

DOI:10.16136/j.joel.2022.02.0349

基于 YOLOv4 的铜带表面缺陷识别研究

王紫玉¹, 张 果^{1*}, 杨 奇², 尹丽琼²

(1. 昆明理工大学 自动化学院, 云南 昆明 650500; 2. 武钢集团昆明钢铁股份有限公司安宁公司, 云南 昆明 650302)

摘要: 本文提出一种基于 YOLOv4 铜板带材表面缺陷检测模型, 针对铜金属板带材生产过程中产生的表面缺陷形式多样、位置随机而导致难以快速定位和识别的问题, 采用大数据驱动的深度学习策略, 以铜带表面缺陷图像为训练样本, 对 YOLOv4 目标检测模型进行训练, 实验结果表明, 改进的模型识别铜带表面缺陷的全类别平均精度均值 (mean average precision, mAP) 为 93.37%, 高于原始 YOLOv4 模型的全类别平均精度 91.46%, 检测速度达到 49 帧/秒, 与双阶段的检测模型更快地 R-CNN (faster region-based convolutional neural network, Faster R-CNN) 相比, 在保证检测精度的同时提升检测速度, 能够满足在线检测需要, 适合完成铜带工业生产过程中缺陷检测任务。

关键词: 深度学习; 缺陷识别; YOLOv4; 模式识别

中图分类号: TP391 **文献标识码:** A **文章编号:** 1005-0086(2022)02-163-08

Research on surface defect recognition of copper strip based on YOLOv4

WANG Ziyu¹, ZHANG Guo^{1*}, YANG Qi², YIN Liqiong²

(1. Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, Yunnan 650500, China; 2. WISCO Group Kunming Iron & Steel Co., Ltd., Anning Company, Kunming, Yunnan 650302, China)

Abstract: This paper proposes a surface defect detection model on a copper plate and strip based on YOLOv4. Aiming at the problem of surface defects in the production process of copper metal plate and strip that are difficult to locate and identify due to their various forms and random positions, a big data-driven deep learning strategy is adopted. Using the copper strip surface defect image as the training sample, the YOLOv4 target detection model is trained. The experimental results show that the improved model recognizes the copper strip surface defect with a full-category mean average precision (mAP) of 93.37%, which is higher than the original YOLOv4. The model has an average accuracy of 91.46% for all categories and a detection speed of 49 frames per second. Compared with the two-stage detection model faster region-based convolutional neural network (Faster R-CNN), it can improve the detection speed while ensuring the detection accuracy, which can meet the needs of online detection. Defect detection task in industrial production process is suitable for completing copper strips.

Key words: deep learning; defect identification; YOLOv4; pattern recognition

1 引 言

长期以来, 表面质量检测均依靠质检员目视进行判断, 当铜带以某个固定速度进行生产时, 质检员会因视觉疲劳、生产速度过快等原因造成缺陷漏检或误判, 传统的缺陷检测方式可以分为涡

流检测技术^[1]、红外检测技术^[2]、漏磁检测技术^[3]等, 但这些检测技术部署昂贵, 在应对不同类型的缺陷检测任务时存在诸多不足, 而基于机器视觉的深度学习目标检测技术在成本降低的同时, 能够稳定高效地应对不同类型的缺陷信息^[4-7]。

目前的深度学习模型可以被分为两类, 一类

* E-mail: 21426717@qq.com

收稿日期: 2021-05-25 修订日期: 2021-06-23

基金项目: 国家重点研发计划(2017YFB0306405)、国家自然科学基金(61364008)和云南省重点研发计划(2018BA070)资助项目

是采用双阶段的目标检测模型,如更快地 R-CNN (faster region-based convolutional neural network, Faster R-CNN) 模型^[8,9],能够将缺陷检测分为特征提取与特征分类,另一类采用单阶段目标检测模型,如 YOLO 家族模型、应用在工业界的检测模型 SSD 等^[10],将缺陷检测当作回归问题,文献[11]使用 Faster R-CNN 模型检测汽车轮毂表面出现的缺陷,达到 91.2% 的正检率,文献[12]采用 YOLOv2 模型检测绝缘子缺陷,准确率可达 94.4%,文献[13]使用 YOLOv3 模型检测钢材表面的压痕与划痕,检测精度达到 92% 和 90%,相比于 Faster R-CNN 这类双阶段目标检测模型,单阶段的 YOLO 目标检测模型没有复杂检测流程,更适合工业生产在线检测环节。

为了解决铜带表面缺陷的识别与分类问题,基于 YOLOv4 网络模型^[14],采用实验室制样得到铜带缺陷样本对网络模型进行训练,得到模型权重,然后将模型应用在铜带金属表面缺陷检测任务中。

2 基于改进 YOLOv4 的铜带缺陷检测模型

2.1 模型理论

YOLOv4 在 YOLOv3 的基础上改进而成,

YOLOv4 整体网络结构如图 1 所示。相较于 YOLOv3, YOLOv4 将激活函数由 LeakyReLU 修改为式(1)所示的 Mish 函数,使用 CSPDarknet-53 作为主干网络,使用 DarknetConv2D_BN_Mish 处理输入数据维度问题,然后使用 SSP 结构对主干网络的输出进行 3 次 DarknetConv2D_BN_Leaky 卷积,再使用 4 个不同大小的池化核进行最大池化处理,它们能够极大地增加感受野,分离出上下文特征,最后将处理完毕的结果使用 PANet 结构对特征进行提取, PANet 网络结构如图 2 所示。PANet 结构首先完成传统特征金字塔的由上到下的采样过程,然后进入(b)结构完成由下到上的采样,即达到反复对特征进行提取的效果, YOLOv4 将 PANet 结构应用在 3 个有效的特征层上,分别位于中间层,偏下层和最下层,输出尺度为 76×76 、 38×38 、 19×19 大小的缺陷特征,最后使用 YOLO Head 得到预测类别、置信度、预测框信息等。Mish 函数计算式为:

$$\text{Mish} = x \times \tanh(\ln(1 + e^x)), \quad (1)$$

式中, x 为输入数据。

2.2 基于 IoU 的 k-means 聚类改进

单阶段检测模型 YOLOv4 借鉴了双阶段检测模型 Faster R-CNN 的先验框思想,在训练中将目标的检测任务等价于锚框内是否存在目标的回归问题,在网络训练过程中,不断修正预测框与真实框之间

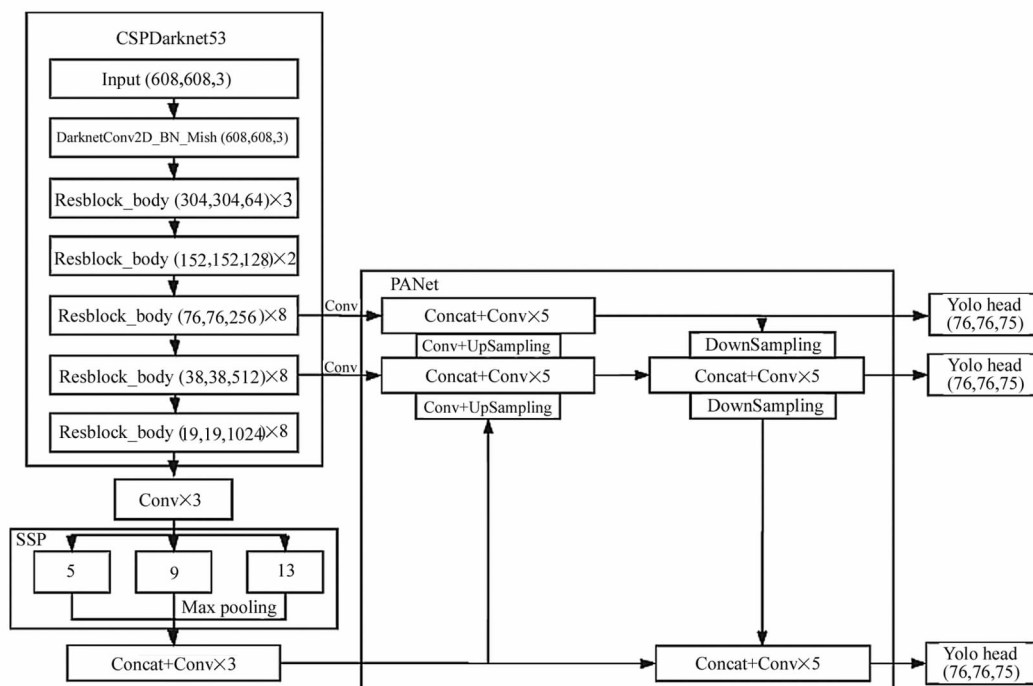


图 1 YOLOv4 整体网络结构

Fig. 1 YOLOv4 network structure

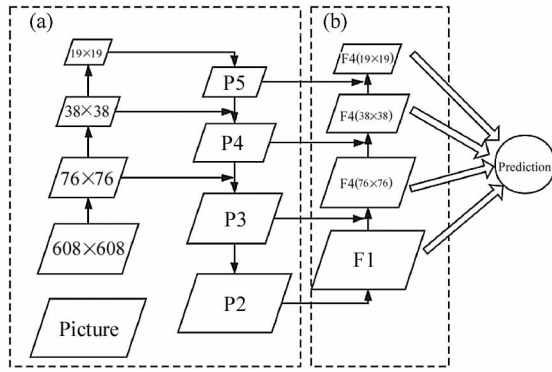


图2 PANet网络结构

Fig. 2 PANet network structure

距离, YOLOv4 模型在检测头模块共设置了 9 个先验框, 训练之前, 原始先验框是根据 VOC 数据集的真实标注得到, 不同数据集中目标真实框的位置与大小均存在一定差异, 而先验框选取的准确性对于网络初始收敛影响较大, 所以原始的先验框并不适用于铜带金属表面缺陷检测任务。为了能够使网络学习到缺陷特征与目标位置, 采用改进的 k-means 算法对收集到的缺陷数据集真实边框进行聚类, 得到适用于缺陷检测任务的新先验框。传统 k-means 算法使用欧几里得距离公式计算误差, 采用这种方法会忽略大边框计算时产生的额外误差, 所以采用改进的式(2)对距离进行计算, 即预测框与真实框的交并比(intersection over union, IoU), 计算距离式如式(3)所示。

$$IoU = \frac{Box^{pred} \cap Box^{gt}}{Box^{pred} \cup Box^{gt}}, \quad (2)$$

$$D(box, centroid) = 1 - IoU(box, centroid), \quad (3)$$

式中, Box^{pred} 代表预测框, Box^{gt} 代表真实框。

使用缺陷数据集中真实标定框进行改进的 k-means 聚类, 最终得到适合数据的先验框。

2.3 训练过程改进

对于 YOLOv4 模型训练过程中的学习率设定, 普遍做法是设置好初始学习率 LR (learning rate) 进行训练, 学习率会随着训练过程的收敛而不断降低, 在模型训练过程中, 大学习率能够加快模型学习速度, 跳出局部最优值, 但会导致模型训练难以收敛, 总是使用大学习率会使训练出的模型不够精确。小学习率能够帮助模型收敛, 有助于提高模型精度, 但却无法跳出局部最优值, 从而收敛缓慢, 学习率的大小与它们对应的功能是矛盾的, 为了调和这种训练时因学习率产生的矛盾, 使模型快速且高效收敛, 这里使用结合了热重启的余弦退火函数对学习率进行

控制, 学习率调整式为:

$$\eta_t = \eta_{\min} + \frac{1}{2}(\eta_{\max} - \eta_{\min})(1 + \cos(\frac{T_{\text{cur}}}{\lambda_{\text{mult}} T_{\text{max}}} \pi)), \quad (4)$$

式中, η_t 表示当前学习率, $\eta_{\min}$ 表示学习率的最小值, $\eta_{\max}$ 表示学习率的最大值, T_{cur} 表示上一次开始到目前经历的 epoch 个数, λ_{mult} 为学习率周期扩大因子, T_{max} 代表固定的周期参数。

当 $\lambda_{\text{mult}} > 1$ 时, 每完成一次热重启后, 新周期都会扩大为前一次的 λ_{mult} 倍, 学习率会随着训练次数增多而衰减得越来越缓慢, 当 $T_{\text{cur}} = \lambda_{\text{mult}} T_{\text{max}}$ 时,

$\cos(\frac{T_{\text{cur}}}{\lambda_{\text{mult}} T_{\text{max}}} \pi) = -1$, 此时模型的学习率大小回归为

$\eta_t = \eta_{\min}$, 即完成了一次学习率由最大值经余弦退火函数下降为最小值的过程。热重启在神经网络的训练过程中, 通过增大学习率, 使网络跳出局部最优的过程, 传统的余弦退火学习率衰减策略在整个训练过程中持续衰减学习率, 直到学习率为 0, 当损失函数值陷入局部最优时, 不断衰减的学习率显然难以跳出局部最优值。热重启很好地解决了这个问题, 热重启是基于模型在上一步收敛的参数开始进行的, 而不是重新开始执行, 它通过训练过程中不断改变的训练批量实现, 可以将这个过程总结为, 热重启采用线性上升方式增大学习率, 当学习率达到最大值 $\eta_{\max}$ 时, 余弦退火函数工作, 使学习率下降到最小值 $\eta_{\min}$, 此时触发热重启, 更新热重启周期。在此基础上提升学习率到最大值 $\eta_{\max}$, 不断往复此过程, 直到训练结束。基于热重启的学习率余弦退火函数如下图 3 所示, 横坐标为训练步长, 纵坐标为学习率。

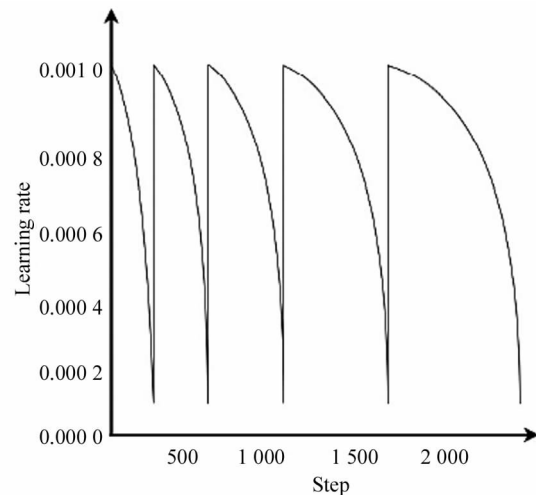


图3 基于热重启的学习率余弦退火函数

Fig. 3 Cosine decay with warm up

2.4 评价指标

铜带金属表面缺陷识别任务需要解决回归问题与分类问题,可以采用平均精度(average precision, AP)、每秒模型处理帧数(frames per second,FPS)、Loss回归曲线对模型质量进行评判,精确率与召回率定义式为:

$$P = \frac{T_P}{T_P + F_P}, \quad (5)$$

$$R = \frac{T_P}{T_P + F_N}, \quad (6)$$

式中, T_P 代表模型正确识别的缺陷个数, F_P 代表模型将非缺陷区域识别为缺陷的个数, F_N 为未模型未识别出的缺陷个数。

以召回率 R 和准确率 P 绘制 PR 曲线对模型进行多指标评价, PR 曲线围成的面积即为 AP 值, AP 值越大代表模型分类效果越好, FPS 越高代表模型在线检测的性能越好, $Loss$ 曲线能反映出模型在训练时以及最终的收敛情况。

3 实验结果与分析

3.1 数据集制作

训练使用实验室制样得到的铜带缺陷样本,样本库中收集了3种不同材质的金属铜合金板材,分别是白铜(cup)、黄铜(bra)和紫铜(red),共1847张,缺陷共有3种,每种材质的不同缺陷约为200张,分别为黑点(black_spot)、划痕(scratch)、油斑(oil_drop),不同类型的缺陷样本如图4所示,训练集与验证集按照8:2比例进行划分,1477张图像为训练集,370张作为测试集。

使用Labelimg工具对图片中存在的缺陷进行矩形框标注,标注格式参照VOC数据集格式,将标注好的文件按xml格式进行存储,以便于接下来对改进的YOLOv4模型进行训练,处理结果如图4(b)所示。YOLOv4中还采用了Mosaic数据增强方式对图像内的信息进行增强处理,将随机4张不同语义图像进行缩放、翻转后拼接,得到融合了多种信息的

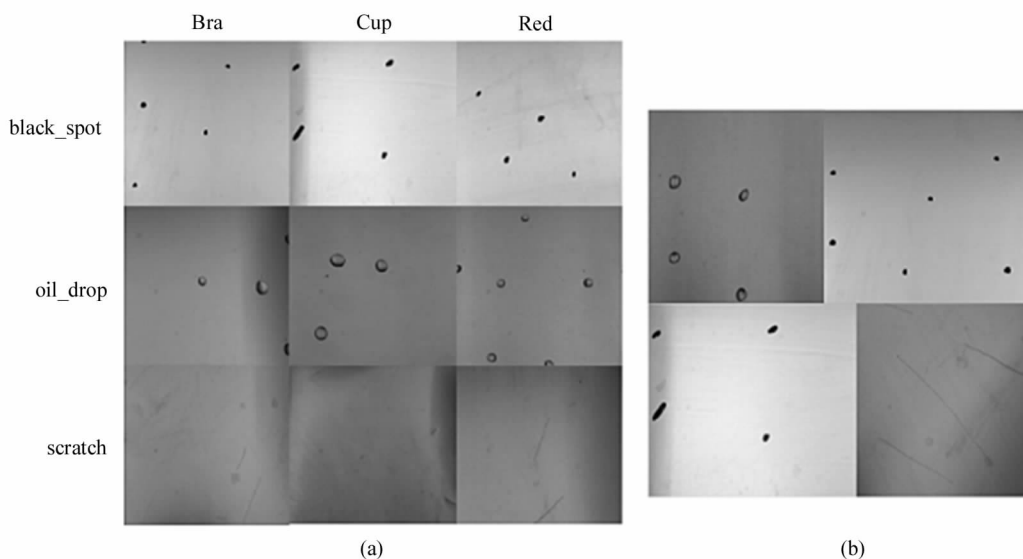


图4 不同类型缺陷样本:(a) 样本库缺陷类型;(b) Mosaic数据增强

Fig. 4 Samples of different types of defects: (a) Defect type of sample library; (b) Mosaic data enhancement

训练图片,如图4(b)所示,使用这种策略对模型进行训练,大大增强了模型对于复杂语境的鲁棒性。

3.2 实验运行环境

本实验环境为Win10操作系统下使用深度学习框架Pytorch1.2.0,Python3.8进行实验,CPU型号:AMD R7-4800H和GPU型号:RTX2060(6GB显存),内存32GB,神经网络加速依赖为Cuda10.0和Cudnn7.4.1.5。

3.3 超参数选取

改进模型使用VOC数据集预训练权重对YOLOv4超参数进行初步设定,对于超参数的选取,YOLOv4预设的9个先验框为(12,16);(19,39);(40,28);(36,75);(76,55);(72,146);(142,110);(192,243);(459,401)。使用改进的k-means聚类后得到适用于铜带缺陷的新先验框为(5,6);(11,14);(14,23);(19,20);(23,29);(30,38);(41,

115);(91,55);(135,147)。采用新的先验框进行模型训练,整个训练过程设定为 85 个 epoch,前 50 个 epoch 采用冻结训练方式提高模型收敛速度,保护初期权值,batch_size 设为 4,训练时开启 Mosaic 数据增强,使用基于热重启的余弦退火函数改进训练,设置学习率的重启周期 $T_{\max}=5$ 次,周期扩大因子 $\lambda_{\text{mult}}=2$,设置前 50 次 epoch 最大学习率 $\eta_{\max}=1\times 10^{-3}$,最小学习率 $\eta_{\min}=1\times 10^{-5}$,设置后 35 次 epoch 最大学习率 $\eta_{\max}=1\times 10^{-4}$,最小学习率保持不变,最后设置模型的预测类别分为 3 类,分别是 black_spot、scratch 和 oil_drop,然后进行训练。

3.4 实验结果

使用 Tensorboard 可视化工具对训练过程进行观察,训练过程的真实损失函数收敛曲线对比如图 5 所示。图 5 中直线为代表使用等间距调整学习率策略的训练结果,学习率初始值与改进方法保持一致,设置学习率调整步长为 1,每经一个 epoch 训练学习率进行调整一次,得到最后 10 次 epoch 平均损失为 5.303。图 5 中带有原点直线代表使用 k-means 聚类算法得到先验框后采用基于热重启和余弦退火函数的学习率调整策略得到的训练结果,最后 10 次 epoch 平均损失为 4.355。

从图 5 中可以看出,训练开始时采用 k-means

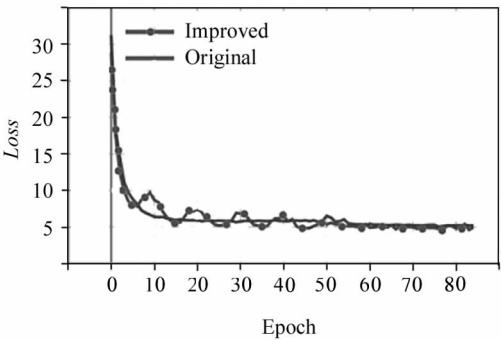


图 5 损失函数收敛曲线对比

Fig. 5 Convergence curve comparison of loss function

进模型会得到更小的 Loss 值,迭代进行 10 次后 Loss 值下降趋势变得缓慢,改进模型采用的学习率优化策略使学习率随着训练次数的增加而不断进行调整,在实际训练过程中表现为第 10 到 50 次迭代过程中 Loss 值在升高与降低间不断循环,最终在第 75 到 85 次迭代区间内 Loss 曲线趋于稳定,训练结束后保存收敛模型的权重,使用该收敛模型对表面缺陷进行识别,设定预测框与真实框的 IoU 大于 0.5 即为成功预测缺陷位置,部分预测结果如图 6 所示。

改进的 YOLOv4 模型与原始 YOLOv4 模型的预测集结果对比如表 1 所示。

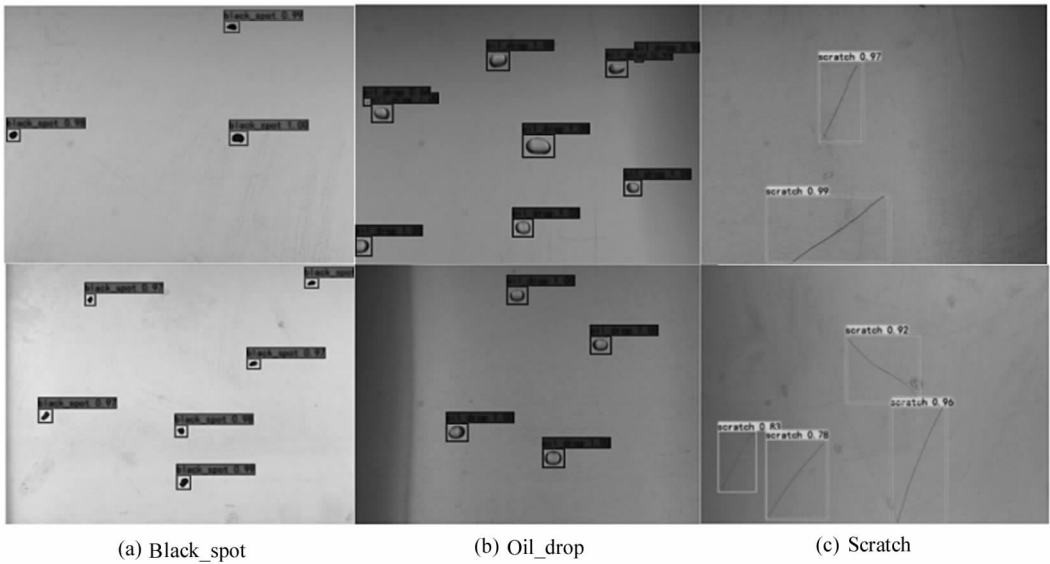


图 6 部分预测结果:(a) 黑点;(b) 油斑;(c) 划痕

Fig. 6 Partial prediction results:(a) Black spot; (b) Oil drop; (c) Scratch

以上数据根据式(5)—(6)计算识别模型的准确率、召回率并绘制不同模型识别缺陷的 PR 曲线,如图 7 所示。横坐标为召回率(recall),纵坐标为准确

率(precision),通过计算图 7 中阴影面积可得原始模型识别黑点(black_spot)的 AP 为 96.48%,改进模型模型识别黑点的 AP 为 97.17%。原始模型识别

油斑(oil_drop)的 AP 为 95.97%,改进模型识别油斑的 AP 为 96.13%。原始模型识别划痕(scratch)的 AP 为 81.94%,改进模型识别划痕的 AP 为 86.81%,提升最为明显。

图 7 中第一行图(a)、(b)、(c)为原始模型对黑

点、油斑、划痕识别得到的 PR 曲线,图 7 中第二行图(d)(e)(f)为使用 k-means 聚类算法结合基于热重启的学习率退火衰减得到模型对黑点、油斑、划痕等识别得到的 PR 曲线,为了更加直观地观察不同模型在不同缺陷上的检测精度如图 8 所示。

表 1 测试集结果对比

Tab.1 Comparison of test set results

Target	True positive		False positive		False negative	
Model	Improved	Original	Improved	Original	Improved	Original
Black_spot	405	403	10	8	11	12
Oil_drop	602	601	65	144	39	33
Scratch	499	484	163	270	160	220

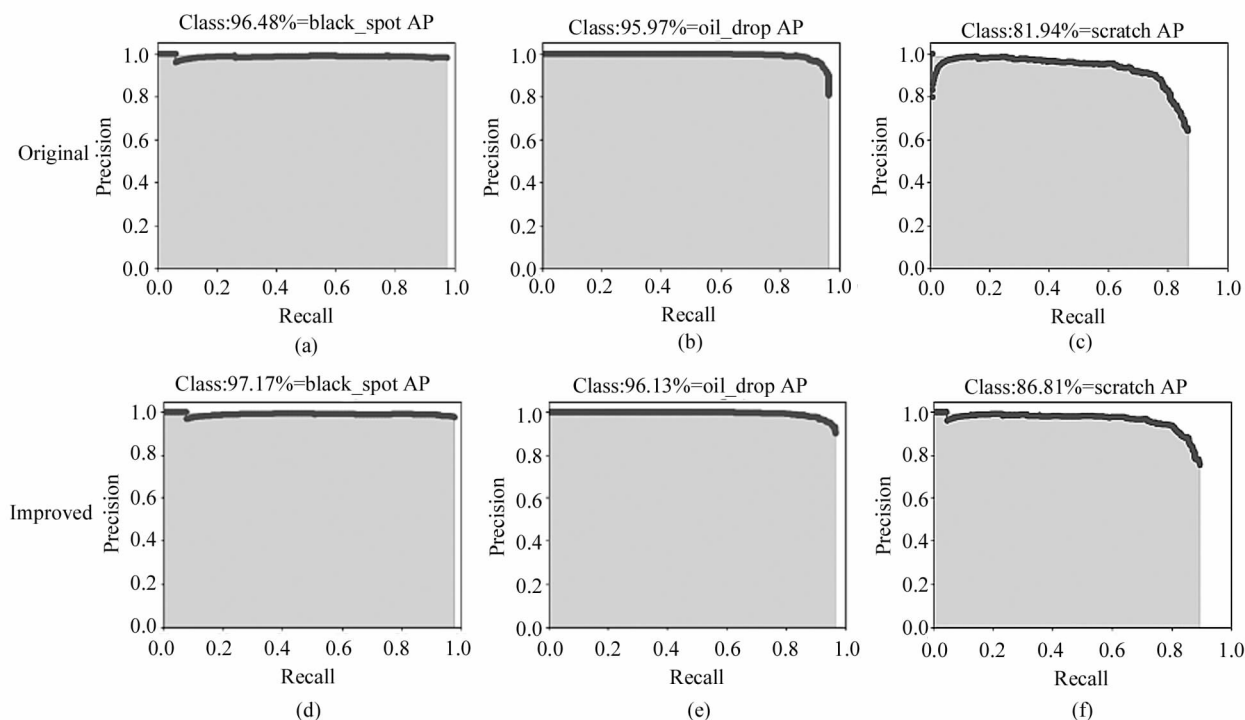


图 7 不同模型识别缺陷的 PR 曲线

Fig. 7 Precision-Recall curves of different models to identify defects

图 8 中,横坐标代表平均检测精度(average precision),纵坐标代表不同缺陷类型,较深颜色矩形代表改进后模型(improved),较浅色矩形代表未改进的模型(original),可以得出改进后模型对划痕缺陷的检测精度提升 5%,对黑点缺陷检测精度提升 1%,性能方面,针对同一测试数据集,使用原始 YOLOv4、改进的 YOLOv4 模型和 Faster R-CNN 模型进行性能测试,计算模式采用 GPU 加速计算,将模型性能对比如表 2 所示。

观察表 2 在检测速度方面,单阶段检测模型

YOLOv4 相较于双阶段检测模型 Faster R-CNN 速率提升超过两倍,改进的 YOLOv4 模型相比于原始模型提升 3 FPS。检测精度方面,原始 YOLOv4 模型相比于 Faster R-CNN 提升 5.07%,改进 YOLOv4 模型相比于 Faster R-CNN 提升 6.98%,原始 YOLOv4 模型平均精度均值(mean average precision, mAP)为 91.46%,改进后 YOLOv4 模型平均检测精度为 93.37%,提升 1.91%。所以,改进的 YOLOv4 模型对铜带缺陷识别在保证检测速率的前提下,提高了缺陷识别的准确性,适合在工业生产过

程中完成在线检测任务。

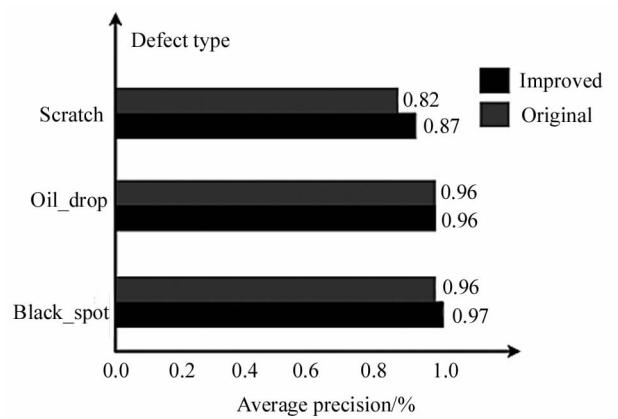


图8 不同模型在不同缺陷上的检测精度
Fig.8 Detection accuracy of different models on different defects

表2 模型性能对比

Tab.2 Comparison model performance

Model	FPS	mAP/%
YOLOv4	45	91.46
YOLOv4 Improved	49	93.37
Faster R-CNN	22	86.39

4 结 论

本文提出一种基于改进 YOLOv4 的铜带表面缺陷识别模型,使用实验室制样得到缺陷样本作为数据集,训练并测试该改进模型,并与未经改进的 YOLOv4 识别模型进行对比试验,结果表明,基于改进的 YOLOv4 铜带缺陷识别模型识别缺陷的平均精度为 93.37%,与原始模型的结果相比较精度提升了 1.9%,基于硬件条件不变,FPS 提升了 3 FPS,改进的 YOLOv4 模型相比于原始 YOLOv4 模型,能够在训练过程中优先跳出局部最优,使模型快速收敛,同时在准确性方面高于未经改进的原始模型。接下来的研究以提升模型的速度为主,在不损失精度的前提下,尽可能提升模型的识别速度。

参考文献:

[1] ZHANG Y L,ZHANG H C,ZHAO J X,et al. Review of non-destructive testing for remanufacturing of high-end equipment[J]. Journal of Mechanical Engineering, 2013, 49(7):80-90.
张元良,张洪潮,赵嘉旭,等. 高端机械装备再制造无损检测综述[J]. 机械工程学报,2013,49(7):80-90.

[2] MO Z X,CHEN Y J. Research and prospects of the do-

mestic infrared thermography technology[J]. Laser & Infrared, 2014, 44(12):1300-1305.

莫朝霞,陈沅江. 我国红外热像检测技术的研究及发展展望[J]. 激光与红外, 2014, 44(12):1300-1305.

[3] SHEN G T,ZHANG W L. Survey of nondestructive testing techniques for special equipment[J]. Nondestructive Testing, 2006, (1):34-39.

沈功田,张万岭. 特种设备无损检测技术综述[J]. 无损检测, 2006, (1):34-39.

[4] TANG B,KONG J Y,WU S Q. Review of surface defect detection based on machine vision[J]. Journal of Image and Graphics, 2017, 22(12):1640-1663.

汤勃,孔建益,伍世度. 机器视觉表面缺陷检测综述[J]. 中国图象图形学报, 2017, 22(12):1640-1663.

[5] WANG Y N,CHEN T J,HE Z D,et al. Review on the machine vision measurement and control technology for intelligent manufacturing equipment[J]. Control Theory & Applications, 2015, 32(3):273-286.

王耀南,陈铁健,贺振东,等. 智能制造装备视觉检测控制方法综述[J]. 控制理论与应用, 2015, 32(3):273-286.

[6] TAO X,HOU W,XU D. A survey of surface defect detection methods based on deep learning[J]. Acta Automatica Sinica, 2021, 47(5):1017-1034.

陶显,侯伟,徐德. 基于深度学习的表面缺陷检测方法综述[J]. 自动化学报, 2021, 47(5):1017-1034.

[7] LU R S,WU A,ZHANG T D,et al. Review on automated optical (visual) inspection and its applications in defect detection[J]. Acta Optica Sinica, 2018, 38(8):23-58.

卢荣胜,吴昂,张腾达,等. 自动光学(视觉)检测技术及其在缺陷检测中的应用综述[J]. 光学学报, 2018, 38(8):23-58.

[8] SHEN X H,LI Z H,LI M,et al. Aluminum surface-defect detection based on multi-task deep learning[J]. Laser & Optoelectronics Progress, 2020, 57(10):283-292.

沈晓海,栗泽昊,李敏,等. 基于多任务深度学习的铝材表面缺陷检测[J]. 激光与光电子学进展, 2020, 57(10):283-292.

[9] FENG X Y,MEI W,HU D S. Aerial target detection based on improved faster R-CNN[J]. Acta Optica Sinica, 2018, 38(6):250-258.

冯小雨,梅卫,胡大帅. 基于改进 Faster R-CNN 的空中目标检测[J]. 光学学报, 2018, 38(6):250-258.

[10] ZHOU X Y,WANG K,LI L Y. Review of object detection based on deep learning [J]. Electronic Measurement Technology, 2017, 40(11):89-93.

周晓彦,王珂,李凌燕. 基于深度学习的目标检测算法综述[J]. 电子测量技术, 2017, 40(11):89-93.

- [11] ZHU C P, YANG Y B. Online detection algorithm of automobile wheel surface defects based on improved faster-RCNN model[J]. Surface Technology, 2020, 49(6): 359-365.
朱超平, 杨永斌. 基于改进的 Faster-RCNN 模型的汽车轮毂表面缺陷在线检测算法研究[J]. 表面技术, 2020, 49(6): 359-365.
- [12] LAI Q P, YANG J, TAN B D, et al. Insulator automatic identification and defect diagnosis model based on YOLOv2 network[J]. Electric Power, 2019, 52(7): 31-39.
赖秋频, 杨军, 谭本东, 等. 基于 YOLOv2 网络的绝缘子自动识别与缺陷诊断模型[J]. 中国电力, 2019, 52(7): 31-39.
- [13] FANG Y X, GAN P, CHEN L. Improved YOLOv3 algorithm for detection of metal surface defect[J]. Mechanical Science and Technology for Aerospace Engineering, 2020, 39(9): 1390-1394.
方叶祥, 甘平, 陈俐. 金属表面缺陷检测的改进 YOLOv3 算法研究[J]. 机械科学与技术, 2020, 39(9): 1390-1394.
- [14] Bochkovskiy A, Wang C Y, Liao H Y M. Yolov4: Optimal speed and accuracy of object detection[EB/OL]. (2020-04-23) [2021-05-25]. <https://arxiv.org/abs/2004.10934>.

作者简介:

张 果 (1976—), 男, 工学博士, 副教授, 硕士生导师, 主要从事智能信息处理方面的研究。