

基于改进 YOLOv3 算法的带钢表面缺陷检测

李维刚^{1,2}, 叶欣¹, 赵云涛¹, 王文波³

- (1. 武汉科技大学冶金自动化与检测技术教育部工程研究中心, 湖北武汉 430081;
2. 武汉科技大学高温材料与炉衬技术国家地方联合工程研究中心, 湖北武汉 430081;
3. 武汉科技大学理学院, 湖北武汉 430081)

摘 要: 针对热轧带钢表面缺陷检测中存在的检测速度慢、检测精度低等问题, 提出了一种改进的 YOLOv3 算法模型. 使用加权 K-means 聚类算法来优化确定先验框参数, 提高先验框 (priors anchor) 与特征图层 (feature map) 的匹配度; 同时, 调整 YOLOv3 算法的网络结构, 融合浅层特征与深层特征, 形成新的大尺度检测图层, 提高网络对带钢表面缺陷的检测精度. 实验结果表明, 改进后的 YOLOv3 算法在 NEU-DET 数据集上平均精度均值达到了 80%, 较原有的 YOLOv3 算法提高了 11%; 同时检测速度保持在 50fps, 优于目前其它深度学习带钢表面缺陷检测算法.

关键词: 目标检测; 带钢表面缺陷; YOLOv3; 加权 K-means

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 0372-2112 (2020)07-1284-09

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2020.07.006

Strip Steel Surface Defect Detection Based on Improved YOLOv3 Algorithm

LI Wei-gang^{1,2}, YE Xin¹, ZHAO Yun-tao¹, WANG Wen-bo³

- (1. *Engineering Research Center for Metallurgical Automation and Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan, Hubei 430081, China;*
2. *National-Provincial Joint Engineering Research Center of High Temperature Materials and Lining Technology, Wuhan University of Science and Technology, Wuhan, Hubei 430081, China;*
3. *College of Science, Wuhan University of Science and Technology Wuhan, Hubei 430081, China*)

Abstract: To solve the problem of slow speed and low accuracy in the surface defect detection of hot rolled strips, an improved YOLOv3 algorithm is proposed. Firstly, the weighting K-means clustering algorithm is put forward to optimize priors anchor's parameters, which can improve the match between priors anchor and feature map. Secondly, the improved network structure of the YOLOv3 algorithm is proposed to improve the detection accuracy, whose shallow features and deep features are combined to form the new large-scale inspection layer. The experiments are carried out on the NEU-DET dataset, the results show that the average accuracy of the improved YOLOv3 algorithm is 80%, which is 11% higher than that of the original algorithm; the detection speed is 50fps, which is faster than other strip surface defect detection algorithms based on deep learning.

Key words: object detection; strip steel surface defect; YOLOv3; weighting K-means

1 引言

带钢热连轧生产过程中, 由于生产工艺的限制, 带钢表面会产生麻点、夹杂、刮伤等缺陷^[1,2]. 这些缺陷对热轧带钢产品的抗腐蚀性和疲劳强度等性能影响较大. 近年来, 随着卷积神经网络等深度学习方法的流行^[3], 使用基于深度学习的目标检测算法对带钢表面

缺陷进行检测的思路逐渐取代了传统机器学习方法. 因此, 研究一种高速、准确的端到端带钢表面缺陷检测算法对于提升热轧带钢表面质量很有意义.

对于带钢表面缺陷问题, 传统方法是使用手工提取特征因子后再对缺陷进行分类^[4,5], 这一方法在检测精度、抗扰动性及实时性等方面难以满足实际生产需求. 近年来, 随着计算机技术、深度学习的发展, 深度神

神经网络已逐渐成为表面检测任务的主要方法。目前,带钢表面缺陷检测普遍使用卷积神经网络提取特征并分类^[6]。例如,文献[7]提出了一种端到端的卷积神经网络,实现了带钢表面缺陷的高精度分类;文献[8,9]使用对抗神经网络生成大量的未标注数据后再进行缺陷分类,解决了缺陷样本不足难以训练的问题。但以上检测模型只解决了带钢表面缺陷分类问题,并没有解决难度更大的缺陷定位问题。文献[10]使用 Defect Detection Network (DDN) 实现了端到端的带钢表面缺陷检测并取得了 70%~80% 的准确率,这一方法虽然具有较好的检测精度,但在检测速度上稍有不足,难以满足工业生产实时检测的需求。

目标检测的任务是找出图像中所有感兴趣的目标,确定它们的位置、大小以及类别信息^[11,12]。目前,基于深度学习的检测方法逐渐取代了传统的检测方法^[13]。现有的深度学习目标检测算法主要分为两类,一类是以 Faster R-CNN^[14]、Mask R-CNN^[15] 为代表的 two-stage 检测算法;另一类是以 YOLO^[16-18]、SSD^[19] 为代表的 one-stage 检测算法。two-stage 检测算法将检测任务分为两个阶段,首先使用区域候选网络 (RPN) 产生候选区域,然后使用检测网络检测候选区域的类别、位置,这种方法的准确度较高但检测速度稍慢^[20]。one-stage 检测算法不需要 RPN 阶段,可以直接得到检测结果,所以检测速度较快,但检测精度较低^[21]。这两类检测算法目前的应用都比较多^[22,23]。YOLO 算法的检测精度与检测速度之间较为平衡,最新的 YOLOv3 算法当输入为 416×416 时,在 COCO 数据集上检测精度 mAP50 为 55%,检测速度达到 34fps,是目前最为优秀的检测算法之一。

但是,作为一种 one-stage 算法,YOLOv3 存在检测精度不高的问题,本文采用改进的 YOLOv3 算法来进行带钢表面缺陷检测研究。使用加权 K-means 算法代替原来的 K-means 算法对目标样本进行聚类:首先分配聚类中心并设置每一个样本的权重参数,再进行聚类计算,避免了缺陷尺寸不平衡对聚类结果的影响,提高了先验框与特征图层的匹配度;同时调整 YOLOv3 的网络结构,增加了残差单元和一个大尺度的特征图层输出,提高了算法对带钢表面缺陷的检测能力。最后,使用 NEU-DET^[10] 数据集进行验证,实验结果表明:改进后的 YOLOv3 算法实现了端到端的带钢表面缺陷检测,具有较高的准确度和较好的实时性。

2 YOLOv3 算法原理

2.1 检测原理

YOLOv3 检测算法采用特征金字塔 (Feature Pyra-

mid Networks, FPN) 方式,提取多个不同尺度的特征图层 (feature map) 分别进行检测,提高算法对不同大小目标的检测能力。如图 1 所示,检测时首先将输入图片分割成 $S \times S$ 个网格,网格的宽、高记为 c_x, c_y ,在每个网格中输出对应于网格的相对中心点坐标 $(\sigma(t_x), \sigma(t_y))$ 和相对宽 t_w 、高 t_h ,通过所在网格的实际位置、宽、高得到最终的预测框 (对应红色实线)。

图 1 中虚线框为先验框, p_w 和 p_h 为先验框的宽、高,以图 1 中第 2 行第 2 列的网格为例,在该网格上输出的值通过式 (1) 得到预测框的中心点位置 (b_x, b_y) 和宽 (b_w) 、高 (b_h) :

$$\begin{aligned} b_x &= \sigma(t_x) + c_x & b_y &= \sigma(t_y) + c_y \\ b_w &= p_w e^{t_w} & b_h &= p_h e^{t_h} \end{aligned} \quad (1)$$

$$\text{IoU} = \frac{\text{Box}_1 \cap \text{Box}_2}{\text{Box}_1 \cup \text{Box}_2} \quad (2)$$

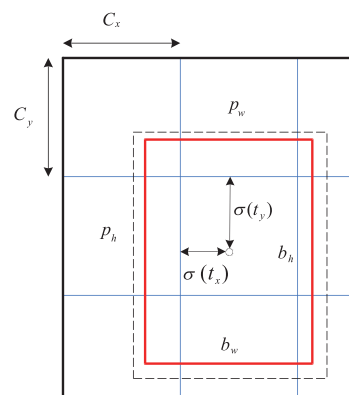


图1 预测框示意图

最终参与损失函数计算的值包括预测框的宽、高,中心点的位置;预测框的置信度以及分类信息。预测框的置信度是指预测框与标注框之间的交并比 (IoU),一般设置 $\text{IoU} > 0.7$ 的预测框为正例,即成功预测到目标; $\text{IoU} < 0.3$ 的目标为反例,即背景;忽视其它预测框,使用正例和负例计算损失函数,同时尽量保持正负例之间数量平衡。分类信息是指预测框中含有某一类别目标的概率。交并比计算公式如式 (2) 所示,其中, Box_1 为预测框, Box_2 为标注框。

2.2 损失函数

目标检测问题既要得到目标的类别,也要得到目标的位置信息。two stage 检测算法将损失函数分为两部分:对目标位置的回归与对目标类别的分类,这一做法不仅增加了训练难度,且无法实现端到端的检测。而 YOLOv3 只使用一个损失函数同时解决了分类与位置两个问题,实现了端到端的检测,同时加快了算法检测速度。损失函数具体如式 (3) 所示:

$$\text{Loss} = \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B l_{ij}^{\text{obj}} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2]$$

$$\begin{aligned}
& + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B l_{ij}^{\text{obj}} [(\sqrt{\omega_i} - \sqrt{\hat{\omega}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2] \\
& + \sum_{i=0}^{S^2} \sum_{j=0}^B l_{ij}^{\text{obj}} (C_i - \hat{C}_i)^2 + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B l_{ij}^{\text{noobj}} (C_i - \hat{C}_i)^2 \\
& + \sum_{i=0}^{S^2} l_i^{\text{noobj}} \sum_{c \in \text{classes}} [p_i(c) - \hat{p}_i(c)]^2
\end{aligned} \quad (3)$$

其中, S 为图像的划分系数, B 为每个网格中所预测的预测框个数, C 为总分类数, p 为类别概率, x_i, y_i, w_i, h_i 为第 i 个网格中预测框的中心点横、纵坐标以及宽、高, λ_{coord} 为权重系数, λ_{noobj} 为惩罚权重系数。

3 YOLOv3 算法的改进

3.1 加权 K-means 聚类算法

在原始的 YOLOv3 算法中, anchors 参数由一组不同大小的先验框组成, 先验框的宽、高通过使用 K-means 算法对训练集中所有检测目标进行聚类得到, 一般设置聚类中心个数为 9, 将聚类结果按照大小顺序分配到 3 个 feature map 上作为先验框. 使用聚类中心与样本之间的最大交并比来评价聚类结果, 目标函数 D 如式(4)所示:

$$D = \min \sum_{i=1}^n \sum_{j=1}^k \left[1 - \frac{\text{box}_i \cap \text{cen}_j}{\text{box}_i \cup \text{cen}_j} \right] \quad (4)$$

其中, box_i 为样本的实际宽、高, cen_j 为第 j 个聚类中心, n 为样本数, k 为选取的聚类中心个数。

YOLOv3 算法使用多尺度特征来检测目标, 使用小尺度的 feature map 检测大目标, 大尺度的 feature map 检测小目标. 在分配先验框时也是遵循这一原则, 将最大的 3 个先验框分配给最小尺度的 feature map. 不同尺度 feature map 上的先验框应具有明显的大小差距, 以便于检测不同大小的目标. 但分配先验框是在聚类完成之后, 在聚类过程中并没有考虑先验框与 feature map 的对应关系, 由于数据集不平衡, 得到的聚类结果不能与 feature map 匹配, 降低了算法的检测能力。

本文采用加权 K-means 聚类算法, 更改聚类算法步骤的顺序, 先确定每个聚类中心所属的检测尺度, 再进行聚类. 在聚类过程中, 设置每一个样本的权重, 将其参与目标函数的计算, 降低样本不平衡对聚类结果的影响. 新的目标函数 D' 如式(5)所示:

$$D' = \min \sum_{i=1}^n \sum_{j=1}^k \left[1 - \omega_{ij} \frac{\text{box}_i \cap \text{cen}_j}{\text{box}_i \cup \text{cen}_j} \right] \quad (5)$$

其中, ω_{ij} 为第 i 个样本与第 j 个聚类中心的权重, l 为聚类中心所属的中心点。

以原始 YOLOv3 算法为例, 设置 9 个聚类中心, 加权 K-means 算法具体的实现步骤如下:

(1) 设置 3 个 feature map 的中心点, 每 1 个中心点对应 3 个聚类中心。

(2) 求出每一个样本与中心点之间的距离, 距离计算公式如式(6)所示:

$$d_{il} = \sqrt{[\text{box}_i(1) - C_l(1)]^2 + [\text{box}_i(2) - C_l(2)]^2} \quad (6)$$

其中, $\text{box}_i(1)$ 、 $\text{box}_i(2)$ 为第 i 个样本的横、纵坐标, $C_l(1)$ 、 $C_l(2)$ 为第 l 个中心点的横、纵坐标。

(3) 求权重矩阵, 即每个样本对应每个聚类中心的权重, 属于同一个中心点的聚类中心共用相同的权重. 权重为样本到聚类中心距离 z-score 标准化后的相反数, 将距离较远的样本权重直接设置为 0, 降低数据不平衡样本的干扰. 式(7)为所有样本到某一中心点距离的均值, 式(8)为权重的计算公式:

$$\bar{d}_l = \frac{1}{n} \sum_{i=1}^n d_{il} \quad (7)$$

$$\omega_{il} = \max \left(0, -\frac{d_{il} - \bar{d}_l}{\sqrt{\frac{1}{n} \sum_{i=1}^n (d_{il} - \bar{d}_l)^2}} \right) \quad (8)$$

其中, l 为中心点, d_{il} 为样本 i 与中心点 l 的距离, \bar{d}_l 为所有样本与中心点 l 的平均距离。

(4) 迭代更新所有聚类中心, 直到聚类中心不再发生变化。

3.2 网络结构改进

YOLOv3 算法使用 Darknet-53 作为特征提取网络, Darknet-53 网络结构如表 1 所示, 网络借鉴了 Resnet^[24] 的思想, 交替使用 1×1 和 3×3 的卷积核构成残差块, 避免了网络层数增加后梯度消失的问题. 在每个卷积层后使用 Batch Normalization 和 Leaky relu, 提高网络的泛化性同时使网络易于训练. 整个网络中使用步长为 2 的卷积层代替池化层来调整输出尺寸, 避免了使用池化后小目标特征丢失。

相比前两个版本的 YOLO 算法, YOLOv3 使用多个尺度的 feature map, 可以对不同尺度的目标进行检测. feature map 的宽、高分别为原始输入的 $1/32$ 、 $1/16$ 、 $1/8$. 每个 feature map 通过特征图融合同时兼顾了浅层特征与深层特征, 提高了网络的检测能力. 这些做法在一定程度上解决了对小目标不敏感的问题, 但相比 two stage 算法(如 Faster R-CNN、Mask R-CNN 等), 在小目标的检测上还是存在不足. 在带钢表面缺陷检测的问题上, 不同种类的缺陷在形态、大小等方面差异很大, 若使用原始的 YOLOv3 网络进行检测, 微小缺陷的检测效果较差. 以大小为 256×256 的图片为例, 原始的 YOLOv3 网络最终划分的网格大小分别为 32×32 、 16×16 、 8×8 , 能够得到的最大预选框数量仅有 1344 个, 难以满足检测需求。

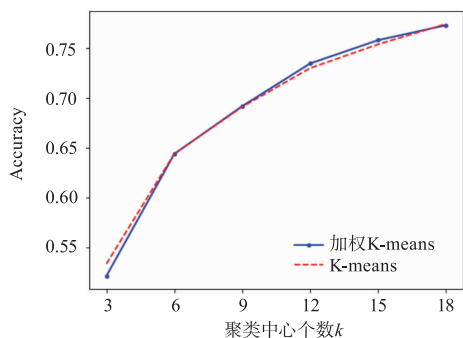


图3 先验框聚类结果

有统一的划分标准,增加了位置信息的检测难度.

4.2 先验框参数聚类

原始 YOLOv3 的先验框是由 VOC 数据集聚类得到,在图片大小、类型上都有很大的区别,有必要重新计算.首先计算每个样本与所有聚类中心的交并比,计算时将样本与聚类中心的左上角点坐标都设为(0,0);然后取每个样本与聚类中心交并比最大值的均值,作为最终评价聚类结果的准确率(Accuracy),计算公式如式(9)所示.

$$\text{Accuracy} = \frac{1}{n} \sum_{i=1}^n \max \left(\frac{\text{box}_i \cap \text{cen}_1}{\text{box}_i \cup \text{cen}_1}, \frac{\text{box}_i \cap \text{cen}_2}{\text{box}_i \cup \text{cen}_2}, \dots, \frac{\text{box}_i \cap \text{cen}_k}{\text{box}_i \cup \text{cen}_k} \right) \quad (9)$$

采用 K-means 聚类和加权聚类两种算法得到先验框参数,图 3 为两种算法在不同个数聚类中心下的对比,分别计算了聚类中心个数为 3、6、9、12、15、18 共 6 种情况下两种算法的准确率.可以看出,两种算法的结果差距不大,在聚类中心个数为 12、15 时,加权 K-means 算法的准确率稍高.在不平衡的样本中,单纯比较 Accuracy 并不能说明聚类好坏,我们希望聚类结果能够覆盖更多的面积,即先验框之间差距更大.表 2 为在聚类中心个数设置为 9 时,两种算法聚类结果在面积、宽度和高度上的对比.将聚类结果按大小分给 3 个 feature map 后,比较每个 feature map 之间先验框的差异性.可以看出,使用加权 K-means 算法后各 feature map 的先验框之间差异性更大,尤其是在尺度为 32×32 的 feature map 上,先验框的面积更小,对于小目标的检测能力更强.

表 2 先验框聚类结果对比图

特征图尺度	平均面积		平均宽度		平均高度	
	K-means	加权 K-means	K-means	加权 K-means	K-means	加权 K-means
32×32	1462	1184	31	27	48	46
16×16	5391	5123	51	52	117	109
8×8	16997	17649	124	116	124	156

4.3 实验结果分析

实验采用的硬件配置为 Core i7-9700K 处理器, RTX 2080 Ti 显卡,软件环境为 CUDA10.0 和 cuDNN9.1.在训练中,使用 Adam 优化器优化网络参数,共迭代 10000 代.在所有的检测中,使用实际目标与预测目标之间的交并比(IoU)来评价是否成功预测目标的位置,认为 $\text{IoU} > 0.5$ 的预测目标是成功预测到了目标位置.评价网络模型性能时需要同时兼顾精确率(Precision)和召回率(Recall),目标检测中一般使用平均精度均值(Mean Average Precision, mAP)来评价网络模型的性能.式(10)为准确率的计算公式,式(11)为召回率的计算公式:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (10)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (11)$$

其中,TP 为成功预测的正例,FP 为被模型误判为正例的负例,FN 表示被模型错误预测为负例的正例.

平均精度(Average Precision, AP)的定义为不同召回率下精确率的平均值,一般用来评价某一类的

检测精度. mAP 是所有目标类别检测精度的均值,一般用来评价网络模型的整体性能,计算公式如式(12)所示. fps 为网络模型每秒能够检测的图片数量,用来评价网络模型的检测速度,计算公式如式(13)所示.

$$\text{mAP} = \frac{\sum_{i=0}^n \text{AP}(i)}{n} \quad (12)$$

$$\text{fps} = \frac{\text{NumFigure}}{\text{TotalTime}} \quad (13)$$

其中 $\text{AP}(i)$ 为某一类的检测精度, n 为类别数, NumFigure 为检测图片总数, TotalTime 为检测总时间.

首先对比两种聚类算法,聚类得到先验框参数后,采用原始的 YOLOv3 算法进行检测,图 4 为两种聚类算法的结果对比.

可以看出,使用传统的 K-means 算法后不同缺陷的检测精度相差很大,麻点(PS)、斑块(Pa)和划痕(Sc)三类缺陷的平均精度较高,可以较好地检测出;但是裂纹(Cr)类缺陷的平均精度很低,大部分此类缺陷无法检测出.使用加权 K-means 算法后,大部分缺

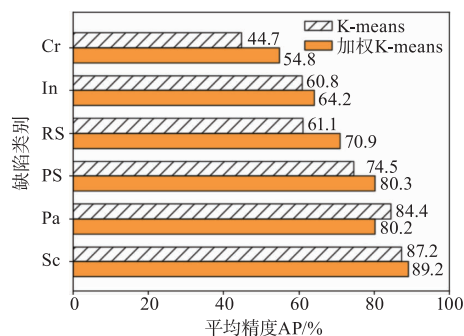


图4 不同聚类算法在原始YOLOv3下的精度对比

陷类别的平均精度都有所提升,尤其是裂纹(Cr)类从44.7%提高到54.8%,压入氧化铁皮(RS)类从60.8%提高到70.9%,对这些检测效果差的缺陷类别提升十分明显.使用加权K-means调整anchor参数后,由于先验框间尺寸差异加大,对于斑块(Pa)这类尺寸适中的缺陷的检测能力稍有下降,但是对于一些尺寸偏大或偏小的缺陷的检测能力则有明显加强.整体上来说调整后算法的检测精度比调整前更高;同时,这一改进有效解决了部分缺陷检测精度太低的问题,增强了算法的适用性.

图5为使用原始YOLOv3算法和改进后YOLOv3算法的检测结果对比,二者都选用了加权K-means算法进行聚类,改进后的YOLOv3算法需要设置12个先验框.从图中可以看出,使用改进后的YOLOv3算法进行检测时,由于feature map由3层增加到4层,候选框的数量也随着增加,所有类别缺陷的检测精度相较于原始YOLOv3算法都有所上升.尤其对于裂纹(Cr)类缺陷,检测精度由54.8%上升到70.0%,提高了15个百分点,有效解决了这一类缺陷检测精度低的问题;而且,所有类别缺陷的检测精度都超过了70%,最高的划痕(Sc)类缺陷的检测精度达到了91.7%.相比于原始YOLOv3算法,改进后的YOLOv3算法能够较准确的检测出所有类别缺陷,能够满足工业生产中带钢表面缺陷检测的精度要求.

表3为算法整体性能对比,先验框的聚类方法由K-means算法换成加权K-means算法后,检测精度mAP有一定提升.而改进后的YOLOv3算法在网络结构上做出了改变,在检测精度上有了很大的提高.相比原始的YOLOv3算法,改进后的算法在mAP上提高了7~9个百分点.最终,同时使用加权K-means聚类算法和改进的YOLOv3算法,在NEU-DET数据集上的平均精度均值达到80.1%,检测速度为50fps,能够快速、准确的检测出各类带钢表面缺陷.

图6为各类别缺陷样本检测结果图,图中蓝色框为人工标签,绿色框为预测框.

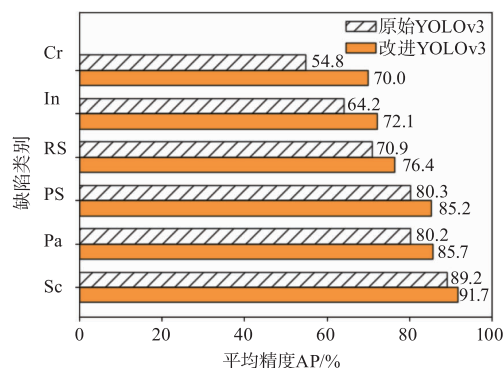


图5 加权K-means聚类后改进模型的精度对比

表3 YOLOv3 算法改进性能对比

网络结构	聚类算法	fps	mAP (%)
原始 YOLOv3	K-means	55	69.1
原始 YOLOv3	加权 K-means	55	73.3
改进 YOLOv3	K-means	50	77.4
改进 YOLOv3	加权 K-means	50	80.1

表4为不同算法在NEU-DET数据集上的性能对比,DDN^[10](Defect Detection Network)是一种应用于带钢表面缺陷的端到端two-stage检测算法,有特征提取网络分别为ResNet34和ResNet50的两个版本,是目前该数据集上表现最好的算法.比较各种算法在每一类缺陷上的检测精度,可以发现:

(1)相比原始的YOLOv3算法,改进后的算法在所有缺陷类别上的检测精度都有提高;其中Cr类缺陷的mAP从44.7%提高到70.0%,提升巨大.

(2)改进后的YOLOv3算法与以ResNet50为特征提取网络的DDN算法相比,Cr类缺陷的检测精度更高,但对于In类缺陷的检测精度则相对较低.

主要原因在于:相较于Faster R-CNN、DDN等two-stage算法,YOLOv3作为one-stage算法直接产生物体的类别概率和位置坐标值,在位置精度上相对较差,对于较为集中的夹杂(In)类缺陷容易产生误检,影响了最终的检测精度.

从各种算法的检测速度fps与整体检测精度mAP可以看出:(1)原始的YOLOv3算法在速度上相较于two stage算法优势较大,但其检测精度较低;(2)改进后的YOLOv3算法继承了原始算法的快速性,同时检测精度达到了80%,弥补了原始YOLOv3算法在精度上的不足;(3)改进后的YOLOv3算法的检测速度是DDN的3倍左右,而检测精度仅相差2%,有效提高了检测效率.

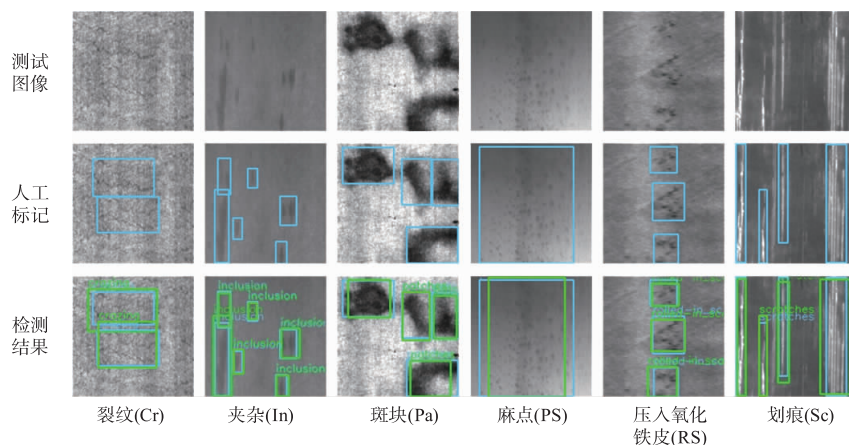


图6 各类别检测结果图

表4 不同算法的检测性能对比

算法	fps	mAP(%)	AP(%)					
			Cr	In	RS	PS	Pa	Sc
Faster R-CNN	10	70.4	46.2	69.7	65.4	75.2	84.6	81.5
DDN ^[10] + ResNet34	<20	74.8	48.0	75.9	68.4	78.3	87.4	90.8
DDN ^[10] + ResNet50	<10	82.3	62.4	84.7	76.3	89.7	90.3	90.1
YOLOv3	55	69.1	44.7	60.8	61.1	74.5	84.4	87.2
改进 YOLOv3	50	80.1	70.0	72.1	76.4	85.2	85.7	91.7

5 结论

针对热轧带钢表面缺陷检测问题,本文提出一种改进的 YOLOv3 算法,首先采用加权 K-means 算法代替传统 K-means 算法对先验框参数进行聚类,接着对 YOLOv3 算法网络结构进行优化调整,从而获得一种快速、准确的端到端带钢表面缺陷检测算法。

通过实验得到了以下结论:(1)对 YOLOv3 算法中的先验框参数聚类算法进行改进,采用加权 K-means 算法代替传统 K-means 算法,获得的 mAP 值提高了 4%;(2)对 YOLOv3 算法网络结构进行优化调整,增加了残差单元和一个大尺度的特征图层输出,以提高算法对小缺陷样本的检测能力,mAP 值相对于原始 YOLOv3 提高了 7%;(3)对改进 YOLOv3 算法在 NEU-DET 数据集上进行验证,6 类带钢表面缺陷的平均精度均值达到 80%,同时检测速度保持 50fps,实现了对带钢表面缺陷的端到端快速、准确检测。

与其他算法的性能对比表明,改进 YOLOv3 算法拥有较快的检测速度,优于目前其它深度学习带钢表面缺陷检测算法;同时算法也拥有不错的检测精度,不弱于其它检测算法。但是,文中算法通过预先设置 anchor 参数中先验框尺寸来降低训练难度同时提高检测精度,这一方法给模型带来了更多的参数同时影响了检

测速度。下一步,将尝试借鉴 Anchor-Free 的思想来改进目标检测算法,摆脱对先验框的依赖,减小模型大小同时进一步加快检测速度。

参考文献

- [1] 屈尔庆,崔月皎,徐森,等.改进的 Gabor 滤波器带钢表面缺陷显著性检测[J].华中科技大学学报(自然科学版),2017,45(10):12-17.
QU Er-qing, CUI Yue-jiao, XU Sen. Saliency defect detection in strip steel by improved Gabor filter[J]. Hua zhong Univ of Sci & Tech (Natural Science Edition), 2017, 45(10):12-17. (in Chinese)
- [2] 任海鹏,马展峰.基于复杂网络特性的带钢表面缺陷识别[J].自动化学报,2011,37(11):1407-1412.
REN Hai-Peng, MA Zhan-Feng. Strip steel surface defect recognition based on complex network characteristics[J]. Acta Automatica Sinica, 2011, 37(11):1407-1412. (in Chinese)
- [3] ZHU S G, DU J P, REN N. A novel simple visual tracking algorithm based on hashing and deep learning[J]. Chinese Journal of Electronics, 2017, 26(05):1073-1078.
- [4] XU K, XU Y, ZHOU P, et al. Application of RNAMlet to surface defect identification of steels[J]. Optics and Lasers in Engineering, 2018, 105(06):110-117.

- [5] SONG K C, YAN Y H. A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects [J]. *Applied Surface Science*, 2013, 285 (part_PB): 858 – 864.
- [6] HE D, XU K, ZHOU P. Defect detection of hot rolled steels with a new object detection framework called classification priority network [J]. *Computers & Industrial Engineering*, 2019, 128(03): 290 – 297.
- [7] FU G Z, SUN P Z, ZHU W B, et al. A deep-learning-based approach for fast and robust steel surface defects classification [J]. *Optics and Lasers in Engineering*, 2019, 121(10): 397 – 405.
- [8] HE Y, SONG K C, MENG Q G, et al. Semi-supervised defect classification of steel surface based on multi-training and generative adversarial network [J]. *Optics and Lasers in Engineering*, 2019, 122(11): 294 – 302.
- [9] HE D, XU K, ZHOU P, et al. Surface defect classification of steels with a new semi-supervised learning method [J]. *Optics and Lasers in Engineering*, 2019, 117(06): 40 – 48.
- [10] HE Y, SONG K C, MENG Q G, et al. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features [J]. *IEEE Transactions on Instrumentation and Measurement*, 2019, PP(99): 1 – 1.
- [11] 姜维, 张重生, 殷绪成. 基于深度学习的场景文字检测综述 [J]. *电子学报*, 2019, 47(5): 1152 – 1161.
JIANG Wei, ZHANG Chong-sheng, YIN Xu-cheng. Deep learning based scene text detection: A survey [J]. *Acta Electronica Sinica*, 2019, 47(5): 1152 – 1161. (in Chinese)
- [12] 张慧, 王坤峰, 王飞跃. 深度学习在目标视觉检测中的应用进展与展望. *自动化学报*, 2017, 43(08): 1289 – 1305.
ZHANG Hui, WANG Kun-Feng, WANG Fei-Yue. Advances and perspectives on applications of deep learning in visual object detection [J]. *Acta Automatica Sinica*, 2017, 43(08): 1289 – 1305. (in Chinese)
- [13] 罗会兰, 童康, 孔繁胜. 基于深度学习的视频中人体动作识别进展综述 [J]. *电子学报*, 2019, 47(5): 1162 – 1173.
LUO Hui-lan, TONG Kang, KONG Fan-sheng. The progress of human action recognition in videos based on deep learning: A review [J]. *Acta Electronica Sinica*, 2019, 47(5): 1162 – 1173. (in Chinese)
- [14] 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.
- [15] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN [A]. *Proceedings of the IEEE International Conference on Computer Vision [C]*. Venice: IEEE Press, 2017. 2961 – 2969.
- [16] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [A]. *Computer Vision and Pattern Recognition [C]*. USA: IEEE, 2016. 779 – 788.
- [17] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [A]. *Computer Vision and Pattern Recognition [C]*. USA: IEEE, 2017. 6517 – 6525.
- [18] REDMON J, FARHADI A. YOLOv3: An incremental improvement [A]. *Computer Vision and Pattern Recognition [C]*. arXiv Preprint, 2018, arXiv:1804.02767.
- [19] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [A]. *Proceedings of the IEEE European Conference on Computer Vision [C]*. Amsterdam, The Netherlands: IEEE, 2016. 21 – 37.
- [20] 石杰, 周亚丽, 张奇志. 基于改进 Mask RCNN 和 Kinect 的服务机器人物品识别系统 [J]. *仪器仪表学报*, 2019, 40(04): 216 – 228.
SHI Jie, ZHOU Ya-li, ZHANG Qi-zhi. Service robot item recognition system based on improved mask RCNN and kinect [J]. *Chinese Journal of Scientific Instrument*, 2019, 40(04): 216 – 228. (in Chinese)
- [21] 裴伟, 许晏铭, 朱永英, 等. 改进的 SSD 航拍目标检测方法 [J]. *软件学报*, 2019, 30(03): 738 – 758.
PEI Wei, XU Yan-Ming, ZHU Yong-Ying, et al. The target detection method of aerial photography images with improved SSD [J]. *Journal of Software*, 2019, 30(03): 738 – 758. (in Chinese)
- [22] 黄继鹏, 史颖欢, 高阳. 面向小目标的多尺度 Faster-RCNN 检测算法 [J]. *计算机研究与发展*, 2019, 56(02): 319 – 327.
HUANG Ji-peng, SHI Ying-huan, GAO Yang. Multi-scale faster-RCNN algorithm for small object detection [J]. *Journal of Computer Research and Development*, 2019, 56(02): 319 – 327. (in Chinese)
- [23] 王俊强, 李建胜, 周学文, 等. 改进的 SSD 算法及其对遥感影像小目标检测性能的分析 [J]. *光学学报*, 2019, 39(06): 373 – 382.
Wang Jun-qiang, Li Jian-sheng, Zhou Xuewen, et al. Improved SSD algorithm and its performance analysis of small target detection in remote sensing images [J]. *Acta Optica Sinica*, 2019, 39(06): 373 – 382. (in Chinese)
- [24] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [A]. *Computer Vision and Pattern Recognition [C]*. USA: IEEE, 2016. 770 – 778.

作者简介



李维刚 男,1977 年出生于湖北通城. 博士,武汉科技大学信息科学与工程学院教授、博导,主要研究方向:冶金过程控制、人工智能和深度学习新方法.

E-mail: liweigang_luck@foxmail.com



叶欣 男,1997 年出生于湖北宜昌. 武汉科技大学信息科学与工程学院硕士研究生,主要研究方向:工业过程建模与仿真、深度学习.

赵云涛 男,1982 年出生于内蒙古赤峰. 博士,武汉科技大学信息科学与工程学院副教授,主要研究方向:计算机仿真.

王文波 男,1978 年出生于湖北襄阳. 博士,武汉科技大学信息与计算科学系教授,主要研究方向为多尺度分析与电力系统信号处理、深度学习.