

基于深度学习的表面缺陷检测方法综述

陶显¹ 侯伟^{1,2} 徐德^{1,2}

摘要 近年来, 基于深度学习的表面缺陷检测技术广泛应用在各种工业场景中. 本文对近年来基于深度学习的表面缺陷检测方法进行了梳理, 根据数据标签的不同将其分为全监督学习模型方法、无监督学习模型方法和其他方法三大类, 并对各种典型方法进一步细分归类 and 对比分析, 总结了每种方法的优缺点和应用场景. 本文探讨了表面缺陷检测中三个关键问题, 介绍了工业表面缺陷常用数据集. 最后, 对表面缺陷检测的未来发展趋势进行了展望.

关键词 深度学习, 表面缺陷检测, 机器视觉, 卷积神经网络

引用格式 陶显, 侯伟, 徐德. 基于深度学习的表面缺陷检测方法综述. 自动化学报, 2021, 47(5): 1017–1034

DOI 10.16383/j.aas.c190811

A Survey of Surface Defect Detection Methods Based on Deep Learning

TAO Xian¹ HOU Wei^{1,2} XU De^{1,2}

Abstract In recent years, surface defect detection techniques based on deep learning have been widely used in various industrial scenarios. This paper reviews the latest works on deep learning based surface defect detection methods. They are classified into three categories: full-supervised learning model method, unsupervised learning model method and other methods. The typical methods are further subdivided and compared. The advantages and disadvantages of these methods and their application scenarios are summarized. This paper analyzes three key issues in surface defect detection and introduces common data sets for industrial surface defects. Finally, the future development trend of surface defect detection is predicted.

Key words Deep learning, surface defect detection, machine vision, convolutional neural network (CNN)

Citation Tao Xian, Hou Wei, Xu De. A survey of surface defect detection methods based on deep learning. *Acta Automatica Sinica*, 2021, 47(5): 1017–1034

表面缺陷检测是机器视觉领域中非常重要的一项研究内容, 也称为 AOI (Automated optical inspection) 或 ASI (Automated surface inspection), 它是利用机器视觉设备获取图像来判断采集图像中是否存在缺陷的技术. 目前, 基于机器视觉的表面缺陷装备已经在各工业领域广泛替代人工肉眼检测, 包括 3C、汽车、家电、机械制造、半导体及电子、化工、医药、航空航天、轻工等行业. 传统的基于机器视觉的表面缺陷检测方法, 往往采用常规图像处理算法或人工设计特征加分类器方式. 一般来说, 通常利用被检表面或缺陷的不同性质进行成像方案

的设计, 合理的成像方案有助于获得光照均匀的图像, 并将物体表面缺陷明显地体现出来. 一种方式是针对被检表面颜色选择光源, 例如文献 [1] 中选择复合白色光源成像彩色布匹表面缺陷. 另一种常见的方式是依据被检测表面反射性质选择不同成像方案, 主要包括明场成像、暗场成像和混合成像等. 例如, Chen 等^[2] 针对金属易拉罐凹凸底部的表面缺陷检测, 设计了两个同心放置的圆锥环形明场光源, 用于同时照亮易拉罐底部的中央和外围区域. Tao 等^[3] 采用了暗场成像对大口径光学元件表面微弱划痕进行检测. 虽然精心构造的成像方案能够大大减轻经典检测算法设计的难度, 但也增加了检测系统的应用成本. 同时在很多开放式的工业环境下 (如图 1(b) 和图 1(d) 所示的自然场景), 期待设计的成像系统完全消除场景或者被检材料等变化对检测系统的影响, 往往不太现实. 在真实复杂的工业环境下, 表面缺陷检测往往面临诸多挑战, 例如存在缺陷成像与背景差异小、对比度低、缺陷尺度变化大且类型多样, 缺陷图像中存在大量噪声, 甚至缺陷在自然环境下成像存在大量干扰等情形, 如图 1 所示, 此时经典方法往往显得束手无策, 难以取得较

收稿日期 2019-11-27 录用日期 2020-03-25

Manuscript received November 27, 2019; accepted March 25, 2020

国家自然科学基金 (61703399, 61703398, 61973302, 61673383) 资助

Supported by National Natural Science Foundation of China (61703399, 61703398, 61973302, 61673383)

本文责任编辑 金连文

Recommended by Associate Editor JIN Lian-Wen

1. 中国科学院自动化研究所精密感知与控制研究中心 北京 100190 2. 中国科学院大学人工智能学院 北京 101408

1. Research Center of Precision Sensing and Control, Institute of Automation, Chinese Academy of Sciences, Beijing 100190 2. School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 101408

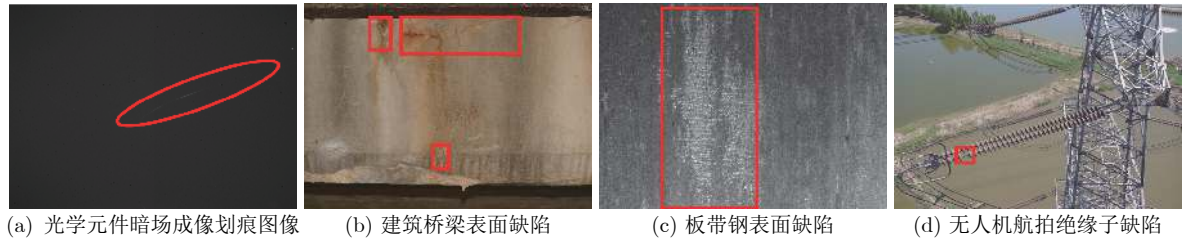


图 1 复杂工业环境下的表面缺陷图像

Fig.1 Images of surface defects in complex industrial environment ((a) Scratch image of dark field image of optical component; (b) Surface defect of building bridge; (c) Strip surface defect; (d) Unmanned aerial vehicle insulator defect)

好的检测效果.

近年来, 随着以卷积神经网络 (Convolutional neural network, CNN) 为代表的深度学习模型在诸多计算机视觉 (Computer vision, CV) 领域成功应用, 例如人脸识别、行人重识别、场景文字检测、目标跟踪和自动驾驶等, 不少基于深度学习的缺陷检测方法也广泛应用在各种工业场景中, 甚至国内外一些公司开发出多种基于深度学习的商用工业表面缺陷检测软件, 如表 1 所示. 全球传统工业视觉及其部件的市场规模在 2025 年将达到 192 亿美元^[4], 其中中国占比约为 30%, 并保持 14% 的年度平均增长率, 这一领域正在逐步被新一代基于深度学习的工业视觉技术替代. 同时我国在《中国制造 2025》白皮书中提出“推广采用先进成型和加工方法、在线检测装置、智能化生产和物流系统及检测设备等, 使重点实物产品的性能稳定性、质量可靠性、环境适应性、使用寿命等指标达到国际同类产品先进水平”. 因此, 基于深度学习的表面缺陷检测方法不仅具有重要的学术研究价值, 同时有着非常广阔的市场应用前景.

表 1 商用基于深度学习的缺陷检测软件
Table 1 Commercial deep learning based defect detection software

软件名称	公司	国家	年份
VIDI	已被康耐视收购	瑞士	2016
Halcon17 以上版本	MVTec	德国	2017
SuaKIT 数优-AI 深度学习缺陷检测软件	已被康耐视收购	韩国	2017
ALFA 深度学习外观缺陷检测软件	东莞埃法智能科技有限公司	中国	2018
AiDitron 人工智能软件	杭州谛创科技有限公司	中国	2018

鉴于目前国内还没有全面细致论述基于深度学习表面缺陷检测方法的综述文献, 本文通过对 2014 年 ~ 2019 年相关文献进行归纳梳理, 旨在帮助研究人员快速和系统地了解该领域相关方法与技术. 本文内容安排如下: 第 1 节给出缺陷检测问题

的定义. 第 2 节重点对最近几年相关方法进行详细介绍、细分和对比. 第 3 节分析基于深度学习的表面缺陷检测中三个关键问题. 第 4 节介绍工业领域公开的缺陷检测数据集. 最后, 对未来可能的研究焦点和发展方向进行了展望.

1 缺陷检测问题的定义

1) 缺陷的定义. 在机器视觉任务中, 缺陷倾向于人类经验上的概念, 而不是一个纯粹的数学定义. 对缺陷模式认知的不同, 会导致两种截然不同的检测手段. 以布匹表面缺陷检测为例, 如图 2 所示, 第 1 种是有监督的方法, 体现在利用标记了标签 (包括类别、矩形框或逐像素等) 的缺陷图像输入到网络中进行训练. 此时“缺陷”意味着标记过的区域或者图像. 因此, 该方法更关注缺陷特征, 例如在训练阶段将包含大片黑色范围的区域或者图像标记为“异色”缺陷用于网络训练. 在测试阶段, 当布匹图像中检测到大片黑色的特征时, 即认为出现了“异色”缺陷. 第 2 种是无监督的缺陷检测方法, 通常只需要正常无缺陷样本进行网络训练, 也称为 one-class learning. 该方法更关注无缺陷 (即正常样本) 特征, 当缺陷检测过程中发现未见过的特征 (异常特征) 时, 即认为检测出缺陷. 此时“缺陷”意味着异常, 因此该方法也称作异常检测 (Anomaly detection).

2) 缺陷检测的定义. 对比计算机视觉中明确的分类、检测和分割任务, 缺陷检测的需求非常笼统. 实际上, 其需求可以划分为三个不同的层次, 即“缺陷是什么”、“缺陷在哪里”和“缺陷是多少”. 第 1 阶段: “缺陷是什么”, 对应计算机视觉中的分类任务, 如图 2 中分类三种缺陷类别: 异色、空洞和经线, 这一阶段的任务可以称为“缺陷分类”, 仅仅给出图像的分类信息. 第 2 阶段: “缺陷在哪里”, 对应计算机视觉中的定位任务, 这一阶段的缺陷定位才是严格意义上的检测. 不仅获取图像中存在哪些类型的缺陷, 而且也给出缺陷的具体位置, 如图 2 中

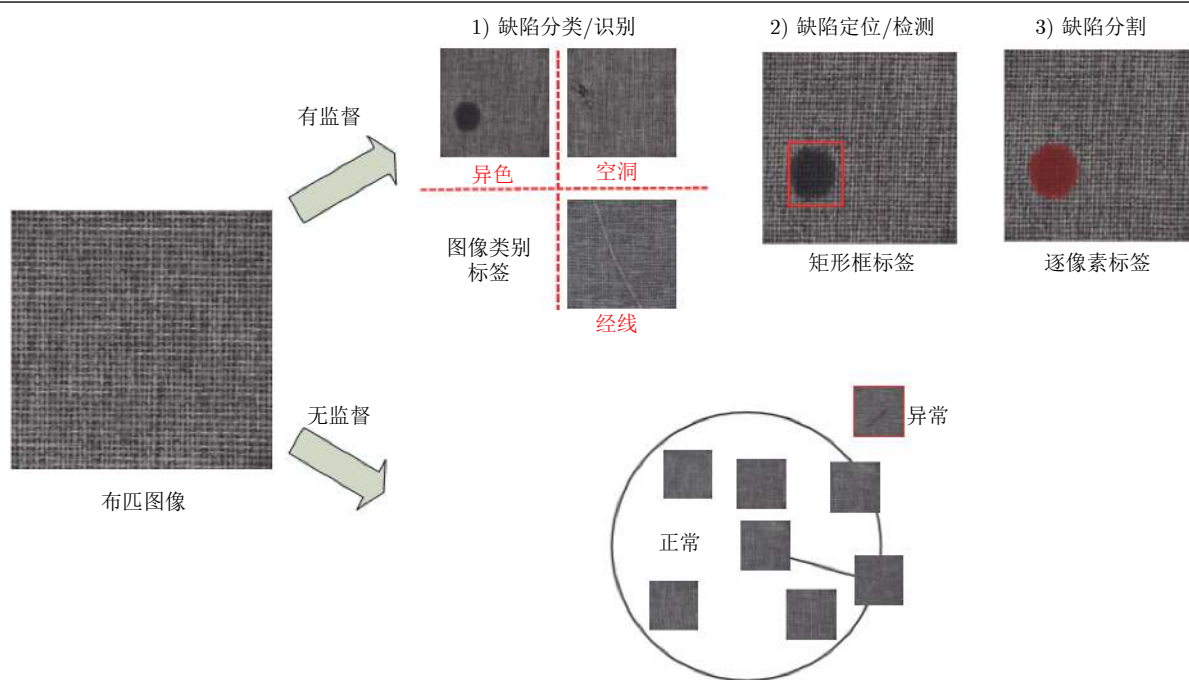


图2 缺陷检测的问题定义

Fig.2 Definition of defect detection problem

将异色缺陷用矩形框标记出来. 第3阶段:“缺陷是多少”,对应计算机视觉中的分割任务,如图2中缺陷分割的区域所示,将缺陷逐像素从背景中分割出来,并能进一步得到缺陷的长度、面积、位置等一系列信息,这些信息能辅助产品高一级的质量评估,例如优劣等级的判断.虽然缺陷检测的这三个阶段的功能需求和目标不同,但实际上三个阶段互相包含且能相互转换.例如第2阶段“缺陷定位”包含第1阶段“缺陷分类”这一过程,第3阶段“缺陷分割”同时也能完成第2阶段“缺陷定位”.第1阶段“缺陷分类”也能通过一些方法实现第2阶段和第3阶段的目标.因此,在后文中还是按照传统工业习惯统称为缺陷检测,只是在针对不同网络结构和目标功能时,才有所区分.

2 表面缺陷检测深度学习方

本节总结概述基于深度学习的表面缺陷检测方法.如图3所示,依据数据标签的不同,将其整体分为全监督学习模型、无监督学习模型和其他方法(半监督学习模型和弱监督学习模型).在全监督模型中,依据输入图像方式和损失函数的差异,分为基于表征学习和度量学习的方法.在表征学习中,根据网络结构的不同可以进一步细分为分类网络、检测网络和分割网络.目前大量的研究工作都是着眼于全监督学习方向,但无监督学习同样是一个值得研究的方向.从图3中可以看出,本文按照每类

方法的处理特点又细分为若干种不同的子方法.

2.1 表征学习

现阶段大部分基于深度学习的表面缺陷检测是基于有监督的表征学习方法.表征学习的本质是将缺陷检测问题看作计算机视觉中的分类任务,包括粗粒度的图像标签分类或区域分类,以及最精细的像素分类.由于实现的目标与计算机视觉任务完全一致,因此基于表征学习的缺陷检测方法可以看作是其相关经典网络在工业领域的一个应用.

2.1.1 分类网络

在真实的工业生产中,检测对象形状、尺寸、纹理、颜色、背景、布局和成像光照的巨大差异使复杂环境下的缺陷分类成为一项艰巨的任务.由于CNN强大的特征提取能力,采用基于CNN的分类网络目前已成为表面缺陷分类中最常用的模式.通常CNN分类网络的特征提取部分由级联的卷积层+pooling层组成,后面连接全连接层(或average pooling层)+softmax结构用于分类.一般来说,现有表面缺陷分类的网络常常采用计算机视觉中现成的网络结构,包括AlexNet^[5],VGG^[6],GoogLeNet^[7],ResNet^[8],DenseNet^[9],SENet^[10],ShuffleNet^[11],MobileNet^[12]等.或者针对实际问题搭建简易的网络结构,通过输入一幅测试图像到分类网络中,网络输出该图像的类别和其类别的置信度.依据分类网络方法实现任务的差异,我们将其细分为三个小类:

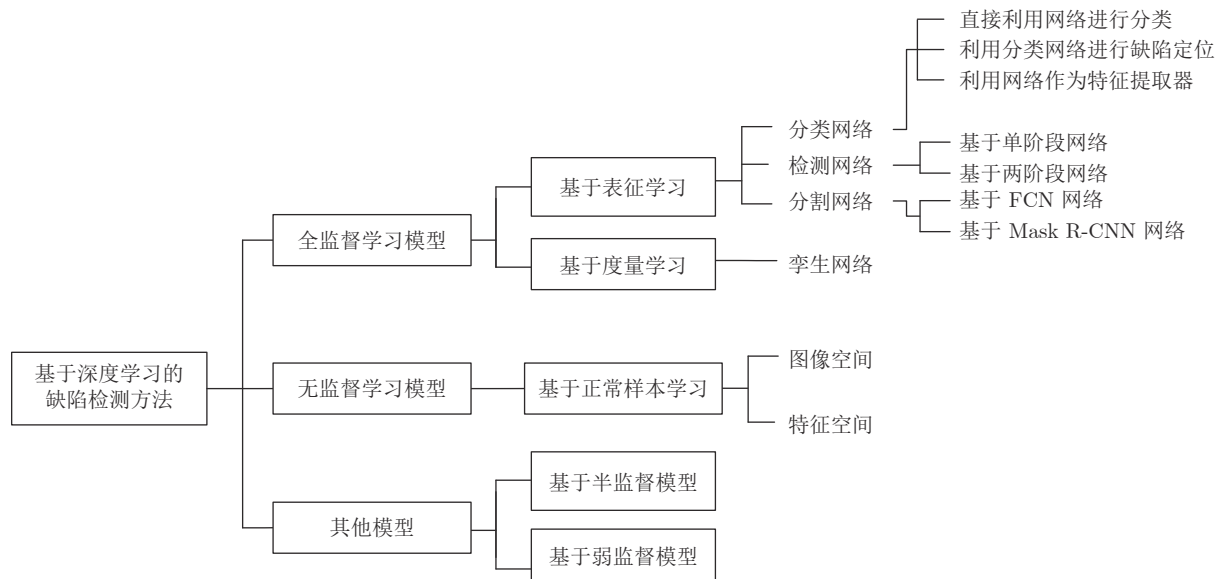


图3 缺陷检测方法框架图

Fig.3 Framework of defect detection methods

直接利用网络做分类、利用网络做缺陷定位和利用网络作为特征提取器。

1) 直接利用网络进行分类

直接利用分类网络做缺陷的分类任务是 CNN 最早应用于表面缺陷检测中常用的手段。根据研究工作的特点,可以进一步将其细分为原图分类、定位感兴趣区域 (Region of interest, ROI) 后分类和多类别分类三种。a) 原图分类。即直接将收集的完整缺陷图像放入网络进行学习训练。2014 年,奥地利科技研究所^[13]最早采集光度立体图像训练 CNN 网络来实现轨道表面空洞缺陷分类,整个网络共包含两个卷积层和两个池化层以及最后一个全连接层,在钢轨表面数据集上最终达到的错误识别率为 1.108%。Park 等^[14]设计了一种简易 CNN 分类网络,用于自动检测表面零件上的污垢、划痕、毛刺和磨损等缺陷。该方法在实验缺陷数据集上的平均检测正确率为 98%,其检测速度为 5 285 样本/min (图像分辨率为 32×32 像素)。Kyeong 等^[15]提出了一种卷积神经网络框架对半导体行业的晶圆仓图 (Wafer bin map, WBM) 中的混合类型缺陷模式进行分类。文献^[16]采用修改的 VGG19 网络用于识别 300×300 分辨率的太阳能面板图像缺陷,网络的准确率达到 88.42%,超过多种手工设计特征 (包括 KAZE^[17]、SIFT (Scale-invariant feature transform)^[18]、SURF (Speeded-up robust feature)^[19] 和支撑向量机 (Support vector machine, SVM) 方法的效果。Liang 等^[20]提出了一种基于 ShuffleNetV2 网络分类复杂背景下的瓶子喷墨码缺陷,所提出的

方法在塑料容器行业的在线喷墨码检测设备上获得了 99.88% 的分类正确率。直接利用原图分类的方法应用非常广泛,可用于多个领域的缺陷分类,例如焊接缺陷分类^[21]、聚合物锂电池水泡缺陷的分类^[22]和印刷电路板 (Printed circuit board, PCB) 缺陷的分类^[23]等。b) 定位 ROI 后分类。这种分类方法在许多工业应用中较为常见。通常来说,针对获取到的整幅图像,人们常常只关注某个固定区域中是否存在缺陷,因此往往预先获取到感兴趣的区域 (ROI),然后将 ROI 输入网络进行缺陷类别的判断。Shang 等^[24]提出了一种两阶段的铁轨缺陷识别算法,首先利用 Canny 算子和直线拟合算法在整个原始图像上对铁轨区域进行裁剪。然后将裁剪的图像放入 Inception V3 网络中提取特征以进行轨道图像分类。文献^[25]中通过级联的目标检测网络对高铁接触网螺栓区域进行获取,然后将裁剪的螺栓图像输入到 CNN 网络中进行缺陷分类。Li 等^[26]首先利用基于局部二值模式 (Local binary pattern, LBP) 特征的级联目标检测器实现扫描隧道显微镜成像材料待检测区域定位,然后采用 CNN 模型来获取表面缺陷的具体类型。c) 多类别分类。当待分类的缺陷类型超过两类时,常规的缺陷分类网络与原图分类方法一样,即网络的输出节点为缺陷类型的数目+1 (包括正常类别)。但多类别分类方法往往先采用一个基础网络进行缺陷与正常样本二分类,然后在同一个网络上共享特征提取部分,修改或者增加缺陷类别的分类分支。通过该方式相当于给后续的多目标缺陷分类网络准备了一个预训练权重参

数, 这个权重参数通过正常样本与缺陷样本之间二分类训练得到. Xie 等^[27] 首先训练第 1 个 ND (normal-defective)-CNN 模型进行二分类 (正常图像和所有其他缺陷图像), 缓解了数据不平衡的问题. 在训练好 ND-CNN 模型后, 将输出向量更改为 6 维向量来训练 ID (interdefect)-CNN 模型, 以使其适合于多类缺陷标签问题. 该模型在 ND-CNN 权重的基础上使用缺陷图像进行微调, 从而减少了样本量需求并节省了训练时间. Nagata 等^[28] 提出了一种 seeNet (Net with SVMs to classify sample images) 网络, 该网络有两个分类分支, 第 1 个二分类分支用来分类正常样本和 NG 样本, 网络的模型采用 AlexNet 进行特征提取, 其分类器采用 SVM; 第 2 个分支用于 7 类别的缺陷分类. 多类别分类采用这种二分支结构, 可以充分利用缺陷样本与正常样本数目不均衡的特点, 挖掘两者特征之间的差异.

2) 利用网络进行缺陷定位

一般认为, 分类网络只能完成图像标签级别的分类, 实际上结合不同的技巧和方式, 分类网络也可以实现缺陷的定位与逐像素的分类. 根据采用的手段不同, 可进一步将其分为滑动窗口、热力图 (heatmap) 和多任务学习网络三种形式. a) 滑动窗口. 是最简单和直观的实现缺陷粗定位的方法. 一般工业表面缺陷检测处理的图像分辨率较大, 通过较小尺寸的窗口在原始图像上进行冗余滑动, 将滑动窗口中的图像输入到分类网络中进行缺陷识别. 最后将所有的滑动窗口进行连接, 即可获得缺陷粗定位的结果. 2017 年, Cha 等^[29] 最早采用基于滑动窗口的 CNN 分类网络实现了裂纹表面缺陷定位, 两种滑动窗口冗余路径结合实现图像全覆盖, 如图 4 所示, 图 4(a) 为滑动窗口路径示意图, 图 4(b) 为裂纹定位的结果图. 文献 [30–32] 采用相同的方法应用于金属表面裂纹、路面表面裂纹和城市建筑物表面缺陷定位. b) 热力图 (heatmap). 是一种反映图像中各区域重要性程度的图像, 颜色越深代表越重要. 在缺陷检测领域, 热力图中颜色越深的区域代表其属于缺陷的概率越大. 2018 年, Ren 等^[33] 采用图像分块和特征迁移方式获取每个分块所对应属于缺陷的置信度, 并将其转换为热力图. 在热力图基础上运用 Otsu 法^[34] 和图割算法进一步得到准确的缺陷轮廓区域. 在计算机视觉领域, 常采用 CAM (Class activation mapping)^[35] 和 Grad-CAM^[36] 方法获得热力图, 其本质上是通过对加权特征图, 确定网络模型是通过哪些像素作为依据来判断输入图片所属的类别. Lin 等^[37] 采用 CAM 获取热力图, 并利用 Otsu 二值化方法分割热力图, 实现 LED 灯图像中划痕或线缺陷的定位. Zhou 等^[38]

采用 grad-CAM 方法获取热力图, 并采用 Otsu 算法分割得到表面缺陷的准确区域. c) 多任务学习网络. 单纯的分类网络若不加其他技巧, 一般只能实现图像级别的分类. 因此, 为了精细定位缺陷位置, 往往设计的网络会加上额外的分割分支, 两个分支共享特征提取的骨架 (backbone) 结果, 这样网络一般有分类和分割两个输出, 构成多任务学习网络. 它兼顾两个网络特点, 对于分割网络分支, 图像中每个像素都能被当作训练样本来训练网络. 因此, 多任务学习网络不仅利用分割分支输出缺陷具体的分割结果, 而且可以大大减少分类网络对样本的需求. Racki 等^[39] 设计了一个紧凑的一体化多任务 CNN 结构用于表面缺陷的分割和分类. 所提出的网络在各种表面纹理缺陷数据集 DAGM 2007 上取得了最好的效果. Tabernik 等^[40] 也提出了一个融合分类和分割分支的多任务缺陷检测网络, 整个网络只用了 50 幅 $1\,408 \times 512$ 分辨率的缺陷图像进行训练, 取得的缺陷准确识别率超过了商业 Cognex ViDi Suite 软件.

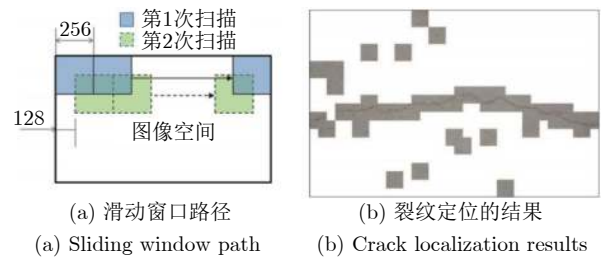


图 4 基于滑动窗口的裂纹定位
Fig. 4 Crack location based on sliding window

3) 利用网络做特征提取器

在早期基于深度学习的缺陷分类方法中, 不少文献利用 CNN 特征提取的强大功能, 先将图像输入到预训练网络中获取图像表征特征, 再将获取的特征输入到常规的机器学习分类器 (例如 SVM 等) 中进行分类. 例如 Zhao 等^[41] 采用 CNN 对电力航拍绝缘子图像块进行特征抽取, 将特征输入到 SVM 中进行二分类, 同时作者还比较了从不同全连接层抽取特征分类效果的差异. Malekzadeh 等^[42] 采用由 VGG-DNN 中的 fc6 层的输出作为特征, 输入到 SVM 中实现飞机表面缺陷分类, 缺陷分类精度达到 96% 以上.

综上, 基于分类网络的方法在实际应用中非常广泛, 同时, 不同的子方法本身也各有其自身优劣, 具体如表 2 所示.

2.1.2 检测网络

目标定位是计算机视觉领域中最基本的任务之

表 2 分类网络各子方法优缺点对比

Table 2 Comparison of advantages and disadvantages of each sub-method of classification network

代表子方法	优点	缺点
直接分类	结构经典,也是其他分类网络子方法的基础,可参考诸多现成网络	缺陷在图像中需要占一定比例,否则其特征容易被池化掉,同时一般一幅图像中只容许存在一种类别的缺陷(多标签分类除外)
定位 ROI 后分类	获取 ROI 的缺陷信息	需借助其他方法获取 ROI
多类别分类	一定程度上解决样本不平衡问题	网络采用二级训练
滑动窗口	在大图中实现缺陷的粗定位	滑动窗口尺寸需要准确选择,且只能获得较粗位置,遍历滑动速度慢
热力图	得到较为精准的缺陷区域	缺陷精确定位效果依赖网络分类性能
多任务学习	联合其他网络同时获取缺陷精确位置和类别,也能减少所需训练样本数目	网络结构相对复杂,在添加分割分支时,需要逐像素的标签
做特征提取器	获取有效的缺陷特征	依赖其他分类器才能获得最终分类结果

一,同时它也是与传统意义上缺陷检测最接近的任务,其目的是获得目标精准的位置和类别信息.目前,基于深度学习的目标检测方法层出不穷,一般来说,基于深度学习的缺陷检测网络从结构上可以划分为:以 Faster R-CNN (Region-CNN)^[43] 为代表的两阶段 (Two stage) 网络和以 SSD (Single shout multibox detector)^[44] 或 YOLO (You only look once)^[45] 为代表的一阶段 (One stage) 网络.两者的主要差异在于两阶段网络需要首先生成可能包含缺陷的候选框 (Proposal),然后再进一步进行目标检测.一阶段网络则直接利用网络中提取的特征来预测缺陷的位置和类别.

1) 基于两阶段的缺陷检测网络

两阶段检测网络 (Faster R-CNN) 的基本流程是首先通过 Backbone 网络获取输入图像的特征图,利用区域生成网络 (Region proposal network, RPN) 计算锚框 (anchor box) 置信度,获取 Proposal 区域.然后对 Proposal 区域的特征图进行 ROIpooling 后输入网络,通过对初步检测结果进行精细调整,最终得到缺陷的定位和类别结果.因此,针对缺陷检测的特点,常用方法往往针对 Backbone 结构或其特征图、锚框比例、ROIpooling 和损失函数等方面进行改进.2018 年,Cha 等^[46] 最早将 Faster R-CNN 直接应用在桥梁表面缺陷定位,其 Backbone 网络被替换为 ZF-net,在包含 2366 幅 500×375 像素大小的 5 类桥梁建筑数据集中,其 mAP (Mean average precision) 值达到 87.8%.Zhong 等^[47] 提出了一种基于三阶段 PVANET++ 的部件缺陷定位系统,用于高速铁路的悬链支撑装置—开口销的松动和缺失检测,PVANET++ 是一个改进版的 Faster R-CNN,相比于原版 Faster R-CNN,在最高层特征图进行 Proposal 提取,其改进点在于将低层两组特征图进行降采样与高层特征图上采样后联结形成新的超特征图用于 Proposal 提取.另外,

作者采用了 54 个不同比例的锚框,而不是原始的 9 个锚框.2020 年,Tao 等^[48] 设计了一个两阶段的 Faster R-CNN 网络用于无人机电力巡检中绝缘子缺陷定位,第 1 阶段用于自然场景下的绝缘子区域定位;第 2 阶段实现绝缘子区域中的缺陷定位.Xue 等^[49] 基于改进的 Faster R-CNN 实现盾构隧道中衬砌缺陷的检测,其 Backbone 采用改进的 Inception 全卷积网络得到特征图,同时增加两个锚框比例,使用位置敏感 ROIpooling 替代传统的 ROIpooling,所提出模型在单幅图像 48 ms 的测试时间速度下,实现了检测精度超过 95%.Ding 等^[50] 提出了一种针对 PCB 表面缺陷检测网络 (TDD (Ting defect detection)-Net),该方法通过使用 K 均值聚类设计合理的锚框大小;其次引入多尺度金字塔网络 (Feature pyramid networks, FPN) 到 Faster R-CNN 中,加强了来自底层结构信息的融合,使得网络适应微小的缺陷检测.最后,考虑到小数据集和样本不平衡的特点,在训练阶段采用了在线困难样本挖掘 (Online hard example mining, OHEM) 技术.该方法在 PCB 缺陷数据集上达到了 98.90% 的 mAP.He 等^[51] 提出了基于 Faster R-CNN 的带钢表面缺陷检测网络,该网络的改进在于将 Backbone 中多级特征图组合为一个多尺度特征图.在缺陷检测数据集 NEU-DET 上,提出的方法在采用 ResNet-50 的 Backbone 下实现了 82.3% 的 mAP.利用 Faster R-CNN 的检测方法也广泛应用在隧道^[52]、液晶面板偏振片表面^[53]、热成像绝缘子缺陷^[54]、铝型材表面^[55] 和轮胎轮毂^[56] 等缺陷检测领域.

2) 基于单阶段的缺陷检测网络

单阶段检测网络分为 SSD 和 YOLO 两种,两者都是利用整幅图作为网络的输入,直接在输出层回归边界框 (Bounding box) 的位置及其所属的类别.SSD 的特点在于引入了特征金字塔检测方式,从不同尺度的特征图中预测目标位置与类别.它使

用 6 个不同特征图检测不同尺度的目标,一般底层特征图用于预测小目标,高层特征图预测大目标. Chen 等^[25]采用改进的 SSD 网络进行接触网支撑装置上的紧固件缺陷区域定位,其主要改进部分在于采用不同层的特征图进行目标检测. Li 等^[57]提出了一种基于 MobileNet-SSD 的灌装生产线容器密封表面缺陷检测方法,通过 MobileNet 优化了 SSD 的 Backbone 结构,以简化检测模型参数. Liu 等^[58]同样采用基于 MobileNet-SSD 网络来定位高铁接触网支撑组部件,相比于原始 SSD 网络,其改进点在于: 1) 采用 MobileNet 为 Backbone; 2) 只使用 4 个不同的特征图来加速目标检测. 在测试的数据集上,其目标达到 25 帧/s 的检测速度和 94.3% 的 mAP. Zhang 等^[59]采用最新 YOLOv3 版本应用于桥梁表面缺陷定位,相比于原始 YOLOv3 网络,引入了预训练权重、批再规范化 (Batch renormalization) 和 Focal loss,进一步提高了缺陷检测率.

随着计算机视觉中目标检测网络的发展,相信在今后会有越来越多新的检测模型应用于表面缺陷检测中. 综上,在现阶段追求检测速度的缺陷检测领域中,基于一阶段的模型应用较多,在强调检测精度的缺陷检测领域中,基于两阶段的模型采用较多.

2.1.3 分割网络

分割网络将表面缺陷检测任务转化为缺陷与正常区域的语义分割甚至实例分割问题,它不但能精细分割出缺陷区域,而且可以获取缺陷的位置、类别以及相应的几何属性 (包括长度、宽度、面积、轮廓、中心等). 按照分割功能的区别,其大致可以分为: 全卷积神经网络 (Fully convolutional networks, FCN)^[60]方法和 Mask R-CNN^[61]方法.

1) FCN 方法

FCN 是图像语义分割的基础,目前几乎所有的语义分割模型都是基于 FCN. FCN 首先利用卷积操作对输入图像进行特征提取和编码,然后再通过反卷积操作或上采样将特征图逐渐恢复到输入图像尺寸大小. 依据 FCN 网络结构的差异,其缺陷分割方法可以进一步细分为常规 FCN、Unet^[62]和 SegNet^[63]三种方法. a) 常规 FCN 方法. Wang 等^[64]提出一种基于 FCN 的轮胎 X 射线图像缺陷分割方法,相比于原始 FCN 方法,该方法通过融合多尺度采样层的特征图来细化分割轮胎图像中的缺陷. Yu 等^[65]提出了一个基于 FCN 的两阶段表面缺陷分割模型,第 1 阶段采用一个轻量级的 FCN 快速获取粗略缺陷区域,然后,第 1 阶段的输出作为第 2 阶段 FCN 的输入用于细化缺陷分割结果,该方法在公共数据集 DAGM2007 上取得了 95.9934% 的

平均像素准确率. Dung 等^[66]采用基于 VGG16 编码器的 FCN 网络对混凝土表面裂缝进行分割,其平均像素准确率达到 90%. b) Unet 方法. Unet 不仅是一种经典的 FCN 结构,同时也是典型的编码器—解码器 (Encoder-decoder) 结构. 它的特点在于引入了跳层连接,将编码阶段的特征图与解码阶段的特征图进行融合,有利于分割细节的恢复. Huang 等^[67]提出了一个 MCuePush Unet 的模型用于磁瓦表面缺陷的显著性检测. 其 Unet 网络的输入为 MCue 模块生成三通道图像,包括一个显著性图像和两个原始图像. Li 等^[68]提出了一种基于改进 Unet 网络的混凝土结构表面缺陷分割方法. 在编码器采用 Dense Block 模块,同时跳层连接没有采用原始的 Concat 操作,而是逐像素求和. 该方法在包含 2750 幅图像 (504×376 像素) 4 种缺陷的混凝土结构数据库上取得了 91.59% 的平均像素准确率和 84.53% 的平均交并比 (Intersection over union, IoU). Liu 等^[69]提出了一种基于 ResNet 模块的 Unet 分割网络,用于 TFT-LCD (Thin film transistor liquid crystal display) 制造过程中导电颗粒的检测. c) SegNet 方法. 它也是一种经典的编码器—解码器结构. 其特点在于解码器中的上采样操作利用了编码器中最大池化操作的索引. Dong 等^[70]提出的 FL-SegNet 方法将原始的 SegNet 网络与 Focal loss 损失函数进行结合应用于分割隧道衬砌中的多种缺陷. Roberts 等^[71]设计了一种基于 SegNet 的网络,用于扫描隧道显微镜成像下材料晶体缺陷的语义分割. 在少量高质量的钢缺陷图像上三种缺陷的像素分类正确率为: 位错为 $91.60 \pm 1.77\%$, 沉淀物为 $93.39 \pm 1.00\%$, 空隙为 $98.85 \pm 0.56\%$. Zou 等^[72]在 SegNet 的编码器—解码器体系结构上构建 DeepCrack 网络,用于地面裂纹检测. 在该模型中,将分层卷积阶段学习的多尺度深度卷积特征融合在一起,以捕获精细的裂纹结构. DeepCrack 在三个具有挑战性的数据集上获得了平均 87% 以上的 F 值.

目前,基于深度学习的分割网络还在不断地提出,例如,LinkNet^[73]、DeepLabv3^[74]、PSPNet^[75]等. 很多最新模型中的模块,例如空洞卷积^[76-77]和金字塔 Pooling^[78]也被添加到 FCN 框架中,广泛应用于各种场景的缺陷分割. 近年来,生成对抗网络 (Generative adversarial network, GAN^[79]) 在计算机视觉领域得到广泛应用,常常用来生成图像. GAN 由生成器和判别器模型构成. 在结合 GAN 的缺陷分割方法中,生成器往往直接采用 FCN 网络,判别器通过分类模型来区分生成器的结果和 Groundtruth,

通过生成器和判别器的不断博弈, 让生成器的输出结果逐渐接近 Groundtruth. 结合 GAN 的分割方法已经用于在手机盖板玻璃信号孔缺陷分割^[80]和道路裂纹缺陷分割^[81]中.

2) Mask R-CNN 方法

Mask R-CNN 是目前最常用的图像实例分割方法之一, 可以看作是一种基于检测和分割网络相结合的多任务学习方法. 当多个同类型缺陷存在粘连或重叠时, 实例分割能将单个缺陷进行分离并进一步统计缺陷数目, 然而语义分割往往将多个同类型缺陷当作整体进行处理. 目前大部分文献都是直接将 Mask R-CNN 框架应用于缺陷分割, 例如路面缺陷分割^[82]、工业制造缺陷^[83]、螺栓紧固件缺陷^[84]和皮革表面缺陷^[85].

相比分类和检测网络方法, 分割方法在缺陷信息获取上有其优势. 但与检测网络一样, 需要大量的标注数据, 其标注信息是逐像素, 因此往往需花费大量的标注精力和成本.

2.2 度量学习

度量学习是使用深度学习直接学习输入的相似性度量. 在缺陷分类任务中, 往往采用孪生网络 (Siamese network) 进行度量学习. 不同于表征学习输入单幅图像转化为分类任务, 孪生网络的输入通常为两幅或多幅成对图像, 通过网络学习输出图片的相似度, 判断其是否属于同一类. 孪生网络损失函数的核心思想是让相似的输入距离尽可能小, 不同类别的输入距离尽可能大.

一般原始孪生网络的输入是两幅成对的图像, 网络的“连体”是通过共享权值来实现的. Kim 等^[86]设计了一个基于 CNN 结构的孪生网络对钢表面缺陷图像进行分类, 首先将两幅图像输入到共享权值的 CNN 中完成特征提取, 然后利用基于相似度函数的对比损失计算两个特征之间的差异程度. 在 NEU 钢表面缺陷数据集^[51]上, 该方法只通过 5 幅和 10 幅少量样本图像进行网络学习, 在 9 种缺陷类别上网络的分类准确度分别为 85.1% 和 86.5%. Wu 等^[87]提出了一种基于孪生网络的相似度度量方法来分类纽扣缺陷. 网络设计了新的损失函数用于自动特征提取和样本的相似性度量, 所提出的方法在包括凹痕、裂纹、污点、孔洞、凹凸不平等多种缺陷的纽扣数据集上进行了评估, 达到的分类精度为 98%. Liu 等^[88]基于改进的孪生网络对大型紧固件缺陷进行同时定位和缺陷分类. 其孪生网络后接两个分支: 第 1 个分支利用对比损失来实现紧固件定位; 第 2 个分支采用 softmax 损失对紧固件缺陷进

行分类. 在测试的数据集上, 紧固件定位的平均检测率达到 99.36%, 紧固件分类的平均检测率达到 92.69%. Staar 等^[89]将原始的孪生网络扩展为 Triplet 网络, 它具有三个输入, 包括两个相同类别和一个不同类别, 其原理与孪生网络一致. 除了缺陷分类, 孪生网络也能用于缺陷定位与检测. Tang 等^[90]提出了一个基于深度学习模型的 PCB 缺陷检测模型, 模型的前端将无缺陷的正常样本和有缺陷的图像对输入到孪生网络后进行特征差分, 将差分结果采用群组金字塔合并模块进行缺陷定位. 在 DeepPCB 数据集上, 该方法以 62 帧/s 达到 98.6% 的 mAP. Liu 等^[91]将孪生网络输出的特征图与其他子网络输出的结果进行特征融合来提高缺陷目标的显著性.

度量学习可以近似看作为学习样本在特征空间进行聚类; 表征学习可以近似看作为学习样本在特征空间的分界面. 相比于表征学习, 度量学习的方法应用在表面缺陷定位中不太多, 目前大部分都是应用在缺陷分类任务中. 在缺陷定位方面, 输入孪生网络的图像对需要具有统一的内容形式, 要求比较严格, 现阶段还无法适应复杂的工业环境.

2.3 正常样本学习

目前, 最常用于表面缺陷检测的无监督学习模型是基于正常样本学习的方法. 由于只需要正常无缺陷样本进行网络训练, 该方法也常称为 One-class learning. 正常样本学习的网络只接受正常 (无缺陷) 样本进行训练, 使得其具备强大的正常样本分布的重建和判别能力. 因此, 当网络输入的样本存在缺陷时, 往往会产生与正常样本不同的结果. 相比于有监督学习模型, 它能够检测到偏离预期的模式或没有见过的模式, 这些模式都可以称为缺陷或者异常. 依据处理空间的不同, 本文将该缺陷检测方法分为基于图像空间和特征空间两种. 通常该方法采用的网络模型为自编码器 (Autoencoder, AE) 和 GAN.

2.3.1 基于图像空间的方法

基于图像空间的方法是在图像空间上对缺陷进行检测. 因此, 该方法不仅能实现图像级别的分类和识别, 也可以获取到缺陷的具体位置. 该方法常用的手段主要有以下两种.

1) 利用网络实现样本重建与补全. 其原理类似去噪编码器, 当输入任意样本图像到网络中, 都可以得到其重建后对应的正常 (无缺陷) 样本, 因此, 网络可以看作具备自动修复或者补全缺陷区域的能力. 用输入图像分别减去这些重建或修复图像可以

获得残差图像, 这些残差图像也称为重建误差. 它能作为判断待检测样本是否异常的指标. 当重建误差过大时, 可以认为输入图像存在缺陷, 差异过大的区域即为缺陷区域. 当重建误差很小时, 即认为输入图像是正常样本. Mei 等^[92]通过使用卷积自编码器网络在不同的高斯金字塔等级上重建图像子块, 并合成来自这些不同分辨率通道的重建结果, 其网络的重建误差采用均方误差 (mean-square error, MSE) 损失. 该方法在多个数据集上取得了相比传统图像处理方法好的效果. Haselmann 等^[93]设计了一个基于全卷积的自动编码器网络, 基于正样本图像补全和重建的误差实现装饰性塑料零件表面缺陷检测. Kang 等^[94]设计了一个深度去噪编码器, 基于其重建误差来检测绝缘子图像中的缺陷, 其重建损失采用 L2 损失. Youkachan 等^[95]使用卷积自动编码器 (Convolutional autoencoder, CAE) 用于图像重建, 通过锐化处理重建误差获取最终的热轧带钢表面缺陷分割结果, 其重建误差采用 MSE. Zhao 等^[96]利用 GAN 和 CAE 实现表面缺陷图像的重建, 该方法将输入图像和重建图像输入到 LBP 算法处理后再做差. Bergmann 等^[97]首次将传统图像处理中的结构相似度 (Structural similarity, SSIM) 指标作为重建损失, 引入到基于自动编码器的图像重建中, 实验结果表明相比于传统 L2 损失, SSIM 损失能大幅提高表面缺陷检测的效果. Yang 等^[98]提出了一种基于无监督多尺度特征聚类的全卷积自动编码器 (Multi-scale feature-clustering-based fully convolutional autoencoder, MS-FCAE) 方法, 该方法利用处于不同比例等级的多个 FCAE 子网来重建若干纹理背景图像. 为了最大程度地提高效率和编码特征图的判别能力, 每个 FCAE 子网采用全卷积神经网络和特征聚类方法, 所提出的 MS-FCAE 方法在多个纹理表面检测数据集上进行了评估, 缺陷检测精度达到 92.0%.

2) 利用网络实现异常区域分类. 这类网络通常采用 GAN 的判别器. 该方法原理是训练生成对抗网络 GAN 以生成类似于正常表面图像的伪图像, 这意味着训练好的 GAN 可以在潜在特征空间中很好地学习正常样本图像. 因此, GAN 的判别器可以自然地用作分类器, 用于分类缺陷和正常样本. Zhai 等^[99]采用多尺度融合策略融合 GAN 鉴别器的三个卷积层的响应, 然后利用 Otsu 方法在融合特征响应图上进一步分割出缺陷位置. 该方法与文献 [33] 的原理类似, 但是利用了正常样本进行训练, 且其网络模型是 GAN 的判别器. Hu 等^[100]提出了一种基于深度卷积生成对抗网络 (Deep convolutional GAN, DCGAN) 的自动检测织物缺陷的新

型无监督方法. 该模型包括两个部分: 第 1 部分利用模型中 GAN 的判别器生成一个缺陷分布似然图, 其中每个像素值都表示该位置出现缺陷的概率; 第 2 部分通过引入编码器到标准 DCGAN, 实现重建检测图像的重建. 当从原始图像中减去重建图像时, 可以创建残差图以突出显示潜在的缺陷区域. 联合残差图和似然图以形成增强的融合图. 在融合图上采用阈值分割算法进一步获取准确的缺陷位置, 该方法在各种真实纺织物样品上进行了评估和验证.

2.3.2 基于特征空间的方法

基于特征空间的方法是在特征空间中, 通过正常样本与缺陷样本特征分布之间的差异来进行缺陷检测. 特征之间的差异也称为异常分数, 当异常分数高于某个值时, 即可认为出现缺陷. 2017 年, Schlegl 等^[101]最早提出了深度卷积生成对抗网络 AnoGAN, 实现检测图像从图像空间到潜在空间的映射, 其异常分数由图像空间中的差异和 GAN 网络中判别器最后一层特征图之间的差异来计算. Lai 等^[102]基于上述映射方法, 在潜在空间中直接采用 Fréchet 距离来实现缺陷与正常样本的区分, 在太阳能面板数据集上取得了 93.75% 的分类正确率. Soukup 等^[103]采用变分自编码器 (Variational autoencoder, VAE) 网络在潜在空间中实现缺陷检测. Liu 等^[104]设计了一个基于 GAN 和单类分类器 (One-class classifier) 的表面缺陷分类模型, GAN 中生成器 G 采用编码解码器, 其编码得到隐空间的特征输入到 SVM 分类器中进行缺陷分类, 提出的方法在钢材表面多缺陷数据集上获得的平均分类精度为 94%. 2019 年, Schlegl 等^[105]针对之前的 AnoGAN 工作提出了改进 f-AnoGAN, 其异常分数由引入编码器模块实现的图像样本重建损失和 GAN 网络中判别器最后一层特征图之间的差异来计算. Akcay 等^[106]提出了一种 GANomaly 用于图像异常检测, 该模型的创新点在于引入了编码解码和再编码模块. 其异常分数由三个部分组成: 图像样本重建差异和 GAN 网络中判别器最后一层特征图之间的差异, 以及编码器隐空间的特征与再编码后的特征之间的差异. 虽然文献 [101, 105–106] 应用对象不是表面缺陷, 但是其方法完全适用于表面缺陷检测. 传统上认为基于特征空间的方法往往只能实现图像级别的分类或识别, 无法获取像素级别的缺陷位置, 实际上, 通过 AE 和 GAN 模块也能实现与图像空间检测方法类似的缺陷精确定位.

综上, 目前基于正常样本学习方法常用于简单统一的纹理表面缺陷检测, 在复杂的工业检测环境下, 相比于监督学习的方法, 其检测效果还不太理想.

2.4 弱监督与半监督学习

相对于全监督和无监督的方法,目前弱监督和半监督方法在表面缺陷检测中的应用相对较少。

通常基于弱监督的方法是指采用图像级别类别标注(弱标签)来获取分割/定位级别的检测效果。Marino 等^[107]采用一种基于峰值响应图(Peak response maps, PRM)的弱监督学习方法^[108]来对马铃薯表面缺陷进行分类、定位和分割,从而实现质量控制任务的自动化。Mayr 等^[109]在原始 ResNet-50 分类网络上,通过删除原始全连接层和平均池化层,在网络最后添加两个 1×1 卷积来获取缺陷响应特征图,实现只用图像标签完成太阳能板上裂纹缺陷初步检测。Niu 等^[110]提出一种基于 GAN 的弱监督学习缺陷检测方法。通过 CycleGAN^[111]实现输入测试图像到其对应无缺陷图像的转化,比较输入图像和生成的无缺陷图像的差异,进而实现表面缺陷检测。半监督学习通常会使用大量的未标记数据和少部分有标签的数据用于表面缺陷检测模型的训练。He 等^[112]提出一个基于半监督 GAN 网络的方法^[113]应用于钢材表面缺陷分类,在设计 CAE-GAN 缺陷检测网络中,采用一个基于 CAE 的编码器并馈入 softmax 层以形成鉴别器。鉴别器不是直接预测输入图像的真假二分类,而是预测 $N + 1$ 类,其中 N 代表缺陷种类的数量,额外的类表示输入图像是来自真实数据集还是来自生成器。该方法在热轧板图像分类上取得了 96.5% 正确率。He 等^[114]提出一种多重训练的半监督学习方法应用于钢表面缺陷分类。该方法使用 cDCGAN 生成大量未标记的样本。为了利用未标记的样本,该模型提出了利用 cDCGAN 和 ResNet-18 的多训练融合算法用于未标记样本的类别标签预测。将预测分配类别标签的样本加入训练集中进一步训练。重复上述操作来逐步优化模型。在 NEU-CLS 缺陷数据集^[51]上的大量实验表明,即使原始样本有限,该方法对于缺陷分类也非常有效。Gao 等^[115]也提出了一种使用卷积神经网络的半监督学习方法来分类钢表面缺陷,通过采用伪标签改进了分类 CNN 的性能。在钢表面缺陷识别基准数据集上的实验结果表明,所提出的方法在有限的标记数据下可以实现良好的性能,达到 90.7% 的精度。目前,基于半监督的方法大部分用于解决缺陷分类或识别任务,还没有广泛应用到定位与分割任务中。

3 关键问题

3.1 小样本

目前深度学习方法广泛应用于各种计算机视觉

任务中,表面缺陷检测一般被看作是其在工业领域的具体应用。在传统的认识中,深度学习方法无法直接应用在表面缺陷检测中的原因是在真实的工业环境中,所能提供的工业缺陷样本太少。相比于 ImageNet 数据集中 1400 多万张样本数据,表面缺陷检测中面临的最关键的问题是样本小问题,在很多真实的工业场景下甚至只有几张或几十张缺陷图片。实际上,针对工业表面缺陷检测中关键问题之一的小样本问题,目前有以下 4 种不同的解决方式:

1) 数据扩增、合成与生成。最常用的缺陷图像扩增方法是对原始缺陷样本采用镜像、旋转、平移、扭曲、滤波、对比度调整等多种图像处理操作来获取更多的样本。例如 Tao 等^[116]、Wei 等^[117]和 Huang 等^[67]采用上述方法将扩增缺陷数据应用到深孔零件铜丝、纺织物表面和磁瓦缺陷检测中。另外一种较为常见的方法是数据合成,常常将单独缺陷融合叠加到正常(无缺陷)样本上构成缺陷样本。文献[48]利用分割网络将带缺陷的绝缘子从自然背景中分割出来,然后通过图像融合将其叠加到正常样本上。Haselmann 等^[118]采用骨架和纹理生成的方式得到逼真的缺陷并融合到正常样本上,形成对应的缺陷图像。最近,由于 GAN 在图像生成上强大的能力,不少基于 GAN 的工作也应用于表面缺陷样本生成上,如 Liu 等^[119]、Huang 等^[120]、Zhang 等^[121]和 Chou 等^[122]的方法。

2) 网络预训练或迁移学习。一般来说,由于深度学习网络参数较多,直接采用小样本训练网络很容易导致过拟合,但在预训练模型中存在一些比较共性的特征数据与权重信息。因此,基于预训练网络或迁移学习是目前针对样本少最常用的方法之一。2018 年, Ren 等^[33]最早将迁移学习应用于表面缺陷检测,其预训练模型采用 ImageNet 预训练模型。Yang 等^[123]、Zhang 等^[124]、Badmos 等^[125]和 Sun 等^[126]都采用迁移学习方法分别应用于液晶面板 Mura 缺陷、PCB 板缺陷、锂电池电极缺陷和金属零件表面缺陷检测。Kim 等^[127]在 DAGM 缺陷数据集上对比了基于微调(Fine-tuning)的迁移学习和从头开始训练网络的效果,证明基于迁移学习方法的性能优于从头开始训练网络的性能。

3) 合理的网络结构设计。通过设计合理的网络结构也可以大大减少样本的需求。例如在文献[40]中, Tabernik 等设计了一个融合分类和分割分支的多任务缺陷检测网络,两个分支共享特征提取的 Backbone,在分割网络分支,图像中每个像素都能被当作训练样本来训练网络。因此,该方式大大减少了网络对样本的需求。整个网络仅采用 50 张缺陷样本训练就能取得很好的效果。Wei 等^[117]提出了

一种将压缩感知与卷积神经网络相结合的缺陷检测方法, 基于压缩采样定理来压缩和扩充小样本数据, 使用 CNN 直接对压缩采样的数据特征进行分类. 相比于原始的图像输入, 通过对输入进行压缩采样能大大降低网络对样本的需求. 此外, 基于孪生网络的表面缺陷检测方法也可以看作是一种特殊的网络设计, 能够大幅减少样本需求.

4) 采用无监督与半监督模型方法. 这两种方式都可以减少样本的需求. 在无监督模型中, 只利用正常样本进行训练, 因此不需要缺陷样本. 半监督方法可以利用没有标注的样本来解决小样本情况下的网络训练难题.

3.2 实时性

基于深度学习的缺陷检测方法在工业应用中包括三个主要环节: 数据标注、模型训练与模型推断. 在实际工业应用中的实时性更关注模型推断这一环节. 目前大多数缺陷检测方法都集中在分类或识别的准确性上, 而很少关注模型推断的效率. 有不少方法用于加速模型, 例如模型权重量化^[128]和模型剪枝^[129]等. 在文献^[130]的工作中, 提出了一种端到端的手机屏幕表面缺陷检测模型. 为了提高模型计算过程的效率以满足实际的工业需求, 在建立低精度版本网络而又不损失准确性的前提下, 对提出的模型中的权重执行权重量化, 其缺陷分析的效率提高了 16%. 在传统的基于图像处理的表面缺陷检测方法中, FPGA 并行加速架构常用于加速图像处理算子. Pan 等^[131]将 FPGA 加速傅里叶重构算子用于纹理表面缺陷分割, 其 FPGA 并行加速架构比同类服务器 CPU 快三倍, 从而将整个 8.5 代 LCD 面板的扫描检测时间缩短至 8.5 s. 文献^[132]开发了一种实时热轧扁钢双表面检测系统, 提出的算法在 FPGA 上并行实现, 其并行图像处理技术将轧制速度从 5 m/s 提升为 20 m/s. 虽然现有深度学习模型使用 GPU 作为通用计算单元, 但随着技术发展, 相信 FPGA 会成为一个具有吸引力的替代方案.

3.3 与传统基于图像处理的缺陷检测方法对比

为了更好地说明基于深度学习的缺陷检测方法的特点, 本小节将基于深度学习的缺陷检测方法与传统基于图像处理的缺陷检测方法进行对比分析, 从方法、本质、所需条件和适应性等四个方面进行比较. 其中传统基于图像处理的缺陷检测的相关方法参考文献^[133]. 详细对比结果如表 3 所示.

4 缺陷检测数据集

缺陷检测数据集是开展研究工作的基础, 相比于计算机视觉任务中的 ImageNet^[134]、PASCAL-VOC2007/2012^[135]和 COCO^[136]等数据集, 缺陷检测并没有一个大量和统一的数据集, 在不同的检测对象和场景下, 研究往往基于不同的缺陷数据集.

本节按照缺陷检测方法进行分类, 提供了多种表面缺陷检测数据集的相关链接. 如表 4 所示, 涵盖多个工业应用领域, 具体包括: 金属表面^[40]、太阳能板^[109]、木材表面^[137]、磁瓦表面^[67]、铸件 X 射线图像^[138]、钢材和钢轨表面^[51, 139]、建筑物和桥梁表面缺陷^[140-141]、PCB 板^[90]、两个无监督学习的缺陷检测数据集: MVTec AD^[142]和 NanoTWICE^[143], 以及 DGAM2007 纹理缺陷数据集^[144].

由于目前基于深度学习的表面缺陷检测方法大多数是在特定的数据集上进行应用, 很多数据集没有公开, 难以统一比较. 为了更好地比较一些典型算法在同一缺陷数据上的应用效果, 我们选取了两个非常典型的数据集作为示例数据集: DAGM 和 GDxRay Casting.

DAGM 为纹理缺陷数据集, 它包含 6 个类别的图像样本, 大小为 512×512 像素. 每个类别包括 1 000 个无缺陷图像和 150 个有缺陷图像, 常用于缺陷分类的应用场景中. 典型缺陷分类方法在 DAGM 上的效果对比如表 5 所示, 表 5 按照时间顺序进行排序, 可以看出滑动窗口使用较多, 同时基于 CNN 的方法在纹理缺陷分类上取得很好的性能. 基于滑

表 3 传统图像处理与基于深度学习的缺陷检测方法的比较

Table 3 Comparison between traditional image processing and deep learning-based defect detection methods

对比项目	传统基于图像处理的方法	深度学习方法
方法	1) 结构法: 边缘、骨架、形态学等	基于卷积神经网络 CNN
	2) 统计法: 直方图、局部二值化特征 LBP、纹理特征、灰度共生矩阵 GLCM 等	
	3) 滤波法: 空间滤波、频域滤波 (傅里叶、gabor、小波) 等	
	4) 模型法: 随机场模型、反散射模型和分形体等	
本质	人工设计特征 + 分类器 (或规则)	从大量数据中自动学习特征
所需条件	相对苛刻的成像环境要求, 缺陷和非缺陷区域之间的高对比度, 少噪声	足够的学习数据和高性能运算单元
适应性	差 (成像环境变化或缺陷类型变化时往往需要更改阈值或重新设计算法)	相对强 (能够应对一定的工业检测环境变化)

表 4 工业表面缺陷检测常用数据集
Table 4 Common data sets for industrial surface defect detection

方法	应用场景	数据集名称	链接
分类	钢材表面	NEU-CLS ^[51]	http://faculty.neu.edu.cn/yunhyan/NEU_surface_defect_database.html
	太阳能板	elpv-dataset ^[100]	https://github.com/zae-bayern/elpv-dataset
	金属表面	KolektorSDD ^[40]	http://www.vicos.si/Downloads/KolektorSDD
	木材表面	wood defect database ^[137]	http://www.ee.oulu.fi/olli/Projects/Lumber.Grading.html
定位	钢材表面	NEU-DET ^[51]	http://faculty.neu.edu.cn/yunhyan/NEU_surface_defect_database.html
	铸件X射线图像	GDXray Casting ^[138]	https://domingomery.ing.puc.cl/material/gdxray/
分割	磁瓦表面	Magnetic-tile-defect-datasets ^[67]	https://github.com/abin24/Magnetic-tile-defect-datasets
	钢轨表面	RSDDs dataset ^[130]	http://icn.bjtu.edu.cn/Visint/resources/RSDDs.aspx
	地面裂纹	Crack_Dataset ^[140]	https://drive.google.com/drive/folders/1cplcUBmgHfD82YQTWnn1dssK2Z_xRpjx
	桥梁裂缝	Bridge Cracks ^[141]	https://github.com/maweifei/BridgeCrack_Image_Data
孪生网络	PCB 板	PCB Dataset ^[90]	https://github.com/tangsanli5201/DeepPCB
无监督学习	多种材质缺陷	MVTec AD ^[142]	http://www.mvtec.com/company/research/datasets
	扫描隧道显微镜成像 SEM 材料表面	NanoTWICE ^[143]	http://www.mi.imati.cnr.it/ettore/NanoTWICE/
弱监督学习	纹理缺陷	DAGM 2007 ^[144]	https://hci.iwr.uni-heidelberg.de/node/3616

表 5 典型缺陷分类方法在 DAGM 数据集上性能比较
Table 5 Effect comparison of defect classification methods on DAGM

方法	基础网络描述	是否预训练	分类准确率 (%)	发表时间
Weimer ^[145]	滑动窗口分类网络, 窗口大小 32×32	否	99.2	2016
Wang ^[146]	基于滑动窗口和多类别分类相结合的网络, 128×128大小的滑动窗口, 搭建的11层 CNN 网络	否	99.8	2017
Wu ^[147]	基于多尺度滑动窗口和分类网络, 8 层网络, 包括 inception 结构	否	98.6	2017
Yu ^[65]	两阶段 FCN 网络	是	95.9934	2017
Racki ^[30]	多任务分类网络, 包括分类分支和分割分支	否	99.655	2018

动窗口和多类别分类相结合的网络已经达到了 99.8% 的分类准确率, 也是目前基于深度学习的缺陷分类方法在该数据集上的最高水平. Yu 等^[65]采用的两阶段 FCN 方法是比较具有代表性的分割方法, 其在 DAGM 数据集上取得了 95.334% 像素分类准确率.

GDXray Casting 数据集包含 2 727 个 X 射线图像, 这些图像主要来自汽车零件, 包括铝制车轮和转向节. 每个图像中的铸件缺陷都用边界框 (Bounding box) 标记. 数据集中图像的大小范围为 256×256 像素到 768×572 像素. 不同的缺陷定位方法在 DAGM 上的效果对比如表 6 所示, 其中一阶段以 VGG-16 为 Backbone 的 SSD 网络在单幅图像获得了最快的检测速度, 达到了 0.025 s/幅 (单 GPU); 而以 ResNet-101 为 Backbone 的两阶

段网络 Faster R-CNN 则取得了最高的检测正确率.

5 总结与展望

缺陷检测是一个宽泛的应用领域, 具体包括缺陷分类、缺陷定位和缺陷分割等环节, 相比于传统图像处理方法分多个步骤和环节处理缺陷检测任务, 基于深度学习的方法将其统一为端到端的特征提取和分类. 虽然表面缺陷检测技术已经不断地从学术研究走向成熟的工业应用, 但是依然有一些需要解决的问题.

1) 网络结构设计. 第 2 节综述的方法大部分具有不同的网络结构, 这些网络都是人工设计的, 从模型包含多少层到每一层的详细结构都是一个漫长设计和调参的过程. 这种网络很难说是最优的, 只能说这些手工设计的网络在当前缺陷检测数据集上

表 6 典型缺陷定位方法在 GDXray casting 数据集上性能比较

Table 6 Effect comparison of defect location methods on GDXray casting

方法	mAP	使用GPU评估的时间/单幅图像 (s)
滑动窗口	0.461	0.231
Faster R-CNN (VGG-16) ^[43]	0.865	0.438
Faster R-CNN (ResNet-101) ^[43]	0.921	0.512
R-FCN (ResNet-101) ^[148]	0.875	0.375
SSD (VGG-16) ^[44]	0.697	0.025
SSD (ResNet-101) ^[44]	0.762	0.051

大致满足需求. 类似于传统手工设计的缺陷检测特征, 例如文献 [16] 中提到的三种手工特征 (KAZE^[17]、SIFT^[18]、SURF^[19]), 相比 CNN 网络自身学到的特征, 其缺陷分类效果逊色不少. 因此, 随着自动机器学习 and 网络架构搜索技术的兴起, 相信会有越来越多机器搜寻和自动生成的网络逐步替代人工设计的网络, 这些网络不仅能够大幅度的减少手工设计网络参数, 同时在检测的正确率上也会领先.

2) 网络训练学习. 人工在进行工业缺陷的目视检测时, 很难收集到所有缺陷类型的样本, 很多时候只有良品数据 (正样本). 然而目前大部分基于深度学习的表面缺陷检测方法是基于大量的缺陷样本的有监督学习. 深度学习的网络学习是一个“黑匣子”, 需要大量标注好的训练样本端到端进行学习, 可解释性差. 因此, 如何利用类脑 (受脑启发的) 计算与仿人视觉认知模型这些先验知识来指导缺陷检测网络的训练和学习, 也是一个值得思考的方向.

3) 异域数据联邦学习. 单个表面缺陷检测数据集往往都很少, 虽然小样本问题可以通过第 3 节介绍的方法缓解相关问题, 但是实际上不同工业行业和领域中, 真实工业表面缺陷数据是非常多的, 一些缺陷种类也是共同的, 例如划痕广泛存在于金属、液晶屏幕、太阳能电池板、玻璃等等一系列材质表面. 同时, 人类也会将统一类型的缺陷进行标记, 并不会因为检测领域的不同而产生差异. 但是由于涉及隐私敏感, 不同检测领域之间数据并没有有效结合和利用. 如何利用不同工业领域的缺陷数据集来进行网络学习, 也是表面缺陷检测的一种重要研究方向. 因此, 基于异域数据的联邦学习将会成为一个趋势, 它能够打破不同应用场景之间的壁垒, 充分学习不同领域之间数据来提升网络性能.

随着人工智能技术的发展, 目前基于机器视觉的表面缺陷检测的研究焦点已经从经典的图像处理 and 机器学习方法转移到深度学习方法, 在很多工业

场景下解决了以往传统方法无法解决的难题. 本文系统地总结、对比和分析了深度学习算法在表面缺陷检测领域的研究进展, 同时对基于深度学习的表面缺陷检测的研究趋势进行了展望, 以期对相关研究人员提供详实和有效参考.

References

- Jing J F, Liu S M, Li P F, Zhang L. The fabric defect detection based on CIE L*a*b* color space using 2-D Gabor filter. *The Journal of the Textile Institute*, 2016, **107**(10): 1305–1313
- Chen T J, Wang Y N, Xiao C Y, Wu Q M J. A machine vision apparatus and method for can-end inspection. *IEEE Transactions on Instrumentation and Measurement*, 2016, **65**(9): 2055–2066
- Tao X, Zhang Z T, Zhang F, Xu D. A novel and effective surface flaw inspection instrument for large-aperture optical elements. *IEEE Transactions on Instrumentation and Measurement*, 2015, **64**(9): 2530–2540
- China Industry Information Network. With the maturity of modern industrial automation technology, the global machine vision industry market will reach US\$ 12.5 billion by 2020, and it will exceed US\$ 19.2 billion by 2025. [Online], available: <http://www.chyxx.com/industry/201901/705852.html>, October 20, 2019
(中国产业信息网. 随着现代工业自动化技术日趋成熟到 2020 年全球机器视觉行业市场规模将达到 125 亿美元 2025 年将超过 192 亿美元. [Online], available: <http://www.chyxx.com/industry/201901/705852.html>, October 20, 2019)
- Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks. In: Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, Nevada, USA: Curran Associates Inc., 2012. 1097–1105
- Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv: 1409.1556, 2014.
- Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE, 2015. 1–9
- He K M, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016. 770–778
- Huang G, Liu Z, Van Der Maaten L, Weinberger K Q. Densely connected convolutional networks. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017. 4700–4708
- Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In: Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018. 7132–7141
- Zhang X Y, Zhou X Y, Lin M X, Sun J. ShuffleNet: An extremely efficient convolutional neural network for mobile devices. In: Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018. 6848–6856
- Howard A G, Zhu M L, Chen B, Kalenichenko D, Wang W J, Weyand T, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv: 1704.04861, 2017.
- Soukup D, Huber-Mörk R. Convolutional neural networks for

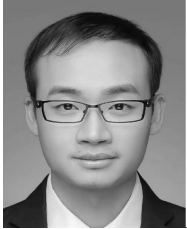
- steel surface defect detection from photometric stereo images. In: Proceedings of the 10th International Symposium on Visual Computing. Las Vegas, USA: Springer, 2014. 668–677
- 14 Park J K, Kwon B K, Park J H, Kang D J. Machine learning-based imaging system for surface defect inspection. *International Journal of Precision Engineering and Manufacturing-Green Technology*, 2016, **3**(3): 303–310
 - 15 Kyeong K, Kim H. Classification of mixed-type defect patterns in wafer bin maps using convolutional neural networks. *IEEE Transactions on Semiconductor Manufacturing*, 2018, **31**(3): 395–402
 - 16 Deitsch S, Christlein V, Berger S, Buerhop-Lutz C, Maier A, Gallwitz F, et al. Automatic classification of defective photovoltaic module cells in electroluminescence images. *Solar Energy*, 2019, **185**: 455–468
 - 17 Alcantarilla P F, Bartoli A, Davison A J. KAZE features. In: Proceedings of the 12th European Conference on Computer Vision. Florence, Italy: Springer, 2012. 214–227
 - 18 Lowe D G. Object recognition from local scale-invariant features. In: Proceedings of the 7th IEEE International Conference on Computer Vision. Kerkyra, Greece: IEEE, 1999. 1150–1157
 - 19 Bay H, Ess A, Tuytelaars T, Van Gool L. Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 2008, **110**(3): 346–359
 - 20 Liang Q K, Zhu W, Sun W, Yu Z, Wang Y N, Zhang D. In-line inspection solution for codes on complex backgrounds for the plastic container industry. *Measurement*, 2019, **148**: Article No. 106965
 - 21 Zhang Z F, Wen G R, Chen S B. Weld image deep learning-based on-line defects detection using convolutional neural networks for Al alloy in robotic arc welding. *Journal of Manufacturing Processes*, 2019, **45**: 208–216
 - 22 Ma L Y, Xie W, Zhang Y. Blister defect detection based on convolutional neural network for polymer lithium-ion battery. *Applied Sciences*, 2019, **9**(6): Article No. 1085
 - 23 Deng Y S, Luo A C, Dai M J. Building an automatic defect verification system using deep neural network for PCB defect classification. In: Proceedings of the 4th International Conference on Frontiers of Signal Processing (ICFSP). Poitiers, France: IEEE, 2018. 145–149
 - 24 Shang L D, Yang Q S, Wang J N, Li S B, Lei W M. Detection of rail surface defects based on CNN image recognition and classification. In: Proceedings of the 20th International Conference on Advanced Communication Technology (ICACT). Chuncheon-si Gangwon-do, Korea (South): IEEE, 2018. 45–51
 - 25 Chen J W, Liu Z G, Wang H R, Núñez A, Han Z W. Automatic defect detection of fasteners on the catenary support device using deep convolutional neural network. *IEEE Transactions on Instrumentation and Measurement*, 2018, **67**(2): 257–269
 - 26 Li W, Field K G, Morgan D. Automated defect analysis in electron microscopic images. *NPJ Computational Materials*, 2018, **4**(1): Article No. 36
 - 27 Xie Q, Li D W, Xu J X, Yu Z G, Wang J. Automatic detection and classification of sewer defects via hierarchical deep learning. *IEEE Transactions on Automation Science and Engineering*, 2019, **16**(4): 1836–1847
 - 28 Nagata F, Tokuno K, Nakashima K, Otsuka A, Ikeda T, Ochi H, et al. Fusion method of convolutional neural network and support vector machine for high accuracy anomaly detection. In: Proceedings of the 2019 IEEE International Conference on Mechatronics and Automation (ICMA). Tianjin, China: IEEE, 2019. 970–975
 - 29 Cha Y J, Choi W, Büyükoztürk O. Deep learning-based crack damage detection using convolutional neural networks. *Computer-Aided Civil and Infrastructure Engineering*, 2017, **32**(5): 361–378
 - 30 Chen F C, Jahanshahi M R. NB-CNN: Deep learning-based crack detection using convolutional neural network and Naïve Bayes data fusion. *IEEE Transactions on Industrial Electronics*, 2018, **65**(5): 4392–4400
 - 31 Park S, Bang S, Kim H, Kim H. Patch-based crack detection in black box images using convolutional neural networks. *Journal of Computing in Civil Engineering*, 2019, **33**(3): Article No. 04019017
 - 32 Feng C, Liu M Y, Kao C C, Lee T Y. Deep active learning for civil infrastructure defect detection and classification. In: Proceedings of the 2017 International Workshop on Computing in Civil Engineering (IWCCE). Seattle, Washington, USA: American Society of Civil Engineers (ASCE), 2017. 298–306
 - 33 Ren R X, Hung T, Tan K C. A generic deep-learning-based approach for automated surface inspection. *IEEE Transactions on Cybernetics*, 2018, **48**(3): 929–940
 - 34 Otsu N. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 1979, **9**(1): 62–66
 - 35 Zhou B L, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning deep features for discriminative localization. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016. 2921–2929
 - 36 Selvaraju R R, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual explanations from deep networks via gradient-based localization. In: Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017. 618–626
 - 37 Lin H, Li B, Wang X G, Shu Y F, Niu S L. Automated defect inspection of LED chip using deep convolutional neural network. *Journal of Intelligent Manufacturing*, 2019, **30**(6): 2525–2534
 - 38 Zhou F, Liu G H, Xu F, Deng H. A generic automated surface defect detection based on a bilinear model. *Applied Sciences*, 2019, **9**(15): Article No. 3159
 - 39 Racki D, Tomazevic D, Skocaj D. A compact convolutional neural network for textured surface anomaly detection. In: Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). Lake Tahoe, NV, USA: IEEE, 2018. 1331–1339
 - 40 Tabernik D, Šela S, Skvarč J, Skočaj D. Segmentation-based deep-learning approach for surface-defect detection. *Journal of Intelligent Manufacturing*, 2020, **31**(3): 759–776
 - 41 Zhao Z B, Xu G Z, Qi Y C, Liu N, Zhang T F. Multi-patch deep features for power line insulator status classification from aerial images. In: Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN). Vancouver, BC, Canada: IEEE, 2016. 3187–3194
 - 42 Malekzadeh T, Abdollahzadeh M, Nejati H, Cheung N M. Aircraft fuselage defect detection using deep neural networks. arXiv preprint arXiv: 1712.09213, 2020.
 - 43 Ren S Q, He K M, Girshick R, Sun J. 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
 - 44 Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C Y, et al. SSD: Single shot multibox detector. In: Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands: Springer, 2016. 21–37
 - 45 Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016. 779–788

- 46 Cha Y J, Choi W, Suh G, Mahmoudkhani S, Büyükoztürk O. Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types. *Computer-Aided Civil and Infrastructure Engineering*, 2018, **33**(9): 731–747
- 47 Zhong J P, Liu Z G, Han Z W, Han Y, Zhang W X. A CNN-based defect inspection method for catenary split pins in high-speed railway. *IEEE Transactions on Instrumentation and Measurement*, 2019, **68**(8): 2849–2860
- 48 Tao X, Zhang D P, Wang Z H, Liu X L, Zhang H Y, Xu D. Detection of power line insulator defects using aerial images analyzed with convolutional neural networks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020, **50**(4): 1486–1498
- 49 Xue Y D, Li Y C. A fast detection method via region-based fully convolutional neural networks for shield tunnel lining defects. *Computer-Aided Civil and Infrastructure Engineering*, 2018, **33**(8): 638–654
- 50 Ding R W, Dai L H, Li G P, Liu H. TDD-net: A tiny defect detection network for printed circuit boards. *CAAI Transactions on Intelligence Technology*, 2019, **4**(2): 110–116
- 51 He Y, Song K C, Meng Q G, Yan Y H. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features. *IEEE Transactions on Instrumentation and Measurement*, 2020, **69**(4): 1493–1504
- 52 Cheng J C P, Wang M Z. Automated detection of sewer pipe defects in closed-circuit television images using deep learning techniques. *Automation in Construction*, 2018, **95**: 155–171
- 53 Lei H W, Wang B, Wu H H, Wang A H. Defect detection for polymeric polarizer based on faster R-CNN. *Journal of Information Hiding and Multimedia Signal Processing*, 2018, **9**(6): 1414–1420
- 54 Zhao Z B, Zhen Z, Zhang L, Qi Y C, Kong Y H, Zhang K. Insulator detection method in inspection image based on improved faster R-CNN. *Energies*, 2019, **12**(7): Article No. 1204
- 55 Neuhauser F M, Bachmann G, Hora P. Surface defect classification and detection on extruded aluminum profiles using convolutional neural networks. *International Journal of Material Forming*, 2020, **13**(4): 591–603
- 56 Sun X H, Gu J N, Huang R, Zou R, Palomares B G. Surface defects recognition of wheel hub based on improved faster R-CNN. *Electronics*, 2019, **8**(5): Article No. 481
- 57 Li Y T, Huang H S, Xie Q S, Yao L G, Chen Q P. Research on a surface defect detection algorithm based on MobileNet-SSD. *Applied Sciences*, 2018, **8**(9): Article No. 1678
- 58 Liu Z G, Liu K, Zhong J P, Han Z W, Zhang W X. A high-precision positioning approach for catenary support components with multiscale difference. *IEEE Transactions on Instrumentation and Measurement*, 2020, **69**(3): 700–711
- 59 Zhang C B, Chang C C, Jamshidi M. Concrete bridge surface damage detection using a single-stage detector. *Computer-Aided Civil and Infrastructure Engineering*, 2020, **35**(4): 389–409
- 60 Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA: IEEE, 2015. 3431–3440
- 61 He K M, Gkioxari G, Dollár P, Girshick R. Mask R-CNN. In: Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017. 2961–2969
- 62 Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich, Germany: Springer, 2015. 234–241
- 63 Badrinarayanan V, Kendall A, Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **39**(12): 2481–2495
- 64 Wang R, Guo Q, Lu S M, Zhang C M. Tire defect detection using fully convolutional network. *IEEE Access*, 2019, **7**: 43502–43510
- 65 Yu Z Y, Wu X J, Gu X D. Fully convolutional networks for surface defect inspection in industrial environment. In: Proceedings of the 11th International Conference on Computer Vision Systems. Shenzhen, China: Springer, 2017. 417–426
- 66 Dung C V, Anh L D. Autonomous concrete crack detection using deep fully convolutional neural network. *Automation in Construction*, 2019, **99**: 52–58
- 67 Huang Y B, Qiu C Y, Guo Y, Wang X N, Yuan K. Surface defect saliency of magnetic tile. In: Proceedings of the 14th International Conference on Automation Science and Engineering (CASE). Munich, Germany: IEEE, 2018. 612–617
- 68 Li S Y, Zhao X F, Zhou G Y. Automatic pixel-level multiple damage detection of concrete structure using fully convolutional network. *Computer-Aided Civil and Infrastructure Engineering*, 2019, **34**(7): 616–634
- 69 Liu E Y, Chen K P, Xiang Z Y, Zhang J. Conductive particle detection via deep learning for ACF bonding in TFT-LCD manufacturing. *Journal of Intelligent Manufacturing*, 2020, **31**(4): 1037–1049
- 70 Dong Y N, Wang J, Wang Z F, Zhang X, Gao Y, Sui Q M, et al. A deep-learning-based multiple defect detection method for tunnel lining damages. *IEEE Access*, 2019, **7**: 182643–182657
- 71 Roberts G, Haile S Y, Sainju R, Edwards D J, Hutchinson B, Zhu Y Y. Deep learning for semantic segmentation of defects in advanced STEM images of steels. *Scientific Reports*, 2019, **9**(1): Article No. 12744
- 72 Zou Q, Zhang Z, Li Q Q, Qi X B, Wang Q, Wang S. DeepCrack: Learning hierarchical convolutional features for crack detection. *IEEE Transactions on Image Processing*, 2019, **28**(3): 1498–1512
- 73 Chaurasia A, Culurciello E. LinkNet: Exploiting encoder representations for efficient semantic segmentation. In: Proceedings of the 2017 IEEE Visual Communications and Image Processing (VCIP). St. Petersburg, FL, USA: IEEE, 2017. 1–4
- 74 Chen L C, Zhu Y K, Papandreou G, Schroff F, Adam H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany: Springer, 2018. 801–818
- 75 Zhao H S, Shi J P, Qi X J, Wang X G, Jia J Y. Pyramid scene parsing network. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017. 2881–2890
- 76 Tao X, Zhang D P, Ma W Z, Liu X L, Xu D. Automatic metallic surface defect detection and recognition with convolutional neural networks. *Applied Sciences*, 2018, **8**(9): Article No. 1575
- 77 Anantrasirichai N, Bull D. DefectNET: Multi-class fault detection on highly-imbalanced datasets. arXiv preprint arXiv: 1904.00863, 2019.
- 78 Choi W, Cha Y J. SDDNet: Real-time crack segmentation. *IEEE Transactions on Industrial Electronics*, 2020, **67**(9): 8016–825
- 79 Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. In: Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Quebec, Canada: MIT Press, 2014. 2672–2680

- 80 Yuan Z C, Zhang Z T, Su H, Zhang L, Shen F, Zhang F. Vision-based defect detection for mobile phone cover glass using deep neural networks. *International Journal of Precision Engineering and Manufacturing*, 2018, **19**(6): 801–810
- 81 Mei Q P, Gül M. A conditional wasserstein generative adversarial network for pixel-level crack detection using video extracted images. arXiv preprint arXiv: 1907.06014, 2019.
- 82 Singh J, Shekhar S. Road damage detection and classification in smartphone captured images using mask R-CNN. arXiv preprint arXiv: 1811.04535, 2018.
- 83 Ferguson M K, Ronay A K, Lee Y T T, Law K H. Detection and segmentation of manufacturing defects with convolutional neural networks and transfer learning. *Smart and Sustainable Manufacturing Systems*, 2018, **2**(1): 137–164.
- 84 Taheritanjani S, Schoenfeld R, Bruegge B. Automatic damage detection of fasteners in overhaul processes. In: Proceedings of the 15th International Conference on Automation Science and Engineering (CASE). Vancouver, BC, Canada: IEEE, 2019, 1289–1295
- 85 Liong S T, Gan Y S, Huang Y C, Yuan C A, Chang H C. Automatic defect segmentation on leather with deep learning. arXiv preprint arXiv: 1903.12139, 2019.
- 86 Kim M S, Park T, Park P. Classification of steel surface defect using convolutional neural network with few images. In: Proceedings of the 12th Asian Control Conference (ASCC). Kitakyushu-shi, Japan, : IEEE, 2019. 1398–1401
- 87 Wu S L, Wu Y B, Cao D H, Zheng C Y. A fast button surface defect detection method based on Siamese network with imbalanced samples. *Multimedia Tools and Applications*, 2019, **78**(24): 34627–34648
- 88 Liu J B, Huang Y P, Zou Q, Tian M, Wang S C, Zhao X X, et al. Learning visual similarity for inspecting defective railway fasteners. *IEEE Sensors Journal*, 2019, **19**(16): 6844–6857
- 89 Staar B, Lütjen M, Freitag M. Anomaly detection with convolutional neural networks for industrial surface inspection. *Proceedia CIRP*, 2019, **79**: 484–489
- 90 Tang S L, He F, Huang X L, Yang J. Online PCB defect detector on a new PCB defect dataset. arXiv preprint arXiv: 1902.06197, 2019.
- 91 Liu R, Yao M. Surface defect detection based on deep convolutional Siamese networks and defect saliency. *Acta Microscopica*, 2019, **28**(2): 199–207
- 92 Mei S, Yang H, Yin Z P. An unsupervised-learning-based approach for automated defect inspection on textured surfaces. *IEEE Transactions on Instrumentation and Measurement*, 2018, **67**(6): 1266–1277
- 93 Haselmann M, Gruber D P, Tabatabai P. Anomaly detection using deep learning based image completion. In: Proceedings of the 17th IEEE International Conference on Machine Learning and Applications (ICMLA). Orlando, Florida, USA : IEEE, 2018. 1237–1242
- 94 Kang G Q, Gao S B, Yu L, Zhang D K. Deep architecture for high-speed railway insulator surface defect detection: Denoising autoencoder with multitask learning. *IEEE Transactions on Instrumentation and Measurement*, 2019, **68**(8): 2679–2690
- 95 Youkachan S, Ruchanurucks M, Phatrapomnant T, Kaneko H. Defect segmentation of hot-rolled steel strip surface by using convolutional auto-encoder and conventional image processing. In: Proceedings of the 10th International Conference of Information and Communication Technology for Embedded Systems (IC-ICTES). Bangkok, Thailand: IEEE, 2019. 1–5
- 96 Zhao Z X, Li B, Dong R, Zhao P. A surface defect detection method based on positive samples. In: Proceedings of the 15th Pacific Rim International Conference on Artificial Intelligence. Nanjing, China: Springer, 2018. 473–481
- 97 Bergmann P, Löwe S, Fauser M, Sattlegger D, Steger C. Improving unsupervised defect segmentation by applying structural similarity to autoencoders. arXiv preprint arXiv: 1807.02011, 2019.
- 98 Yang H, Chen Y F, Song K Y, Yin Z P. Multiscale feature-clustering-based fully convolutional autoencoder for fast accurate visual inspection of texture surface defects. *IEEE Transactions on Automation Science and Engineering*, 2019, **16**(3): 1450–1467
- 99 Zhai W, Zhu J, Cao Y, Wang Z F. A generative adversarial network based framework for unsupervised visual surface inspection. In: Proceedings of the 2018 International Conference on Acoustics, Speech and Signal Processing (ICASSP). Calgary, AB, Canada: IEEE, 2018. 1283–1287
- 100 Hu G H, Huang J F, Wang Q H, Li J R, Xu Z J, Huang X B. Unsupervised fabric defect detection based on a deep convolutional generative adversarial network. *Textile Research Journal*, 2020, **90**(3-4): 247–270
- 101 Schlegl T, Seeböck P, Waldstein S M, Schmidt-Erfurth U, Langs G. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In: Proceedings of the 25th International Conference on Information Processing in Medical Imaging. Boone, NC, USA: Springer, 2017. 146–157
- 102 Lai Y T K, Hu J S. A texture generation approach for detection of novel surface defects. In: Proceedings of the 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC). Miyazaki, Japan: IEEE, 2018. 4357–4362
- 103 Soukup D, Pinetz T. Reliably decoding autoencoders' latent spaces for one-class learning image inspection scenarios. In: Proceedings of the 2018 OAGM Workshop. Hall/Tyrol, Austria: Verlag der TU Graz, 2018. 90–93
- 104 Liu K, Li A M, Wen X, Chen H Y, Yang P. Steel surface defect detection using GAN and one-class classifier. In: Proceedings of the 25th International Conference on Automation and Computing (ICAC). Lancaster, United Kingdom: IEEE, 2019. 1–6
- 105 Schlegl T, Seeböck P, Waldstein S M, Langs G, Schmidt-Erfurth U. f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks. *Medical Image Analysis*, 2019, **54**: 30–44
- 106 Akcay S, Atapour-Abarghouei A, Breckon T P. GANomaly: Semi-supervised anomaly detection via adversarial training. In: Proceedings of the 14th Asian Conference on Computer Vision. Perth, Australia: Springer, 2018. 622–637
- 107 Marino S, Beausery P, Smolarz A. Weakly-supervised learning approach for potato defects segmentation. *Engineering Applications of Artificial Intelligence*, 2019, **85**: 337–346
- 108 Zhou Y Z, Zhu Y, Ye Q X, Qiu Q, Jiao J B. Weakly supervised instance segmentation using class peak response. In: Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018. 3791–3800
- 109 Mayr M, Hoffmann M, Maier A, Christlein V. Weakly supervised segmentation of cracks on solar cells using normalized Lp norm. In: Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP). Taipei, China: IEEE, 2019. 1885–1889
- 110 Niu S L, Lin H, Niu T Z, Li B, Wang X G. DefectGAN: Weakly-supervised defect detection using generative adversarial network. In: Proceedings of the 15th International Conference on Automation Science and Engineering (CASE). Vancouver, BC, Canada: IEEE, 2019. 127–132
- 111 Zhu J Y, Park T, Isola P, Efros A A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Pro-

- ceedings of the 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017. 2223–2232
- 112 He D, Xu K, Zhou P, Zhou D D. Surface defect classification of steels with a new semi-supervised learning method. *Optics and Lasers in Engineering*, 2019, **117**: 40–48
- 113 Odena A. Semi-supervised learning with generative adversarial networks. arXiv preprint arXiv: 1606.01583, 2016.
- 114 He Y, Song K C, Dong H W, Yan Y H. Semi-supervised defect classification of steel surface based on multi-training and generative adversarial network. *Optics and Lasers in Engineering*, 2019, **122**: 294–302
- 115 Gao Y P, Gao L, Li X Y, Yan X G. A semi-supervised convolutional neural network-based method for steel surface defect recognition. *Robotics and Computer-Integrated Manufacturing*, 2020, **61**: Article No. 101825
- 116 Tao X, Wang Z H, Zhang Z T, Zhang D P, Xu D, Gong X Y, et al. Wire defect recognition of spring-wire socket using multi-task convolutional neural networks. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 2018, **8**(4): 689–698
- 117 Wei B, Hao K R, Tang X S, Ding Y S. A new method using the convolutional neural network with compressive sensing for fabric defect classification based on small sample sizes. *Textile Research Journal*, 2019, **89**(17): 3539–3555
- 118 Haselmann M, Gruber D. Supervised machine learning based surface inspection by synthesizing artificial defects. In: Proceedings of the 16th IEEE international conference on machine learning and applications (ICMLA). Cancun, Mexico: IEEE, 2017. 390–395
- 119 Liu L Z, Cao D H, Wu Y B, Wei T R. Defective samples simulation through adversarial training for automatic surface inspection. *Neurocomputing*, 2019, **360**: 230–245
- 120 Huang C C, Lin X P. Study on machine learning based intelligent defect detection system. In: Proceedings of the 3rd International Conference on Inventions (ICI). Taiwan, China: MATEC Web of Conferences, 2018. 1–10
- 121 Zhang H D, Chen Z Z, Zhang C Q, Xi J T, Le X J. Weld defect detection based on deep learning method. In: Proceedings of the 15th International Conference on Automation Science and Engineering (CASE). Vancouver, BC, Canada: IEEE, 2019. 1574–1579
- 122 Chou Y C, Kuo C J, Chen T T, Horng G J, Pai M Y, Wu M E, et al. Deep-learning-based defective bean inspection with GAN-structured automated labeled data augmentation in coffee industry. *Applied Sciences*, 2019, **9**(19): Article No. 4166
- 123 Yang H, Mei S, Song K Y, Tao B, Yin Z P. Transfer-learning-based online Mura defect classification. *IEEE Transactions on Semiconductor Manufacturing*, 2018, **31**(1): 116–123
- 124 Zhang C, Shi W, Li X F, Zhang H J, Liu H. Improved bare PCB defect detection approach based on deep feature learning. *The Journal of Engineering*, 2018, **2018**(16): 1415–1420
- 125 Badmos O, Kopp A, Bernthaler T, Schneider G. Image-based defect detection in lithiumion battery electrode using convolutional neural networks. *Journal of Intelligent Manufacturing*, 2020, **31**(4): 885–897
- 126 Sun J, Wang P, Luo Y K, Li W Y. Surface defects detection based on adaptive multiscale image collection and convolutional neural networks. *IEEE Transactions on Instrumentation and Measurement*, 2019, **68**(12): 4787–4797
- 127 Kim S, Kim W, Noh Y K, Park F C. Transfer learning for automated optical inspection. In: Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN). Anchorage, USA: IEEE, 2017. 2517–2524
- 128 Krishnamoorthi R. Quantizing deep convolutional networks for efficient inference: A whitepaper. arXiv preprint arXiv: 1806.08342, 2018.
- 129 Han S, Pool J, Tran J, Dally W J. Learning both weights and connections for efficient neural network. In: Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, Quebec, Canada: MIT Press, 2015. 1135–1143
- 130 Lei J, Gao X, Feng Z L, Qiu H M, Song M L. Scale insensitive and focus driven mobile screen defect detection in industry. *Neurocomputing*, 2018, **294**: 72–81
- 131 Pan Y F, Lu R S, Zhang T D. FPGA-accelerated textured surface defect segmentation based on complete period Fourier reconstruction. *Journal of Real-Time Image Processing*, 2020, **17**(5): 1659–1673
- 132 Luo Q W, He Y G. A cost-effective and automatic surface defect inspection system for hot-rolled flat steel. *Robotics and Computer-Integrated Manufacturing*, 2016, **38**: 16–30
- 133 Luo Jing, Dong Ting-Ting, Song Dan, Xiu Chun-Bo. A review on surface defect detection. *Journal of Frontiers of Computer Science and Technology*, 2014, **8**(9): 1041–1048 (罗菁, 董婷婷, 宋丹, 修春波. 表面缺陷检测综述. 计算机科学与探索, 2014, **8**(9): 1041–1048)
- 134 Deng J, Dong W, Socher R, Li L, Li K, Li F F. Imagenet: A large-scale hierarchical image database. In: Proceedings of the 2009 Conference on Computer Vision and Pattern Recognition. Anchorage, USA: IEEE, 2009. 248–255
- 135 Everingham M, Van Gool L, Williams C K I, Winn J, Zisserman A. The PASCAL visual object classes (VOC) challenge. *International Journal of Computer Vision*, 2010, **88**(2): 303–338
- 136 Lin T Y, Maire M, Belongie S, Hays J, Perona P, Ramanan D, et al. Microsoft COCO: Common objects in context. In: Proceedings of the 13th European Conference on Computer Vision. Zurich, Switzerland: Springer, 2014. 740–755
- 137 Silvén O, Niskanen M, Kauppinen H. Wood inspection with non-supervised clustering. *Machine Vision and Applications*, 2003, **13**(5-6): 275–285
- 138 Ferguson M, Ak R, Lee Y T T, Law K H. Automatic localization of casting defects with convolutional neural networks. In: Proceedings of the 2017 IEEE International Conference on Big Data. Boston, MA, USA: IEEE, 2017. 1726–1735
- 139 Gan J R, Li Q Y, Wang J Z, Yu H M. A hierarchical extractor-based visual rail surface inspection system. *IEEE Sensors Journal*, 2017, **17**(23): 7935–7944
- 140 Yang X C, Li H, Yu Y T, Luo X C, Huang T, Yang X. Automatic pixel-level crack detection and measurement using fully convolutional network. *Computer-Aided Civil and Infrastructure Engineering*, 2018, **33**(12): 1090–1109
- 141 Li Liang-Fu, Ma Wei-Fei, Li Li, Lu Cheng. Research on detection algorithm for bridge cracks based on deep learning. *Acta Automatica Sinica*, 2019, **45**(9): 1727–1742 (李良福, 马卫飞, 李丽, 陆铖. 基于深度学习的桥梁裂缝检测算法研究. 自动化学报, 2019, **45**(9): 1727–1742)
- 142 Bergmann P, Fauser M, Sattlegger D, Steger C. MVTec AD — A comprehensive real-world dataset for unsupervised anomaly detection. In: Proceedings of the 2019 IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA: IEEE, 2019. 9592–9600
- 143 Carrera D, Manganini F, Boracchi G, Lanzarone E. Defect detection in SEM images of nanofibrous materials. *IEEE Transactions on Industrial Informatics*, 2017, **13**(2): 551–561
- 144 DAGM 2007 Datasets. Weakly supervised learning for industrial optical inspection [Online], available: <https://hci.iwr.uni-heidelberg.de/node/3616>, February 27, 2018

- 145 Weimer D, Scholz-Reiter B, Shpitalni M. Design of deep convolutional neural network architectures for automated feature extraction in industrial inspection. *CIRP Annals*, 2016, **65**(1): 417–420
- 146 Wang T, Chen Y, Qiao M N, Snoussi H. A fast and robust convolutional neural network-based defect detection model in product quality control. *The International Journal of Advanced Manufacturing Technology*, 2018, **94**(9-12): 3465–3471
- 147 Wu X J, Cao K, Gu X D. A surface defect detection based on convolutional neural network. In: Proceedings of the 11th International Conference on Computer Vision Systems. Shenzhen, China: Springer, 2017. 185–194
- 148 Dai J F, Li Y, He K M, Sun J. R-FCN: Object detection via region-based fully convolutional networks. In: Proceedings of the 30th International Conference on Neural Information Processing Systems. Barcelona, Spain: Curran Associates Inc., 2016. 379–387



陶 显 中国科学院自动化研究所副研究员. 2016 年获得中国科学院自动化研究所博士学位. 主要研究方向为机器视觉, 缺陷检测和深度学习. 本文通信作者.

E-mail: taoxian2013@ia.ac.cn

(**TAO Xian** Associate professor at the Institute of Automation, Chinese Academy of Sciences. He received his Ph.D. degree at the Institute of Automation, Chinese Academy of Sciences in 2016. His research interest covers machine vision, defect detection, and deep learning. Corresponding author of this paper.)



侯 伟 中国科学院大学人工智能学院博士研究生. 2009 年和 2014 年分别获得兰州大学学士和硕士学位. 主要研究方向为缺陷检测, 计算机视觉, 图像处理和机器学习.

E-mail: houwei2018@ia.ac.cn

(**HOU Wei** Ph.D. candidate at the School of Artificial Intelligence, University of Chinese Academy of Science. He received his bachelor degree and master degree from Lanzhou University in 2009 and 2014, respectively. His research interest covers defect detection, computer vision, image processing, and machine learning.)



徐 德 中国科学院自动化研究所研究员. 1985 年和 1990 年分别获得山东工业大学学士和硕士学位. 2001 年获得浙江大学博士学位. 主要研究方向为机器人视觉测量, 视觉控制, 智能控制, 视觉定位, 显微视觉, 微装配.

E-mail: de.xu@ia.ac.cn

(**XU De** Professor at the Institute of Automation, Chinese Academy of Sciences. He received his bachelor degree and master degree from Shandong University of Technology in 1985 and 1990, respectively, and received his Ph.D. degree from Zhejiang University in 2001. His research interest covers robotics and automation such as visual measurement, visual control, intelligent control, visual positioning, microscopic vision, and microassembly.)