

基于机器视觉的透明方杯位姿与缺陷检测方法^{*}

王龙达 刘强 任志贵 王军利 刘文帅

(陕西理工大学机械工程学院 汉中 723000)

摘要: 针对医用透明方杯人工检测效率与精确率低、对产品有磨损的问题,引入视觉检测技术,结合图像处理与深度学习,设计开发一种医用透明方杯位姿与缺陷检测系统。针对方杯位姿检测,首先,通过基于改进动态阈值准确分割方杯图像边缘区域。然后,改进 Canny 算法的高低阈值设计,检测出更多细微幅值差别的边缘,设定共线边缘合并的阈值条件,进而拟合得到两边缘直线,基于两边缘计算出虚拟中轴线。最后,通过最小外接矩形粗定位中心,粗定位点到轴线的垂足作为方杯精定位坐标,根据仿射变换将方杯图像矫正到参考位姿。缺陷检测方面,对图像局部缺陷进行灰度拼接,构建 3 种缺陷类间均衡的图像数据集。基于 SqueezeNet、Inception-V3 与 ResNet-50 三种神经网络迁移学习训练后,综合评估发现 SqueezeNet 模型的性能表现最优,在测试集平均精度达 98.6%,识别精确率和召回率分别达 99.8% 与 98.8%。实验验证结果表明对单张图像位姿与缺陷检测速度分别为 770.5 ms 与 553.1 ms,改进后位姿检测精度更高,检测缺陷准确率达到 94%,具有良好的实时性与稳定性,可满足检测要求。该研究可为医用透明方杯的位姿与缺陷检测提供技术支持。

关键词: 机器视觉;图像处理;位姿检测;目标检测;缺陷检测

中图分类号: TP391.4; TN919.8 **文献标识码:** A **国家标准学科分类代码:** 520.60; 460.4030

Transparent square cup pose and defect detection method based on machine vision

Wang Longda Liu Qiang Ren Zhigui Wang Junli Liu Wenshuai

(School of Mechanical Engineering, Shaanxi University of Technology, Hanzhong 723000, China)

Abstract: Aiming at the problems of low efficiency and accuracy of manual inspection of medical transparent square cups and wear on products, visual inspection technology is introduced, combining image processing and deep learning, a medical transparent square cup posture and defect detection system is designed and developed. For square cup posture detection, firstly, the edge area of the square cup image is accurately segmented based on the improved dynamic threshold. Then, the high and low threshold design of the Canny algorithm is improved to detect more edges with subtle amplitude differences, and the threshold condition for merging collinear edges is set, and then the two edge straight lines are fitted, and the virtual central axis is calculated based on the two edges. Finally, the center is roughly located by the minimum circumscribed rectangle, and the foot of the rough positioning point to the axis is used as the square cup precise positioning coordinate, and the square cup image is corrected to the reference posture according to the affine transformation. In terms of defect detection, the local defects of the image are grayscale spliced to construct an image data set that is balanced between the three defect classes. After training three neural networks, SqueezeNet, Inception-V3 and ResNet-50, through comprehensive evaluation, it was found that the SqueezeNet model had the best performance, with an average accuracy of 98.6% in the test set, and recognition accuracy and recall rates of 99.8% and 98.8% respectively. The experimental verification results show that the detection speeds for the pose and defects of a single image are 770.5 ms and 553.1 ms respectively. After the improvement, the pose detection accuracy is higher, and the defect detection accuracy rate reaches 94%, which has good real-time performance and stability and can meet the detection requirements. This research can provide technical support for the pose and defect detection of medical transparent square cups.

Keywords: machine vision; image processing; pose detection; object detection; defect detection

0 引言

在医疗领域中,医用透明方杯广泛应用于临床实验、诊

断、药物研发等关键领域。随着机器视觉与图像处理技术的广泛应用^[1-2],视觉检测的效率与精度进一步提升^[3-4]。

在医用透明方杯的生产与运输存取过程中,方杯位姿会由

于各种原因发生变化,引起末端执行器,如机械手的定位不准确,所以需要进行精确的位姿检测^[5]。方杯中披峰、起雾、污点等缺陷将直接对光学检测与实验结果的准确性产生不良影响,为提高光学检测的稳定性,需要有效筛选出瑕疵产品,保障交付产品合格。人工筛查缺陷的方式成本高、效率低,且依赖检测人员经验。传统的视觉检测需人工设计缺陷的特征描述,适用于缺陷特征明显且固定的应用。而基于深度学习的方法能够有效提高缺陷检测的速度与准确性^[6],由于医用透明方杯缺陷形貌纹理等具有一定差别,固定特征难以准确描述缺陷,因此,采用基于深度学习的方法来识别方杯缺陷,缺陷类型可进行统计,通过先验知识汇总并标注数据训练模型,降低手工设计特征的成本,且检测精度大幅提高。

位姿检测方面,罗振威等^[7]提出一种基于边缘信息拟合的芯片定位算法,能够实现芯片中心的亚像素级定位,通过去除芯片倒角后拟合边缘直线,提升了检测精度。王朕等^[8]提出一种基于 Fréchet 距离相似度的轮廓提取方法,检测出的边缘更准确,利用矩形度联和边缘特征实现位姿检测。刘建春等^[9]利用差分多线结构光技术降低了光条中心线出现异常点情况,改进拟合算法提升边缘与角点的提取精度。刘茂霖^[10]基于改进 Randon 变换检测直线,实现缩小变换域并去除噪点,结合迭代加权最小二乘检测边缘,有效提升检测精度。陈鹏等^[11]基于改进的 Blob 分析定位算法,结合图像灰度取反、直方图均衡和分水岭分割算法,增强在黑色托盘图像中提取圆心坐标的可靠性。神经网络与缺陷检测方面,王文秀等^[12]通过引入注意力模块增强模型的特征提取性能,将样本数据平衡,提高数据不足缺陷类别检测准确率。郭子宁等^[13]基于 SqueezeNet 模型,借鉴 DenseNet 思想,改善训练前期梯度退化的问题,调整 Fire 模块的卷积内核通道参数,使得模型在训练过程中能获取更多图像细节,提高处理的能力。蒲秋梅等^[14]提出一种基于改进 Inception-V3 网络的肺炎检测方法,引入残差连接、空间注意力与通道注意力机制以及通道混洗,提升 X 光胸片在肺炎诊断中的准确度和泛化能力。常青等^[15]提出一种结合改进 ResNet-50 与迁移学习的竹片缺陷识别方法,通过引入局部二值法(local binary patterns, LBP)和 OTSU

(注,OTSU 为提出此算法的人名,中文名译为大津)算法进行特征提取,利用预训练模型提高识别准确性,并搭建了分选系统验证其可行性。王莉等^[16]在干香菇分级识别中将传统的 ResNet18 中 Stem 模块的 7×7 卷积层替换为 3 个 3×3 卷积层串联,减少参数计算量,并引入非对称卷积和 h-swish 激活函数,学习更深层的特征。刘兰等^[17]利用 ResNet18 作为基础模型,扩展特征金字塔网络(feature pyramid network, FPN),包含小目标特征模块,强化微小缺陷的检测能力。Ai 等^[18]在 YOLOv 5s 网络中引入通道注意力机制、双向特征金字塔网络(bidirectional feature pyramid network, BiFPN)模块与深度可分离卷积,实现了检测精度与速度的提升。总的来说,在透明方杯位姿检测中,需要高鲁棒性与更加适应检测任务的方法,在缺陷检测方面,得益于卷积神经网络强大的特征提取和高维数据的表征能力,研究者将卷积神经网络(convolutional neural network, CNN)的特征提取应用在目标检测中,通过优化算法提升性能,但由于方杯材质透明,使得缺陷目标与背景相似,检测任务挑战巨大,仍需进一步研究。

基于此,本文将机器视觉与深度学习技术应用在医用透明方杯位姿与缺陷检测中,以提升检测精度与效率。利用待分割区域与背景灰度的差值进行局部动态阈值分割,改进 Canny 算法高低阈值设计,检测出更多微小的边缘轮廓,设置阈值条件将其线轮廓合并,去除伪边缘后得到两条边缘直线,进而拟合出虚拟中轴线,利用最小外接矩形粗定位中心,再利用粗定位点到轴线的垂足作为最终精定位坐标,计算出方杯位姿。自制类间均衡的缺陷图像数据集,基于迁移学习方法检测方杯缺陷。通过搭建检测系统实验验证,为医用透明方杯检测提供可靠的方法与技术支撑。

1 检测系统设计

医用透明方杯检测框图如图 1 所示,硬件由双远心镜头、CMOS 相机、背光源、光源控制器、电动旋转平台、支架底座与计算机等组成。计算机运行位姿检测算法程序实现医用透明方杯位姿检测与纠偏;运行迁移学习算法程序实现医用透明方杯的缺陷检测。使用的计算机配置与视觉系统硬件参数分别如表 1 和表 2 所示。

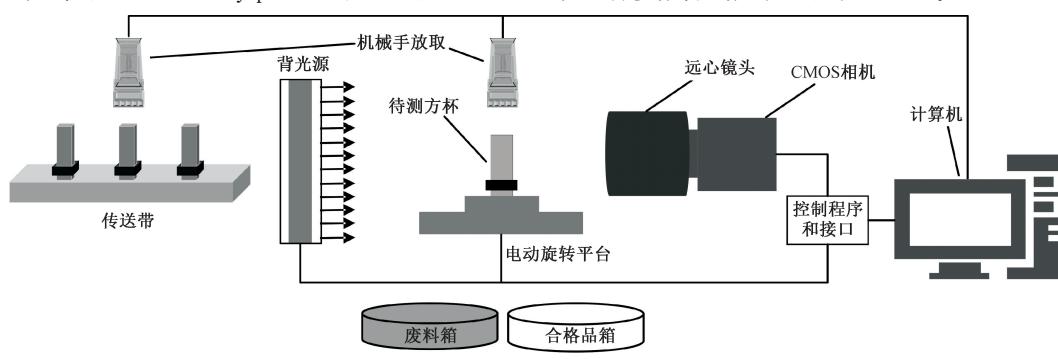


图 1 方杯检测系统框图

Fig. 1 Block diagram of square cup detection system

表 1 实验平台参数

Table 1 Experimental platform parameters

类型	型号	内存	平台	版本
CPU	Core(TM)i9-13900KF	3.0 GHz	HALCON	22.05×64
GPU	RTX4080	16 G	DLTool	23.11
RAM	DDR4	64 G	Visual Studio	22

表 2 硬件参数

Table 2 Hardware parameters

硬件	型号	参数
相机	MER-1220-32U3M	分辨率 4024×3036 , 像元 $1.85 \mu\text{m} \times 1.85 \mu\text{m}$
镜头	DTCM118-64H-AL	放大倍率 0.141 倍, 物距 158 mm
光源	DH-GY-90X90W	发光面积 $90 \text{ mm} \times 90 \text{ mm}$

整体检测流程如图 2 所示,首先,采集图像时机械手将方杯放置在平台上,相机与方杯长轴方向呈 90° 。当计算机获取机械手信号后启动相机,开始运行位姿检测程序并将定位后的图像编号保存。图像满足数据集条件后结束采集,拍照间隔 300 ms。接着,基于灰度值拼接实现方杯图像缺陷数据集类间均衡,标注均衡后的数据集。然后,基于迁移学习训练并评估 SqueezeNet、InceptionV-3 和

ResNet-50 三种神经网络的最优检测模型。最后,将算法部署到人机交互软件中,实现相机实时采集并检测方杯位姿与缺陷,判断产品是否合格,利用机械手抓取放置在合格箱或废料箱。最终将检测结果保存以便后续查看。

2 方杯位姿检测

在机械装置存放吸取方杯一系列动作后,方杯易与参考位姿产生偏差,一是由于机械手的机械误差导致的方杯中心偏移产生位置偏差,二是由于方杯的角度偏转产生与参考角度的偏差,所以需要检测医用透明方杯的位姿(位置与方向),即确定方杯在像素坐标系的位置与角度,以便进行定位与抓取、精密测量等工作。通过图像处理的方法对方杯图像进行特征提取后分割感兴趣区域(region of interest, ROI),基于 ROI 内亚像素边缘拟合两条直线,基于轴线计算出虚拟中轴轴线的角度,计算最小外接矩形中心点坐标后,计算中心点垂直于轴线的垂足坐标,作为方杯最终中心坐标。经过一系列处理后,得到方杯位姿。机械手利用计算的位姿信息完成方杯的平移与旋转来完成纠偏。同时计算方杯位姿与标准图像位姿之间变换矩阵,对采集的方杯图像进行射影变换,可将方杯图像变换到需要的位置与角度。使图像以固定的位姿进行后续检测与取放工作,方杯图像位姿定位的流程如图 3 所示。

2.1 图像增强

为了图像的边缘特征显得更加突出且清晰,对图像进行去噪与增强。首先,采用双边滤波对图像去噪,可以在去除噪声的同时保留边缘细节。双边滤波由两个高斯滤波函数乘积构成,在空间距离权重的基础上引入像素值域权重,在图像中边缘处像素值差异大,权重系数趋于 0,对边缘的滤波影响很小,双边滤波输出为:

$$\mathbf{I}_p^{bf} = \frac{1}{W_p^{bf}} \sum_{q \in S} \mathbf{G}_{\sigma_s}(\|p - q\|) \mathbf{G}_{\sigma_r}(\|\mathbf{I}_p - \mathbf{I}_q\|) \mathbf{I}_q \quad (1)$$

式中: W_p^{bf} 为归一化因子, \mathbf{G}_{σ_s} 为空间域权重, \mathbf{G}_{σ_r} 为像素值域权重, p 为计算的像素坐标, q 为 p 的邻域。

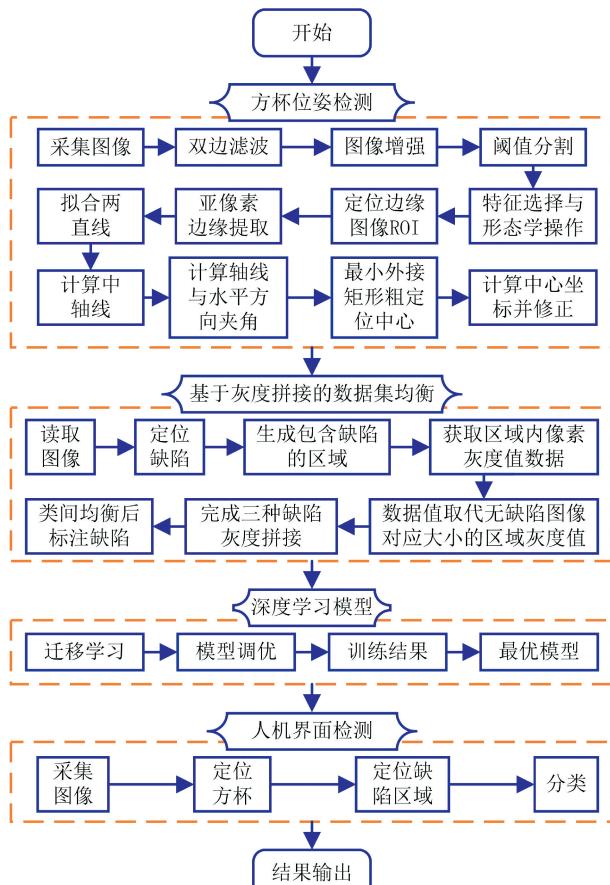


图 2 检测流程图

Fig. 2 Detection flow chart

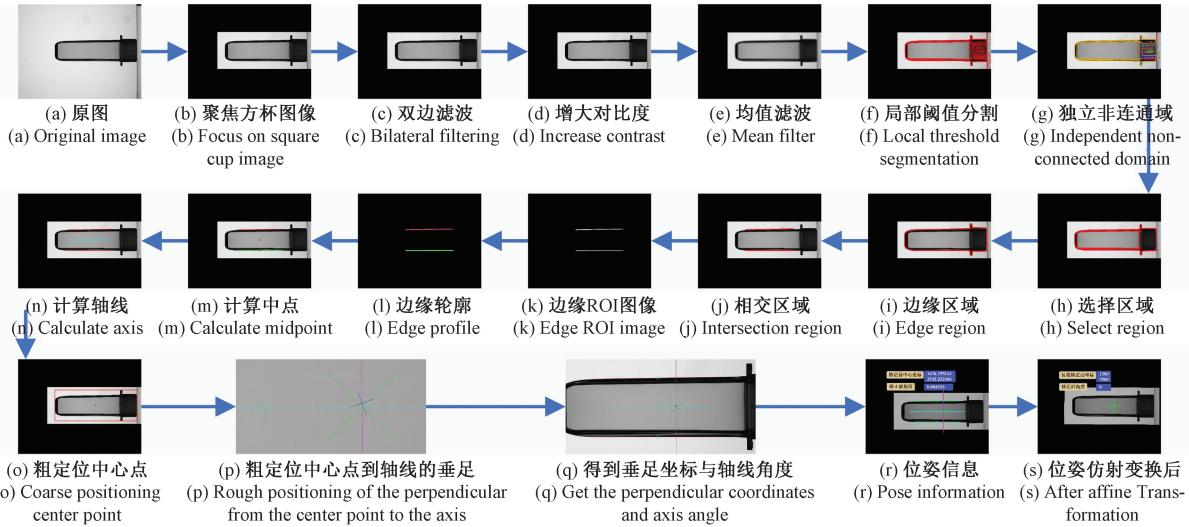


图 3 方杯位姿定位流程

Fig. 3 Square cup posture positioning process

接着,增大图像的对比度。首先对双边滤波后的图像使用模板为 7×7 的均值滤波,接着对图像灰度进行缩放,得到增强后的图像。图像增强后的灰度计算公式如下:

$$res = round((orig - mean) \times factor) + orig \quad (2)$$

式中:*orig* 和 *mean* 分别为增强之前与经过均值滤波后图像的灰度值; *factor* 为增强因子,用来控制对比度增强的幅度,设为 1.2; *round* 表示经过计算的结果转换为最接近的整数值; *res* 为经过增强后的灰度值。

通过对像素灰度值进行调整,使得高于局部平均值的像素变得更亮,低于局部平均值的像素变得更暗,增强后的结果如图 3(d)所示。

2.2 阈值分割

在确定 ROI 图像时,根据图像灰度直方图选取灰度阈值,将每个像素点的灰度值跟选取的阈值作比较,根据比较结果进行分类,用不同的数值标记不同类别的像素,从而生成具有固定阈值的图像区域。阈值分割可定义为:

$$S = \{(r, c) \in R \mid g_{min} \leq f_{r,c} \leq g_{max}\} \quad (3)$$

式中: g_{min} 、 g_{max} 分别为设置的最小与最大阈值, $f_{r,c}$ 为在图像中某一行列像素的灰度值, S 为阈值范围下的所有像素的区域的集合。

如图 4 所示,利用方杯灰度直方图进行图像分割时,在方杯灰度直方图中有 4 个峰值,3 个谷底,选择 3 个谷底作为分割图像的阈值。图中有两个明显谷底灰度值为 71 与 180,有不明显的 99 的谷底,这 3 个灰度阈值将图像划分为 4 个区域。分割后的二值图像定义为:

$$g(x) = \begin{cases} a, & f(x,y) \leq T_1 \\ b, & T_1 < f(x,y) \leq T_2 \\ c, & T_2 < f(x,y) \leq T_3 \\ d, & f(x,y) > T_3 \end{cases} \quad (4)$$

式中:*a*、*b*、*c*、*d* 分别对应阈值内的分割结果。 T_1 、 T_2 、 T_3

为 3 个阈值,分别为 71、99、180。

如图 4(a)所示,分割得到包含方杯边缘的区域,其灰度值为 0~71。

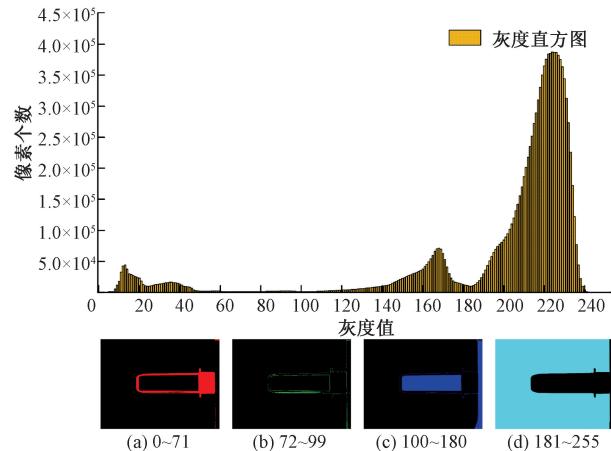


图 4 根据灰度直方图阈值分割

Fig. 4 Segmentation based on grayscale histogram threshold

在实际情况下,由于方杯位置偏移,表面产生反光,存在图像局部出现多种灰度范围,使用基于直方图确定的全局固定阈值来分割图像容易将真正 ROI 边缘排除在外,影响后续边缘提取的准确性。因此,本文采取局部阈值分割出更准确的图像边缘 ROI,两种分割算法分割效果如图 5 所示。

图 5(a)为灰度直方图分割,可以看出这种算法将真正的方杯边缘排除在外,图 5(b)为局部阈值分割,将有灰度变化的方杯边缘精确分割出。图 5(c)和(d)为对这两种方法分割出的区域进行边缘提取的结果,可以看出局部阈值分割后进行边缘提取更精确。图 5(e)显示图 5(c)沿 A-A 的灰度分布,出现亮-暗-亮-暗-亮的变化。在这些局部区域

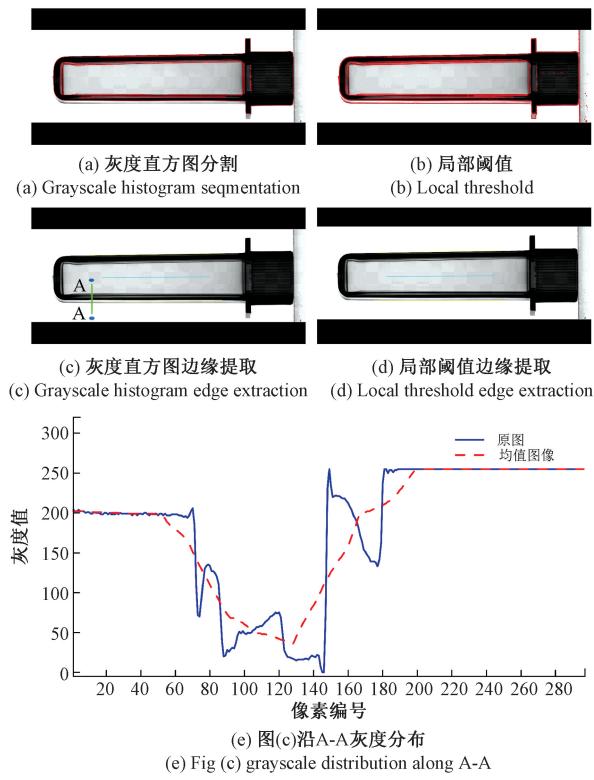


图 5 分割效果对比

Fig. 5 Segmentation effect comparison

中,感兴趣区域与背景的灰度值的区别仍然存在。因此使用改进的局部阈值分割与均值滤波结合,利用邻域的局部灰度对比,可以有效解决问题。首先,对图像采取均值滤波平滑,本文中设置均值模板的大小为 39×39 ,使用均值滤波后的输出作为对局部背景灰度的估计。通过比较原始图像与均值滤波后图像的灰度差值来分割目标区域。在分割图像中较暗区域时,感兴趣区域的灰度值满足的公式如下:

$$g_o \leq g_i - offset \quad (5)$$

式中: g_o 为原始图像灰度, g_i 为经过均值滤波处理的图像灰度, $offset$ 为应用于均值滤波图像上的偏移量,一般这个值不应过小,否则分割的区域会包含过多噪声,本文设为 5。

在灰度发生多次亮暗变化时,局部阈值明显优于灰度直方图分割,提取的边缘也是真正方杯边缘,后续可以更加准确计算中轴线。图 5(e)显示了原图与均滤波值处理后的沿 A-A 分布的灰度值,均值处理后的结果作为局部背景,待分割的暗区域比背景要暗,将均值图像对应灰度值减去 5 后比原图的灰度还大的区域分割出。

2.3 拟合虚拟轴线与定位中心坐标

边缘作为图像的重要信息,一般用梯度幅值局部最大定义图像的边缘。精确提取边缘能够实现定位与测量目标物尺寸等任务。常用的边缘提取算法有一阶微分法,如

Roberts、Sobel;二阶微分 Laplace;Canny 算法。边缘提取结果如图 6 所示,图 6(b)~(e)为像素级边缘,图 6(f)是经过插值处理的亚像素边缘。

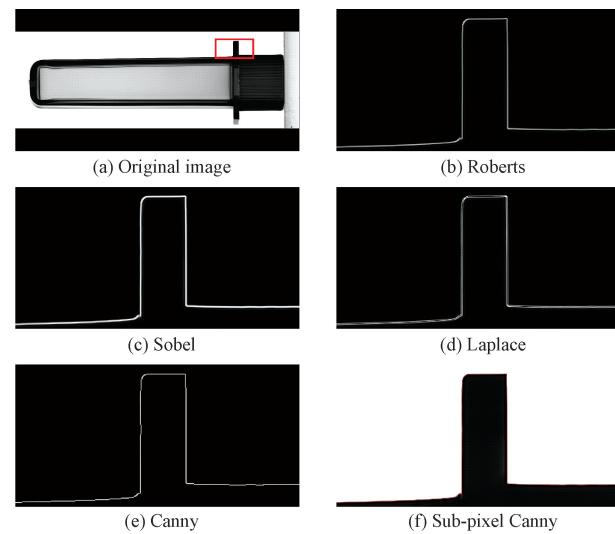


图 6 方杯图像局部边缘提取效果

Fig. 6 Local edge extraction effect of square cup image

从图中可以看出 Canny 算法的效果最优,其原理为:

1)利用高斯滤波器平滑图像;2)用一阶偏导的有限差分计算梯度的幅值和方向;3)利用第二步的梯度强度进行非极大值抑制,在梯度方向上查找局部极大值;4)利用双阈值操作,在高低阈值之间的像素,查找 8 邻域内有无已确定的强边缘,若有,则将其作为边缘,若无则舍弃。

本文基于传统的 Canny 算法改进高低阈值的设计,通过将高阈值与阈值之差设置较低,检测出更多细微幅值差别的边缘。设置相邻的两条边缘轮廓合并的距离、角度与重叠度阈值条件,对检测到的与回归线方向一致的细小边缘按照设置的条件合并连接,在合并共线的所有边缘轮廓中选择距离最远的一对作为方杯的两侧边缘,最终利用二次多项式插值确定亚像素边缘位置。

如图 7 所示,在合并共线轮廓时将 c_1 作为参考轮廓,将输入轮廓按照设置阈值条件合并,将轮廓线最大投影距离设为 30 pixel,轮廓线最大垂直距离设为 5 pixel,轮廓线最大角度设为 11.46° ,轮廓线最大重叠距离 $\min(t_1, t_2)$ 距离设为 30 pixel。

如图 8 所示,基于边缘拟合计算得到中轴线。首先提取多组细微的边缘轮廓,接着合并共线轮廓,合并效果如图 8(b)所示。然后根据轮廓长度过滤掉较短的轮廓,选择最外侧轮廓,利用基于图基权重过滤与回归线较远的干扰点后拟合出最终两边缘直线,如图 8(e)所示。最后根据两边缘直线计算出虚拟中轴线。且在光照强度有增强与减弱时,对不同位姿的方杯检测轮廓边缘的误差波动较小,证明光照变化并不明显影响本方法的稳定性。

如图 9 所示,拟合的两条直线分别为 L_1, L_2 ,据两条

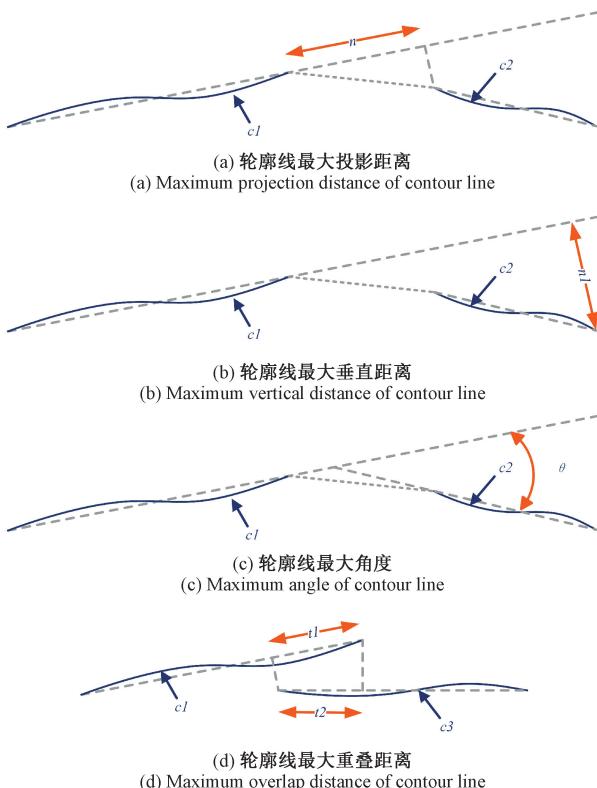


图 7 轮廓合并的条件设置

Fig. 7 Conditional settings for contour merging



图 8 基于边缘的轴线计算

Fig. 8 Edge-based axis calculation

直线段的端点坐标 $M_1(x_{M1}, y_{M1})$ 、 $N_1(x_{N1}, y_{N1})$ 与 $M_2(x_{M2}, y_{M2})$ 、 $N_2(x_{N2}, y_{N2})$ 。可得两条直线段的中点坐标分别为 $P_1(x_{P1}, y_{P1})$ 、 $P_2(x_{P2}, y_{P2})$ ，则中轴线一定经过

P_1 、 P_2 的中点 $P(x_p, y_p)$ ，根据点 P 的坐标与轴线斜率 k 可解得轴线 L 。

$$L_1 = k_1 x + b_1, L_2 = k_2 x + b_2 \quad (6)$$

$$x_{P1} = \left(\frac{x_{M1} + x_{N1}}{2} \right), y_{P1} = \left(\frac{y_{M1} + y_{N1}}{2} \right) \quad (7)$$

$$x_{P2} = \left(\frac{x_{M2} + x_{N2}}{2} \right), y_{P2} = \left(\frac{y_{M2} + y_{N2}}{2} \right) \quad (8)$$

$$x_p = \left(\frac{x_{P1} + x_{P2}}{2} \right), y_p = \left(\frac{y_{P1} + y_{P2}}{2} \right) \quad (9)$$

$$k = \frac{k_1 + k_2}{2} \quad (10)$$

$$L = kx + b \quad (11)$$

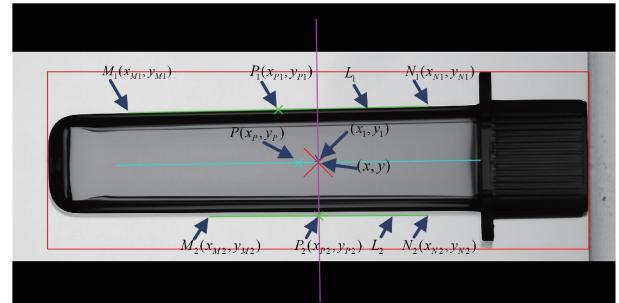


图 9 计算方杯中心坐标

Fig. 9 Calculating the center coordinates of the square cup

轴线 L 斜率为 k ，经过粗定位点的直线的法线斜率为 $-\frac{1}{k}$ ，将 $L = kx + b$ 转换为一般式 $Ax + By + C = 0$ ，其中 $A = k$ ， $B = -1$ ， $C = b$ 。

$$Ax + By + C = 0 \quad (12)$$

$$\frac{B}{A} = -\frac{1}{k} = \frac{y - y_1}{x - x_1} \quad (13)$$

$$x = \frac{B^2 x_1 - AB y_1 - AC}{A^2 + B^2} \quad (14)$$

$$y = \frac{A^2 y_1 - AB x_1 - BC}{A^2 + B^2} \quad (15)$$

式中： (x_1, y_1) 为方杯的粗定位中心，是利用阈值分割区域的最小外接矩形的中心坐标； (x, y) 为垂足(方杯精定位中心)坐标。

2.4 仿射变换

如图 10 所示，假设方杯中心坐标为 (x, y) ，方杯相对于 x 轴的偏转角度为 θ ，旋转中心坐标为 (x_m, y_m) ，将方杯相对于旋转中心旋转 $-\theta$ 度可实现纠偏角度，纠偏后的方杯中心坐标 (x', y') 可利用式(16)计算出。

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & (1 - \cos\theta)x_m - y_m \sin\theta \\ -\sin\theta & \cos\theta & (1 - \cos\theta)y_m + x_m \sin\theta \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (16)$$

在对方杯图像仿射变换时使用双三次插值方式对像素插值，可以得到更加清晰的图像边缘，保证图像质量。

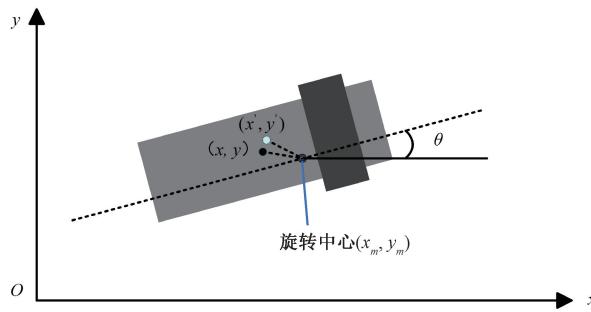


图 10 方杯纠偏原理图

Fig. 10 Schematic diagram of square cup correction

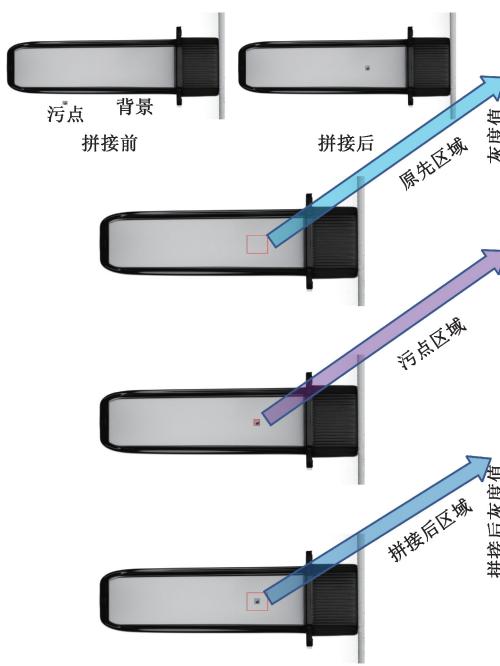
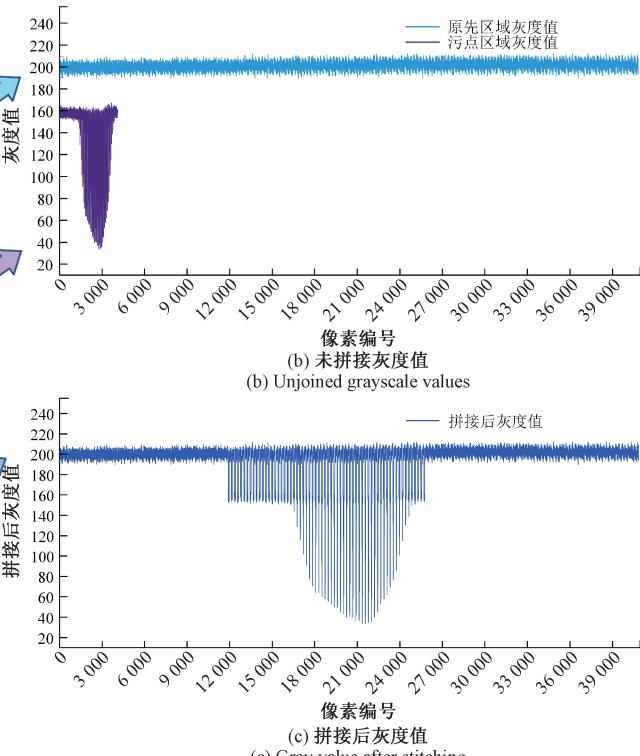
对坐标为 $(i + u, j + v)$ (u, v 为 0~1 的小数) 的像素利用双三次插值计算灰度值, 计算公式如下:

$$S(w) = \begin{cases} \frac{3}{2} |w|^3 - \frac{5}{2} |w|^2 + 1, & |w| \leq 1 \\ -\frac{1}{2} |w|^3 + \frac{5}{2} |w|^2 - 4|w| + 2, & 1 < |w| \leq 2 \\ 0, & |w| > 2 \end{cases} \quad (17)$$

$$A = [S(1+u) \quad S(u) \quad S(1-u) \quad S(2-u)] \quad (18)$$

$$B = \begin{bmatrix} f(i-1, j-1) & f(i, j-1) & f(i+1, j-1) & f(i+2, j-1) \\ f(i-1, j) & f(i, j) & f(i+1, j) & f(i+2, j) \\ f(i-1, j+1) & f(i, j+1) & f(i+1, j+1) & f(i+2, j+1) \\ f(i-1, j+2) & f(i, j+2) & f(i+1, j+2) & f(i+2, j+2) \end{bmatrix} \quad (19)$$

$$C = [S(1+v) \quad S(v) \quad S(1-v) \quad S(2-v)]^T \quad (20)$$

(a) 图像后灰度拼接
(a) Image grayscale stitching(c) 拼接后灰度值
(c) Gray value after stitching

$$f(i + u, j + v) = ABC \quad (21)$$

式中: 矩阵 B 中的项分别为需要插值的坐标点 16 邻域的像素值; $S(w)$ 为双三次插值核。

3 基于灰度拼接的数据集均衡

医用透明方杯待检测的缺陷有披峰、起雾与污点 3 种, 在利用深度学习技术检测方杯缺陷时, 需要大量高质量缺陷数据, 而收集并构建这些数据不容易。在采集图像时发现, 一张图上存在有多个同种缺陷, 有一张图像只有一种缺陷或某两种。使用这些图片构建数据集, 会使得某类缺陷数据过多, 而其他数据相对偏少, 由此会造成训练数据类间分布不均衡。缺陷样本少的种类会对训练结果产生过拟合现象, 标注数据多的会导致模型在预测少数类别时忽略少数类别比重, 导致检测精度降低, 因此对采集的图像进行基于灰度拼接的均衡操作, 使每张图像上均有 3 类缺陷。

图像灰度拼接如图 11 所示, 首先, 对采集图像中缺陷部分进行定位, 分析方杯的缺陷特征, 确定包含披峰、起雾和污点区域大小分别为 64×128 、 512×512 和 64×64 像素。生成包含缺陷区域, 然后获取这些区域每个像素行、列坐标, 接着获得每个坐标的灰度值, 将这些灰度值取代无缺陷的图像对应区域灰度值, 遍历所有缺陷区域像素, 最终完成图像缺陷数据类间均衡。拼接之前, 背景灰度值稳定在 200 左右, 拼接后整个原先区域的灰度值降低, 如图 11(c) 所示。拼接 3 种缺陷如图 11(d) 与 (e) 所示, 能清晰看到灰度值在缺陷区域的变化。

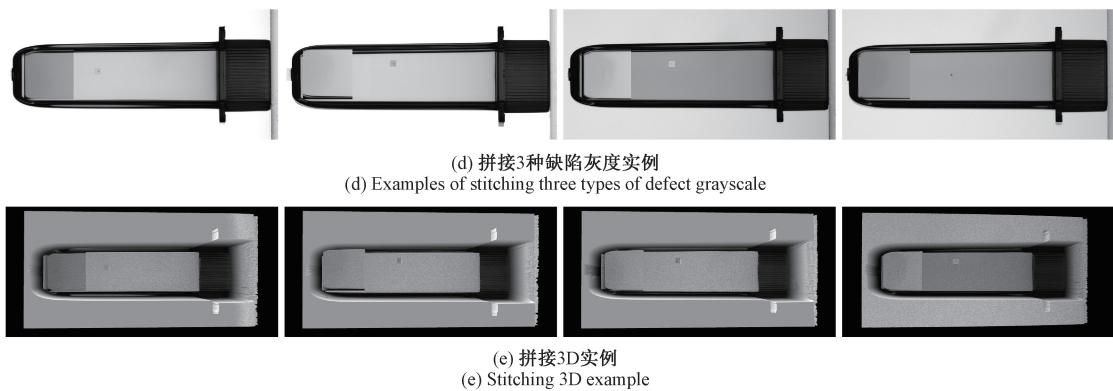


图 11 图像灰度值拼接

Fig. 11 Image grayscale value stitching

将方杯缺陷类间均衡后,首先构建了 543 张缺陷图像,3 类缺陷实例共计 1 629 个,经过水平镜像、垂直镜像、旋转 180° 的操作将缺陷图像扩充至 1.5 倍达 815 张,缺陷实例共 2 444 个,接着对扩充后的数据集图像按 7:1.5:1.5 的比例划分为训练集(570 张),验证集(122 张),测试集(123 张),以此来提高神经网络的泛化能力。在研究中使用 Deep Learning Tool 工具来标注数据集;标注过程中,使用矩形框对缺陷区域标注完成图像数据集制作;标注完成后每张图像含有 3 种缺陷标签结果,标注示例如图 12 所示。

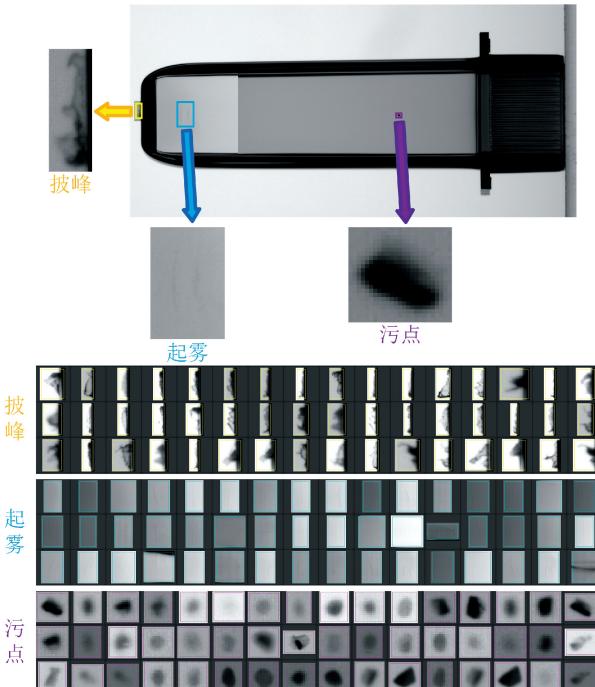


图 12 标注数据示例

Fig. 12 Annotated data example

4 神经网络模型训练与评估

4.1 网络模型

缺陷目标检测原理如图 13 所示,通过生成特征图、特

征图组合,最后将特征图对应层输入附加网络进行定位与分类。首先,由预训练的分类网络组成主要骨干部分,删除分类层后生成编码信息不同的特征图,在生成特征图模块,通过自下而上的路径,使用卷积模块构建深层金字塔,在上层得到语义丰富的特征。然后,具有相同宽高的特征图视为同一种级别,在组合特征图模块通过自上而下路径,用上采样的方式增加分辨率,从语义丰富特征层来重建更高分辨率的层,保留重建层语义信息,不同级别特征图经过组合,获得含有低与高级别信息的特征图,与骨干层结合构成特征金字塔网络,重建特征图与原始特征图融合,使得丰富语义信息并提高定位精度,同时训练网络更加容易。最后,在定位与分类模块通过附加网络获取特征图作为输入信息,重复操作生成后续预测层,实现定位和分类功能,最终完成目标检测任务^[19-20]。

在实现缺陷检测任务中需要充分提取图像信息、提升参数效率与有效避免网络过拟合。将 CNN 的特征提取应用在目标检测中,具有代表性的神经网络包括 SqueezeNet、Inception-V3 和 ResNet-50。SqueezeNet 是一种轻量级卷积神经网络,具有较小的模型大小和计算复杂度,它通过使用 1×1 卷积核和全局平均池化层来减少参数数量,在保证精度的同时,其参数比 AlexNet 少 50 倍。Inception-V3 采用 Inception 模块,通过不同尺度的卷积核并行处理输入特征图,从而提取多尺度特征信息,能够有效提高网络的感受野和特征表达能力,Inception-V3 根据初始 Inception 模块结构,通过将 5×5 卷积核改成 2 个 3×3 的卷积核,降低了计算参数。ResNet-50 通过引入残差模块,成功解决深层网络训练时的梯度消失,使得训练深层的网络更容易。

为提高网络模型训练效率与效果,本研究在预训练神经网络模型上进行迁移学习,考虑表现卓越的 SqueezeNet、InceptionV-3 和 ResNet-50 网络。这些模型在投入使用前,已经通过大量图像数据进行充分训练,确保在后续能够更准确地执行医用透明方杯缺陷检测任务。

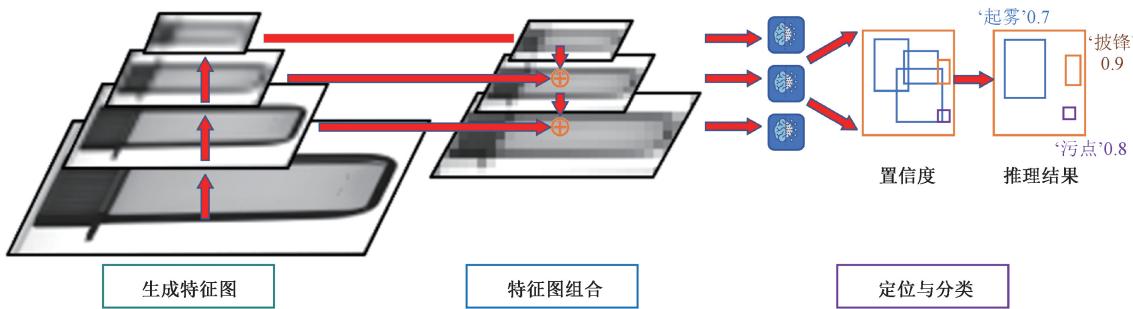


图 13 目标检测原理图

Fig. 13 Schematic diagram of object detection

4.2 训练参数

在数据集的训练过程中,设置合理的训练参数能确保训练过程的高效性与稳定性。主要涉及的参数包括 batch size(批次大小)、epoch(整个训练集的完整训练次数)和 learning rate(学习率)。经过系列试验,发现当 epoch 值达到 50 次后,平均精确率并无进一步提升。因此,为了充分利用硬件条件,在训练过程中将 batch size 设为 1, epoch 设为 50, SqueezeNet 与 Inception-V3 初始学习率设为 0.000 5, ResNet-50 的初始学习率设为 0.000 1, 每进行 5 轮次训练学习率降为原先的 0.9, 求解器类型选择 adam(自适应矩估计),使得训练可自适应调整超参数,加快训练速度并提高收敛速度。

4.3 性能评估指标

一般而言,对训练完成的模型评估性能,主要是以精确率($P_{\text{precision}}$)、召回率(R_{recall})、平均精度(m_{AP})来衡量。精确率即预测正确类别在所有预测结果中所占比例,反映模型预测结果的准确性;召回率,即被正确预测的目标与所有实际正确目标的比值,体现模型检测的全面性;平均精度,即针对交并比(intersection over union, IOU)阈值下每个类别的平均精确率进行平均,更全面反映模型在不同类别上的检测性能;计算方法如式(22)~(25)所示。

$$P_{\text{precision}} = \frac{T_p}{T_p + F_p} \quad (22)$$

$$R_{\text{recall}} = \frac{T_p}{T_p + F_N} \quad (23)$$

$$AP = \sum_{n=1}^N (r_{n+1} - r_n) \rho(r_{n+1}) \quad (24)$$

$$m_{\text{AP}} = \frac{1}{K} \sum_{q=1}^K AP_q \quad (25)$$

式中: T_p 代表实际正类预测为正类的数量,即正确认别出缺陷区域的数量; F_p 为误检数量,即实际为负类却预测为正类的数量,表示模型错误地将背景区域识别为缺陷区域的数量; F_N 为漏检数量,指已标注但模型未找到的目标数量,即实际为正类但预测为负类的数量; AP 为单个类别平均精度; r_n 代表第 n 个召回率的取值, $\rho(r)$ 代表在召回率下对应精确率的取值。 K 表示验证集总数, q 代表单个验证集。

4.4 训练与评估结果分析

设定相应模型参数后,分别使用 3 种神经网络对数据集进行 50 轮次训练,迭代 2 550 次后,利用测试集对网络模型进行评价,包含精确率、召回率、平均精度、损失值以及训练时长。在评估过程中,IOU 阈值选择 0.5,当预测框与真值框的 IOU 大于 0.5 为正确检测,否则为错误检测。

损失值曲线如图 14 所示,训练与评估结果如表 3 所示。3 种网络模型训练结果中,SqueezeNet 损失最低,达到 0.21, Inception-V3 和 ResNet-50 则分别降到 0.24 与 0.28,在平均精度值上,SqueezeNet 与 Inception-V3 模型精度较高,都达到 98.6% 以上。在训练时间与推断时间上,SqueezeNet 网络所需时间最短,Inception-V3 与 ResNet-50 网络模型更复杂,ResNet-50 在 3 者中耗时最长。

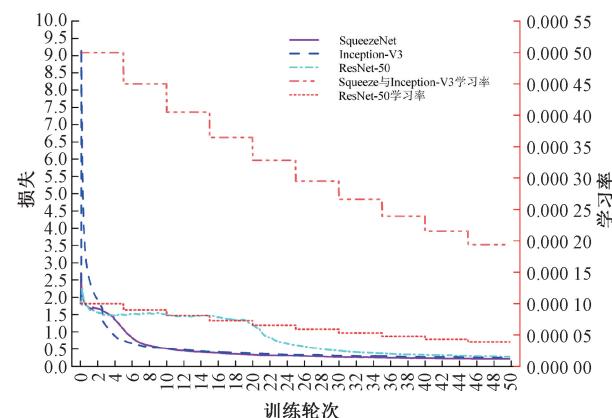


图 14 3 种网络损失曲线图

Fig. 14 Three network loss curves

表 3 3 种网络模型在测试集的训练与评估结果

Table 3 Training and evaluation results of three network models on the test set

网络模型	训练时间	检测时间/ms	平均精度/%
SqueezeNet	13 h 18 m	553.1	98.6
Inception-V3	30 h 22 m	793.4	99.3
ResNet-50	42 h 29 m	3481.3	96.6

3 种网络在测试集的检测评估结果如图 15 所示。(左右两侧分别为精确率、误检率与召回率、漏检率)。

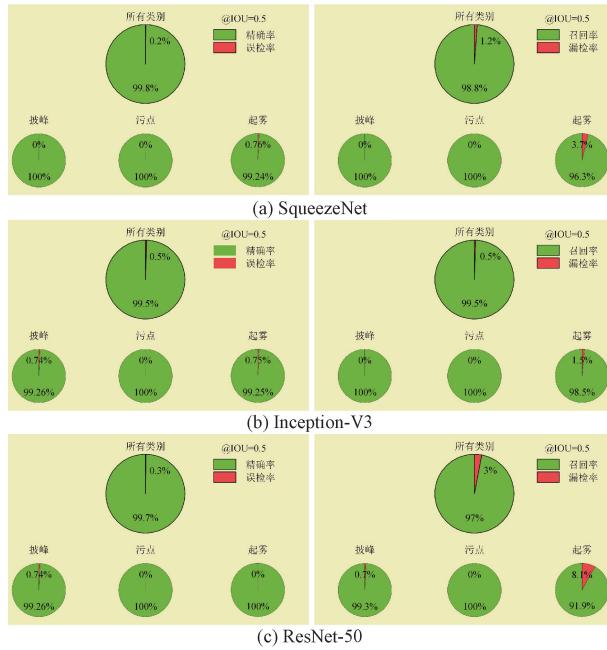


图 15 3 种网络在测试集的评估结果

Fig. 15 Evaluation results of three networks on the test set

从图中可以看到精确率、召回率均高于 97%，相比于 Inception-V3 和 ResNet-50，SqueezeNet 网络综合表现最优，检测一张图像时间最短。主要原因是 SqueezeNet 具有较少的参数量，并采用深度压缩模型，它通过 squeeze 层和 expand 层构建了 fire 模块的多分支结构，使用较少参数又能保证精度。

5 实验

为进一步验证系统检测位姿与缺陷的实际效果，对不同光照与相对位置条件下采集的 200 张方杯图像位姿与缺陷进行检测。

对于位姿检测，使用基于直方图分割、原 Canny 边缘提取与使用改进局部阈值分割、改进亚像素 Canny 边缘检测算法定位坐标与角度。根据图 5 可以看出改进后算法检测的边缘更贴合图像边缘，因此计算的位姿也更准确。在实际情况下，由于无法得知方杯真实位姿信息，无法体现改进后算法的有效性，所以利用计算机仿真出五张中心坐标与角度已知的图像进行算法验证。仿真图像与实际采集图像具有相同灰度变化，最终边缘提取聚焦到边缘 ROI 上，将边缘 ROI 旋转一定角度，中心坐标移动一定距离，其中一张如图 16 所示。

将仿真出的图像在原先未改进与改进后的算法上验证，同时使用与前文相同的图像增强与滤波条件。5 张仿真图像在改进算法前后检测结果如图 17 所示。改进后算法对图像分割与边缘提取的准确度有所提升，使得仿真图

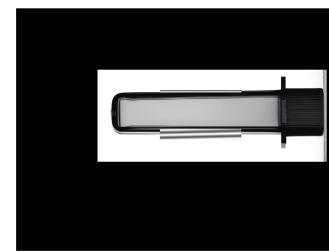


图 16 已知位姿的方杯仿真图

Fig. 16 Simulation of a square cup with known pose

实际位姿与改进算法检测的位姿更加接近。改进阈值分割算法利用 ROI 与背景灰度之差来分割图像，能够更准确分割出灰度发生多次明暗变化的方杯边缘区域。改进 Canny 算法的高阈值与阈值之差设置较小，能够检测出更多细微边缘，合并共线的轮廓后，根据这些边缘选择最外侧边缘轮廓来拟合方杯边缘。在检测时间上与原算法的检测时间大致相等，约为 770.5 ms。

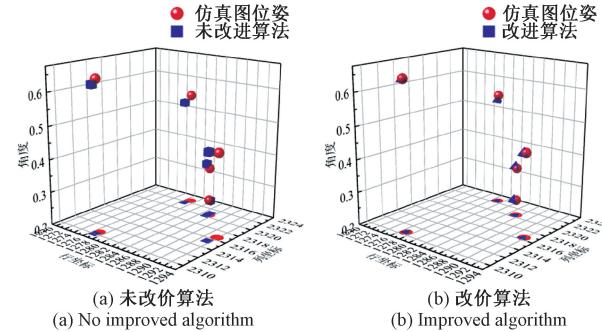
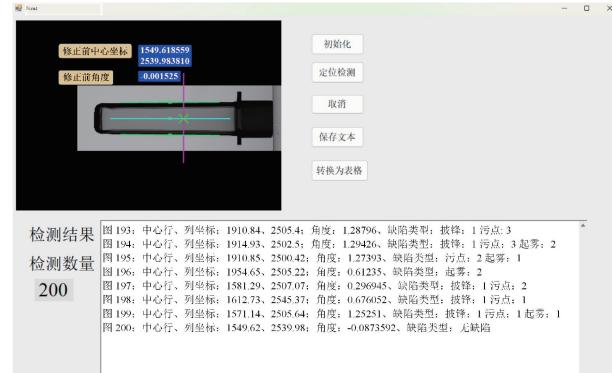


图 17 仿真方杯位姿检测结果对比

Fig. 17 Comparison of simulated square cup pose detection results

对于缺陷检测，使用 SqueezeNet 网络模型在不同光照强度、相机与方杯相对距离发生微小变化情况下验证检测精度与效率，采集新图像进行推断，平均检测一张图像需要时间 553.1 ms，精确率相对于在测试集的评估结果稍微降低，但仍然可以达到 94%。

上位机界面与部分检测效果如图 18 所示，其中图 18(a)为上位机界面，图 18(b)为位姿检测示例，图 18(c)为缺陷检测示例，缺陷检测实验结果如表 4 所示。



(a) 上位机界面
(a) Host computer interface

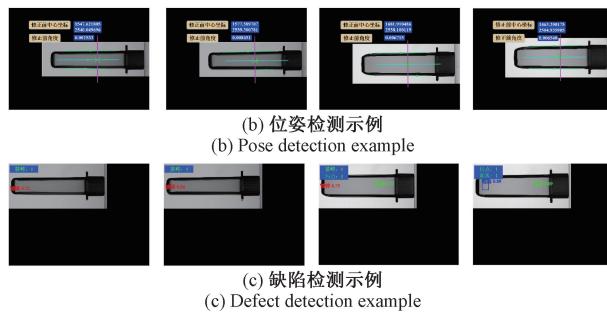


图 18 上位机界面及部分检测示例

Fig. 18 Host computer interface and some detection examples

表 4 缺陷检测结果

Table 4 Defect detection results

缺陷类别	总数	准确检测数	准确率/%
披峰	161	152	94.4
起雾	65	61	93.8
污点	134	126	94.0

总的来说,该检测系统及方法在方杯检测任务中能保证足够的精度,并且在光照与相对位置有差异的条件下,也具有较好效果,相比人工检测,可有效提高检测效率与精确率,更适合快速大批量方杯位姿与缺陷检测场景。

6 结 论

本文设计一种基于图像处理与深度学习的医用透明方杯位姿与缺陷检测系统。首先提出一种基于边缘信息的医用透明方杯位姿检测方法。基于改进均值滤波与局部动态阈值相结合的阈值分割方法,能准确分割出图像边缘区域;进一步将 Canny 算法的高阈值与阈值之差设定较低,检测出更多细微边缘,按照设定阈值的条件合并共线边缘,拟合得到两条亚像素精度边缘直线,进一步计算虚拟中轴线;最小外接矩形粗定位中心后计算到轴线的垂足作为方杯中心,定位精度有所提高,检测一张图像时间为 770.5 ms,根据仿射变换矩阵矫正图像。然后基于灰度拼接的方法自制类间均衡缺陷图像数据集,能够有效提升模型训练的精确率。最后通过在 SqueezeNet、Inception-V3 与 ResNet-50 三种神经网络上迁移学习,训练评估后,得出 SqueezeNet 网络在医用透明方杯缺陷检测任务中识别效果最佳,在测试集的评估结果中,精确率、召回率和平均精度分别达到 99.8%、98.8% 与 98.6%。训练同样 570 张数据集图像花费时间最短,约为 13 时 18 分,检测一张图像时间为 553.1 ms。

基于搭建的医用透明方杯位姿与缺陷检测系统进行实验验证,改进后位姿检测算法相比未改进检测准确度有所提升,缺陷的检测准确率可以达到 94% 左右,本系统可为医用透明方杯位姿与缺陷检测提供高效、准确的视觉检测技术支持。

参考文献

- [1] 王一博,韩巧丽,张曦文,等.基于机器视觉的风力机叶片损伤检测系统[J].科学技术与工程,2022,22(12):4879-4886.
WANG Y B, HAN Q L, ZHANG X W, et al. Blade of wind turbine based on machine vision damage detection system [J]. Science Technology and Engineering, 2022, 22(12): 4879-4886.
- [2] 增日高.基于机器视觉的多楔带缺陷检测系统设计[D].青岛:青岛科技大学,2023.
ZANG R G. Design of multi-wedge belt defect detection system based on machine vision [D]. Qingdao: Qingdao University of Science and Technology, 2023.
- [3] 罗佳睿,黄沁元,黄家雄,等.面向航空发动机叶片榫齿轮廓的尺寸测量技术[J].电子测量技术,2023,46(24):139-148.
LUO J R, HUANG Q Y, HUANG J X, et al. Dimensional measurement technology for the tenon tooth profile of aero-engine blades [J]. Electronic Measurement Technology, 2023,46(24):139-148.
- [4] 胡天圣华.基于机器视觉的动力电池缺陷检测系统设计与实现[D].杭州:杭州电子科技大学,2023.
HU T SH H. Design and implementation of power battery defect detection system based on machine vision [D]. Hangzhou: Hangzhou Dianzi University, 2023.
- [5] 王诗宇.智能化工业机器人视觉系统关键技术研究[D].沈阳:中国科学院大学(中国科学院沈阳计算技术研究所),2021.
WANG SH Y. Research on key technologies of vision system for intelligent industrial robot[D]. Shenyang: University of Chinese Academy of Sciences(Shenyang Institute of Computing Technology, Chinese Academy of Sciences), 2021.
- [6] 罗东亮,蔡雨萱,杨子豪,等.工业缺陷检测深度学习方法综述[J].中国科学:信息科学,2022,52(6):1002-1039.
LUO D L, CAI Y X, YANG Z H, et al. Survey on industrial defect detection with deep learning [J]. Scientia Sinica (Informationis), 2022, 52 (6): 1002-1039.
- [7] 罗振威,丁跃浇,甘玉坤,等.基于边缘线性拟合的芯片亚像素定位算法[J].软件,2020,41(6):204-207.
LUO ZH W, DING Y J, GAN Y K, et al. Chip sub-pixel localization algorithm based on edge linear fitting [J]. Software, 2020,41(6):204-207.
- [8] 王朕,叶文华,陈煜昊,等.复杂背景下退役圆柱锂电池

- 轮廓精确提取与位姿检测方法[J]. 电子测量与仪器学报, 2024, 38(5): 119-129.
- WANG ZH, YE W H, CHEN Y H, et al. Accurate contour extraction and pose detection method for retired cylindrical lithium batteries in complex backgrounds[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(5): 119-129.
- [9] 刘建春, 刘轩, 蒋炳翔, 等. 复杂纹理工件的位姿检测方法研究[J]. 机床与液压, 2024, 52(2): 50-56.
- LIU J CH, LIU X, JIANG B X, et al. Research on pose detection of complex texture workpiece [J]. Machine Tool & Hydraulics, 2024, 52(2): 50-56.
- [10] 刘茂霖. 面向多功能贴片机的机器视觉定位关键技术研究[D]. 杭州: 浙江大学, 2023.
- LIU M L. Research on key technologies of machine vision positioning for multi-functional placement machine[D]. Hangzhou: Zhejiang University, 2023.
- [11] 陈鹏, 陈彦秋, 彭俊, 等. 基于改进 Blob 分析的机器人批量搬运纠偏研究[J]. 仪器仪表学报, 2023, 44(8): 41-50.
- CHEN P, CHEN Y Q, PENG J, et al. Research on correction of robot batch handling based on the improved Blob analysis [J]. Chinese Journal of Scientific Instrument, 2023, 44(8): 41-50.
- [12] 王文秀, 郑鹏, 徐颖杰, 等. 基于改进 SqueezeNet 的棒状物表面缺陷识别[J]. 电子测量与仪器学报, 2023, 37(4): 240-249.
- WANG W X, ZHENG P, XU Y J, et al. Rods surface defect identification based on improved SqueezeNet[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(4): 240-249.
- [13] 郭子宁, 黄海龙, 高培根. 基于改进 SqueezeNet 模型的药片表面缺陷检测算法[J]. 机电技术, 2022(4): 2-4, 8.
- GUO Z N, HUANG H L, GAO P G. Tablet surface defect detection algorithm based on improved SqueezeNet model [J]. Mechanical & Electrical Technology, 2022(4): 2-4, 8.
- [14] 蒲秋梅, 田景龙, 邢容畅, 等. 基于改进 Inception-v3 网络的肺炎检测方法[J]. 东北师大学报(自然科学版), 2023, 55(4): 67-76.
- PU Q M, TIAN J L, XING R CH, et al. Pneumonia detection method based on improved Inception-v3 network[J]. Journal of Northeast Normal University (Natural Science Edition), 2023, 55(4): 67-76.
- [15] 常青, 郑世祺, 邓宇书, 等. 基于改进 ResNet50 和迁移学习的竹片表面缺陷检测方法[J/OL]. 控制与决策, 1-9 [2024-10-18] <https://doi.org/10.13195/j.kzyjc.2023.1783>.
- CHANG Q, ZHENG SH Q, DENG Y SH, et al. Detection method of bamboo sheets based on improved ResNet50 and transfer learning[J/OL]. Control and Decision, 1-9 [2024-10-18] <https://doi.org/10.13195/j.kzyjc.2023.1783>.
- [16] 王莉, 董鹏豪, 王瞧, 等. 基于改进 ResNet18 的干香菇等级识别[J]. 国外电子测量技术, 2024, 43(1): 117-125.
- WANG L, DONG P H, WANG Q, et al. Dried shiitake mushroom grade recognition based on improved ResNet18 [J]. Foreign Electronic Measurement Technology, 2024, 43(1): 117-125.
- [17] 刘兰, 叶芸, 郭太良. 基于拓展型特征金字塔的 OLED 像素缺陷检测方法[J]. 光学学报, 2023, 43(2): 115-123.
- LIU L, YE Y, GUO T L. Detection method for OLED pixel defects based on extended feature pyramid[J]. Acta Optica Sinica, 2023, 43(2): 115-123.
- [18] AI SH, CHEN Y T, LIU F, et al. Pill defect detection based on improved YOLOv5s network[J]. Instrumentation, 2022, 9(3): 27-36.
- [19] 王正家, 谷峰, 曾臻. 针对低像质轨道车辆车顶螺栓松动检测方法研究[J]. 机械设计与研究, 2024, 40(2): 220-224.
- WANG ZH J, GU F, ZENG ZH. Research on the detection method of loose roof bolts of low-quality rail vehicles [J]. Machine Design & Research, 2024, 40(2): 220-224.
- [20] 李冲, 王光辉, 陆志恒, 等. 基于深度学习的树种识别系统设计与试验[J]. 森林工程, 2022, 38(6): 88-95.
- LI CH, WANG G H, LU ZH H, et al. Design and experiment of tree species recognition system based on deep learning[J]. Forest Engineering, 2022, 38(6): 88-95.

作者简介

王龙达, 硕士研究生, 主要研究方向为机器视觉。

E-mail: 3263319212@qq.com

刘强(通信作者), 博士, 讲师, 主要研究方向为色散共焦精密测量技术及应用研究、机器视觉算法及应用研究。

E-mail: liuqiang@snut.edu.cn