

DOI: 10.19650/j.cnki.cjsi.J2108078

基于卷积神经网络的热轧钢条表面实时缺陷检测*

刘艳菊¹, 王秋霖², 赵开峰², 刘彦忠²

(1. 南京特殊教育师范学院数学与信息科学学院 南京 210038;

2. 齐齐哈尔大学计算机与控制工程学院 齐齐哈尔 161000)

摘要: 热轧钢条的表面质量对成品至关重要,因此必须要严格控制热轧钢条的表面出现的缺陷。针对当前 YOLOv4 算法检测精度不高、对小范围信息表现较差等问题,提出一种改进 YOLOv4 自动检测方法。首先,将 YOLOv4 中特征提取网络 CSPDarknet53 换为轻量级深层神经网络 MobileNetv3 来提高检测速度,并且加强对检测目标特征提取以及减少梯度消失问题。其次,采用 K-Means 聚类生成适合本实验的先验框,有效提高学习效率,加快收敛速度。最后,对置信度损失进行重新定义,提出一种能够适应多尺度的损失函数,来解决因正负样本不平衡而导致检测效果差的问题。实验结果表明,该方法较原 YOLOv4 模型在热轧钢条的表面缺陷检测上的均值平均精度值提高约 7.94%,速度提升约 4.52 f/s,在保证检测速度的基础上有效提高了精确度。

关键词: 缺陷检测; YOLOv4; MobileNetv3; K-Means 聚类

中图分类号: TP391.41 TH878 **文献标识码:** A **国家标准学科分类代码:** 520.20

Real-time defect detection of hot rolling steel bar based on convolution neural network

Liu Yanju¹, Wang Qiuji², Zhao Kaifeng², Liu Yanzhong²

(1. School of Mathematics and Information Science, Nanjing Normal University of Special Education, Nanjing 210038, China;

2. School of Computer and Control Engineering, Qiqihar University, Qiqihar 161000, China)

Abstract: It is important for the surface quality of hot rolled steel strips to make final product. Therefore, it is necessary to strictly control the defects on the surface of hot rolled steel strips. The current you only look once (YOLO) v4 algorithm has low detection accuracy and poor performance on small-scale information. To address these issues, an improved YOLOv4 automatic detection method is proposed. First, to improve detection speed, enhance detection target feature extraction and reduce gradient vanishing, the feature extraction network CSPDarknet53 in YOLOv4 is replaced with the lightweight deep neural network MobileNetv3. Secondly, to improve the learning efficiency and accelerate the convergence speed, the K-Means clustering is utilized to generate a prior box to suit for this experiment. Finally, the confidence loss is redefined and a loss function is proposed that can adapt to the multi-scale to solve the problem of poor detection effect due to the imbalance of positive and negative samples. Compared with the original YOLOv4 model for the surface defect detection of the hot rolled steel strip, experimental results show that the proposed method enhance the mean average precision and the speed about 7.94% and 4.52 f/s, respectively. The accuracy of this model is improved effectively while ensuring the detection speed.

Keywords: defect detection; YOLOv4; MobileNetv3; K-Means clustering

0 引 言

热轧钢条在汽车、电器制造、工业以及日常生活等领

域都有着重要应用。热轧钢条的表面质量受到多个因素的影响,尽管通过各种控制方法能够减少缺陷,但热轧钢条表面还是不可避免地会出现不同类型的缺陷。热轧钢条不同的表面缺陷对生产质量影响也是不同的,因此需要

收稿日期:2021-06-11 Received Date: 2021-06-11

* 基金项目:国家自然科学基金青年基金项目(61403222)资助

对热轧钢条表面缺陷进行分类,以减少它们的不利影响。

目前对热轧钢条还是以人工的方式做缺陷检测,来进行质量跟踪和保证。然而随着生产能力的不断提高和生产节奏的逐渐加速,通过肉眼检测的热轧钢条质量不稳定等缺点逐渐暴露,并且提高热轧钢条表面缺陷分类的精度,降低缺陷分类的人工干预程度,还可以带来显著的经济和社会效益。因此,需要建立一个准确、实时的检测装置来充分应用于热轧钢条的生产过程,以便及时识别和准确地判断缺陷信息,分析缺陷原因,并及时进行调整,从而提高热轧钢条的表面质量。

近年来,机器视觉和深度学习已经成为一个流行的话题。经典的神经网络算法的预选框是基于选择性搜索的,在机器学习中使用支持向量机和卷积神经网络(convolution neural networks, CNN)^[1],在一定程度上提高了识别精度。然而,选择性的搜索不仅耗时,并且传统分类器的检测效果也不好,无法满足当前的工业需求,因此在实际的目标检测实验中应用是不可能的。目前,基于深度学习的目标检测算法有两种框架类型:两阶目标检测器和单阶目标检测器。朱阳光等^[2]详细地讲述了深度学习模型的渗透性、识别过程和各自的优势。对于两阶段的检测框架, Girshick 等^[3]提出了第一个基于 CNN 的目标检测框架 RCNN (region convolution neural network)^[4]。与传统算法相比,在 PASCAL VOC2012 测试集上实验得出的均值平均精度 (mean average precision, mAP) 值相较于传统算法高出 30%。随后,一系列优秀的两阶目标检测器相继出现,如 Fast RCNN、Faster RCNN 和 Mask RCNN。尽管两阶目标检测器检测精度较高,但通常速度很慢,并不能在实际的工业生产过程中使用。对于单阶目标检测器, Liu 等^[5]采用了一种单阶多层检测器 SSD (single shot multibox detector)^[6],使用多层特征图进行检测分类与回归。Redmon 等^[7]提出的 YOLO (you only look once)^[8]将目标检测过程作为回归任务,从而大大加快了速度。之后, Redmon 等对 YOLO 进行改进并提出了 YOLOv2^[9] 版本,大大提高了检测速度,但并不适用于小目标检测。因此, Redmon 等人又基于 SSD 的思想提出了 YOLOv3^[10] 版本, YOLOv3 算法相较于 YOLOv2 进行了充分的优化,提高了对小目标的检测能力。YOLOv4^[11] 在 YOLOv3 的基础上又增加了一些实用的技巧,使精度与速度有了进一步提升。

综上所述,为了在检测精度和速度之间达到良好的平衡,本文选取 YOLOv4 作为基础网络,并为解决目前的 YOLOv4 检测精度不高、对小范围信息的表现较差等问题,提出了一种针对热轧钢条表面缺陷的改进目标实时检测算法。与原 YOLOv4 相比,本文提出的该方法具有保证较快的检测速度和较高的 mAP 值,可以准确地识别缺陷的位置和形状,能够满足当前工业生产需求。

1 YOLOv4 算法

YOLOv4 模型是一种端到端的目标检测算法,相较于 YOLOv3 模型吸收了近些年来 CNN 领域中很多优秀的优化策略,包括在主干网络、数据处理、激活函数、网络训练、损失函数等方面,很好的结合了精度与速度,使网络性能得到了明显改善。YOLOv4 主要由 3 大部分组成,即主干特征提取网络、外加模块与特征处理层。

1.1 主干特征提取网络

YOLOv4 在 YOLOv3 模型的基础上将其主干提取网络 DarkNet53 替换为 CSPDarknet53, CSPDarknet53 与 DarkNet53 相比,在保证准确率的同时很大程度地减少了计算量。对于输入维度为 $416 \times 416 \times 3$ 的图像,每次通过残差块处理后,特征图的通道数会变为原来的两倍,而尺寸会缩小为原本的 $1/2$ 。为了避免特征图像通过多次卷积后丢失语义信息, CSPDarknet53 最后以 3 个不同维度的特征图来作为输出,这 3 个维度分别是 $13 \times 13 \times 1024$ 、 $26 \times 26 \times 512$ 和 $52 \times 52 \times 256$ 。

1.2 外加模块

为了使先验框提取出更有效的特征用于分类与回归任务上, YOLOv4 在 CSPDarknet53 的基础上增加了空间金字塔池化网络 (spatial pyramid pooling networks, SPPNet)^[12] 和路径聚合网络 (path aggregation network, PANet)^[13] 两个模块。SPP 是通过 4 个大小分别是 13×13 、 9×9 、 5×5 和 1×1 的池化核对输入图像进行了最大池化,并将每个池化后的结果进行堆叠后再输出。PAN 对深层特征进行上采样,将特征图维度扩大为原来的两倍后与上一层同纬度的特征进行堆叠,另外对浅层特征进行下采样,将特征图缩小为原来的 $1/2$ 后与下一层同纬度的特征进行堆叠。最后,以输入维度为 $416 \times 416 \times 3$ 为例,在通过 SPP 和 PAN 这两个外加模块处理后,模型最终输出的特征维度大小分别是 $13 \times 13 \times 75$ 、 $26 \times 26 \times 75$ 和 $52 \times 52 \times 75$ 。

1.3 特征处理层

特征处理层的主要作用则是通过对输出特征进行解码来得到预测结果。在经过 YOLOv4 模型的处理后,输出了 3 个 $n \times n \times 75$ 维的特征进行预测,然后将获取到的图片分成 $n \times n$ 个网格,之后通过确认每个网格是否真的存在物体再进行下一步操作。对于存在物体网格则构成先验框,然后根据特征提取出的目标信息再对先验框进行调整。最后,将每个特征层提取的目标信息进行解码,选取其中置信度最高的作为最后的预测结果。

2 改进的 YOLOv4 算法

本文以 YOLOv4 算法为基础,为了提高模型的性能,解决 YOLOv4 网络对目标特征提取不够充分以及梯度消失,从而导致易漏检、误检等问题,本文以更轻量

级深层神经网络 MobileNetv3 为基础网络替换了原 YOLOv4 中的原主干网络 CSPDarknet53。改进后的 YOLOv4 网络结构图如图 1 所示。其中,CBM 是由卷积层(CONV)、BN(batch normalization)以及 Mish 激活函数组成;CBL 由 CONV、BN 以及 Leaky ReLU 激活函数组成。

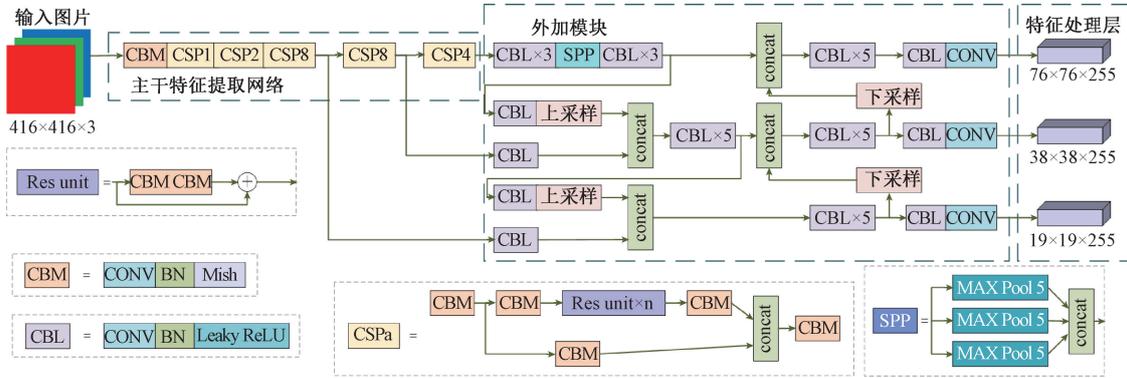


图 1 改进后的 YOLOv4 网络结构

Fig. 1 The improved YOLOv4 network architecture

除此之外,本文还使用了 K-Means 聚类算法进一步调整先验框的尺寸,使模型能够对小目标检测更加友好,提高对热轧钢条不同缺陷类型的检测精度。并且为了抑制正负样本输出差值,本文重新定义了置信度损失函数来解决过拟合的问题,最终能够使得网络模型的泛化能力更强。

2.1 替换特征提取网络

本文将 YOLOv4 主干网络 CSPDarknet53 替换为 MobileNetv3, MobileNetv3 是根据深度可分离卷积 MobileNet 模型改进的第 3 版,与 YOLOv4 的主干网络 CSPDarkNet53 相比程度上减小了体积,但依旧能够保持较强的特征提取能力,是一种更加轻量级的深层神经网络。

MobileNetv3 继承了 MobileNetv1 的深度可分离卷积,并且与 MobileNetv2 一样都具有线性瓶颈的残差结构,为了更好地适应不同场景,MobileNetv3 还推出了 large 与 small 两种版本。MobileNetv3 改进的内容具体有以下几个部分。

1) 修改尾部结构

MobileNetv3 利用平均池化将特征图大小由 7x7 降到了 1x1 之后再提高维度,如此便能减少了 7x7 即 49 倍的计算量,一定程度上加快了检测速度。除此之外,MobileNetv3 还去除了前面纺锤型卷积中的 3x3 和 1x1 卷积,使得计算量进一步减少,再次提高了模型的检测效率。

2) 更改非线性变换

MobileNetv3 使用 h-swish 将 swish 进行替换,将原来的 sigmoid 函数替换为 $\text{ReLU}_6(x+3)/6$,不仅能够减少由于网络层数的增加而造成网络梯度消失等问题,还能够

有效减少参数增长,很大程度地减少了运算量,提高模型性能,加快了模型检测效率,在速度上做了进一步优化,尤其是在深层网络中表现更加明显,在量化模式下大约能够提高 15% 的效率。

2.2 使用 K-Means 聚类算法调整先验框

YOLOv4 是采用先验框(anchor)的方式来进行目标位置预测,因此对先验框的设置是至关重要的。YOLOv4 原本的先验框是针对 COCO 数据集中 80 种物体类别设计的,然而本文数据集中目标缺陷的宽高比差异较大,默认的先验框并不适用于本文热轧钢条表面缺陷,最终会导致检测效果较差。

因此,为了提高模型对热轧钢条表面缺陷检测的精确度,本文使用 K-means 聚类算法对先验框尺寸进行调整。K-means 算法是目前机器学习中常用的聚类算法,它的主要思想是在给定 K 值和 K 个初始类簇中

心点的情况下,把每个点分到离其最近的类簇中心点所代表的类簇中,在所有点分配完之后,根据一个类簇内的所有点重新计算该类簇的中心点,然后再迭代的进行分配点和更新类簇中心点的步骤,直至类簇中心点的变化很小,或者达到指定的迭代次数^[14]。

K-means 聚类算法计算公式如式(1)所示。

$$E = \sum_{i=1}^k \sum_{a \in C_i} \|a - \mu_i\|^2 \quad (1)$$

其中, $\mu_i = \frac{1}{|C_i|} \sum_{a \in C_i} a$ 是簇 C_i 的均值向量, E 值越小代表簇内样本是相似度越高^[15]。

通过重新使用 K-means 聚类算法对本实验的数据集进行聚类分析,并且对于深层特征图,使用较大的先验框进行检测;对于浅层特征图,使用较小的先验框进行检测。最终得到的先验框尺寸如表 1 所示。

表 1 改进的 YOLOv4 特征图先验框尺寸
Table 1 The prior frame dimension of the improved YOLOv4

特征图尺寸	感受野	先验框
19×19	大	(37,77), (46,160), (64,392)
38×38	中	(92,73), (98,141), (137,293)
76×76	小	(174,156), (319,401), (366,54)

2.3 重新定义置信度损失函数

YOLOv4 的损失函数主要分为 3 个部分,分别是边界框回归损失、置信度损失以及分类损失。相较于 YOLOv3, YOLOv4 对边界框回归损失进行了改进,将 YOLOv3 原始的均方误差 (mean square error, MSE) 替换为 CIoU (complete intersection over union),然而在置信度损失与分类损失上并没有做实质性的变动。为了解决正负样本不平衡的问题,本文对置信度损失进行了重新定义。

由于 YOLOv4 最后是生成 3 个大小不同的先验框进行预测,然而不同尺度的先验框测出的正负样本比是不一样的,使用相同的损失系数 α 和 γ 在不同尺度的先验框下预测,效果不尽相同,导致模型很难得到最优的参数值,因此本文提出一种多尺度损失函数来适用于 YOLOv4 中不同尺度的先验框。本文是以先验框尺度为 38×38 为参照点,对尺寸为 19×19 和 76×76 的损失系数增加不同的尺度平衡系数 β ,假设 A_1, A_2, A_3 是 3 个不同尺度下的正负样本比, β 的取值如式(2) 所示。

$$\beta = \begin{cases} \frac{A_2}{3A_1}, & M = 76 \\ 1, & M = 38 \\ \frac{3A_2}{A_3}, & M = 19 \end{cases} \quad (2)$$

由式(2)可知, β 是通过正负样本比自适应选取来平衡不同的尺度。本文通过在正样本置信度损失前增加样本平衡系数 α , 在负样本置信度损失前增加系数 $(1 - \alpha)$, 从而解决正负样本极其不平衡的问题,其中, α 取值越大,正样本权重越高,则负样本权重降低; α 取值越小,负样本权重越高,则正样本权重降低。除此之外,为了解决训练集中简单样本过多从而导致对复杂样本识别率不高的问题,本文还对正样本增加调制系数 $(1 - C_x^y)^\gamma$, 对负样本增加调制系数 $(C_x^y)^\gamma$ 。假设 M 表示

网格数,每个尺度的特征图均划分为 M^2 个网格, A 则是每个网格产生的先验框个数, I_{xy}^{obj} 表示是否为正样本, I_{xy}^{noobj} 表示是否为负样本, \hat{C}_x^y 代表真实的置信度, C_x^y 代表预测的置信度,改进后的置信度损失如式(3) 所示。经实验证明,当 α 取值为 0.6, γ 取值为 2 时,本文的实验结果最佳。

$$Loss_c = - \sum_{x=0}^{M^2} \sum_{y=0}^N I_{xy}^{obj} [\hat{C}_x^y \beta \alpha (1 - C_x^y)^\gamma \log(C_x^y) + (1 - \hat{C}_x^y) \log(1 - C_x^y)] - \sum_{x=0}^{M^2} \sum_{y=0}^N I_{xy}^{noobj} [\hat{C}_x^y \log(C_x^y) + (1 - \hat{C}_x^y) (1 - \beta \alpha) (C_x^y)^\gamma \log(1 - C_x^y)] \quad (3)$$

3 实验结果与分析

3.1 实验数据与平台

本文数据集采用了东北大学宋教授等公开的钢表面缺陷数据集,主要包括了 NEU-CLS^[16] 和 NEU-DET^[17] 两部分,前者是用于缺陷分类,而后者用于缺陷识别。该数据集包含了结疤 (rolled-in_scale), 补丁 (patches), 点蚀 (pitted_surface), 杂质 (inclusion) 与划痕 (scratches) 这 5 种缺陷类型,对于每种缺陷类型都有 300 张图片可用于目标检测。

本文的实验运行环境如表 2 所示。

表 2 实验运行环境
Table 2 Experiment running environment

类别	环境条件	类别	环境条件
CPU	Intel (R) Core i7-9750H@2.60 GHz	CUDA 版本	CUDA 9
显卡	Nvidia GeForce GTX 1660 Ti	深度学习框架	PyTorch
内存	16 GB	运行环境	PyCharm
操作系统	Windows 1 064 位	脚本语言	Python 3.6

3.2 评价指标

mAP 是目标检测中常用的评价指标,本文则以 mAP 和 Fps 作为本实验的衡量精度指标。 $P-R$ 曲线图是以查全率 R (recall) 为横轴、查准率 P (precision) 为纵轴作图,而 mAP 值则是通过得到各类别的 $P-R$ 曲线图中面积 AP 的总和再进行平均的过程。 mAP 的计算流程如下所示,查准率 P 与查全率 R 的计算分别如式(4)、(5) 所示; $P(R)$ 为查准率 P 关于查全率 R 的 $P-R$ 曲线函数,其计算如式(6) 所示; AP 的计算如式(7) 所示以及 mAP 的计算如式(8) 所示。

$$P = \frac{TP}{TP + FP} \quad (4)$$

$$R = \frac{TP}{TP + FN} \quad (5)$$

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

$$AP = \int_0^1 P(R) dR \quad (7)$$

$$mAP = \frac{\text{sum}(AP)}{n} \quad (8)$$

其中, TP 为真正例, FN 为假反例, FP 为假正例, TN 为真反例, $\int_0^1 P(R) dR$ 为 P - R 曲线图下面积, $\text{sum}(AP)$ 为所有类别 AP 值之和, n 为所有类别总数, 显然有 $TP + TN + FP + FN =$ 样本总数。查准率 P 与查全率 R 通常都是矛盾的, 查准率 P 高时, 查全率 R 通常偏低; 但查全率 R 高时, 查准率 P 通常偏低。然而在评判模型性能时, 用单一的交并比 (intersection over union, IoU) 来衡量其精度并不准确, 因此, 就需要计算 P - R 曲线下的面积 AP 来求得本文所用的评价指标 mAP 。因为本文数据集中缺陷类别数为 5, 所以 n 这里取值为 5。

3.3 实验分析

本文以 8:1:1 的比例将数据集划分为训练集、测试集与验证集。训练中的 epochs 设为 100, batch size 与 mini-batch size 分别设置为 4 和 2, 最初的学习率设置为 0.001, 最终的学习率设置为 0.000 01, 训练时长约 300 min, 以及测试中的置信度阈值取值为 0.5。最终训练的损失函数曲线图如图 2 所示。

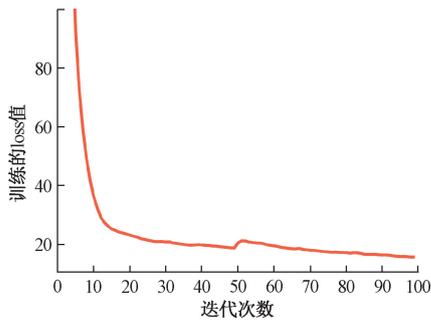


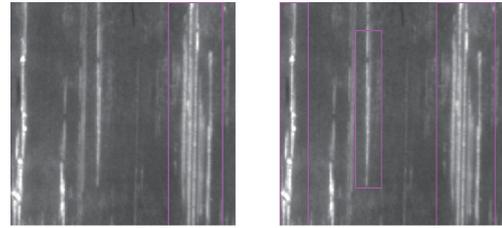
图 2 训练的 loss 变化曲线图

Fig. 2 Loss variation curve for training

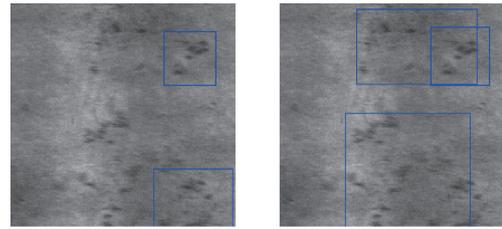
从图 2 中可以看出, 训练总共迭代了 100 次, 随着训练的进行, 本文算法模型的 loss 值在迭代 20 次之前持续下降, 在迭代超过 20 之后基本趋于稳定, 检测模型达到收敛。

本文从每个缺陷类别中各选取部分测试集作为检测效果图, 如图 3 所示。图 3 中根据缺陷类别数共分为 5 组数据, 这 5 组数据分别代表不同的缺陷类型, 其中每

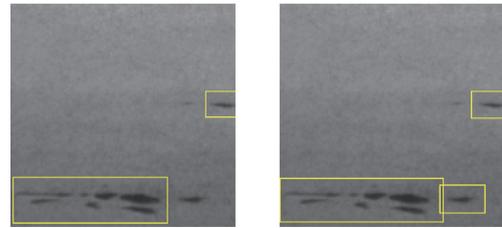
组数据的左侧展示的是原 YOLOv4 的检测效果图, 右侧展示了改进后 YOLOv4 算法的检测效果图。



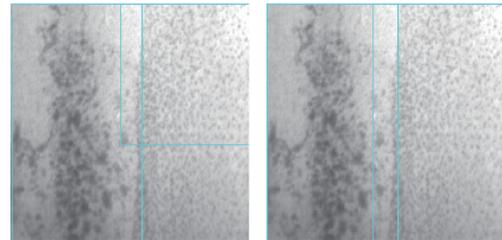
(a) 划痕
(a) Scratches



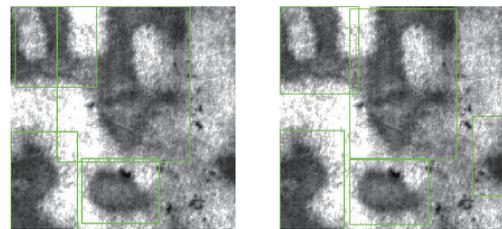
(b) 结疤
(b) Rolled-in_scale



(c) 杂质
(c) Inclusion



(d) 点蚀
(d) Pitted_surface



(e) 补丁
(e) Patches

图 3 检测效果对比

Fig. 3 Comparison of test results

从图 3 显示的检测效果图中可以看出, 原 YOLOv4 算法的检测结果在 (a)、(c) 以及 (e) 组中均出现了漏检的情况并且对缺陷的识别度较低, 在 (b)、(d) 组中还出

现了不能把存在缺陷的部分完全检测出来的情况,而改进后的 YOLOv4 算法检测效果相较于原算法有明显的改善。改进后的 YOLOv4 算法不仅对小目标检测更加友好,并且对中大型目标的识别率也有显著提高,对目标定位也更加准确,很少会出现误检与漏检的情况。

图 4 则显示了算法改进前后各类别的 P - R 曲线图。在图 4 中,图 4(a) 组中数据展示的是原 YOLOv4 算法测出的各类别 P - R 曲线图,而图 4(b) 组中数据展示的是改进后 YOLOv4 算法测出的各类别 P - R 曲线图。从图 4 中可以看出,改进后的算法相较于原 YOLOv4 算法,各缺陷

类型的 AP 值都有不同程度的提升,尤其是对于 rolled-in_scale 类别, AP 值提升了 19.52%,表明本文算法对于结疤类缺陷的识别能力大幅提升,极大程度的减少了各缺陷类别检测的精度差,模型性能得到显著提高。本文对 inclusion、pitted_surface、scratches 这 3 种类别缺陷的识别能力都有很大程度的提升, AP 值分别提升了 5.7%、6.02% 和 6.85%,相比于原算法模型均提升了 5% 以上。最后,本文对于 patches 类别缺陷的识别能力也有一定的提升, AP 值从 92.53% 提升到了 94.14%,达到了较高的识别精度。

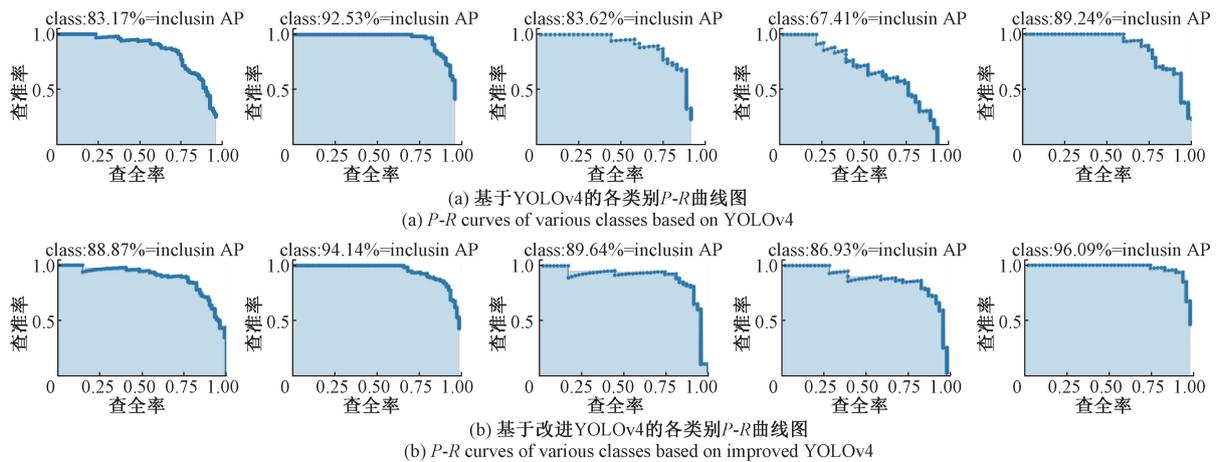


图 4 P - R 曲线图对比

Fig. 4 P - R curve chart comparison

为了进一步验证本文方法对 YOLOv4 算法进行改进的有效性,通过使用不同的改进策略来进行多组实验对比,最终得到的实验对比结果如表 3 所示。从表 3 中可以明显看出,模型 1 为原 YOLOv4 算法;模型 2 是将 YOLOv4 主干网络替换为 MobileNetv3 后的算法,通过替换主干网络,模型的精度与速度都有一定程度的提高;模型 3 是在模型 2 的基础上修改了 MobileNetv3 的尾部结构,减少了模型的计算量,很大程度地提升了模型检测效

率;模型 4 在模型 3 的基础上将 MobileNetv3 更改非线性变换后的算法,不仅减少了参数的增长,加快了速度,并且对精度也有很大的提高,尤其是对于 rolled-in_scale 类缺陷, AP 值从 69.08% 提升到 79.55%;模型 5 是在模型 4 的基础上增加了 K-means 聚类算法。实验表明,调整先验框后,模型更加适用于本实验数据集,能够有效提升检测效率,提高精度;而模型 6 又将模型 5 的置信度损失函数进行重新定义,通过增加不同尺度的平衡系数以及

表 3 各模型实验结果对比

Table 3 Comparison experimental of each model

模型	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
Inclusion/%	83.17	86.27	86.39	88.13	86.81	88.87
patches/%	92.53	91.38	92.87	90.55	93.16	94.14
pitted_surface/%	83.62	85.24	82.19	90.63	87.31	89.64
rolled-in_scale/%	67.41	71.76	69.08	79.55	77.26	86.93
scratches/%	89.24	91.05	93.65	93.74	94.18	96.09
mAP/%	83.19	85.14	86.28	88.52	89.66	91.13
Fps/(f·s ⁻¹)	21.87	22.16	24.43	26.15	26.37	26.39

正负样本调制系数,在保证检测速度的情况下,极大地提高了检测精度,改善了模型性能。通过表 3 的对比实验可知,与原 YOLOv4 算法相比,通过改进后的 YOLOv4 算法的 mAP 值从最初的 83.19% 增长到了 91.13%, Fps 值从 21.87 f/s 增长到 26.39 f/s,有明显的提升,基本能够满足当前的工业需求。

通过本文算法与传统 YOLOv4 算法相比,传统的 YOLOv4 算法对热轧钢条图像目标缺陷检测的检测率很低,很多小目标物体容易被漏检,并且也会存在误检的现象。而本文算法对图像中的目标缺陷检测效果远优于传统的 YOLOv4 算法,极大地避免了误检甚至漏检的情况。

本文将改进后的 YOLOv4 算法模型应用到型号为 S500MC 高强度结构钢的实际工业生产中,S500MC 钢材如图 5 所示。S500MC 产品通常用于制造汽车大梁、横梁、滚型车轮、汽车传动轴管、汽车桥壳等结构,其产品的表面质量直接影响到产品使用的安全性和美观程度,因此确保 S500MC 钢材的表面质量是非常有必要的。在 S500MC 型号高强度结构钢的生产过程中,经常会产生 pitted_surface、patches 等类型的缺陷,严重影响到产品表面质量,导致用户退货率增大甚至影响到企业形象。针对 S500MC 钢材表面缺陷,通过在生产现场进行详细调研,本文采集部分典型的缺陷样本,检测效果如图 6 所示。

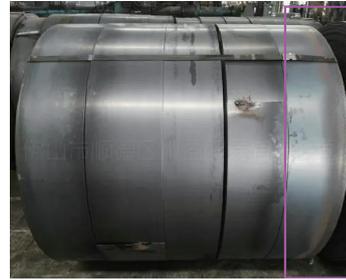


图 5 S500MC 钢材示意图

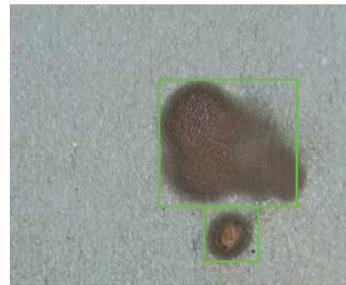
Fig. 5 Schematic diagram of S500MC steel

通过在生产现场对 S500MC 钢材样品的表面缺陷进行观察,片状划伤的宏观形貌如图 6(a) 组样本 1 所示,通过改进后的 YOLOv4 算法模型识别出主要缺陷位置并判别为 scratches 类型缺陷。图 6(b) 组样本 2 表面有明显一大一小两块深色块状缺陷,通过模型最终判定为 patches 类型缺陷并将这两块缺陷具体位置圈出。图 6(c) 组样本 3 表面多处呈条带状密集分布,并无裂纹延伸倾向,是典型的杂质类缺陷,因此模型以实线框将缺陷位置圈出并判断出此部分缺陷为 inclusion 类型。从图 6 的检测效果看,漏检率还有很大的提升空间;从检测

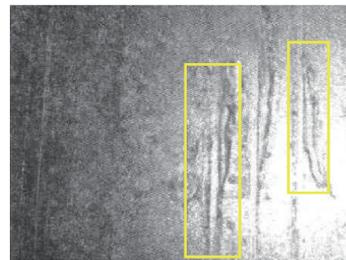
结果上看, mAP 值为 91.13%,由于检测检测特定材质的钢材,缺陷类型相对固定,所以检测效率相对提升,因此本文算法能够满足实时检测的生产需要。



(a) 样本1
(a) Sample one



(b) 样本2
(b) Sample two



(c) 样本3
(c) Sample three

图 6 S500MC 钢材表面缺陷样本检测效果

Fig. 6 S500MC test effect steel surface defect sample

4 结 论

本文针对热轧钢条表面缺陷检测问题,提出了基于传统 YOLOv4 的改进算法。为了加快检测速度、加强对检测目标特征提取以及减少梯度消失的问题,本文以更轻量级深层神经网络 MobileNetv3 替换了 YOLOv4 原本的基础网络 CSPDarknet53,并且采用 K-Means 聚类算法调整先验框尺寸以及重新定义了置信度损失函数,进一步提高了模型的准确率和检测效率,减少漏检和误检的现象。

然而个别缺陷类型容易受环境光照影响或者背景剧烈变化,导致的识别精度依旧不高,因此在后续的工作中,可以进一步扩充数据集,并对网络模型继续优化,提升模型的识别精度和速度。

参考文献

- [1] KHAN A, CHEFRANOV A, DEMIREL H. Image scene geometry recognition using low-level features fusion at multi-layer deep CNN [J]. Neurocomputing, 2021(440): 111-126.
- [2] 朱阳光,刘瑞敏,黄琼桃. 基于深度神经网络的弱监督信息细粒度图像识别 [J]. 电子测量与仪器学报, 2020,34(2): 115-122.
- ZHU Y G, LIU R M, HUANG Q T. Fine-grained image recognition of weak supervisory information based on deep neural network [J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(2): 115-122.
- [3] GIRSHICK R, DONAHUE J, DARRELL T, et al. Region-based convolutional networks for accurate object detection and segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016,38(1): 142-58.
- [4] ZHOU C H, ZHOU J Y, YU C, et al. Multichannel sliced deep RCNN with residual network for text classification [J]. Chinese Journal of Electronics, 2020, 29(5): 880-886.
- [5] LIU X, LI Y, SHUANG F, et al. ISSD: Improved SSD for insulator and spacer online detection based on UAV system [J]. Sensors, 2020, 20(23): 6961.
- [6] JIN L, LIU G D. An approach on image processing of deep learning based on improved SSD [J]. Symmetry, 2021, 13(3):495.
- [7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE Computer Society, 2016: 779-788.
- [8] QING Y H, LIU W Y, FENG L Y, et al. Improved YOLO network for free-angle remote sensing target detection [J]. Remote Sensing, 2021, 13 (11): 2171-2171.
- [9] XU K, WANG X Y, LIU X Y, et al. A dedicated hardware accelerator for real-time acceleration of YOLOv2 [J]. Journal of Real-Time Image Processing, 2020, 18(3): 481-492.
- [10] 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. Using active thermography for defect detection of aerospace electronic solder joint base on the improved Tiny-YOLOv3 network [J]. Chinese Journal of Scientific Instrument, 2020, 41(11): 42-49.
- [11] ZHANG Z, XIA S H, CAI Y X, et al. A soft-YOLOv4 for high-performance head detection and counting [J]. Mathematics, 2021, 9(23): 3096.
- [12] XIN H J, CHEN Z B, WANG B Y. PCB Electronic component defect detection method based on improved YOLOv4 algorithm [J]. Journal of Physics: Conference Series, 2021, 1827(1):012167.
- [13] YAO G, SUN Y J, WONG M P, et al. A real-time detection method for concrete surface cracks based on improved YOLOv4 [J]. Symmetry, 2021, 13 (9): 1716-1726.
- [14] LIU X, SANG X F, CHANG J X, et al. The water supply association analysis method in Shenzhen based on k-means clustering discretization and apriori algorithm [J]. PloS One, 2021, 16(8): e0255684.
- [15] HIEU L N, TRANG N T H, XUAN P N, et al. MCCVA: A new approach using SVM and k-means for load balancing on cloud [J]. International Journal on Cloud Computing: Services and Architecture, 2020, 10(3): 1-14.
- [16] LUO Q, FANG X X, SU J J, et al. Automated visual classification of frequency for flat steel surface: A survey, IEEE Transactions on Instrumentation and Measurement, 2020, 69(12): 9329-9349.
- [17] LUO Q, FANG X X, LIU L, et al. Automated visual defect detection for flat steel surface: A survey [J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(3): 626-644.

作者简介



刘艳菊,1998年于哈尔滨工业大学获得学士学位,2007年于齐齐哈尔大学获得硕士学位,2013年于哈尔滨理工大学获得博士学位,现为南京特殊教育师范学院教授、硕士生导师,主要研究方向为机器学习,缺陷检测和三维重建。

E-mail: yanjuliu@qqhru.edu.cn

Liu Yanju received her B. Sc. degree from Harbin Institute of Technology in 1998, received her M. Sc. degree from Qiqihar University in 2007, and received her Ph. D. degree from Harbin University of Science and Technology in 2013. She is currently a professor and a graduate advisor at Nanjing Normal University of Special Education. Her main research interests include machine learning, defect detection and three-dimensional reconstruction.



王秋霁, 2018 年于江苏师范大学文学院获得学士学位, 现为齐齐哈尔大学硕士研究生, 主要研究方向为机器学习, 计算机视觉和目标检测。

E-mail: wangqiuji727@163.com

Wang Qiujie received her B. Sc. degree from Jiangsu Normal University Kewen College in 2018. She is currently a master student at Qiqihar University. Her main research interests include machine learning, computer vision and object detecting.



赵开峰, 2018 年于济宁医学院获得学士学位, 现为齐齐哈尔大学硕士研究生, 主要研究方向为图像识别和图像分类。

E-mail: zkf969696@163.com

Zhao Kaifeng received his B. Sc. degree from Jining Medical University in 2018. He is

currently a master student at Qiqihar University. His main research interests include image recognition and image classification.



刘彦忠 (通信作者), 1994 年于齐齐哈尔大学获得学士学位, 2009 年于齐齐哈尔大学获得硕士学位, 现为齐齐哈尔大学副教授, 主要研究方向为机器学习, 误差分析和缺陷检测。

E-mail: 15146692464@163.com

Liu Yanzhong (Corresponding author) received his B. Sc. degree from Qiqihar University in 1994, and received his M. Sc. degree from Qiqihar University in 2009. He is currently an associate professor at Qiqihar University. His main research interests include machine learning, error analysis and defect detecting.