

基于卷积神经网络的轮胎缺陷 X 光图像分类

崔雪红 刘云 王传旭 李辉

(青岛科技大学信息科学技术学院 青岛 266061)

摘要: 轮胎缺陷的类型直接决定着轮胎是否为残次品或废品,对于轮胎定级具有重要参考价值,探索高性能的轮胎缺陷分类方法至关重要。采用卷积神经网络技术,提出一个端到端的图像自动分类算法。首先,从国内某轮胎生产线上通过现场运行的轮胎 X 光射线缺陷检测系统采集胎侧异物缺陷、胎冠异物缺陷、气泡缺陷、胎冠劈缝、胎侧开根 5 种最常见缺陷类型和 1 种正常胎侧图像作为分类目标,并且依据缺陷图像的缺陷尺度,将每幅图像缩放到 127×127 像素的统一大小;然后,设计含有 5 个卷积层、3 个池化层、3 个全连接层的网络结构。最后,用采集的缺陷样本对所设计的深度网络进行训练学习与测试。并将该算法和大量传统分类算法进行实验对比,取得更好的分类效果,平均测试识别率高达 96.51%。

关键词: 卷积神经网络;池化;轮胎缺陷;图象分类;全连接

中图分类号: TP391;TN081 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Defect classification for tire X-ray images using convolutional neural network

Cui Xuehong Liu Yun Wang Chuanxu Li Hui

(The College of Information Science and Technology, Qingdao University of Science and Technology, Qingdao 266061, China)

Abstract: Type of tire defects directly determines whether the tire is defective products or waste, which has important reference value for tire grading, it is vital to explore high performance tire defect classification method. First, collecting five common types of defects and a normal images from a typical tire manufacturing, namely belt-foreign-matter, sidewall-foreign-matter, belt-joint-open, cords-distance, bulk-sidewall and normal-cords, was used to perform the tire defect classification experiments. And down-sampling or up-sampling the images in the dataset to a fixed resolution of 127×127 . And then designing depth network which contains 5 convolutional layers, 3 max-pooling layers, 3 fully-connected layers. Finally, training and testing the designed depth network with defect samples collected. Experimental results showed that the method proposed has higher recognition rates for tire defects than other algorithms, the averaged rate of recognition is high to 96.51%.

Keywords: CNN; pooling; tire defect; image classification; fully connected

1 引言

在轮胎生产中,轮胎缺陷对轮胎质量有很大的影响,轮胎的质量直接决定了轮胎的使用寿命,甚至关系到人车的安全。轮胎在连续、高速的自动化工业生产过程中会出现各种各样的质量问题,诸如胎体钢丝帘线缺陷、带束层缺陷和异物缺陷等问题,这些质量问题一旦出现,将导致大量次品及废品,会大大增加原材料的消耗和成本。若能及时检测出轮胎缺陷并识别出类型,可以及时发现问题、调整工艺、提高产品质量。此外,根据轮胎生产企业生产线对轮胎缺陷检测的要求,缺陷检测需要能够根据缺陷的类型、缺陷的数量及缺陷程度等因素,自动对被检测轮胎进行定级。

而且射线无损检测技术已得到广泛的研究和使用^[1]。因此,对检测到的轮胎缺陷图像进行模式识别(pattern recognition)是十分必要的而且是可行的研究内容。

目前的轮胎缺陷检测及分类方法一般是用 X 光机^[1]对轮胎进行 X 光成像,工人通过观察图像判断轮胎是否有缺陷并对缺陷归类。这种判读是主观判读,很容易受到外界多种因素的干扰,并且工人工作量和强度均很大,容易产生用眼疲劳而错判误判,不利于轮胎缺陷的检测及分类,因而导致不合格产品出厂或合格产品误判为废品。近年来,我国也提出了许多轮胎缺陷检测算法^[2-6],但大部分不带自动分类功能,在后续的工作中还需要人

工再次确认和归类检测出来的缺陷。人没有完全解脱

出来,使得轮胎判级任务带有很大的主观性。目前许多学者对轮胎缺陷分类也有一些研究,例如:张小丽、冯霞、朱越、邵明红等对原始图像进行预处理、阈值变换^[7]、细化等操作后提取各种缺陷的特征参数,并据此识别胎侧帘线缺陷,这些方法只能对胎侧帘线缺陷进行识别分类。青岛科技大学张岩从轮胎缺陷的形状、灰度以及纹理变化3个方面表征缺陷共提取22个缺陷特征量,利用径向基核函数RBF的支持向量机对轮胎缺陷分类,分类正确率达到94.2%,此算法可以对各种区域的缺陷分类,也能达到高的分类率,但是提取特征繁琐,需要人的参与选择。由于轮胎缺陷的类内特征变化较大,而类间特征却又存在较大的相似性,这种情况对特征选择和分类器设计提出了挑战,往往需要启发式方法和非常专业的知识^[8],很大程度上依靠个人经验,而且分类模型的设计取决于人工提取的特征^[9],适应性较差。

近年来,无需人工提取特征、也无须任何的预处理输入图像即可直接作用于卷积神经网络方法^[10],在图像分类领域中的应用日渐受到关注。该文针对目前轮胎缺陷识别方法需要人的参与来提取缺陷特征的缺点,提出了用卷积神经网络对轮胎缺陷进行分类的新方法,并和大量算法进行实验对比,此类算法分类精度较高且复杂度较小,具有很好的鲁棒性。

2 卷积神经网络

卷积神经网络是一种特殊的深层神经网络,它的特殊性体现在两个方面:局部连接和权重共享,这两个特性使网络更类似于生物神经网络,降低了网络模型的复杂度,减少了权值的数量。其特征映射层与特征提取层这两种特有的特征提取结构使网络在识别分类时对输入图像有较高的畸变容忍能力。它较一般深层神经网络在图像处理方面有如下优点:1)原图像可以作为输入数据直接输入无须进行复杂的图像预处理操作;2)权值共享使网络的训练参数极大的减少,网络结构变得更简单,对视觉可以具有更好的泛化能力;3)特征提取和模式分类同时进行,避免了数据重建的复杂度,提高了效率。

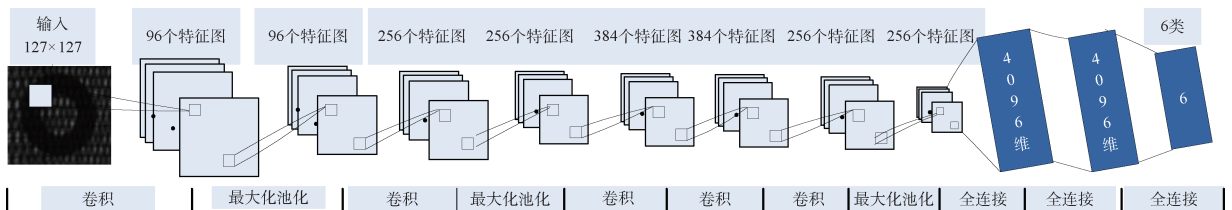


图3 卷积神经网络结构

2.2 卷积层

CNN更适用于图像处理操作,卷积和图像的局部特征相对应。卷积层的前一层(一般是输入层或降采样层)

在传统的图像分类^[11-13]上,如图1所示,首先要输入图像,然后人工参与复杂的预处理来进行特征提取^[14],接着把提取的特征 f_1, f_2, \dots, f_n 输入分类器,最后在输出端给出分类结果。手工设计特征需要大量的经验,需要对这个领域和数据特别了解,此外还需要选择比较合适的分类器算法,只有两者结合达到最优,才能获得较好的分类效果,这几乎是不可能的。传统图像分类把特征提取和分类器设计分开来做,然后在应用时再结合在一起,此分类方法耗时并且定义的视觉特征是利用一个区域内像素级别的数据信息综合表现出来,这与人类视觉过程并不十分吻合。而卷积神经网络的分类模型不需要手动设计特征,也不需要挑选分类器,它能同时自学习特征和分类器。即输入某一神经网络时,输入的只是二维图片,输出的是其自己的标签,如图2所示。与传统分类模型相比,更符合人类视觉过程。这一点已被神经生物学家David Hubel证明大脑的可视皮层是分级的,人的视觉功能是一个抽象迭代的过程。

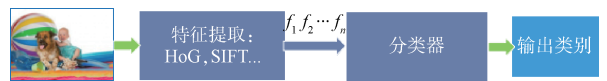


图1 传统分类模型



图2 卷积神经网络分类模型

2.1 网络结构

卷积神经网络是一种前馈式神经网络,前向传播是特征提取的过程,反向传播是权值修正的过程。如图3卷积神经网络结构,包括1个输入层、5个卷积层、3个池化层和3个全连接层共12层。图像从输入层Input输入,卷积层(Conv)提取特征、最大池化层进行降采样(Max-pooling)、最后的全连接层(Fc)输出分类结果。卷积层和抽样层中有若干个特征图(即特征平面),每个特征图都与其前一层特征图相连接。

的特征图与若干个可学习的核进行卷积,然后卷积的结果经过激活函数(一般为Sigmoid函数或Tanh函数)输出形成这一层的特征图。图3中,卷积层为Conv1~Conv5,

与降采样层大部分间隔出现，卷积层每一个生成的特征图是前一层的一个或几个特征图的输入，然后与卷积核卷积组合的结果。即第 l 层卷积层的第 j 个特征图为：

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l\right) \quad (1)$$

式中： M_j 为选择输入的前一层特征图的集合， $f(\cdot)$ 表示激励函数； k_{ij}^l 是可学习的卷积核，每个输入特征图可以有不同卷积核，可以输出不同的特征图 x_j^l ； b_j^l 是卷积特征图 x_j^l 的加性偏置。通过训练，卷积核可以提取出某些有意义的特征，例如第一个卷积层的卷积核类似于 Gabor 或 ICA 滤波器，可以提取边缘、角等信息，(例如图 4(b) 是学习得到的 96 个 11×11 大小的卷积核，测试图像(图 4(a))可以使用这组已学习的卷积核提取 96 个相应的特征，如图 4(c)，从图 4(c) 可以看到一些明显的边缘轮廓)。但优于 Gabor 滤波器，由于本文卷积核是依据输入的数据集自学习而获得的，能够适应图像的变化，提取的特征更能表征轮胎的缺陷信息，而 Gabor 滤波器是事先人工设计的，不能适应缺陷信息的变化而变化。多层的网络结构可以对输入的图像进行逐层的抽象表示，获得更高层次的分布式特征表达。

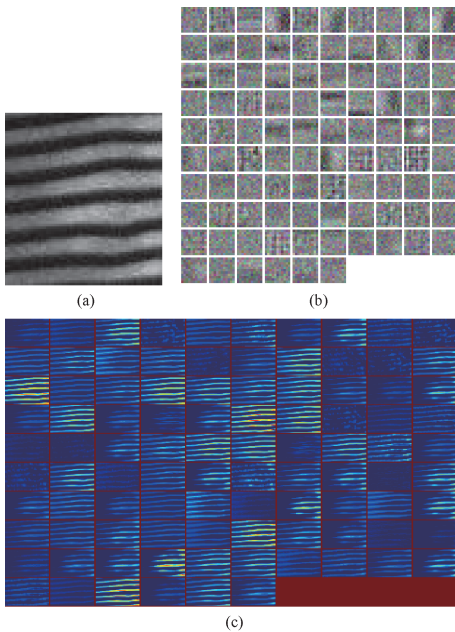


图 4 (a) 测试图像；(b) 96 个卷积核；
(c) 相应的 96 个特征

2.3 池化层

池化(pooling)是计算机视觉与机器学习领域中的常见操作。就是将不同位置的特征进行聚合。CaffNet 模型框架采用最大池化的形式，该网络模型中的 conv1、conv2 以及 conv5 层之后均连接了一个最大池化层，其池化大小均为 3×3 像素，池化步长均为 2 像素。通过池化，不仅可以降低特征的维数，还可以保持平移不变性，提高特征的

鲁棒性。最大化池化(Max-pooling)公式如下：

$$x_{rc}^l = \max_{i,j \in \{0,1,\dots,m\}} x_{(r+i-1)(c+j+1)}^{l-1} \quad (2)$$

其中， x_{rc}^l 是在特征图位置 $\{r,c\}$ 点上的输出， m 为池化区域大小， $x_{(r+i-1)(c+j+1)}^{l-1}$ 是在 l 层特征图位置 $(r+i-1)(c+j+1)$ 上的输入。如池化步长 $stride=1$ 等同于使用所有可能的池化窗口，步长 $stride>1$ 等同于使用所有非重叠池化窗口。图 5(a) 是池化步长为 1、池化窗口大小为 10×10 的示意图；图 5(b) 是池化步长为 10、池化窗口大小为 10×10 的示意图 Scherer 等人^[15]探讨了不同的池化结构，并指出最大化池化已获得了最好的结果。由于特征图上的独立激活单元位于池化窗口内，因此池化层提高了模型的鲁棒性，卷积池化这两次特征提取使网络在识别时对输入样本有较高的畸变容忍能力。根据经验，池化操作有助于提高目标识别率^[16]。

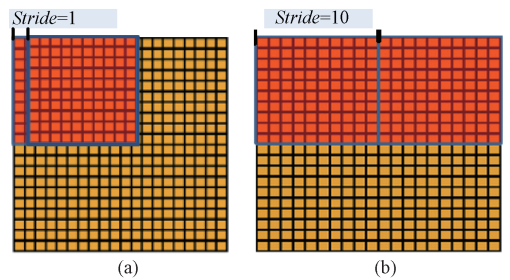


图 5 池化示意图

2.4 全连接层

相比于卷积层的局部连接方式，全连接层的全连接方式将会带来更多的网络参数。全连接公式如下：

$$x_{rc}^l = \sigma\left\{\sum_{i=1}^m \sum_{j=1}^m W_{ij}^{l;rc} x_{(r+i-1)(c+j+1)}^{l-1}\right\} \quad (3)$$

式中： $w_{ij}^{l;rc}$ 是与 l 层相连的权重， x_{rc}^l 是在 l 层位置 i,j 上的输出单元 r,c ， $x_{(r+i-1)(c+j+1)}^{l-1}$ 作为 l 层位置 $(r+i-1)(c+j+1)$ 上的输入。

CaffeNet 模型的最后 3 层为全连接层(fc6~fc8)。由于之前的卷积层及池化层已经将特征的维数降低至可接受的大小，因而使用全连接层不会导致特别严重的计算负担。在 CaffeNet 的 3 个全连接层中，fc6，(图 6) 和 fc7 层

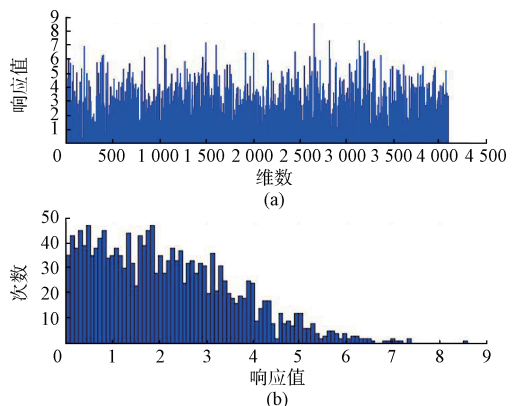


图 6 fc6 全连接层统计

均为包含 4 096 个神经元的隐藏层,而 fc8 层为 6 路的 softmax 输出层(图 7)。图 6 是对 fc6 全连接层的统计,图 6(a)的横坐标是神经元的个数即维数(4 096 维),纵坐标是每个神经元的响应值。图 6(b)是直方图统计,横轴是神经元的响应值,纵轴是每个响应值出现的次数。图 7 统计输入图像属于哪一类的概率,从图 7 可以看出输入图像(图 4(a))属于第(2+1=3)类,即气泡缺陷类。

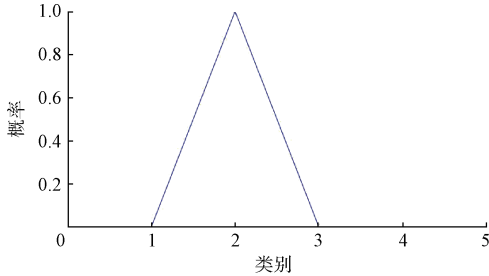


图 7 fc8 全连接层分类结果统计

3 实 验

为了验证本文方法的性能,将 PCA+BP 神经网络、具有空间位置关系的稀疏编码分类算法(ScSPM、LLC)、空间金字塔匹配法(KSPM)以及 LeNet 网络等最广泛使用的先进算法和本文算法进行对比分析。

3.1 网络参数优化

CaffeNet 包含 5 个卷积层(conv1~conv5),其卷积核大小(KernelSize)分别为 11×11 、 5×5 、 3×3 、 3×3 、 3×3 像素,输出的特征图个数分别为 96、256、384、384、256 个,卷积的步长(Stride)分别为 4、1、1、1、1 像素。CaffNet 模型采用最大池化的方式,该模型中的 conv1、conv2 以及 conv5 层之后均连接了一个最大池化层,其池化尺寸均为 3×3 像素,池化步长均为 2 像素。本文网络使各隐层的空间分辨率逐层递减,而使所含特征平面的数量却逐层递增,以能检测出更多类型的缺陷特征信息。在 CaffeNet 的 3 个全连接层中,fc6 和 fc7 层均为包含 4 096 个神经元的隐藏层,而 fc8 层为 6 路的 softmax 输出层。每层卷积层使用 ReLU 激活函数,它较 sigmoid 和 tanh 具有更快的收敛速度。其网络各层参数如表 1。

该文以 127×127 图像作为网络的输入,最后输出 6 类结果;测试网络验证数据时,由于数据量小,以防误差下降太快,初始的学习率(base-lr)设为 0.001;由于共有 246 张测试样本,测试批量(batch)设为 246,测试批次(test-iter)设为 1,;测试间隔(test-interval)设为 200,即每迭代 200 次测试一次,显示一次分类准确率结果;使用 CPU 求解模式(slover_mode)求解;其余参数使用 caffenet 优化模型的默认参数。

表 1 网络各层参数

Layer	Type	Maps & neurons	Kernel	Stride
0	Input	3 maps of 127×127 neurons		
1	conv1	96 maps of 30×30 neurons	11×11	4
2	Pooling1	96 maps of 15×15 neurons	3×3	2
3	conv2	256 maps of 15×15 neurons	5×5	1
4	Pooling2	256 maps of 7×7 neurons	3×3	2
5	conv3	384 maps of 7×7 neurons	3×3	1
6	conv4	384 maps of 7×7 neurons	3×3	1
7	conv5	256 maps of 7×7 neurons	3×3	1
8	Pooling3	256 maps of 3×3 neurons	3×3	2
9	fc6	4096 neurons	1×1	1
10	fc7	4096 neurons	1×1	1
11	fc8	6 neurons	1×1	1

3.2 数据集

本文从国内某轮胎生产线上通过现场运行的轮胎 X 光射线缺陷检测系统采集的胎侧异物缺陷(339)、胎冠异物缺陷(266)、气泡缺陷(226)、胎冠劈缝(261)、胎侧开根(273)5 种最常见缺陷类型和 1 种正常胎侧图像(217),括号里的数字是每类图像的数量,共计 1 582 幅图像,每类缺陷样本所占比例与该类缺陷出现频率一致。这些缺陷图像是包含缺陷的最小矩形图像,由于生产中产生缺陷的不确定性,缺陷图像的长宽和大小各异,分布在 $50 \times 50 \sim 200 \times 500$ 像素。为了使缺陷图像的大小统一以满足算法的要求,并能使其最大程度上表征图像缺陷,同时兼顾降低计算复杂度,该文首先将每幅图像缩放到 127×127 像素大小,然后用这些图像组成轮胎缺陷库。其中随机选取缺陷库中每类缺陷的 20% 作为验证集,随机选取每类缺陷的 20% 作为测试集,剩余的图像作为训练集,重复 10 次这样的样本选取,取 10 次分类率的平均值作为最终的识别率。各算法在该文采集的数据集上分类准确率如表 2 所示,其中 KSPM-200-3 与 KSPM-400-2 中的 200、400 分别是指字典的大小,3 与 2 是指金字塔层数;LeNet 与本文算法中,括号里的识别率是验证识别率。

表 2 本文数据集上性能比较

算法	分类准确率/%
PCA+BP 神经网络	69.44
ScSPM09 ^[17]	95.56
LLC10 ^[18]	94.85
KSPM-200-3 ^[19]	92.77
KSPM-400-2	92.37
LeNet ^[13]	91.89(93.46)
本文算法	96.51(98.37)

由表2的数据可以看出,本文验证识别率较其他算法分类准确率有了较大的提高,分别提升了28.93%、2.81%、3.52%、5.6%、6%、4.91%。测试识别率也比其他算法识别率高,分别提高了26.08%、0.22%、0.67%、2.75%、3.15%、3.89%。导致测试识别率较验证识别率低的因素可能有二,1)训练样本不足;2)没有给予最优的参数设置。由于轮胎的结构及其制造过程过于复杂,轮胎缺陷具有各种大小各种形态各种角度各种不同的破损位置等特点,又因CNN具有大量的网络参数,所以卷积神经网络要求大量的训练样本,才能训练出泛化能力强的模型。ScSPM09和LLC10是考虑到位置信息的稀疏编码算法,也是目前被广泛研究、成功的算法,均取得了较好的识别率。ScSPM09、LLC10、KSPM算法的缺点是需要人工参与图像特征的提取和分类器的选择,选取的好坏直接影

响分类的效果。在本文中这3种算法均提取的是sift特征,前两种算法使用的是线性SVM分类器,后一种算法使用的是非线性SVM分类器。

除了比较6种算法的总体识别率外,本文还对每种缺陷类型的识别情况进行了分析,表3所示为本文算法对测试样本的详细识别结果。

由表3可以看出,1)气泡的识别率比较低,其主要被误识为正常胎侧。通过观察被误识的气泡样本图像发现,弱边缘的气泡缺陷图像与正常胎侧的图像特征非常相似,所以导致错分。2)除气泡外,其余所有缺陷类型的识别率都非常高。这是由于此算法是一个端到端的图像自动分类算法,能够自动学习出表征能力强的深层缺陷图像特征,这与人类视觉过程十分吻合。

表3 测试样本的识别结果

	胎侧异物	胎冠异物	胎冠劈缝	帘线异常	气泡	正常胎侧	正确个数	样本总数
胎侧异物	68	0	0	0	0	0	68	68
胎冠异物	0	53	0	0	0	0	53	53
胎冠劈缝	0	2	50	0	0	0	50	52
帘线异常	0	0	0	0	1	1	53	55
气泡	0	0	0	0	0	5	40	45
正常胎侧	0	0	0	0	2	41	41	43
总样本							305	316

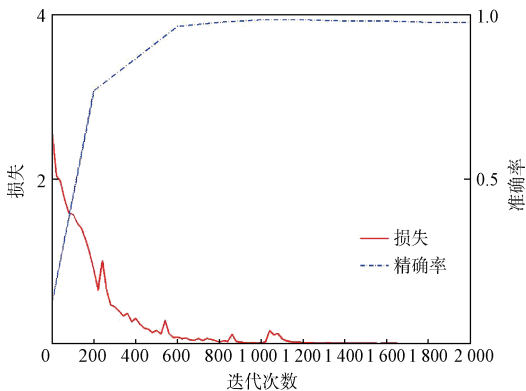


图9 Loss和Accuracy的曲线

此外,该文还分析了损失(Loss)和准确率(Accuracy),图9所示为该算法Loss和Accuracy的曲线图,其横坐标表迭代次数,左纵坐标代表损失函数值(loss function, LF),右纵坐标表示验证识别率。损失函数是用于衡量最优的策略,LF越小越好,表示越接近于真实,即越接近正确的结果误差越小,识别准确率越高。由图9可以看出Loss曲线几乎接近于0,而Accuracy曲线接近于1,正符合优化目标的要求。

4 结论

为了避免现行的缺陷分类算法需人工提取缺陷特征、人为选取分类器以及需要非常专业的知识的问题,本文探索了一种轮胎缺陷分类的新思路,提出用卷积神经网络对轮胎缺陷进行端到端的图像自动分类的方法,此方法无需手动提取特征,是一个自学习特征的过程;同时局部野和权值共享技术减少了参数空间,大幅度降低了算法的复杂度;卷积和图像的局部特征相对应,能更好的表征缺陷特征;次抽样技术增强了网络鲁棒性,能容忍图像一定程度的畸变。通过实验证明,采用卷积神经网络来分类轮胎缺陷,识别率高达98.37%,其识别率明显优于其它现行的分类算法,具有很好的应用前景。从实验中可以看出方法识别率

参考文献

- [1] 邬冠华,熊鸿建.中国射线检测技术现状及研究进展[J].仪器仪表学报,2016,37(8):1683-1695.
- [2] 张传海.纹理无关的裂纹缺陷检测算法[D].济南:山东大学,2013.
- [3] 章玲.基于图像放缩算法的轮胎缺陷检测系统研究与实现[D].济南:山东大学,2015.

- [4] LIU Q, WANG G, GUO Q. Tire defect detection based on radon transform [J]. *Journal of Computational Information Systems*, 2015, 11(21): 7841-7848.
- [5] 向媛媛. 基于图像字典表示的缺陷检测算法[D]. 济南:山东财经大学, 2015.
- [6] XIANG Y, ZHANG C, GUO Q. A dictionary-based method for tire defect detection [C]. *IEEE International Conference on Information and Automation*, IEEE, 2014.
- [7] 滕秋霞,沈天飞,杨金霄. 基于多肤色模型的人脸检测系统研究[J]. *电子测量技术*, 2015, 38(9):47-51.
- [8] ANAND R S A, KUMAR P. Flaw detection in radiographic weldment images using morphological watershed segmentation technique [J]. *Ndt & E International*, 2009, 42 (1):2-8.
- [9] 白丰,张明路,张小俊,等. 局部二进制特征描述算法综述 [J]. *电子测量与仪器学报*, 2016, 30 (2): 165-178.
- [10] 余永维,杜柳青,曾翠兰,等. 基于深度学习特征匹配的铸件微小缺陷自动定位方法 [J]. *仪器仪表学报*, 2016, 37(6): 1364-1370.
- [11] 王剑. 自由立体显示系统中的人眼跟踪定位方法 [J]. *电子测量技术*, 2015, 38(9):52-54.
- [12] 黄志鸿,毛建旭,王耀南,等. 基于机器视觉的啤酒瓶口缺陷检测分类方法研究 [J]. *电子测量与仪器学报*, 2016, 30(6):873-879.
- [13] SHUANG L, JIAYUAN M, XIN Z, et al. Cross-task emotion recognition using eeg measures; first step towards practical application [J]. *Instrumentation*, 2014, 1(3):17-24.
- [14] 鲁萌萌,赵凤军,李宁. 基于词包模型的高分辨率 SAR 图像特征提取 [J]. *国外电子测量技术*, 2015, 34 (6):62-69.
- [15] SCHERER D, LLER A, BEHNKE S. Evaluation of pooling operations in convolutional architectures for object recognition [C]. *International Conference on Artificial Neural Networks*, Springer-Verlag, 2010: 92-101.
- [16] BENGIO Y, COURVILLE A. Deep Learning of Representations [C]. *Handbook on Neural Information Processing*. Berlin Heidelberg: Springer, 2013:1-28.
- [17] YANG J, YU K, GONG Y, et al. Linear spatial pyramid matching using sparse coding for image classification [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2009:1794-2223.
- [18] WANG J, YANG J, YU K, et al. Locality-constrained Linear Coding for image classification [C]. *IEEE Computer Society Conference on Computer Vision & Pattern Recognition*, 2010:3360-3367.
- [19] LAZEBNIK S, SCHMID C, PONCE J. Beyond bags of features; spatial pyramid matching for recognizing natural scene categories [C]. *IEEE Computer Society Conference on Computer Vision & Pattern Recognition*, 2006:2169-2178.

作者简介

崔雪红,工学硕士,讲师,主要研究方向为计算机视觉,图像处理。

E-mail: cuixuehongzhe@163.com