

**本科毕业设计（论文）**

**基于深度学习的点云补全算法研究**

|  |  |
| --- | --- |
| 学院（部、中心）： | 软件学院 |
| 专 业： | 软件工程 |
| 班 级： | 2104 |
| 学生姓名： | 余华威 |
| 学 号： | 2214414307 |
| 指导教师： | 赵加坤 |

2025年06月

摘 要

随着三维感知技术的迅速发展，点云作为三维几何信息的核心载体，在自动驾驶、工业检测、虚拟现实等领域展现出重要的应用价值。然而，受限于传感器精度、遮挡等因素，实际采集的点云数据常常存在缺失、稀疏及噪声等问题，严重影响了下游任务的可靠性。点云补全技术旨在从残缺观测中恢复完整且高保真的三维形状，是三维视觉领域的关键研究方向。传统方法多依赖几何插值或模板匹配，难以应对复杂缺失场景；而基于深度学习的方法通过端到端特征学习实现语义驱动的形状生成，但在极稀疏场景下仍面临几何信息不足和局部细节丢失等挑战。

针对极稀疏场景下的点云补全任务，本文提出了一种基于 Transformer 的双路特征增强的渐进式补全框架。主要创新点包括以下几个方面。

1. 本文提出了一种全新的点云补全网络模型，采用分层式Transformer编码器-解码器架构，将补全任务转化为集合到集合的转换问题，并通过自注意力与交叉注意力机制，有效提取点代理序列的上下文语义信息。
2. 设计了双路并行特征提取模
3. 块，结合局部特征分支与全局特征分支，实现局部细节与全局语义的平衡融合，通过交叉注意力实现局部几何细节与全局形状先验的自适应融合，显著提升了特征表达能力。
4. 提出多阶段渐进式细化网络，优化点云从粗到精的重建流程，从而实现高保真的点云补全。

为了验证所提出三维点云补全算法的有效性，本文在ShapeNet-55数据集上进行了系统实验。实验结果表明，本文提出的点云补全模型在多个主流评估指标上均优于现有方法，所生成的点云不仅表面分布更加均匀，局部结构更加精细，整体形状也更加接近真实物体。此外，通过消融实验进一步验证了各创新模块的有效性，充分证明了本文方法在提升点云补全质量方面的显著作用和实际应用价值。

**关 键 词**：点云补全；极稀疏点云；双路特征增强；多阶段点云细化

ABSTRACT

With the rapid development of 3D perception technology, point clouds, as a core carrier of 3D geometric information, have demonstrated significant application value in fields such as autonomous driving, industrial inspection, and virtual reality. However, due to limitations of sensor accuracy and occlusion, the collected point cloud data often suffer from incompleteness, sparsity, and noise, which severely affect the reliability of downstream tasks. Point cloud completion aims to recover complete and high-fidelity 3D shapes from partial observations, making it a key research direction in 3D vision. Traditional methods mainly rely on geometric interpolation or template matching, which are inadequate for handling complex missing scenarios. Deep learning-based approaches achieve semantic-driven shape generation through end-to-end feature learning, but still face challenges such as insufficient geometric information and loss of local details in extremely sparse cases.

To address point cloud completion in extremely sparse scenarios, this paper proposes a progressive completion framework based on Transformer with dual-branch feature enhancement. The main innovations are as follows:

1. A novel point cloud completion network is proposed, which adopts a hierarchical Transformer encoder-decoder architecture to formulate the completion task as a set-to-set transformation problem. Contextual semantic information of point proxies is effectively extracted through self-attention and cross-attention mechanisms.
2. A dual-branch parallel feature extraction module is designed, combining local and global feature branches to achieve balanced fusion of local details and global semantics. Cross-attention is used to adaptively integrate local geometric details and global shape priors, significantly enhancing feature representation capability.
3. A multi-stage progressive refinement network is introduced to optimize the reconstruction process from coarse to fine, thereby achieving high-fidelity point cloud completion.

To validate the effectiveness of the proposed 3D point cloud completion algorithm, extensive experiments are conducted on the ShapeNet-55 dataset. Experimental results show that the proposed model outperforms existing methods on multiple mainstream evaluation metrics. The generated point clouds exhibit more uniform surface distribution, finer local structures, and more accurate overall shapes. Furthermore, ablation studies confirm the effectiveness of each innovative module, fully demonstrating the significant improvements and practical value of the proposed method in point cloud completion quality.

**KEY WORDS**: point cloud completion; extremely sparse point cloud; dual-branch feature enhancement; multi-stage point cloud refinement

目 录

摘 要.............................................................................................................................................................I

ABSTRACT................................................................................................................................................II

[1 绪论 1](#_Toc153986581)

[1.1 课题背景及研究意义](#_Toc153986582) **[错误！未定义书签。](#_Toc153986582)**

[1.2 国内外研究现状 3](#_Toc153986582)

[1.2.1 传统三维补全算法标题 3](#_Toc153986583)

[1.2.2 基于深度学习的点云补全模型 5](#_Toc153986583)

[2 基础理论 7](#_Toc153986584)

[2.1 点云数据基础 7](#_Toc153986585)

[2.1.1 定义 7](#_Toc153986583)

[2.1.2 核心特性 7](#_Toc153986583)

[2.1.3 点云数据获取与表示 7](#_Toc153986583)

[2.1.4 采样方法 9](#_Toc153986583)

[2.1.5 稀疏理论 1](#_Toc153986583)0

[2.2 点云质量评估指标 1](#_Toc153986585)1

[2.2.1 倒角距离 1](#_Toc153986583)1

[2.2.2 推土距离 1](#_Toc153986583)2

[2.2.3 稀疏场景下评估指标的适应性 1](#_Toc153986583)2

[2.3 点云深度学习方法基础 1](#_Toc153986585)3

[2.3.1 基础网络架构 1](#_Toc153986583)3

[2.3.2 Transformer网络架构 1](#_Toc153986583)7

[2.3.3 动态图卷积网络 1](#_Toc153986583)8

[2.3.4 生成式模型 1](#_Toc153986583)9

[2.4 本章小结 2](#_Toc153986585)1

[3 面向极稀疏点云的双路特征增强渐进式补全框架 2](#_Toc153986610)2

[3.1 模型总体架构 2](#_Toc153986611)2

[3.2 点云数据预处理 2](#_Toc153986611)2

[3.2.1 稀疏点云生成策略 2](#_Toc153986612)2

[3.2.2 零点填充与不完整点云模拟 2](#_Toc153986612)3

[3.3 双路并行特征提取模块 2](#_Toc153986611)3

[3.3.1 局部特征分支设计 2](#_Toc153986612)3

[3.3.2 全局特征分支实现 2](#_Toc153986612)5

[3.3.3 特征融合与增强策略 2](#_Toc153986612)5

[3.4 Transformer编码-解码核心架构 2](#_Toc153986611)6

[3.4.1 自注意力编码器设计 2](#_Toc153986612)7

[3.4.2 动态Query生成器 2](#_Toc153986612)8

[3.4.3 基于形状先验的解码器 2](#_Toc153986612)8

[3.5 全局特征增强与融合模块 2](#_Toc153986611)9

[3.6 多阶段渐进式点云细化网络 2](#_Toc153986611)9

[3.7 损失函数设计 3](#_Toc153986611)0

[3.7.1 多阶段监督机制 3](#_Toc153986612)0

[3.8 本章小结 3](#_Toc153986611)1

[4 实验设计与结果分析 3](#_Toc153986613)2

[4.1 实验环境设计 3](#_Toc153986614)2

[4.1.1 数据集介绍 3](#_Toc153986615)2

[4.1.2 训练设置 3](#_Toc153986615)2

[4.2 模型对比 3](#_Toc153986614)3

[4.2.1 度量指标对比 3](#_Toc153986615)3

[4.2.2 可视化对比 3](#_Toc153986615)5

[4.3 本章小结 3](#_Toc153986611)5

[5 结论与展望 3](#_Toc153986637)6

[5.1 结论 3](#_Toc153986638)6

[5.2 展望 3](#_Toc153986638)6

[致 谢 3](#_Toc153986640)7

[参考文献 3](#_Toc153986641)8

[附 录 4](#_Toc153986642)0

# 绪论

## 课题背景及研究意义

随着计算机视觉领域的迅速扩充，三维感知技术也快速发展，那么3D形状就出现了各式各样的表示形式，其中包括点云、体素网格、多边形网格和多视图表示等，不同表示形式如图1-1所示，而点云因其能够精确描述物体三维几何特性并且表示简单易懂，已成为3D形状的核心数据载体，在自动驾驶、工业检测、虚拟现实（VR/AR）、文化遗产保护等领域展现出巨大应用潜力。根据行业分析，全球三维传感市场规模从2024年到2030年的复合年增长率即将达到10%，激光雷达、深度相机等设备的广泛应用催生了海量点云数据处理需求。然

而在实际应用中，由于传感器的物理特性、视角单一以及物体遮挡等问题，实际采集的点云普遍存在以下这些缺陷：数据不完整性、几何结构模糊及噪声干扰严重等。这些问题不仅影响了下游的三维重建和计算机视觉任务，还可能导致错误的环境感知和决策，严重制约了三维感知系统的可靠性。例如，在自动驾驶场景中，部分遮挡的点云数据可能导致障碍物检测的不准确，从而影响车辆的路径规划和安全决策。[2]在文化遗产保护领域，受损文物的点云数据不完整，可能影响文物的修复和数字化保护。[3]因此，研究点云补全具有重要的应用价值。

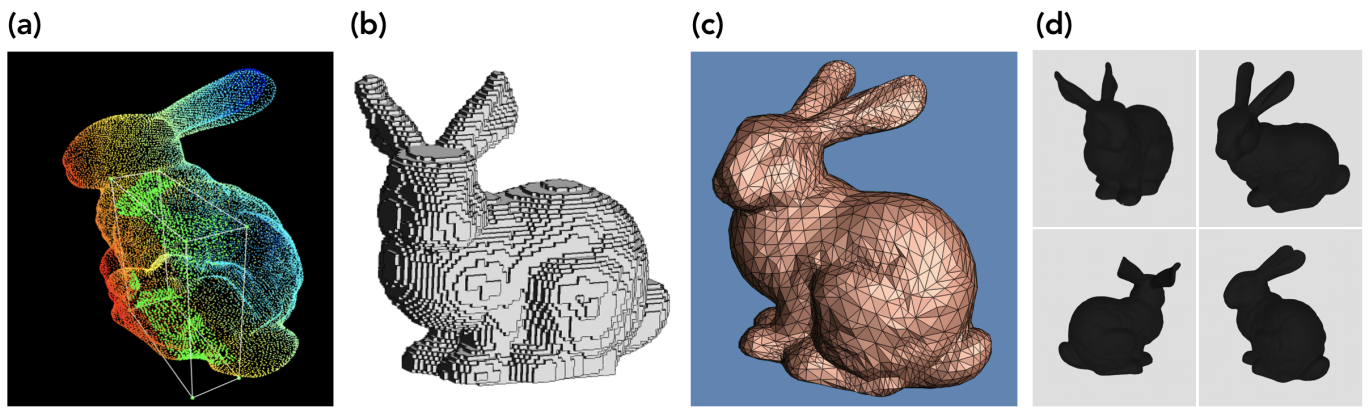


图 1-1 3D 数据的不同表示形式。（a）点云（来源：加州理工学院），（b）体素网格（来源：印度理工学院），（c）三角形网格（来源：华盛顿大学），（d）多视图表示（来源：斯坦福大学）[1]

点云补全技术作为三维视觉领域的核心研究方向，近年来随着深度学习技术的突破取得了显著进展。其目标是从残缺、稀疏或不完整的点云数据中恢复出完整且高保真的三维形状，这一过程不仅涉及几何推理与特征建模，更需要结合语义理解与物理约束。那么传统方法高度依赖几何插值或模板匹配，就难以处理复杂场景下的非结构化缺失，而基于深度学习的点云补全算法应运而生，通过端到端的特征学习，能够从局部观测中推理全局几何结构，实现语义驱动的形状生成。例如，在自动驾驶场景中，激光雷达采集的车辆周围点云常因动态物体遮挡导致关键区域缺失，先进的补全算法可准确重建被遮挡的行人或障碍物轮廓，为环境感知系统提供更完整的三维信息，从而提升决策系统的安全性与可靠性。这种从“被动修复”到“主动推理”的技术跨越，标志着三维视觉系统向更高层级的空间认知能力演进。

从产业应用维度看，点云补全技术正在重塑多个领域的生产模式与效率。在智能制造领域，高精度点云补全被广泛应用于工业质检与逆向工程。例如，航空发动机叶片的检测需通过三维扫描获取表面点云，但复杂曲面的自遮挡会导致数据缺失。基于层次化特征融合的补全算法能够从局部扫描数据中重建完整曲面，结合几何约束将补全精度提升至亚毫米级，使检测效率提高30%以上，大幅降低人工复检成本。在文化遗产保护领域，针对文物碎片的数字化修复需求，传统方法难以处理断裂面的复杂拓扑关系，而结合图卷积网络与注意力机制的补全算法可实现语义级形状重建。敦煌莫高窟风化壁画的数字化修复工程中，此类技术已成功还原破损浮雕的结构细节，为文化遗产的永续留存提供了创新解决方案。

研究点云补全算法，也驱动着三维视觉技术生态的完善。其核心挑战在于平衡几何细节保真度、算法计算效率与生成结果的物理合理性。现有方法虽已取得突破，但仍面临诸多瓶颈：一方面，点云的无序性与非结构化特征使传统卷积神经网络难以有效提取局部几何模式；另一方面，Transformer等自注意力机制虽能建模长距离依赖，但其平方级计算复杂度限制了在大规模点云场景中的应用。针对这些问题，学界提出了多尺度特征融合、轻量化注意力机制、微分渲染约束等创新方法。

在深度学习驱动下的点云补全技术发展历程中，研究人员对这个问题展开了深入的研究，研究范式经历了显著的演变轨迹。随着PointNet[4]的突破。摒弃了体素化的传统3D形状数据，深度学习直接处理无序点云成为可能，其通过最大池化实现置换不变性，而PointNet++[5]进一步引入层次化采样与局部特征聚合，奠定多尺度分析基础。该技术广泛应用于后继者研究点云补全任务中的许多开创性工作中，创新性地提出采用编码器-解码器架构来处理点云数据。

由于点云补全本质是从残缺观测中推理完整三维结构，这要求模型具备结构化的建模能力、几何一致性约束与全局形状感知能力。然而，点云的无序性迫使模型必须对输入点的排列顺序完全鲁棒（置换不变性），这与需要显式建模空间关系的结构化学习目标相冲突。同时，非结构化特性使得传统用于规则数据（如图像网格）的局部操作（如卷积）难以直接应用，导致几何特征提取效率低下。因此需要探索基于Transformer的点云补全机制，利用自注意力机制建模点云全局结构关系，突破传统方法对局部卷积核的依赖；还需要增强模型的几何感知能力，显式建模三维形状的几何先验，提升补全结果的细节保真度。

点云补全在三维视觉领域具有重要应用价值，然而现有方法在处理极稀疏点云时普遍存在性能下降问题。极稀疏场景下，点云几何信息严重不足，导致重建困难，特别是复杂形状的恢复尤为挑战。那么，在攻克这些挑战的道路上，研究人员投入了大量精力探索解决方案，但该领域仍存在诸多亟待突破的难点。点云补全技术的突破不仅将提升三维数据的完整性，更对自动驾驶、机器人感知、数字孪生等下游任务的高效处理与深度应用具有决定性意义。

## 国内外研究现状

三维视觉技术作为数字化时代的关键支撑，在智能驾驶、工业检测、机器人导航及数字孪生等领域展现出显著的应用价值，其技术突破直接推动了目标检测、语义分割与三维重建等核心任务的性能提升。点云作为三维几何表征的重要信息载体，通过离散采样点精确描述物体表面拓扑结构，已成为三维场景理解的基础数据形式。然而，受限于传感器分辨率、物体遮挡、镜面反射及动态环境干扰，原始点云常存在局部缺失、噪声冗余及密度不均等问题，严重制约了下游任务（如逆向工程重建、高精度地图生成）的准确性与鲁棒性。在此背景下，点云补全技术应运而生，其核心目标是通过算法重构缺失区域的几何拓扑与语义一致性，旨在提升点云数据的完整性、可用性及对复杂场景的适应性，进而为三维视觉系统的全链条处理提供高质量数据基础。当前点云补全研究围绕传统几何驱动方法与数据驱动的深度学习模型两大技术范式展开，其核心目标均是通过算法恢复缺失区域的几何与语义信息，提升点云数据的完整性与可用性。

### 传统三维补全算法

传统点云补全方法作为三维几何处理领域的重要分支，其技术框架建立在数学建模与经验规则基础之上，旨在通过几何推理与局部特征分析实现对缺失区域的结构化重建。此类方法的核心思想可概括为：基于物体内在的几何规律性假设，通过人工设计的特征提取与规则推导，恢复因遮挡、传感器局限或环境干扰导致的点云缺损。

传统方法主要围绕三大类方法展开：几何对称法、表面重建法、模板匹配法。几何对称性补全通过检测模型的对称轴或对称平面，利用镜像、旋转等变换将已知几何部分映射至缺失区域，并优化接合处的几何连续性以实现完整重建。Mitra等人[6]（2013）在三维几何对称性检测领域构建了首个系统性分类框架，通过多维度对比方法特征、输入输出形式及算法机理，揭示了不同检测范式的内在关联与适用边界，为跨领域研究建立了统一的理论框架；其阐明的对称性信息在形状匹配、修复等几何处理中的普适性应用范式，有效推动了算法研究与工程实践的深度融合。然而，该研究在非刚性对称建模、动态层次化检测等复杂场景的理论支撑仍显薄弱，对数据噪声与部分遮挡的鲁棒性分析不足，虽前瞻性指出这些开放问题的发展路径，但未能突破传统刚性变换的局限，为后续深度学习与拓扑分析方法的融合研究留下了创新空间。表面插值与拟合重建技术则基于邻域点集的分布特性，通过数学工具构建连续表面模型，其中插值法依赖局部点云密度生成平滑表面，但易因采样稀疏导致细节丢失或过度平滑，而拟合法通过参数化模型逼近全局几何结构，虽能提升曲面连续性，却面临初始参数敏感性与高维优化计算复杂度的双重挑战。温佩芝等人[7]（2020）提出的三维重建网格模型缺陷孔洞修复方法，创新性地采用特征线匹配桥接技术构建基曲面，并引入无约束三角剖分结合各向异性网格优化策略，将传统O(n³)复杂度的动态规划剖分简化为O(logn)的高效处理。该工作的实验表明该方法在几何特征还原度和网格质量上优于传统方法，但其特征线检测依赖曲率估计的准确性，在强噪声或非连续特征孔洞场景中可能存在匹配失效风险，且各向异性细化对初始基曲面拓扑结构的完整性要求较高，可能限制其在多连通复杂孔洞修复中的应用范围。模型匹配驱动三维补全则通过匹配输入残缺点云与数据库中的完整三维模型或局部部件，预定义模板库与目标区域的局部特征对齐实现缺失替换，利用几何相似性填补缺失区域。Li等人[8]（2015）设计了融合局部几何细节与全局支撑基元的关键点星座描述符，基于数据库对3D物体进行实时扫描检索，能够在未分割、含噪声的扫描数据中实现鲁棒匹配，有效解决了传统方法对精确模板匹配和场景分割的依赖；然而，该方法对微观几何特征和低纹理透明物体的识别仍存在局限，且模型替换的准确性依赖于数据库的几何相似性覆盖度，在缺乏语义关联的异构数据场景下可能产生次优匹配。以上诸类经典的传统三维补全方法为解决三维点云或网格模型因采集设备局限、环境遮挡或目标自身破损导致的几何缺失问题提供了基础框架：几何对称性方法基于物体结构的重复性特征构建缺失区域的几何映射关系，在规则化工业部件或标准建筑结构的局部缺损修复中展现了计算效率高、物理可解释性强的优势；表面重建方法则通过参数化插值或非参数化拟合构建连续曲面，能够针对小尺度孔洞或平滑表面缺失实现快速补全；模板匹配方法基于先验知识库，通过特征对齐与形状替换策略实现缺损区域的语义化补全，在文物修复、医学器官重建等具有明确形态先验的领域表现出一定可靠性。

然而，这些传统方法在复杂场景下的技术效能存在系统性局限：其一，几何对称性方法高度依赖目标物体的理想化对称假设，难以适应实际场景中普遍存在的非对称结构或动态遮挡干扰，导致补全结果出现拓扑断裂或几何畸变；其二，表面重建方法受限于其数学模型的固有约束，插值法在处理高曲率区域时易因非均匀采样产生表面振荡或细节平滑化，而拟合法则因参数敏感性易陷入局部最优解，难以实现复杂曲面的高精度重建；其三，模板匹配方法的性能受制于模板库的完备性与匹配算法的鲁棒性，面对未知形状类别或大规模几何缺失时易出现特征误匹配或形状失真，且其计算复杂度随模板规模呈指数级增长，难以满足实时性需求。更为本质的局限性在于，传统方法的“规则驱动”特性导致其缺乏对点云全局语义关联的建模能力：人工设计的几何规则难以覆盖现实场景中多样化的几何形态与拓扑结构，而局部优化策略无法建立缺损区域与完整结构的全局一致性关联，导致补全结果在整体结构合理性与局部细节保真度之间难以平衡。例如，在文化遗产数字化修复中，传统方法虽能基于对称性补全陶器口沿的规则缺失，却无法重建非对称铭文或破损浮雕的微观几何特征；在动态人体扫描补全任务中，模板匹配法易因姿态差异导致肢体连接处出现非物理性形变。尽管传统方法凭借低计算开销与强可解释性，在工业逆向工程、产品质量检测等对实时性与确定性要求较高的场景中仍具应用价值，但其“先验强依赖—局部弱关联”的技术范式已难以满足高复杂度、强噪声环境下的三维补全需求。当前研究趋势正尝试通过机理融合或架构创新突破传统方法性能边界，但其在跨尺度几何生成、多模态语义理解等方向仍需理论突破与算法革新。

### 基于深度学习的点云补全模型

基于深度学习的点云补全网络通过数据驱动范式突破了传统几何方法的固有局限，其核心在于利用神经网络自动挖掘点云数据中隐含的全局语义关联与局部几何细节，从而实现缺失区域的高保真重建。随着卷积神经网络、图神经网络及Transformer架构的迭代优化，以及GPU并行计算能力的指数级提升，深度学习模型逐步取代传统手工设计规则，成为点云补全领域的主导技术路线。

此类方法通过端到端的特征学习机制，将点云补全任务转化为隐空间中的概率分布建模问题，能够自适应地处理复杂几何变形、非均匀采样及多尺度缺失等挑战，显著提升了补全结果的几何一致性与语义合理性。

早期研究集中于单模态表征学习，通过不同数据转换策略构建补全网络架构。体素化方法[9,10]将点云离散化为规则体素网格，利用三维卷积提取多层次特征，但受限于体素分辨率的量化误差与内存消耗问题。多视图投影方法[11]通过多视角二维投影结合2D CNN提取特征，再经反投影与融合生成完整点云，然而投影过程中的几何信息损失导致细节重建质量受限。早期的点云数据处理方法主要依赖于手工特征设计或中间格式转换以应对点集的无序性挑战，这些方法普遍面临几何信息损失、计算效率低下和架构灵活性不足等核心问题，无法在保持点云原生无序性、刚性变换不变性的前提下实现端到端特征学习，从而制约了三维几何深度学习的性能上限与应用边界。2017年，Qi等人[4]首次提出了一种全新的直接处理点云数据的方法，标志着深度学习直接处理原始点云数据的突破性进展。这一工作解决了点云无序性、刚性变换不变性等关键挑战，为后续三维几何深度学习研究奠定了基础，之后该团队进一步进行研究，对该方法进行改进升级，提出了PointNet++[5]方法，通过局部区域采样与多层感知机堆叠实现细粒度特征提取。此类方法虽避免了量化误差，但对全局拓扑结构的建模能力不足，易产生结构畸变。

点云数据因其非规则采样特性兼具低存储开销与高保真几何细节表征能力，近年来逐渐成为三维几何建模的重要范式。然而，传统卷积神经网络的局部聚合操作本质依赖于欧氏空间下的规则网格结构，其固有的平移不变性与局部感受野构建机制难以适配点云数据的拓扑非规则性及置换不变性需求。这种结构化归纳偏置与点集无序性之间的本质矛盾，导致基于体素化或多视图投影的经典三维分析方法在点云处理场景下面临特征表达效率低下与几何信息损失的双重挑战，亟待发展新型几何推理架构以实现原生点集的高效语义理解。随着PointNet系列的提出，也有更多的新型网络应运而生。Yuan等人[12]（2018）利用PointNet，首次设计出一种于 Encoder-Decoder 框架的深度学习架构直接处理原始点云数据，从而避免了体素化的内存开销和几何信息损失，采用多阶段解码器架构在保持较少参数量的同时生成高密度细节的完整点云。实验表明该方法在合成与真实数据上均能有效处理不同程度的不完整性和噪声，并能提升下游任务如点云配准的精度。然而，PCN在处理由多个不连接部件组成的物体时可能错误连接部件，且对细薄结构的恢复能力有限。同年，Yang等人[13]提出了FoldingNet，这是一种基于深度网格变形的点云自动编码器，利用二维网格变形重构三维点云，显著减少了参数量的同时实现了高效重建和高判别性表征，在无监督分类任务中达到接近监督学习的精度。然而，该方法在处理复杂拓扑结构时仍存在局限性，可能因二维网格的固有约束导致精细结构丢失，且对高复杂度表面重构需依赖更多采样点或更深层网络，限制了其在极端几何场景的应用效果。Wang等人[23]（2019）提出了DGCNN及其核心模块EdgeConv，通过动态构建局部邻域图并融合全局坐标与局部几何特征，显著提升了点云分类、分割等任务的性能。该方法在保持排列不变性的同时，利用特征空间动态更新邻域关系，使得深层网络能够捕捉长距离语义关联。而其局限性在于动态图计算带来的额外开销，处理大规模点云时效率受限；此外，特征空间的邻域构建对初始特征质量敏感，在极端噪声或稀疏场景下的鲁棒性仍需验证。

尽管现有方法围绕点云的无序拓扑结构和刚性变换不变性特征构建了持续演进的深度学习框架，但多数模型仍侧重于单点或局部几何属性的提取，未能从全局视角建模局部语义单元间的空间-语义关联关系，导致补全结果在跨部件结构连贯性上存在明显局限。Yu等人[15]（2021）提出的PointTr模型首次将NLP 领域中的Transformer架构引入点云补全任务，通过将点云表示为具有位置嵌入的点代理序列，构建编码器-解码器框架捕捉全局结构关联，并设计几何感知模块强化局部几何关系，在多个基准测试中显著超越现有方法，同时构建了更贴近真实场景的多样化基准数据集；然而该模型在点代理特征抽象过程中可能弱化了局部区域的细粒度细节表达，导致生成点云虽具备完整结构但表面分布的均匀性和微观几何精度仍有提升空间，反映出全局语义关联与局部特征保真之间的权衡难题。深度学习点云补全技术的持续演进，那么如何在有效建模点云局部区域间语义关联性的基础上，同步实现高精度细节还原与表面均匀分布的完整点云生成，这构成了当前点云补全领域亟需突破的关键技术瓶颈。深度学习点云补全技术的推进也将为三维重建、数字孪生与智能感知系统提供更为鲁棒的数据基础。

## 文章研究内容

本文的主要研究内容是三维点云数据的补全方法，针对现有方法在几何信息严重不足的极稀疏点云场景下存在的性能瓶颈，提出了一种基于深度学习的双路特征增强渐进式补全框架。

首先，针对极稀疏点云局部几何信息缺失严重的问题，本文提出了一种双路并行特征提取模块。该模块包括局部特征分支和全局特征分支：局部特征分支采用改进的偏移注意力机制，有效捕获点云的微观几何结构；全局特征分支则通过层次化特征聚合，建立物体的宏观形状先验。两条分支通过交叉注意力机制实现自适应融合，从而获得兼具局部细节敏感性与全局结构一致性的增强特征表示。

其次，为了对点云中长距离的几何依赖关系进一步建模，本文使用Transformer编码器-解码器架构作为基础框架。该架构引入动态Query生成器和几何感知解码器，将稀疏点云映射为点代理序列，并通过自注意力机制有效捕捉跨区域的几何关联。该方法充分考虑了点云已观测部分与缺失部分之间的上下文关系，显著提升了点云补全的整体准确性和结构一致性。

最后，考虑到单阶段生成方法易导致细节丢失的问题，本文提出了多阶段渐进式细化网络。该网络通过级联折叠操作，实现从粗到细的分层重建：第一阶段恢复点云的拓扑骨架，第二阶段优化各部件之间的连接关系，第三阶段进一步细化表面细节。各阶段之间通过残差连接保持几何信息的一致性，最终输出高保真的完整点云。

为了验证本文提出方法的有效性，本研究将本文的方法与近年来多个主流的方法进行实验对比。通过多项度量指标和可视化结果，横向展示本文模型的优势。同时，设计消融实验，纵向评估本文模型的性能。

## 文章组织结构

本文共分为五个章节，各章节内容安排如下：

第一章为绪论。阐述点云补全的研究背景与意义，系统分析传统方法与深度学习方法的技术特点与发展脉络，明确本文研究内容与创新点，并说明全文组织结构。

第二章为基础理论。详细介绍点云数据的数学表征、获取方法及核心特性，深入分析稀疏性的信息熵表征与补全挑战，系统梳理点云质量评估指标，并阐释深度学习模型的关键组件及其在点云处理中的适应性改进。

第三章为本文的主要内容。提出双路特征增强渐进式补全框架，详细阐述模型总体架构，包括数据预处理策略、双路特征提取模块设计、Transformer编码-解码核心架构、全局特征增强机制以及多阶段细化网络的实现方法，并给出损失函数的优化目标。

第四章为实验分析与验证。介绍实验环境配置与数据集特性，通过定量指标对比和可视化结果分析验证模型性能，并设计消融实验探究各模块的贡献度，全面评估方法的有效性与鲁棒性。

第五章为结论与展望。总结本文研究成果与创新点，指出当前方法的局限性，并对未来研究方向提出建议。

# 基础理论

## 点云数据基础

### 定义

点云是在同一空间参考系下通过三维扫描技术获取的、以离散点集合形式表征目标物体表面几何形态与属性特征的数据结构。每个采样点至少包含三维笛卡尔坐标

（2-1）

以精确描述空间位置，点云数据也可扩展携带更多复杂信息，如激光反射强度、RGB颜色信息、法向量以及时间戳等多元属性，用于描述物体或场景的表面几何形态与属性特征。以下给出点云的形式化定义：

（2-2）

式中：——欧氏空间坐标；——第k类属性特征；*N*——点云基数。该集合无序且满足，即不存在固有排列次序。

### 核心特性

#### 无序性

点云是一组没有特定数据的点，集合内部不存在固定排列顺序，任何点序列的置换操作均不改变其表征的几何语义。点云数据缺乏规则化的数据组织形式，既无法直接映射为二维图像中的栅格矩阵，也不具备三维网格模型中的顶点-边-面拓扑连接关系。那么，这一特性要求算法必须满足置换不变性，例如对称函数等。

#### 点间交互性

三维空间中的点并非孤立存在，而是来自具有距离度量的空间，通过局部邻域形成具有语义关联的几何结构。因此，模型需要能够从附近的点中捕获局部结构，以及局部结构之间的组合交互。

#### 变换不变性

该特性包含刚性变换不变性与非刚性形变鲁棒性，即在变换后不改变其几何本质。这一特性通过空间变换网络学习仿射矩阵实现特征空间对齐，前者要求算法输出不随坐标系旋转而变化，后者需应对物体弹性形变。

### 点云数据获取与表示

#### 点云数据的获取机制

点云数据的获取机制依托于多模态三维传感技术的协同发展，主要可划分为主动式激光测距与被动式光学成像两大技术体系，具体包括激光雷达扫描（LiDAR）[16]、结构光扫描技术[18]、摄影测量法[19]和深度相机技术[20]等，四种技术的获取原理如图2-1所示。

在主动式测距领域，激光雷达扫描（LiDAR）通过纳秒级脉冲激光发射与回波相位差解算，将激光角分辨率、视场角、反射率、点频、偏差值和物体颜色信息等进行数值转化，存储为对目标物体的各个属性具有表征能力的三维点云数据；从而实现亚厘米级空间定位精度。结构光扫描技术则通过编码光栅投影至物体表面，利用双目相机捕捉形变图案，结合三角测量原理测量计算三维坐标，从而实现存储亚毫米级重建精度的点云数据。该技术虽具备百万级点/帧的高分辨率与RGB色彩融合能力，但易受环境光干扰导致点云缺失。而在被动式成像体系中，摄影测量法基于多视几何理论与光束法平差优化，通过特征匹配从二维图像序列重建稀疏点云，再经稠密匹配生成2.5D表面模型。深度相机技术融合了主动红外与被动双目视觉，将具有结构特征的光线投影到物体上，再利用红外摄像头采集反射光获取深度信息，最后通过实时深度图和透视投影模型转换生成动态点云。

|  |  |
| --- | --- |
| 34d69cc2baaeb2e459e005329f80f2a  (a)激光雷达扫描原理[17] | (b)结构光扫描原理[18] |
| a36b5388b2a434a65b1af526c698cac  (c)摄影测量法原理[19] | 3f722ac3fccfc84c20b83e977744a66  (d)深度相机技术原理（ToF）[20] |

图 2-1 点云数据的获取机制原理示意图

#### 点云数据的多维表征

点云数据的表示方法作为三维视觉领域的关键技术，其本质在于如何高效、准确地描述离散三维点的空间分布与属性特征。当前主流的表示方法可归纳为以下五类：点表示法、体素表示法、图表示法、多视图投影法以及混合表示法。点表示法直接以原始点云数据作为输入，通过多层感知机点提取特征。其优势在于最大限度保留几何细节，且无需预处理，适用于高精度场景。然而，点云的无序性与非结构性导致计算复杂度高，需引入置换不变性设计以解决排列敏感问题，比如引入对称函数。体素表示法将三维空间离散化为规则立方体网格，利用3D卷积神经网络处理体素化的数据。该方法通过量化坐标实现数据规整化，显著提升计算效率，尤其适用于实时性要求高的自动驾驶场景。但体素分辨率与计算资源的矛盾突出，体素数据的高分辨率导致内存爆炸，而低分辨率又会引发几何细节丢失。图表示法根据图神经网络构建点间拓扑关系，利用边卷积或动态图更新机制来捕捉局部邻域特征。该方法天然适应点云的不规则分布特性，能够建模长程依赖关系，在点云分割任务中表现优异。多视图投影法将三维点云投影至多视角二维平面生成RGB-D图像[21]，利用成熟的2D CNN提取特征。此方法在物体分类任务中取得较高准确率，但因投影过程引入视角遮挡与几何信息损失，在复杂场景中鲁棒性受限。改进策略包括视角自适应加权融合与几何一致性约束。混合表示法则结合点与体素的双模态优势，例如结合点与体素的层次化表征，在粗粒度体素层实现快速区域提案，再通过点级特征精细化回归，兼顾效率与精度。

如今，随着深度学习技术的发展，点云表示方法正朝着多尺度特征融合、动态场景适应以及轻量化部署方向演进，为自动驾驶、虚拟现实等领域的应用奠定了基础。

### 采样方法

点云数据的采样方法的核心目标在于通过优化点云密度与分布来平衡计算效率与几何特征保留。根据采样方向的不同，主要分为下采样与上采样两类方法。

#### 下采样方法（Downsampling）

点云数据的下采样方法是指通过降低点云密度以减少数据冗余和计算量，同时保留关键几何特征的技术。

##### 体素化网格采样

该方法将点云空间划分为规则立方体网格，即体素，对每个体素内的点取质心或加权平均值作为代表点。该方法效率高、分布均匀，且通过调整体素尺寸可间接控制点间距，但量化误差易导致高频细节丢失，且无法精确控制输出点数。

##### 均匀采样

该方法基于空间分布均匀性准则选择代表性点，典型算法如最远点采样，通过迭代选取距离已有采样点最远的点，确保全局均匀性。其采样点分布优于体素方法，但时间复杂度随点云规模呈指数增长，需通过分治策略优化效率。

##### 几何采样

该方法以局部几何特征为权重进行非均匀采样。通过计算邻域法线夹角近似曲率，划分特征区与非特征区并差异化采样密度，可有效保留尖锐边缘与复杂结构。然而，曲率计算的耗时性需通过GPU加速或邻域预计算缓解。

##### 随机采样

该方法通过随机剔除点实现快速降维，虽能精确控制输出点数，但随机性可能导致关键信息丢失，故常作为预处理步骤与其他方法联用。

#### 上采样方法（Upsampling）

点云数据的上采样方法是通过插值或生成模型增大点云密度或增加点的特征维度以恢复细节信息的技术。

##### 基于插值的增采样

该方法采用滑动最小二乘法或线性插值对稀疏区域进行点云扩充。MLS通过局部曲面拟合生成平滑过渡点，适用于点云去噪与表面光顺处理，但拟合精度受邻域半径参数影响显著，过度平滑可能导致细节损失。

##### 深度学习方法

该方法基于神经网络的生成模型通过特征扩展与坐标重建实现高密度点生成。此类方法采用分片特征提取、多尺度特征融合及对抗性训练策略，可在稀疏输入下生成几何合理的稠密点云。近年来，标准化流模型通过潜在空间自适应插值优化生成点的多样性与分布均匀性，显著提升了复杂曲面的重建精度。

### 稀疏性理论

#### 点云稀疏性的定义及其统计特性

点云的稀疏性是指激光雷达等设备所采集的三维点云数据相对于其所表达的实际场景空间而言，采样点分布稀少且不均匀的特性，其点云数量通常只有完整点云的5%-10%。

稀疏点云由于其点云数量极少，呈现出在空间分布上的不均匀性、几何信息的不确定性、拖布结构碎片化以及高维特征分布的稀疏性。

#### 点云稀疏性的来源

根据2.1.4小节所述的点云采样方法，可以得知点云稀疏性主要来源于传感器物理限制、距离因素、遮挡效应、材质特性以及环境因素。

##### 传感器物理限制

##### 采样设备的物理结构会导致稀疏点云的产生。如激光雷达的光束数量和扫描频率有限，导致采样点间存在固有间隔。再者随着距离增加，相邻激光束之间的间距呈发散状增大，导致采集到的点云数据进一步稀疏。

##### 距离因素

点云密度随距离平方反比快速下降，远处物体的采样点相较于近处的采样点明显稀疏。采样设备发出的激光能量随距离衰减，导致远处目标反射信号弱，出现大部分点云数据丢失的现象。

##### 遮挡效应

近处物体遮挡远处物体，造成"激光雷达阴影"区域；物体自身结构导致的部分表面被遮挡，无法被激光束照射到；移动物体产生的临时遮挡，导致点云采集不完整；以上多种原因都有可能导致采集到的点云出现稀疏性。

##### 材质特性

扫描的物体材质也会导致设备采集到稀疏点云，如高反光表面导致激光束偏转，无法返回传感器；或玻璃、水面等材质导致激光穿透或散射；还有黑色、暗色物体吸收激光能量，返回信号微弱；导致稀疏点云的产生。

##### 环境因素

点云采集时所处的环境也会影响采集点云的实际效果。例如，雾、雨、雪、尘埃等影响激光传播和反射；强光源可能干扰激光雷达的接收系统；高速移动物体或传感器自身运动导致的采样失真都可能导致在实际采集活动中采集到的点云数据极稀疏。

#### 信息理论视角下的点云稀疏性分析

##### 点云分布的空间信息熵表征

点云在三维空间中的稀疏分布可视为一种低熵状态。由于传感器采样的物理局限，远距离或遮挡区域的点云密度显著降低，导致局部信息熵急剧下降。这种现象可量化表示为：

（2-3）

式中：——体素中包含点的概率。稀疏点云的空间熵通常较低，表明其信息分布不均衡，大部分区域不包含有效信息。

##### 点云分布的吉尼系数(Gini coefficient)量化

稀疏点云中，信息集中于少量非空区域，因此形成"信息孤岛"，导致香农熵在空间分布上呈现高度不均衡状态。这种不均衡可用吉尼系数(Gini coefficient)进行描述：

（2-4）

式中：——表示区域的点密度。

#### 稀疏点云的补全挑战

稀疏点云最根本的挑战是信息不足，导致关键几何细节缺失。尤其当采样点间距大于物体的特征尺度时，重要结构特征可能完全缺失，导致形状歧义性显著增加。再者在稀疏条件下，建立有效的局部特征与全局结构间的映射变得极为困难。传统的基于k近邻的局部特征提取方法失效，因为k近邻可能跨越实际不相邻的表面区域，引入错误的几何关联。在极稀疏情况下，同一组稀疏点可能对应多种合理的完整形状，形成"一对多"的映射关系。例如，仅10个点可能对应球体、立方体或多种不同的形状。这种本质的歧义性使得点云补全成为一个病态问题。补全算法面临细节保真度与全局一致性的固有矛盾——过度强调与稀疏输入的一致性可能导致不合理的局部几何结构；而过度强调全局平滑性则可能丢失输入中蕴含的局部细节信息。

稀疏点云的这些挑战构成了点云补全研究的理论基础，也为新型算法设计提供了重要指导方向，特别是启发了融合多尺度特征提取和渐进式细化策略的双路特征增强补全框架的发展。

## 点云质量评估指标

### 倒角距离

3D点云中的倒角距离（Chamfer Distance，）是衡量两组点云之间几何相似性的核心指标，主要用于点云重建或者3D重建工作。如果该距离较大，则说明两组点云区别较大；如果距离较小，则说明重建效果较好，匹配精度较高。一般来说，该距离可用作3D点云补全模型的损失函数。

倒角距离通过计算两组点云间双向最近邻距离的均值，量化几何匹配度。给定点云集合和，其数学表达式为：

（2-5）

该指标包含两部分：前向距离：中每个点到的最近邻点距离平方的平均值；反向距离：中每个点到的最近邻点距离平方的平均值。

### 推土机距离[22]

3D点云中的推土机距离（Earth Mover’s Distance，EMD），也叫作Wasserstein距离。这是一种基于最优传输理论的几何相似性度量方法，通过计算两个点集之间最小运输成本来评估其分布匹配程度。相较于其他距离指标，EMD不仅关注局部几何偏差，更强调点云整体分布的全局一致性，因此在三维重建、点云配准与生成模型评估中具有独特优势。

给定点云集合和，EMD定义为两集合间最优双射映射的最小运输代价之和：

（2-6）

式中：——双射映射，满足每个点与唯一目标点对应。

该指标通过求解线性规划问题，寻找使总移动代价最小的点对匹配关系。其物理意义可类比为将一堆“沙土”（点云）搬运至目标位置（点云）所需的最小工作量，推土机距离因此而得名。

### 稀疏场景下评估指标的适应性

点云补全任务来到稀疏点云场景下，评估指标需适应稀疏点云固有的低密度、非均匀分布及噪声干扰等特性而兼顾到局部几何误差的鲁棒性与全局分布一致性。本小节针对倒角距离（CD）与推土机距离（EMD）在稀疏点云场景下的适应性进行对比分析。

#### 倒角距离（CD）的适应性

由于CD需计算双向最近邻距离均值量化点云匹配度，CD的计算效率高且对局部几何偏差敏感。在稀疏点云补全中，由于真实点云与补全点云均存在密度不足问题，点云间可能存在较大的局部缺失区域。那么CD的最近邻匹配机制在此场景下表现出对稀疏点云的鲁棒性和对计算的高效性。

#### 推土机距离（EMD）的适应性

在稀疏点云补全任务中，EMD因其全局最优传输的特性，对分布一致性的评估具有独特价值，但其在稀疏场景下真实点云的低密度特性与补全点云的高密度输出导致规模非对称，直接应用需依赖重采样或插值，可能引入评估偏差；同时，稀疏关键点的微小位置偏移会因全局传输代价累积而放大误差，降低评估稳定性。

## 点云深度学习方法基础

### 基础网络架构

#### 多层感知机（Multilayer Perceptron，MLP）

MLP的核心在于通过引入一个或多个隐藏层实现对复杂非线性关系的建模，其神经网络结构如图2-2所示。隐藏层位于输入层与输出层之间，通过多层叠加结构扩展网络深度，从而提升模型对高维特征空间的计算能力。在架构设计上，每层神经元均与上一层的所有神经元形成全连接关系，即每个神经元的输入由前一层全部神经元的输出构成。对于单个神经元而言，其输入向量为，输出的计算过程可通过公式2-5表达：

（2-7）

式中：——输入维度；——对应维度的权重参数；——偏置项。

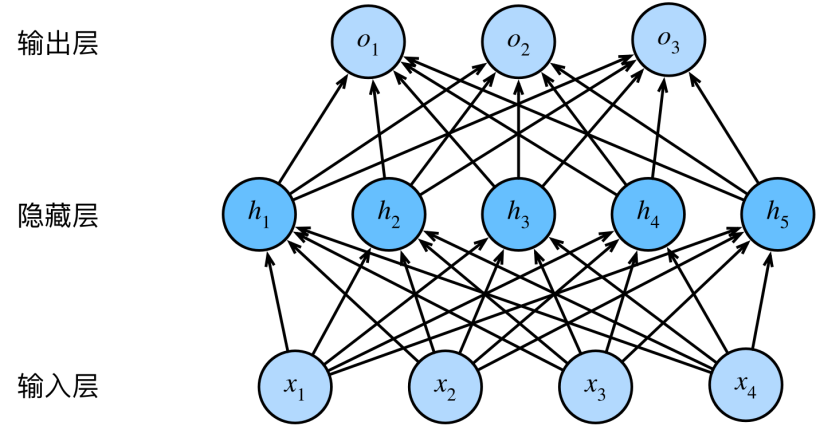


图 2-2 一个单隐藏层的多层感知机，具有5个隐藏单元

MLP作为深度学习的核心模块，通过堆叠全连接层与非线性激活函数实现对原始点云数据的逐层抽象，在点云处理任务中通过非线性映射实现特征抽象与信息传递。在点云处理中，MLP的核心作用包括：

逐点特征映射：将每个点的三维坐标及其可能的附加属性通过线性变换映射至高维特征空间。例如PointNet中采用共享权重的MLP对输入点云进行逐点特征提取，从而保留局部几何信息。

全局特征聚合：通过多层非线性变换，MLP能够逐步融合邻域或全局上下文信息。例如，PointNet++通过层次化MLP结构，在多尺度上提取点云特征，兼顾局部细节与整体形状。

参数共享设计：在基于点的网络（如PointNet）中，MLP参数在所有点上共享，确保模型对点云无序性的鲁棒性。

局部几何特征提取：通过逐点映射将原始坐标信息转换为高维语义特征。例如，PoinTr模型中采用轻量级DGCNN网络提取局部区域的点代理特征，其核心由多层MLP构成，通过聚合邻域点信息生成具有判别性的特征向量。

特征维度适配：MLP可通过调整输出维度实现不同层级特征的空间对齐，比如在点云下采样过程中，MLP将不同分辨率的特征映射到统一维度以支持后续的跨尺度信息融合。

1. 激活函数设计

多层感知机隐藏层的计算无法解决非线性问题，因此通过引入激活函数来实现神经网络的非线性计算，其设计直接影响模型的收敛速度、特征抽象能力及对复杂几何关系的建模效果；而在点云补全任务中，激活函数的选择需兼顾计算效率、梯度稳定性与对局部几何细节的敏感性。

激活函数作用于神经元的加权和输出，通过非线性映射突破线性模型的局限性，使神经网络能够拟合复杂的数据分布。数学上，在引入激活函数后，对于单个神经元的输出的计算过程可通过公式2-6完整表达：

（2-8）

式中：——激活函数。

引入激活函数的核心作用是实现非线性建模，将线性组合转换为非线性输出，支持网络学习点云中的曲面结构、空洞边界等复杂几何关系；其次调控梯度的传播，通过激活函数的导数特性影响反向传播中误差梯度的计算，避免梯度消失或爆炸问题；最后保持数值稳定性，规范化神经元输出范围，缓解深层网络训练中的数值溢出的风险。

目前常用的激活函数有三种，分别是 Sigmoid 函数、ReLU（Rectified Linear Unit）函数、Softmax函数，计算公式分别如2-7、2-8、2-9所示。

（2-9）

（2-10）

（2-11）

Sigmoid函数是早期神经网络中广泛使用的激活函数之一，该激活函数的函数和导数图像如图2-3所示，该函数将输入值映射到（0,1）区间内，对输出为0和1的阶跃函数进行了平滑处理。Sigmoid函数的输出恒为正值，且在输入接近正/负无穷时分别趋近于1和0，中间区域呈现平滑的S形曲线。这一特性使其曾广泛用于概率建模。

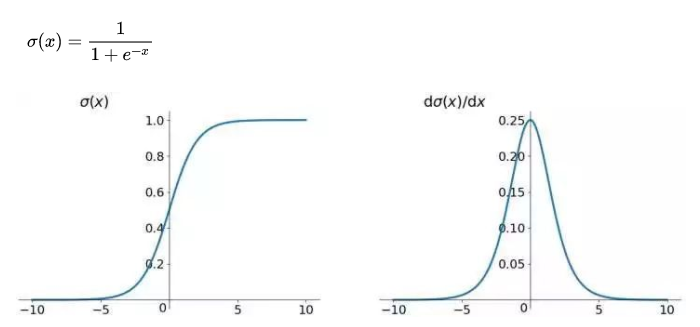


图 2-3 Sigmoid 激活函数及其导数图像

ReLU函数也是深度学习中应用最广泛的激活函数之一，属于目前最常用的激活函数，该激活函数的函数和导数图像如图2-4所示，呈现出典型的单侧抑制特性：输入为负时输出恒为0，输入为正时保持线性增长。正区间导数为1，负区间导数为0。其核心设计以简洁性和高效性著称。

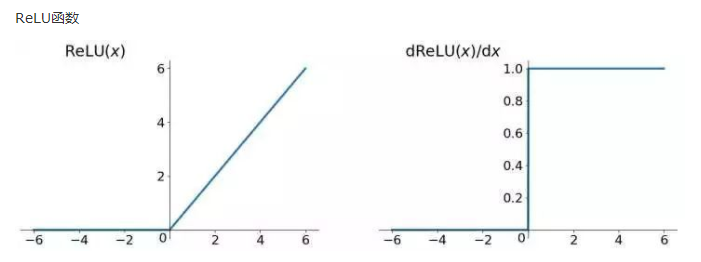


图 2-4 ReLU激活函数及其导数图像

Softmax函数则是深度学习中处理多分类问题的核心激活函数，该激活函数的函数图像如图2-5所示，其设计理念与Sigmoid、ReLU等激活函数一脉相承，但在功能和应用场景上具有显著差异。该函数的所有输出值位于(0,1)区间，总和为1，适合表示多分类概率，通过归一化指数机制，将多分类输出转化为概率分布，是深度学习模型的核心组件之一。

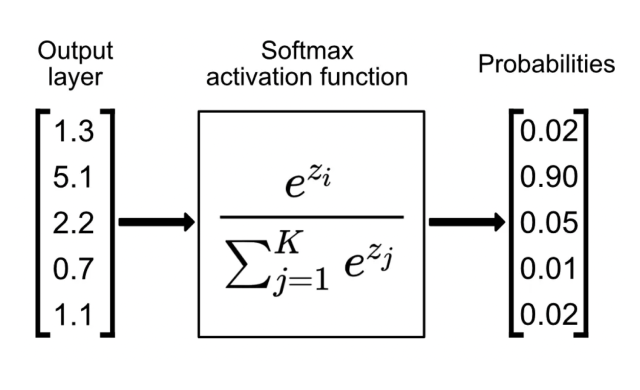


图 2-5 Softmax分类函数图像

1. 批量归一化（Batch Normalization，BN）

批量归一化在深度神经结构中通过标准化神经网络中间层的输入分布，缓解梯度消失或梯度爆炸的问题，同时引入隐式正则化效果；具体而言，BN通过逐层规范化神经元激活值，使其近似服从均值为0、方差为1的标准正态分布，从而有效缓解深层网络训练中普遍存在的内部协变量偏移问题。从而达到加速训练、提升模型稳定性的预期。

对于输入数据（式中：——批量大小；——通道数；——高度；——宽度），BN分别对每个通道计算均值和方差：

（2-12）

标准化后通过可学习参数（缩放）和（偏移）恢复模型表达能力：

（2-13）

式中：——防止除零的小常数（如1e-5）。

而在点云补全模型中，BN除了起到经典的效果如稳定训练过程外，还能增强多尺度特征的一致性。针对点云数据固有的非均匀密度分布与局部几何结构复杂性，BN通过引入跨层特征分布对齐机制，有效解决了多尺度特征融合过程中的表征失配问题。其次，还允许深度网络调整超参数，获得更高学习率。通过减少内部协变量偏移，BN使得深层网络可采用更激进的学习策略。  
 上述优化机理的协同作用，使BN成为现代点云补全模型的核心组件，该原理通过建立跨层级、跨尺度的动态特征校准机制，在保持模型表征能力的前提下，实现训练过程的显式可控性。综上，批量归一化通过标准化中间层输入分布，成为深度神经网络训练的“加速器”与“稳定器”。尽管存在对小批量的敏感性等问题，但其在图像处理等领域的表现仍不可替代。

1. 残差连接（Residual Connection）

残差连接通过跳跃路径将输入直接传递至深层网络输出端，解决梯度消失与网络退化问题，实现跨层级特征融合与梯度稳定传播。残差连接通过引入残差映射替代直接映射，将输入与非线性变换后的输出相加：其数学形式如2-12所示：

（2-14）

式中：——残差函数。

残差函数作为深度神经网络的核心设计范式，在点云补全任务中通过多层次信息融合与动态梯度调控机制，显著提升了模型对复杂几何形变和局部细节的建模能力。在点云补全模型中，残差函数对梯度传播进行了优化，残差路径通过构建跨层恒等映射，如同为反向传播提供“高速公路”，为反向传播过程建立低衰减梯度通路，有效解决深层网络退化问题，使得模型可以接受更深层次的网络。其次，残差函数的恒等映射特性还实现了原始几何信息与深层抽象特征的协同复用。该机制通过恒等映射保留原始几何信息，避免补全过程中细节丢失。例如，在动态点云序列补全中，残差连接可融合时序特征与空间结构。最后，其还实现了对称性打破与形变感知机制；残差块将强制网络学习输入与输出间的残差函数，从而增强模型对点云局部结构变化的敏感性。

残差连接的技术通过跨层跳跃式传递机制，有效缓解了深层网络训练中的梯度消失与特征退化问题，为复杂几何结构的逐级抽象提供了稳定的梯度流。其在点云补全模型中的引入，不仅保障了全局与局部特征的高效融合，更通过保留原始几何信息增强了细节恢复的鲁棒性。

1. 池化层（Pooling Layer）

池化层是在深度学习网络模型中一种用于缩小特征图空间尺寸的操作，模仿人的视觉系统对数据进行降维，用更高层次的特征表示图像。它通过局部区域特征聚合操作对输入特征图进行压缩，提取具有平移不变性与抗噪性的高层语义信息同时减少计算复杂度。池化层通常用于卷积神经网络中，以降低参数数量和计算量，并在一定程度上防止过拟合。

池化层通过局部特征聚合与降维操作实现多层次特征抽象，其变体涵盖从规则数据到非结构化点云的多样化设计，在保留关键信息的同时显著提升模型的计算效率与几何感知鲁棒性。常见的池化层类型包括平均池化（Average Pooling）如图2-6（a）所示、最大池化（Max Pooling）如图2-6（b）所示、重叠池化（Global Pooling）等；在点云补全模型中，最远点采样池化（Farthest Point Sampling Pooling，FPS）和局部特征聚合池化（EdgeConv Pooling）最为常见。

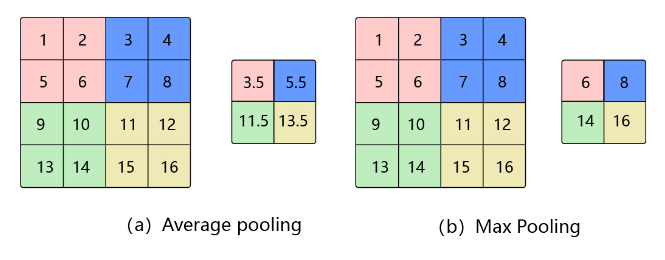


图 2-6 平均池化和最大池化

平均池化是深度学习中经典的池化操作之一，该操作对局部区域内所有特征值取平均从而实现空间信息的压缩和平滑化。平均池化操作实现了对区域特征的平滑，抑制了噪声，适用于需要全局上下文信息的任务，同时也可以减少模型的参数量。

最大池化是最常见、使用最广泛的池化操作，旨在通过局部区域特征聚合降低特征图的空间维度，同时保留显著激活响应。最大池化操作是对输入特征图的局部邻域进行取最大值运算，从而逐层减少特征图分辨率，有效控制模型参数量，缓解过拟合风险。

在深度学习网络模型中，池化层通过降维和特征筛选提升模型鲁棒性，起到了至关重要的作用；其中最大池化因其对显著特征的强保留能力，在点云处理中尤为关键。在点云补全模型中，池化层被用于聚合局部点云的几何信息，同时结合自适应权重或注意力机制，以平衡全局结构与局部细节的恢复。这类设计不仅增强了模型对稀疏输入的容错性，还为复杂几何关系的建模提供了高效的特征抽象基础，成为点云补全任务中不可或缺的组成部分。

### Transformer网络架构[26]

Transformer是一种用于自然语言处理（NLP）和其他序列到序列（seq-to-seq）任务的深度学习模型架构，由Vaswani等人（2017）首次提出。其作为一种革命性的架构，Transformer完全摒弃了传统的循环或卷积结构，而是完全基于注意力机制来处理序列数据，从而在机器翻译等任务上取得了突破性的进展。Transformer的标准结构由编码器(Encoder)和解码器(Decoder)两大部分组成，实现了高效的并行计算，解决了循环神经网络(RNN)在处理长序列时的梯度消失和计算效率问题。

#### 自注意力机制(Self-Attention)

自注意力机制是Transformer的最核心的创新，它使模型能够同时考虑输入序列中的所有位置，捕捉序列内部的依赖关系。它引入了查询(Query)、键(Key)和值(Value)的概念，对于输入序列中的每个元素，通过线性变换生成三个向量。并由此计算与所有键向量的点积，得到注意力分数。随后对注意力分数进行缩放和归一化处理，与向量相乘，得到加权和。其核心公式如2-15所示：

（2-15）

这种机制允许模型关注输入序列中与当前位置相关的其他位置，不受位置距离的限制，有效捕获长距离依赖关系。为进一步增强模型的表示能力，Transformer还引入了多头注意力机制。该机制将输入投影到多个不同的子空间，并行执行多个自注意力计算，然后将结果拼接并投影回原始维度。

#### 位置编码(Positional Encoding)

由于Transformer架构本身不包含任何关于序列中元素位置的信息，因此需要额外添加位置编码来表示序列的顺序信息。Transformer采用了正弦和余弦函数的组合来生成位置编码，这样的编码方式保证了每个位置编码唯一，并且编码之间的相对关系确定，因此可以直接与输入向量相加，为模型提供位置信息。

#### 编码器-解码器结构

Transformer架构具体分为编码器和解码器，二者的内部结构相似，紧密协作以完成序列到序列的转换任务。其内部整体架构大致分为多头自注意力层、残差连接与层归一化和前馈神经网络层，共同构成了一个端到端的深度学习模型。

### 动态图卷积网络

传统卷积依赖规则网格和固定邻域，而点云数据具有无序性、非结构化和局部几何敏感性的核心特性，因此传统卷积难以直接应用。图卷积网络通过构建动态邻域图实现点云数据的局部几何结构建模，利用可学习的边卷积算子聚合邻域点特征，克服传统卷积对规则网格结构的依赖。那么在点云补全的网络模型中，该类网络通过多层次图特征传递与自适应邻域更新机制，有效融合局部细节与全局形状先验，显著提升了复杂拓扑缺失区域的几何连贯性恢复能力。

Wang等人[23]于2019年首次提出了一种新的神经网络：动态图卷积网络（DGCNN），搭载着其核心模块EdgeConv。与传统图卷积仅聚合节点特征不同，EdgeConv显式地对点云中局部邻域内点与点之间的边特征关系进行了建模。EdgeConv在每个顶点发出的所有边的关联边特征上执行通道对称聚合的操作（例如：或），因此EdgeConv在第*i*个顶点的输出如公式2-16所示。

（2-16）

式中：——中心点特征；——邻域点特征；——共享的多层感知机；——对称聚合操作。

此操作计算了中心点与相邻点的相对坐标差异，从而捕捉局部几何模式，再与中心点绝对坐标拼接，最终进行非线性组合以生成具有判别性的边特征，在保留局部几何细节的同时引入全局位置信息。这种双路特征融合机制使得网络能够同时感知局部点云的形态变化与整体空间分布，通过多层堆叠逐步构建从细节到整体的层次化表征。

动态图构建是EdgeConv区别于传统静态图卷积的核心创新，进一步强化了模型的语义感知能力。在点云处理的过程中，点之间的语义关系随网络层次变化，在浅层网络中，邻接关系主要反映原始空间的几何邻近性，有助于基础几何特征的提取；随着网络深度的增加，特征空间的度量逐渐转变为语义相似性导向，使得空间距离较远但语义相关的点能够建立连接。那么DGCNN便采用逐层更新的动态k近邻策略进行图结构的构建，其具体实现流程可参考[23]。

其次，在特征传递过程中，该网络利用跨层连接的策略将浅层的高分辨率几何特征与深层的抽象语义特征进行通道拼接，形成64+64+128+256的多尺度特征向量。这种设计使得网络能够同时捕获点云的局部微分结构和全局拓扑特性，并通过动态更新的邻接关系实现特征空间的语义驱动重构。最终，通过全局最大池化层将多层次特征聚合为紧凑的全局描述子，实现局部几何结构与全局语义信息的协同优化，显著提升了模型对复杂三维形状的表征能力。

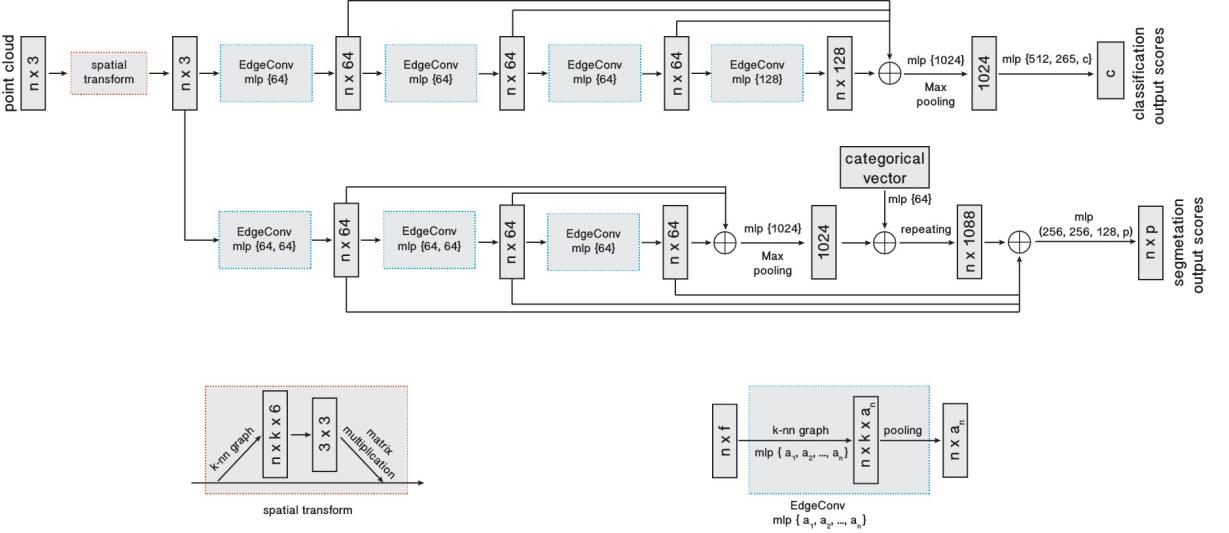


图 2-7 动态图卷积神经网络的模型架构：用于分类（顶部分支）和分割（底部分支）所使用的模型架构。

### 生成式模型

生成式模型在点云补全任务中通过编码器-解码器架构构建部分观测到完整形状的映射关系，这些模型通过对抗训练策略与多尺度特征融合等技术，在进行拓扑合理性构建的同时恢复细粒度几何细节，挑战克服点云无序性带来的结构歧义，并平衡生成结果的多样性保真度与局部几何一致性。

#### FoldingNet的网格折叠机制[13]

在三维点云处理领域中，基于深度学习的生成式模型面临的核心挑战在于如何有效捕捉与重建不规则点集的内在几何结构。Yang等人提出了FoldingNet，创新性地引入了网格折叠机制，为点云自编码器提供了一种高效且具有强表达能力的解码范式，实现了从2D种子点到3D完整曲面的转化。该机制的理论基础源于微分几何中的流形嵌入思想，通过将三维物体表面建模为二维流形的非线性变换结果，实现了对点云生成过程的几何约束。

FoldingNet的关键——折叠机制认识到三维物体表面本质上是二维流形的非线性嵌入，从而利用MLP实现从二维规则网格到三维点云的可微分形变。该网络的具体工作流程如图2-8所示：

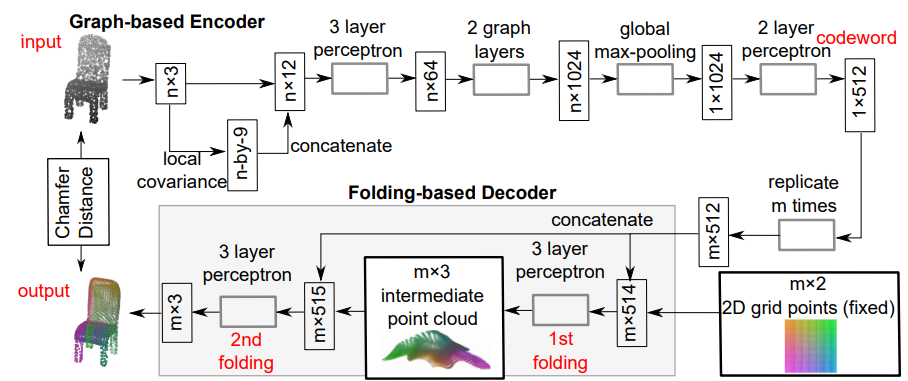


图 2-8 FoldingNet工作架构：基于编码器-解码器的级联式双阶段形变架构

输入点云通过编码器得到局部-全局的特征和潜在码字，随后进入模型的关键解码器通过双重折叠操作将二维规则网格变形为目标点云。第一阶段将均匀采样的二维网格点与编码器提取的潜在码字 拼接，经过一个MLP 映射至三维基础曲面。第二阶段在此基础上进行精细化调整，将的中间结果再次与拼接，通过第二个MLP 实现局部几何特征的增强。这种级联式折叠架构的迭代细化策略在数学上等效于复合函数，其中首次折叠建立全局拓扑框架，二次折叠则聚焦于局部曲率调整。值得注意的是，潜在码字在两次折叠中均作为条件参数输入，确保了形变过程与输入几何特征的强关联性。

这个新型的流程在不同数据集上表现出了杰出性能，证明了网格折叠机制在几何表征学习中的优越性。

#### PCN的粗-细两阶段补全策略[12]

基于生成式模型的点云补全方法通过构建从部分观测到完整形状的概率映射实现几何推理。Yuan等人提出了一种新型点云补全网络PCN，装载了一种创新的粗-细两阶段生成策略，有效解决了高分辨率点云生成中的参数效率与几何细节平衡问题。该网络采用了一种新的级联式解码结构，实现了先构建全局几何再细化局部细节的渐进式补全模式。该网络的具体工作流程如图2-9所示：

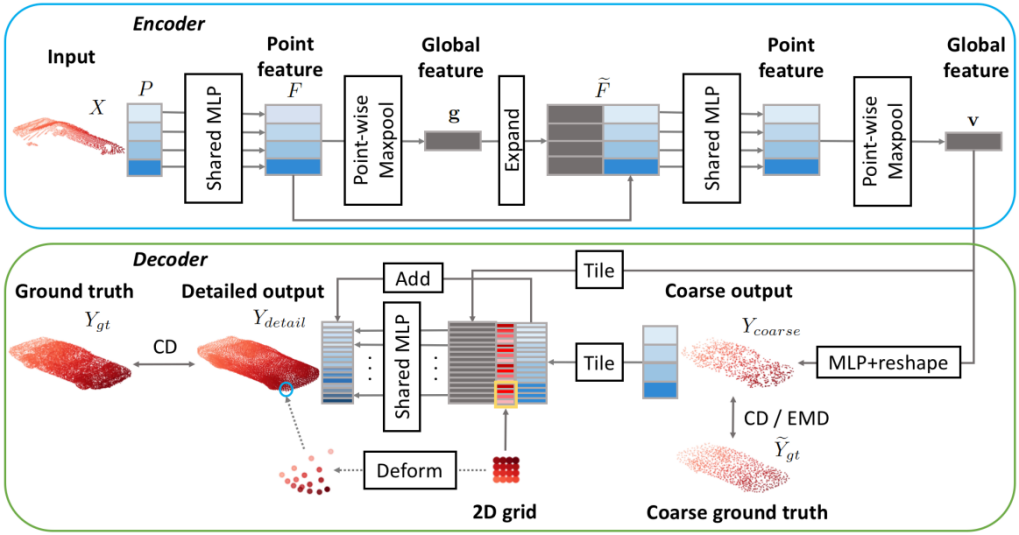


图 2-9 PCN 架构：由粗到细的两阶段补全模式

编码器将输入点云抽象为一个特征向量，首先将其送入一个全连接解码器生成一个粗略的稀疏点集，此阶段捕捉出形状的拓扑骨架。随后采用参数化折叠操作，以每个粗粒度点为中心生成局部曲面，从而构成一个详细的点云输出。该方法有效结合了全连接解码器与折叠操作的互补优势，在保证参数效率的同时实现了高分辨率输出。

作者团队通过实验，表明该网络具有在粗阶段通过低维流形约束保证拓扑合理性，在细阶段利用局部坐标系的仿射不变性增强细节重建以及通过共享参数有效抑制训练过拟合的优势。这种层级生成范式为点云补全任务提供了新的建模框架，在保持生成质量的同时显著降低计算复杂度。

## 本章小结

本章系统构建了点云补全的理论框架与技术体系。首先对三维点云数据的基础信息和相关理论展开介绍，其中包括其核心特性、点云数据的来源与表示形式，指出了在点云领域常用的评估指标，并重点探讨了稀疏点云的数学表征与信息熵特性，从信息论视角量化了稀疏性对几何重建的固有影响，为算法设计提供了理论指导。然后详细介绍了在作者设计的点云处理模型中会用到的基础方法，从深度学习视角切入，详细解释了Transformer的模型结构，重点阐释了MLP、DGCNN及FoldingNet的核心思想与处理流程，通过剖析网络结构设计揭示了点云特征学习与拓扑重建的内在机制。本章内容为后续提出的点云补全算法奠定了理论框架与技术支撑。

# 面向极稀疏点云的双路特征增强渐进式补全框架

## 模型总体架构

本研究针对极稀疏点云场景补全任务，提出了一种高效的补全框架双路特征增强渐进式补全框架，该模型在Transformer架构基础上引入了双路特征提取和多阶段渐进式细化策略，有效解决了极稀疏点云重建中的几何信息不足问题。模型整体架构由四个关键模块组成：双路特征提取器、Transformer编码-解码器、全局特征增强融合模块以及多阶段渐进式细化网络。

首先，利用一个双路特征提取器接受输入的稀疏点云，通过并行的局部和全局特征提取分支，同时捕获点云的局部几何结构和全局形状信息，生成增强特征。再使用轻量级边缘特征提取器DGCNN 提取输入点云周围的局部特征，并将中心点与其特征融合成一组点代理向量。其次，基于Transformer的编码-解码架构处理这些点代理向量。编码器采用自注意力机制建立点之间的全局关联，而解码器则利用可学习的查询嵌入和交叉注意力机制生成一个粗糙的中心点云坐标以及预测的缺失位置点代理向量。随后，全局特征增强模块根据缺失位置点代理提取出全局描述向量，并将其与点代理特征、粗糙点云坐标以及双路特征提取的增强特征融合，生成包含丰富语义和几何信息的重建特征。最后，多阶段渐进式细化网络基于折叠操作原理，逐步细化点云，将网络初始输入与多阶段生成的点云作为补全结果一起输出。整体网络架构如图 3-1所示。



图 3-1 网络结构图：双路特征增强渐进式补全框架

## 点云数据预处理

本研究创新性地提出了双路特征增强补全框架应用于极稀疏的点云场景下，但由于现有公开数据集缺乏真实的极稀疏点云样本，作者设计了一套灵活的采样率控制机制，通过对标准数据集(ShapeNet、PCN、KITTI)实施受控下采样来模拟极稀疏场景。

### 稀疏点云生成策略

采样率参数化设计实现对稀疏场景点云进行模拟，同时允许精确控制输入点云的稀疏程度，便于系统性评估模型在不同稀疏度下的性能表现。具体而言，作者引入了可配置的稀疏比率参数sparse\_ratio，利用该参数来调整保留原始点云的比例。

本研究采用FPS算法作为核心采样策略生成稀疏场景的点云输入。FPS算法通过迭代选择距离已选点集合最远的点，确保采样点均匀分布于原始形状表面，有效保留了原本数据集部分点云的轮廓和主要结构信息。进行采样后的稀疏点云输入效果图如图3-2所示。

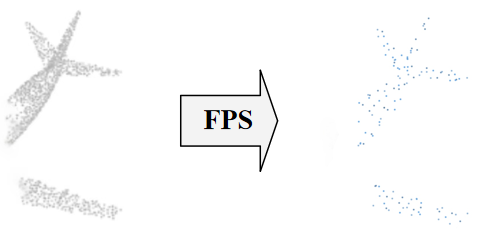


图 3-2 稀疏点云生成示例

### 零点填充与不完整点云模拟

深度学习模型通常要求固定维度的输入张量，这对点云处理带来挑战，特别是在处理形状各异、点数不一的点云时。本研究的模型输入同样也要求点云数量固定，为有效处理极稀疏点云数据，作者提出了零点填充机制，该方法不仅解决了输入数据不一致的问题，还为模型提供了对缺失区域的明确标识。零点填充机制的具体措施是在对下采样后的稀疏点云中，添加坐标值全为零的点，使输入点云维持固定大小，同时明确标记数据缺失区域。该机制保证了输入维度的一致性，并由此来模拟稀疏场景下的原始点云输入。

## 双路并行特征提取模块

针对极稀疏点云补全任务中信息严重不足的挑战，本研究设计了一个双路并行特征提取模块，采用了双路并行结构，分别从全局和局部两个不同角度提取点云特征，并通过跨分支注意力机制实现特征融合，最大化利用极稀疏点云中的有限几何信息。

### 局部特征分支设计

局部特征分支将偏移注意力机制(Offset Attention)[24]装载在两层堆叠架构[25]中捕获有限的点云当中的局部几何结构信息，将输入点云坐标转换为高维特征表示，为后续点云补全任务提供丰富的局部几何知识。局部特征分支从输入点云坐标开始，通过一系列特征转换和增强操作，构建多层次的特征表示，该分支的具体架构如图3-3所示。

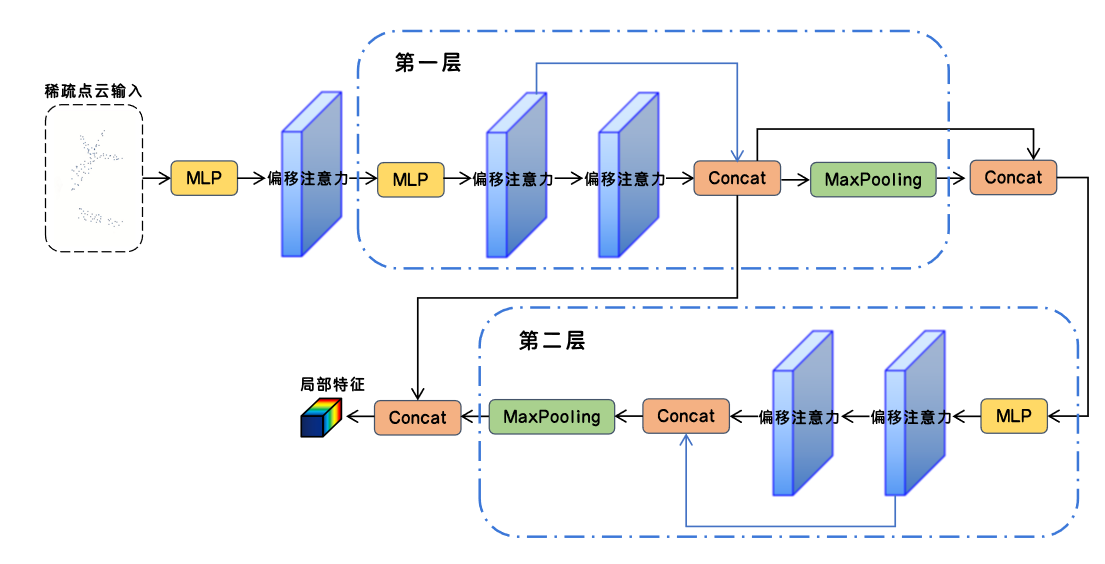


图 3-3 局部特征分支结构图

首先，原始的3D点云坐标采用初始卷积映射到高维特征空间，随后立即为初始特征向量应用偏移注意力实现局部几何结构特征的增强得到。通过这一操作，网络建立了点之间的初始关联性，使每个点能够感知其局部邻域的几何环境。随后特征经过两层卷积网络进一步进行特征提取得到。接下来，将特征送入一个创新性的级联注意力处理策略当中，连续应用两层偏移注意力机制网络，将这两层注意力的特征输出进行拼接整合后，形成增强特征表示。这种级联的设计结构使网络能够捕获不同范围的局部关系——第一层主要关注直接邻域，第二层则扩展到更广的上下文。拼接操作保留了这两种互补的局部结构信息，丰富了特征表达。下一步为了增强局部特征与全局上下文的关联，将该送入全局池化层中提取全局描述符，并扩展到每个点与与局部特征融合得到。这一设计使每个点的特征既包含局部几何信息，又融合了全局形状上下文，特别有利于极稀疏场景下的形状理解，因为全局信息可以弥补局部信息不足的缺陷。此后再重复上一步骤将特征维度进一步拉伸。最后将此高层全局特征和中层特征拼接，形成多尺度的输出特征。这种多尺度表示融合了不同抽象层次的几何信息——中层特征捕获基本局部结构，高层特征编码更抽象的语义概念。通过整合这两类信息，网络获得了既包含具体几何细节又包含抽象语义知识的综合表示，为点云补全任务提供了全面的特征基础。

该分支创新性地引入了偏移注意力模块从而增强局部结构的感知能力，其通过计算特征与注意力映射的偏移量突出局部几何变化。输入特征后，进行注意力映射计算，具体公式如3-1至3-6所示：

（3-1）

（3-2）

（3-3）

（3-4）

（3-5）

（3-6）

式中：——注意力能量；——归一化权重；( )——注意力输出。

做足如上准备后即可进行注意力偏移的特征变换映射工作，具体公式如3-7所示：

（3-7）

式中：( )——。；——1D卷积（核大小=1）；——批归一化；——GELU激活函数。

经过此偏移注意力函数，对稀疏点云的局部集合差异进行显式建模，构建了对极稀疏点云局部几何结构高度敏感的特征表示，特别在捕获边缘、曲率等关键几何特征方面表现优异。

### 全局特征分支实现

全局特征分支则专注于捕获点云的整体形状信息，同样也采用了类似局部特征分支的两层堆叠架构，有效提取极稀疏点云的全局几何特征。该分支同样以点云坐标为输入，但侧重于提取全局上下文信息。

该分支大体的工作流程分为：局部编码-全局聚合-特征融合-高维提取的模式。在第一层处理中，输入的稀疏点云通过MLP映射至高维特征空间得到。随后的处理与局部分支不同，全局分支不包含注意力模块，而是直接通过最大池化操作获取初始全局特征。那么此分支的核心在于将局部和全局的特征进行融合，将全局特征扩展复制到每个点特征，形成增强的特征矩阵。随后，通过第二层MLP将特征扩展到更高维度。最终通过最大池化提取高层全局描述符，并与中层特征拼接形成多尺度的输出特征。该分支具有高效的特征聚合和多尺度表示能力的核心优势，能够在极稀疏条件下有效捕获点云的整体形状信息。相比于局部分支，全局分支更侧重于提取点云的宏观结构和语义特征。

### 特征融合与增强策略

为了结合全局形状信息和局部集合细节，在双路特征并行提取完成后，全局特征向量和局部特征向量需要进行融合，本研究采用交叉注意力机制实现两路特征的自适应融合，特征融合示意图如图3-4所示。

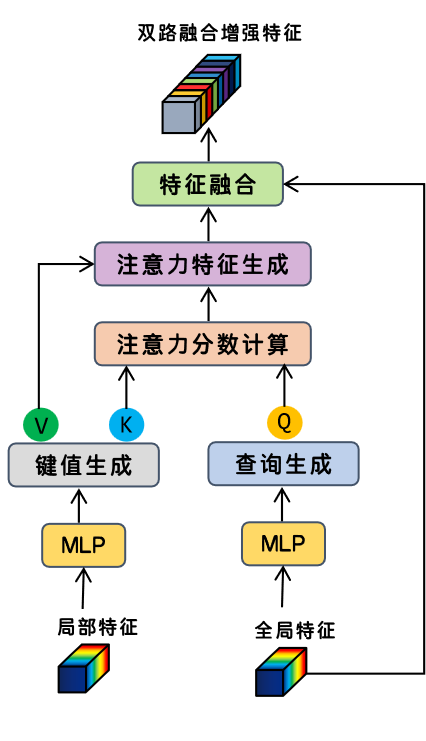


图 3-4 双路特征融合与增强流程图

其中交叉注意力机制则是将全局特征向量作为查询，局部特征向量作为键和值，实现跨分支的信息交互，最后融合输出同时携带全局-局部信息的双路融合增强特征。这种融合策略不仅关注到了点云的宏观结构，还吸收了边缘、曲率等关键局部特征。通过交叉注意力机制，模型能够根据任务需求自适应地关注不同特征分支的信息，生成既保持全局一致性又包含局部细节的综合特征表示。

## Transformer编码-解码核心架构

Transformer编码-解码核心架构是本研究的中枢组件，负责将稀疏点云输入的三维坐标转换为包含丰富语义和几何信息的点代理序列。该架构基于Transformer的序列转换思想，首先将稀疏的点云输入转换为一系列点代理序列，这些特征序列代表点云中的各个局部区域，自然地将点云补全任务重构为一个集合到另一个集合的转换问题——观测特征序列到点代理序列的映射问题，有效利用了注意力机制在建立远距离依赖关系方面的优势。

Transformer编码-解码架构处理的是一个集合到另一个集合的映射，那么首要工作就是将稀疏的点云输入抽象成一组携带点云几何信息的向量集合——点代理序列。首先对输入点云使用KNN算法确定其局部邻域的点集，随后采用类似DGCNN中EdgeConv的轻量级边缘特征提取机制，计算邻域点相对于中心点的偏移量，使用共享的三层MLP处理这些边缘特征，最终点代理序列通过残差连接融合语义与位置信息，以此作为Transformer模块的输入：

（3-8）

式中：——三层MLP。

与传统Transformer不同，本架构引入了针对点云处理的几何感知增强设计，包括边缘特征聚合、相对位置编码和形状先验约束，显著提升了模型对三维形状的理解能力。整体架构接收双路特征提取的输出，生成点代理特征和粗糙点云坐标，为后续的多阶段细化网络提供结构化的形状表示，实现从极稀疏观测到完整点云的高质量重建。

### 自注意力编码器设计

自注意力编码器由六块自注意力层的堆叠构建而成，每个块包含多头自注意力、前馈网络及残差连接，负责从点代理序列中提取全局语义特征和几何关系。通过自注意力机制建立点代理间的全局关联，同时融合几何敏感的设计，总结出三维点云的空间结构特性。

如图3-5所示，与传统的Transformer编码器不同，本研究的编码器引入了并行的局部几何处理分支，其输入分为两部分，一是编码器的传统输入，通过自注意力机制捕获全局语义关系；二是加入的点代理序列的邻域特征，学习局部几何结构。

随后将二者进行拼接后通过前向反馈网络，输出经过几何增强处理后的点代理特征序列。这些特征既包含通过自注意力捕获的全局语义关系，也保留了通过几何处理分支获取的局部结构信息。

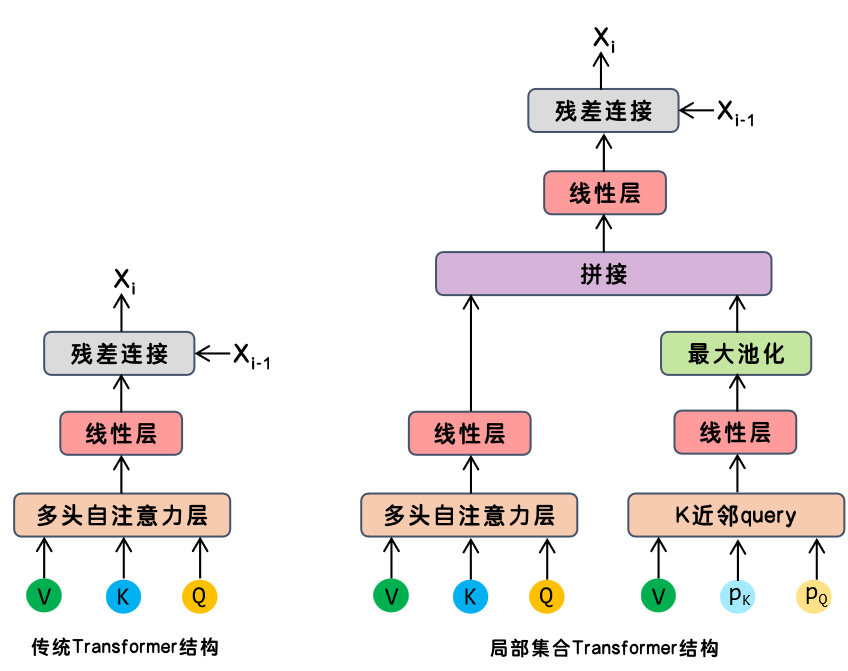


图 3-5 两种Transformer结构对比图

### 动态Query生成器

在编码器输出与解码器输入之间，本研究使用动态查询生成器进行过渡处理，生成流程如图3-6所示。该生成器首先对编码器输出特征进行池化提取全局特征，然后将该全局特征通过一个MLP映射为初始的粗糙点云坐标，以此作为待预测的缺失点云的中心点坐标。最后，将这些坐标与扩展的全局特征拼接，通过MLP生成初始查询向量，作为解码器的输入，具体计算过程如3-9所示：

（3-9）

这一设计使查询向量能够自适应地依据输入形状调整，而非使用固定查询。

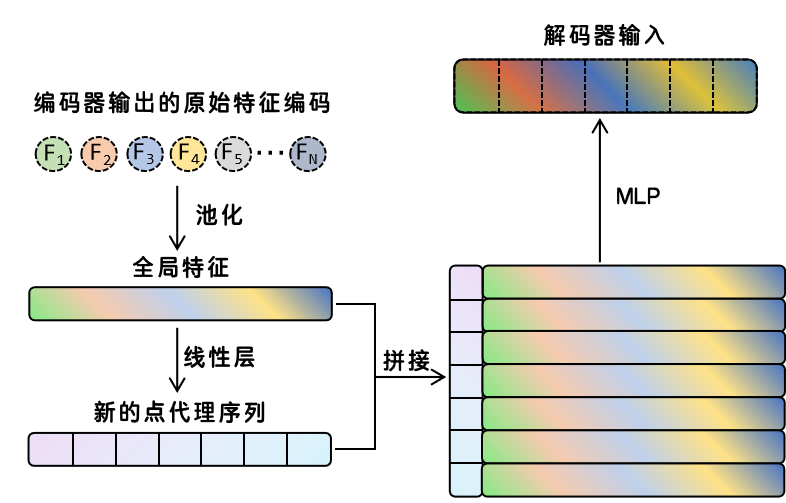


图 3-6 动态Query生成流程图

### 基于形状先验的解码器

解码器则采用八层的深度设计，接受两类输入：一是来自动态查询生成器的初始查询向量，作为点代理的初始状态；二是编码器输出的特征序列，用于交叉注意力机制中提供上下文信息。解码器在自注意力阶段维护可学习的位置编码，并通过交叉注意力融合编码器特征：

（3-10）

（3-11）

（3-12）

（3-13）

式中：——解码器层的输入查询向量。

（3-14）

（3-15）

（3-16）

（3-17）

式中：——解码器层的输入查询向量；——编码器的输出特征。

在此之后，解码器引入了几何处理分支，利用查询向量的坐标估计构建KNN邻域，提取局部几何特征，侧重于捕获查询间的空间关系。最后，解码器将自注意力、交叉注意力和几何处理的结果融合，生成更新的查询向量。然后通过前馈网络进一步转换特征，同时采用残差连接和层归一化来稳定训练过程，使查询向量通过层层细化逐步转变为准确的点代理表示。解码器的最后一部分为两个独立的MLP，分别生成更新的点代理序列和粗糙的点云坐标。点代理序列包含丰富的语义和几何信息，用于后续的点云生成；粗糙点云坐标则是对缺失区域形状的初步估计，提供了基本的空间结构框架。

## 全局特征增强与融合模块

全局特征增强与融合模块是连接Transformer编码-解码架构与后续点云细化网络的重要桥梁，采用了创新地多元特征融合的策略，将Transformer解码器输出、与双路特征提取模块的结果进行深度整合，构建更全面的重建特征表示。

首先，将粗略点云的坐标通过MLP拉伸到更高维度，随后，通过最大池化操作从该高维特征中提取全局描述向量。接下来，将、、、拼接，该操作相比于原来的维度有了大大的提升，提供了更丰富的特征表示。随后通过MLP映射回1024的维度，保持特征维度一致性的同时保留多源信息。

该模块实现了跨阶段特征互补，将Transformer解码阶段的高级语义特征与双路特征提取阶段的低级几何特征进行融合。这种设计创造了一个信息流闭环，使后期重建过程能够直接访问初始输入点云的几何信息，特别是在极稀疏场景下，这种回溯连接对保持几何细节至关重要。然后，这个模块的创新策略增强了模型的容错能力。当Transformer解码过程中可能出现信息丢失或偏差时，来自双路特征提取的自适应特征提供了额外的几何约束，帮助模型生成更符合原始形状特性的点云。这种冗余设计显著提高了模型在处理复杂形状时的稳定性。这种多角度、多层次的综合表示为后续的点云细化提供了理想基础。

## 多阶段渐进式点云细化网络

有了如上的一系列前置工作，现可以对稀疏点云进行补全工作。本研究设计的多阶段渐进式点云细化网络是基于传统FoldingNet的创新扩展，通过分阶段细化策略实现从粗糙形状到精细点云的渐进式重建，接收重建特征和粗糙点云坐标作为输入，通过多阶段折叠操作，每个阶段负责特定精度的形状恢复，最后将各阶段的点云进行拼接，生成高质量的完整点云。

第一阶段是初步形状的构建。网络接收融合模块输出的重建特征和粗糙点云，以粗糙点云为基础，通过fold\_step×fold\_step的二维网格为每个点代理生成局部点集。这一阶段主要恢复物体的整体结构和主要组件，生成约M×fold\_step²个点，形成初步形状。

第二阶段完成结构优化与加密。传入上一阶段的输出，在保持几何一致性的同时进一步细化点云分布。该阶段增加点密度，改善表面连续性，并开始恢复一些中等尺度的几何细节，如明显的边缘和拐角。

最后阶段作为最终细化步骤，恢复精细的细节。该阶段在小尺度上的表面变化、纹理和微小结构方面进行重建，生成最终的高质量点云。

这样的模式相比单阶段生成，分阶段细化使网络能够先专注于形状的主要结构，再逐步恢复细节，降低了学习难度，使各阶段形成一个连贯的信息流，保持全局一致性的同时逐步细化局部细节，特别适合从极少量观测点推断完整形状的困难任务。

## 损失函数设计

### 多阶段监督机制

多阶段监督机制为整个模型的学习运转提供了全面且层次化的训练信号，充分考虑了点云重建过程中粗细两个阶段的精度要求，通过多尺度监督策略实现了对网络参数的高效优化。由于多阶段渐进式点云细化网络的输出特性，作者以3D点云中的倒角距离为基础，倒角距离的数学定义如3-18所示：

（3-18）

设置了包含粗糙监督和多阶段细化监督的联合损失函数，具体定义如3-19所示：

（3-19）

式中：——对Transformer解码器输出的粗糙点云的监督；——对多阶段细化网络各阶段输出的综合监督；——对应的权重系数。

接下来将对以上每项内容做出详细解释：

是来自于Transformer生成的初始点云骨架与目标点云之间的距离度量，这一损失函数促使模型学习目标形状的整体结构和主要几何特征，具体定义如3-20所示：

（3-20）

式中：——通过Transformer编码-解码器生成的粗糙点云，其中包含了M个点；——目标点云。

针对的是折叠操作后生成的高分辨率点云，这一损失函数引导模型关注局部细节重建的精确性，确保生成点云在形状细节上的准确表达，具体定义如3-21所示：

（3-21）

式中：——通过折叠网络细化后的点云，其中包含了N个点。

## 本章小结

本章主要提出了一种应用于极稀疏点云的基于双路特征增强的渐进式点云补全网络，并展开了详细介绍。本模型的创新性主要体现在以下三个方面：

##### （1）提出了一种新的用于点云补全任务的网络模型。模型使用分层式Transformer编码器-解码器架构为基础框架，将点云补全任务转换为已有集合到缺失集合的转换问题，使用自注意力机制和交叉注意力机制对点代理序列进行特征提取，并通过结合点云数据的上下文信息，帮助模型对点云的语义信息做更有效的特征提取工作，实现了高质量点云重建的目标。

（2）设计了双路特征并行提取模块。该模块同时从局部和全局两个不同的视角分析点云数据。局部特征提取通过改进的点卷积网络捕获点与点之间的局部几何关系，保留精细的表面细节；全局特征提取则应用Transformer编码器提取全局上下文特征，理解整体形状结构。这两路特征通过特征融合策略有机结合，形成了既包含局部细节又具备全局一致性的综合表示。这种双路并行设计显著增强了模型对复杂几何形状的理解能力，尤其在处理具有丰富细节的点云数据时表现出色。

（3）采用多阶段渐进式点云细化网络。不同于传统的一步到位生成策略，模仿PCN由粗到细的补全模式，本模型通过分阶段、渐进式地细化点云，实现了从粗糙到精细的形状重建过程。每个细化阶段都利用前一阶段的输出以及原始不完整点云的特征信息，确保生成的点云在保持全局形状一致性的同时，逐步恢复精细的局部细节。这种渐进式细化策略有效降低了点云补全任务的复杂度，显著提高了重建质量，在处理信息缺乏的极稀疏点云时，展现出明显优势。

# 实验设计与结果分析

为了验证本文提出的点云补全网络的有效性，本研究在ShapeNet-55数据集上进行了全面的实验评估，并与多个主流方法进行了比较，同时通过对度量指标进行数值评估和对补全点云进行可视化来分析模型的性能。

## 实验环境设计

### 数据集介绍

为了全面评估模型的性能，本研究选择了ShapeNet-55数据集[27]进行实验，并利用3.2节所述的数据预处理方法在所有数据集上将原始缺失点云数据下采样模拟稀疏点云的输入场景。

ShapeNet-55数据集是目前点云补全领域最全面的基准数据集，它基于ShapeNetCore构建，总共有55个类别的3D点云模型，包含了ScanNet[28]和S3DIS[29]等无法从真实世界数据集中获得的完整物体模型。ShapeNet-55包含总计52470个点云形状，采用4:1的比例对数据集进行分割，得到包含41952个点云形状的训练集和包含10518 个点云形状的测试集。与现有的点云补全数据集（如PCN）相比，ShapeNet-55考虑了更多样化的任务，包括点云上采样和补全。其次，它覆盖了更广泛的类别，从8个类别扩展到55个类别。第三，它采用了更多样化的视角，从8个固定视角扩展到所有可能的视角。最后，它包含了更多样化的不完整程度，缺失真实点云的25%至75%的点。

为了全面评估模型的性能，我们选择了10个具有代表性的类别进行详细分析。这些类别根据训练样本数量分为两类：样本充分类别（训练样本超过2500个）和样本不充分类别（训练样本少于80个）。样本充分类别包括桌子、椅子、飞机、汽车和沙发，这些类别具有丰富的训练数据；样本不充分类别包括禽舍、背包、遥控器、键盘和火箭，这些类别训练数据相对稀缺。这种选择有助于我们分析样本数量对模型补全效果的影响。

### 训练设置

#### 硬件环境设置

本实验使用2张显存为24GB的英伟达（NVIDIA）RTX 4090显卡进行训练和测试。由于点云补全任务需要大量计算资源，特别是对于ShapeNet-55等大规模数据集，RTX 4090的24GB显存能够支持较大的批处理大小，从而加速训练过程并提高模型收敛效率。

#### 软件环境设置

软件方面在Linux（版本Ubuntu 22.04）操作系统下使用Python（Python版本3.9）作为编程语言，采用Pytorch（Torch版本 1.13.1）深度学习框架。PyTorch是一个基于Python的深度学习框架，可以作为 Python 库来使用。它提供了一套 Python 接口，包括各种模型、层、优化器、损失函数等，通过它可以方便地用Python语言构建和训练深度学习模型。同时，使用并行计算模型CUDA（Compute Unified Device Architecture，CUDA 版本11.8）来加速张量计算。所有模型训练都在相同的硬件和软件环境中进行，以确保实验结果的公平性和可比性。

#### 训练参数设置

神经网络参数设置对深度学习模型的训练效果、泛化能力和训练效率有着重要的影响。在以上的硬件和软件环境下，本实验采用以下训练配置：模型采用AdamW优化器进行训练，初始学习率设置为0.0005，权重衰减系数为0.0005，这种配置能够有效防止模型过拟合。学习率调度采用LambdaLR策略，每21个epoch进行一次衰减，衰减率为0.76，最低学习率限制为初始学习率的0.02倍。Batch Normalization层采用动态调整策略，每21个epoch进行一次衰减，衰减率为0.5，动量系数为0.9，最低衰减限制为0.01。本研究的模型目的是应用在稀疏的点云场景中，为了模拟稀疏点云的输入场景，设置相比于原本残缺点云输入的稀疏度为0.5。训练过程中采用64的批次大小，每个批次更新一次参数，最大训练轮数设置为200轮。这些参数来源于先验知识，可以改善模型在ShapeNet-55数据集上的点云补全效果，为点云补全任务提供了更优的解决方案。

## 模型对比

在本实验中，作者将多个主流方法在ShapeNet-55数据集上进行训练和测试，与本文提出的新方法进行对比分析补全效果。

### 度量指标对比

正如第二章所介绍，倒角距离（CD）是点云补全任务中最常用的评估指标。它测量生成的点云与真实点云之间的平均最近点距离。较低的CD值表示生成的点云更接近真实点云。作者同时使用了一阶和二阶倒角距离（CDL1、CDL2），其计算公式可见2.2.1节公式2-5。

此外，本实验还使用了F-Score作为参考指标。F-Score是精确度和召回率的调和平均值，用于评估生成点云的完整性和准确性。F-score的取值范围是0到1，较高的F-Score表示模型生成的点云既准确又完整。

下列三表展示了本研究的改进模型与多个主流模型（PoinTr、PCN、FoldingNet、GRNet）在ShapeNet-55数据集上10个代表性类别的性能对比。

#### CDL1

表4-1 本文模型与其他方法对比表现表-CDL1

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 样本充分类别 | | | | | | 样本不充分类别 | | | | | | 全部类别 |
| 类别  模型 | 桌  子 | 椅  子 | 飞  机 | 汽  车 | 沙  发 | CDL1-Avg | 耳  机 | 帽  子 | 鸟  窝 | 键  盘 | 火  箭 | CDL1-Avg | CDL1-Avg |
| FoldingNet | 25.641 | 28.070 | 23.070 | 26.218 | 26.894 | 25.9786 | 40.550 | 30.198 | 34.125 | 21.538 | 21.721 | 29.6264 | 27.735 |
| PCN | 22.854 | 24.440 | 19.880 | 23.973 | 24.337 | 23.0968 | 28.895 | 26.736 | 31.368 | 18.111 | 19.773 | 24.9766 | 24.500 |
| GRNet | 19.914 | 20.989 | 17.791 | 21.480 | 21.581 | 20.351 | 23.837 | 20.748 | 24.201 | 17.308 | 17.733 | 20.7654 | 21.259 |
| PoinTr | 14.919 | 16.127 | 13.295 | 17.032 | 16.161 | 15.5068 | 19.493 | 15.325 | 18.904 | 12.440 | 13.836 | 15.9996 | 16.215 |
| **Mine** | 13.025 | 13.789 | 11.259 | 14.937 | 13.992 | 13.4004 | 16.661 | 13.766 | 15.775 | 11.188 | 11.252 | 13.7284 | 13.981 |

#### CDL2

表4-2 本文模型与其他方法对比表现表-CDL2

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 样本充分类别 | | | | | | 样本不充分类别 | | | | | | 全部类别 |
| 类别  模型 | 桌  子 | 椅  子 | 飞  机 | 汽  车 | 沙  发 | CDL2-Avg | 耳  机 | 帽  子 | 鸟  窝 | 键  盘 | 火  箭 | CDL2-Avg | CDL2-Avg |
| FoldingNet | 2.412 | 2.746 | 1.877 | 2.239 | 2.399 | 2.3346 | 7.512 | 3.507 | 3.863 | 1.430 | 1.633 | 3.589 | 2.735 |
| PCN | 1.994 | 2.117 | 1.429 | 1.924 | 1.960 | 1.8848 | 3.048 | 2.612 | 3.558 | 1.068 | 1.290 | 2.3152 | 2.221 |
| GRNet | 1.376 | 1.485 | 1.075 | 1.413 | 1.422 | 1.3542 | 2.143 | 1.332 | 1.807 | 0.849 | 1.025 | 1.4312 | 1.618 |
| PoinTr | 1.006 | 1.099 | 0.760 | 1.090 | 1.045 | 1 | 1.653 | 0.960 | 1.515 | 0.560 | 1.031 | 1.1438 | 1.153 |
| **Mine** | 0.750 | 0.809 | 0.544 | 0.822 | 0.774 | 0.7398 | 1.207 | 0.808 | 0.999 | 0.406 | 0.575 | 0.799 | 0.859 |

#### F-Score

表4-3 本文模型与其他方法对比表现表-F-Score

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 样本充分类别 | | | | | | 样本不充分类别 | | | | | | 全部类别 |
| 类别  模型 | 桌  子 | 椅  子 | 飞  机 | 汽  车 | 沙  发 | F-Score  Avg | 耳  机 | 帽  子 | 鸟  窝 | 键  盘 | 火  箭 | F-Score  Avg | F-Score  Avg |
| FoldingNet | 0.128 | 0.094 | 0.168 | 0.096 | 0.094 | 0.116 | 0.061 | 0.073 | 0.066 | 0.149 | 0.186 | 0.107 | 0.108 |
| PCN | 0.183 | 0.160 | 0.256 | 0.137 | 0.139 | 0.175 | 0.125 | 0.128 | 0.103 | 0.234 | 0.275 | 0.173 | 0.172 |
| GRNet | 0.219 | 0.208 | 0.296 | 0.172 | 0.178 | 0.2146 | 0.189 | 0.182 | 0.138 | 0.227 | 0.313 | 0.2098 | 0.212 |
| PoinTr | 0.376 | 0.345 | 0.434 | 0.298 | 0.324 | 0.3554 | 0.280 | 0.339 | 0.289 | 0.419 | 0.454 | 0.3562 | 0.350 |
| **Mine** | 0.426 | 0.406 | 0.514 | 0.343 | 0.378 | 0.4134 | 0.329 | 0.403 | 0.340 | 0.456 | 0.521 | 0.4098 | 0.408 |

表中选取了10种具有代表性的类别展示结果。其中，桌子、椅子、飞机、汽车和沙发类别在训练集中样本数量超过500个，而耳机、帽子、鸟窝、键盘和火箭类别的样本数量不到10个。本文选取这些类别分别代表样本充分和样本不充分的情况，从而探究样本数量对补全效果的影响。三个表格分别展示了CDL1、CDL2和F-Score三个描述各类别补全效果的重要指标，以及选取的代表性类型的平均结果和所有类别的平均结果。从以上三表可以看出，改进模型在所有类别上都取得了显著的提升。

首先从倒角距离的两种计算范式对模型性能进行评估（表4-1、4-2）。在CDL1指标中，本文模型以平均13.981的显著优势领先，较次优模型PoinTr（16.215）降低12.6%。对于CDL2，模型以0.859的平均值优于对比方法，较PoinTr（1.153）降低了25.5%。这种差异源于双路特征提取模块对几何误差分布特性的适应能力：CDL1对局部大尺度偏差敏感的特性（如飞机的机翼部位的11.259对比PoinTr 13.295），然而CDL2更关注整体的结构分布（如沙发的曲面1.045对比PoinTr 0.774）的优化，这个结果同时也验证了渐进式细化网络对误差传播的有效控制。

其次如表4-3所示，本文模型在F-Score指标上达到0.408的平均值，较最优对比模型（PoinTr 0.350）提升16.5%，对比其他主流模型也表现出了优异的补全能力。对于样本充分和不充分的类别都表现出了领先的补全效果，这得益于多阶段细化网络对部件连接关系的补全方法。

综合以上实验结果，本研究的改进模型相较于其他主流方法，在点云补全任务中展现出了显著的优势，表现出了更好的性能和更强的鲁棒性。这些结果表明，本研究的改进策略有效地提升了模型的特征提取能力和几何推理能力，使其能够更好地处理不同场景下的点云补全任务。

### 可视化对比

为了直观展示不同模型的补全效果，最终选择了ShapeNet-55数据集中的三类具有代表性的物体进行可视化对比：具有复杂几何结构的物体（如飞机）、缺失关键信息的物体（如椅子）和表面光滑的物体（如沙发），具体补全结果如图4-1所示：

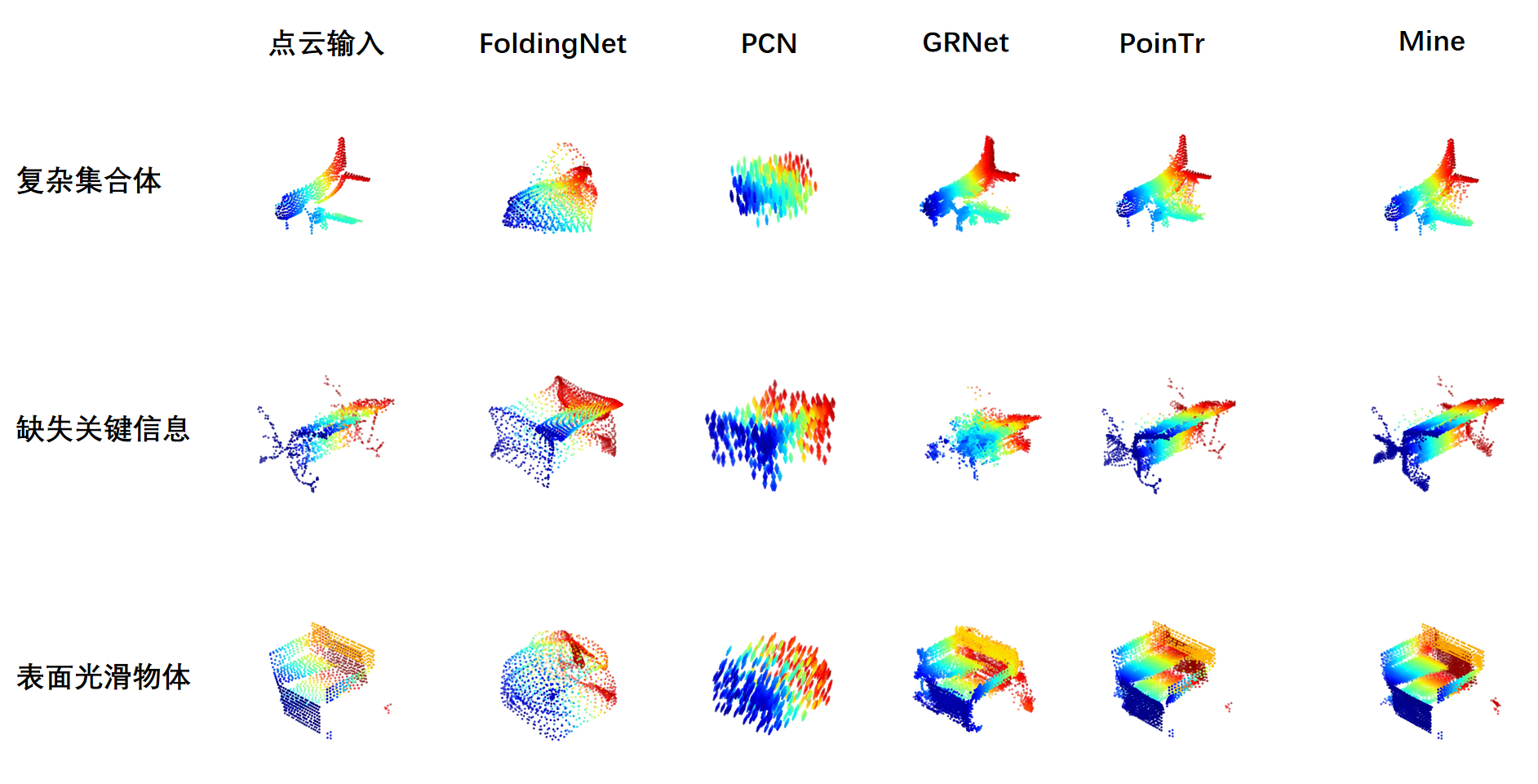


图 4-1 部分补全结果二维示意图

以上展示结果印证了量化分析的结论，图中从左至右依次展示了点云输入、FoldingNet、PCN、GRNet、PoinTr以及本文的方法（Mine）的补全结果。从实验结果（图 4-1）可以看出，FoldingNet和PCN无法适应稀疏点云的工作场景，那么之后这两个模型的补全结果不再参与对比讨论。其次本文方法的补全效果显著优于主流算法 GRNet、PoinTr。

在处理结构复杂的点云时，GRNet和PoinTr能够较好地恢复整体形状，但在局部细节上仍有缺失。相比之下，本文的方法能够更准确地还原复杂结构，点云分布更加均匀，细节表现更为丰富。当输入点云缺失关键结构时，显然GRNet无法处理缺乏关键信息的情况，输出的点云较为混乱。而PoinTr在整体形态上有所改善，但仍有部分细节缺失。本文的方法能够有效推断缺失部分，补全结果与真实结构更为接近，表现出更强的鲁棒性。对于表面光滑的物体，GRNet和PoinTr在点云密度和表面连续性上有所提升。本文的方法在点云分布的均匀性和表面光滑性方面表现最佳，能够更好地还原物体的真实表面特征。

总而言之，通过结果对比可以得出相较于其他主流方法，本文的方法对极稀疏场景的点云补全具有更强的适应能力，能够处理输入点云更少的极端情况，生成更为完整且细节丰富的点云，验证了方法的有效性和泛化能力。

## 消融实验

为了验证本文提出的双路特征提取模块和多阶段渐进式点云细化网络的有效性，作者进行了全面的消融实验来评估模块对模型的影响。具体设计如下：

1. 移除双路特征提取模块，仅使用原始的特征提取器。
2. 移除多阶段渐进式点云细化网络，替换为传统的轻量型FoldingNet结构。

上述实验方案基于原本的Transformer框架，在原本的ShapeNet-55数据集上进行训练，并同样采用CDL1、CDL2和F-Score作为评估指标，得到如表4-4所示的量化对照结果，以及如图4-2所示的可视化对比图。

表 4-4 消融实验结果表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 双路特征提取模块 | 多阶段渐进式点云细化网络 | CDL1 | CDL2 | F-Score |
| 实验1 | √ |  | 15.221 | 1.024 | 0.373 |
| 实验2 |  | √ | 14.878 | 1 006 | 0.385 |
| 实验3 | √ | √ | 13.981 | 0.859 | 0.408 |

从实验结果可以看出，仅使用双路特征提取模块时，模型的CDL1和CDL2分别为15.221和1.024，F-Score为0.373，说明该模块能够有效提升点云补全的准确性和细节还原能力。仅使用多阶段渐进式点云细化网络时，CDL1和CDL2分别下降至14.878和1.006，F-Score提升至0.385，表明该模块有助于进一步细化点云结构，提高点云的完整性和分布均匀性。同时引入双路特征提取模块和多阶段渐进式点云细化网络后，模型在所有指标上均取得最佳表现，CDL1和CDL2分别降至13.981和0.859，相较于以上两个情况，CDL1分别降低了8.9%和6.4%，CDL2分别降低了16.1%和14.6%；F-Score提升至0.408，相较于前两个情况分别提高了9.4%和6.0%，显著优于单独使用任一模块的结果。

由此可见，双路特征提取模块和多阶段渐进式点云细化网络均对点云补全性能有积极作用，二者结合能够进一步提升模型的补全精度和点云质量，验证了所提模块设计的有效性。

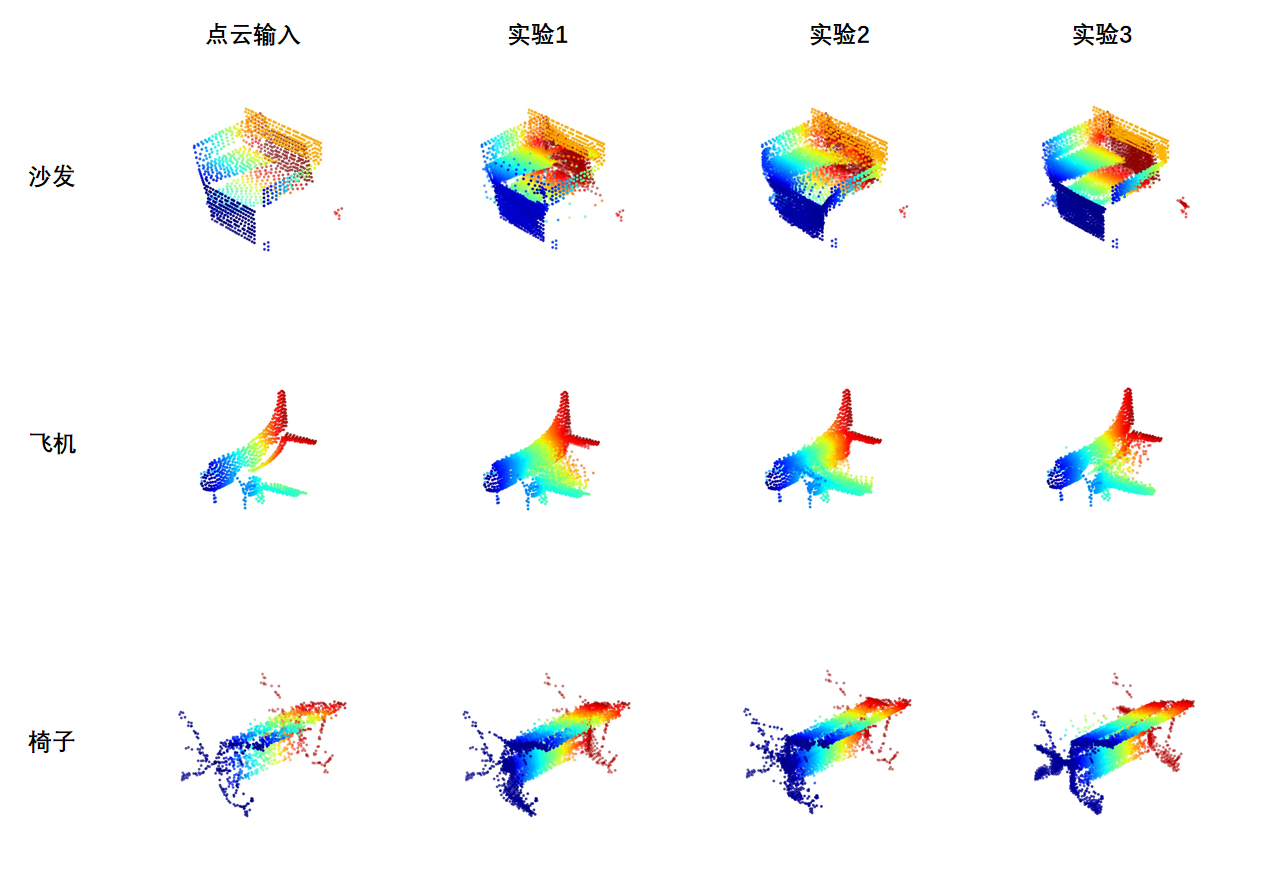


图 4-2 消融实验补全结果二维示意图

## 本章小结

本章的主要工作是对本研究提出的面向极稀疏点云的双路特征增强渐进式补全框架的改进效果进行实验论证。首先详细介绍了实验环境，包括硬件配置和软件环境，并针对ShapeNet-55数据集的特点设计了统一的训练策略，然后通过指标量化和可视化的方式对比了本文方法与原始模型在点云补全方面的表现。

# 结论与展望

## 结论

三维点云补全任务是计算机视觉领域的一个关键研究课题，其目标是利用有限的观测数据，对缺失或不完整的三维点云模型进行完整重建，从而实现对现实世界场景的精确建模与深入理解。本研究围绕以深度学习为基础的点云补全技术，系统性地构建了从基础理论到算法优化的研究框架，创新性地提出了面向极稀疏点云的基于Transformer的双路特征增强渐进式补全框架。本模型首先通过双路特征提取器并行捕获局部几何结构与全局形状语义。再使用DGCNN提取输入点云的边缘特征，结合多层感知机构建的全局分支生成增强特征表达。通过将稀疏的点云输入与邻域特征融合为点代理向量，模型构建了连接微观细节与宏观拓扑的中间表征。在此基础上，基于Transformer的编码-解码架构利用自注意力机制建立点间全局依赖关系，解码器通过可学习查询嵌入与交叉注意力机制生成粗糙中心点云坐标及缺失区域点代理向量，实现了局部观测与全局结构的跨域对齐。全局特征增强模块进一步融合点代理特征、粗糙坐标及双路分支的增强特征，生成包含多粒度几何语义的重建特征。最终通过多阶段折叠网络逐级优化点云分布，以迭代式坐标更新融合初始稀疏输入与各阶段生成结果，形成从粗粒度到细粒度的高质量补全输出。

综合实验结果表明，我们的改进模型在点云补全任务中展现出了显著的优势。无论是在处理简单几何结构还是复杂几何结构时，无论是在样本充分还是样本不充分的类别中，改进模型都表现出了更好的性能和更强的鲁棒性。这些结果充分验证了我们提出的改进策略的有效性，为点云补全领域的研究提供了有价值的参考。

本文的创新性工作为点云补全任务的研究提供了新的思路和方法，对点云处理领域具有重要的意义。

## 展望

本文提出的双路特征增强渐进式补全框架在极稀疏点云场景下能够有效地完成补全工作，相较于其他主流方法也取得了显著的性能提升，但实验结果表明，补全效果与完整点云仍存在一定差距，尤其在复杂几何结构和高频细节的恢复上仍有改进空间，未来可从以下几个方面展开深入研究：

1. 本文的模型计算复杂度随点云规模呈平方级增长，限制了其在移动设备或实时系统中的部署。针对这一问题，未来的研究可以进一步探索预训练和迁移学习在点云补全领域的应用。通过引入迁移学习策略，可以将在大规模相关任务上预训练得到的模型参数迁移到点云补全任务中，不仅有望显著缩短模型的训练时间，还能提升模型在不同数据分布下的泛化能力。此外，预训练模型能够为点云补全任务提供更丰富的先验知识，有助于模型更好地理解和还原复杂的三维结构。因此，结合预训练与迁移学习的方法将成为提升点云补全效率和性能的重要研究方向。
2. 可探索更高效的全局与局部特征融合机制，进一步增强模型对复杂几何结构的表达能力；还可以引入多模态信息（如图像、语义标签等）辅助点云补全，有望提升模型对缺失区域的推理能力。
3. 推动点云补全方法在真实场景下的应用落地，如自动驾驶、机器人感知等，也是未来值得关注的方向。

点云补全仍面临诸多挑战，本文的工作为稀疏点云补全提供了新的思路和方法，后续将继续致力于提升补全质量和实际应用价值。

# 致 谢

# 参考文献

1. Mihir Garimella, Prathik Naidu. Beyond the pixel plane: sensing and learning in 3D[J/OL]. The Gradient,2018[2025-03-26].https://thegradient.pub/beyond-the-pixel-plane-sensing-and-learning-in-3d.
2. YINT W， ZHOUX Y， KRÄHENBÜHLP. Center-based 3D object detection and tracking［C］. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition （CVPR）. Nashville， TN， USA. IEEE， 2021： 11779-11788. doi: 10.1109/cvpr46437.2021.01161
3. ASSAELY， SOMMERSCHIELDT， PRAGJ. Restoring ancient text using deep learning：a case study on Greek epigraphy［C］. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing （EMNLP-IJCNLP）.Hong Kong， China. Stroudsburg， PA， USA： Association for Computational Linguistics， 2019. doi: 10.18653/v1/d19-1668
4. 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.
5. Qi, C., Yi, L., Su, H., & Guibas, L.J. (2017). PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. ArXiv, abs/1706.02413.
6. Mitra N J, Pauly M, Wand M, et al. Symmetry in 3D geometry: Extraction and applications[C]//Computer graphics forum. 2013, 32(6): 1-23.
7. 温佩芝,雷永庆,孙梦龙.三维重建网格模型的缺陷孔洞识别与修复方法[J].计算机应用研究,2020,37(04):1234-1238.DOI:10.19734/j.issn.1001-3695.2018.09.0773.
8. Li Y, Dai A, Guibas L, et al. Database‐assisted object retrieval for real‐time 3d reconstruction[C]//Computer graphics forum. 2015, 34(2): 435-446.
9. Sharma A, Grau O, Fritz M. Vconv-dae: Deep volumetric shape learning without object labels[C]//Computer Vision–ECCV 2016 Workshops: Amsterdam, The Netherlands, October 8-10 and 15-16, 2016, Proceedings, Part III 14. Springer International Publishing, 2016: 236-250.
10. Dai A, Ruizhongtai Qi C, Nießner M. Shape completion using 3d-encoder-predictor cnns and shape synthesis[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 5868-5877.
11. Hu T, Han Z, Shrivastava A, et al. Render4Completion: Synthesizing multi-view depth maps for 3D shape completion[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops. 2019: 0-0.
12. Yuan W, Khot T, Held D, et al. Pcn: Point completion network[C]//2018 international conference on 3D vision (3DV). IEEE, 2018: 728-737.
13. Yang, Y., Feng, C., Shen, Y., & Tian, D. (2017). FoldingNet: Interpretable Unsupervised Learning on 3D Point Clouds. ArXiv, abs/1712.07262.
14. Wang Y, Sun Y, Liu Z, et al. Dynamic graph cnn for learning on point clouds[J]. ACM Transactions on Graphics (tog), 2019, 38(5): 1-12.
15. Yu X, Rao Y, Wang Z, et al. Pointr: Diverse point cloud completion with geometry-aware transformers[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2021: 12498-12507.
16. Jelalian A V. Laser radar systems[C]//EASCON 1980; Electronics and Aerospace Systems Conference. 1980: 546-554.
17. Bula J, Derron M H, Mariethoz G. Dense point cloud acquisition with a low-cost Velodyne VLP-16[J]. Geoscientific Instrumentation, Methods and Data Systems, 2020, 9(2): 385-396.
18. Geng J. Structured-light 3D surface imaging: a tutorial[J]. Advances in optics and photonics, 2011, 3(2): 128-160.
19. MUBANGA K. What is Photogrammetry?[EB/OL]. [2025-02-28]. <https://www.artec3d.cn/learning-center/what-is-photogrammetry.>
20. Foix S, Alenya G, Torras C. Lock-in time-of-flight (ToF) cameras: A survey[J]. IEEE Sensors Journal, 2011, 11(9): 1917-1926.
21. Endres F, Hess J, Sturm J, et al. 3-D mapping with an RGB-D camera[J]. IEEE transactions on robotics, 2013, 30(1): 177-187.
22. Rubner Y, Tomasi C, Guibas L J. The earth mover's distance as a metric for image retrieval[J]. International journal of computer vision, 2000, 40: 99-121.
23. Wang Y, Sun Y, Liu Z, et al. Dynamic graph cnn for learning on point clouds[J]. ACM Transactions on Graphics (tog), 2019, 38(5): 1-12.
24. Guo M H, Cai J X, Liu Z N, et al. Pct: Point cloud transformer[J]. Computational visual media, 2021, 7: 187-199.
25. Wen X, Xiang P, Han Z, et al. PMP-Net++: Point cloud completion by transformer-enhanced multi-step point moving paths[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(1): 852-867.
26. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.
27. 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.
28. 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. 2017: 5828-5839.
29. Armeni I, Sax S, Zamir A R, et al. Joint 2d-3d-semantic data for indoor scene understanding[J]. arXiv preprint arXiv:1702.01105, 2017.

# 附 录