

# 基于 Mask R-CNN 的触摸屏玻璃疵病检测与识别

张 博<sup>1,2</sup>, 周 军<sup>1</sup>, 王 芳<sup>1,2</sup>, 韩 森<sup>1,2</sup>

(1. 上海理工大学 光电信息与计算机工程学院, 上海 200093; 2. 苏州慧利仪器有限责任公司, 江苏 苏州 215123)

**摘 要:**为快速有效地检测与识别触摸屏玻璃疵病,提出了一种基于 Mask R-CNN 技术的检测与识别方法。Mask R-CNN 是 Faster R-CNN 技术的扩展,其在边界框识别的现有分支上添加一个并行的用于预测对象掩膜的分支,是基于 R-CNN 系列、FPN、FCIS 等工作之上的一种技术,相对于传统的目标检测和分割算法有很大提升。使用暗场散射成像、多图片堆栈等方法采集数据样本,同时采用几何变换方法扩增样本库。为提升训练速率和检测精确率,根据训练样本差异性,通过测试不同的 VGG16、ResNet50+FPN 以及 ResNet101+FPN 主干网络提取图像特征。实验表明,ResNet50+FPN 的识别精确率较高,达到 96.7%。基于 Mask R-CNN 的检测与识别方法不仅能对疵病快速检测与识别,还能提高识别精确率。

**关键词:**Mask R-CNN; 疵病检测; 暗场散射; 多图片堆栈; 触摸屏玻璃

**DOI:**10.11907/rjdk.181934

**中图分类号:**TP319

**文献标识码:**A

**文章编号:**1672-7800(2019)002-0064-04

## Detection and Identification of Glass Defects on the Touch Screen Based on Mask R-CNN

ZHANG Bo<sup>1,2</sup>, ZHOU Jun<sup>1</sup>, WANG Fang<sup>1,2</sup>, HAN Sen<sup>1,2</sup>

(1. School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China; 2. Suzhou H&L Instruments LLC, Suzhou 215123, China)

**Abstract:** In order to detect and identify glass defects on the touch screen quickly and effectively, a screen glass defect detection and identification method based on Mask R-CNN technology is proposed. Mask R-CNN is Faster R-CNN technical extension, on the existing branches of applied to identify boundary box a parallel branch is added to predict object mask based on R-CNN series, FPN, FCIS. Compared with traditional target detection and segmentation algorithm, the algorithm has been greatly improved. This paper uses dark field scattering imaging, multiple image stacks and other methods to collect data samples, and uses geometric transformation methods to expand the sample library. In order to improve the training rate and detection accuracy, according to the differences of the training sample, this paper tests different backbone networks of VGG16, ResNet50+FPN, ResNet101+FPN to extract the image features, and finds that the recognition accuracy is higher and the accurate rate reaches 96.7%. As result, the detection and recognition method based on Mask R-CNN can not only detect and identify defects quickly and effectively, but also improve the recognition accuracy rate.

**Key Words:** Mask R-CNN; disease identification; dark field scattering; multiple image stack; touch screen

## 0 引言

智能手机、平板电脑的普及,触摸屏玻璃产量大增。触摸屏玻璃质量要求很高,人工检测玻璃表面缺陷(如斑点、异物、裂纹和划痕)很困难。目前世界各国表面疵病检

测标准依据主要有两个:疵病面积等级及疵病可见度等级,所以针对疵病检测的方法也主要基于疵病散射光的特性对疵病面积以及疵病可见度进行测量。基于可见度的检测方法就是在一定的光照条件下通过人眼对玻璃样品进行观察,该方法简单但效率低,而且随着工作时长的增加,大脑和眼睛的负担随之加重,导致误检率增高。对疵

收稿日期:2018-06-12

基金项目:国家重大科学仪器设备开发专项项目(2016YFF0101903)

作者简介:张博(1992-),男,上海理工大学光电信息与计算机工程学院硕士研究生,研究方向为光学检测技术;韩森(1961-),男,博士,上海理工大学光电信息与计算机工程学院教授、博士生导师,研究方向为光学干涉计量检测。本文通信作者:韩森。

万方数据

病面积的检测主要通过成像法或能量法检测,成像法包括目视法以及滤波成像法等,能量法则利用角度积分探测器获取待测元件散射光能量,对应求出疵病面积。其中成像法系统结构简单,但检测精度低,易受杂散光影响。能量法对检测区域的疵病面积等级检测较精确,但无法获取表面疵病的数量、位置信息与外形特征,孙佳明<sup>[1]</sup>的论文对此进行了介绍。目前传统的识别方法是通过图像灰度分布情况,采用图像增强技术如滤波、边缘检测、腐蚀膨胀、噪声抑制等方法实现玻璃缺陷的检测与识别<sup>[2]</sup>,这些传统方法主要问题有:特征提取的鲁棒性较差,不能反映光照的变化、背景多样化等;区域选择没有针对性,时间复杂度高、窗口冗余。目标检测是计算机视觉和模式识别的重要研究方向,深度学习作为机器学习的新领域,其特点是能自动学习到有效特征。本文通过使用暗场散射成像、多图片堆栈等方法采集数据样本,结合Mask R-CNN技术对疵病进行检测与识别。该方法大大减少了复杂的人工提取特征过程,对多样性疵病种类特征实现自动提取,大大提高了疵病检测精度。

Mask R-CNN<sup>[3]</sup>是一个小巧、灵活的通用对象实例分割框架,它不仅可对图像中的目标进行检测,还能对每个目标给出一个高质量的分割结果。本文主要针对斑点、划痕、崩边、污点等疵病进行识别检测,暂不考虑其它疵病类型。

## 1 网络结构模型

### 1.1 Faster R-CNN

Faster R-CNN<sup>[4]</sup>是Fast R-CNN<sup>[5]</sup>的改进,针对Fast R-CNN的一些不足,如使用Selective Search提取RP(region proposal)(耗时2~3s,而特征提取仅仅耗时0.32s)无法满足实时应用,无法实现端到端训练和测试等。虽然利用了GPU,但是RP方法是在CPU上完成的。Faster R-CNN可简单看作“区域生成网络(RPN)+Fast R-CNN”系统,其网络结构如图1所示。Faster R-CNN使用VGG16Net<sup>[6]</sup>特征提取,相对于LeNet<sup>[7]</sup>和AlexNet<sup>[8]</sup>有着深度更深、特征提取效果更好等特点。

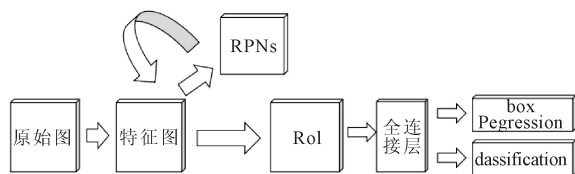


图1 Faster R-CNN 网络结构

区域建议网络(RPN)是Faster R-CNN的一大亮点,它是一个全卷积网络,特点是可利用GPU生成高质量高速率的建议框,它和检测的网络共享全图卷积特征,最大的特点是用区域生成网络代替Fast R-CNN中的Selective Search方法,大大提升了检测速率。区域生成网络如图2所示。

万方数据

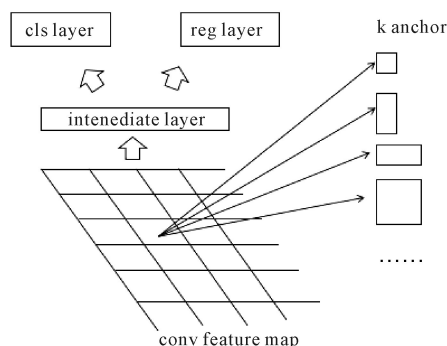


图2 区域建议网络(RPN)

在区域建议网络(RPN)中设置了3×3的滑窗,其中心点位置对应预测输入图像的3种尺度,即3种长宽比的RP(1:1,1:2,2:1),这种映射机制称为anchor,产生k个anchor(一般我们取k=9),每个3×3区域可产生9个RP,后面连接到两个全连接层,即cls layer和reg layer,分别用来分类和边框回归。cls layer包含两个元素,用来判别目标和非目标,而reg layer包含4个坐标元素(x,y,w,h),用来找到目标位置,最后根据RP的得分高低选取前300个RP作为Fast R-CNN输入目标进行检测。

### 1.2 Mask R-CNN

Mask R-CNN与Faster R-CNN很相似,其网络结构如图3所示。

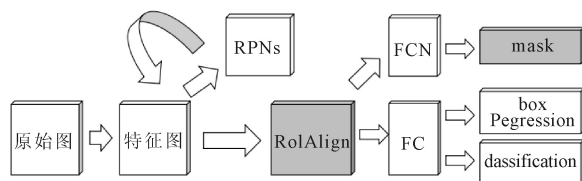


图3 Mask R-CNN 网络结构

Mask R-CNN在Faster R-CNN基础上的主要改进如下:

(1)添加Mask层。Faster R-CNN中每个候选对象有两个输出:Box Regression和Classification。Kaiming He等在每个候选对象上添加了第3个分支,即输出Object Mask,从而得到Mask R-CNN。值得注意的是在Mask R-CNN中,Faster R-CNN中的CNN特征顶部添加一个完全卷积网络(FCN)用来生成掩码(分割输出)。

(2)提出RoIAlign层,重新对齐RoI Pool。RoI Pool是针对每个RoI提取一个小尺度特征图(比如7×7)的标准操作,用来解决不同尺度的RoI提取成相同尺度的特征大小问题。RoI Pool首先将浮点数值RoI量化成离散颗粒的特征图,然后将量化的RoI分成几个空间小块,对每个小块进行Max Pooling操作生成最后结果。例如当特征图的步长为16时,通过计算 $\lceil x/16 \rceil$ 在连续坐标x上进行量化, $\lceil \cdot \rceil$ 表示四舍五入。这些量化引入了RoI与提取到的特征不对准问题。由于分类问题对平移问题比较鲁棒,所以影响较小,但是这在预测像素级精度的掩模时会产生非常大的负面影响。提出RoIAlign层解决这个问题,将提取到的特征

与输入对齐,避免对 RoI 的边界或块作任何量化,例如直接使用  $x/16$  代替  $[x/16]$ 。使用双线性插值在每个 RoI 块中 4 个采样位置上计算输入特征的精确值,并将结果聚合(使用 max 或者 average)。

(3)重新定义 ROI 损失函数。依旧采用多任务损失函数,针对每个 RoI 定义为:

$$L = L_{cls} + L_{box} + L_{mask} \quad (1)$$

## 2 疵病训练样本库建立

### 2.1 疵病样本图像采集

本文采用暗视场散射<sup>[9-10]</sup>方法采集图像。反射集光镜反射的光线投射在被测玻璃上,因其倾斜角度极大,由玻璃上反射的光线仍以极大的倾斜角度向反方向反射,不可能进入物镜,所以在目镜筒内只能看到漆黑一片。只有被测面凹洼之处才能有光射进物镜,玻璃面上的凹洼之处以亮白影像衬映在漆黑的视域内,称为“暗视场”(暗场)照明,如图 4(a)所示。

暗场散射优点:①入射光束倾斜角度极大,使物镜的有效数值孔径随之增大,提高了物镜的鉴别能力;②入射光不进入物镜,显著减少了由于光线多次通过玻璃-空气界面形成的反射光与散射光,从而提高了最终影像的衬度;③消除了试样抛光表面强烈反射光的干扰。

由于不同方向、不同角度的平行光打在待测面上会检测出不同的疵病信息,因此在采集装置上作了一些改动,即用不同方向、不同角度的平行光打在待测面上,得到该待测面的不同疵病信息暗场图片。因为相机的位置是固定不变的,所以为检测到更完整的疵病图像<sup>[11-13]</sup>,可将不同方向不同角度拍摄的图片堆栈到一幅图上,这样堆栈图就包含了所有方向和角度的疵病<sup>[14-21]</sup>信息,经过灰度化处理后再进行检测和识别,改进如图 4(b)所示。

### 2.2 疵病样本数据准备

本文人工标注了常见的 4 种玻璃缺陷类型图像  $800 \times 600$  共 300 张。深度学习训练模型识别准确率依赖大量样本,但没有公共数据集,人工采集和标注需要大量的时间成本。为弥补训练样本数量的不足,本文通过对原有训练样本进行几何变换扩充样本,例如旋转 90 度、镜像、平移等,最终扩充为 1 200 张,其中训练集为 1 000 张,200 张为测试集。

标注图像如图 5 所示。使用标注工具生成的 label 会保存为 json 文件,所以还需将该文件解析出 mask 并生成掩码标签,生成的标签 label.png 为 16 位存储,而 OPENCV 默认读取 8 位,所以还需将 16 位转换成 8 位。

## 3 实验与分析

### 3.1 模型训练

模型训练和测试流程如图 6 所示。模型训练阶段主

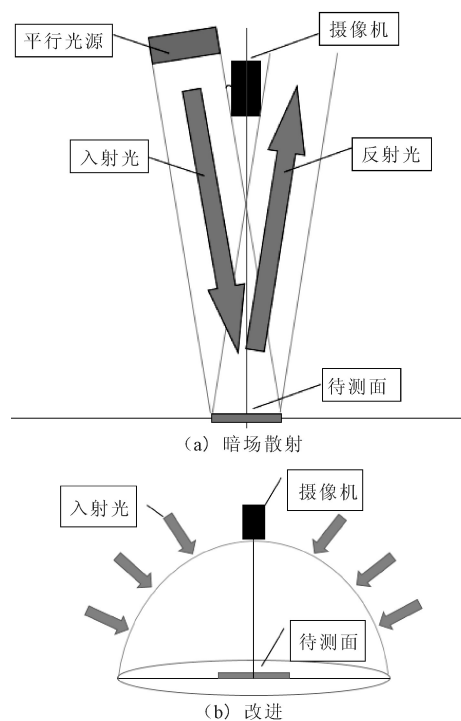


图 4 图像采集

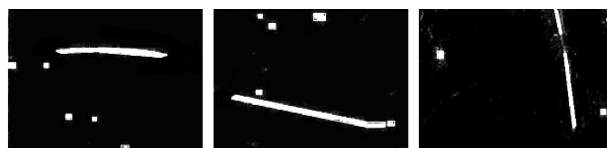


图 5 图像标注

要进行数据预处理、超参数调优、更换不同的主干网络等工作;模型测试阶段主要利用测试图片对训练所得的模型进行有效性检验。

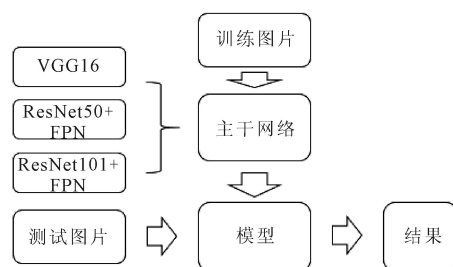


图 6 模型训练与测试流程

训练 Mask R-CNN 使用的数据集为 CoCo,大约有 80 000 张图片,81 个类别,而本实验训练数据集为笔者采集、标注,经过样本扩增有 1 200 张图片,4 个类别,背景较为简单一。相对于 CoCo 数据集,本文的数据集较小,所以在特征提取阶段,考虑到不同网络对实验结果的影响较大,所以本文测试 3 种特征提取网络:VGG16Net、ResNet50+FPN、ResNet101+FPN。

实验环境:GPU:1080ti、8G 内存,Ubuntu16.04、Python=3.4、TensorFlow=1.3.0、Keras=2.0.8 的实验环境,配置 Mask R-CNN,训练 40 000 次,所对应的损失变化如图 7 所示。

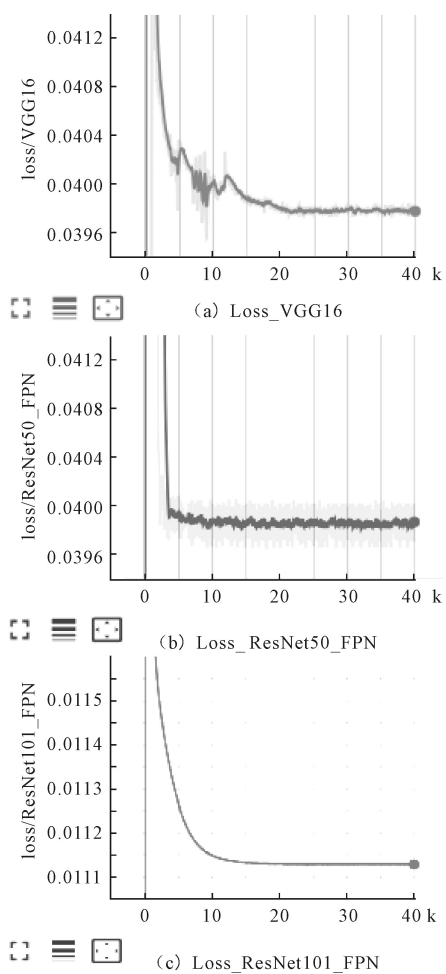


图 7 各主干网络损失变化曲线

3.2 结果分析

为评估 Mask R-CNN 在检测疵病问题时的有效性,本文使用召回率  $R=(TP)/(TP+FP)$  和精确率  $P=(TP)/(TP+FN)$  衡量模型性能。其中 200 张图片中含有不同的疵病数共 1191 个,测试结果如表 1 所示。

表 1 3 种不同主干网络测试结果

主干网络	测试数	误检数	漏检数	召回率(%)	精确率(%)
VGG16Net	1035	77	156	86.9	93.1
ResNet50+FPN	1042	35	149	87.4	96.7
ResNet101+FPN	1045	42	146	87.7	96.1

实验对比发现,ResNet101+FPN 有着更高的召回率,ResNet50 + FPN 有着更好的精确率,但是召回率和 ResNet101+FPN 相差不多,而 VGG16Net 的召回率和精确率较低。使用 ResNet50+FPN 为主干网络检测效果如图 8 所示,图中目标疵病处显示的是该目标的类别名和置信度,使用字母 A、B、C、D 代替斑点、划痕、崩边、污点等类别,掩膜和框图相对完整,能够检测图中各种尺寸、形状的疵病目标。

4 结语

本文以 Mask R-CNN 为基础模型,利用暗场散射、多  
万方数据

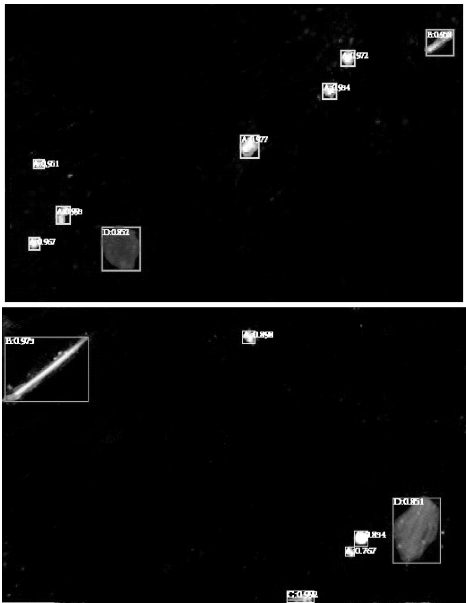


图 8 ResNet50+FPN 检测效果

图片堆栈的方法获取样本数据,然后进行处理、标注、样本扩增。接着用该数据集对基础模型进行训练,合理调节参数,并对比 3 种不同的主干网络,结果表明:ResNet50+FPN 的主干网络检测效果较好,精确率高,相对于传统方法具有良好的鲁棒性。但是实验也存在很多不足,如样本量少,没能实现更多的疵病种类识别,没有对比更多的主干网络,功能过少,没能实现疵病面积、周长、坐标等信息的获取等,这些有待在后续工作中进一步改进。

参考文献:

[1] 孙佳明. 激光陀螺反射镜基片表面疵病暗场散射显微检测系统研究[D]. 南京:南京理工大学,2015.

[2] 罗菁,董婷婷,宋丹,等. 表面缺陷检测综述[J]. 计算机科学与探索,2014(9):1041-1048.

[3] HE K,GKIOXARI G,DOLLÁR P,et al. Mask R-CNN[C]. IEEE International Conference on Computer Vision. IEEE,2017:2980-2988.

[4] REN S,HE K,GIRSHICK R,et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]. International Conference on Neural Information Processing Systems. MIT Press, 2015: 91-99.

[5] GIRSHICK R. Fast R-CNN [C]. IEEE International Conference on Computer Vision, 2015:1440-1448.

[6] SIMON YAN K,ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. arXiv preprint, 2014 (1): 1409-1556.

[7] HAYKIN S,KOSKO B. Gradient based learning applied to document recognition[M]. Hoboken:Wiley-IEEE Press,2009: 306-351.

[8] KRIZHEVSKY A,SUTSKEVER I,HINTON G E. ImageNet classification with deep convolutional neural networks[C]. Proceedings of the 25th International Conference on Neural Information Processing Systems,2012: 1097-1105.

[9] 张晋宽,李攀,刘耀红,等. 极低散射光学表面微观形貌评价技术研究[J]. 航空精密制造技术,2013 (3):1003-5451.



@RequestMapping 注册请求的路径,通过@ResponseBody 封装返回的数据对象,经过上述处理实现控制层功能。

(4)数据访问层。本系统采用的 Mysql 数据库与 dao 层进行数据交互,实现数据增删改查,在 Mybatis 源码中,Mybatis 框架首先创建 SQL 会话工厂实例,其作用相当于一个数据库连接池,然后 SQL 会话工厂通过配置的 xml 文件获取一个 SQL 会话。而执行 SQL 语句所需方法,都包含在 SQL 会话中,所以底层通过采用 SQL 会话完成对数据库的操作,同时对事物提交、回滚,用完之后需要释放数据库连接和关闭会话功能。

## 4 结语

针对当前后台管理亟待解决的问题,通过对该系统后台的开发,提出了基于 SpringBoot 微服务框架的后台管理系统方案,运用该框架提高开发效率增强、减少冗余配置、增强实用性。相关人员可以通过本系统进行后台的管理,包括售药机各个模块信息添加、删除、修改、查询以及售药机维护。在一般情况下,只需打开售药机后台管理就能了解售药机的情况,如需详细了解可以通过手机 APP 进行买药、出药,十分快捷。经过系统的运行测试,预期效果良好。但是使用该框架还存在很多不足,虽然将很多需要手动书写的代码集成在配置文件中,但原理上的认识存在缺陷,因此希望后续开发中能更深层次地理解核心源码。

### 参考文献:

- [1] FAIN, RA. Enterprise Web development[M]. Sebastopol O, Reily Media, 2017.
- [2] 刘天真,董彦军. 高校科研项目管理系统开发研究[J]. 中国高校科技与产业化, 2009(12): 52-53.
- [3] 李刚. 轻量级 J2EE 企业应用实战[M]. 第4版. 北京: 电子工业出版社, 2015.
- [4] 肖睿,禹晨. java 核心 API 编程[M]. 北京: 中国水利水电出版社, 2012.
- [5] DECK P. Spring MVC 学习指南[M]. 林仪明崔毅,译. 北京: 人民邮

电出版社, 2015.

- [6] 王艳清,陈红. 基于 SSM 框架的智能 Web 系统研发设计[J]. 计算机工程与设计, 2012, 33(12): 4751-4757.
- [7] 陈夫真. 基于 SSM 的某高校教室管理信息系统的设计与实现[D]. 苏州: 苏州大学, 2012.
- [8] 张俊萍,朱小冬. 基于 SSM 的软件体系结构开发过程研究[J]. 计算机测量与控制, 2011, 19(8): 2029-2032.
- [9] 任晓鹏,赵文兵,张春平. 基于框架的 Web 系统开发研究[J]. 计算机工程与设计, 2010, 31(4): 772-773.
- [10] 李洋. SSM 框架在 Web 应用开发中的设计与实现[J]. 计算机技术与发展, 2016(12): 190-194.
- [11] 张建勇,胥斌,尚艳英,等. 基于 SSH 框架的固定资产管理系统研究与设计[J]. 山西煤炭管理干部学院学报, 2013, 26(1): 112-113, 116.
- [12] 张朋. 基于 Struts2+Hibernate+Spring 框架整合的 xx 边防支队资产管理系统的研究与开发[D]. 北京: 北京化工大学, 2008.
- [13] 李勉. 基于 SSH 框架的企业固定资产管理系统设计与实现[D]. 北京: 北京邮电大学, 2011.
- [14] 刘婷婕. 基于 SSH 框架的固定资产管理系统的设计与原型实现[D]. 成都: 电子科技大学, 2011.
- [15] 王钱,王蓉,张利. 基于 iBatis 的通用数据持久层的研究与设计[J]. 微计算机信息, 2007, 23(4-3): 172-173.
- [16] 王威. MySQL 数据库源代码分析及存储引擎的设计[D]. 南京: 南京邮电大学, 2012.
- [17] 谈华宇,吴昶成,邱小平. 基于 Bootstrap 框架的动态表单设计与实现[J]. 无线互联科技, 2015(3): 99-100.
- [18] WALLS C. Spring Boot in Action[M]. America: Manning Publications Company, 2008.
- [19] 韩陆. Java RESTful Web Service 实战[M]. 第2版. 北京: 机械工业出版社, 2016.
- [20] BIBEALUT B, KATZ Y. jQuery in action[M]. New York: Manning Publications Company, 2010.
- [21] 董源,李培军,许舒人. 基于 AOP 的数据库应用安全控制的设计与实现[J]. 计算机系统应用, 2015, 24(12): 74-80.
- [22] NORMAN D. The Design of Everyday Things[M]. America: Mit Press, 2013.
- [23] 高兰,倪远平. 采用 Java 技术开发可复用的 Web 表示层构件[J]. 计算机工程与设计, 2006, 27(18): 3483-3485.

(责任编辑:江 艳)

(上接第67页)

- [10] XIE X H. A review of recent advances in surface defect detection using texture analysis techniques[J]. Electronic Letters on Computer Vision and Image Analysis, 2008, 7(3): 1-22.
- [11] ZHOU W J, FEI M U, ZHOU H Y, et al. A sparse representation based fast detection method for surface defect detection of bottle caps[J]. Neurocomputing, 2014(123): 406-414.
- [12] ZHAO J, PENG Y K, DHAKAL S, et al. On-line detector of apple surface defect based on image processing method[J]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(1): 260-263.
- [13] SUN X C, JIANG X N, FU Y, et al. Surface detection system for camshaft based on computer vision[J]. Infrared and Laser Engineering, 2013, 42(6): 1648-1653.
- [14] 陈璐. 光学系统中光学零件表面疵病测试系统研究[D]. 长春: 长春理工大学, 2013.

万方数据

- [15] 果宝智. 光学零件表面疵病的标识[J]. 激光与红外, 2000(9): 46-49.
- [16] 戴名奎,徐德衍. 光学元件的疵病检验与研究现状[J]. 激光与红外, 2001(18): 3-4.
- [17] 王甲甲. 扫描式手机屏疵病检测仪信号处理技术研究[D]. 西安: 西安工业大学, 2016.
- [18] 吕剑维. 基于积分散射理论的光学疵病快速检测方法[J]. 价值工程, 2016(11): 169-173.
- [19] 朱聪. 光学镜片表面疵病检测算法研究[D]. 成都: 西南交通大学, 2014.
- [20] 朱聪,于广婷,李柏林,等. 一种新的精密光学镜片表面疵病宽度测量方法[J]. 计算机应用与软件, 2014(12): 63-68.
- [21] 刘旭. 光学元件表面疵病检测扫描拼接的误差分析[D]. 杭州: 浙江大学, 2008.

(责任编辑:杜能钢)