

基于 YOLOv5s 改进的机油瓶表面缺陷检测方法

李文辰, 张亚慧

(上海第二工业大学 智能制造与控制工程学院, 上海 201209)

摘要: 针对机油瓶表面瑕疵点目标过小、ROI 候选框鲁棒性较差导致液位线定位不准等问题, 提出一种改进 YOLOv5s 的缺陷检测算法; 利用 K-Means++ 代替 K-Means 进行聚类中心的初始化, 使得生成的先验框更加接近检测目标的真实形状和大小; 同时在主干网络中引入可变形卷积, 提高特征提取的灵活性, 并引入 SE 注意力机制, 对特征图不同通道进行权重调整; 此外, 在颈部网络中使用 BiFPN 代替原有的 PANet, 实现对不同尺度信息的自适应特征融合; 实验结果表明, 改进的 YOLOv5s 算法 mAP 达到了 96.9%, 较 YOLOv5s 算法提升了 6%, 准确率提升了 4%; 实验验证了改进后的 YOLOv5s 在检测准确率方面优于原始的 YOLOv5s 算法, 解决了小目标漏检及液位线定位存在偏差的问题。

关键词: 机油瓶; YOLOv5s; 缺陷检测; 可变形卷积; 注意力模块

Improving the Defect Detection Method of YOLOv5s for Oil Bottle Surfaces

LI Wenchen, ZHANG Yahui

(School of Intelligent Manufacturing and Control Engineering, Shanghai Second Polytechnic University, Shanghai 201209, China)

Abstract: To address the issues of small target size of surface defects in engine oil bottles and poor robustness in ROI candidate boxes leading to inaccurate liquid level positioning, an improved defect detection algorithm for YOLOv5s is proposed. Firstly, K-Means++ is used to initialize the clustering centers instead of K-Means, making the generated prior boxes closer to the real shape and size of the detection targets. Then, the variable convolution is introduced into the backbone network to improve the flexibility of feature extraction. And then, the SE attention mechanism is introduced to adjust the weights of different channels in the feature maps. Additionally, the bi-directional feature pyramid network (BiFPN) is used in the neck network instead of the original path aggregation network (PANet) to achieve adaptive feature fusion of different scale information. Experimental results show that the improved YOLOv5s algorithm achieves an mAP of 96.9%, which is 6% higher than that of the YOLOv5s algorithm, and the accuracy is increased by 4%. Experimental results verify that the improved YOLOv5s algorithm outperforms the original YOLOv5s algorithm in terms of detection accuracy, solving the problems of missing small targets and inaccurate liquid level positioning.

Keywords: oil bottles; YOLOv5s; defect detection; variable convolution; attention module

0 引言

机油瓶作为典型的注塑工艺制品, 承载着储存和运输机油的重要任务, 是机械行业不可或缺的零部件之一。在机油瓶生产过程中, 成品机油瓶往往会存在一些

缺陷, 如液位线不完整、瓶口变形、瓶身凹陷、合模线密封不良、瓶身瑕疵点等, 若不及时检测和修复, 成品机油瓶可能在充装机油后出现泄露或爆炸等安全问题。此外, 将存在缺陷的机油瓶销售到市场还会损害消费者对品牌的信任度。因此, 对机油瓶缺陷进行检测有一定

收稿日期:2024-01-23; 修回日期:2024-03-01。

作者简介:李文辰(1989-),男,博士,讲师。

通讯作者:张亚慧(1998-),女,硕士研究生。

引用格式:李文辰,张亚慧. 基于 YOLOv5s 改进的机油瓶表面缺陷检测方法[J]. 计算机测量与控制,2025,33(4):32-39.

的现实意义。

最近十几年, 机器视觉技术逐渐产业化, 其在实时在线检测中的应用逐渐成为常态。许多研究者提出了一系列基于图像处理技术的产品外观缺陷检测算法或系统。如张鲲等人^[1]提出一种传统图像处理结合深度学习的方法, 采用图像滤波与灰度变换对图像进行预处理, 再将图像传递给卷积神经网络处理, 提升了柔性印制电路板的缺陷检测准确率, 然而该方法对光照强度有一定的依赖性。孟陆等人^[2]通过对 MapReduce 存储结构的改进, 解决大量小文件存储导致效率低下的问题, 且完成基于 MapReduce 的并行化玻璃缺陷图像阈值分割算法。龙智帆等人^[3]运用多线程并行处理技术, 针对瓶盖中常见的高盖和斜盖缺陷搭建视觉缺陷检测系统, 完成了对瓶盖缺陷的检测, 但产线传输的不稳定会导致误检。麦国铭^[4]以 LabVIEW 和 Matlab 混合编程的方式, 提出了基于连通区域个数的反馈式阈值算法, 实现了塑料制品缺陷的完整分割, 然而该方法的检测速度有待提高。

随着深度学习的不断发展, 越来越多的深度学习模型被应用到目标检测领域。王莉等人^[5]提出了基于 YOLOv5s 进行垃圾分类和预测, 通过提取不同种类垃圾图像的特征和位置信息, 实现垃圾的分类和预测, 但不同角度的拍摄对同一目标的检测存在较大影响。许思昂等人^[6]提出基于改进的 YOLOv5 进行 PCB 裸板缺陷检测, 通过对特征融合网络的改进及 SE 机制的引入提高检测准确率, 但参数量有所提升。朱豪等人^[7]针对木材表面缺陷的复杂多样性和特征提取困难的问题, 在原模型的基础上引入了坐标注意力机制和混合空间金字塔池化结构。通过改进的曲线高效交叉联合作为训练函数, 对模型进行提升, 但模型的召回率有所降低。乔炎等人^[8]针对当前目标检测模型存在结构复杂、计算量大、检测准确率低等问题, 提出了一种改进型 YOLOv5 安全帽佩戴算法。通过在 YOLOv5 中引入轻量级网络 ShuffleNetv2 和 CBAM_H 注意力机制, 以及对 C3 模块进行优化, 实现了更轻量级的运行, 但对较远的目标会存在漏检, 对特殊姿势和特别形态的物体存在部分错检的情况。

在上述改进算法的基础上, 本文针对机油瓶瓶身上的液位线及瑕疵点的检测提出以下改进: 通过使用 K-Means++ 初始化聚类中心, 以获得更好的先验框适应不同大小的目标检测, 并在主干网络引入 DCNv2, 对卷积核在空间上进行微小采样位置的改变来适应不同目标的形状和结构。在主干网络的最后一层引入 SE 注意力机制, 使其能够自适应地加强或减弱特征图的不同通道之间的影响。同时在颈部网络引入 BiFPN 网络, 加

强对图像特征的融合。通过改进算法来提高产品缺陷检测的准确性和可靠性, 以满足产品的安全性和美观度的双重要求。

1 YOLOv5s 算法

YOLOv5s 是一种单阶段目标检测算法^[9], 具有网络结构简单明了、拓展性强、场景适应性广等特点, 在工业缺陷检测领域表现出色, 因此本文选择采用 YOLOv5s 模型作为机油瓶表面缺陷检测任务的基本框架。该算法主要由输入端、主干网络、颈部网络和头部网络 4 个部分组成^[10], 分别负责输入图像的处理、特征提取、特征融合和检测结果输出, 其网络结构如图 1 所示。

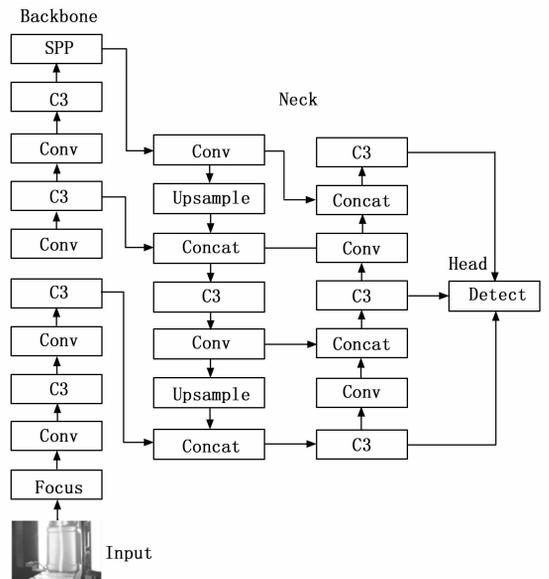


图 1 YOLOv5s 网络结构图

输入端主要用来对图像进行一些预处理、数据增强等。主干网络是提取输入图像信息特征的部分, 使用 CSPDarknet53 网络, 主要包含 Focus、Conv、C3、SPP 等模块^[11]。Focus 模块通过对输入图像进行切片操作使网络能够聚焦于图像的不同部分, 从而在扩大网络感受野的同时能够在更细粒度的层面上获取图像信息。Conv 模块用于提取局部空间信息, 其通过卷积操作, 能够有效地捕捉到输入数据中的局部特征, 进一步增强网络的识别能力。为了提高感受野, 网络采用了 C3 模块, 该模块通过增加网络的深度和复杂度, 能够扩大网络对目标的感知范围, 提高目标检测的准确性。SPP 模块则主要用来将不同尺寸的特征图转换为固定大小的特征向量, 来减少信息损失, 各个模块的网络结构如图 2 所示。

颈部网络层采用了 PANet 特征融合模块, 它通过增加路径聚合来捕捉不同尺度的目标信息, 通过融合来

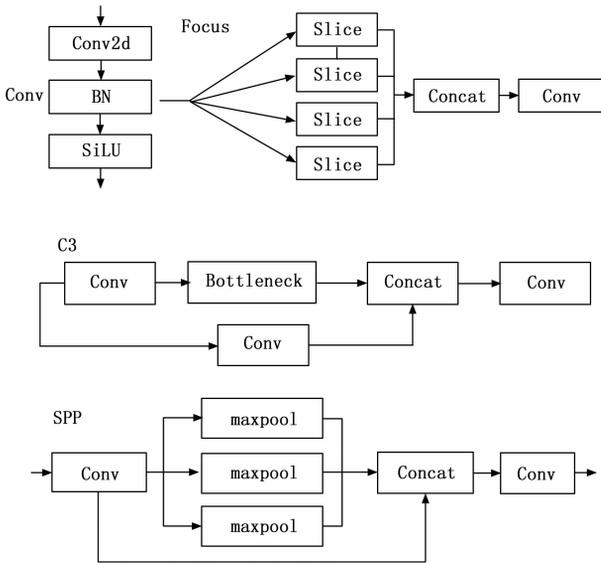


图 2 各模块网络结构图

随机选择的优势^[13]。K-Means++工作流程如图 3 所示。

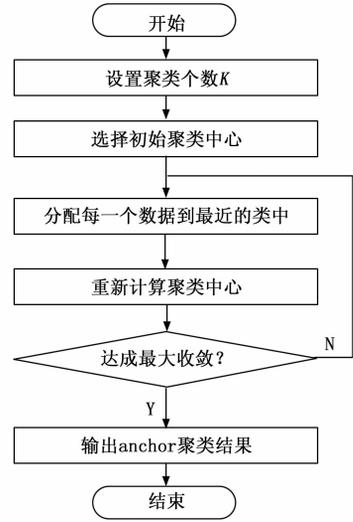


图 3 K-Means++工作流程图

自不同层级的目标信息，使模型更好地学习图像内容，增强模型的识别和定位方面的表现。

头部网络通过利用主干网络提取的特征和颈部网络的特征融合的处理，对图片进行目标类别预测及结果输出。

2 改进的 YOLOv5s 算法

2.1 K-Means++ 算法

在 YOLOv5s 中，采用了 K-Means 算法对检测数据集进行聚类。以训练集的边界为基准，将 3 种不同尺寸的特征映射设置为 3 个锚框，将锚框作为先验框辅助预测目标尺寸。K-Means 算法在初始聚类时的聚类中心是随机选取的，这可能会导致初始聚类中心与最优聚类中心位置相距甚远，不仅会影响模型的收敛速度，还会导致检测效果不佳^[12]。此外，YOLOv5s 的预设锚框是基于 COCO 数据集生成的，这些锚框的大小未必适用于本文自主构建的数据集。因此，在训练过程中，选择符合数据集中真实标记框大小的锚框是实现训练收敛的关键^[13]。

因此，本文采用 K-Means++ 算法来替代传统的 K-Means 算法。K-Means++ 算法通过改进初始聚类中心的选择方式，提高了聚类结果的稳定性和准确性。它利用特定的概率选择方式来选取初始聚类中心，从而避免了传统 K-Means 算法中可能出现的不稳定和不准确的聚类结果。在目标检测中，锚框数量通常很多，使用 K-Means++ 算法可以提高聚类的效果和效率，从而更准确地预测目标的位置和大小，并提高目标检测的准确率和效率。相比于 K-Means 聚类算法，K-Means++ 聚类算法具有减少初始聚类中心选择依赖和优化聚类中心

2.2 可变形卷积

针对机油瓶瓶身表面的不规则形状、大小不固定以及位置偏移等瑕疵点特征，传统的固定位置卷积操作可能无法完全捕捉到这些复杂特征。为了更准确地捕捉这些不规则瑕疵的特征，本文引入了 DCNv2 模块。它引入了 DSC (Depthwise Separable Convolution)，允许卷积核相对于其标准位置进行微小的偏移，并引入了 Modulated DCN，以更好地适应目标在图像中的真实位置、尺寸和形状变化^[14]。

传统的卷积操作在特征提取时使用固定位置的卷积核，这些卷积核在预定义位置对输入特征图进行采样，然后通过权重求和生成输出特征图。以一个二维卷积核为例^[15]，定义为 R ，大小为 3×3 。对于输出特征映射 y 的每个位置 p_0 ，如式 (1) 所示：

$$y(p_0) = \sum p_{n \in R} \omega(p_n) \cdot x(p_0 + p_n) \quad (1)$$

其中： $\omega(p_n)$ 是卷积核权重， $x(p_0 + p_n)$ 是输入特征图。

不同于传统方法，DCNv2 引入了一种可变形卷积，允许卷积核在不同位置上进行采样，可表示为以下形式：

$$y(p_0) = \sum p_{n \in R} \omega(p_n) \cdot x(p_0 + p_n + \Delta p_n) \cdot \Delta m_n \quad (2)$$

其中： Δp_n 表示根据目标特征分布情况计算的偏移量， Δm_n 为权重系数，其余变量与传统卷积操作中的变量相同。偏移量 Δp_n 为输入特征图被应用到单独的卷积层产生的， Δm_n 的范围是 $[0, 1]$ ，由于 $p_0 + p_n + \Delta p_n$ 是小数，按照文献 [16] 所述，利用双线性插值法进行 $x(p_0 + p_n + \Delta p_n)$ 的计算。因此，可变形卷积的公式就

转化如下:

$$X(p) = \sum_q G(q, p) \cdot x(q) \quad (3)$$

其中: q 为经过偏移调整后的采样点的位置, p 为整数网格点。 $G(q, p)$ 为通过双线性插值得到的采样点位置。可变形卷积的结构如图 4 所示。

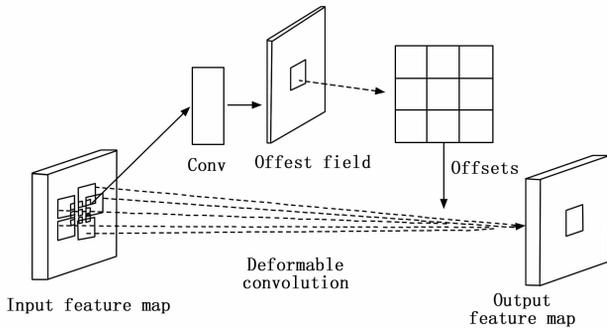


图 4 可变形卷积结构图

2.3 SE 注意力模块

由于机油瓶在拍摄时面临工业环境的不稳定性、传感器触发的不确定性以及不同的光照强度等背景因素, 导致拍摄的图像存在一定的差异, 给液位线的定位及瑕疵点的识别带来一定的问题。传统的液位线定位方法通常采用固定的 ROI 候选框截取, 但对于不同形状和大小的机油瓶, 选择合适的 ROI 候选框需要考虑机油瓶瓶身的几何形状和液位线的相对位置。如果 ROI 候选框过小或过大, 会出现定位不准等问题, 可能无法准确地检测到液位线的完整信息。而瓶身瑕疵点检测则依靠瑕疵点与瓶身颜色的差异进行检测, 但瑕疵点过小且仅出现在瓶身的局部区域, 极易造成因光照原因导致的漏检。

基于上述原因, 本文引入 SE 注意力机制, 其将输入特征图中的每个通道表示为一个向量, 并将这些向量传递给一个注意力层, 在该层中, 每个通道的向量都会被乘以一个权重, 然后将所有乘积的结果进行加权求和, 得到每个通道的增强特征, 并将这些增强特征重新注入网络中, 增强模型对图像有用信息的提取。然而, 若引入全局注意力机制可能会带来参数增加的问题。为了保持网络轻量级, 因此将 SE 模块限制在主干网络的最后一层。在增强局部特征学习的同时, 控制了模型的参数数量, 以确保计算资源的有效利用和实际应用的可行性。

SE 机制主要通过以下步骤进行: 首先, 对输入特征图的维度为 $H \times W \times C$ 进行全局均值池化, 对特征图在空间维度进行降维, 得到一个 $1 \times 1 \times C$ 的全局感受野, 该方式^[17] 允许模型跨越整个特征图, 捕捉到跨位置的特征相关性。其次, 通过全连接层和 Sigmoid 激活

函数, 获取通道之间的相关性信息, 进而构建模型, 将得到的结果作为权重, 与输入特征进行相乘。

在上述步骤中存在两个关键步骤^[18]: 挤压 (Squeeze) 和激励 (Excitation)。挤压负责将全局信息嵌入, 而激励则用于根据通道关系进行自适应的重新校准。给定输入 $X = [x_1, x_2, \dots, x_c] \in R^{C \times H \times W}$, 第 c 个通道的挤压步骤可表述如下:

$$z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j) \quad (4)$$

其中: z_c 是与第 c 个通道相关的输出。挤压操作对主干网络输出的 $C \times H \times W$ 输入进行降维压缩, 使得输出的维度与特征通道数相匹配, 并且从全局范围内捕捉了信息, 从而形成全局感受野。激励可表述为:

$$\hat{X} = X \cdot \sigma\{T_2[\text{ReLU}(T_1(z))]\} \quad (5)$$

其中: \hat{X} 为通道乘法, σ 为 sigmoid 函数, T_1 和 T_2 为两个线性变换, 激励操作利用通过全连接层计算得到的通道权重, 对特征通道进行加权, 弱化不重要通道的信息。通过将 SE 模块集成到主干网络的最后一层, 增强网络对于背景复杂的机油瓶图像中液位线、瑕疵点缺陷特征的提取能力。SE 通道注意力模块如图 5 所示。

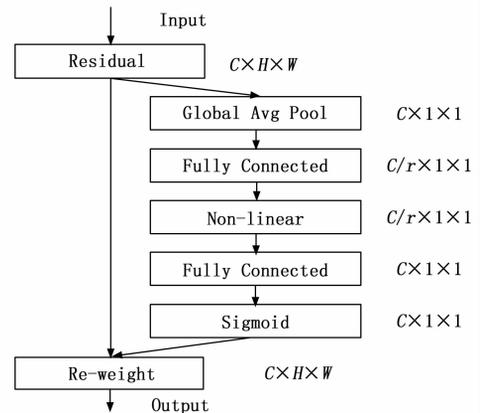


图 5 SE 通道注意力模块结构

2.4 BiFPN 双向特征金字塔网络

目标检测中的一个主要难点是有效地表示和处理多尺度特征。在早期工作中, STDL (Scale-Transferrable Object Detection with Long Short-Term Dependency)^[19] 提出了一个尺度转换模块, 该模块能有效地提取跨尺度特征; 特征金字塔网络 FPN (Feature Pyramid Network) 提出了一种自顶向下的路径来组合多尺度特征^[20], 在这个思路的基础上, PANet (Path Aggregation Network) 在 FPN 之上添加了一个自底向上的路径聚合网络^[21]。尽管 PANet 在一定程度上改善了多尺度特征融合的效果, 但它也存在一些缺点和问题, PANet 的自底向上路径聚合只依赖于从上层到下层的信息传

递, 缺乏跨尺度的信息交流, 可能会导致低分辨率的特征图被忽略或过度压缩, 导致信息在不同层级之间传播不均衡。

随着神经网络层次的加深, 不同网络层级所产生的特征呈现明显差异。为了提高模型对微小瑕疵点等小目标检测的准确性, 并从高层级特征中获取更多细节信息, 因此本文在 YOLOv5s 颈部网络中采用 BiFPN (Bi-directional Feature Pyramid Network) 结构^[22]代替 PANet 结构, 其三者的网络结构对比如图 6 所示。BiFPN 引入了加权特征融合的机制, 可以根据特征图的重要性对不同尺度的特征图进行加权融合, 从而提高了特征图的质量, 其结构中每个层级的特征图都能够自适应地聚合来自上下两个方向的特征信息, 实现高效的双向跨尺度连接与加权特征图融合, 便于适应任务需求, 提高模型的可适应性。

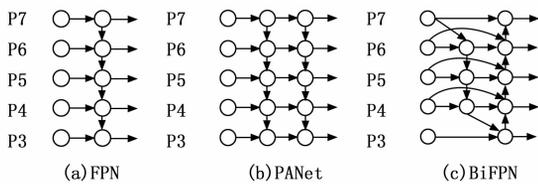


图 6 网络结构对比图

3 实验与分析

3.1 实验环境

YOLOv5s 模型训练实验环境配置如表 1 所示。

表 1 实验环境配置

配置对象	配置内容
操作系统	Windows11 专业版
处理器	NVIDIA GeForce RTX 3060 显卡
开发环境	PyCharm
深度学习框架	Pytorch 1.13.1
CUDA	CUDA 11.7
Optimizer	SGD
相机	BasleracA1600-60gc 工业相机

3.2 数据集

本文使用机油瓶图像数据集进行训练, 该数据集有 1 930 幅图像, 大小为 1 600 × 1 200, 其均在实际工业生产环境中获取。在训练中, 数据集按照 7 : 2 : 1 的比例随机分为训练集 1 351 张、训练集 386 张、验证集 193 张图像, 且使用标注软件 LabelImg 以 YOLO 格式标注数据集的缺陷位置和类别, 共有两种标注类别, 分别是瑕疵点和液位线区域, 标注以 PASCAL VOC 格式保存为 XML 文件。在标注过程中瓶身表面的瑕疵点及液位线区域以矩形框形式标注, 其反映了检测目标的坐标和类别。机油瓶液位线、瑕疵点如图 7 所示。

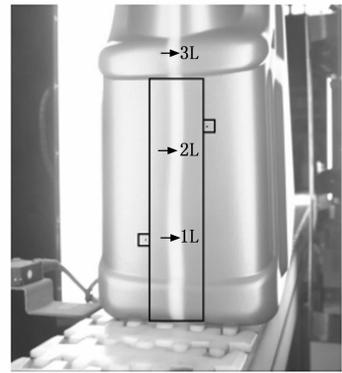


图 7 机油瓶瓶身图像样例

3.3 评估指标

本文采用 4 种常见的评估指标: 准确率 *Precision*、平均精度均值 *mAP* (mean Average Precision)、召回率 *Recall* 和 F_1 。其中, 准确率是指在所有被模型识别为正样本中, 真正为正样本的占比, 所以用预测对的正样本数除以模型预测的正样本总数。*mAP* 值代表模型在多个类别上的平均值, 是对整个目标检测系统性能的一个总体评价。*mAP* 值越大, 算法对目标的检测和识别能力越强, 算法的性能也越好。召回率是指在所有实际为正样本的样本中, 被模型正确识别为正样本所占的比例, 所以用预测为正样本的样本数除以实际共有的正样本总数。 F_1 是准确率和召回率的调和平均数, 值越大表明模型性能越好。上述每种指标的计算公式如下:

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

$$AP = \int_0^1 PdR \quad (7)$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (8)$$

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

$$F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (10)$$

式中, *TP* 表示本来是正样本且模型同样识别为正样本的数量, *FP* 表示本来是负样本但模型识别为正样本的数量, *FN* 表示本来是正样本但模型识别为负样本的数量。*mAP* 是通过对所有类别的 *AP* 进行平均得到的。*N* 表示检测的类别总数。

3.4 算法有效性分析

3.4.1 K-Means++ 聚类实验结果

为了验证使用 K-Means++ 初始化聚类中心对后续模型训练的影响, 本文在原始 YOLOv5s 算法基础上, 利用 K-Means++ 聚类方法取代传统 K-Means 聚类方

法进行实验。此外,为更好地评估这一改进,本文在百度飞桨平台上选用了 3 个公开数据集,分别为 Fire-Detection、Helmet-Detection 以及 Mask-Detection,上述 3 种数据集检测对象皆为不同大小的目标,并且这些目标出现位置随机。由于这些特征与本文研究中机油瓶瓶身前侧的瑕疵点特征相类似,因此引入上述 3 个数据集来验证模型性能。首先利用引入的 K-Means++ 算法对其进行实验,具体实验结果如表 2 所示。

从表 2 中可以看到,在 Fire-Detection 数据集上,原始 YOLOv5s 算法的 mAP 为 76.3%,而引入 K-Means++ 机制后算法的 mAP 达到了 79.7%。对于 Helmet-Detection 数据集,传统 YOLOv5s 算法获得 89% 的 mAP ,而引入 K-Means++ 的改进算法则提升至 94.9%。在 Mask-Detection 数据集上,原始 YOLOv5s 算法的准确率为 92.2%,而引入 K-Means++ 机制的改进算法上升至 95.4%。上述 3 个数据集直观地展示了在不同数据集中引入 K-Means++ 的实验结果以及其在不同数据集上的一致表现,因此本文引入了 K-Means++ 来初始化聚类中心,为后续的模型训练提供良好的初始条件。

3.4.2 可变形卷积实验结果

为了验证引入 DCNv2 的性能,在上述 3 个公开数据集上进行测试,评估 DCNv2 在不同数据集下的性能,实验结果如表 3 所示。

由表 3 可知,引入 DCNv2 后的模型在 3 个公开数据集上的性能都有所提升,尤其在火灾检测数据集中,模型的准确率从 81.8% 增长至 91.6%,同时 mAP 值也增加了 7.3%。在安全帽检测数据集中,模型的检测准确率提高了 4%。此外,在口罩检测数据集中,准确率也得到了显著提升,达到了 94.3%。由以上结果来看,引入 DCNv2 模块对于原始的 YOLOv5s 模型在目标识别方面产生了一定的积极作用。

3.4.3 SE 注意力机制实验结果

为了比较不同的注意力模块在本文数据集检测任务

中的效果并且找出适合的注意力模块,本文在 YOLOv5s 算法中引入了多个不同的注意力模块,并进行了一系列实验来进行比较。在实验设置中选取了几个常用的注意力模块,以便在本文所使用的数据集上评估它们的效果。SE 模块被认为是通道注意力模块的经典代表,通过学习通道之间的关联关系,实现对重要特征的强化。CBAM 模块融合了通道注意力和空间注意力,被认为是融合型注意力模块的代表之一。NAM 模块是一种轻量级的融合型注意力模块,它建立在 CBAM 模块的基础之上,能够在保持高效性的同时融合通道和空间信息。CA 模块着重关注特征图中通道之间的关联性,通过动态调整通道权重,提升了模型对关键特征的感知能力。表 4 详细呈现了每个注意力模块在本文数据集上的实验结果。

由表 4 可知,引入不同的注意力模块到 YOLOv5s 模型中,对在本文所使用的数据集的目标检测任务都产生了不同程度的提升。具体来说,在 YOLOv5s 中引入 NAM 模块使得实验指标都有所提升,在一定程度上提高了检测性能。模型引入 CA、CBAM 模块后更是在准确率和 F_1 分数方面表现良好。在引入 SE 模块以后模型检测性能更是取得了大幅度提升, mAP 值从原先的 90.9% 提升到 97.6%, F_1 分数值从原先的 93% 提升到 96.8%,有效地提升了模型的性能。综合各个模块的表现,因此针对本文所涉及的数据集识别任务选用 SE 注意力模块进行缺陷检测。

3.4.4 双向特征金字塔网络实验结果

为了验证在 YOLOv5s 中引入 BiFPN 前后的效果差异,在本文数据集上进行测试,实验结果如表 5 所示。

从表 5 中可以清晰地看出,引入 BiFPN 框架后,模型识别准确率由原先的 92.6% 提升至 96.6%,召回率由原先的 93.2% 提升至 97.2%,该结果也进一步体现了 BiFPN 对于特征融合的有效性。

3.5 消融实验

为了进一步观察在 YOLOv5s 模型中引进各个模块

表 2 引入 K-Means++ 算法与原始 YOLOv5s 的检测结果比较

Algorithm	Fire-Detection			Helmet-Detection			Mask-Detection		
	mAP	P	R	mAP	P	R	mAP	P	R
Yolov5s	76.3	81.8	73.8	89	92.6	81	96.4	92.2	92.6
Yolov5+K-Means++	79.7	90.5	75.3	94.9	95	90.8	96.3	95.4	88.7

表 3 不同数据集引入 DCNv2 模块结果对比

Algorithm	Fire-Detection			Helmet-Detection			Mask-Detection		
	mAP	P	R	mAP	P	R	mAP	P	R
YOLOv5s	76.3	81.8	73.8	89	92.6	81	96.4	92.2	92.6
YOLOv5s+DCNv2	83.6	91.6	74.1	92.1	96.6	86.1	95.5	94.3	92

表 4 引入不同注意力机制检测结果对比 %

Algorithm	<i>mAP</i>	<i>P</i>	<i>R</i>	<i>F₁</i>
YOLOv5s	90.9	92.6	93.2	93
YOLOv5s+NAM	92.2	94.2	94.6	94.4
YOLOv5s+CA	93.2	95.2	94.4	94.8
YOLOv5s+CBAM	92.7	94	94.6	94.3
YOLOv5s+SE	97.6	97.6	95.9	96.7

表 5 引入 BiFPN 前后检测结果对比 %

Algorithm	<i>mAP</i>	<i>P</i>	<i>R</i>	<i>F₁</i>
YOLOv5s	90.9	92.6	93.2	93
YOLOv5s+BiFPN	96.3	96.6	97.2	96.9

的效果,在本文机油瓶数据集的基础上进行一系列实验,并对结果进行了比较。图 8 展示了引入相应模块后随着不断迭代,相应的损失函数的下降趋势。由图 8 可知,通过引入上述模块,模型训练的收敛速度得到了提升,并且较快地呈现稳定趋势。

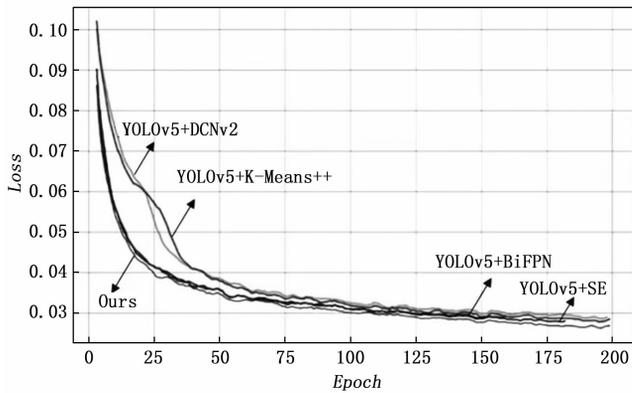


图 8 损失曲线对比图

为了更清晰地比较引入各个模块后的模型相关数值,评估不同改进方法对模型的影响,列出了如表 6 所示的对比结果。

表 6 引入不同模块效果实验对比 %

Algorithm	<i>mAP</i>	<i>P</i>	<i>R</i>	<i>F₁</i>
YOLOV5s	90.9	92.6	93.2	93
YOLOv5s+DCNv2	94.9	94	95.8	94.9
YOLOv5s+K-Means++	95.4	95.2	95.9	95.5
YOLOv5s+BiFPN	96.3	96.6	97.2	96.9
YOLOv5s+SE	97.6	97.6	95.9	96.7
Improved YOLOv5s	96.9	96.6	97.3	97

从表 6 中可以看出,当单独引入 DCNv2 模块,模型的目标检测能力有一定程度的增强,其 *mAP* 值达到了 94.9%。引入 BiFPN 模块进一步增强了模型的特征融合能力,使得模型的召回率提升至 97.2%。而引入 SE 注意力模块则使得模型具有较高的召回率和准确率,

对特征的提取起到了关键作用。改进后的 YOLOv5s 算法相较于原始 YOLOv5s 算法,其漏检率有一定程度的下降,算法检测性能有所提升。

3.6 实验分析

为了验证本文模型的有效性,将本文算法、文献 [7] 的改进算法与其他算法在本文数据集上进行训练与测试,并应用上述评价指标进行模型性能的对比,对比结果如表 7 所示。从表 7 可以看出,本文所提出的改进 YOLOv5s 后的模型的召回率达到了 97.3%,比原始 YOLOv5s 提升了 4.1%。相比其他算法,改进后的算法在检测准确率和召回率上均有所优化。

表 7 不同目标检测算法对比实验 %

Algorithm	<i>mAP</i>	<i>P</i>	<i>R</i>	<i>F₁</i>
YOLOv5l	85.7	89.9	90.5	90.2
YOLOv5m	86.8	91	91.9	91.4
YOLOv5n	89.5	92.6	91.9	92.2
YOLOv7	88.6	93.1	92	92.5
YOLOv5s	90.9	92.6	93.2	93
文献[7]	92.3	94.9	92.6	93.7
YOLOv6	95.8	94.6	93.9	94.2
Improved YOLOv5s	96.9	96.6	97.3	97

为了验证改进后模型的可行性,选取部分图片进行测试,图 9 为改进前后模型检测结果对比图。图 (a) 中使用原始 YOLOv5s 算法对目标检测的置信度普遍偏低,在后续处理中容易被滤除,影响检测准确度;图 (b) 为使用改进后算法检测结果,瑕疵点在被正确检测的前提下置信度都有所提升。这表明算法对小目标的检测能力有所提高。图 (c) (d) 分别是同一张图片经过改进前和改进后模型的测试结果。由图可知,液位线的

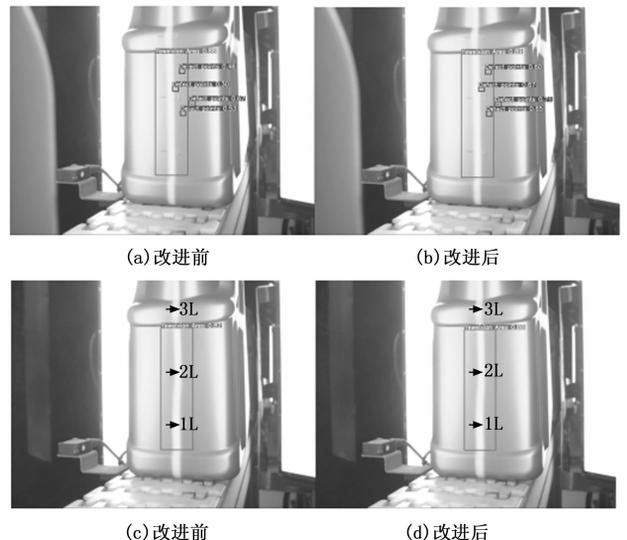


图 9 模型改进前后检测结果对比图

检测置信度有一定提升,从0.82提升至0.88,这一数据的提高表明模型在液位线检测和定位性能有所进步,为后续液位线裁剪提供了更高的准确性,降低出现偏差的可能性。

4 结束语

本文提出了一种基于改进YOLOv5s算法的机油瓶瓶身缺陷检测方法。首先利用K-Means++代替K-Means初始化聚类中心,生成更加准确和具有代表性的先验框;其次在主干网络中引入DCNv2模块来更加灵活地调整卷积核和池化操作,从而更好地捕捉缺陷的特征信息;第三步在主干网络的最后一层引入SE注意力机制,通过全局感受野的线性变换,将输入特征图的每个位置都映射到一个新的特征表示,增强模型在局部区域内的特征提取能力;最后在颈部网络中引入BiFPN网络,通过对不同尺度的特征进行融合,更好地提取图像信息。尽管模型总体参数量有所增加,但改进方法对于机油瓶瓶身微小瑕疵点特征的提取及液位线的检测能力均有所增强,且模型部署在工业检测现场达到了产线检测要求。该方法为注塑产品瓶身外观缺陷的检测提供了新的解决思路。

参考文献:

- [1] 张 鲲, 邓明星. 基于传统图像处理联合深度学习的FPCB自动缺陷检测算法[J]. 智能计算机与应用, 2023, 13 (7): 53-57.
- [2] 孟 陆, 金 永. 基于分布式的玻璃缺陷检测技术研究及性能优化[J]. 计算机测量与控制, 2019, 27 (12): 47-51.
- [3] 龙智帆. 基于机器视觉的瓶盖缺陷检测算法设计及可重构方案的研究[D]. 杭州: 杭州电子科技大学, 2011.
- [4] 麦国铭. 表面光洁的注塑制品外观缺陷视觉检测方法研究[D]. 广州: 华南理工大学, 2012.
- [5] 王 莉, 何牧天, 徐 硕, 等. 基于YOLOv5s网络的垃圾分类和检测[J]. 包装工程, 2021, 42 (8): 50-56.
- [6] 许思昂, 李艺杰, 梁桥康, 等. 基于改进YOLOv5算法的PCB裸板缺陷检测[J]. 包装工程, 2022, 43 (15): 33-41.
- [7] 朱 豪, 周顺勇, 曾雅兰, 等. 基于改进YOLOv5s的木材表面缺陷检测模型[J]. 木材科学与技术, 2023, 37 (2): 8-15.
- [8] 乔 炎, 甄 彤, 李智慧. 改进YOLOv5的安全帽佩戴检测算法[J]. 计算机工程与应用, 2023, 59 (11): 203-211.
- [9] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C] //Proceedings of the IEEE Conference on Computer Vision and

Pattern Recognition, 2016: 779-788.

- [10] 王淑青, 顿伟超, 黄剑锋, 等. 基于YOLOv5的瓷砖表面缺陷检测[J]. 包装工程, 2022, 43 (9): 217-224.
- [11] ZHOU F, ZHAO H, NIE Z. Safety helmet detection based on YOLOv5 [C] //2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA). IEEE, 2021: 6-11.
- [12] WANG Z, WU L, LI T, et al. A smoke detection model based on improved YOLOv5 [J]. Mathematics, 2022, 10 (7): 1190.
- [13] 胡 欣, 周运强, 肖 剑, 等. 基于改进YOLOv5的螺纹钢表面缺陷检测[J]. 图学学报, 2023, 44 (3): 427-437.
- [14] ZHU X, HU H, LIN S, et al. Deformable convnets v2: More deformable, better results [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 9308-9316.
- [15] GUO J, XIAO X, MIAO J, et al. Design and experiment of a visual detection system for Zanthoxylum-harvesting robot based on improved YOLOv5 Model [J]. Agriculture, 2023, 13 (4): 821.
- [16] DAI J, QI H, XIONG Y, et al. Deformable convolutional networks [C] //Proceedings of the IEEE International Conference on Computer Vision, 2017: 764-773.
- [17] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [18] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722.
- [19] ZHOU P, NI B, GENG C, et al. Scale-transferrable object detection [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 528-537.
- [20] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117-2125.
- [21] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [22] TAN M, PANG R, LE Q V. Efficientdet: scalable and efficient object detection [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10781-10790.