

DOI: 10.13382/j.jemi.B2104171

改进 YOLOv4 的内丝接头密封面缺陷检测算法*

单明陶 高玮玮

(上海工程技术大学机械与汽车工程学院 上海 201620)

摘要:针对传统目标检测算法对内丝接头密封面的缺陷识别率不高的问题,提出利用改进的 YOLOv4 算法对其进行检测。首先使用 K-means++ 聚类算法对目标样本进行先验框参数优化,提高先验框与特征图的匹配度;其次在主干网络嵌入 SENet 注意力机制模块,强化图像关键信息,抑制图像背景信息,提高不易识别缺陷的置信度;然后在网络颈部增加 SPP 模块,增强主干网络输出特征的接受域,分离出重要的上下文信息;最后采用收集的内丝接头密封面缺陷数据集分别对改进前后的 YOLOv4 进行训练,并分别测试模型效果。实验结果表明,YOLOv4 检测内丝接头密封面缺陷的性能较好,但有部分小目标漏检;改进后的模型对小目标缺陷的检测表现优异,均值平均精度 (mAP) 达到了 87.47%,相比于原始 YOLOv4 提升了 10.2%,平均检测时间为 0.132 s,实现了对内丝接头密封面缺陷的快速准确检测。

关键词: 目标检测;YOLOv4;SENet;SPP;K-means++

中图分类号: TP391.4;TH16 **文献标识码:** A **国家标准学科分类代码:** 520.20

Improved YOLOv4's algorithm for detecting defects on the sealing surface of inner wire joints

Shan Mingtao Gao Weiwei

(College of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, Shanghai 201620, China)

Abstract: Aiming at the problem of low recognition rate of traditional target detection algorithm for inner wire joint sealing surface defects, an improved YOLOv4 algorithm was proposed to detect the defects. Firstly, k-means++ clustering algorithm is used to optimize the parameters of the anchor frame of the target sample, and improve the matching degree between the anchor frame and the feature map; Secondly, the SENet attention mechanism module is introduced into the backbone network to strengthen the key information of the image, suppress the background information of the image, and improve the confidence of the defect that is not easy to identify; after that, the SPP module is added to the neck of the network to enhance the acceptance domain of the backbone network output features and separate the important context information; Finally, using the collected data set of inner wire joint sealing surface defects to train the original YOLOv4 and the improved YOLOv4, and the performance of models were tested respectively on test set. The experimental results show that the performance of YOLOv4 is good, but some small targets are missed; The improved model has excellent detection performance for small target defects, the mean average accuracy (mAP) reaches 87.47%, which is 10.2% higher than the original YOLOv4, and the average detection time is 0.132 s, which realize the rapid and accurate detection of inner wire joint sealing surface defects.

Keywords: target detection; YOLOv4; SENet; SPP; K-means++

0 引言

内丝接头作为刹车器件中的重要零件,其经过锻造、冲孔、车削、磨削、热处理等多道工序加工后,表面

可能会出现各种缺陷(划痕、磨损等),这些表面缺陷不仅会对产品外观产生影响,还会对产品质量产生影响,对产品使用寿命也会产生影响,严重的时候甚至会造事故,故需对其进行缺陷检测。传统的人工检测方法存在劳动强度大、检测效率低等问题,而且容易受工

人的检测经验、疲劳程度等因素的影响。因此,研究一种智能且高效的内丝接头表面缺陷的检测方法具有重要意义。

研究人员提出了各种表面缺陷检测方法,具体可分为传统目标检测法和深度学习目标检测法两类。传统的目标检测主要使用基于手工特征配合机器学习的方法(即 Feature-Based 方法)^[1]。伍玉琴等^[2]将齿轮图像进行二值化后,再与标准模板进行对比,以检测待测齿轮缺陷;郭慧等^[3]提出了由计算机算法处理得到缺陷的特征样本,再利用支持向量机(SVM)模型对钢板表面缺陷进行类型识别。这类方法的优点在于检测速度较快,缺点为可检测缺陷类型单一,泛化能力弱。

近年来,国家鼓励大力发展人工智能技术,深度学习的一系列方法也被引入到表面缺陷检测中。深度学习目标检测先通过卷积神经网络提取缺陷图像的特征,然后对特征进行学习,再将学习后的模型应用到相似缺陷的检测中。这种方法检测精度高、速度快且泛化能力强,能够适应不同类型的缺陷。现在主流的深度学习目标检测算法分为 two-stage 和 one-stage 两类。two-stage 需要先产生目标候选框,然后再对候选框做分类与回归,这类算法的代表为 Faster-RCNN^[4]、Mask-RCNN^[5]; one-stage 仅仅使用一个卷积神经网络 CNN 直接预测不同目标的类别与位置,这类算法的代表为 YOLO^[6]、SSD^[7]。随着深度学习的快速发展,结合深度学习目标检测的金属表面缺陷检测也得到了广泛应用^[8]。Wang 等^[9]设计了多层卷积神经网络对 6 种金属缺陷进行缺陷检测,使用滑窗法在原图像中进行特征采样,后对每一类图像采样的小图像块进行二分类,识别效果优于传统方式。张广世等^[10]采用改进 YOLOv3 对齿轮缺陷进行检测取得良好效果。

虽然深度学习方法被运用于金属工件表面缺陷检测的例子不少,但对于内丝接头密封面上的小缺陷检测效果并不理想,因此基于当前最优的 one-stage 目标检测算法 YOLOv4^[12],提出一种改进 YOLOv4 的算法对内丝接头密封面缺陷进行检测。

1 实验材料及设备

内丝接头工件来源于上海某汽车零部件制造公司,源图像由 CCD 相机拍摄,共收集 20 00 张分辨率为 2 592×1 944 pixels 的内丝接头密封面缺陷图像,分为磨损(缺陷 1,如图 1(b)所示)、划痕(缺陷 2,如图 1(c)所示)、密封面内含铁丝(缺陷 3,如图 1(d)所示),密封面内大面积损伤(缺陷 4,如图 1(e)所示),每类缺陷图像各 500 张。对图像进行人工标注,将图像宽高和深度、缺陷种类及缺陷位置信息存入 xml 文件。为保证训练样本与测试

样本的独立性,从 2 000 张图像中随机抽取 1 600 张作为训练集,将剩下的 400 张图像作为测试集,训练集和测试集占比为 8 : 2。

实验平台配置如下:计算机处理器为 Intel Core i5-10300H,运行内存为 16.00 GB,独立显卡为 RTX2060,显存为 6 GB。运行软件为 Anaconda 虚拟环境下的 Pycharm;python 版本为 3.7;Pytorch 版本为 1.2.0;安装 Cuda10.0 和 Cudnn7.51 进行 GPU 加速运算。

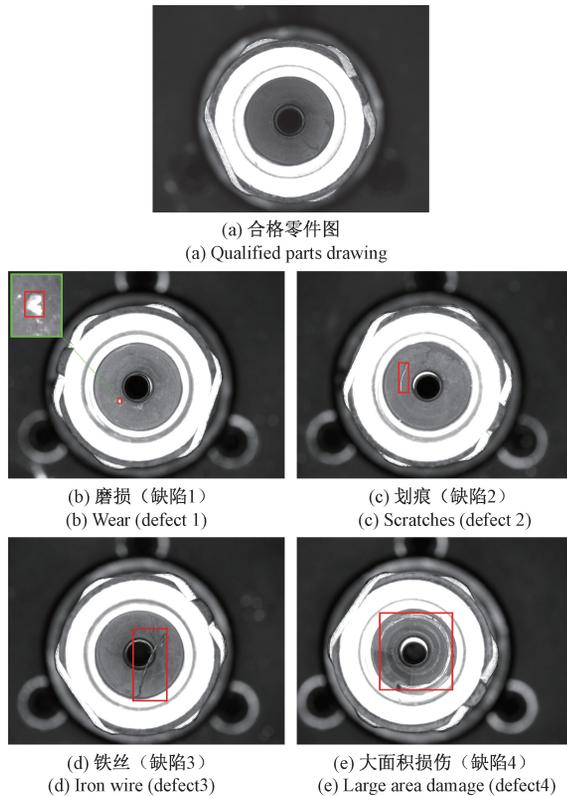


图 1 部分密封面图像

Fig. 1 Partial sealing surface image

2 方法

2.1 改进的 YOLOv4 算法

YOLOv4 算法由 YOLOv3^[11]算法改进而来,由特征提取的主干网络(backbone)、特征融合的颈部(neck)和进行分类与回归的检测头(head)三部分组成。网络结构如图 2 所示,主干网络在 DarkNet-53 的基础上融合 CSPNet^[13]思想,形成 CSPDarkNet-53,降低网络计算量的同时也可保证准确率;颈部采用融合特征金字塔网络(feature pyramid networks, FPN)、空间金字塔池化^[14](spatial pyramid pooling, SPP)的路径聚合网络^[15](path aggregation network, PANet),解决了 FPN 网络由浅层向深层传递时浅层特征丢失的问题;检测头仍然使用

YOLOv3 检测头。YOLOv4 的损失函数在 YOLOv3 的基础上使用 CIOU^[17] (complete intersection over union) 误差

代替均方误差作为回归框预测误差 L_{loc} , 置信度误差 $L_{c_{conf}}$ 和分类误差 L_{cls} 并未改变。

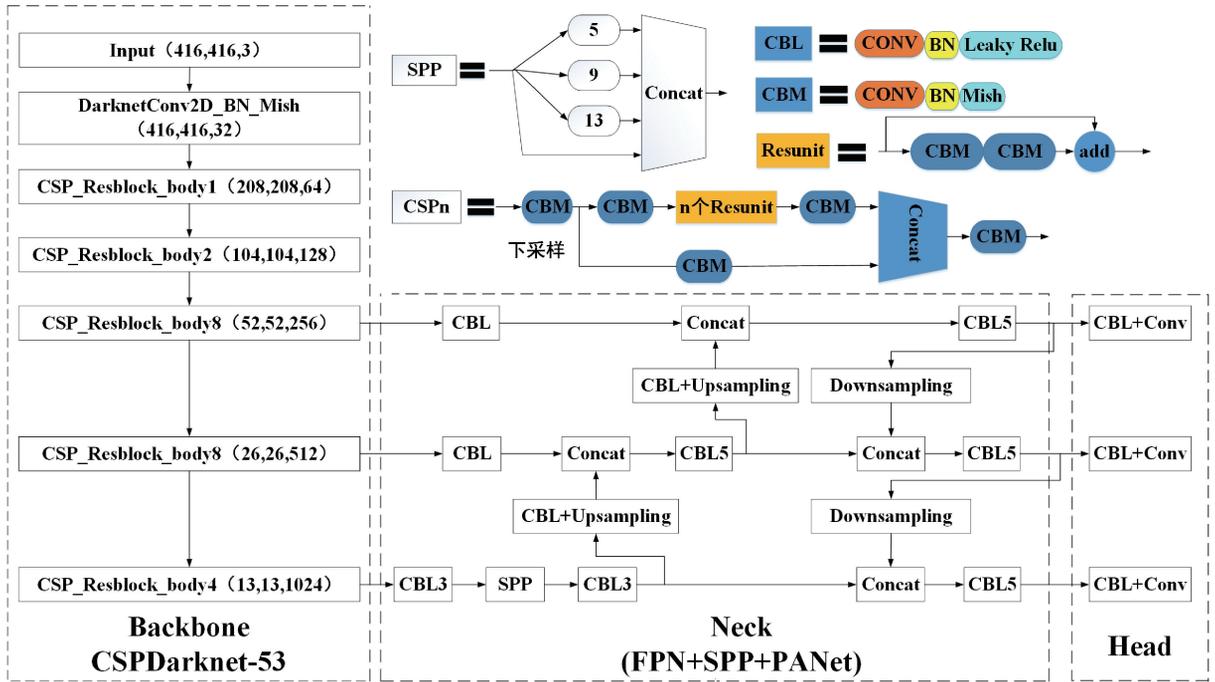


图 2 YOLOv4 网络结构

Fig. 2 YOLOv4 network structure

YOLOv4 主干网络包含 5 个 CSP 模块,各模块的卷积核大小为 3×3 ,步幅为 2,能够进一步增强网络学习能力。在网络颈部采用了 SPP 结构和 PANet 结构。SPP 结构是对最后一个特征层经过 3 次卷积后的结果进行最大化池化。池化过程中共使用 4 种不同尺度的池化层进行处理,池化核大小分别为 1×1 、 5×5 、 9×9 、 13×13 。经过 SPP 处理后能有效增加感受野,分离出显著的上下文特征。PANet 结构是由卷积操作、上采样操作、特征层融合、下采样操作构成的循环金字塔结构。经过 PANet 结构后,不同特征层之间充分融合,能有效提高目标的特征提取能力。YOLOv4 Head 利用 PANet 处理后的 3 个特征层进行结果预测,与 YOLOv3 Head 的原理一样,进行大小为 3×3 和 1×1 的两次卷积操作,对每个特征层的 3 个先验框进行判别,判断其内部是否包含目标以及目标种类。CBM 和 CBL 模块是在卷积层后加入了批量归一化层 (batch normalization, BN),且分别使用 Mish 和 Leaky ReLU 激活函数,CSP 模块由 CBM 模块和残差块 Resunit 堆叠而成。

在一定程度上,YOLOv4 解决了对小目标不敏感的问题,但在内丝接头密封面缺陷检测上,由于缺陷在形态、大小等方面存在很大差异,直接使用 YOLOv4 网络检测密集小目标缺陷,检测效果并不理想。由图 1 可以看出,部分磨损缺陷相比于整幅图像占比较小,故针对所检

测缺陷的特征,对 YOLOv4 算法进行改进。首先采用 K-means++ 进行目标框优化以提高先验框 (anchor box) 与目标数据集的匹配度;在此基础上,通过在主干网络中嵌入 SENet^[16] 结构来提高网络对于密集小目标缺陷的置信度;此外,还在网络的颈部增加了 SPP 模块以实现丰富检测头的信息特征表达能力的目标。

1) K-means++ 进行目标框优化

在 YOLOv4 算法中,先验框是采用 K-means 聚类算法对数据集中标注好的检测目标聚类得到的。原始的先验框是针对 VOC 数据集预设的,使用其会使在计算交并比 IOU 时筛选出不合适的边界框 (bounding box),影响模型性能。

采用 K-means++ 聚类算法对收集到的数据集进行聚类,K-means++ 与 K-means 相比,优化了初始点的选择,能显著降低分类的误差,以获得更适合目标数据集的先验框的大小,提高检测精度。首先从样本数据集中随机选择一个目标框区域作为初始聚类中心,然后计算每个样本 x_i 与当前已有聚类中心之间的最短距离 $D(x)$ (即与最近的一个聚类中心的距离),并计算每个样本被选为下一个聚类中心的概率 $P(x)$:

$$P(x) = D(x)^2 / \sum D(x)^2 \quad (1)$$

最后用轮盘法选出下一个聚类中心。重复计算 $D(x)$ 和 $P(x)$,直到找到 k 个聚类中心。选出初始点后,

就继续使用标准的 K-means 算法更新聚类中心,直到先验框的大小不再发生变化。

2) 嵌入 SENet 结构的 YOLOv4 算法

SENet 是一种利用全局信息来有选择地强调信息特征并抑制无用特征的网络。该网络结构如图 3 所示,它是由常规快捷连接层 (Shortcut) 和 SE 结构组成的重复单元。Squeeze 操作使用全局平均池化来显示该图层中通道特征图的数值分布,也称为全局信息。Excitation 操作使用门控机制和 Sigmoid 激活函数来描述张量中每个通道特征图的权重。两个全连接层 (FC) 的功能是融合每个通道的特征图信息。

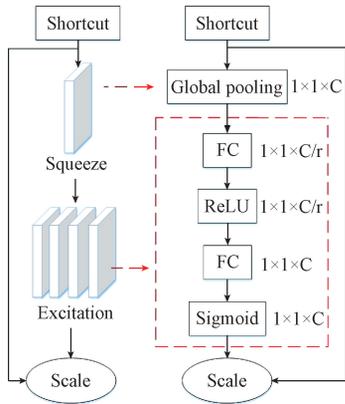


图 3 SENet 结构

Fig. 3 SENet structure

在内丝接头密封面缺陷图像中,原始的 YOLOv4 检测密集小目标缺陷往往会误检或漏检,这是由于置信分布不平衡造成的错误识别。为了使网络学习全局特征,提高密集小目标缺陷的检测精度,对各个通道的权值进行自动标定,选择在 YOLOv4 网络中嵌入 SENet 结构,在大量减少参数的情况下,保留全局信息可以增强模型的鲁棒性。

主干网络中的特征图语义信息并不丰富,保留着目标在中、浅层的纹理信息和轮廓信息,这些信息对小目标的检测极其重要;而在网络深层的颈部和头部,特征图表现的语义信息更丰富,感受野更大。注意力机制模块 (SENet) 难以从特征信息高度融合的小尺度特征图中有效区分重要特征,因此选择在主干网络中嵌入 SENet 模块。主干网络中的 CSP 模块嵌入 SENet 模块如图 4 所示,将改进主干网络的算法称为 YOLOv4-a。

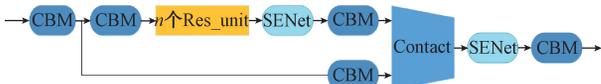


图 4 嵌入 SE 模块的 CSP 模块

Fig. 4 CSP module embedded with SE module

3) 增加 SPP 模块的 YOLOv4 算法

借鉴 SPPNet, YOLOv4 模型在主干网络后加入了 SPP 模块,结构如图 2 所示。SPP 结构对主干网络最后的特征层经过 3 次卷积后的结果进行最大池化,使用 4 种不同尺度的池化层处理,池化核大小为 1x1、5x5、9x9、13x13。最大池化层在保持特征图平移不变性的同时扩大感受野,而 SPP 结构将使用不同大小池化核得到的特征图局部区域感受野和接近全局的感受野信息进行特征融合。这种融合不同尺度感受野的操作能有效丰富特征图的表达能力,增强主干网络输出特征的接受范围,分离出重要的上下文信息,有效提升模型的检测性能。

在原始 YOLOv4 中,SPP 模块位于特征图大小为 13x13 的主干网络之后,经过 SPP 模块提取出的特征信息送入特征图大小为 13x13 的检测头进行预测。为了丰富输入到 26x26 和 52x52 尺度的检测头的特征图特征信息,实现更好的检测效果,在 YOLOv4 第 1 次上采样后和第 2 次上采样后加入 SPP 模块,改进后的 YOLOv4 颈部如图 5 所示,将改进 YOLOv4 颈部的算法称为 YOLOv4-b。

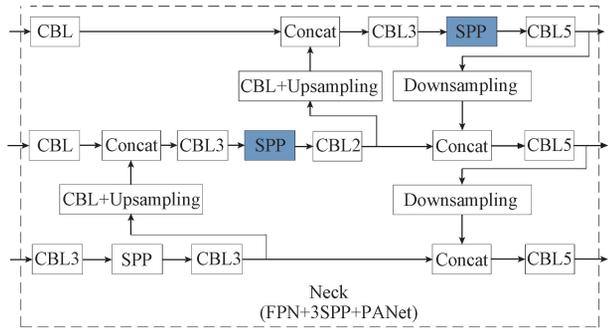


图 5 YOLOv4 颈部修改

Fig. 5 YOLOv4 neck modification

2.2 基于改进 YOLOv4 的内丝接头密封面缺陷检测方法

缺陷检测方法流程如图 6 所示,在模型训练阶段,首先对训练样本进行数据增强,来提高模型泛化能力;其次采用增强的数据训练原始 YOLOv4 网络;同时为了提高改进网络模型对边界框的预测准确率,对收集到的所有样本标注采用 K-means++ 聚类获取先验框的初始大小;然后将先验框参数应用于不同网络结构改进的算法;最后经过训练迭代,得到不同的检测模型。在模型测试阶段,将测试样本作为输入分别使用各算法模型检测,判断图像中是否存在缺陷,若存在缺陷则输出缺陷种类及位置,并比较不同模型的性能。

为了提高模型的泛化能力,对内丝接头图像进行在线数据增强,包括随机旋转、水平翻转,竖直翻转,中值滤波,平滑滤波,调整饱和度 (0.9~1.3 倍)、曝光度 (0.9~1.3 倍) 等。为保证预测结果的精度,同时避免过多的先

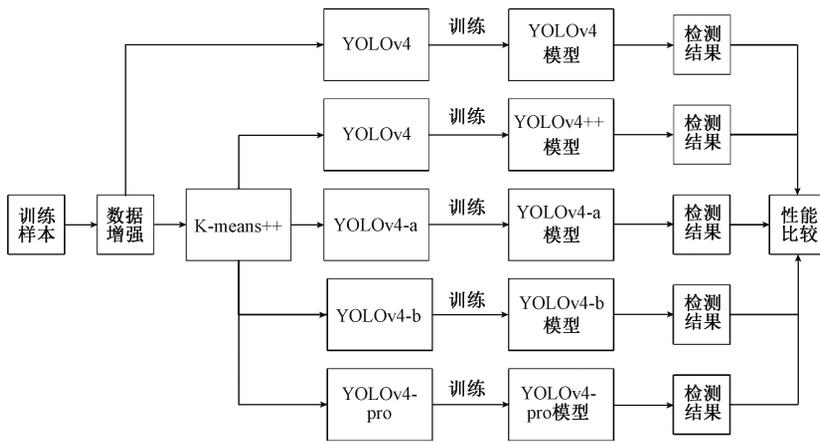


图 6 方法流程

Fig. 6 Method flow chart

验框带来过大的运算量,仍然选取聚类中心 $k=9$,其宽和高分别为(6, 8)、(7, 12)、(9, 10)、(10, 14)、(11, 19)、(12, 10)、(14, 14)、(25, 27)、(128, 160)。

采用带动量因子的小批量随机梯度下降法 (MSGD) 进行训练。MSGD 是介于全批量梯度下降法 (BGD) 和随机梯度下降法 (SGD) 之间的折中方法。BGD 每一次迭代都需要计算所有样本,训练过程很慢。SGD 每次迭代一个样本来更新参数,训练速度加快,但可能会收敛到局部最优。MSGD 可以减少 SGD 中的杂乱样本数量,且比 BGD 更高效。为使梯度稳定更新,采用 Momentum 作为优化器对其进行梯度下降。为了增强模型鲁棒性,采用 L2 正则化来防止过拟合的出现。

学习率的大小对随机梯度下降法的性能影响较大,直接决定了参数移动到最优值时的速度。训练开始时应该设置较大的学习率,能够使模型快速收敛;随着训练迭代次数的增加,应该逐渐降低学习率,以此找出最优解。

选择在 VOC 预训练权重的基础上设定 YOLOv4 算法的超参数。训练的 epoch 经过尝试 20、40、60 等常用设置,并根据 loss 不再下降而选择为 60。Momentum 动量参数常用值为 0.55、0.9、0.99,经实验比较得出 0.9 最优。若 L2 正则化权重衰减系数过大,会把所有的参数最小化,造成欠拟合现象;若权重衰减系数过小,则可能无法避免过拟合问题,因此 L2 正则化权重衰减系数通常设置为 0.9。常用的学习率衰减方式有分段衰减、指数衰减、余弦衰减,经过测试分段衰减效果最佳,初始学习率一般为 0.001,经过实验发现在 30 个 epoch 时衰减为初始学习率 1/10 效果最佳。批大小 (batchsize) 一般设置为 1、2、4、8,若批大小过小,会导致训练迭代不收敛且训练速度慢;若批大小过大,则会导致网络收敛到局部最优点。考虑到数据集及硬件设备性能,将批量大小设置为 2,每个 epoch 迭代 800 次,整个训练过程迭代

48 000 次。YOLOv4-pro 算法的整个训练过程的损失函数变化如图 7 所示,训练损失收敛值为 0.042。

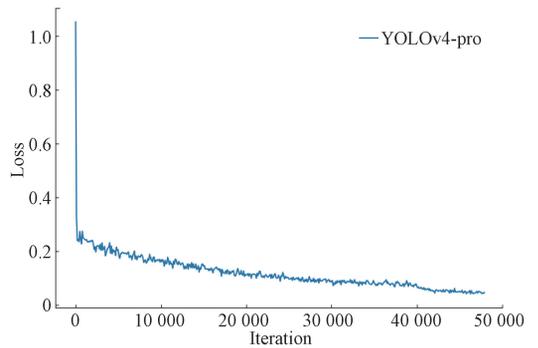


图 7 损失函数变化曲线

Fig. 7 Loss function change curve

3 实验结果与分析

3.1 性能指标

评价网络性能需要兼顾召回率 (recall) 和精确率 (precision),目标检测中通常用平均精度均值 (mean average Precision, mAP) 来评价网络模型性能。式 (2) 为精确率的计算公式,式 (3) 为召回率的计算公式。平均精度 (average precision, AP) 的含义是不同 Recall 下的 Precision 值的均值 (AP 就是 P-R 曲线的积分),通常用来评价某一类的检测精度。mAP 表示的是所有类别检测精度的均值,计算公式如式 (4)。采用准确率 (accuracy) 来评价检测结果的好坏,计算公式如式 (5)。检测时间 (time) 为网络每检测一张图像的时间,用来评估网络模型的检测速度。

$$P = TP / (TP + FP) \tag{2}$$

$$R = TP / (TP + FN) \quad (3)$$

$$mAP = \sum_{i=1}^n AP(i) / n \quad (4)$$

$$Acc = (TP + TN) / (TP + FP + TN + FN) \quad (5)$$

式中: P 为精确率; R 为召回率; Acc 为准确率; TP 为成功预测的正例; TN 为成功预测的负例; FP 为被误判为正例的负例; FN 为被错误预测为负例的正例; $AP(i)$ 为某一类的检测精度; n 为类别数。

3.2 实验结果与分析

对比使用 K-means++ 聚类算法优化先验框后, 在内丝接头密封面数据集上, YOLOv4 算法的性能变化。采用未经优化与优化先验框后的 YOLOv4 网络在数据集上训练, 并在测试集上进行验证, 结果如表 1 所示。

表 1 K-means++ 优化结果对比

Table 1 Comparison of K-means++ optimization results

Network Structure	K-means++	mAP/%
YOLOv4	×	77.27
YOLOv4++	√	80.19

表 2 不同改进 YOLOv4 模型在测试集上的性能评估

Table 2 Performance evaluation of different improved YOLOv4 models on the test set

Network structure	SENet Added	SPP Added	AP/%				mAP/%	Acc/%	Time/s
			1	2	3	4			
Baseline	×	×	77.47	72.02	81.14	90.13	80.19	91.75	0.1250
YOLOv4-a	√	×	81.67	77.91	83.38	90.76	83.43	97.5	0.1284
YOLOv4-b	×	√	79.64	78.54	82.73	89.69	82.65	95.5	0.1295
YOLOv4-pro	√	√	85.16	83.14	88.69	92.89	87.47	98.5	0.1320

表 3 不同算法的检测性能对比

Table 3 Comparison of detection performance of different algorithms

Network Structure	mAP/%	Acc/%	Time/s
Faster-RCNN (VGG16)	66.73	88.75	0.3754
Faster-RCNN (ResNet50)	74.82	92	0.6012
YOLOv3	73.43	89.5	0.1236
SE-IYOLOv3	76.24	96.75	0.1280
YOLOv4	77.14	91.25	0.1250
YOLOv4-pro	87.47	98.5	0.1320

由表 3 可知, 与 Faster-RCNN (VGG16) 和 Faster-RCNN (ResNet50) 相比, YOLOv4-pro 在 mAP 上分别提升了 20.74% 和 12.65%, 检测时间也大大缩短。主要原因在于 Faster-RCNN 的先验框的长宽比是预先设定的, 无法根据数据集目标大小进行自由调整, 而划痕和铁丝目标框多为细长型, 磨损这类缺陷的目标尺寸较小, 导致 Faster-RCNN 容易遗漏目标, 整体检测效果并不好; YOLOv4-pro 中的先验框是由 K-means++ 算法聚类获得的, 先验框尺寸与数据集中目标大小更接近, 更适合检测

表 1 中 YOLOv4 是未经优化的原始网络, YOLOv4++ 为使用 K-means++ 优化先验框的网络。从表 1 结果可以看出, 优化先验框使其适合内丝接头密封面数据集的框定能明显提高网络的检测精度, mAP 提升了近 3%。

由表 2 可以看出, 每种网络改进策略均在不同程度上提高了 Baseline (用 K-means++ 优化先验框的 YOLOv4) 的检测性能。采用主干网络嵌入 SE 模块的 YOLOv4-a 模型, mAP 提升了 3.24%; 采用颈部增加 SPP 模块的 YOLOv4-b 模型, mAP 提升了 2.46%; 采用主干网络和颈部皆改进的 YOLOv4-pro 模型, mAP 提升了 7.28%。

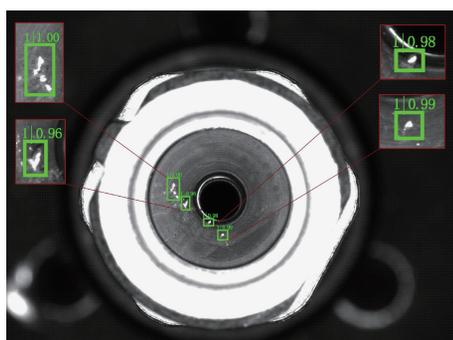
Faster-RCNN 为目前主流的二阶段检测网络, VGG16 和 ResNet50 是其常用的主干网络, YOLOv3 为目前常用的一阶段检测网络。采用相同的网络训练策略和超参数去训练 Faster-RCNN 网络、YOLOv3 网络以及 Deng 等^[17] 提出的 SE-IYOLOv3 网络, 并在内丝接头密封面缺陷测试集上进行性能评估, 具体结果如表 3 所示。

内丝接头密封面缺陷。与 YOLOv3 和 Deng 等^[18] 提出的 SE-IYOLOv3 相比, YOLOv4-pro 算法在检测精度上优于二者, 检测速度略慢, 但仍能够满足工程实际要求。与原始 YOLOv4 相比, YOLOv4-pro 在保持检测时间基本持平的情况下增加了检测精度, 提升了网络的整体性能。这是因为 YOLOv4 是在 YOLOv3 的基础上改进而来的, 检测精度优于 YOLOv3, 而 YOLOv4-pro 又在 YOLOv4 基础上进一步提升了性能。

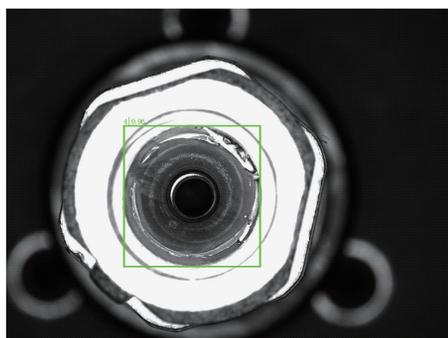
内丝接头密封面缺陷检测结果如图 8 所示, 图 8(a) 对应磨损 (缺陷 1), 图 8(b) 对应密封面内大面积损伤 (缺陷 2), 图 8(c) 中包含磨损 (缺陷 1) 和密封面内含有铁丝 (缺陷 3), 图 8(d) 包含磨损 (缺陷 1) 和划痕 (缺陷 2)。由图 8 看出, 该检测方法不仅能够检测出上述缺陷, 而且能够对不同缺陷进行精确分类和定位。

4 结论

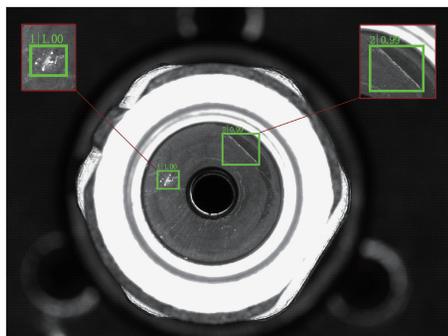
为了实现内丝接头密封面缺陷检测的自动化和智能化, 针对密封面内密集小缺陷难以检出的问题, 以现阶段



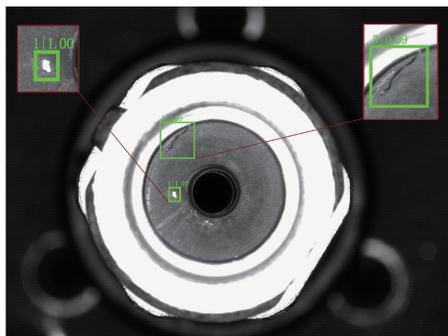
(a) 缺陷1
(a) Defect 1



(b) 缺陷4
(b) Defect 4



(c) 缺陷1和缺陷2
(c) Defect 1 and defect 2



(d) 缺陷1和缺陷3
(d) Defect 1 and defect 3

图 8 密封面缺陷检测效果

Fig. 8 Detection effect of sealing surface defects

类对样本进行先验框大小的优化,有利于提高先验框尺寸与目标大小的匹配;此外在主干网络嵌入 SENet 结构扩大网络的感知域,有利于提高不易识别缺陷的置信度;最后在颈部增加 SPP 结构,增强主干网络输出特征的接受域,有利于分离重要的上下文信息。YOLOv4-pro 模型对每种缺陷的检测精度均有提升,在 mAP 上提升了 10.2%,实现了对内丝接头密封面缺陷的快速准确检测,满足了工程项目对检测精度和速度的要求,对 YOLOV4 应用于工程实际具有重要的现实意义。

参考文献

- [1] 李章维,胡安顺,王晓飞. 基于视觉的目标检测方法综述[J]. 计算机工程与应用, 2020, 56(8): 1-9.
LI ZH W, HU A SH, WANG X F. Review of vision based object detection methods [J]. Computer Engineering and Applications, 2020, 56(8): 1-9.
- [2] 伍玉琴,徐海元,程晓军. 基于机器视觉的齿轮缺陷研究[J]. 煤矿机械, 2019, 40(4): 170-172.
WU Y Q, XU H Y, CHENG X J. Research on gear defects based on machine vision [J]. Coal Mining Machinery, 2019, 40(4): 170-172.
- [3] 郭慧,徐威,刘亚菲. 基于支持向量机的钢板表面缺陷检测[J]. 东华大学学报(自然科学版), 2018, 44(4): 635-639.
GUO H, XU W, LIU Y F. Steel plate surface defect detection based on support vector machine[J]. Journal of Dong Hua University (Natural Science Edition), 2018, 44(4): 635-639.
- [4] REN S Q, HE K M, GIRSHICK R B, 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.
- [5] HE K M, GEORGIA G, DOLLAR P, et al. Mask R-CNN[C]. IEEE International Conference on Computer Vision, 2017.
- [6] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. Computer Vision and Pattern Recognition, 2016: 779-788.
- [7] LIU W, ANGELOV D, ERHAN D, et al. SSD: single shot multibox detector [C]. Proceedings of the IEEE European Conference on Computer Vision, 2016: 21-37.
- [8] 蓝金辉,王迪,申小盼. 卷积神经网络在视觉图像检测的研究进展[J]. 仪器仪表学报, 2020, 41(4): 167-182.
LAN J H, WANG D, SHEN X P. Research progress of convolutional neural network in visual image detection[J].

最优的一阶段目标检测算法 YOLOv4 为基础,提出一种改进 YOLOv4 的缺陷检测算法。首先采用 K-means++聚

- Chinese Journal of Scientific Instrument, 2020, 41(4): 167-182.
- [9] WANG T, CHEN Y, QIAO M, et al. A fast and robust convolutional neural network-based defect detection model in product quality control [J]. International Journal of Advanced Manufacturing Technology, 2017, 94 (9) : 3465-3471.
- [10] 张广世,葛广英,朱荣华,等. 基于改进 YOLOv3 网络的齿轮缺陷检测 [J]. 激光与光电子学进展, 2020, 57(12): 153-161.
- ZHANG G SH, GE G Y, ZHU R H, et al. Gear defect detection based on improved YOLOv3 network [J]. Laser & Optoelectronics Progress, 2020, 57(12): 153-161.
- [11] REDMON J, DIVVALA S, GIRSHICK R, et al. YOLOv3: An incremental improvement [C]. Computer Vision and Pattern Recognition, 2018.
- [12] BOCHKOVSKIY A, WANG C Y, LIAO H Y. YOLOv4: Optimal speed and accuracy of object detection [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2020.
- [13] WANG C Y, LIAO H Y, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of cnn [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020: 390-391.
- [14] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [15] 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.
- [16] 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.
- [17] ZHENG Z, WANG P, LIU W, et al. Distance-IOU Loss: Faster and better learning for bounding box regression [C]. AAAI Conference on Artificial Intelligence, 2020: 12993-13000.
- [18] DENG Z R, YANG R, LAN R, et al. SE-IYOLOV3: An accurate small scale face detector for outdoor security [J]. Mathematics, 2020, 8(1): 93.
- [19] 何静,余昊宇,张昌凡,等. 基于 Canny-YOLOv3 的列车轮对踏面损伤检测 [J]. 电子测量与仪器学报, 2019, 33(12): 25-30.
- HE J, YU H Y, ZHANG CH F, et al. Train wheel tread damage detection based on Canny-YOLOv3 [J]. Journal of Electronic Measurement and Instrumentation, 2019, 33(12): 25-30.
- [20] 韩航迪,徐亦睿,孙博,等. 基于改进 Tiny-YOLOv3 网络的航天电子焊点缺陷主动红外检测研究 [J]. 仪器仪表学报, 2020, 41(11): 42-49.
- HAN H D, XU Y R, SUN B, et al. Research on active infrared detection of solder joint defects in aerospace electronics based on improved Tiny-YOLOv3 network [J]. Chinese Journal of Scientific Instrument, 2020, 41(11): 42-49.
- [21] 向宽,李松松,栾明慧,等. 基于改进 Faster RCNN 的铝材表面缺陷检测方法 [J]. 仪器仪表学报, 2021, 42(1): 191-198.
- XIANG K, LI S S, LUAN M H, et al. Aluminum surface defect detection method based on improved Faster RCNN [J]. Chinese Journal of Scientific Instrument, 2021, 42(1): 191-198.

作者简介



单明陶, 2018 年于太原工业学院获得学士学位, 现为上海工程技术大学硕士研究生, 主要研究方向为计算机视觉、数字图像处理、深度学习。

E-mail: 17836210737@163.com

Shan Mingtao received his B. Sc. degree in 2018 from Taiyuan Institute of Technology. Now he is a M. Sc. candidate at Shanghai University of Engineering Science. His main research interests include Computer vision, digital image processing and deep learning.



高玮玮 (通信作者), 2014 年于南京航空航天大学获得博士学位, 现为上海工程技术大学讲师, 主要研究方向为机器视觉、图像处理、模式识别。

E-mail: gww03020234@sina.com

Gao Weiwei (Corresponding author) received her Ph. D. degree in 2014 from Nanjing University of Aeronautics and Astronautics, now she is a lecturer in Shanghai University of Engineering Science. Her main research interests include machine vision, image processing and pattern recognition.