

DOI: 10.19650/j.cnki.cjsi.2007109

基于改进 Faster RCNN 的铝材表面缺陷检测方法^{*}

向宽,李松松,栾明慧,杨莹,何慧敏

(大连海洋大学信息工程学院 大连 116023)

摘要: 针对传统检测算法对工业铝材表面缺陷识别率不高、对于小缺陷定位不准确等问题,提出改进的 Faster RCNN 深度学习网络对于铝材表面 10 种缺陷进行检测。首先,对数据进行增强后,在主干网络加入特征金字塔网络(FPN)结构以加强对小缺陷的特征提取能力,随后用感兴趣区域校准(ROI Align)算法来代替粗糙的感兴趣区域池化(ROI Pooling)算法,获得更准确的缺陷定位信息,最后加入 K-means 算法对缺陷数据进行聚类,得出更适应铝材缺陷的锚框。实验表明,改进后的网络对铝材表面缺陷检测的平均精度均值(mAP50)为 91.20%,比原始的 Faster RCNN 网络提高了 16%,并且对铝材小缺陷的检测能力也得到明显的提高。

关键词: 铝材检测;深度学习;Faster RCNN;特征金字塔网络;感兴趣区域校准;K-means

中图分类号: TP391.4 TH16 **文献标识码:** A **国家标准学科分类代码:** 520.2060

Aluminum product surface defect detection method based on improved Faster RCNN

Xiang Kuan, Li Songsong, Luan Minghui, Yang Ying, He Huimin

(School of Information Engineering, Dalian Ocean University, Dalian 116023, China)

Abstract: Aiming at the problems of low recognition rate of surface defects in industrial aluminum product and inaccurate location of small defects, etc. of the traditional detection algorithm, an improved deep learning network called Faster RCNN is proposed to detect 10 kinds of aluminum product surface defects. Firstly, after the data is enhanced, the feature pyramid network (FPN) structure is added to the backbone network to enhance the feature extraction ability for small defects, and then the ROI Align algorithm is used to replace ROI Pooling algorithm to obtain more accurate defect location information. Finally, the K-means algorithm is added to cluster the defect data to obtain an anchor more suitable for aluminum product defects. The experiment shows that the mean of the average precision (mAP50) of the improved network for aluminum product surface defect detection is 91.20%, which is 16% higher than that of the original Faster RCNN network, and the detection ability of aluminum product small defects is also improved obviously.

Keywords: aluminum product detection; deep learning; Faster RCNN; feature pyramid network; ROI Align; K-means

0 引言

我国是有色金属产能大国、产量大国、消费大国和出口大国^[1],铝材作为日常生活中不可或缺的一部分,更是与中国航空航天、高铁、飞机、汽车制造业等国家支柱性产业密不可分^[2]。因此铝材质量的检测具有重大意义。

目前针对铝板的表面缺陷识别研究较晚,对于金属(如钢材)的表面缺陷识别研究早已受到业界广泛的关注,美国学者 Pernkopf^[3]根据光学断层扫描的方法获取

钢锭表面数据范围,利用相关三维特性绘制出深度图像后送入贝叶斯网络进行特征提取与分类,对于分割表面的分类精度超过 98%。沈阳工业大学的杨理践等^[4]提出一种基于平衡电磁技术的钢板裂纹检测方法。验证了在交流励磁条件下 U 形传感器利用平衡电磁技术能够对钢板表面横纵向裂纹进行有效检测,能够有效区分裂纹类型的钢材缺陷。周鹏等^[5]提出一种基于剪切波和小波特征融合的金属表面缺陷识别方法,分别对 3 种金属的 6 种缺陷进行检测,并取得良好的效果,验证了该方法能够对不同种类的金属缺陷进行检测。哈尔滨工业大学的程

收稿日期:2020-11-12 Received Date: 2020-11-12

^{*} 基金项目:国家自然科学基金(51778104)、辽宁省渔业厅资助项目(201723)资助

星振^[6]依据缺陷无损检测的原理,仿真并搭建了金属材料缺陷多模态无损检测系统,该系统对铝材、钢板的表面和浅表面裂纹缺陷进行检测,能够对金属缺陷的种类、尺度、深度等信息进行全面精确检测。齐鲁工业大学的冯序文^[7]设计了一套铝板表面缺陷检测系统,结合中值滤波进行降噪,自适应阈值进行分割铝板缺陷等方法对缺陷图像进行预处理后,提取特征数据,分类器采用的是支持向量机(support vector machine, SVM),该方法不仅能识别常见的4种铝板表面缺陷并取得不错的结果,而且能识别出新的缺陷类型。

随着深度学习的蓬勃发展,结合深度学习的金属表面缺陷检测也得到了应用^[8]。如:Wang等^[9]设计了多层的卷积神经网络(convolutional neural network, CNN)对6种金属缺陷进行缺陷检测,使用滑窗法方法在原图进行特征采样,后对每一类图像采样的小图像块进行二分类,并与传统方式进行对比,识别效果优于传统方式。Mei等^[10]提出一种结合图像金字塔层次结构思想和卷积去噪自编码器(convolutional denoising auto-encoders, CDAE)对纹理图像缺陷的检测方法,该方法在布匹丝织物这种重复性背景纹理很强的图集上效果很好,相对于其他传统方式,检测效果更优,但在金属表面加工部件表面数据集效果不理想。

综上,目前对金属表面缺陷的研究主要包括基于单一机理的无损检测方法,传统机器视觉方法和深度学习方法。然而无损检测方法对检测人员的专业性要求较高,必须对检测试件的物理指标进行接收后处理,对检测材料以及检测的缺陷种类有局限性,传统机器视觉在需要人工进行特征的提取,利用图像分割技术对图像背景与检测缺陷进行分割,检测器的选取等繁琐步骤,使得检测周期过长,适应性较差。深度学习算法拥有强大的特征提取能力,能将缺陷特征提取,图像分割、缺陷分类在一个网络中完成,能够利用GPU^[11](graphics processing unit)强大的计算能力,模型统一化程度高,训练过程人工参与量小,从长远来看有取代传统检测方法的趋势^[12]。鉴于目前深度学习算法对于工业铝材表面缺陷检测研究较少,仅有模型对铝材表面缺陷识别效果不佳等现状,为提高我国对铝材的质量检测需求,本文提出改进的Faster RCNN(faster region-based convolutional neural networks)深度学习检测网络,在不过多损失检测速度的同时兼顾精度的提升,对铝材表面缺陷进行检测。

1 数据获取与分析

1.1 数据获取

基于深度学习的铝材表面缺陷识别方法的实验数据集选取自阿里云天池大数据创新大赛,均为实际生产中

有缺陷的铝型材监测影像数据,其中缺陷数据包含单缺陷和多缺陷数据,共计3 005张图片。缺陷图片采用矩形框进行标注,采用utf-8的编码格式,2 560×1 960的分辨率。缺陷数据提供的铝材表面缺陷共10类,如图1所示。

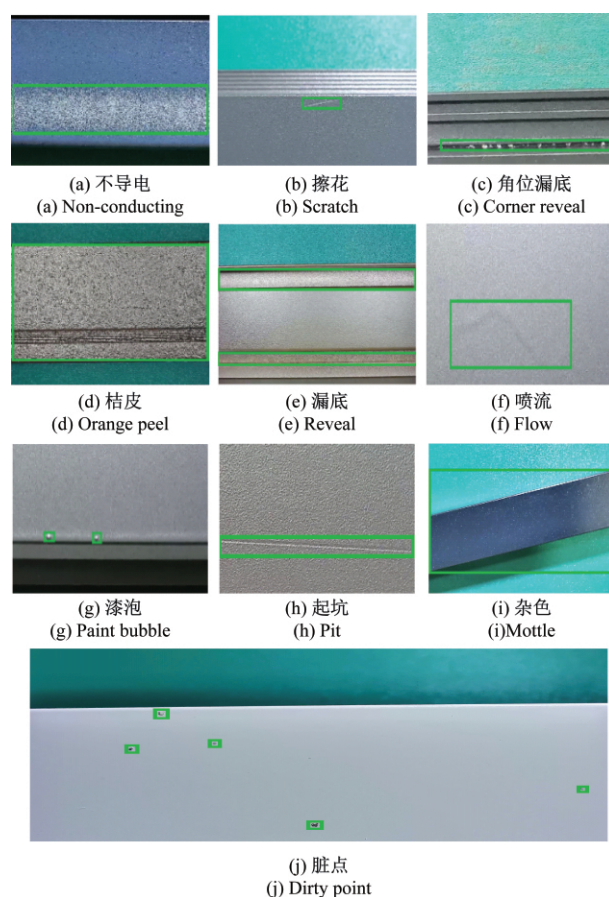


图1 铝材表面缺陷

Fig. 1 Aluminum product surface defect

1.2 数据分析

从1.1节可知,铝材表面缺陷图片分辨率高,缺陷的尺寸多样、缺陷宽高比及大小差异大。为更加了解数据特点,训练出合适的深度检测网络,本文对真实缺陷面积与图片面积占比进行统计,缺陷的面积占比如表1所示。

表1 缺陷面积占比统计

Table 1 Statistics of the proportion of defect area %

缺陷面积占比	百分比个数
<1	21.9
1~5	22.0
5~10	24.4
>10	31.7

由表 1 可知, 缺陷面积与图片面积之比 $< 1\%$ 的占缺陷数据总数的 21.9%, 而缺陷面积占比 $< 10\%$ 的缺陷占缺陷总数的 69% 左右, 可见对于获取的铝材数据中表面小缺陷及中小缺陷(如: 脏点、漆泡、喷流等)的占比巨大, 这就要求算法的识别能力更强。

2 改进 Faster RCNN 结构

2.1 Faster RCNN 网络结构

Faster RCNN^[13] 可以看做由 Fast RCNN^[14] 和区域建议网络(region proposal network, RPN) 两个网络组成, 主要包括以下 4 个部分, Faster RCNN 基本结构如图 2 所示。

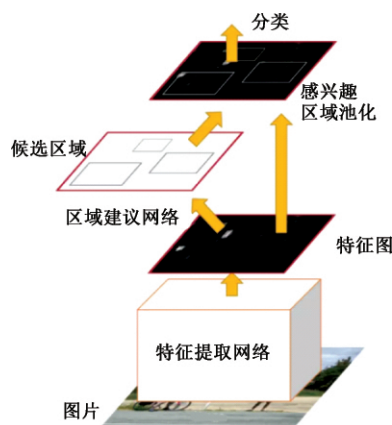


图 2 Faster RCNN 基本结构

Fig. 2 Faster RCNN basic structure

1) 特征提取网络: 作为一种 CNN 网络目标检测方法, Faster RCNN 首先使用一个卷积神经网络(如: ZFNET^[15]、VGG16^[16]、ResNet^[17]等)提取图像的特征。该特征图被共享用于后续 RPN 层和全连接层。

2) 区域建议网络: RPN 网络用于生成候选区域。该层判断锚框属于前景或者背景, 再利用边框回归修正锚框获得精确的预选框。

3) 感兴趣区域池化: 该层收集输入的特征图和预选框, 综合这些信息后提取特征图, 将相应区域池化为固定尺寸的特征图后送入后续全连接层判定目标类别与位置。

4) 分类: 利用特征图计算预选框的类别, 同时再次利用边框回归获得检测框最终的精确位置。

由 1.2 节数据分析可知, 铝材数据中表面缺陷复杂多样, 小缺陷占比巨大, 用基础的 Faster RCNN 架构检测铝材表面缺陷会有以下两个弊端。

1) 对小缺陷的识别能力差。Faster RCNN 网络仅仅在最后一层的特征图上做缺陷检测, 而随着网络深度的加深, 经过不断的卷积与池化操作, 特征图的分辨率大小会被不断减小, 在顶层的特征图会缺少甚至完全丢失细节信息, 导致对小缺陷的识别能力变差。

2) 对复杂缺陷的检测能力弱。RPN 对于锚框(anchor) 的设置默认的 9 种(默认为 3 种尺度 $\{32^2, 64^2, 128^2\}$ 3 种宽高比 $\{1:2, 1:1, 2:1\}$), 这对于复杂多样的 10 类铝材缺陷是不足的, 不能够很好的适应铝材表面缺陷导致检测效果不佳; 本文针对以上处出现的缺点, 提出改进的 Faster RCNN 结构。

改进后的 Faster RCNN 网络结构如图 3 所示, 铝材图像输入 ResNet-50 网络结合 FPN 结构提取不同尺度的特征图并融合, 结合 K-means 聚类算法获取定制的铝材表面缺陷的 anchors, 经区域建议网络得到候选区域, 将得到的不同大小的特征图和候选区域输入 ROI(region of interest) Align 进行区域校准池化, 得到准确的 7×7 大小的特征图后送入两个 1 024 维全连接层对其进行分类和边框回归。下面对改进的 Faster RCNN 结构进行具体分析。

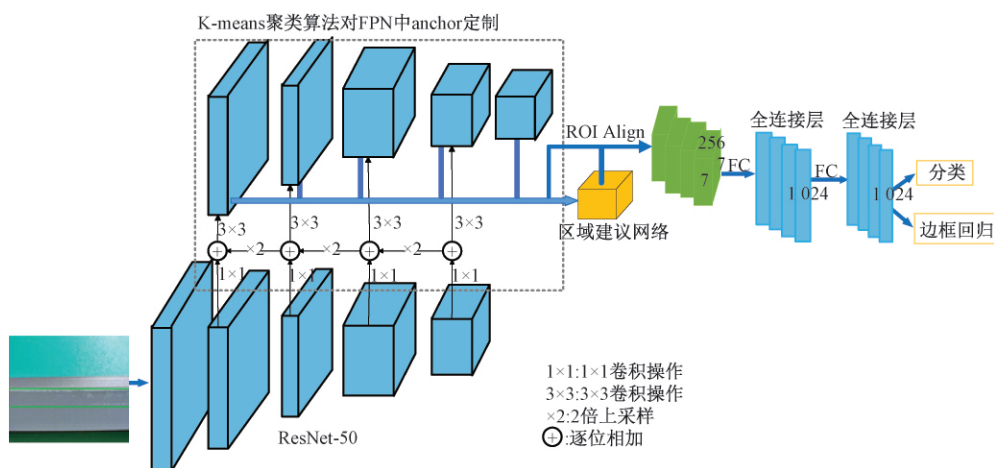


图 3 改进的 Faster RCNN 网络整体结构

Fig. 3 The overall structure of the improved Faster RCNN network

2.2 加入 FPN 的特征提取网络

特征金字塔 FPN^[18] 由一个自底向上的线路,一个自顶向下的线路,以及横向连接组成。如图4所示,其中自底向上的线路为原特征提取网络前向过程,特征图在此过程中的语义信息逐渐丰富而细节信息逐渐丢失。

Faster RCNN 加入 FPN 本质上是修改了特征提取部分,可以将深层的高语义特征与浅层特征图的细节特征相融合,使得检测网络对于小缺陷的检测能力变强。特征图经过横向连接将深层特征图2倍上采样后与相连浅层特征图进行自顶向下融合,其中 1×1 的卷积的作用是减少输出特征图的通道数,并在融合后的每个特征图上进行预测。

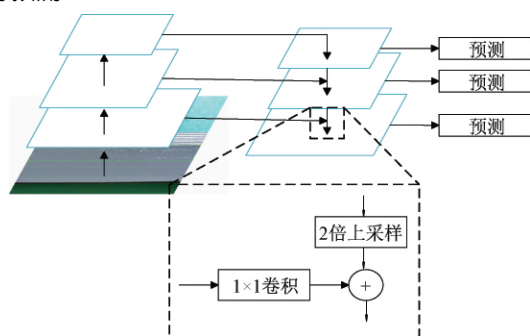


图4 FPN 算法结构

Fig. 4 FPN algorithm structure

本文的骨干网络为 ResNet-50,结合 FPN 的特征提取结构如图5所示,在 ResNet-50 的 {Conv2、Conv3、Conv4、Conv5} 卷积层进行自底向上特征的提取,并对每一层最深层的特征图进行 1×1 的卷积,使其特征图为256维后进行2倍上采样并与浅层自顶向下融合,将横向连接融合后的特征图与256个 3×3 的卷积核进行卷积运算,以减少融合过程中产生的混叠效应,加上对P5特征图的最大池化下采样后,最终得到{P2、P3、P4、P5、P6}5层特征图。与原 Faster RCNN 不同,对该5个特征图分别设置单尺度的5种锚框,分别为 $\{32^2, 64^2, 128^2, 256^2, 512^2\}$,锚框的宽高比仍然为 $\{1:2, 1:1, 2:1\}$ 种类增加到了15种。

候选区域生成后需要映射回单尺度特征图进行感兴趣区域池化,对于{P2、P3、P4、P5、P6}5层不同大小的特征图,不同大小的候选区域与特征图的映射关系如下:

$$k = k_0 + \log_2(\sqrt{w \times h}/224) \quad (1)$$

式中:224为ImageNet预训练图像大小; w 、 h 为ROI的宽和高; $k_0=5$ 表示特征图的映射种类。

结合 FPN 的特征提取网络,能够生成更为丰富的锚框,使得每个尺度利用到丰富的细节信息,因此能使模型对于缺陷的尺寸多样化和小缺陷的识别表现的更好。

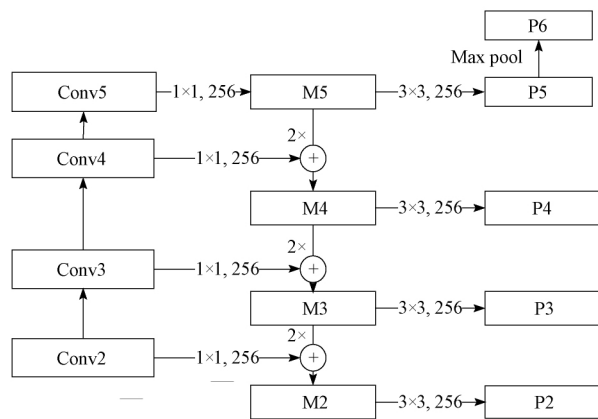


图5 ResNet + FPN 结构

Fig. 5 ResNet + FPN structure

2.3 感兴趣区域校准(ROI Align)

在原 Faster RCNN 中利用感兴趣区域池化(ROI Pooling)将不同尺寸的候选区域粗略的池化为相同尺寸特征图,此过程会产生两次量化误差,步骤如下:

1) 根据输入图像,将候选区域映射回特征图对应位置,映射后的浮点型坐标会被向下取整为矩阵坐标值;

2) 对获得的区域平均划分为 $k \times k$ (7×7)个单元(bin),对浮点单元坐标进行量化取整。

产生的两次量化误差会使得候选区域对于原始回归后的特征图上的位置出现较大偏差,此偏差对于小缺陷的检测影响是巨大的。如:对于脏点表面缺陷的像素大小往往在20以下,经过过本网络 ResNet-50 预训练网络提取特征后,缺陷和图像都被缩放为原图的 $1/32$,仅仅0.7 pixel的偏差都会导致该表面缺陷信息的丢失,也叫区域不匹配(misalignment)问题^[19]。

本文采用 ROI Align 来代替粗糙的 ROI Pooling,解决这种区域不匹配(misalignment)问题。

ROI Align 与 ROI Pooling 不同,不是简单的量化后池化,而是采用一种区域特征聚集方式将其转化为一个连续的操作,如图6所示。

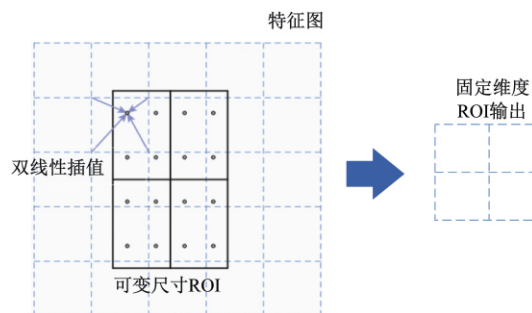


图6 ROI Align 原理

Fig. 6 Principle of ROI Align

1) 遍历所有候选区域,并保持映射后的候选区域浮点坐标不做量化。

2) 将候选区域划分为 $k \times k$ (图中为 2×2) 个单元,每个单元也不做量化。

3) 在各单元中确定 4 个采样点的位置,对采样点的浮点坐标使用双线性插值方法计算出这 4 个位置的值,然后进行最大池化操作得到固定维度的 ROI 输出。

其中,大的虚线网格代表特征图 (5×5),实线网格为可变尺寸的 ROI,即候选区域。每个单元内确定的 4 个采样点,是通过将单元平均划分为 4 个小方块,分别取方块中心点得到的。经过上述处理,能避免区域不匹配问题,使得模型能获得更加准确的候选特征区域,检测网络对于细小的缺陷检测能力进一步加强,获得更高精度。

2.4 基于 K-means 聚类算法的 anchor 定制

在基于 anchor 的检测任务中,结合 FPN 结构的 Faster RCNN 中 5 种尺度 3 种宽高比的 anchors 都是通过人工进行设计的先验框,对于本文数据集缺陷宽高比差异较大的情况,默认的先验框不能很好的适应本文铝材表面缺陷数据,则会导致最终的检测效果不佳。

本文通过 K-means 聚类算法^[20-21]训练出具有针对性的 anchors,对铝材表面缺陷数据集中待检测的 10 种标注缺陷框聚类出 anchor 的宽高维度,使得检测网络更容易学习到准确的缺陷检测 anchor,提高检测准确度。K-means 聚类算法是一种无监督学习,算法步骤如下:

1) 确定聚类簇个数 k ,随机初始化聚类中心;

(W_i, H_i) $i \in \{1, 2, 3, \dots, k\}$

2) 对所有的真实标注框 (w_i, h_i) 计算与 k 个聚类中心的距离 $d = 1/4 \times \text{IOU}$ (标注框,聚类中心),并将各标注框其归类给距离最近的聚类中心,形成 k 个簇;

3) 对步骤 2) 中的 k 个簇由式 (2) 更新聚类中心;

$$W'_i = \frac{1}{N} \sum w_i, H'_i = \frac{1}{N} \sum h_i \quad (2)$$

4) 重复步骤 2) 和 3) 直到标注框的聚类中心基本不发生变化或达到最大迭代次数。

由 K-means 聚类算法分析数据集标注数据的分布,令 $k = 15$ 时聚类得出的效果与结果分别如图 7、8 所示。

根据 K-means 聚类出的铝材表面缺陷 15 个 anchors 坐标大小以及宽高比结果,本文对 FPN 生成的锚框尺度重新定制为 $\{32^2, 96^2, 192^2, 398^2, 640^2\}$ 5 种。对于过大的 anchor 宽高比如: 7.1、10.5、20 等,新增 5:1 的 anchor 宽高比,最终 anchor 的 4 种宽高比为 $\{1:2, 1:1, 2:1, 5:1\}$ 结合 5 种尺度,共定制出 20 种 anchors。由 K-means 聚类算法定制出的 anchors,针对铝材表面缺陷数据集更加合理,能够使缺陷检测网络更快收敛,使模型获得更好的检测性能。

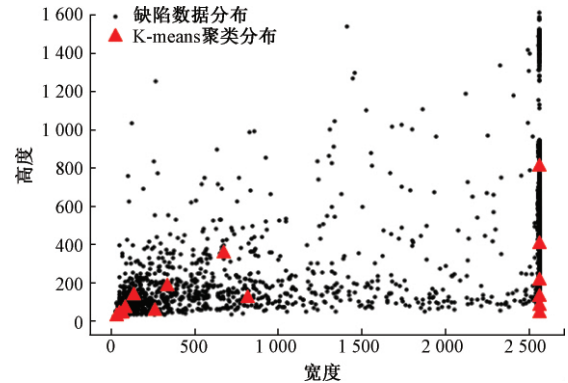


图7 K-means 聚类效果

Fig. 7 K-means clustering effect

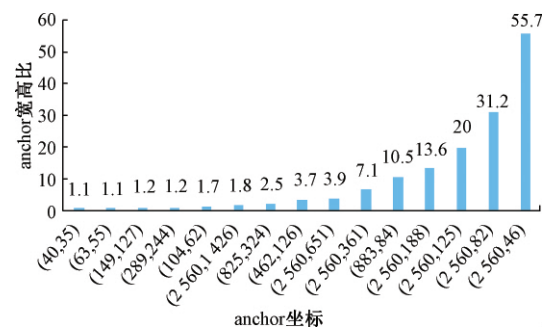


图8 K-means 聚类结果

Fig. 8 K-means clustering result

3 实验与结果分析

3.1 数据集增强

所获取的缺陷数据为 3 005 张,对于 10 类表面缺陷检测是不足的,而数据的缺少会导致训练过拟合、检测精度不高以及泛化能力差等问题,为了解决这些问题,本文采用水平翻转、垂直翻转、水平垂直翻转对原始数据进行增强。

数据增强示例如图 9 所示,增强后的数据扩增为 12 020 张缺陷图片,数据量提升到原来的 4 倍。将丰富后的图片按固定的批次输入深度检测网络中,仅会增加图片的数据量,而不会增加训练的运算量而且能增加模型的鲁棒性。

3.2 实验环境与超参数

本文实验的硬件环境为: AMD Ryzen5 3600 @ 3.6 Hz 六核处理器, GeForce RTX 2070 SUPER 的显卡, 250 G 的 860EVOSATA 机械硬盘; 软件环境为: 基于 64 位的 windows10 操作系统, 基于 pytorch 的 detectron2 框架, 基于 CUDA10.1, OpenCV2 库及 PyCharm 集成开发环境。

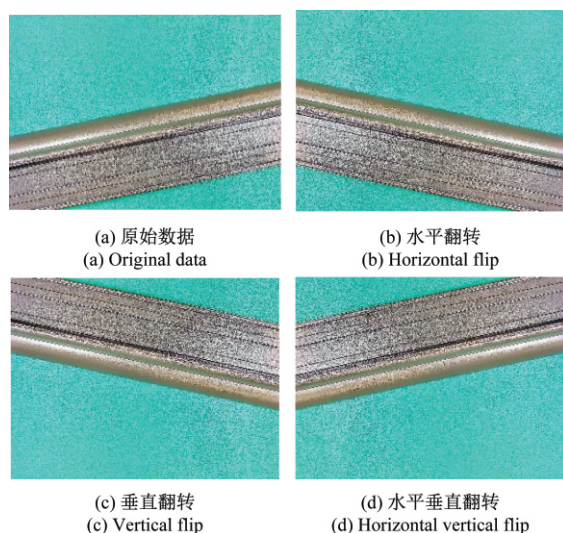


图9 数据增强示例

Fig. 9 Data enhancement examples

经过反复调参得到最终超参数设置为: 优化器为随机梯度下降算法, 优化器动量 momentum 为 0.9, 初始学习率为 0.002, 权值衰减 (weight decay) 为 0.000 1, 设置迭代次数为 40 000, 学习率迭代到 15 000 次进行衰减, 权值衰减倍数 (gamma) 为 0.1, 在刚开始训练时, 学习率慢慢增加到初始学习率, 热身迭代 10 次。

3.3 实验训练

训练时将 $2\,560 \times 1\,960$ 分辨率的铝材图片 resize 到 $1\,780 \times 980$, 将经过增强后的 12 020 张铝材图片划分为包含 10 220 张图片的训练集和 1 800 张图片的测试集。对于 Faster RCNN 检测网络的损失函数为位置回归 smooth_{L1} 函数和预测分类交叉熵函数, 检测网络整体损失函数如式 (3) 所示。

$$L(\{p_i, p_i^*\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (3)$$

式中: i 为 anchor 索引; p_i 和 t_i 分别为预测的概率和位置参数; p_i^* 和 t_i^* 分别为真实标注框的概率标签与位置参数; 当 anchor 与真实标注框的 IOU 超过设定值 (默认为 0.7) 时, 认为该 anchor 是正样本, 此时 $p_i^* = 1$, 当 anchor 与真实标注框的 IOU 低于设定值 (默认为 0.3) 时, 认为该 anchor 是负样本, 此时 $p_i^* = 0$; λ 为回归权重默认为 10; N_{cls} 为 mini-batch 大小; N_{reg} 为 anchor 位置的数量; L_{reg} 为位置回归 smooth_{L1} 函数如下:

$$\text{smooth}_{L1}(x) = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & \text{其他} \end{cases} \quad (4)$$

L_{cls} 为分类交叉熵损失函数:

$$L_{cls}(p_i, p_i^*) = -\log[p_i^* p_i + (1 - p_i^*)(1 - p_i)] \quad (5)$$

改进后的 Faster RCNN 网络每迭代 50 次纪录一次损失函数值, 将数据送入 coco 预训练模型进行训练, 最终总损失如图 10 所示。

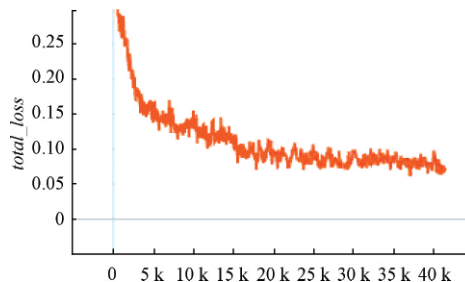


图10 改进模型总损失图

Fig. 10 The total loss diagram of the improved model

由图 10 可知, 训练总共迭代了 40 k 次, 随着训练的进行, 铝材表面检测网络总损失函数在迭代 20 k 之前持续下降, 检测网络模型迭代超过 20 k 之后基本趋于稳定, 检测模型达到收敛。

3.4 实验结果对比

为了直观地比较改进网络对于铝材表面缺陷的检测性能, 将最终 3 次改进后的 Faster RCNN 网络和基础 Faster RCNN 网络对于 10 种类别表面缺陷的平均精度 (average precision, AP) 以折线图的形式给出, 如图 11 所示。

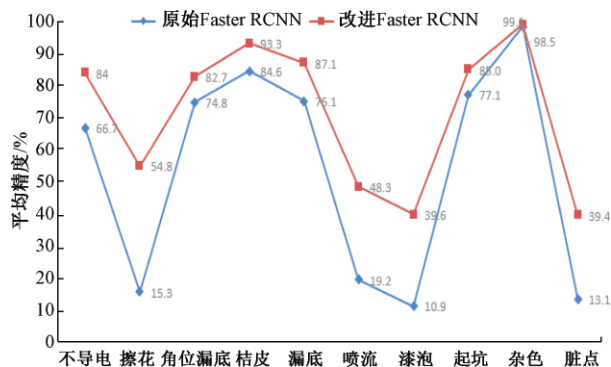


图11 各类别缺陷平均精度

Fig. 11 Average precision of various types of defects

最终改进后的 Faster RCNN 网络相较于原始网络在不导电、角位漏底、桔皮、漏底、起坑、杂色识别效果提高不明显。而在擦花、喷流、漆泡、脏点等中小型缺陷的类别上, 平均准确率都有较大提高, 可见改进算法对于中小型缺陷定位更准确, 对于 10 种类别的缺陷识别具有更好的效果。

为了验证本文对 Faster RCNN 改进算法的有效性以及对小缺陷的检测能力, 在相同的实验环境下, 分别在 1 800 张图片测试集上对逐步改进后的 4 种模型进行比较, 比较指标为各个模型对于 10 种表面缺陷检测的平均

精度均值(mean average precision, mAP) 和检测速度。其中 mAP^{IOU50} 选取 IOU 阈值为 0.5, 检测面积不限; mAP^{small} 的 IOU 阈值为 0.5:0.05:0.95, 检测面积为小于 32×32 像素的小缺陷, 比较结果如表 2 所示。

表 2 各模型性能对比

Table 2 Performance comparison of various models

	Faster RCNN	Faster RCNN + FPN	Faster RCNN + FPN + ROI Align	Faster RCNN + FPN + ROI Align + K-means
$mAP^{IOU50} / \%$	75.20	87.05	90.06	91.20
$mAP^{small} / \%$	16.9	21.2	33.9	39.1
速度(s/图)	0.257	0.288	0.296	0.299

由表可以看出, Faster RCNN 模型的 mAP^{IOU50} 为 75.2%, mAP^{small} 为 16.9%, 检测速度为 0.257 s/图, 随着对 Faster RCNN 基础网络的改进, mAP^{IOU50} 和 mAP^{small} 相比改进前模型都有一定程度的提升, 其中加入 FPN 结构后 mAP^{IOU50} 提高较为明显, 提高了 11.85%。改进 ROI Align 结构后 mAP^{small} 提升较大, 提高了 12.7%。最终改进的网络性能 mAP^{IOU50} 为 91.20%, mAP^{small} 为 39.1%, 检测速度为 0.299 s/图。 mAP^{IOU50} 和 mAP^{small} 的提高验证了经过 3 次改进方案的有效性, 表明改进后的网络对铝材小缺陷检测能力更强, 对铝材缺陷检测效果更佳。

最终改进的 Faster RCNN 网络对工业铝材的单缺陷、多缺陷的部分识别效果如图 12、13 所示。

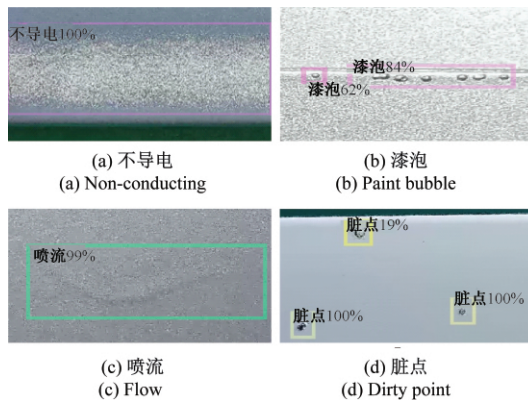


图 12 单缺陷检测效果

Fig. 12 Single defect detection effect

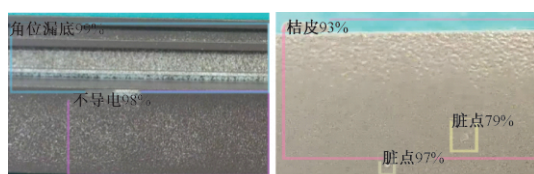


图 13 多缺陷检测效果

Fig. 13 Multi-defect detection effect

4 结 论

本文使用改进 Faster RCNN 的深度学习网络对铝材表面缺陷进行识别, 对 Faster RCNN 网络有以下 3 方面的改进。

1) 将 Faster RCNN 网络于图像金字塔 FPN 结构, 形成多尺度的特征融合模型, 增加对铝材表面小缺陷的识别能力。

2) 利用 ROI Align 代替 ROI Pooling, 以消除 ROI Pooling 产生的量化误差, 提供更加准确的候选框, 获得更高精度。

3) 通过 K-means 聚类算法得到针对铝材表面缺陷数据的定制 anchors, 使得模型收敛更快, 检测能力加强。

最终经过与原始 Faster RCNN 算法在测试集进行试验对比, 验证了最终改进后的 Faster RCNN 网络在检测速度相差不大的情况下, 整体 mAP^{IOU50} 从 75.2% 提升到 91.2%、对于 32×32 像素以下的铝材小缺陷, 识别精度 mAP^{small} 从 16.9% 提升到 39.1%, 相比原始 Faster RCNN 的铝材缺陷检测效果更佳, 为工业铝材质量检测提供一种有效的应用参考价值。

参考文献

- [1] 刘静安, 李党军, 单海珍. 我国铝镁合金挤压工艺装备现状及技术创新与发展趋向[J]. 铝加工, 2015(1): 51-54.
- [2] LIU J AN, LI D J, SHAN H ZH. The current situation, technological innovation and development trend of aluminum-magnesium alloy extrusion process equipment in my country [J]. Aluminum Processing, 2015(1): 51-54.
- [3] 张海伦. 铝板表面缺陷在线检测系统关键技术研究[D]. 济南: 齐鲁工业大学, 2014.
- [4] ZHANG H L. Research on key technologies of on-line inspection system for aluminum plate surface defects[D]. Jinan: Qilu University of Technology, 2014.
- [5] PERNKOPF F. Detection of surface defects on raw steel blocks using Bayesian network classifiers [J]. Pattern Analysis & Applications, 2004, 7(3): 333-342.
- [6] 杨理践, 郑文学, 高松巍, 等. 基于平衡电磁技术的钢板裂纹缺陷检测方法[J]. 仪器仪表学报, 2020, 41(10): 196-203.
- [7] YANG L J, ZHENG W X, GAO S W, et al. Steel plate crack defect detection method based on balanced electromagnetic technology [J]. Chinese Journal of Scientific Instrument, 2020, 41(10): 196-203.
- [8] 周鹏, 徐科, 刘顺华. 基于剪切波和小波特征融合的金属表面缺陷识别方法[J]. 机械工程学报, 2015,

- 51(6):98-103.
- ZHOU P, XU K, LIU SH H. Metal surface defect recognition method based on the feature fusion of shear wave and wavelet [J]. Chinese Journal of Mechanical Engineering, 2015, 51(6):98-103.
- [6] 程星振. 基于多模态信号的金属材料缺陷无损检测方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2016.
- CHENG X ZH. Research on non-destructive testing methods for metal material defects based on multi-modal signals[D]. Harbin: Harbin Institute of Technology, 2016.
- [7] 冯序文. 铝板表面复杂缺陷识别及分类方法研究[D]. 济南: 齐鲁工业大学, 2018.
- FENG X W. Research on the recognition and classification of complex defects on the surface of aluminum plates [D]. Jinan: Qilu University of Technology, 2018.
- [8] 蓝金辉, 王迪, 申小盼. 卷积神经网络在视觉图像检测的研究进展[J]. 仪器仪表学报, 2020, 41(4): 167-182.
- LAN J H, WANG D, SHEN X P. Research progress of convolutional neural network in visual image detection[J]. Chinese Journal of Scientific Instrument, 2020, 41(4): 167-182.
- [9] WANG T, CHEN Y, QIAO M, et al. A fast and robust convolutional neural network-based defect detection model in product quality control [J]. International Journal of Advanced Manufacturing Technology, 2017, 94(9): 3465-3471.
- [10] MEI S, YANG H, YIN Z. An unsupervised-learning-based approach for automated defect inspection on textured surfaces [J]. IEEE Transactions on Instrumentation and Measurement, 2018: 1266-1277.
- [11] 钟俊杰, 邵珩, 聂中原, 等. GPU 加速的路径依赖数字图像相关法[J]. 电子测量与仪器学报, 2020, 34(10): 17-24.
- ZHONG J J, SHAO H, NIE ZH Y, et al. GPU-accelerated path-dependent digital image correlation method [J]. Journal of Electronic Measurement and Instrument, 2020, 34(10): 17-24.
- [12] 蔡超鹏. 基于深度学习的金属轴表面缺陷检测与分类研究[D]. 杭州: 浙江工业大学, 2019.
- CAI CH P. Research on detection and classification of metal shaft surface defects based on deep learning [D]. Hangzhou: Zhejiang University of Technology, 2019.
- [13] 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.
- [14] GIRSHICK R. Fast R-CNN [C]. Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [15] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks [C]. European Conference on Computer Vision Springer, Cham, 2014: 818-833.
- [16] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [C]. International Conference on Learning Representations (ICLR), 2015: 1-14.
- [17] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) IEEE, 2016: 770-778.
- [18] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 936-944.
- [19] KAIMING H, GEORGIA G, PIOTR D, et al. Mask R-CNN [C]. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017: 2961-2969.
- [20] TAPAS K, DAVID M M, NATHAN S N C, et al. A local search approximation algorithm for K-means clustering [J]. Computational Geometry, 2002, 28(2-3): 89-112.
- [21] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]. IEEE Conference on Computer Vision & Pattern Recognition, 2017: 6517-6525.

作者简介



向宽 2014 年于湖北工程学院获得学士学位, 现为大连海洋大学硕士研究生, 主要研究方向为无损检测技术及图像处理技术。
E-mail: 1604133880@qq.com

Xiang Kuan received bachelor degree from Hubei Institute of Technology in 2014, and he is currently a master student at Dalian Ocean University. His main research direction is non-destructive testing technology and image processing technology.



李松松(通信作者) 2006 年毕业于日本富山县立大学, 获工学博士学位。现为大连海洋大学教授, 硕士生导师, 主要研究方向为超声无损检测技术, 无损检测信号及图像处理技术。
E-mail: lisongsong@dlou.edu.cn

Li Songsong (Corresponding author) received Ph. D. from Toyama Prefectural University, Japan in 2006. Now, she is a professor and master student supervisor in Dalian Ocean University. Her main research interest is ultrasonic non-destructive testing technology, non-destructive testing signal and image processing technology.