

# 控制与决策

Control and Decision

## 几何先验引导的堆叠点云抓取位姿联合预测方法

李小晗, 张哲戩, 徐胜军, 孟月波, 杨玥, 武经, 苏晓杰, 陈义明

引用本文:

李小晗, 张哲戩, 徐胜军, 等. 几何先验引导的堆叠点云抓取位姿联合预测方法[J]. *控制与决策*, 2026, 41(5): 1403-1414.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2025.0696>

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### [基于改进RRT\\*FN算法的机器人路径规划](#)

Robot path planning based on improved RRT\*FN algorithm

控制与决策. 2021, 36(8): 1834-1840 <https://doi.org/10.13195/j.kzyjc.2019.1713>

#### [基于两阶段迭代优化的空天观测资源协同任务规划方法](#)

A two-stage iterative optimization method for the coordinated task planning of space and air observation resources

控制与决策. 2021, 36(5): 1147-1156 <https://doi.org/10.13195/j.kzyjc.2019.1193>

#### [基于16方向24邻域改进蚁群算法的移动机器人路径规划](#)

Mobile robots path planning based on 16-directions 24-neighborhoods improved ant colony algorithm

控制与决策. 2021, 36(5): 1137-1146 <https://doi.org/10.13195/j.kzyjc.2019.0600>

#### [尺度自适应的多特征融合相关滤波目标跟踪算法](#)

Scale adaptation and multi-feature fusion correlation filtering object tracking algorithm

控制与决策. 2021, 36(2): 429-435 <https://doi.org/10.13195/j.kzyjc.2019.0445>

#### [基于 \$\pm 3\sigma\$ 正态概率区间分族遗传蚁群算法的移动机器人路径规划](#)

Path planning of mobile robot based on  $\pm 3\sigma$  normal probability interval population division using genetic ant-colony algorithm

控制与决策. 2021, 36(12): 2861-2870 <https://doi.org/10.13195/j.kzyjc.2020.0745>

# 几何先验引导的堆叠点云抓取位姿联合预测方法

李小晗<sup>1†</sup>, 张哲戩<sup>1</sup>, 徐胜军<sup>1</sup>, 孟月波<sup>1</sup>, 杨玥<sup>1</sup>, 武经<sup>2</sup>, 苏晓杰<sup>3</sup>, 陈义明<sup>4</sup>

(1. 西安建筑科技大学 信息与控制工程学院, 西安 710055; 2. 南方科技大学 机械与能源工程学院, 广东 深圳 518055; 3. 重庆大学 自动化学院, 重庆 400044; 4. 南洋理工大学 机械与宇航工程学院, 新加坡 639798)

**摘要:** 为提升机器人在非结构化堆叠场景中的抓取识别精度与执行稳定性, 提出一种融合几何先验建模与姿态质量评估机制的六自由度抓取预测算法. 首先, 构建点云识别网络 Point-LaKan, 通过增强输入点云的局部几何特征, 设计由局部聚合模块与高维非线性映射模块构成的 LAKAN 特征提取结构, 提升对堆叠抓取区域的结构表征能力; 其次, 设计方向向量约束下的抓取姿态估计策略, 通过最小化初始与目标姿态间的空间差异, 提升姿态生成的可执行性与可解释性; 最后, 构建融合方向约束、碰撞检测与质心评分的抓取姿态筛选机制, 实现候选姿态的多因素评估与排序, 增强算法在复杂环境下的执行鲁棒性. 为验证算法性能, 自主构建多类别堆叠物体仿真点云抓取数据集, 分别在 CoppeliaSim 仿真平台与真实机器人系统中开展实验. 结果表明: 在模型参数量减少 4.69%、推理速度提升 37.19% 的条件下, 抓取区域识别准确率提升了 25.26%; 真实抓取成功率与任务完成率最高可提升 29.40% 与 18.39%.

**关键词:** 机器人抓取; 几何先验特征; 多尺度点云特征融合; 抓取目标识别网络; 6-DOF 位姿估计; 方向约束  
**中图分类号:** TP39; TH86 **文献标志码:** A

**DOI:** 10.13195/j.kzyjc.2025.0696

**引用格式:** 李小晗, 张哲戩, 徐胜军, 等. 几何先验引导的堆叠点云抓取位姿联合预测方法 [J]. 控制与决策, 2026, 41(5): 1403-1414.

## A joint grasp pose prediction method for stacked point clouds guided by geometric priors

LI Xiao-han<sup>1†</sup>, ZHANG Zhe-jian<sup>1</sup>, XU Sheng-jun<sup>1</sup>, MENG Yue-bo<sup>1</sup>, YANG Yue<sup>1</sup>, WU Jing<sup>2</sup>, SU Xiao-jie<sup>3</sup>, CHEN Yi-ming<sup>4</sup>

(1. School of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China; 2. School of Mechanical and Energy Engineering, Southern University of Science and Technology, Shenzhen 518055, China; 3. School of Automation, Chongqing University, Chongqing 400044, China; 4. School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 639798, Singapore)

**Abstract:** To improve the grasp recognition accuracy and execution stability of robots in unstructured stacked scenes, this paper proposes a 6-DOF grasp prediction algorithm that combines geometric prior modeling and a pose quality evaluation mechanism. Firstly, a point cloud recognition network named Point-Lakan is constructed. By enhancing the local geometric features of the input point cloud, a LAKAN feature extraction structure consisting of a local aggregation module and a high-dimensional nonlinear mapping module is designed to improve the structural representation ability of the stacked grasp area. Secondly, a grasp pose estimation strategy incorporating directional vector constraints is designed by minimizing the spatial difference between the initial and target poses to improve the executability and interpretability of pose generation. Finally, a grasp posture screening mechanism that combines direction constraints, collision detection, and centroid score is constructed to realize multi-factor evaluation and ranking of candidate postures, thereby enhancing the execution robustness of the algorithm in complex environments. To verify

收稿日期: 2025-07-01; 录用日期: 2025-11-12.

基金项目: 国家自然科学基金项目 (62503377, 62306228, 62441237); 陕西省高等学校重点实验室项目 (24JS024); 陕西省自然科学基金基础研究计划一般项目 (2025JC-YBQN-857); 陕西省重点研发计划项目 (2024GX-YBXM-132).

责任编辑: 贾英民.

<sup>†</sup>通信作者. E-mail: lixiaohan272@gmail.com.

本文附带电子附录文件, 可登录本刊官网该文“资源附件”区自行下载阅览.

the performance of the algorithm, this paper conducts thorough evaluations using the CoppeliaSim simulation platform as well as a real-world robot system based on a self-constructed dataset of multi-category stacked object simulation point clouds for grasping tasks. The results indicate that the accuracy of grasping area recognition increases by 25.26%, while decreasing the model parameters by 4.69% and boosting the inference speed by 37.19%. In real-world grasping tasks, the success rate and task completion rate increase by up to 29.40% and 18.39%, respectively.

**Keywords:** robotic grasping; geometry-prior feature; multi-scale point cloud feature fusion; grasp target recognition network; 6-DOF pose estimation; direction constraints

## 0 引言

基于深度学习的抓取位姿估计已成为机器人领域的研究热点之一,其核心在于利用神经网络模型预测目标物体的抓取位置与姿态,以实现高精度的自主抓放操作<sup>[1]</sup>.以往方法<sup>[2-3]</sup>依赖 2D 或 2.5D 视觉信息,仅支持平面抓取,限制了机器人在非结构化环境下的操作灵巧性与通用性<sup>[4]</sup>.为解决上述局限,研究者逐渐转向利用三维点云数据开展六自由度 (6-DOF) 抓取姿态估计.

在 6-DOF 抓取位姿估计中,点云识别技术是关键基础,方法主要分为两类:基于传统算法的识别方法与基于深度学习的识别方法.前者通常依赖物体几何特征和启发式规则,通过点云配准与特征匹配实现目标识别<sup>[5]</sup>.例如, Han 等<sup>[6]</sup>提出了 K-D Tree-ICP 算法,在迭代最近点过程中嵌入 K-D 树快速最近邻搜索,显著提升了检测场景的配准效率与精度.尽管传统方法在规则物体或小规模场景中表现良好,但在复杂噪声环境下易失效.相较而言,基于深度学习的点云识别方法展现出更强的泛化能力与鲁棒性. Qi 等<sup>[7]</sup>提出的 PointNet 网络利用共享多层感知器 (MLP) 实现对无序点集的直接处理.然而,由于未考虑局部邻域结构,其对细粒度几何信息的建模能力有限.为此, Qi 等<sup>[8]</sup>进一步提出了 PointNet++ 网络,通过引入多尺度几何感知模块对局部邻域特征进行聚合编码,显著增强了局部结构建模能力.随后, Qian 等<sup>[9]</sup>提出了 PointNeXt 网络,利用更深层次的 MLP 对多尺度领域特征进行高维映射,以获得更丰富的局部语义表达.尽管深层 MLP 架构已广泛应用于点云处理任务,但仍存在以下不足: 1) MLP 对点云的空间拓扑不敏感,描述点间几何依赖的能力有待提升; 2) 随着网络加深,关键特征在多层映射中逐步衰减<sup>[10]</sup>,导致模型性能提升受限.因此,探索更具结构建模能力的点云特征表达机制,成为突破基于 MLP 架构瓶颈、提升抓取姿态估计精度的重要方向.

为提升机器人在复杂场景中的抓取性能,基于深度学习的 6-DOF 抓取姿态估计算法主要分为一阶段抓取策略和二阶段抓取策略两类.前者通过点云识别网络直接从整体点云中提取局部特征,并生

成抓取姿态. Ni 等<sup>[11]</sup>提出的 PointNet++ Grasping 模型利用 PointNet++ 网络增强了局部特征提取能力,提升了抓取准确性; Alliegro 等<sup>[12]</sup>提出直接从点云中采样抓取动作的方法,有效简化了预处理流程并提高了推理速度; Zhao 等<sup>[13]</sup>提出的 REGNet 模型通过聚合抓取区域特征而非单点特征实现更稳定的姿态预测.尽管上述方法在抓取姿态估计中具备端到端建模的效率优势,但其普遍依赖精确且多样的标注数据,对数据质量高度敏感,限制了模型在开放环境中的泛化能力.此外,由于识别与姿态生成过程高度耦合,无法实现对关键子任务的独立调控与性能解释.

相较于端到端的一阶段策略,两阶段抓取方法通常先识别高抓取置信度区域,再估算抓取姿态.例如: Liang 等<sup>[14]</sup>提出的 PointNetGPD 网络通过在目标点云中随机生成抓取候选姿态,再将对应的局部点云输入分类网络,对其抓取可行性进行评分; Fang 等<sup>[15]</sup>提出的 AnyGrasp 网络采用高维局部特征进行抓取区域检测,再估计相应的候选抓取姿态; 蔡子豪等<sup>[16]</sup>则提出一种基于几何分析的六自由度抓取策略,先通过几何特征引导候选采样,再结合几何约束与力学封闭条件进行可行性评估,生成最优抓取姿态.上述方法支持算法架构的模块化设计与阶段性优化.然而,其第 1 阶段仍依赖点云处理网络完成抓取目标识别,主要通过局部聚合 (LA) 与多层感知机 (MLP) 堆叠提取特征.尽管具备一定表达能力,但深层结构在提升模型容量的同时会增加计算复杂度与训练成本,限制了在非结构化环境下对关键局部特征的高效感知,进而影响抓取策略的鲁棒性.

针对上述问题,本文提出一种融合几何先验建模与姿态质量评估机制的两阶段点云 6-DOF 抓取方法,主要工作包括:

1) 构建几何先验信息引导的点云识别网络 Point-LaKan, 通过增强输入点云的局部几何特征,设计由局部聚合模块与高维非线性映射模块构成的 LAKAN 特征提取结构,提升对堆叠抓取区域的结构感知能力.

2) 设计结合目标物体姿态与末端执行器抓取方

向的联合推理策略,引入方向向量约束以优化初始姿态与目标抓取姿态之间的空间差异,从而提高生成抓取姿态的执行可行性。

3) 利用融合方向约束与质量评估的姿态筛选机制,通过对高置信度抓取姿态的候选生成、碰撞过滤与优选排序,提升算法在复杂环境下的可解释性以及稳健性。

4) 自主构建覆盖多类别目标的仿真点云抓取数据集,并在 CoppeliaSim 仿真环境与真实机器人系统中完成算法部署与抓取实验,从执行层面验证所提出方法的有效性与环境适应能力。

## 1 抓取算法总体框架

针对非结构化场景中对堆叠目标抓取姿态高效

精准预测问题,本文提出一种面向非结构化堆叠场景的两阶段 6-DOF 点云抓取算法框架。如图 1 所示,在第 1 阶段,构建融合局部特征聚合模块与高维特征重构模块的点云特征提取模块 LAKAN,并提出一种几何先验信息引导的抓取区域识别网络 Point-LaKan,用于得到抓取置信度区域集合  $H = \{h_1, h_2, \dots, h_n\}$ 。在第 2 阶段,算法围绕上述区域执行点云细化筛选,获得局部关键点集  $P_e = \{p_1, p_2, \dots, p_n\}$ ,并结合各点的空间位置和方向信息,生成姿态候选集  $G_p = \{g_1, g_2, \dots, g_n\}$ 。在姿态构建过程中引入方向约束机制以保持操作合理性,同时结合碰撞检测策略剔除不可行解。最终,通过置信度评估与排序选出最优姿态  $g_h$  执行抓取任务。

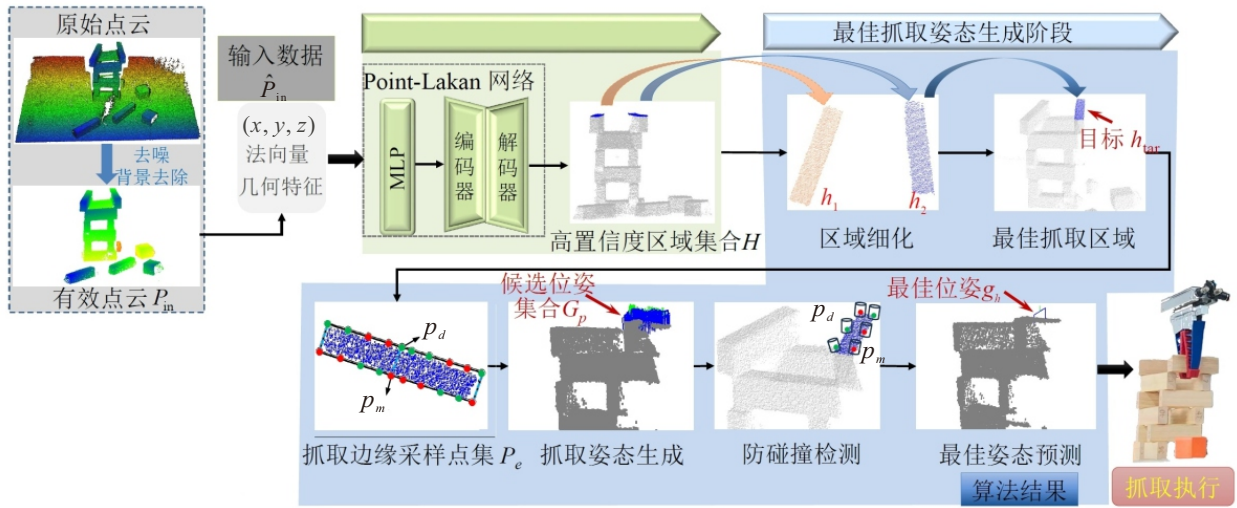


图1 抓取算法框架

## 2 几何先验引导的抓取区域识别网络

### 2.1 融合几何属性的点云输入通道扩展

现有点云识别网络多以原始空间坐标  $P_m = \{p_i | i = 1, 2, \dots, n\}$  作为输入,其中  $p_i = (x, y, z)_i \in \mathbb{R}^3$ ,未能充分挖掘点云中蕴含的结构性信息,限制了模型对复杂抓取区域的判别能力<sup>[17]</sup>。为弥补上述不足,如图 2 所示,本文通过提取  $P_m$  的邻域几何信息,对输入数据维度进行扩展。首先,为提高模型对不同尺度及非均匀分布点云数据的适应性,对输入数据进行归一化与中心化处理,有

$$\bar{p}_i = \frac{p_i}{\max \|p_i\|}, \quad (1)$$

$$\hat{p}_i = \bar{p}_i - \frac{1}{N} \sum_{i=1}^N \bar{p}_i. \quad (2)$$

其中:  $\max \|p_i\|$  为输入点云的几何中心到点云中最远点的距离;  $\bar{p}_i$  为归一化后的点云坐标;  $\hat{p}_i$  为消除坐标偏移影响后的点云坐标。为增强模型对几何特征的

学习能力,逐点提取点云的相对高度特征  $\Gamma_h$  和领域高度  $\Gamma_b$ , 即

$$\begin{aligned} \Gamma_h &= h_{\hat{p}_i} - h_{\hat{p}_{\min}}, \\ \Gamma_b &= \frac{1}{k} \sum_{\hat{p}_j \in \hat{P}} |h_{\hat{p}_i} - h_{\hat{p}_j}|. \end{aligned} \quad (3)$$

其中:  $\hat{P}$  表示点  $\hat{p}_i$  的  $k$  个最近邻点集合;  $h_{\hat{p}_i}$  表示点  $\hat{p}_i$  沿重力方向(即  $z$  轴)的高度值;  $h_{\hat{p}_{\min}}$  表示该邻域内的最小高度值;  $\Gamma_h$  描述点相对最低邻域点的高度差;  $\Gamma_b$  表示该点与其邻域点在高度上的平均差异,描述了局部高度变化程度。基于上述处理,将点云坐标、法向量信息以及  $\Gamma_h$  与  $\Gamma_b$  进行数组拼接,构成新的网络输入数据形式,即  $\hat{P}_m = \{\hat{p}_i | i = 1, 2, \dots, n\}$ , 其中  $\hat{p}_i = [x, y, z, n_x, n_y, n_z, \Gamma_h, \Gamma_b]_i \in \mathbb{R}^8$ 。

### 2.2 抓取点云识别网络 Point-LaKan

1) 主干网络。为提升对点云抓取区域的几何特征表达能力与建模效率,本文提出一种以 Point-LaKan 为核心的抓取区域识别网络。该网络通过增

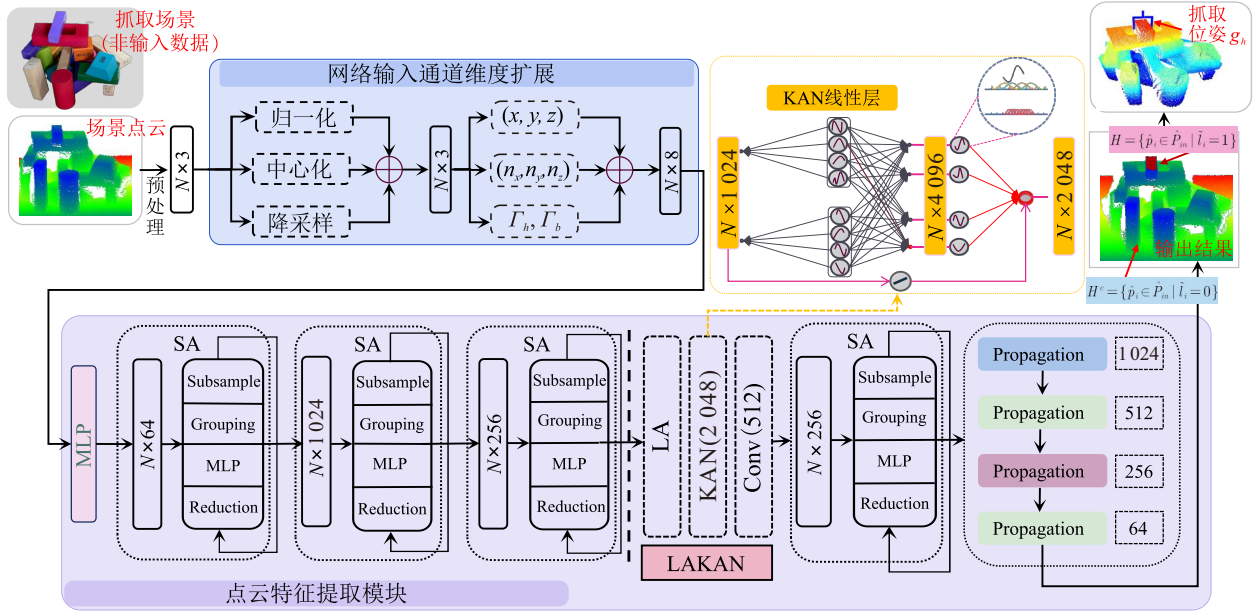


图2 结合几何先验增强的 Point-LaKan 网络结构

强局部领域特征, 对非结构化场景中的高置信度抓取区域  $H = \{\hat{p}_i \in \hat{P}_m | \tilde{l}_i = 1\}$  进行有效预测. 主干网络结构延续了点云学习框架的模块化设计思想, 同时对传统深层 MLP 堆叠结构进行了替代与优化: 在浅层采用 Set Abstraction (SA) 模块以完成基本特征下采样与空间聚合, 并移除冗余 MLP 层; 在中深层引入融合式特征增强模块, 提升语义层对复杂几何结构的建模能力, 缓解深层特征衰减问题.

2) 局部特征聚合模块  $F_{LA}$ . 首先构建以点  $\hat{p}_i$  为中心的多尺度球形邻域, 其定义为

$$\mathfrak{N}(\hat{p}_i) = \{\hat{p}_j | \|\hat{p}_j - \hat{p}_i\|_2 \leq r_{base} s^l\}. \quad (4)$$

其中:  $\mathfrak{N}(\hat{p}_i)$  为当前层级  $l$  下的邻域点集;  $r_{base}$  为基础半径;  $s$  为比例缩放因子, 用于跨层调节感受野大小. 在每个点  $\hat{p}_i$  的邻域  $\mathfrak{N}(\hat{p}_i)$  内, 局部特征通过特征拼接与最大池化操作完成聚合, 即

$$f_i^{agg} = \max_{\hat{p}_j \in \mathfrak{N}(\hat{p}_i)} \{f(\Delta p_{i-j} \oplus f(\hat{p}_j))\} \in R^{b \times d_{out}}. \quad (5)$$

其中:  $\Delta p_{i-j}$  为点  $\hat{p}_j$  相对于中心点  $\hat{p}_i$  的相对位置编码,  $f(\hat{p}_j)$  为邻域点的初始特征表示,  $\oplus$  为特征拼接操作,  $\max\{\cdot\}$  为在邻域范围内的最大池化操作. 聚合后每个点的输出维度为  $d_{out} = 8 + d_l$ ,  $d_l$  为该层生成的新特征通道数. 因此, 所有采样点的聚合特征在 batch 维度下重组为高维特征张量, 即

$$F_{LA} = [f_1^{agg}, f_2^{agg}, \dots, f_{\tilde{n}}^{agg}]_b \in R^{b \times \tilde{n} \times d_{out}}. \quad (6)$$

其中:  $b$  表示 batch 大小,  $\tilde{n}$  表示每个 batch 的点数.

3) 高维非线性映射模块  $F_{KAN}$ . 式 (6) 生成的聚合特征  $F_{LA}$  具有较高维度, 若直接采用对拓扑结构缺乏感知能力的 MLP 进行拟合, 则会导致网络结构

加深, 关键信息在多层传递过程中逐渐衰减, 从而影响模型的收敛速度与特征保留能力. 为提升模型对高维非线性关系的拟合能力, 本文借鉴 KAN 结构设计思想<sup>[18]</sup>. 但若直接将聚合特征  $F_{LA}$  输入双路 KAN 模块, 则会因缺乏显著性调控而导致对局部几何结构建模不充分, 易受噪声或离群点干扰. 为此, 本文提出一种全局-局部特征融合机制, 以增强模型对关键几何信息的表达能力. 具体而言, 首先对  $F_{LA}$  进行线性变换, 以保留整体语义结构并形成全局基准特征  $F_{base}$ , 即

$$F_{base} = b(F_{LA}^T) W_b. \quad (7)$$

其中:  $b(\cdot)$  为线性基函数,  $W_b$  为可学习参数矩阵. 随后, 对每个局部特征  $f_i^{agg}$  进行非线性函数映射, 提取具备细粒度几何结构的局部特征  $F_{spline}$ , 即

$$F_{spline} = \sum_{i=1}^N s_i \left( \sum_{m=1}^8 c_m B_m(f_i^{agg}) W_{s[m]} \right). \quad (8)$$

其中:  $B_m(\cdot)$  为三阶 B 样条函数;  $c_m$  为可学习系数;  $W_{s[m]}$  为连接样条输出的线性变换矩阵. 并设计可学习的缩放因子  $s_i$ , 实现点级别的自适应权重调制, 从而增强模型对关键局部区域的关注能力. 通过融合线性映射得到的全局特征  $F_{base}$  与非线性映射后的特征  $F_{spline}$ , 得到新的重构特征

$$F_{KAN} = F_{base} + F_{spline}. \quad (9)$$

综上, 结合式 (6) 和 (9) 可以得到一个融合局部几何聚合机制和非线性函数映射的中深层特征提取模块——LAKAN, 作为主干网络的核心模块. 该模块在保持局部几何一致性建模能力的同时, 引入全局函数拟合机制, 可以实现对抓取区域不同尺度特

征的深度融合与表达增强。

### 2.3 结合最高点抓取策略的损失函数构造

为提升模型对目标顶端区域的响应能力,本文在损失函数设计中引入描述相对高度的几何先验特征  $\Gamma_{h_i}$ . 抓取区域提取任务本质上为点级二分类回归问题,因此需要拟合出输入点云  $\hat{P}_m$  与抓取候选区域集合  $H$  之间的非线性映射函数

$$F: \hat{P}_m \rightarrow H = \{h_i \subseteq \hat{P}_m | i = 1, 2, \dots, \mathfrak{S}\}. \quad (10)$$

模型训练的核心目标在于优化函数  $F$  的逼近能力,使预测值与真实标签  $\tilde{l} \in \{0, 1\}$  之间的误差最小化. 在堆叠场景中,目标点云往往存在多个高置信度抓取区域. 根据经验规律,相较于侧向抓取,顶端抓取具有更小的碰撞风险、更短的运动路径与更高的相机可见性. 因此,为引导网络更加关注具有物理抓取优势的区域,本文在损失函数中引入高度先验因子

$$\text{Loss} = -\lg(p_t) \left( (1 - p_t)^\mu \right) e^{\nu \Gamma_{h_i}}. \quad (11)$$

其中:  $p_t$  用来描述网络预测概率  $p$  的置信度值,使损失函数在正负样本上表现出一致的梯度形式,有

$$p_t = \tilde{l}p + (1 - \tilde{l})(1 - p); \quad (12)$$

$\mu$  为调节非平衡样本的困难样本加权指数因子,本文取  $\mu = 2.0$ ;  $\nu$  为高度权重因子,控制几何先验对损失权重的调整幅度,本文取  $\nu = 3.0$ ;  $\Gamma_{h_i}$  为点  $\hat{p}_i$  与局部邻域最低点的高度差. 该损失函数通过对高位点赋予更高的损失权重,引导网络在训练过程中优先聚焦于具有更高操作可达性和稳定性的抓取位置.

设  $\tilde{l}_i$  表示每个点  $\hat{p}_i$  经过网络预测得到的可行性标签,对应的高置信度抓取区域定义为  $H = \{\hat{p}_i \in \hat{P}_m | \tilde{l}_i = 1\}$ , 其补集  $H^c = \{\hat{p}_i \in \hat{P}_m | \tilde{l}_i = 0\}$  表示不可抓取区域. 该划分实现了对点云中抓取候选区与背景干扰区的显式分离,为后续抓取姿态生成提供基础.

## 3 6-DOF 抓取位姿预测

### 3.1 抓取区域细化

在密集堆叠场景下,上一节得到的  $H = \{h_i | h_1, h_2, \dots, h_{\mathfrak{S}}\}$  可能包含多个相邻物体区域(如图1所示),导致抓取定位存在多义性. 为此,本文在姿态预测之前设计一种基于几何高度与点云密度联合评分的区域细化策略,旨在从候选区域中筛选出最具可行性的单一抓取目标. 针对每个候选区域  $h_i \in H$ , 定义评分函数如下:

$$S_i = \alpha z_i + \beta N(h_i). \quad (13)$$

其中:  $z_i$  表示区域  $h_i$  内所有点在重力方向 ( $z$  轴) 上

的平均高度;  $N(h_i)$  表示区域  $h_i$  中的点数,即点云密度指标;  $\alpha$  与  $\beta$  为可调的加权因子,本文分别取  $\alpha = 0, \beta = 1.5$ . 根据评分函数  $S_i$  对候选区域进行排序,选取得分最高的区域  $h_{\text{tar}}$  作为最佳抓取点簇.

### 3.2 融合方向向量约束的抓取姿态生成

设定 6-DOF 末端抓钳的通用位姿表达式为  $g = (p_g, r_g)$ . 其中:  $p_g = (x_g, y_g, z_g)$  表示抓取位置,  $r_g = (\vec{r}_x, \vec{r}_y, \vec{r}_z)$  为旋转矩阵. 为实现稳定抓取,本文认为抓取位姿需满足以下两个条件: 首先,抓取目标的空间姿态需与末端执行器的预期姿态高度一致,可以引导末端执行器沿近似垂直于物体表面的方向进行吸附或抓握,有效降低碰撞概率及掉落风险; 其次,目标初始姿态与所需抓取姿态之间的位姿差异越小,路径调整难度越低,有助于提高整体抓取效率. 基于此,如图3所示,提出一种融合方向向量约束的 6-DOF 姿态生成方法,其步骤如下:

1) 采用 RANSAC 点云平面拟合方法得到两个候选平面子点集,分别记为内点集  $P_{\text{ran1}}\{p_{c1}(x_1, y_1, z_1), \mathbf{n}_1\}$  和外点集  $P_{\text{ran2}}\{p_{c2}(x_2, y_2, z_2), \mathbf{n}_2\}$ ,  $p_{c1}$  和  $p_{c2}$  是两个子区域的中心点,  $\mathbf{n}_1$  和  $\mathbf{n}_2$  是方向特征向量.

2) 提取  $h_{\text{tar}}$  的 3D 定向边界框,获取长边端点  $p_1(x_{11}, y_{11}, z_{11})$  和  $p_2(x_{12}, y_{12}, z_{12})$  以及边界框中心点  $p_{\text{box}}(x_{\text{box}}, y_{\text{box}}, z_{\text{box}})$ , 并计算长边方向向量  $\mathbf{l} = p_1 - p_2$ .

3) 评估两个法向量  $\mathbf{n}_1$  和  $\mathbf{n}_2$  与参考向量  $\mathbf{n}_{\text{ref}} = (0, 0, -1)$  的余弦相似度,以判定其垂直程度,即

$$\cos \theta_i = \frac{\mathbf{n}_{\text{ref}} \cdot \mathbf{n}_i}{\|\mathbf{n}_{\text{ref}}\| \cdot \|\mathbf{n}_i\|}, \quad i \in \{1, 2\}. \quad (14)$$

其中  $\cos \theta_i \in [0, 1]$ , 值越大表示法向量越接近垂直于地面的方向.

4) 根据上述相似度,选取最适合作为抓取面的候选区域

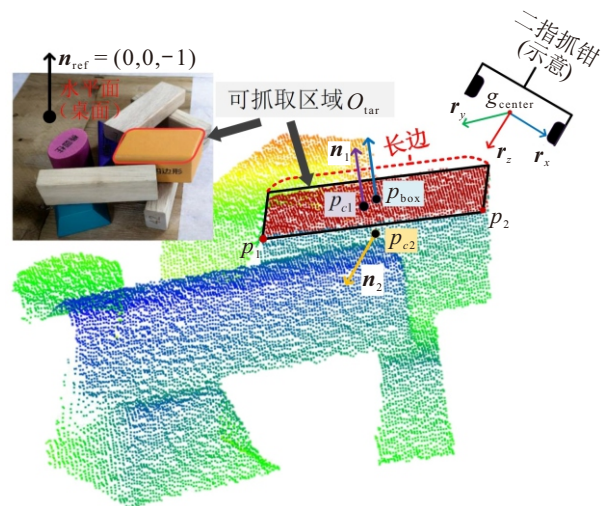


图3 6-DOF 末端位姿估计示意图

$$P_{\text{ran}} = \begin{cases} P_{\text{ran}1}, & \cos \theta_1 \geq \cos \theta_2; \\ P_{\text{ran}2}, & \text{otherwise.} \end{cases} \quad (15)$$

5) 估算最终 6-DOF 抓取位姿  $g = (p_g, r_g)$ , 其中抓取点  $p_g(x_g, y_g, z_g)$  的坐标由候选面中心点与边界框中心点共同确定, 有

$$\begin{cases} X_g = \frac{x_{\text{ran}} + x_{\text{box}}}{2}, \\ Y_g = \frac{y_{\text{ran}} + y_{\text{box}}}{2}, \\ Z_g = z_{\text{ran}} + \mu. \end{cases} \quad (16)$$

其中  $\mu$  为末端执行器夹具中心到目标表面的经验偏移量。

6) 构建抓取姿态的旋转矩阵  $r_g = [r_x, r_y, r_z]$ , 其中

$$r_z = -n, \quad (17)$$

$$r_y = l - \frac{l \cdot r_z}{\|r_z\|^2} \cdot r_z, \quad (18)$$

$$r_x = r_y \times r_z. \quad (19)$$

其中:  $\times$  表示向量叉乘运算, 用于构建正交基底, 确保旋转矩阵的几何一致性与物理可行性。

在实际抓取任务中, 若目标点云的法向量 (如图 3 中的  $n_2$ ) 趋近于水平, 上述步骤所生成的抓取姿态可能引发机械臂关节的运动学约束超限。为提升姿态估计的物理可行性, 算法在步骤 3) 中引入参考向量  $n_{\text{ref}} = (0, 0, -1)$ , 通过限制抓取姿态在  $Z$  轴方向的自由度, 引导夹持器主轴趋近垂直于地面。需要注意的是, 当步骤 4) 中任一候选区域的方向相似度  $\cos \theta_1$  或  $\cos \theta_2$  小于 0.01 时, 表明估算方向近似平行于接触面, 夹爪难以建立有效接触。为此, 采用方向打断机制, 将此时的方向向量替换为固定的垂直向量  $n_{\text{fix}} = (0, 0, 1)$ , 确保末端执行器以垂直姿态完成抓取, 从而增强抓取的稳定性与鲁棒性。

### 3.3 位姿可行性评估与筛选策略

本节提出一种针对高置信度抓取区域的 6-DOF 位姿筛选策略, 整体流程如图 4 所示。该方法旨在通过几何分析与可行性评估联合剔除存在碰撞风险的抓取姿态, 从最佳抓取区域  $h_{\text{tar}}$  中得到高可行性抓取点集  $g_h$ 。

首先, 从  $h_{\text{tar}}$  中提取边界点集  $P_e = \{p_1, p_2, \dots, p_m\}$  作为潜在接触点。针对二指夹爪结构, 以每个边界点  $p_m$  为单侧接触点, 基于几何对称性在  $h_{\text{tar}}$  中搜索其对称点  $p_d$ , 构成夹爪两侧接触点对  $\{p_m, p_d\}$ 。抓取方向由式 (17) ~ (19) 中计算得到的旋转矩阵  $r_g = [r_x, r_y, r_z]$  决定, 构成完整的姿态六元组  $g = (p_g, r_g)$ ,

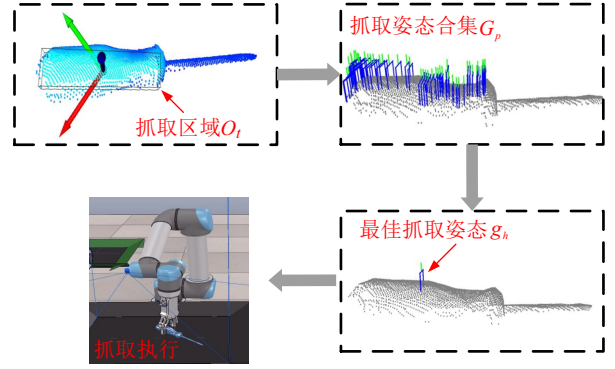


图4 最佳抓取姿态估计

其中  $p_g = (p_m + p_d)/2$ 。为确保抓取动作的无碰撞性, 对候选姿态集  $G_p$  引入碰撞检测机制: 以  $p_g$  为圆心, 夹爪开距为直径构建圆柱体碰撞模型, 并检测其包络空间中是否存在障碍点云, 仅保留满足无碰撞条件的姿态进入下一阶段。为筛选最优抓取姿态, 引入基于空间几何中心的评分机制, 定义评分函数

$$\tilde{S}(g_i) = -|p_g^{(i)} - p_c|^2, \quad (20)$$

其中  $p_c$  表示  $h_{\text{tar}}$  的几何中心。最终抓取姿态  $g_h$  通过最大化评分函数选出, 即  $g_h = \arg \max_{g_i \in G_p} \tilde{S}(g_i)$ 。为确保抓取姿态预测的唯一性与稳定性, 本文在评分并列情况下采用固定索引优先的确定性选择策略。

## 4 实验与分析

### 4.1 实验准备

#### 4.1.1 机器人抓取平台搭建

为验证所提出的 6-DOF 抓取识别算法在多场景下的通用性与实用性, 如图 5 所示, 本文分别在 CoppeliaSim 仿真环境与真实机器人平台上构建完整的实验系统。在仿真环境中, 构建典型工业抓取场景, 包括输送带模块、随机堆叠物料筐与抓取作业区。集成的虚拟 RGB-D 相机具有  $H50^\circ \times V30^\circ$  视场角, 图像输出分辨率为  $640 \times 480/\text{pixel}$ 。在真实实验

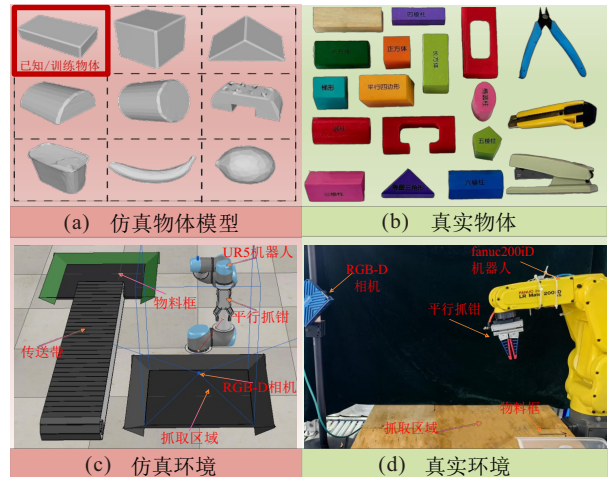


图5 仿真与真机实验环境中的抓取物体以及平台

中,实验平台采用 Fanuc 200Di 六轴工业机械臂,末端执行器配置为双指并联式气动夹爪.视觉感知模块选用 Revopoint 公司的 Trancer P1 型 3D 工业相机,其视场角为  $H64^\circ \times V50^\circ$ ,支持  $960 \times 600/\text{pixel}$  分辨率的深度图与稠密点云数据采集,并提供完整的内参与外参标定参数,以保证抓取场景的空间几何信息精确重建.系统各组件间通过标准 TCP/IP 通信协议连接,控制指令与感知数据可实现高速交互.

#### 4.1.2 抓取点云数据集

为提升网络模型在抓取任务中对目标点云特征的泛化能力,本文构建一个多物体抓取点云数据集 (MPDG),包含两个子数据集:基于仿真环境生成的 MPDG-Sim 和基于真实场景采集的 MPDG-Real,覆盖 20 余类具有代表性的常见物体模型,旨在支撑本文提出的类无关、6 自由度的抓取策略学习.其与其他公开抓取数据集的特性对比见表 1.数据集中每组点云采用随机姿态摆放的多物体混合堆叠场景,确保物体种类、数量、几何形状及空间姿态的差异性,为模型提供丰富的训练样本.

表1 MPDG 与公开抓取数据集的特性对比

数据集	类别数	数据格式	6-DOF	场景
YCB <sup>[19]</sup>	21种	RGB-D	×	真实
VPG <sup>[20]</sup>	6种以上	深度图	×	仿真
MPDG	29种	稠密点云	✓	仿真+真实

1) MPDG-Sim. 包含 1000 组仿真点云数据,用于模型训练与初步验证,其中训练集中仅包含一种矩形方块物体(作为先验模板),用于建立基础抓取策略;测试集则引入来自公开的 YCB 数据集<sup>[19]</sup>和 VPG 数据集<sup>[20]</sup>的 10 余种三维物体模型,以验证所提出方法在未知物体上的泛化能力.

2) MPDG-Real. 包含 200 组工业 3D 相机采集的真实点云数据,仅用于算法测试.评估网络模型在真实且未知环境中的 6-DOF 抓取识别性能.该子数据集不提供物体类别标签,训练与测试物体无交集.

#### 4.1.3 评价指标

为客观衡量多物体堆叠场景下的抓取目标识别效果,本文参考文献[21]中的评估方法,使用准确率 (Precision)、召回率 (Recall)、交并比 (IoUs) 及综合评价指标 (F-Score) 作为评价指标.此外,使用抓取成功率 (SR) 和抓取完成率 (CR) 作为指标系统评估机器人执行效率,将物体抓取并抬高 5 cm 视为抓取成功,定义

$$\text{SR} = \frac{N_s}{N_t} \times 100\%, \quad (21)$$

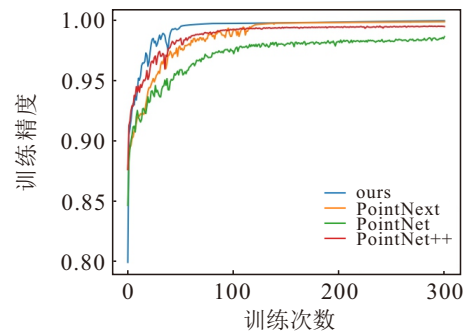
$$\text{CR} = \frac{N_c}{N_a} \times 100\%. \quad (22)$$

其中:  $N_s$  为成功抓取的次数,  $N_t$  为总抓取尝试次数,  $N_c$  为成功移除的物体数量,  $N_a$  为总任务尝试次数.

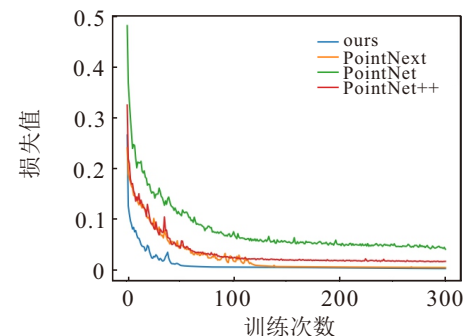
#### 4.2 抓取目标识别实验分析

为验证第 2 节提出的 Point-LaKan 网络在提升训练效率方面的有效性,选取 3 种典型的点云识别模型 (PointNet<sup>[7]</sup>、PointNet++<sup>[8]</sup> 与 PointNext<sup>[9]</sup>) 构建对比实验.所有网络模型均基于 PyTorch 平台实现,采用 Adam 优化器,在 Nvidia GTX 3070 GPU 上进行训练.具体参数设置如下: batch size 为 4, 初始学习率为 0.001, 学习率衰减因子设为 0.1, 最大训练轮次为 300.为保证对比的公平性与统计的合理性,在计算处理平台为统一配置的情况下,在 MPDG-Real 数据集中选取 100 组点云样本用于测试,并对所有样本进行统一的降采样处理(点数保持一致).每种模型均进行 10 次独立推理任务,记录其推理耗时,最终以平均值作为统计结果.

1) 网络模型收敛性能验证.图 6 展示了各模型在训练精度 (Train Accuracy) 与损失函数 (Loss) 方面的收敛性能对比.由图 6(a) 可知, PointNet 作为基础点云特征学习网络,其最高训练精度达到 98.55%,但由于缺乏对局部几何结构的刻画能力,其收敛速度和精度提升空间均受到限制. PointNext 在 PointNet++ 的基础上进一步增强了特征提取模块的层次性与感受野,将训练精度提升至 99.71%.然而,由于其网络



(a) 精度变化曲线



(b) 损失函数变化曲线

图6 不同网络模型的训练性能对比



深度的增加,训练过程中参数更新复杂度提高,导致模型需约 150 个轮次才能完成有效收敛。相比之下,本文提出的 Point-LaKan 网络通过高效的几何敏感特征融合策略,显著提升了对抓取区域局部空间结构的建模能力,仅需约 80 个轮次即可达到收敛,并实现最高训练精度 99.89%。因此,Point-LaKan 的训练稳定性、收敛效率优于主流网络模型,表现出其在点云抓取任务中的结构设计优势。

2) 网络模型参数量对比。为测试本文提出的网络模型 Point-LaKan 在参数规模与推理效率方面的综合性能,选取 PointNet++、PointMLP、PointNext 作为对比基准,并在统一的测试数据集与硬件平台下,对各模型的网络参数量与单次平均推理时间进行实验测试。对比结果如表 2 所示,Point-LaKan 网络虽然提升了对点云特征的表达能力,但相应地也使整体结构较 PointNet++ 更为复杂,导致参数规模有所增加。尽管 PointNext 同样采用了可分离式 MLP 架构,单次推理耗时为 1.28 s;而 Point-LaKan 在实现相似特征建模能力的同时,将推理时间控制在 0.804 s,实现了约 37.2% 的速度提升,展现出良好的执行效率与计算资源适应性。综上所述,Point-LaKan 网络在保持较强语义建模能力的前提下,兼顾模型复杂度与运行速度,体现出对高精度与高效率抓取识别任务的良好适应性。

表2 网络参量对比结果

网络模型	网络参量/M	推理时间/s
PointNet++ <sup>[8]</sup>	1.734	0.36
PointMLP <sup>[22]</sup>	16.754	0.974
PointNext <sup>[9]</sup>	16.275	1.28
Point-LaKan	15.512	0.804

3) 抓取识别性能验证。为系统评估所提出的 Point-LaKan 网络在目标识别精度与跨域识别能力方面的综合表现,在构建的复杂堆叠测试点云上开展对比实验。如表 3 所示,与 PointNet 相比,PointNet++ 通过分层特征聚合显著增强了局部几何建模能力,使 Recall 和 IoU 分别提升 23.77%、27.50%;PointNext 在 Precision、F-Score、IoU 上分别提升 0.20%、3.18%、14.65%。然而,这些方法仍依赖多层 MLP 堆叠结构,在处理复杂点云时面临训练收敛缓慢与局部特征提取不足的问题。相比之下,Point-LaKan 在顶部抓取目标识别任务中表现出更优性能:Precision 分别较 PointNet、PointNet++ 提升 25.26%、19.89%;Recall 达到 90.62%,相比 PointCNN、KPConv 分别提升 26.60%、21.67%;F-Score 达到 92.79%,相较 PointNext 提升

15.46%。这些结果表明,Point-LaKan 在大体积点簇结构中展现出更强的结构敏感性与特征判别能力,有效支撑了非结构化场景下的高精度抓取策略。

表3 不同点云处理网络的性能比较 %

网络模型	Precision	Recall	IoUs	F-Score
PointNet <sup>[7]</sup>	70.46	49.11	40.06	52.82
PointNet++ <sup>[8]</sup>	75.83	58.67	46.92	61.77
PointCNN <sup>[23]</sup>	80.14	64.02	52.31	66.39
KPConv <sup>[24]</sup>	84.32	68.95	57.06	70.81
PointMLP <sup>[22]</sup>	86.74	72.13	60.85	74.52
PointNext <sup>[9]</sup>	88.56	75.48	63.77	77.33
Point-LaKan	95.72	90.62	86.82	92.79

在多物体密集堆叠场景中(见图 7),PointNet 与 PointNet++ 缺乏多尺度上下文建模能力,导致识别结果易受遮挡与邻近干扰影响,表现为抓取区域定位不稳定,难以提取唯一有效的目标区域。PointNext 通过模块化特征增强提升了整体感知能力,但在复杂背景中仍存在误识别现象,部分非目标点云被错误判定为可抓取区域。相比之下,所提出的 Point-LaKan 凭借其结构感知与边界判别能力,能稳定提取目标物体的完整抓取表面,并有效抑制背景噪声对识别结果的干扰。特别是在高密度堆叠测试中,Point-LaKan 能够准确定位高置信抓取区域,展现出更强的识别鲁棒性与跨场景泛化能力,为后续姿态估计提供可靠支撑。

### 4.3 机器人抓取实验

#### 4.3.1 仿真场景抓取实验

为验证所提出算法在抓取精度与泛化能力方面的性能表现,在 CoppeliaSim 仿真平台构建多个具有不同复杂度的堆叠场景,并在多种未参与训练的物体上开展抓取实验。实验分别在 UR5 机械臂末端安装真空吸盘与平行抓钳两类执行器,以评估对不同末端操作策略的适应性。实验在 30 组非训练场景中以 15 ~ 40 个物体随机堆叠方式进行测试。

图 8 展示了在上述测试场景中,不同末端执行器在未知场景下对未知物体的抓取示例结果。统计数据显示,所提出算法在真空吸盘执行策略下的 SR 可以达到 99.58%,在平行抓钳配置下的 SR 可以达到 92.11%,均表现出较高的鲁棒性与适应性。特别是在混合非规则物体的高复杂度堆叠场景中(见图 8(a)与图 8(b)),算法仍能准确识别顶部具有高抓取置信度的区域,并输出满足执行需求的 6-DOF 抓取姿态,充分验证了其在非结构化环境下的识别准确性与策略泛化能力。上述结果表明,本文方法不仅

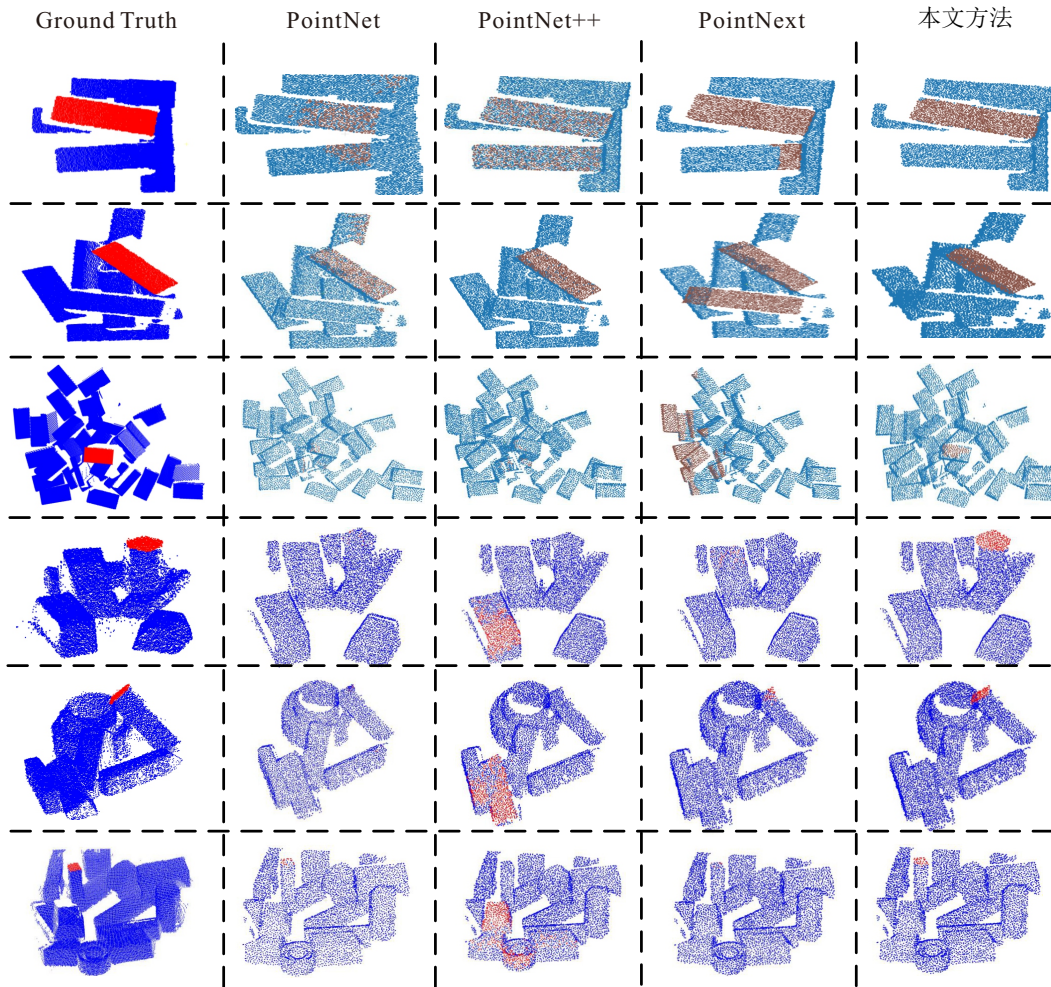
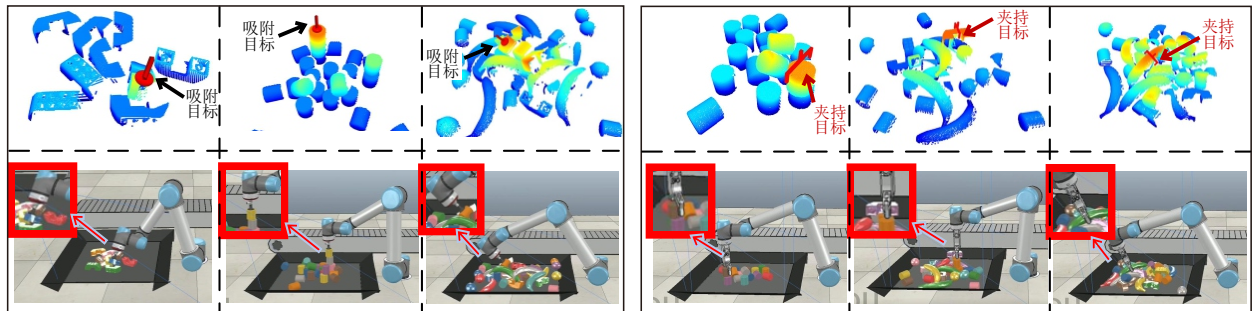


图7 抓取识别结果对比



(a) 真空吸盘作为机器人末端执行器

(b) 二指抓钳作为机器人末端执行器

图8 仿真场景下不同末端执行器的抓取效果示例

适用于常规规则物体抓取场景, 在面向复杂、动态及不可预测物体堆叠情境下亦具备良好的通用性与工程应用潜力.

### 4.3.2 真实场景抓取实验

为系统评估所提出算法在真实场景中的抓取性能与执行稳定性, 本节在 MPDG-Real 数据集选取 50 组点云样本作为验证集, 其中包含 25 组单一物体点云场景和 25 组多物体堆叠点云场景. 需特别指出, 表 4 所列的性能数据基于 25 组多物体堆叠场景样本. 针对每种姿态估计算法, 分别独立运行 10 次推

理流程, 并记录每次的抓取成功率 SR 和完成率 CR,

表4 不同抓取位姿检测方法对比结果 %

网络模型	SR	CR
SE-ResUNet <sup>[25]</sup>	50.77	74.94
PointNetGPD <sup>[6]</sup>	67.01	84.17
GQ-CNNs <sup>[26]</sup>	65.00	83.98
PointNet++ Grasping <sup>[11]</sup>	71.43	91.43
Anygrasp <sup>[15]</sup>	72.31	89.16
Generalizing-Grasp <sup>[27]</sup>	76.92	92.20
本文算法	80.17	93.33

最终统计指标取平均值. 并引入上限机制避免机械臂产生冗余抓取操作: 当场景中物体数  $\geq 10$  时, 将最大抓取执行次数设定为物体总数的 150%.

为验证各核心模块对抓取性能的具体提升作用, 表 5 展示了消融实验结果. 结果表明, “维度扩展”对应第 2.1 节提出的几何先验建模策略, 显著增强了输入点云的空间结构表达能力, 使抓取区域识别网络在多类场景下的 SR 均超过 90%, CR 维持在接近 90% 的高水平; “LAKAN”对应第 2.2 节提出的特征提取结构, 分别在单物体与堆叠场景下使抓取准确率提升 12.33%、12.2%; “几何约束”对应第 3.2 节与第 3.3 节提出的方向向量融合策略, 在堆叠任务中使任务完成率进一步提升 7.18%, 有效提升了抓取姿态的可执行性与规划合理性. 上述结果表明, 各模块从输入建模、特征提取到姿态生成环节均对抓取性能产生正向影响, 其协同作用显著提升了整体算法在非结构化复杂场景中的执行效率与稳定性.

表5 所提出算法的消融实验结果

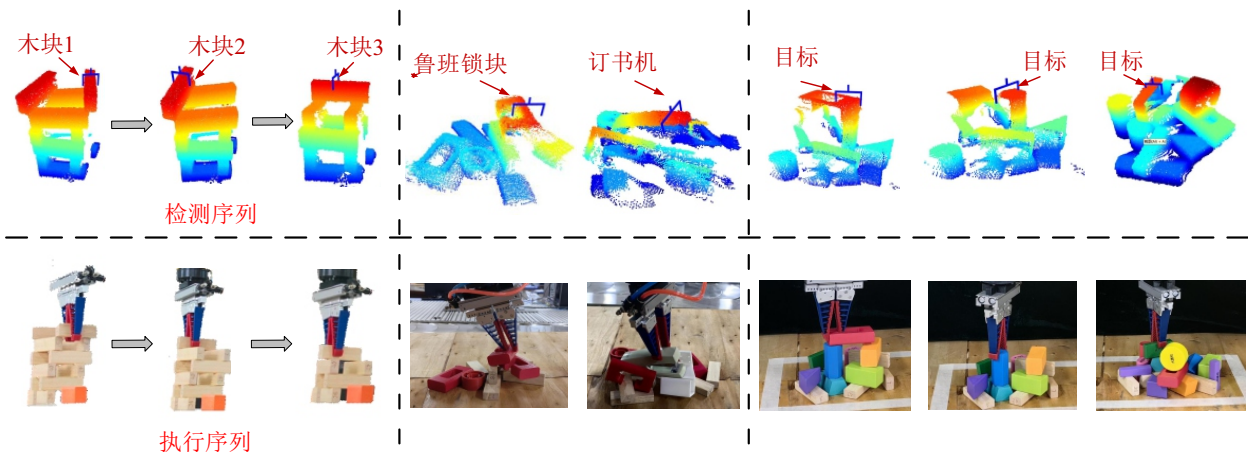
维度扩展	LAKAN	几何约束	单一物体场景		堆叠物体场景	
			SR/%	CR/%	SR/%	CR/%
			67.69	59.09	49.23	36.53
		✓	69.15	59.21	50.05	43.71
	✓	✓	80.02	67.75	61.43	46.35
✓	✓	✓	90.31	91.18	90.48	89.47

为进一步验证所提出算法在点云抓取识别任务中的综合性能, 选取 6 种近年具有代表性的基准方法进行对比评估, 如表 4 所示. 结果显示, 所有对比方法的 SR 均未超过 77%, 任务完成率 CR 普遍低于 93%, 在单次抓取的准确性与任务执行的稳定性方面存在明显不足. 相比之下, 本文方法在识别精度与姿态生成可靠性方面表现优异, 抓取成功率最高

提升 29.40%, 任务完成率最高提升 18.39%, 在各评价指标上均取得显著优势. 综上, 本文方法在面对多物体堆叠的复杂场景时, 展现出优于现有主流方法的识别性能与执行鲁棒性, 验证了其在实际抓取任务中的有效性.

此外, 本实验所使用的网络模型完全基于 CoppeliaSim 平台生成的仿真点云数据进行训练. 实验结果表明, 该方法可在“未见真实数据”的前提下稳定迁移至实际环境, 体现了所提出 Point-LaKan 网络在 Sim-to-Real 迁移中的可行性与泛化潜力. 图 9 进一步展示了算法在真实场景下的抓取实例: 图 9(a) 记录了算法在混合堆叠的多木块场景中连续完成多目标抓取的全过程, 验证了其在实际搬运与码垛任务中的应用能力; 图 9(b) 和图 9(c) 则呈现了算法在未知物体与未知场景中的 6-DOF 抓取姿态生成结果, 表现出良好的稳定性与执行精度, 有效支撑了跨物体、跨场景的识别与策略泛化能力. 综合来看, 本文方法能够在复杂非结构化场景中准确识别高置信度抓取区域, 并生成稳定、可执行的抓取姿态, 体现出其在实际应用中的可行性、鲁棒性与应用价值.

图 10 展示了所提出算法在异形物体上的抓取实验结果, 测试对象包括螺丝刀、钳子等常见生活工具类物体. 这类物体通常具有非对称结构、细长形态或局部几何突变, 传统方法在此类场景下常面临抓取点识别困难、姿态估计误差大等挑战. 实验结果表明, 本文算法能够准确识别适宜的抓取区域并引导机械臂完成稳健操作, 单次抓取成功率 (SR) 达 100%. 该结果进一步表明, 本文方法不仅在堆叠场景中具备较强表现力, 在面对高几何复杂度与非规则结构物体时亦具有较高的鲁棒性与适应性, 具备良好的实际部署前景.



(a) 连续性任务下的检测和执行序列 (b) 非训练集物体的抓取示例 (c) 非训练集堆叠场景下的抓取示例

图9 真实场景下的抓取效果示例

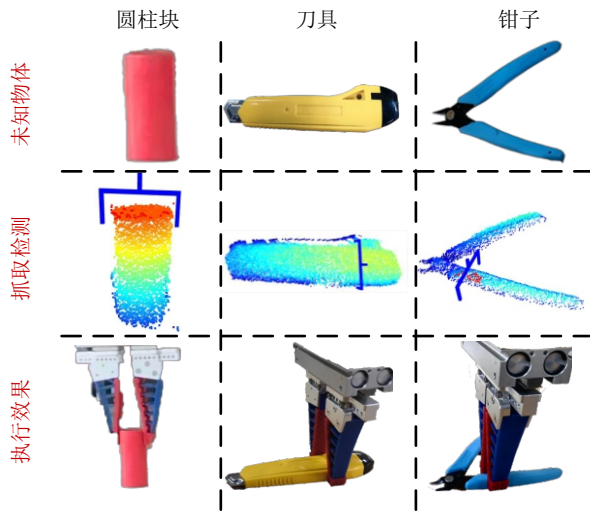


图10 规则以及异形物体的抓取示例

## 5 结论

本文提出了一种融合几何先验建模与姿态质量评估机制的点云 6-DOF 抓取姿态联合预测方法,旨在提升非结构化堆叠场景中目标识别的几何感知能力与姿态生成的执行稳定性. 首先,构建几何先验引导的点云识别网络 Point-LaKan,有效提升了对堆叠抓取区域的结构表征能力;其次,设计结合方向向量约束的抓取姿态估计策略,增强抓取姿态的可执行性与策略可解释性. 进一步地,提出融合方向约束、碰撞检测与质心评分的多因素姿态筛选机制,实现高置信度候选姿态的可执行性排序. 通过自主构建的多类别堆叠物体点云抓取数据集,并开展充分的实验验证,结果表明所提出方法在识别准确率与执行成功率等方面均优于现有主流方法,展现出良好的工程实用性与跨域适应能力. 未来的研究将重点围绕以下两个方向开展:

1) 进一步提升模型对遮挡干扰、物体形变等复杂场景的适应能力,增强其在实际操作中的稳定性;

2) 引入多模态感知机制,实现更强的语义理解与场景感知能力,以提升抓取策略在开放环境下的生成效率与通用性.

## 参考文献 (References)

[1] Du G G, Wang K, Lian S G, et al. Vision-based robotic grasping from object localization, object pose estimation to grasp estimation for parallel grippers: A review[J]. *Artificial Intelligence Review*, 2021, 54(3): 1677-1734.

[2] 王勇, 李邑灵, 苗夺谦, 等. 基于 Transformer-CNN 混合架构的跨模态融合抓取检测[J]. *控制与决策*, 2024, 39(11): 3607-3616.  
(Wang Y, Li Y L, Miao D Q, et al. Cross-modal fusion grasp detection based on a Transformer-CNN hybrid architecture[J]. *Control and Decision*, 2024, 39(11): 3607-3616.)

[3] 楚红雨, 冷齐齐, 张晓强, 等. 融入注意力机制的多模态特征机械臂抓取位姿检测[J]. *控制与决策*, 2024, 39(3): 777-786.  
(Chu H Y, Leng Q Q, Zhang X Q, et al. Robotic arm grasp pose detection with multi-modal features incorporating an attention mechanism[J]. *Control and Decision*, 2024, 39(3): 777-786.)

[4] Liang H Z, Ma X J, Li S, et al. PointNetGPD: Detecting grasp configurations from point sets[C]. 2019 International Conference on Robotics and Automation. Montreal, 2019: 1809-1822. [LinkOut].

[5] Wang C J, Zheng Z, Zha B T, et al. Fast robust point cloud registration based on compatibility graph and accelerated guided sampling[J]. *Remote Sensing*, 2024, 16(15): 2789.

[6] Han Z, Xie X Y, Tang G J, et al. A K-dimensional tree-iterative closest point algorithm for overbreak and underbreak assessment of mountain tunnels[J]. *Applied Sciences*, 2025, 15(2): 566.

[7] Qi C R, Su H, Mo K, et al. PointNet: Deep learning on point sets for 3D classification and segmentation[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, 2017: 652-660.

[8] Qi C R, Yi L, Su H, et al. Deep hierarchical feature learning on point sets in a metric space[C]. *Advances in Neural Information Processing Systems (NeurIPS)*. Long Beach, 2017: 5099-5108.

[9] Qian G, Li Y, Peng H, et al. PointNeXt: Revisiting PointNet++ with improved training and scaling strategies [C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, 2022: 6916-6925.

[10] Zhang Z Y, Wang Q W, Zhang Y X, et al. Physics-informed neural networks with hybrid Kolmogorov-Arnold network and augmented Lagrangian function for solving partial differential equations[J]. *Scientific Reports*, 2025, 15: 10523.

[11] Ni P Y, Zhang W G, Zhu X X, et al. PointNet++ grasping: Learning an end-to-end spatial grasp generation algorithm from sparse point clouds[C]. 2020 IEEE International Conference on Robotics and Automation. Paris, 2020: 3619-3625.

[12] Alliegro A, Rudorfer M, Frattin F, et al. End-to-end learning to grasp via sampling from object point clouds[J]. *IEEE Robotics and Automation Letters*, 2022, 7(4): 9865-9872.

[13] Zhao B, Zhang H, Lan X, et al. REGNet: Region-based grasp network for single-shot grasp detection in point clouds[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(11): 4019-4032.

[14] Liang H Z, Ma X J, Li S, et al. PointNetGPD: Detecting grasp configurations from point sets[C]. International Conference on Robotics and Automation. Montreal, 2019: 3629-3635.

[15] Fang H S, Wang C X, Fang H J, et al. AnyGrasp: Robust and efficient grasp perception in spatial and temporal

- domains[J]. *IEEE Transactions on Robotics*, 2023, 39(5): 3929-3945.
- [16] 蔡子豪, 杨亮, 黄之峰. 基于点云采样权重估计的未知物体抓取位姿生成方法[J]. *控制与决策*, 2023, 38(10): 2859-2866.  
(Cai Z H, Yang L, Huang Z F. Grasping pose generation method for unknown objects based on point cloud sampling weight estimation[J]. *Control and Decision*, 2023, 38(10): 2859-2866.)
- [17] Anvekar T, Bazazian D. GPr-net: Geometric prototypical network for point cloud few-shot learning[C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Vancouver, 2023: 4179-4188.
- [18] Liu Z, Wang Y, Vaidya S, et al. Kolmogorov-Arnold networks[J/OL]. 2024, arXiv: 2404.19756.
- [19] Calli B, Singh A, Walsman A, et al. The YCB object and model set: Towards common benchmarks for manipulation research[C]. 2015 International Conference on Advanced Robotics. Istanbul, 2015: 510-517.
- [20] Zeng A, Song S R, Welker S, et al. Learning synergies between pushing and grasping with self-supervised deep reinforcement learning[C]. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems. Madrid, 2018: 4238-4245.
- [21] Li X H, Zhang X Z, Zhou X, et al. UPG: 3D vision-based prediction framework for robotic grasping in multi-object scenes[J]. *Knowledge-Based Systems*, 2023, 270: 110491-110503.
- [22] Ma X, Guo Y, Guo Y. PointMLP: An MLP-based architecture for point cloud classification[C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, 2022: 963-973.
- [23] Li Y, Bu R, Sun M, et al. PointCNN: Convolution on X-transformed points[C]. *Advances in Neural Information Processing Systems*. Montreal, 2018: 820-830.
- [24] Thomas H, Qi C R, Deschaud J E, et al. KPConv: Flexible and deformable convolution for point clouds[C]. 2019 IEEE/CVF International Conference on Computer Vision. Seoul, 2019: 6411-6420.
- [25] Yu S, Zhai D H, Xia Y Q, et al. SE-ResUNet: A novel robotic grasp detection method[J]. *IEEE Robotics and Automation Letters*, 2022, 7(2): 5238-5245.
- [26] Mahler J, Matl M, Satish V, et al. Learning ambidextrous robot grasping policies[J]. *Science Robotics*, 2019, 4(26): eaau4984.
- [27] Ma H X, Shi M D, Gao B Y, et al. Generalizing 6-DoF grasp detection via domain prior knowledge[C]. 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, 2024: 18103-18109.

## 作者简介

李小明 (1993-), 女, 助理教授, 博士, 硕士生导师, 主要研究方向为智能机器人技术、具身智能, E-mail: [lixiaohan272@gmail.com](mailto:lixiaohan272@gmail.com);

张哲戡 (1998-), 男, 硕士生, 主要研究方向为视觉引导的机器人抓取, E-mail: [zjzhang@xauat.edu.cn](mailto:zjzhang@xauat.edu.cn);

徐胜军 (1976-), 男, 教授, 博士, 博士生导师, 主要研究方向为人工智能与智能化系统、模型仿真智能控制理论, E-mail: [duplin@sina.com](mailto:duplin@sina.com);

孟月波 (1979-), 女, 教授, 博士, 博士生导师, 主要研究方向为智能感知、理解与智能化系统以及建筑智能化技术, E-mail: [mengyuebo@163.com](mailto:mengyuebo@163.com);

杨玥 (1986-), 女, 副教授, 博士, 硕士生导师, 主要研究方向为智能无人系统的安全控制, E-mail: [yangyue@xauat.edu.cn](mailto:yangyue@xauat.edu.cn);

武经 (1989-), 女, 副教授, 博士, 硕士生导师, 主要研究方向为机械设计、串并联机械臂运动学分析, E-mail: [wuj@sustech.edu.cn](mailto:wuj@sustech.edu.cn);

苏晓杰 (1985-), 男, 教授, 博士, 博士生导师, 主要研究方向为智能控制系统及其在无人系统中的应用, E-mail: [suxiaojie@cqu.edu.cn](mailto:suxiaojie@cqu.edu.cn);

陈义明 (1964-), 男, 教授, 博士, 博士生导师, 主要研究方向为机器人与自动化、模块化可重构机器人系统设计, E-mail: [michen@ntu.edu.sg](mailto:michen@ntu.edu.sg).