

基于深度森林的人机协同工业制品表面缺陷识别^①



阎昊¹, 刘奕阳²

¹(武汉数字工程研究所, 武汉 430074)

²(郑州大学 计算机与人工智能学院, 郑州 450001)

通信作者: 刘奕阳, E-mail: 202012172013083@gs.zzu.edu.cn

摘要: 针对工业制品缺陷分类存在的样本图像少、分类准确性不足和模型训练耗时长等问题, 提出了一种基于深度森林的人机协同分类模型. 该模型首先通过深度森林对样本图像进行初步识别, 经多粒度扫描模块和级联森林模块提取特征, 得到初始预测结果并分离出识别困难的样本图像; 然后采用人机协同的策略, 采用人工方式随机标注部分识别困难的样本, 再利用 K 近邻算法对剩余识别困难的样本进行再分类. 通过在公开数据集以及生产线实际采集的真实数据上的实验结果表明, 改进的分类模型在工业制品表面缺陷数据集上的性能优于基线算法.

关键词: 工业制品缺陷; 缺陷分类; 深度森林; 人机协同

引用格式: 阎昊, 刘奕阳. 基于深度森林的人机协同工业制品表面缺陷识别. 计算机系统应用, 2022, 31(12): 280-286. <http://www.c-s-a.org.cn/1003-3254/8868.html>

Human-machine Collaborative Identification of Industrial Product Surface Defects Based on Deep Forest

YAN Hao¹, LIU Yi-Yang²

¹(Wuhan Digital Engineering Institute, Wuhan 430074, China)

²(School of Computer and Artificial Intelligence, Zhengzhou University, Zhengzhou 450001, China)

Abstract: A human-machine collaborative classification model based on the deep forest is proposed to solve the problems encountered in the classification of defects in industrial products, such as sample image shortage, insufficient classification precision, and time-consuming model training. For this purpose, sample images are preliminarily identified with deep forest, and their features are extracted by the multi-granularity scanning module and the cascaded forest module. The initial estimation result is thereby obtained, and sample images difficult to identify are separated out. Then, the human-machine collaboration strategy is employed. Specifically, some of the sample images difficult to identify are randomly labeled manually, and the remaining ones are reclassified with the K-nearest neighbor algorithm. The experimental results on the public dataset and the real data collected from the production line indicate that the improved classification model offers performance superior to that of the baseline algorithm on the dataset of industrial product surface defects.

Key words: industrial product defect; defect classification; deep forest; human-machine collaboration

在工业制品生产过程中, 由于生产环境不理想、生产设备老化及故障、操作工艺不规范等原因, 制品难免出现各种缺陷. 这些缺陷不仅会对制品的外观、功能、质量造成不良影响, 还有可能造成巨大经济损

失甚至带来安全隐患. 工业制品的缺陷检测是现代工业生产流程中不可或缺的重要环节, 缺陷分类是缺陷检测的基本任务之一, 旨在解决“缺陷是什么”的基本问题. 目前工业生产中仍有大量检测任务由人工实现,

① 收稿时间: 2022-04-13; 修改时间: 2022-06-01; 采用时间: 2022-06-06; csa 在线出版时间: 2022-08-12

依靠人眼进行缺陷的识别。近年来,为了降低工业制品检测的成本、提高检测的精度、效率和一致性,基于图像处理、机器视觉等技术的工业制品缺陷自动检测技术开始得到应用^[1]。

工业制品缺陷分类主要是针对工业制品表面缺陷的分类,即根据被检制品表面的颜色、反射特性等,选择或设计不同类型的成像方案,再利用图像处理、机器学习等技术自动识别工业产品表面的裂痕、划伤、污损,外包装的打皱、斜封、压纹不良、标签缺失等缺陷。实际工业生产环境的复杂性、开放性以及工业制品成像条件的可变性、不确定性增加了工业制品缺陷分类的难度^[2]。近年来,随着数据资源的丰富以及计算机算力的不断提升,深度学习以其强大的特征学习能力在计算机视觉、语音识别、人机对弈、自动驾驶等诸多应用领域取得了突破性的进展,基于深度学习的工业制品缺陷分类也被成功应用于不同的工业领域^[3]。然而,基于深度学习的工业制品缺陷分类在理论研究和应用落地仍然存在亟待解决的瓶颈难题。首先,深度学习检测模型的训练依赖大量高质量的标注样本,而在实际生产线上能够获得的缺陷样本一般数量较为有限。其次,基于深度神经网络的检测模型存在拓扑设计难,学习调试超参数难的问题。另外,基于深度神经网络的检测模型的训练成本较高,尤其当需要对模型的结构和控制参数进行反复调试时,会消耗更多的计算资源。针对上述问题,本文提出了一种基于深度森林的人机协同工业制品缺陷分类模型,将深度森林应用到缺陷分类中进行预分类,使分类模型不受样本数量规模和特征向量规模的限制;引入人机协同思想,对样本图像预分类的结果进行纠错。

1 相关工作

1.1 传统工业制品缺陷分类方法

传统工业制品缺陷分类方法一般基于图像处理算法,采用人工设计的特征,通过训练缺陷类别分类器实现。常用的分类器包括贝叶斯分类器、支持向量机、随机森林、神经网络等^[4]。例如, Hanzaei 等^[5]通过统计图像局部方差进行缺陷边缘检测,然后提取瓷砖局部特征,采用支持向量机进行瓷砖表面缺陷分类。Sun 等^[6]通过改进的高斯混合模型提取焊接缺陷的特征区域,构建决策树分类器实现焊接缺陷的分类。Pasadas 等^[7]将白噪声激励方法与支持向量机结合,对多层铝结构

中地下裂纹的位置和深度进行分类。Li 等^[8]采用支持向量机对电缆缺陷进行分类识别,并使用粒子群优化算法提升分类性能。总体来说,传统方法需要领域专家为特定的产品专门设计特征,通用性较差,不利于推广落地。

1.2 基于深度学习的工业制品缺陷分类方法

基于深度学习的缺陷分类方法一般不需要人工设计特征,能够实现由数据驱动的、端到端的自动缺陷分类。目前大部分基于深度学习的缺陷分类模型采用计算机视觉中比较成熟的网络结构,包括 AlexNet、VGG、ResNet、ShuffleNet 等。例如, Masci 等^[9]提出了最大汇集卷积神经网络方法对钢材缺陷进行分类,相比于传统的分类算法,降低了超过一半的分类错误率。Cheon 等^[10]提出了一种基于深度学习的 ADC 方法,可自动对各种类型的晶圆表面损伤进行分类。张光建等^[11]使用 CNN 网络模型,采用早停技术检测网络训练,实现小数据样本的手机外壳表面划痕缺陷分类。Saqlain 等^[12]提出了一种卷积神经网络用于晶圆缺陷自动分类,使用正则化方法提高网络模型的性能。Deutsch 等^[13]采用修改的 VGG19 网络用于识别太阳能面板图像缺陷。Aydin 等^[14]提出了一种精度较高的融合模型,通过融合 SqueezeNet 和 MobileNetV2 的特征对钢轨表面缺陷进行分类。Sassi 等^[15]通过全连接神经网络识别多类焊接缺陷,在 306 个训练样本的小数据集中达到了 97.22% 的分类准确率。

1.3 深度森林

深度森林可视为基于决策树的集成学习模型,是由决策树森林集成串联构建的多层有监督学习模型。深度森林主要包括级联森林和多粒度扫描两个相互独立的模块^[16]。

1.3.1 级联森林

级联森林是一个通过加入新特征来对原始特征进行扩展的结构,由多个级联层组成,每个级联层包括随机森林和完全随机森林。级联森林在最后一级对输出的概率向量取均值,得到预测结果。级联森林的层数是自适应确定的,使得级联森林能够适合不同规模的数据集。级联结构增加了模型的深度而不引入额外的参数,超参数较少也有利于整体模型的训练和理论分析。

1.3.2 多粒度扫描

深度森林使用多粒度扫描对级联森林进行增强。

将样本图像输入后,使用类似 CNN 的滑动窗口,利用多尺度滑窗在数据集图片上滑动取值以获得多种特征子样本.特征子样本通过多个随机森林转换为表征向量,将多个表征向量拼接得到本层的输出^[16].

2 人机协同分类

本文提出的基于深度森林的人机协同缺陷分类方法主要包括 4 个步骤:(1) 图像预处理.对图像样本进行去噪处理及数据增广等;(2) 基于深度森林的缺陷预

分类.图像样本训练多粒度级联森林模型,使用滑动窗口提取特征向量,用级联森林训练特征向量后得到初始预测结果.根据初始分类预测概率,设置可变动阈值分离出识别困难的样本图像;(3) 人机协同纠错.人工识别部分困难的样本图像,利用困难样本在特征空间内分布上的相似性,通过 K 最近邻算法对剩余识别困难的样本图像进行纠错;(4) 集成分类.综合样本图像的初始预测,得出最终的预测结果.分类方法的主要过程如图 1 所示.

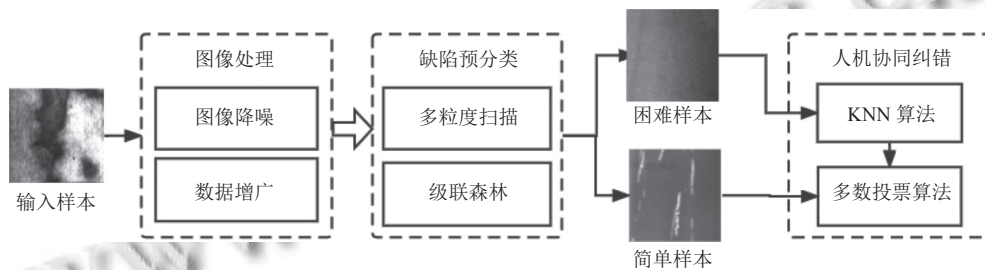


图 1 基于深度森林的分类过程

2.1 图像处理

2.1.1 图像降噪

在实际生产线上采集制品图像的过程中,工业相机、传输信道等会随机产生椒盐噪声.为此,首先采用中值滤波对采集到的图像样本进行降噪预处理,减少椒盐噪声对后续缺陷分类带来的不利影响.中值滤波的去噪效果如图 2 所示,其中图 2(a) 为去噪前的图像样本,图 2(b) 为去噪后的图像样本.

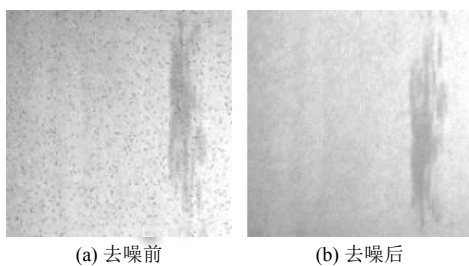


图 2 中值滤波降噪

2.1.2 数据增广

实际生产线上能够获得的缺陷样本一般数量较为有限,与此同时,工业相机采集到的图像一般尺寸较大.为此,我们采用随机裁剪的方法对图像样本进行增广,不仅能够增加缺陷样本的数量和多样性,还能够压缩模型规模,降低计算复杂性.同时,随机裁剪相当于建立局部特征与相应类别的权重关系,减弱背景(或噪

声)因子的权重,因此也能起到抑制噪声与增加模型稳定性的作用.随机裁剪过程如图 3 所示.

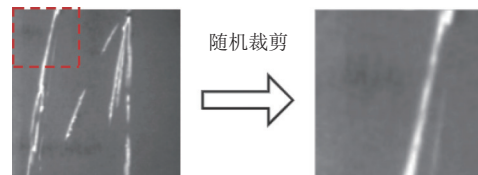


图 3 样本图像的随机裁剪

2.2 缺陷预分类

工业制品缺陷的样本图像首先输入深度森林模型进行预分类,预分类过程如图 4 所示.在多粒度扫描阶段,根据随机分割后的图像大小,设置了 3×3、6×6、9×9 三种滑动窗口,利用滑动窗口获取样本图像的特征子样本.特征子样本通过多个随机森林转换为特征的概率向量,将多个向量拼接作为级联森林的输入.

在级联森林阶段,扫描得到的缺陷特征经过每个级联层后的输出与原特征组合成扩展后的特征作为下一级联层的输入,每层的随机森林在分裂时随机选择 \sqrt{d} 个特征作为候选特征 (d 是输入特征的数量)^[16].该模型在每个级联层结束后进行分类性能的测试,然后继续生成下一级.当扩展一个新的级联层后,将整个模型的性能在验证集中进行测试,如果没有显著的性能增益,训练过程终止,确定级联层数.在最后一层上,级

联森林计算向量的平均值,输出具有最大概率那一类的标签,作为缺陷预分类结果。

经过深度森林预分类后,根据预测概率将样本图像分为容易识别和识别困难两种。识别困难的样本图

像将作为新的数据集,参与之后标注预测过程。为了提高后续人工标注的效率,以及考虑工厂实际对于产品质量控制的要求,设置一个可以根据初始预测结果及实际需要而变动的阈值,用以分离识别困难的样本图像。

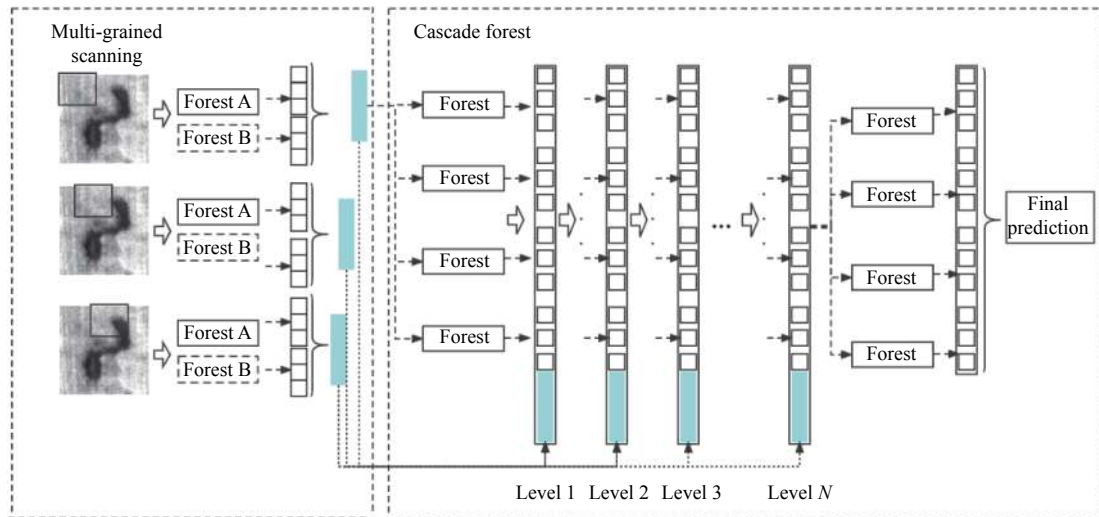


图4 缺陷预分类过程

2.3 人机协同纠错

经过缺陷预分类得到困难样本,人工标注部分识别困难的样本。根据识别困难的样本图像数量及实际产品质量要求,辅助控制样本标注比例以达到减少劳动量的目的。

标注后的样本图像作为 KNN 算法的训练集。在 KNN 算法中,如果一个样本在特征空间中的 K 个最相似(即特征空间中最邻近)的样本中的大多数属于某一个类别,则该样本也属于这个类别。人工标注的样本图像构成了特征空间的样本集,剩余识别困难的样本图像作为 KNN 算法的测试集,计算与测试集之间的欧氏距离。找到与测试集样本相距最近的 K 个训练集样本,选择出现次数最多的类别作为测试集样本的分类结果。

分离出的困难样本图像,经 KNN 算法得到新的预测结果,用该预测结果替换缺陷预分类结果。统计裁剪前同一图像中各个小图像的预测结果,根据多数投票算法 (majority voting) 将预测结果数量最多的类别作为该图像的最终缺陷分类结果。

3 实验

3.1 实验数据

我们使用两个常用的缺陷分类数据集以及从实际

生产线上采集的真实数据评估所提模型的有效性。两个公开的缺陷分类数据集分别为热轧钢带表面缺陷数据集 (NEU)、制品焊接缺陷数据集 (WELD)。实验使用的 3 个缺陷数据集的基本信息如表 1 所示。

表 1 数据集的基本信息

数据集	样本数	类别
NEU	1 800	6
WELD	140	4
CLS	170	4

NEU 是由东北大学发布的热轧钢带表面缺陷数据集,包含热轧钢带的 6 种典型表面缺陷,即轧制氧化皮 (Rs), 斑块 (Pa), 开裂 (Cr), 点蚀表面 (Ps), 内含物 (In) 和划痕 (Sc)。NEU 数据集样本图像如图 5 所示。为了评估算法在小样本条件下的分类效果,实验中我们在每类缺陷中随机选取了 30 个样本。制品焊接缺陷数据集 (WELD) 收集了 4 种焊接缺陷,包括条形缺陷 (Ba)、未焊透 (Pe)、未熔合 (Fu)、圆形缺陷 (Ro), 样本图像分辨率为 128 像素×128 像素。为了评估本文所提模型在真实数据上的有效性,我们使用如图 6 所示中原内配集团股份有限公司汽车气缸套生产线上实际采集的图像构建了气缸套缺陷分类数据集 (CLS)。该数据集包含气缸套的 4 种典型表面缺陷,即擦痕 (sc)、疤

疤 (bl)、裂缝 (fi)、砂眼 (blo), 每种类型的缺陷图像样本均为 50 个. 4 种类别缺陷样本的实例分别如图 7 所示, 样本分辨率为 100 像素×100 像素. 该数据集的难点在于同类缺陷在外观上存在较大差异, 而不同类别缺陷具有一定的相似性. 例如, 疤痕呈现圆形和条形等不同形状, 同时与擦痕又具有相似的视觉特征.

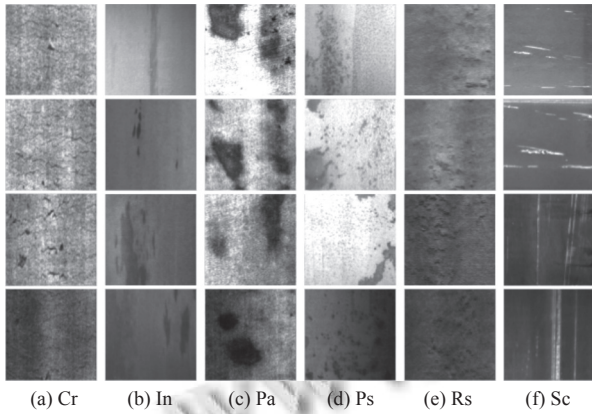


图 5 NEU 数据集样本图像



图 6 汽车气缸套生产线及数据采集系统

3.2 评价标准

实验采用准确率 (*Accuracy*)、精确度 (*Precision*)、召回率 (*Recall*) 这 3 个分类性能评价指标. 计算如下:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

其中, *TP*, *FP*, *FN*, *TN* 分别表示预测为正的样本、预测为正的负样本、预测为负的正样本、预测为负的负样本.

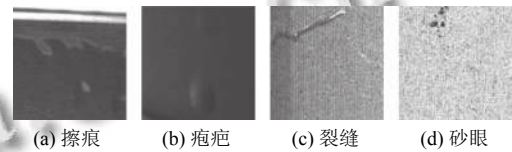


图 7 4 种缺陷样本图像

3.3 实验结果与分析

3.3.1 分类方法比较

实验使用具有代表性的 SVM、深度卷积神经网络以及多粒度级联森林 gcForest 作为基本方法, 与本文提出的方法进行比较. 其中, 深度卷积神经网络采用具有密集连接的 DenseNet 模型^[17]. NEU、WELD、CLS 三个数据集的可信度阈值分别设置为 0.8、0.5、1, 人工标注比例为 0.25. 表 2 是 4 个数据集在几种分类模型上的 *Accuracy* 值、*Precision* 值和 *Recall* 值. 通过这些指标的对比, 可以看出本文方法相较于传统方法 SVM、gcForest 及 CNN 具有更高的准确性, 分类效果较好.

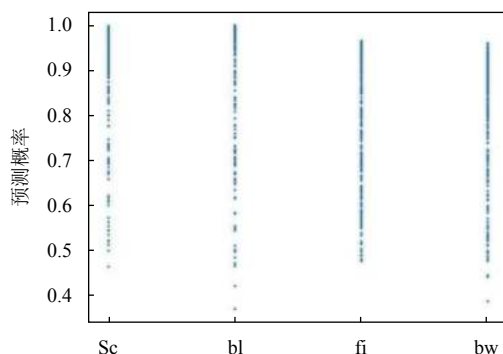
表 2 各方法的指标统计结果 (%)

方法	<i>Accuracy</i>			<i>Precision</i>			<i>Recall</i>		
	CLS	NEU	WELD	CLS	NEU	WELD	CLS	NEU	WELD
DenseNet	87.8	70.0	67.1	85.9	67.8	60.4	84.9	70.0	59.6
gcForest	65.9	68.3	74.3	75.8	76.2	65.4	63.3	70.0	74.3
SVM	64.7	66.7	71.4	63.9	83.5	59.8	68.8	68.1	50.1
Ours	90.1	91.7	96.7	91.8	94.4	96.7	90.1	91.7	96.7

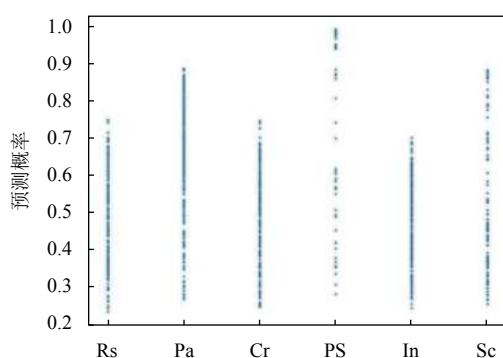
3.3.2 参数敏感性分析

样本图像经过深度森林识别后, 得到预测的每个类别的概率值, 概率值最大的类别即为初始预测结果. 根据最大概率值的分布情况, 可以初步判定预测效果. 如图 8 所示两个数据集的最大概率值的分布, 图 8(a) 中数据集的预测效果要明显好于图 8(b). 技术人员可以根据初始预测效果, 控制参数可行性阈值的变动.

本文所提缺陷分类方法的准确率主要受可行度阈值和标注比例两个参数影响, 两个参数变动对准确率的影响如图 9 所示. 从图 9 中可以看出, 初始预测结果好的数据集, 受标注比例影响较小, 受阈值影响较大; 初始预测结果差的数据集, 受标注比例影响较大, 受阈值影响较小. 基于这种现象, 技术人员在实际生产中可以根据初始预测结果确定最佳的标注比例.

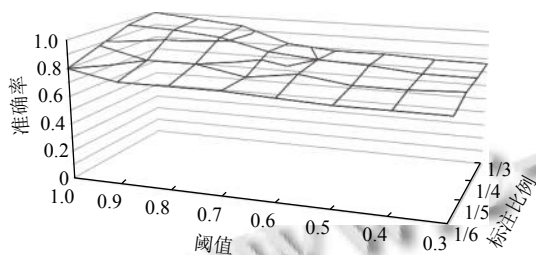


(a) CLS 数据集

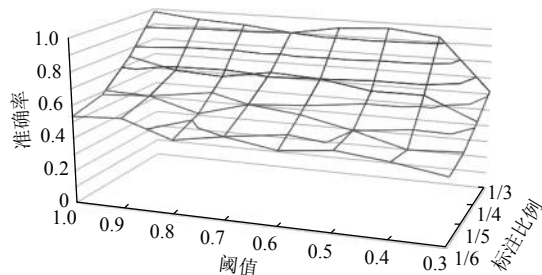


(b) NEU 数据集

图8 最大概率值的分布



(a) NEU 数据集



(b) WELD 数据集

图9 准确率的变动状况

4 总结

针对目前应用比较广泛的一些分类模型算法对工业缺陷小样本数据分类时的不足,本研究提出基于深度森林的人机协同模型.本研究使用了3个小样本数据集,将这些数据集在4种模型下实验,通过 *Accuracy*、*Precision*、*Recall* 三种指标对比模型在此数据集上的分类性能.经过验证,改进后的模型具有较好的分类性能.

未来将继续提高模型的识别性能,并考虑模型的实际应用.对模型资源要求高的缺点进行优化,减小模型算法的计算消耗,提高稳定性.

参考文献

- 1 He Y, Song KC, Meng QG, *et al.* 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. [doi: [10.1109/TIM.2019.2915404](https://doi.org/10.1109/TIM.2019.2915404)]
- 2 Dong HW, Song KC, He Y, *et al.* PGA-net: Pyramid feature fusion and global context attention network for automated surface defect detection. *IEEE Transactions on Industrial Informatics*, 2020, 16(12): 7448–7458. [doi: [10.1109/TII.2019.2958826](https://doi.org/10.1109/TII.2019.2958826)]
- 3 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](https://doi.org/10.16383/j.aas.c190811)]
- 4 谢芬. 基于机器视觉的表面缺陷检测技术研究 [硕士学位论文]. 南京: 南京航空航天大学, 2019. [doi: [10.27239/d.cnki.gnhhu.2019.001376](https://doi.org/10.27239/d.cnki.gnhhu.2019.001376)]
- 5 Hanzaei SH, Afshar A, Barazandeh F. Automatic detection and classification of the ceramic tiles' surface defects. *Pattern Recognition*, 2017, 66: 174–189. [doi: [10.1016/j.patcog.2016.11.021](https://doi.org/10.1016/j.patcog.2016.11.021)]
- 6 Sun J, Li C, Wu XJ, *et al.* An effective method of weld defect detection and classification based on machine vision. *IEEE Transactions on Industrial Informatics*, 2019, 15(12): 6322–6333. [doi: [10.1109/TII.2019.2896357](https://doi.org/10.1109/TII.2019.2896357)]
- 7 Pasadas DJ, Ramos HG, Feng B, *et al.* Defect classification with SVM and wideband excitation in multilayer aluminum plates. *IEEE Transactions on Instrumentation and Measurement*, 2020, 69(1): 241–248. [doi: [10.1109/TIM.2019.2893009](https://doi.org/10.1109/TIM.2019.2893009)]
- 8 Li XK, Guo YC, Li YM. Particle swarm optimization-based SVM for classification of cable surface defects of the cable-

- stayed bridges. *IEEE Access*, 2020, 8: 44485–44492. [doi: [10.1109/ACCESS.2019.2961755](https://doi.org/10.1109/ACCESS.2019.2961755)]
- 9 Masci J, Meier U, Ciresan D, *et al.* Steel defect classification with max-pooling convolutional neural networks. *The 2012 International Joint Conference on Neural Networks (IJCNN)*. Brisbane: IEEE, 2012. 1–6. [doi: [10.1109/IJCNN.2012.6252468](https://doi.org/10.1109/IJCNN.2012.6252468)]
- 10 Cheon S, Lee H, Kim CO, *et al.* Convolutional neural network for wafer surface defect classification and the detection of unknown defect class. *IEEE Transactions on Semiconductor Manufacturing*, 2019, 32(2): 163–170. [doi: [10.1109/TSM.2019.2902657](https://doi.org/10.1109/TSM.2019.2902657)]
- 11 张光建, 朱婵. 卷积神经网络自动分类手机外壳划痕. *计算机系统应用*, 2020, 29(11): 271–275. [doi: [10.15888/j.cnki.csa.007674](https://doi.org/10.15888/j.cnki.csa.007674)]
- 12 Saqlain M, Abbas Q, Lee JY. A deep convolutional neural network for wafer defect identification on an imbalanced dataset in semiconductor manufacturing processes. *IEEE Transactions on Semiconductor Manufacturing*, 2020, 33(3): 436–444. [doi: [10.1109/TSM.2020.2994357](https://doi.org/10.1109/TSM.2020.2994357)]
- 13 Deitsch S, Christlein V, Berger S, *et al.* Automatic classification of defective photovoltaic module cells in electroluminescence images. *Solar Energy*, 2019, 185: 455–468. [doi: [10.1016/j.solener.2019.02.067](https://doi.org/10.1016/j.solener.2019.02.067)]
- 14 Aydin I, Akin E, Karakose M. Defect classification based on deep features for railway tracks in sustainable transportation. *Applied Soft Computing*, 2021, 111: 107706. [doi: [10.1016/J.ASOC.2021.107706](https://doi.org/10.1016/J.ASOC.2021.107706)]
- 15 Sassi P, Tripicchio P, Avizzano CA. A smart monitoring system for automatic welding defect detection. *IEEE Transactions on Industrial Electronics*, 2019, 66(12): 9641–9650. [doi: [10.1109/TIE.2019.2896165](https://doi.org/10.1109/TIE.2019.2896165)]
- 16 Zhou ZH, Feng J. Deep forest. *National Science Review*, 2019, 6(1): 74–86. [doi: [10.1093/nsr/nwy108](https://doi.org/10.1093/nsr/nwy108)]
- 17 Huang G, Liu Z, van der Maaten L, *et al.* Densely connected convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 2261–2269.

(校对责编: 孙君艳)