

图像异常检测研究现状综述

吕承侃^{1,2} 沈飞^{1,2,3} 张正涛^{1,2,3} 张峰^{1,2,3}

摘要 图像异常检测是计算机视觉领域的一个热门研究课题,其目标是在不使用真实异常样本的情况下,利用现有的正常样本构建模型以检测可能出现的各种异常图像,在工业外观缺陷检测、医学图像分析、高光谱图像处理等领域有较高的研究意义和应用价值.本文首先介绍了异常的定义以及常见的异常类型.然后,本文根据在模型构建过程中有无神经网络的参与,将图像异常检测方法分为基于传统方法和基于深度学习两大类型,并分别对相应的检测方法的设计思路、优点和局限性进行了综述与分析.其次,梳理了图像异常检测任务中面临的主要挑战.最后,对该领域未来可能的研究方向进行了展望.

关键词 图像异常检测, 计算机视觉, 深度学习, 神经网络, 背景重构

引用格式 吕承侃, 沈飞, 张正涛, 张峰. 图像异常检测研究现状综述. 自动化学报, 2022, 48(6): 1402–1428

DOI 10.16383/j.aas.c200956

Review of Image Anomaly Detection

LV Cheng-Kan^{1,2} SHEN Fei^{1,2,3} ZHANG Zheng-Tao^{1,2,3} ZHANG Feng^{1,2,3}

Abstract Image anomaly detection is a hot research topic in the field of computer vision, which aims at training a model by normal images to detect potential anomalous images without requiring any real anomalous images for training. Therefore, it has high research and application value in the fields of industrial surface defect detection, medical image analysis, and hyperspectral image processing. This paper firstly introduces the definition of anomaly and its common type. Secondly, image anomaly detection methods are divided into traditional and deep learning-based methods according to whether the neural network is involved. Meanwhile, the design ideas, advantages and limitations of these methods are summarized and analyzed respectively. Thirdly, the major challenges faced in image anomaly detection are introduced. Finally, the future directions of image anomaly detection are discussed.

Key words Image anomaly detection, computer vision, deep learning, neural network, background reconstruction

Citation Lv Cheng-Kan, Shen Fei, Zhang Zheng-Tao, Zhang Feng. Review of image anomaly detection. *Acta Automatica Sinica*, 2022, 48(6): 1402–1428

异常检测是机器学习领域中一项重要的研究内容.它是一种利用无标注样本或者正常样本构建检测模型^[1],检测与期望模式存在差异的异常样本的方法.异常检测在各种领域中都有广泛的应用,如网络入侵检测、信号处理、工业大数据分析、异常行为检测和图像与视频处理等.

早期的异常检测算法大多应用于数据挖掘领域,而近年来随着计算机视觉和深度学习等相关技术的发展,许多相关工作将异常检测引入到图像处

理领域来解决样本匮乏情况下的目标检测问题.

传统的目标检测算法中很大一部分方法属于监督学习的范畴,即需要收集足够的目标类别样本并进行精确的标注,比如图像的种类、图像中目标的位置以及每一个像素点的类别信息等^[2-3].然而,在许多应用场景下,很难收集到足够数量的样本.例如,在表面缺陷检测任务当中,实际收集到的图像大部分属于正常的无缺陷样本,仅有少部分属于缺陷样本,而需要检测的缺陷类型又十分多样,这就使得可供训练的缺陷样本的数量十分有限^[4].又比如在安检任务当中,不断会有新的异常物品出现^[5].而对于医学图像中病变区域的识别任务,不仅带有病变区域的样本十分稀少,对样本进行手工标注也十分耗时^[6].在这些情况下,由于目标类别样本的缺乏,传统的目标检测和图像分割的方法已不再适用.

而异常检测无需任何标注样本就能构建检测模型的特点,使得其十分适用于上述几种情况^[7].在图像异常检测当中,收集正常图像的难度要远低于收集异常图像的难度,能显著减少检测算法在实际应

收稿日期 2020-11-18 录用日期 2021-06-25

Manuscript received November 18, 2020; accepted June 25, 2021

中国科学院青年创新促进会 (2020139) 资助

Supported by Youth Innovation Promotion Association, Chinese Academy of Sciences (2020139)

本文责任编辑 白翔

Recommended by Associate Editor BAI Xiang

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 100049 3. 中科慧远视觉技术(洛阳)有限公司 洛阳 471000

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 100049 3. CASI Vision Technology CO., LTD., Luoyang 471000

用中的时间和人力成本. 而且, 在异常检测中模型是通过分析与正常样本之间的差异来检测异常样本, 这使得异常检测算法对各种类型甚至是全新的异常样本都具有检测能力. 虽然标注样本的缺失给图像异常检测带来了许多问题和挑战, 不过由于上述各种优点, 如表 1 所示, 已经有许多方法将图像异常检测应用在各种领域中. 因此, 图像异常检测问题具有较高的研究价值和实际应用价值.

随着对异常检测研究的深入, 大量研究成果不断涌现, 也有许多学者开展了一些综述性工作. 如 Ehret 等^[20] 根据不同的图像背景, 对大量图像异常检测方法进行了综述, 不过对基于深度学习的方法还缺乏一定的梳理. Pang 等^[21] 和 Chalapathy 等^[22] 则是从更为广阔的角度对基于深度学习的异常检测方法进行了梳理, 不过由于数据类型的多样性, 这些工作对异常检测在图像中的应用还缺乏针对性. 陶显等^[23] 对异常检测在工业外观缺陷检测中的应用进行了一些总结, 不过重心落在有监督的检测任务上, 对无监督的异常检测方法欠缺一定的整理和归纳. 而本文则针对无监督的图像异常检测任务, 以工业、医学和高光谱图像作为具体应用领域, 对传统和基于深度学习的两大类方法进行梳理. 上述 3 种应用领域都有相同的特点即可使用的带标注异常样本数量稀少, 因此有许多工作针对这几个领域内的异常目标检测问题开展了研究. 本文整体结构安排如下: 第 1 节将介绍异常的定义以及常见的形态. 第 2 节根据模型构建过程中有无神经网络的参

与, 将现有的图像异常检测算法分为传统方法和基于深度学习两大类并分别进行综述与分析. 第 3 节将介绍图像异常检测中常用的数据集. 第 4 节将介绍在图像异常检测当中面临的主要挑战. 第 5 节将综合图像异常检测的研究现状对未来可能的发展方向进行展望. 最后第 6 节将对本文内容进行总结.

1 异常检测的定义

1.1 异常的定义及类型

异常, 又被称为离群值, 是一个在数据挖掘领域中常见的概念^[24], 已经有不少的工作尝试对异常数据进行定义^[25-26]. Hawkins 等^[25] 将异常定义为与其余观测结果完全不同, 以至于怀疑其是由不同机制产生的观测值. 一般情况下, 会将常见的异常样本分为 3 个类别^[1]: 点异常、上下文异常和集群异常.

点异常一般表现为某些严重偏离正常数据分布范围的观测值, 如图 1(a) 所示的二维数据点, 其中偏离了正常样本点的分布区域 (N_1 , N_2) 的点 (O_1 , O_2 和 O_3) 即为异常点.

上下文异常则表现为该次观测值虽然在正常数据分布范围内, 但联合周围数据一起分析就会表现出显著的异常. 如图 1(b) 所示, t_2 点处的温度值虽然依然在正常范围内, 但联合前后两个月的数据就能发现该点属于异常数据.

而集群异常, 又称为模式异常, 是由一系列观测结果聚合而成并且与正常数据存在差异的异常类

表 1 图像异常检测的应用领域

Table 1 Applications of image anomaly detection

应用领域	具体应用
缺陷检测	各种产品表面缺陷检测, 包括布匹 ^[8] 、玻璃 ^[9] 、钢板 ^[10] 、水泥 ^[11] 等纹理表面以及印制电路板 ^[12] 、酒瓶 ^[13] 等物体表面缺陷的检测
医学影像分析	在核磁共振图像 ^[14] 、虹膜图像 ^[15] 、眼底视网膜图像 ^[16] 等医学图像中检测可能的病变区域
高光谱图像处理	海面船舶检测 ^[17] 、地面异常区域检测 ^[18] 、机场飞机定位 ^[19] 等

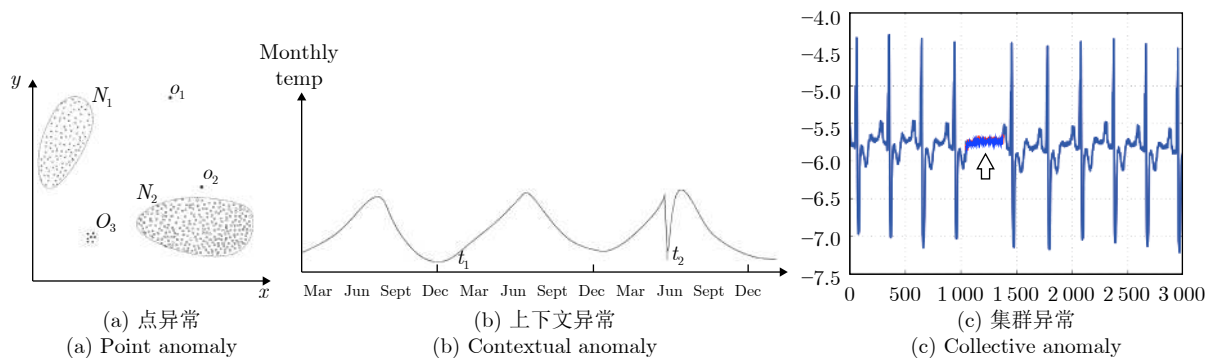


图 1 异常的类型^[1]

Fig.1 The type of anomaly^[1]

型. 该类异常中, 可能单独看其中任意一个点都不属于异常, 但是当一系列点一起出现时就属于异常, 如图 1(c) 箭头所指区域内单独看每一个点的值都在正常范围内, 但这些点聚合在一起就形成了与正常信号模式完全不同的结构.

1.2 图像数据中的异常

图像数据中每一个像素点上的像素值就对应着一个观测结果. 由于图像内像素值的多样性, 仅仅分析某一个点的像素值很难判断其是否属于异常. 所以在大部分图像异常检测任务中, 需要联合分析图像背景以及周围像素信息来进行分类, 检测的异常也大多属于上下文或者模式异常. 当然, 这 3 种异常类型之间并没有非常严格的界限. 例如, 有一部分方法就提取图像的各类特征^[27], 并将其与正常图像的特征进行比较以判断是否属于异常, 这就将原始图像空间内模式异常的检测转换到了特征空间内点异常的检测.

图像异常检测任务根据异常的形态可以分为定性异常的分类和定量异常的定位两个类别.

定性异常的分类, 类似于传统图像识别任务中的图像分类任务, 即整体地给出是否异常的判断, 无需准确定位异常的位置. 如图 2 左上图所示, 左侧代表正常图像, 右侧代表异常图像, 在第 1 行中, 模型仅使用服饰数据集 (Fashion mixed national institute of standards and technology database, Fashion-MNIST)^[28] 中衣服类型的样本进行训练, 则其他类别的样本图像 (鞋子等) 对模型来说都是需要检测的异常样本, 因为他们在纹理、结构和语

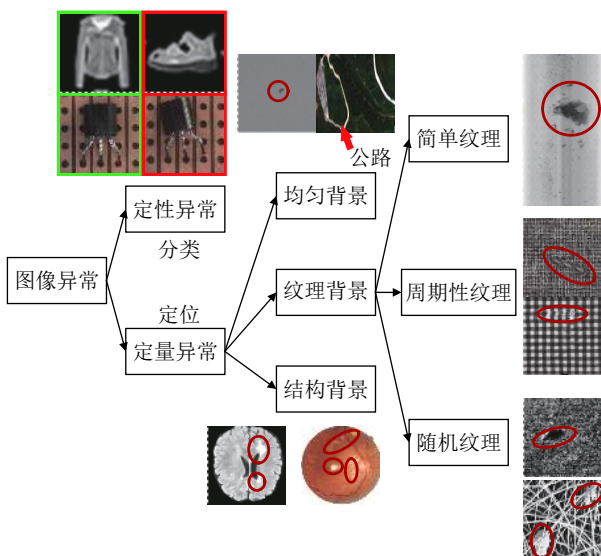


图 2 图像异常分类图

Fig. 2 The classification of image anomalies

义信息等方面都不相同. 又或者如第 2 行所示, 异常图像中的三极管与正常图像之间只是出现了整体的偏移, 而三极管表面并不存在任何局部的异常区域, 难以准确地定义出现异常的位置, 更适合整体地进行异常与否的分类.

而定量异常的定位则类似于目标检测或者图像分割任务, 需要得到异常区域的位置信息. 在这种类型的图像异常检测中, 测试图像中只有一小部分区域出现了异常模式. 而异常定位任务根据具体的图像背景又可分为以下几类, 如图 2 所示, 其中被圈出部分为异常区域.

1) 均匀背景

均匀背景代指一些内容较为单一的场景, 如图 2 中上图所示磨砂玻璃表面对局部缺陷的定位, 或者深色山区图像中对盘山公路的定位. 这一类背景下的异常检测属于相对简单的检测任务.

2) 纹理背景

纹理背景主要出现在工业生产领域中, 根据纹理形态又可以分为简单纹理、周期性纹理和随机纹理 3 种. 其中, 简单纹理代指因光照和材质反光等因素影响, 在原本均匀的物体表面产生的一些非均匀的纹理背景, 如图 2 右上所示的钢板表面图像. 而周期性纹理则代指各类由大量重复单元组成的具有显著周期性的图像, 最具代表性的就是图 2 所示的各类布匹图像. 而随机纹理则代指一些由无规则结构组成的图像背景, 如图 2 右下图所示的声呐和纳米材料图像.

3) 结构背景

结构背景则是一类更为广泛的图像背景, 一般具有结构复杂、个体差异大和语义信息丰富等特点, 需要整体进行分析而无法仅依靠局部信息进行异常检测, 如图 2 左下图所示的各种医学图像. 这类图像背景下的异常检测问题是相对较难的一类检测任务.

2 图像异常检测技术研究现状

一般情况下图像异常检测的目标是通过无监督或者半监督学习的方式, 检测与正常图像不同的异常图像或者局部异常区域. 近年来传统机器学习方法已经在图像异常检测领域有了较多的应用, 而随着深度学习技术的发展, 越来越多的方法尝试结合神经网络来实现图像异常检测. 根据在模型构建阶段有无神经网络的参与, 现有的图像异常检测方法可以分为基于传统方法和基于深度学习的方法两大类. 如图 3 所示, 基于传统方法的异常检测技术大致包含 6 个类别: 基于模板匹配、基于统计模型、

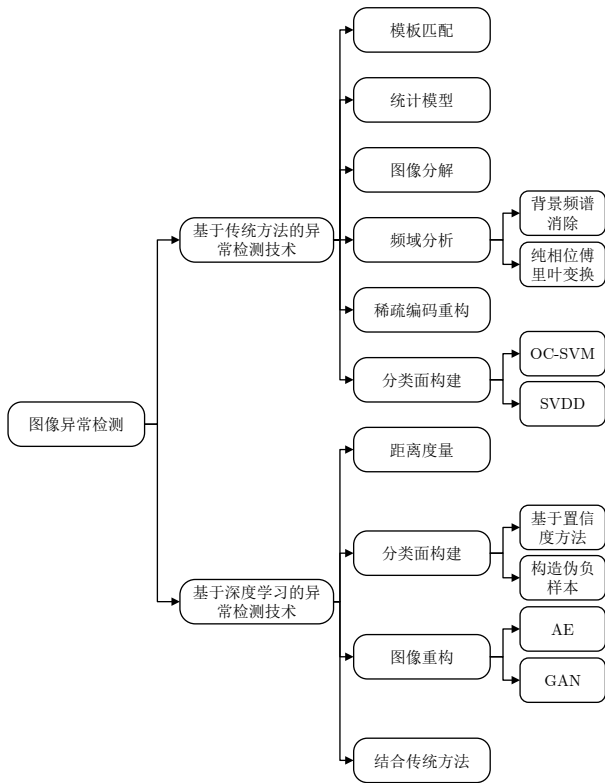


图 3 图像异常检测技术的分类图

Fig. 3 The classification of image anomaly detection methods

基于图像分解、基于频域分析、基于稀疏编码重构和基于分类面构建的异常检测方法。而基于深度学习的方法大致包含 4 个类别：基于距离度量、基于分类面构建、基于图像重构和结合传统方法的异常检测方法。

2.1 基于传统方法的异常检测技术

本文根据检测原理将传统图像异常检测方法分类为以下类别：基于模板匹配、基于统计模型、基于图像分解、基于频域分析、基于稀疏编码重构和基于分类面构建的异常检测方法。传统的图像异常检测算法大多会学习一个模型来描述正常图像，随后在检测阶段根据待检图像与现有模型之间的匹配程度来进行异常检测。

2.1.1 基于模板匹配的异常检测方法

在图像异常检测任务中，最理想的情况是所有的正常图像都高度相似，且异常图像与正常图像之间只会一小部分区域出现区别。此时，模板匹配是非常有效的一类异常检测方法。得到待测图像和模板图像之间的对应关系后，比较两者之间的差异即可实现异常检测。

图像匹配作为一种计算机视觉中的常见任务，

已经得到了充分的发展，根据相关文献的整理^[29]，现有的方法可以大致分为点特征匹配、线匹配和区域匹配 3 大类别。点特征匹配中，常用有诸如 Harris 特征^[30]、尺度不变特征变换 (Scale invariant feature transform, SIFT)^[31] 和仿射尺度不变特征变换 (Affine scale invariant feature transform, ASIFT)^[32] 等人工设计的特征提取方法，也有结合机器学习进行特征点提取的改良的加速分割测试特征 (Features from accelerated segment test-enhanced repeatability, FAST-ER)^[33] 和时间不变学习检测器 (Temporally invariant learned detector, TILDE)^[34] 等算法。不过由于点特征提取过程中往往存在大量的外点，为了估计准确的几何参数还需要额外的筛选流程去除外点。线匹配中，常用的有均值-标准差线描述符 (Mean-standard deviation line descriptor, MSLD)^[35] 和条带描述符 (Line band descriptor, LBD)^[36] 等通过提取直线特征进行匹配的方法，也有 Xia 等^[37] 和 Wolfson 等^[38] 采用的曲线匹配方法。线特征相比于点特征包含更多的场景信息和结构信息，但是也会面临边缘特征不明显和线段断裂等问题。而第 3 类区域匹配的方法包含最为完整的图像信息，相比于另外两种方法具有更好的鲁棒性。早期的方法包括归一化积相关^[39] 和平均绝对差^[40] 等算法，近期较为常用的还包括 Korman 等^[41] 提出的快速匹配算法 (Fast affine template matching, FAST-Match) 和 Jia 等^[42] 提出的适用于彩色图像匹配的彩色快速匹配算法 (Fast affine template matching for colour images, CFAST-Match)。而近年来有一些方法利用深度学习进行区域匹配。如 Han 等^[43] 通过对图像区域进行特征提取和相似性度量以实现图像匹配。也有方法不进行相似性度量，如 Balntas 等^[44] 利用三元损失对网络进行特征提取的训练，相比于传统的特征描述符有更高的提取效率和匹配效果。

而在图像异常检测领域，有许多相关方法利用图像匹配技术进行异常检测。如 Chen 等^[45] 提取了待测图像和模板图像的 Hu 矩作为特征对三维打印零件进行异常与否的分类。Vaikundam 等^[46] 先提取了图像的 SIFT 特征，利用聚类算法进行描述符筛选之后，找到最为匹配的正常图像作为模板来进行异常区域的定位。Herwig 等^[47] 则是通过中值滤波创建模板图像来检测钢材表面的异常区域。考虑到空域模板匹配方法容易受到诸如光照变化和正常图像间微小差异的影响，Tsai 等^[48] 将模板匹配过程迁移到了频域，通过比较待检图像和模板图像经傅里叶变换后的频域分量来定位印制电路板 (Printed circuit board, PCB) 上细微的缺陷。

但模板匹配的方法一般适用于图像采集环境稳定且可控的场景,如图4所示的PCB板表面的异常检测,虽然结构复杂但内容基本保持不变.而更多的情况下,即便是正常图像之间都会存在着较多的差异,难以通过模板匹配实现异常检测.

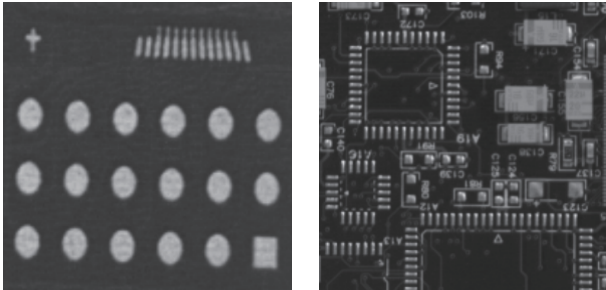


图4 模板匹配的适用场景^[48]

Fig.4 Applicable scenes of template matching^[48]

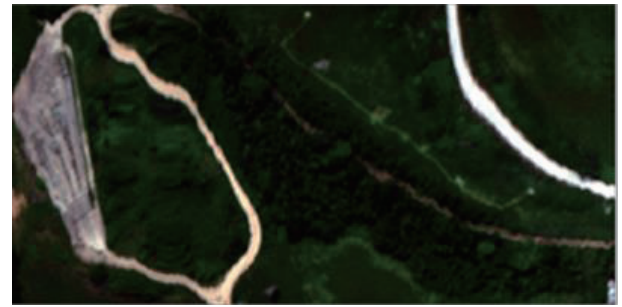
2.1.2 基于统计模型的异常检测方法

这一类方法通常是利用统计模型来描述正常图像中像素值或者特征向量的分布情况,而对于一些远离该分布的图像区域则认定为存在异常.

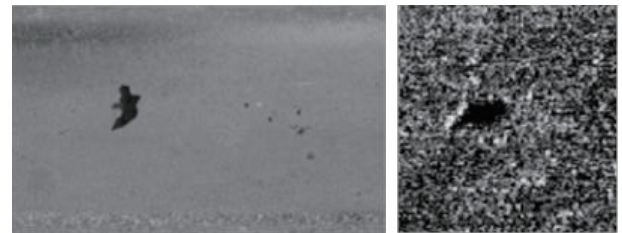
其中较为常见的方式是利用高斯模型进行描述,比如Reed等^[49]提出的RX(Reed-xiao)算法就利用高斯分布函数来描述高光谱图像中像素点内信息的分布情况^[50].然而一个高斯分布模型很难描述更为复杂多变的场景,为了进一步地提升建模效果,Veracini等^[51]通过高斯混合模型对高光谱图像的背景像素进行描述,并通过期望最大化算法来自适应地确定高斯混合模型中子分布模型的数量.Zhang等^[52]则是利用马尔科夫随机场提升了高斯混合模型的鲁棒性,在钢板表面获得了较好的异常定位效果.

上述方法大多仅对像素点内的信息进行了建模而忽略了区域信息,为了让模型对于孤立的噪声像素点有足够的鲁棒性,有方法同时考虑了像素点的邻域信息来进行背景建模.Goldman等^[27]首先对局部图像片提取特征,然后再通过构造高斯模型的方式检测异常区域.

此类方法为了估计模型的参数需要一定数量的正常样本,而且这些方法都预先对图像数据的分布做了假设,这在一定程度上降低了该方法的通用性.基于统计模型的方法较为适合图5(a)所示的高光谱图像这种一个像素点包含大量信息的特殊图像类型,因此在高光谱图像中的异常检测问题上衍生出了许多基于RX算法的改进方法,如子空间RX算法(Subspace RX, SSRX)^[53]和局部RX算法(Local RX, LRX)^[54]等.而对于普通图像类型而言,统计



(a) 山区公路图像
(a) Mountain road image



(b) 钢板图像
(b) Steel image
(c) 声呐图像
(c) Sonar image

图5 统计模型的适用场景^[27, 51-52]

Fig.5 Applicable scenes of statistical model^[27, 51-52]

模型的方法在一些背景较为简单的图像中有较好的检测效果,如图5(b)所示的钢板表面图像,而结合特征提取算法也能应用在一些具有一定纹理背景的图像中,如图5(c)所示的声呐图像.但是结构更为复杂的图像往往难以预先假设其数据分布,在模型参数的估计上也有较高的难度,无法保证检测效果^[55].

2.1.3 基于图像分解的异常检测方法

基于图像分解的方法大多针对的是周期性纹理表面小面积异常区域的检测任务.由于异常区域一般是随机出现的,其周期性较弱,这一特点使其可以与周期性的背景纹理进行区分.

较为常用的方法主要利用了周期性背景纹理低秩性的特点,采用低秩分解将原始待检测图像分解成为代表背景的低秩矩阵和代表着异常区域的稀疏矩阵^[8]:

$$\min_{L, S} \|L\|_* + \lambda \|S\|_1, \text{ s.t. } F = L + S \quad (1)$$

其中, F 代表原始图像矩阵, L 和 S 分别代表着低秩矩阵和稀疏矩阵. $\|\cdot\|_*$ 代表着矩阵的秩, $\|\cdot\|_1$ 代表 L1 范数来近似描述矩阵的稀疏度, λ 是一个人工变量.不过,直接对原始图像进行低秩分解容易受到正常背景变化的影响,导致其定位精度较低.因此,一般会先对原始图像进行特征提取,然后在特征向量构成的矩阵上进行低秩分解以区分正常和异常区域.

其大致流程如图6所示,图中分解得到的稀疏

矩阵 S_i 就对应着可能存在异常的区域. Li 等^[8] 将原始图像切分成不同图像区域, 并利用 Gabor 变换与方向梯度直方图 (Histogram of oriented gradient, HOG) 提取图像区域的特征, 再利用低秩分解来寻找对应着异常区域的图像块. 杨恩君等^[56] 则结合布匹图像的基元特征和 HOG 特征作为辅助信息来进一步优化低秩分解的效果. Li 等^[57] 联合多通道特征提取和低秩分解算法来进行布匹表面缺陷的检测. Wang 等^[10] 在多特征融合的基础上, 结合背景先验知识挖掘和超像素分割的方法, 在钢材表面缺陷分割任务中获得了更为细致的分割结果.

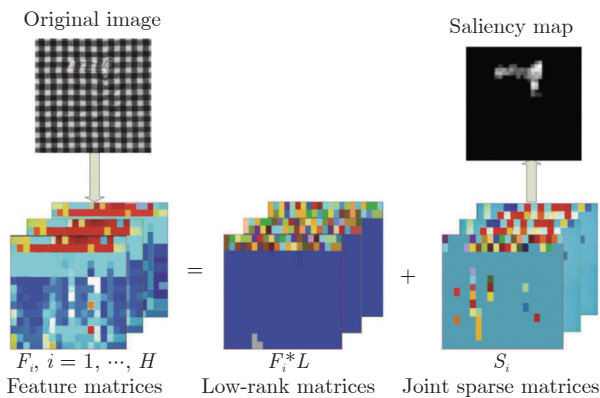


图 6 基于低秩分解的图像异常检测示意图^[57]

Fig. 6 Illustration of anomaly detection based on low-rank decomposition^[57]

但是, 原始的低秩分解方法只是将原始图像分解成了低秩矩阵和稀疏矩阵两部分, 这一过程没有考虑可能出现的加性噪声, 这些噪声同样被保留在了分解出的稀疏矩阵中, 影响后续异常定位过程. 因此, Zhou 等^[58] 提出的分解算法 (Go decomposition, GoDec) 将原始矩阵分解成低秩矩阵、稀疏矩阵和噪声矩阵 3 个部分, 相比于原始的低秩分解方法有更强的鲁棒性. Zhang 等^[59] 在 GoDec 的基础上, 充分利用了背景区域的先验知识, 结合马氏距离来检测高光图像中的异常区域. Wang 等^[60] 在分解布匹图像的过程中考虑了异常区域的局部连续性以保证缺陷形态的完整, 还提出了一种积分机制来进一步提升稀疏矩阵中缺陷的完整性并减少噪声的残留. 然而, 在稀疏矩阵中也存在着部分背景区域, 同样也会影响检测的过程. 针对这一问题, Yang 等^[61] 结合正交子空间投影和自适应权重选取算法来抑制稀疏矩阵中的背景噪声.

基于低秩分解的异常检测算法, 其优点在于完全不需要任何训练样本, 直接可以在待检图像上进行异常区域的检测. 不过这一类方法由于涉及到优化过程, 计算量较大, 其检测速度相对较慢, 难以进

行实时检测. 由于其对背景图像的低秩性假设, 这一类方法适用于如图 7(a) 所示的布匹^[8] 等各种规则纹理表面的缺陷检测问题, 同时, 低秩分解在图 7(b) 所示的更为复杂的高光谱图像的异常检测中也有应用^[59, 61], 不过需要将原始 $h \times w \times b$ 的三维图像数据重新排列成 $(h \times w) \times b$ 的二维矩阵再进行后续的分析, 其中 b 代指频段数. 但此类方法对于无周期性或者纹理更加复杂的普通图像并不适用.

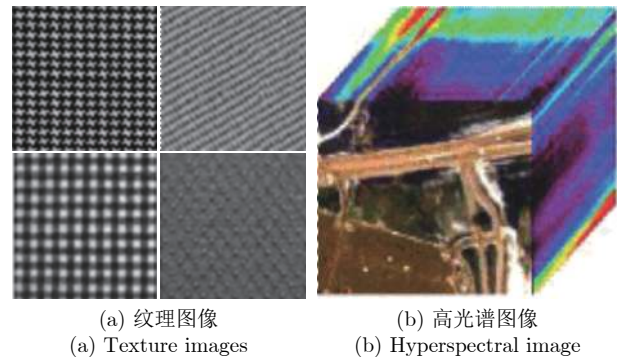


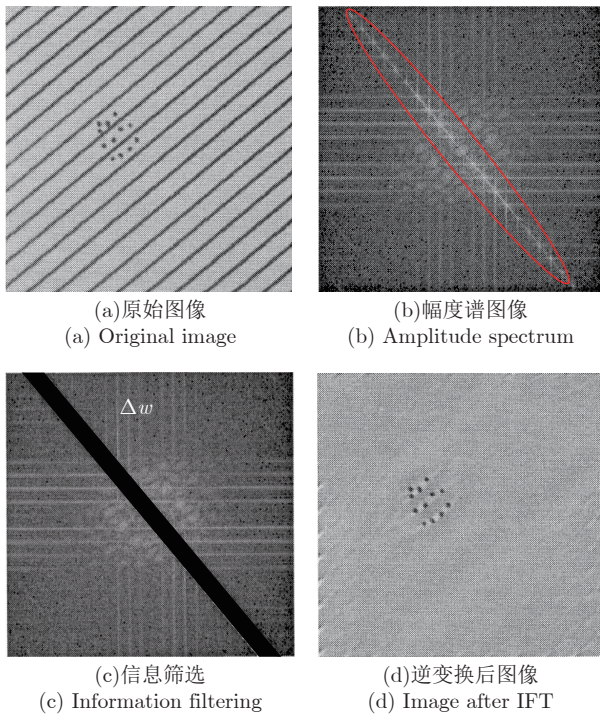
图 7 图像分解的适用场景
Fig. 7 Applicable scenes of image decomposition

2.1.4 基于频域分析的异常检测方法

基于频域分析的方法主要针对的也是规则纹理表面异常区域的检测. 不过这一类方法主要是对图像的频谱信息进行编辑, 尝试消除周期性背景纹理以凸显异常区域. 其中常用的方法包含背景频谱消除和纯相位傅里叶变换法 (Phase only fourier transform, POFT)^[62] 两类, 前者通过消除背景的频谱信息来凸显异常区域, 而后者则尝试在逆傅里叶变换时仅利用相位谱以消除重复背景.

1) 早期的方法如图 8 所示, 在检测过程中首先利用傅里叶变换 (Fourier transform, FT)^[63] 将原始图像转换到频域空间, 在幅度谱中将对应着周期性背景纹理的频谱分量去除之后, 通过逆傅里叶变换 (Inverse fourier transform, IFT) 来得到异常区域的位置信息. 可以看到, 图 8(a) 中的斜向纹理在经过傅里叶变换之后会在图 8(b) 中被圈出的部分得到较高的频谱分量, 按照图 8(c) 的方式将这些区域消除后再进行逆变换, 从图 8(d) 中可以看到原始的纹理背景已经被消除, 而属于缺陷的黑色圆点依旧被保留了下来.

Liu 等^[64] 将待检图像与正常图像的频谱作差来进行异常检测, 不过该方法更适合背景较为固定的场景, 当目标出现旋转等变换时其性能就会受到严重的影响. Tsai 等^[65-66] 通过删除幅度谱中某些特定的信息来进行纹理图像表面的异常检测. 不过该方法只适合于一些均匀或者有方向性的纹理图像, 因

图 8 通过背景频谱消除来进行异常检测^[66]Fig.8 Anomaly detection based on subtraction of background spectral^[66]

为这类图像其背景对应的频谱分量比较容易确定. 但对于较为复杂的图像, 背景区域频谱分量的分布更为复杂, 无法直接利用滤波^[65]等操作来去除. 后续的方法尝试用更加自适应的方式来进行异常检测, 比如 Zhang 等^[63]就结合谱减法来尝试检测玻璃表面的缺陷. 谱减法^[67]首先使用特定的滤波器处理幅度谱, 随后计算其与原始幅度谱之间的差值, 并以该差值进行逆傅里叶变换来得到图像中异常区域. Li 等^[68]在谱减法的基础上, 利用重构后图像的熵来自适应地选择最优滤波器, 以此来实现对不同大小目标的检测. 不过, 谱减法的检测性能与使用的滤波器核有较强的相关性^[68], 需要在模型设计阶段投入大量的精力来设计合适的滤波器核.

2) 第二类比较常用的方法, 则是采用 POFT 来消除背景区域. POFT 在对图像进行傅里叶变换之后, 抛弃了幅度谱的信息而仅采用其相位谱信息来进行逆傅里叶变换^[69]:

$$F = \mathcal{F}\{I\} \quad (2)$$

$$F'(u, v) = \frac{F(u, v)}{M(u, v)}, \forall u \in [0, m), \forall v \in [0, n) \quad (3)$$

$$I' = \mathcal{F}^{-1}\{F'\} \quad (4)$$

其中, I 代表着二维图像矩阵, \mathcal{F} 和 \mathcal{F}^{-1} 分别代表着傅里叶变换和逆傅里叶变换, $F(u, v)$ 代表着傅里

叶变换后的结果 F 在坐标 (u, v) 处的值, 而 $M(u, v)$ 代表着该点处的幅度值. m 和 n 则分别代表着图像矩阵的行数和列数. 处理后的 F' 中仅保留了相位谱的信息, I' 则代表着逆变换后得到的图像. 在此过程中能够去除原始图像中全部的周期性纹理部分而保留异常区域, 这一点已经由 Bai 等^[12]进行了数学证明. Aiger 等^[69]在 POFT 的基础上结合自适应阈值的方式自动在逆变换后的图像中分割异常区域. POFT 比背景频谱消除的方法更为便捷, 但对于长线条或划痕这种面积较大且具有周期性的缺陷却无法进行检测.

上述这两种方法由于其本身的特点导致其无法应用在集成电路等无周期性的图像, 有一种解决思路是人工地构造周期性. 如图 9(a) 所示, Bai 等^[12]将一系列待测样本以矩阵的形式进行排列, 人工构造出了正常图像的周期性, 由于异常一般不会在一个区域反复出现, 因此不会呈现周期性, 在逆变换后依然能被保留下来. 除了在二维空间中排列图像来构建周期性的方法, 也有 Tsai 等^[70]通过将多张图像进行堆叠, 使得正常背景在创建出的图像轴上能够体现出周期性, 如图 9(b) 所示.

相比于别的方法, 基于频域分析的异常检测方法其优点在于计算速度快, 而且不需要事先构建模型, 直接可以在待测图像上进行检测. 这使得该方法非常适合实时性要求较高的场景. 这一类方法对有周期性且纹理较为简单的图像有较好的检测效果, 比如磨砂玻璃表面^[63]以及图 7 中展示过的布匹图像等, 但依然对图像背景有一定的限制. 虽然有一些方法尝试利用人工构造周期性的策略来检测不规则图像, 但是其泛化性能并不好, 仅适用于图像采集环境严格可控的一些工业生产环境中. 当样本出现光照变化或者图像内目标出现平移和旋转等变化导致周期性较弱时, 这些方法往往表现不佳, 在正常目标附近也会出现许多干扰, 如图 10 所示.

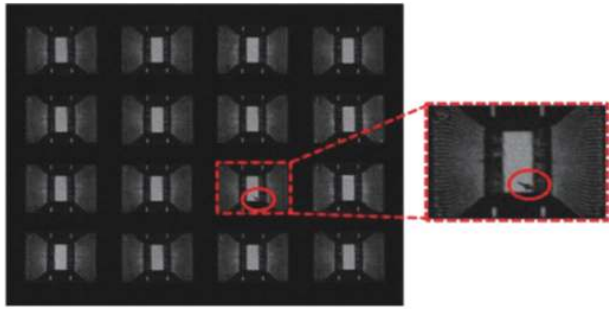
2.1.5 基于稀疏编码重构的异常检测方法

这一类方法通常是借助稀疏编码^[71-73]的方式对图像进行重构, 并在此过程中学习一个字典来表示正常图像, 然后在测试阶段从重构差异和稀疏度等角度进行异常检测.

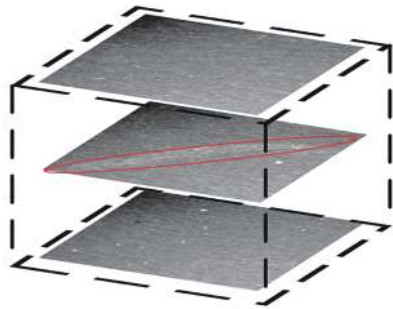
稀疏编码在模型训练阶段需要学习一个过完备的字典来存储有代表性的特征, 并通过线性组合字典中存储的元素来重构输入图像. 其特点在于, 需要尽可能地选择少的元素的线性组合来表示正常样本, 即:

$$\min_{D, \alpha} \frac{1}{2} \|D\alpha - x\|_2^2 + \lambda \|\alpha\|_1 \quad (5)$$

其中, D 为字典, α 是线性组合中的系数向量, x 代



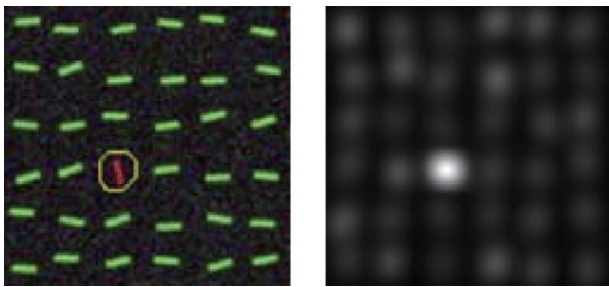
(a) 在二维平面中构建周期性
(a) Constructing periodicity in 2D plane



(b) 在三维空间中构建周期性
(b) Constructing periodicity in 3D space

图 9 人工构造周期性示意图

Fig.9 Illustration of artificial periodicity



(a) 原始图像
(a) Original image
(b) 检测结果
(b) Detection result

图 10 周期性较弱时的检测效果^[62]

Fig.10 Detection result in weakly periodic scene^[62]

表训练集中的正常样本, $\|D\alpha - \mathbf{x}\|_2$ 描述了字典对样本的重构误差, $\|\alpha\|_1$ 近似描述了系数向量 α 的稀疏度. 这种约束使得字典需要去不断精炼学习到的特征来提升编码过程的稀疏度. 字典学习的过程一般会先将原始问题转换成凸优化问题, 然后采用 Boyd 等^[74] 提出的交替方向乘子法来求解. 图 11 展示了在纳米材料上进行字典训练和稀疏编码的过程, 对于具有线状结构的纳米材料, 字典中存储了各种用来描述边缘区域的元素以有效地对正常样本进行重构. 而在测试阶段, 用学习到的字典对异常样本进行重构时, 由于字典仅学习了正常图像的表达方法, 对于异常样本就会体现出较大的重构差异,

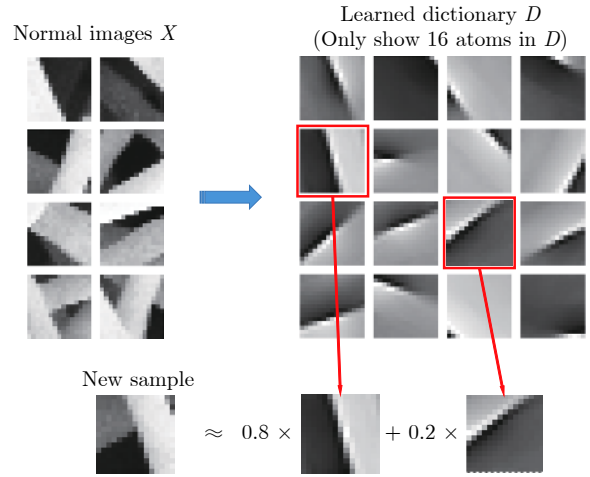


图 11 稀疏编码中的字典^[71]

Fig.11 The dictionary learned in sparse encoding^[71]

编码向量的稀疏性也较弱.

Liang 等^[72] 依据编码向量的稀疏度来进行触摸屏表面的异常检测. Boracchi 等^[71] 在稀疏度的基础上联合考虑重构误差来进行纳米材料表面的异常检测. Carrera 等^[75] 则在传统的稀疏编码中结合了多尺度的策略, 对于原始图像进行不同比例的缩放并分别构建字典, 通过这种方式来提升对不同大小异常区域的检测性能. Chen 等^[76] 在考虑稀疏度和重构精度的同时, 对图像局部区域的平滑性同样进行了约束以得到更高质量的字典和重构过程. 与稀疏编码比较接近的还有独立成分分析 (Independent component analysis, ICA)^[77]. Sezer 等^[78] 借助 ICA 寻找构成正常样本的独立元素作为字典, 在测试阶段计算待检图像的编码向量到正常图像的编码向量之间的距离来进行异常检测.

基于稀疏编码的方法相比于前 3 种方法而言, 不需要预先对数据的分布做假设, 仅使用足够的样本就能很好地学习到正常图像的表达方法, 这使得稀疏编码比之前的方法拥有更为广泛的应用场景. 在检测各类周期性图像的同时, 也能处理具有随机结构的图像, 特别是在如图 12 所示的纳米材料表面异常检测中有许多应用.

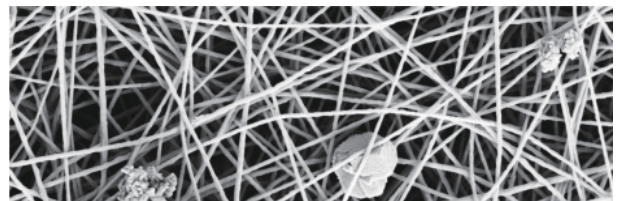


图 12 纳米材料图像^[71]

Fig.12 Image of nanofibres^[71]

不过这类方法依然存在一些不足之处.

首先, 稀疏编码的效率较为低下. 在测试阶段由于需要通过迭代优化的方式寻找最优的编码向量, 在 Carrera 等^[71]的实验中, 对于一张 1024×696 像素的图像, 迭代求解编码向量的过程需要约 50 秒的时间, 这使得这类方法很难直接应用到生产线上进行在线检测. 不过也有一些方法尝试结合目标的特点来进行优化, Zhou 等^[13]针对瓶盖表面异常检测的问题建立了可能出现的异常区域的字典来加速迭代寻优的过程.

其次, 稀疏编码需要较多的空间来保存字典. 为了获得稀疏的编码向量, 一般字典都是过完备的, 即字典中元素的数量要超过样本的维度. 但对于图像数据而言, 无论是局部图像区域还是特征向量其维度都比较高, 过大的字典不仅会占用较多的存储空间, 也会降低算法的运行速度.

2.1.6 基于分类面构建的异常检测方法

基于分类面构建的方法大多是希望能够在正常图像分布区域外构建一个足够紧致的分类面以区分正常样本和潜在的异常样本. 较为常用的两类方法为单类支持向量机方法 (One-class support vector machines, OC-SVM)^[79-80]和支持向量数据描述方法 (Support vector data description, SVDD)^[81]. OC-SVM 通过在高维空间创建超平面来分割正常样本和潜在异常样本, 而 SVDD 则创建超球面来包裹绝大部分正常样本以实现异常检测.

图 13 分别展示了 OC-SVM 和 SVDD 构建分类面的示意图, 其中实心与空心的点分别代表已知的正常样本和潜在的异常样本. 如图 13(a) 所示, OC-SVM 采用的策略是将原始数据映射到核函数对应的特征空间之后, 在创建分类超平面时最大化超平面到原点的距离并确保大部分正常样本都落在超平面的另一侧. 而在检测阶段, 落在超平面靠近原点一侧的图像就对应着异常样本. 其训练阶段的目标函数可表示为^[82]:

$$\min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|_2^2 - b + c \sum_{i=1}^n \xi_i, \text{ s.t. } \mathbf{w} \cdot \Phi(\mathbf{x}_i) \geq b - \xi_i \quad (6)$$

其中, \mathbf{w} 和 b 分别代表了超平面的法向量和截距, ξ_i 是松弛变量, n 代表训练样本的数量, c 则是对应的惩罚因子, Φ 代表投影函数, \mathbf{x}_i 代表正常样本. OC-SVM 模型简单而且有较为成熟的优化算法, 不过随着样本维度的上升, 其效果会受到显著的影响, 因此现有方法大多都会先对图像进行特征提取并对得到的特征向量进行异常检测. Amraee 等^[80]提取图像中每一个区域的 HOG 和局部二值模式 (Local binary pattern, LBP) 等特征, 然后借助 OC-

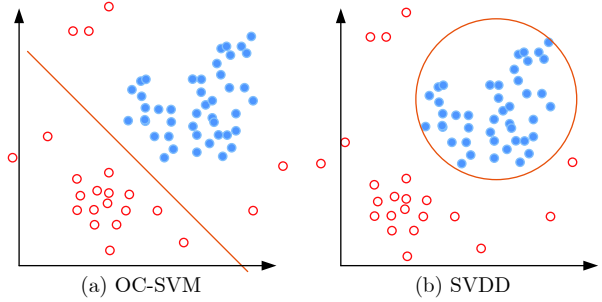


图 13 OC-SVM 和 SVDD 的示意图^[82]

Fig. 13 Illustration of OC-SVM and SVDD^[82]

SVM 对图像中每一个区域进行异常与否的分析. Azami 等^[83]对脑部核磁共振图像提取多种特征之后, 结合 OC-SVM 进行异常区域的检测.

不过 OC-SVM 仅考虑构建一个超平面来进行异常检测, 这样一个半开放的决策边界并没有很好地包围正常样本的分布区域, 这限制了 OC-SVM 对于异常样本的检出能力. 如图 13(a) 所示, 对于落在与正常样本同一侧的异常样本就无法进行良好的检测. 因此另外一类常用方法, SVDD 在分类面的构建过程中进行了更强的约束. SVDD 的基本思想为, 利用一个超球面来包裹住全部或者绝大部分的正常样本, 并且希望该超球面的半径越小越好, 对应的目标函数可表示为:

$$\min_{\mathbf{m}, r} \frac{1}{2} r^2 + c \sum_{i=1}^n \xi_i, \text{ s.t. } \|\Phi(\mathbf{x}_i) - \mathbf{m}\|_2^2 \leq r^2 - \xi_i \quad (7)$$

其中 \mathbf{m} 为聚类中心, r 为超球面的半径, 对应的超球面就是区分正常与异常样本的决策边界, 如图 13(b) 所示. 相比于 OC-SVM, SVDD 使用的超球面进一步提升了分类面的紧致性, 获得了相比于 OC-SVM 更加良好的异常检测效果. Zhang^[84]等和 Gurram 等^[85]都将 SVDD 引入到了高光图像谱图像的异常检测任务中, Liu 等^[86]提出了加速 SVDD 以实现液晶屏表面细微缺陷的实时检测.

不过, SVDD 对数据的分布又做了一个假设, 即所有正常样本都满足一个单峰的分布模式, 可以用一个超球面良好地进行表示. 这是一个比较强的假设, 虽然也可以通过引入核方法来提升对不同形态分布的拟合能力, 但是对于许多图像而言, 不同区域的像素值或者特征向量的分布模式完全不同, 导致原始数据表现为多分布融合的结果, 很难用一个超球面来进行描述. 如图 13(b) 中的样本点, 在超球面的内部还是存在着许多代表着未知样本的空区域. 如果异常图像落在这些区域内就很有可能会被判定为正常图像. 为了解决这一问题, Liu 等^[87]尝试集成多个 SVDD 模型来描述这种样本分布更为

复杂的情况, 并利用模糊 c 均值聚类的方式确定子 SVDD 模型的数量及对应的参数。

基于分类面构建的方法, 其优点在于对待测图像类型没有如图像分解或频谱分析那样较高的限制, 能够适用于各种类型的图像, 而且计算复杂度的问题也有相关的文献工作给出了解决思路。不过也存在一些问题, 在处理复杂图像时, 核函数的选择可能会成为一个问题。如果选择使用如 Liu 等^[87]提出的多超球面的方法, 其各项参数的选择也需要较为精心的设计。而且本质上本类方法属于图像分类算法, 无法直接实现异常区域的定位, 更适合于一些识别定性异常的场景, 面对定量异常, 大多需要通过区域划分的方式进行定位, 反而降低了算法的处理效率。

表 2 总结了上述各类基于传统方法的图像异常检测方法的设计思路和优缺点。其中, 基于模板匹配的方法十分适合工业生产这类环境可控且目标高度一致的场景, 实现了较高的检测精度, 但不适用于采集环境多变或者正常图像间存在较大差异的场景。基于统计模型的方法虽然速度很快, 但需要一定的训练样本来评估背景模型的参数。基于图像分解的方法则适合训练样本稀缺的场合, 可以直接在待测样本上检测异常区域, 不过该方法速度较慢。而基于频域分析的方法则兼顾了检测速度和对训练样本的低依赖性, 但是由于对背景图像有一定的限制, 因此通用性较差, 这也是前 3 类方法的一个通病。而后两类基于稀疏编码和分类面构建的方法适用于各种类型的图像, 具有更高的通用性。但稀疏编码方法在重构阶段非常耗时, 而分类面构建的方法也面临着参数选择困难和定位精度的问题。

2.2 基于深度学习的异常检测技术

近年来, 深度学习在计算机视觉中的各个领域

内都得到了长足的发展。相比于传统的方法, 深度学习由于其无需人工设计特征, 算法通用性更高等优点, 已经被广泛引入到了图像异常检测任务当中。现有的方法大致可以分为以下几类: 基于距离度量的方法、基于分类面构建的方法、基于图像重构的方法和与传统方法相结合的方法。

2.2.1 基于距离度量的异常检测方法

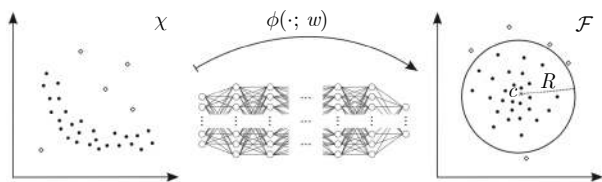
基于距离度量的方法, 其核心思想在于训练一个神经网络作为特征提取器, 让正常图像所提取到的特征向量的分布尽量紧凑, 即尽可能地减小样本的类内距离。而在测试阶段, 大多数方法计算待测样本的特征与正常特征之间的距离作为度量来进行异常检测。

Ruff 等^[88]提出的深度支持向量数据描述 (Deep support vector data description, Deep SVDD) 是这一类型中较为常用的一种方法。作者首先在特征空间中人为指定了一个点作为特征中心, 然后以正常样本到该点的距离之和作为损失函数的主体来进行特征提取网络的训练, 如图 14 所示, 经过训练网络能够将原始图像空间中的正常样本都映射到特征中心点附近, 而异常样本其对应的特征就可能远离该中心点, 因此根据距离就能判断待测样本是否属于异常样本。

但这种方式存在着较多的限制。首先, 该方法假设正常样本在特征空间中属于一个单峰分布, 即都落在指定的中心点附近。这一假设在一些较为复杂多变的数据集上可能并不成立。这一点从 Ruff 等的实验结果中也能看出, Deep SVDD 对于手写数字数据集 (Mixed national institute of standards and technology database, MNIST)^[89] 的效果十分优秀, 但是对于加拿大高等研究院 (Canadian institute for advanced research, CIFAR) 构建的 CI-

表 2 基于传统方法的图像异常检测技术的分类和特点
Table 2 The classification and characteristic of traditional image anomaly detection methods

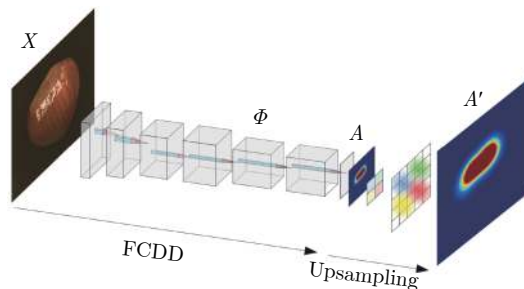
方法类别	设计思路	优点	缺点	参考文献
模板匹配	建立待测图像和模板图像之间的对应关系, 通过比较得到异常区域	方法简单有效, 对于采集环境高度可控的场景有很高的检测精度	不适用于多变的场景或目标	[29-48]
统计模型	通过统计学方法构建背景模型	具有详实的理论基础和推导过程, 检测速度快	需要大量的训练样本, 且仅适用于一些简单背景下的异常检测	[49-54]
图像分解	将原始图像分解成代表背景的低秩矩阵和代表异常区域的稀疏矩阵	具有详实的理论基础且无需训练过程	速度较慢, 而且不适合在结构复杂的图像中进行异常检测	[56-61]
频域分析	通过编辑图像的频谱信息来消除图像中重复的背景纹理部分以凸显异常区域	无需训练过程, 检测速度很快	还需更详实的理论论证, 且仅适用于一些有重复性纹理的图像, 通用性较差	[64-70]
稀疏编码重构	借助稀疏编码和字典学习等方式学习正常样本的表示方法, 从重构误差和稀疏度等角度检测异常	适用于各种类型的图像, 通用性很好	检测时间长, 而且需要额外的空间保存过完备的字典。	[71-78]
分类面构建	建立分类面将现有的正常样本和潜在的异常样本进行区分	通用性较好, 且速度较快	各项参数的选择过程较为复杂	[81-87]

图 14 Deep SVDD 的原理示意图^[88]Fig.14 Illustration of Deep SVDD^[88]

FAR-10^[90] 这种更为复杂的自然图像数据集其性能较为有限. 此外, 该方法容易出现模型退化的问题, 即学习到的模型会把所有图像都映射到同一点上, 因此该方法在训练过程中还有许多限制, 比如网络结构中不能有偏置项, 特征中心点需要人工指定并且不可修改, 而且在检测阶段还涉及到了对特征向量的编辑过程, 这些都限制了 Deep SVDD 的应用场景.

不过可以通过额外的训练目标来优化特征提取过程以避免 Deep SVDD 遇到的各种限制^[91]. Perera 等^[92] 并没有人工地指定特征中心, 而只是以减小正常样本特征之间的距离为训练目标. 不过作者在正常样本的基础上, 引入了 ImageNet 数据集中的样本来构建一个分类的子任务. 网络不仅需要减小正常样本特征的类内离散度, 还需要有一定的类间离散度来保证分类性能, 网络需要有较强的语义信息提取能力而不仅仅是将样本映射到某一区域内, 这使得该方法对于复杂的自然图像数据也有较好的性能. 而 Wu 等^[93] 增加了解码模块将特征解码成与输入样本近似的图像, 以此来保证所提取到的特征向量有足够的语义信息, 避免出现 Deep SVDD 中的模型退化问题. Bergman 等^[94] 在 Deep SVDD 的基础之上, 利用多种几何变换方法将原本仅有一类的正常图像数据扩增成了多类数据集, 并且为每种变换对应的类别都指定了一个特征中心. 作者在训练过程中采用了三元损失 (Triplet loss)^[95] 来强调不同变换方式对应的特征向量之间的可分性. 这种方式既避免了模型退化的问题还构建出了多个特征中心点作为分类参考, 相比于原始的 Deep SVDD 方法有了显著的性能提升. 不过依然需要更加通用的几何变换方式来适应具有周期性或者纹理类的图像^[94]. 近期, Liznerski 等^[96] 提出的全卷积数据描述 (Fully convolutional data description, FCDD) 利用神经网络中的偏置项替代 Deep SVDD 中的聚类中心, 并且利用全卷积网络提取图像的特征图来直接定位异常区域, 如图 15 所示.

除了用待测样本的特征与正常特征之间的距离

图 15 FCDD 的原理示意图^[96]Fig.15 Illustration of FCDD^[96]

来衡量样本的异常程度, 还有方式尝试结合距离的方差来进行异常检测. Bergmann 等^[97] 同时训练多个结构相同的特征提取网络. 而在检测阶段, 由于训练过程的随机性, 不同的网络会得到不同的特征, 根据特征分布的方差来实现对异常样本的检测. 此外还有方式尝试利用半监督的方式提升检测性能, 因为在实际检测任务中标注少量的异常样本是可行的. Ruff 等^[98] 就将少量经过标注的异常样本加入到训练过程中, 并且以尽量增大异常样本特征向量分布的熵作为目标来训练特征提取网络. 实验表明这种基于熵的损失函数要优于传统的基于距离的损失函数.

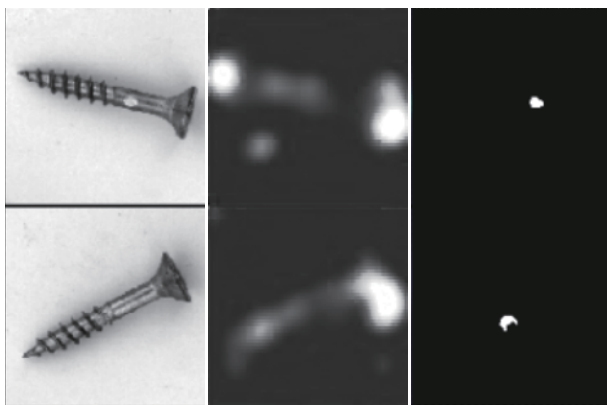
上述这些方法的优点在于简单高效, 不过这种方法大多需要事先人工指定特征中心, 而且为了避免模型退化需要在训练阶段设计额外的任务. 仅设定一个全局特征中心点的方式在一定程度上对图像背景产生了一些约束, 在医学图像或者自然图像这些较为多变的场景中, 可能难以在保证泛化能力的情况下将全部的图像映射到同一目标点附近. Wu 等^[93] 采用的为每种图像区域分别设置中心点的方式并不适合所有图像, 而 FCDD^[96] 以偏置项作为中心点的方式或许也不适用于待测目标经常出现平移和旋转等变化的场景下的异常定位任务, 如图 16 所示.

2.2.2 基于分类面构建的异常检测方法

基于分类面构建的方法, 其核心思想在于将单类正常样本转换成多类别样本以训练分类器, 通过这种方式来在图像空间中构建分类面, 实现对正常样本和潜在的异常样本的分类. 常用的基于分类面构建的方法大致包含以下两个类别:

1) 第一类方法将原始单类样本通过几何变换得到多类别样本, 并结合在分布外检测 (Out-of-distribution, OOD)^[99] 任务当中比较常见的基于置信度的方法来进行异常检测.

OOD 检测任务与异常检测任务目标非常相似,



(a) 原始图像 (b) 检测结果 (c) 真值图
 (a) Original images (b) Detection results (c) Ground truths

图 16 在旋转目标上的检测效果^[96]

Fig. 16 Detection results on targets with rotation^[96]

同样也需要模型对训练过程中未出现过的样本有检测能力, 但 OOD 的特点在于训练样本中包含了多个类别, 所以可以直接在训练样本上进行多分类的训练. Hendrycks 等^[100]指出, 由于异常样本落在正常样本的分布之外, 分类器对异常样本输出的最大 softmax 值往往会低于正常样本的最大 softmax 值, 所以可以通过设定阈值的方式区分正常样本和异常样本. 由于异常检测中训练集仅包含了一类样本, 所以相关工作就尝试通过对正常样本进行变换的方式来构建多类别的训练集.

如图 17 所示, Golan 等^[101]采用了以翻转、平移和旋转为基础的一共 72 种几何变换方式来处理原始图像, 每一种变换方式下得到的图像即为一类样本, 以此构建了一个 72 类的分类数据集来训练一个分类网络. 在检测阶段, 对待检样本进行全部的 72 种变换并分别进行分类, 异常样本经过变换之后, 网络会无法确定其对应的变换类别, 使得分类时输出的最大 softmax 值降低, 以此来进行异常图像的检测. Hendrycks 等^[102]对该方法进行了改进, 将原本直接的 72 分类任务转换成了一个多任务 (Multi-task) 模型, 在提取到的特征图上进行额外的角度和平移量的分类任务, 以此来提升网络特征提取和分类的能力.

第一类方法虽然在 CIFAR-10 等自然图像数据集上取得了优异的成绩, 不过存在着一些限制, 如图 18 所示, 当检测目标为斑马时, 网络能够识别图像之间存在旋转的关系. 但如果检测目标为具有对称结构或者没有方向信息的纹理图像, 比如图 18 第二行所示斑马表面的黑白纹理, 网络无法直接从纹理图像中感知到足够大的差异实现旋转角的预测, 此时强迫网络输出相应的角度反而会

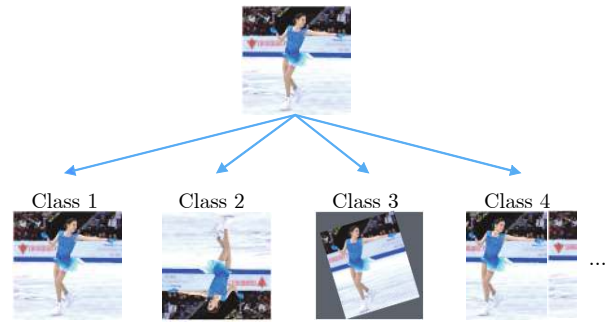


图 17 将单类样本转换成多类样本^[101]

Fig. 17 Transforming one-class samples into multi-class samples^[101]

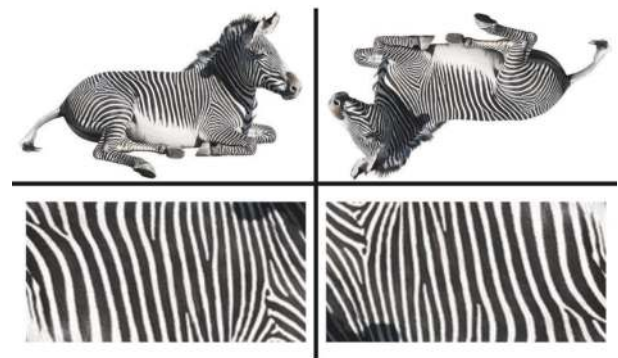


图 18 不同图像上旋转效果对比^[102]

Fig. 18 Comparison of rotation on different images^[102]

的训练过程^[101]. 类似地, 当检测目标本身存在角度上的显著变化时也难以应用上述方法. 因此这类方法还需要设计更为通用的变换方式以扩展其应用领域, 充分发挥其高精度的优势. 此外, 也有相关的研究指出, 在 OOD 检测任务中, 分类模型对异常样本也有可能产生很高的概率值, 影响异常检测的过程^[103], 这同样是一个值得进一步研究和改进的内容.

2) 第二类方法则考虑结合传统方法中 OC-SVM 或者 SVDD 的思路, 构建尽可能贴合正常样本分布的分类面来进行异常检测. 这类方法大多将正常样本当作正样本, 并采用额外的辅助样本作为负样本, 以此在图像空间中构建正常和异常图像间的分类面.

Oza 等^[104]利用预训练好的神经网络提取正常图像的特征作为正样本, 同时在特征空间中使用以原点为中心的随机高斯噪声向量作为负样本, 以此来训练一个分类网络. 不过这种方法仅使用高斯噪声作为负样本, 负样本比较单一且容易分类, 容易出现过拟合的问题导致网络无法检测新的异常样本. Hendrycks 等^[105]注意到了随机生成的简单噪声样本的不足, 因此从别的图像数据集中选取图像作

为负样本来训练分类模型。

然而上述方法在选取负样本时没有考虑到正常样本的分布情况,在这种情况下训练得到的分类效果就无法保证.当选择的负样本距离正常样本较远时,网络容易出现过拟合的现象,导致其无法对真正的异常样本进行分类.而如果选择的样本与正常样本的分布过于相似,也可能出现网络无法训练的问题。

因此,许多方法尝试在正常样本分布区域附近通过生成式模型创建负样本.而生成式对抗网络(Generative adversarial network, GAN)^[106]是近年来备受关注的生成式模型,有许多方法就结合 GAN 来进行分类器的训练. GAN 的结构如图 19 所示,在 GAN 的训练过程中,目标函数一般可以表示为:

$$\min_G \max_D E_X [\ln D(X)] + E_z [\ln(1 - D(G(z)))] \quad (8)$$

其中, G 和 D 分别代表生成器和判别器, X 是真实样本, z 是随机的噪声向量, $G(z)$ 即为生成的样本。

为了对抗不断优化的判别器,生成器所生成的图像质量会不断提升,最终能生成清晰且真实的样本图像,而 Sabokrou 等^[107-108]注意到在 GAN 训练的中期,生成的依然是低质量的图像,鉴于其与正常图像之间存在一定的相似性,此时的判别器正好可以作为一个分类器进行图像异常检测.作者将自编码器作为生成器,并且附加了一个分类网络作为判别器.在自编码器训练过程中作者采用了 Early-stop^[109]的策略,即当重构误差小于特定阈值时就停止自编码器的训练避免生成过于真实的样本,如果过度训练反而会影响判别器对真实异常样本的分类性能^[110].利用此时的自编码器处理正常样本就能得到大量接近正常样本的合成图像.而判别器为了区分这些样本和正常样本,就需要在正常样本周围构建分类面,足够贴合的分类面为该方法提供了良好的异常检测性能. Yang 等^[110]和 Chatillon 等^[111]也采取了相似的策略,不同的是, Yang 等是通过分析

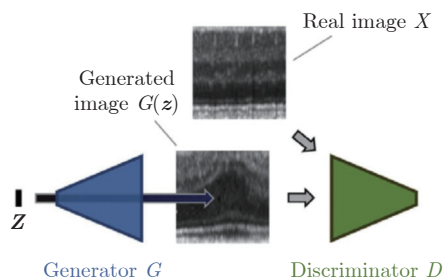


图 19 GAN 结构示意图^[15]

Fig. 19 The structure of GAN^[15]

损失函数曲线的方式人工设定 Early-stop 的节点. Chatillon 等是在 GAN 训练完毕之后,选择一部分训练中期的生成器权重,用来生成低质量样本作为分类时使用的负样本,并且还给出了一部分数学上的证明以表达该方法的有效性. Zaheer 等^[112]同样将原始 GAN 中的生成器换成了自编码器,并且将判别器的任务从区分真实图像和生成图像转换成了区分正常图像和异常图像,其中在训练阶段使用的异常图像同样是借助未充分训练的生成器得到的。

这类方法都采用了 Early-stop 来避免生成器过度优化,为此还需要进行反复的实验来确定最优的超参数,这使得这些方法在应用到新的图像类型中时需要较长的调参过程.为了让模型能够生成低质量样本又避免反复调参,部分方法开始主动探索对应着低质量样本的区域来生成负样本. Lim 等^[113]尝试在特征空间中生成低质量样本,作者在 GAN 的基础上,对隐变量的分布进行了约束,迫使其分布在原点周围.训练完毕后,距离原点越远的特征向量,其解码后的图像质量就越差.因此,可以通过在该区域内进行采样和插值等方式得到低质量样本以进行分类器的训练.而 Liu 等^[114]则通过修改生成器训练目标来直接生成低质量样本,该方法用一些正常样本中最容易被判定为异常图像的低质量本来训练生成器,在这种情况下生成器的性能就会被限制,导致其只能生成低质量的样本.不过由于并不是每一批训练样本中都包含低质量样本,所以可能会存在训练不稳定的问题.针对这一点, Ngo 等^[115]改进了生成器训练阶段使用的损失函数,对于判别器能以较高置信度判别正常与否的样本都进行惩罚,迫使生成器生成与正常样本尽量相似又有所区别的低质量样本. Schlachter 等^[116]也采用了类似的思路,不过 Schlachter 并没有利用 GAN,而是通过分析正常样本特征之间的距离将训练样本分成了典型和非典型两类,典型样本特征之间的距离较小而非典型样本特征之间的距离较大.随后在两种样本之间构建分类面,以区分位于非典型样本区域之外的异常样本。

Goyal 等^[117]提出的鲁棒单类别分类 (Deep robust one-class classification, DROCC) 是这一类方法的最新工作,作者仅采用了一个分类器进行异常检测,并通过梯度上升自动生成最适合现有数据的异常样本作为负样本.在分类器训练的初期,仅使用正常样本进行训练并希望分类器对全部样本输出相同的结果.而在后续的训练过程中,对于每一个正常样本,都通过梯度上升的方式,在以正常样本

为中心半径为 r 的图像空间中寻找潜在的负样本并进行二分类的训练. 该方法在各种类型的数据集上都表现出优异的性能. 不过该方法依然存在一个需要调节的超参数 r , 而且从实验结果来看不同的半径 r 会对分类器的检测性能产生较大的影响, 因此同样需要反复实验来得到最合适的半径.

图 20 展示了上述几种需要构建负样本来训练分类网络的方法示意图. 其中 (a)~(d) 分别展示了利用随机噪声^[104]、利用随机图像^[105]、利用 GAN^[110-115] 以及利用梯度上升^[117] 来创建负样本的方法. 其中实心与空心的圆点分别代表了正常样本与生成的负样本. 实线与虚线则分别代表了分类器对正常样本和异常样本的决策边界. 而点划线代表着两者之间的不确定区域. 曲线上的数字代指分类器的输出值, 越高表示分类器越肯定该图像为正常图像, 反之亦然. 对于 (a) 和 (b) 方法, 由于随机的噪声或者图像与正常样本之间并没有关联, 所以构建出的分类面中存在大量的不确定区域, 分类器无法很好地鉴别位于这些区域内的图像进而影响检测性能. 而后两种方法都主动在正常图像周围生成低质量图像作为负样本, 而且梯度上升的方法更是显式地约束了生成图像到正常图像的距离, 这些方法由于采用的负样本更加贴近正常图像, 因此训练得到的分类器能够产生更加紧密的决策边界和更小的不确定区域, 异常检测精度更高.

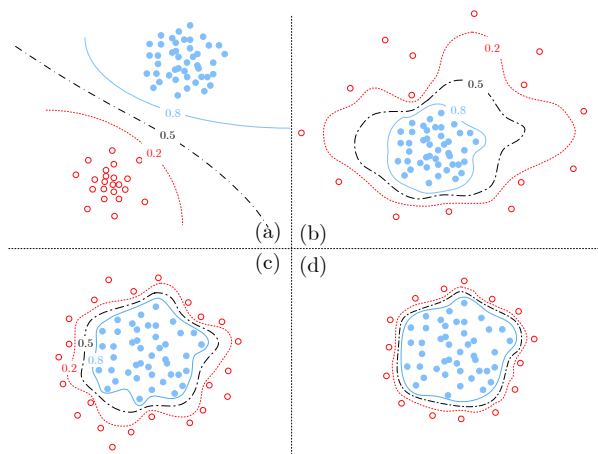


图 20 需要构建负样本的几种方法的示意图

Fig. 20 The graphical illustration of the methods based on creating fake-negative samples

然而, 这些利用负样本构建分类面的方法存在较多的假设和随机性, 大多难以直观地控制所生成负样本的图像质量和分布, 虽然 DROCC 尝试通过梯度上升的方式生成负样本, 但调参过程较为繁琐. 而且这些方法所生成样本的分布特性大多只是在二

维或三维空间进行了可视化验证, 在更高维的复杂图像或者特征空间中, 或许就无法得到理想的紧致分类面, 这或许就是第二种方法在精度上略低于第一种几何变换类的方法的原因^[102, 117]. 所以, 如何更直观有效地设计负样本生成算法使其紧密分布在正常样本周围依然是一个需要解决的问题.

2.2.3 基于图像重构的异常检测方法

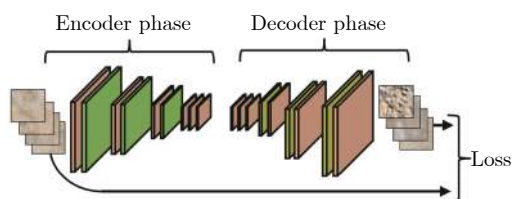
基于图像重构的方法, 其核心思想在于对输入的正常图像进行编解码, 并以重构输入为目标训练神经网络, 以此来学习正常图像的分布模式. 然后在检测阶段通过分析重构前后图像之间的差异来进行异常检测. 根据采取的训练模式, 常用的基于图像重构的方法大致包含基于自编码器和基于生成式对抗网络 (Generative adversarial networks, GAN) 两种类型.

1) 在基于图像重构的方法中, 最为常用的网络结构为自编码器 (Autoencoder, AE)^[118-119]. 仅利用正常样本训练得到的自编码器, 在测试阶段能够良好地重构正常图像. 而对于存在异常的图像, 在图像编码以及后续的重构过程中都会与正常图像产生较大的差异, 而差异的大小即为衡量待测样本异常程度的指标.

自编码器的结构如图 21 所示, 其一般由一个编码器和一个解码器组成, 且两者的网络结构一般是对称的. 其中, 编码器在网络前向传播过程中不断缩小特征图的尺寸, 以此来删除冗余的信息. 而解码器负责对特征进行解码, 得到与输入图像相同大小的图像, 通过计算重构前后图像之间的差异来训练网络. 而在此过程中最为常用的损失函数就是均方误差 (Mean square error, MSE)^[120]. MSE 用重构前后图像中所有像素点上的像素值之差的平方均值来衡量图像重构的质量. 训练结束后, 由于瓶颈结构的存在, 对于一些异常区域面积较小的样本, 自编码器能够在图像编解码的过程中消除异常区域的影响, 重构出一张正常图像作为参考, 随后可以通过逐像素比较的方式得到异常区域.

很多方法利用自编码器来进行异常检测^[121-123]. Mei 等^[119] 利用降噪自编码器来进行纹理图像的异常区域定位, 作者将样本切分成一系列小的图像片并分别进行重构. 此外, Mei 等还表明结合多尺度的策略^[124], 即在多个不同尺度下对图像进行重构可以有效地提升异常的定位精度.

除了重构误差, 近期有一些方法开始利用图像重构中得到的梯度信息来进行异常检测. Kwon 等^[125] 指出, 重构误差一般仅衡量了在中间层特征空间以及最后的图像空间内的差异. 但如果分析梯度差异,

图 21 自编码器的结构^[119]Fig.21 The structure of autoencoder^[119]

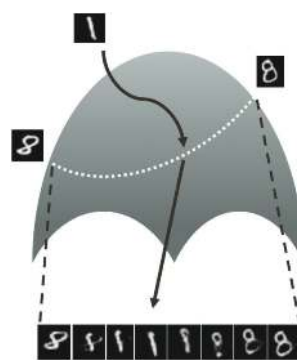
则可以从重构网络的任意一层处获取并进行比较,这使得基于梯度的方法能从更全面的角度衡量待测图像的异常程度.此外,Zimmerer等^[126]表明在异常样本上计算得到的梯度还提供了额外的方向信息辅助分类.对于图像而言,异常图像上计算得到的梯度其实就代表着朝向正常样本的优化方向.这些优点使得梯度向量相比于简单的重构损失能提供更加完备的信息.Zimmerer等^[126]首先将梯度信息引入到了异常检测任务中,通过计算变分自编码器(Variational autoencoder, VAE)中证据下界(Evidence lower bound, ELBO)相对于输入图像的梯度来进行异常检测.Kwon等^[125]利用余弦相似度来计算正常样本梯度向量之间的角度,以此来构建正常样本梯度向量的方向一致性约束.在测试阶段出现不满足该一致性约束的梯度向量时,则认为该图像为异常图像.Chu等^[127]则从一个全新的角度进行异常图像检测,作者们发现自编码器在以正常样本为主的训练集上训练时,正常图像的重构误差会稳步下降,但异常图像的重构误差则会出现波动.因此,作者通过分析损失函数的变化曲线(Loss profile)以实现未标记样本中异常图像的检测.

不过,这种基于传统自编码器的图像重构方法存在重构后图像比较模糊的问题.自编码器常用的MSE统计了每一个像素点上的重构差异,但由于瓶颈结构的存在,在编解码的过程中会有信息的丢失,导致从理论上来说自编码器无法保证每一个像素点上的像素值都不变.如果迫使MSE接近于0,则会得到一个大致而平均的结果,这就使得重构后的图像容易模糊,在边缘区域会出现较多的差异从而影响后续的异常定位过程.

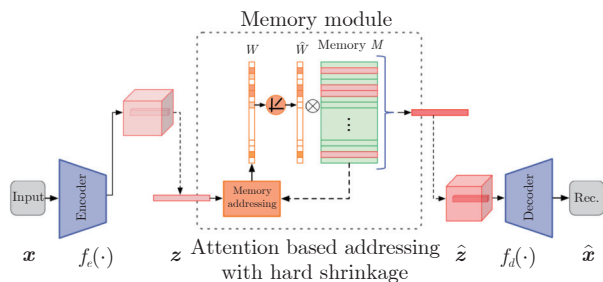
对于自编码器重构较为模糊的问题,早期方法会尝试采用VAE并结合编码后的特征辅助异常检测^[128],或者修改重构阶段的损失函数^[120]来进行改进,不过效果都比较有限.Abati等^[129]则是在图像重构的基础上,引入了自回归过程以学习潜在表示的概率分布,并通过最小化隐变量分布的微分熵来训练网络.相比于VAE,这种更加灵活的隐变量分布提升了图像重构的质量和异常检测的能力.

此外有方法尝试通过修改网络结构的方式优化重构的质量,Zhou等^[16]通过额外增加一条用来提取图像结构的支路来尽可能地保留输入样本的信息.作者采用预先训练好的网络提取输入样本的结构特征,并且将此特征融合到图像重构的训练过程中.此外,还通过比较重构前后图像结构信息的方式来确保重构图像的质量.

除了重构模糊的问题,基于自编码器的方法还存在着无法保证完全消除输入图像中的异常区域的问题^[11].当训练样本比较多样化时,自编码器会体现出强大的学习能力并对潜在的异常样本产生过强的适应能力.如图22所示,自编码器仅使用了数字8进行训练.在编码异常图像数字1时,虽然1对应的特征向量在特征空间中会远离正常特征向量的分布区域,但重构网络依然重构出了与输入接近的图像,导致其重构前后的差异反而较小,使得一些异常样本被误判成正常样本.针对这一问题,大多方法采用对隐变量进行编辑的方式来解决.Tian等^[131]尝试利用一个全连接层优化待测图像的特征向量,迫使其接近于正常样本的特征向量,以此来保证重构后的图像中不存在异常区域.此外,如图23所示,Gong等^[132]提出的记忆强化自编码器(Memory-augmented deep autoencoder, MemAE)在自编码器的基础上,增加了一个记忆模块,用来存储最具有代表性的特征向量以提升图像重构的稳定度.

图 22 异常样本的重构示意图^[130]Fig.22 The reconstruction of anomalous images^[130]

不过该方法存在一个问题就是需要大量的空间存储训练得到的记忆模块,文章中一共采用了1000个记忆向量来保证特征重构的效果.针对这一问题,Park等^[133]在MemAE的基础上,尝试提升记忆向量学习的有效性.在记忆向量的训练阶段,结合聚类中类内紧致类间可分的思想,最小化归属于同一记忆向量的特征的类内距离,同时提升不同记忆向量之间的可分度.通过增加这两项损失函数,作者

图 23 隐变量编辑示意图^[132]Fig. 23 The editing of latent vector^[132]

将记忆模块的容量从原先的 1000 个降到了 10 个, 显著提升了算法的效率. 除了编辑隐变量, Dehaene 等^[134] 采用直接在图像空间内进行迭代优化的方式来消除重构图像中的异常结构, 相比于优化隐变量再解码的方式, 这种直接在图像空间内优化的方式能更好的保留图像的细节. 不过, 上述几种方法在检测阶段的优化过程较为烦琐且耗时. 而 Yang 等^[11] 在隐变量层增加了基于聚类损失的正则项来避免重构出异常区域, 在纹理图像上的效果较好, 但缺乏理论上的保证.

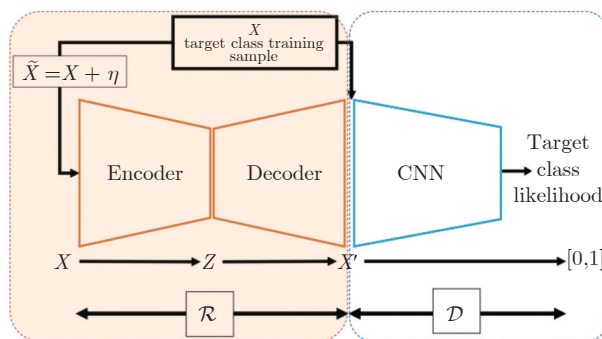
2) 基于 GAN 的异常检测方法大多是利用 GAN 能生成逼真图像的特点^[106], 采用 GAN 来重构出比自编码器更加清晰的图像. 而根据其具体重构方式, 又可分为直接利用 GAN 重构以及结合 AE 与 GAN 重构两个类型.

直接利用 GAN 重构的方法考虑到原始的 GAN 仅创建了从隐空间到图像空间的映射关系, 因此采用迭代优化的方式获得重构图像. Schlegl 等^[15] 提出的基于 GAN 的异常检测模型 (Anomaly detection with generative adversarial network, AnoGAN) 从某个随机变量开始, 计算该变量生成的图像和待测图像之间的差异, 通过梯度下降的方式迭代优化该随机变量, 使得生成的图像逐渐接近待测图像. 由于生成器仅使用正常样本进行训练, 所以理论上仅能生成正常样本. 当待测图像中存在异常区域时, 生成器会生成与其尽量接近但属于正常类别的图像作为参考, 通过计算待测图像和生成图像之间的差异来进行异常检测. Deecke 等^[135] 则在此基础上进行了改进, 从多个点开始尝试对待测图像进行重构, 而且在迭代优化过程中, 同时优化隐变量和生成器内部参数, 以此来提升图像重构的质量. AnoGAN 已经在实际检测任务中有了相关的应用^[136], 不过其存在算法效率上的不足. AnoGAN 在检测阶段需要进行多次的迭代优化来生成合适的正常图像作为参考, 迭代优化的过程会显著增加算法的执行时间, 在一些需要实时检测的任务当中就无法应用.

而更多的方法会将 GAN 与 AE 结合, 将 GAN 中对抗训练的思想引入到传统 AE 的训练框架中, 以此来提升自编码器的重构质量.

在异常检测领域, 常见的结合 AE 与 GAN 的方法如图 24 所示, 在自编码器的基础上增加一个判别器, 用来区分重构后的图像和输入的真实图像, 通过判别器和重构网络的对抗训练来提升重构图像的质量^[137-139]. Baur 等^[140] 在 VAE 的基础上, 增加了对抗损失来对医学图像进行更为真实的重构. Akcay 等^[141] 在传统重构误差的基础上, 还统计了重构前后图像经编码器和判别器提取到的特征之间的差异, 通过多角度的约束迫使重构后的图像在特征空间和图像空间中都能尽量地与原始图像接近. Schlegl 等^[142] 直接用 GAN 中训练好的生成器替换了原始 AE 中的解码器, 以一种更为直接的方式利用 GAN 强大的图像生成能力来进行图像重构. Tang 等^[143] 将原本仅用来对图像进行二分类的判别器替换成了一个辅助的图像重构网络, 原本判别器执行的二分类任务也转换成了一个重构任务, 通过这种更为直观的约束来提升自编码器的重构质量. Venkataramanan 等^[144] 表示, 人们会需要观察整张图像来觉察到哪里是有异常的部分, 意味着图像中所有的区域都有助于进行异常检测. 因此, 作者在结合了 VAE 与 GAN 进行图像重构的基础上, 计算出了重构过程中的注意力图^[145], 并期望注意力图对于所有正常图像区域都能输出高值. 而在检测阶段, 当图像中出现没有训练过的异常图像模式时, 就会表现出低注意力的情况.

不过在实验过程中, 即便是使用正常样本训练得到的 GAN, 同样也会生成一些低质量的图像^[146], 导致在重构图像中可能会残留异常区域. 虽然其概率相比于基于自编码器的方法要小, 但依然不容忽视. Perara 等^[130] 发现, 自编码器的这一问题主要是

图 24 结合自编码器和 GAN 进行图像重构^[107]Fig. 24 Image reconstruction based on autoencoder and GAN^[107]

由于中间层对应的特征空间中, 还存在着能解码出其他类型图像的区域. 所以, 作者们提出的单类 GAN(One-class GAN, OCGAN) 通过一种梯度上升的方法, 主动地探索并消除中间层特征空间内能解码成其他图像的区域. 但毕竟中间层特征向量的空间维度较高, 通过这种探索式的策略很难全部覆盖到, 这种利用梯度上升的方法可能仅能优化原始特征向量分布区域附近的空间. 而且 OCGAN 在 CIFAR-10 数据集上的实验表明该方法在较为复杂的图像异常检测任务中优势并不明显.

此外, 在某些正常样本数量也极少的情况下, 经典的重构网络可能会出现过拟合等问题. 而 GAN 也可能无法获得生成多样化数据的能力, 因为 GAN 本身就是一个需要大量数据进行训练的模型, 当数据量不足时会严重影响其图像生成的能力. 因此, Lu 等^[147]借助了小样本学习中元学习^[148]的思想来进行异常图像帧的检测. 作者预先在多个数据集上进行元训练, 来得到一组有良好通用性的重构网络权重. 而对于新的异常检测环境, 仅需要少量正常图像进行几次梯度下降优化就能得到最适合目标场景的检测模型.

图像重构类方法借助重构后的图像, 无需采用滑窗或者逐区域分析的方法就可以高效地实现异常区域的定位.

不过, 图像重构类方法也有许多值得进一步研究的内容, 如正常区域重构误差的问题. 由于瓶颈结构的存在, 在重构过程中很容易丢失图像细节, 导致在重构前后图像比较过程中在正常图像区域出现较大的差异, 现有方法往往难以解决这一问题, 这一点在结构复杂多样的医学图像中尤为显著, 如图 25 所示虹膜和眼底图像, 矩形框代表着存在异常的区域. 可以看到重构前后的图像在正常区域内也存在较大的差异, 特别是在具有细致纹理的区域. 因此也有方法不直接从像素级别的重构误差进行分析, Xia 等^[149]针对的是图像语义分割中的异常检测任务, 在利用 GAN 将分割结果重构成正常图像后, 通过比较重构前后图像特征间的余弦距离来定位异常目标.

此外还有如何进行大尺度下精细重构的问题. 上述许多方法在处理大分辨率图像时大多都会将原图缩放至较小的分辨率, 此时对于异常较为明显的自然图像而言影响不大, 但对于如图 26 所示面积占比极小的异常而言, 缩放后不仅在原图中难以观测, 而且定位过程容易受到正常区域重构差异的影响. 而这种小面积异常区域的检测在工业领域内又有较高的需求, 因此如何实现大尺度下图像的精细

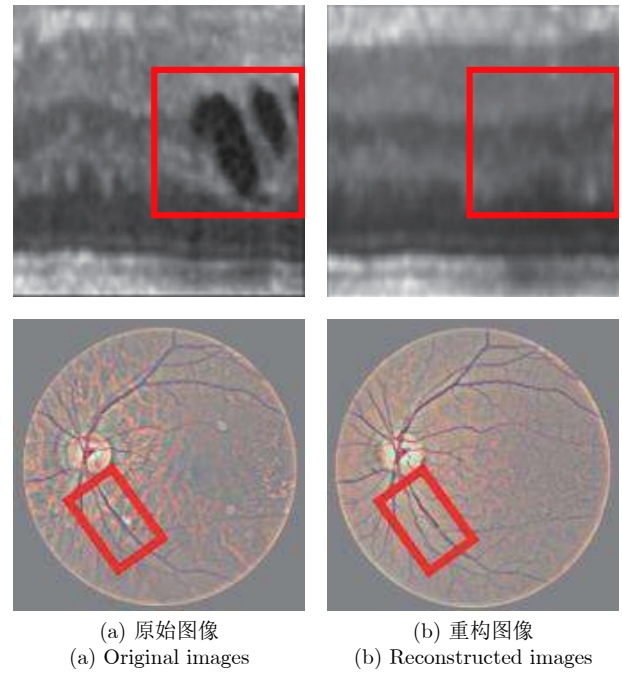


图 25 医学图像的重构^[16, 142]

Fig. 25 Reconstruction of medical images^[16, 142]

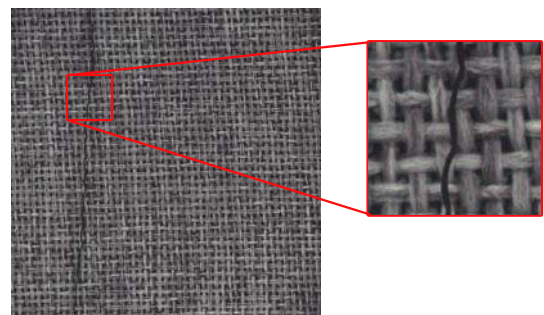


图 26 工业图像中的微小异常^[146]

Fig. 26 Tiny anomaly in industrial image^[146]

重构或许是值得进一步研究的内容.

2.2.4 结合传统方法的异常检测方法

结合传统方法和深度学习来进行异常检测的方法, 大部分是借助神经网络来实现特征提取, 然后通过传统方法对图像特征进行快速分类以实现图像异常检测. Gupta 等^[150]借助用 ImageNet 预训练的网络作为特征提取器, 然后借助 OC-SVM 实现对异常图像的分类. Napoletano 等^[151]对网络提取到的特征利用主成分分析^[152]和聚类算法构建特征向量的字典来进行异常检测. Wang 等^[82]在利用预训练网络进行特征提取的基础上, 使用多个超平面和子空间以获得比 OC-SVM 更加贴合正常样本分布的决策边界.

上述这些方法相比于人工设计特征具有更好的

效果, 不过在自然图像数据集上预训练的网络或许并不适合其他类型比如工业图像和医学图像, 在这种情况下就需要利用目标类型的数据集对网络进行特征提取的训练. 常用的方式是先利用自编码器进行特征提取, 然后计算与正常样本特征向量之间的欧氏距离^[153]以进行异常检测. Sun 等^[154]利用稀疏编码来学习变分自编码器隐藏层特征向量的表示方法, 通过特征向量的重构误差来进行异常检测. Alaverdyan 等^[155]在自编码器的基础上结合暹罗网络的思想进行训练, 在重构输入图像的同时减小隐变量之间的距离, 随后在特征图上构建 OC-SVM 以实现异常区域的定位. Burlina 等^[156]提出了一种更为通用的检测框架, 除了使用预训练的网络, 还结合 GAN 作为生成式特征提取器, 而对于得到的图像特征, 则采用 OC-SVM 或者孤立森林^[157]等传统方法进行异常图像检测. Kozerawski 等^[158]则是利用迁移学习的方法学习了从图像特征向量到支持向量机 (Support vector machine, SVM) 决策面的映射, 然后仅使用一张正常图像就可以得到 SVM 的分类面实现异常检测.

大部分上述方法在决策阶段采用了传统的方式, 在检测速度上会优于深度学习的方法, 相比于传统异常检测方法则有更高的精度和更好的通用性. 然而, 这类方法在检测精度上要逊于最新的深度学习方法, 所以有一种方式是在现有深度学习方法的基础上, 在模型训练过程中结合传统方法来进一步提升检测的精度^[159], 如 Nie 等^[160]采用高斯混合模型对自编码器隐变量的分布进行建模, 在并不显著影响算法效率的同时提升了模型对全局异常的检测能力.

表 3 总结了各类基于深度学习的图像异常检测方法的设计思路和优缺点. 整体来看, 这些方法得益于神经网络强大的学习能力, 可以适用于各种纹理和结构下图像的异常检测, 因此在检测精度和通用性上要明显优于传统方法. 不过相对应的, 这些方法也更为复杂, 需要设计各种策略来保证网络的顺利训练.

3 图像异常检测数据集

图像异常检测相关的研究方兴未艾, 目前有许多识别定性异常的相关文章是在传统图像分类数据集上开展的, 诸如 MNIST^[89]、Fashion-MNIST^[28]、CIFAR-10^[90]等等. 而对于定量异常的检测任务, 所使用的数据集就与具体的应用领域相关, 如表 4 所示.

在工业外观检测领域, 对于织布检测, 常用的有 TILDA (Textile texture database)^[161]和 PFID (Patterned fabric image database)^[162]等数据集. TILDA 是最为常用的织布图像数据集之一, 包含 8 种代表性的纹理图像总计 3200 张, 每种图像提供了正常图像和 7 种缺陷图像, 不过没有提供像素级的标注. PFID 则是由香港大学提供的一个包含 3 种花纹织布的图像数据集, 每种图像都提供了数十张正常和异常图像, 并且进行了像素级的标注. 对于金属表面检测, 有 MT (Magnetic tile defect datasets)^[163]、RSDDs (Rail surface discrete defects datasets)^[164]和 NEU (Northeastern university surface defect database)^[165]等. MT 数据集包含 5 类从不同光照条件下采集到的磁瓦表面缺陷图像, 每类包含数十张缺陷图像并且提供了像素级

表 3 基于深度学习的图像异常检测技术的分类和特点

Table 3 The classification and characteristic of deep learning based image anomaly detection

方法类别	设计思路	优点	缺点	参考文献
距离度量	将正常图像映射到指定区域内, 并减小正常特征之间距离, 根据待测图像的特征到聚类中心的距离进行异常检测	模型结构简单, 适用范围广	模型可能出现退化, 需要设计额外的辅助任务, 且无法准确定位异常区域	[88-98]
分类面构建	通过几何变换扩增现有数据, 直接训练分类模型并利用置信度来检测异常	模型训练较为简单, 语义信息提取能力更强, 异常检测精度很高	几何变换的操作在纹理图像等场景下并不适用	[101-102]
	寻找与正常样本近似的图像作为负样本来训练二分类网络, 构建正常图像与潜在异常图像间的分类面	应用场景广泛, 异常检测精度高	需要精心设计损失函数和生成的负样本, 模型设计复杂	[104-117]
图像重构	利用自编码器等模型学习正常图像的表达方式, 并根据待测图像的重构误差来进行异常检测	训练阶段无需引入额外的样本, 且应用场景广泛, 速度较快	一般的方法重构结果较为模糊, 且缺乏更为高效可靠的方法避免重构出异常区域	[118-134]
	利用GAN来获得更为清晰的图像重构效果	应用场景广泛, 异常区域定位精度高	模型训练复杂, 而且缺乏理论上的保证	[135-147]
结合传统方法	利用预训练的网络或者自编码器模型对图像进行特征提取, 在决策阶段利用传统方法进行异常检测	相比传统方法精度更高通用性更好, 且速度较快	在检测精度上略有不足	[150-160]

表 4 图像异常检测常用数据集

Table 4 Common datasets for image anomaly detection

应用场景	数据集名称	参考文献
工业	布匹	TILDA [161]
		PFID [162]
		MT [163]
	金属	RSDD [164]
		NEU [165]
	纳米材料	NanoTWICE [71]
	综合	MVTec AD [146]
医学	大脑	BraTS [167]
	视网膜	AMD [168]
高光谱	混合	AVIRIS [169]
		ABU [170]

的标注,同时提供了大量的正常图像作为参考. RSDD 则是一个钢轨表面缺陷数据集,包含两类图像共计 195 张并且提供了像素级标注. NEU 包含了 6 类图像,每类有 300 张热轧带钢表面缺陷图像,不过仅以边界框 (bounding box) 的形式提供标注, Zhou 等^[166]则是为其提供了像素级的标注. 而 NanoTWICE (Nanocomposite nanofibres for treatment of air and water by an industrial conception of electrospinning)^[71]是一个常用的纳米材料图像数据集,包含 45 张利用扫描电子显微镜得到的纳米材料图像,其中 5 张为正常图像,其余 40 张则带有各种缺陷并且提供了像素级的标注. MVTEC AD (MVTEC anomaly detection dataset)^[146]是一个综合了工业生产中各种常见材质的图像异常检测数据集,包含了 5 种纹理图像和 10 种物体图像,每种图像包含 60 至 320 张正常图像以及几十张带有异常区域的测试图像,并且提供了像素级的人工标注,常用于异常定位方法的验证. 而在医学领域,有 BraTS (Brain tumor image segmentation benchmark)^[167]和 AMD (Age-related macular degeneration)^[168]等. BraTS 包含一共 65 张低级别和高级别神经胶质瘤的多模态核磁共振图像,并且提供了精细的人工标注. AMD 是杜克大学整理的一个针对老年性黄斑病变的大型图像数据集,包含来自 115 个正常眼部和 269 个患者眼部的总计 38400 张谱域光学相干层析扫描图像 (Spectral domain optical coherence tomography, SD-OCT),并且提供了精细的像素级标注. 而对于高光谱图像则有 AVIRIS (Airborne visible infrared imaging spectrometer)^[169]、ABU (Airport-beach-urban)^[170]等常用数据,囊括了海岸、城市和机场等各种场景的高光谱图像.

除了上述这些数据集,还有很多相关的数据集,但限于篇幅等原因此处不再展示. 而 MVTEC AD 虽然刚刚于 CVPR2019 上提出但已经有许多相关方法在其上进行了实验,如表 5 所示.

表中受试者工作特性曲线下面积 (Area under receiver operating characteristic curve, AUROC) 和区域重叠分数 (Per-region-overlap score, PRO-score) 是常用的两个衡量异常定位效果的指标^[171]. AUROC 中的 ROC (Receiver operating characteristic curve, ROC) 是模型在不同分类阈值下真阳性率和假阳性率的变化曲线,而 AUROC 是一个整体的评价指标,越高说明其模型分类效果越好. 不过有文献^[97]指出 AUROC 对于一些面积较大的缺陷会比较宽容,因此提出了 PRO-score. PRO-score 同样也是在一系列阈值下构建性能曲线,并以曲线下面积作为综合评估指标. 不同的是,PRO-score 统计的是不同阈值下的区域重叠率 (Per-region-overlap, PRO),PRO 是以二值化后连通域和真值图之间的相对重叠率作为每一个阈值下的模型分类性能. 从表 5 中可以看到,许多方法采用图像重构或者距离度量的方式进行异常区域的定位. 虽然前者在速度上更有优势,但精度上往往不如距离度量类的方法,这可能是源于图像重构对图像细节的丢失,也有可能是自编码器容易残留异常区域的问题所导致的. 而基于距离度量类的方法则没有上述这些潜在的问题,特别是 Defard 等^[171]所提出的方法在两个指标上都实现了最高的性能. 但从实际的检测结果来看,该方法速度较慢,而且会更倾向于提高召回率,在精准率上的优势并不明显,特别是在一些微小缺陷的检测上往往因为异常区域的响应值较低而导致较多的误检现象,这也表明基于距离度量的方法对于一些微弱异常的检测效果还有待提升.

4 图像异常检测问题面临的挑战

异常检测一般是在没有真实异常样本的情况下进行模型训练,这种特点使得异常检测任务面临着不小的挑战.

1) 异常样本的未知性. 在异常检测当中,一般仅有正常样本可供使用,由于异常样本的未知性,传统的基于监督学习的目标识别算法难以直接应用到异常检测领域当中. 这使得研究人员需要设计新的模型建立方法或者网络训练方法来进行异常检测. 而且仅利用正常样本训练得到的异常检测模型对实际异常样本的检测还存在一定的风险,依然可能会遗漏一些人眼认为较为显著的异常目标.

表 5 各图像异常定位方法在 MVTEC AD 上的性能
Table 5 Performance of image anomaly localization methods on MVTEC AD

方法	大致思路	定位性能	
		AUROC	PRO-score
AE ^[146]	利用自编码器进行图像重构	0.817	0.790
AnoGAN ^[145]	利用GAN中的生成器进行图像重构	0.743	0.443
Iterative Projection ^[134]	在图像重构基础上采用迭代优化寻找最优的正常图像	0.893	—
AESc ^[172]	利用蒙特卡洛对重构网络进行Dropout并利用预测不确定性进行异常定位	0.86	—
P-Net ^[146]	在图像重构过程中添加对纹理结构的约束	0.89	—
Uninformed Students ^[97]	联合考虑待测图特征到目标特征之间的距离和方差进行异常定位	—	0.857
CAVGA ^[144]	在图像重构的基础上采用注意力定位异常区域	0.93	—
FCDD ^[96]	利用全卷积网络提取特征并以偏置项作为特征映射中心	0.96	—
Patch SVDD ^[173]	计算待检图像片和最近似的正常图像片之间的距离进行异常定位	0.957	—
PaDiM ^[171]	用预训练的网络进行特征提取, 利用多维高斯模型进行异常定位	0.975	0.921
SPADE ^[174]	寻找待测样本的K-近邻正常图像作为参考, 再通过距离度量进行异常检测	0.965	0.917

2) 异常定义的不清晰. 由于仅拥有正常样本, 对于异常的定义存在一定的难度, 比如异常程度到多少为异常, 如果设定太过严苛, 可能会导致很多因噪声而产生的误检出现, 而如果太过宽松又会使得一些较微弱的异常项被判定成正常. 但又缺乏足够的真实异常样本来辅助这一决策过程, 使得现阶段检测方法往往较为严苛, 容易出现较多的误检区域.

3) 微弱异常的定位. 如第 2 节所说, 图像异常检测一般有分类和定位两个类型. 对于异常图像分类任务来说, 异常样本和正常样本之间存在明显差异. 利用人工设计的特征或者预训练好的神经网络进行特征提取就有望将两者的特征向量区分开. 但是对于异常定位任务而言, 图像中一般只有一部分区域出现了异常, 而且经常会出现面积较小的目标, 比如在工业外观检测过程中可能会出现宽度仅有 7 个像素的细微异常区域, 也可能出现一些对比度较弱的异常区域^[146]. 在高光谱图像异常检测或者医学图像中病变区域的定位中, 目标区域的面积一般都只占整张图像很小的比例, 使得异常区域的定位较为困难.

4) 维数灾难. 异常检测是一个从数据挖掘领域中发展而来的概念, 因此早期的方法也大多是针对低维数据设计的^[79, 81], 而这些方法在面临高维数据时其检测性能会受到严重影响. 而图像数据是一个典型的高维数据, 即便是最为基础的 Mnist 数据集, 如果仅仅是直接地将其转换成向量也会形成长达 784 维的向量, 这使得一些在数据挖掘中常用的异常检测算法很难直接用于图像数据.

5) 算法的通用性. 不同类型的图像数据差别很大, 其实际检测的目标也不尽相同, 导致现阶段许

多异常检测算法是针对某一类图像而开发的. 较低的通用性使得现有算法难以应用到新的图像类型当中.

5 展望

本文对近年来图像异常检测方法的发展状况进行了回顾, 可以看到针对这一问题已经有了一定数量的研究. 关于未来可能的研究方向, 我们认为可以从以下几个角度进行考虑:

1) 构建更为高效的异常检测算法. 对于异常检测而言, 不仅仅需要对待检图像进行正常与否的判断, 往往还需要对异常区域进行定位. 比如工业图像表面的缺陷检测, 医学图像中病变区域的定位等等. 然而, 由于在训练阶段没有任何关于异常区域的标注信息, 传统的目标检测或者图像分割的方法无法直接应用到异常检测任务中. 因此, 现有的方法大多采用的是将待检图像切分成一系列的图像块, 然后分块进行异常与否的二分类来进行异常区域的定位. 而且, 为了获得异常区域的准确轮廓, 这种切分的步长一般较小, 会显著影响算法的效率. 现有的一些方法比如频谱分析虽然能够同时处理整张图像以实现高效的定位, 但该方法对于图像有一定的要求. 而基于深度学习的图像重构方法虽然没有上述约束, 但重构图像中残留的异常区域会影响后续的检测. 因此, 如何兼顾检测精度和实时性仍需进一步的探索.

2) 小样本/半监督学习. 现阶段的异常检测方法大部分仅利用正常样本来训练模型. 但是在实际检测任务中, 并不是完全无法获取真实的异常样本. 比如在工业外观检测任务中, 少量的缺陷样本是可以获取的. 而且对几张缺陷图像进行标注并不会显

著地增加训练成本. 而且相关文献^[96]初步尝试了在训练过程中引入一张真实异常图像并且获得了一定的效果提升. 因此可以考虑结合小样本学习的方法, 利用大量正常样本和几张真实的异常样本来进行模型训练以提高性能. 而有些异常检测任务面临的是严格无监督的环境^[98], 连所用样本是否属于正常样本也不可预知, 此时训练样本中存在的少量异常样本就会对模型的训练产生性能上的影响, 如果采用半监督的训练方式, 对少量正常和异常样本进行标注, 可以有效提升模型对潜在异常样本的检测能力. 但是这种方法还是会面临一些问题, 比如采集到的异常样本显然不可能囊括所有类别, 如何让模型兼顾对已知类型和未知类型异常样本的检测能力, 也是一个待研究的任务.

3) 更自适应的样本合成方法. 在许多相关的文献中^[105, 108, 110]都已经证明了在模型训练阶段, 引入各种人工构造的异常图像能有效地提升检测性能. 即便构造的异常图像与真实的异常图像并不相同, 额外增加的异常图像可以提升分类面的贴合度或者背景重构的稳定度, 这都可以增加模型对潜在异常图像的检测能力. 但相关文献表明这些额外的异常样本越接近与正常样本模型的性能越好^[105]. 然而, 相关方法中额外使用的异常图像大多是采集自别的数据集, 这些图像一般与正常样本的分布之间存在较为明显的差异. 虽然有方法尝试采用梯度上升的方式合成异常图像, 但该方法在更为复杂的图像上的结果还有待论证. 因此, 如何针对各种正常图像自适应地合成异常样本也是一个有待解决的问题.

4) 轻量化网络设计. 基于深度学习的异常检测方法得益于神经网络强大的学习能力往往能得到比传统方法更优秀的性能, 但代价是需要更多的计算量和更长的处理时间. 对于一张待测图像, 需要利用深层神经网络提取特征向量以区分正常和异常样本, 而且重构类的方法还需要再次经过第二个深层神经网络来重构输入图像. 因此, 更为轻量化的网络设计能够减少方法的运行时间. 此外, 大多数方法在验证时硬件条件较好, 而实际生产现场要部署同等算力的设备会需要较高的成本, 因此, 轻量化的网络设计在减少计算量的同时, 还能降低对硬件设备的需求, 降低在实际应用中的成本. 针对这一问题, 现阶段常用的有两类方法: a) 轻量模型设计, 设计更为高效的网络计算方法以实现减小模型体积的同时保持性能不变, 例如 MobileNet^[175]等. 也可以采用知识蒸馏的方式, 用复杂网络的输出作为目标来训练一个更为简单的网络; b) 模型压缩, 有通过剪枝的方式剔除冗余的权重以减小模型大小, 也

有通过网络量化的方式, 以牺牲一定精度为代价减小网络参数所占空间, 其中二值化模型具有突出的压缩性能, 更利于模型部署.

5) 更高精度的异常定位方法. 对于异常定位任务, 现有的方法大多会采用滑窗的方式将原始图像分解成一系列小的图像区域, 然后再利用异常分类的方式对每一个区域进行异常与否的分析. 这种分块分析的方式无法精准地定位异常区域, 处于异常纹理与正常纹理的交界处的图像区域也很有可能被误判为异常. 而对于能直接定位异常的图像重构类方法, 又会因自身重构精度的限制, 在正常纹理区域也会出现差异, 这也会影响对一些微弱异常区域的定位效果. 在医学和工业等领域内异常目标的检测中, 不仅要关注召回率, 异常检测的精准率也十分重要. 但从现有方法的效果看, 许多方法主要在召回率方面性能优异, 因为在实际应用领域中漏检的危害性远高于误检. 但如果能够在保证召回率的同时提高精准率, 尽可能减少后续人工或者算法的二次处理, 异常检测方法将能更广泛地应用在相关领域中. 因此, 如何精准定位异常区域并减少对正常图像区域的误判情况, 同样也是一个值得研究的问题.

6 结论

图像异常检测的目标是在真实异常样本难以获取的情况下, 构建分类模型对潜在的异常图像进行检测. 本文首先介绍了异常的定义和常见的类型. 然后按照具体的检测思路, 分别对传统的和基于深度学习的图像异常检测方法的发展现状进行了综述和分析, 并且给出了未来可行的研究方向. 现阶段对于图像异常检测方法已经有了较多的研究, 但依然可以在算法的检测效率、小样本/半监督学习以及更自适应的样本合成方法等方面有进一步的发展.

References

- 1 Chandola V, Banerjee A, Kumar V. Anomaly detection: A survey. *ACM Computing Surveys*, 2009, **41**(3): 15
- 2 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
- 3 Tian Z, Shen C H, Chen H, He T. FCOS: Fully convolutional one-stage object detection. In: Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea: IEEE, 2019. 9626-9635
- 4 Choi J, Kim C. Unsupervised detection of surface defects: A two-step approach. In: Proceedings of the 19th IEEE International Conference on Image Processing (ICIP). Orlando, USA: IEEE, 2012. 1037-1040
- 5 Griffin L D, Caldwell M, Andrews J T A, Bohler H. "Unexpected item in the bagging area": Anomaly detection in X-ray security images. *IEEE Transactions on Information Forensics*

- and Security*, 2019, **14**(6): 1539–1553
- 6 Ahn E, Kumar A, Feng D G, Fulham M, Kim J. Unsupervised deep transfer feature learning for medical image classification. In: Proceedings of the 16th International Symposium on Biomedical Imaging (ISBI). Venice, Italy: IEEE, 2019. 1915–1918
 - 7 Agyemang M, Barker K, Alhadj R. A comprehensive survey of numeric and symbolic outlier mining techniques. *Intelligent Data Analysis*, 2006, **10**(6): 521–538
 - 8 Li C L, Gao G S, Liu Z F, Huang D, Xi J T. Defect detection for patterned fabric images based on GHOG and low-rank decomposition. *IEEE Access*, 2019, **7**: 83962–83973
 - 9 Lv C K, Zhang Z T, Shen F, Zhang F, Su H. A fast surface defect detection method based on background reconstruction. *International Journal of Precision Engineering and Manufacturing*, 2020, **21**(3): 363–375
 - 10 Wang J Z, Li Q Y, Gan J R, Yu H M, Yang X. Surface defect detection via entity sparsity pursuit with intrinsic priors. *IEEE Transactions on Industrial Informatics*, 2020, **16**(1): 141–150
 - 11 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
 - 12 Bai X L, Fang Y M, Lin W S, Wang L P, Ju B F. Saliency-based defect detection in industrial images by using phase spectrum. *IEEE Transactions on Industrial Informatics*, 2014, **10**(4): 2135–2145
 - 13 Zhou W J, Fei M R, Zhou H Y, Li K. A sparse representation based fast detection method for surface defect detection of bottle caps. *Neurocomputing*, 2014, **123**: 406–414
 - 14 Baur C, Wiestler B, Albarqouni S, Navab N. Bayesian skip-autoencoders for unsupervised hyperintense anomaly detection in high resolution brain mri. In: Proceedings of the 17th International Symposium on Biomedical Imaging (ISBI). Iowa City, USA: IEEE, 2020. 1905–1909
 - 15 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 (IPMI). Boone, USA: Springer, 2017. 146–157
 - 16 Zhou K, Xiao Y T, Yang J L, Cheng J, Liu W, Luo W X, et al. Encoding structure-texture relation with P-net for anomaly detection in retinal images. In: Proceedings of the 16th European Conference on Computer Vision (ECCV). Glasgow, UK: Springer, 2020. 360–377
 - 17 Wang N, Li B, Xu Q Z, Wang Y H. Automatic ship detection in optical remote sensing images based on anomaly detection and SPP-PCANet. *Remote Sensing*, 2019, **11**(1): 47
 - 18 Jiang T, Li Y S, Xie W Y, Du Q. Discriminative reconstruction constrained generative adversarial network for hyperspectral anomaly detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, **58**(7): 4666–4679
 - 19 Jiang K, Xie W Y, Li Y S, Lei J, He G, Du Q. Semisupervised spectral learning with generative adversarial network for hyperspectral anomaly detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, **58**(7): 5224–5236
 - 20 Ehret T, Davy A, Morel J M, Delbracio M. Image anomalies: A review and synthesis of detection methods. *Journal of Mathematical Imaging and Vision*, 2019, **61**(5): 710–743
 - 21 Pang G S, Shen C H, Cao L B, Van Den Hengel A. Deep learning for anomaly detection: A review. *ACM Computing Surveys*, 2022, **54**(2): 38
 - 22 Chalapathy R, Chawla S. Deep learning for anomaly detection: A survey [Online], available: <https://arxiv.org/abs/1901.03407>, January 23, 2019
 - 23 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
(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)
 - 24 Pimentel M A F, Clifton D A, Clifton L, Tarassenko L. A review of novelty detection. *Signal Processing*, 2014, **99**: 215–249
 - 25 Hawkins D M. *Identification of Outliers*. London: Chapman and Hall, 1980. 1–12
 - 26 Cook A, M\imaths\imathrl\imath G, Fan Z. Anomaly detection for IoT time-series data: A survey. *IEEE Internet of Things Journal*, 2020, **7**(7): 6481–6494
 - 27 Goldman A, Cohen I. Anomaly detection based on an iterative local statistics approach. *Signal Processing*, 2004, **84**(7): 1225–1229
 - 28 Xiao H, Rasul K, Vollgraf R. Fashion-MNIST: A novel image dataset for benchmarking machine learning algorithms [Online], available: <https://arxiv.org/abs/1708.07747>, August 25, 2017
 - 29 Jia Di, Zhu Ning-Dan, Yang Ning-Hua, Wu Si, Li Yu-Xiu, Zhao Ming-Yuan. Image matching methods. *Journal of Image and Graphics*, 2019, **24**(5): 677–699
(贾迪, 朱宁丹, 杨宁华, 吴思, 李玉秀, 赵明远. 图像匹配方法研究综述. 中国图象图形学报, 2019, **24**(5): 677–699)
 - 30 Loog M, Lauze F. The improbability of Harris interest points. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(6): 1141–1147
 - 31 Lowe D G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004, **60**(2): 91–110
 - 32 Morel J M, Yu G S. ASIFT: A new framework for fully affine invariant image comparison. *SIAM Journal on Imaging Sciences*, 2009, **2**(2): 438–469
 - 33 Rosten E, Porter R, Drummond T. Faster and better: A machine learning approach to corner detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(1): 105–119
 - 34 Verdie Y, Yi K M, Fua P, Lepetit V. TILDE: A temporally invariant learned detector. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA: IEEE, 2015. 5279–5288
 - 35 Wang Z H, Wu F C, Hu Z Y. MSLD: A robust descriptor for line matching. *Pattern Recognition*, 2009, **42**(5): 941–953
 - 36 Zhang L L, Koch R. An efficient and robust line segment matching approach based on LBD descriptor and pairwise geometric consistency. *Journal of Visual Communication and Image Representation*, 2013, **24**(7): 794–805
 - 37 Xia M H, Liu B. Image registration by “Super-curves”. *IEEE Transactions on Image Processing*, 2004, **13**(5): 720–732
 - 38 Wolfson H J. On curve matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, **12**(5): 483–489
 - 39 Pratt W K. Correlation techniques of image registration. *IEEE Transactions on Aerospace and Electronic Systems*, 1974, **AES-10**(3): 353–358
 - 40 Leese J A, Novak C S, Clark B B. An automated technique for obtaining cloud motion from geosynchronous satellite data using cross correlation. *Journal of Applied Meteorology and Climatology*, 1971, **10**(1): 118–132
 - 41 Korman S, Reichman D, Tsur G, Avidan S. FasT-match: Fast affine template matching. In: Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Portland, USA: IEEE, 2013. 2331–2338
 - 42 Jia D, Cao J, Song W D, Tang X L, Zhu H. Colour FAST (CFAST) match: Fast affine template matching for colour images. *Electronics Letters*, 2016, **52**(14): 1220–1221
 - 43 Han X F, Leung T, Jia Y Q, Sukthankar R, Berg A C. Match-

- Net: Unifying feature and metric learning for patch-based matching. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA: IEEE, 2015. 3279–3286
- 44 Balntas V, Johns E, Tang L L, Mikolajczyk K. PN-Net: Conjoined triple deep network for learning local image descriptors [Online], available: <https://arxiv.org/abs/1601.05030v1>, January 16, 2016
- 45 Chen W J, Ho J H, Mustapha K B, Chai T Y. A vision based system for anomaly detection and classification in additive manufacturing. In: Proceedings of the 2019 IEEE Conference on Sustainable Utilization and Development in Engineering and Technologies (CSUDET). Penang, Malaysia: IEEE, 2019. 87–92
- 46 Vaikundam S, Hung T Y, Chia L T. Anomaly region detection and localization in metal surface inspection. In: Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP). Phoenix, USA: IEEE, 2016. 759–763
- 47 Herwig J, Leßmann S, Bürger F, Pauli J. Adaptive anomaly detection within near-regular milling textures. In: Proceedings of the 8th International Symposium on Image and Signal Processing and Analysis (ISPA). Trieste, Italy: IEEE, 2013. 113–118
- 48 Tsai D M, Huang C K. Defect detection in electronic surfaces using template-based Fourier image reconstruction. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 2019, **9**(1): 163–172
- 49 Reed I S, Yu X. Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1990, **38**(10): 1760–1770
- 50 Du B, Zhang L P. Random-selection-based anomaly detector for hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 2011, **49**(5): 1578–1589
- 51 Veracini T, Matteoli S, Diani M, Corsini G. Fully unsupervised learning of Gaussian mixtures for anomaly detection in hyperspectral imagery. In: Proceedings of the 9th International Conference on Intelligent Systems Design and Applications. Pisa, Italy: IEEE, 2009. 596–601
- 52 Zhang H, Jin X T, Wu Q M J, Wang Y N, He Z D, Yang Y M. Automatic visual detection system of railway surface defects with curvature filter and improved Gaussian mixture model. *IEEE Transactions on Instrumentation and Measurement*, 2018, **67**(7): 1593–1608
- 53 Schaum A. Joint subspace detection of hyperspectral targets. In: Proceedings of the 2004 IEEE Aerospace Conference Proceedings (IEEE Cat. No.04TH8720). Big Sky, USA: IEEE, 2004. Article No. 1824
- 54 Molero J M, Garzón E M, García I, Plaza A. Analysis and optimizations of global and local versions of the RX algorithm for anomaly detection in hyperspectral data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2013, **6**(2): 801–814
- 55 Mishne G, Cohen I. Multiscale anomaly detection using diffusion maps. *IEEE Journal of Selected Topics in Signal Processing*, 2013, **7**(1): 111–123
- 56 Yang En-Jun, Liao Yi-Hui, Liu An-Dong, Yu Li. Detection for fabric defects based on low-rank decomposition. *Journal of Textile Research*, 2020, **41**(5): 72–78
(杨恩君, 廖义辉, 刘安东, 俞立. 基于低秩分解的织物疵点检测. 纺织学报, 2020, **41**(5): 72–78)
- 57 Li C L, Liu C D, Gao G S, Liu Z F, Wang Y P. Robust low-rank decomposition of multi-channel feature matrices for fabric defect detection. *Multimedia Tools and Applications*, 2019, **78**(6): 7321–7339
- 58 Zhou T Y, Tao D C. GoDec: Randomized low-rank & sparse matrix decomposition in noisy case. In: Proceedings of the 28th International Conference on Machine Learning (ICML). Bellevue, USA: PMLR, 2011. 33–40
- 59 Zhang Y X, Du B, Zhang L P, Wang S G. A low-rank and sparse matrix decomposition-based Mahalanobis distance method for hyperspectral anomaly detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, **54**(3): 1376–1389
- 60 Wang J Z, Li Q Y, Gan J R, Yu H M. Fabric defect detection based on improved low-rank and sparse matrix decomposition. In: Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP). Beijing, China: IEEE, 2017. 2776–2780
- 61 Yang Y X, Zhang J Q, Song S Z, Zhang C, Liu D L. Low-rank and sparse matrix decomposition with orthogonal subspace projection-based background suppression for hyperspectral anomaly detection. *IEEE Geoscience and Remote Sensing Letters*, 2020, **17**(8): 1378–1382
- 62 Guo C L, Ma Q, Zhang L M. Spatio-temporal Saliency detection using phase spectrum of quaternion Fourier transform. In: Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Anchorage, USA: IEEE, 2008. 1–8
- 63 Zhang H X, Guo Z D, Qi Z G, Wang J G. Research of glass defects detection based on DFT and optimal threshold method. In: Proceedings of the 2012 International Conference on Computer Science and Information Processing (CSIP). Xi'an, China: IEEE, 2012. 1044–1047
- 64 Liu H X, Zhou W, Kuang Q W, Cao L, Gao B. Defect detection of IC wafer based on spectral subtraction. *IEEE Transactions on Semiconductor Manufacturing*, 2010, **23**(1): 141–147
- 65 Tsai D M, Huang T Y. Automated surface inspection for statistical textures. *Image and Vision Computing*, 2003, **21**(4): 307–323
- 66 Tsai D M, Hsieh C Y. Automated surface inspection for directional textures. *Image and Vision Computing*, 1999, **18**(1): 49–62
- 67 Hou X D, Zhang L Q. Saliency detection: A spectral residual approach. In: Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Minneapolis, USA: IEEE, 2007. 1–8
- 68 Li J, Levine M D, An X J, Xu X, He H G. Visual saliency based on scale-space analysis in the frequency domain. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(4): 996–1010
- 69 Aiger D, Talbot H. The phase only transform for unsupervised surface defect detection. In: Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, USA: IEEE, 2010. 295–302
- 70 Tsai D M, Kuo C C. Defect detection in inhomogeneously textured sputtered surfaces using 3D Fourier image reconstruction. *Machine Vision and Applications*, 2007, **18**(6): 383–400
- 71 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
- 72 Liang L Q, Li D, Fu X, Zhang W J. Touch screen defect inspection based on sparse representation in low resolution images. *Multimedia Tools and Applications*, 2016, **75**(5): 2655–2666
- 73 Boracchi G, Carrera D, Wohlberg B. Novelty detection in images by sparse representations. In: Proceedings of the 2014 IEEE Symposium on Intelligent Embedded Systems (IES). Orlando, USA: IEEE, 2014. 47–54
- 74 Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, 2011, **3**(1): 1–122
- 75 Carrera D, Boracchi G, Foi A, Wohlberg B. Scale-invariant anomaly detection with multiscale group-sparse models. In: Proceedings of the 2016 IEEE International Conference on Image

- Processing (ICIP). Phoenix, USA: IEEE, 2016. 3892–3896
- 76 Chen Y, Nasrabadi N M, Tran T D. Sparse representation for target detection in hyperspectral imagery. *IEEE Journal of Selected Topics in Signal Processing*, 2011, **5**(3): 629–640
- 77 Hyvärinen A, Oja E. Independent component analysis: Algorithms and applications. *Neural Networks*, 2000, **13**(4–5): 411–430
- 78 Sezer O G, Ercil A, Ertuzum A. Using perceptual relation of regularity and anisotropy in the texture with independent component model for defect detection. *Pattern Recognition*, 2007, **40**(1): 121–133
- 79 Schölkopf B, Williamson R, Smola A, Shawe-Taylor J, Platt J. Support vector method for novelty detection. In: Proceedings of the 12th International Conference on Neural Information Processing Systems. Denver, USA: MIT Press, 1999. 582–588
- 80 Amrae S, Vafaei A, Jamshidi K, Adibi P. Abnormal event detection in crowded scenes using one-class SVM. *Signal, Image and Video Processing*, 2018, **12**(6): 1115–1123
- 81 Tax D M J, Duin R P W. Support vector data description. *Machine Learning*, 2004, **54**(1): 45–66
- 82 Wang J, Cherian A. GODS: Generalized one-class discriminative subspaces for anomaly detection. In: Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea: IEEE, 2019. 8200–8210
- 83 Azami M E, Bouet R, Jung J, Hammers A, Lartzien C. Combining multi-parametric MR images for the detection of epileptogenic lesions. In: Proceedings of the 12th International Symposium on Biomedical Imaging (ISBI). Brooklyn, USA: IEEE, 2015. 122–125
- 84 Zhang L Y, Sun Y H, Meng D, Li X J. Anomaly detection for hyperspectral imagery based on incremental support vector data description. In: Proceedings of the 2010 International Conference on Multimedia Technology. Ningbo, China: IEEE, 2010. 1–4
- 85 Gurram P, Kwon H. Hyperspectral anomaly detection using an optimized support vector data description method. In: Proceedings of the 3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS). Lisbon, Portugal: IEEE, 2011. 1–4
- 86 Liu Y H, Liu Y C, Chen Y J. Fast support vector data descriptions for novelty detection. *IEEE Transactions on Neural Networks*, 2010, **21**(8): 1296–1313
- 87 Liu Y H, Lin S H, Hsueh Y L, Lee M J. Automatic target defect identification for TFT-LCD array process inspection using kernel FCM-based fuzzy SVDD ensemble. *Expert Systems with Applications*, 2009, **36**(2): 1978–1998
- 88 Ruff L, Vandermeulen R A, Göernitz N, Deecke L, Siddiqui S A, Binder A, et al. Deep one-class classification. In: Proceedings of the 35th International Conference on Machine Learning. Stockholm, Sweden: PMLR, 2018. 4390–4399
- 89 Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998, **86**(11): 2278–2324
- 90 Krizhevsky A. Learning Multiple Layers of Features from Tiny Images. Technical Report TR-2009, University of Toronto, Canada, 2009.
- 91 Kitamura S, Nonaka Y. Explainable anomaly detection via feature-based localization. In: Proceedings of the 28th International Conference on Artificial Neural Networks (ICANN). Munich, Germany: Springer, 2019. 408–419
- 92 Perera P, Patel V M. Learning deep features for one-class classification. *IEEE Transactions on Image Processing*, 2019, **28**(11): 5450–5463
- 93 Wu P, Liu J, Shen F. A deep one-class neural network for anomalous event detection in complex scenes. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, **31**(7): 2609–2622
- 94 Bergman L, Hoshen Y. Classification-based anomaly detection for general data. In: Proceedings of the 8th International Conference on Learning Representations (ICLR). Addis Ababa, Ethiopia: OpenReview.net, 2020.
- 95 He X W, Zhou Y, Zhou Z C, Bai S, Bai X. Triplet-center loss for multi-view 3D object retrieval. In: Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, USA: IEEE, 2018. 1945–1954
- 96 Liznerski P, Ruff L, Vandermeulen R A, Franks B J, Kloft M, Müller K R. Explainable deep one-class classification. In: Proceedings of the 9th International Conference on Learning Representations. Vienna, Austria: OpenReview.net, 2021.
- 97 Bergmann P, Fauser M, Sattlegger D, Steger C. Uninformed students: Student-teacher anomaly detection with discriminative latent embeddings. In: Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE, 2020. 4182–4191
- 98 Ruff L, Vandermeulen R A, Görnitz N, Binder A, Müller E, Müller K R, Kloft M. Deep semi-supervised anomaly detection. In: Proceedings of the 8th International Conference on Learning Representations (ICLR). Addis Ababa, Ethiopia: OpenReview.net, 2020.
- 99 Lee K, Lee K, Lee H, Shin J. A simple unified framework for detecting out-of-distribution samples and adversarial attacks. In: Proceedings of the 2018 Advances in Neural Information Processing Systems (NIPS). Montréal, Canada: Curran Associates Inc., 2018. 7167–7177
- 100 Hendrycks D, Gimpel K. A baseline for detecting misclassified and out-of-distribution examples in neural networks. In: Proceedings of the 5th International Conference on Learning Representations (ICLR). Toulon, France: OpenReview.net, 2017.
- 101 Golan I, El-Yaniv R. Deep anomaly detection using geometric transformations. In: Proceedings of the 32nd International Conference on Neural Information Processing Systems. Montréal, Canada: MIT Press, 2018. 9781–9791
- 102 Hendrycks D, Mazeika M, Kadavath S, Song D. Using self-supervised learning can improve model robustness and uncertainty. In: Proceedings of the 33rd International Conference on Neural Information Processing Systems. Vancouver, Canada: MIT Press, 2019. Article No. 1403
- 103 Perera P, Patel V M. Deep transfer learning for multiple class novelty detection. In: Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019. 11536–11544
- 104 Oza P, Patel V M. One-class convolutional neural network. *IEEE Signal Processing Letters*, 2019, **26**(2): 277–281
- 105 Hendrycks D, Mazeika M, Dietterich T G. Deep anomaly detection with outlier exposure. In: Proceedings of the 7th International Conference on Learning Representations (ICLR). New Orleans, USA: OpenReview.net, 2019.
- 106 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. Montréal, Canada: MIT Press, 2014. 2672–2680
- 107 Sabokrou M, Khalooei M, Fathy M, Adeli E. Adversarially learned one-class classifier for novelty detection. In: Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, USA: IEEE, 2018. 3379–3388
- 108 Sabokrou M, Fathy M, Zhao G Y, Adeli E. Deep end-to-end one-class classifier. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, **32**(2): 675–684
- 109 Yao Y, Rosasco L, Caponnetto A. On early stopping in gradient descent learning. *Constructive Approximation*, 2007, **26**(2):

- 289–315
- 110 Yang Y, Hou C P, Lang Y, Yue G H, He Y. One-class classification using generative adversarial networks. *IEEE Access*, 2019, **7**: 37970–37979
- 111 Chatillon P, Ballester C. History-based anomaly detector: An adversarial approach to anomaly detection. *Intelligent Systems and Applications*. Cham: Springer, 2020: 761–776
- 112 Zaheer M Z, Lee J H, Astrid M, Lee S I. Old is gold: Redefining the adversarially learned one-class classifier training paradigm. In: Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020. 14171–14181
- 113 Lim S K, Loo Y, Tran N T, Cheung N M, Roig G, Elovici Y. DOPING: Generative data augmentation for unsupervised anomaly detection with GAN. In: Proceedings of the 2018 IEEE International Conference on Data Mining (ICDM). Singapore: IEEE, 2018. 1122–1127
- 114 Liu Y Z, Li Z, Zhou C, Jiang Y C, Sun J S, Wang M, et al. Generative adversarial active learning for unsupervised outlier detection. *IEEE Transactions on Knowledge and Data Engineering*, 2020, **32**(8): 1517–1528
- 115 Ngo P C, Winarto A A, Kou C K L, Park S, Akram F, Lee H K. Fence GAN: Towards better anomaly detection. In: Proceedings of the 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI). Portland, USA: IEEE, 2019. 141–148
- 116 Schlachter P, Liao Y W, Yang B. Deep one-class classification using intra-class splitting. In: Proceedings of the 2019 IEEE Data Science Workshop (DSW). Minneapolis, USA: IEEE, 2019. 100–104
- 117 Goyal S, Raghunathan A, Jain M, Simhadri H, Jain P. DROCC: Deep robust one-class classification. In: Proceedings of the 37th International Conference on Machine Learning. Virtual: PMLR, 2020. 3711–3721
- 118 Hinton G E, Zemel R S. Autoencoders, minimum description length and Helmholtz free energy. In: Proceedings of the 6th International Conference on Neural Information Processing Systems. Denver, USA: Morgan Kaufmann Publishers Inc., 1993. 3–10
- 119 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
- 120 Bergmann P, Löwe S, Fauser M, Sattlegger D, Steger C. Improving unsupervised defect segmentation by applying structural similarity to autoencoders. In: Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. Prague, Czech Republic: SciTePress, 2019. 372–380
- 121 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, USA: IEEE, 2018. 1237–1242
- 122 Sakurada M, Yairi T. Anomaly detection using autoencoders with nonlinear dimensionality reduction. In: Proceedings of the MLSDA 2nd Workshop on Machine Learning for Sensory Data. Gold Coast, Australia: ACM, 2014. 4–11
- 123 Zhou C, Paffenroth R C. Anomaly detection with robust deep autoencoders. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Halifax, Canada: ACM, 2017. 665–674
- 124 Lin T Y, Dollár P, Girshick R, He K M, Hariharan B, Belongie S. Feature pyramid networks for object detection. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE, 2017. 936–944
- 125 Kwon G, Prabhushankar M, Temel D, AlRegib G. Backpropagated gradient representations for anomaly detection. In: Proceedings of the 16th European Conference on Computer Vision (ECCV). Glasgow, UK: Springer, 2020. 206–226
- 126 Zimmerer D, Petersen J, Kohl S A A, Maier-Hein K H. A case for the score: Identifying image anomalies using variational autoencoder gradients [Online], available: <https://arxiv.org/abs/1912.00003>, November 28, 2019
- 127 Chu W H, Kitani K M. Neural batch sampling with reinforcement learning for semi-supervised anomaly detection. In: Proceedings of the 16th European Conference on Computer Vision (ECCV). Glasgow, UK: Springer, 2020. 751–766
- 128 Soukup D, Pinetz T. Reliably decoding autoencoders' latent spaces for one-class learning image inspection scenarios. In: Proceedings of the OAGM Workshop 2018. Hall/Tyrol, Austria: Verlag der TU Graz, 2018. 90–93
- 129 Abati D, Porrello A, Calderara S, Cucchiara R. Latent space autoregression for novelty detection. In: Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019. 481–490
- 130 Perera P, Nallapati R, Xiang B. OCGAN: One-class novelty detection using GANs with constrained latent representations. In: Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019. 2893–2901
- 131 Tian H, Li F. Autoencoder-based fabric defect detection with cross-patch similarity. In: Proceedings of the 16th International Conference on Machine Vision Applications (MVA). Tokyo, Japan: IEEE, 2019. 1–6
- 132 Gong D, Liu L Q, Le V, Saha B, Mansour M R, Venkatesh S, et al. Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection. In: Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea: IEEE, 2019. 1705–1714
- 133 Park H, Noh J, Ham B. Learning memory-guided normality for anomaly detection. In: Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020. 14360–14369
- 134 Dehaene D, Frigo O, Combexelle S, Eline P. Iterative energy-based projection on a normal data manifold for anomaly localization. In: Proceedings of the 8th International Conference on Learning Representations (ICLR). Addis Ababa, Ethiopia: OpenReview.net, 2020.
- 135 Deecke L, Vandermeulen R, Ruff L, Mandt S, Kloft M. Image anomaly detection with generative adversarial networks. In: Proceedings of the 2018 Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Dublin, Ireland: Springer, 2018. 3–17
- 136 Lai Y T K, Hu J S, Tsai Y H, Chiu W Y. Industrial anomaly detection and one-class classification using generative adversarial networks. In: Proceedings of the 2018 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM). Auckland, New Zealand: IEEE, 2018. 1444–1449
- 137 Pidhorskyi S, Almoheisen R, Adjeroh D A, Doretto G. Generative probabilistic novelty detection with adversarial autoencoders. In: Proceedings of the 32nd International Conference on Neural Information Processing Systems. Montréal, Canada: Curran Associates Inc., 2018. 6823–6834
- 138 Tuluptceva N, Bakker B, Fedulova I, Konushin A. Perceptual image anomaly detection. In: Proceedings of the 5th Asian Conference on Pattern Recognition. Auckland, New Zealand: Springer, 2019. 164–178
- 139 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 (PRICAI). Nanjing, China: Springer, 2018. 473–481
- 140 Baur C, Wiestler B, Albarqouni S, Navab N. Deep autoencod-

- ing models for unsupervised anomaly segmentation in brain MR images. In: Proceedings of the 4th International MICCAI Brainlesion Workshop. Granada, Spain: Springer, 2018. 161–169
- 141 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 (ACCV). Perth, Australia: Springer, 2018. 622–637
- 142 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
- 143 Tang T W, Kuo W H, Lan J H, Ding C F, Hsu H, Young H T. Anomaly detection neural network with dual auto-encoders GAN and its industrial inspection applications. *Sensors*, 2020, **20**(12): 3336
- 144 Venkataramanan S, Peng K C, Singh R V, Mahalanobis A. Attention guided anomaly localization in images. In: Proceedings of the 16th European Conference on Computer Vision (ECCV). Glasgow, UK: Springer, 2020. 485–503
- 145 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 (ICCV). Venice, Italy: IEEE, 2017. 618–626
- 146 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/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019. 9584–9592
- 147 Lu Y W, Yu F, Reddy M K K, Wang Y. Few-shot scene-adaptive anomaly detection. In: Proceedings of the 16th European Conference on Computer Vision (ECCV). Glasgow, UK: Springer, 2020. 125–141
- 148 Finn C, Abbeel P, Levine S. Model-agnostic meta-learning for fast adaptation of deep networks. In: Proceedings of the 34th International Conference on Machine Learning (ICML). Sydney, Australia: PMLR, 2017. 1126–1135
- 149 Xia Y D, Zhang Y, Liu F Z, Shen W, Yuille A L. Synthesize then compare: Detecting failures and anomalies for semantic segmentation. In: Proceedings of the 16th European Conference on Computer Vision (ECCV). Glasgow, UK: Springer, 2020. 145–161
- 150 Gupta K, Bhavsar A, Sao A K. Detecting mitotic cells in HEp-2 images as anomalies via one class classifier. *Computers in Biology and Medicine*, 2019, **111**: 103328
- 151 Napoletano P, Piccoli F, Schettini R. Anomaly detection in nanofibrous materials by CNN-based self-similarity. *Sensors*, 2018, **18**(1): 209
- 152 Candès E J, Li X D, Ma Y, Wright J. Robust principal component analysis? *Journal of the ACM*, 2011, **58**(3): 11
- 153 Mujeeb A, Dai W T, Erdt M, Sourin A. One class based feature learning approach for defect detection using deep autoencoders. *Advanced Engineering Informatics*, 2019, **42**: 100933
- 154 Sun J Y, Wang X Z, Xiong N X, Shao J. Learning sparse representation with variational auto-encoder for anomaly detection. *IEEE Access*, 2018, **6**: 33353–33361
- 155 Alaverdyan Z, Jung J, Bouet R, Lartizien C. Regularized siamese neural network for unsupervised outlier detection on brain multiparametric magnetic resonance imaging: Application to epilepsy lesion screening. *Medical Image Analysis*, 2020, **60**: 101618
- 156 Burlina P, Joshi N, Wang I J. Where's wally now? Deep generative and discriminative embeddings for novelty detection. In: Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019. 11499–11508
- 157 Liu F T, Ting K M, Zhou Z H. Isolation forest. In: Proceedings of the 8th IEEE International Conference on Data Mining. Pisa, Italy: IEEE, 2008. 413–422
- 158 Kozerawski J, Turk M. CLEAR: Cumulative LEARning for one-shot one-class image recognition. In: Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, USA: IEEE, 2018. 3446–3455
- 159 Zong B, Song Q, Min R Q, Cheng W, Lumezamu C, Cho D K, et al. Deep autoencoding Gaussian mixture model for unsupervised anomaly detection. In: Proceedings of the 6th International Conference on Learning Representations (ICLR). Vancouver, Canada: OpenReview.net, 2018. 1–19
- 160 Nie L H, Zhao L P, Li K Q. Glad: Global and local anomaly detection. In: Proceedings of the 2020 IEEE International Conference on Multimedia and Expo (ICME). London, UK: IEEE, 2020. 1–6
- 161 Workgroup on Texture Analysis of DFG, TILDA textile texture database [Online], available: <https://lmb.informatik.uni-freiburg.de/resources/datasets/tilda.en.html>, August, 1996
- 162 Ngan H Y T. Patterned fabric database [Online], available: <https://ytingan.wordpress.com/codes/>, April 3, 2017.
- 163 Huang Y, Qiu C, Guo Y, Wang X, Yuan K. Surface defect saliency of magnetic tile. In: Proceedings of 14th International Conference on Automation Science and Engineering (CASE). Munich, Germany: IEEE, 2018. 612–617
- 164 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
- 165 Song K C, Yan Y H. A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects. *Applied Surface Science*, 2013, **285**: 858–864
- 166 Zhou S Y, Wu S Q, Liu H G, Lu Y, Hu N Z. Double low-rank and sparse decomposition for surface defect segmentation of steel sheet. *Applied Sciences*, 2018, **8**(9): 1628
- 167 Menze B H, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, et al. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Transactions on Medical Imaging*, 2015, **34**(10): 1993–2024
- 168 Farsiu S, Chiu S J, O'Connell R V, Folgar F A, Yuan E, Izatt J A, et al. Quantitative classification of eyes with and without intermediate age-related macular degeneration using optical coherence tomography. *Ophthalmology*, 2014, **121**(1): 162–172
- 169 Baumgardner M, Biehl L, Landgrebe D. 220 band AVIRIS hyperspectral image data set: June 12 1992 indian pine test site 3 [Online], available: <https://purr.purdue.edu/publications/1947/citations?v=1>, September 30, 2015
- 170 Kang X. Airport-beach-urban dataset (ABU) [Online], available: <http://xudongkang.weebly.com/data-sets.html>, July 3, 2017
- 171 Defard T, Setkov A, Loesch A, Audigier R. PaDiM: A patch distribution modeling framework for anomaly detection and localization. In: Proceedings of the International Conference on Pattern Recognition. Springer, 2020. 475–489
- 172 Collin A S, De Vleeschouwer C. Improved anomaly detection by training an autoencoder with skip connections on images corrupted with stain-shaped noise. In: Proceedings of the 25th International Conference on Pattern Recognition (ICPR). Milan, Italy: IEEE, 2021. 7915–7922
- 173 Yi J H, Yoon S. Patch SVDD: Patch-level SVDD for anomaly detection and segmentation. In: Proceedings of the 16th Asian Conference on Computer Vision (ACCV). Kyoto, Japan: Springer, 2020. 375–390
- 174 Cohen N, Hoshen Y. Sub-image anomaly detection with deep pyramid correspondences [Online], available: <https://arxiv.org/abs/2005.02357>, February 3, 2021.

- 175 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 [Online], available: <https://arxiv.org/abs/1704.04861>, April 17, 2017.

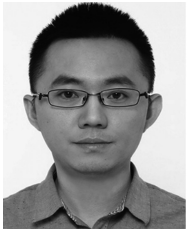


吕承侃 中国科学院自动化研究所精密感知与控制研究中心博士研究生. 2017 年获得山东大学学士学位. 主要研究方向为神经网络, 计算机视觉与图像处理.

E-mail: lvchengkan2017@ia.ac.cn

(**LV Cheng-Kan** Ph. D. candidate

at the Research Center of Precision Sensing and Control, Institute of Automation, Chinese Academy of Sciences. He received his bachelor degree from Shandong University in 2017. His research interest covers neural networks, computer vision and image processing.)



沈 飞 中国科学院自动化研究所精密感知与控制研究中心研究员. 2012 年获得中国科学院自动化研究所博士学位. 主要研究方向为视觉检测, 机器人视觉控制与微装配.

E-mail: fei.shen@ia.ac.cn

(**SHEN Fei** Professor at the Re-

search Center of Precision Sensing and Control, Institute of Automation, Chinese Academy of Sciences. He received his Ph. D. degree from Institute of Automation, Chinese Academy of Sciences in 2012. His research interest covers visual inspection, robot vision

control and micro-assembly.)



张正涛 中国科学院自动化研究所精密感知与控制研究中心研究员. 2010 年获得中国科学院自动化研究所博士学位. 主要研究方向为视觉测量, 微装配与自动化. 本文通信作者.

E-mail: zhengtao.zhang@ia.ac.cn

(**ZHANG Zheng-Tao** Professor at

the Research Center of Precision Sensing and Control, Institute of Automation, Chinese Academy of Sciences. He received his Ph. D. degree from Institute of Automation, Chinese Academy of Sciences in 2010. His research interest covers visual measurement, micro-assembly and automation. Corresponding author of this paper.)



张 峰 中国科学院自动化研究所精密感知与控制研究中心副研究员. 2012 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人控制, 机器人视觉控制与微装配.

E-mail: feng.zhang@ia.ac.cn

(**ZHANG Feng** Associate profess-

or at the Research Center of Precision Sensing and Control, Institute of Automation, Chinese Academy of Sciences. He received his Ph. D. degree from Institute of Automation, Chinese Academy of Sciences in 2012. His research interest covers robot control, robot vision control and micro-assembly.)