

文章编号: 2095-2163(2022)03-0022-06

中图分类号: TP242.6

文献标志码: A

# 基于改进的 YOLOv4-tiny 钢卷端面缺陷检测

吴奎<sup>1,2</sup>, 向峰<sup>1,2</sup>, 周顺<sup>1,2</sup>, 张雪荣<sup>3</sup>, 李红军<sup>4</sup>, 张驰<sup>4</sup>

(1 武汉科技大学 冶金装备及其控制教育部重点实验室, 武汉 430081; 2 武汉科技大学 机械传动与制造工程湖北省重点实验室, 武汉 430081; 3 武钢有限技术中心, 武汉 430083; 4 武汉纺织大学 机械自动化学院, 武汉 430200)

**摘要:** 针对带钢生产过程中钢卷端面出现的缺损和拉丝缺陷, 本文提出了一种改进的 YOLOv4-tiny 检测方法。首先, 在主干网络中加入了注意模块, 增强检测模型对缺陷特征的聚焦能力, 增加一个上采样层, 优化了深度特征和浅层特征的特征融合; 其次, 使用 Focal 损失函数替换置信度和分类的二分交叉熵损失函数, 解决分类过程中存在正负类样本分布不均衡问题; 最后, 利用加权 K-means 聚类算法重新聚类得到新的先验框。实验结果表明, 本文改进后的模型参数数量和检测速度与原模型相当, 但检测精确度上取得了更好的效果, 更适用带钢生产的实时检测任务。

**关键词:** Yolov4-tiny; 注意模块; K-means 聚类; 缺陷检测

## Defect detection of steel coil based on improved Yolov4-Tiny

WU Kui<sup>1,2</sup>, XIANG Feng<sup>1,2</sup>, ZHOU Shun<sup>1,2</sup>, ZHANG Xuerong<sup>3</sup>, LI Hongjun<sup>4</sup>, ZHANG Chi<sup>4</sup>

(1 Key Laboratory of Metallurgical Equipment and Control of Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, China; 2 Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science and Technology, Wuhan 430081, China; 3 Wuhan Iron and Steel Co., Ltd. Technology Center, Wuhan 430080, China; 4 School of Mechanical Engineering and Automation, Wuhan Textile University, Wuhan 430200, China)

**[Abstract]** For strip steel edge of production process of edge defect and drawing defects, this paper puts forward an improved YOLOv4-tiny detection method. Firstly, an attention module is added to the backbone network to enhance the ability of the detection model to focus on defect features. An upper sampling layer is added to optimize the feature fusion of depth features and shallow features. Then, the Focal loss function was used to replace the confidence and dichotomous cross entropy loss function of classification to solve the problem of unbalanced distribution of positive and negative samples in the classification process. Finally, the weighted K-means clustering algorithm is used to get a new prior box. The experimental results show that the number of parameters and detection speed of the improved model are similar to the original model, but the detection accuracy is better, and it is more suitable for the real-time detection task of strip production.

**[Key words]** Yolov4-tiny; attention module; K-means clustering; defect detection

## 0 引言

热轧带钢在卷取过程中由于受轧制工艺及系统控制等因素的影响, 钢卷两侧端面存在缺损和拉丝缺陷。端面缺陷不仅会影响钢卷的美观, 还会降低钢卷的成材率和性能, 影响下游企业加工使用。传统的带钢缺陷检测方法一般可分为两种, 一种是基于边缘检测的方法; 另一种是基于形态学的方法。例如: 基于 Canny 的检测器, 基于 Sobel 算子和 LoG 算子的边缘检测方法。然而, 这些方法都需要提取特征因子, 再对缺陷进行分类, 检测过程不仅复杂, 而且速度较慢, 在检测精度和实时性等方面不能很好的满足实际生产需求<sup>[1]</sup>。

近年来, 随着机器视觉、深度学习技术的不断发

展, 神经网络已逐渐成为工业生产中检测产品缺陷的主要方法。神经网络能够自动从图像中提取特征, 避免了手工提取特征可能存在的主观性错误, 在目标检测与分类问题中得到了广泛的运用<sup>[2]</sup>。目前, 带钢缺陷检测利用卷积神经网络提取缺陷特征并进行分类, 例如: HE 等人<sup>[3]</sup>使用对抗神经网络获得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导致网络难以训练的问题; FU 等人<sup>[4]</sup>提出的一种端到端的卷积神经网络, 实现了带钢缺陷的高精度分类。但这些检测模型只是解决了带钢缺陷分类问题, 并不能解决缺陷定位问题<sup>[5]</sup>, HE<sup>[6]</sup>使用 Defect Detection Network (DDN) 实现了端到端的带钢缺陷检测, 并取得了 80% 的准确率, 这一方法虽然取得了较好的检测精度, 但检测速

**作者简介:** 吴奎 (1995-), 男, 硕士研究生, 主要研究方向: 机器学习; 向峰 (1983-), 男, 博士, 副教授, 主要研究方向: 数字孪生、机器视觉; 周顺 (1996-), 男, 硕士研究生, 主要研究方向: 机器视觉、数字孪生。

收稿日期: 2021-11-15

哈尔滨工业大学主办 ◆ 学术研究与应用

度较慢,无法达到工业生产实时检测的要求。

缺陷检测的目的是找出图像中所有感兴趣的目标,确定目标的位置、大小以及类别等信息<sup>[7]</sup>。目前,基于深度学习的缺陷检测方法逐步取代了传统的检测方法。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、Mask R-CNN 为代表的两阶式检测算法;另一类是以 YOLO、SSD 为代表的 one-stage 检测算法。两阶式检测算法的检测过程分为两个阶段,第一阶段使用区域候选网络(RPN)产生候选区域,然后使用检测网络确定候选区域的类别、位置,这种检测算法有较高的准确度,但检测速度稍慢。单阶式检测算法可以直接得到检测结果,不需要产生候选区域阶段,所以检测速度较快,但检测准确度较低。本文采用 YOLOv4 算法的轻量化版本 YOLOv4-tiny,减少了模型参数和计算量,检测精度在满足要求的同时拥有更高的检测速度。同时, YOLOv4-tiny 还具有占用内存小、易部署的特点,非常适合实际生产过程中的钢卷端面缺陷检测<sup>[8]</sup>。

为了提高 YOLOv4-tiny 的检测精度,本文对其进行了改进,提出了一种改进后的 YOLOv4-tiny 算法。实验结果表明,改进后的 YOLOv4-tiny 算法在钢卷数据集上平均精度均值为 78.8%,检测速度达到了 94 fps,满足了带钢生产实时检测的要求。

## 1 YOLOv4-tiny 算法

YOLOv4-tiny 是在 YOLOv4 的基础上改进而来的,是 YOLOv4 的轻量化版本,适用于嵌入式平台的轻量级实时检测。与 YOLOv4 相比,检测精度有所下降,但在实现了模型压缩,提高了模型检测速度。YOLOv4-tiny 由骨干网络(CSP Darknet53-Tiny)、特征金字塔(Feature Pyramid Network, FPN)和 YOLO 检测头(YOLO Head)组成。其骨干网络主要包括 CSP(Cross Stage Partial)结构和下采样 CBL(Conv+Bn+Leaky-relu 组成)结构,CSP 结构将低层的特征映射划分为两部分,通过跨层连接将其合并,使卷积神经网络有了更强的学习能力,减少了计算量的同时也保证了检测的准确率;下采样 CBL 结构中,每个卷积核大小为 3×3,步长为 2,主要对图像进行下采样处理。FPN 结构可以融合不同网络层之间的特征,既可以保证深层网络丰富的语义信息,又可以获得低层网络的特征细节信息,以此来加强对特征的提取能力。YOLO Head 利用得到的特征信息进行最后的预测,最终形成 13×13 和 26×26 两个预测尺度。

## 2 YOLOv4-tiny 算法的改进

### 2.1 网络结构的改进

在原始主干网络中,残差网络模块使用 3×3 卷积核提取特征,其感受野大小也是 3×3。较小的感受野虽然可以提取更多的局部信息,但丢失了全局信息,影响了目标检测的准确性。为了提取更多的全局特征,使用两个连续的相同的 3×3 卷积来获得辅助残差网络块中的 5×5 感受野。辅助网络将提取的全局信息传输到主干网络,主干网络结合较大感受野获得的全局信息和较小感受野获得的局部信息,以获得更多的目标信息。此外,随着网络深度的增加,语义信息也变得更加高级。注意机制可以集中处理和传递有效特征,通道抑制无效特征。因此,本文在辅助网络模块中引入了通道注意力模块和空间注意力模块,以获取更有效的特征信息。卷积块注意模块(Convolutional Block Attention Module, CBAM)将通道注意力模块和空间注意力模块串联结合组成的注意力模块,直接使用 CBAM 来同时实现通道注意和空间注意<sup>[9]</sup>。

通道注意力模块分别通过最大池化层(Max-Pooling)和平均池化层(Average-Pooling)对输入特征图的空间维度进行压缩,输出两个特征描述符,将两个特征描述符发送到一个共享网络,使用元素求和来合并并输出特征向量。最后,通过 sigmoid 函数激活特征向量,得到通道注意力图。通道注意力模块结构如图 1 所示。

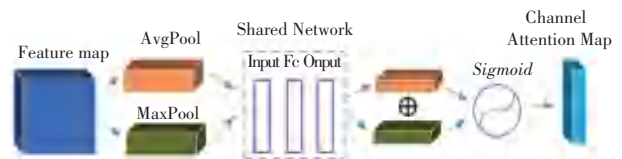


图 1 通道注意力模块结构图

Fig. 1 Channel attention module structure

空间注意力模块对输入特征图进行平均池化和最大池化操作,将其连接起来生成一个有效的特征描述符,应用两个卷积层来强调描述符区域,再由 sigmoid 函数进行激活,得到空间注意力图。空间注意力模块结构如图 2 所示。

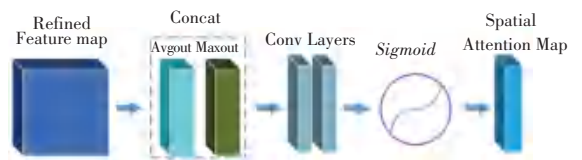


图 2 空间注意力模块结构图

Fig. 2 Spatial attention module structure

卷积块注意模块(CBAM)是将通道注意力模块和空间注意力模块串联相结合组成的注意力模块,CBAM的表达式如式(1)和式(2)所示,其组成的辅助网络结构如图3所示。

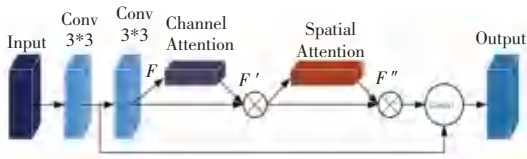


图3 辅助网络结构图

Fig. 3 Auxiliary network structure

$$F' = M_c(F) \otimes F \quad (1)$$

$$F'' = M_s(F') \otimes F' \quad (2)$$

其中,  $F \in R^{C \times H \times W}$  表示输入特征图;“ $\otimes$ ”表示元素式乘法;  $F''$  为最终输出特征图;  $M_c(\cdot)$  和  $M_s(\cdot)$  分别为通道注意力和空间注意力图。

$M_c(F)$  的计算公式(3)为:

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \quad (3)$$

其中,  $AvgPool(\cdot)$  和  $MaxPool(\cdot)$  分别表示平均池化操作和最大池化操作;  $MLP(\cdot)$  表示多层感知

器网络;  $\sigma(\cdot)$  是 *sigmoid* 函数。

$M_s(F')$  的计算公式(4)为:

$$M_s(F') = \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)])) \quad (4)$$

其中,  $f^{7 \times 7}$  表示一个滤波器大小为  $7 \times 7$  的卷积运算,“ $[\cdot]$ ”表示连接操作。

在YOLOv4-tiny中,利用FPN结构对输出的两个有效特征层进行简单的特征融合。其过程如下:对最后一个有效特征层进行卷积,然后上采样,一方面通过yolo\_head处理输出特征尺度为  $13 \times 13$  的预测结果,另一方面与前一个输出的有效特征层进行叠加,通过yolo\_head处理输出特征尺度为  $26 \times 26$  的预测结果。YOLOv4-tiny网络输出的  $13 \times 13$  和  $26 \times 26$  两个预测尺度,对被检测图像中的大目标的检测效果较好,对小目标而言检测效果较差。为了提高网络对小目标的检测能力,在原网络结构上增加了上采样2,在通道维度上连接了CPSBlock层和上采样2,并增加了检测尺度。特征金字塔网络由原来的  $13 \times 13$  和  $26 \times 26$  两个预测尺度增加为  $13 \times 13$ 、 $26 \times 26$  和  $52 \times 52$  3个预测尺度,有助于提高检测网络对小目标的检测精度。改进的YOLOv4-tiny网络结构如图4所示。

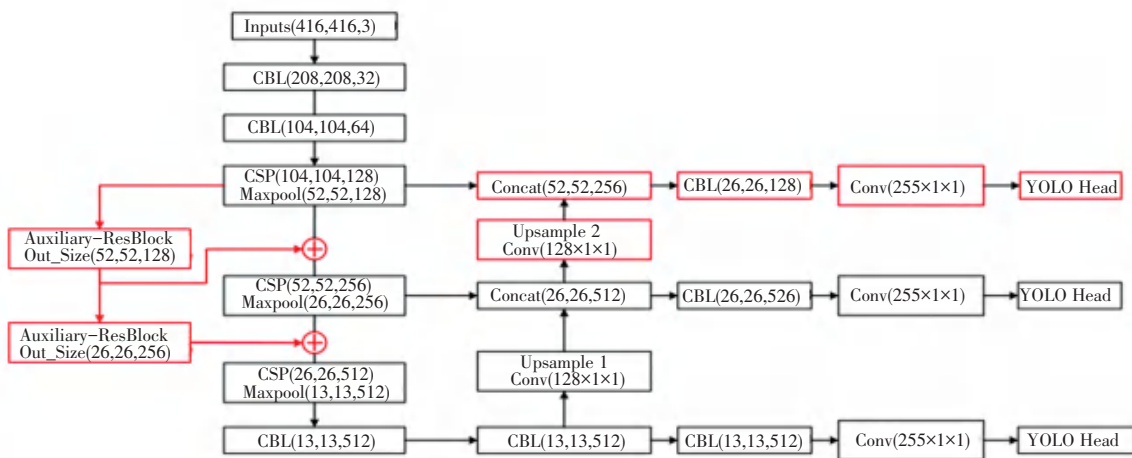


图4 改进的YOLOv4-tiny网络结构图

Fig. 4 Improved Yolov4-tiny network structure diagram

## 2.2 损失函数的改进

YOLOv4-tiny的损失函数由边界框回归损失、置信度损失和分类损失3部分组成。边界框回归损失函数采用  $CIoU$  函数,置信度损失和分类损失采用与YOLOv4相同的交叉熵函数。Rezatofighi等<sup>[10]</sup>提出  $GIoU$  损失,在  $IoU$  损失基础上增加了包含检测框和真实框的最小矩形框的  $C$  检测框,解决了检测框和真实框不重叠的问题,但是当真实框和检测框之间出现包含关系的时候,  $GIoU$  和  $IoU$  效果相同;Zheng等<sup>[11]</sup>针对这一问题,提出  $DIoU$  损失,将预测

框和真实框都包含进  $C$  检测框,但是  $DIoU$  计算的不是框之间的交并,而是计算每个检测框之间的欧氏距离。 $CIoU$  损失考虑了中心点距离、长宽比和重叠面积,在  $DIoU$  损失的基础上加入长宽比的惩罚项,这样预测框就会更加的接近真实框。 $CIoU$  损失的计算公式(5)~公式(8):

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (5)$$

$$IoU = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|} \quad (6)$$

$$\alpha = \frac{v}{(1 - IoU) + v} \quad (7)$$

$$v = \frac{4}{\pi^2} \arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \quad (8)$$

其中,  $B$  为预测框;  $B^{gt}$  为真实框;  $b$  和  $b^{gt}$  分别表示预测框和真实框的中心点;  $w$  和  $h$  为预测框的宽和高;  $w^{gt}$  和  $h^{gt}$  为真实框的宽和高;  $c$  表示预测框和真实框的最小外界矩形的对角线距离;  $\rho(\cdot)$  表示欧式距离;  $\alpha$  是用于协调比例参数;  $v$  用来衡量长宽比的一致性。

为了增强模型对缺损检测的能力,使用 Focal 损失函数替换置信度和分类的二分交叉熵损失函数。Focal 损失是针对分类过程中存在正负类样本分布不均衡问题提出的,通过减少易分类样本的权重,使模型在训练时更注意难以分类的样本。Focal 损失计算公式(9):

$$L_{fl} = \begin{cases} -\alpha(1-y')^{\gamma} \log y', & y = 1 \\ -(1-\alpha)y'^{\gamma} \log(1-y'), & y = 0 \end{cases} \quad (9)$$

其中,  $y$  为真实样本的标签;  $y'$  为经过 sigmoid 的预测输出,在  $[0,1]$  之间取值;  $\gamma$  的作用是减少易分类样本的损失,使损失函数更加关注难以分类和误分类的样本;  $\alpha$  为平衡因子,其作用是平衡正负样本的不均匀比例,防止易分类类别的损失函数过小。 $\gamma$  和  $\alpha$  共同调节达到相对平衡。

### 2.3 先验框的改进

为了适应钢卷端面缺陷的尺寸,还需要对数据集进行聚类,得到合适的先验框。合适的先验框不仅能降低模型最后的损失值,还可以加快模型的收敛速度。网络原有的先验框是在 VOC 数据集上聚类得到的,而 VOC 数据集包含的种类较多,情况复杂。对钢卷端面缺陷检测而言,只包含端面缺损与端面拉丝两个主要缺陷目标,和 VOC 数据集相比差异较大,直接使用网络原有的先验框不能很好地满足识别要求,所以利用加权 K-means 聚类算法对钢卷端面缺陷数据集的宽和高进行聚类,得到新的先验框。加权 K-means 聚类算法使用聚类中心与样本之间的最大交并比来评价聚类结果,其目标函数  $D$  如式(10)所示:

$$D = \min \sum_{i=1}^n \sum_{j=1}^k \frac{c_{ij}}{c_i} - \omega_{ij} \frac{box_i \cap cen_j}{box_i \cup cen_j} \quad (10)$$

其中,  $box_i$  为样本的实际长、宽;  $cen_j$  为第  $j$  个聚类中心;  $n$  为样本数;  $k$  为选取的聚类中心个数;  $\omega_{ij}$  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。

加权 K-means 算法具体实现步骤如下:

- (1) 分析数据集中的样本,随机划分  $k$  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(11)所示:

$$d_{il} = \sqrt{[box_i(1) - C_l(1)]^2 + [box_i(2) - C_l(2)]^2} \quad (11)$$

其中,  $C_l(1)$ 、 $C_l(2)$  为第  $l$  个中心点的横、纵坐标,  $box_i(1)$ 、 $box_i(2)$  为第  $i$  个样本的横、纵坐标;

- (3) 求权重矩阵,即每个样本所对应的聚类中心的权重,属于同一个中心点的聚类中心拥有相同的权重。权重是样本到聚类中心距离标准化后的相反数,距离较远的样本权重为 0,降低不平衡样本的干扰。式(12)为所有样本到某一中心点距离的均值,式(13)为权重的计算公式。

$$\bar{d}_j = \frac{1}{n} \sum_{i=0}^n d_{il} \quad (12)$$

$$\omega_{ijl} = \max \left\{ 0, -\frac{d_{il} - \bar{d}_l}{\sqrt{\frac{1}{n} \sum_{i=1}^n (d_{il} - \bar{d}_l)^2}} \right\} \quad (13)$$

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

- (4) 迭代更新所有的聚类中心,直到聚类中心不再发生改变。

对于钢卷端面数据集,运用上述方法重新进行聚类分析后,得到的先验框见表 1。

表 1 钢卷端面数据集的先验框值

Tab. 1 The prior box values of the data set of steel coil end faces

特征层	特征图尺寸	先验框值
特征层 1	13×13	(34,28)(49,45)(62,53)
特征层 2	26×26	(26,22)(24,18)(20,16)
特征层 3	52×52	(13,10)(15,12)(19,11)

## 3 实验结果与分析

本文所有的实验均在 Windows10 系统下进行,硬件环境:CPU 为 Intel i5-11400F 8G;GPU 为 NVIDIA GTX1080Ti;开发环境:Tensorflow-gpu2.2.0;CUDA 版本号 为 CUDA10.1;OpenCV3.4;Python3.6。

### 3.1 实验数据集

本文数据集是通过某热轧厂生产现场采集得到,共采集缺陷图片 1 000 张。为了丰富数据集,获得更优的训练效果,使用随机转换的方式增强数据集,包括图像顺时针旋转、逆时针旋转、镜像翻转、尺

度变换等对数据集进行扩充,由原来的1 000张图像扩充到5 000张。根据VOC2007数据集的格式制作标准的钢卷端面数据集,使用LabelImg缺陷标注软件对图片中的缺陷进行标注。得到缺陷区域的边界框以及缺陷的类别标签,并生成对应的XML文件,完成VOC2007数据集格式的转换,之后将数据集按8:2的比例随机生成了训练验证集和测试集,再进行模型训练及测试。

### 3.2 算法评价指标

为了更加合理有效地对最后的检测结果进行判别,根据检测速度和检测精度要求,本文采用精准率  $P$  (Precision) 和召回率  $R$  (Recall) 作为缺陷识别的评价指标,精准率是所有被正确识别的缺陷占有所有被识别到的缺陷的比例;而召回率指的是所有被正确识别的缺陷占有所有应该被识别到的缺陷的比例,式(14)和式(15)。为了同时考虑精准率和召回率,权衡两者之间的平衡关系引入了  $F_1$  值,计算公式(16)。

$$P = \frac{TP}{TP + FP} \quad (14)$$

$$R = \frac{TP}{TP + FN} \quad (15)$$

$$F_1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (16)$$

其中,  $TP$  为成功预测的缺陷目标;  $FP$  为被模型误判为缺陷的非缺陷目标;  $FN$  表示被模型错误预测为非缺陷的缺陷目标。

由于精确率和召回率受到置信度的影响,单独采用精确率与召回率作为评价指标会有局限性,所以在实验中加入平均精度均值 (mean Average Precision,  $mAP$ ) 和检测速度作为评价指标。平均精度 (Average Precision,  $AP$ ) 为不同召回率下精确率的平均值,用来评价某一缺陷的检测精度,  $mAP$  是所有缺陷类别检测精度的均值,一般用来评价网络模型的整体检测性能,计算公式如式(17)所示。目标检测算法的另一个重要性能指标是检测速度,评估检测速度的常用指标是每秒帧率 (frame per second,  $FPS$ ), 即网络模型每秒能够检测的图片数量,计算公式如式(18)所示。

$$mAP = \frac{\sum_{i=0}^n AP(i)}{n} \quad (17)$$

$$FPS = \frac{NumFigure}{TotalTime} \quad (18)$$

其中,  $AP(i)$  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。

### 3.3 实验结果分析

本文实验中训练网络模型的参数设置为:每次迭代训练样本数为64,分16个批次,动量因子设置为0.949,权重衰减系数为0.0005,最大迭代次数为5 000,学习率为0.001;训练迭代次数达到3 000和4 000时,将学习率分别降至初始学习率的0.1和0.01。

本文以相同的实验数据分别使用YOLOv3、YOLOv3-tiny、YOLOv4、YOLOv4-tiny与本文改进后的YOLOv4-tiny模型进行训练,在相同测试集下进行测试对比,测试其在  $mAP$  和  $FPS$  中的性能。对比结果见表2。

表2 模型实验结果对比

检测模型	$mAP$ /%	FPS	模型大小/MB
YOLOv3	82.6	31	235
YOLOv3-tiny	65.4	86	36
YOLOv4	86.7	42	258
YOLOv4-tiny	74.2	96	24
改进的YOLOv4-tiny	78.8	94	23

从表2中可以看出,YOLOv3和YOLOv4模型的  $mAP$  数值高于其他模型,但这两个模型的检测速度却远远小于其他模型。因为YOLOv3和YOLOv4检测网络结构复杂,参数较多,使得其相较轻量化的模型可以取得更好的检测精度,但检测速度相应的会低于轻量化的模型。YOLOv3-tiny、YOLOv4-tiny和本文改进的YOLOv4-tiny方法属于轻量级深度学习方法,其网络结构相对简单,参数较少,因此在检测速度上表现的更好。本文改进后的YOLOv4-tiny方法相较YOLOv3-tiny和YOLOv4-tiny,在  $mAP$  数值上分别提升了13.4%和4.6%。由于改进后的YOLOv4-tiny在检测网络中加入注意模块和一个上采样,增加了网络的参数,检测速度较YOLOv4-tiny稍慢一些,但可以满足实际生产的检测需求。表明本文改进后的YOLOv4-tiny检测精度提升的同时又保证了检测的速度。

由于YOLOv3-tiny、YOLOv4-tiny和本文改进后的YOLOv4-tiny方法都属于轻量级深度学习方法,所以采用本文改进后的YOLOv4-tiny方法、YOLOv3-tiny和YOLOv4-tiny方法,分别对缺陷检测后的精准率  $P$ 、召回率  $R$ 、 $F_1$  值进行比较,模型的实验结果对比如图5所示。改进后的YOLOv4-tiny实验检测效果如图6所示。

由图5可以看出,本文改进后的YOLOv4-tiny对缺陷的检测精准率  $R$  和召回率  $P$  比YOLOv3-tiny提升12.3%和3.3%,比YOLOv4-tiny提升8.2%和

1.5%, 体现模型综合性能的  $F1$  值分别高出 9% 和 5.6%; 对拉丝的检测精准率  $R$  和召回率  $P$  比 YOLOv3-tiny 提升 5.5% 和 5.8%, 比 YOLOv4-tiny 提升 4.3% 和 2.1%,  $F1$  值分别高出 6.5% 和 3.5%; 对缺

损+拉丝的检测精准率  $R$  和召回率  $P$  比 YOLOv3-tiny 提升 5.5% 和 5.8%, 比 YOLOv4-tiny 提升 4.3% 和 2.1%,  $F1$  值分别高出 6.5% 和 3.5%。本文改进后的 YOLOv4-tiny 方法综合性能得到一定的提升。

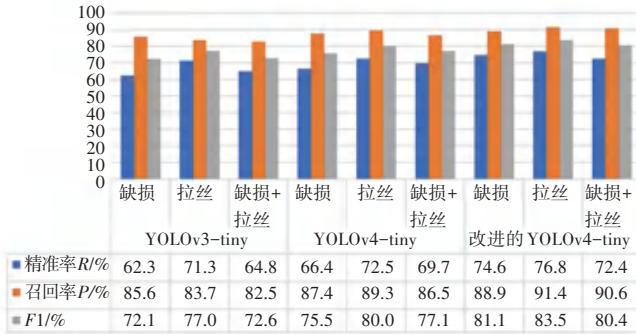


图 5 模型实验结果对比

Fig. 5 Comparison of model experiment results

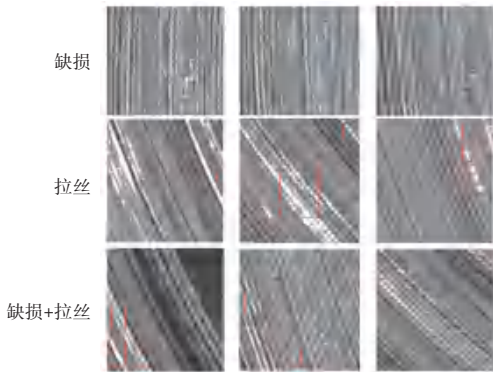


图 6 改进后的 YOLOv4-tiny 检测效果图

Fig. 6 Improved detection effect of Yolov4-Tiny

### 4 结束语

针对带钢生产过程中钢卷端面出现的端面缺损和拉丝缺陷, 本文提出了一种改进的 YOLOv4-tiny 检测方法。首先, 在骨干网络中加入了注意模块, 增强了检测模型对缺陷特征的表达和对缺陷区域的聚焦能力; 同时又增加了一个上采样层, 优化了深度特征和浅层特征的特征融合, 改善了感兴趣区域的特征表达; 改进模型的损失函数, 使用 *Focal* 损失函数替换置信度和分类的二分交叉熵损失函数, 解决分类过程中存在正负类样本分布不均衡问题; 最后, 利用加权 K-means 聚类算法对钢卷端面缺陷数据集的宽和高进行聚类, 得到新的先验框, 以适应钢卷端面缺陷的尺寸。改进的 YOLOv4-tiny 网络模型通过实验结果验证与分析可知, 相比较改进前的 YOLOv4-tiny 网络模型, 检测速度接近,  $mAP$  提升了 4.6%, 保证检测速度的同时检测能力得到了一定的提高, 满足了带钢生产的端面缺陷检测要求。

### 参考文献

- [1] 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 (6): 110-117.
- [2] 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.
- [3] 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.
- [4] 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.
- [5] 李维刚, 叶欣, 赵云涛, 等. 基于改进 YOLOv3 算法的带钢表面缺陷检测 [J]. 电子学报, 2020, 48(7): 1284-1292.
- [6] 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 (99): 1.
- [7] 陈辉东, 丁小燕, 刘艳霞. 基于深度学习的目标检测算法综述 [J]. 北京联合大学学报, 2021, 35(3): 39-46.
- [8] 陈科圻, 朱志亮, 邓小明, 等. 多尺度目标检测的深度学习研究综述 [J]. 软件学报, 2021, 32(4): 1201-1227.
- [9] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]// Proceedings of the 2018 European Conference on Computer Vision, LNCS 11211. Cham: Springer, 2018: 3-19.
- [10] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression [C]// Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA. 2019. 658-666.
- [11] ZHENG ZH, WANG P, LIU W, et al. Distance-IoU Loss: Faster and better learning for bounding box regression [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 34(7): 12993-13000.