

基于深度学习的 X 射线鞋底异物检测系统^①



肖力炆, 李 伟, 高 荣, 申 浩, 王 孟

(长安大学 信息工程学院, 西安 710064)

通讯作者: 肖力炆, E-mail: liyang_xiao@163.com

摘 要: 在服饰鞋厂的加工生产过程中经常会出现断针现象, 残留在鞋子里的多余断针等金属异物会威胁人们的人身安全. 针对这一问题, 本文提出了一种基于深度学习的鞋底金属异物检测系统. 首先, 将鞋子依次放在传送带上送入检针机, 经过 X 光照射采集图像. 之后对采集到的图像进行预处理操作, 使金属异物变得清晰. 最后通过深度学习网络模型识别当前图像是否含有金属异物, 并检测异物所处位置. 实验结果表明, 经过图像预处理和微调标注框的做法, 能有效提高模型识别的精度. 本文提出模型的平均精度为 97.6%, 该结果表明此模型可以有效检测遗留在各种鞋类中不同形状的金属异物, 具有很好的商业价值.

关键词: 深度学习; Faster R-CNN; 断针检测; 异物检测

引用格式: 肖力炆, 李伟, 高荣, 申浩, 王孟. 基于深度学习的 X 射线鞋底异物检测系统. 计算机系统应用, 2021, 30(3): 88-94. <http://www.c-s-a.org.cn/1003-3254/7820.html>

X-Ray Detection System for Foreign Bodies in Sole of Shoes Based on Deep Learning

XIAO Li-Yang, LI Wei, GAO Rong, SHEN Hao, WANG Meng

(School of Information Engineering, Chang'an University, Xi'an 710064, China)

Abstract: Broken needles are frequently seen in the production process of clothing and shoe factories. This study proposes a detection system of metal foreign bodies in sole of shoes based on deep learning since those residual bodies such as broken needles in shoes will threaten people's safety. Firstly, shoes are put on a conveyor belt in turn and sent to a needle detector, and the images are collected by X-ray irradiation. After that, the images are preprocessed to highlight the small metal foreign bodies. Finally, metal foreign bodies and their positions are detected with a deep learning network model. Experimental results show that preprocessing images and fine-tuning the label box can make metal foreign bodies clearer, and the average precision of the model is 97.6%. It proves that the model can effectively detect the metal foreign bodies with different shapes left in footwear, presenting great commercial potential.

Key words: deep learning; Faster R-CNN; broken needle detection; foreign body detection

随着人们生活水平的不断提高, 人们对生活用品的质量要求也日益提升, 对产品的安全更是重视. 在服饰的生产缝制过程中常常出现断针现象, 残留在衣服和鞋子里的断针、多余针头等金属异物会对人们的生命安全造成很大的伤害. 上世纪 70 年代日本曾发生过一起纺织品断针残留导致儿童受伤致死的事件, 后来

日本立法《检针法》明令所有纺织品以及与人体皮肤接触的所有穿戴品进入市场前必须进行尖锐金属异物检测^[1,2].

随着服饰类加工行业的技术复杂度不断提高, 各种服饰内残留断针的风险也会增大, 对服饰进行金属异物的检测也因此必不可少. 对服饰的异物检测通常

① 基金项目: 国家自然科学基金面上项目 (51978071)

Foundation item: General Program of National Natural Science Foundation of China (51978071)

收稿时间: 2020-07-03; 修改时间: 2020-07-30; 采用时间: 2020-08-19; csa 在线出版时间: 2021-03-03

是对其 X 光图像进行检测, X 光具有穿透性^[3,4], 因此可以避免开包检查成箱的服饰及鞋盒. 传统的服饰异物检测手段主要是利用人工目检进行识别. 目前市面上主要存在两种检测机, 一种是基于电磁感应方式, 这种检测方式无法感应微小断针, 且不能进行成箱级检测^[5]; 另一种是基于射线的人工目检方法, 通过人工分析经过检针机的 X 光图像, 判断是否隐藏断针, 这种方法具有较强的主观性, 且人工目检效率低, 容易出现漏判和错判现象^[6]. 针对这些问题, 本文采用深度学习的方法对服饰和鞋底异物进行目标检测, 让机器代替人工去辨别异物图像, 抛弃了传统检测方法的缺点.

深度学习技术近几年来发展迅速, 其在图像识别、目标检测等领域都取得了很好的应用效果^[7-10]. 近年来出现了 Fast R-CNN^[11]、SSD^[12]、YOLO^[13]、Faster R-CNN^[14] 等一系列目标检测算法, 这些算法在一些公开数据集中都取得了不错的检测效果. 因此也为自动判断 X 光图像中是否存在金属断针提供了可能.

由于现有的服饰类检测的相关产品和研究都较少, 且目前没有 X 光服饰金属异物自动检测方面的研究, 因此本文的主要工作是提出一种基于深度学习的 X 光

服饰异物检测系统, 对生产线上的异物金属 X 光图像进行检测和预警. 在采集服饰断针的 X 光图像数据的同时使用深度学习算法迅速有效地筛选出含有异物金属的服饰, 并将这些异物图像标注出来, 通过系统画面展示给用户, 同时通过蜂鸣器及闪烁的报警灯提示用户发现异物金属. 经过验证, 此检测系统的正确率能够达到 95% 以上, 可以有效检测鞋饰内残留的断针, 该方法可以用于检测遗留在不同类型鞋子中的订书针、细铁屑、断针、鞋钉、钢珠等不同形状的铁质金属异物, 有效降低了异物检针机的使用成本, 提高了对残留在服饰鞋类中断针的检测效率.

1 异物检测系统介绍

本文提出了一种基于深度学习的 X 射线鞋底异物检测系统. 首先将鞋子依次放在传送带上送入检针机, 经过 X 光照射采集图像, 之后通过深度学习网络模型识别当前图像是否含有金属异物, 并检测异物所处位置, 如果判断为含有异物则通过蜂鸣器报警, 最后将检测结果通过图像界面展示出来. 整体流程框图如图 1 所示.

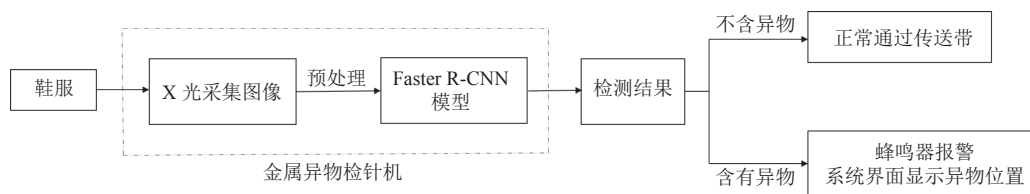


图 1 金属异物检测流程框图

1.1 图像预处理

图 2(a) 为残留在鞋子和鞋盒中的订书针、断针、鞋钉等金属大异物, 图 2(b) 为残留的钢珠等小异物. 由于采集到的图像本身存在噪声干扰, 因此需要对拍摄的 X 光图像进行图像预处理.

本文采用 Sobel 梯度算子锐化来增强图像中的异物. 图像锐化处理的目的是为了增强图像的边缘及灰度跳变的部分, 使图像的细节变得清晰. 它可以突出图像上地物的边缘、轮廓, 或某些线性目标要素的特征, 提高地物边缘与周围像元之间的反差, 因此也被称为边缘增强. 通常采用微分运算可以达到该效果.

Sobel 算子有两个, 一个是检测水平边缘的; 另一个是检测垂直边缘的. Sobel 算子对于像素的位置的影

响做了加权, 可以降低边缘模糊程度, 因此效果更好. 本文采用的是 3×3 模板, 水平方向和垂直方向梯度公式如式 (1) 和式 (2) 所示, 模板如图 3 所示.

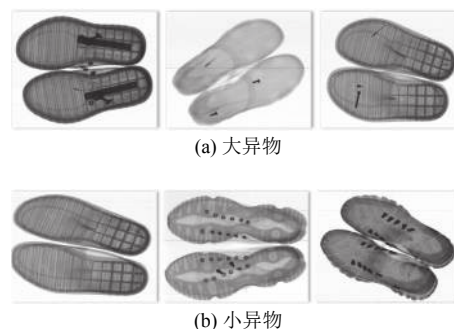


图 2 不同形状的金属异物

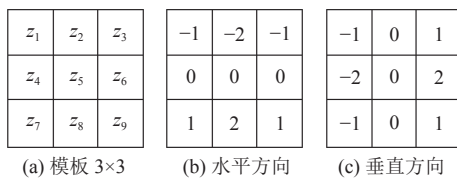


图3 Sobel 梯度算子

$$g_x = \frac{\partial f}{\partial x} = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3) \quad (1)$$

$$g_y = \frac{\partial f}{\partial y} = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7) \quad (2)$$

经过锐化后的图像再进行中值滤波,消除图像中的椒盐噪声的同时可以保留跳变点的尖锐度.图4为图像为处理前和处理后的结果,从图中可以看出,预处理之后的这些金属异物变得更加清晰可见,有利于后续模型的训练和检测.

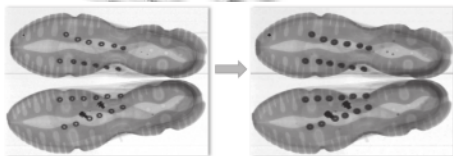


图4 图像预处理后的效果

1.2 Faster R-CNN 算法

目标检测领域的算法从 RCNN 开始,通过引入卷积神经网络取得了很大的进展,但由于选择性搜索(selective search)产生候选区域这一步,需要耗费大量的时间,还是未能摆脱传统的区域建议算法的限制^[15].后来提出的 Fast R-CNN 算法^[11]中提到如果去除区域建议的算法,网络就可以接近实时进行检测,其中的原因是,卷积神经网络运行在 GPU 上,而选择性搜索是运行在 CPU 上的.为了进一步减少检测网络的运行时间,微软的任少卿等提出了最新的目标检测方法 Faster R-CNN^[14],这种方法从提高区域建议的速度出发提出了区域建议网络 RPN (Region Proposal Network) 并通过 GPU 实现了快速区域建议的运算,RPN 的出现代替了之前的选择性搜索方法,它和检测网络共享全图的卷积特征,使得区域建议检测几乎不花时间.Faster R-CNN 在生成区域建议上的改进,使得目标检测的效率得到了极大的提升.现如今, Faster R-CNN 不仅用于物体的检测,还可以用于许多商业的领域,如服装检测,人脸识别等,已经成为了通用的检测框架.因

此本文采用 Faster R-CNN 模型作为鞋服金属异物的检测算法.

Faster R-CNN 的网络结构如图5所示.待检测图像没有固定的尺寸要求,网络结构中的 ROI Pooling 层用来将输入图像的尺寸转换为固定尺寸,以满足后续全连接网络层的要求.待检测图像首先通过深度神经网络来生成图像特征,这些特征既被候选的 RPN 使用,又被后续的检测过程使用,即实现了两个不同目的的特征共享,这也就是图中被称为共享神经网络的原因.

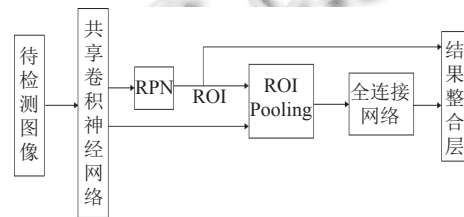


图5 Faster R-CNN 算法的网络结构

Faster R-CNN 提出了一种加速区域建议计算的方法,即建立 RPN 网络.RPN 是 Faster R-CNN 中的核心部分,RPN 的网络结构如图6所示.其输入数据是来自共享卷积网络的输出,维度是 $1 \times 512 \times 38 \times 50$,即1张图片,512个特征图(feature map),每个feature map的尺寸是 38×50 .用一个滑动窗口(3×3 的卷积核)在一个卷积特征图上进行滑动,这个卷积特征图是由共享卷积网络得到,滑动操作可以获得一个512维的向量,将这个向量送入到两个平行的全连接层,即分类层(cls layer)和位置回归层(reg layer),用来得到分类信息和位置信息.

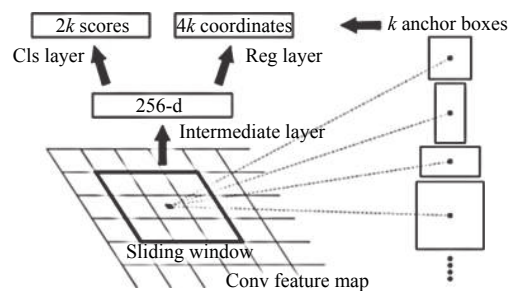


图6 RPN 结构模型

每个滑动窗口的中心会对应 k 个锚点(anchor).每个 anchor 都会对应不同种的尺寸和长宽比.RPN 使用3种尺寸 $\{128 \times 128, 256 \times 256, 512 \times 512\}$ 和3种长宽比 $\{1:1, 1:2, 2:1\}$,所以每个滑动窗口就有 $3 \times 3 = 9$ 个 anchor.

相应的,在每个滑动窗口的位置同时预测9个区域建议,所以位置回归层有 4×9 个输出,代表9个区域建议框的位置坐标;分类层有 2×9 个输出,代表对每个建议框是否是目标的概率。

为了训练RPN,对于每个anchor需要先定义一个二值类别标签(判断是否是异物)。本文规定,正标签为:(1)与真实窗口有最高IoU重叠率的anchor;(2)与任意真实窗口的IoU大于0.7的anchor。负标签为:与所有真实窗口的IoU都小于0.3的anchor。舍弃那些既不是正标签也不是负标签(IoU在0.3到0.7之间)的anchor。

本文对损失函数定义为式(3):

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (3)$$

$$L_{cls}(p_i, p_i^*) = -\log[p_i^* p_i + (1 - p_i^*)(1 - p_i)] \quad (4)$$

$$L_{reg}(t_i, t_i^*) = R(t_i - t_i^*) \quad (5)$$

$$smooth_{L1}(x) = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & \text{其他} \end{cases} \quad (6)$$

其中, i 代表anchor的索引, p_i 代表anchor预测为目标的概率,如果该anchor为正标签时 $p_i^* = 1$,为负标签时 $p_i^* = 0$ 。 $t_i = \{t_x, t_y, t_w, t_h\}$ 是一个向量,表示预测框的4个坐标, t_i^* 是与正标签对应的真实窗口的坐标向量。式(4) $L_{cls}(p_i, p_i^*)$ 是两个类别(是否是目标)的对数损失;式(5) $L_{reg}(t_i, t_i^*)$ 是回归损失;式(5)中 R 是Smooth L1函数,即式(6)。 $p_i^* L_{reg}$ 意味着只有当 $p_i^* = 1$ (即判断为目标时)才有回归损失。

N_{cls} 和 N_{reg} 都是设定值, λ 为平衡因子,用来权衡分类损失和回归损失, λ 值越大越重视回归, λ 值越小越重视分类(文中 λ 取10)。

1.3 实现细节

本文使用的数据是由检针机中内置的X光相机拍摄的鞋子X射线图像(如图7所示),由于X光相机采集到的图像尺寸很大^[16-18],且为了提高运行效率,数据在进入模型训练之前必须要resize将尺寸变小,由于断针钢珠等金属异物本身就是小目标,resize之后会变得更小,因此需要在resize的同时对标注文件中每个异物点的标注框也进行相应的扩大,才能保证模型训练时的匹配。

本文使用VIA标注工具进行鞋底异物的标注,对于每一个标注文件,在进入模型训练之前采用程序对

标注框的4个顶点坐标进行扩大,使得图像resize之后仍然可以完整包围异物点,图8中的内框为原始标注框,外框为扩大之后的框。

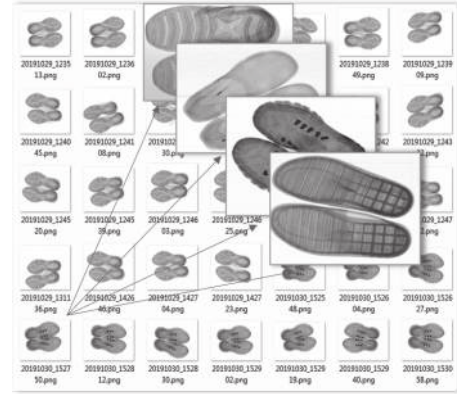


图7 采集不同种类的鞋底X光图像

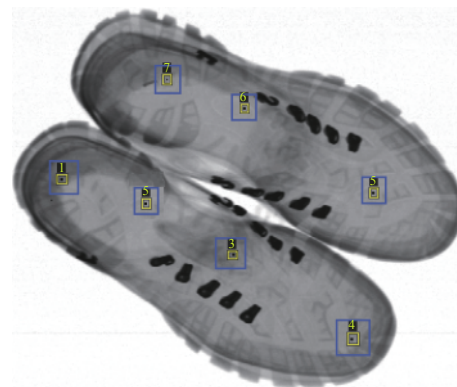


图8 微调标注框的结果

2 实验结果与分析

2.1 数据采集及环境平台

图9为本文开发的鞋底金属异物检针机,表1为该检针机的技术参数。为了模拟鞋服中残留金属异物的情况(订书针、细铁屑、断针、鞋钉、钢珠等),如图10所示,在进入X射线拍摄之前人为的在每双鞋子中放入不同种类和个数的金属异物,按顺序依次送入检针机,经过内置的X射线相机照射采集鞋子X光图像并保存到计算机中,用于模型的训练和测试。为了适用于多种版型的检测对象,本实验使用的数据有板鞋、休闲鞋、运动鞋、拖鞋等多种鞋型以及装箱的鞋子数据。本实验采集的X光图像共计9851张,按9:1的比例将样本数据随机划分为训练集和测试集,即训练集8866张,测试集985张,每张图像的尺寸为 5000×1024 像素。



图9 鞋服金属异物检测机

表1 鞋底异物检针机的技术参数

参数	值
传送带速率	<40 m/min
检测能力	直径1.0 mm的金属物
传送带最大负荷	5 kgs
分辨力	AWG40
功率损耗	1.0 KVA
X射线发射角	80°
X射线传感器	0.4 mm
X射线传感器	光电二极管阵列单能探测器
照射方向	由上向下
冷却方式	密闭油冷
图像处理	24位真彩色实时处理



图10 模拟鞋子生产中残留的金属异物

本文使用 MMDetection 深度学习目标检测工具箱来实现 Faster R-CNN 算法, MMDetection 是一个基于 Pytorch 的开源对象检测工具箱, 支持 Faster R-CNN, Mask R-CNN, Fast R-CNN 等主流的目标检测框架^[19,20], 其优点是训练速度快、所需显存小. 本文实验所使用的实验设备以及软件版本的详细信息如表2所示.

为有效训练 Faster R-CNN 模型, 本文设置初始学习率为 0.02, 动量参数为 0.9, 权重衰减为 0.0001, 总共迭代 100 次, batch size 设为 1, 训练过程中的 loss 曲线如图 11 所示, 可知, 迭代 50 次以后 loss 基本趋于平稳.

表2 实验环境配置详情

设备名称	版本
PC机操作系统	Linux
GPU	1080ti
CUDA	Cuda_9.0.176_windows
CUDNN	Cudnn-7.0-windows
MMDetection	V2.0
Python	3.0
PyTorch	1.5
Anaconda	Anaconda3-5.0.1-Windows-x86_64

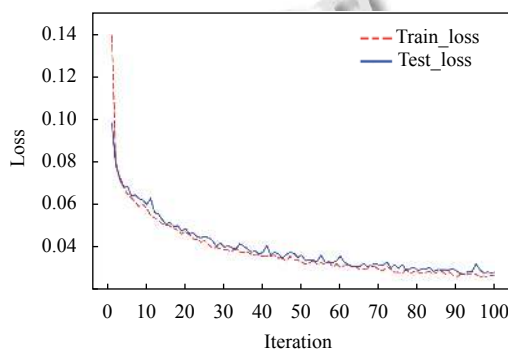


图11 异物检测模型损失函数 loss 曲线

2.2 实验对比

为了验证该方法的有效性, 本文采用精度 (Pr)、召回率 (Re) 和平均精度 (AP) 对模型进行评价, 计算公式如下.

$$Re = \frac{TP}{TP + FN} \tag{7}$$

$$Pr = \frac{TP}{TP + FP} \tag{8}$$

$$AP = \int_0^1 PrRe dRe \tag{9}$$

其中, TP 表示正确提取异物信息的个数, TN 表示正确识别背景信息的个数, FN 表示未被检测出来的异物信息个数, FP 表示错误提取异物信息的个数. AP 是 Pr-Re 曲线下的面积, 可以反映模型的性能好坏, 该值越大代表模型性能越好.

实验结果的 PR 曲线如图 12 所示, 由图可知, 加了预处理和调整标注框之后的模型性能最优. 表 3 是模型训练结果的 AP 值, 分别对是否经过图像预处理和调整标注框进行了对比, 从表中可以看出, 仅使用 Faster R-CNN 模型对异物进行识别, 检测精度只有 13.9%, 只能识别到部分的大异物, 像钢珠点等小异物根本无法检测到. 而经过图像锐化和滤波操作之后, 异物变得清

晰可见,检测精度也有了提升.同时针对图像本身尺寸大、异物目标小的问题,对标注文件进行改动(扩大标注框),可以防止resize缩小尺寸之后目标看不见的情况.经过这些改进,检测精度提升到了97.3%,因此本文提出的模型可以有效检测遗留在鞋底中的金属异物.

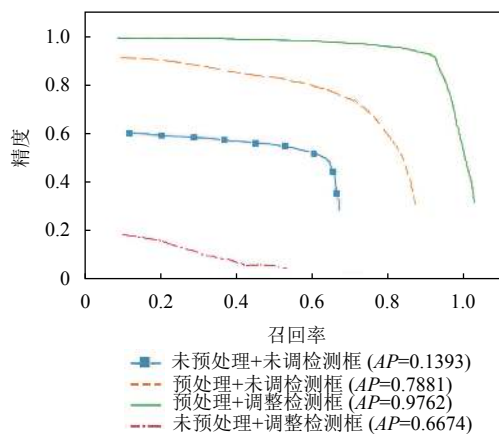


图12 异物检测模型的P-R曲线

表3 鞋底数据模型训练结果

编号	实验对比	AP (%)
1	未预处理+未调标注框	13.9
2	未预处理+调整标注框	66.7
3	预处理+未调标注框	78.8
4	预处理+调整标注框	97.6

2.3 检测结果

图13展示了一些鞋底异物的检测效果,从图中可以看出,该方法可以正确区分鞋子的鞋带环和金属异物,对于复杂的异物形态,以及不同种类的金属异物均可以做到精准的检测.

实验结果表明,该系统对于不同种类以及有复杂背景的鞋子,都能达到准确的区分,具有可靠的精度.且通过此深度学习训练出来的模型不需要对每一款鞋型进行专用的模版设计,有效降低了异物检测机的使用成本,提高了对残留在服饰鞋类中断针的检测效率.

3 结束语

本文提出了一种基于深度学习的X射线鞋底异物检测系统,在采集鞋底断针X光图像数据的同时使用深度学习算法迅速有效的筛选出含有异物金属的服饰,并将这些异物定位.通过实验表明,采用图像预处理和调整标注框之后有效提高了模型的检测精度.经过验证,此检测系统的平均精度能够达到97%以上,可

以有效检测遗留在不同类型鞋子中的订书针、细铁屑、断针、鞋钉、钢珠等铁质金属异物,具有非常重要的商业价值.



图13 不同鞋底异物检测结果(包含钢珠、断针、铁钉等)

参考文献

- 顾虎, 顾伟, 胡惠强. 服装生产过程中金属针的控制与检测. 国外丝绸, 2009, 24(4): 32-33.
- 黄勇. 金属探测器的研究与设计 [硕士学位论文]. 广州: 华南理工大学, 2010.
- 郑金州, 鲁绍栋. X射线技术在安检领域的应用. CT理论与应用研究, 2012, 21(2): 357-364.
- 龙钧宇, 余爱民, 余红. 一种基于模糊MSR的医学X光图像增强算法. 计算机技术与发展, 2014, 24(1): 235-237, 241.
- 陈思, 林锦纯. 一种服装断针检测系统的设计. 价值工程, 2018, 37(5): 185-186. [doi: 10.14018/j.cnki.cn13-1085/n.2018.05.074]
- 王宇, 邹文辉, 杨晓敏, 等. 基于计算机视觉的X射线图像异物分类研究. 液晶与显示, 2017, 32(4): 287-293.
- 姚韵. 基于深度学习的图像识别技术与应用研究 [硕士学位论文]. 西安: 西安工程大学, 2017.
- 英红, 丁海明, 侯新月, 等. 基于BP神经网络的沥青路面裂缝识别方法研究. 河南理工大学学报(自然科学版), 2018, 37(4): 105-111. [doi: 10.16186/j.cnki.1673-9787.2018.04.16]

- 9 于恒, 梅红岩, 许晓明, 等. 基于深度学习的图像压缩算法研究综述. 计算机工程与应用, 2020, 56(15): 15–23. [doi: [10.3778/j.issn.1002-8331.2003-0294](https://doi.org/10.3778/j.issn.1002-8331.2003-0294)]
- 10 姚群力, 胡显, 雷宏. 深度卷积神经网络在目标检测中的研究进展. 计算机工程与应用, 2018, 54(17): 1–9. [doi: [10.3778/j.issn.1002-8331.1806-0377](https://doi.org/10.3778/j.issn.1002-8331.1806-0377)]
- 11 Girshick R. Fast R-CNN. Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile. 2015. 1440–1448.
- 12 Liu W, Anguelov D, Erhan D, *et al.* SSD: Single shot MultiBox detector. Proceedings of the 14th European Conference on Computer Vision. Amsterdam, the Netherlands. 2016. 1–17.
- 13 Redmon J, Divvala S, Girshick R, *et al.* You only look once: Unified, real-time object detection. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 779–788.
- 14 Ren SQ, He KM, Girshick R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137–1149. [doi: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031)]
- 15 Uijlings JRR, van de Sande KE, Gevers T, *et al.* Selective search for object recognition. International Journal of Computer Vision, 2013, 104(2): 154–171. [doi: [10.1007/s11263-013-0620-5](https://doi.org/10.1007/s11263-013-0620-5)]
- 16 姜翼. 地铁屏蔽门激光扫描异物检测方法研究. 科技创新导报, 2019, 16(27): 65–66. [doi: [10.16660/j.cnki.1674-098X.2019.27.065](https://doi.org/10.16660/j.cnki.1674-098X.2019.27.065)]
- 17 周雯, 史天运, 李平, 等. 基于深度学习的动车组运行安全图像异物检测. 交通信息与安全, 2019, 37(6): 48–55.
- 18 洪潮, 安桂香, 潘孟良. 胶条 x 光机异物自动检测剔除设备. 橡塑技术与装备, 2020, 46(3): 35–40. [doi: [10.13520/j.cnki.rpte.2020.03.007](https://doi.org/10.13520/j.cnki.rpte.2020.03.007)]
- 19 王建鹏. 一种基于异构串联卷积神经网络的铁路异物检测模型. 测试技术学报, 2020, 34(4): 344–348. [doi: [10.3969/j.issn.1671-7449.2020.04.012](https://doi.org/10.3969/j.issn.1671-7449.2020.04.012)]
- 20 孙雄峰, 林浒, 王诗宇, 等. 基于改进 Faster RCNN 的工业机器人分拣系统. 计算机系统应用, 2019, 28(9): 258–263. [doi: [10.15888/j.cnki.csa.007074](https://doi.org/10.15888/j.cnki.csa.007074)]