基于图像处理的轮胎X光图像缺陷检测

姜 明

（青岛科技大学 信息科学技术学院，山东 青岛 266061）

摘要：在图像处理的基础上使用深度学习方法对轮胎X光图像缺陷进行自动检测。轮胎X光图像具有高分辨率、形状狭长、缺陷目标较小的特点，通过将每张X光图像按照640×640像素进行切割，对切割出的每个区域进行标注，把存在缺陷的区域划分到训练集，对训练集进行直方图均衡化以增强图像前景与背景的对比度，进而继续对训练集进行数据增强以提高模型的泛化能力，最后在Faster R-CNN深度学习缺陷检测模型上训练出最优权重。在模型推理阶段，完整的X光图像会被送入模型，缺陷范围被框出，重组为原始X光图像；若某个缺陷具有多个选框，则将所有邻近的选框合成为一个选框。该方法能够有效降低小缺陷目标的漏检率，提高检测的准确率，间接解决了原始X光图像特征丢失的问题。

关键词：轮胎X光图像；图像处理；深度学习；缺陷检测

中图分类号：TQ336.1 **文章编号：**1006-8171(2024)04-0250-06

文献标志码：A **DOI：**10.12135/j.issn.1006-8171.2024.04.0250

OSID开放科学标识码（扫码与作者交流）

随着工业互联网模式的发展，轮胎X光图像缺陷自动检测取代人工检测已成为必然趋势。但由于轮胎X光图像具有高分辨率、过于狭长的形状和相对较小的缺陷目标等特点，典型的深度学习缺陷检测模型在训练之前的对图像重置大小的操作容易对X光图像造成特征损失。使用损失后的X光图像特征进行模型训练几乎是没有意义的，导致缺陷目标的特征权重不高，继而产生缺陷目标漏检的问题。

目前，工业缺陷检测问题高度依赖于人工智能技术，尤其是深度学习。王英玫等[1]通过X光图像纹理的周期性而设计小波函数，实现了对异物和钢丝帘线等缺陷的精确检测。冯霞等[2]采用傅里叶变换产生频谱，通过对频谱进行图像处理来判断轮胎有无缺陷，但无法识别缺陷类型。刘宏贵[3]采用图像处理技术，对比合格图像，检测胎侧气泡和裂缝两类缺陷。顾乃杰等[4]利用边缘分析的图像分割算法对钢丝帘线稀疏缺陷进行检测。吴则举等[5]采用改进的Faster R-CNN模型，引入在线难例挖掘算法来提高轮胎缺陷检测的准确率。李明达等[6]采用Faster R-CNN模型，研究轮胎缺陷检测算法。陈亮等[7]采用Efficient-Net和FPN结合的方式，对轮胎X光图像进行检测，有效提高了轮胎缺陷检测的精度。王鹏辉等[8]采用YOLOv5模型，显著提高轮胎面缺陷智能检测速度。

本工作使用自制的轮胎X光图像缺陷检测数据集，将Faster R-CNN作为基线模型，并对原始图像数据集进行相应的图像处理，既扩充了数据集，又提高了图像特征提取的有效性，在模型推理阶段，原始X光图像将被分割后再送入模型，输出时进行组合，将同一缺陷的所有邻近选框组合成1个选框，虽然会牺牲检测速度，但大幅提高了检测的准确度。

1 缺陷检测算法

缺陷检测是将图像中的缺陷目标检测出来，并识别缺陷的种类。传统的缺陷检测方法是使用统计学来提取图像的特征，这些统计学方法的代表有HOG[9]和SIFT[10]等，然后再依靠机器学习的分类器，如SVM[11]，AdaBoost[12]和DPM[13]等，将图像特征进行缺陷检测，但是其有很强的局限性，因为并不是所有种类的图像都能找到有效的特征提取方法。深度学习则很好地解决了此问题，成为缺陷检测问题的主流算法。

基金项目：国家自然科学基金资助项目(61702295)；泰山学者基金资助项目(tshw201502042)；山东省重点研发计划项目(2017CXGC0607和2017GGX30145)

作者简介：姜明(1996—)，男，山东烟台人，青岛科技大学在读硕士研究生，主要从事计算机视觉、工业互联网和图像处理等研究。

E-mail：jiangminsky@foxmail.com

基于深度学习的缺陷检测算法主要分为两类，即One-stage类和Two-stage类，二者各有其优缺点。One-stage类算法是将图像送入神经网络后，直接检测出图像中缺陷类别和位置，这类算法有SSD[14]和YOLO[15]等。Two-stage类算法是先将图像提取候选框，再将候选框的特征送入神经网络，进而检测出缺陷类别和位置，这类算法有R-CNN[16]，Fast R-CNN[17]和Faster R-CNN[18]等。One-stage类算法在检测速度上优于Two-stage类算法，可以做到实时检测；Two-stage类算法虽然检测速度慢，但具有很高的检测准确率，尤其是擅长小缺陷目标的检测，能够显著降低小缺陷目标的漏检率。由于在生产线上，具有缺陷的轮胎X光图像的数量极少，且其缺陷检测过程是脱离生产线的，不需要做到实时监测，最重要的是检测出缺陷类别和位置，尤其不能漏检，因此本研究选择Two-stage类算法中的Faster R-CNN模型算法作为基线算法，并对其进行相应改进，以提高针对轮胎X光图像中的缺陷检测准确率。

2　缺陷检测算法设计

轮胎X光图像中的缺陷检测算法[19]的流程如图1所示，主要包括准备数据、搭建模型、训练模型和试验分析4个模块。在准备数据方面，需要对图像进行640×640像素切割，对切割完产生的区域进行数据标注，将标注缺陷的区域图片组成模型训练所需的数据集，然后对数据集进行直方图均衡化和数据增强处理。在搭建模型方面，采用Faster R-CNN作为基线模型，ResNet-50[20]作为骨干网络用来提取图像的特征。

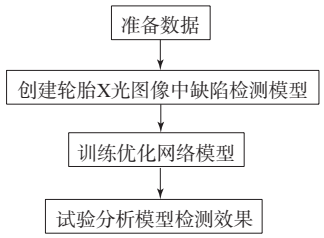


图1　轮胎X光图像中的缺陷检测算法的流程

2.1　数据

本研究所用数据来自真实轮胎生产企业提供的轮胎X光图像，其中部分数据图像如图2所示。

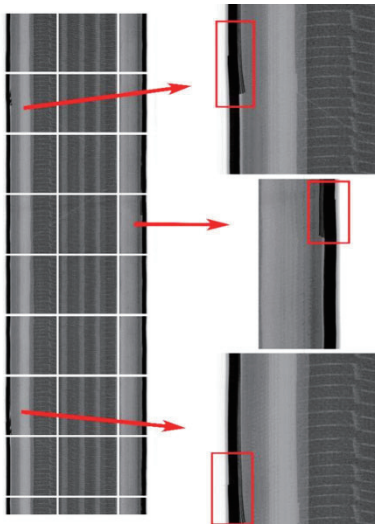


图2　部分数据图像

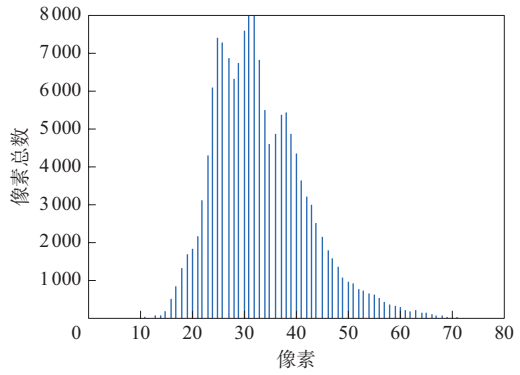
将每张原始X光图像（约1 728×5 300像素）按照640×640像素切割成多个区域图像，然后对每一个区域图像使用LabelImg工具进行缺陷人工标注，凡是被标注的区域图像都作为模型训练的数据集，未被标注的则舍弃，最后得到标注区域图像一共325张。

由于原始图像前景与背景的对比度较小，因此使用直方图均衡化来增强图像的对比度，如图3所示。由于数据较少，容易出现过拟合问题，因此使用数据增强来对数据集进行扩增，数据增强方法有90°旋转、180°旋转、水平翻转、椒盐噪声、高斯噪声和变暗等，扩增后的数据集有2 275张图像。将数据集按照90%和10%的比例随机划分为训练集和验证集。

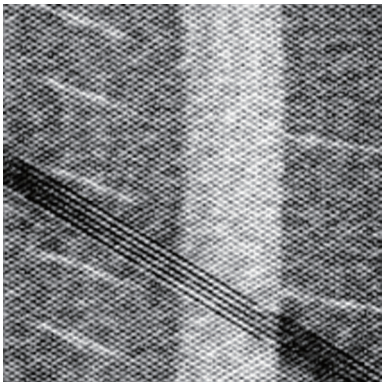
2.2　网络结构

本研究所使用的Faster R-CNN模型网络结构如图4所示，即主要由ResNet-50作为骨干网络对图像进行特征提取，输入图像设置为640×640像素，经过骨干网络提取的特征图被后续的RPN层和全连接层共享。Faster R-CNN第1阶段所需要的精确的区域候选框通过RPN层生成。在RPN层中，首先通过Softmax分类器判断特征图上所有的锚框（anchor）各自属于前景（positive）或背景（negative），锚框是根据事先设置的矩阵大小和长宽比这两个参数计算得到，前景即锚框内容含有目标缺陷，反之即背景，然后通过回归算法将锚框的位置进行精确，最后将同一缺陷类别的

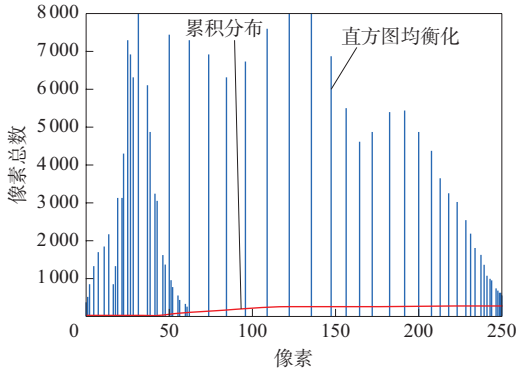
(a) 原始图像



(b) 原始图像对应的直方图



(c) 均衡化图像



(d) 直方图均衡化和累积分布

图3 原始图像和直方图均衡化后图像对比

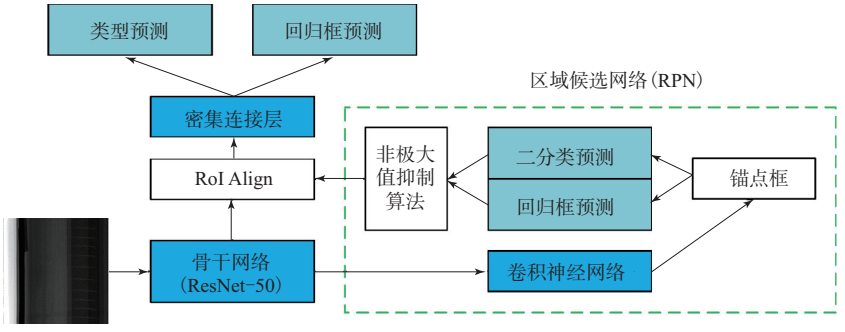


图4 Faster R-CNN模型网络结构示意

多个邻近锚框合并成1个。同时共享特征图的全连接层即RoI Align[21]，通过该层得到最终的选框（经过NMS[22]筛选）相对应的选框特征值，其将被送入后续的全连接层，进而判断选框所属的类别种类。

2.3 损失函数

Faster R-CNN的损失[$L(\{p_i\},\{t_i\})$]主要分为两部分，即RPN层的损失和剩余部分即Fast R-CNN的损失，且二者均包含分类损失[$L_c(p_i,p_i^*)$]和回归损失[$L_r(t_i,t_i^*)$]，如式(1)所示。

$$L(\{p_i\},\{t_i\})=\frac{1}{N_c}\sum_i L_c(p_i,p_i^*)+\lambda\frac{1}{N_r}\sum_i p_i^* L_r(t_i,t_i^*) \tag{1}$$

$$L_c(p_i,p_i^*)=-\lg[p_i^* p_i+(1-p_i^*)(1-p_i)] \tag{2}$$

$$L_r(t_i,t_i^*)=R(t_i-t_i^*) \tag{3}$$

$$R=\mathrm{smooth}_{L_1}(x)=\begin{cases}0.5x^2\times\frac{1}{\sigma^2},|x|<\frac{1}{\sigma^2}\\|x|-0.5\end{cases} \tag{4}$$

$$p_i^*=\begin{cases}0,\text{negative}\\1,\text{positive}\end{cases} \tag{5}$$

$$x = t_i - t_i^* \tag{6}$$

$$\lambda = \text{ceil}\left(\frac{N_r}{N_c}\right) \tag{7}$$

$$t_i = \{t_x + t_y + t_w + t_h\} \tag{8}$$

式中：p_i是anchor预测为目标的概率；p_i^*是真值（ground truth，GT）的标签；t_i是1个向量，表示anchor在回归分析中预测的偏移量；t_i^*（与t_i维度相同的向量）表示anchor在回归分析中相对于GT的实际偏移量；N_c是在训练RPN时选择的anchor数量；N_r是回归分析时的特征图像的数量；λ的存在是为了让分类损失与回归损失的权重基本相同；当RPN训练时，$\sigma=3$，当Fast R-CNN训练时，$\sigma=1$；t_x和t_y以及t_w和t_h分别代表anchor的中心坐标（横坐标值与纵坐标值）以及anchor的宽度和高度。

使用$\text{smooth}_{L_1}(x)$损失函数的原因是：对于选框的预测是一个回归问题，针对回归问题一般选择平方损失函数（也称L_2损失），但是这个损失函数对于出现比较大误差的情况时会有非常高的惩罚，因此使用相对缓和的绝对损失函数（即L_1损失）。L_2损失是平方增长，L_1损失是线性增长，但是在0点处导数不存在，影响收敛。本研究使用分段函数来解决这一问题，即在0点处使用平滑L_1损失函数，通过参数σ来控制平滑的区域。

2.4　模型训练

试验硬件配置为：Intel Core i7-8700 CPU，NVIDIA GeForce GTX 1080 Ti GPU，64GB内存。软件配置为：Ubuntu 16.04（64位）操作系统，Python 3.7语言及编程环境，Pytorch 1.6.0和CUDA 10.1深度学习工具集。

使用ResNet-50对图像进行特征提取时，使用预训练模型权重，预训练权重由在COCO数据集上进行的预训练处理得到。Faster R-CNN模型训练的参数设置如表1所示，在整个训练过程中，采用冻结学习的方法来更新模型参数。模型的整个训练轮次为100个epochs，其中前50个epochs为冻结阶段，在冻结阶段ResNet-50的参数将不被更新，此阶段设置初始学习率为0.001，批处理为32；在后续的解冻阶段ResNet-50的参数也将会被更新，此阶段设置初始学习率为0.000 1，批处理为16。采用Adam梯度下降算法来更新模型的权重。

表1　模型训练的参数设置

参数类型	参数名称	参数值
基本参数	控制参数Beta 1	0.9
	控制参数Beta 2	0.99
	权重衰减	0.000 5
冻结阶段参数	批处理	32
	初始学习率	0.001
	迭代轮次	50
解冻阶段参数	批处理	16
	初始学习率	0.000 1
	迭代轮次	50

2.5　试验

在缺陷检测任务中，通常使用平均精度（Average Precision，AP）评价指标，故本研究试验中也使用AP评价指标，AP值越大，表示目标检测算法的效果越好。针对图像切割过程中对缺陷目标切割范围的随机性（可能会出现切割区域含有非常小的缺陷目标范围的问题），调整不同的置信度，分别设置AP0.1，AP0.5和AP0.9。

图5示出了3张切割后的原始区域图像数据及

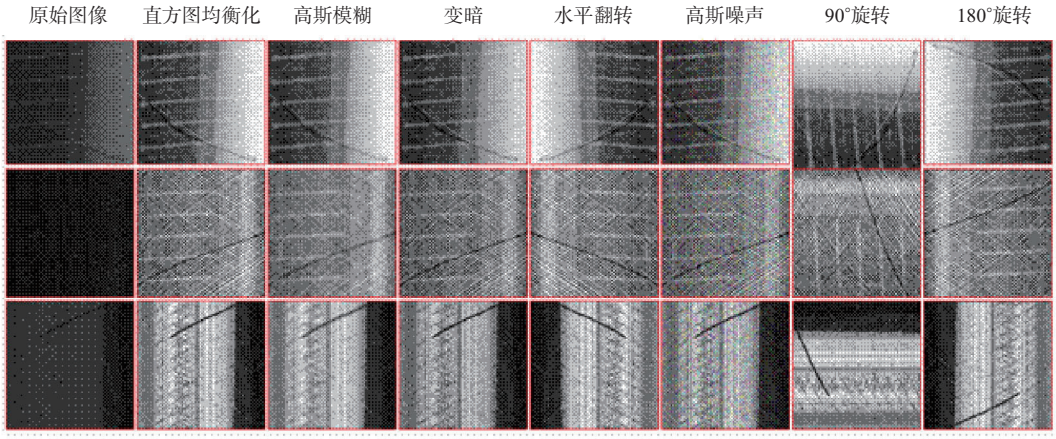


图5　部分切割后的原始区域图像数据及其经过图像增强后的图像数据

其经过7种图像增强后的数据，第1列是原始数据，第2—8列分别是经过直方图均衡化、高斯模糊、变暗、镜像变换、高斯噪声以及90°和180°旋转后的图像数据。

由于模型训练时使用了被分割后的图像，故模型推理过程也有相应的改变，在图像送入推理之前，同样需要被切割，即模型预测的是切割后的区域图像，待所有的区域图像被预测完后，重组成原始图像并且将同一缺陷目标的选框合并，合并过程依据式(9)—(14)，针对同一个缺陷的N个选框，取左上角坐标最大值和右下角坐标最大值，组成新的选框。

$$b_i = (x_{\min}^{\ i}, y_{\min}^{\ i}, x_{\max}^{\ i}, y_{\max}^{\ i}), i \in N \tag{9}$$

$$B = (x_{\min}, y_{\min}, x_{\max}, y_{\max}) \tag{10}$$

$$x_{\min} = \max(x_{\min}^{\ i}) \tag{11}$$

$$y_{\min} = \max(y_{\min}^{\ i}) \tag{12}$$

$$x_{\max} = \max(x_{\max}^{\ i}) \tag{13}$$

$$y_{\max} = \max(y_{\max}^{\ i}) \tag{14}$$

式中，b_i代表第i个选框的坐标参数(左上角坐标，右下角坐标)，B代表最终的合成选框的坐标参数。

模型推理流程如图6所示。

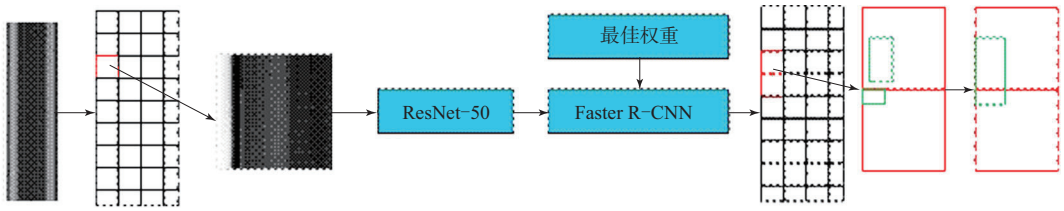


图6 模型推理流程图

本研究模型(图像处理后的Faster R-CNN模型)与原Faster R-CNN，YOLOv3[23]和YOLOv4[24]模型的AP值进行比较，结果见表2。

表2 模型性能对比

方法	AP0.1	AP0.5	AP0.9
YOLOv3	0.75	0.62	0.34
图像处理后的YOLOv3	0.82	0.71	0.59
YOLOv4	0.78	0.68	0.47
图像处理后的YOLOv4	0.88	0.75	0.58
Faster R-CNN	0.85	0.76	0.52
图像处理后的Faster R-CNN	0.98	0.86	0.81

从表2可以看出，本研究模型的AP值最大，AP0.1，AP0.5和AP0.9分别为0.98，0.86和0.81，表明其目标检测算法的效果最好。由于与整个X光图像相比，缺陷的尺寸非常小，而YOLO系列相较于Faster R-CNN具有显著的速度优势，但是对小目标的检测却不擅长，Faster R-CNN则与之相反。训练经过切割后的图像，能够充分提取和学习图像特征，对小目标的检测做到无所遁形，最后合成的选框坐标参数也与真值更加契合。

3 结论

本研究针对轮胎X光图像中存在的缺陷(跳线)检测问题，提出了一种基于Faster R-CNN模型，通过对数据图像进行相关处理并改进模型推理部分以达到更好的检测效果的方法。相较于直接使用YOLOv3，YOLOv4和Faster R-CNN模型，图像处理后的Faster R-CNN模型的AP0.1值达到0.98，目标检测算法的效果非常好。该方法也对尺寸较大、形状狭长、缺陷目标很小且与背景相似的数据提供了试验经验，但也存在不足之处，后续将尝试进一步进行试验和研究，尤其是增强其泛化性，以期该方法能早日发挥其价值。

参考文献：

[1] 王英玫，王传旭. 基于计算机视觉的轮胎缺陷检测[J]. 信息技术与信息化，2013(6)：101-103.

[2] 冯霞，石超，丁文波，等. 基于傅里叶变换的频谱分析法在X射线轮胎检测中的应用[J]. CT理论与应用研究，2014，23(3)：453-458.

[3] 刘宏贵. 轮胎X射线图像缺陷检测算法研究[J]. 传感器世界，2014，20(11)：14-18.

[4] 顾乃杰，张宇翔，张孝慈. 一种基于图像处理的轮胎X光图像缺陷检测方法[P]. 中国：CN 110827272A，2020-02-21.

[5] 吴则举，焦翠娟，陈亮. 基于改进Faster R-CNN的轮胎缺陷检测方法[J]. 计算机应用，2021，41(7)：1939-1946.

[6] 李明达，姜静. 基于深度学习的轮胎缺陷检测算法[J]. 信息技术与信息化，2021(2)：52-53.

[7] 陈亮，白文涛. 基于Efficient-Net的轮胎X光片缺陷检测技术研究[J]. 沈阳理工大学学报，2021，40(2)：8-14.

[8] 王鹏辉,王旭飞,刘怡帆,等. 基于YOLOv5网络的轮胎面缺陷检测分析[J]. 汽车实用技术,2022,47(17):25-30.

[9] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]. Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego: IEEE,2005:886-893.

[10] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2):91-110.

[11] CRISTIANINI N,SHAWE-TAYLOR J. An introduction to support vector machines and other kernel-based learning methods[M]. Cambridge:Cambridge University Press,2000.

[12] SUN X W, ZHOU H B. Experiments with two new boosting algorithms[J]. Intelligent Information Management, 2010, 2(6): 386-390.

[13] FELZENSZWALB P F, GIRSHICK R B, MCALLESTER D, et al. Object detection with discriminatively trained part-based models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2010,32(9):1627-1645.

[14] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]. Proceedings of the 14th European Conference on Computer Vision. Amsterdam:ECCV,2016:21-37.

[15] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas:IEEE,2016:779-788.

[16] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014. 580-587.

[17] LI J N,LIANG X D,SHEN S M,et al. Scale-aware fast R-CNN for pedestrian detection[J]. IEEE Transactions on Multimedia,2018,20(4):985-996.

[18] REN S Q, HE K M, GIRSHICK R. 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.

[19] 何红,陈增云,张亚茹,等. 橡胶复合材料中炭黑微观结构图像的拟合算法[J]. 橡胶工业,2023,70(1):68-74.

[20] HE K M,ZHANG X Y,REN S Q,et al. Deep residual learning for image recognition[C]. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas:IEEE,2016: 770-778.

[21] HE K M, GKIOXARI G, DOLLÁR P. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2020,42(2):386-397.

[22] LIU S, HUANG D, WANG Y. Adaptive NMS: Refining pedestrian detection in a crowd[C]. Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach,CA:CVPR, 2019:6452-6461.

[23] REDMON J, FARHADI A. YOLOv3: An incremental improvement[EB/OL]. [2018-04-08]. https://arxiv.org/abs/1804.02767.

[24] BOCHKOVSKIY A,WANG C Y,LIAO H Y M. YOLOv4:Optimal speed and accuracy of object detection[EB/OL]. [2020-04-23]. https://arxiv.org/abs/2004.10934.

收稿日期:2023-11-16

Defect Detection of Tire X-ray Images Based on Image Processing

JIANG Ming

(Qingdao University of Science and Technology, Qingdao　266061, China)

Abstract: The defects in tire X-ray images were detected automatically by using deep learning methods based on image processing. Tire X-ray images had the characteristics of high resolution, narrow shape and small defect targets. By cutting each X-ray image into 640×640 pixels and annotating each cutted region, the defective regions were divided into a training set, and the training set was histogram balanced to enhance the contrast between the foreground and background of the image. Further data augmentation was performed on the training set to improve the model's generalization ability. Finally, the optimal weights were trained on the Faster R-CNN deep learning defect detection model. In the model inference stage, the complete X-ray image would be fed into the model, the defect range would be framed, and reassembled into the original X-ray image, and if a defect had multiple boxes, all adjacent boxes were combined into one box. The method could effectively reduce the missed detection rate of small defect targets, improve the accuracy of detection, and indirectly solve the problem of feature loss in the original X-ray image.

Key words: tire X-ray image; image processing; deep learning; defect detection