

DOI:10.19651/j.cnki.emt.2211521

基于改进 Faster RCNN 的轮对踏面缺陷检测*

刘应桃¹ 郭世伟¹ 付孟新¹ 张青松²

(1.西南交通大学机械工程学院 成都 610031; 2.西南交通大学牵引动力国家重点实验室 成都 610031)

摘要: 针对日前传统图像处理算法对踏面缺陷检测存在效率不高、对环境鲁棒性不足等问题,本文提出基于改进 Faster RCNN 的踏面缺陷检测方法。改进的网络首先使用 Resnet50 作为特征提取网络,并在特征金字塔层(FPN)特征融合输出部分加入自注意力机制,加强了检测网络对小缺陷的检测能力,最后使用 K-means++ 聚类算法对踏面缺陷数据集锚框进行聚类,并通过聚类结果定制出更适合轮对踏面缺陷的锚框。实验结果表明,改进后的 Faster RCNN 网络对轮对踏面缺陷检测的平均检测速度为 68 ms,平均精度(mAP)达到了 97.3%,对小目标缺陷的检测精度(mAP^{small})达到了 39.3%。

关键词: Faster RCNN;踏面缺陷;特征金字塔;自注意力机制;K-means++

中图分类号: U216.3 **文献标识码:** A **国家标准学科分类代码:** 580.3099

Defect detection of wheelset tread based on improved Faster RCNN

Liu Yingtao¹ Guo Shiwei¹ Fu Mengxin¹ Zhang Qingsong²

(1. School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China;

2. State Key Laboratory of Traction Power, Southwest Jiaotong University, Chengdu 610031, China)

Abstract: To address the problems of inefficiency and lack of robustness to the environment in the current traditional image processing algorithms for tread surface defect detection, this paper proposes an improved tread surface defect detection method based on the Faster RCNN. The improved network first uses Resnet50 as the feature extraction network, and adds a self-attention mechanism to the feature fusion output part of the Feature Pyramid Network to enhance the detection ability of the detection network for small defects, and finally uses the K-means++ clustering algorithm to cluster the anchor frames of the tread defect dataset, and uses the clustering results to customize anchor frames that are more suitable for wheel-to-tread defects. The experimental results show that the improved Faster RCNN network has an average detection speed of 68 ms, an average accuracy (mAP) of 97.3% and an accuracy of 39.3% for the detection of small target defects.

Keywords: Faster RCNN;wheel-to-tread defects;feature pyramid network;self-attentive mechanism;K-means++

0 引言

轮对是列车行走部中最重要的部件之一,其踏面长时间与轮轨接触会出现如擦伤、剥离、疲劳裂纹、压入性凹坑等不规则缺陷^[1-2]。轮对的健康状态将会直接影响到列车的行驶安全,因此针对轮对的缺陷检测具有重大意义。

铁路行业检测量巨大,如果全部依赖人工进行检测将会耗费大量的人力物力,且人工检测具有很强的主观性,人工长时间进行缺陷检测工作很难保持高度的注意力,将造成某些缺陷漏检误检,因此利用机器视觉进行轮对踏面缺陷图像检测变得很有意义。目前针对机器视觉的踏面缺陷特征提取方式可分为传统基于人工的特征提取方式^[3]以及

基于深度学习的目标检测方法^[4]。西北工业大学田丽丽等^[5]对图像进行基于平稳小波和 Canny 融合算法的边缘检测对踏面区域进行提取,利用基于踏面边缘线搜索缺陷区域的方法实现对踏面擦伤、剥离的定位,定位精度分别达到了 97.8%与 96.7%。王永胜等^[6]提出了一种基于畸变矫正的车轮踏面区域提取方法,采用改进的彩色图像标记分水岭算法分割踏面曲面图像,然后根据曲面曲线横向和纵向矫正,最终得到了矫正过后的踏面区域图像,为更加精确定位踏面缺陷区域与计算缺陷面积打下了基础。兰州交通大学张志腾^[7]利用改进 canny 算法结合灰度共生矩阵与 BP 神经网络较好的识别与定位了缺陷的位置。以上传

收稿日期:2022-09-26

* 基金项目:博士后科学基金(2020M682506)、成都科技局重点研发项目(2019-YF05-01823-SN)资助

统图像处理方法虽然能以较高的精度定位出踏面缺陷,但是鲁棒性不高,由于 canny 算法与阈值分割等传统算法都是需要人为设定特征,在背景较为复杂如踏面油污、环境亮度、车体零部件干扰等不利条件下都会对识别精度造成一定影响。近年来,随着基于卷积神经网络的目标检测算法不断发展,出现了以区域卷积神经网络^[8](region convolutional neural network, RCNN)、Fast RCNN^[9]、Faster RCNN^[10]等为代表的双阶段目标检测算法,以及以 SSD^[11](single shot multi box detector, SSD)、YOLO^[12]等为代表的单阶段目标检测算法,文献[13]采用 Faster RCNN 网络为基础实现对踏面二维图像踏面缺陷区域的识别和定位,其识别精度达到了 85%。文献[14]采用 YOLOv3-tiny 实现踏面区域粗定位,利用傅里叶变换、带阻滤波器等传统算法构建缺陷的尺寸测量模型,最后精确计算出缺陷的位置和尺寸,缺陷平均识别精度达到了 89.4%。文献[15]提出了一种 Canny-YOLOv3 的踏面检测方法,采用 Canny 边缘检测算法提取出踏面区域,然后采用 YOLOv3 算法检测出踏面图像中的缺陷。

综上,目前针对轮对踏面缺陷的检测方法主要包括人工静态检测、传统图像处理方法检测以及基于深度学习的目标检测方法,人工静态检测费时费力且易受人的主观性影响,传统图像处理方法主要依赖人工进行特征提取,每个方法针对具体应用和场景,泛化能力不强,同时对复杂环境鲁棒性不高。基于深度学习的目标检测算法改善了传统图像处理方法存在的一些问题,但仍然存在着小目标、数据量不足、数据分布不均等问题。针对这些问题,本文通过改进原始 Faster RCNN 网络以提升对轮对踏面缺陷检测效果,并在不过多损失检测效率的条件下,提升网络的检测精度。

1 实验数据获取与分析

1.1 实验数据采集

本文在列车进库低速运行(0~15 km/h)过程中利用高

强度补光光源搭配高速工业相机进行踏面图像采集,数据采集系统示意图以及现场采集设备如图 1、2 所示。

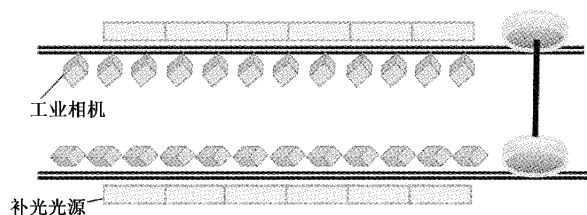


图 1 图像采集系统示意图

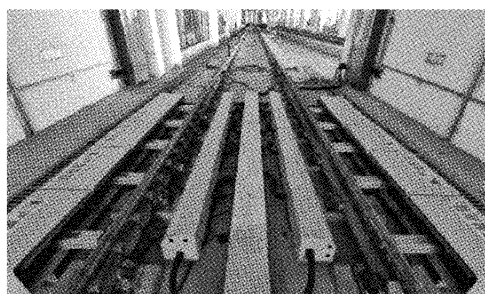


图 2 踏面缺陷采集设备

图像采集原理为:轨道内侧两边各布置 12 台工业相机,轨道外侧两边布置可伸缩光源,当列车靠近时,轨道外侧的补光光源点亮,轮对靠近并激活相机触发装置,与之匹配的工业相机被触发并进行图像采集。每台相机拍摄 30°左右轮周踏面,通过 12 台相机后刚好拍摄完整个轮周踏面,采集到的图像像素为 1 920 pixel×1 200 pixel。

通过实验模拟以及实际过库对不同车次的采集,最终获取的轮对踏面缺陷图像共 656 张,其中踏面擦伤图片 207 张,踏面裂纹图片 136 张,踏面压入性凹坑图片 313 张。3 种缺陷类型示意图如图 3 所示。

1.2 数据集标注与数据分析

本文数据集格式为 PASCAL VOC 数据集格式^[16],使

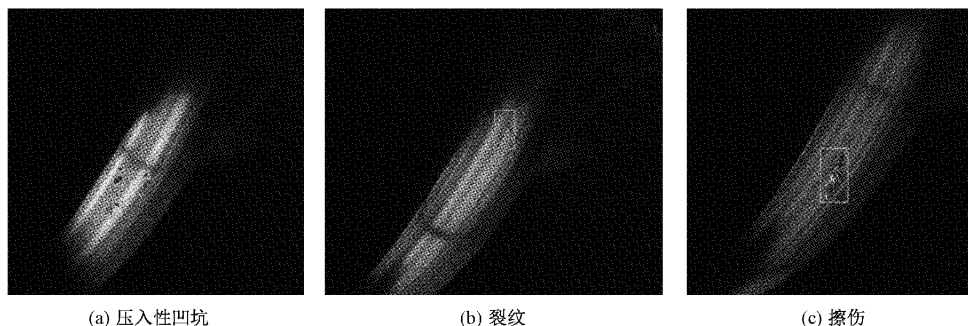


图 3 踏面缺陷类型示意图

用 LabelImg 数据标注工具进行标注,每张图像标注完成之后将生成一个 XML 文件,其中记录了图片的存储路径、文件名、标注框的坐标信息与缺陷类型等。

为了更好的分析踏面缺陷数据特点,训练出更合适的

检测模型,通过对标注框信息进行统计,分析不同大小的缺陷在整个数据集中所占的比例,各尺寸缺陷分布及所占比例如图 4 所示。

目标检测领域一般将像素面积小于 32 pixel×32 pixel

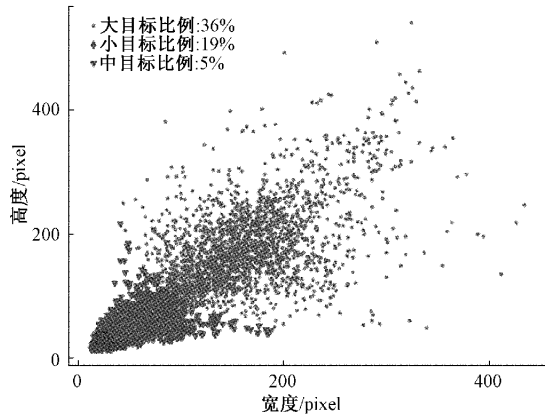


图 4 各尺寸缺陷分布图

的目标定义为小目标,将面积在 $32 \text{ pixel} \times 32 \text{ pixel}$ 与 $96 \text{ pixel} \times 96 \text{ pixel}$ 之间的目标定义为中目标,将像素面积大于 $96 \text{ pixel} \times 96 \text{ pixel}$ 的目标定义为大目标。由图 4 可知,数据集中小目标占比为 19%,中目标占比为 45%,大目标占比为 36%。

2 改进 Faster RCNN 模型构建

2.1 Faster RCNN 网络结构介绍

原始 Faster RCNN 网络结构如图 5 所示,主要由卷积层、区域建议网络(region proposal network,RPN)、感兴趣区域池化(region of proposals pooling,ROI Pooling)、分类与回归层 4 个部分组成。首先使用一系列卷积层如 ZFnet^[17],VGG16^[18],Resnet50^[19] 等特征提取网络对输入的图像进行特征提取并生成特征图,随后将生成的特征图输入到 RPN 层,输出为多个候选区域,具体为目标概率值和坐标信息,其中概率值反应锚框中是否有目标,坐标值

定义目标的位置。将二分类和坐标回归损失统一起来训练 RPN 网络,获取精确的预选框,之后将第一步生成的特征图以及预选框信息在 ROI Pooling 层进一步提取特征图,并将相应的区域以池化的方式获得固定尺寸的特征图,最后将候选区域特征图通过全连接层进行边框回归以及类别分类。

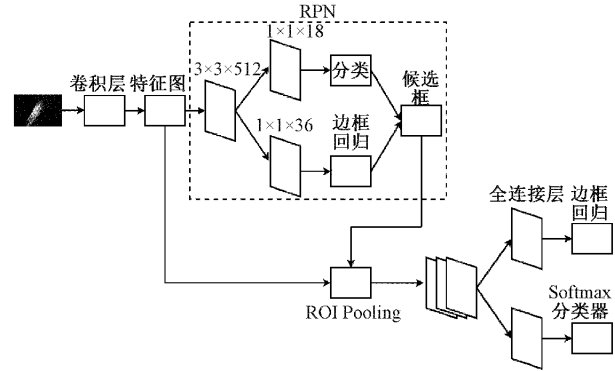


图 5 Faster RCNN 网络结构

改进后的 Faster RCNN 网络结构如图 6 所示。将踏面缺陷图像输入 Resnet50 特征提取网络中进行特征提取,随后为了使网络关注重要特征并抑制非重要特征,在特征金字塔网络^[20](feature pyramid networks,FPN)特征融合阶段加入自注意力机制提升特征表现力;使用 K-means++ 聚类算法对标注框进行聚类,定制出更适合踏面缺陷的锚框;通过 RPN 后得到候选区域,将获取到的特征图输入到 ROI Align^[21]进行候选区域校准,之后再将其输入到后续的全连接层进行类别分类以及边框回归,得到最终的类别、类别分数以及目标区域坐标。下面将具体分析改进的 Faster RCNN 网络。

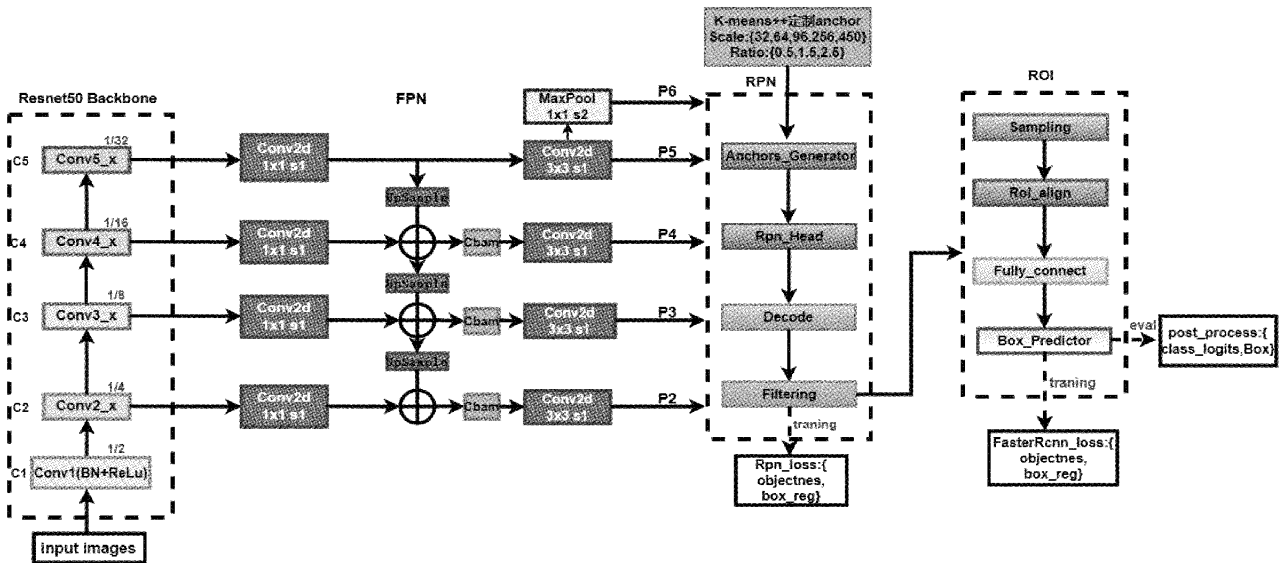


图 6 改进 Faster RCNN 网络整体结构

2.2 FPN 层加入 CBAM 自注意力机制

为降低在训练过程中复杂或与缺陷相似背景的干扰,如踏面外车体零部件、踏面区域局部过曝光等导致缺陷显著度降低,让网络在训练过程中更加关注需要检测的缺陷区域,在 FPN 层上采样融合输出部分加入 CBAM (convolutional block attention module) 自注意力机制^[22]。加入 CBAM 注意力机制后的 FPN 结构如图 7 所示。

CBAM 注意力机制由通道注意力模块与空间注意力模块两部分组成,其结构如图 8 所示。图 8 中, M_c 表示通道注意力机制, M_s 表示空间注意力机制,给定的特征图 $F \in F^{C \times h \times w}$, 其中 C 为通道数, $h \times w$ 为特征图大小。

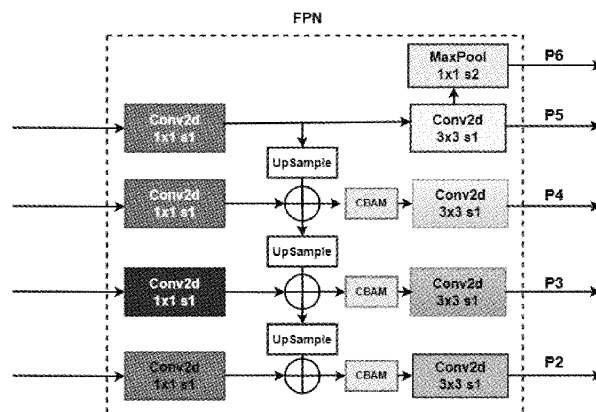


图 7 FPN+CBAM 结构

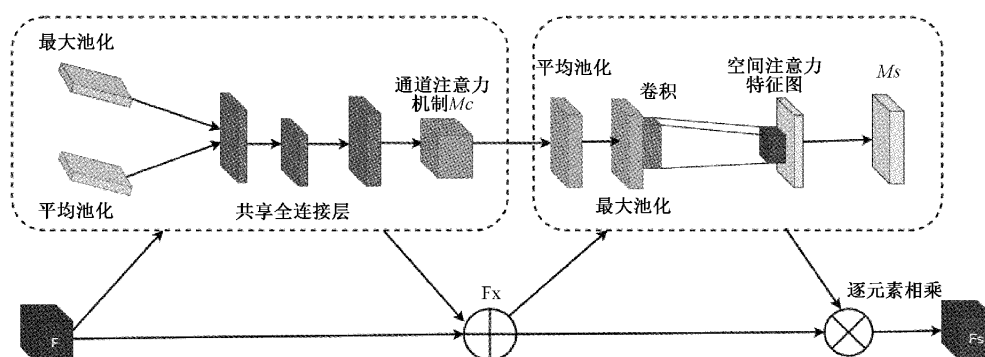


图 8 CBAM 自注意力模块

通道注意力模块的实现又可分为两个部分,首先对输入进来的单个特征图 F 分别进行全局平均池化和全局最大池化,对两个结果利用共享全连接层进行处理并进行相加,之后通过 sigmoid 获得输入特征层每一个通道在 $0 \sim 1$ 之间的权值,再将这个权值乘以原输入特征层从而得到通道注意力特征 $M_c(F)$, 其计算公式如式(1)所示。

$$M_c(F) = \sigma(MLP(AvgPool(F))) + MLP(MaxPool(F)) = \sigma(W_1(W_0(F_{avg}^c))) + W_1(W_0(F_{max}^c)) \quad (1)$$

式中: $\sigma(\cdot)$ 表示 sigmoid 函数; MLP (multilayer perceptron) 为共享全连接层, $AvgPool(\cdot)$ 和 $MaxPool(\cdot)$ 分别表示平均池化操作和最大池化操作; F_{avg}^c 和 F_{max}^c 分别表示在通道注意力机制中的全局平均池化和最大平均池化操作。

空间注意力模块对输入的特征图 F_x 在每一个特征点的通道上取最大值和平均值,再将两个结果进行堆叠生成空间特征图 $F_{avg}^s \in R^{1 \times w \times h}$ 和 $F_{max}^s \in R^{1 \times w \times h}$, 并进行一次通道数为 1 的卷积调整通道数,之后通过 sigmoid 得到输入特征层每一个通道在 $0 \sim 1$ 之间的权值,再用获得的权值与 F_x 逐元素相乘,得到空间注意力特征图 F_s , 其计算公式如式(2)所示。

$$F_s(F) = \sigma(Conv(Cat(F_{avg}^s, F_{max}^s))) \otimes F_x = \sigma(f^{7 \times 7}([F_{avg}^s; F_{max}^s])) \quad (2)$$

式中: Cat 表示连接操作; $f^{7 \times 7}$ 表示 7×7 的卷积运算; F_{avg}^s

和 F_{max}^s 表示空间注意力机制的全局平均池化和最大平均池化运算。

在基础的 Faster RCNN 网络提取特征过程中无注意力偏好,使用的是相同的加权方式对不同特征进行加权。本文通过在 FPN 三次上采样融合位置加入 CBAM 模块,以解决原始网络无注意力偏好的问题,使网络在检测过程中更关注感兴趣区域。

2.3 基于 K-means++ 聚类算法定制锚框方案

基础的 Faster RCNN 中对于锚框的设置是根据 Pascal VOC 数据集人工进行设计的,该数据集的标注框大小并不能很好的代表轮对缺陷数据,在生成候选区域框时会进行大量的回归操作从而影响检测效率,需要针对轮对缺陷数据集的特点对锚框进行定制。本文通过分析轮对缺陷数据集标注框的分布情况,使用 K-means++ 聚类算法^[23]生成锚框方案,首先通过轮对踏面缺陷数据集中的标注框获取标注框的高宽信息,然后设定聚类数目并对宽高信息进行聚类,最后根据得到的聚类结果定制出锚框大小和尺度。

根据 K-means++ 聚类算法分析轮对缺陷数据集标注框的分布,选择聚类中心数量为 15 时的聚类效果如图 9 所示,对聚类中心宽高比分析如图 10 所示。根据 K-means++ 聚类得到的 15 个锚框的宽高以及分析得到的锚框宽高比,本文对锚框尺度重新定制为 $\{32^2, 64^2, 96^2, 196^2, 456^2\}$, 锚框比例定制为 $\{1:2, 3:2, 5:2\}$ 总共 15 种锚框。通过

K-means++ 聚类算法针对轮对踏面缺陷数据集定制出的锚框更能代表轮对踏面缺陷数据的特点,能使检测网络减少边框回归次数,加速检测网络的收敛。

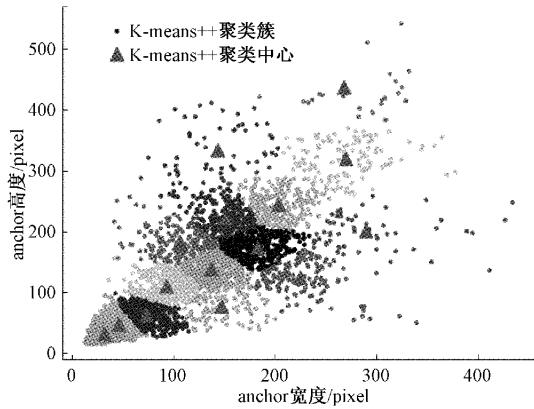


图 9 K-means++ 聚类结果

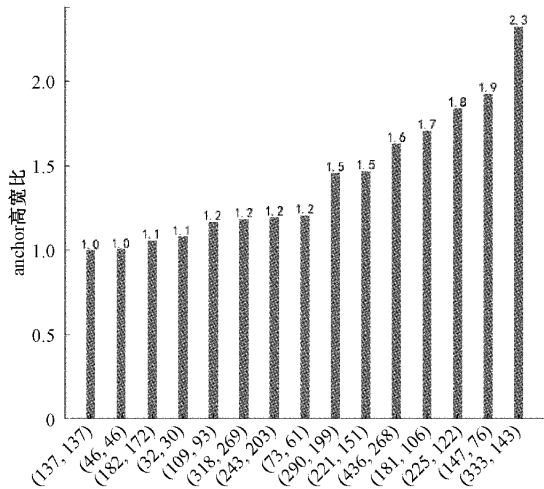


图 10 K-means++ 聚类结果宽高比分析

3 实验与结果分析

3.1 实验环境与超参数配置

本文实验的硬件环境为: Intel(R) Core(TM) i5-10600KF@4.10 GHz 六核处理器, NVIDIA GeForce RTX 3060 12 G 显存显卡, 内存为 16 G, 1 TB 的 ST1000DM010-2EP102 机械硬盘; 软件环境为基于 64 位 window10 操作系统, 深度学习框架为 Pytorch1.10.1 版本, 基于 CUDA11.3, Open CV2 库及 Pycharm 集成开发环境。

由于目前踏面缺陷数据集数量的限制, 本文采用迁移学习的训练方法, 将在 COCO 数据集上训练好的模型作为预训练模型加载到本文的改进网络中, 对踏面缺陷数据进行训练。COCO 数据集中包含了 80 个类别的物体, 数据分布较广, 其中部分图片与本文图片背景具有一定的相似度从而在轮对踏面缺陷数据集上具有较好的泛化能力, 使得模型初始化性能较高且能加快模型的训练速度, 使网络

在较少的训练周期便能获得较好的训练结果。

经过参数微调最终得到的超参数配置为: 优化器采用随机梯度下降(SGD)算法, 优化器动量(momentum)设置为 0.9, 初始化学率(learning rate)为 5×10^{-3} , 权重衰减(weight decay)设置为 5×10^{-4} , 学习率权值衰减倍数(gamma)设置为 0.33, 迭代次数(epoch)为 50, 训练的批大小(batchsize)设置为 8。

3.2 数据集增强

656 张缺陷图像对于 3 类踏面缺陷的检测能力是不足的, 数据量过少会导致模型训练过拟合、模型泛化能力不足以及检测能力下降等问题, 为解决这些问题, 本文针对列车在运行过程中轮对踏面背景的特点, 采用随机变换亮度、随机添加椒盐噪声、随机旋转、随机平移、镜像对称共五种数据增强操作, 将数据扩增为原来的 6 倍并去除一些增强后不符合要求的数据, 最终数据集总共 3 636 张缺陷图片。

数据增强示例如图 11 所示, 增强后的数据将按设定的固定批次输入到网络中进行训练, 只增加了图片的数据量, 不会增加训练的运算量还能增加整个模型的鲁棒性。

3.3 模型训练

将数据增强后的数据集划分为训练集和验证集, 其中训练集 2 727 张图片, 验证集 909 张图片, 在训练时将分辨率为 1 920 pixel × 1 200 pixel 的图像尺寸 resize 到 800 ~ 1 333 之间。将训练集和验证集送入改进后的 Faster RCNN 网络, 并基于加载的预训练模型训练 50 个 epoch, 每迭代一次记录一次结果, 最终总体平均精度(mean average precision, mAP)达到 97.3%, 平均精度总损失以及训练过程中的 mAP 曲线如图 12、13 所示。

由图 12 可知, 随着训练的进行, 当迭代到 24 个 epoch 时损失下降曲线开始趋于稳定, 模型开始收敛。

3.4 评价指标

本文使用 IOU 阈值为 0.5 时的类别平均精度(average precision, AP)与总体平均精度作为轮对踏面缺陷检测结果的评价指标, 其数值越高代表算法的检测精度越高。类别平均精度是每个缺陷类别平均精度, 总体平均精度为所有类别的平均精度的均值, 两者的计算公式为:

$$AP = \int_0^1 P(R) dR \quad (3)$$

式中: P 是精确率; R 是召回率, 两者的计算公式为:

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

式中: TP 是预测正确的正样本数目(true positive), FP 是预测错误的正样本数目(false positive), FN 为预测错误的负样本数目(false negative)。

$$mAP = \frac{\sum_{i=1}^{N_c} AP_i}{N_c} \quad (5)$$

式中: N_c 为检测的类别数目。

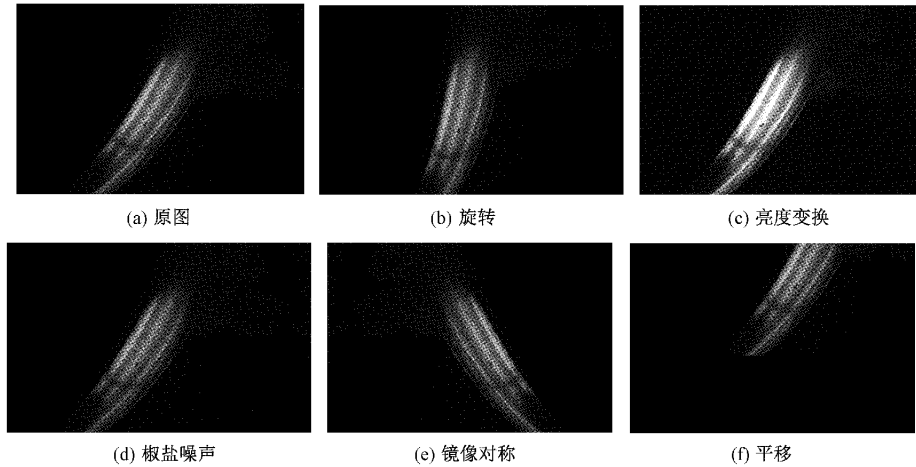


图 11 数据增强实例

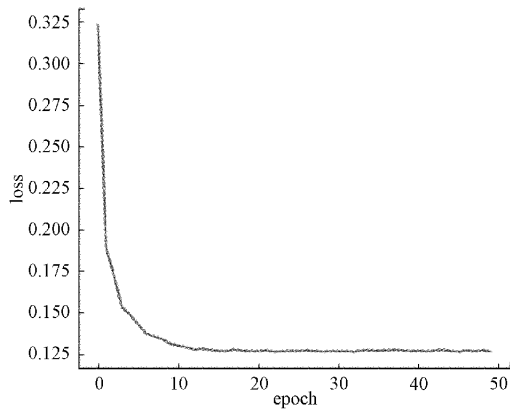


图 12 总损失下降曲线

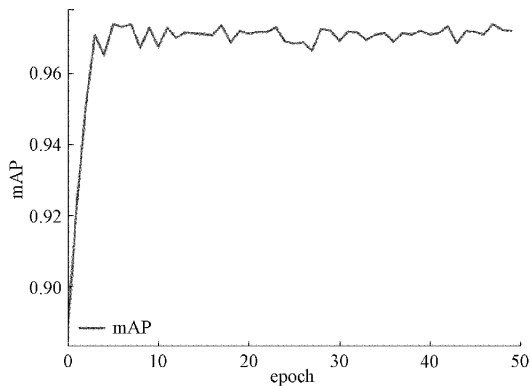


图 13 总体平均精度曲线

3.5 实验结果对比

为了直观的表现出改进后网络的性能,将逐步改进后的网络模型在 3 种类别的平均精度上进行对比,其形式如图 14。

由图 14 可以看出,改进后网络在擦伤类别上的精度提升不太明显,在裂纹、压入性凹坑上面的精度提升较为明显,由此可见改进后网络对于中目标以及小目标缺陷识别具有更好的性能。

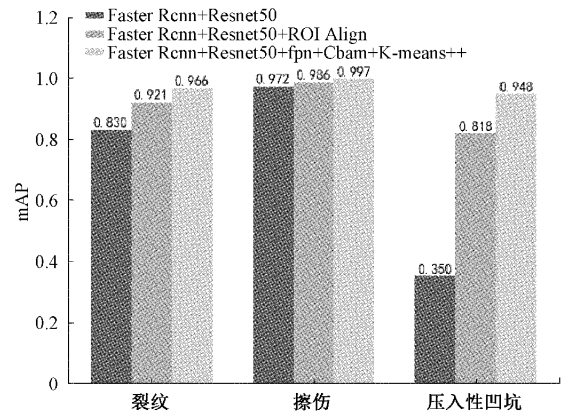


图 14 各类别平均精度比较

为了进一步验证本文改进 Faster RCNN 算法的有效性,以及对小目标缺陷的检测能力,在相同实验条件下分别在训练集以及验证集上对逐步改进的 Faster RCNN 模型进行训练与测试,并比较各模型在 3 种踏面缺陷上的总体平均精度、 mAP^{small} 以及检测速度进行对比,其中 mAP^{small} 为 IOU 阈值取 0.5 : 0.05 : 0.95 时对小目标缺陷检测结果计算的 mAP。比较结果如表 1 所示。

由表 1 可以看出,当 Faster RCNN 采用 Resnet50 为主干网络且采用的感兴趣区域池化方式为原始的 ROI Pooling 时,模型的 $mAP^{0.5}$ 为 71.7%, mAP^{small} 为 9.8%, 检测速度为 84 ms/图,随着对 Faster RCNN 网络的逐步改进, $mAP^{0.5}$ 和 mAP^{small} 都有了一定程度的提升,其中将 ROI Pooling 改为 ROI Align 后提升较为明显,其 $mAP^{0.5}$ 提升了 19.1%, mAP^{small} 提升了 20%。加入 FPN 与 CBAM 后 $mAP^{0.5}$ 提升了 5.7%, mAP^{small} 提升了 6.4%。最终的模型 $mAP^{0.5}$ 为 97.3%, mAP^{small} 为 39.3%, 检测速度为 68 ms/图,满足踏面缺陷检测系统的精度和速度要求。 $mAP^{0.5}$ 和 mAP^{small} 的不断提高验证了对原始 Faster RCNN 改进的有效性,表明改进后的网络在小缺陷检测上具有更好的检测能力与检测效果。

表 1 逐步改进模型性能对比

序号	主干 Resnet50	+ROI Align	+FPN+Cbam	+K-means++	mAP ^{0.5} /%	mAP ^{small} /%	速度/ms
1	✓				71.7	9.8	84
2	✓	✓			90.8	29.8	55
3	✓	✓	✓		96.5	36.2	65
4	✓	✓	✓	✓	97.3	39.3	68

最终改进的 Faster RCNN 模型对轮对踏面缺陷结果 如图 15 所示。

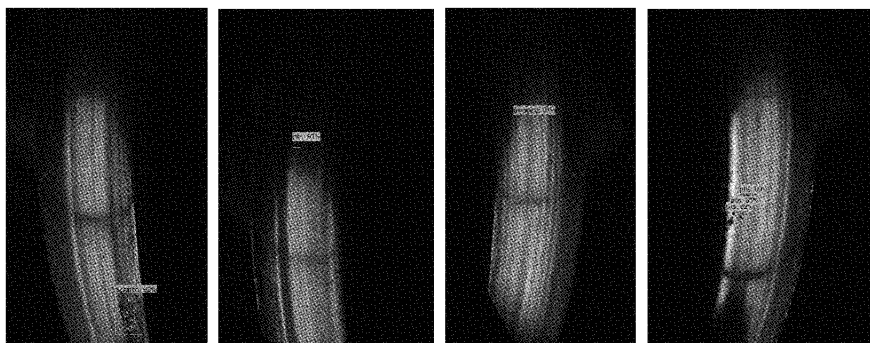


图 15 踏面缺陷检测结果

4 结 论

以运行过程中的列车轮对踏面缺陷为检测对象,对基础 Faster Rcnm 目标检测算法进行改进,改进的网络采用 Resnet50 作为基础特征提取网络,为进一步提高小缺陷的检测精度,在特征金字塔结构中加入 CBAM 自注意力机制增强小缺陷的表现力,最后通过 K-means++ 聚类算法对踏面缺陷数据集进行聚类,并根据聚类结果定制 RPN 网络中的锚框大小与比例,使其更符合轮对踏面缺陷的特点,提高缺陷检测精度与效率。最终结果表明,改进后的检测网络对轮对踏面缺陷平均检测精度达到了 97.3%,对小目标缺陷的检测精度(mAP^{small})达到了 39.3%,相较于基础 Faster Rcnm 分别提升了 25.6%与 29.5%。

改进后的 Faster Rcnm 网络模型在网络分类和回归时使用的全连接层占用了一大部分参数量,并且所有保留的 ROI Align 处理过的 ROI 都会经过这个全连接层并单独计算,因此网络模型复杂度高于传统图像处理算法。后续将尝试采用全卷积网络而去除全连接层、使用正则化等不同策略以达到降低模型复杂度的目的。

参考文献

- [1] 张斌,付秀琴. 铁路车轮、轮箍踏面剥离的类型及形成机理[J]. 中国铁道科学,2001,22(2):73-78.
- [2] 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伤损因素分析[J]中国铁道科学,2017,38(5):93-99.
- [3] 张渝,王黎,高晓蓉,等. 国内外车轮踏面损伤检测技术综述[J]. 机车车辆工艺,2002(1):1-4,8.
- [4] 陶显,侯伟,徐德. 基于深度学习的表面缺陷检测方法

综述[J]. 自动化学报,2021(5):1017-1034.

- [5] 田丽丽,方宗德,赵勇. 铁路货车车轮踏面伤损检测中剥离与擦伤定位方法[J]. 铁道学报,2009,31(5):31-36.
- [6] 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.
- [7] 张志腾. 基于图像处理的列车车轮踏面损伤的识别研究[D]. 兰州:兰州交通大学,2018.
- [8] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA: IEEE, 2014: 580-587.
- [9] GIRSHICK R. Fast R-CNN[C]. Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [10] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [C]. Advances in Neural Information Processing Systems,2015: 91-99.
- [11] LIU W, AAGUELOV D, ERHAN D, et al. SSD: Single shot multi box detector [C]. European Conference on Computer Vision, Springer, Cham, 2016: 21-37.
- [12] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2016: 779-788.
- [13] 曾大为. 基于光电的轮对踏面缺陷检测关键技术研

- 究[D]. 南京:南京航空航天大学,2018.
- [14] 金楷荣,王俊平,陈胜蓝. 基于 YOLOv3-tiny 的列车轮对踏面缺陷定位与测量方法[J]. 控制与信息技术: 1-7.
- [15] 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.
- [16] EVERINGHAM M, VAN GOOL L, WILLIAMS C K I, et al. The pascal visual object classes (VOC) challenge [J]. International Journal of Computer Vision,2010,88(2):303-338.
- [17] ZEILER M D, FER G R. Visualizing and understanding convolutional networks [C]. European Conference on Computer Vision. Springer, Cham, 2014:818-833.
- [18] SIMON Y K, ZISSER M A. Very deep convolutional networks for large-scale image recognition [J]. Computer Science,2014,1409(15):1556-1563.
- [19] HE K, ZHAN G X, REN S, et al. Deep residual learning for image recognition[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016: 770-778.
- [20] 向宽,李松松,栾明慧,等. 基于改进 Faster RCNN 的铝材表面缺陷检测方法[J]. 仪器仪表学报,2021, 42(1): 191-198.
- [21] HE K, GKIOXARI G, DOLLAR P, et al. Mask R-CNN [C]. Proceedings of the IEEE International Conference on Computer Vision IEEE,2017:2961-2969.
- [22] 夏浩宇,索双富,王洋,等. 基于 Keypoint RCNN 改进模型的物体抓取检测算法[J]. 仪器仪表学报,2021, 42(4):236-246.
- [23] 夏禹,肖金球,翁玉尚. 基于改进 Faster-RCNN 的偏光片表面缺陷检测[J]. 光学技术,2021,47(6):695-702.

作者简介

刘应桃,硕士研究生,主要研究方向为深度学习图像处理技术。

E-mail: lytao11@163.com

郭世伟(通信作者),博士,副教授,主要研究方向为机电系统测控技术与深度学习故障诊断技术。

E-mail: guoswall@163.com

付孟新,硕士研究生,主要研究方向为数字图像处理技术。

E-mail: fmengxin@foxmail.com