

DOI: 10.7672/sgjs2026030044

# 建筑打磨机器人视觉系统研究\*

季元吉<sup>1</sup>, 郑昊<sup>1</sup>, 金艳萍<sup>1</sup>, 应晔<sup>1</sup>, 何芊熙<sup>2</sup>, 高润东<sup>1</sup>

(1. 上海市建筑科学研究院有限公司上海市工程结构安全重点实验室, 上海 200032; 2. 同济大学, 上海 200092)

**[摘要]** 随着既有建筑加固改造工程规模的不断扩大, 混凝土表面预处理作为加固改造施工的关键环节, 其自动化需求日益迫切。对建筑打磨机器人视觉系统进行研究, 基于机器视觉技术, 采用 YOLOv8 算法识别墙面标记点。首先在不同环境下制备数据集, 利用 Roboflow 平台(计算机视觉开发平台)完成数据标注, 并通过多种数据增强手段模拟实际打磨施工环境, 完成模型训练后通过图像后处理模块判断图像有效性并分类, 为机器人打磨提供依据。研究表明, 经系列优化后的 YOLOv8s 模型在检测精度、鲁棒性和泛化能力上均得到显著提升, 关键指标达到了较高水平, 能够满足实际工程应用对打磨机器人实时性和准确性的基本要求, 但模型在复杂施工现场的综合适配性与长期运行稳定性有待实际工程进一步检验。

**[关键词]** 既有建筑; 加固; 打磨; 机器人; 视觉系统; 深度学习; 数据集

**[中图分类号]** TU741

**[文献标识码]** A

**[文章编号]** 2097-0897(2026)03-0044-10

## Research on the Vision System of Building Grinding Robot

JI Yuanji<sup>1</sup>, ZHENG Hao<sup>1</sup>, JIN Yanping<sup>1</sup>, YING Ye<sup>1</sup>, HE Qianxi<sup>2</sup>, GAO Rundong<sup>1</sup>

(1. Shanghai Key Laboratory of Engineering Structure Safety, Shanghai Research Institute of Building Sciences Co., Ltd., Shanghai 200032, China; 2. Tongji University, Shanghai 200092, China)

**Abstract:** With the continuous expansion of the scale of reinforcement and renovation projects for existing buildings, the demand for automation of concrete surface pretreatment, a key link in reinforcement and renovation construction, has become increasingly urgent. This paper conducts a research on the vision system of building grinding robots, which adopts the YOLOv8 algorithm to identify wall marking points based on machine vision technology. Firstly, datasets are prepared under different environmental conditions, and the Roboflow platform (a computer vision development platform) is used to complete data annotation. Various data augmentation methods are applied to simulate the actual grinding construction environment. After the completion of model training, an image post-processing module is used to judge and classify the image validity, which provides a basis for robot grinding operations. The research results show that the YOLOv8s model after a series of optimizations has achieved significant improvements in detection accuracy, robustness and generalization ability, with its key indicators reaching a high level, which can meet the basic requirements of real-time performance and accuracy for grinding robots in practical engineering applications. However, the comprehensive adaptability and long-term operational stability of the model in complex construction sites need to be further verified in practical engineering.

**Keywords:** existing buildings; reinforcement; grinding; robot; vision system; deep learning; datasets

## 0 引言

当前, 随着我国城市化进程的不断推进, 城市建设与发展已逐渐从增量时代过渡到存量时代, 城

市更新类项目与日俱增, 既有建筑加固改造作为城市更新的重要环节, 其技术水平和施工效率直接影响着城市空间品质提升与建筑可持续利用。

在既有建筑加固改造工程中, 混凝土表面预处理是确保新旧材料协同工作、提升加固效果的关键工序。传统的人工打磨方式存在劳动强度大、作业环境恶劣、效率低、质量一致性难以保证

\*上海市住房和城乡建设管理委员会 2024 年度科研项目(沪建科 2024-002-007)

[作者简介] 季元吉, 工程师, E-mail: jiyuanji@sribs.com

[收稿日期] 2025-10-28

等问题,易引发安全事故。与此同时,全球制造业与建筑业发展促使智能化、自动化成为行业核心驱动力,机器人技术尤其是打磨机器人在多领域取得显著进展。因此,采用机器人技术实现混凝土表面打磨自动化与智能化,已成为行业发展的必然趋势。

机器人自动化打磨的核心在于其视觉系统,高效、精确的视觉系统能够快速识别并定位打磨区域,进而为机器人打磨路径规划与磨削参数自适应控制提供决策依据。然而,既有建筑混凝土表面状况复杂多变,受表面不平整、识别点被遮挡、表面存在干扰识别点及光线亮度变化等因素影响,尤其是在打磨作业过程中能见度较低的情况下,对视觉系统识别精度、鲁棒性和实时性提出了严峻挑战。

因此,本研究聚焦于建筑打磨机器人视觉系统,旨在通过构建面向实际工程应用的机器人视觉感知系统,以提升混凝土表面打磨的自动化水平、作业质量与整体效率,为推动既有建筑加固改造领域智能化升级提供技术支撑。

## 1 建筑打磨机器人视觉模块设计

### 1.1 视觉模块构成与工作原理

建筑打磨机器人系统中的视觉模块是自主识别打磨区域、辅助决策控制的核心,主要通过自动检测识别施工墙面图像标记点,为打磨路径规划和任务执行提供依据。该模块基于深度学习设计,融合目标检测算法与图像后处理策略,实现标记点精准提取和区域有效性判断,由图像采集设备、目标检测模型、图像后处理模块构成。

#### 1) 图像采集设备

机器人搭载的摄像头实时采集墙面图像,考虑到打磨现场高粉尘、遮挡、光照变化等干扰,需选用高分辨率、抗干扰能力强的摄像头。

#### 2) 目标检测模型

系统运用 YOLOv8 目标检测算法完成图像标记点检测,这种算法模型架构精简、运行速度快,在边缘设备计算资源受限的情况下也可稳定部署,可精准输出各标记点位置、类别及置信度数据,为后续图像分析判断筑牢数据基础。

#### 3) 图像后处理模块

图像后处理模块通过分析目标检测结果判断图像有效性。当图像中同类标记点数量 $\geq 4$ ,且置信度均高于设定阈值时,认定为有效图像,否则为无效图像。图像识别工作流程:机器人通过摄像头采集墙面图像→将图像输入检测模型,提取目标类

别、位置与置信度→图像后处理模块对结果进行筛选与计数→根据标记点数量与类别将图像分类→输出分类结果,供打磨机器人使用。

### 1.2 目标检测算法设计

YOLO 算法是基于深度学习的目标检测方法,核心是将图像划分为多个网格,由各网格预测区域内边界框和目标类别,凭借较快检测速度和较高检测精度在目标检测任务中广泛应用,尤其适合实时性要求高的工业场景。本研究选用 YOLO 算法,因其在多目标检测中表现优异,可在复杂背景下精准识别图像中多个打磨标记点,输出类别和位置信息,助力打磨机器人判定有效打磨区域,提高作业精度。Ultralytics 官方发布的 YOLOv8 是企业认可的 YOLO 系列算法较新稳定开源版本,整体结构由输入层、骨干网络、颈部网络和检测头组成,其中输入层主要功能为进行图像尺寸调整、归一化及数据增强,提升训练效果;骨干网络主要功能为提取图像深层次特征,采用 C2f 模块提高特征表达能力;颈部网络主要功能为通过 FPN 和 PAN 结构融合不同尺度的特征信息;检测头主要功能为输出检测框位置、类别及置信度。YOLOv8 相比早期版本在精度、速度与整体效率方面大幅度提升,其特征提取能力强、实时性高,在保持轻量化的同时提升了检测精度与推理速度,且工程接口完善,尤其适合本系统中对实时性和高效性要求较高的场景。YOLO 模型输出检测目标类别和置信度,后处理阶段据此判断图像中是否有符合要求的有效打磨区域,使打磨机器人依据准确图像信息高效决策,保障打磨作业可靠性与质量。

本系统采用 Python+PyTorch 作为开发平台,开发语言为 Python3.11,深度学习框架为 PyTorch,目标检测算法为 YOLOv8,操作系统为 Windows10。根据系统需求,视觉模块需实现以下输出目标:①准确检测图像中所有存在的标记点;②基于检测结果判断图像中标记点类别及数量,用于判定图像是否为有效图像,有效图像指包含 4 个相同类别标记点的图像。

### 1.3 数据采集与标注

数据采集、标注、增强及数据集划分方法对于本研究而言较关键。为确保视觉模块目标检测算法训练效果,需高质量、具有代表性的图像数据。本研究应用工业相机采集图像(见图 1),其拍摄灵活、分辨率高且便于操作,能够快速适配施工现场多样环境,采集时注重保持画面清晰稳定,防止模糊、抖动影响数据质量。



图 1 构建数据集的原始图像

Fig. 1 Original image for constructing the dataset

为提升模型的泛化能力,采集的数据覆盖了多种环境变化条件,图像拍摄光照变化条件包括日光、阴天、人工灯光等,角度变化条件包括正面、斜侧面、俯视、仰视等,干扰因素变化条件包括引入背景杂物、遮挡等。

数据标注使用 Roboflow 平台(计算机视觉开发平台)中的 Roboflow Annotate 工具(数据标注工具),用矩形框标注,并按 YOLO 格式保存标签信息,包括目标类别(有效标记点/无效标记点)、标注框中心点  $x$  坐标(归一化)、标注框中心点  $y$  坐标(归一化)、标注框宽度(归一化)、标注框高度(归一化)。标注质量直接影响模型训练效果,标注时严格遵循统一标准,并经人工复查确保标注质量。为提升模型鲁棒性,对原始数据进行图像旋转(随机角度旋转)、亮度调整(模拟不同光照变化)、图像扭曲(轻微变形)、加噪声(模拟高粉尘浓度环境)等多种增强处理,处理完成后最终将数据集按 8:1:1 划分为训练集、验证集和测试集。

#### 1.4 打磨目标识别(后处理)算法设计

打磨目标识别算法核心是对输入图像进行目标检测,精准找出标记点,分析其类别与数量,判断有无 4 个相同类别标记点,以确定图像是否为有效打磨区域。整体识别分目标检测和后处理阶段,目标检测模块工作流程:数据输入(将采集的图像输入至 YOLOv8 模型)→算法检测(YOLOv8 模型执行推理,输出图像中所有检测到的目标位置信息和类别标签)→检测结果输出(输出每个目标的边界框坐标和类别)。

YOLOv8 检测模型兼具高精度与高速度,可同时定位多个标记点并分类,提升检测准确率与实时性。初步目标检测后,图像后处理模块筛选结果,确定打磨有效区域。图像后处理模块工作流程:检测结果接收(接收 YOLOv8 输出的所有检测目标边界框及类别)→目标统计(统计每个类别对应的检测框数量)→有效性判断(判断是否存在某一类别对应检测框数量为 4)→输出分类结果(若存在分类为有效,否则分类为无效)。

## 2 基于图像视觉的打磨识别算法优化

### 2.1 算法优化路径

为提升算法效果,从数据集和训练参数方面进行优化,使训练集更接近真实场景,帮助算法在各种条件下均可精准识别打磨区域。

首先通过多种数据增强手段增强数据集丰富度、真实性及适应性,包括常规增强、引入噪声和亮度扰动。常规增强指对图像进行翻转、旋转,避免算法仅适应特定角度,提高通用性。引入噪声指在图像中加入类似粉尘的高斯噪声,增强算法在模糊、杂乱环境下的识别能力。亮度扰动指随机改变图像亮度,模拟室内外不同光照条件,减少光线对识别的干扰。

此外,在训练时通过多次试验调整关键参数,以实现训练参数优化,包括调整学习率和使用早停策略。通过尝试不同的学习率,提升训练稳定性与收敛速度。使用早停策略防止模型过拟合,提高泛化能力。

### 2.2 模型训练与算法性能提升

为评估打磨目标识别算法性能,本研究构建了 5 组标注数据集,探究图像采集环境、数据增强策略、模型参数设置对识别精度的影响。

#### 2.2.1 基本模型构建

第 1 次训练集图像均由工业相机拍摄,标注对象是图像中 4 个形状大小一致的标记点,每张图像含 4 个有效目标,部分图像加入颜色差异或存在遮挡的手部、杂物等干扰物,模拟真实打磨环境下的识别干扰情况(见图 2)。

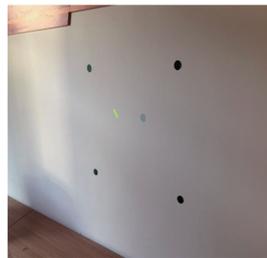


图 2 第 1 次训练集图像

Fig. 2 The first training dataset image

为提升模型鲁棒性,本次训练采用 Roboflow 平台对数据进行了自动预处理与增强。自动预处理时自动调整图像朝向,并统一输入尺寸(640 像素×640 像素)。数据增强包括随机水平翻转、轻微旋转变换(角度 $-10^{\circ} \sim 10^{\circ}$ )、模拟模糊情况(采用标准差 $\leq 0.7$  像素的模糊核对图像做卷积运算,施加不会破坏墙面标记点的核心视觉特征轻微模糊)。本次训练所用数据集称为数据集①,共包含 188 张图

像,训练集、验证集、测试集分别包含 150, 19, 19 张图像。

本次训练使用轻量化模型 YOLOv8n (该模型包含约 300 万个可训练参数),训练轮数为 100 (未设置早停),学习率 (每次迭代中模型参数更新幅度) 默认为 0.01,输入图像尺寸为 640 像素×640 像素 (与数据增强保持一致),每次迭代样本数为 16,使用 Ultralytics 默认优化器 (SGD)。

完成训练后,分析验证阶段各项指标变化,包括定位误差、分类误差、边界框回归损失值、精确率、召回率、平均精度 (包括交并比 IoU 阈值为 0.5 及 0.5~0.95 的情况),从多个维度评估模型在目标检测任务中的表现,如图 3 所示。基于训练结果,可初步判断模型存在以下主要问题。

### 1) 类别识别受样本分布影响,验证损失局部不稳定

由于第 1 次训练验证集样本数量少,训练集与验证集随机划分易造成特征分布差异大,使验证损失不稳定、检测指标波动大,同时数据集数据增强方式简单,现有的增强策略无法模拟亮度变化、噪声等真实干扰情况。

### 2) 模型过拟合问题明显

模型在训练集上学习效果良好,但在验证数据方面泛化性能不足,存在过拟合。这可能是由于数据集规模小、图像分布集中,使模型对训练集记忆性强却泛化能力弱。

### 3) 边界框预测稳定性不高,训练精度仍受约束

采用默认初始学习率 0.01 时模型参数更新强烈,损失值多轮震荡起伏且下降趋势不明显,边界框回归精度提升缓慢。

#### 2.2.2 引入噪声扰动与早停

针对第 1 次训练的问题,第 2 次训练对配置进行了优化,采用参数量更大的 YOLOv8s 模型,该模型包含约 500 万个可训练参数。对于模型过拟合倾向问题,引入早停机制 (容忍度参数设定为 20),自动终止性能无提升的训练,并采用更强的数据增强策略,以提升模型在粉尘干扰等复杂场景下的鲁棒性。对于学习率过高问题,将初始学习率下调至 0.005,使模型训练更平稳,有助于提升平均精度 (IoU=0.5~0.95) 等对定位精度要求更高的评估指标。对于验证集代表性不足问题,重新划分训练集与验证集,增加验证样本量,确保数据分布一致,以提升评估结果的代表性与稳定性,并引入亮度扰动与噪声扰动等增强策略,提升模型泛化能力。

第 2 次训练集以首次训练集为基础,新增更多

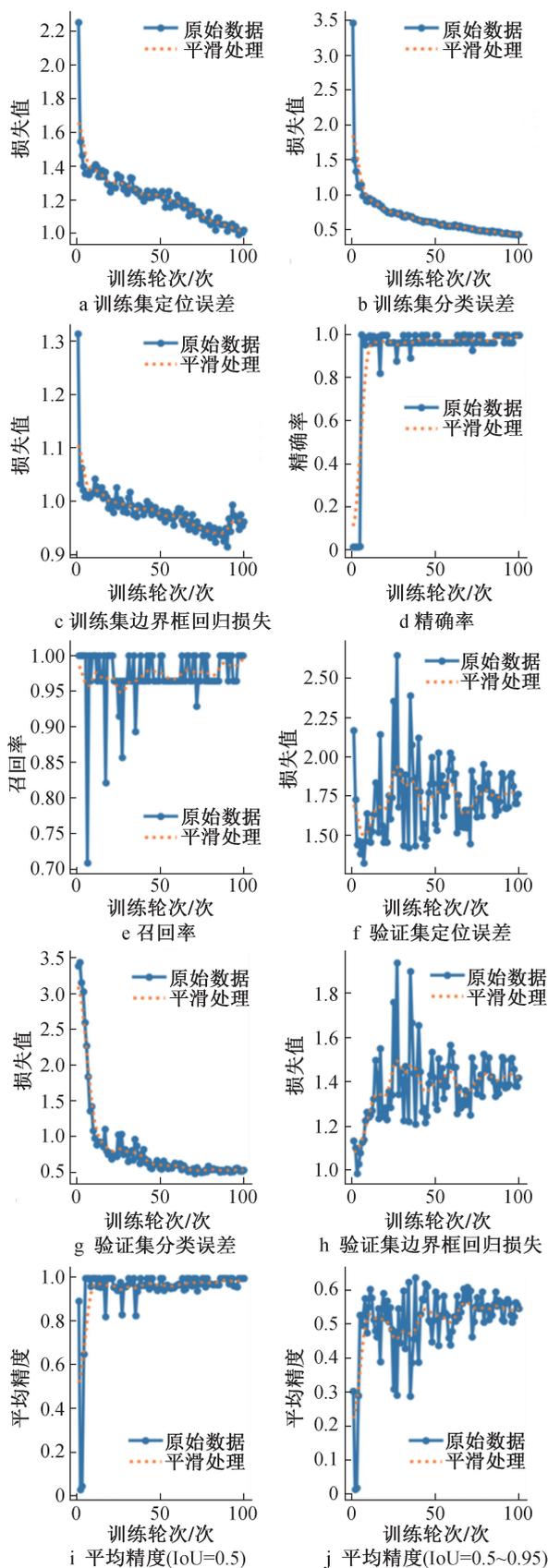


图 3 第 1 次训练结果

Fig. 3 The first training results

实际测试中常见干扰因素 (见图 4),将训练集数据

量进行了扩充。所有图像仍由工业相机拍摄,每张含 4 个形状大小相同的标记点,部分图像加入不同颜色或遮挡的粉尘、模糊背景等干扰物,模拟更复杂打磨现场环境,用于测试模型在高难度条件下的性能表现。

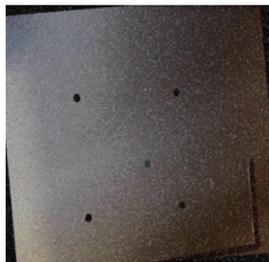


图 4 第 2 次训练集图像

Fig. 4 The second training dataset image

为提升模型在干扰环境下的鲁棒性,本次训练在 Roboflow 平台对数据集进行了进一步的自动预处理与增强。自动预处理时自动调整图像朝向,统一输入尺寸(640 像素×640 像素)。数据增强包括随机水平翻转、随机垂直翻转、随机旋转(角度 $-10^{\circ}\sim 10^{\circ}$ )、模拟模糊情况(采用标准差 $\leq 0.7$  像素的模糊核对图像做卷积运算)、调整图像亮度(随机降低 $0\%\sim 25\%$ 的图像亮度)和模拟高浓度粉尘环境(在训练图像上随机添加粉尘的视觉效果,且粉尘覆盖的像素点占整张图像总像素的比例 $\leq 5\%$ )。

最终数据集由第 1 次训练所用数据集①与本次训练数据集合并得到,称为数据集②,共包含 302 张图像。本次训练轮数增至 150,每次迭代样本数同样为 16(受限于 CPU 计算资源,与第 1 次训练保持一致),同样使用 Ultralytics 默认优化器(SGD)。

第 2 次训练结果如图 5 所示,综合各指标变化趋势,可知第 2 次训练仍存在以下问题。

#### 1) 类别损失异常波动,训练过程存在异常点

验证集分类损失值在第 12 轮出现极端峰值,推测原因是某轮验证集中类别标签错配或特定类别图像占比过大。

#### 2) 目标边界回归稳定性有待提升

边界框回归损失值整体下降但高频波动,表明模型边界拟合精度易受个别图像噪声影响,应在后续引入更精细的特征增强或注意力机制提升稳定性。

#### 3) 检测鲁棒性波动明显,召回能力仍有限

召回率曲线多次骤降,模型在复杂场景下易发生漏检,需在训练集中增加多样化干扰条件,提升模型泛化能力。

相比于第 1 次训练,第 2 次训练多项指标显著

提升,如平均精度(IoU=0.5)和召回率,然而训练过程稳定性及边界框精度仍有待提升。

### 2.2.3 引入负样本与复杂标记

基于工程实际需求,提升模型对真实施工场景多变标记识别能力的目标,第 3~5 次训练从数据标注样式和数据集组成方面进行改进。

#### 1) 标签样式复杂化提升(环境区分度提升)

真实施工现场图像背景存在纹理复杂、颜色相近的墙面或施工痕迹,易干扰标记点检测。改进后的标记点在原有彩色圆片上贴覆白色胶带形成“×”字形图案,增加结构复杂性与纹理特征,助力模型在复杂背景下精准提取目标特征,降低误检率并增强鲁棒性。

#### 2) 不完全有效标记点图像增加(模型通用性增强)

在第 1,2 次训练中,训练样本主要为理想条件下的标准图像,即图像中严格包含 4 个颜色形状一致的有效标记点。而在实际施工过程中,由于机器人拍摄图像时存在标记点遮挡、部分缺失、图像边缘裁切等情况,导致部分图像中仅包含 0/1/2/3 个有效标记点或包含与有效标记点颜色/形状部分相似的干扰物。如果模型仅在标准样本上训练,将难以正确处理此类非理想情况,降低实际应用效果。因此,从第 3 次训练起新增了大量不完全有效图像样本,包括但不限于以下几类(见图 6):①图像中仅包含 3 个颜色形状均相同的有效标记点;②图像中出现 3 个有效标记点和 1 个颜色不同的相似干扰点;③图像中仅包含 1 个或 2 个有效标记点;④图像中完全无有效标记点,甚至仅包含干扰目标。

第 3 次训练使用的数据集③由 2 个小规模数据集整理合并得到。2 个数据集均在不同程度上引入干扰因素,以提升模型泛化能力。图像全部由工业相机拍摄,并统一尺寸为 640 像素×640 像素。为提升模型在复杂环境下的鲁棒性,本次训练同样依托 Roboflow 平台完成了图像自动预处理与数据增强。数据集③共包含 493 张图像,均采自复杂背景墙面,部分图像存在照度不足、纹理干扰强烈、边缘遮挡等因素(见图 7)。

第 3 次训练沿用 YOLOv8s 模型开展目标检测,运行环境为 Python 3.11,PyTorch 2.6.0,训练代码与第 2 次训练相同。本次训练依旧使用 Ultralytics 默认优化器(SGD),对超参数进行了合理设置,以平衡训练时长与模型性能。本次训练轮数设定为 100,每次迭代样本数同样为 16,初始学习率为 0.005,容忍度参数为 20,以提高训练稳定性,防止

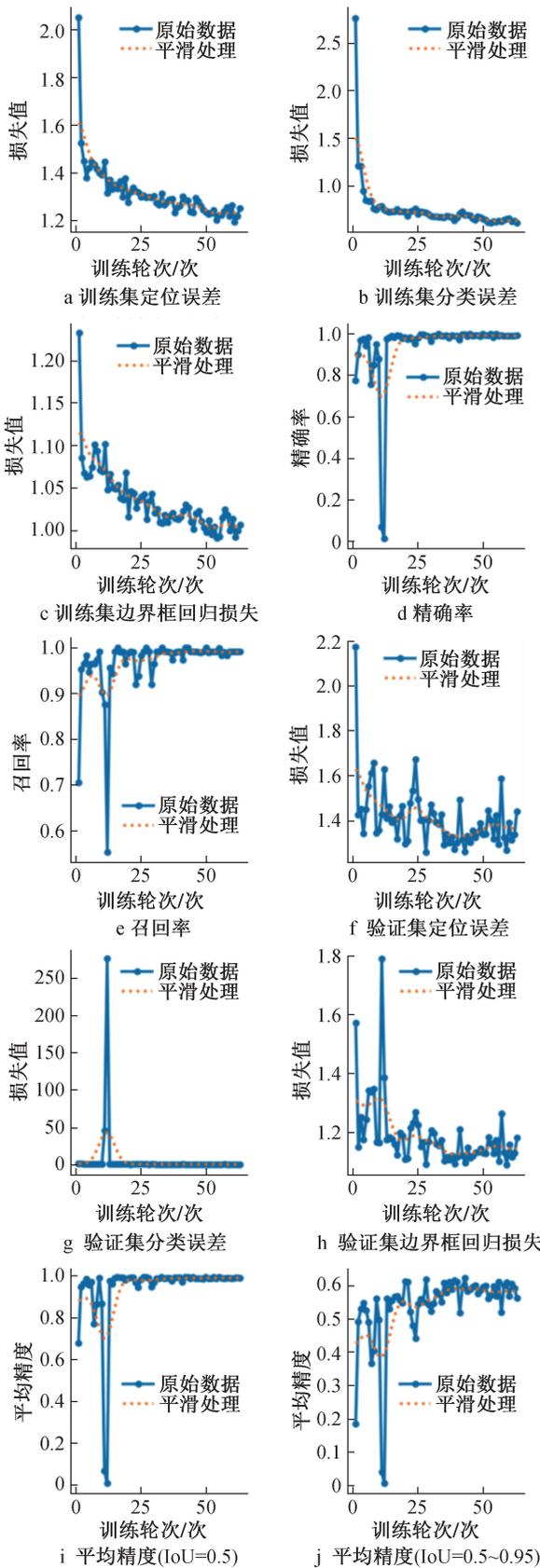


图5 第2次训练结果

Fig. 5 The second training results

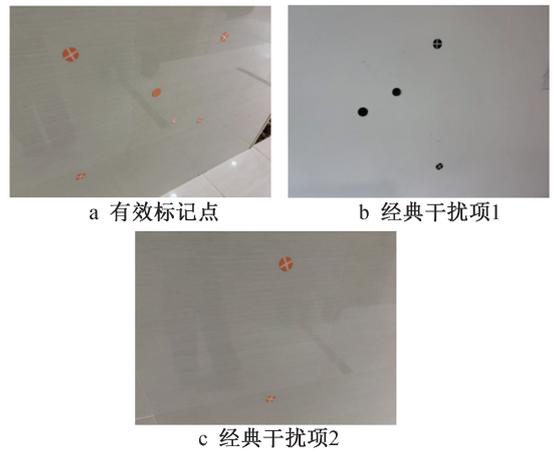


图6 第3次训练原始图像

Fig. 6 Original images of the third training

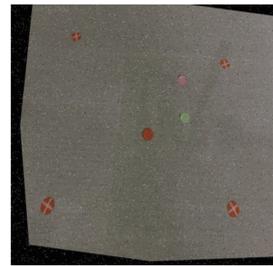


图7 第3次训练集图像

Fig. 7 The third training dataset image

第3次训练结果如图8所示,结合指标趋势变化分析,可知本次训练存在以下问题。

1) 训练初期不稳定性强

前10~15轮多项指标波动大,验证集分类误差损失值和精确率数据出现极端值,这表明模型未充分收敛时受数据扰动或梯度问题干扰,需优化初始参数与正负样本比例。

2) 类别判别误差突出但逐步修正

验证集分类误差损失值极高峰值可能由于标签偏移、类别不平衡、异常样本引起,其后期快速回落说明模型可自我修正。

3) 边界框精度有待优化

验证集边界框回归损失值和平均精度(IoU=0.5~0.95)指标提升不明显,反映模型高精度目标检测时边界回归不足,可通过提高边界标签精度、调整回归损失函数权重加以优化。

4) 模型泛化能力较强但鲁棒性略显不足

最终平均精度(IoU=0.5)、精确率、召回率等检测精度达较高水平,表明模型在标准环境下泛化能力良好,不过指标波动显示其在扰动条件下稳定性仍待加强。

过拟合。

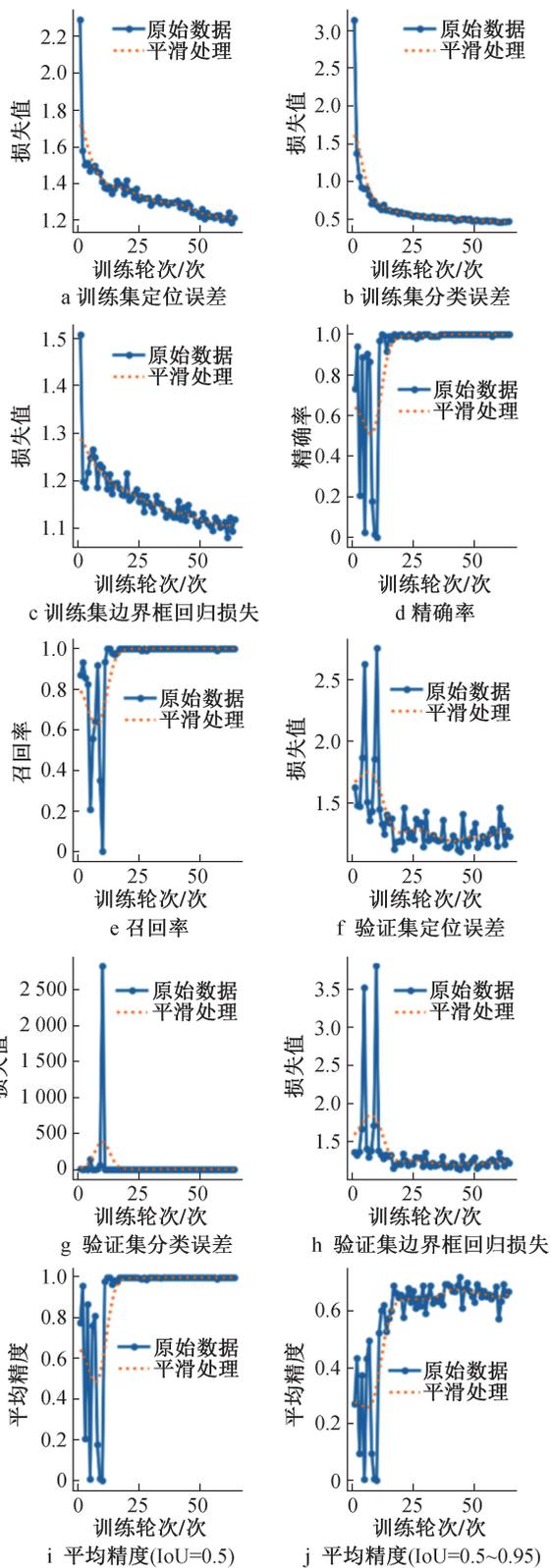


图 8 第 3 次训练结果  
Fig. 8 The third training results

2.2.4 多维干扰下的鲁棒性强化

为测试模型在恶劣环境下的稳定性和高强度数据增强对检测结果的影响,第 4 次训练仅使用数

据集④,其包含 319 张图像,由于本次数据增强方案特殊、操作强度高,为避免不同数据间相互干扰,未混合其他数据集(见图 9)。



图 9 第 4 次训练集图像  
Fig. 9 The fourth training dataset image

数据集④中非标准样本占比明显提高,图像中的标记点普遍存在数量不足、形状缺失、被遮挡、颜色混淆等复杂情况。这些设计并非偶然,而是刻意模拟真实施工现场中可能遇到的异常情况,旨在提升模型对各类不确定因素的适应能力,使其具备更强的鲁棒性和泛化能力。

与前几次训练一致,数据预处理时先统一尺寸拉伸,再叠加多种增强方式,包括亮度扰动、随机遮挡、颜色抖动、高斯噪声干扰等,以提高模型对异常样本的学习能力和稳定性表现。本次训练每张原图扩充为 3 张增强图像(对训练集中每张原始标注样本,通过随机数据增强操作,自动生成并输出 3 张不同的增强图像,将其用于模型训练),随机旋转角度为 $-10^{\circ} \sim 10^{\circ}$ ,进行 $\pm 5^{\circ}$ 水平及竖直剪切,随机降低或增加 15%亮度,模拟相机因环境波动产生的轻微曝光偏差,对训练图像曝光度在 $-5\%$ (轻微欠曝) $\sim 5\%$ (轻微过曝)随机调整,采用标准差 $\leq 0.1$ 像素的模糊核对图像做卷积运算,随机添加图像噪声像素,其占整张图像总像素的比例 $\leq 3.03\%$ 。

第 4 次训练沿用轻量化 YOLOv8s 模型,维持与第 3 次相同的训练结构和超参数设置,以保证结果可比性。训练结果如图 10 所示。综合来看,第 4 次训练仍存在以下关键问题。

1) 类别损失在前期出现极端异常,训练初期不稳定

验证集分类误差损失值前期超高峰值显著偏离其他轮次,可能由于个别验证图像标签缺失、错误或单类别图像占比失衡,需检查验证集标签完整性并控制样本分布。

2) 目标框位置损失收敛较弱,边界拟合精度仍有提升空间

验证集定位误差损失值波动大且未收敛,表明

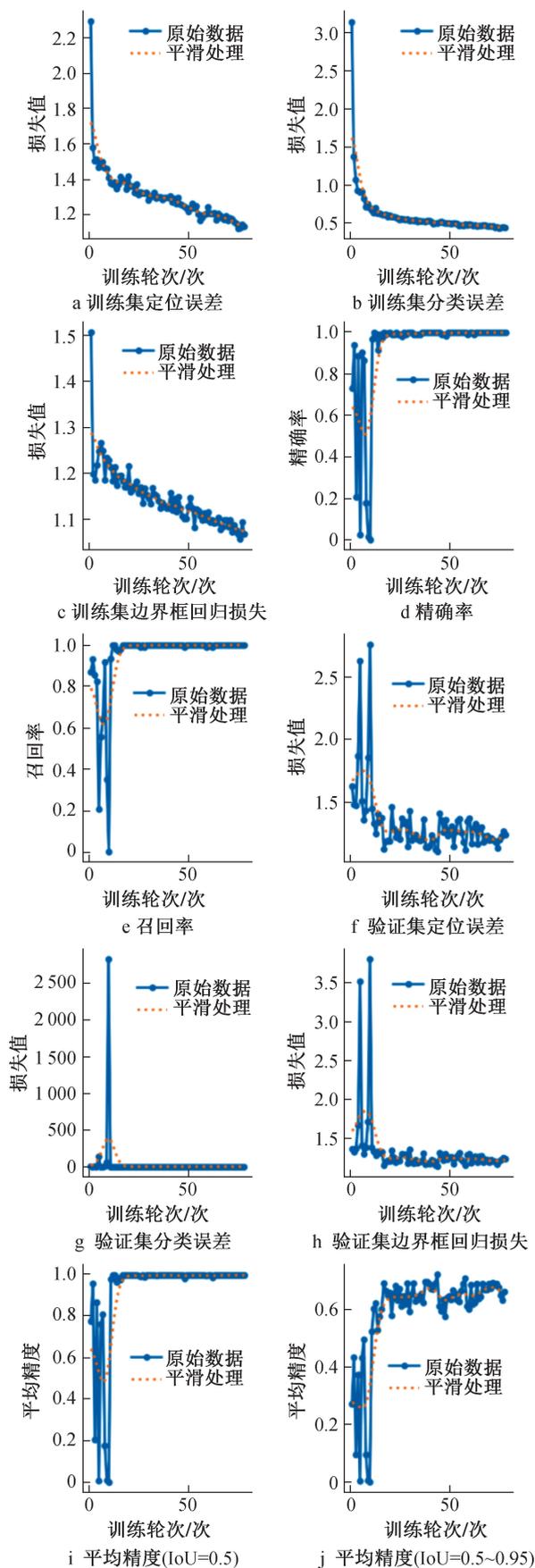


图 10 第 4 次训练结果

Fig. 10 The fourth training results

模型实际目标定位稳定性欠佳或因图像分辨率、标记点位置差异小等因素制约,应采用更细粒度特征提取方法优化。

3) 召回能力稳定性有待加强,仍存在轻微漏检风险

召回率虽整体较高,但局部仍有下滑,推测模型在标记点数量不足或干扰背景下的判别力有限,应适当提升训练数据中的非完整图像占比,以增强模型在边缘条件下的判断能力。

### 2.2.5 大规模样本扩展

为提升模型综合泛化能力,第 5 次训练采用前 4 次训练所有数据集整合而成的数据集⑤,该数据集共有 976 张图像,包含多种标准与非标准背景,模拟复杂施工场景及不同干扰强度场景,较之前训练更具代表性与多样性。在对前 4 次训练中发生的问题进行优化后,本次训练数据增强策略及超参数配置沿用第 4 次训练标准流程,保证横向比较一致性与代表性。

第 5 次训练结果如图 11 所示。本次训练中关键指标表现良好,模型精度与稳定性趋于成熟,达成预期目标。个别问题实际可控,继续训练收益有限,当前模型性能已较可靠实用。

1) 目标框收敛已基本完成,定位精度稳中有进,定位误差损失值轻微波动下显著收敛,模型对复杂样本处理良好,基本满足部署要求。

2) 召回能力已达应用要求,鲁棒性可接受

召回率整体水平较高,对非完整图像的识别能力已足够支撑应用。

3) 检测精度持续提升,综合性能基本满足任务需求

平均精度 ( $IoU=0.5$  及  $IoU=0.5\sim 0.95$ ) 呈明显上升态势,高精度标准下可精准拟合目标边界,为打磨任务中目标区域判定提供有力支撑。

通过第 5 次训练,模型获得了在多样化条件下的综合训练机会,为后续模型评估和真实施工部署提供了更稳健的基础。

### 2.2.6 图像后处理分析

在完成目标检测模型训练的基础上,本研究设计并实现图像后处理算法,对测试集图像进行有效性分类,通过统计每幅图像中置信度高于阈值的标记点数量,依据 4 个有效标记点的判断标准自动划分有效图像与无效图像。

利用 Python 编写图像后处理脚本,调用 Ultralytics YOLOv8 模型推理接口及训练完成的权重文件,逐张检测测试图像。处理流程:①加载

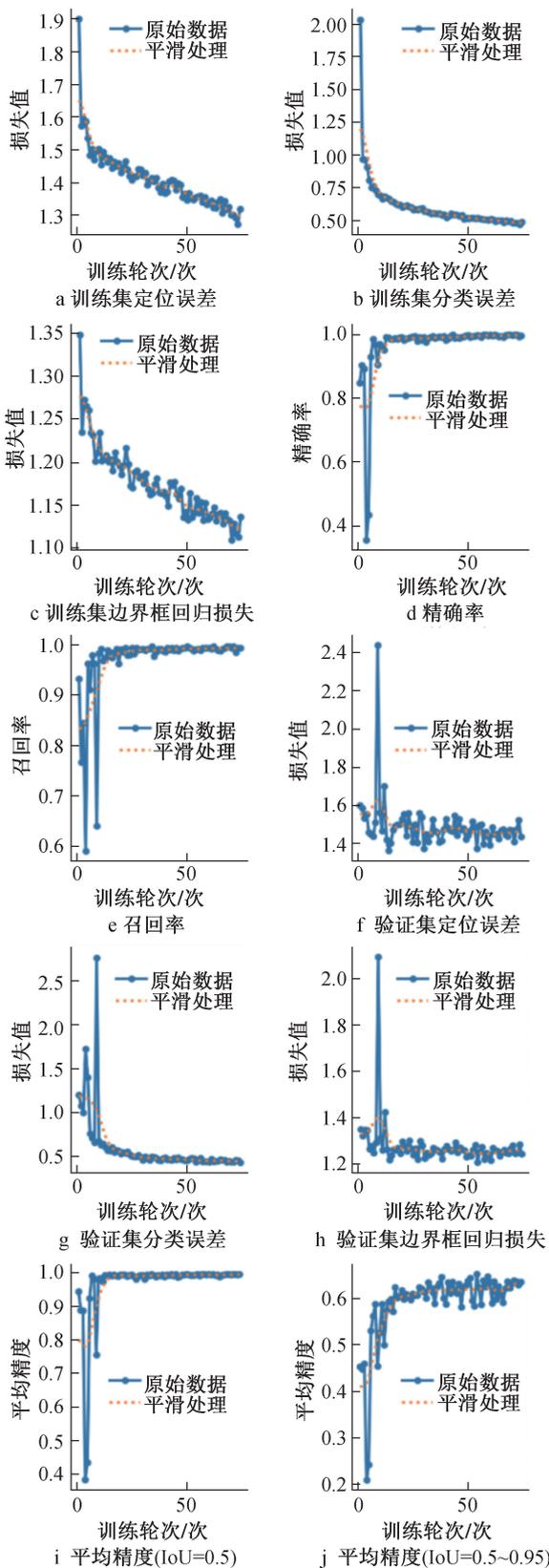


图 11 第 5 次训练结果

Fig. 11 The fifth training results

YOLOv8 训练所得的最佳模型权重;②依次读取测试集图像,进行推理并提取检测结果;③设定置信

度阈值(设为 0.5),筛选可靠目标;④如果图像中识别出的有效标记点数量为 4,则判定为有效图像,反之为无效;⑤根据判断结果将图像分别存入有效/无效图像文件夹中;⑥输出处理总耗时。

该算法自动化与处理效率较高,第 5 次训练测试结果表明,图像后处理程序可准确识别并分类大部分图像,即使存在标记点部分缺失或视角受限问题,只要检测到 4 个有效目标仍可正确判定为有效图像,展现出良好的实用性。

此外,在测试中发现适配性问题,即第 4,5 次训练时图像后处理程序将早期无图案圆片标记点图像全部判定为无效。经分析,这是因为目标检测模型对初版无特征标记点识别能力弱,导致标记点识别数量不足 4 个而被归类为无效。

图像后处理方案在识别定位新版带白色“x”图案圆片标记点数据上表现稳定、分类准确,可支撑打磨区域筛选。但对初版标记点图像分类准确性明显下降,说明检测模型对标记点样式差异具有敏感性。

### 3 结论

本研究围绕既有建筑加固改造领域智能化升级需求,聚焦于建筑打磨机器人视觉感知系统,开展一系列深入的研究与试验。针对混凝土表面打磨作业中环境复杂、传统识别方法精度与鲁棒性不足的挑战,本研究设计并实现了基于 YOLOv8 目标检测算法的机器人视觉模块。该模块通过图像采集、目标检测与图像后处理的协同工作,能够精准识别墙面标记点,并有效判断打磨区域,为机器人自主路径规划与作业控制提供了可靠的感知基础。

通过 5 次迭代训练与优化,本研究系统探讨了数据集构建策略、数据增强方法及模型训练参数对算法性能的影响。从最初使用小规模标准数据集构建基础模型到逐步引入噪声扰动、亮度变化、早停机制,再到纳入包含不完全有效标记点及复杂干扰的负样本,最终整合形成大规模、高多样性的综合数据集,模型检测精度、鲁棒性和泛化能力得到了显著且持续提升。试验结果表明,经优化后的 YOLOv8s 模型在复杂多变的模拟施工环境下,能够稳定实现标记点精准定位与分类,其精度和召回率等关键指标均达到了较高水平,基本满足了实际工程应用对于实时性、准确性和稳定性的要求。

然而,本研究仍存在一定局限性,尽管通过数据增强模拟了多种干扰,但与真实打磨现场高浓度的粉尘、强烈光照变化及复杂结构遮挡相比,实验室环境下的测试条件仍有一定差距,这有待后续实

际工程进一步应用与检验。

### 参考文献:

- [ 1 ] 陈翀,李星,邱志强,等. 建筑施工机器人研究进展[J]. 建筑科学与工程学报,2022,39(4):58-70.  
CHEN C,LI X,QIU Z Q,et al. Research progress of construction robots[J]. Journal of architecture and civil engineering,2022,39(4):58-70.
- [ 2 ] 吴庆杰,张红伟,陈少林,等. 基于计算机视觉的盾构隧道管片错台自动测量方法[J]. 施工技术(中英文),2025,54(17):40-44.  
WU Q J,ZHANG H W,CHEN S L,et al. Automatic measurement method of segmental dislocation for shield tunnels based on computer vision [J]. Construction technology, 2025, 54 ( 17 ) : 40-44.
- [ 3 ] LIU Z H,LIU Q,XU W J,et al. Robot learning towards smart robotic manufacturing: A review [J]. Robotics and computer-integrated manufacturing,2022,77:102360.
- [ 4 ] ZHOU Y,LI X,YUE L Z,et al. Global vision-based impedance control for robotic wall polishing [C]//2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS),2019.
- [ 5 ] 蔡志勇,吕金彪,赵骥洲,等. 基于计算机视觉的钢筋加工智能感知方法研究[J]. 施工技术(中英文),2025,54(5):39-45.  
CAI Z Y,LÜ J B,ZHAO J Z,et al. Research on intelligent perception method for steel rebar processing based on computer vision[J]. Construction technology,2025,54(5):39-45.
- [ 6 ] YANG Z J,CHEN F,ZHAO J,et al. A novel vision localization method of automated micro-polishing robot[J]. Journal of bionic engineering,2009,6(1):46-54.
- [ 7 ] 王飞. 粗糙金属表面非导电覆盖层厚度非接触测量技术[J]. 仪表技术与传感器,2025(2):96-99,108.  
WANG F. Non-contact measurement technology for non-conductive coating thickness on rough metal surfaces [J]. Instrument technique and sensor,2025(2):96-99,108.
- [ 8 ] 于海洋,李海生,彭伟,等. 基于机器视觉的预制叠合板智能检测关键技术[J]. 施工技术(中英文),2024,53(20):36-41.  
YU H Y,LI H S,PENG W,et al. Key technology of intelligent detection for precast composite panels based on machine vision [J]. Construction technology,2024,53(20):36-41.
- [ 9 ] 张岩,赵玉涛,李光旭,等. 图像的边缘检测技术研究进展[J]. 计算机与数字工程,2020,48(5):1176-1184,1232.  
ZHANG Y,ZHAO Y T,LI G X,et al. Research progress on edge detection technology of images [J]. Computer & digital engineering,2020,48(5):1176-1184,1232.
- [ 10 ] 翟浩博,任宝双,陈洪敏,等. 房建施工机器人的应用及展望[J]. 施工技术(中英文),2023,52(23):20-26.  
ZHAI H B,REN B S,CHEN H M,et al. Application and prospect of building construction robot[J]. Construction technology,2023,52(23):20-26.
- [ 11 ] 杨萍,赵丹,李鹤岐,等. 专用焊接打磨机器人视觉系统的研究[J]. 兰州理工大学学报,2004,30(3):41-44.  
YANG P,ZHAO D,LI H Q,et al. Investigation of visual system on special-purpose welding-polishing robot [J]. Journal of Lanzhou University of Technology,2004,30(3):41-44.
- [ 12 ] 张兰尧,陈晓玲,张达敏,等. ValidFlow:基于标准化流的无监督图像缺陷检测[J]. 数据采集与处理,2023,38(6):1445-1457.  
ZHANG L Y,CHEN X L,ZHANG D M,et al. ValidFlow: Unsupervised image defect detection based on normalizing flows [J]. Journal of data acquisition and processing,2023,38(6):1445-1457.
- [ 13 ] 李盼,温卫军,周家豪,等. 基于地面找平机器人的找平机构调平性能研究[J]. 施工技术(中英文),2025,54(13):12-17.  
LI P,WEN W J,ZHOU J H,et al. Research on leveling performance of leveling mechanism based on ground leveling robot [J]. Construction technology,2025,54(13):12-17.
- [ 14 ] 王洋,施保昌. 水面常见目标边缘检测与识别方法[J]. 船舶,2025,36(2):60-66.  
WANG Y,SHI B C. Edge detection and recognition methods for common water surface targets [J]. Ship & boat,2025,36(2):60-66.
- [ 15 ] 张振海,季坤,党建武. 基于机器视觉的隧道裂缝检测方法研究[J]. 重庆大学学报,2024,47(12):83-91.  
ZHANG Z H,JI K,DANG J W. Crack detection method for tunnels based on machine vision [J]. Journal of Chongqing University,2024,47(12):83-91.