

# 基于改进 YOLOv5 的轮对踏面缺陷检测

孙耀泽<sup>1\*</sup>, 高军伟<sup>2</sup>

<sup>1</sup>青岛大学自动化学院, 山东 青岛 266071;

<sup>2</sup>山东省工业控制技术重点实验室, 山东 青岛 266071

**摘要** 针对高速列车轮对踏面缺陷检测精度低、效率慢的问题,提出一种改进的 YOLOv5 算法实现快速准确检测。引入卷积注意力机制,在通道维度和空间维度上进行特征优化,使重要的目标特征在网络处理中占比更大,从而增强网络对目标区域的特征学习能力;根据踏面缺陷类别的大小对 Neck 区结构进行精简,保留适合检测中小目标的特征图分支,从而降低模型复杂度;将边框回归的损失函数改为有效交并比(EIoU),融合更多边界框信息,提高预测的准确率。实验结果表明:相较于原始的 YOLOv5,改进后的 YOLOv5 在测试集上的平均精度均值(mAP)提高了 5.5 个百分点,检测速度提升了 2.8 frame/s,面对复杂场景具有较强的泛化能力。

**关键词** 图像处理; 缺陷检测; YOLOv5 算法; 卷积注意力机制; 损失函数

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP202259.2215003

## Defect Detection of Wheel Set Tread Based on Improved YOLOv5

Sun Yaoze<sup>1\*</sup>, Gao Junwei<sup>2</sup>

<sup>1</sup>School of Automation, Qingdao University, Qingdao 266071, Shandong, China;

<sup>2</sup>Shandong Key Laboratory of Industrial Control Technology, Qingdao 266071, Shandong, China

**Abstract** Aiming at the problems of low accuracy and low efficiency of wheel set tread defect detection of high-speed trains, an improved YOLOv5 algorithm is proposed to realize fast and accurate detection. A convolution attention mechanism is introduced to optimize the features in the channel and spatial dimensions so that the essential target features occupy a greater proportion in the network processing to enhance feature learning ability in the target region. According to the size of the tread defect category, the structure of the Neck area is simplified, and the characteristic graph branches suitable for detecting small- and medium-sized targets are retained to reduce the model's complexity. The loss function of the bounding box regression is changed to efficient intersection over union (EIoU) to integrate more bounding box information and improve prediction accuracy. Compared with the original YOLOv5 algorithm, experimental results demonstrate that the mean average precision (mAP) of the enhanced YOLOv5 algorithm on the test set is increased by 5.5 percentage points, and the detection speed is improved by 2.8 frames/s, which has a strong generalization ability in complex scenes.

**Key words** image processing; defect detection; YOLOv5 algorithm; convolution attention mechanism; loss function

## 1 引言

随着铁路运营速度增加,车轮磨耗日益严重,车轮在服役过程中容易出现踏面损伤,如踏面擦伤、裂纹、剥离等,影响铁路车辆行车安全<sup>[1-2]</sup>。因此,快速准确地检测出列车走行过程中容易出现的踏面缺陷损伤对于列车安全运行是非常重要的。

基于机器视觉和超声探伤的检测方法可以有效代替人工。文献[3]对 Canny 算子进行改进,通过定位轮

对踏面损伤的边缘信息完成缺陷检测任务。文献[4]采用随机抽样一致性的轮缘线拟合方法,实现机车轮对踏面的提取,再通过灰度共生矩阵对踏面部分的纹理信息进行采集,完成对踏面损伤的识别。文献[5]提出一种基于 electromagnetic acoustic transducer (EMAT) 的轮对在线自动探伤方法,该方法利用电磁超声表面波回波实现踏面缺陷检测,可完成对踏面剥离缺陷的检测。但上述方法存在识别准确率较低、面对复杂环境泛化性能较差等问题。

收稿日期: 2021-12-15; 修回日期: 2021-12-25; 录用日期: 2022-01-11

基金项目: 山东省自然科学基金(ZR2019MF063)、山东省重点研发计划(2017GGX10115)

通信作者: \*486695900@qq.com

随着深度学习的发展,出现了经典的 Faster RCNN 算法<sup>[6]</sup>、YOLO 算法<sup>[7-8]</sup>及 SSD 算法<sup>[9-10]</sup>,基于深度卷积神经网络(DCNN)的目标检测技术被有效运用于众多领域<sup>[11]</sup>。文献[12]利用对抗网络扩大踏面损伤数据集,并利用 YOLOv2 完成对踏面损伤的检测。文献[13]提出一种基于 Canny-YOLOv3 的踏面损伤图像检测算法,该算法提取经 Canny 边缘的轮对踏面图像后对剔除背景的踏面图像进行损伤检测。文献[14]将两种目标检测算法结合,先通过 SSD 算法对踏面区域进行提取,再通过改进的 YOLOv3 对缺陷进行检测,增加了检测速度。

本文将高速列车运行过程中常见的三类缺陷(踏面擦伤、踏面硌伤、踏面剥离)作为检测目标,在 YOLOv5 算法的基础上进行改进,在 Backbone 区域引入卷积注意力机制,加大重要目标特征在网络的占比,从通道维度和空间维度上加强对目标特征的学习能力;根据缺陷目标的大小特征对 Neck 区网络结构进行适当调整,去掉 19×19 的特征检测层,降低模型的复杂度;将有效交并比(EIoU)作为边框的损失函数,融合边界框宽高比的尺度信息,在保证精度的同时提升了检测速度。

## 2 YOLOv5 算法

YOLOv5 作为一种单步检测算法,相较于

YOLOv4<sup>[15-16]</sup>,虽然检测性能稍弱,但因其具有更强的灵活性与速度,更适合模型的快速部署。YOLOv5 网络结构由输入端、Backbone、Neck 区、Prediction 组成。

输入端由 Mosaic 数据增强、自适应图片缩放、自适应锚框计算 3 部分组成。Mosaic 数据增强通过随机缩放、随机裁剪、随机排布的方式对图像进行拼接,丰富了检测数据集;自适应图片缩放是针对图片长宽比不同进行的一种改进,通过减少图像上的边界填充以加快推理速度;自适应锚框计算通过不断迭代参数,计算最佳的锚框值。Backbone 主要包含了 Focus 模块和 cross stage partial (CSP)。Focus 结构对图片进行切片操作,对图像相邻的四个位置进行堆叠,提高通道中每个点的感受野,保留更多原始信息,减少计算量并增加速度;CSP 将梯度信息集成到特征图中,减少模型的参数量,在保证推理速度和准确率提升的基础上还能使模型具有更小的尺寸。Neck 区域为特征融合部分,基于特征金字塔网络(FPN)<sup>[17]</sup>增添自底向上的加强,这样顶层特征图会拥有更多的位置特征,使检测性能更优。Prediction 部分包括边界框预测、损失函数计算及非极大值抑制,其输出端由三层检测层组成,不同尺寸的特征图用于检测不同尺寸的目标对象。整 YOLOv5 模型结构如图 1 所示。

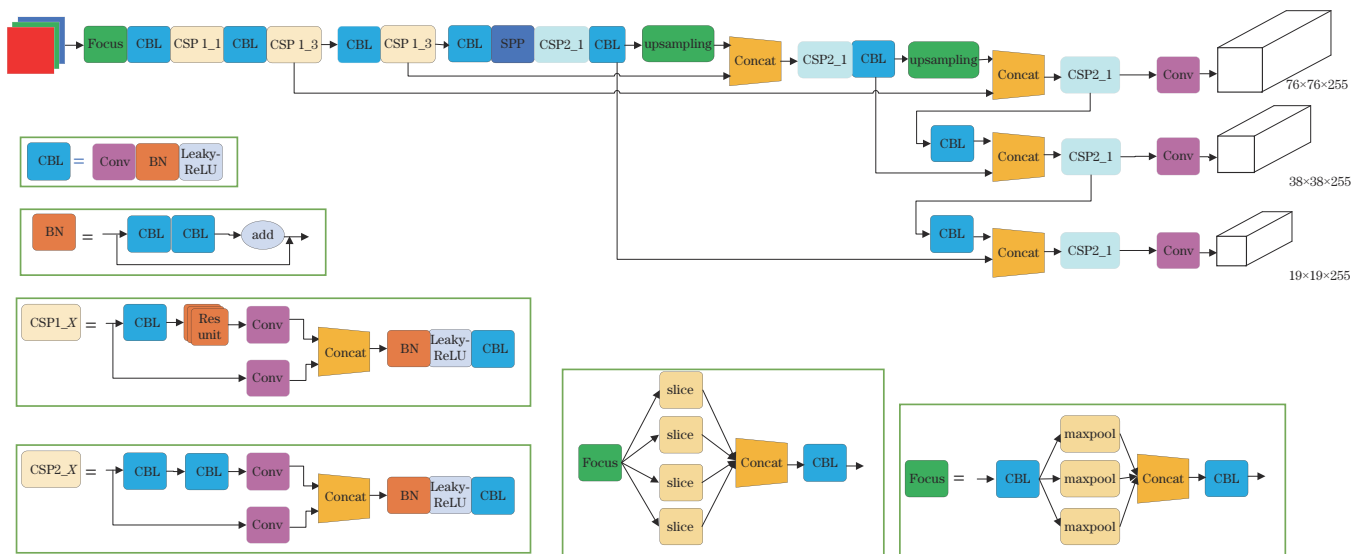


图 1 YOLOv5 模型结构

Fig. 1 YOLOv5 model structure

## 3 改进的 YOLOv5

### 3.1 卷积块注意力机制

注意力机制主要分为通道注意力机制、空间注意力机制及混合注意力机制三种。在模型中引入注意力机制,可以加强模型对重要特征的关注度,从而起到提高模型检测精度的效果。卷积块注意力机制(CBAM)<sup>[18]</sup>

包含通道注意力机制和空间注意力机制,其中通道注意力模块负责关注通道信息,判断哪个通道具有目标的主要特征;空间注意力模块则关注空间位置并确定哪个位置包含目标的主要信息。卷积注意力机制的结构如图 2 所示。

通道注意模块主要关注 feature map 中有效通道的权重。为了更好地计算每个特征通道的重要程度,首先

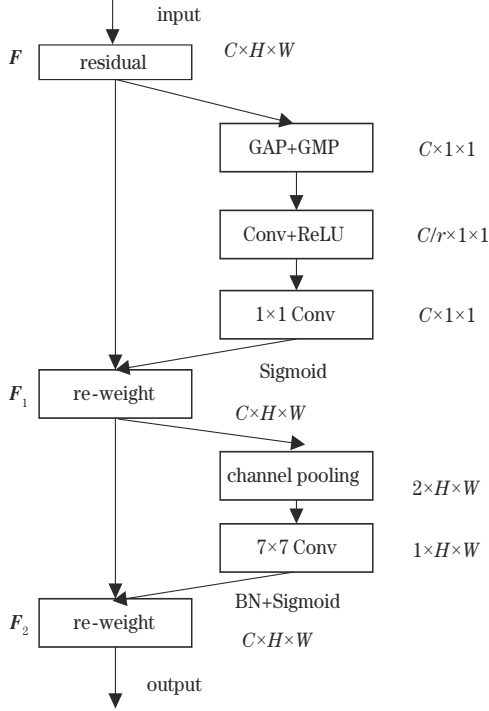


图 2 卷积注意力机制的结构

Fig. 2 Structure of convolution attention mechanism

对输入尺寸为  $C \times H \times W$  的特征图  $F$  进行全局空间最大池化和平均池化,压缩得到两个  $C \times 1 \times 1$  的通道描述;随后将其输入到一个包含多层感知机的共享神经网络,得到两个输出特征图像;再将特征图像堆叠后输入激活函数,得到通道权值,最后将权值与特征图  $F$  的像素值逐一相乘,得到新的特征图  $F_1$ 。空间注意模块将重点放在区域位置上,强调信息更丰富的区域。为了得到更多有意义区域特征的权值系数,首先对输入为  $C \times H \times W$  的特征图  $F_1$  在通道维度进行一个全局空间最大池化和平均池化,得到两个  $1 \times H \times W$  的空间描述子;随后将它们合并以获得更高效的特征描述子,在经过一个  $7 \times 7$  的卷积层后,将 feature map 输入 Sigmoid 激活函数后得到空间权值;最后将空间权值与特征图  $F_1$  的像素值逐一相乘,得到最终的注意力特征  $F_2$ 。

### 3.2 改进的结构

踏面缺陷的尺寸相对整张图像过小,使得 Neck 区域对较大目标检测的部分变得冗余,占用计算资源。为提高模型检测速度,对 YOLOv5 模型的 Neck 区域进行适当精简,将其中拥有最大感受野、适合检测较大尺寸对象的  $19 \times 19$  特征图分支删除,如图 3 所示,从而降低模型复杂度并提高检测实时性。

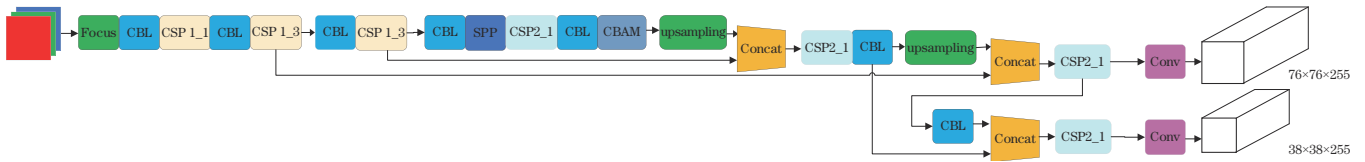


图 3 改进后的 YOLOv5 结构图

Fig. 3 Structure diagram of improved YOLOv5

### 3.3 边界框损失函数

YOLOv5 的损失函数包含边框回归损失、置信度损失及分类概率损失三部分。在边框回归损失中采用完全交并比(CIoU)损失函数实现预测,公式为

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

$$\alpha = \frac{v}{(1 - R_{IoU}) + v}, \quad (2)$$

$$v = \frac{4}{\pi^2} \left( \arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2, \quad (3)$$

$$R_{IoU} = \frac{A \cap B}{A \cup B}, \quad (4)$$

式中:  $\mathbf{b}$  和  $\mathbf{b}^{gt}$  分别代表预测框和真实框的中心点;  $\rho$  代表两个中心点之间的欧氏距离;  $c$  表示能够同时包含预测框和真实框的最小闭包区域的对角线距离;  $w$  和  $w^{gt}$  分别代表预测框和真实框的宽度;  $h$  和  $h^{gt}$  分别代表预测框和真实框的高度;  $R_{IoU}$  是预测框和真实框之间的交集和并集之比。CIoU 损失考虑了边界框回归的重叠面积、中心点距离和纵横比,但忽略了宽高分别与其置信度的真实差异,阻碍了模型优化的有效性。针对

这一问题, Zhang 等<sup>[19]</sup> 在 CIoU 损失的基础上提出了有效交并比(EIoU)损失函数, EIoU 损失的惩罚项将纵横比的影响因子拆开, 分别计算目标框和锚框的长和宽。EIoU 损失函数包含重叠损失、中心距离损失和宽高损失三部分, 其中宽高损失直接使目标盒与锚盒的宽度和高度之差最小。EIoU 损失公式为

$$L_{EIoU} = L_{IoU} + L_{dis} + L_{asp} = 1 - R_{IoU} + \frac{\rho^2(\mathbf{b}, \mathbf{b}^{gt})}{c^2} + \frac{\rho^2(w, w^{gt})}{c_w^2} + \frac{\rho^2(h, h^{gt})}{c_h^2}, \quad (5)$$

式中:  $c_w$  和  $c_h$  是覆盖预测框和真实框的最小外接框的宽度和高度。在边框回归损失中, EIoU 损失中的宽高损失使得收敛速度更快, 精度更高, 因此本文采用性能更优的 EIoU 损失中的边框回归损失函数。

## 4 实验与结果分析

### 4.1 实验环境与方案设计

采用的实验环境: Win10 操作系统, Python3.8, NVIDIA GTX1660Ti GPU, Intel(R) Core(TM) i5-10300H CPU@2.50 GHz 处理器, CUDA11.1, 调用了 Pytorch 环境库。在模型训练中, 按照 8:1:1 划分训练

集、测试集和验证集,采用 COCO 预训练模型进行参数初始化,训练批次为 8,学习率为 0.01,模型迭代 200 次。

#### 4.2 实验数据处理

实验数据包括轮对踏面擦伤、踏面硌伤及踏面剥离三种缺陷,如图 4 所示。

为了得到更加丰富的踏面缺陷数据集,在训练任务执行之前通过水平翻转、目标裁剪、曝光度增强、高

斯噪声引入等方式进行数据增强,得到由 2730 张图片组成的数据集。通过 LabelImg 软件对图像中的各类缺陷进行标注,随后对训练集、测试集及验证集按照 8:1:1 进行划分,训练集中得到 2184 张图片、测试集中得到 273 张图片及验证集中得到 273 张图片,其中擦伤类缺陷的数量最多,硌伤类的数量较少。表 1 展示了数据集中各类缺陷的数量。

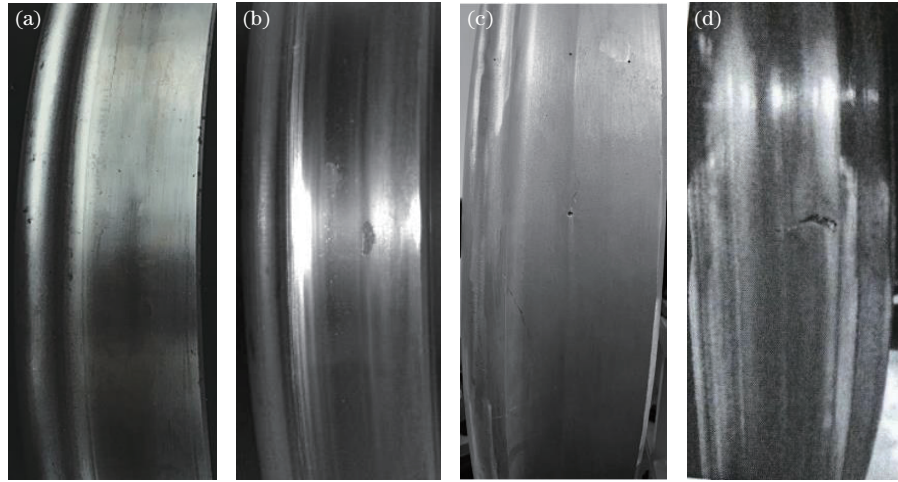


图 4 踏面缺陷类型。(a)正常;(b)擦伤;(c)硌伤;(d)剥离

Fig. 4 Tread defect type. (a) Normal; (b) scratch; (c) laceration injury; (d) peel

表 1 数据中的目标数量

Table 1 Target quantity in data

Defect type	Training set	Validation set	Test set	Total
Scratch	2074	254	266	2594
Laceration injury	1037	143	130	1310
Peel	1425	174	183	1782

#### 4.3 模型评估

采用平均精度均值(mAP)、检测速度来作为主要评价指标,检验模型效果。mAP的计算公式为

$$P_{\text{mAP}} = \frac{1}{C_{\text{class}}} \sum_{i=1}^{C_{\text{class}}} \int_0^1 P(R) dR, \quad (6)$$

式中: $P$ 和 $R$ 分别表示精确度和召回率。它们的计算公式分别为

$$P = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}}, \quad (7)$$

$$R = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}}, \quad (8)$$

式中: $N_{\text{TP}}$ 代表检测结果中正确目标的个数; $N_{\text{FP}}$ 代表检测结果中错误目标个数; $N_{\text{FN}}$ 代表正确目标中漏检的个数。

#### 4.4 实验设计

为了验证所提算法对踏面缺陷检测的有效性,设计了 2 组实验。第 1 组实验验证不同的改进部分对模型性能的影响,另 1 组实验是对所提模型与主流目标检测模型进行性能对比。改进的 YOLOv5 在训练过

程中的损失变化如图 5 所示,可以看出:损失值在开始时下降较快,且波动较大;随着训练轮数的增加,在 50 轮之后波动起伏开始变小,损失曲线逐渐降低并趋于平稳;当轮数达到 200 左右时算法损失降至稳定,模型收敛,训练过程中未出现过拟合现象。

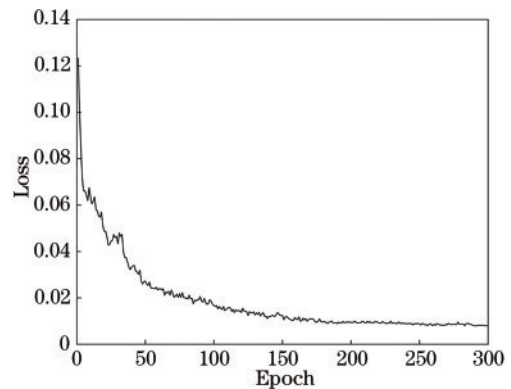


图 5 损失变化曲线

Fig. 5 Loss variation curve

##### 4.4.1 改进方法对模型性能的影响

首先在 YOLOv5 的基础上分析不同改进对网络性能的影响,设计 3 组实验对结果进行分析,每组实验使用相同的训练参数,其中“+”代表在模型中使用了对应的改进策略。表 2 为不同方法给模型检测性能带来的影响结果,表 3 为不同方法对模型复杂度的影响。对表 2 和表 3 进行分析可知:改进 1 在主干网络中引入卷积块注意力机制,使得模型关注重要特

表 2 不同改进的实验结果

Table 2 Experimental results of different improvements

Model	CBAM	Reduce layer	EIoU	Precision / %	mAP / %	Speed / (frame·s <sup>-1</sup> )
YOLOv5				86.3	87.1	45.3
Improvement 1	+			89.2	91.6	44.1
Improvement 2	+	+		88.5	90.8	48.1
Improvement 3	+	+	+	90.7	92.6	48.1

表 3 模型复杂度对比

Table 3 Comparison of model complexity

Method	Model size / MB	Parameters / 10 <sup>6</sup>	FLOPs / 10 <sup>9</sup>
YOLOv5s	13.6	7.03	15.9
Improvement 1	14.0	7.04	16.2
Improvement 2	10.3	5.28	15.5
Improvement 3	10.3	5.28	15.5

征,进而提高模型的检测精度,在原模型基础上 mAP 提升了 4.5 个百分点,但模型尺寸增大 2.9%,模型参数量以及浮点运算量也都有所上升。改进 2 在此基础上对 Neck 区域网络结构进行精简,去掉冗余特征分支, mAP 仅有 0.8 个百分点的损耗,模型大小减小 26.4%,参数量与浮点运算量也有所下降,速度得到 4.0 frame/s 的提升。改进 3 在此基础上将 EIoU 作为边框回归的损失函数,在预测边框回归时考虑到两框的重叠部分、中心点距离和长宽比,使模型优化更加有效,提高模型边界框的定位精度,改进后的 YOLOv5 相较于原始 YOLOv5,精度提高了 4.4 个百分点, mAP

提高了 5.5 个百分点,速度提高了 2.8 frame/s。

#### 4.4.2 主流目标检测模型性能对比

为了验证改进的 YOLOv5 算法对踏面缺陷各个类别的检测性能,对改进算法、主流目标检测算法、加入改进策略的 YOLOv5 进行对比,对比结果如表 4 所示,其中 YOLOv5s\_A 在 Backbone 区引入通道注意力机制, YOLOv5s\_B 将 MobileNet-V2 作为主干特征提取网络。

由表 4 分析可得:改进的 YOLOv5 算法的 mAP 均高于其他算法,比 SSD 高 21.8 个百分点,比 Fast-RCNN 高 4.4 个百分点,比 YOLOv5s\_A、YOLOv5s\_B 及原模型 YOLOv5 分别高 3.4 个百分点、6.4 个百分点及 5.5 个百分点,相比其他主流目标检测网络模型,具有更好的检测精度,同时模型的速度也有一定的提升;相较于原始 YOLOv5 算法,各类别缺陷检测精度中硌伤检测的效果提升较大,擦伤检测的 AP 值由 87.1% 提升到 92.1%,剥离检测的 AP 值由 89.1% 提升到 93.3%,这两类缺陷的检测精度均提升超过 4 个百分点。综合来说,改进的 YOLOv5 检测性能要优于其他算法。

表 4 主流目标检测算法性能对比

Table 4 Performance comparison of mainstream target detection algorithms

Model	AP( $R_{IoU}=0.5$ ) / %			mAP / %	Speed / (frame·s <sup>-1</sup> )
	Tread scratch	Tread bruise	Tread peel		
SSD	72.4	67.9	71.9	70.8	55.4
Fast-RCNN	89.3	88.2	87.2	88.2	2.2
YOLOv5	87.1	85.2	89.1	87.1	45.3
YOLOv5s_A	88.8	88.7	89.9	89.2	45.1
YOLOv5s_B	85.9	84.2	88.6	86.2	52.3
Proposed model	92.1	92.3	93.3	92.6	48.1

#### 4.5 结果分析

图 6 为部分检测结果可视化。可以看出:踏面擦伤和踏面硌伤的检测效果都有所提升,对每个目标的分类有更高的置信度得分;在踏面剥离缺陷的检测中,原始 YOLOv5 出现了错检,而改进后的 YOLOv5 准确地检测出缺陷并正确分类。对于部分缺陷,原始模型存在误检,而改进后的模型能够改善这种情况,并且置信度得分也有所提升,表明改进后的模型具有较强的泛化能力。

高速列车在实际运行中出现的损伤可能因为夹杂、表面粘附其他材料而出现其他表现特征,如杂物包含在损伤中、损伤与表面粘附物存在一定交集等。以踏面擦伤缺陷为例,图 7 展示了在同一张图像的基础上分别添加两类附加特征后的检测结果可视化对比,可以看到添加杂物附着后,所提模型虽然置信度有所下降,但是仍能检测出损伤,具有较强的应对复杂场景的能力。

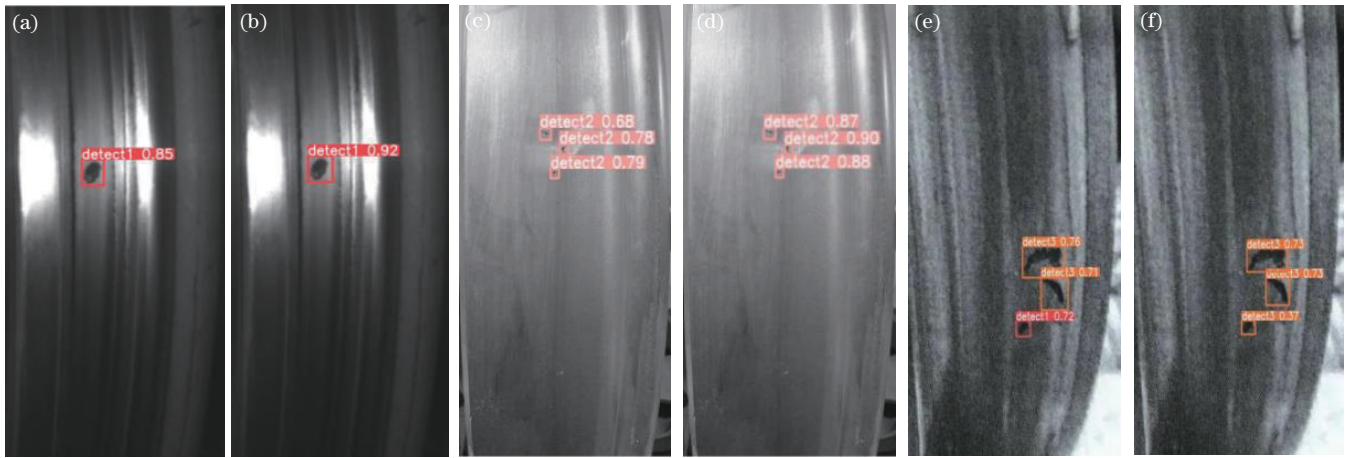


图 6 YOLOv5与改进 YOLOv5检测结果对比。(a)(c)(e)YOLOv5的检测结果;(b)(d)(f)改进 YOLOv5的检测结果  
Fig. 6 Comparison of detection results between YOLOv5 and improved YOLOv5. (a) (c) (e) Detection results of YOLOv5;  
(b) (d) (f) detection results of the improved YOLOv5

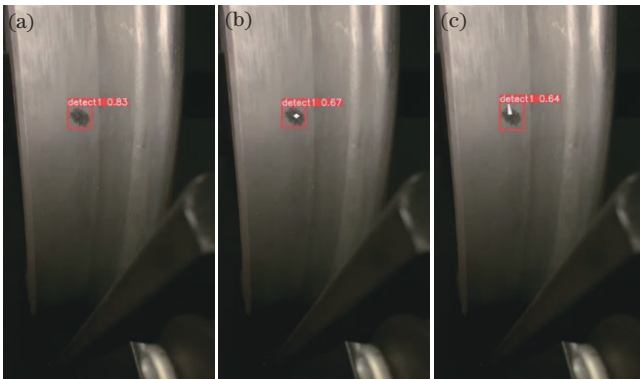


图 7 夹杂材料检测结果对比。(a)擦伤;(b)包含类擦伤;  
(c)交并类擦伤  
Fig. 7 Comparison of inclusion detection results. (a) Scratch;  
(b) including scratch; (c) confluent scratch

## 5 结 论

针对传统目标检测方法不能快速准确地检测踏面缺陷的问题,提出一种基于改进 YOLOv5 的轮对踏面缺陷检测算法。通过旋转、裁剪、增加曝光度、添加噪声等图像增强方法扩充轮对踏面缺陷检测数据集,使所提算法能够学习到更多的缺陷特征,提高泛化能力;通过在主干特征提取网络引入卷积注意力机制,网络更加注重目标特征,从而提高对目标区域的学习能力,同时也会起到一个结合全局信息的作用;通过对缺陷大小类别的分析,对特征图分支进行精简,在保证精度的前提下,进一步提升检测速度;最后将 EIoU 作为边框的损失函数,融合边界框宽高比的尺度信息,提高预测框回归精度。与现有目标检测算法进行性能对比,所提算法对踏面缺陷图像分类识别的准确率最高。这里只对踏面缺陷中常见的三种损伤类别进行分类检测,高速列车在实际运作过程中由于环境气候等不同条件,同一类损伤也会出现其他表现特征,后续还需要进一步扩充数据,进行更多种踏面缺陷类型的分类识别。

## 参 考 文 献

- [1] 丛韬, 韩建民, 张关震, 等. 铁路车轮轮辋疲劳裂纹和踏面剥离掉块的微观损伤因素分析[J]. 中国铁道科学, 2017, 38(5): 93-99.  
Cong T, Han J M, Zhang G Z, et al. Analysis of micro damage factors of shattered rim and tread shelling of railway wheel[J]. China Railway Science, 2017, 38(5): 93-99.
- [2] 金学松, 刘启跃. 轮轨摩擦学[M]. 北京: 中国铁道出版社, 2004.  
Jin X S, Liu Q Y. Tribology of wheel and rail[M]. Beijing: China Railway Publishing House, 2004.
- [3] 侯涛, 张志腾. 改进 Canny 算子在列车车轮踏面损伤检测中的应用[J]. 铁道科学与工程学报, 2018, 15(8): 2107-2112.  
Hou T, Zhang Z T. Improve the application of the Canny algorithm in the detection of wheel tread damage[J]. Journal of Railway Science and Engineering, 2018, 15(8): 2107-2112.
- [4] 吕洪波. 机器视觉系统下的机车轮对踏面损伤检测研究[D]. 南京: 南京航空航天大学, 2017.  
Lü H B. Research on damaged tread detection of train wheel set under machine vision[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2017.
- [5] 宗大公, 梅劲松, 蒋银男. 机车轮对跟随式探伤系统的研究[J]. 电子测量技术, 2018, 41(17): 128-132.  
Zong D G, Mei J S, Jiang Y N. Research on the locomotive wheel to follow-type flaw detection system[J]. Electronic Measurement Technology, 2018, 41(17): 128-132.
- [6] Ren S Q, He K M, Girshick R, 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.
- [7] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu,

- HI, USA. New York: IEEE Press, 2017: 6517-6525.
- [8] 来文豪, 周孟然, 王锦国, 等. 多光谱波段筛选的煤矸石快速定位[J]. 中国激光, 2021, 48(16): 1611001.  
Lai W H, Zhou M R, Wang J G, et al. Fast location of coal gangue based on multispectral band selection[J]. Chinese Journal of Lasers, 2021, 48(16): 1611001.
- [9] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9905: 21-37.
- [10] 郭瑞鸿, 张莉, 杨莹, 等. 基于改进 SSD 的 X 光图像管制刀具检测与识别[J]. 激光与光电子学进展, 2021, 58(4): 0404001.  
Guo R H, Zhang L, Yang Y, et al. X-ray image controlled knife detection and recognition based on improved SSD[J]. Laser & Optoelectronics Progress, 2021, 58(4): 0404001.
- [11] 范丽丽, 赵宏伟, 赵浩宇, 等. 基于深度卷积神经网络的目标检测研究综述[J]. 光学精密工程, 2020, 28(5): 1152-1164.  
Fan L L, Zhao H W, Zhao H Y, et al. Survey of target detection based on deep convolutional neural networks[J]. Optics and Precision Engineering, 2020, 28(5): 1152-1164.
- [12] 肖海鹏. 基于深度学习的机车轮对踏面损伤检测[D]. 南京: 南京航空航天大学, 2019.  
Xiao H P. Wheel tread damage detection based on deep learning[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2019.
- [13] 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 C F, et al. Damage detection of train wheelset tread using Canny-YOLOv3[J]. Journal of Electronic Measurement and Instrumentation, 2019, 33(12): 25-30.
- [14] 张力, 黄丹平, 廖世鹏, 等. 基于目标检测网络的轮对踏面缺陷检测方法[J]. 激光与光电子学进展, 2021, 58(4): 0410020.  
Zhang L, Huang D P, Liao S P, et al. Wheelset tread defect detection method based on target detection network[J]. Laser & Optoelectronics Progress, 2021, 58(4): 0410020.
- [15] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23)[2021-02-04]. <https://arxiv.org/abs/2004.10934>.
- [16] 李彬, 汪诚, 吴静, 等. 改进 YOLOv4 算法的航空发动机部件表面缺陷检测[J]. 激光与光电子学进展, 2021, 58(14): 1415004.  
Li B, Wang C, Wu J, et al. Surface defect detection of aeroengine components based on improved YOLOv4 algorithm[J]. Laser & Optoelectronics Progress, 2021, 58(14): 1415004.
- [17] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 936-944.
- [18] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer Vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 3-19.
- [19] Zhang Y F, Ren W Q, Zhang Z, et al. Focal and efficient IOU loss for accurate bounding box regression[EB/OL]. (2021-01-20)[2021-02-05]. <https://arxiv.org/abs/2101.08158>.