

# 基于改进 Faster RCNN 的马克杯缺陷检测方法

李东洁\*, 李若昊

哈尔滨理工大学自动化学院, 黑龙江 哈尔滨 150080

**摘要** Faster RCNN 在小目标的检测方面准确度和鲁棒性性能不够理想。为此, 提出一种改进 Faster RCNN 对马克杯缺陷进行检测, 将 Faster RCNN 和特征金字塔网络(FPN)相结合, 增加使用细节化的浅层特征, 使其对小目标有更好的检测效果。分别使用改进前后的 Faster RCNN 在 Caffe 上进行仿真。结果表明, Faster RCNN 检测马克杯缺陷的性能出色, 但是遗漏了部分小目标; 改进后的 Faster RCNN 对缺口和划痕缺陷的检测准确率最多提升 2.485 个百分点, 并且在小目标识别方面也有更优异的表现。

**关键词** 机器视觉; 深度学习; Faster RCNN; 特征金字塔网络; 缺陷检测

中图分类号 TP391

文献标志码 A

doi: 10.3788/LOP57.041515

## Mug Defect Detection Method Based on Improved Faster RCNN

Li Dongjie\*, Li Ruohao

School of Automation, Harbin University of Science and Technology, Harbin, Heilongjiang 150080, China

**Abstract** Faster RCNN has poorer performance in terms of accuracy and robustness for detecting small targets. For this reason, an improved Faster RCNN was proposed to detect the defects in mugs. The Faster RCNN and feature pyramid network (FPN) were combined to increase the use of detailed shallow features, so as to achieve better detection effect for small targets. Faster RCNNs before and after improvement were used to conduct simulation on Caffe. The results show that Faster RCNN performs well in defect detection for mugs, but it misses some small targets. The improved Faster RCNN increases the detection accuracy by 2.485 percent at most for gaps and scratches and performs better in small target recognition.

**Key words** machine learning; deep learning; Faster RCNN; feature pyramid network; defect detection

**OCIS codes** 150.1135; 100.2000; 100.3008

## 1 引言

在生产制造马克杯的过程中, 生产材料、工艺和设备等客观因素会导致马克杯的杯口出现划痕、缺口和斑点等缺陷, 这些缺陷直接影响了消费者的使用体验和生产企业的企业形象。为了避免不合格的马克杯流入市场, 早期许多马克杯的生产厂家采用人工抽检来检测缺陷, 但此类检测方法无法保证检测强度和精度。目前的工业检测主要采用机器学习方法<sup>[1-2]</sup>, 如神经网络<sup>[3]</sup>和支持向量机等<sup>[4-5]</sup>, 但这些方法都需要人工设计缺陷特征, 其带来的突出问题就是人工设计的缺陷无法满足缺陷的多样性, 无法

应对缺陷种类大小、数量繁多的情况, 鲁棒性不足<sup>[6-7]</sup>。

近年来, 随着图形处理器(GPU)性能的不断提升, 深度学习领域蓬勃发展。2006 年, Hinton 等<sup>[8]</sup>提出了深度学习这一概念, 运用卷积神经网络(CNN)从数据中学习特征。2012 年, 在 ImageNet 图像分类竞赛中, Krizhevsky 团队<sup>[9]</sup>首次提出了深度 CNN AlexNet, 以超过第二名 15.3% 的准确率获得冠军, 使人们对 CNN 在视觉任务中的应用有了进一步认识。Girshick 等<sup>[10]</sup>提出了区域卷积神经网络(RCNN)模型, 该模型利用选择性搜索方法选取候选区域, 使用多个支持向量机对特征进行分类,

收稿日期: 2019-05-30; 修回日期: 2019-07-25; 录用日期: 2019-08-15

基金项目: 国家自然科学基金青年基金(51105117)、黑龙江省普通高校基本科研业务费专项基金(LGYC2018JQ017)

\* E-mail: dongjieli2013@163.com

从而完成目标检测。2015 年, Girshick<sup>[11]</sup>提出了 Fast RCNN, 这是 RCNN 的改进版本, 采用 RoI Pooling 来共享计算量较大的部分, 提升整个模型的工作效率。后来针对提取候选特征区域运行时间较长的问题, Ren 等<sup>[12]</sup>在原有网络基础上再次加以改进, 引入区域推荐网络 (RPN), 设计了 Faster RCNN 模型并取得了很好的效果。相比于人工设计的特征, CNN 提取的特征鲁棒性更好且语义信息更强, 在人脸识别<sup>[13-14]</sup>、目标检测和语音识别<sup>[15]</sup>等计算机视觉领域中均取得了巨大成果<sup>[16]</sup>。Zhao 等<sup>[17]</sup>提出了一种基于 CNN 的绝缘子航拍图像表面缺陷检测方法, 该方法可实现对绝缘子缺陷的种类识别及定位。姚明海等<sup>[18]</sup>提出了一种基于深度主动学习的磁片表面缺陷检测方法以解决磁片表面对比度低, 磨痕纹理干扰和缺陷块小且亮度变化大等难点。洪伟等<sup>[19]</sup>提出了一种基于区域全卷积网络结合残差网络的深度学习方法对火焰进行检测识别, 提升了在火焰视频数据集上的平均识别精度。韩燮等<sup>[20]</sup>提出了一种基于 CNN 的棋子识别方法, 该方法可应用于不同字体的棋子识别, 并在定位误差、平均定位时间和平均棋子识别准确率等方面都有优秀表现。欧攀等<sup>[21]</sup>融合了空间变换网络的 CNN 所提取的特征, 提出的方法具有更好的旋转不变特性, 从而使遥感图像检测达到更高的检测精度。

马克杯的缺陷在马克杯图像中占比较小, 若单独使用 Faster RCNN 可能无法取得理想的效果。因此本文分别对 Faster RCNN 和特征金字塔网络 (FPN) 进行研究, 从中找到合适的方法将二者进行结合以得到更理想的检测结果。

## 2 基本原理

### 2.1 Faster RCNN 和 FPN 融合原理

对于 CNN 而言, 浅层特征提供更多细节信息且语义信息比较稀疏, 是细节化的特征; 深层特征提供更多整体信息且语义信息比较丰富, 是全局化的特征。

原始 Faster RCNN 中 RPN 仅采用网络的最后一层特征, 这是因为最高层的特征具有最丰富的语义信息。该思想对于大目标的检测很有效, 但是对于小目标的检测效果不佳。采集马克杯图像, 发现部分缺陷在图片中占比较小, 为了使最后的缺陷检测结果更为准确, 将选择一种金字塔 RPN 结构来代替原有 RPN 结构。

金字塔 RPN 结构相比于原始 Faster RCNN 中的 RPN 结构可更好地利用浅层特征提供的细节信息, 更多的细节特征使其对小目标的识别更具效果。最顶层的特征通过上采样与相邻低层进行特征融合, 并且每一层都进行独立预测。整个算法大致分为三个部分: 自底向上、自顶向下和横向连接。具体步骤如下。

**自底向上:** 普通的 CNN 前馈过程, 如果同一阶段的不同层产生相同大小的输出映射时, 则选择每个阶段最后一层的输出作为待处理特征图。

**自顶向下:** 将自底向上过程得到的特征图进行上采样, 其目的是使上采样后的特征图和相邻低层的特征图大小相同, 从而可以进行特征图的相加。

**横向连接:** 将上采样的结果和自底向上生成的特征图进行融合。

图 1 为 FPN+RPN 算法网络结构图。

经 FPN 融合更多浅层特征的特征图, 会在 RPN 作用下生成一系列候选框, 也就是目标的定位, 最后会生成两个得分, 用于过滤效果不佳的候选框和修正候选框, 以便更好地拟合预测的目标。经过 RPN 处理后得到的目标候选框会映射到原始特征图上, 固定尺寸后会连接到 Softmax 用于最终目标分类。

Faster RCNN 使用目标推荐框和特征图作为输入。但加入 FPN 算法后, 会得到多个特征图, 此时感兴趣区域 (ROI) 会根据特征图尺度的不同来选择不同层级的金字塔层, 具体选择方式为

$$k = k_0 + \log_2 \left( \frac{\sqrt{wh}}{224} \right), \quad (1)$$

式中:  $k_0 = 5$ ;  $w, h$  分别为 ROI 的宽和高; 224 为预训练数据集图像大小。

### 2.2 构造数据集

采集好马克杯图像后, 通过翻转、裁剪和旋转等方式扩充数据集, 使用 labelImg 软件对数据集进行标注, 其中训练集为 1532 张, 测试集为 1470 张。标注时需要注意的一点是只需要标注缺陷本身, 如果标注区域过大, 包括了一些非缺陷区域, 则会影响最终训练效果。训练样本的多寡同样会影响训练效果: 如果训练样本过少, 会导致过拟合; 训练样本越多, 网络的泛化能力越强, 最后得到的模型精度越高。部分训练样本如图 2 所示。其中图 2(a) 中包含四个缺陷, 即缺口、划痕、两个斑点; 图 2(b) 中包含一种缺陷, 即斑点; 图 2(c) 中包含两个缺口缺陷。

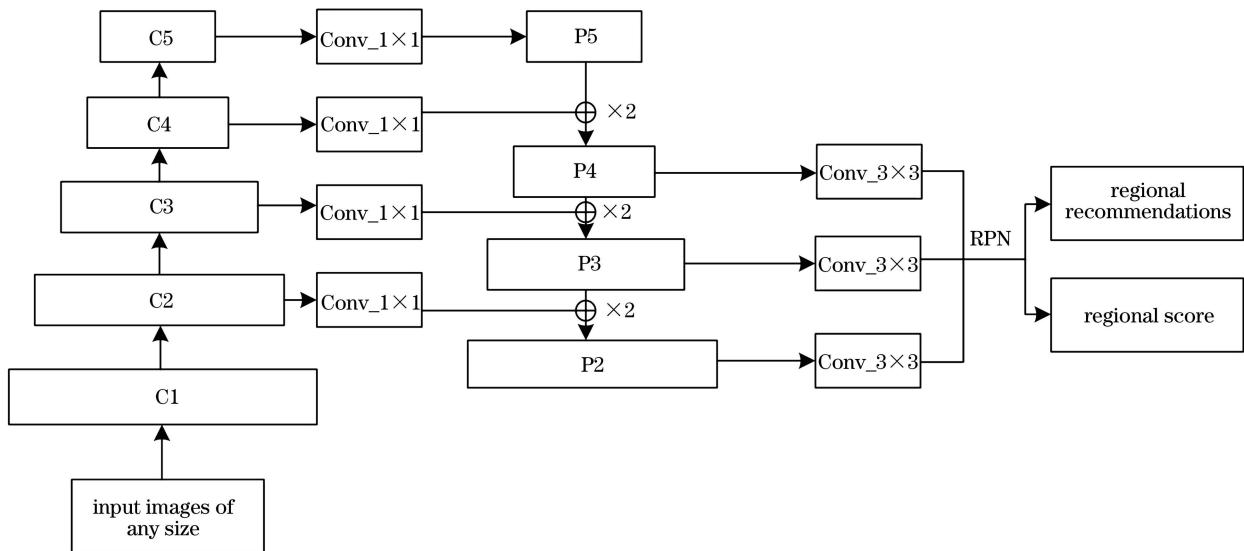


图 1 FPN+RPN 算法网络结构图

Fig. 1 FPN+RPN algorithm network structure diagram

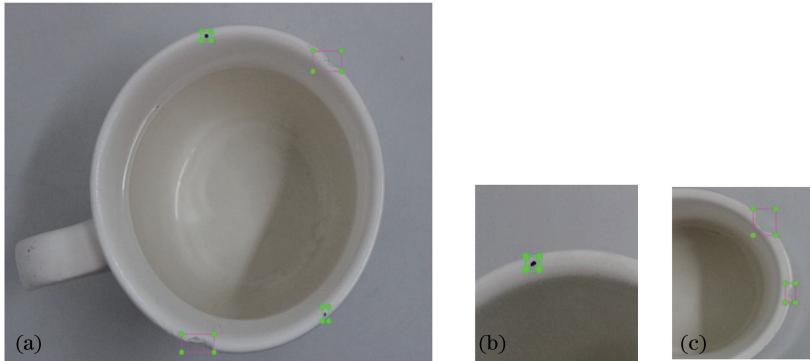


图 2 部分训练样本。(a)包含四个缺陷,即缺口、划痕、两个斑点;(b)包含一种缺陷,即斑点;(c)包含两个缺口缺陷

Fig. 2 Partial training samples. (a) With four defects, one gap, one scratch, and two speckles; (b) with one speckle defect; (c) with two gap defects

为保持训练样本和测试样本的独立性,选择与训练样本不同的测试样本进行测试,这样做可以使

实验结果更加有说服力。部分测试样本如图 3 所示。

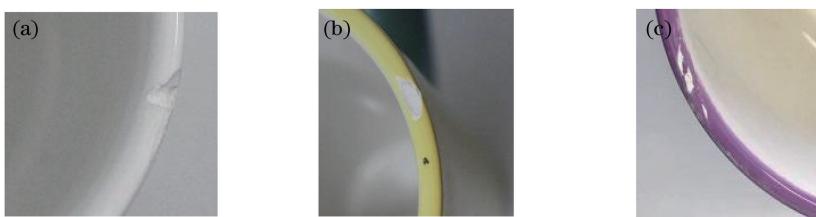


图 3 部分测试样本。(a)包含一个缺口缺陷;(b)包含两个缺陷,即缺口和斑点;(c)包含一个划痕缺陷

Fig. 3 Partial test samples. (a) With one gap defect; (b) with two defects, one gap and one speckle; (c) with one scratch defect

### 2.3 CNN 的选取

ZF(Zeiler & Fergus)网络与经典的 AlexNet 相比,其前面的层使用更小的卷积核和步长,使提取的特征更加精确;与 VGG-16 网络相比,ZF 网络对

GPU 性能要求更低,更加符合实验要求。但 ZF 网络也有缺点,其准确度较低,若直接使用 ZF 网络很难使马克杯检测分类达到预定要求。因此,适当改进 ZF 网络,在保留其优点的同时增加准确度。表 1

表 1 改进前后的 ZF 网络结构  
Table 1 ZF network structure before and after improvement

Type of layers		Number of convolution kernels		Step size	
Before	After	Before	After	Before	After
Conv_1/1	Conv_1/2	96	64	$7 \times 7 / 2$	$3 \times 3 / 1$
Max pooling/1	Max pooling/1	96	64	$3 \times 3 / 2$	$2 \times 2 / 2$
Conv_2/1	Conv_2/2	256	128	$5 \times 5 / 2$	$3 \times 3 / 1$
Max pooling/1	Max pooling/1	256	128	$3 \times 3 / 2$	$2 \times 2 / 2$
Conv_3/3	Conv_3/2	384/384/256	256	$3 \times 3 / 1$	$3 \times 3 / 1$
Max pooling/1	Max pooling/1	256	256	$3 \times 3 / 2$	$2 \times 2 / 2$
	Conv_4/3		384		$3 \times 3 / 1$
	Max pooling/1		384		$2 \times 2 / 2$
	Conv_5/3		512		$3 \times 3 / 1$

为改进前后的 ZF 网络结构。

原始 ZF 网络包含 5 个卷积层和 3 个最大池化层, 改进后的 ZF 网络在原有 ZF 网络的基础上增加了 7 个卷积层和 1 个最大池化层, 并减小了卷积层的大小和数量, 将步长缩短为 1。

为了更加直观地对比改进前后的 ZF 网络结构的性能, 分别采用两种网络结构作为特征提取层, 在 PASCAL VOC2007 数据集上, 使用扩充后的数据

集对 Faster RCNN 进行训练, 绘制 4 个阶段的损失曲线图。其中改进前 ZF 网络的损失曲线如图 4 所示, 整个训练过程共进行了 4 个阶段的  $1.8 \times 10^5$  次迭代, 迭代次数分别是  $6 \times 10^4$ ,  $3 \times 10^4$ ,  $6 \times 10^4$ ,  $3 \times 10^4$ 。从图 4 可以看出, 虽然改进前 ZF 网络的损失曲线最后都可以收敛到一个范围, 但是整体效果不甚理想, 其中 RPN 和 Faster RCNN 第一阶段训练损失曲线均处于振荡幅度过大的状态。

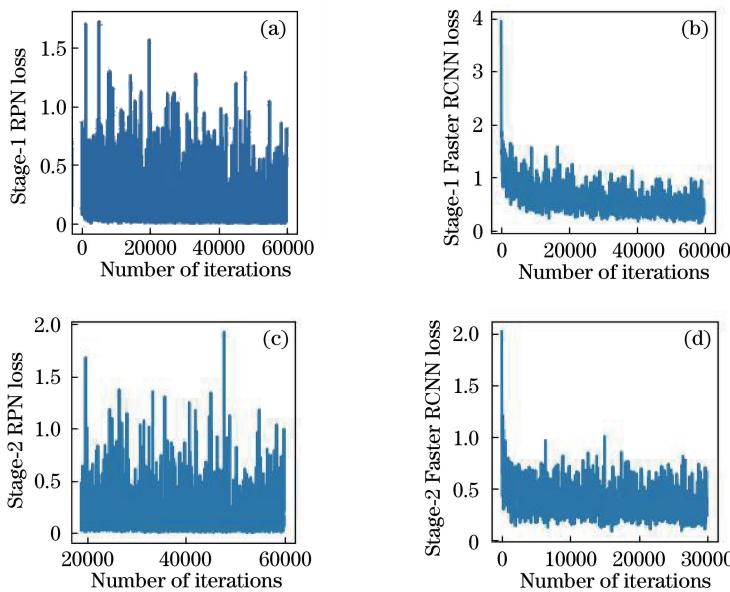


图 4 基于 ZF 网络的训练损失。(a) RPN 第一阶段训练损失; (b) Faster RCNN 第一阶段训练损失;  
(c) RPN 第二阶段训练损失; (d) Faster RCNN 第二阶段训练损失

Fig. 4 Training loss based on ZF network. (a) Stage-1 training loss of RPN; (b) stage-1 training loss of Faster RCNN; (c) stage-2 training loss of RPN; (d) stage-2 training loss of Faster RCNN

图 5 为改进后 ZF 网络的损失曲线图, 迭代次数与改进前相同。相比于改进前 ZF 网络, 改进后的 RPN 和 Faster RCNN 第一阶段训练损失曲线振荡范围过大的缺点均有所改善, 第二阶段训练损失曲线均处于一个值域更低的范围内。总体来说, 改

进后 ZF 网络的训练效果优于改进前。

## 2.4 网络训练

网络训练流程如下: 1) 初始化权值和偏置; 2) 将批量训练样本送入网络; 3) 经前向传播及反向传播更新参数; 4) 迭代后得到最终的模型及准确率。

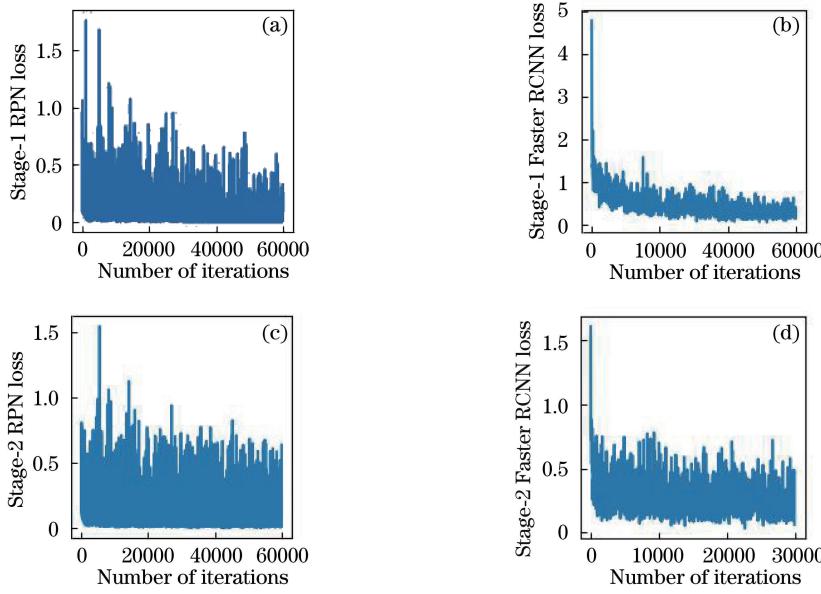


图 5 改进后 ZF 网络的训练损失。(a) RPN 第一阶段训练损失;(b) Faster RCNN 第一阶段训练损失;  
(c) RPN 第二阶段训练损失;(d) Faster RCNN 第二阶段训练损失

Fig. 5 Training loss based on improved ZF network. (a) Stage-1 training loss of RPN; (b) stage-1 training loss of Faster RCNN; (c) stage-2 training loss of RPN; (d) stage-2 training loss of Faster RCNN

Caffe(快速特征嵌入的卷积体系结构)中的Filler层可以对权值和偏置进行初始化。其中, Xavier 和 MSRA 都是不错的初始化方法, 考虑到本文的激活函数, 选择 MSRA 初始化权值和偏置。批量大小的选择同样至关重要: 如果选择过小, 会导致损失函数不收敛且训练速度慢; 如果选择过大, 会导致网络收敛到一些效果不佳的局部最优点。考虑到数据集大小及内存容量大小, 将批量大小设置为 32。

前向传播的计算公式为

$$\mathbf{Y} = s_n \{ \cdots \{ s_2 [s_1 (\mathbf{X}\boldsymbol{\omega}_1 + \mathbf{b}_1) \boldsymbol{\omega}_2 + \mathbf{b}_2] \cdots] \boldsymbol{\omega}_n + \mathbf{b}_n \}, \quad (2)$$

式中:  $\mathbf{X}, \mathbf{Y}$  分别为网络的输入和输出矩阵;  $\boldsymbol{\omega}_i$  为权值矩阵;  $\mathbf{b}_i$  为偏置矩阵;  $s_i$  为每个神经元的激活函数;  $i=1, 2, \dots, n$ 。激活函数除了最后一层为计算分类得分的 Softmax 函数外, 其余均采用 Relu 函数。这是因为 Relu 函数在激活状态下的梯度能维持在较大且一致的水平, 减小训练过程中出现梯度消失和梯度爆炸等问题的概率。

完成一次前向传播后, 计算损失函数的值, 根据权值更新算法更新权值, 重新进行前向传播, 直至损失函数取得最小值。选择适用于分类问题的交叉熵损失函数<sup>[22]</sup>, 其表达式为

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^k y_i \log(p_j), \quad (3)$$

式中:  $y_i$  为期望输出值;  $k$  为样本总类别数;  $N$  为本批

次训练样本总数;  $p_j$  为 Softmax 函数。但是单一的交叉熵损失函数的网络泛化能力一般, 有可能出现过拟合, 最后得到的效果不佳。为了防止过拟合, 除了增加训练样本的数量和多样性之外, 还可以引入 L2 范数正则化项, 即在原损失函数后面添加一个 L2 范数正则项来完成对权重的正则化, 其表达式为

$$r(\mathbf{w}) = \frac{1}{2} \|\mathbf{w}\|_2^2, \quad (4)$$

式中:  $\mathbf{w}$  为网络的权重。带有 L2 范数正则化项的损失函数为

$$\tilde{L}(\mathbf{w}; \mathbf{X}, \mathbf{Y}) = L(\mathbf{w}; \mathbf{X}, \mathbf{Y}) + \lambda \frac{1}{2} \|\mathbf{w}\|_2^2 = L(\mathbf{w}; \mathbf{X}, \mathbf{Y}) + \frac{\lambda}{2} \mathbf{w}^\top \mathbf{w}, \quad (5)$$

式中:  $\lambda \in [0, \infty)$  表示正则化项系数。则损失函数对  $\mathbf{w}$  的梯度为

$$\nabla_{\mathbf{w}} \tilde{L}(\mathbf{w}; \mathbf{X}, \mathbf{Y}) = \nabla_{\mathbf{w}} L(\mathbf{w}; \mathbf{X}, \mathbf{Y}) + \lambda \mathbf{w}. \quad (6)$$

执行一次梯度下降后  $\mathbf{w}$  为

$$\mathbf{w} = \mathbf{w} - \alpha \nabla_{\mathbf{w}} \tilde{L}(\mathbf{w}; \mathbf{X}, \mathbf{Y}) = \mathbf{w} - \alpha [\nabla_{\mathbf{w}} L(\mathbf{w}; \mathbf{X}, \mathbf{Y}) + \lambda \mathbf{w}], \quad (7)$$

即

$$\mathbf{w} = (1 - \alpha \lambda) \mathbf{w} - \alpha \nabla_{\mathbf{w}} L(\mathbf{w}; \mathbf{X}, \mathbf{Y}), \quad (8)$$

式中:  $\alpha$  为学习率。同理可得, 未添加 L2 范数正则化项的原损失函数使用梯度下降法,  $\mathbf{w}$  为

$$\mathbf{w} = \mathbf{w} - \alpha \nabla_{\mathbf{w}} L(\mathbf{w}; \mathbf{X}, \mathbf{Y}). \quad (9)$$

对比(8)式和(9)式可以看到, 添加 L2 范数正则

化项后,在执行每步梯度下降前会使前一步的  $w$  收缩,这使得添加 L2 范数正则化项后的损失函数最小时的解  $w^*$  会沿着一定方向衰减,并且沿着此下降方向的参数会保留,而沿着不重要方向的参数会逐渐衰减,这有助于提升网络训练效率和泛化能力。

选择带动量随机梯度下降法(SGD)作为网络的权值更新算法。之所以未选择普通的 SGD,是因为普通 SGD 的学习率是固定的,如果学习率的初始值设定不合适,就会导致损失函数的振荡或不收敛。引入动量后,其作用是在函数下降阶段积累参数之前更新时刻的梯度方向的速度,其他方向的速度会被限制,使其快速下降,并在函数上升时提供反向作用力使其停下并收敛到最小值。

学习率影响权值更新速度,其大小变化与权值更新速度呈正相关。假设学习率共进行  $\vartheta$  次迭代,则第  $v$  次迭代时学习率的值为

$$\alpha_v = \left(1 - \frac{v}{\vartheta}\right)\alpha_0 + \frac{v}{\vartheta}\alpha_\vartheta, \quad (10)$$

式中:初始学习率  $\alpha_0 = 0.1$ 。在  $\vartheta$  次迭代后,学习率会保持一个恒定的值,这样会增加网络训练收敛的概率。

### 3 分析与讨论

实验采用 Caffe 深度学习平台, GPU 型号为 GTX1060ti。图 6 为使用 labelImg 标注训练集的软件界面。图 7 为基于 Faster RCNN 和 Faster RCNN+FPN 的检测结果。由图 7(a)可以看出, Faster RCNN 在面对大目标时,识别率还是可观的,但是在面对小目标时,会出现漏检的情况;而图 7(b)使用 Faster RCNN+FPN 就很好地避免了这一情况,准确度较 Faster RCNN 有提升,而且识别出 Faster RCNN 未能识别的小目标。

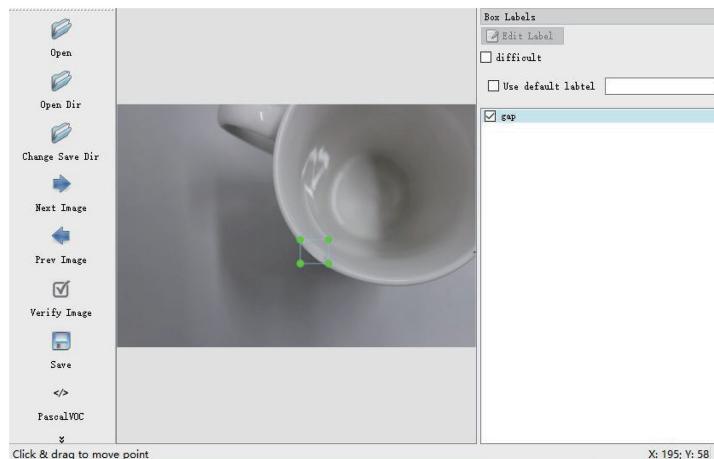


图 6 使用 labelImg 标注数据集

Fig. 6 Labeling data set with labelImg

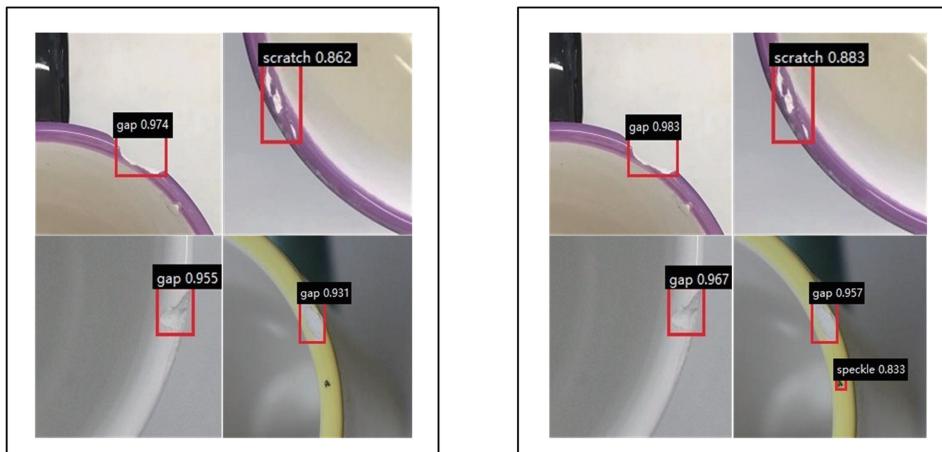


图 7 马克杯缺陷检测结果对比图。(a)原始 Faster RCNN 检测结果;(b)加入 FPN 后的 Faster RCNN 检测结果

Fig. 7 Comparison of mug defect inspection results. (a) Original Faster RCNN; (b) Faster RCNN after FPN addition

对于 Faster RCNN + FPN 来讲,融合不同层数的底层特征,最后的检测表现也不同。为了验证是否

融合层数越多越好,使用相同的数据集对不同模型进行检测,结果见表 2,AP 表示某一类缺陷的准确率。

表 2 不同网络结构的分类性能

Table 2 Comparison of various network structures on classification performance

Network structure	AP / %			Average detection time / s
	Scratch	Speckle	Gap	
Faster RCNN	85.32	0	97.03	0.094
Faster RCNN and two layers of FPN	87.86	81.11	98.30	0.123
Faster RCNN and three layers of FPN	88.32	82.36	98.51	0.135
Faster RCNN and four layers of FPN	88.56	82.98	98.76	0.149

从表 2 可以看出,在使用传统 Faster RCNN 检测时,其准确率尚可,平均检测时间较短,但也有不足,在 Faster RCNN 的斑点 AP 为 0 即意为其无法检测出较小的缺陷;而使用 Faster RCNN + 若干 FPN 层时,添加的 FPN 层数越多,即结合的特征尺度数量越多,对应缺陷的 AP 也越高,说明该网络结构的性能越好。增加到 4 层 FPN 时,检测性能达到最佳。缺口和划痕缺陷的 mAP(即所有类别缺陷的平均准确率)为 93.66%,相比于原始 Faster RCNN 的 91.175%,提升了 2.485 个百分点。但是,Faster RCNN+FPN 使用了额外的浅层特征,导致其比 Faster RCNN 的检测时间长 0.029~0.055 s。

## 4 结 论

以 Faster RCNN 为基础,提出了一种改进 Faster RCNN 的马克杯缺陷检测方法。原始 Faster RCNN 仅采用网络的最后一层特征,无法提供足够的细节特征,而引入 FPN 后,使得更多的浅层特征参与到整个算法中,利用浅层特征细节信息丰富的特点,将浅层特征和深层特征相融合,丰富了特征的语义信息。最后在 Caffe 框架上分别构建 Faster RCNN 和 Faster RCNN+FPN 网络,利用标记好的数据集对两个网络进行训练,最后得到融合 4 层网络的 Faster RCNN+FPN 对缺口和划痕缺陷的检测准确率为 93.66%,相比于单独使用 Faster RCNN 提升了 2.485 个百分点,并且检测出 Faster RCNN 未能识别的直径小于 1 mm 的斑点缺陷。检测准确率的提升,将减小合格马克杯检测为缺陷马克杯的概率。马克杯小缺陷的准确识别,避免更多的小缺陷马克杯流入市场,对建立良好的企业形象和提升消费者用户体验均有裨益。

## 参 考 文 献

[1] Liu M Z, Ma J, Zhang M, et al. Online operation method for assembly system of mechanical products

- based on machine vision [J]. Computer Integrated Manufacturing Systems, 2015, 21(9): 2343-2353.
- 刘明周, 马靖, 张森, 等. 基于机器视觉的机械产品装配系统在线作业方法 [J]. 计算机集成制造系统, 2015, 21(9): 2343-2353.
- [2] Zhou X E, Wang Y N, Zhu Q, et al. Research on defect detection method for bottle mouth based on machine vision [J]. Journal of Electronic Measurement and Instrumentation, 2016, 30 (5): 702-713.
- 周显恩, 王耀南, 朱青, 等. 基于机器视觉的瓶口缺陷检测方法研究 [J]. 电子测量与仪器学报, 2016, 30(5): 702-713.
- [3] Wang P, Zhu L, Zhu Q J, et al. An application of back propagation neural network for the steel stress detection based on Barkhausen noise theory[J]. NDT & E International, 2013, 55: 9-14.
- [4] Xie L J, Huang R, Gu N, et al. A novel defect detection and identification method in optical inspection[J]. Neural Computing and Applications, 2014, 24(7/8): 1953-1962.
- [5] Halfawy M R, Hengmeechai J. Automated defect detection in sewer closed circuit television images using histograms of oriented gradients and support vector machine [J]. Automation in Construction, 2014, 38: 1-13.
- [6] LeCun Y, Boser B, Denker J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural Computation, 1989, 1 (4): 541-551.
- [7] Dai J F, Li Y, He K M, et al. R-FCN: object detection via region-based fully convolutional networks[C] // Proceedings of the 30th International Conference on Neural Information Processing Systems, December 5-10, 2016, Barcelona, Spain. San Diego: NIPS, 2016: 379-387.
- [8] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets [J]. Neural

- Computation, 2006, 18(7): 1527-1554.
- [9] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks [C] // Advances in Neural Information Processing Systems, December 3-6, 2012, Lake Tahoe, Nevada, United States. San Diego: NIPS, 2012: 1097-1105.
- [10] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, Ohio. New York: IEEE, 2014: 580-587.
- [11] Girshick R. Fast R-CNN [C] // 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 1440-1448.
- [12] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39 (6): 1137-1149.
- [13] Sun Y, Liang D, Wang X G, et al. Deepid3: face recognition with very deep neural networks[EB/OL]. (2015-02-03) [2019-05-29]. <https://arxiv.org/abs/1502.00873>.
- [14] Hafemann L G, Sabourin R, Oliveira L S. Learning features for offline handwritten signature verification using deep convolutional neural networks[J]. Pattern Recognition, 2017, 70: 163-176.
- [15] Abdel-Hamid O, Mohamed A R, Jiang H, et al. Convolutional neural networks for speech recognition [J]. ACM Transactions on Audio, Speech, and Language Processing, 2014, 22(10): 1533-1545.
- [16] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 779-788.
- [17] Zhao Z B, Liu N, Wang L. Localization of multiple insulators by orientation angle detection and binary shape prior knowledge [J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2015, 22 (6): 3421-3428.
- [18] Yao M H, Chen Z H. Deep active learning in detection of surface defects on magnetic sheet [J]. Computer Measurement & Control, 2018, 26 (9): 29-33.
- 姚明海, 陈志浩. 基于深度主动学习的磁片表面缺陷检测[J]. 计算机测量与控制, 2018, 26(9): 29-33.
- [19] Hong W, Li C F. Flame detection method based on regional fully convolutional networks with residual network [J]. Laser & Optoelectronics Progress, 2018, 55(4): 041011.
- 洪伟, 李朝锋. 基于区域全卷积网络结合残差网络的火焰检测方法[J]. 激光与光电子学进展, 2018, 55 (4): 041011.
- [20] Han X, Zhao R, Sun F S. Methods for location and recognition of chess pieces based on convolutional neural network [J]. Laser & Optoelectronics Progress, 2019, 56(8): 081007.
- 韩燮, 赵融, 孙福盛. 基于卷积神经网络的棋子定位和识别方法[J]. 激光与光电子学进展, 2019, 56 (8): 081007.
- [21] Ou P, Zhang Z, Lu K, et al. Object detection in of remote sensing images based on convolutional neural networks [J]. Laser & Optoelectronics Progress, 2019, 56(5): 051002.
- 欧攀, 张正, 路奎, 等. 基于卷积神经网络的遥感图像目标检测[J]. 激光与光电子学进展, 2019, 56 (5): 051002.
- [22] Liu C. Research on surface defects detection of micro parts based on convolution neural network [D]. Harbin: Harbin University of Science and Technology, 2019.
- 刘聪. 基于卷积神经网络的微小零件表面缺陷检测技术研究[D]. 哈尔滨: 哈尔滨理工大学, 2019.