiscale end-to-end network	Learn point cloud features by sparse convolution	Noise can affect performance
Voxel-based	[59]	Variational autoencoders based on neural networks	Apply stacked 3D convolutions in a variational autoencoder structure	Convolution efficiency needs to be improved
Auto-Encoder	[62]	Encoding method based on CNN	Extend deep learning coding methods	Long encoding and decoding time
	[63]	Deep autoencoders with hierarchical	Multi-scale layered encoder to obtain	Can only handle small and fixed

		structure	features at each level	size point clouds
	[66]	Convolutional autoencoders	Enhanced encoding robustness and more flexible decoding	Rate distortion
	[67]	Compression with spatial and temporal redundancy	Increased compression ratio and compression speed	Computational cost is high
Other methods	[69]	Folding-based network	Fold the 3D manifold onto the image	Unsuitable for point clouds with complex geometries
	[73]	End-to-end TransPCC framework	Learn complex relationships between points by self-attention structure	Computational efficiency needs to be improved
	[74]	Multi-scale local self-attention mechanism	Capture high-level feature in dynamic local neighborhoods	Model running speed still needs to be optimized
	[75]	Transformer network model based on attention mechanism	Use the Transformer to enhance point space feature perception	Long encoding and decoding time

3.3 点云超分辨率

点云超分辨率是一种提升分辨率的点云增强技术,能够对稀疏 LR 点云数据进行上采样,恢复由环境和设备影响而缺失的信息,有效解决了雷达扫描的稀疏性问题。同时,生成的密集 HR 点云数据包含更丰富的细节,便于后续的三维重建和数据理解。基于深度学习的点云超分辨率方法根据网络类型可以分为以下 4 种:基于卷积神经网络、基于图卷积神经网络、基于生成对抗网络和其他结构。

3.3.1 基于卷积神经网络的方法

卷积神经网络是三维点云超分辨率任务最常用的一种网络结构。PU-Net^[76]使用卷积单元学习点云块中每个点的多级特征,通过联合损失函数使生成的点均匀分布在对象的表层。为了提高地形特征描述的准确性,文献[77]提出一种点云密度增强网络,将三维激光点云数据转换为深度图,借助 SRCNN(Super-Resolution Convolutional Neural Network)提升深度图的清晰度,再将超清晰的深度图转化为三维点云数据。考虑到低成本激光雷达收集到的点云密度有限,文献[78]从原始稀疏点云中获取深度图,利用多层 CNN 学习深度图的不连续性来恢复密集点云。文献[79]设计了一种两阶段点云超分辨率框架,在第一阶段使用局部插值方法丰富稀疏特征以提高目标点云的密度和均匀性,在第二阶段采用外积神经网络来调整插入点的坐标,进而提升点云超分辨率的精度。文献[80]提出了基于信道注意力的融合网络,采用圆形填充方法恢复投影到距离图像后的边界点,进而得到更高精度的高分辨率点云。为了实现高效的大型点云超分辨率,文献[81]使用基于稀疏卷积的深度神经网络来预测逐点占用率,减少了体素化点云的失真,同时使用动态阈值避免生成具有缺失点的区域。

基于卷积神经网络的方法采用卷积操作获取并学习点云特征,旨在增强点云密度以提升点云超分辨率性能。

3.3.2 基于图卷积神经网络的方法

针对点云数据缺乏规则内部结构的问题,部分研究人员提出构建图结构,采用 GCN 捕获拓扑图的特征信息。由于大多数点云超分辨率方法缺乏局部信息,PU-GCN^[82]利用多尺度点特征提取器提取特征,通过图卷积网络聚合点的邻域信息,能够生成细粒度细节。文献[83]则提出了一种动态残差图卷积网络 PSR-DRGCN,将动态邻域空间与图卷积操作结合,增大感受野以获取局部几何信息,点云的生成质量提升。先前的方法仅使用几何属性来预测密集的几何坐标,文献[84]设计了一种联合处理几何和颜色属性的双通道图卷积网络,应用高维属性特征的相似性构造图结构,进而得到准确的上下文信息。

基于图卷积神经网络的方法在特征提取部分应用 GCN 保留了丰富的有效细节信息,但是构建图结构不可避免地增加了计算成本,这也是今后继续优化的一个研究重点。

3.3.3 基于生成对抗网络的方法

基于生成对抗网络的点云超分辨率方法的核心是生成器和判别器之间的博弈,生成器需要尽可能生成判别器难以区分的模拟点云数据,判别器需要尽可能准确地判断生成的点云数据是真实的还是模拟的,二者形成动态对抗关系。为了提高点云分布的均匀性,PU-GAN^[85]引入了对抗性框架来训练上采样生成器学习潜在空间的点特征并得到分布均匀的点集,利用判别器惩罚偏离的输出,结合自注意单元融合特征。考虑到 PU-Net^[76]忽略了 LR 和 HR 点云之间的相似性且训练困难,文献[86]设计了一种对抗残差图网络 AR-GCN,借助图卷积学习点云的局部相似性以及输入和输出之间的相似性,引入图对抗损失函数自动获取点云特征。由于现有方法难以适应所有尺寸和形状结构,文献[87]提出了一种整体水平的内部上采样框架 ZSPU(“Zero-Shot” Point Cloud Upsampling),将 LR 点云输入生成器并按比例逐步上采样,在鉴别器中连接全局和逐点特征以理解点云的细节和整体形状,训练时间大大减少。

基于生成对抗网络的方法利用生成器学习点特征,结合鉴别器进一步优化细节信息,有效提升了点云超分辨率性能。

3.3.4 基于其他结构的方法

点云超分辨率方法还可以借助编解码器等结构捕获更多的点云特征。由于 PU-Net^[76]难

以处理大尺度或精细结构的点，文献[88]设计了一种渐进式点集上采样网络 MPU，通过多步上采样单元学习不同级别的特征，采用密集连接来共享特征，保留了较多的局部细节信息。考虑到先前的超分辨率方法主要适用于几何物体，文献[89]提出了一个两阶段 3D 人脸点云超分辨率网络，从低分辨率三维人脸点云数据中预测高分辨率数据，结合正面人脸特征，将点云超分辨率转换为 z 坐标预测问题，使用级联自编码器保留不同面部的边界信息。PU-Transformer 方法^[90]在点云上采样过程引入 Transformer 模型，采用多头自注意力结构提取丰富的细粒度点特征，利用位置融合模块学习位置信息并进行编码，进一步提升了点云超分辨率的准确性。为了提高从雷达传感器获得的点云质量，文献[91]提出了一种基于移位窗口的编码器-解码器神经架构，编码器结合了视觉转换器，解码器引入了全连接条件随机场，模型能够捕获更丰富的局部和全局语义特征，点云超分辨率的准确性得到有效提升。

基于其他结构的方法将二维超分辨率算法的思想与 PointNet^[92]的结构相结合，扩展了三维点云超分辨率的方法，但目前相关的研究还比较少。

3.3.5 点云超分辨率方法小结

点云超分辨率方法通过各种神经网络学习点云的结构信息和细节特征，基于 CNN 和 GCN 的方法获取了丰富的细粒度特征，但会占用大量的计算资源；基于 GAN 的方法能够学习点云样本的细节信息，但是对抗样本会干扰网络使其做出错误判断；基于其他结构的方法成功将二维应用较好的模型扩展到三维，但还需要深入的研究以获取更好的点云超分辨率性能。基于深度学习的点云超分辨率方法对比如表 5 所示。

表 5 基于深度学习的点云超分辨率任务主要方法对比

Table 5 Comparison of main methods for point cloud super-resolution tasks based on deep learning

Type	Method	Specific structure	Contribution	Limitation
CNN-based	[76]	Multi-level feature aggregation	Good anti-noise performance	Inability to fill large holes
	[77]	Point cloud density enhanced convolutional network	Enhancing point cloud density with SRCNN	Point cloud density increase is small
	[78]	Based on single LiDAR	Eliminates dependency on camera	Sensitive to outliers
	[80]	Channel-based attention network	Use circular fills to solve edge recovery issues	Need more reasonable evaluation indicators
GCN-based	[82]	Graph convolutional network	Fewer network parameters	Computational cost increases
	[83]	Dynamic residual graph convolutional networks	Learn local geometric features by multilayer graph convolution	Sensitive to rotating point clouds
	[84]	Double-channel graph convolutional network	Apply feature similarity to construct local graphs of point clouds	Computational complexity increases
GAN-based	[85]	Based on GAN	Robust to noise and sparse point clouds	Unsuitable for filling large gaps
	[86]	Adversarial residual graph	Obtain features by graph confrontation	Cannot repair large holes or missing

		network	loss function	parts
	[87]	“Zero-shot” point cloud upsampling network	Training time is reduced	Complex regions are still mismatched
Other structure methods	[88]	Progressive point set upsampling network	The generated point cloud is smoother and more complete	Difficult to handle sparse low-quality point clouds
	[89]	Face point cloud super-resolution network	Predict high-resolution data from low-resolution data	The preprocessing stage is not in the super-resolution network
	[90]	Based on the Transformer	Can upsample different types of data	Consume more network parameters

3.4 点云修复、补全和重建

点云修复、补全和重建是提升点云质量的关键步骤，主要包含基于图像、基于采样和基于补全这 3 种深度学习方法。

3.4.1 基于图像的方法

基于图像的方法通过二维图像重建三维对象。由于采集的深度图像分辨率低且边缘不准确，文献[93]设计了一种基于深度图像纹理边缘引导的重建方法，通过上采样和基于结构化学习的边缘检测算法获取图像边缘，结合深度信息提取正确的边缘，利用插值填充错误边缘和正确边缘间的不可靠区域，重建点云的精度提升。PCDNet^[94]将图像中提取的全局特征与点云编码特征混合，通过 GraphX 层进行点云的随机变形，生成最终的对象点云。为了提高图像重建三维形状的扩展能力，Pix2Pc^[95]将单个图像作为输入，利用 CNN 提取图像的形状特征，并根据形状特征对随机生成的点云进行变形，得到最终表示。由于目前口腔修复方法未考虑缺失牙体的多样性，文献[96]提出了基于高分辨的生成对抗网络，使用基于图像熵的正交投影方法得到包括咬合面、颊侧面和舌侧面等形态的高分辨率牙体图像，借助生成对抗的思想充分学习牙体特征，缺失牙冠的重建质量得到改善。

基于图像的方法都需要提取二维图像特征，通过变形将特征信息传递到三维点云中，节省了计算成本。

3.4.2 基于采样的方法

基于采样的方法通过学习给定点云数据的特征信息，生成更密集且均匀的点云，进而实现点云修复。为了进一步提高点云几何细节的修复性能，文献[88]提出了一种多步上采样网络，利用密集连接提取点云块特征，通过双边特征插值进行层间特征传播，实现了对不同细节层次信息的自适应学习。针对点云极其稀疏、丢失关键结构的情况，Zhang 等人^[97]设计了基于数据驱动的采样算法，将整个点云模型作为输入，利用深度网络捕获潜在特征并辅助采

样,生成密集点云。由于点云修复的质量与点云密度和曲面法线估计密切相关,Sharma 等人^[98]从点云块中提取特征,利用特征重塑进行采样,通过复合损失函数预测法线,得到更准确的点云修复结果。考虑到特征点匹配不足或错误时,点云重建精度会降低,文献^[99]将MVSNet(Multi-View Stereo Network)网络模型应用到三维重建中,使用多尺度卷积网络充分学习深度图像的高级特征,减轻了在纹理稀疏或重复等不利环境下对特征点的依赖。

基于采样的修复方法直接处理点云数据,依赖于采样网络,如何生成密集且完整的点云是修复的重点问题。

3.4.3 基于补全的方法

基于补全的方法一般借助深度神经网络进行点云修复。为了降低运行内存,减少原始信息丢失,PCN 方法^[100]直接对原始点云进行补全,利用编码器-解码器网络结构生成密集完整的点云。TopNet 方法^[101]设计了一种基于树结构的解码器,根节点从编码器获取点云集的特征向量,叶子节点嵌入所有后代的集合,遵循分层树结构生成完整的结构化点云,解码器可以学习任意拓扑结构,灵活直观。考虑到 PCN 的输入点云只与编码器部分相关,文献^[102]利用折叠解码器结合输入点云信息,再通过多级损失的网络架构学习全局特征,在缺失区域生成真实的结构。PF-Net^[103]以不完整的点云作为输入,直接预测缺失的点云,通过多层感知提取器捕获多尺度特征,利用两种损失函数优化补全网络,生成的点云不仅能保留输入轮廓,还具有丰富的特征信息。虽然现有的点云补全方法得到了较好的结果,但对细节部分的补全效果一般,文献^[104]、SA-Net^[105]和 MRC-Net^[106]针对该问题对网络进行了调整。文献^[104]设计了一种级联细化网络 CRN,通过对抗性训练学习准确的点分布,利用部分输入的细节来补全缺失部分,同时保留了原始点云的现有细节。SA-Net^[105]利用跳跃注意机制将无序点云的局部特征传递到解码器,通过分层折叠逐步细化点云并分层保留结构细节,进而补全完整的点云形状。MRC-Net^[106]是一种多分辨率补全网络,引入多尺度退化机制降低预训练阶段 GAN 生成的 3 种不同分辨率点云,通过多尺度判别器逐步优化生成器,点云的细节特征得以更好地保留。由于大部分方法直接从不完整点云提取全局特征,忽略了局部特征信息,文献^[107]、DPG-Net^[108]和 AGFA-Net^[109]对编码器和解码器进行了改进。文献^[107]采用多尺度特征提取模块获取全局和局部特征,通过跨区域注意模块构建局部特征图进行图卷积,利用两阶段编码器生成精细点云。DPG-Net^[108]则基于生成对抗方法,利用多分辨率编码器从点云中学习并捕获有用的特征信息,再通过渐进式增长编码器恢复完整的点云结构。文献^[109]设计了一种自适应全局特征增强网络(AGFA-Net),采用空间注意模块计算逐点注意值以自适应地聚合局部和全局特征,使用通道注意模块增强有用的几何细节特征,通过添加

跳跃连接在点云补全过程中传递几何特征，进而生成更加精细的点云形状。文献[110]和 Tensorformer^[111]借助 Transformer 结构捕获局部特征。文献[110]在编码器部分借助特征嵌入模块捕获点云局部特征信息，使用 Transformer 多尺度注意力模块提取特征，同时结合生成对抗思想进一步优化补全后的点云结构。Tensorformer^[111]是一种归一化矩阵注意力 Transformer，利用矩阵注意模块同步整合来自不同通道和邻域的特征信息，结合线性归一化方法避免梯度消失，点云局部几何形状的重建能力增强。鉴于机载雷达采集建筑物点云时受地物遮挡的限制，难以得到完整的立面，文献[112]设计了一种基于几何特征的隐式重建网络 GEOP-Net，通过几何编码模块提取高维特征，利用隐式网络学习占用概率构造建筑物曲面，从而实现从稀疏和不完整点云预测完整模型结构。

基于补全的修复方法能够利用部分输入点云填补缺失的结构，但是目前不完整点云的结构细节和局部信息对整体补全的影响较大。

3.4.4 点云修复、补全和重建方法小结

深度学习方法可以有效修复不完整点云，尽可能还原三维模型。基于图像的方法实现简单，但在训练中拓扑结构不易改变。采样和补全的修复方法都对点云数据进行处理，简单直接，是目前主流的修复方法。基于深度学习的点云修复、补全和重建方法对比如表 6 所示。

表 6 基于深度学习的点云修复、补全和重建任务主要方法对比
Table 6 Comparison of main methods for point cloud restoration, completion and reconstruction tasks based on deep learning

Type	Method	Specific structure	Contribution	Limitation
Image-based	[94]	Point cloud deformation network	Invariant to disordered point clouds	Lack of some details
	[95]	CNN	Efficient and scalable	Lack of projection information
Sampling-based	[88]	Multi-step upsampling network	Robust to noisy and sparse inputs	Unsuitable for sparse point clouds
	[97]	Data driven	Generate more accurate upsampling with less chamfer loss	Sampling of unknown features degrades
	[98]	Feature reshaping	The generated point cloud is smoother and more complete	Difficult to handle sparse input
Completion-based	[100]	Learning-based shape completion methods	Robust to occlusion and noise	Not sure if the output preserves the input points
	[103]	Multi-scale generative network based on feature points	The spatial arrangement of the point cloud is preserved	Only a part of the point cloud missing area is predicted
	[104]	Cascade refinement network	Remain more details	Occlusion leads to large errors
	[105]	Skip-attention network	High-quality point cloud restoration	Calculation efficiency still needs to be optimized
	[111]	Normalized matrix attention Transformer	Integrate features from different channels and neighborhoods	High computational complexity

4 数据集及评价指标

4.1 数据集

近年来,多种类型的点云公开数据集极大地促进了点云处理任务的深入研究。表 7 整理了在三维点云处理领域常用的 22 种数据集,给出了每种数据集适用的点云处理任务和网址,“√”表示数据集适用该类任务。表格中 PCD 表示点云去噪, PCC 表示点云压缩, PCSR 表示点云超分辨率, PCR 表示点云重建。

表 7 基于深度学习的点云处理任务常用数据集
Table 7 Common datasets for point cloud processing tasks based on deep learning

Dataset	Year	Task				Website
		PCD	PCC	PCSR	PCR	
KITTI ^[113]	2012	√	√	√	√	http://www.cvlibs.net/datasets/kitti
Paris-rue-Madame ^[114]	2014	√	—	—	√	https://people.cmm.minesparis.psl.eu/users/serna/rueMadameDataset.html
SHREC15 ^[115]	2015	—	—	√	√	https://www.icst.pku.edu.cn/zlian/representa/3d15/index.htm
ModelNet ^[116]	2015	√	—	√	√	http://modelnet.cs.princeton.edu/
ShapeNet ^[117]	2015	√	√	√	√	https://shapenet.org/
vKITTI ^[118]	2016	—	—	—	√	https://europe.naverlabs.com/Research/Computer-Vision/Proxy-Virtual-Worlds/
ShapeNet Part ^[119]	2016	—	√	—	—	https://cs.stanford.edu/~ericyi/project_page/part_annotation/
S3DIS ^[120]	2016	—	√	—	√	http://buildingparser.stanford.edu/dataset.html
MVUB	2016	—	√	—	—	http://plenodb.jpeg.org/pc/microsoft/
8iVFB	2017	—	√	—	—	http://plenodb.jpeg.org/pc/8ilabs/
3DMatch ^[121]	2017	—	—	—	√	http://3Dmatch.cs.princeton.edu/#rgbd-reconstruction-datasets
ScanNet ^[122]	2017	√	√	√	√	http://www.scan-net.org/
Matterport3D ^[123]	2017	—	—	—	√	https://niessner.github.io/Matterport/
PU-Net ^[76]	2018	√	—	√	—	https://drive.google.com/file/d/1R21MD1O6q8E7ANui8FR0MaABkKc30PG4/view
PCN ^[100]	2018	—	—	—	√	https://drive.google.com/drive/folders/1M_IJN14Ac1RtPtEQxNICV9e8pom3U6Pa
PU-GAN ^[85]	2019	—	—	√	—	https://drive.google.com/file/d/1BNqjidBVWP0_MUDMTeGy1wZiR6fqyGmC/view?pli=1
SemanticKITTI ^[124]	2019	—	√	√	—	http://semantic-kitti.org/
MPEG PCC ^[125]	2019	—	√	—	—	https://mpeg-pcc.org/
nuScenes ^[126]	2020	—	—	—	√	https://nuscenes.org/
Waymo ^[127]	2020	—	—	—	√	https://waymo.com/open/
PCNet ^[35]	2020	√	—	—	—	https://nuage.lix.polytechnique.fr/index.php/s/xSRrTNmtgqgeLGa
PU1K ^[82]	2021	√	—	√	—	https://drive.google.com/file/d/1oTAX34YNbL6GDwHYL2qqvjmYtTVWcELg/view

PCNet 数据集^[35]包含 18 个训练形状和 10 个测试形状，包括简单的几何形状和复杂的对象形状。每个干净点云包含从原始三角形网格均匀采样的十万个点。PU-Net^[76]收集了来自 Visionair 存储库的 60 个不同模型，其中训练模型 40 个，测试模型 20 个，包含 Bunny 等非刚性物体和椅子等刚性物体。PU1K 数据集^[82]由 1147 个 3D 网格模型组成，其中 1020 个训练样本和 127 个测试样本。大多数样本来自 ShapeNetCore 数据集，涵盖 50 个不同类别，包括简单形状和复杂形状。PCN 数据集^[100]是从 ShapeNet 数据集选取了飞机、橱柜、汽车、椅子、台灯、沙发、桌子、容器 8 种类别构建的，共 30974 个模型。其中测试数据 150 个，验证数据 100 个，剩余的都为训练数据。

Microsoft Voxelized Upper Bodies(MVUB)数据集包含 Andrew、David、Phil、Ricardo 和 Sara 共 5 个主体的动态体素化点云序列，提供两种空间分辨率：一个 $512 \times 512 \times 512$ 和一个 $1024 \times 1024 \times 1024$ 体素的立方体，分别标记为深度 9 和深度 10。8i Voxelized Full Bodies(8iVFB)数据集包含 longdress、loot、redandblack 和 soldier 这 4 个动态体素化点云序列，提供一个标记为深度 10 的空间分辨率。

4.2 评价指标

点云处理任务涉及多个类别，每一类处理任务所采用的性能指标也不统一。因此本节整理了常用的 4 类评价指标，希望能够对点云处理方法的性能进行客观公正的评价。

4.2.1 精确度、空间复杂度和误差参数

精确度是评估点云处理性能的关键指标，其中精确率(Precision)、召回率(Recall)和 F-Score 是最常用的 3 种指标。精确率表示正确预测为正的点云样本占有所有预测为正的点云样本的比例，召回率表示正确预测为正的点云样本占有所有真实值为正的点云样本的比例。理想情况下精确率和召回率都是越高越好，但在实际点云处理中二者相互制约，需要综合考虑。F-score 是精确率和召回率的调和值，调和值越大表明预测点云的质量越好。公式分别为：

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$F - score = \frac{(1 + \beta^2) \cdot Precision \cdot Recall}{\beta^2 \cdot Precision + Recall} \quad (3)$$

其中，TP 为正样本被正确识别的数量，FP 为误报的负样本数量，FN 为漏报的正样本数量。

β 为调和参数， $\beta=1$ 表示精确率和召回率权重相同，称为 F1-score。 $\beta>1$ 表示召回率更重

要， $\beta < 1$ 则表示精确率更重要。

空间复杂度是算法运行所消耗的存储空间，反映了点云处理任务算法性能的优劣与否，通常借助模型内存和参数数量来评估深度学习性能。

误差评估参数反映了处理后得到的预测点云偏离真实点云的程度，常用的均方根误差(Root Mean Square Error, RMSE)与平均绝对误差(Mean Absolute Error, MAE)的公式分别如下：

$$\text{RMSE} = \sqrt{\text{MSE}} = \sqrt{\frac{1}{m} \sum_{i=1}^m (h(x_i) - y_i)^2} \quad (4)$$

$$\text{MAE} = \frac{1}{m} \sum_{i=1}^m |h(x_i) - y_i| \quad (5)$$

其中 m 是样本总数， $h(x_i)$ 是预测点云值， y_i 是真实点云值。

4.2.2 点云距离度量

点云距离度量包含倒角距离、陆地移动距离和豪斯多夫距离，反映了预测点云与真实点云的相似程度，距离指标越小，相似度越高。

倒角距离(Chamfer Distance, CD)表示预测点云 P 中每个点到真实点云 Q 中最近点的平均距离，陆地移动距离(Earth-Mover's Distance, EMD)表示将 P 中的点移动至 Q 中对应位置所需的最小总工作量，豪斯多夫距离(Hausdorff Distance, HD)表示 P 中的点到 Q 的最近距离的最大值。计算公式分别如下：

$$\text{CD}(P, Q) = \frac{1}{|P|} \sum_{p \in P} \min_{q \in Q} \|p - q\|_2 + \frac{1}{|Q|} \sum_{q \in Q} \min_{p \in P} \|p - q\|_2 \quad (6)$$

$$\text{EMD}(P, Q) = \min_{\phi: P \rightarrow Q} \sum_{p \in P} \|p - \phi(p)\|_2 \quad (7)$$

$$\text{HD}(P, Q) = \max[h(P, Q), h(Q, P)] \quad (8)$$

$$\begin{cases} h(P, Q) = \max_{p \in P} \min_{q \in Q} \|p - q\| \\ h(Q, P) = \max_{q \in Q} \min_{p \in P} \|q - p\| \end{cases} \quad (9)$$

其中 P 是预测点云， Q 是真实点云， p 是 P 内的 3D 点， q 是 Q 内的 3D 点。 $\|\cdot\|_2$ 表示欧式范数的平方， ϕ 是一个双射。HD 要求每个输出点在真实值上都有唯一对应的点，对异常值敏感；而 CD 不要求输出点和真实值之间一一对应，对点云的不规则分布不太敏感。

4.2.3 点云相似性度量

点云相似性度量包含结构相似性和峰值信噪比，反映了预测深度图像或者预测点云的失真程度，取值越高，失真越少，预测结果越接近原始数据。

结构相似性(Structural Similarity Index, SSIM)通过深度、深度对比度和结构来衡量两幅深度图像的相似度, 峰值信噪比(Peak Signal-to-Noise Ratio, PNSR)是评估预测点云信息失真程度的指标。公式分别如下:

$$SSIM(P, Q) = \frac{(2\mu_p\mu_q + C_1)(2\sigma_p\sigma_q + C_2)(\sigma_{pq} + C_3)}{(\mu_p^2 + \mu_q^2 + C_1)(\sigma_p^2 + \sigma_q^2 + C_2)(\sigma_p\sigma_q + C_3)} \quad (10)$$

$$PSNR = 10 \lg \left(\frac{M^2}{MSE} \right) \quad (11)$$

其中 μ_p 和 μ_q 分别表示图像 P 和 Q 的像素平均值, σ_p 和 σ_q 分别表示图像 P 和 Q 的像素标准值, σ_{pq} 表示图像 P 和 Q 的协方差, C_1 、 C_2 和 C_3 都为常数, 避免计算时出现分母为 0 的情况。峰值 M 为点云边界框的对角线距离, MSE 为均方误差。

4.2.4 其他参数

除了上述共用的点云评估参数外, 针对不同点云处理任务, 还有一些特定的评估参数。

点云去噪评估参数包括点到网格距离和噪声指标。点到网格距离(Point-to-Mesh Distance, P2M)通过计算去噪点云和干净三角网格之间距离的平均值来评估点云去噪效果。点云去噪和滤波任务还会引入各项同性高斯噪声指标来模拟不同程度的点云噪声, 公式如下:

$$p(x; s) = \frac{1}{\sqrt{2\pi}s} e^{-\frac{x^2}{2s^2}} \quad (12)$$

其中 s 为尺度参数, 在三维中可以设置为边界球半径的 1%、2% 和 3% 来模拟不同程度噪声。

比特位数(Bits Per Point, BPP)反映了点云压缩任务中每个点存储所占用的有效比特数。BPP 越小, 点云压缩效率越高。点云编码和解码时间也是衡量点云压缩性能的重要参数。

点云超分辨率性能的评估参数还有点到面距离和标准化均匀系数。点到面距离(Point-to-Surface Distance, P2F)为超分辨率处理后得到的点到点云表面之间距离的平均值。标准化均匀系数(Normalized Uniformity Coefficient, NUC)反映了点云分布的均匀性, NUC 值越小, 点集分布越均匀^[76], 公式如下:

$$avg = \frac{1}{KD} \sum_{k=1}^K \sum_{i=1}^D \frac{n_i^k}{N^k p} \quad (13)$$

$$NUC = \sqrt{\frac{1}{KD} \sum_{k=1}^K \sum_{i=1}^D \left(\frac{n_i^k}{N^k p} - avg \right)^2} \quad (14)$$

其中, K 表示测试对象数量, D 表示圆盘数量, n_i^k 表示第 k 个对象的第 i 个圆盘内的点数, N^k 表示第 k 个对象上的总点数。表 8 总结了 4 种点云处理任务常用的评估指标, 表格中 PCD、PCC、PCSR 和 PCR 分别表示点云去噪、点云压缩、点云超分辨率和点云重建。

表 8 常用的点云处理任务评估指标

Table 8 Common evaluation indicators for point cloud processing tasks

Tasks	Evaluation indicators			
	Accuracy	Distance	Similarity	Others
PCD	Precision, Recall, F-Score, RMSE, MAE	CD, EMD, HD	PSNR	P2M
PCC	Precision, Recall, F-Score, RMSE, MAE	CD, EMD, HD	PSNR	BPP, time
PCSR	Precision, Recall, F-Score, RMSE, MAE	CD, EMD, HD	SSIM, PSNR	P2F, NUC
PCR	Precision, Recall, F-Score, RMSE, MAE	CD, EMD, HD	PSNR	—

4.3 性能分析

由于大部分基于深度学习的点云处理方法采用自主采集的点云数据,部分点云处理方法虽然使用了公共数据集,但是涉及的数据集种类繁多且评价指标差异较大,难以在统一的指标下对比每一类处理方法的性能。本文仅整理了常用数据集上的几种经典方法的实验结果。

表 9 给出了经典的点云去噪方法在 PU-Net 和 PCNet 数据集上的性能对比结果,使用泊松盘采样从网格中采样分辨率为 10K 和 50K 的点,各向同性高斯噪声水平分别为 1%、2% 和 3%,评价指标采用 CD 和 P2M,表格中的数值都乘以 10^4 以便对比。PCNet 方法^[35]无法处理低分辨率点,常需要稍微膨胀去噪结果来减轻潜在的形状收缩;GPDNet 方法^[19]难以去除高水平噪声;Score-based 方法^[39]不存在形状收缩问题,PSR 方法^[40]可以在早期迭代产生高质量的去噪点云,但这两种方法得到的去噪后的点云分布不均匀;而 GPCD++方法^[42]能够生成更均匀、干净点云,在 CD 和 P2M 指标上都优于先前的方法。

表 9 经典的点云去噪方法在 PU-Net 和 PCNet 数据集上的性能对比

Table 9 Performance comparison of classic point cloud denoising methods on PU-Net and PCNet datasets

Points Noise		10K(Sparse)						50K(Dense)					
Dataset	Methods	1%		2%		3%		1%		2%		3%	
		CD	P2M	CD	P2M	CD	P2M	CD	P2M	CD	P2M	CD	P2M
PU-Net	PCNet ^[35]	3.515	1.148	7.467	3.965	13.067	8.737	1.049	0.346	1.447	0.608	2.289	1.285
	GPDNet ^[19]	3.780	1.337	8.007	4.426	13.482	9.114	1.913	1.037	5.021	3.736	9.705	7.998
	DMR ^[46]	4.482	1.722	4.982	2.115	5.892	2.846	1.162	0.469	1.566	0.800	2.432	1.528
	Score ^[39]	2.521	0.463	3.686	1.074	4.708	1.942	0.716	0.150	1.288	0.566	1.928	1.041
	PSR ^[40]	2.353	0.306	3.350	0.734	4.075	1.242	0.649	0.076	0.997	0.296	1.344	0.531
	GPCD++ ^[42]	1.881	0.251	2.728	0.654	3.433	1.161	0.505	0.073	0.852	0.303	1.198	0.534
PCNet	PCNet ^[35]	3.847	1.221	8.752	3.043	14.525	5.873	1.293	0.289	1.913	0.505	3.249	1.076
	GPDNet ^[19]	5.470	1.973	10.006	3.650	15.521	6.353	5.310	1.716	7.709	2.859	11.941	5.130
	DMR ^[46]	6.602	2.152	7.145	2.237	8.087	2.487	1.566	0.350	2.009	0.485	2.933	0.859
	Score ^[39]	3.369	0.830	5.132	1.195	6.776	1.941	1.066	0.177	1.659	0.354	2.494	0.657
	PSR ^[40]	2.873	0.783	4.757	1.118	6.031	1.619	1.010	0.146	1.515	0.340	2.093	0.573
	GPCD++ ^[42]	2.813	0.759	4.195	0.893	5.385	1.333	0.857	0.132	1.344	0.331	1.920	0.530

表 10 给出了经典的点云无损压缩方法在 MVUB 和 8iVFB 数据集上的比特位数对比结

果。基于几何的点云压缩方法(Geometry based PCC, G-PCC)^[128]是一种基于八叉树模型的传统点云压缩方法，多用于静态点云压缩。在评估基于深度学习的点云无损压缩方法时，通常会与 G-PCC 方法的性能进行对比，以此分析深度学习方法的优点。与传统 G-PCC 方法相比，VoxelDNN 方法^[53]使用深度生成模型来顺序估计占用概率，结合分区策略优化速率，但是上下文中缺少被占用的体素，在稀疏点云上的压缩性能下降，且编解码速度比 G-PCC 方法慢。MSVoxelDNN 方法^[54]通过消除同组体素间的部分依赖关系来实现并行处理，编解码速度相较于 VoxelDNN 方法也有了很大提升，但是依赖关系的破坏会降低生成上下文模型的精确度。OctAttention 方法^[51]采用具有固定长度上下文窗口的八叉树结构扩大了感受野，编码时间有效减少，鲁棒性也得到增强。

表 10 经典的点云无损压缩方法的比特位数结果

Table 10 Average bits per point (bpp) results of classic point cloud lossless compression methods					
Point Cloud	Frames	G-PCC ^[128]	VoxelDNN ^[53]	MSVoxelDNN ^[54]	OctAttention ^[51]
Microsoft Voxelized Upper Bodies (MVUB) dataset					
Phil9	245	1.23	0.92	—	0.83
Phil10	245	1.07	0.83	1.02	0.79
Ricardo9	216	1.04	0.72	—	0.72
Ricardo10	216	1.07	0.75	0.95	0.72
Average	—	0.95	0.81	0.99	0.76
8i Voxelized Full Bodies (8iVFB) dataset					
Loot10	300	0.95	0.64	0.73	0.62
Redandblack10	300	1.09	0.73	0.87	0.73
Boxer9/10	1	0.96/0.94	0.76/—	—/0.70	0.60/0.59
Thaidancer9/10	1	0.99/0.99	0.81/—	—/0.85	0.64/0.65
Average	—	0.99	0.73	0.79	0.64

表 11 给出了不同点云有损压缩方法在 8iVFB、KITTI 和 MVUB 数据集上的平均编解码时间对比结果，单位为秒。Learned-PCGC 方法^[59]应用了密集的三维卷积，编解码时间大大增加。PCGCv2 方法^[58]的编码时间与 G-PCC 八叉树编码器相当，但是解码过程需要处理基于二值分类的分层重建产生的多余体素，导致解码时间变长。SparsePCGC^[72]采用稀疏卷积减少了大量的空运算，编码时间与 G-PCC 方法相差不大，但是解码时间慢两倍。PCGFormer 方法^[74]的编解码时间略慢于 PCGCv2 方法，验证了应用 Transformer 结构压缩点云的合理性。

表 12 和表 13 分别给出了几种经典的点云超分辨率方法在 PU-Net 和 PU1K 数据集上的性能对比结果，评估指标包括 CD、HD、P2F、参数数量等。表 12 中的时间参数是算法的训练时间，而表 13 中的时间是指每次采样的平均推理时间。PU-Net 方法^[76]在处理高细节的性能较差，参数数量较多。MPU 方法^[88]采用基于块的架构，参数数量明显减少。AR-GCN 方

法^[86]和 PU-GAN 方法^[85]都引入了对抗性损失。AR-GCN 方法的 CD 和 HD 指标明显优于 PU-Net 方法，但是使用的参数量相近。PU-GAN 方法在距离指标上优于 PU-Net、MPU 和 AR-GCN 方法，训练时间和 MPU 相当，但是模型参数比 PU-Net 和 AR-GCN 方法少。PU-GCN 方法^[82]缩减了模型尺寸和参数数量，CD 和 HD 指标也优于 PU-Net 和 MPU，与 PU-GAN 相比，PU-GCN 方法的训练时间大幅减少。ZSPU 方法^[87]进一步优化算法框架，避免为各种场景选择超参数，训练时间明显少于先前的方法。PU-Transformer 方法^[90]在 CD、HD 和 P2F 三个指标上都优于先前的方法，但是 Transformer 结构会带来模型尺寸和参数数量的增加。

表 11 不同点云有损压缩方法的平均编解码时间对比

Table 11 Comparison of average encoding and decoding time for different point cloud lossy compression methods

Methods	8iVFB dataset		KITTI dataset		MVUB dataset	
	Encoding	Decoding	Encoding	Decoding	Encoding	Decoding
	time/s	time/s	time/s	time/s	time/s	time/s
G-PCC(octree) ^[128]	1.6	0.6	—	—	0.73	0.07
G-PCC(trisoup) ^[128]	8.1	6.6	—	—	2.06	1.10
G-PCC v8 ^[128]	—	—	1.30	0.55	—	—
Learned-PCGC ^[59]	9.3	9.5	—	—	—	—
PCGCv2 ^[58]	1.6	5.4	—	—	0.53	0.18
SparsePCGC ^[72]	—	—	1.44	1.32	—	—
PCGFormer ^[74]	—	—	—	—	0.87	0.51

表 12 不同点云超分辨率方法在 PU-Net 数据集上的性能对比

Table 12 Performance comparison of different point cloud super-resolution methods on PU-Net dataset

Methods	CD	HD	P2F (10^{-3})		NUC	Epoch	Time	Parameter
	(10^{-3})	(10^{-3})	μ	σ	0.4% (10^{-3})			(10^3)
PU-Net ^[76]	0.38	3.67	8.19	6.65	6.36	120	4.5h	814
AR-GCN ^[86]	0.23	1.78	3.02	3.52	1.29	120	6.2h	822
MPU ^[88]	0.21	1.90	1.72	2.21	1.32	400	27h	304
PU-GAN ^[85]	0.17	1.76	1.05	1.92	0.55	100	25h	684
PU-GCN ^[82]	0.26	2.62	2.15	3.01	1.75	100	9h	542
ZSPU ^[87]	0.19	1.11	2.12	2.21	2.24	50	96s	310

表 13 不同点云超分辨率方法在 PU1K 数据集上的性能对比

Table 13 Performance comparison of different point cloud super-resolution methods on PU1K dataset

Methods	CD	HD	P2F	Epoch	Time	Parameter	Model
	(10^{-3})	(10^{-3})	(10^{-3})		(10^{-3} s)	(10^3)	
PU-Net ^[76]	1.155	15.170	4.834	100	8.4	812.0	10.1
MPU ^[88]	0.935	13.327	3.551	100	8.3	76.2	6.2
PU-GCN ^[82]	0.585	7.577	2.499	100	8.0	76.0	1.8
PU-Transformer ^[90]	0.451	3.843	1.277	100	9.9	969.9	18.4

表 14 给出了不同的点云修复、补全和重建方法在 PCN 和 ShapeNet 数据集上的性能对比结果，评估指标分别为倒角距离的平均值和倒角距离，值越低越好。在 PCN 数据集上，

PCN 方法^[100]的结果总体优于 TopNet 方法^[101]，而在 ShapeNet 数据集上 TopNet 方法的性能更好，但是 PCN 和 TopNet 方法都无法恢复 3D 物体的精细细节。CRN 方法^[104]在所有类别的平均 CD 值都有所下降，但能够获得更精细的局部形状。SA-Net 方法^[105]通过跳跃注意机制充分利用不同分辨率的局部区域，在 ShapeNet 数据集上的 6 个类别的性能得到提升。AGFA-Net 方法^[109]采用空间注意模块替代 K 最近邻运算，能够提取更有效的局部特征，在 PCN 数据集的所有类别中表现最佳，恢复的点云表面更加光滑准确。

表 14 不同的点云修复、补全和重建方法的性能对比

Table 14 Performance comparison of different point cloud restoration, completion and reconstruction methods									
Methods	Average	Airplane	Cabinet	Car	Chair	Lamp	Sofa	Table	Vessel
PCN dataset (Mean Chamfer Distance per point $\times 10^3$)									
PCN ^[100]	9.64	5.50	10.63	8.70	11.00	11.34	11.68	8.59	9.67
TopNet ^[101]	9.89	6.24	11.63	9.83	11.50	9.37	12.35	9.36	8.85
CRN ^[104]	8.51	4.79	9.97	8.31	9.49	8.94	10.69	7.81	8.05
AGFA-Net ^[109]	6.76	3.89	9.03	7.68	7.18	5.52	8.72	6.18	5.91
ShapeNet dataset (Chamfer Distance per point $\times 10^4$)									
PCN ^[100]	14.72	8.09	18.32	10.53	19.33	18.52	16.44	16.34	10.21
TopNet ^[101]	9.72	5.50	12.02	8.90	12.56	9.54	12.20	9.57	7.51
SA-Net ^[105]	7.74	2.18	9.11	5.56	8.94	9.98	7.83	9.94	7.23

5 总结与展望

随着计算机硬件设备和传感器采集技术的不断升级，基于深度学习的三维点云处理方法受到了广泛研究，但由于点云数据的特性以及实际应用场景的限制，仍存在以下 4 个问题：

(1) **噪声和异常值。**受扫描设备精度和障碍物遮挡的影响，点云在采集过程中不可避免地存在噪声和离群点。而噪声点和离群点会极大地影响表面法线、曲率等局部点云特征，进而影响点云压缩、超分辨率、修复、重建等任务的结果，因此点云去噪和滤波处理十分重要。

(2) **点云的不规则性、无序性和旋转性。**受采集设备和区域的影响，采集到的点云数据分布不均匀，稀疏和密集区域兼而有之，且点的距离也是随机的；点云是按一系列点集的形式存储的，点存储先后次序的改变并不会影响点云所表示的场景；点云旋转或平移后的位置坐标会变化，但表示的点云仍旧不变。综上，点云是无序、非结构化的，而深度学习适合处理有序、结构化的数据，直接处理点云数据十分困难。

(3) **依赖硬件和采集设备。**深度网络的运行需要解决复杂的矩阵计算，而 CPU 和 GPU 能够处理大量的计算，缩短训练时间，但目前除 CPU 和 GPU 外暂时没有其他能够高效处理复杂计算的配置。此外，点云采集设备的精度还有待提升，设备易受周围环境的干扰，且采

集设备的成本较高。

(4) 多样化场景。现有的基于深度学习的点云处理方法大多是针对某类场景或者某种数据集进行处理，训练时场景较单一，对未知新场景的测试效果不一定好，算法在实际工程应用时的泛化性还可优化。

综上，虽然目前基于深度学习的三维点云处理方法取得了很多优异的研究成果，但是仍存在很大的改进空间，具体可以从以下 4 个方面深入研究：

(1) 多种处理任务结合。在点云实际处理过程中，为了提升点云质量，常常会同时处理多种任务。点云去噪和修复都需要去除噪声和异常值，将两种任务结合起来可以得到更干净且完整的点云。点云压缩、超分辨率和补全任务都涉及编码和采样操作，三者结合可以获取丰富的特征信息。多种处理任务结合是点云处理未来发展的一个重要方向，综合了各种任务的优势，有望实现更优异的性能。

(2) 点云数据特性处理。鉴于采集的点云数据存在遮挡、旋转以及稀疏区域，虽然不会影响点云的表示，但使用深度学习方法处理起来十分复杂。因此，如何在预处理阶段搜索密集且有效的点云区域，以及如何借助旋转操作降低目标之间的相互遮挡是未来实现高效三维点云处理值得尝试的新方向。

(3) 网络模型和硬件设备。由于大部分基于深度学习的点云处理方法在追求精度提升的过程中，忽视了模型的复杂度和计算效率。为了保证在实际应用中的时效性，轻量化的网络模型是一个不错的选择。目前点云处理任务在自动驾驶、机器人、遥感、文物修复、口腔医疗等方面得到了应用和推广，但是由于三维采集设备的成本较高，难以普及到人们的实际生活中，未来还需要进一步探索低成本的点云采集设备。

(4) 数据集和评价指标。目前同一种处理任务所采用的数据集不一致，不同任务的数据集也不统一，采用的评价指标也是多种多样，这给分析比较基于深度学习的三维点云处理方法带来了很大的困难。因此，未来研究人员可以重点研究构建一个丰富且适用性强的数据集，深入探索适合对比各类点云处理任务性能的评价指标。

点云处理任务在自动驾驶、机器人、遥感等领域的应用越来越广，采用深度学习方法处理点云不仅提高了效率，还实现了优秀的性能。本文总结了近十年来国内外基于深度学习的三维点云处理方法的研究进展，重点阐述了点云去噪和滤波、点云压缩、点云超分辨率以及点云修复、补全和重建这 4 种任务的深度学习方法，详细对比并分析了各种方法的优势和局限性。之后介绍了常用数据集和评估指标，对未来点云处理的研究工作进行了展望，期望通过本文的总结和分析，能为三维点云处理研究提供一定的参考。

参考文献

- [1] 郑少武,李巍华,胡坚耀.基于激光点云与图像信息融合的交通环境车辆检测[J].仪器仪表学报, 2019, 40(12): 143-151.
Zheng S W, Li W H, Hu J Y. Vehicle detection in the traffic environment based on the fusion of laser point cloud and image information[J]. Chinese Journal of Scientific Instrument, 2019, 40(12): 143-151.
- [2] 余洪山,付强,孙健,等. 面向室内移动机器人的改进 3D-NDT 点云配准算法[J]. 仪器仪表学报, 2019, 40(9): 151-161.
Yu H S, Fu Q, Sun J, et al. Improved 3D-NDT point cloud registration algorithm for indoor mobile robot[J]. Chinese Journal of Scientific Instrument, 2019, 40(9): 151-161.
- [3] Espinosa N, Lenz A, Gross W, et al. Towards fast 3D reconstruction of urban areas from aerial nadir images for a near real-time remote sensing system[C]. IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 2018: 6472-6475.
- [4] 张子健,程效军,曹宇杰,等.结合激光与视觉点云的古遗迹三维重建应用[J].中国激光, 2020, 47(11): 1110001.
Zhang Z J, Cheng X J, Cao Y J, et al. Application of 3D reconstruction of relic sites combined with laser and vision point cloud[J]. Chinese Journal of Lasers, 2020, 47(11): 1110001.
- [5] 尚业华,张光强,孟志军,等.基于欧氏聚类的三维激光点云田间障碍物检测方法[J].农业机械学报, 2022, 53(1): 23-32.
Shang Y H, Zhang G Q, Meng Z J, et al. Field obstacle detection method of 3D LiDAR point cloud based on Euclidean clustering[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(1): 23-32.
- [6] 张贝,戴宁,田素坤,等.结合稀疏八叉树卷积神经网络的牙齿预备体颈缘线提取方法[J].计算机辅助设计与图形学学报, 2019, 31(12): 2129-2135.
Zhang B, Dai N, Tian S K, et al. Extraction of tooth preparation line based on sparse octree convolutional neural network[J]. Journal of Computer-Aided Design & Computer Graphics, 2019, 31(12): 2129-2135.
- [7] 麻卫峰,王成,王金亮,等.点云数据温度变化环境下输电线形态模拟[J].电力系统保护与控制, 2021, 49(9): 90-96.
Ma W F, Wang C, Wang J L, et al. Shape simulation for a transmission line under different temperatures based on point cloud data[J]. Power System Protection and Control, 2021, 49(9): 90-96.
- [8] 谷帆,张长伦,郭志光,等.提取桥梁点云的组合滤波算法[J].激光与光电子学进展, 2023, 60(2): 0211001.
Gu F, Zhang C L, Guo Z G, et al. Combined filtering algorithm for extracting bridge point cloud[J]. Laser & Optoelectronics Progress, 2023, 60(2): 0211001.
- [9] 李维刚,梅洋,樊响,等.基于车载激光点云的铁路轨道检测[J].中国激光, 2022, 49(4): 0410002.
Li W G, Mei Y, Fan X, et al. Railway track detection based on vehicle laser point cloud[J]. Chinese Journal of Lasers, 2022, 49(4): 0410002.
- [10] Liu W, Lai B, Wang C, et al. Learning to match 2D images and 3D LiDAR point clouds for outdoor augmented reality[C]. 2020 IEEE Conference on Virtual Reality and 3D User

Interfaces Abstracts and Workshops (VRW), Atlanta, GA, USA, 2020: 654-655.

- [11] Wirth F, Quehl J, Ota J, et al. PointAtMe: Efficient 3D point cloud labeling in virtual reality[C]. 2019 IEEE Intelligent Vehicles Symposium (IV), Paris, France, 2019: 1693-1698.
- [12] 解则晓,李美慧.机器学习在基于点云的三维物体识别领域的研究综述[J].中国海洋大学学报(自然科学版), 2021, 51(6): 125-130.
Xie Z X, Li M H. A survey on machine learning in recognition of 3D object based on point cloud[J]. Periodical of Ocean University of China, 2021, 51(6): 125-130.
- [13] Deng L, Yu D. Deep learning: Methods and applications[J]. Foundations and Trends in Signal Processing, 2014, 7(3-4): 197-387.
- [14] Zhou L, Sun G, Li Y, et al. Point cloud denoising review: from classical to deep learning-based approaches[J]. Graphical Models, 2022, 121: 101140.
- [15] 张卉冉,董震,杨必胜,等.点云压缩研究进展与趋势[J].武汉大学学报(信息科学版), 2023, 48(2): 192-205.
Zhang H R, Dong Z, Yang B S, et al. Progress and perspectives of point cloud compression[J]. Geomatics and Information Science of Wuhan University, 2023, 48(2): 192-205.
- [16] 毕勇,潘鸣奇,张硕,等.三维点云数据超分辨率技术[J].中国光学, 2022, 15(2): 210-223.
Bi Y, Pan M Q, Zhang S, et al. Overview of 3D point cloud super-resolution technology[J]. Chinese Optics, 2022, 15(2): 210-223.
- [17] 刘彩霞,魏明强,郭延文.基于深度学习的三维点云修复技术综述[J].计算机辅助设计与图形学学报, 2021, 33(12): 1936-1952.
Liu C X, Wei M Q, Guo Y W. 3D point cloud restoration via deep learning: A comprehensive survey[J]. Journal of Computer-Aided Design & Computer Graphics, 2021, 33(12): 1936-1952.
- [18] Roveri R, Öztireli A C, Pandele I, et al. PointProNets: Consolidation of point clouds with convolutional neural networks[J]. Computer Graphics Forum, Delft, Netherlands, 2018, 37(2): 87-99.
- [19] Pistilli F, Fracastoro G, Valsesia D, et al. Learning robust graph-convolutional representations for point cloud denoising[J]. IEEE Journal of Selected Topics in Signal Processing, 2020, 15(2): 402-414.
- [20] Chen Z, Li P, Wei Z, et al. GeoGCN: Geometric dual-domain graph convolution network for point cloud denoising[C]. ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, 2023: 1-5.
- [21] Lu D, Lu X, Sun Y, et al. Deep feature-preserving normal estimation for point cloud filtering[J]. Computer-Aided Design, 2020, 125: 102860.
- [22] Jang J W, Kwon Y C, Lim H, et al. CNN-based denoising, completion, and prediction of whole-body human-depth images[J]. IEEE Access, 2019, 7: 175842-175856.
- [23] Heinzler R, Piewak F, Schindler P, et al. CNN-based LiDAR point cloud de-noising in adverse weather[J]. IEEE Robotics and Automation Letters, 2020, 5(2): 2514-2521.
- [24] Seppänen A, Ojala R, Tammi K. 4DenoiseNet: Adverse weather denoising from adjacent point clouds[J]. IEEE Robotics and Automation Letters, 2022, 8(1): 456-463.
- [25] Zhou H, Chen K, Zhang W, et al. DUP-Net: Denoiser and upsampler network for 3D adversarial point clouds defense[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea (South), 2019: 1961-1970.
- [26] Liu X, Liu X, Liu Y S, et al. SPU-Net: Self-supervised point cloud upsampling by coarse-to-fine reconstruction with self-projection optimization[J]. IEEE Transactions on Image

Processing, 2022, 31: 4213-4226.

- [27] Qian Y, Hou J, Kwong S, et al. PUGeo-Net: A geometry-centric network for 3D point cloud upsampling[C]. Proceedings of the European Conference on Computer Vision, Glasgow, UK, 2020: 752-769.
- [28] Qian Y, Hou J, Kwong S, et al. Deep magnification-flexible upsampling over 3D point clouds[J]. IEEE Transactions on Image Processing, 2021, 30: 8354-8367.
- [29] Long C, Zhang W X, Li R, et al. PC2-PU: Patch correlation and point correlation for effective point cloud upsampling[C]. Proceedings of the 30th ACM International Conference on Multimedia, Lisboa Portugal, 2022: 2191-2201.
- [30] Han B, Zhang X, Ren S. PU-GACNet: Graph attention convolution network for point cloud upsampling[J]. Image and Vision Computing, 2022, 118: 104371.
- [31] Yu L, Li X, Fu C W, et al. EC-Net: An edge-aware point set consolidation network[C]. Proceedings of the European Conference on Computer Vision, Munich, Germany, 2018: 398-414.
- [32] Duan C, Chen S, Kovacevic J. 3D point cloud denoising via deep neural network based local surface estimation[C]. ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 2019: 8553-8557.
- [33] Hermosilla P, Ritschel T, Ropinski T. Total denoising: Unsupervised learning of 3D point cloud cleaning[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea (South), 2019: 52-60.
- [34] Guerrero P, Kleiman Y, Ovsjanikov M, et al. PCPNet learning local shape properties from raw point clouds[C]. Computer Graphics Forum. 2018, 37(2): 75-85.
- [35] Rakotosaona M J, La Barbera V, Guerrero P, et al. PointCleanNet: Learning to denoise and remove outliers from dense point clouds[C]. Computer Graphics Forum. 2020, 39(1): 185-203.
- [36] Zhang D, Lu X, Qin H, et al. Pointfilter: Point cloud filtering via encoder-decoder modeling[J]. IEEE Transactions on Visualization and Computer Graphics, 2020, 27(3): 2015-2027.
- [37] Chen S, Wang J, Pan W, et al. Towards uniform point distribution in feature-preserving point cloud filtering[J], 2022, arXiv:2201.01503.
- [38] Wang X, Fan X, Zhao D. PointFilterNet: A filtering network for point cloud denoising[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(3): 1276-1290.
- [39] Luo S, Hu W. Score-based point cloud denoising[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 2021: 4583-4592.
- [40] Chen H, Luo S, Hu W. Deep point set resampling via gradient fields[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022: 1-17.
- [41] Zhao Y, Zheng H, Wang Z, et al. Point cloud denoising via momentum ascent in gradient fields[J], 2022, arXiv:2202.10094.
- [42] Xu T X, Guo Y C, Yang Y L, et al. Gradient-based point cloud denoising with uniformity[J], 2022, arXiv:2207.10279.
- [43] 徐小微. 基于深度学习的三维点云去噪算法研究[D]. 长春: 吉林大学, 2022: 15-22.
Xu X W. Research on 3D point cloud denoising algorithm based on deep learning[D]. Changchun: Jilin University, 2022: 15-22.
- [44] Zhu X, Ma S, Chen D, et al. MSaD-Net: A mix self-attention networks for 3D point cloud denoising[J]. IEEE Photonics Journal, 2023, 15(3): 1-7.
- [45] Chen H, Wei Z, Li X, et al. RePCD-Net: Feature-aware recurrent point cloud denoising

- network[J]. *International Journal of Computer Vision*, 2022, 130(3): 615-629.
- [46] Luo S, Hu W. Differentiable manifold reconstruction for point cloud denoising[C]. *Proceedings of the 28th ACM International Conference on Multimedia*, Seattle, WA, USA, 2020: 1330-1338.
 - [47] Mao A, Du Z, Wen Y H, et al. PD-Flow: A point cloud denoising framework with normalizing flows[J], 2022, arXiv:2203.05940.
 - [48] Regaya Y, Fadli F, Amira A. Point-Denoise: Unsupervised outlier detection for 3D point clouds enhancement[J]. *Multimedia Tools and Applications*, 2021, 80(18): 28161-28177.
 - [49] Hou G, Qin G, Sun M, et al. NoiseTrans: Point cloud denoising with Transformers[J], 2023, arXiv:2304.11812.
 - [50] Huang L, Wang S, Wong K, et al. OctSqueeze: Octree-structured entropy model for LiDAR compression[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, 2020: 1313-1323.
 - [51] Fu C, Li G, Song R, et al. OctAttention: Octree-based large-scale contexts model for point cloud compression[C]. *Proceedings of the AAAI Conference on Artificial Intelligence*, Vancouver, Canada, 2022, 36(1): 625-633.
 - [52] Que Z, Lu G, Xu D. VoxelContext-Net: An octree based framework for point cloud compression[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Nashville, TN, USA, 2021: 6042-6051.
 - [53] Nguyen D T, Quach M, Valenzise G, et al. Learning-based lossless compression of 3D point cloud geometry[C]. *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, ON, Canada, 2021: 4220-4224.
 - [54] Nguyen D T, Quach M, Valenzise G, et al. Multiscale deep context modeling for lossless point cloud geometry compression[C]. *2021 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, Shenzhen, China, 2021: 1-6.
 - [55] Nguyen D T, Kaup A. Lossless point cloud geometry and attribute compression using a learned conditional probability model[J]. 2023, arXiv:2303.06519.
 - [56] Wang J, Ding D, Ma Z. Lossless point cloud attribute compression using cross-scale, cross-group, and cross-color prediction[C]. *2023 Data Compression Conference (DCC)*, Snowbird, UT, USA, 2023: 228-237.
 - [57] Quach M, Valenzise G, Dufaux F. Improved deep point cloud geometry compression[C]. *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP)*, Tampere, Finland, 2020: 1-6.
 - [58] Wang J, Ding D, Li Z, et al. Multiscale point cloud geometry compression[C]. *2021 Data Compression Conference (DCC)*, Snowbird, UT, USA, 2021: 73-82.
 - [59] Wang J, Zhu H, Ma Z, et al. Learned point cloud geometry compression[J], 2019, arXiv:1909.12037.
 - [60] Yan W, Liu S, Li T H, et al. Deep autoencoder-based lossy geometry compression for point clouds[J], 2019, arXiv:1905.03691.
 - [61] Quach M, Valenzise G, Dufaux F. Learning convolutional transforms for lossy point cloud geometry compression[C]. *2019 IEEE International Conference on Image Processing (ICIP)*, Taipei, Taiwan, 2019: 4320-4324.
 - [62] Guarda A F R, Rodrigues N M M, Pereira F. Point cloud coding: Adopting a deep learning-based approach[C]. *2019 Picture Coding Symposium (PCS)*, Ningbo, China, 2019: 1-5.

- [63] Huang T, Liu Y. 3D point cloud geometry compression on deep learning[C]. Proceedings of the 27th ACM International Conference on Multimedia, Nice, France, 2019: 890-898.
- [64] He Y, Ren X, Tang D, et al. Density-preserving deep point cloud compression[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 2022: 2333-2342.
- [65] Wen X, Wang X, Hou J, et al. Lossy geometry compression of 3D point cloud data via an adaptive octree-guided network[C]. 2020 IEEE International Conference on Multimedia and Expo (ICME), London, UK, 2020: 1-6.
- [66] Milani S. A syndrome-based autoencoder for point cloud geometry compression[C]. 2020 IEEE International Conference on Image Processing (ICIP), Abu Dhabi, United Arab Emirates, 2020: 2686-2690.
- [67] Feng Y, Liu S, Zhu Y. Real-time spatio-temporal LiDAR point cloud compression[C]. 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 2020: 10766-10773.
- [68] Sheng X, Li L, Liu D, et al. Deep-PCAC: An end-to-end deep lossy compression framework for point cloud attributes[J]. IEEE Transactions on Multimedia, 2021, 24: 2617-2632.
- [69] Quach M, Valenzise G, Dufaux F. Folding-based compression of point cloud attributes[C]. 2020 IEEE International Conference on Image Processing (ICIP), Abu Dhabi, United Arab Emirates, 2020: 3309-3313.
- [70] Zhao L, Ma K K, Lin X, et al. Real-time LiDAR point cloud compression using bi-directional prediction and range-adaptive floating-point coding[J]. IEEE Transactions on Broadcasting, 2022, 68(3): 620-635.
- [71] 赵志,马燕新,许可,等.一种优化的可拓展激光雷达点云可学习二值量化网络[J].光学学报, 2022, 42(12): 1212005.
Zhao Z, Ma Y X, Xu K, et al. Optimized scalable and learnable binary quantization network for LiDAR point cloud[J]. Acta Optica Sinica, 2022, 42(12): 1212005.
- [72] Wang J, Ding D, Li Z, et al. Sparse tensor-based multiscale representation for point cloud geometry compression[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022: 1-18.
- [73] Liang Z, Liang F. TransPCC: Towards deep point cloud compression via Transformers[C]. Proceedings of the 2022 International Conference on Multimedia Retrieval, Newark, USA, 2022: 1-5.
- [74] Liu G, Wang J, Ding D, et al. PCGFormer: Lossy point cloud geometry compression via local self-attention[C]. 2022 IEEE International Conference on Visual Communications and Image Processing (VCIP), Suzhou, China, 2022: 1-5.
- [75] 刘舸昕,章骏腾,丁丹丹.基于 Transformer 的点云几何有损压缩方法[J/OL].北京航空航天大学学报: 1-10[2023-02-21].
Liu G X, Zhang J T, Ding D D. Lossy point cloud geometry compression based on Transformer[J/OL]. Journal of Beijing University of Aeronautics and Astronautics: 1-10[2023-02-21]. <https://doi.org/10.13700/j.bh.1001-5965.2022.0412>.
- [76] Yu L, Li X, Fu C W, et al. PU-Net: Point cloud upsampling network[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018: 2790-2799.
- [77] Meng H, Han Y, Chen J, et al. A point cloud density enhancement method based on super-

- resolution convolutional neural network[C]. 2019 IEEE International Conference on Unmanned Systems and Artificial Intelligence (ICUSAI), Xi'an, China, 2019: 8-12.
- [78] Yue J, Wen W, Han J, et al. 3D point clouds data super resolution-aided LiDAR odometry for vehicular positioning in urban canyons[J]. IEEE Transactions on Vehicular Technology, 2021, 70(5): 4098-4112.
- [79] Wang G, Xu G, Wu Q, et al. Two-stage point cloud super resolution with local interpolation and readjustment via outer-product neural network[J]. Journal of Systems Science and Complexity, 2021, 34(1): 68-82.
- [80] Chen K, Liu C, Ou Y. Channel attention based network for LiDAR super-resolution[C]. 2021 China Automation Congress (CAC), Beijing, China, 2021: 5458-5463.
- [81] Matsuzaki K, Komorita S. Efficient deep super-resolution of voxelized point cloud in geometry compression[J]. IEEE Sensors Journal, 2023, 23(2): 1328-1342.
- [82] Qian G, Abualshour A, Li G, et al. PU-GCN: Point cloud upsampling using graph convolutional networks[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 2021: 11683-11692.
- [83] 钟帆, 柏正尧. 采用动态残差图卷积的 3D 点云超分辨率[J]. 浙江大学学报(工学版), 2022, 56(11): 2251-2259.
- Zhong F, Bai Z Y. 3D point cloud super-resolution with dynamic residual graph convolutional networks[J]. Journal of Zhejiang University (Engineering Science), 2022, 56(11): 2251-2259.
- [84] Zhou Y, Zhang X, Wang S, et al. Multi-Attribute Joint Point Cloud Super-Resolution with Adversarial Feature Graph Networks[C]. 2022 IEEE International Conference on Multimedia and Expo Workshops (ICMEW), Taipei City, Taiwan, 2022: 1-6.
- [85] Li R, Li X, Fu C W, et al. PU-GAN: A point cloud upsampling adversarial network[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea (South). 2019: 7203-7212.
- [86] Wu H, Zhang J, Huang K. Point cloud super resolution with adversarial residual graph networks[J], 2019, arXiv:1908.02111.
- [87] Zhou K, Dong M, Arslanturk S. "Zero-Shot" point cloud upsampling[C]. 2022 IEEE International Conference on Multimedia and Expo (ICME), Taipei, Taiwan, 2022: 1-6.
- [88] Wang Y, Wu S, Huang H, et al. Patch-based progressive 3D point set upsampling[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 2019: 5958-5967.
- [89] Li J, Zhu F, Yang X, et al. 3D face point cloud super-resolution network[C]. 2021 IEEE International Joint Conference on Biometrics (IJCB), Shenzhen, China, 2021: 1-8.
- [90] Qiu S, Anwar S, Barnes N. PU-Transformer: Point cloud upsampling Transformer[C]. Proceedings of the Asian Conference on Computer Vision, Macau, China, 2022: 2475-2493.
- [91] Wang S, Wang H, She S, et al. Swin-T-NFC CRFs: An encoder-decoder neural model for high-precision UAV positioning via point cloud super resolution and image semantic segmentation[J]. Computer Communications, 2023, 197: 52-60.
- [92] 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, Honolulu, HI, USA, 2017: 77-85.
- [93] 李宇翔, 邓慧萍, 向森, 等. 纹理边缘引导的深度图像超分辨率重建[J]. 中国图象图形学报, 2018, 23(10): 1508-1517.

- Li Y X, Deng H P, Xiang Sen, et al. Depth map super-resolution reconstruction based on the texture edge-guided approach[J]. *Journal of Image and Graphics*, 2018, 23(10): 1508-1517.
- [94] Nguyen A D, Choi S, Kim W, et al. Graph X-convolution for point cloud deformation in 2D-to-3D conversion[C]. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Seoul, Korea (South), 2019: 8628-8637.
- [95] Choi S, Nguyen A D, Kim J, et al. Point cloud deformation for single image 3D reconstruction[C]. *2019 IEEE International Conference on Image Processing (ICIP)*, Taipei, Taiwan, 2019: 2379-2383.
- [96] 郭闯,戴宁,田素坤,等.高分辨率深度生成网络的缺失牙体形态设计[J].*中国图象图形学报*, 2020, 25(10): 2249-2258.
- Guo C, Dai N, Tian S K, et al. Morphological design of missing tooth driven by high-resolution deep generation network[J]. *Journal of Image and Graphics*, 2020, 25(10): 2249-2258.
- [97] Zhang W, Jiang H, Yang Z, et al. Data-driven upsampling of point clouds[J]. *Computer-Aided Design*, 2019, 112: 1-13.
- [98] Sharma R, Schwandt T, Kunert C, et al. Point cloud upsampling and normal estimation using deep learning for robust surface reconstruction[J], 2021, arXiv:2102.13391.
- [99] 王思启,张家强,李丽圆,等.MVSNet 在空间目标三维重建中的应用[J].*中国激光*, 2022, 49(23): 2310003.
- Wang S Q, Zhang J Q, Li L Y, et al. Application of MVSNet in 3D reconstruction of space objects[J]. *Chinese Journal of Lasers*, 2022, 49(23): 2310003.
- [100] Yuan W, Khot T, Held D, et al. PCN: Point completion network[C]. *2018 International Conference on 3D Vision (3DV)*, Verona, Italy, 2018: 728-737.
- [101] Tchapmi L P, Kosaraju V, Rezaeifighi H, et al. TopNet: Structural point cloud decoder[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, 2019: 383-392.
- [102] Huang H, Chen H, Li J. Deep neural network for 3D point cloud completion with multistage loss function[C]. *2019 Chinese Control and Decision Conference (CCDC)*, Nanchang, China, 2019: 4604-4609.
- [103] Huang Z, Yu Y, Xu J, et al. PF-Net: Point fractal network for 3D point cloud completion[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, 2020: 7662-7670.
- [104] Wang X, Ang Jr M H, Lee G H. Cascaded refinement network for point cloud completion[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, 2020: 790-799.
- [105] Wen X, Li T, Han Z, et al. Point cloud completion by skip-attention network with hierarchical folding[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, 2020: 1939-1948.
- [106] Long J, Zhu Q, He H, et al. 3D point cloud completion based on multi-scale degradation[C]. *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, Greece, 2023: 1-5.
- [107] Wu H, Miao Y. Cross-regional attention network for point cloud completion[C]. *2020 25th International Conference on Pattern Recognition (ICPR)*, Milan, Italy, 2021: 10274-10280.
- [108] Li J, Guo S, Meng X, et al. DPG-Net: Densely progressive-growing network for point cloud completion[J]. *Neurocomputing*, 2022, 491: 1-13.

- [109]Liu X, Ma Y, Xu K, et al. AGFA-Net: Adaptive global feature augmentation network for point cloud completion[J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 1-5.
- [110]刘心溥,马燕新,许可,等.嵌入 Transformer 结构的多尺度点云补全[J].中国图象图形学报,2022,27(2):538-549.
- Liu X P, Ma Y X, Xu K, et al. Multi-scale Transformer based point cloud completion network[J]. Journal of Image and Graphics, 2022, 27(2): 538-549.
- [111]Tian H, Qin Z, Yi R, et al. Tensorformer: Normalized matrix attention Transformer for high-quality point cloud reconstruction[J]. IEEE Transactions on Multimedia, 2023. doi: 10.1109/TMM.2023.3277271.
- [112]Yan Y, Wang Z, Xu C, et al. GEOP-Net: Shape reconstruction of buildings from LiDAR point clouds[J]. IEEE Geoscience and Remote Sensing Letters, 2023. doi: 10.1109/LGRS.2023.3277717.
- [113]Geiger A, Lenz P, Urtasun R. Are we ready for autonomous driving? the KITTI vision benchmark suite[C]. 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, USA, 2012: 3354-3361.
- [114]Serna A, Marcotegui B, Goulette F, et al. Paris-rue-Madame database: a 3D mobile laser scanner dataset for benchmarking urban detection, segmentation and classification methods[C]. Proceedings of the 4th International Conference on Pattern Recognition, Applications and Methods, Angers, France, 2014: 1-6.
- [115]Lian Z, Zhang J, Choi S, et al. Non-rigid 3D shape retrieval[C]. Proceedings of the Eurographics Workshop on 3D Object Retrieval, Zurich, Switzerland, 2015: 107-120.
- [116]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, Boston, MA, USA, 2015: 1912-1920.
- [117]Chang A X, Funkhouser T, Guibas L, et al. ShapeNet: An information-rich 3D model repository[J], 2015, arXiv:1512.03012.
- [118]Gaidon A, Wang Q, Cabon Y, et al. Virtual worlds as proxy for multi-object tracking analysis[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 2016: 4340-4349.
- [119]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.
- [120]Armeni I, Sener O, Zamir A R, et al. 3D semantic parsing of large-scale indoor spaces[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 2016: 1534-1543.
- [121]Zeng A, Song S, Nießner M, et al. 3DMatch: Learning local geometric descriptors from RGB-D reconstructions[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 2017: 1802-1811.
- [122]Dai A, Chang A X, Savva M, et al. ScanNet: Richly-annotated 3D reconstructions of indoor scenes[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 2017: 5828-5839.
- [123]Chang A, Dai A, Funkhouser T, et al. Matterport3D: Learning from RGB-D data in indoor environments[C]. 2017 International Conference on 3D Vision (3DV), Qingdao, China, 2017: 667-676.
- [124]Behley J, Garbade M, Milioto A, et al. SemanticKITTI: A dataset for semantic scene

- understanding of LiDAR sequences[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea (South), 2019: 9297-9307.
- [125]Schwarz S, Preda M, Baroncini V, et al. Emerging MPEG standards for point cloud compression[J]. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2019, 9(1): 133-148.
- [126]Caesar H, Bankiti V, Lang A H, et al. nuScenes: A multimodal dataset for autonomous driving[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 2020: 11621-11631.
- [127]Sun P, Kretzschmar H, Dotiwalla X, et al. Scalability in perception for autonomous driving: Waymo open dataset[C]. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, Seattle, WA, USA, 2020: 2446-2454.
- [128]Graziosi D, Nakagami O, Kuma S, et al. An overview of ongoing point cloud compression standardization activities: video-based (V-PCC) and geometry-based (G-PCC)[J]. APSIPA Transactions on Signal and Information Processing, 2020, 9: 1-17.

Research Progress of 3D Point Cloud Processing Methods Based on Deep Learning

Wu Yiquan^{*}, Chen Huixian, Zhang Yao

College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, Jiangsu, China

Abstract

Significance With the continuous development of three-dimensional (3D) acquisition equipment like 3D Lidar, 3D point cloud data has recently become more accurate and easier to obtain. As the most important representation of 3D data, 3D point cloud is widely used in vision tasks in the fields of autonomous driving, robotics, remote sensing, cultural relic restoration, augmented reality, virtual reality, and so on. Due to the huge amount of original point cloud data and the fact that the acquisition process is easily mixed with noise and outliers, the effect of directly using the original point cloud data is not good. Therefore, it is of great practical significance to study the processing methods of 3D point cloud.

A point cloud is a collection of spatial sampling points of the target surface properties in the same coordinate system. Sampling points contain geometric information such as 3D coordinates and size, as well as attribute information such as object color and texture features. Traditional 3D point cloud processing methods are based on the principle of geometric analysis. Point cloud data is processed by estimating geometric information such as normal vector, curvature, and density of the point cloud and combining traditional feature descriptors. Although the accuracy is high, it is not suitable for complex point cloud scenes such as large rotations, and the function calculation is extremely cumbersome. Classical machine learning methods can process 3D data and learn effective feature information, but machine learning is highly dependent on accurate manual identification features. Massive 3D data not only increases the number of manual labels but also makes labeling much more difficult than two-dimensional (2D) images. The deep learning methods can train and calculate large-scale data, autonomously learn the latent space features and advanced laws in the input information, and are suitable for processing massive point cloud data. Although the deep learning methods spend a lot of time training samples to learn parameter information, the test time is much less than that of machine learning methods, and the prediction results are more accurate. In view of the irregular, sparse, and uneven internal structure of 3D point cloud, how to efficiently implement 3D point cloud processing based on deep learning has become the focus of researchers in recent years. Therefore, this paper reviews the research progress of deep learning-based 3D point cloud processing methods in the past six years and looks forward to future research trends, hoping to provide inspiration and ideas for researchers in point cloud processing.

Progress In this paper, we focus on deep learning-based 3D point cloud processing tasks, and give the development route of the most commonly used deep learning methods for four point cloud processing tasks in the past six years (Fig.1). The 3D point cloud processing tasks mainly include the following four types: point cloud denoising and filtering, point cloud compression, point cloud super-resolution, point cloud restoration, completion and reconstruction.

Deep learning methods for point cloud denoising and filtering task can be divided into five types: CNN-based, upsampling-based, filter-based, gradient-based, and other methods. PointProNet and GeoGCN learn feature differences based on convolutional networks to remove noise, but point cloud information is lost in the preprocessing stage. DUP-Net, PUGeo-Net, and PU-GACNet are classic upsampling-based denoising methods that denoise by modifying the feature extractor and feature expander but ignore some local features. NPD and PointCleanNet combine filtering ideas with deep learning and can simultaneously achieve noise removal and point cloud geometric feature retention. The Score-based method constructs the gradient field according to the distribution characteristics of the noise point cloud, and the robustness is enhanced, but there are still relatively few studies. NoiseTrans draws on the idea of Transformer to realize the effective extraction and retention of fine features in point clouds. Table 2 makes a comparison of the advantages and disadvantages of common methods.

Then, deep learning methods for point cloud compression task are generalized. According to lossless and lossy compression, they are divided into two categories and analyzed (Table 3 and Table 4). Point cloud lossless compression methods OctSqueeze and VoxelDNN improve the accuracy of point cloud probability prediction, but part of the point cloud information will be lost. PCGCv2, TransPCC and SparsePCGC are typical point cloud lossy compression methods. The point cloud feature is learned through the network structure, which avoids the loss of detailed information and improves the quality of the reconstructed point cloud.

Subsequently, deep learning methods for the point cloud super-resolution task are outlined. The classification and comparative analysis are carried out in four parts: Convolutional Neural Network (CNN), Graph Convolutional Neural Network (GCN), Generative Adversarial Network (GAN) and other structures (Table 5). PU-Net and PU-GCN extract rich detailed features based on CNN and GCN respectively, but the amount of calculation is large. PU-GAN exploits the dynamic adversarial optimization details of the generator and discriminator. MPU and PU-Transformer combine the idea of 2D super-resolution algorithm with the PointNet structure, which is a new idea worth trying.

The deep learning methods for point cloud restoration, completion and reconstruction task include three aspects: image-based, sampling-based and completion-based, and a comparative analysis is carried out (Table 6). PCDNet reduces computation by extracting 2D image features and deformations. Sampling-based methods utilize networks to generate dense and complete point cloud. PCN, TopNet, and SA-Net are able to fill in missing structures with input point cloud, but completion-based methods are susceptible to the structure of incomplete point cloud.

In recent years, KITTI, PCN, nuScenes and other public point cloud datasets and performance indicators such as CD, P2M, and RMSE have greatly promoted the in-depth research of point cloud processing tasks (Table 7 and Table 8).

Conclusions and Prospects 3D point cloud processing methods based on deep learning have gradually become an important research direction in the field of computer vision. Although many good achievements have already been attained, there is still much room for development. The following aspects may be taken into account while conducting in-depth research: the combination of multiple processing tasks, point cloud data feature processing, low-cost network models and hardware devices, and adaptable datasets.

Key words 3D point cloud processing; deep learning; point cloud denoising; point cloud compression; point cloud restoration

网络首发:

标题: 基于深度学习的三维点云处理方法研究进展

作者: 吴一全, 陈慧娴, 张耀

收稿日期: 2023-06-19

录用日期: 2023-08-11

DOI: 10.3788/CJL230924

引用格式:

吴一全, 陈慧娴, 张耀. 基于深度学习的三维点云处理方法研究进展[J]. 中国激光, 2024, 51(05): 0500002.

网络首发文章内容与正式出版的有细微差别, 请以正式出版文件为准!

您感兴趣的其他相关论文:

基于金字塔长程Transformer的OCT图像超分辨率重建

芦焱琦 陈明惠 秦楷博 吴玉全 尹志杰 杨政奇

上海理工大学健康科学与工程学院, 上海介入医疗器械工程技术研究中心, 教育部医学光学工程中心, 上海 200093

中国激光, 2023, 50(15): 1507107

利用深度学习扩展双光子成像视场

李迟件 姚靖 高玉峰 赖溥祥 何悦之 齐苏敏 郑炜

曲阜师范大学网络空间安全学院, 山东 济宁 273100

中国激光, 2023, 50(9): 0907107

基于三维点云处理的交联聚乙烯电缆接头反应力锥参数测量算法

邓磊 刘桂华 邓豪 黄军杰 周炳宏

西南科技大学信息工程学院, 四川 绵阳 621010

中国激光, 2023, 50(6): 0604003

基于AlexNet的近岸水生植物光谱分类及特征光谱分析

郑宗生 刘贝 卢鹏 王振华 邹国良 赵家惠 李云飞

上海海洋大学信息学院, 上海 201306

中国激光, 2023, 50(2): 0211001

快速宽场三维显微技术研究进展

任婧荣 傅相达 王孟瑞 赵天宇 汪召军 冯坤 梁言生 王少伟 雷铭

教育部物质非平衡合成与调控重点实验室, 陕西 西安 710049

中国激光, 2023, 50(3): 0307104