

# 基于深度自编码器网络的压盖缺陷检测

张洪波, 隋文涛, 袁 林, 李长安, 逯海滨  
(山东理工大学 机械工程学院, 山东 淄博 255000)

**摘 要:**口服液压盖过程,会出现压盖不良等情况,瓶盖可能会出现划痕、刮花、表面卷曲、压盖破损等缺陷,为保证食品药品安全必须在出厂前进行检测。在基于深度学习的口服液瓶盖缺陷检测的研究过程中,使用传统卷积神经网络对口服液压盖缺陷数据集进行训练,需要进行人工标注,效率较低。为有效解决上述问题,设计出一种无监督学习的深度卷积去噪自编码器网络模型用于口服液瓶盖质量检测,并使用结构相似性 SSIM 作为损失函数。针对口服液压盖质量图像进行预处理,建立合格产品图像数据集,然后构建一种以卷积神经网络为基础,结合多层感知器的去噪自编码器网络模型。该模型仅以无缺陷产品图像进行训练并学习无缺陷产品特征,通过将缺陷图像重构为无缺陷图像,再与缺陷图像相减,获得包含缺陷信息的残差图。实验结果表明:该方法能够很好地识别口服液瓶盖缺陷,准确率达到 95.2%,且具有较好的泛化能力和鲁棒性。

**关键词:**压盖质量;缺陷检测;自编码器;神经网络;多层感知器

**中图分类号:**TP391

**文献标识码:**A

**文章编号:**1673-629X(2022)02-0143-05

doi:10.3969/j.issn.1673-629X.2022.02.023

## Capping Defect Detection Based on Deep Autoencoder Network

ZHANG Hong-bo, SUI Wen-tao, YUAN Lin, LI Chang-an, LU Hai-bin  
(School of Mechanical Engineering, Shandong University of Technology, Zibo 255000, China)

**Abstract:** In the process of oral hydraulic cap, there will be problems such as poor capping, and the bottle cap may have defects such as scratches, scratches, surface curling and capping damage. To ensure the safety of food and drugs, it must be tested before leaving the factory. In the research process of oral liquid bottle capping defect detection based on deep learning, the traditional convolutional neural network is used to train the oral hydraulic cap defect data set, which requires manual annotation and has low efficiency. In order to effectively solve the above problems, an unsupervised learning deep convolution denoising autoencoder network model is designed for the quality detection of oral liquid bottle caps, and the structural similarity SSIM is used as the loss function. The quality image of the oral hydraulic cover is preprocessed to establish a qualified product image data set, and then a denoising autoencoder network model based on a convolutional neural network combined with a multi-layer perceptron is constructed. The model only uses non-defective product images for training and learns the features of non-defective products. The defect images are reconstructed into non-defective images and then subtracted from the defective images to obtain a residual image containing defect information. The experiment shows that the proposed method can well identify the capping defects of oral liquid bottles, with an accuracy rate of 95.2%, and has ideal generalization ability and robustness.

**Key words:** capping quality; defect detection; autoencoder; neural network; multilayer perceptron

## 0 引言

口服液是一种中药制剂,在国内得到了众多消费者的青睐,有巨大的消费市场。由于口服液产品在生产中设计多道工序,受生产工艺的影响,不可避免地会产生一些缺陷产品。如果这些缺陷产品在生产环节未能检测出来,不仅会影响产品质量,更有甚者会严重危害消费者身体健康,其中,压盖质量尤为重要。在口服

液灌装过程中,药液灌入口服液瓶中,由压盖机自动贴合至瓶口,由于存在机械系统不稳定等因素,压盖过程会出现一定偏差,造成压盖不良等情况,瓶盖可能会出现划痕、刮花、表面卷曲、压盖破损等缺陷,因此需要将它们检测出来并剔除。现如今,口服液生产厂家使用的口服液瓶盖一般为铝塑撕开盖,整体分为两个部分,上部分为塑料材质的撕开盖,下部分为铝制的压盖部

收稿日期:2021-02-07

修回日期:2021-06-11

基金项目:山东省重点研发计划(重大科技创新工程)项目(2018CXGC0602)

作者简介:张洪波(1995-),男,硕士,研究方向为计算机视觉;隋文涛,博士,教授,研究方向为故障诊断与模式识别。

分。如图 1 所示,图(a)为正常压盖图像,图(b)为划痕缺陷,图(c)为压盖卷曲缺陷,图(d)为压盖破损缺陷,图(e)为压盖不良缺陷,图(f)为完全未压盖缺陷。

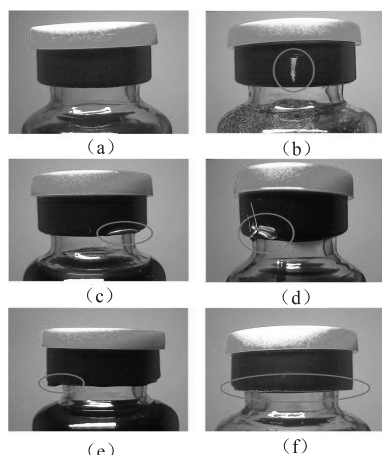


图 1 缺陷示意图

最传统检测方法一般采用人工检测,但随着传统机器视觉检测<sup>[1]</sup>方法的迅猛发展,很快被淘汰。因机器视觉检测具有成本低、效率高等优点而被广泛应用于口服液生产检测环节。近些年,一些国外、国内公司生产的灯检机大多也基于传统机器视觉检测。湖南大学刘学兵<sup>[2]</sup>基于传统机器视觉检测方法,提出了一种基于水平截距投影的方法提取瓶盖下边缘点作为缺陷检测信息,并提出了边缘点均值和方差两种数据特征作为缺陷判断数据。虽然传统机器视觉检测虽说已经取得了不错的检测成果,但只能针对特定的一类产品,若企业更新包装则需重新编写特定程序。近几年,深度学习<sup>[3]</sup>在计算机视觉、自然语言处理、语音识别等领域都取得了较好的成果,在缺陷检测领域,基于深度学习的检测方法也在迅猛发展,Allen Zhang 等人<sup>[4]</sup>使用深度学习网络在沥青表面自动进行像素级路面裂缝检测;CV Dung 等人<sup>[5]</sup>提出了基于深度学习技术的基于视觉的混凝土表面裂缝自动检测。这些深度学习检测方法虽然在检测方面很有效,但其都需要人工进行数据标注。在工业生产中,一是缺陷产生具有不确定性,二是缺陷产品相比于正常产品比例极低。因此,异常检测<sup>[6]</sup>中的无监督检测逐渐受到关注。

Bergmann P<sup>[7]</sup>提出了一种基于结构相似性与自编码器结合的无监督缺陷分割,用于织物缺陷检测,经验证该方法可以有效地检测织物缺陷;合肥工业大学罗月童等<sup>[8]</sup>提出了一种卷积去噪自编码器检测芯片表面弱缺陷,实验验证了该方法在芯片表面弱缺陷检测中有良好效果。综上所述,深度学习在缺陷检测方面取得了良好的效果,但还未有基于无监督学习的检测方法应用在口服液瓶盖质量检测中,另外口服液压盖缺陷具有缺陷与背景对比度低、缺陷较小等特点,又

给传统检测方法带来了挑战。因此,该文设计一种无监督学习<sup>[9]</sup>的深度卷积去噪自编码器网络模型用于口服液瓶盖质量检测,为口服液压盖质量检测提出一种新思路、新方法。

## 1 方法概述

### 1.1 自动编码器原理

自编码器<sup>[10-12]</sup> (Autoencoder, AE) 是一种无监督学习方法。所谓无监督学习,就是不需要对目标提前标注的深度学习方式。自编码器是通过训练来达到重构数据自身的一种人工神经网络,即一种典型的输入等于输出的无监督神经网络模型。自编码器通常包含三个部分:编码器 (Encoder)、隐藏层、解码器 (Decoder)。在自编码器的训练过程中,输入经过编码这一过程后,编码器对输入数据进行编码压缩,并且将输入数据压缩映射为一特征向量,该特征向量虽然维度较低,但其包含着输入数据的隐藏特征信息,解码器会根据特征向量重构出输入数据,先编码再解码这一过程即可学习到输入数据的密集表征。在自编码器三个部分中,隐藏层是最重要的部分,因为数据的密集表征远比输入数据维度低,相当于用更少的特征去表达输入数据,这使得自编码器可以用于降维压缩,这在图像分类上有重要意义,其原理图如图 2 所示。

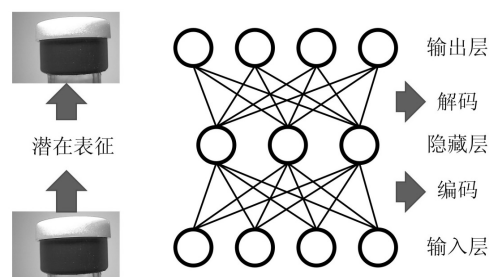


图 2 自编码器原理图

编码过程就是编码器通过函数  $f$  将原始输入数据  $x = [x_1, x_2, \dots, x_n]^T$  映射到隐含层  $h$ , 解码过程是解码器通过解码函数  $g$  将隐含层作为输入, 得到重构数据  $z = [z_1, z_2, \dots, z_r]^T$ , 其公式表达如下:

$$h = f(W_e x + b_e) \quad (1)$$

$$z = g(W_d h + b_d) \quad (2)$$

式中,  $W_e$  和  $b_e$  分别是编码器的权重和偏置,  $W_d$  和  $b_d$  是解码器的权重和偏置。自编码器在训练过程中,使得重构数据与原始输入尽量逼近,从而表明隐含层学到了原始数据另外一种潜在的表达方式。因此,给定一个数据集  $x = [x_1, x_2, \dots, x_n]^T$ , 自编码器通过最小化以下函数来优化模型参数  $\theta$ :

$$\theta = \arg \min_{\theta} \frac{1}{N} \sum_{i=1}^N L(x_i, g(f(x_i))) \quad (3)$$

其中,  $L$  为重构函数,一般使用平方误差损失函数或者

交叉熵损失函数。

## 1.2 模型介绍

上文介绍了简单的自编码器的原理,该文使用卷积去噪自编码器进行口服液盖质量检测。卷积去噪自编码器(convolutional denoising autoencoder, CDAE)中的去噪就是在自编码器的基础上,在编码器与解码器中使用卷积神经网络并人为添加噪声对输入数据进行训练。其中输入数据为  $x = [x_1, x_2, \dots, x_n]^T$ , 首先对输入数据引入噪声:

$$\begin{cases} \tilde{x} = x + \varepsilon \\ \varepsilon \sim E(\vartheta) \end{cases} \quad (4)$$

式中,  $\tilde{x}$  为添加噪声后的输入数据,  $E(\vartheta)$  为噪声的分布类型,  $\vartheta$  为已知参数。此时,对去噪自编码器网络而言,输入数据经人工添加噪声已被损坏,期望输出为没有噪声的数据,相较于前文所述的普通自编码器,在极大程度上增强了其鲁棒性。去噪自编码器通过最小化以下函数来优化模型参数  $\theta$ :

$$\theta = \arg \min_{\theta} \frac{1}{N} \sum_{i=1}^N L(\tilde{x}, x) \quad (5)$$

式中,  $L$  为重构函数,通常情况下,一般为逐像素误差测量函数,如  $L_2$  损失函数,其具体公式如下:

$$L_2(x, \hat{x}) = \sum_{i=0}^h \sum_{j=0}^w [x(i, j) - \hat{x}(i, j)]^2 \quad (6)$$

式中,数据  $\hat{x}$  是大小为  $h \times w$  的重构图像,  $x$  是大小为  $h \times w$  的输入图像。通过计算输入图像  $x$  与重构图像  $\hat{x}$  之间像素点灰度值差值的平方,实现模型优化。

$L_2$  损失函数计算的是两幅图像之间的像素距离,其缺陷十分明显,当重建图像中,目标物体的边缘具有定位误差时,原始图像与重构图像之间的像素距离增大,计算两者之间的残差图时,会出现大量残差,但此时重构图像中目标物体结构特征与原始图像中目标物体相同。针对上述问题,该文提出使用基于结构相似性(structural similarity, SSIM)作为优化模型的损失函数。SSIM 通过比较两幅图像之间的三个参数进行计算,包括:亮度(Luminance)、对比度(Contrast)和结构(Structure),三个参数计算公式如下:

$$l(p, q) = \frac{2\mu_p\mu_q + c_1}{\mu_p^2 + \mu_q^2 + c_1} \quad (7)$$

$$c(p, q) = \frac{2\sigma_p\sigma_q + c_2}{\sigma_p^2 + \sigma_q^2 + c_2} \quad (8)$$

$$s(p, q) = \frac{2\sigma_{pq} + c_3}{2\sigma_p\sigma_q + c_3} \quad (9)$$

式中,  $p$  和  $q$  是两个大小为  $M \times N$  的图像块,  $\mu_p$  为  $p$  的均值,  $\mu_q$  为  $q$  的均值,  $\sigma_p$  为  $p$  的方差,  $\sigma_q$  为  $q$  的方差,  $\sigma_{pq}$  为  $p$  和  $q$  的协方差,  $c_1$  与  $c_2$  为两个常数,避免公式作除零运算,一般取  $c_3 = c_2/2$ 。每次计算的时候都从

图片上取一个  $M \times N$  的窗口,然后不断滑动窗口进行计算,最后取平均值作为全局的 SSIM。

$$\text{SSIM}(p, q) = [l(p, q)]^\alpha \cdot [c(p, q)]^\beta \cdot [s(p, q)]^\gamma \quad (10)$$

将  $\alpha, \beta, \gamma$  设置为 1 时,可得到:

$$\text{SSIM}(p, q) = \frac{(2\mu_p\mu_q + c_1)(2\sigma_{pq} + c_2)}{(\mu_p^2 + \mu_q^2 + c_1)(\mu_p^2 + \mu_q^2 + c_2)} \quad (11)$$

SSIM 通过考虑亮度、对比度和结构信息,而不是简单地比较单个像素值,相对于  $L_2$  函数,在重建过程中对小的定位误差不太敏感,并且 SSIM 函数作为损失函数,检测出的缺陷表现在目标物体结构的变化,而不是像素强度之间的差异。

卷积去噪自编码器结合了卷积神经网络<sup>[13-15]</sup>中的卷积层池化层操作来进行特征提取,能够很好地保留图像的空间信息,又能够进行无监督式学习,训练速度更快。该文使用基于卷积神经网络(CNN)和多层感知器(MLP)组合的去噪自编码器,其结构如图3所示。

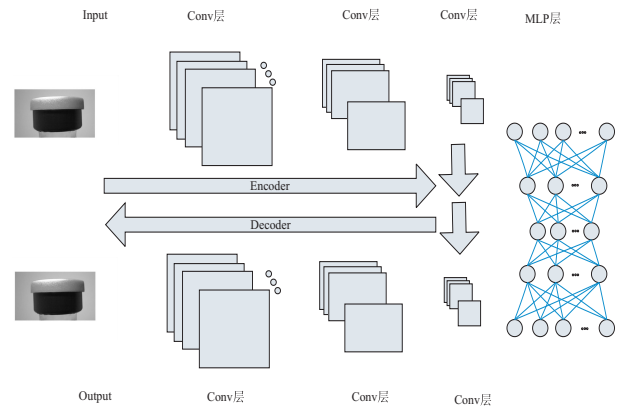


图3 模型结构

模型具体参数如下:

(1)输入为  $128 \times 128 \times 1$  像素单通道 PNG 格式灰度图像。

(2)编码器:包含 4 个卷积层、4 个最大池化层和 4 个多层感知器。池化层统一为最大池化层,可以最大限度保留瓶盖上的缺陷。编码器具体结构见表 1。

(3)解码器:解码器与编码器的相反版本。包含 4 个多层感知器、4 个上采样层和 4 个卷积层。4 个卷积层中,卷积核 Kernel 都设为  $3 \times 3$ 。

(4)输出为  $128 \times 128 \times 1$  像素单通道 PNG 灰度图像。

在编码器与解码器中,除了解码器最后一层 Conv4-Deconder 使用 Sigmoid 作为激活函数,其余卷积层与 MLP 层皆使用 Relu 函数作为激活函数。另外在训练中,为防止过拟合,分别在编码器 Conv2 与 Conv4 后,以及解码器 Conv1 与 Conv3 后,增加 Dropout 层,其值设置为 0.15。

模型超参数设置如表 2 所示,Epoch 设置为 250,



Batch size 设置为 64。经多次试验发现,当优化器设置为 Adam,学习率为 0.002,每次参数更新后学习率衰减值设置为 0.000 01 时,训练效果最好。

表 1 编码器结构信息

Layer	Output Size	Kernel
Input	128×128×1	3×3
Conv1	128×128×64	3×3
MaxPooling1	64×64×64	3×3
Conv2	64×64×32	3×3
MaxPooling2	32×32×32	3×3
Conv3	32×32×16	3×3
MaxPooling3	16×16×16	3×3
Conv4	16×16×8	3×3
MaxPooling4	8×8×8	3×3
Flatten	512	
MLP1	512	
MLP2	256	
MLP3	128	
MLP4	64	

表 2 模型超参数信息

Parameter	Value
Batch size	64
Epoch	250
Optimizer	Adam
Learning rate	0.002
Learning rate decay	0.000 01

## 2 实验结果与分析

### 2.1 数据集及处理

#### 2.1.1 实验数据集

本次实验在采集图像过程中,每个样品每旋转 30°采集一张图像。如图 4 所示,总计采集 5 700 幅图片,其中 4 700 幅合格产品图片,1 000 幅次品图片。后期给出了图片的样例,所提供测试的图片包含了常见的缺陷,如:划痕、刮花、表面卷曲、压盖破损等,测试

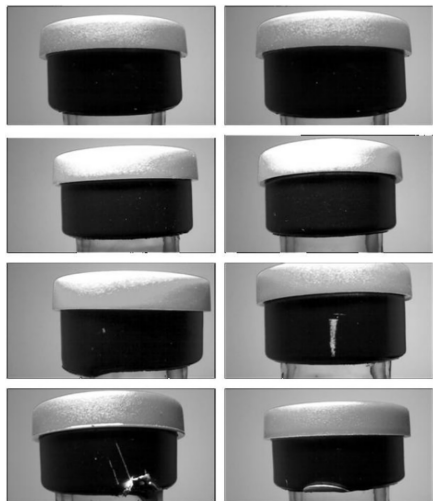


图 4 数据集图像

图像为专业人员挑选,确保涵盖压盖过程中常见缺陷,具有代表性。如图 4 所示,前两行为合格产品,后两行为缺陷产品,从左往右依次为压盖不良、划痕、压盖破损、压盖卷曲缺陷。

#### 2.1.2 数据增强

使用数据增强从现有的数据集中生成更多的训练数据,包括随机水平翻转,在 0°~30°范围内随机旋转,随机添加噪声、随机错切角度等,将数据集扩充到 10 000 幅图像,包括 8 000 幅合格产品图像,2 000 幅缺陷图像。

#### 2.1.3 图像预处理

该文采取常用图像预处理方式对数据集进行预处理,流程如图 5 所示。原始图像使用高速相机拍摄的产品照片,为 1 600×1 200 像素 24 位彩色图像,首先对其进行图像预处理。为了提高算法运算速度,将图片转化为灰度图,对图像进行高斯滤波,以消除图像采集过程中不可避免的噪声,然后经尺寸变换转换为像素值 128×128×1 的灰度图像。经检验,瓶盖图像在 128×128×1 像素时,图像不失真且能最大程度保留缺陷信息。

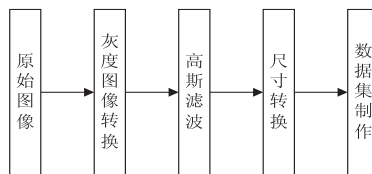


图 5 图像预处理流程

#### 2.1.4 网络训练

在网络训练阶段用于训练的图像全部为合格产品图像,在去噪自编码器中随机添加高斯噪声,并将输入的像素值部分随机设为零。经多次实验验证发现高斯噪声幅度为 0.15 时效果较好。添加对比试验组,使用标注后的数据集对卷积神经网络模型进行训练。

### 2.2 结果分析

#### 2.2.1 损失函数对比试验

如表 1 介绍的去噪自编码器,该文在此分别使用 SSIM 函数及  $L_2$  函数作为卷积去噪自编码器的优化模型损失函数,通过在数据集上进行性能比较,结果如图 6 所示。

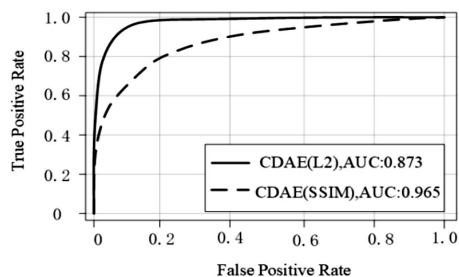


图 6 ROC 曲线

图中展示了 SSIM 与  $L_2$  的 CDAE 模型在压盖数据集上 ROC 曲线及其各自的 AUC 值,使用 SSIM 作为损失函数比使用  $L_2$  作为损失函数的自编码器性能要好。仅通过改变损失函数,在压盖图像数据集上,实现的 AUC 从 0.873 提高到 0.966。

### 2.2.2 实验对比

本次实验采用 VGG16 卷积神经网络与卷积去噪自编码器 CDAE 进行对比实验。卷积神经网络与卷积去噪卷积自编码器在验证集上的准确率对比如图 7 所示。

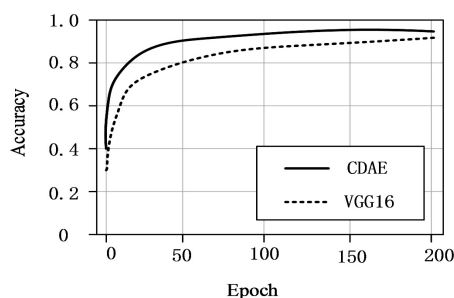


图7 CDAE与VGG16准确率对比

在VGG16卷积神经网络进行训练时,在训练集中的表现为:30次迭代之前,正确率提升明显,在160次迭代时,准确率达到89.6%。在验证集中的表现为:在30次迭代附近,准确率迅速提高,随后趋于稳定。卷积去噪自编码器在训练集上的表现为:在30次迭代前,准确率提升较快,在120次迭代准确率达到95.2%,相较于卷积神经网络准确率提升5.6%,且收敛速度更快。

### 2.2.3 重构图像残差图

经卷积去噪自编码器重构的缺陷图片输出后已不包含缺陷,与输入图片相减,即可得到压盖缺陷残差图。如图8所示,第一列为输入图像,为压盖缺陷图像,第二列为残差图与输入图像叠加图像(标注区域为其对应缺陷区域),第三列为缺陷残差图。

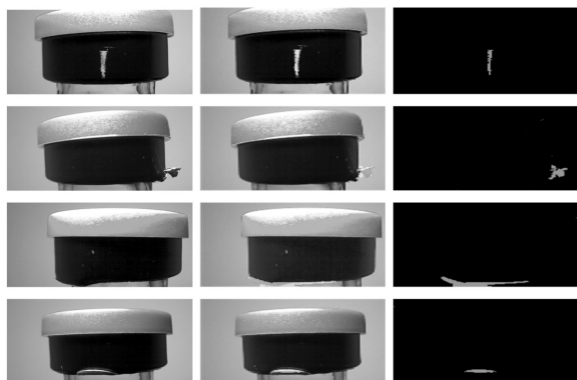


图8 重构图像及残差图

## 3 结束语

针对口服液瓶盖压盖质量检测,该文提出了一种

卷积去噪自编码方法代替卷积神经网络方法对压盖中产生的缺陷进行识别与定位,通过实验对比分析得出以下结论:

基于SSIM的卷积去噪自编码器网络在口服液小瓶压盖质量图像数据集上的准确率达到95.2%,相比于卷积神经网络提升了5.6个百分点,且收敛性更好。同时,卷积去噪自编码器模型可以应用在更多不同的领域,且鲁棒性更强,可以尝试将该方法应用在其他缺陷检测领域。

### 参考文献:

- [1] BO T, JIANYI K, SHIQIAN W. Review of surface defect detection based on machine vision[J]. Journal of Image and Graphics, 2017, 22(12): 1640-1663.
- [2] PANG G, SHEN C, CAO L, et al. Deep learning for anomaly detection: a review[J]. arXiv:2007.02500, 2020.
- [3] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [4] ZHANG A, WANG K C P, FEI Y, et al. Deep learning - based fully automated pavement crack detection on 3D asphalt surfaces with an improved CrackNet[J]. Journal of Computing in Civil Engineering, 2018, 32(5): 04018041.
- [5] DUNG C V, ANH L D. Autonomous concrete crack detection using deep fully convolutional neural network[J]. Automation in Construction, 2019, 99: 52-58.
- [6] CHANDOLA V, BANERJEE A, KUMAR V. Anomaly detection: a survey[J]. ACM Computing Surveys, 2009, 41(3): 1-58.
- [7] BERGMANN P, LÖWE S, FAUSER M, et al. Improving unsupervised defect segmentation by applying structural similarity to autoencoders[J]. arXiv:1807.02011, 2018.
- [8] 罗月童, 卞景帅, 张 蒙, 等. 基于卷积去噪自编码器的芯片表面弱缺陷检测方法[J]. 计算机科学, 2020, 47(2): 118-125.
- [9] 胡 聪. 基于自编码器和生成对抗网络的图像识别方法研究[D]. 无锡: 江南大学, 2019.
- [10] LITJENS G, KOOI T, BEJNORDI B E, et al. A survey on deep learning in medical image analysis[J]. Medical Image Analysis, 2017, 42: 60-88.
- [11] 曲之琳, 胡晓飞. 基于改进激活函数的卷积神经网络研究[J]. 计算机技术与发展, 2017, 27(12): 77-80.
- [12] 高友文, 周本君, 胡晓飞. 基于数据增强的卷积神经网络图像识别研究[J]. 计算机技术与发展, 2018, 28(8): 62-65.
- [13] 刘华春. 卷积神经网络在车牌识别中的应用研究[J]. 计算机技术与发展, 2019, 29(4): 128-132.
- [14] 周飞燕, 金林鹏, 董 军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.
- [15] 李彦冬, 郝宗波, 雷 航. 卷积神经网络研究综述[J]. 计算机应用, 2016, 36(9): 2508-2515.