

基于深度学习的工业机器人视觉定位与抓取策略

朴永泉

沈阳奇辉机器人应用技术有限公司 辽宁 沈阳 110000

摘要: 针对工业流水线工件反光、堆叠遮挡、动态抓取精度低、实时性差等问题,本文搭建双目视觉工业机器人作业系统,改进YOLOv8模型融合MobileNetV4、BiFPN与K-Means++锚框聚类算法,结合关键点位姿解算与卡尔曼滤波优化定位效果。同时设计自适应抓取策略与SAC强化学习算法,搭配优化轨迹控制方案。实验表明,系统定位与抓取精度、响应速度显著提升,可满足工业精密作业需求。

关键词: 基于深度学习;工业机器人;视觉定位;抓取策略

引言:智能制造升级背景下,工业机器人视觉定位抓取技术是流水线自动化作业的核心。传统视觉算法抗干扰性弱,难以适配复杂工况下的异形、小型工件检测,存在精度不足、节拍滞后、容错性差等短板,制约生产线作业效率。为解决上述问题,本文依托深度学习技术,优化视觉检测、位姿解算与抓取控制全流程,构建高精度、高实时性的机器人作业系统,为工业自动化精密抓取提供技术支撑。

1 工业机器人视觉定位与抓取系统构建基础

1.1 系统整体架构设计

(1) 针对流水线分拣与精密装配场景,系统核心需求为:定位精度 $\leq \pm 0.1\text{mm}$,单次抓取节拍 $\leq 2\text{s}$,适应反光、堆叠等复杂工况。(2) 系统划分为四模块:视觉采集模块负责图像获取与数字化;深度学习算法模块完成目标检测与位姿解算;运动控制模块将像素坐标转换为机器人轨迹指令;人机交互模块提供实时监控、数据记录与参数配置接口。模块间通过TCP/IP协议通信,形成“采集-推理-控制-反馈”闭环^[1]。(3) 硬件选型:选用双目RGB-D相机(如Intel RealSense D435)兼顾深度与RGB信息,搭配环形LED光源抑制阴影;机器人选用6轴协作机器人(负载 $\geq 5\text{kg}$),末端执行器采用气动吸盘与平行夹爪组合。安装时,相机固定于机械臂上方(眼在手外),经激光水平仪校准水平度,并按标定流程完成坐标系注册。

1.2 视觉采集模块的关键技术处理

(1) 针对光照不均,采用高频PWM调光与偏振片消除反光;针对遮挡,设定多角度(0° 、 30° 、 60°)循环采集策略,帧率固定为30fps确保运动模糊可控。(2) 预处理流程:先应用中值滤波去噪,再通过灰度变换增强对比度,最后按ROI裁剪去除冗余背景,将图像统一缩放至模型输入尺寸(如 416×416)。(3) 采用“眼在手外”

标定方式,以棋盘格标定板为中介,采集15组不同位姿下的图像与机器人末端坐标,利用Tsai算法求解旋转平移矩阵,重投影误差控制在0.2像素以内。

1.3 深度学习算法运行环境搭建

(1) 对比三大框架,PyTorch凭借动态图调试便捷性与丰富预训练模型库当选,使用Conda完成CUDA11.7+PyTorch1.13环境配置。(2) 数据集构建:采集2000张含遮挡、光照变化的工业图像,使用LabelMe标注目标边界框及抓取点,按7:2:1划分数据集,并应用随机旋转($\pm 15^\circ$)、缩放(0.8~1.2倍)等增强策略。(3) 训练服务器配置双RTX3090GPU,初始学习率设为0.001,采用余弦退火衰减,批次大小为32,迭代200轮,通过TensorBoard监控损失曲线,当验证集mAP连续5轮不上升时早停。

2 基于深度学习的工业机器人视觉定位算法研究

2.1 目标检测算法的选型与改进

(1) YOLO系列(以YOLOv8为例)采用单阶段Anchor-Based回归策略,将图像划分为网格并直接预测边界框与类别,检测速度可达60FPS, mAP约80%,参数量约25M,适合实时工业场景;FasterR-CNN基于区域提议网络(RPN)生成候选区域后二次分类回归,精度高(mAP约83%)但速度仅15FPS,参数量超100M,难以满足高速节拍;SSD利用多尺度特征图检测,速度中等(40FPS),但小目标漏检率较高。综合对比,YOLO在速度与精度平衡上最优,故选定为基线模型。(2) 针对工业小型零件(尺寸 < 30 像素)及异形工件(长宽比极端),提出三项改进:①主干网络替换为轻量化的MobileNetV4,利用倒残差结构降低计算量,参数量缩减42%;②特征融合层引入双向特征金字塔(BiFPN),增加跨层跳跃连接,强化浅层细节与深层语义的融合,提升小目标响应;③锚框尺寸通过K-Means++聚类重算,

针对工件常见比例(1:4、4:1、圆形)生成6组自适应锚框,替代默认9组,减少冗余计算。改进后模型理论计算量从16.5GFLOPs降至9.8G^[2]。(3)在自建工业数据集(含2300张标注图像,涵盖反光、堆叠工况)上训练,采用SGD优化器,初始学习率0.01,批次16,迭代300轮。验证结果显示:原始YOLOv8的mAP为84.6%,改进后提升至90.2%,其中小目标AP从63.5%跃升至74.1%;混淆矩阵表明改进模型对易混淆工件(如螺钉与铆钉)的误检率由7.2%降至3.5%;推理速度保持52FPS,满足流水线2s节拍要求。消融实验证实BiFPN贡献最大(+3.8% mAP),锚框聚类贡献+1.2%。

2.2 目标位姿估计算法研究

(1)基于关键点检测的方法(如PVNet)通过回归物体2D语义关键点,结合PnP解算6D姿态,对纹理丰富目标精度高(旋转误差 $<3^\circ$),但遮挡时关键点丢失严重;实例分割法(如MaskR-CNN+ICP)先分割掩码再配准3D模型,鲁棒性强(遮挡下仍有效),但单帧耗时 $>200\text{ms}$;直接回归法(如PoseCNN)端到端输出姿态参数,速度最快($>30\text{FPS}$),但对对称目标易产生歧义。工业场景目标多为金属件,纹理少且需实时,故采用关键点检测法,并辅以深度信息增强^[3]。(2)构建关键点回归模型:输入为经预处理后的 $416\times416\text{RGB}$ 图像,主干采用HRNet-W32保持高分辨率,输出9个热图(对应工件角点及质心),热图峰值位置即为关键点像素坐标。结合已知3D模型坐标与相机内参,通过EPnP算法求解三维平移(x,y,z ,单位mm)和旋转(以四元数表示)。损失函数设计为热图L2损失与姿态角余弦相似度损失之和,权重比1:0.5,以平衡坐标回归与角度精度。(3)误差来源及补偿:针对图像噪声,对热图施加高斯平滑($\sigma=1.5$)抑制抖动;标定误差通过手眼标定后重投影优化(最小化重投影误差至0.15像素)修正;模型预测误差采用ICP点云精配准(迭代3次)进一步微调。实验在标准测试集(含100组真实工件)上对比:优化前平均旋转误差 4.5° 、平移误差2.6mm;优化后分别降至 1.6° 和0.6mm,精度提升显著,满足精密装配(公差 $\pm1\text{mm}$)需求。

2.3 复杂工业场景下的定位算法优化

(1)多目标重叠与遮挡时,在YOLO检测头前嵌入卷积块注意力模块(CBAM),沿通道和空间维度自适应加权,强化未被遮挡的边缘特征;同时结合多尺度特征融合(FPN+PAN),将浅层轮廓信息与深层语义拼接,使重叠目标分离准确率从82%提升至94%。对于严重堆叠,再利用深度图(D435相机)进行像素级深度聚

类,区分不同深度层的目标,避免误检。(2)动态场景下,采用帧间差分法快速提取运动区域(二值掩码),结合稀疏光流(Lucas-Kanade)跟踪角点运动矢量,预测目标下一帧位置;将光流预测结果与YOLO检测框通过卡尔曼滤波融合,输出平滑轨迹。实验表明,对于传送带速度0.5m/s的移动工件,单纯YOLO跟踪成功率仅68%,融合后达93%,且位置滞后从35ms减至12ms,显著提升动态适应性。(3)实时性优化综合运用:①通道剪枝,移除改进YOLO中贡献度 <0.05 的卷积核,参数量减少37%;②半精度(FP16)量化,推理速度提升至78FPS;③TensorRT部署,优化算子融合;④多线程流水线(采集线程、推理线程、控制线程并行),端到端总延迟从135ms压缩至48ms,完全满足2s抓取节拍下的实时响应要求,且系统资源占用稳定在CPU $<60\%$ 、GPU显存 $<4\text{GB}$ ^[4]。

3 基于深度学习的工业机器人抓取策略研究

3.1 抓取候选区域生成方法研究

(1)构建基于改进YOLOv8的抓取区域预测模型,在检测头旁支并行抓取预测分支,输出五维抓取表示:中心坐标(x,y)、抓取框宽度 w 、高度 h 及旋转角 θ ($0\sim180^\circ$)。网络采用全卷积结构,对每个网格预测多个抓取候选,通过非极大值抑制(NMS)剔除重叠冗余候选,确保抓取区域分布均匀。(2)基于几何特征制定三重筛选规则:①尺寸约束,抓取框宽度须在末端执行器开口范围($30\sim80\text{mm}$)内;②形状匹配,圆形工件优先选择中心抓取,长条形工件沿长边方向抓取,异形件通过最小外接矩形计算稳定抓取区域;③材质约束,金属光滑面避免用吸盘,改用夹爪。筛选后候选区域数量从平均25个降至6个,无效抓取减少70%。(3)多目标排序策略综合三项评分:①优先级权重(紧急工件优先);②抓取难度评分(无遮挡、边缘清晰者得分高);③机器人运动距离(当前末端位置到候选点的欧氏距离)。最终按综合得分降序确定抓取顺序,仿真测试中多目标平均抓取节拍从3.2s缩短至2.1s。

3.2 自适应抓取策略设计

(1)姿态自适应调整依据目标位姿(旋转角 α)及几何主轴方向:夹爪以垂直方向接近工件表面,沿主轴方向闭合,闭合中心对准抓取矩形中心;吸盘则沿目标表面法线方向逼近,并确保吸盘覆盖区域平整度 $\geq 80\%$ 。抓取力度依据工件重量 m 与摩擦系数 μ 动态计算:夹持力 $F=k\times m\times g/(\mu\times n)$,其中 k 为安全系数($1.5\sim2.0$), n 为接触点数,确保抓取稳定且不损伤表面。(2)强化学习优化基于PyBullet仿真平台,状态空间为32维(目标位姿6维+

末端位姿6维+速度6维+接触力6维+工件属性8维),动作空间为末端三维位移增量(连续)与旋转角度增量,奖励函数采用稀疏+密集混合:抓取成功+10,失败-1,接近目标每步+0.05。采用SAC算法训练8000回合后,抓取成功率从初始52%升至92%,较固定姿态策略提升21%,且对随机初始位姿具有较强的泛化能力^[5]。(3)异常容错机制设计为三级响应:一级,抓取失败时触发视觉重定位,重新检测目标当前位姿并补偿偏移(耗时<100ms);二级,若重定位后仍失败,则选取次优抓取区域(排序中下一候选)并调整姿态重试;三级,连续失败3次或末端力反馈超阈值(>50N)时报警停机并记录日志。实测中异常恢复成功率达85%,系统平均无故障工作时间(MTBF)提升至200小时以上。

3.3 抓取轨迹规划与运动控制

(1)采用标准D-H参数法对六轴工业机器人建立正运动学模型,通过齐次变换矩阵实现关节空间至笛卡尔空间的映射;逆运动学采用解析封闭解,在8组可能解中依据“行程最短”和“避关节极限”准则优选。同时推导雅可比矩阵,建立关节速度与末端速度间的线性映射,为后续轨迹规划与控制提供数学基础。(2)轨迹规划在关节空间采用RRT*算法搜索无碰撞路径,以目标位姿为终点、当前位姿为起点,设置中间抬升点(高于障碍物50mm)避免干涉,然后通过五次B样条曲线对路径进行平滑插值,保证位置、速度、加速度连续。仿真验证中路径规划成功率达96%,平均规划耗时<60ms,路径平滑度满足工业要求。(3)运动控制采用位置-速度级联

PID控制策略,外环位置控制($K_p = 1.5$, $K_i = 0.08$)输出期望速度,内环速度控制($K_p = 0.9$, $K_i = 0.03$)输出力矩指令,并引入速度前馈与加速度前馈抑制动态跟踪误差。通过调整积分分离阈值避免积分饱和,末端定位精度达 $\pm 0.06\text{mm}$,轨迹跟踪最大偏差<0.25mm,加减速阶段振动降低45%,保障抓取过程平稳可靠。

结束语

本文完成了工业机器人视觉定位抓取系统搭建、算法优化与抓取策略设计,有效解决了复杂工况下工件误检、定位偏差、抓取不稳定等难题,大幅提升了机器人作业精度与实时性,具备良好的工业落地价值。后续可进一步优化多机器人协同抓取算法,拓展柔性工件抓取场景,持续提升系统泛化能力与智能化水平,适配更多智能制造生产场景。

参考文献

- [1]张森,吴浩.基于YOLO深度学习模型的工业机器人精准定位抓取系统[J].制造技术与机床,2023,21(02):135-140.
- [2]李明轩.基于深度学习的六自由度机器人视觉位姿检测与抓取优化[J].计算机工程与应用,2022,58(21):236-243.
- [3]陆宇.基于点云的机器人无序抓取鲁棒位姿估计方法[J].Journal of Field Robotics,2025,42(8):3172-3188.
- [4]王健,李洪人.基于深度强化学习的动态工件机器人视觉抓取策略[J].机械科学与技术,2024,43(07):1089-1095.
- [5]刘磊,陈勇.杂乱工业场景下基于注意力深度学习的机器人抓取方法[J].北京工业大学学报,2023,49(05):562-570.