基于计算机视觉与深度学习技术的汽车白车身工艺智能检测系统

王家琦 石阳 安进 李佳鑫 王思澄

（北京奔驰汽车有限公司，北京 100176）

摘要：为解决传统汽车白车身连接工艺质量控制中人工目检效率低、漏检率高的问题，构建基于计算机视觉与深度学习技术的智能检测系统。通过搭建工业相机成像系统并优化光学参数，采用交互式图像分割自动掩码技术完成图像预处理，结合端到端目标检测模型（DETR）实现目标检测，搭建涵盖采集、处理、存储与可视化的全栈数据管理架构。试验表明，系统检测准确率稳定在99%以上，任务完成时间缩短80%，漏检率降低67%，检测效率与结果一致性大幅提升，可满足白车身批量检测需求。

关键词：白车身 连接工艺 计算机视觉 深度学习 目标检测

中图分类号：U465.5；TP391.4；TP181 **文献标志码：**B **DOI:** 10.19710/J.cnki.1003-8817.20250288

Intelligent Detection System for Automotive White Body Process Based on Computer Vision and Deep Learning Technologies

Wang Jiaqi, Shi Yang, An Jin, Li Jiaxin, Wang Sicheng

(Beijing Benz Automotive Co., Ltd., Beijing 100176)

Abstract: To address the challenges of low efficiency, high miss rate, and insufficient imaging stability associated with manual visual inspection in traditional automotive Body-In-White (BIW) joining processes, an intelligent inspection system based on computer vision and deep learning technologies is developed. An industrial camera imaging system is constructed with optimized optical parameters, and image preprocessing is realized using automatic masking techniques based on interactive image segmentation. Object detection is performed by incorporating an end-to-end object detection model (DETR). A full-stack data management architecture encompassing data acquisition, processing, storage, and visualization is established. Experimental results demonstrate that the system achieves a detection accuracy of over 99%, reduces task completion time by 80%, and lowers the miss rate by 67%. Both inspection efficiency and result consistency are significantly improved, satisfying the requirements for batch inspection of BIWs.

Key words: Body-In-White (BIW), Joining processes, Computer vision, Deep learning, Object detection

1 前言

随着汽车制造业的快速发展，车身制造质量检测成为保证汽车安全性和可靠性的关键环节。白车身生产制造作为汽车制造领域的重要环节，其生产效率与质量控制水平对汽车整体性能和市场竞争力有着深远影响。传统人工目检方法存在效率低下、主观性强、漏检率高等问题，且难以匹配现代化生产线的节拍[1]。视觉检测技术检测精度高、检测速度快，能够对白车身零部件的尺寸、

作者简介：王家琦（1989—），男，工程师，硕士学位，研究方向为人工智能、智能制造，wangjq05@bbac.com.cn。

通信作者：王思澄（1994—），男，工程师，硕士学位，研究方向为计算机视觉、深度学习、目标检测与识别，wangsch01@bbac.com.cn。

参考文献引用格式：

王家琦，石阳，安进，等. 基于计算机视觉与深度学习技术的汽车白车身工艺智能检测系统[J]. 汽车工艺与材料，2026(6)：23-30.

WANG J Q, SHI Y, AN J, et al. Intelligent Detection System for Automotive White Body Process Based on Computer Vision and Deep Learning Technologies[J]. Automobile Technology & Material, 2026(6): 23-30.

缺陷等进行实时、准确检测。通过对大量图像数据进行学习与分析，视觉检测系统能够快速识别不同焊接质量问题，有效提高质量检测的效率与准确性，减少因质量问题导致的生产延误与成本浪费。

计算机视觉技术与深度学习算法的快速发展为工业质检带来了变革，但在汽车制造领域仍面临复杂光照条件下的图像采集精度低、微小缺陷特征提取困难、检测速度与精度难以平衡等问题[2]。本文开发基于计算机视觉与深度学习技术的汽车白车身工艺智能检测系统，以期解决上述问题。

2 人工智能与视觉技术

2.1 计算机视觉与OpenCV

如图1所示，计算机视觉作为人工智能的重要分支，致力于使计算机从图像或视频中获取信息并作出决策。开源计算机视觉库（Open Source Computer Vision Library，OpenCV）提供了丰富的算法支持，包括图像预处理、特征提取、目标检测和相机校准等功能[3]。在汽车工艺检测系统的开发中，OpenCV用于图像去噪、形态学操作，以增强缺陷特征，以及基于交互式图像分割的自动掩码生成，精确提取有效检测区域。尽管传统计算机视觉方法在特定场景下表现良好，但对复杂缺陷类型的处理仍存在局限性，而深度学习技术可有效弥补这一不足[4]。

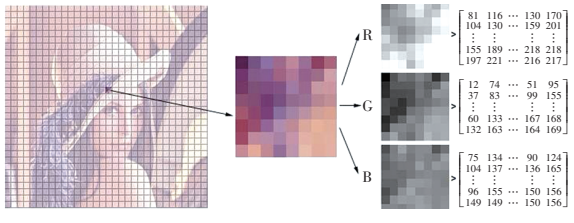


图1 计算机视觉与OpenCV

2.2 深度学习技术

2.2.1 卷积神经网络

卷积神经网络（Convolutional Neural Network，CNN）是深度学习中最成功的架构之一，特别适用于图像处理任务。如图2所示，CNN通过多层卷积运算提取图像的局部特征，通过池化层降低特征维度，最终通过全连接层实现分类或回归。典型的CNN结构包括卷积层、池化层、全连接层以及激活函数[5]。

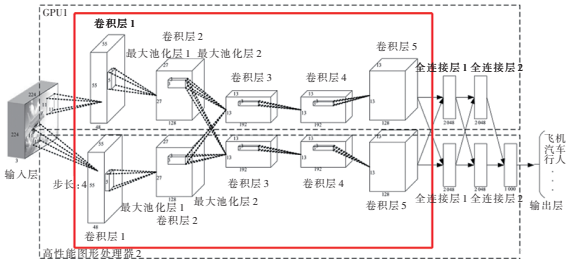


图2 卷积神经网络

2.2.2 目标检测算法

YOLO（You Only Look Once）是一种流行的实时目标检测算法，其最新版本如YOLOv11、YOLOv10、YOLOv8等在精度和速度上均达到业界领先水平。相比于传统的两阶段检测器如R-CNN[6]、Fast R-CNN[7]、Faster R-CNN[8]，YOLO采用单阶段检测策略，将目标检测视为回归问题，显著提高了检测速度[9]。如图3所示，YOLO的核心思想包括网格划分、锚框机制以及多尺度预测[10]。

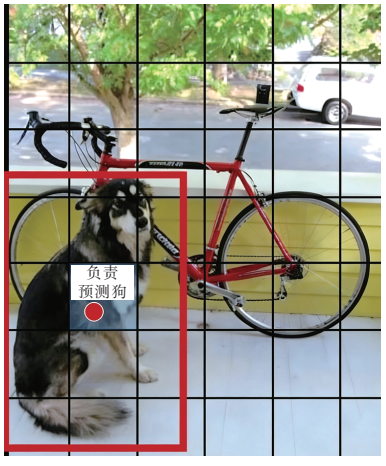


图3 YOLO目标检测

2.2.3 端到端目标检测器

基于Transformer的端到端目标检测器（Detection Transformer，DETR）是一种基于Transformer架构的端到端目标检测方法[11]。如图4所示，DETR将传统目标检测任务建模为一个集合预测问题，依托Transformer架构的全局特征建模能力，可直接完成目标类别判定与边界框回归预测，显著简化了传统目标检测算法的流程与后处理环节，算法核心优势体现在架构简洁性与端到端推理特性[12]。总体而言，DETR为目标检测提供了全新的研究思路，推动了基于Transformer架构的视觉任

务进展,同时在多目标跟踪、实例分割等延伸视觉任务中均展现出广泛的应用潜力[13]。

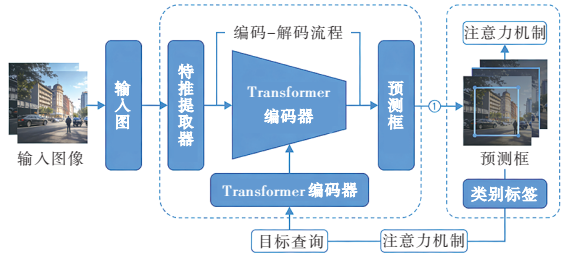


图4　DETR目标检测

3　视觉检测系统开发与应用

3.1　硬件系统设计

3.1.1　高精度成像系统构建

图5所示为本文开发的汽车工艺智能检测系统整体架构。为满足汽车白车身检测的高精度需求,选用海康威视工业相机(600万像素,CMOS卷帘快门,帧率19.1 帧/s,靶面尺寸1/1.8,彩色CS060-10GC接口),搭配多焦距定焦镜头(可选焦距有6 mm、8 mm、12 mm、16 mm、25 mm),可根据白车身不同检测部位的距离与精度要求灵活切换镜头。成像系统的安装支架采用三维不锈钢支架与欧标4040铝型材搭建刚性固定结构,有效抑制检测过程中的振动干扰,确保成像的稳定性与一致性。数据传输与供电环节采用PoE千兆交换机(TP-Link TL-SG2024MP),该设备可实现24路工业相机的同步供电与高速数据传输,将单帧图像采集时间控制在50 ms以内,保障检测的实时性。操控终端选用联想5080MT电脑主机与触摸显示屏,实现硬件设备与检测软件的协同,完成图像采集、参数设置、结果显示等全流程视觉识别操控。

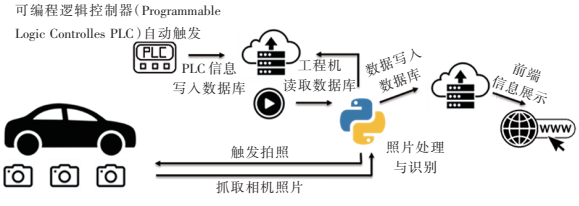


图5　基于计算机视觉与深度学习技术的汽车工艺智能检测系统开发

3.1.2　光学参数优化

结合汽车白车身静止工件的检测特性,为避免自动曝光模式下的成像波动,关闭相机自动曝光功能,采用F2.8~F16可调光圈配合手动曝光补偿模式,实现光学参数的精准调控。通过多组景深对比测试确定最佳光圈F值,由于本系统需对采集图像像素内的白车身连接工艺细节进行精准判别,对成像景深要求较高,因此需尽量增大光圈F值(对应进光量减小),以获得更大的成像景深,确保车身连接部位的上、下细节均能清晰成像。同时,通过合理延长曝光时间来弥补单位时间内进光量的不足,在确保白车身连接工艺细节成像清晰、无模糊、无过曝/欠曝的前提下,有效提升系统拍摄效率,满足汽车白车身批量检测的生产需求。

3.2　智能图像采集与预处理

3.2.1　多相机协同控制

基于海康工业相机机器视觉系统(Machine Vision System, MVS)软件开发工具包(Software Development Kit, SDK),开发Python多线程取流系统。该系统可同时调用多台相机取流,支持16路相机并行触发,通过千兆以太网实现个人计算机(Personal Computer, PC)与多台相机的连接。通过设置cam.MV_CC_GetOneFrameTimeout超时参数(需大于最大曝光时间),避免取流错误;使用MVS软件更改相机设备名称,利用语句camNamehuoqu.SpecialInfo.stGigEInfo.chUserDefinedName输出自定义的设备名称字符并定义取流的名称,使每次拍照采集的图像均按照固定名称保存,确保跨工位数据的一致性;同时启用巨帧模式,以提升数据包传输速度,优化系统的整体取流速度。

3.2.2　基于交互式图像分割的自动掩码技术

在汽车工艺智能检测系统中,工业相机采集的原始图像通常包含大量与检测无关的背景信息。为提高后续视觉检测模型的训练效率和准确率,采用OpenCV设计了一套完整的图像掩码处理流程(图6)。该流程通过智能化的区域选择与裁剪,实现对多摄像机视角图像的批量区域分割与处理,精确提取车身表面的有效检测区域。该技术采用OpenCV库构建,通过人机交互方式定义感兴趣区域(Region of Interest, ROI),生成透明度通道掩码,并自动应用于同视角序列图像,实现高效的非ROI滤除。

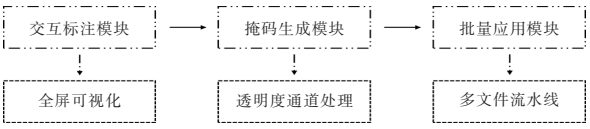


图 6 基于交互式图像分割的自动掩码分层处理架构

基于交互式图像分割的自动掩码技术核心技术路线如下：

交互式 ROI 标注系统采用事件驱动架构实现交互式多边形标注功能；

全屏可视化界面通过 cv2.WINDOW_NORMAL 和 cv2.WND_PROP_FULLSCREEN 参数创建自适应显示环境，优化用户体验；

鼠标事件三重态响应机制采用 EVENT_LBUTTONDOWN 进行初始化绘制点集，采用 EVENT_MOUSEMOVE 进行实时绘制半透明轨迹线（RGB(0,255,0)），采用 EVENT_LBUTTONUP 闭合多边形并填充（cv2.fillPoly）；

操作栈管理支持通过'c'键指令撤销最近操作，维护 rects 点集堆栈；

四通道掩码生成算法采用 RGBA 色彩空间构建透明度掩码，关键技术特征包括：Alpha 通道分离存储，仅保存掩码的透明度通道（单通道 PNG）；掩码持久化，采用无损压缩格式存储，支持后续批量处理调用；模板匹配机制基于 Cam{i}-template.png 生成掩码，确保空间一致性。具体 Python 代码为：

“

```
mask=np.zeros((h,w,4),dtype=np.uint8) # 初始化全透明基底
for roi in rois:
cv2.fillPoly(mask,[roi],(255,255,255,255)) # ROI 区域设为不透明
```

”

批量图像处理可实现多模态图像的统一处理框架：

输入适配层：灰度图像→COLOR_GRAY2BGRA 转换；三通道 BGR→COLOR_BGR2BGRA 扩展；四通道 BGRA→直接处理。

尺寸验证机制具体 Python 代码如下：

“

```
if image.shape[:2] != mask.shape[:2]:
  raise SizeMismatchWarning
```

”

像素级合成运算具体 Python 代码如下：

Alpha 通道替换：image[:,:,3] = mask[:,:,3]；背景透明化：image[mask==0] = [0,0,0,0]；输出标准化：强制 PNG 格式保存，文件名标记规则：masked-{basename}.png。

基于交互式图像分割的自动掩码应用后的图像预处理效果示意如图 7 所示。

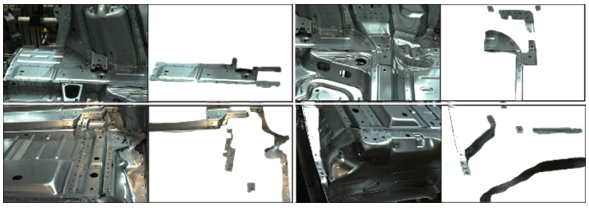


图 7 基于交互式图像分割的自动掩码生成与应用

3.3 深度学习模型工程化落地

3.3.1 数据标注标准化

在服务器或本地计算机上部署计算机视觉标注工具（Computer Vision Annotation Tool，CVAT），完成数据库、后端服务及前端界面的配置，如图 8 所示。根据实际需求设置用户权限，支持多人协同标注。如图 9 所示，基于 CVAT 平台制定焊点、螺柱焊等 12 类车身连接工艺的标注规范。数据集包含 1 500 张高分辨率图像，涵盖不同光照条件、拍摄角度及连接工艺状态（正常与异常），能够清晰呈现连接工艺的细节特征。本文遵循项目创建、任务分配、标注操作、审核修改、数据导出的标准化流程，完成了全部白车身工艺图像的标注工作。

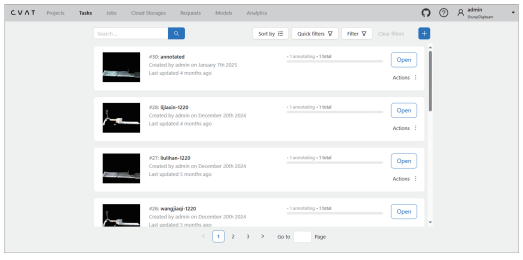


图 8 CVAT 协同标注界面

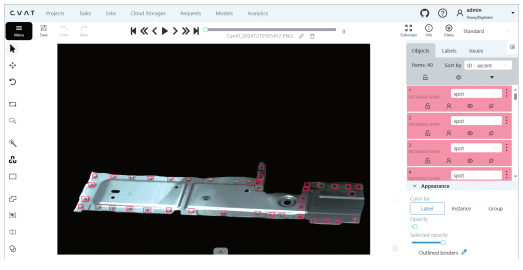


图 9 工艺信息标注

3.3.2　模型训练与调优

在配置Python环境的计算机上,通过pip安装YOLOv8所需的依赖库(如Pytorch)。为确保硬件配置满足模型训练需求,配备高性能图形处理器(Graphics Processing Unit,GPU)加速训练进程。如图10所示,将从CVAT导出的数据按YOLOv8的数据结构组织,划分为训练集、验证集和测试集。对图像进行归一化处理并调整尺寸至模型输入大小,同时对标注数据进行格式转换,确保与YOLOv8输入格式兼容。根据白车身连接工艺检测任务特点配置YOLOv8模型的超参数,包括学习率、批次大小、迭代次数等。图11所示为启动训练程序后,模型从训练集加载数据进行学习,并在每个迭代结束后在验证集上评估性能,计算损失函数值(如交叉熵损失)及检测指标(如精确率、召回率、平均精度均值等)。如图12所示,训练过程中需密切关注上述指标的变化,监测模型是否存在过拟合或欠拟合现象。若出现过拟合,可采用增加数据增强操作、调整正则化参数等方式优化;若出现欠拟合,则可考虑增加训练数据量或调整模型结构等措施。训练完成后,即可获得一个能够识别不同车身工艺的视觉检测模型。

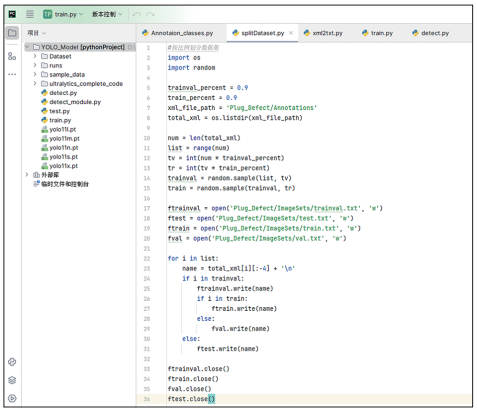


图10　训练集与验证集分配

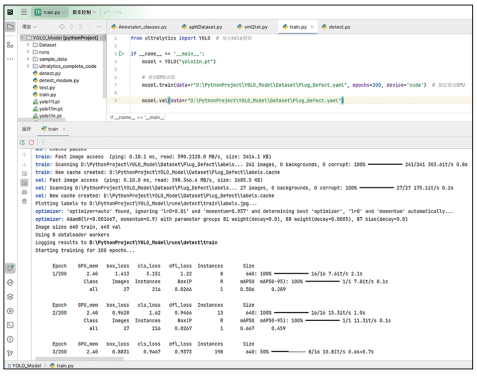


图11　车身工艺视觉检测模型训练

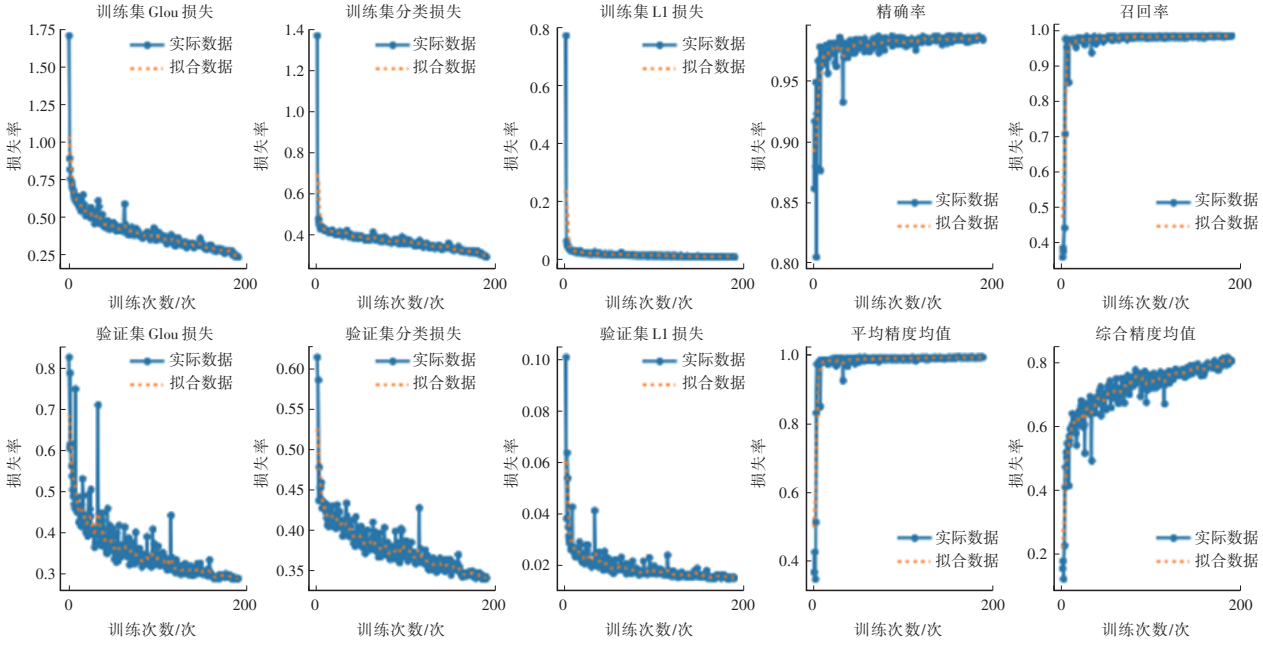


图12　模型性能评估:损失函数值&检测指标

3.3.3　模型评估与优化

为全面评估不同深度学习模型在汽车白车身连接工艺视觉检测任务中的综合性能,本研究选取YOLOv8系列的不同规模模型YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l、YOLOv8x及DETR开展对比评估试验。所有模型均采用相同的白车身连接工艺标注数据集完成训练,验证阶段选取5张涵盖焊接、铆接等多种连接工艺典型场景的代

表性白车身图像构建验证集，确保评估结果的客观性与针对性。各模型在验证集上的检测准确率对比数据如表 1 所示。

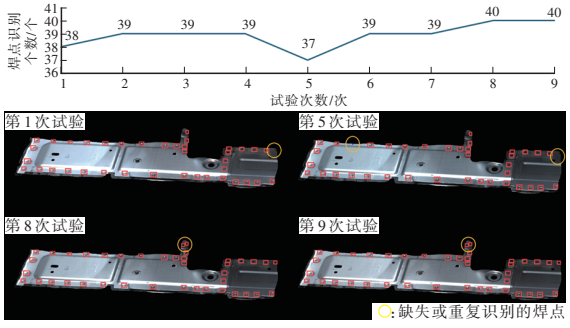
表 1 不同模型在验证集中的准确率 %

模型名称	图 1	图 2	图 3	图 4	图 5	平均值
YOLOv8n	82.0	83.5	81.0	84.0	82.5	82.3
YOLOv8s	85.0	86.0	85.5	86.5	87.0	85.6
YOLOv8m	88.5	89.0	89.5	88.0	89.0	88.9
YOLOv8l	90.0	90.5	91.0	90.5	90.0	90.2
YOLOv8x	91.5	92.0	92.5	91.5	92.0	91.5
DETR	99.0	98.5	99.0	98.0	99.5	98.8

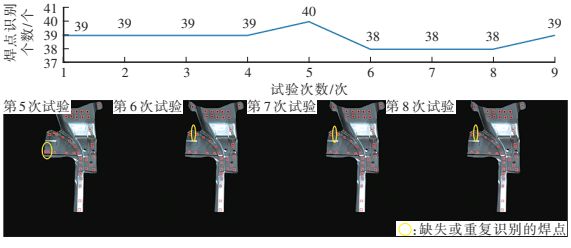
由表 1 可知，在白车身连接工艺检测任务中，DETR 模型的平均检测准确率达 98.8%，显著优于 YOLOv8 系列模型。YOLOv8 系列模型的检测准确率随模型参数量与规模的增大呈梯度提升趋势，其中 YOLOv8x 模型平均准确率为 91.5%，为该系列最优。综合考量检测精度、模型部署复杂度及工业现场应用需求，本文选定 DETR 模型作为汽车白车身连接工艺视觉检测的核心模型。

为进一步提升模型在工业场景下的检测精度与重复精度，验证其检测结果的稳定性与可靠性，本文选取 15 组涵盖不同光照条件、背景环境及连接工艺类型的白车身图像，每组分别开展 9 次重复性验证试验。试验结果显示：11 组图像的检测准确率为 100%，表明模型在该类场景下可稳定识别所有连接工艺特征，无漏判、误判现象；其余 4 组图像的检测准确率存在微小波动，如图 13 所示，经溯源分析，波动成因为白车身表面的油污、灰尘等脏污区域被模型误识别为连接工艺特征，导致检测精度下降。

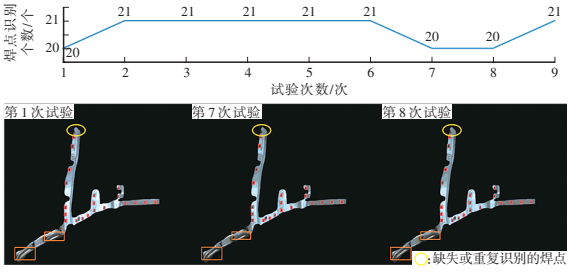
针对脏污干扰问题，本文采用 3.2.2 节提出的基于交互式图像分割的自动掩码技术，对 4 组存在干扰的图像进行预处理：通过精准标注脏污区域并生成掩码，裁剪移除图像中的无关干扰区域，仅保留白车身连接工艺的有效检测区域。预处理后对该 4 组图像重新验证，结果显示，其检测准确率提升至 99% 以上，有效消除了脏污对模型检测的干扰。在此基础上，对全部 15 组图像开展全量验证，优化后模型的平均检测准确率达 99.2%，且检测结果的重复精度显著提升，准确率标准差降低至 0.3%。上述结果表明，优化后的 DETR 模型兼具高检测精度与优异的稳定性，满足汽车白车身批量生产过程中对连接工艺检测的要求。



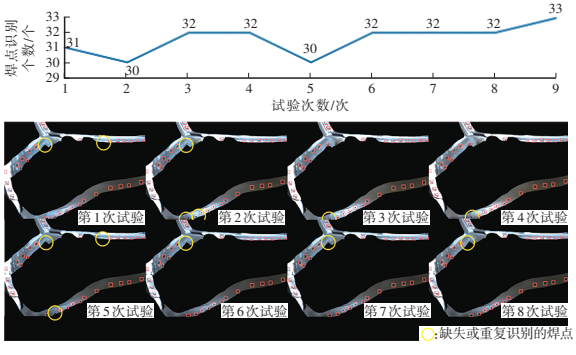
(a)第 0 号相机测试结果(焊点识别目标数:39)



(b)第 5 号相机测试结果(焊点识别目标数:39)



(c)第 9 号相机测试结果(焊点识别目标数:21)



(d)第 14 号相机测试结果(焊点识别目标数:33)

图 13 模型对白车身连接工艺特征检测的重复性验证试验

3.4 全栈式数据管理系统

3.4.1 后端数据处理与存储架构

本系统后端数据库采用 MySQL 与 PostgreSQL 双数据库架构设计，构建四层关联数据表，即零件基础信息表(Part Info)、连接工艺统计量表(Joint

Stats)、缺陷信息表(Defect Info)、图像存储表(Image Storage),并针对各表核心字段(如零件编号、VIN码、缺陷类型)建立B+树索引,提升数据检索效率。基于Node-RED低代码开发平台实现检测结果的对象表示法(JavaScript Object Notation,JSON)数据解析,完成坐标系统一化转换(从像素坐标系映射至工件坐标系),同时引入置信度阈值动态调整机制及时序关联算法,实现检测数据的时空维度匹配。通过索引优化策略、Redis缓存机制与聚合查询语句优化,实现后端数据处理性能的全面优化。

3.4.2　前端交互与可视化设计

通过前端核心功能模块设计,如表2所示,对每台抽样车身以及历史一段时间内的相关质量信息进行统计与分析,针对焊装车间质量门检查白车身的连接工艺数量及位置信息,建立蓝领友好界面,实现车身连接工艺缺陷照片实时展示、焊接类型占比展示、交互式图像查看等数据可视化功能,如图14所示。

表2　前端核心功能模块

模块名称	功能描述
实时看板	展示当前产线连接技术分布、缺陷数量趋势
历史查询	按时间/零件编号/缺陷类型筛选检测记录
图像标注	点击数据条目时叠加显示原图检测框与缺陷标记
报表导出	生成PDF/Excel格式的日/周/月报(含统计图表)
摄像头控制	一键启动/停止采集、调整焦距与曝光参数

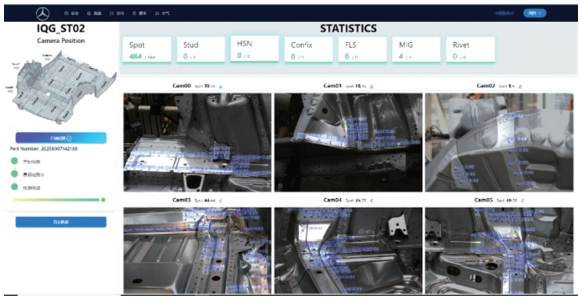


图14　汽车工艺智能检测系统用户界面布局

3.4.3　用户测试与反馈迭代

为验证系统在真实工业生产场景下的可靠性、易用性与检测效能,选取焊装车间车身后底板工位开展为期2周的实地验证测试。测试团队由5名工艺工程师(负责缺陷标定、数据统计与算法效果评估)及13名一线蓝领操作工(负责系统实际操作与使用体验反馈)组成,覆盖车间每日两班制生产模式,完成累计64台白车身的全流程连接工艺检测。测试聚焦后底板工位核心连接工艺缺陷类型,即点焊、螺栓连接、熔化极惰性气体保护焊(Metal Inert Gas Welding,MIG)缺陷,重点验证系统对传统人工检测难度高的场景(如狭小空间焊点、螺栓浮钉)的适配能力,并建立"功能优化→体验提升→效率改进"的闭环迭代体系,按月度开展系统版本升级,具体优化措施如下:

a. 针对操作工反馈的"检测结果显示不够明显,屏幕距离较远"的问题,新增语音播报功能,包含缺陷类型(如"焊点缺失""螺栓浮钉")、缺陷位置(如"后底板左侧第3颗螺栓")等信息的实时语音提醒,无需人工持续关注屏幕即可同步获取检测结果。

b. 针对工程师反馈的"数据导出流程繁琐、筛选维度单一"问题,优化了UI界面的报表生成模块,支持按缺陷类型、生产班次、车身VIN码的快速筛选导出,数据统计效率提升30%。

c. 针对复杂工况下系统偶发的轻微误判,迭代了AI视觉算法模型,新增"飞溅干扰库"及"脏污异常样本库",模型识别准确率提升至99.9%。

优化后的系统不仅解决了历史版本在缺陷检测方面的痛点,更通过人机协同设计适配车间实际生产需求,为焊装车间连接质量的精准管控提供了可靠支撑。

4　结束语

本文以汽车白车身连接工艺检测的工业需求为导向,融合计算机视觉与深度学习技术,研发了一套汽车白车身工艺智能检测系统。该系统通过高精度成像硬件搭建、交互式图像分割掩码预处理、DETR目标检测算法优化及全栈式数据管理系统构建,有效解决了工业场景下复杂光照干扰、微小连接工艺缺陷识别、检测速度与精度平衡等核心技术问题。实际应用验证表明,该系统对白车身连接工艺的检测准确率稳定提升至99%以上,检测效率远超人工目视检测,且检测结果具备客观、可重复、无主观偏差的特性;为期两周的车间实际测试结果显示,相较于传统人工检测模式,系

统的单台车身检测平均完成时间缩短80%，漏检率降低67%，显著提升了汽车焊装车间的生产效率与质量管控水平。

在生产端，该系统可大幅优化车间人员配置，减少人工目视检测岗位的数量需求；在成本端，通过降低缺陷漏检率与返工率，预计每年可节省人工成本、质量返工成本数百万元，同时有效降低终端客户投诉率，强化品牌市场信誉。目前该系统已实现跨域部署，可全面覆盖冲压、焊装、涂装、总装等汽车制造核心工艺环节，推动整车制造全流程质量管控的标准化、数字化与一致性。此外，系统运行过程中积累的海量检测数据可反哺生产工艺优化：通过挖掘缺陷高频发生区域与类型，针对性调整焊接参数、工装定位精度等，实现从“事后检测”到“事前预防”的质量管控模式升级。

参考文献：

[1] 曹立硕. 基于深度学习的乘用车轮毂防错视觉检测方法研究[D]. 广州: 华南理工大学, 2024.

[2] 童金武, 王希, 邓明洋, 等. 缺陷检测中小样本问题的研究进展[J]. 桂林电子科技大学学报, 2026, 46(1): 10-19.

[3] 大卫·米兰·埃斯克里瓦. OpenCV 4计算机视觉项目实战：原书第2版[M]. 北京: 机械工业出版社, 2019.

[4] 王华溢, 黄要诚, 蔡波. 基于传统方法与深度学习方法的图片相似度算法比较[J]. 计算机系统应用, 2024, 33(2): 253-264.

[5] 阿斯顿·张, 扎卡里·C. 立顿, 李沐, 等. 动手学深度学习：PyTorch版[M]. 北京: 人民邮电出版社, 2023.

[6] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[C]// Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014: 580-587.

[7] GIRSHICK R. Fast R-CNN[C]// Proceedings of the 2015 IEEE International Conference on Computer Vision. Santiago: IEEE, 2015: 1440-1448.

[8] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

[9] 徐彦威, 李军, 董元方, 等. YOLO系列目标检测算法综述[J]. 计算机科学与探索, 2024, 18(9): 2221-2238.

[10] 杨建华, 李瑞峰. 目标检测与YOLO算法原理及应用[M]. 北京: 人民邮电出版社, 2023.

[11] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end Object Detection with Transformers[C]// Computer Vision-ECCV 2020: 16th European Conference. Glasgow: Springer, 2020: 213-229.

[12] ZHAO Y, LV W, XU S, et al. DETRs beat YOLOs on Real-Time Object Detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2024: 16965-16974.

[13] 李沂杨, 陆声链, 王继杰, 等. 基于Transformer的DETR目标检测算法研究综述[J]. 计算机工程, 2026, 52(4): 62-81.

AT&M