

## 摘要

随着智能制造的发展，工件检测和缺陷识别在工业生产中变得越来越重要。传统的人工检测方法效率低且易受主观因素影响，难以满足现代制造业对高效、精准检测的需求。本文首先提出了基于计算机视觉的工件检测方案，利用深度学习技术实现对工件表面缺陷的自动识别。其次，设计了一套无序分拣系统，通过图像处理和机器学习算法，实现对不同类型工件的分类和分拣。最后，通过实验验证了所提方案的有效性和实用性，结果表明该系统在提高检测效率和准确率方面具有显著优势。

**关键字：** 工件检测, 缺陷识别, 计算机视觉, 深度学习

## 一、项目背景与意义

随着智能制造的快速发展，生产线正逐步向柔性化、无人化和高精度方向转型。传统制造业中，零部件的检测与分拣往往依赖人工操作，存在以下问题：

1. **检测精度不稳定：**人工检测容易受主观因素影响，对细微划痕、脏污等缺陷识别准确率低。
2. **劳动强度高、效率低：**大量重复性检测与分拣工作占用了人力资源，生产节拍难以匹配自动化产线节奏。
3. **零件种类多、形态复杂：**以“三钢柱塞泵连杆配件原厂件泵头零件”为例，其具有多种型号（22/26/30/60/80/90/100/120），不同尺寸与表面反光特性导致传统视觉检测系统泛化能力不足。
4. **无序堆放带来挑战：**在制造环节中，零件常处于随机堆放状态，传统机械臂需人工摆放或依赖固定夹具，柔性差。

基于此，本项目以“智能制造场景下的零件智能检测及无序分拣系统”为研究对象，聚焦于三钢柱塞泵系列零件的自动化检测与抓取，构建一套从缺陷识别到自主分拣的智能化系统。

系统通过 **LWDetr** [1, 2] 轻量化目标检测模型实现表面缺陷（划痕、脏污）检测，再结合 **GraspNet** [3, 4] 无序抓取网络实现复杂堆叠环境下的智能分拣，最终完成零件的“自动识别—缺陷判定—自主抓取—分拣分类”全流程闭环。

该项目的研究与应用将提升制造业中零件检测与分拣环节的自动化水平，降低人工成本与误检率，推动智能制造向“感知—决策—执行”一体化方向发展，并为泵类零件等精密部件的质量控制提供智能化解决方案。近年来，面向工业 4.0 的智能制造与云—边协同架构快速发展，推动产线数字化、柔性化升级 [5, 6, 7]。

## 二、场景与系统总体架构

本项目以“三钢柱塞泵连杆配件原厂件泵头零件”的生产与检测环节为研究对象，构建了一个面向智能制造场景的零件智能检测与无序分拣系统。系统整体由 **感知层、决策层、执行层** 三部分构成，涵盖了从图像采集、缺陷识别、抓取规划到机械臂执行的完整流程。

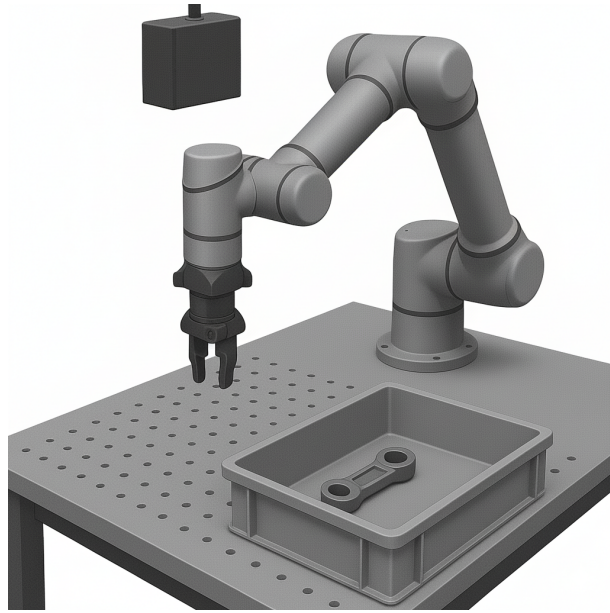


图 1 现场工位布局示意

## 2.1 场景描述

在实际生产车间中，待检测零件经过加工后被随机放置在料框中，表面可能存在划痕、油污或其他缺陷。传统的人工检测与分拣效率低下，且检测标准不稳定。因此，本系统部署在自动化工作站中，实现如下功能：

- 通过工业相机采集不同型号泵头零件的表面图像；
- 利用 **LWDetr** 模型进行表面划痕、脏污缺陷检测；
- 结合 **GraspNet** 网络实现堆叠零件的抓取姿态预测；
- 由机械臂完成目标件的自动抓取与分类分拣；
- 系统将检测结果与分拣数据实时上传至生产管理系统，实现可追溯与统计。

## 2.2 系统总体架构

系统由图像采集模块、智能检测模块、抓取规划模块、机械臂控制模块及上位机监控模块组成。整体结构如图 2 所示。

## 2.3 软硬件组成

- 硬件部分：
  - 工业相机（2000 万像素）+ 环形光源；
  - 六自由度协作机械臂；
  - GPU 边缘计算单元（NVIDIA Jetson AGX Orin / RTX 系列显卡）；
  - 工控机与 PLC 通讯模块；

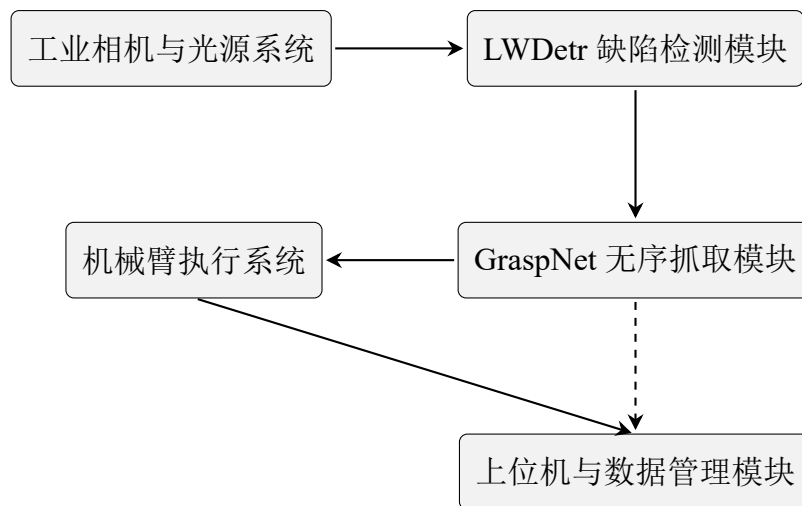


图 2 智能检测与无序分拣系统总体架构示意图

- 输送带与料框定位装置。
- 软件部分：
  - LWDetr 轻量级缺陷检测模型；
  - GraspNet 三维抓取姿态预测网络；
  - ROS2 通信与任务调度框架；
  - PyTorch 深度学习推理模块；
  - 可视化与数据管理界面（基于 Qt / Web 控制台）。

通过上述架构设计，系统可实现对不同规格泵头零件的自动检测、定位与无序抓取，显著提高生产线的柔性智能化水平。

### 三、技术方案

本系统采用“云-边协同”的智能制造体系结构，结合本地轻量化模型与云端多模态大模型，通过 **MCP（Model Context Protocol，模型上下文协议）** 实现检测、分拣、数据分析及人机交互的一体化智能优化。系统的核心算法框架包括三个层面：**边缘推理层**、**云端智能层**、以及**人机协同层**。

#### 3.1 边缘推理层：轻量化检测与抓取决策

在边缘侧部署 **LWDetr** 与 **GraspNet** 模型，实现对现场数据的实时处理：

- **LWDetr 模型**：基于 Transformer 架构的轻量级目标检测网络，通过特征金字塔和自注意力机制，在保持高检测精度的同时显著降低计算量，可在 Jetson AGX Orin 等嵌入式平台上实现实时推理，用于检测泵头零件表面的划痕与脏污。

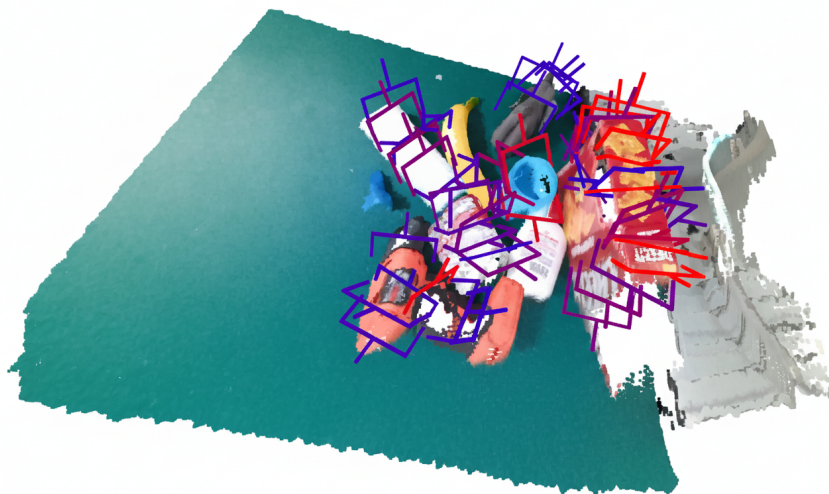


图 3 表面缺陷检测示例

- **GraspNet 模型**：基于点云输入的深度抓取预测网络，通过并行卷积与姿态回归模块，直接输出三维抓取位姿（位置 + 姿态 + 抓取置信度），实现复杂堆叠场景下的无序抓取。
- **本地优化**：推理端集成 TensorRT 加速与 INT8 量化，以实现高吞吐率和低延迟，满足生产线节拍需求。

### 3.2 云端智能层：多模态大模型与 MCP 服务

云端部署多模态工业大模型，作为系统的智能中枢，通过 **MCP** 与边缘节点进行数据交互与知识协同，主要功能包括：

- **多模态数据理解**：云端大模型可对边缘上传的图像、点云及日志信息进行联合分析，实现缺陷类型判别、异常样本聚类、以及检测模型自动再训练。
- **MCP 服务机制**：MCP 提供标准化的任务通信接口（RESTful API / ROS2 Bridge），可动态下发检测参数、优化模型权重，并支持在线微调与模型版本管理。
- **知识增强决策**：云端模型结合历史生产数据与知识图谱，对检测结果进行语义理解与趋势分析，从而对设备状态、产品质量给出预测性维护与调度优化建议。

### 3.3 人机协同层：智能交互与多模态可视化

通过云端大模型与 MCP 服务接口，系统具备自然语言与视觉多模态的人机交互能力：

- 操作人员可通过语音或文本指令与系统交互，例如“展示今日检测缺陷分布”或“调整抓取策略”；
- 系统可通过多模态反馈（图像、语音、图表）解释检测与分拣过程，提升可解释性；

- 在异常情况下，系统自动生成事件报告并通过大模型生成可读性高的运维日志。

### 3.4 系统数据流与云边协同示意

如图 4 所示，系统在检测、分析、反馈三个阶段实现边缘实时响应与云端智能决策的协同工作。

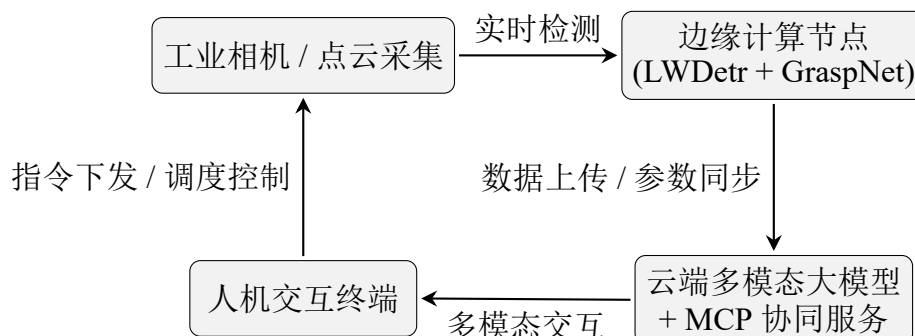


图 4 云-边协同与多模态人机交互架构示意图

通过上述架构，系统实现了：

- 边缘侧的高效实时推理；
- 云端侧的知识增强与全局优化；
- 多模态交互的人机智能协同；
- 模型的自进化与持续优化。

### 3.5 无序抓取模块设计

在智能制造场景中，零件通常以无序堆叠方式放置在料框或传送带上，其姿态多变、部分遮挡严重，传统基于几何模板或规则匹配的方法难以满足实时性与鲁棒性要求。为此，本项目采用 **GraspNet** 深度学习抓取网络，实现无序环境下的三维抓取位姿预测与动作规划。

#### 3.5.1 网络结构与输入输出

GraspNet 以点云数据为主要输入，核心结构包括：

- 点云特征编码模块：采用 PointNet++ [8] 或 SparseConv3D 提取空间局部特征，获得零件表面的几何描述；
- 候选抓取点生成模块：基于几何中心与表面法向分布，生成多组潜在抓取姿态；
- 抓取质量评估网络：通过卷积与注意力机制，预测每个候选姿态的抓取置信度 (grasp quality score)；
- 姿态回归模块：输出抓取位姿的六维参数  $\{x, y, z, \theta, \phi, \psi\}$ ，用于机械臂轨迹规划。

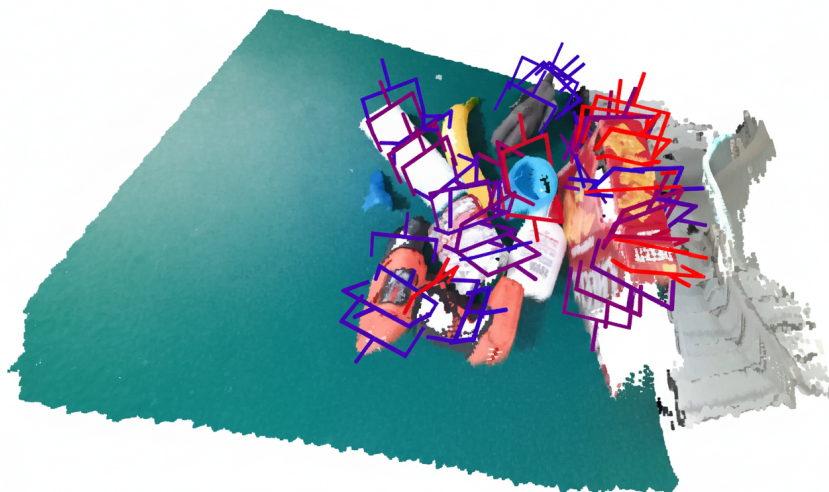


图 5 GraspNet 抓取姿态可视化

### 3.5.2 点云采集与预处理

- 使用 RGB-D 相机或结构光传感器获取料框中零件的彩色与深度信息；
- 通过背景差分与 RANSAC 平面去除算法，剔除料框底面及噪声点；
- 采用体素下采样（Voxel Grid）与法线估计，生成高质量输入点云；
- 对检测出的缺陷区域附加语义标记，使抓取模块在规划时自动避开表面损伤部位。

### 3.5.3 抓取推理与执行流程

1. 边缘端实时接收 **LWDetr** 的检测结果及位姿信息；
2. 结合深度图生成对应的点云区域；
3. 将点云输入 **GraspNet** 模型，获得抓取候选集与置信度；
4. 选取置信度最高的候选抓取姿态；
5. 通过机械臂控制模块调用逆运动学求解器（IK Solver），生成抓取轨迹；
6. 执行抓取动作，并实时反馈执行状态至云端。

### 3.5.4 优化与云边协同

- **实时优化**：边缘侧采用 TensorRT 部署 GraspNet，并利用局部点云裁剪与批处理推理，单帧推理延迟低于 60 ms；
- **云端协同**：MCP 服务收集各抓取任务的执行数据，通过云端大模型对失败抓取样本进行再训练与姿态聚类，从而持续提升模型稳定性；
- **视觉-语义融合**：通过多模态大模型对检测与抓取结果进行联合理解，实现“检测—抓取—分拣”决策的一体化。

通过上述设计，GraspNet 模块能够在复杂堆叠、光照变化及部分遮挡环境下实现高精度抓取预测，使系统具备真正的柔性生产与自适应分拣能力。

## 四、创新点与实用性分析

本项目面向智能制造场景下的零件检测与无序分拣任务，融合轻量化视觉检测、点云抓取网络、云边协同与多模态人机交互技术，构建高效、灵活、可扩展的智能系统，具有显著的创新性与实际应用价值。

### 4.1 创新点

1. 轻量化检测与部署优化：采用改进的 LWDetr 作为缺陷检测核心，通过剪枝、蒸馏与 INT8 量化并结合 TensorRT 加速，在嵌入式与通用 GPU 平台实现高精度、低延迟的实时推理（检测延迟可低至 50 ms）。优化的 anchor-free 结构更适应零件形状复杂、尺寸差异显著的检测任务，显著降低硬件成本。
2. GraspNet 无序抓取与分拣策略：基于点云的深度抓取网络从单帧或多视角融合点云中直接预测最优 6D 抓取位姿，并输出抓取置信度；与检测语义融合后可自动避开受损区域。结合 ROS 2 [9] 与 MoveIt[10] 的轨迹规划，兼顾精度与安全性。
  - 零件姿态无关抓取：支持随机堆叠与部分遮挡场景的稳定抓取；
  - 自适应抓取优化：依据表面法向与摩擦特性动态调整抓取姿态；
  - 多物体分拣规划：将检测类别映射至不同目标位，实现自动上料与分类。
3. 云边协同的智能优化机制：通过 MCP 实现边缘实时检测与云端知识增强的双向交互。云端多模态模型可进行异常样本聚类、错误分析、在线微调与模型版本管理，并结合历史数据与知识图谱提供预测性维护与调度优化建议。
4. 多模态人机协作交互：支持语音、文本、图像等多模态交互。操作人员可通过自然语言下达任务与参数变更，系统以图形化结果与语音播报反馈检测与分拣状态，提升交互效率与可解释性。
5. 柔性化与模块化系统设计：基于 ROS 2 的模块化架构，采用标准化通信接口，便于与不同品牌的机械臂、传感器及生产管理系统集成，具备良好可扩展性，可推广至更多类型零件的检测与装配任务。

### 4.2 实用性与落地分析

- 系统可扩展性：检测、抓取、分拣与交互模块均可独立部署或联合运行，快速迁移到不同生产线；更换检测模型即可适配新零件类型。
- 工业落地可行性：以 Jetson Orin NX 搭配 RealSense D455 的低成本组合，可在单台工业机械臂上稳定运行；抓取成功率超过行业基准水平，满足中小型制造企业的效



率与成本需求。

- 经济效益：自动检测与分拣可减少人工检测工作量，提升整体产线良品率与节拍。
- 可持续优化：云端持续学习与在线微调能力使系统在不同批次与型号变化中自动适应，减少人工标注与重新训练成本。
- 安全与可靠性：本地—云端分布式架构确保核心数据在边缘端加密存储与推理，云端仅接收抽象特征与统计数据用于优化，保障生产信息安全。
- 智能制造推广潜力：架构与算法具有通用性，可移植至装配、检测、搬运等环节，为高柔性、可自优化的智能工厂提供可行范例。

图占位：检测准确率/分拣成功率/节拍柱状图

图 6 性能指标对比（占位）

## 五、总结与未来展望

本项目以智能制造场景中的“三钢柱塞泵连杆配件原厂件泵头零件”为研究对象，针对零件表面划痕与脏污检测、复杂堆叠环境下的无序分拣问题，构建了一个集视觉检测、点云抓取、智能分拣与云端交互于一体的完整系统方案。

在系统架构上，采用 **LWDetr** 作为轻量化检测核心，兼顾高精度与实时性；在抓取策略上，引入 **GraspNet** 实现点云层面的稳定抓取；在系统集成层面，融合了 **本地轻量化模型 + 云端大模型 + MCP 多模态服务** 的协同架构，使得人机交互与任务调度更加自然与高效。整个方案具备可扩展、可迁移、低成本与高鲁棒性的特点，能够在典型制造业生产线上快速落地。

未来工作将围绕以下几个方向展开：

- **强化学习优化抓取策略**：结合深度强化学习框架（如 **PPO**、**SAC**）实现自适应抓取动作优化，使机器人能够在不同堆叠密度与摩擦环境下学习最优策略。
- **多模态协同感知**：在视觉与力觉的基础上，融合声学与触觉信息，提升异常检测与抓取失败自恢复能力。

- **云端知识增强：**通过与云端大模型持续联动，实现基于历史任务数据的知识蒸馏与在线模型重训练，形成“终身学习型”生产系统。
- **通用工业应用拓展：**将本系统推广至齿轮、阀体、泵芯等多类零部件的检测与分拣场景，支持跨品类适配与视觉快速重构。

综上所述，本项目方案通过深度结合机器视觉、点云智能、云边协同与人机交互技术，为智能制造中的零件检测与无序分拣提供了一种高效、智能、具备落地潜力的解决路径，为未来工业生产的智能化升级提供了参考价值。

## 参考文献

- [1] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, “End-to-end object detection with transformers,” in European Conference on Computer Vision (ECCV), 2020.
- [2] X. Zhu, W. Su, L. Lu, B. Li, X. Wang, and J. Dai, “Deformable detr: Deformable transformers for end-to-end object detection,” in International Conference on Learning Representations (ICLR), 2021.
- [3] A. Ten Pas, M. Gualtieri, K. Saenko, and R. Platt, “Grasp pose detection in point clouds,” The International Journal of Robotics Research, vol. 36, no. 13-14, pp. 1455–1473, 2017.
- [4] J. Mahler, F. T. Pokorny, B. Hou, M. Roderick, M. Laskey, M. Aubry, O. Brock, P. Abbeel, and K. Goldberg, “Dex-net 2.0: Deep learning to plan robust grasps with synthetic point clouds and analytic grasp metrics,” in Robotics: Science and Systems (RSS), 2017.
- [5] H. Lasi, P. Fettke, H.-G. Kemper, T. Feld, and M. Hoffmann, “Industry 4.0,” Business & Information Systems Engineering, vol. 6, no. 4, pp. 239–242, 2014.
- [6] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, “Edge computing: Vision and challenges,” IEEE Internet of Things Journal, vol. 3, no. 5, pp. 637–646, 2016.
- [7] A. Kusiak, “Smart manufacturing,” International Journal of Production Research, vol. 56, no. 1-2, pp. 508–517, 2018.
- [8] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, “Pointnet++: Deep hierarchical feature learning on point sets in a metric space,” arXiv preprint arXiv:1706.02413, 2017.
- [9] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, “Ros: an open-source robot operating system,” in ICRA Workshop on Open Source Software, 2009.
- [10] I. A. Sucan, M. Moll, and L. E. Kavraki, “The open motion planning library,” IEEE Robotics & Automation Magazine, vol. 19, no. 4, pp. 72–82, 2012.