



计算机工程与应用
Computer Engineering and Applications
ISSN 1002-8331, CN 11-2127/TP

《计算机工程与应用》网络首发论文

题目: 改进 Faster RCNN 的瓷砖表面瑕疵检测研究
作者: 赵楚, 段先华, 苏俊楷
网络首发日期: 2022-05-24
引用格式: 赵楚, 段先华, 苏俊楷. 改进 Faster RCNN 的瓷砖表面瑕疵检测研究[J/OL]. 计算机工程与应用.
<https://kns.cnki.net/kcms/detail/11.2127.tp.20220524.1039.009.html>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

改进 Faster RCNN 的瓷砖表面瑕疵检测研究

赵楚, 段先华, 苏俊楷

江苏科技大学 计算机学院, 江苏 镇江 212100

摘要: 针对瓷砖表面瑕疵中存在极小瑕疵目标, 瑕疵形态差异较大, 易出现漏检、准确率低等问题, 本文提出了一种改进 Faster RCNN 的瓷砖表面瑕疵检测算法。首先在 Faster RCNN 的特征提取网络 resnet101 的后三个阶段引入可变形卷积, 自适应地学习瑕疵特征。其次, 优化区域建议网络, 通过对瓷砖数据集的分析, 改进锚点生成参数, 使得生成的锚框更加契合目标尺度, 定位更加准确。最后, 优化损失函数, 引入 Rank & Sort Loss, 减少超参数数量, 提高模型性能, 使其对训练中类别不平衡问题更加鲁棒。实验结果表明, 改进后算法的 mAP 为 76.3%, 比原始 Faster RCNN 算法提高了 17.9%, 可以更好地检测小目标瑕疵, 满足瓷砖表面瑕疵检测的要求。

关键词: 目标检测; 瓷砖表面瑕疵; Faster RCNN; Rank & Sort Loss; 可变形卷积

文献标志码: A **中图分类号:** TP391 **doi:** 10.3778/j.issn.1002-8331.2203-0414

Research On Ceramic Tile Surface Defect Detection By Improved Faster RCNN

ZHAO Chu, DUAN Xianhua, SU Junkai

School of Computer Science, Jiangsu University of Science and Technology, Zhenjiang, Jiangsu 212100, China

Abstract: Aiming at the problems of minimal defect target, large difference of defect shape, easy missing detection and low accuracy in ceramic tile surface defects, an improved ceramic tile surface defect detection algorithm based on Faster RCNN is proposed. Firstly, based on the original Faster RCNN, resnet101 is selected as the feature extraction network, and deformable convolution networks is introduced in the last three stages of resnet101 to adaptively learn the defect features. Secondly, the regional proposals network is optimized, and the anchor generation parameters are improved through the analysis of ceramic tile data set, so that the generated anchors are more consistent with the target scale and the positioning is more accurate. Finally, the loss function is optimized and Rank & Sort loss is introduced to reduce the number of super parameters and improve the performance of the model, making it more robust to the class imbalance problem in training. Experimental results show that the average detection accuracy of the improved Faster RCNN is 76.3%, which is 17.9% higher than that of Faster RCNN. It can better detect small target defects and meet the requirements of ceramic tile surface defect detection.

Key words: target detection; ceramic tile surface defect; Faster RCNN; Rank & Sort Loss; deformable convolution networks

瓷砖作为建筑行业的重要装饰材料之一, 在我们的日常生活中有着广泛的应用^[1]。但瓷砖制作工艺复

基金项目: 国家自然科学基金(61806087); 江苏省研究生科研与实践创新计划项目 (KYCX21_3489)。

作者简介: 赵楚(1996—), 女, 硕士研究生, 研究方向为计算机视觉与图像处理, E-mail: 17865311776@163.com; 段先华(1965—), 男, 博士, 教授, 研究方向为计算机视觉与图像处理等; 苏俊楷(1997—), 男, 硕士研究生, 研究方向为计算机视觉与图像处理。

杂,在生产过程中表面难免会有各种瑕疵,例如粉团、角裂、缺角等。严重的瑕疵会影响瓷砖产品的质量和美观程度,从而导致相关企业的巨大损失。传统的人工检测方法^[2]依靠受检人员的主观经验检测瑕疵缺陷,检测效率不高,耗费人力物力,对人工依赖大,并且容易造成漏检错检,因此开发出一种快速、高精度的检测机制,对解决陶瓷产业在工业制造场景下的瓶颈问题具有重要意义。

目前常用的表面缺陷检测方法主要分为两种:一是传统的基于机器视觉的表面缺陷检测方法^[3],例如,李光亚等人^[4]提出基于 Gabor 变换与区域生长的瓷砖表面缺陷检测方法,采用双向积分投影实现缺陷区域标注。李军华等人^[5]将多特征融合算法应用于瓷砖缺陷检测,使用 SVM 分类器对瓷砖缺陷进行分类。虽然这类方法可以有效地降低对人工的依赖,提高检测效率,但是机器视觉技术通常采用常规的图像处理算法和人工设计特征加分类器方式,应用成本高,场景适用性低,难以应用到真实的工业环境中。二是基于深度学习的表面缺陷检测方法,其主要分为以 SSD^[6]、YOLO^[7, 8, 9, 10]系列为代表的 One-Stage 算法和以 Fast RCNN^[11]、Faster RCNN^[12]、Mask RCNN^[13]为代表的 Two-Stage 算法。一般情况下,Two-Stage 算法在准确度和小目标检测上有较高水平,One-Stage 算法在速度上有一定优势。例如,陈学仕等人^[14]提出的基于 Faster RCNN 的印刷电路板瑕疵检测算法,将 smoothL1 损失函数替换为 DIoU 损失,加速了模型的收敛,但是没有解决样本不平衡问题。程婧怡等人^[15]改进 YOLOv3 的网络结构,使用 K-Means++ 聚类,检测金属表面缺陷,虽然速度较快,但检测精度不高。安萌等人^[16]引入特征金字塔增加浅层特征,用 Soft-NMS 算法代替非极大值抑制算法,改进 Faster RCNN 模型检测面料瑕疵,在一定程度上可以减少漏检情况。随着计算机视觉技术的发展,深度学习的方法在工业产品表面缺陷检测中的应用也越来越广泛,但是对于瓷砖表面瑕疵检测而言,现阶段大都是采用传统的机器视觉方法,很少有研究使用深度学习模型来进行瓷砖表面瑕疵检

测,因此,本文选用在工业界应用较为广泛的 Faster RCNN 作为初始模型来检测瓷砖表面瑕疵。目前,现有的基于 Faster RCNN 的改进方法,主要是在网络层上进行改进,例如引入特征金字塔结构提高特征提取能力,改进损失函数帮助模型收敛等。这类改进方法确实对提高模型的检测效果有一定的帮助,但是在瓷砖瑕疵检测中,由于数据集目标的特点,图像分辨率大,瑕疵多变,同时存在 8 个像素左右的极小目标瑕疵和狭长目标瑕疵,从而容易出现漏检、误检问题,所以需要对 Faster RCNN 模型进行针对性改进。

针对上述瓷砖表面瑕疵检测中的问题,本文提出了一种改进的 Faster RCNN 瓷砖表面瑕疵检测方法。主要从以下三个方面进行改进:第一,改进网络结构,引入可变形卷积,使模型可以学习到瑕疵的更多形态,解决瓷砖表面瑕疵多变的问题。第二,优化区域建议网络,改进锚点生成参数,使生成的锚框更加契合瑕疵目标尺度,解决狭长类目标瑕疵、较小目标瑕疵容易漏检以及原始锚框定位不够准确的问题。第三,优化损失函数,引入一种基于排序的损失函数,Rank & Sort(RS) Loss^[17],简化模型训练的复杂性,减少参数数量,由于 RS Loss 不仅实现了正负样本之间的排序,还可以在正样本间进行排序,使正样本在训练时有不同的优先级,因此不需要添加任何样本均衡策略就可以很好地解决类别不平衡问题。

1 Faster RCNN 相关介绍

Faster RCNN 的基本结构如图 1 所示,首先模型通过一系列卷积层例如 VGG、ResNet 等特征提取网络对输入图像进行特征提取并生成特征图,其次将特征图输入到区域建议网络中(Region Proposal Networks, RPN)以生成候选区域。在 RPN 中,通过一组 3x3 的卷积核遍历输入的特征图,在特征图上的每个位置采用 k 个不同尺度的锚框,通过全连接层对锚框进行二分类用来判断该锚框属于目标的概率,同时利用边框回归来修正锚框,使其更加符合目标尺度。然后,将 RPN 层生成的候选区域和卷积层生成的最后一层特征图一起输入到感兴趣区域池化层(ROI Pooling),将候选区域特征图规格化为固定大小。最后,将这些候选区域特征图再次通过全连接层实现类别分类

以及边框回归获得更加准确的边框预测位置。

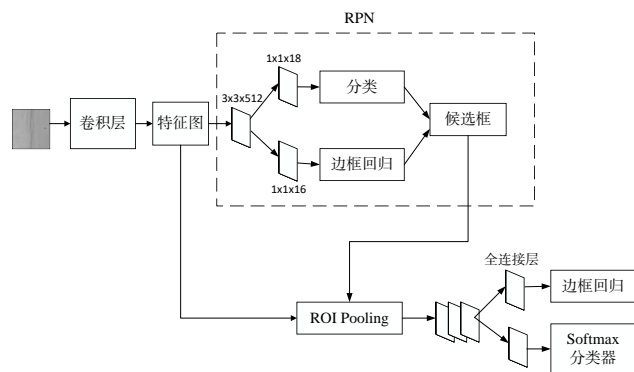


图1 Faster RCNN 基本结构

Fig.1 Basic structure of Faster RCNN

2 改进的 Faster RCNN 网络

2.1 改进网络结构

本文模型采用目前性能较好的 ResNet101^[18]+FPN^[19]作为基准 baseline。ResNet101 网络由 100 个卷积层和一个全连接层组成,通过卷积步长改变图像尺度,如图 2 所示,分为五个阶段,每经过一个阶段图片下采样 2 倍,第一个阶段由一个 7x7 的卷积构成,其余四个阶段分别引入了 3、4、23、3 个以 1x1、3x3 和 1x1 的卷积方式组成的残差结构,并在残差块中对卷积相关参数进行设置,避免了深层网络的梯度消失和退化问题。特征金字塔网络 (Feature Pyramid Networks, FPN) 将浅层特征和深层特征进行融和,增加了小目标的映射分辨率,可以有效改善小目标的检测效果。虽然 ResNet101+FPN 可以有效提高小目标检测的准确度,但是由于本文所研究的瓷砖表面瑕疵形状往往很不规则,存在较多方向多变的细长型瑕疵目标,标准的卷积核有着固定的尺寸和大小,只能在特征图的固定位置提取瑕疵特征,对未知的变化适应性差,泛化能力不强,从而会提取到很多无用信息,因此本文引入可变形卷积^[20] (Deformable Convolution Networks, DCN) 来适应不同形状的瑕疵。

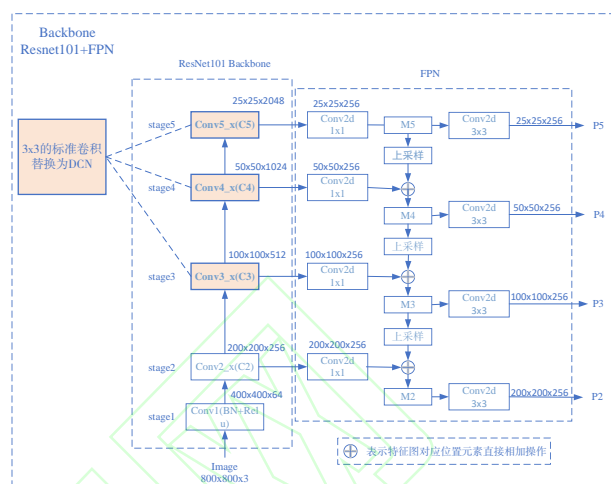


图2 改进后的特征提取网络结构图

Fig.2 Improved feature extraction network structure diagram

如图 2 所示,本文具体的改进方法是将可变形卷积结构嵌入到 ResNet101 中,将 ResNet101 中的后三个阶段即 stage3~stage5 中所有的 3x3 传统卷积替换成可变形卷积。

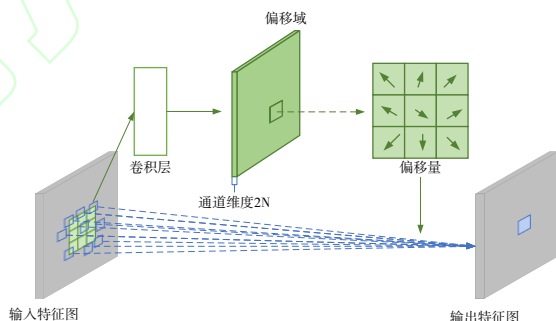


图3 可变形卷积实现过程

Fig.3 Implementation process of deformable convolution

可变形卷积的实现过程如图 3 所示,可变形卷积是在原有卷积上附加一层 3x3 卷积以获得 N 个 x, y 方向上的二维偏移量,其中 N 为采样点个数,然后将偏移量作用在原有的卷积核上来学习瑕疵的特征信息。由于偏移量的加入,采样点可以自由变换,因此在特征提取的过程中可以根据瑕疵目标的形态自适应的调整卷积核的形状,从而更好的学习瑕疵特征。

传统二维卷积首先利用固定尺寸卷积核 R 对特征图 x 进行采样,然后将各采样点进行加权求和。例如, $R = \{(-1, -1), (-1, 0), \dots, (0, 1), (1, 1)\}$, 其卷积过程如公式 1 所示:

$$y(p_0) = \sum_{p_n \in R} w(p_n) x(p_0 + p_n) \quad (1)$$

式中: p_n 为 R 中所列位置的枚举值, p_0 为输出特征图上的位置, $w(*)$ 为采样点权重。

可变形卷积是在同样的采样位置处增加了偏移量 Δp_n (第 n 个采样点的偏移), 此时, 特征图采样位置变为 $p_n + \Delta p_n$, 可变形卷积过程如公式 2 所示:

$$y(p_0) = \sum_{p_n \in R} w(p_n) x(p_0 + p_n + \Delta p_n) \quad (2)$$

2.2 区域建议网络优化

锚框机制是区域建议网络 RPN 的核心, 锚框的大小和比例对候选区域生成部分有着十分重要的影响, 合适的锚框可以更多地检测出待测目标, 若锚框大小和目标大小相差较大, 会使生成的候选区域不准确, 从而严重影响模型性能和检测效果, 因此合适的锚框参数对模型的检测效果至关重要。原始 Faster RCNN 模型中以 $[0.5, 1, 2]$ 三种长宽比和 $[8, 16, 32]$ 三种缩放比例组合生成如图 4 所示的一组锚框, 即一组由 9 个 anchor boxes 组成。

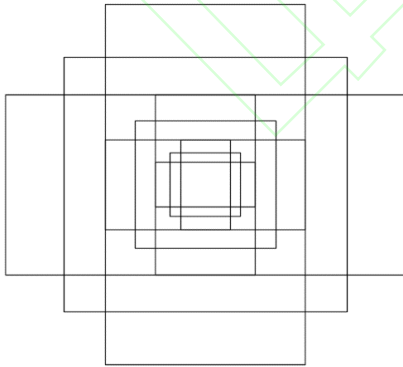


图 4 原始锚框

Fig.4 Original anchor boxes

本文对瓷砖瑕疵数据集进行分析, 瓷砖瑕疵数据集中瑕疵目标的长宽比以及面积分布如图 5、图 6 所示。根据数据统计可得, 瓷砖表面瑕疵数据集中瑕疵目标的长宽比大都集中在 1.0 左右, 最大的长宽比达到 19.63, 同时也有一部分瑕疵目标长宽比位于 3 到

8 之间。数据集中存在 8 个像素左右的小目标, 以及长宽比差异较大的瑕疵目标, 由此可得原始生成锚框的尺度不能很好地匹配瓷砖表面瑕疵数据集中所检测的目标尺度, 因此需要对锚框生成方式进行优化, 本文在原始三种长宽比的基础上增加了 0.2, 5 两种长宽比, 剔除了 16, 32 两种缩放比例, 增加了缩放比例 2, 最终每个像素点以 $[0.2, 0.5, 1.0, 2.0, 5]$ 五种长宽比和 $[2, 8]$ 两种缩放比例生成如图 7 所示的 10 个 anchor boxes。实验表明, 本文设置的锚框参数可以更好地匹配瓷砖瑕疵目标, 检测狭长瑕疵目标和较小瑕疵目标, 有效地提升模型性能。

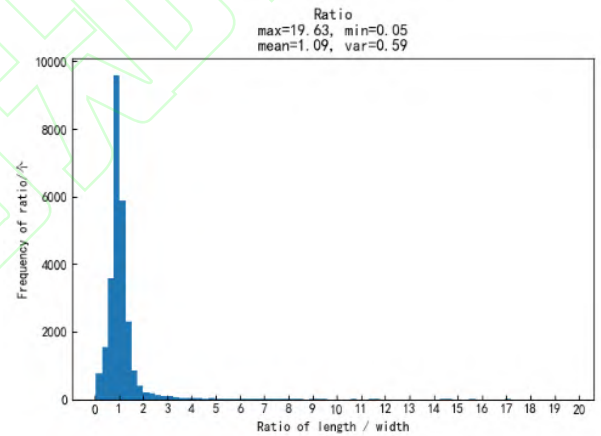


图 5 瑕疵长宽比分布

Fig.5 Defect aspect ratio distrib

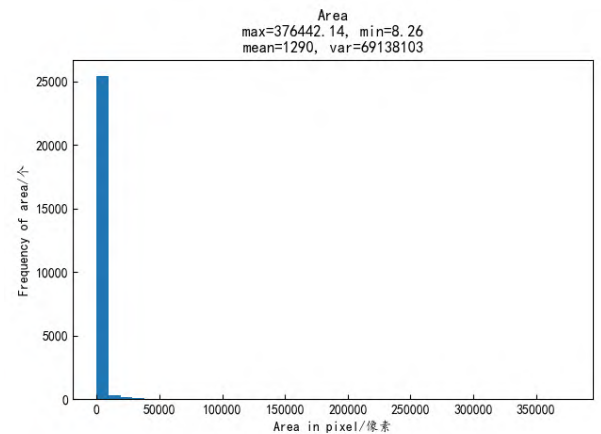


图 6 瑕疵面积分布

Fig.6 Defect area distribution

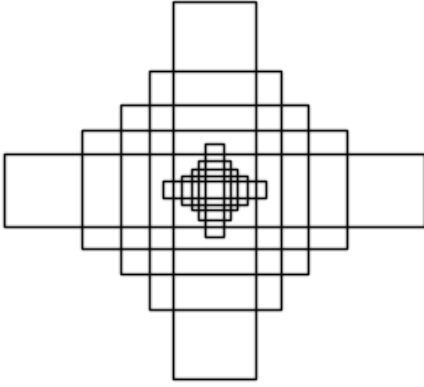


图7 改进后的锚框

Fig.7 Improved anchor boxes

2.3 损失函数优化

Faster RCNN 模型中包含了 classification, box regression 两个子任务, 其总损失为分类损失和回归损失两个部分的加权和。由于多任务训练中损失部分的超参量和任务数量成正比, 因此在 Faster RCNN 模型的训练中损失部分会产生较多超参数, 从而需要不断的调参来获得较好的模型性能, 但是调参的过程浪费了大量的时间和资源。在本文瓷砖瑕疵数据集中, 浅色块瑕疵、边异常、光圈瑕疵这三类的的数据量相对较少, 从而导致数据分布不平衡现象, 致使模型检测性能不佳。针对上述问题, 本文对损失函数部分进行优化改进, 引入了 2021 年 K.Oksuz 等人提出的 Rank & Sort (RS) Loss^[17], 以便简化模型训练的复杂性, 使模型可以达到更好的性能。

RS Loss 总共由 Ranking 和 Sorting 两部分组成。Ranking 指的是根据分类得分区分出正负样本, 使所有正样本都排在负样本之上, Sorting 指的是根据预测框和真实框之间的交并比 IoU 得到 0-1 范围内的数值对正样本进行降序排序, 使得不同正样本在训练时有不同的优先级。RS Loss 不仅将正负样本进行了排序, 还在正样本间进行了排序, 这一特性使得在训练过程中可以有效解决极端的类别不平衡问题, 而不需要加入任何样本均衡策略。

RS Loss 的计算过程主要分为以下几步:

(1) 先根据 Faster RCNN 模型边界框回归输出的预测框, 求出预测框和真实框之间的交并比 IoU 记为 y , 再将 Faster RCNN 模型分类回归输出的分类得分 s 和 y 共同作为 RS Loss 的输入。

(2) 根据以下公式依次计算出当前 ranking 误差

$\ell_R(i)$, 当前 sorting 误差 $\ell_S(i)$, 目标 ranking 误差 $\ell_R^*(i)$, 目标 sorting 误差 $\ell_S^*(i)$ (注: 当所有正样本排在所有负样本之上时, $N_{FP}(i) = 0$, 所以 $\ell_R^*(i) = 0$)。

(3) 将当前 ranking 误差与目标 sorting 误差相减即得到 ranking Loss, $\ell_R(i) - \ell_R^*(i)$, 将当前 sorting 误差与目标 sorting 误差相减即得到 sorting Loss, $\ell_S(i) - \ell_S^*(i)$, 将两者相加即得到总损失, $(\ell_R(i) + \ell_S(i)) - (\ell_R^*(i) + \ell_S^*(i))$, 其中 $\ell_R(i) + \ell_S(i)$ 为当前 RS 误差 $\ell_{RS}(i)$, $\ell_R^*(i) + \ell_S^*(i)$ 为目标 RS 误差 $\ell_{RS}^*(i)$ 。

(4) 将 (3) 中求得的总损失除以正样本总数即可得到 RS Loss, 其公式定义如公式 8 所示。

$$\ell_{RS}(i) := \underbrace{\frac{N_{FP}(i)}{\text{rank}(i)}}_{\ell_R(i)} + \underbrace{\frac{\sum_{j \in pos} H(x_{ij})(1 - y_j)}{\text{rank}^+(i)}}_{\ell_S(i)} \quad (3)$$

$$\ell_{RS}^*(i) = \ell_R^*(i) + \underbrace{\frac{\sum_{j \in pos} H(x_{ij})[y_j \geq y_i](1 - y_j)}{\sum_{j \in pos} H(x_{ij})[y_j \geq y_i]}}_{\ell_S^*(i)} \quad (4)$$

$$N_{FP}(i) = \sum_{j \in neg} H(x_{ij}) \quad (5)$$

$$\text{rank}(i) = \sum_{j \in pos \cup neg} H(x_{ij}) \quad (6)$$

$$\text{rank}^+(i) = \sum_{j \in pos} H(x_{ij}) \quad (7)$$

式中: i, j 为样本编号; pos 为正样本的集合; neg 为负样本的集合; $N_{FP}(i)$ 为所有负样本中大于等于该正样本分类得分的数量; $\text{rank}(i)$ 为所有正样本和负样本中大于等于该正样本分类得分的数量; $\text{rank}^+(i)$ 为所有正样本中大于等于该正样本分类得分的数量; x_{ij} 为分类得分 s_i 和 s_j 之差即 $x_{ij} = s_j - s_i$; $H(x)$ 为单元阶跃函数, 如果 $x \geq 0$ 则 $H(x) = 1$, 反之, $H(x) = 0$; y 为预测框和真实框之间的交并比 IoU 的值; $\ell_R(i)$ 为当前 ranking 误差; $\ell_S(i)$ 为当前 sorting 误差; $\ell_R^*(i)$ 为目标 ranking 误差, 当每个正样本排在所有负样本之上时, $N_{FP}(i) = 0$, 所以目标 ranking 误

差 $\ell_R^*(i)=0$; $\ell_S^*(i)$ 为目标 sorting 误差; $[y_j \geq y_i]$ 表示当 $y_j \geq y_i$ 为真时, $[y_j \geq y_i]=1$, 当 $y_j \geq y_i$ 为假时, $[y_j \geq y_i]=0$ 。

$$\mathcal{L}_{RS} := \frac{1}{|pos|} \sum_{i \in pos} (\ell_{RS}(i) - \ell_{RS}^*(i)) \quad (8)$$

式中: i 为样本编号; pos 为正样本的集合; $\ell_{RS}(i)$ 为当前 ranking 误差和当前 sorting 误差的总和, 即当前 RS 误差; $\ell_{RS}^*(i)$ 为目标 ranking 误差和目标 sorting 误差的总和, 即目标 RS 误差。

本文中的具体改进方法为用 RS 损失代替分类损失中的交叉熵损失, 回归损失采用 GIOU 损失, 改进后模型回归损失的加权参数为 RS 损失除以回归损失。采用这种加权参数设置方法有效的减少了超参数的数量, 在模型训练过程中不需要反复调节超参数, 只需要调整学习率就可以提高模型性能, 避免了复杂的调参过程和一个损失占据主导地位的情况。

3 实验分析

3.1 实验平台搭建

实验硬件平台为 Unbutu 操作系统。GPU 为 NVIDIA GeForce RTX 2080, 12G 显存。利用 CUDA10.1 和 CUDNN, 加快 GPU 运算。采用基于 pyTorch 实现的 MMDetection 深度学习目标检测框架, 代码运行环境为 Python3.7, 学习率为 0.0012, 由于本文的缺陷检测任务需要在测试集上的泛化能力较好, 自适应优化器在测试集上的泛化能力远不如带动量的随机梯度优化器, 因此本文优化器采用带动量的随机梯度优化器 (SGD), 动量为 0.9, 权值衰减系数设置为 0.0001, batch_sizes 设置为 2, 每个实验均训练 36 个 epoch。

3.2 实验数据集

本实验数据集采用广东工业智造创新大赛提供的瓷砖表面瑕疵数据集。原始数据集共有约 12000 张图片, 图片为白板瓷砖, 花色简单, 原始分辨率为 8192*6000, 数据集类别分为: 边异常、角异常、白色点瑕疵、浅色块瑕疵、深色点瑕疵、光圈瑕疵六类, 如图 8 所示。

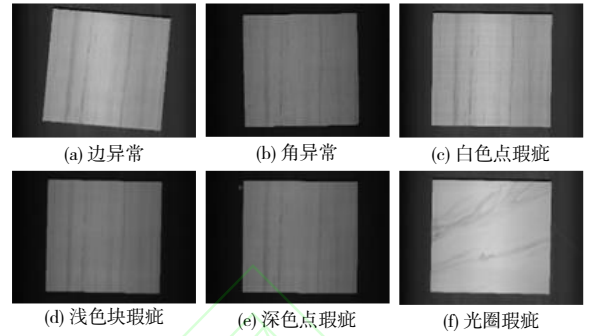


图 8 原始瓷砖缺陷类别图

Fig.8 Defect category of original ceramic

由于原始数据集图片分辨率过大, 同时存在 8 个像素左右的小目标, 导致训练时间长, 检测难度大, 因此为了增强模型鲁棒性和训练效果, 本实验对原始数据集进行了切片处理, 首先检测图像瓷砖边缘, 切掉一些黑边, 对切边后的图像进行滑窗切图, 步长为 [640,640], 处理后的图片大小为 800*800, 一共 20724 张图片, 以 8:2 的比例随机划分为训练集和测试集, 即训练集 16580 张, 测试集 4144 张。处理后的类别图样例如图 9 所示。从图 9 中可以看出, 处理后的图片瑕疵目标更加凸显。

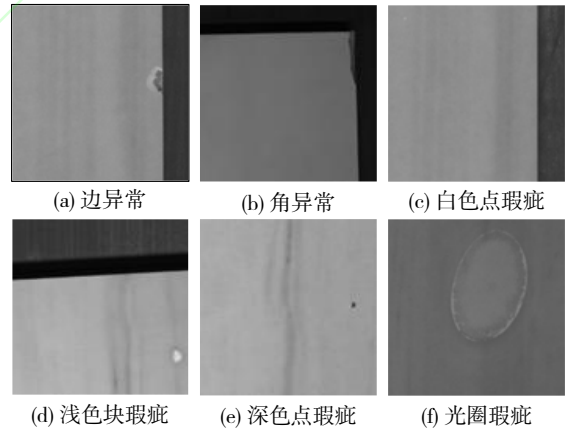


图 9 处理后瓷砖缺陷类别图

Fig.9 Ceramic defect category diagram after treatment

3.3 模型评估方法

本文采用目标检测算法中常用性能评价指标平均精度 (Average Precision, AP)、平均精度均值 (mean Average Precision, mAP) 对训练好的模型进行评估, 并采用同一硬件环境下检测速率帧每秒 (frame per second, FPS) 来衡量模型检测速度。

AP 表示在不同召回率 (Recall, R) 下精确率 (Precision, P) 的平均值即 PR 曲线与坐标轴围成的面积, mAP 表示每个类别的 AP 的算术平均值。P、

R 、 AP 、 mAP 的计算公式如下所示:

$$P = \frac{TP}{TP + FP} \quad (9)$$

$$R = \frac{TP}{TP + FN} \quad (10)$$

$$AP = \int_0^1 P(R) dR \quad (11)$$

$$mAP = \frac{\sum_{n=0}^N AP(n)}{N} \quad (12)$$

其中, TP 为正样本被正确识别为正样本; FP 为负样本被错误识别为正样本; FN 为正样本被错误识别为负样本; R 为召回率; $P(R)$ 为 PR 曲线; n 为类别编号; N 为检测类别数。

3.4 实验结果与分析

3.4.1 消融实验

为论证上述所提算法的检测性能, 探究各个改进方法的有效性, 本文将使用上述的改进方法在瓷砖表面瑕疵数据集上对模型进行训练和测试, 每组实验使用的超参数和训练技巧相同, 实验结果如表 1 所示, 其中 \checkmark 表示引入了该模块。

表 1 不同改进方法实验结果

组别	改进 锚框	RS LOSS	DCN	mAP(%)	FPS(f/s)
1				58.4	13.1
2	\checkmark			66.8	13.0
3		\checkmark		62.1	15.0
4			\checkmark	61.0	13.4
5	\checkmark	\checkmark		68.1	15.2
6	\checkmark	\checkmark	\checkmark	76.3	17.7

从表 1 中可以看出, 采用改进的锚点生成参数之后, 模型的 mAP 为 66.8%, 比原始模型高出 8.4 个百分点, 而模型的检测速度几乎没有变化, 表明本文所提出的锚点生成参数比原始锚点生成参数更加契合瑕疵目标尺度, 定位更加准确, 可以有效降低漏检误检情况, 提高模型精度。引入 RS Loss 后, 模型精度较原始模型提高了 3.7 个百分点, 检测速度也有所提升, 表明 RS Loss 可以有效提高模型性能。加入可变形卷积 DCN 后, 虽然模型在速度上的提升不太明显, 但在精度上较原始模型提高了 2.6 个百分点, 表明可变形卷积的加入对模

型检测速度几乎没有什么影响, 但是可以使模型精度得到有效提升, 使卷积学习到更多瑕疵特征。总的来看, 本文提出的每个改进方法都有效地提升了模型精度, 而将各个改进方法累加后, 模型的精度也在逐渐提高, 例如, 同时引入改进锚框和 RS Loss 之后, 模型的 mAP 达到 68.1%, 比原始模型提高了 9.7 个百分点, 高于单独引入这两个模块的 mAP 值, FPS 也提高了 2.1 f/s, 表明将各个改进方法累加算法性能更好。最终, 将本文所提出的三个改进方法都引入后, 模型得到最优结果, 其 mAP 值可以达到 76.3%, 较原始模型提高了 17.9 个百分点, FPS 为 17.7 f/s, 模型的检测速度也有较大提升, 表明本文算法有效地提升了模型性能, 能够更为准确地检测瓷砖表面瑕疵。

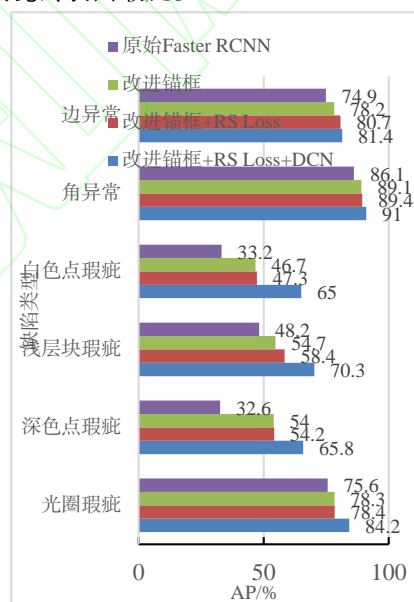


图 10 不同改进方法各瑕疵 AP 值对比

Fig.10 Comparison of AP values of defects in different improvement methods

图 10 为不同改进方法下各类型瑕疵在测试集上的平均精度 AP 值。从图 10 中可以看出, 各个改进方法对每个瑕疵类别的检测精度都有所提升, 尤其在小瑕疵目标类别上的检测精度提升更加显著。例如, 原始 Faster RCNN 模型对白色点瑕疵、浅色块瑕疵、深色点瑕疵这三种小瑕疵类别的检测效果较差, 只有 33.2%、48.2%、32.6%, 而加入本文改进方法后, 最终本文算法对这三种类别的检测精度分别为 65%、70.3%、65.8%, 比原始模型分别提高了 31.8%、22.1%、33.2%。

为了更加清晰地看出本文改进后的算法对小目标瑕疵的检测效果,本文从原始 Faster RCNN 模型和改进后的模型在测试集上的可视化检测图中各选取了四张可视化图作对比如图 11 和图 12 所示,从图中可以直观地看出,改进后的模型可以更多更准确的检测出小目标瑕疵,实验结果表明,本文的改进方法对小目标瑕疵的检测具有一定的效果,可以有效地降低瑕疵漏检率、错检率。

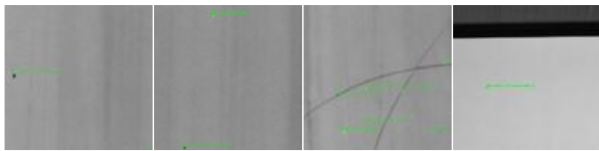


图 11 原始 Faster RCNN 检测效果

Fig.11 Original Faster RCNN detection effect

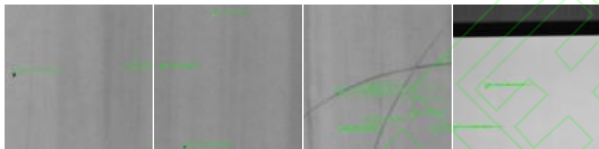


图 12 改进后的 Faster RCNN 检测效果

Fig.12 Improved Faster RCNN detection effect

3.4.2 复杂度分析

本文采用计算量(Floating point Operations, FLOPs)和参数量(Parameters, Params)来衡量模型的复杂度, FLOPs、Params 越高,复杂度越高。不同方法的参数量和计算量对比如表 2 所示。

表 2 不同方法参数量与计算量对比

Table 2 Comparison of parameter quantity and calculation quantity of different methods

	参数量 (M)	计算量 (GFLOPs)
原始 Faster RCNN	60.61	284.07
改进锚框	60.15	283.53
改进锚框+RS Loss	41.16	207.45
改进锚框+RS Loss+DCN	44.64	192.22

从表 2 中可以看出,改进锚框后的模型参数量和计算量比原始 Faster RCNN 模型变化不大,说明改进锚框的方法对模型复杂度影响不大。引入 RS Loss 后

模型的参数量比原始模型降低了 32.09%,计算量降低了 26.97%,说明 RS Loss 的引入可以大大减少参数量,有效降低模型的复杂度。在引入改进锚框和 RS Loss 之后,再引入 DCN,模型的参数量和计算量较没引入 DCN 之前,变化不大,说明引入 DCN 对模型复杂度影响较小。最终,本文改进后的模型参数量和计算量比原始模型分别降低 26.34%、32.33%,说明本文算法和原始算法相比,复杂度更低,对内存硬件的要求更低,更利于实际应用,本文算法满足在工业任务中的部署需求。

3.4.3 其他检测算法对比

为了验证本文中的改进方法对瓷砖表面瑕疵的检测性能,本文将改进后的算法与 Faster RCNN、SSD300、YOLOv3、YOLOv4、Cascade^[21]模型进行实验对比,实验采用 AP、mAP 作为评价指标,训练时 IoU 大于等于 0.5 的被视为正样本,小于 0.5 的被视为负样本,实验结果如表 3 所示。其中 YOLOv3、YOLOv4 是通过在 YOLO 上进行一系列改进得到的一阶段模型, Cascade 为二阶段模型。从表 3 中可以看出,本文算法在白色点瑕疵、浅色块瑕疵、深色点瑕疵上的 AP 值均优于上述模型,在其他类别上的检测效果也较好,这是因为本文算法充分考虑到了瓷砖瑕疵数据集中目标的特点,对数据集中存在的图片分辨率大,形态多变,小目标问题进行了针对性改进。同时,本文算法的 mAP 值为 76.3%,均优于原始 Faster RCNN 模型、SSD 模型、YOLOv3 模型、YOLOv4 模型,满足对瓷砖表面瑕疵检测精度更高的要求。

表 3 不同算法的对比

Table 3 Comparison of different algorithms

	Faster RCNN	SSD 300	YOLO v3	YOLO v4	Cascade	本文算法
边异常	74.9	47.6	47.5	82.0	74.6	81.4
角异常	86.1	63.6	59.8	85.0	84.0	91.0
白色点瑕疵	33.2	29.7	32.4	47.0	38.0	65.0
浅色块瑕疵	48.2	17.6	39.6	58.0	49.5	70.3
深色点瑕疵	32.6	34.3	35.8	39.0	37.6	65.8
光圈瑕疵	75.6	46.2	53.1	77.0	76.3	84.2
mAP (%)	58.4	39.8	44.7	64.7	60.0	76.3

3.4.4 模型普适性

为验证本文模型在缺陷检测任务中的普适性,将本文改进后的模型在金属表面缺陷数据集 NEU-DET 上进行训练测试,并与其他检测算法进行对比,采用 mAP 进行评估,实验结果如表 4 所示。其中,文献[13]和文献[14]分别为基于 Faster RCNN 和基于 YOLOv3 改进的金属缺陷检测算法,其实验数据集都为 NEU-DET 数据集。从表 4 中可以看出,本文改进后的模型在金属表面缺陷检测中的 mAP 为 78.5%,均优于文献[13]、文献[14]。实验结果表明,本文模型可以满足金属表面缺陷检测的需要,同样适应于金属表面缺陷检测的任务,在其他缺陷检测任务中具有良好的普适性。

表 4 NEU-DET 上实验对比

Table 4 Experimental comparison on NEU-DET

网络模型	mAP (%)	发表年份
文献[13]	72.4	2021
文献[14]	67.6	2021
本文算法	78.5	

4 结语

本文检测的最大难点在于较大尺寸原图,存在极小瑕疵目标的问题。针对这一问题,本文对 Faster RCNN 模型进行了一系列改进。首先对原始数据集进行了切片处理,有效的降低了原图尺度,并改进了锚框生成方式,增加了较小的锚框尺度,使生成的锚框大小可以更好地契合本文检测的瑕疵目标,从而降低了小瑕疵漏检、错检的概率,提高了模型的检测精度。可变形卷积的引入使网络可以自适应地学习瑕疵的特征。采用 RS Loss,减少了模型超参数的数量,提高了模型性能,使模型对于类别不平衡的问题更加鲁棒。实验结果表明,本文提出的改进方法可以有效地提高模型对小目标瑕疵的检测效果,检测出瓷砖表面瑕疵。但是本实验主要研究的是花色较为简单的瓷砖表面瑕疵,未来将对复杂花色的瓷砖进行研究,并对模型进行轻量化处理,使模型可以更好地应用于瓷砖表面瑕疵检测任务中。

参考文献:

[1] 李治. 陶瓷砖质量发展概述[C]//第四届建筑卫生陶瓷质量大会暨中国硅酸盐学会建筑卫生陶瓷专业委员会

2019 学术年会论文集, 2019: 15-16.

- LI Z. Overview of ceramic tile quality development[C]// Proceedings of the 4th architectural and sanitary ceramics quality conference and 2019 academic annual meeting of architectural and sanitary ceramics Professional Committee of China Silicate Society, 2019: 15-16.
- [2] 肖克来提. 表面缺陷检测应用研究综述[J]. 电子技术, 2020, 49(8): 189-191+194.
- Shockletti. Review on Application of Surface Defect Detection [J]. Electronic Technology, 2020, 49(8): 189-191+194.
- [3] 杨传礼, 张修庆. 基于机器视觉和深度学习的材料缺陷检测应用综述[J/OL]. 材料导报, 2022(16): 1-19.
- YANG C L, ZHANG X Q. Survey of material defect detection applications based on machine vision and deep learning[J/OL]. Materials Reports, 2022(16): 1-19.
- [4] 李光亚, 邹国锋, 傅桂霞. 基于 Gabor 变换与区域生长的瓷砖表面缺陷视觉检测方法研究[J]. 现代计算机, 2019(24): 37-42.
- LI G Y, ZOU G G, FU G X. Research on visual inspection method of ceramic tile surface defects based on Gabor transform and region growth [J]. Modern Computer, 2019(24): 37-42.
- [5] 李军华, 权小霞, 汪宇玲. 多特征融合的瓷砖表面缺陷检测算法研究[J]. 计算机工程与应用, 2020, 56(15): 191-198.
- LI J H, QUAN X X, WANG Y L. Research on Defect Detection Algorithm of Ceramic Tile Surface with Multi-feature Fusion[J]. Computer engineering and application, 2020, 56(15): 191-198.
- [6] Liu W, Anguelov D and Erhan D. SSD: Single shot multi-box detector[C]//European conference on computer vision. Springer, Cham, 2016, 21-37.
- [7] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]// Computer Vision & Pattern Recognition. IEEE, 2016: 779-788.
- [8] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//IEEE Conference on Computer Vision & Pattern Recognition. [S.l.]:[s.n.], 2017: 6517-6525.
- [9] Redmon J, Farhadi A. YOLOv3: an incremental improvement[C]//IEEE Conference on Computer Vision & Pattern Recognition. [S.l.]:[s.n.], 2018: 2767-2773.
- [10] BOCHKOVSKIY A, Wang C Y, Liao H Y M. YOLOv4: Optimal Speed and Accuracy of Object Detection[C]// IEEE Conference on Computer Vision and Pattern Recognition. 2020. arXiv: 2004.10934.
- [11] Girshick R. Fast r-cnn[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2015: 1440-1448.
- [12] Ren S, He K and Girshick R. 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] HE K, GKIOXARI G, PIOTR D, et al. Mask RCNN[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017: 2961-2969.
- [14] 陈学仕, 苏通, 漆为民. 基于改进 Faster RCNN 的印刷

- 电路板瑕疵检测算法[J].江汉大学学报(自然科学版), 2022, 50(1): 87-96.
- CHEN X S, SU T, QI W M. Printed Circuit Board Defect Detection Algorithm Based on Improved Faster RCNN[J]. Jiangnan Univ.(Nat. Sci. Ed.), 2022, 50(1): 87-96.
- [15] 程婧怡, 段先华, 朱伟. 改进 YOLOv3 的金属表面缺陷检测研究[J].计算机工程与应用, 2021, 57(19): 252-258.
- CHENG J Y, DUAN X H, ZHU W. Study on Metal Surface Defect Detection by Improved YOLOv3[J]. Computer engineering and application, 2021, 57(19): 252-258.
- [16] 安萌, 郑颺默, 王诗宇. 一种改进 Faster R-CNN 的面料疵点检测方法[J].小型微型计算机系统, 2021, 42(5): 1029-1033.
- AN M, ZHENG L M, WANG S Y. Fabric Defect Detection Method Based on Improved Faster R-CNN[J]. Journal of Chinese Computer Systems, 2021, 42(5): 1029-1033.
- [17] K.Oksuz, B.C.Cam, E.Akbas, et al. Rank & Sort Loss for Object Detection and Instance Segmentation[C]//2021 IEEE/CVF Inter-national Conference on Computer Vision(ICCV), 2021: 2989-2998.
- [18] He K, Zhang X and Ren S. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2016: 770-778.
- [19] Lin T, Dollar P, Girshick R, et al. Feature pyramid networks for object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2017: 936-944.
- [20] Dai J F, Qi H Z, Xiong Y W, et al. Deformable Convolutional Networks[C]//Proceedings of 2017 IEEE International Conference on Computer Vision. 2017: 764-773.
- [21] CAI Z, VASCONCELOS N. Cascade R-CNN: Delving into high quality object detection[C]//Computer vision and pattern recognition, 2018:6154-6162.