

DOI:10.19651/j.cnki.emt.2210225

面向石英坩埚的小气泡检测算法研究*

赵谦¹ 郑超¹ 马文越¹ 尹怡晨²

(1.西安科技大学通信与信息工程学院 西安 710054; 2.西安地山视聚科技有限公司 西安 712044)

摘要: 针对石英坩埚气泡检测现有方法实时性差及小目标检测能力不足的问题,提出了一种改进YOLOv5的石英坩埚气泡检测算法YOLOv5-QCB。首先,自建石英坩埚气泡数据集,根据气泡尺寸小且分布密集的特点,减少网络下采样的深度,保留丰富的细节特征信息;同时,在颈部使用空洞卷积以增大特征图感受野,实现全局语义特征的提取;最后,在检测层前添加有效通道注意力机制,增强重要通道特征的表达能。实验结果表明,相比于原模型,改进后YOLOv5-QCB能有效降低对小气泡的漏检率,平均检测精度从96.27%提升到98.76%,权重缩减了1/2,能够实现快速、准确检测石英坩埚气泡数量。

关键词: 目标检测;石英坩埚气泡;空洞卷积;注意力机制;YOLOv5

中图分类号: TP391.41 **文献标识码:** A **国家标准学科分类代码:** 510.4

Research on small bubble detection algorithm for quartz crucible

Zhao Qian¹ Zheng Chao¹ Ma Wenyue¹ Yin Yichen²

(1. School of Communication and Information Engineering, Xi'an University of Science and Technology, Xi'an 710054, China;
2. Xi'an Dishan Vision Technology Limited Company, Xi'an 712044, China)

Abstract: To address the problems of poor real-time performance and insufficient small target detection capability of existing methods for quartz crucible bubble detection, a modified YOLOv5 algorithm for quartz crucible bubble detection, YOLOv5-QCB, is proposed. Firstly, a self-built quartz crucible bubble dataset is constructed, and based on the characteristics of small bubble size and dense distribution, the depth of network down-sampling is reduced to retain rich detailed feature information. Meanwhile, the neck using dilated convolution to increase the feature map perceptual field to achieve global semantic feature extraction. Finally, the effective channel attention mechanism is added before the detection layer to enhance the expression of important channel features. The results show that compared with original model, the improved YOLOv5-QCB can effectively reduce the missed detection rate of small bubbles, improve the average accuracy from 96.27% to 98.76%, and reduce the weight by one-half, which can achieve fast and accurate detection of quartz crucible bubble targets.

Keywords: target detection; quartz crucible bubbles; dilated convolution; attentional mechanism; YOLOv5

0 引言

石英坩埚作为拉制单晶硅一次性使用的关键材料^[1],广泛应用在太阳能电池和集成电路的制备中,对所制备产品的质量有直接的重要影响。现有技术所生产出的石英坩埚具有内壁透明层和外壁不透明层双层结构,其透明情况由气泡数量和大小引起。外壁含有大量密集气泡,使得外观呈絮状不透明,可以增加热绝缘性及提供均匀辐射的热量,内壁含有较稀疏的微小气泡,在承受长达50 h的1400℃

高温环境过程中,气泡体积逐渐增大,很容易发生破裂,将气泡中的气体和石英杂质渗入硅液中,破坏晶体结构^[2]。因此,在使用坩埚前,对其内壁透明层气泡的大小及数量检测尤为重要。

传统图像处理算法针对气泡检测一般基于边缘轮廓和圆周曲率进行拟合。文献[3]在气泡图像轮廓上筛选出所有凹点进行重叠气泡的分割,并基于最小二乘圆周拟合方法实现气泡拟合定位。文献[4]首先对气泡图像进行多尺度分解,分别在低频和高频子带上进行边界检测,最后通过

收稿日期:2022-06-06

* 基金项目:国家自然科学基金青年基金项目(51804248)、陕西省科技厅工业攻关(2022GY-115)、西安市碑林区应用技术研发项目(GX2114)资助

多尺度边界融合提取气泡,有效检测了不同尺度大小的气泡。但是传统算法受限于成像时的光照环境,以及小气泡目标本身存在特征信息较少、分辨率较低、边界轮廓不完整等因素,存在泛化能力不足、目标漏检等问题,无法满足工业检测需求。当前基于深度学习的目标检测算法主要分为两阶段目标检测算法和单阶段目标检测算法,其中单阶段目标检测算法结构简单,计算效率更高^[5]。相对于传统图像处理算法,基于深度学习的目标检测算法在大量的数据样本中利用卷积神经网络强大的特征提取能力,获得目标信息丰富的特征图,有效解决了传统算法的难点^[6],且模型具有高度模块化的特点,通过改进不同的结构以应用在工业、医学等领域中视觉测量和缺陷检测的任务。文献[7]提出了一种改进 Faster R-CNN 的电子元件气泡缺陷检测方法,检测准确率大幅度提高,但参数量较大难以部署。文献[8]提出一种改进 YOLOv3 模型对贴装元器件进行检测,重新设计特征金字塔结构,提升了小目标检测精度,但大量的推理计算降低了检测速度。文献[9]为提高输电线路小目标的检测速度,通过减少 YOLOv4 特征层的数量,大幅度提高了网络推理速度,但检测精度有所下降。

针对以上问题,为了对石英坩埚内壁实现实时有效的微小气泡检测,基于目前综合性能较好的 YOLOv5 进行网络结构改进,提出了一种 YOLOv5-QCB 的气泡检测网络。在主干网络,对输入图像经过大量的下采样操作会丢失细节特征,因此本文针对气泡目标尺寸小、大小差异小的特点,减少下采样倍数,将三个尺度预测改为两个尺度,在此基础上,使用 K-means 算法重新聚类目标先验框尺寸;同时为了弥补缺失的深层语义特征,使用空洞卷积增大特征图感受野;颈部利用有效通道注意力机制以增强重要通道特征权重,使网络自适应地去关注与任务相关的通道特征。实验结果表明,本文方法在小气泡检测的精度和速度上有着显著提升。

1 YOLOv5 算法原理

YOLOv5 算法根据网络深度和特征层通道的宽度的不同,目前有 YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x

等不同大小版本,其中 YOLOv5s 宽度和深度最小,在特定的工业检测任务方面,满足检测精度的同时,具有速度快、易于嵌入式部署的优势。其网络结构如图 1 所示。

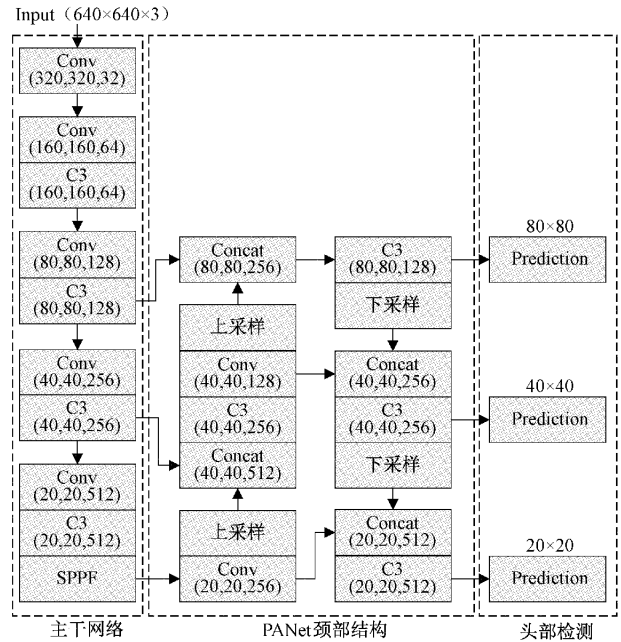


图 1 YOLOv5s 网络结构

主干网络(Backbone):以 640×640 的输入图像为例,主干网络对输入图像不断进行下采样,再通过 C3 模块进行深度的特征提取,最后使用 SPPF 模块扩大特征图感受野,获取重要的上下文信息。其中,C3 模块借鉴轻量化 CSPNet^[10] 结构的思想,可以降低网络计算量,并实现丰富梯度信息的融合,增强了网络学习能力。快速空间金字塔池化(spatial pyramid pooling fast, SPPF)是对空间金字塔池化(spatial pyramid pooling, SPP)的改进结构。改进的 SPPF 为 3 个池化核大小均为 5 的串行池化结构,相比于 SPP 使用的池化核大小分别为 5、9、13 的并行池化结构,可有效利用上层的池化输出,以更小的计算代价实现了多重感受野特征融合。SPPF 结构如图 2(a)所示,SPP 结构如图 2(b)所示。

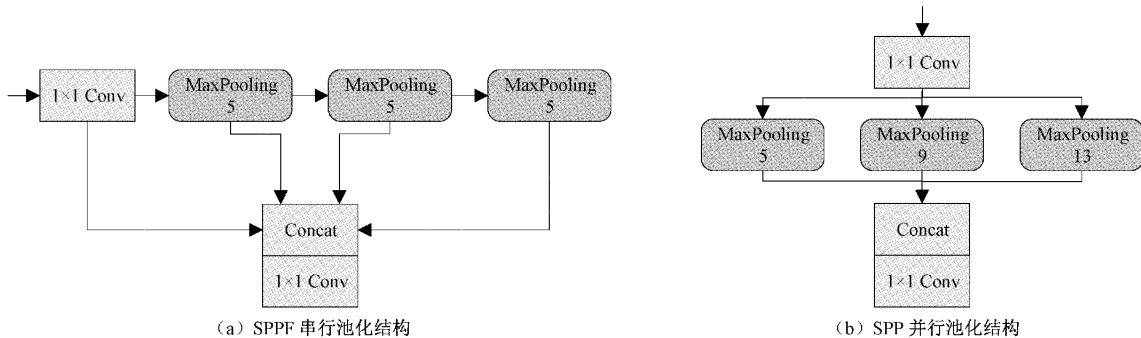


图 2 多重感受野融合结构

颈部网络(Neck):对主干网络经过 8、16、32 倍下采样获得的 80×80 、 40×40 、 20×20 三个尺度特征图通过路径融合结构(path aggregation network, PANet)^[11]进行浅层和深层的特征融合,在获得图像深层特征语义信息的同时,也融合了浅层特征细节信息。

头部检测(Head):由于浅层特征图分辨率高,预测目标框尺寸小且分布密集,能有效检测小目标;中层特征和深层特征图分辨率较低,预测目标框尺寸较大且分布稀疏,用于检测中目标和大目标。头部网络分别对各尺度特征图的每个位置预测 3 个目标框,若网络预测的类别个数为 k ,则每个目标框计算出 $5+k$ 个预测值,其中前 4 个值用于调整目标框的位置,第五个值表示目标框中含有目标的置信度。

2 本文算法

2.1 改进的 YOLOv5-QCB 网络结构

YOLOv5 网络模型作为通用的目标检测框架具有了良好的检测性能和推理速度,但由于显微镜下的图像光照差,气泡边缘轮廓模糊,检测效果并不理想。本文基于 YOLOv5 提出了改进的 YOLOv5-QCB 气泡检测网络,其整体结构如图 3 所示。

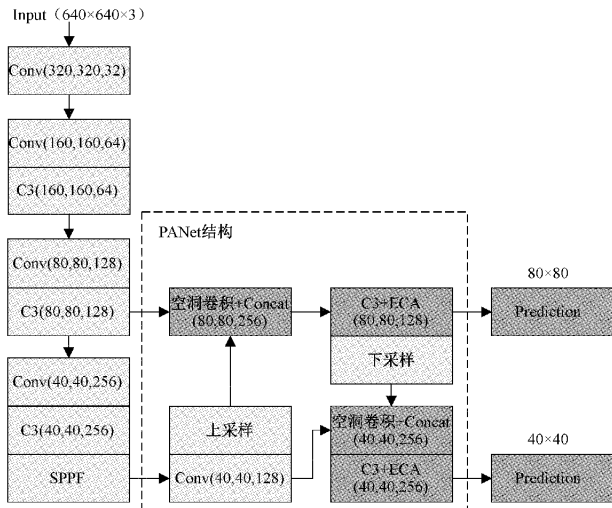


图 3 YOLOv5-QCB 结构

首先,针对数据集气泡目标尺寸小且密集的情况,删减主干网络中 32 倍下采样的相关特征层,对 8 倍下采样的 $(80, 80, 128)$ 浅层特征图和 16 倍下采样的 $(40, 40, 128)$ 中层特征图通过 PANet 网络进行特征信息融合,头部使用 40×40 和 80×80 的两个尺度特征图分别对中尺寸和小尺寸目标进行检测。

其次,由于删减了 32 倍下采样的特征层,会导致网络缺少部分深层语义特征,因此,在 PANet 结构中使用空洞卷积,通过增大卷积核感受野得到全局语义特征。

最后,由于原始头部检测网络对所有通道同等对待的

处理方式,看似公平但不合理,针对小气泡的检测的任务,增强小气泡的细节特征信息是非常重要的。因此在检测网络前,利用有效通道注意力机制,增强重要通道特征权重,使网络更加关注与任务相关的通道特征。

2.2 K-means 聚类锚框

为了提高网络目标框回归能力,使用 K-means 聚类算法在本数据集上重新聚类获得 6 组合理的先验锚框尺寸,并按大小分配给两个尺度的输出特征图,如表 1 所示。

表 1 锚框尺寸分配情况

特征图尺度	锚框尺寸		
80×80	(8,8)	(13,13)	(17,17)
40×40	(22,22)	(28,28)	(38,39)

根据 COCO 数据集对目标大小按绝对尺寸的定义,尺寸小于 32×32 的目标被认为是小目标,尺寸大于 96×96 为大目标,尺寸介于 32×32 与 96×96 为中等目标。从聚类的锚框尺寸可以看出,本数据集中无大尺寸目标,且绝大部分锚框的宽高之比为 1 : 1,符合坍塌透明层气泡特点。

2.3 空洞卷积

随着神经网络的层数加深,通常通过下采样和增加卷积核的大小来扩大特征图感受野,但是下采样会造成特征图分辨率降低,细节信息丢失,而增加卷积核大小会带来大量的参数和计算量,降低网络性能。

空洞卷积即在标准卷积的卷积核中进行零填充,在不增加计算量的情况下来扩大感受野,同时可以维持输出特征图的分辨率,避免了由于下采样导致的细节信息的丢失问题,对于小目标的检测更加有效^[12]。本文为了弥补删除深层特征图所损失的语义特征,采用卷积核大小为 3×3 、步长为 1、膨胀率为 2 的空洞卷积扩大特征图感受野,如图 4 所示。

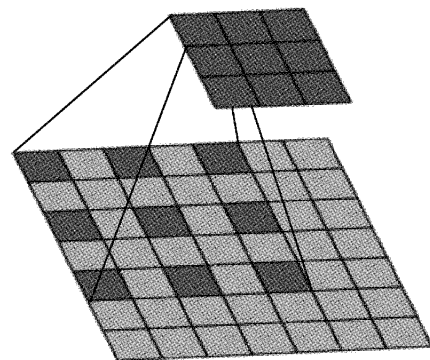


图 4 空洞卷积示意图

2.4 引入有效通道注意力机制

由于本文研究的检测对象均为小尺寸目标,特征信息较少,并且原网络作为通用模型并未针对小目标做出优化

策略,存在特征信息提取能力不足的问题。而注意力机制是在几乎不增加模型大小和计算量的前提下,通过增强重要特征信息提高检测性能。最早提出的通道注意力模块 SE^[13] (squeeze-and-excitation) 被广泛使用,但降维不可避免带来了特征信息的丢失,并且计算得到所有通道间的依赖关系是低效且不必要的。相较于 SE 模块, ECA^[14] (efficient channel attention) 注意力模块并不通过降维的方式减少计算量,在对输入特征图经过全局平均池化后,得到具有全局感受野的 $1 \times 1 \times C$ 的特征图,然后通过卷积核大小为 k 的快速一维卷积对每个通道与其 k 个近邻实现局部跨通道信息交互,其中卷积核大小 k 与输入通道数成正比,最后通过 Sigmoid 函数压缩通道权重在 $0 \sim 1$ 之间后,得到局部通道依赖关系,实现了快速有效的通道特征值加权,使网络更加关注与任务相关的通道特征信息。ECA 模块结构如图 5 所示。

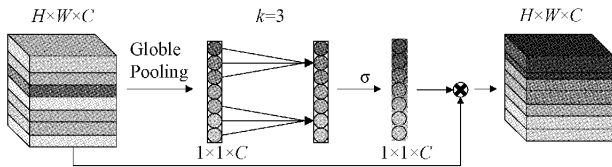


图 5 ECA 模块结构

3 实验结果与分析

3.1 实验环境及相关参数

本文实验均在 Windows 10 操作系统下运行, CPU 为 2.30 GHz 的 i7-11800H, 内存 16 GB, 显卡为 NVIDIA RTX 3050, 显存 4 GB, 使用 CUDA11.1 和 CUDNN8.1 支持 GPU 加速。Python 版本为 3.6, 深度学习框架为 Pytorch1.9.0。模型输入尺寸为 640×640 , 所有网络参数随机初始化并累积计算 200 个 epoch。

3.2 数据采集系统

实验室搭建的视频图像采集系统如图 6 所示, 该系统由主机、相机、定焦镜头等组件构成。由于石英坩埚中的气泡真实存在大小在微米级别, 普通相机无法捕捉到清晰的气泡图像, 因此采用可调节放大倍率的数码显微镜作为图像采集设备, 调节光源以模拟不同光照环境, 为方便数据采集, 使用坩埚碎片作为研究对象。

对采集的视频数据进行抽帧, 得到大量气泡图像, 最终选择 500 张图像作为石英坩埚气泡数据集, 并通过 LabelImg 手工进行标注, 总共标注有 6 190 个气泡目标, 标注类别名称记为 QCB (quartz crucible bubbles), 将手动标注的 xml 文件格式经过程序处理, 读取每个图像对应位置、目标标注的坐标和类别信息, 统一转换为 txt 格式, 以便 YOLOv5 对数据集的读取。将气泡数据集按 $8:2$ 随机划分训练集和验证集, 如表 2 所示, 训练集有图像 400 张, 包含 4 878 个气泡目标, 验证集有图像 100 张, 包含 1 339 个气泡目标。

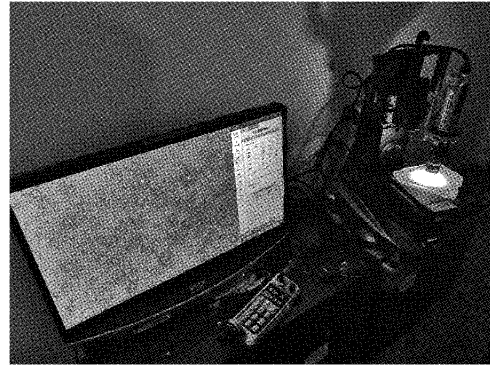


图 6 图像采集系统

表 2 训练集和测试集信息

数据集	类别	图片数量	标注框数量
训练集	QCB	400	4 878
验证集	QCB	100	1 339

3.3 评估指标

对于模型性能的评价, 通常选取检测精度 (precision, P)、召回率 (recall, R)、平均精度 (average precision, AP)、平均精度均值 (mean AP, mAP)、权重大小 (weight) 和帧率 (frames per second, FPS) 作为模型评价指标。其中检测精度用来衡量模型的查准率, 计算公式如式 (1) 所示; 召回率评估模型的查全率, 计算公式如式 (2) 所示; 单一的精度和召回率并不能全面评估网络算法的性能, 而平均精度综合了精度和召回率, 是目标检测算法最主要的性能评价指标之一, 计算公式如式 (3) 所示; mAP 表示所有类别的 AP 平均值, 当检测为单类别时, mAP 等同于 AP; 权重大小作为评估模型复杂度, 权重越小, 模型越轻量化; FPS 指网络每秒处理的图像帧数, 用于衡量网络的检测速度。

$$P = \frac{TP}{TP + FP} \quad (1)$$

$$R = \frac{TP}{TP + FN} \quad (2)$$

$$AP = \int_0^1 P(R) dR \quad (3)$$

其中, TP 表示正确检测的正样本数量, 即检测正样本为正样本, FP 表示检测错误的正样本数量, 即检测负样本为正样本, FN 表示检测错误的负样本数量, 即检测负样本为负样本。AP 表示为 P-R 曲线下的面积, 即在每一个不同的召回率下所有检测精度的平均值。

3.4 各深度特征融合对性能的影响

原始颈部网络是对 8 倍、16 倍、32 倍下采样后的 3 个不同深度特征图进行特征融合, 更利于检测大尺寸和中尺寸目标, 但对于本数据集, 过大的下采样倍数会丢失小气泡目标的细节特征。因此需探究各深度特征融合对模型检测性能的影响, 选取 4 个融合结构进行分组实验。A 表示原网络 20×20 、 40×40 、 80×80 三个尺度的特征融合结

构, B 表示 20×20 、 40×40 深层特征融合结构, C 表示 80×80 、 160×160 高分辨率特征融合结构, D 表示 40×40 、 $80 \times$

80 浅层特征融合结构。模型 A、B、C、D 结构如图 7 所示, 结果如表 3 所示。

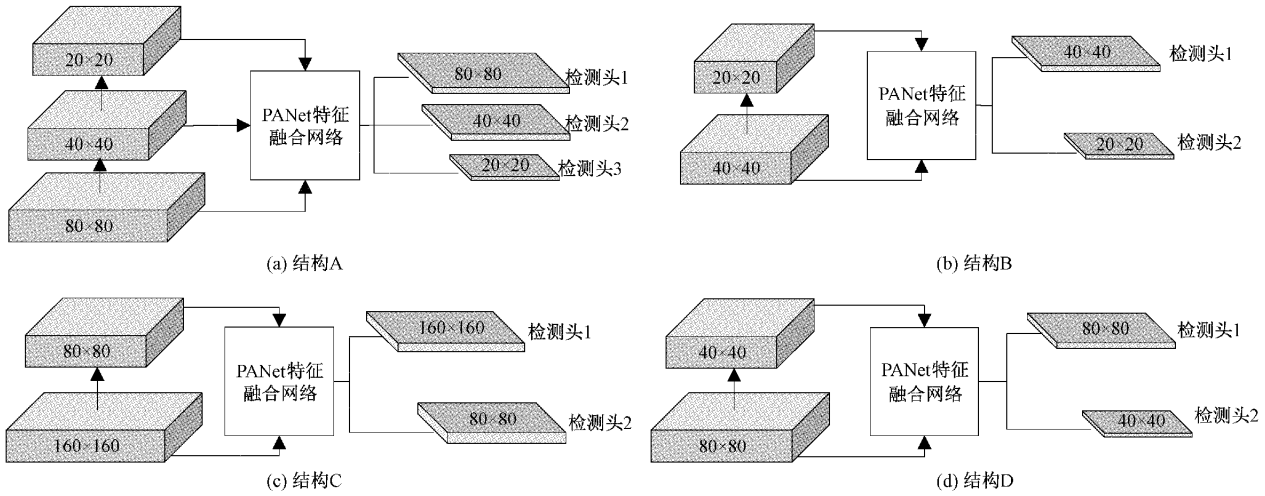


图 7 不同深度特征融合结构

表 3 不同深度特征融合结构对比

模型	精度/ %	召回率/ %	平均 精度/%	权重 大小/MB	检测 速度/fps
A	94.77	93.91	96.27	13.70	65
B	94.73	93.06	95.51	12.50	83
C	94.66	94.69	97.31	1.16	101
D	94.86	95.40	97.66	3.85	89

通过实验可知, 结构 D 性能最佳, AP 为 97.66%, 结构 C 次之, 结构 A 性能较差, 结构 B 最差。这是因为当特征层较浅时, 特征图分辨率较大, 小目标的细节信息较为丰富; 而当特征层数加深, 下采样操作会获取深层语义信息, 但丢失了大量细节信息。由于石英坩埚气泡尺寸小, 结构 B 与原结构 A 相比, 仅利用了深层语义信息, 而缺失大量的细节特征, 导致小目标精度大大降低; 结构 C 与结构 D 相比, 虽然含有丰富的细节特征, 但忽略了深层的语义特征, 检测精度较低; 而结构 D 能更好地融合浅层特征和深层特征, 避免细节特征的丢失, 小目标的精度得到提升, 相较于原结构 A 而言, 不仅精度大大提升, 而且大幅度减少了权重文件大小, 检测速度得到明显提升, 相比于结构 B 和结构 C 的模型性能, 检测速度虽不及结构 C, 但其它各项指标均为最优, 整体性能较强。因此本文针对石英坩埚气泡的检测, 采用结构 D, 删除 32 倍下采样层, 并使用 40×40 、 80×80 两个尺度的输出特征图作为头部检测。

3.5 消融实验

图 8 为 YOLOv5 基准网络和改进后 YOLOv5-QCB 模型训练的 AP 曲线图, 可以看出两者总体呈现上升趋势, 但 YOLOv5-QCB 的 AP 值基本上均高于 YOLOv5 原模型。发现在 50 轮迭代后, 均趋于平缓, YOLOv5-QCB 在 75 轮迭代处出现较大波动后趋于平稳直至收敛, 而

YOLOv5 分别在 110 和 120 轮迭代处仍出现较大波动, 最终在相同迭代轮数下, 改进后的 YOLOv5-QCB 模型的 AP 值明显高于 YOLOv5。

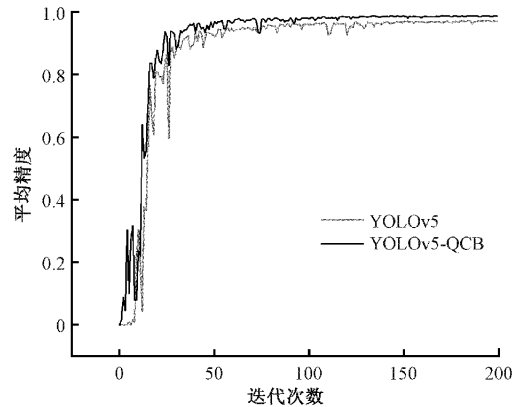


图 8 网络模型训练 AP 曲线

为证明本文改进的有效性, 通过消融实验验证改进的浅层网络结构、使用空洞卷积、添加有效通道注意力机制对性能的提升效果。结果如表 4 所示, 模型 1 使用浅层特征网络结构, 由于删减了大部分深层卷积, 大幅度降低权重文件大小的同时, 较大保留了浅层细节特征, 显著提升了召回率和平均精度。模型 2 在改进的浅层网络结构的基础上, 使用空洞卷积以增大卷积核的感受野, 提取到更加完整的有效全局语义特征信息, 相较于模型 1, 提高了模型检测精度和召回率。模型 3 为引入的有效通道注意力机制模型, 相较于模型 1, 能够自适应地增强重要通道特征, 提高了检测精度。最后模型 4 同时使用三种改进方法, 较原始 YOLOv5, 各项指标得到显著提升, AP 达到 98.76%, 提高了 2.49%, 召回率为 96.18%, 提高了 2.27%, 检测速度达到了 82 帧, 提高了 17 帧, 较模型 1、模

表 4 消融实验

模型	浅层结构	空洞卷积	ECA 模块	精度/%	召回率/%	平均精度/%	权重大小/MB	检测速度/fps
YOLOv5s				94.77	93.91	96.27	13.70	65
1	✓			94.86	95.40	97.66	3.85	89
2	✓	✓		95.35	95.96	98.38	6.11	85
3	✓		✓	95.90	94.69	98.06	3.85	87
4	✓	✓	✓	95.03	96.18	98.76	6.11	82

型 2、模型 3 不同组合的改进方法,在检测速度相差不多的情况下,召回率、平均精度有明显提升。可以看出,三种改进方法相辅相成,对小目标检测起到重要作用。

3.6 不同检测算法对比

为证明本文改进的 YOLOv5-QCB 模型的检测性能,将本文算法同 SSD^[15]、YOLOv3^[16]、YOLOv4^[17] 以及 YOLOv5s 这 4 种主流的目标检测算法在相同数据集下进行测试。实验结果如表 5 所示。

由表 5 可知,针对石英坩埚透明层气泡检测,改进模型 YOLOv5-QCB 在各方面性能指标上表现优异。本文提出的 YOLOv5-QCB 虽然精度略低于 SSD,但是在召回率、平均精度、模型大小上有着显著优势;而相对于速度和精度更加均衡的 YOLOv3、YOLOv4、YOLOv5s 经典算法,YOLOv5-QCB 不仅模型权重最小,而且在识别精度和边框回归能力上均为最优,具有更好的综合性能。

为了更直观地展示改进模型性能,对不同放大倍数下

表 5 不同检测算法性能对比

模型	精度/ %	召回率/ %	平均 精度/%	权重 大小/MB	检测 速度/fps
SSD	96.57	65.79	86.09	90.60	13
YOLOv3	88.24	89.31	92.35	235.00	14
YOLOv4	91.54	91.93	95.77	244.00	11
YOLOv5s	94.77	93.91	96.27	13.70	65
YOLOv5-QCB	95.03	96.18	98.76	6.11	82

的密集气泡图像进行测试,分别使用 YOLOv5-QCB 和性能较优的 YOLOv5s 对石英坩埚气泡进行检测,结果如图 9 所示。从图中可以看出,在不同场景中 YOLOv5-QCB 对气泡目标检测效果均为理想。在弱光照环境下,气泡图像轮廓与背景对比度低,存在大量噪声干扰的情况,YOLOv5-QCB 能够检测到 YOLOv5s 中一些漏检目标,进一步提升了模型的密集多目标检测和小目标检测能力。

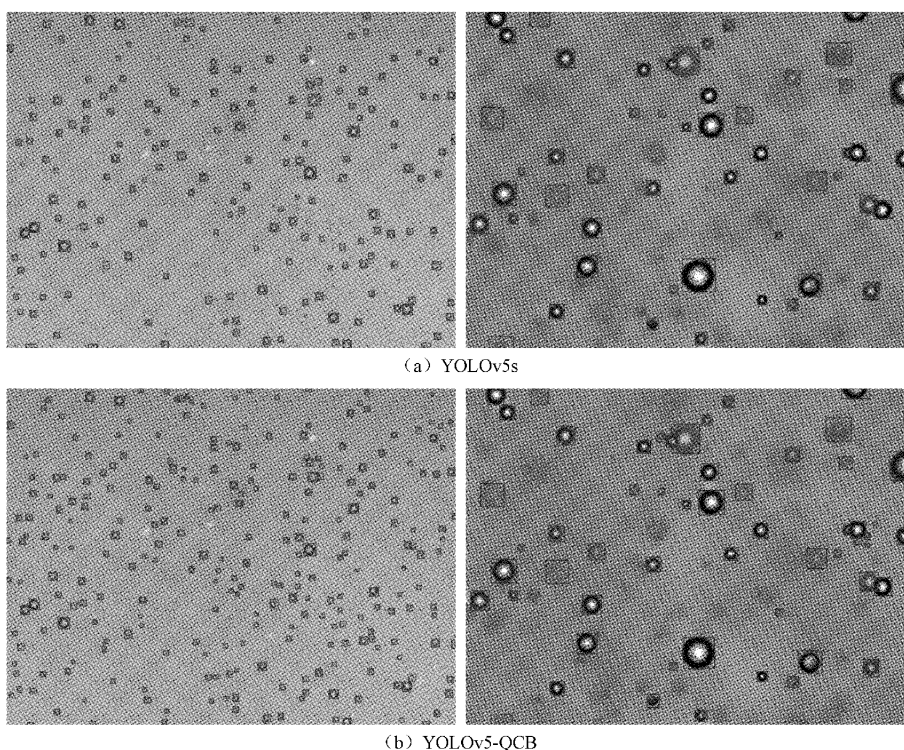


图 9 检测结果对比

4 结 论

本文基于深度学习强大的特征提取能力,完成了石英坩埚气泡的检测任务。针对坩埚透明层小气泡检测的难点,基于构建的坩埚气泡数据集进行训练和验证。首先,通过实验验证了删除 32 倍下采样层能够保留了足够的局部细节特征,提升了小目标的检测性能,同时大幅度减少了参数和计算量。其次,在颈部网络中,使用空洞卷积增大卷积核感受野,实现局部细节特征与全局语义特征的融合。最后,在输出检测层前使用有效通道注意力机制,增强重要通道特征表达能力。通过在坩埚气泡数据集上的检测结果表明,较原 YOLOv5 网络,改进后 YOLOv5-QCB 精度和召回率分别提升了 0.26% 和 2.27%,平均精度达到 98.76%,提升了 2.49%,权重大小为 6.11 MB, FPS 为 82 帧/s,在速度和精度上具有显著优势,满足工业化实时检测的需求,可为石英坩埚质量检测提供数据参考和技术支持。后续将通过研究模型压缩方法,进一步使模型轻量化,以应用在工业检测下的嵌入式开发环境。

参考文献

- [1] 张大虎,田青越,江灵,等. 单晶硅生长用石英坩埚的组成与结构特征[J]. 矿物岩石, 2016, 36(1): 12-16.
- [2] HIRSCH A, TREMPA M, KUPKA I, et al. Investigation of gas bubble growth in fused silica crucibles for silicon Czochralski crystal growth[J]. Journal of Crystal Growth, 2020, 533: 125470.
- [3] 潘志成,赵陆海波,张彪,等. 多尺度气泡尺寸分布数字图像测量方法研究[J]. 仪器仪表学报, 2019, 40(7): 129-137.
- [4] 廖一鹏,王卫星. 基于 Shearlet 多尺度边界检测及融合的浮选气泡提取[J]. 光学学报, 2018, 38(3): 351-359.
- [5] 李科岑,王晓强,林浩,等. 深度学习中的单阶段小目标检测方法综述[J]. 计算机科学与探索, 2022, 16(1): 41-58.
- [6] 邹慧海,侯进. 改进 SSD 算法的道路小目标检测研究[J]. 计算机工程, 2022, 539(5): 281-288.
- [7] 王标,周雅兰,王永红. 改进型 Faster R-CNN 网络在电子元件 LED 气泡缺陷检测中的应用[J]. 电子测量与仪器学报, 2021, 35(9): 136-143.
- [8] 窦子豪,刘新妹,殷俊龄,等. 基于 YOLO v3 的贴装元器件检测技术[J]. 电子测量技术, 2021, 44(13):

127-131.

- [9] 解尧婷,张丕状. 基于改进的 YOLOv4 输电线路小目标检测[J]. 国外电子测量技术, 2021, 40(2): 47-51.
- [10] WANG C Y, LIAO H, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN[C]. IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2020: 1571-1580.
- [11] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [12] 时永刚,李祎,周治国,等. 基于阶梯结构的 U-Net 结肠息肉分割算法[J]. 电子与信息学报, 2022, 44(1): 39-47.
- [13] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [14] WANG Q, WU B, ZHU P. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2020: 11531-11539.
- [15] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. European Conference on Computer Vision, Springer, 2016: 21-37.
- [16] REDMON J, FARHADI A. Yolov3: An incremental improvement [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018, ArXiv: 1804. 02767.
- [17] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2020, ArXiv: 2004. 10934v1.

作者简介

赵谦,博士,副教授,主要研究方向为图像处理、机器视觉、工业无损检测等。

郑超(通信作者),硕士研究生,主要研究方向为图像处理、目标检测。

E-mail:1056389866@qq.com

马文越,硕士研究生,主要研究方向为图像处理、交通灯配时。

尹怡晨,硕士研究生,主要研究方向为图像分割。