

# 基于目标检测网络的轮对踏面缺陷检测方法

张力<sup>1</sup>, 黄丹平<sup>1\*</sup>, 廖世鹏<sup>2</sup>, 于少东<sup>1,3</sup>, 叶建秋<sup>1</sup>, 王鑫<sup>1</sup>, 董娜<sup>1</sup>

<sup>1</sup>四川轻化工大学机械工程学院, 四川 宜宾 644000;

<sup>2</sup>中国科学院成都计算机应用研究所, 四川 成都 610041;

<sup>3</sup>四川大学机械工程学院, 四川 成都 610065

**摘要** 针对传统图像处理算法难以快速、准确识别轮对踏面缺陷的问题,提出一种采用双深度神经网络对轮对踏面缺陷进行检测的算法。该双网络分为踏面提取网络与缺陷识别网络。根据踏面为大目标的特点,分析与测试 SSD 网络,并用该网络提取轮对图像中的踏面区域。为提高踏面缺陷识别效率,在提取出踏面图像后,针对踏面缺陷属于中、小目标的特点,对 YOLOv3 网络结构进行优化得到 M-YOLOv3。实验测试表明:提取踏面区域时,SSD 算法的精度均值(AP)最高,达 99.8%;识别踏面缺陷时,M-YOLOv3 的 AP 达 89.9%,相较于原始 YOLOv3,单张图像计算耗时减少 7.1%,同时 AP 仅有 0.6%的损耗。结果表明,所提算法具有较高的检测准确率。

**关键词** 图像处理; 轮对踏面; 缺陷检测; 深度学习; SSD 网络; YOLOv3 网络

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP202158.0410020

## Wheelset Tread Defect Detection Method Based on Target Detection Network

Zhang Li<sup>1</sup>, Huang Danping<sup>1\*</sup>, Liao Shipeng<sup>2</sup>, Yu Shaodong<sup>1,3</sup>, Ye Jianqiu<sup>1</sup>,  
Wang Xin<sup>1</sup>, Dong Na<sup>1</sup>

<sup>1</sup> School of Mechanical Engineering, Sichuan University of Science & Engineering, Yibin, Sichuan 644000, China;

<sup>2</sup> Chengdu Institute of Computer Application, Chinese Academy of Sciences, Chengdu, Sichuan 610041, China;

<sup>3</sup> School of Mechanical Engineering, Sichuan University, Chengdu, Sichuan 610065, China

**Abstract** It is difficult to quickly and accurately identify wheelset tread defects using traditional image processing algorithms. We propose an algorithm to accomplish this using a dual deep neural network. The dual network is divided into a tread-extraction network and a defect-identification network. Based on the characteristics of the treads as a big target, we analyze and test the SSD network, and apply this network to extract the tread area from wheelset images. To improve the efficiency of tread defect recognition, after the tread image is extracted, we optimize the YOLOv3 network structure to obtain M-YOLOv3 for the characteristics of medium and small tread defect targets. The experimental results show that when extracting tread areas, the average precision (AP) of the SSD algorithm is the highest (99.8%). When identifying tread defects, the AP of the M-YOLOv3 network reaches 89.9%. Compared with the original YOLOv3, the image computing time of the M-YOLOv3 network is reduced by 7.1%, with the AP showing only a 0.6% loss. The results demonstrate the proposed algorithm's high detection accuracy.

**Key words** image processing; wheelset tread; defect detection; deep learning; SSD network; YOLOv3 network

**OCIS codes** 150.0155; 150.1135; 100.4996

收稿日期: 2020-07-10; 修回日期: 2020-07-28; 录用日期: 2020-08-12

基金项目: 国家自然科学基金青年基金(51704199, 51303115)、四川省科技厅重点研发项目(2019YFG0167)

\* E-mail: hdpvx2002@163.com

## 1 引言

轮对是高速列车的重要部件,在高速、长时间运行过程中,容易使踏面产生表面缺陷,若及时发现,将给列车运行带来极大的安全隐患<sup>[1-2]</sup>。因此,快速准确地检测出踏面缺陷对保障列车的安全运行至关重要。

基于机器视觉的缺陷检测技术可代替人工进行繁琐的轮对踏面缺陷检测。文献[3]介绍了基于机器视觉图像处理的踏面缺陷识别方法,该方法通过分区定位可疑目标来实现对缺陷的识别。文献[4]提出一种基于图像畸变校正的车轮踏面区域提取方法,该方法利用分水岭算法实现踏面区域分割。文献[5]利用阈值分割法对轮缘和轮辋裂纹进行识别。但上述所使用的相关传统图像处理方法存在识别不准确、鲁棒性较低等缺点。

随着深度学习的快速发展,出现了经典的 Faster RCNN 算法<sup>[6]</sup>、YOLO 算法<sup>[7]</sup>及 SSD 算法<sup>[8]</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>[9]</sup>。其中文献[10]采用 DCNN 对铁路路牌进行快速识别,综合准确率达到 80.45%。文献[11]利用 YOLOv3 模型<sup>[12]</sup>定位视频流中无人机位置,再利用比例微分积分算法调节摄像头实现无人机的追踪。文献[13]提出一种改进的 YOLOv3 模型,该模型增强了对小目标的识别能力。文献[14]将 VGG16 与 Faster RCNN 结合,对铁路轨道紧固件进行定位和识别。文献[15]提出改进的 SSD 网络模型,提高了水下目标的识别准确率。文献[16]将目标检测技术运用于农业生产,使用改进的 Faster RCNN 实现自然环境下的苦瓜叶部病害检测。文献[17]搭建了多尺度目标检测网络,引入注意力机制,在红外行人数据集上实现了好的识别效果。

为准确识别轮对踏面缺陷,提高算法的鲁棒性,本文提出一种利用双深度神经网络模型对轮对踏面缺陷进行检测的算法。所提算法将 SSD 作为踏面提取网络,根据踏面缺陷目标尺寸的特点,选择 YOLOv3 作为骨干网络,删除其冗余结构,提高算法的检测效率。由于收集的缺陷样本限制,本文只对与正常区域存在一定灰度值差异且边缘明显的踏面磕伤缺陷进行研究。

## 2 轮对踏面缺陷检测系统结构

### 2.1 轮对踏面信息采集

本检查系统采用机器视觉技术检测踏面缺陷,

利用搭配高强度光源的工业相机,可在列车以 20~30 km/h 时速运行时进行动态检测,其现场采集示意图如图 1 所示。

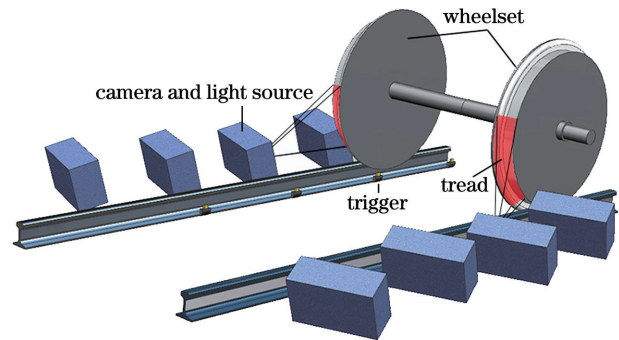


图 1 采集设备示意图

Fig. 1 Schematic of collection equipment

采集原理为:当高速列车轮对激活某个触发器时,通过与之匹配的相机和光源同时工作来进行踏面视觉信息的采集。如图 1 所示,每个相机采集 1/4 圆周的踏面信息,经过 4 对采集设备后可完整获取到一对轮对踏面的图像,并将采集到的轮对踏面信息传输给视觉信息处理单元<sup>[18]</sup>,之后再由后续算法对轮对踏面缺陷进行识别处理。

### 2.2 双网络踏面缺陷检测算法流程

设备采集到的轮对图像如图 2 所示。为检测踏面中的磕伤缺陷,所提在线双网络轮对踏面缺陷检测算法流程如图 3 所示。算法的具体过程:1)对轮对视觉信息进行预处理,对于存在角度偏差的原始轮对图像,进行方向校正;2)根据踏面目标的特点,搭建踏面提取网络,用于在轮对图像中提取踏面区域;3)利用求得的坐标,从原始轮对图像中截取踏面区域,再对踏面图像进行尺寸统一;4)最后根据踏面缺陷的特点,搭建缺陷识别网络,进行缺陷目标的训练与识别。

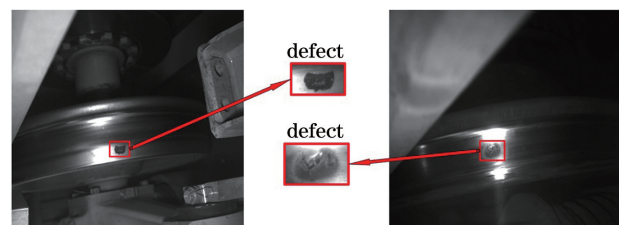


图 2 采集的轮对图像

Fig. 2 Collected wheelset images

整个算法的耗时需满足项目实际要求,以 8 节车厢的动车组列车为例,需要列车经过检测设备后在 1 min 内出具检测报告并进行预警。通过计算,每节车厢有 4 个轮对,每个轮对的单侧需要采集

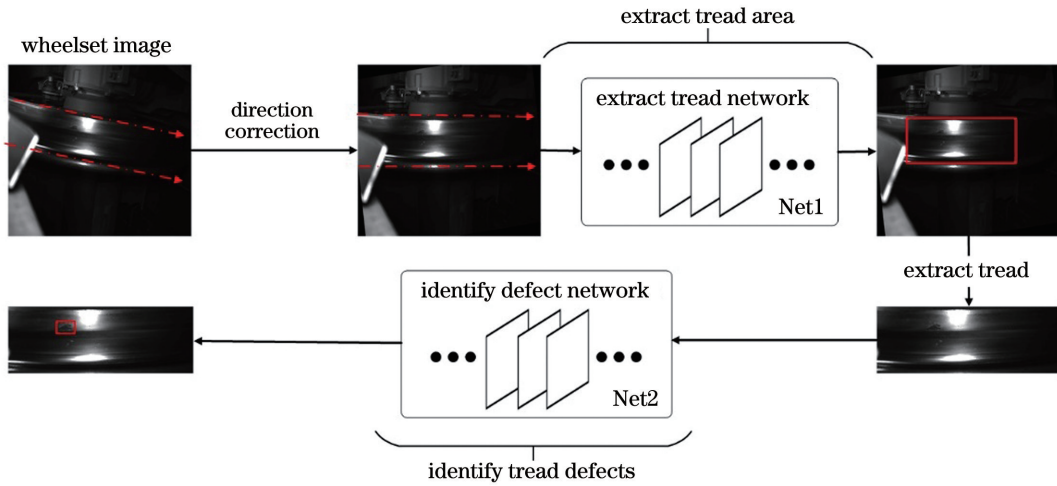


图 3 双网络检测算法的流程

Fig. 3 Flow chart of dual network detection algorithm

4 张图像才能包含完整的轮对踏面区域,所以每节车厢需要采集 32 张图像,整辆列车共需要采集 256 张图像,分配到处处理每张图像的时间为 234.4 ms。但这仅是理论耗时,在实际采集情况下,考虑到数据传输、系统响应延时等因素,分配到每张轮对图像的处理时间将更少。所以在设计深度模型时,在满足识别准确率的基础上,模型计算耗时是本研究重点。

### 3 轮对图像预处理研究

从图 1 可知,由于工业相机与轨道之间的安装位置,采集的轮对图像中踏面区域与水平方向存在一定角度偏差,且不同侧相机采集的踏面区域偏转角度相反,其中踏面区域如图 4 所示。在后续使用深度学习方法识别踏面区域时,方向水平的踏面区域更有利于模型训练与提高识别精度,所以首先应对轮对图像进行方向校正。两组工业相机安装于轨道两侧,分别采集高速列车不同位置的轮对图像,所采集轮对图像可分为两类特征,如图 4 所示。该两类特征分别为轮对附近遮挡物位于图像信息左上角和左下角,由该分类特征可知,图 4 中实线框所示区域的平均灰度值高于虚线框所示区域,根据这一变

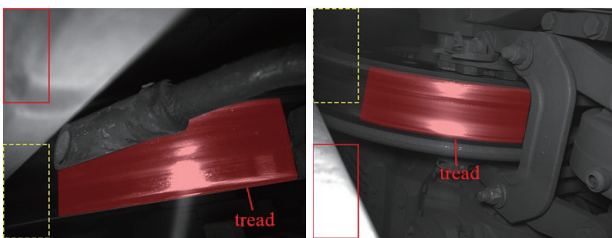


图 4 对比灰度值的区域

Fig. 4 Areas with contrasting gray values

化特征,可快速确定轮对倾斜方向,由此得出图像校正方向。

根据上述原理,轮对方向校正算法为

$$\theta_{\text{rotate}} = \begin{cases} \theta, & M_{\text{up}} > M_{\text{low}} \\ -\theta, & M_{\text{up}} < M_{\text{low}} \end{cases}, \quad (1)$$

式中: $M$  为所选区域的平均灰度值,up 表示左上区域,low 表示左下区域; $\theta_{\text{rotate}}$  为图像旋转角度; $\theta$  为顺时针旋转角度, $\theta=9^\circ$ 。利用所提算法对图 4 所示的原始轮对图像进行方向校正,结果如图 5 所示。

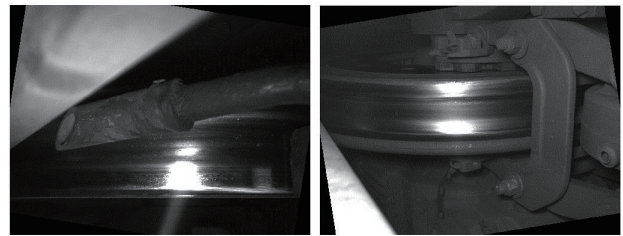


图 5 方向校正后的轮对图像

Fig. 5 Direction-corrected wheelset images

### 4 踏面提取网络

为满足快速性与准确性要求,在识别缺陷目标前,需提取轮对踏面视觉信息。由于踏面视觉信息受到遮挡物、不同时段亮度变化等干扰,快速准确识别轮对踏面区域成为轮对踏面缺陷检测的难点之一。而使用传统图像处理算法识别踏面区域时难以准确定位踏面位置信息,对后续的缺陷检测造成影响。针对上述问题,所提基于深度学习的目标检测方法利用标注的轮对踏面数据集训练网络,使该方法能够在复杂背景中准确找到踏面特征并进行位置预测。

所选目标检测网络应满足实时性要求,通过对

比同为 one-stage 的 SSD 网络 (VGG16) 与 YOLOv3 网络可知:YOLOv3 在 COCO 数据集上的识别准确率更高,对小目标的识别效果更好,但是其模型复杂,耗时较多;由于踏面区域在轮对图像中

属于大目标,SSD 模型较 YOLOv3 模型更简单,对大目标有好的识别效果,并且耗时更少。所以选择 SSD 网络 (VGG16) 作为该阶段的检测模型,其网络模型结构如图 6 所示。

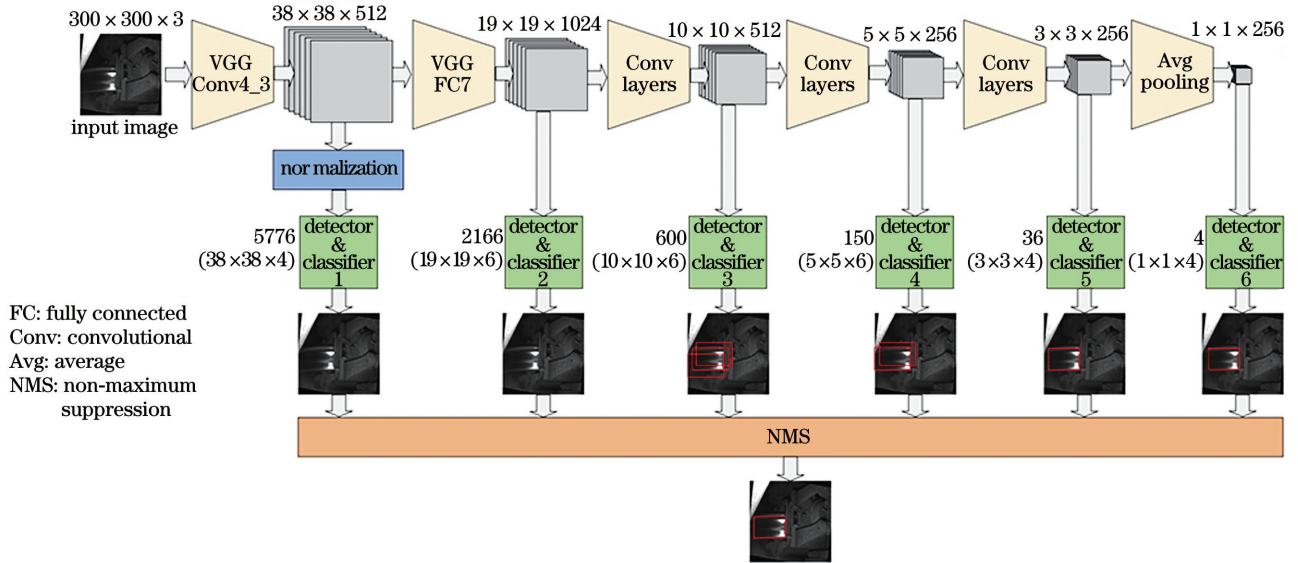


图 6 SSD 网络结构

Fig. 6 SSD network structure

在训练网络之前,为使 SSD 网络计算正确,需将工业相机采集的轮对灰度图像转换为三通道图像,并且图像尺寸也需满足网络要求。网络训练时,尺寸为  $300 \times 300$  的输入轮对图像经卷积神经网络输出特征图,SSD 算法抽取其中的 6 个特征图,并按照特定规则<sup>[19]</sup>对每个 cell 生成不同尺寸的 default box,总共可生成 8732 个 default box。但在 SSD 算法训练过程中,并非所有 default box 都会进行目标的预测,类似于 Faster RCNN 中的 anchor 机制,通过输入图片的 GT (ground truth) 和特征图上对应位置分配的 default box 计算交并比 (IOU),将满足条件的 default box 作为该 GT 的候选框,未被选中的 default box 作为背景的候选框,最后通过优化损失函数来更新网络参数,得到训练好的模型。网络预测时,在最终生成的所有 default box 中,通过设置置信度阈值 (本研究中设置为 0.5),可排除一部分,再对剩余的 default box 运用非极大值抑制 (NMS) 算法,可得出最终的踏面区域预测框。

SSD 算法在训练时采用的损失函数为

$$L(x, c, l, g) = \frac{1}{N} [L_{\text{conf}}(x, c) + \alpha L_{\text{loc}}(x, l, g)], \quad (2)$$

式中:函数分为两部分,计算生成的 default box 与目标类别的置信度损失函数,对应的位置回归损失

函数; $N$  为 GT 满足 IOU 匹配条件的 box 数量,若  $N=0$ ,则将损失设为 0; $\alpha$  为两种损失值之间的平衡系数。

置信度损失函数的定义为

$$L_{\text{conf}}(x, c) = - \sum_{i \in P_{\text{os}}} x_{ij}^p \log \hat{c}_i^p - \sum_{i \in N_{\text{eg}}} \log \hat{c}_i^0, \quad (3)$$

式中:等号右边第 1 项是正样本的损失,即分类类别为目标,等号右边第 2 项是负样本的损失,即分类类别为背景; $x_{ij}^p$  表示第  $i$  个 default box 与第  $j$  个 GT 关于类别  $p$  是否匹配,即 IOU 是否大于设定的阈值, $x_{ij}^p$  等于 1 时表示二者相匹配,等于 0 时表示不匹配;Softmax 损失函数  $\hat{c}_i^p = \frac{\exp c_i^p}{\sum_p \exp c_i^p}$ ,  $\hat{c}_i^p$  为第  $i$  个 default box 判断为类别  $p$  的概率,  $\hat{c}_i^0$  表示判断为背景的概率。

位置回归损失函数的定义为

$$s_{L1}(x) = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & |x| \geq 1 \end{cases}, \quad (4)$$

$$L_{\text{loc}}(x, l, g) = \sum_{i \in P_{\text{os}}} \sum_{m \in \{cx, xy, w, h\}} x_{ij}^k s_{L1}(l_i^m - \hat{g}_j^m), \quad (5)$$

式中:  $\hat{g}_j^{cx} = (g_j^{cx} - d_i^{cx})/d_i^w$ ,  $\hat{g}_j^w = \log \frac{g_j^w}{d_i^w}$ ,  $\hat{g}_j^{cy} =$

$(g_j^{cy} - d_i^{cy}) / d_i^h, \hat{g}_j^h = \log \frac{g_j^h}{d_i^h}; l_i^m$  为网络预测框对 default box 的偏移值;  $g_j^m$  为真实框的值;  $\hat{g}_j^m$  为真实框对 default box 的偏移值;  $(d_i^{cx}, d_i^{cy})$  为匹配到的 default box 的中心坐标,  $(d_i^w, d_i^h)$  为宽和高;  $x_{ij}^k$  表示第  $i$  个预测框与第  $j$  个真实框关于类别  $k$  是否匹配, 取值为 0 和 1。

### 5 缺陷识别网络

经过 SSD 网络计算并识别到踏面区域坐标后, 由于此时踏面图像已被压缩, 为准确检测踏面表面缺陷, 需提取原尺寸踏面图像, 为此, 根据比例变换映射关系, 从原始轮对图像上计算踏面区域坐标并截取轮对踏面视觉信息。在获取踏面缺陷数据集后, 选择合适的网络模型进行训练, 目前深度学习目标检测网络有 SSD、YOLO、Faster RCNN 及 EfficientDet 等。综合考虑检测精度与实时性要求, 选择 YOLOv3 作为缺陷识别网络, 其结构图如图 7 所示。因为相较于 Faster RCNN, YOLOv3 对小目标的检测能力虽然稍弱, 但检测速度更快; 相较于 SSD, YOLOv3 对小目标的检测能力更强; 而 EfficientDet 难以复现论文中的检测效果。但 YOLOv3 网络结构对踏面缺陷检测任务而言存在冗余部分, 故本研究在不改变 YOLOv3 检测精度的前提下, 根据高铁踏面缺陷数据集的特点, 对 YOLOv3 输出结构提出优化方案。

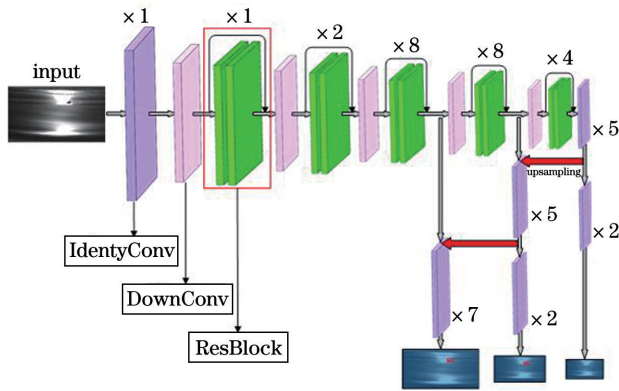


图 7 YOLOv3 网络结构

Fig. 7 YOLOv3 network structure

YOLOv3 借鉴 Faster RCNN 的 Anchor Boxes 方法<sup>[20]</sup>, 帮助模型加快收敛。针对图像尺寸为  $416 \times 416$  的 COCO 数据集, 原始 YOLOv3 网络提供了 9 个 Anchor Boxes 尺寸, 分别为  $(10, 13), (16, 30), (33, 23), (30, 61), (62, 45), (59, 119), (116,$

$90), (156, 198), (373, 326)$ , 由左向右每 3 个尺寸为一组, 分别对应 YOLOv3 输出的  $52 \times 52, 26 \times 26$  及  $13 \times 13$  三个特征张量图, 用于预测小、中、大目标。为得到踏面缺陷尺寸统计量, 将踏面图像尺寸变换为  $416 \times 416$ , 由此可得踏面缺陷目标的尺寸与 Anchor Boxes 分布状态之间的关系, 如图 8 所示。

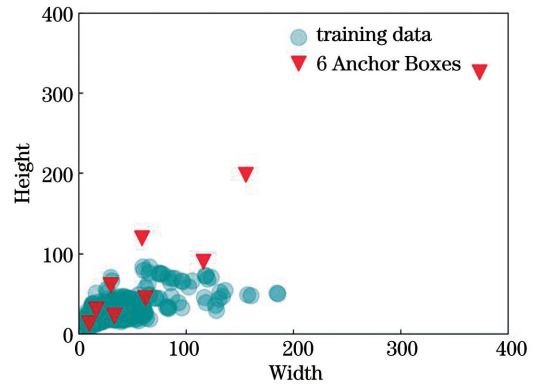


图 8 原始 Anchor Boxes 与目标尺寸分布图

Fig. 8 Distribution chart between original Anchor Boxes and target size

从图 8 可知, 对比这 9 个 Anchor Boxes, 整个踏面缺陷目标尺寸分布在中、小面积区域, 因此用于预测大目标的输出尺度在该数据集上无法发挥作用, 故本研究删除网络结构中的大目标尺度部分以提高计算效率, 该网络记为 M-YOLOv3, 其结构如图 9 所示。

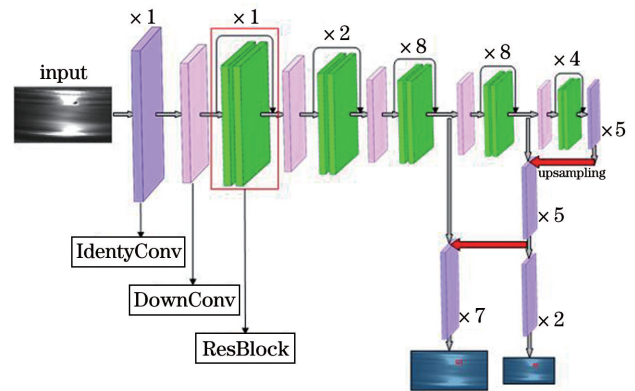


图 9 M-YOLOv3 网络结构

Fig. 9 M-YOLOv3 network structure

M-YOLOv3 只有 2 个输出尺度, 每个输出尺度的特征图上设置 3 个 Anchor Boxes, 所以总共需要 6 个 Anchor Boxes。运行 K-means 算法<sup>[21]</sup>, 在踏面图像尺寸为  $608 \times 320$  时, 聚类生成 6 个 Anchor Boxes, 分别为  $(19, 14), (26, 19), (39, 17), (63, 26), (109, 54), (183, 37)$ , 如图 10 所示。

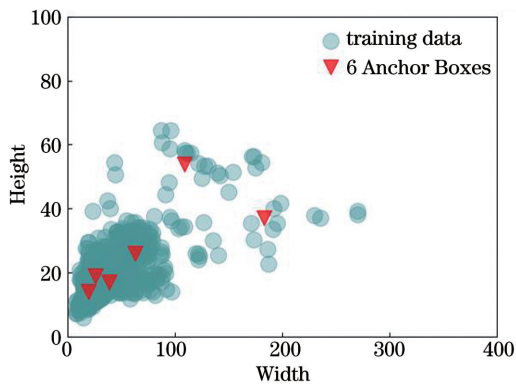


图 10 新的 Anchor Boxes 与目标尺寸分布图

Fig. 10 Distribution chart between new Anchor Boxes and target size

## 6 实验结果与分析

实验条件:操作系统为 Windows10,CPU 为 i5-7500,GPU 为 NVIDIA 1660Ti,深度学习框架为 PyTorch 1.4.0。实验共分为三个部分:第一组实验是 SSD300 与 YOLOv3 系列在识别踏面区域时的对比;第二组实验是 M-YOLOv3 与原始 YOLOv3、YOLOv3-tiny、SSD300、SSD512 在识别踏面缺陷时的对比;第三组实验是原始 YOLOv3 与 M-YOLOv3 加上 SSD300 的组合在轮对图像中识别踏面缺陷时的对比。

在第一组实验中,共选取 2000 张轮对图像,其中验证集 300 张、测试集 500 张、训练集 1200 张,在对训练集进行数据增强后扩充至 3600 张。第二组实验中,在所有的轮对图像中选取包含有缺陷的图像共 712 张,其中验证集 100 张、测试集 200 张、训练集 412 张,在对训练集进行数据增强后扩充至 1648 张。第三组实验的数据设置与第二组实验相同。

由于在两个阶段的检测任务中均只有一类缺陷,所以使用精度均值(AP)<sup>[22]</sup>作为性能度量指标。除此之外,使用网络对每张图片的平均计算耗时来表示计算效率,用时越短表示计算效率越高,模型的计算均在 GPU 上进行。

### 6.1 实验一:踏面提取对比

为验证 SSD300 网络对轮对踏面区域的识别效果,实验对比了 SSD300、YOLOv3 及 YOLOv3-tiny。由于各自网络结构的不同,输入图像的尺寸也不同,对于 SSD300,要求输入图像尺寸为  $300 \times 300$ ;而对于 YOLO 网络,要求输入图像尺寸长宽均为 32 的整数倍,为更准确地与 SSD300 进行对

比,将 YOLO 网络输入尺寸设置为  $320 \times 320$ ,并且轮对图像在输入到每个网络前均需要进行归一化。

实验中,设置 SSD300 网络的学习率为 0.0001,动量参数为 0.9,权重衰减参数为 0.0005,迭代次数为 40000;对于 YOLOv3,设置初始学习率为 0.001,动量参数为 0.9,首先冻结 Darknet53 的参数进行 10000 次迭代,然后解冻 Darknet53 后 20 层的参数,同时学习率更改为 0.0001,再迭代 30000 次;对于 YOLOv3-tiny,设置动量参数为 0.9,以 0.001 的学习率迭代 10000 次后,再以 0.0001 的学习率迭代 30000 次。

三种网络在数据集上的精准率和召回率(P-R)曲线如图 11 所示。

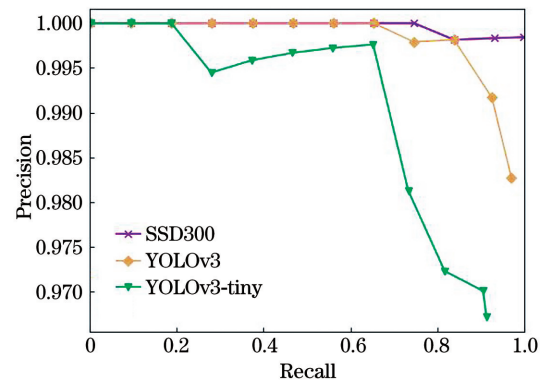


图 11 提取踏面时三个模型的 P-R 曲线

Fig. 11 P-R curves of three models when extracting tread

由于模型预测的踏面区域坐标与踏面区域实际坐标的差异是衡量模型检测踏面准确率的重要标准,所以在检测指标 AP 与耗时的基础上增加预测框与真实框的平均 IOU 指标。三种网络在测试集上的性能如表 1 所示。

表 1 踏面检测性能

Table 1 Tread detection performance

Network	AP / %	Time / ms	IOU / %
SSD300	99.8	31.8	83.2
YOLOv3	96.9	44.3	80.1
YOLOv3-tiny	90.6	13.9	73.6

由表 1 可以看出,在三种网络模型中,SSD300 网络在测试集上的 AP 与 IOU 最高,同时计算单张轮对图像的耗时比 YOLOv3 更少。图 12 为三种模型在部分轮对图像上的识别效果,图中虚线为真实框,实线为预测框,上方的数值为预测框置信度。可以看出,SSD300 在数据集上的泛化能力要比其他

两种网络模型要好。针对上述结果进行分析,由于 YOLOv3 模型参数多,而训练集数量相对较少,不能很好地训练让其收敛到最佳值;YOLOv3-tiny 模型相对较小,无法对轮对踏面特征进行准确提取,导

致识别效果差;而 SSD300 对于轮对踏面这样的大目标能有更好的识别效果,且模型相对简单,能够在小样本数据集上进行训练,并达到一个很好的识别效果。

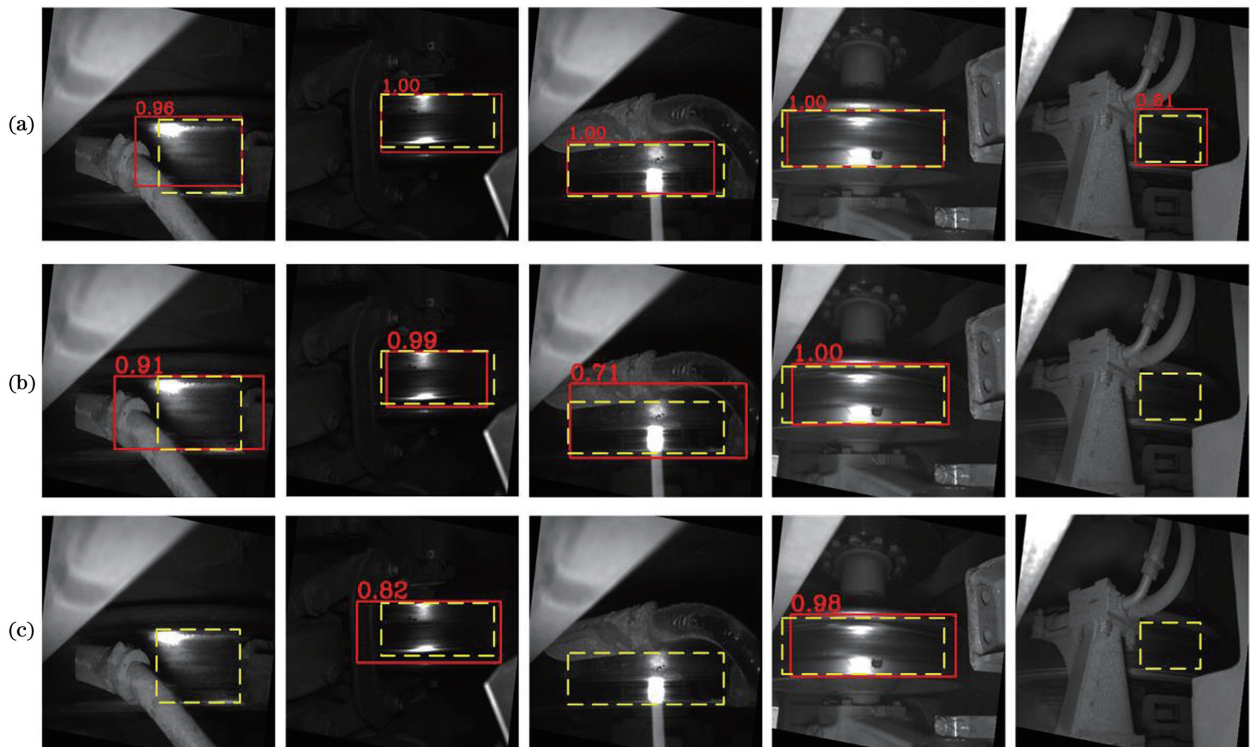


图 12 三种网络的部分识别效果图。(a)SSD300;(b)YOLOv3;(c)YOLOv3-tiny

Fig. 12 Partial identification renderings of the three networks. (a) SSD300; (b) YOLOv3; (c) YOLOv3-tiny

### 6.2 实验二:踏面缺陷识别对比

在利用 SSD 网络求得轮对图像中踏面区域的坐标后,对截取出的踏面图像进行尺寸统一,踏面图像在输入网络前进行归一化。对于 YOLOv3 系列,将图片尺寸变换为  $608 \times 320$ ;对于 SSD300,将图片尺寸变换为  $300 \times 300$ ;对于 SSD512,将图片尺寸变换为  $512 \times 512$ 。其中 SSD300 和 SSD512 网络的超参数设置与实验一中的 SSD300 相同;M-YOLOv3、YOLOv3 网络的超参数设置与实验一中 YOLOv3 相同;YOLOv3-tiny 网络的超参数设置与实验一中 YOLOv3-tiny 相同。

5 种网络在数据集上的 P-R 曲线如图 13 所示。

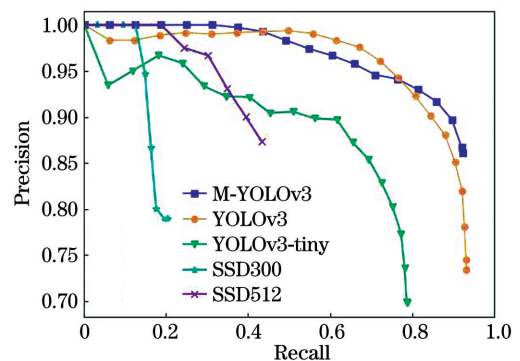


图 13 识别缺陷时五个模型的 P-R 曲线

Fig. 13 P-R curves of five models when identifying defects

表 2 为 5 种网络在测试集上的性能对比。

表 2 缺陷检测性能

Table 2 Defect detection performance

Network	M-YOLOv3	YOLOv3	YOLOv3-tiny	SSD300	SSD512
AP / %	89.9	90.5	72.2	19.3	42.0
Time /ms	58.7	63.2	17.5	24.5	46.4

从表 2 可以看出:原始 YOLOv3 的 AP 在 5 种网络中是最高的;改进的 M-YOLOv3 网络的 AP

与原始 YOLOv3 相比,仅有 0.6% 的损失,并且每张踏面图片的平均耗时减少 7.1%; SSD 算法在

识别踏面区域时,由于踏面区域属于大面积目标,有很高的 AP 值,但在识别踏面缺陷时,缺陷目标在整幅图像中所占像素相对较少,所以 SSD 算法

对小目标检测效果不佳。5 种网络在部分踏面图像上的识别效果如图 14 所示,每个预测框上方数值为置信度。

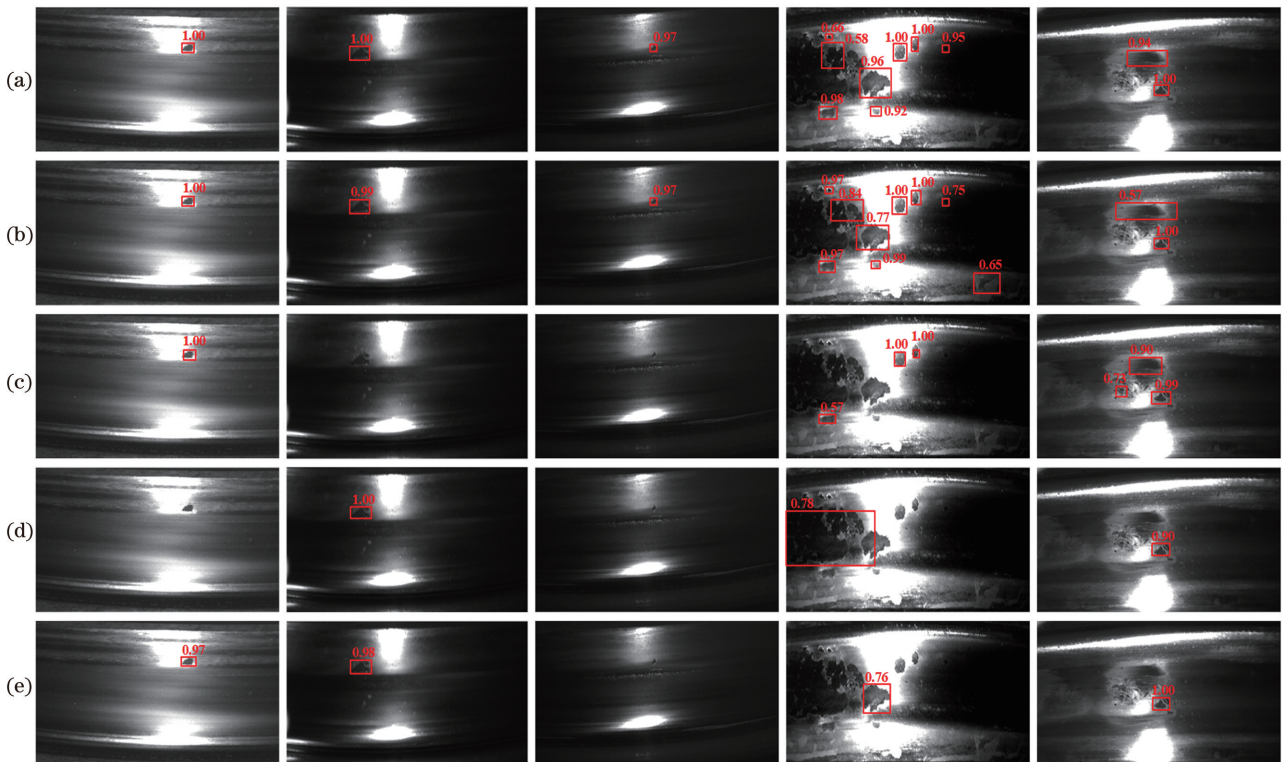


图 14 5 种网络的部分识别效果。(a)M-YOLOv3;(b)YOLOv3;(c)YOLOv3-tiny;(d)SSD300;(e)SSD512

Fig. 14 Partial identification renderings of five kinds of networks. (a) M-YOLOv3; (b) YOLOv3; (c) YOLOv3-tiny; (d) SSD300; (e) SSD512

当缺陷和背景的对比度太低或者缺陷特征不明显时,M-YOLOv3 网络也会出现漏检或误检,如图 15 所示。

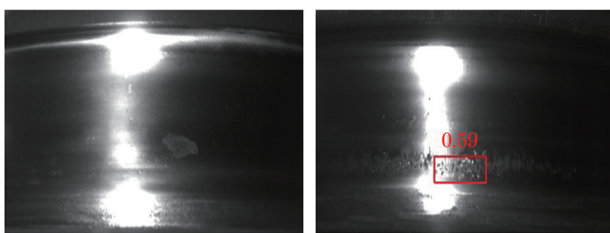


图 15 部分踏面缺陷漏检与误检

Fig. 15 Missing and misdetection of some tread defects

### 6.3 实验三:在轮对图像中直接识别缺陷

当直接在轮对图像中进行缺陷识别时,应保证缺陷在轮对图像中所占像素与在实验二的踏面图像中所占像素接近。对比发现,当轮对图像尺寸为  $960 \times 960$  时,其中踏面区域的尺寸接近于  $608 \times 320$ ,满足实验要求。本节实验使用的 YOLOv3 网络超参数设置与实验一中的 YOLOv3 相同。

对比两个网络模型分步识别缺陷的综合 AP,两者的性能对比如表 3 所示。

表 3 两种检测方法的缺陷检测性能

Table 3 Defect detection performance of two detection methods

Network	AP / %	Time / ms
YOLOv3	73.2	92.7
Exper1 &. Exper2	89.7	92.3

表 3 中 Exper1 &. Exper2 表示 SSD 与 M-YOLOv3 分步识别缺陷的综合性能评价,其中 SSD 识别踏面目标的 AP 为 99.8%,M-YOLOv3 识别踏面缺陷目标的 AP 为 89.9%,所以综合 AP 为 89.7%;平均耗时为两个模型平均耗时的和再加上根据坐标截取踏面区域并进行预处理的用时。

两种方法的 P-R 曲线如图 16 所示。结合表 3 数据分析,训练的 YOLOv3 网络在轮对图像上识别缺陷时,AP 值低于两个模型分步识别缺陷的方法,并且耗时略有增加。



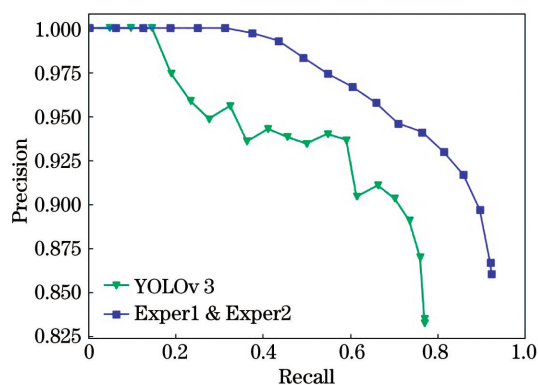


图 16 两种检测方法的 P-R 曲线

Fig. 16 P-R curve of two detection methods

## 7 结 论

针对如何准确、高效检测出轮对踏面缺陷的问题,提出基于双目标检测网络的方法。该方法通过对比区域灰度值实现轮对方向自适应校正,并运用踏面提取网络(SSD)来定位轮对图像中踏面区域的坐标,提取踏面区域,最后根据踏面缺陷特点改进YOLOv3网络,得到缺陷识别网络(M-YOLOv3)。在轮对数据集上进行各网络的对比实验,结果证明,所提检测方法取得了好的效果,并且适合移植到工业现场中。本研究还存在一些问题:1)本研究仅在训练模型阶段对训练集数据进行了数据增强,后面将设计图像增强算法,消除光照不均匀造成踏面左右两侧亮度低的问题,同时修复拍摄原因造成的过曝光与欠曝光情况,提高缺陷与背景的对比如;2)在整个算法设计过程中,由于样本数量的限制,只对主要存在的缺陷类型进行了检测,后续将获取更多的关于其他缺陷类别的数据,来训练网络进行多类别缺陷的预测。

## 参 考 文 献

[1] 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.  
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.

[2] Guo Q, Cui J Y, Zhang Z F. Overview of measurement method for wheels of railway vehicles [J]. Chinese Journal of Scientific Instrument, 2006, 27(S2): 1125-1127.  
郭琼, 崔建英, 张志峰. 铁路车辆轮对检测技术综述 [J]. 仪器仪表学报, 2006, 27(S2): 1125-1127.

[3] Lü H B. Research on damaged tread detection of train wheel set under machine vision [D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2017.  
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.

[4] Wang Y S, Ma Z Q, Song Z B, et al. Wheel tread area extraction based on image distortion correction [J]. Journal of Graphics, 2018, 39(3): 432-439.  
王永胜, 马增强, 宋子彬, 等. 基于图像畸变校正的车轮踏面区域提取 [J]. 图学学报, 2018, 39(3): 432-439.

[5] Pu F P, Zhao J, An Z. Recognition and extraction of wheel cracks based on machine vision [J]. Journal of Railway Science and Engineering, 2018, 15(8): 2113-2122.  
蒲富鹏, 赵军, 安喆. 基于机器视觉的车轮裂纹识别与提取 [J]. 铁道科学与工程学报, 2018, 15(8): 2113-2122.

[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, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection [EB/OL]. (2016-05-09)[2020-07-09]. <https://arxiv.org/abs/1506.02640>.

[8] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector [M] // Leibe B, Matas J, Sbebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9905: 21-37.

[9] 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.  
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.

[10] Meng L, Sun X Y, Zhao B, et al. An identification method of high-speed railway sign based on convolutional neural network [J]. Acta Automatica Sinica, 2020, 46(3): 518-530.  
孟球, 孙霄宇, 赵滨, 等. 基于卷积神经网络的铁路路牌识别方法 [J]. 自动化学报, 2020, 46(3): 518-530.

[11] Tao L, Hong T, Chao X. UAV recognition and location tracking based on YOLOv3 [J]. Journal of Engineering Science, 2020, 42(4): 463-468.  
陶磊, 洪韬, 钞旭. 基于YOLOv3的无人机识别与定

- 位追踪[J]. 工程科学学报, 2020, 42(4): 463-468.
- [12] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08)[2020-07-09]. <https://arxiv.org/abs/1804.02767>.
- [13] Ju M R, Luo H B, Wang Z B, et al. Improved YOLO V3 algorithm and its application in small target detection[J]. Acta Optica Sinica, 2019, 39(7): 0715004.  
鞠默然, 罗海波, 王仲博, 等. 改进的 YOLO V3 算法及其在小目标检测中的应用[J]. 光学学报, 2019, 39(7): 0715004.
- [14] Wei X K, Yang Z M, Liu Y X, et al. Railway track fastener defect detection based on image processing and deep learning techniques: a comparative study [J]. Engineering Applications of Artificial Intelligence, 2019, 80: 66-81.
- [15] Qiang W, He Y Y, Guo Y J, et al. Research on underwater target detection algorithm based on improved SSD[J]. Journal of Northwestern Polytechnical University, 2020, 38(4): 747-754.  
强伟, 贺昱曜, 郭玉锦, 等. 基于改进 SSD 的水下目标检测算法研究[J]. 西北工业大学学报, 2020, 38(4): 747-754.
- [16] Li J H, Lin L J, Tian K, et al. Improved Faster R-CNN for field balsam pear disease detection [J]. Transactions of the Chinese Society of Agricultural Engineering, 2020, 36(12): 179-185.  
李就好, 林乐坚, 田凯, 等. 改进 Faster R-CNN 的田间苦瓜叶部病害检测[J]. 农业工程学报, 2020, 36(12): 179-185.
- [17] Zhao B, Wang C P, Fu Q, et al. Multi-scale infrared pedestrian detection based on deep attention mechanism[J]. Acta Optica Sinica, 2020, 40(5): 0504001.  
赵斌, 王春平, 付强, 等. 基于深度注意力机制的多尺度红外行人检测[J]. 光学学报, 2020, 40(5): 0504001.
- [18] Huang D P, Liao S P, Yu S D, et al. Research on automatic counting system corrugated paper board based on machine vision [J]. Journal of Mechanical Engineering, 2017, 53(10): 79-86.  
黄丹平, 廖世鹏, 于少东, 等. 基于机器视觉瓦楞纸板自动计数系统研究[J]. 机械工程学报, 2017, 53(10): 79-86.
- [19] Zhao X F, Xu M Y, Wang D P, et al. Infrared camouflage detection method for special vehicles based on improved SSD [J]. Infrared and Laser Engineering, 2019, 48(11): 1104003.  
赵晓枫, 徐明扬, 王聘漂, 等. 基于改进 SSD 的特种车辆红外伪装检测方法[J]. 红外与激光工程, 2019, 48(11): 1104003.
- [20] Zhong Y Y, Wang J F, Peng J, et al. Anchor box optimization for object detection[EB/OL]. [2020-07-09]. <https://arxiv.org/pdf/1812.00469.pdf>.
- [21] Kanungo T, Mount D M, Netanyahu N S, et al. An efficient k-means clustering algorithm: analysis and implementation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 881-892.
- [22] Shi W X, Tan D L, Bao S L. Feature enhancement SSD algorithm and its application in remote sensing images target detection [J]. Acta Photonica Sinica, 2020, 49(1): 0128002.  
史文旭, 谭代伦, 鲍胜利. 特征增强 SSD 算法及其在遥感目标检测中的应用[J]. 光子学报, 2020, 49(1): 0128002.