

瓷砖表面非常规尺寸瑕疵的检测^①



陈 凯, 鲁涵统, 程浩宇, 方梦园

(浙江理工大学 信息学院, 杭州 310018)

通信作者: 方梦园, E-mail: myfang@zstu.edu.cn

摘 要: 与普通目标检测任务不同, 瓷砖表面瑕疵检测的困难之处在于检测小尺寸和大长宽比等非常规尺寸的目标。为了解决这两个问题, 本文提出了一种基于改进 Cascade R-CNN 的新型瓷砖表面瑕疵检测算法。为了提高模型对小瑕疵的检测能力, 本文模型利用侧向连接结构进行上下层语义信息的融合, 使用可切换空洞率的空洞卷积来增加模型的感受野; 为了提高模型对于大长宽比瑕疵的检测能力, 本文模型在标准卷积上引入偏移域以更好提取目标特征信息。此外, 本文模型调整 Cascade R-CNN 框架中预选锚框的大小和长宽比例。实验结果表明, 在从瓷砖工厂收集的数据集上, 本文所提出算法的平均精度均值 (mean average precision, mAP) 达到了 73.5%, 比改进前的 Cascade R-CNN 模型提高了 9.7%。本文实验代码可从以下链接获取: <https://github.com/mashibin/Ceramic-tile-defect-detection>。

关键词: 目标检测; Cascade R-CNN; 侧向连接; 空洞卷积; 偏移域; 深度学习; 卷积神经网络

引用格式: 陈凯, 鲁涵统, 程浩宇, 方梦园. 瓷砖表面非常规尺寸瑕疵的检测. 计算机系统应用, 2022, 31(11): 192-198. <http://www.c-s-a.org.cn/1003-3254/8807.html>

Unconventional Size Defects Detection for Tile Surface

CHEN Kai, LU Han-Tong, CHENG Hao-Yu, FANG Meng-Yuan

(School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: Different from ordinary object detection tasks, the difficulty of detecting tile surface defects lies in the detection of unconventional size objects, such as small-sized objects and objects with large aspect ratios. To solve these two problems, this study proposes a new type of tile surface defect detection algorithm based on improved Cascade R-CNN. To improve the detection ability for small defects, the model in this study uses the lateral connection structure to fuse the semantic information of the upper and lower layers and applies the dilated convolution with switchable dilation rates to increase the receptive field of the model. To improve the detection ability for defects with large aspect ratios, the proposed model introduces an offset field on the standard convolution to better extract the object feature information. In addition, the model adjusts the size and length of the pre-selected anchor box in the Cascade R-CNN framework. The experimental results show that on the dataset collected from the tile factory, the mean average precision (mAP) of the proposed algorithm reaches 73.5%, which is 9.7% higher than that of the Cascade R-CNN model before improvement. The experimental code of this study is available at: <https://github.com/mashibin/Ceramic-tile-defect-detection>.

Key words: object detection; Cascade R-CNN; lateral connection; dilated convolution; offset field; deep learning; convolutional neural networks (CNN)

瓷砖厂需要在产品生产出来后进行瑕疵检测过程, 通常采用人工目检的方法。但是这种方法不仅劳动强

度大, 而且无法保证检测的准确性, 特别是难以检测出小尺寸瑕疵。

^① 基金项目: 国家自然科学基金 (61903338); 浙江省自然科学基金 (LQ19F030015); 辽宁省自然科学基金 (2019-KF-23-02)

收稿时间: 2022-03-03; 修改时间: 2022-04-02; 采用时间: 2022-04-15; csa 在线出版时间: 2022-07-15

瓷砖产品的瑕疵种类很多,常见的瑕疵包括边异常、角异常、白色点瑕疵、浅色块瑕疵、深色点块瑕疵、光圈瑕疵等。在这些瑕疵中,其中存在很多小尺寸瑕疵和一定数量的大长宽比瑕疵,给瑕疵检测带来了困难。因此,为了提高检测能力,需要根据小瑕疵和大长宽比瑕疵的特点设计针对此类瑕疵的检测模型。

现有的瓷砖表面瑕疵检测算法可分为两类,即传统的目标检测算法和基于深度学习的检测算法。

传统的目标检测算法一般需要基于一定的规则进行图像特征提取。例如权小霞等^[1]提出了一种基于局部方差加权信息熵的瓷砖表面瑕疵检测算法,与以往的瑕疵检测算法相比,该算法在保留瑕疵图像的细节信息方面有显著的提高,同时对低质量瓷砖图像的识别也能取得较好的效果,但是该算法在检测实时性方面具有一定的不足。Zou等^[2]提出了一种基于图像增强和区域增长的瓷砖表面瑕疵检测算法,所提算法在裂纹、孔洞和色差瑕疵的检测上表现出色。但是所提出的方法具有一定的遗漏和错误检测,特别是对于麻点瑕疵,检测精度较低。陈雪云等^[3]提出一种线性逻辑矢量模式特征用于遥感图像目标检测,能够以较短的训练时长达到高精度及广适应性的双重标准。这些算法的检测精度可达到要求,但是均存在一定的局限性,即需要预先提取待检测图像的特征,而面对如瓷砖表面瑕疵这类复杂图像,需要提取的特征过于复杂。

基于深度学习的目标检测算法可以分为两大类,即以YOLO系列^[4-7]为代表的一阶段目标检测算法和以Faster R-CNN^[8]、Cascade R-CNN^[9]等为代表的两阶段目标检测算法,这两类算法均在工业场景下的瑕疵检测中得以应用。Wei等^[10]以YOLOv3网络结构为主要模型,对磁瓦表面瑕疵实现了实时检测。Li等^[11]提出了一种基于YOLOv5的轻量级遥感旋转目标检测模型,由于其模型轻巧的特点,可轻松部署到相应平台。YOLO虽然具有模型规模相对较小、检测速度快等优点,但对小尺寸目标的检测能力较弱^[4]。

两阶段的目标检测算法将目标检测分为特征提取阶段和RCNN阶段,这类算法虽然与一阶段算法相比计算量较大,但往往具有更高的精度^[8]。该类算法有在特征提取阶段作出改进的,Fang等^[12]使用更灵活的上下文信息集成方法改进了Faster R-CNN,改进后的Faster R-CNN算法在小物体检测的准确率和召回率上都有很好的表现。Li等^[13]将特征金字塔网络和可变形

卷积引入特征提取网络,该算法基本可以满足实时性和准确性的要求,同时对长宽比例与旋转角度非常规的目标检测能力有所提高。

两阶段检测算法也有在RCNN阶段作出改进的,如Zhang等^[14]将Cascade R-CNN多检测头的思想用于检测带一定旋转角度目标的R2CNN,并提出一种新的RoI池化方式(SERoI pooling),在识别船舶光学遥感图像上的小物体表现优异。

考虑到瓷砖表面瑕疵的检测需要对小瑕疵具有较高的检测精度,本文提出以Cascade R-CNN为主要模型结构,即在Faster R-CNN的基础上增加多个检测分支以进一步提高检测能力,并在此基础上针对小尺寸瑕疵和大长宽比瑕疵的检测提出一些改进。

1 本文方法

1.1 方法总览

本文提出一种瓷砖瑕疵检测方法专门用于检测小瑕疵和大长宽比的瑕疵。其中包括针对非常规尺寸瑕疵的特征提取网络与RCNN阶段作出的改进。

如图1所示,特征提取阶段中,输入图像进入Backbone产生特征图。在下采样过程最后3个阶段,我们使用带有偏移域的卷积网络来更好提取大长宽比瑕疵特征信息,为使模型在尽可能不损失太多特征信息的前提下增加感受野,我们引入了空洞卷积,并使用锁机制进行标准卷积与空洞卷积之间的软切换。本文模型还在Backbone中使用侧向连接结构来融合上下层信息,减少下层特征信息的丢失,保留检测小瑕疵的能力,同时模型引入了特征反馈以进一步融合信息。

在RCNN部分,本文通过调整预选框的尺寸与比例来适应不同形状的瑕疵,并使用RoI Align代替常用的RoI Pooling,进一步提升检测精度。

1.2 针对非常规尺寸瑕疵的特征提取网络

1.2.1 多层特征侧向连接与反馈

在原始的特征提取网络中,主干网络经过连续的下采样损失了大量的特征图信息,从而限制了检测小瑕疵的能力。然而下采样得到的上层信息是目标分类所必须的,融合上层与下层信息可以很好地解决这个问题。特征金字塔网络(feature pyramid networks, FPN)^[15]提出带有侧向连接的自顶向下网络结构,本文将带有侧向连接的自顶向下网络结构引入Cascade R-CNN架构,对上层特征进行上采样并与下层特征融合,从而保

留下层的信息, 增强模型检测小瑕疵的能力, 使用 B_i 表示主干网络第 i 阶段的输出, F_i 表示第 i 个 FPN 操作, 主干网络与 FPN 融合后输出用一组特征图表示

$\{f_i | i = 1, \dots, S\}$, 其中 S 表示阶段数, 如图 1 所示, $S = 4$, 输出特征 f_i 可以表示为式 (1), x_0 表示输入图像 并且 $f_{S+1} = 0$.

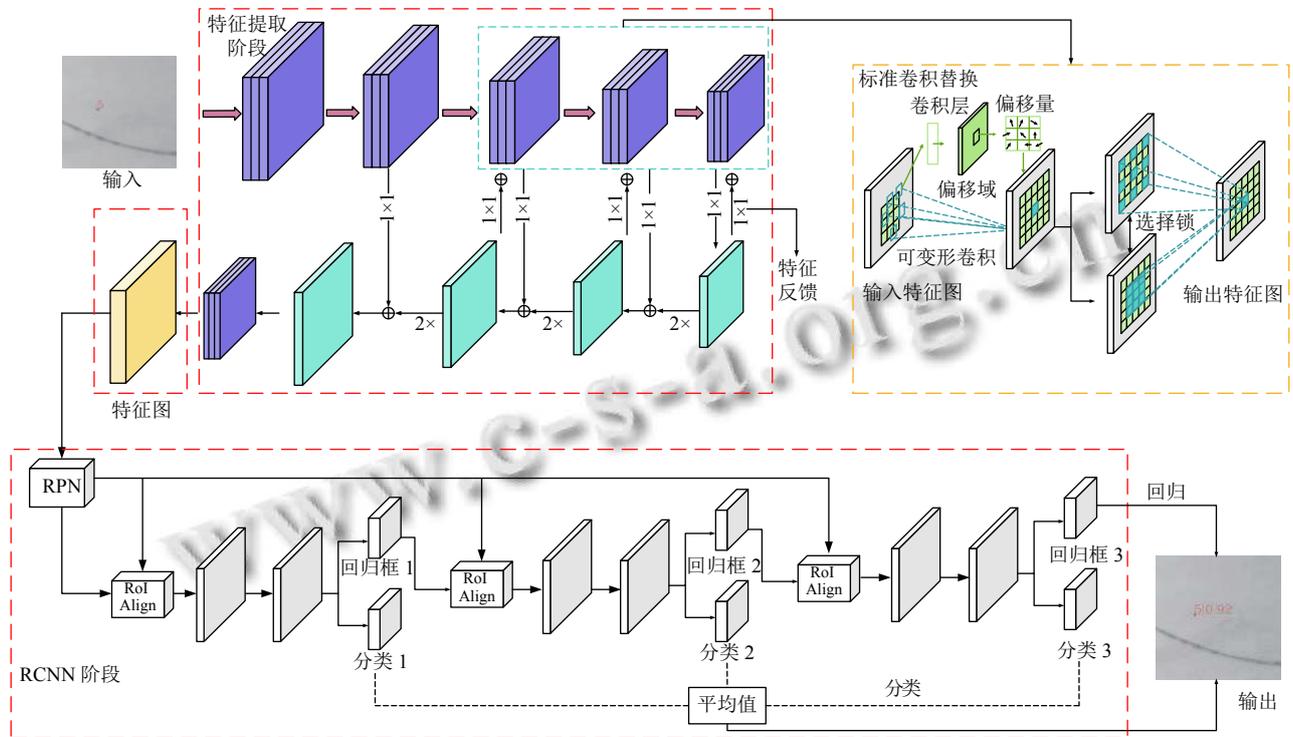


图 1 本文模型架构

$$f_i = F_i(f_{i+1}, x_i), \quad x_i = B_i(x_{i-1}) \quad (1)$$

反馈机制在深度学习中并不少见, Cao 等^[16] 引入了反馈回路, 根据网络的目标 (例如高级语义标签) 来推断隐藏层神经元的激活状态, 反馈网络有助于更好地可视化和理解深度神经网络的工作原理, 并捕捉预期对象的视觉注意力. Qiao 等^[17] 在 FPN 的基础上加入反馈连接, 进一步丰富特征信息, 在 COCO 数据集上作实例分割与全景分割均实现了很好的效果. 如图 1 所示, 本文通过将特征金字塔的语义信息传回主干网络, 使下层特征进一步融合, 提高模型检测小瑕疵的能力. 最终输出特征 f_i 可以表示为式 (2), 使用 R_i 表示反馈之前的融合特征, 其中 $\forall i = 1, \dots, S$.

$$f_i = F_i(f_{i+1}, x_i), \quad x_i = B_i(x_{i-1}, R_i(f_i)) \quad (2)$$

1.2.2 可软切换空洞卷积

原始的特征提取网络通常在主干网络中使用传统的卷积操作. 传统的卷积操作通过加深网络和使用下采样操作来增加特征图的感受野, 从而提高分类的准

确性, 但是却丢失了更多的细节信息, 降低了对小目标的检测能力.

空洞卷积^[18] 是一种在不增加额外参数和计算的情况下可以有效增加模型感受野的方法, 并且不需要过多的下采样操作, 因此它会保持模型对小目标的检测能力. DeepLab 系列模型^[19-22] 中使用多种空洞率的空洞空间金字塔池化 (atrous spatial pyramid pooling, ASPP), 使模型对多尺度的分割获得较好的效果. 但是, 固定空洞率的空洞卷积无法适应不同的目标尺度. Qiao 等^[17] 提出在不同空洞率的卷积网络中加入锁机制, 模型可以根据目标作出自适应调整.

本文将空洞卷积扩展为可以在不同空洞率之间软切换的结构, 这是通过使用并行结构和切换功能实现的 (如图 2 所示). 从一个标准卷积层到可调节空洞卷积的转换过程可以表示为式 (3):

$$\begin{aligned} Conv(x, \omega, 1) &\xrightarrow[SAC]{\text{Convert to}} S(x) \times Conv(x, \omega, 1) \\ &+ (1 - S(x)) \times Conv(x, \omega + \Delta\omega, r) \end{aligned} \quad (3)$$

其中, x 是输入, ω 是标准卷积操作的权重, r 是空洞卷积的空洞率 (也是可调节空洞卷积的超参数), $\Delta\omega$ 表示标准卷积和空洞卷积之间的可训练差异, $S(x)$ 是由 5×5 平均池化层和 1×1 卷积层组成的锁函数. $S(x)$ 是与特

征输入和位置相关的, 因此主干网络能够适应不同的目标尺度. 如图 1 所示, 本文使用可调节空洞卷积替换了 Cascade R-CNN 架构的主干网络中最后 3 个阶段的卷积层.

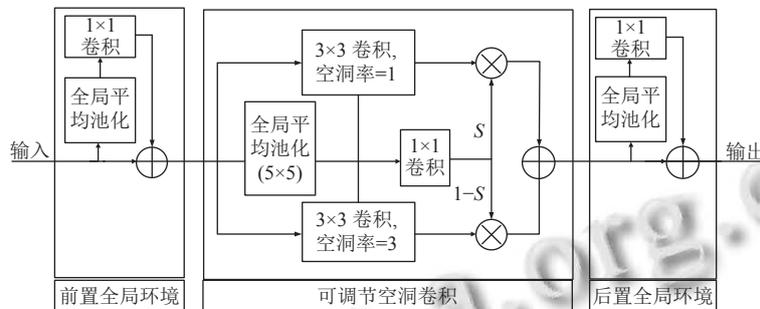


图 2 可调节空洞卷积, 两个全局信息模块通过添加图像层面信息对特征进行处理

1.2.3 带偏移域的卷积网络

瓷砖表面瑕疵中存在一定数量的大长宽比瑕疵, 如边异常瑕疵. 传统的卷积方法存在硬性的规则, 在针对物体放大或旋转时, 不能作出自适应性改变. 由于瓷砖表面瑕疵存在长宽比大的特性, 使用标准卷积提取信息往往不够充分. Dai 等^[23] 在标准卷积采样上加入偏移量, 使模型能针对不规则形状目标进行充分提取信息. 如图 1 所示, 本文在标准卷积层上增加了一个偏移域, 它为卷积核中的每个采样点设置了一个偏移值, 然后通过学习偏移值, 可以根据对象动态调整可变形卷积核的大小和位置. 该卷积过程可以表示为式 (4):

$$y(p_0) = \sum_{p_n \in R} \omega(p_n) \times I(p_0 + p_n + \Delta p_n) \quad (4)$$

其中, p_0 是特征输出图 y 的卷积核采样点, R 表示采样点集合, $\omega(\cdot)$ 表示权重, I 表示输入特征图, p_n 表示枚举集合 R 中每个采样点位置, Δp_n 是偏移值. 通过这种方式可以得到对形状不规则的目标更好的检测精度, 例如大长宽比瑕疵. 本文使用带偏移域的卷积替换了 Cascade R-CNN 架构的主干网络中最后 3 个阶段的标准卷积层.

1.3 RCNN 阶段

本文采用 Cascade R-CNN 作为原型模型, 在 Faster R-CNN 的 R-CNN 检测头之后增加了两个结构相同的检测头 (如图 1 所示), 并通过逐渐增加的方式设置 IoU 阈值. 前一阶段的输出是下一阶段的输入, 这种级联结构使得回归边界框能更准确地定位检测对象. 与 Faster R-CNN 相比, Cascade R-CNN 在检测精度上有

明显的提升.

本文在标准预选框上增加了更多比例和尺度的锚框, 使模型更加容易生成精准回归检测框.

Cascade R-CNN 在 RCNN 阶段使用 RoI Pooling, 对于检测大目标影响不大, 但是对于小目标检测会损失一定精度, 本文选择使用 RoI Align 能够得到更精确的结果^[24].

2 实验及结果分析

2.1 数据处理及参数设置

本文在实验中使用的数据集来自中国佛山的一家瓷砖厂. 数据集包含 7150 张图像, 并将训练集和测试集以 9:1 的比例随机划分, 如图 3 所示图像中包含的瑕疵种类大致可分为: (1) 边异常; (2) 角异常; (3) 白色点瑕疵; (4) 浅色块瑕疵; (5) 深色点块瑕疵; (6) 光圈瑕疵. 原始图像大小为 8192×6000 和 4096×3500 , 为了增加图像中小瑕疵的面积占比, 并且减少模型训练的内存需求, 本文使用步长为 512 的滑动窗口将图像裁剪为 640×640 的大小作为模型的输入. 此外, 本文使用水平翻转、镜像翻转等方法来增强数据规模, 并在训练集图像中加入高斯模糊以增强模型的泛化能力.

在模型参数方面, batch-size 设置为 8, 总共训练了 50 个 epoch, 使用随机梯度下降 (stochastic gradient descent, SGD) 作为优化器, 动量设置为 0.9. 模型的损失函数由两部分组成, 分类的损失函数是交叉熵损失, 回归边界框的损失函数是 L1 损失的平滑版本. 学习率

初始化为 0.005, 在第 25 和第 40 个 epoch 时, 将它降至前一个值的 1/3. 3 个级联检测头的 IoU 阈值分别设置为 0.4, 0.5 和 0.6. 为了进一步提高小瑕疵的检测能力, 本文在训练和测试阶段将原始的 640×640 输入随机放大到 640×640、800×640 和 1000×640 的大小.

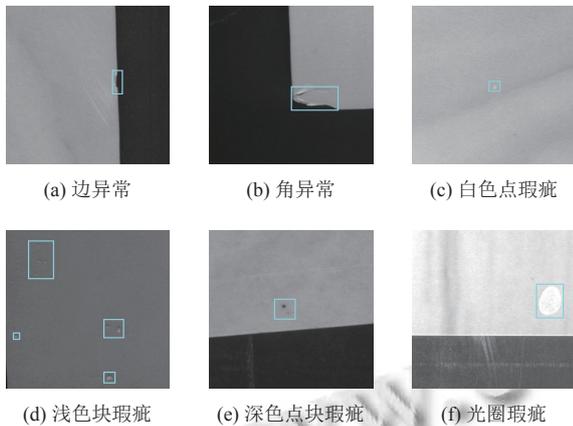


图3 瓷砖表面瑕疵

此外, RPN 中锚框的长宽比默认为 1:1, 1:2, 2:1, 为了进一步提高大长宽比瑕疵的检测能力, 本文将锚框的长宽比调整为 1:5, 1:2, 1:1, 2:1, 5:1, 1:10, 10:1. 此外, 对于锚框的尺度, 除了原来的 64×64 像素尺度之外, 本文增加了两个更小的尺度, 即 32×32 和 16×16, 以提高对小瑕疵的检测能力. 实验结果见表 1, 结果显示此操作 (在表 1 中表示为多尺度锚框) 显著改善了平均精度均值 (mean average precision, mAP) 值. 运行环境为 Ubuntu 18.04, 软件环境为 Python 3.7 和 PyTorch 1.6, GPU 设备是两张 NVIDIA 2080Ti 显卡, CPU 是 i7-1170F.

表 1 各模型检测结果

方法	模型结构					多尺度锚框	mAP (%)	FPS	参数量 (M)
	级联检测头	FPN	DCN	RFP	SAC				
方法1	—	—	—	—	—	—	57.6	24	57.4
方法2	√	—	—	—	—	—	63.8	18	170.6
方法3	√	√	—	—	—	—	65.4	16	228
方法4	√	—	√	—	—	—	66.5	4.8	172.2
方法5	√	—	√	—	—	√	72.8	4.8	172.2
方法6	√	—	—	√	—	—	67.0	2.5	228
方法7	√	—	—	√	√	—	69.5	2.1	228
本文	√	—	√	√	√	√	73.5	2	229.6

2.2 实验结果

本文提出的方法与以下方法进行了比较: 方法 1 为使用 ResNet50 作为骨干网络的 Faster R-CNN; 方

法 2 使用 ResNet50 作为骨干网络的 Cascade R-CNN (这是原始的 Cascade R-CNN); 方法 3 在方法 2 的基础上增加侧向连接; 方法 4 在方法 2 的基础上增加带偏移域的卷积; 方法 5 在方法 2 的基础上, 增加带偏移域的卷积, 调整锚框 (记为“多尺度锚框”); 方法 6 在方法 2 的基础上增加侧向连接与反馈机制; 方法 7 在方法 2 的基础上增加侧向连接与反馈机制以及可软切换的空洞卷积. 具体来说, 本文所提出的方法是在方法 2 的基础上添加侧向连接、反馈机制、可软切换的空洞卷积、带偏移域的卷积和“多尺度锚框”.

结果如表 1 所示, 本文主要从 mAP 方面进行比较. 此外, 本文还给出了每秒帧数 (frames per second, FPS) 和参数量以供参考. 可以看出, 添加带偏移域的卷积、侧向连接与反馈机制, mAP 分别增加了 2.7% 和 3.2%, 通过添加可软切换的空洞卷积, mAP 增加了 2.5%, 锚框的调整带来了 6.3% 的 mAP 增量, 最后本文所提出的方法实现了 73.5% 的 mAP. 此外, 实验还使用 YOLOv5s 进行比较, mAP 为 62.7%, 由于检测效果一般, 故并未放入表 1 中展示.

在表 2 中, 本文分别展示了所有 6 类瑕疵的平均精度 (average precision, AP) 值. 可以看出, 本文所提出的方法在大长宽比瑕疵 (边异常) 和小瑕疵 (白色点瑕疵、深色点块瑕疵) 方面表现出比 Faster R-CNN 和 Cascade R-CNN 更好的性能. 本文所提出的模型在光圈瑕疵上表现出性能下降. 然而, 这种类型的瑕疵既不是小尺寸瑕疵, 也不是具有大长宽比的瑕疵. 我们推测性能下降的原因是它的样本量很小. 由于所提出的方法包含较多的参数, 因此需要更多的样本进行训练.

表 2 模型针对 6 种瑕疵检测的 AP 值 (%)

算法	边异常	角异常	白色点瑕疵	浅色块瑕疵	深色点块瑕疵	光圈瑕疵
Faster R-CNN	70.6	83.1	47.8	37.1	34.0	73.4
Cascade R-CNN	66.0	86.2	54.2	43.2	56.5	76.5
本文	78.8	91.0	73.9	54.7	77.6	65.1

下面我们给出一些检测结果例子, 如图 4、图 5 (大长宽比的边异常), 和图 6、图 7 (小尺寸的深色点块瑕疵). 可以看出, Faster R-CNN 和原始的 Cascade R-CNN 在有些示例中未能检测到瑕疵, 而本文所提出的模型能够正确地检测到这些瑕疵

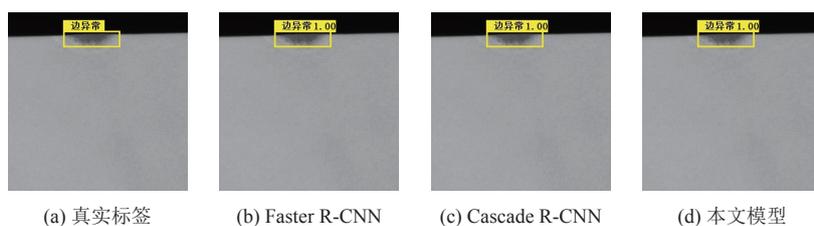


图4 边异常的真实标签及各模型检测结果示例1

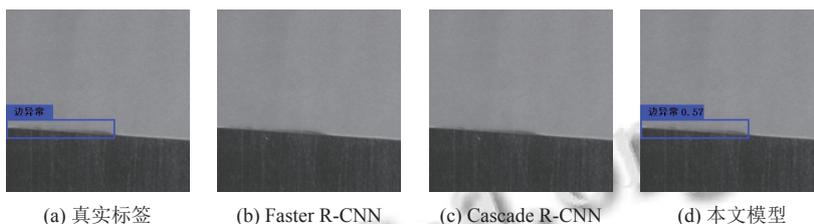


图5 边异常的真实标签及各模型检测结果示例2

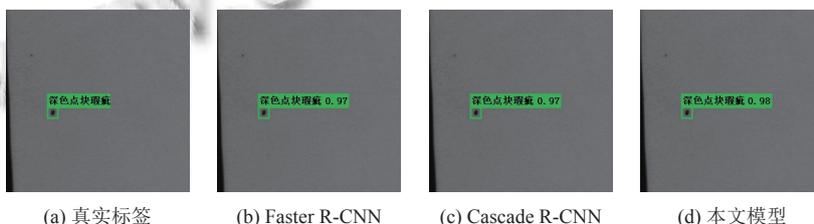


图6 深色点块瑕疵的真实标签及各模型检测结果示例1

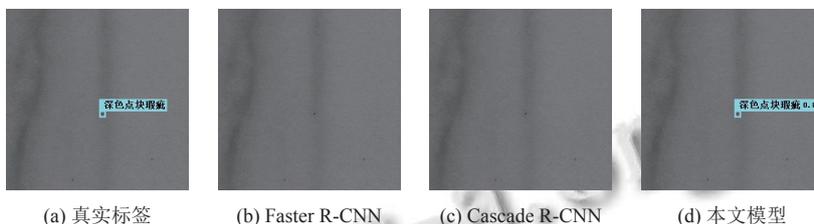


图7 深色点块瑕疵的真实标签及各模型检测结果示例2

3 结论与展望

本文中提出了一种基于 Cascade R-CNN 架构的瓷砖瑕疵检测方法, 其中进行了多项改进, 以增强对小尺寸瑕疵和大长宽比瑕疵的检测能力. 实验结果表明, 所提出的方法显著提高了对所用数据集的检测性能, 尤其是对小瑕疵和大长宽比瑕疵的检测性能. 但是, 该方法的缺点是模型参数较多. 结果导致检测速度下降到 FPS 为 2, 也可能导致模型对于具有小样本瑕疵类别的检测性能下降, 如光圈瑕疵. 实验结果表明, 参数数量的增加主要是来自于模型对于图像特征的再处理操作, 即侧向连接和反馈模块. 进一步优化特征再处理块以及瓷砖瑕疵检测的主干将会是未来的工作.

实验结果还表明, 在对检测速度要求较高的应用中, 可以在原有的 Cascade R-CNN 中只增加带偏移域的卷积和多尺度锚框, 即方法 5. 与本文所提出的方法比较, 该方法的 mAP 相当, 但 FPS 高得多.

另一方面, 本文专注于目标检测框架中主干和特征再处理块的改进. 在未来的工作中, 我们还将探索其他模块, 例如检测头中的损失函数, 希望这将带来进一步的性能改进, 并有助于解决光圈瑕疵上未解决的性能下降问题.

参考文献

- 1 权小霞, 李军华, 汪宇玲. 基于局部方差加权信息熵的瓷砖

- 表面缺陷检测. 中国陶瓷, 2019, 55(10): 46–55.
- 2 Zou GF, Li TT, Li GY, *et al.* A visual detection method of tile surface defects based on spatial-frequency domain image enhancement and region growing. Proceedings of the 2019 Chinese Automation Congress (CAC). Hangzhou: IEEE, 2019. 1631–1636.
 - 3 陈雪云, 黄金汉, 胡子灿, 等. 基于线性逻辑矢量模式的遥感图像目标检测. 浙江大学学报(工学版), 2022, 56(1): 47–55.
 - 4 Redmon J, Divvala S, Girshick R, *et al.* You only look once: Unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Vegas: IEEE, 2016. 779–788.
 - 5 Redmon J, Farhadi A. YOLO9000: Better, faster, stronger. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 6517–6525.
 - 6 Redmon J, Farhadi A. YOLOv3: An incremental improvement. arXiv: 1804.02767, 2018.
 - 7 Bochkovskiy A, Wang CY, Liao HYM. YOLOv4: Optimal speed and accuracy of object detection. arXiv: 2004.10934, 2020.
 - 8 Ren SQ, He KM, Girshick R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks. Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal: ACM, 2015. 91–99.
 - 9 Cai ZW, Vasconcelos N. Cascade R-CNN: Delving into high quality object detection. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 6154–6162.
 - 10 Wei JQ, Zhu PY, Qian X, *et al.* One-stage object detection networks for inspecting the surface defects of magnetic tiles. Proceedings of 2019 IEEE International Conference on Imaging Systems and Techniques (IST). Abu Dhabi: IEEE, 2019. 1–6.
 - 11 Li Y, Li MJ, Li SS, *et al.* Improved YOLOv5 for remote sensing rotating object detection. Proceedings of the 2021 6th International Conference on Communication, Image and Signal Processing (CCISP). Chengdu: IEEE, 2021. 64–68.
 - 12 Fang PC, Shi YJ. Small object detection using context information fusion in faster R-CNN. Proceedings of the 2018 IEEE 4th International Conference on Computer and Communications (ICCC). Chengdu: IEEE, 2018. 1537–1540.
 - 13 Li Y, Wang ZW. Research on textile defect detection based on improved cascade R-CNN. Proceedings of the 2021 International Conference on Artificial Intelligence and Electromechanical Automation (AIEA). Guangzhou: IEEE, 2021. 43–46.
 - 14 Zhang CG, Xiong BL, Kuang GY. Ship detection and recognition in optical remote sensing images based on scale enhancement rotating cascade R-CNN networks. Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium. Brussels: IEEE, 2021. 3545–3548.
 - 15 Lin TY, Dollár P, Girshick R, *et al.* Feature pyramid networks for object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 936–944.
 - 16 Cao CS, Liu XM, Yang Y, *et al.* Look and think twice: Capturing top-down visual attention with feedback convolutional neural networks. Proceedings of the IEEE International Conference on Computer Vision. Santiago: IEEE, 2015. 2956–2964.
 - 17 Qiao SY, Chen LC, Yuille A. DetectoRS: Detecting objects with recursive feature pyramid and switchable atrous convolution. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 10208–10219.
 - 18 Yu F, Koltun V, Funkhouser T. Dilated residual networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 636–644.
 - 19 Chen LC, Papandreou G, Kokkinos I, *et al.* Semantic image segmentation with deep convolutional nets and fully connected CRFs. arXiv: 1412.7062, 2014.
 - 20 Chen LC, Papandreou G, Kokkinos I, *et al.* DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834–848. [doi: 10.1109/TPAMI.2017.2699184]
 - 21 Chen LC, Papandreou G, Schroff F, *et al.* Rethinking atrous convolution for semantic image segmentation. arXiv: 1706.05587, 2017.
 - 22 Chen LC, Zhu YK, Papandreou G, *et al.* Encoder-decoder with atrous separable convolution for semantic image segmentation. Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich: Springer, 2018. 833–851.
 - 23 Dai JF, Qi HZ, Xiong YW, *et al.* Deformable convolutional networks. Proceedings of the IEEE International Conference on Computer Vision. Venice: IEEE, 2017. 764–773.
 - 24 Bai T, Pang Y, Wang JC, *et al.* An optimized faster R-CNN method based on DRNet and RoI align for building detection in remote sensing images. Remote Sensing, 2020, 12(5): 762. [doi: 10.3390/rs12050762]

(校对责编: 牛欣悦)