

基于 YOLOv3 的医药玻璃瓶缺陷检测方法

陈宏彩^{1,2}, 任亚恒^{1,2}, 郝存明^{1,2}, 程煜^{1,2}, 张效玮³

(1.河北省科学院应用数学研究所, 石家庄 050081; 2.河北省信息安全认证工程技术研究中心, 石家庄 050081; 3.石家庄市发展和改革委员会, 石家庄 050011)

摘要: **目的** 为了准确而快速地自动检测医药玻璃瓶的外观缺陷。**方法** 基于 YOLOv3 算法, 建立深度卷积神经网络学习检测模型, 通过将神经网络结构浅层和深层特征向量连接并标准化, 提取玻璃瓶多尺度特征信息。采用 K-means 聚类方式获得锚点框初始大小, 提高模型对边界框预测的准确性; 利用多尺度训练策略, 增强模型对不同尺寸图像的鲁棒性。**结果** 实验结果表明, 提出的医药玻璃瓶缺陷检测方法能够准确检测识别玻璃瓶上的管端残损、气线、气泡、划伤、污渍和结石等缺陷种类。与主流的目标检测方法相比, 提出的方法在处理速度和准确度上都有提高, 缺陷目标检测精确率达到 96.23%, 召回率为 93.82%, 平均精度为 89.35%。**结论** 该方法已经在国内几家大型医药玻璃包装生产公司成功应用, 显著提高了医药玻璃包装产品的质量和合格率, 降低了人工成本。

关键词: 医药玻璃瓶; 缺陷检测; YOLOv3; 卷积神经网络; 多尺度特征

中图分类号: TS206 文献标识码: A 文章编号: 1001-3563(2020)07-0241-06

DOI: 10.19554/j.cnki.1001-3563.2020.07.034

Defect Detection Method for Medical Glass Bottles Based on YOLOv3

CHEN Hong-cai^{1,2}, REN Ya-heng^{1,2}, HAO Cun-ming^{1,2}, CHENG Yu^{1,2}, ZHANG Xiao-wei³

(1.Hebei Academy of Sciences Institute of Applied Mathematics, Shijiazhuang 050081, China; 2.Hebei Authentication Technology Engineering Research Center, Shijiazhuang 050081, China; 3.Shijiazhuang Municipal Development and Reform Commission, Shijiazhuang 050011, China)

ABSTRACT: The work aims to automatically detect the defects on the medicinal glass bottles accurately and quickly. Based on the YOLOv3 algorithm, a learning detection model of deep convolutional neural network was built. The multi-scale feature information of glass bottles was extracted by connecting and normalizing the shallow and deep feature vectors of the neural network structure. To improve the accuracy of the model for the prediction of the bounding box, K-means clustering method was used to obtain the initial size of the anchor box. The multi-scale training strategy was used to enhance the robustness of the model to images of different sizes. The experimental results showed that, the proposed defect detection method of medical glass bottles could accurately detect and identify such defects as damaged tube end, gas lines, bubbles, scratches, stains and calculi on glass bottles. Compared with the mainstream target detection methods, the processing speed and accuracy of the proposed method were improved. The accuracy of defect target detection reached 96.23%, the recall rate was 93.82%, and the average accuracy was 89.35%. The proposed method has been successfully applied to several large medical glass packaging production companies in China, which has significantly improved the quality and qualified rate of medical glass packaging products and reduced the labor costs.

收稿日期: 2019-09-04

基金项目: 河北省科技计划 (2019034288)

作者简介: 陈宏彩 (1981—), 女, 副研究员, 主要研究方向为图像处理、机器视觉。

KEY WORDS: medical glass bottles; defect detection; YOLOv3; convolutional neural network; multi-scale features

药用玻璃瓶是主要的药品包装容器,其质量关系到国民生命健康。为了保障医药产品质量,目前我国药品包装企业普遍采用质检员抽检的方式对药用玻璃瓶进行质量检测。人工抽检存在漏检、误检、效率低下等诸多弊端。随着国家对药包标准的日益规范和严格管控,企业因产品质量问题付出的代价不断加剧,且人工成本逐年提高,企业迫切需要实现从人工抽检到自动化、智能化检测的转型升级。发达国家已经利用计算机技术实现药用玻璃瓶在线自动检测。由于直接引进国外的检测系统不仅价格昂贵,而且面临检测设备与企业现有生产线的调整和衔接困难等一系列问题,因此亟需研制稳定可靠、适合我国药品包装生产线的药用玻璃瓶检测系统。

近年来,随着人工智能的迅猛发展,基于深度学习的图像目标检测、识别技术有了很大的突破,使用卷积神经网络(Convolutional Neural Networks, CNNs)方法提取的特征要远远优于传统人工设计的特征。基于CNN的目标检测与识别方法已经被广泛应用到农业果蔬检测^[1-2]、工业零部件检测^[3-5]、交通领域^[6-7]、医学图像检测与分类^[8-9]等多个领域。

为了克服传统非智能化的医药玻璃瓶缺陷检测方法,提高医药包装产品质量,文中结合深度学习的研究成果,提出基于卷积神经网络技术的医药玻璃瓶缺陷检测方法。

1 相关工作

当前,基于深度CNN的目标检测算法可以分为2类,一类是基于候选区域(Region Proposal)的算法,代表性的算法有R-CNN^[10],该算法检测效果好,但是计算量较大。随后,人们对其进行改进,先后提出了Fast R-CNN^[11],Faster R-CNN等^[12]方法。尽管Faster R-CNN算法处理速度较之前方法有了很大提升,但仍无法满足实时性的应用需求。另一类是基于回归的目标检测算法,代表性的方法有YOLOv1算法^[13],该方法较基于候选区域的方法极大提高了检测速度。YOLOv1算法仅使用 7×7 的网格回归,这不仅导致其预测目标数量有限,而且目标定位也不精准。2016年,Liu等^[14]提出了一种基于回归策略的SSD目标检测模型,SSD方法使用图像各个位置的多尺度区域进行回归,根据锚点直接对目标进行分类和回归边框,既保持了YOLOv1方法速度快的特性,也在一定程度上提高了窗口预测精准度。Redmon等^[15]针对YOLOv1在召回和定位方面存在的问题提出了一个新的网络框架Darknet-

19,即YOLOv2算法,YOLOv2在处理速度和预测精度上较YOLOv1均有所提高。2018年,Redmon等^[16]借鉴残差神经网络的思想,再次对YOLOv2进行改进,提出了YOLOv3算法,YOLOv3是当前目标检测算法中集检测速度和准确率于一身的优秀算法之一。

针对药用玻璃瓶外观缺陷检测问题,文中基于YOLOv3算法,利用深度卷积神经网络技术设计并搭建一个用于医药玻璃瓶外观缺陷检测的模型。该模型权衡缺陷检测精度和处理速度,有效解决药用玻璃瓶缺陷目标大小差距较大、检测尺度单一等问题。

2 基于YOLOv3的药用玻璃瓶缺陷检测方法

2.1 药用玻璃瓶缺陷检测模型设计

为了实现实时高效的药用玻璃瓶外观缺陷检测,文中基于YOLOv3框架,设计并搭建了一个基于卷积神经网络的玻璃瓶外观缺陷检测模型,称为YOLOv3-bottle,网络结构主要包括卷积层、批归一化处理层和跃层连接模块,激活函数采用Leaky ReLU函数。基于YOLOv3的玻璃瓶缺陷检测模型的整体框架结构,见图1。

图1中张量拼接表示连接2个输入层,S2表示卷积操作滑动的步幅是2,残差块是具有残差结构的重复卷积层,数字1,2,4,8等表示重复的残差结构数目。一个残差组件结构见图2。

2.2 多尺度特征向量的联合预测

为了解决玻璃瓶缺陷目标大小差距较大、检测尺度单一等问题,文中设计的基于YOLOv3的玻璃瓶缺陷检测模型采用类似FPN网络的尺度金字塔结构^[16],其基本思想是将输入图像划分为 $S \times S$ 个大小相同的单元格,每个单元格分配3个锚点框,负责预测单元格中心的3个边界框,见图3。为了提升对玻璃瓶细小缺陷目标的检测效果,文中在3个不同尺度大小特征图上进行缺陷目标任务检测。

玻璃瓶缺陷检测模型采用逻辑回归的方法为每个边界框预测。在每一个单元格上,CNN会对每一个边界框预测得到4个结果,边界框中心坐标 (x, y) 以及检测目标的宽 w 和高 h ,分别记为 t_x, t_y, t_w, t_h (c_x, c_y)是目标中心在单元格中相对于图像左上角的位置, p_w, p_h 是预设的锚点框的宽度和高度,则最终得到的边界框的中心坐标和宽高分别为:

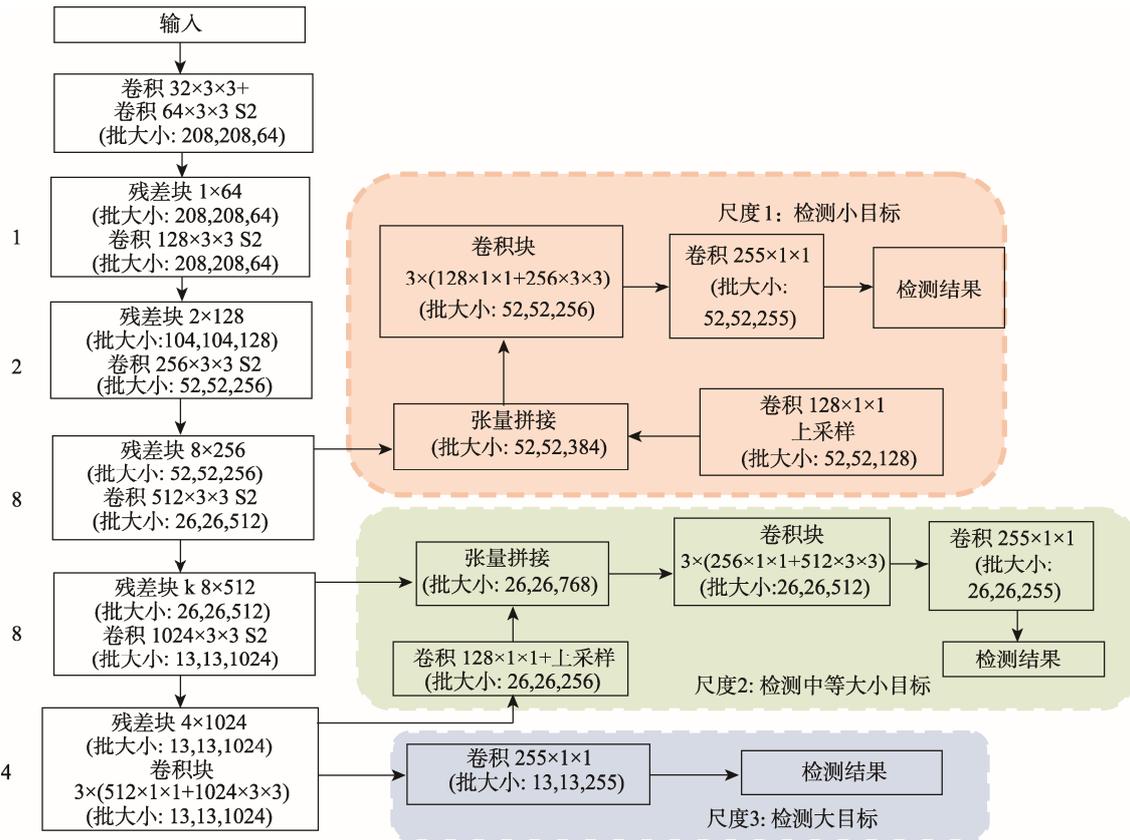


图 1 基于 YOLOv3 的医药玻璃瓶缺陷检测模型
Fig.1 Defect detection model of medical glass bottles based on YOLOv3

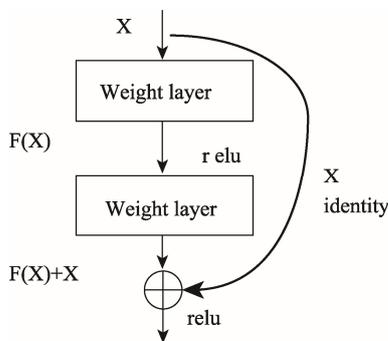


图 2 一个残差组件
Fig.2 A residual component

$$b_x = \sigma(t_x) + c_x$$

$$b_y = \sigma(t_y) + c_y$$

$$b_w = p_w e^{t_w}$$

$$b_h = p_h e^{t_h}$$

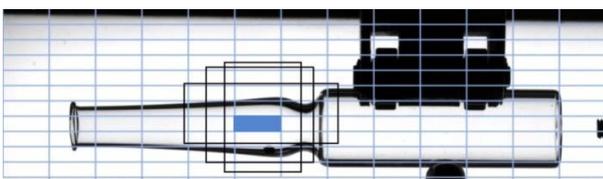


图 3 模型在 13×13 的单元格中的预测边界框示意
Fig.3 Illustration of predicted bounding boxes of the model on 13×13

YOLOv3-bottle 网络模型在进行类别预测时，采用 Logistic 函数作为分类器。然后采用平方和损失函数计算预测框定位误差、IOU 误差以及分类误差作为最终损失。梯度可以通过最小化损失函数求得。

为了实现多尺度联合预测，将 YOLOv3-bottle 模型网络结构中浅层特征与深层特征进行连接，具体操作步骤如下所述。

- 1) 输入图像经过模型的前 7 个卷积层获取第 1 个特征图谱，进行第 1 次预测。
- 2) 获取网络结构倒数第 3 个卷积层的输出，并对结果进行卷积和上采样操作，再将获取的特征与网络结构的第 43 个卷积特征拼接，得到第 2 个特征图谱，进行第 2 次预测。
- 3) 获取网络结构倒数第 3 个卷积层的输出，并对结果做卷积和上采样操作，然后将获取的特征与网络结构的第 26 个卷积特征向量拼接，得到第 3 个特征图谱，并进行第 3 次预测。

每个预测任务得到的特征大小都为 $N \times N \times [3 \times (4+1+class_Num)]$ ，N 为单元格大小，3 为每个单元格得到的边界框数量，4 是边界框坐标数量，1 是目标预测值，class_Num 是预测目标类别数量。

根据对象边界框与锚点框的交并比 (IoU) 来预测目标类别及位置信息，如果最大 IoU 大于预先设置

阈值,则对象与最大IoU的锚点框匹配,否则匹配失败。如果锚点框与某个对象的IoU大于阈值,但没有与对象匹配,则忽略该锚点框的损失。

3 实验与结果

3.1 测试环境与数据

文中在 AMAX 服务器上进行实验测试。实验软件环境 Ubuntu18.04 和 CUDA9.0,实验的硬件 CPU 为 Intel(R)E5-2650 @2.60GHz, GPU 为 Tesla K20c。

药用玻璃瓶数据集是从成都某药品包装公司采集的药用安瓿瓶图像。根据国家食品药品监督管理局《直接接触药品的包装材料和容器标准》的质量管理要求以及低硼硅玻璃药用安瓿用玻管的外观质量检测标准,文中将药用安瓿瓶的外观质量缺陷划分为6类:管端残损、气线、裂纹、划伤、污渍和结石,并进行实验测试。每类缺陷目标训练样本数约为400张至1000张不等,总共5520张训练图像,测试数据集共5500张图像,每张图像的分辨率大小为1596×640。

3.2 药用玻璃瓶缺陷检测训练策略

在药用玻璃瓶缺陷检测的训练阶段,经过反复实验调整,文中采用带动量因子的小批量随机梯度下降法进行模型训练。随机梯度下降法的性能与学习率有直接关系,学习率的大小决定了参数移动到最优值时的速度。在训练过程开始时,初始化的学习率设置较大些,这样能快速收敛,文中初始学习率设置为 $\eta_{lr}=0.001$,设置动量参数为0.9。随着训练过程的进行,逐渐降低学习率,这样有助于找到最优解。为了保持网络模型稳定,随后将学习率衰减方式设置为 $\eta_{learning_rate}=\eta_{lr}\times(N_{batch}/1000)^2$,其中 N_{batch} 为批处理次数,经过1000批次训练后,将学习率设置为0.001。另外,该实验中为了扩充训练样本数量,对玻璃瓶图像样本进行了各种变换,包括样本图片随机旋转 -5° 到 5° 、水平翻转、平移变换,调整饱和度(1/1.5到1.5倍)、曝光度(1/1.5到1.5倍)、色调(变化范围为 $-0.1\sim 0.1$)等。

为了增强模型的鲁棒性,该实验采用多尺度方法进行训练。药用玻璃瓶样本图像尺寸变化范围从 320×320 到 608×608 ,每进行10批次训练,图像尺寸变化1次,图像尺寸采样间隔为32。

为了提高YOLOv3-bottle网络模型对边界框预测的准确率,文中对药用玻璃瓶训练集中的所有样本标注使用K-means聚类的方式获得锚点框的初始大小。通过聚类获取的9个锚点框分别为(16,22),(21,34),(29,41),(31,48),(36,56),(54,88),(76,98),(67,136),(106,187)。

3.3 实验和结果

运用文中设计的卷积神经网络模型YOLOv3-bottle,对采集的药用玻璃瓶样本图像进行了相关训练,训练结束后,测试了模型的准确性及相关性能。

训练数据集随着训练迭代次数的增加误差损失情况见图4。从图4可以看出,文中设计的检测模型随着迭代次数增加,误差损失经过震荡呈下降趋势。当训练迭代批次达到8万次左右,平均损失维持在0.1左右时,终止训练,保存权重文件。

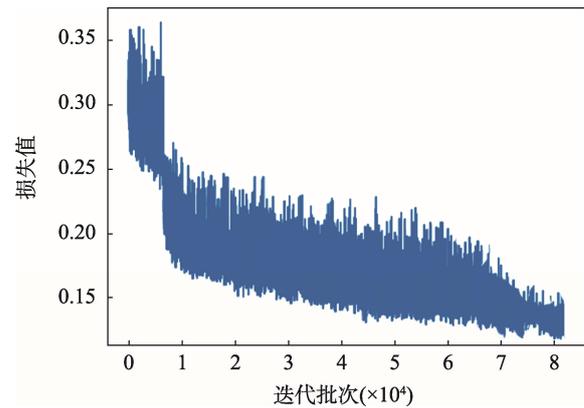


图4 训练迭代损失曲线
Fig.4 Curve of training iterative loss

实验中,使用精确率 P 、召回率 R 、 F_1 分数、IOU和平均精度评价医药玻璃瓶检测模型的性能。

$$\text{精确率: } P = \frac{T_p}{F_p + T_p}$$

$$\text{召回率: } R = \frac{T_p}{F_p + F_N}$$

$$F_1 \text{ 分数: } F_1 = \frac{2P \cdot R}{P + R}$$

式中: T_p 为真正例; F_p 为假正例; F_N 为假反例。采用VOC2007中的标准^[17],求最大精度的平均值便得到平均精度 A_p :

$$A_p = \frac{1}{11} \sum_{R \in (0,0.1 \dots 1)} P_{\text{interp}}(R)$$

最大精度 $P_{\text{interp}}(R)$ 的计算方法为:

$$P_{\text{interp}}(R) = \max_{\tilde{R}, \tilde{R} \geq R} p(\tilde{R})$$

式中: \tilde{R} 为取得最大值精度时的召回率。

一般情况下,处理图像或视频的速度达到每秒30帧(这里速度单位是帧/s),就能够实现实时性,然而,在药用安瓿瓶生产线中,药瓶的最快传输速度达到70帧/s。在文中实验中,为了评估模型的处理速度,测试了单位时间内算法处理的图像数量。文中模型YOLOv3-bottle性能的测试结果见表1,并且与原始目标检测模型YOLOv3^[16],YOLOv3-tiny,Faster R-CNN^[12]和SSD^[14]方法测试结果做了比较,测试指

表 1 药用玻璃瓶缺陷检测模型性能测试结果
Tab.1 Test results of defect detection model performance of medical glass bottles

方法	$P/\%$	$R/\%$	$F_1/\%$	$A_p/\%$	$R_{IOU}/\%$	速度/(帧·s ⁻¹)
Faster R-CNN	92.28	85.34	88.67	80.08	69.43	13
SSD	93.56	80.69	86.65	85.24	67.65	22
YOLOv3	93.47	78.62	85.40	78.94	68.74	50.4
YOLOv3-tiny	90.21	82.65	86.26	78.86	66.87	212
YOLOv3-bottle	96.23	93.82	95.00	89.35	71.45	75

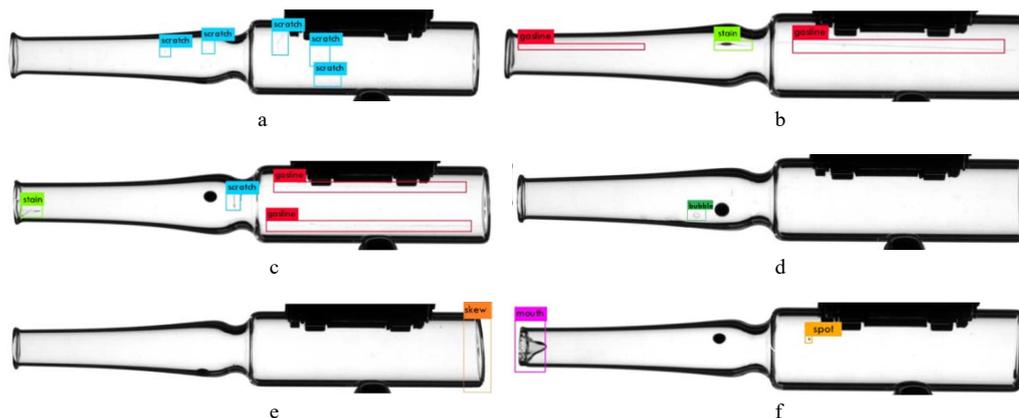


图 5 药用玻璃瓶缺陷检测与识别效果

Fig.5 Defect detection and recognition effect of medical glass bottles

标包括精确率 P 、召回率 R 、 F_1 值、平均精度 A_p 以及测试时间。文中设计的药用玻璃瓶检测方法在 5500 张测试集中得到检测精确率为 96.23%、召回率为 93.82%、平均精度为 89.35%。模型的测试速度为 75 帧/s，达到了药用安瓿瓶生产线的出瓶速度。

药用玻璃瓶缺陷检测的实验结果见图 5。从图 5 中可以看出，文中设计的药用玻璃瓶缺陷检测方法不仅能够检测出药用玻璃瓶的各种外观质量缺陷，如图 5 中矩形框框出的划痕、污渍、气线、结石、气泡、管端嘴部残损、管底歪斜等缺陷，而且能够对不同的缺陷进行精准分类。这为药包生产企业改进设备工艺，增强产品质量，提高企业的市场竞争力提供有力的指导。

4 结语

文中详细阐述了基于 YOLOv3 模型的医药玻璃瓶外观缺陷检测方法，主要包括卷积神经网络模型框架的搭建，多尺度特征向量的联合预测及具体的实施步骤。通过在网络模型中引入多尺度特征向量，采用类似 FPN 网络的尺度金字塔结构，解决了药用玻璃瓶缺陷目标大小差距较大、检测尺度单一等问题。该方法已成功应用在河北、四川、重庆等多家医药包装公司的玻璃瓶生产检测流水线上，显著提高了药用玻璃包装产品的质量和合格率、降低了人工成本，提升了药包企业产品的安全性和市场竞争力。

参考文献：

- [1] 赵德安, 吴任迪, 刘晓洋, 等. 基于 YOLO 深度卷积神经网络的复杂背景下机器人采摘苹果定位[J]. 农业工程学报, 2019, 35(3): 164—173.
ZHAO De-an, WU Ren-di, LIU Xiao-yang, et al. Apple Positioning Based on YOLO Deep Convolutional Neural Network for Picking Robot in Complex Background[J]. Transactions of the Chinese Society of Agricultural Engineering, 2019, 35(3): 164—173.
- [2] ALVARO F, SOOK Y, SANG K, et al. A Robust Deep-learning-based Detector for Real-time Tomato Plant Diseases and Pests Recognition[J]. Sensors, 2017, 17(9): 2022.
- [3] 乔丽, 赵尔敦, 刘俊杰, 等. 基于 CNN 的工件缺陷检测方法研究[J]. 计算机科学, 2017(S2): 248—253.
QIAO Li, ZHAO Er-dun, LIU Jun-jie, et al. Workpiece Defect Detection Method Based on CNN[J]. Computer Science, 2017(S2): 248—253.
- [4] 赵耀霞, 吴桐, 韩焱. 基于卷积神经网络的复杂构件内部零件装配正确性识别[J]. 电子学报, 2018(8): 1983—1988.
ZHAO Yao-xia, WU Tong, HAN Yan. Identifying the Correctness of Fit of Internal Components Based on a Convolutional Neural Network[J]. Acta Electronica Sinica, 2018(8): 1983—1988.
- [5] 张辉, 宋雅男, 刘理, 等. 一种基于深度学习网络的钢轨表面缺陷检测方法: 中国, 109064461A[P].

- 2018-08-06.
ZHANG Hui, SONG Ya-nan, LIU Li, et al. A Method of Rail Surface Defect Detection Based on Deep Learning Network: China, 109064461A[P]. 2018-08-06.
- [6] LUO H L, YI Y, BEI T, et al. Traffic Sign Recognition Using a Multi-task Convolutional Neural Network[J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 19(4): 1100—1111.
- [7] JIE F, YU Z, YAO Y, et al. Fine-grained Vehicle Model Recognition Using a Coarse-to-fine Convolutional Neural Network Architecture[J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 18(7): 1782—1792.
- [8] 王媛媛, 周涛, 陆惠玲, 等. 基于集成卷积神经网络的肺部肿瘤计算机辅助诊断模型[J]. 生物医学工程杂志, 2017(4): 65—73.
WANG Yuan-yuan, ZHOU Tao, LU Hui-ling, et al. Computer Aided Diagnosis Model of Lung Tumor Based on Ensemble Convolutional Neural Network[J]. Journal of Biomedical Engineering, 2017(4): 65—73.
- [9] LIU X L, HOU F, QIN H, et al. Multi-view Multi-scale CNNs for Lung Nodule Type Classification from CT Images[J]. Pattern Recognition, 2018, 77: 262—275.
- [10] GIRSHICK R, DONAHUE J, DARRELL T, et al. Region-based Convolutional Networks for Accurate Object Detection and Segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 38(1): 142—158.
- [11] GIRSHICK R. Fast r-cnn[C]// Proceedings of the IEEE International Conference on Computer Vision, Santiago, 2015: 1440—1448.
- [12] 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, 2015, 39(6): 1137—1149.
- [13] REDMON J, DIVVALA S, GIRSHICK R, et al. You only Look Once: Unified, Real-Time Object Detection[C]// Proceedings of 2016 IEEE Computer Vision and Pattern Recognition, Las Vegas, 2016: 779—788.
- [14] LIU W, ANGELOV D, ERHAN D, et al. SSD: Single Shot Multibox Detector[C]// European Conference on Computer Vision, Amsterdam, 2016: 21—37.
- [15] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger[C]// IEEE Conference on Computer Vision and Pattern Recognition. Puerto Rico, 2017: 6517—6525.
- [16] REDMON J, FARHADI A. YOLOv3: An Incremental Improvement[R/OL]. (2018)[2019-09-04]. <http://arxiv.org/pdf/1804.02767>.
- [17] 戴伟聪, 金龙旭, 李国宁, 等. 遥感图像中飞机的改进 YOLOv3 实时检测算法[J]. 光电工程, 2018, 45(12): 84—92.
DAI Wei-cong, JIN Long-xu, LI Guo-ning, et al. Real-time Airplane Detection Algorithm in Remote-Sensing Images Based on Improved YOLOv3[J]. Opto-Electronic Engineering, 2018, 45(12): 84—92.