

基于改进FCOS的钢带表面缺陷检测算法

黄 颀

同济大学电子与信息工程学院, 上海

收稿日期: 2022年1月3日; 录用日期: 2022年1月21日; 发布日期: 2022年2月7日

摘 要

针对现有钢带表面缺陷检测所存在的检测效率低、适用范围有限等缺陷, 提出一种基于改进FCOS的钢带表面缺陷检测算法。该算法使用含形变卷积的卷积神经网络提取缺陷特征, 使用关键点特征融合增强检测模型输入, 并使用中心采样策略选取训练样本优化模型训练, 最后使用东北大学钢带表面缺陷公共数据集进行训练和评估。本文算法在东北大学钢带表面缺陷公共数据集上平均检测精度为74%, 检测速度为31.4 FPS。

关键词

钢带表面缺陷检测, 形变卷积, 特征融合, 中心采样

Algorithm on Surface Defect Detection of Strip Based on Improved FCOS

Qi Huang

College of Electronics and Information Engineering, Tongji University, Shanghai

Received: Jan. 3rd, 2022; accepted: Jan. 21st, 2022; published: Feb. 7th, 2022

Abstract

Aiming at the defects of low detection efficiency and limited applicable scope in strip surface defect detection, a steel strip surface defect detection algorithm based on improved FCOS was proposed. The algorithm uses convolutional neural network with deformable convolution to extract defect features, uses key point feature fusion to enhance the detection model input and uses the central point sampling strategy to select training samples to optimize model training. Finally, the proposed algorithm is trained and evaluated on public dataset NEU-DET, Northeastern University surface defect database. On NEU-DET dataset, the mean average precision of this algorithm achieves 74% and the detection velocity is 31.4 FPS.

Keywords

Strip Surface Defect Detection, Deformable Convolution, Feature Fusion, Center Sampling

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

作为钢铁产业的重要产品，钢带被广泛应用于各个领域。但由于设备和工艺条件的限制，钢带在生产过程中常出现划痕、裂缝等表面缺陷，不仅影响产品的外观，而且会影响产品的性能。因此，在钢带生产制造中对其进行表面缺陷检测具有重要意义。

目前，常用的钢带表面缺陷检测方法包括人工检测、漏磁检测和基于计算机视觉的检测。人工检测方法受检验人员主观因素影响较大，实时性差，误检率高。漏磁检测方法虽然成熟，但对不同缺陷的敏感度不同，漏检率高，且由于使用条件的限制，难以大规模推广[1]。随着深度学习的飞速发展，基于计算机视觉的检测技术以其智能化、实时性、高精度、可靠性等优点为钢带表面缺陷检测技术提供新的解决方案[2]。王等人提出一种融合多级特征的 Faster R-CNN 算法，解决了钢带表面多样化和随机缺陷的检测问题[3]。戴等人设计了一种基于改进 Faster R-CNN 的缺陷检测算法，以解决工件表面缺陷检测的局限性和精度不高的问题，检测效果相比传统方法得到大幅提高[4]。李等人提出一种改进 YOLOv3 算法，使用加权 K-means 聚类算法设计检测框参数提高先验框与特征图层匹配度，并增加残差单元和大尺度特征图输出，从而增强算法对缺陷检测能力[5]。刘等人提出一种改进 YOLOv4 算法，使用轻量级深层神经网络 MobileNetV3 作为特征提取网络以提高模型检测速度和特征提取能力，并设计多尺度损失函数解决因正负样本不平衡而导致检测效果差的问题，在保证检测速度的基础上有效提升了精度[6]。

然而，上述算法都是基于 Anchor-based 模型的实现，即需要基于先验知识给检测网络预设一组检测框，然后模型输出检测框微调参数，最终通过微调参数和检测框预设参数计算出最终模型输出的检测框，因此预设检测框的尺度和形状直接影响模型训练效果。并且由于预设检测框的尺度和形状与输入数据相关，所以 Anchor-based 模型设计依赖设计者对先验知识的了解，泛化能力有限。针对以上问题，本文提出基于改进 FCOS 的钢带表面缺陷检测算法，通过直接回归检测框参数而非对预设检测框微调实现缺陷检测，避免引入先验知识导致模型泛化能力不够问题。模型使用含形变卷积的卷积神经网络提取网络特征，增强特征提取效果，使用关键点特征融合方式丰富检测分支输入，增强模型检测效果，并使用中心采样策略选取合适的训练样本，优化模型训练效果。基于改进 FCOS 的表面缺陷检测算法在东北大学钢带表面公开数据集上平均检测精度(mean Average Precision, mAP)为 74.0%，检测速度为 31.4 FPS，性能满足当前工业生产需求。

2. FCOS 模型结构

FCOS [7]模型结构如图 1 所示，首先使用深度卷积神经网络提取输入图像特征图，然后使用分类网络、回归网络、中心度网络对特征图上的所有特征点逐个进行检测，分类网络输出原图上以此特征点为中心的区域所含缺陷类别，回归网络输出原图上以此特征点为中心的缺陷区域，中心度网络输出特征点在原图上对应的点为中心点的概率，抑制偏心度较大的低质量检测框的输出。

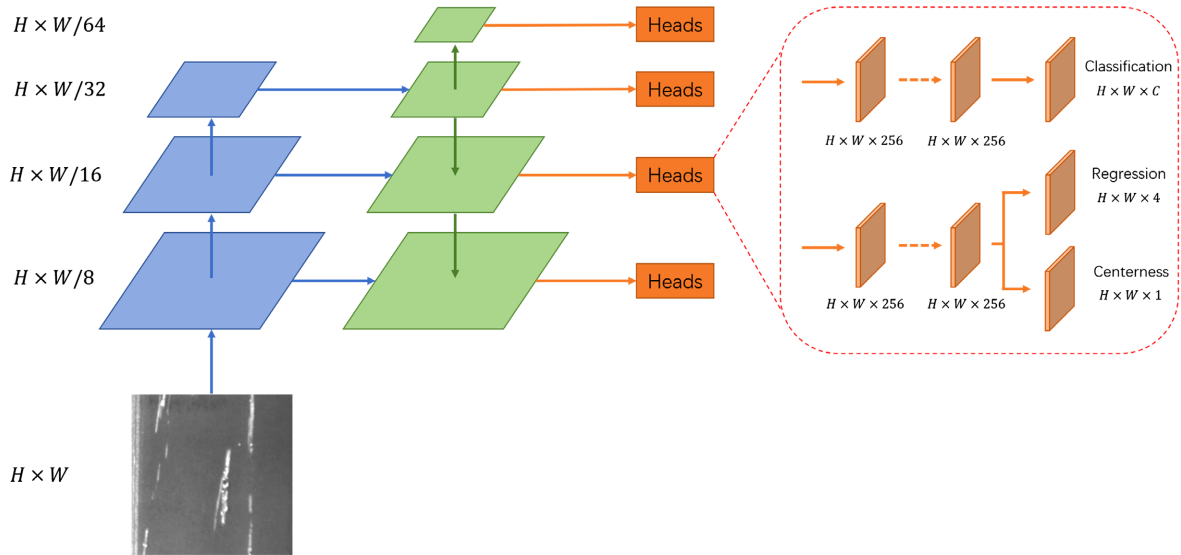


Figure 1. The basic structure of improved FCOS model

图 1. 改进 FCOS 模型基本结构

令 $F_i \in \mathbb{R}^{H \times W \times C}$ 是第 i 层的特征图, s 为该层总的下采样倍率, 输入图像中的真实检测框定义为 $\{B_i\}$, 其中 $B_i = (x_0^{(i)}, y_0^{(i)}, x_1^{(i)}, y_1^{(i)}, c^{(i)}) \in \mathbb{R}^4 \times \{1, 2, \dots, C\}$, 此处 $(x_0^{(i)}, y_0^{(i)})$ 和 $(x_1^{(i)}, y_1^{(i)})$ 分别为真实检测框的左上角点和右下角点的坐标, $c^{(i)}$ 为真实检测框中缺陷的类别, C 为缺陷类别个数。对于特征图 F_i 上的每个位置 (x, y) , 其在输入图像上的映射位置约为 $\left(\left\lfloor \frac{s}{2} \right\rfloor + xs, \left\lfloor \frac{s}{2} \right\rfloor + ys\right)$, 与 Anchor-based 通过回归基于此处预设的

检测框的微调参数从而得到真实检测框的方式不同, FCOS 直接回归该位置的检测框, 即 FCOS 直接将位置视为训练样本, 而不是将生成的检测框视为训练样本。具体来说, 如果坐标为 (x, y) 的这个特征点落在真实检测框中心点一定区域内且特征点的类别标签 c^* 为真实检测框的类别, 则将此特征点为正样本, 否则将其视为背景类负样本, 即 $c^* = 0$ 。除此之外, 检测网络还会输出一个 4 维向量 $t^* = (l^*, t^*, r^*, b^*)$ 作为特征点的回归结果, 此处 l^*, t^*, r^*, b^* 为特征点到边界框四个边的距离。

基于特征点的缺陷检测可以视作对以特征点为几何中心的矩形感受野区域的缺陷检测, 因此当特征点远离目标中心时, 存在特征点对应的感受野无法完全覆盖目标的情况, 基于这类特征点进行目标检测时则会生成很多低质量的检测框, 因此引入中心度概念, 通过中心度分支抑制此类检测框。

中心度描述了从特征点到特征点对应目标中心的归一化距离, 给定 l^*, t^*, r^*, b^* 时, 中心度监督信号定义如下:

$$centerness^* = \sqrt{\frac{\min(l^*, r^*)}{\max(l^*, r^*)} \times \frac{\min(t^*, b^*)}{\max(t^*, b^*)}} \quad (1)$$

中心度数值范围为 0 到 1, 训练时使用交叉熵作为损失函数进行训练, 测试时, 特征点对应缺陷类别的最终分数由分类分数乘中心度获得, 因此, 中心度可以降低远离目标中心的边界框分数, 使得这些低质量边界框在 NMS 后处理中被过滤掉, 从而显著提高检测性能。

3. FCOS 的改进

3.1. 形变卷积

CNN 网络由卷积层和池化层组成, 由于卷积层和池化层都是对固定尺度的区域进行采样, 即在同一

层的 CNN 模块的每个激活单元拥有固定尺度和形状的感受野,然而不同位置的物体的尺度和形状都是不一样的,对于尺度、形状和感受野不匹配的物体,感受野固定的 CNN 无法充分提取物体的特征(如图 2 所示)。因此,本文引入形变卷积(Deformable Convolution Network, DCN) [8]模块增强深度卷积网络特征提取能力。

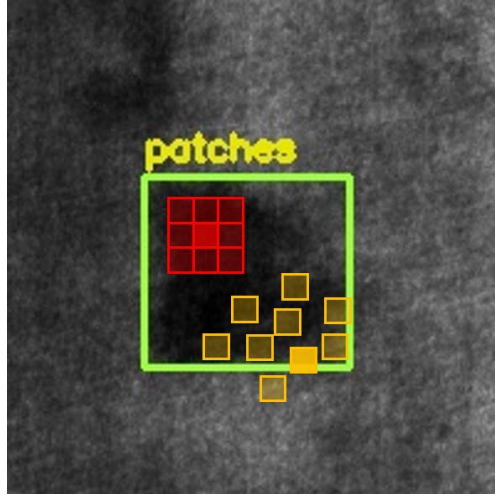


Figure 2. 3×3 convolution receptive field (the red part is the receptive field of conventional convolution, and the yellow part is the receptive field of deformable convolution)

图 2. 3×3 卷积感受野(红色为常规卷积感受野,黄色为形变卷积感受野)

常规卷积由两步组成: 1) 在输入特征图 x 上对矩形网格 \mathcal{R} 区域进行采样; 2) 对采样值使用权重 w 进行加权求和。其中网格 \mathcal{R} 定义了感受野的大小和膨胀比例,如 $\mathcal{R} = \{(-1,-1),(-1,0),\dots,(0,1),(1,1)\}$ 表示一个膨胀比例为 1、大小为 3×3 的卷积核。对于输出特征图 y 上的每个特征点 p_0 , 都有

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

其中 p_n 为 \mathcal{R} 上的每个点的坐标。在形变卷积中,矩形网格 \mathcal{R} 使用偏移量 $\{\Delta p_n | n=1,\dots,N\}$ 表示,其中 $N = |\mathcal{R}|$, 则卷积计算变为

$$y(p_0) = \sum_{p_n \in \mathcal{R}} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \quad (3)$$

此时采样变为非规则且基于偏移量 $p_n + \Delta p_n$, 且由于偏移量 Δp_n 通常都为小数,因此公式(3)中的 $x(p_0 + p_n + \Delta p_n)$ 通过双线性插值实现,公式如下:

$$x(p) = \sum_q G(q, p) \cdot x(q), \quad p = p_0 + p_n + \Delta p_n \quad (4)$$

其中 q 为特征图 x 上的整数空间位置, $G(\cdot, \cdot)$ 为双线性插值核函数,由于 $G(\cdot, \cdot)$ 为 2 维计算,可以拆分成两个单维核函数相乘:

$$G(q, p) = g(q_x, p_x) \cdot g(q_y, p_y), \quad g(a, b) = \max(0, 1 - |a - b|) \quad (5)$$

如图 3 所示,采样点的偏移量通过在相同的输入特征图上使用旁路卷积获得,旁路卷积核的尺度和膨胀比率与当前卷积层相同,且输出的偏移量向量空间分辨率与输入特征图相同,通道数为输入特征图的两倍,与偏移量的两个维度相对应。在训练时,同时学习用于生成输出特征的卷积核与偏移量,梯度通过公式(4)和公式(5)反向传播。

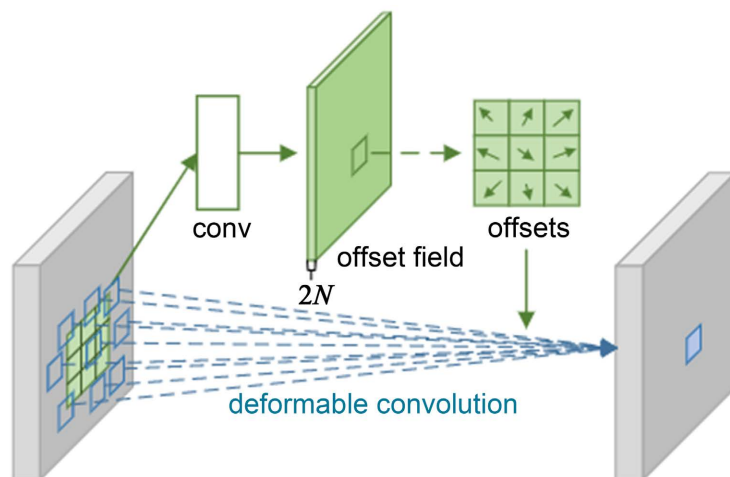


Figure 3. Deformable convolution network

图 3. 形变卷积网络

3.2. 关键点特征融合

在 FCOS 中，检测网络是基于特征点进行检测，也可看作基于以特征点为几何中心的矩形感受野区域的检测，然而由于缺陷尺度、形状各异，特征点又是根据空间映射直接计算出来的，特征点感受野区域和缺陷区域可能存在一定程度的不匹配，使用此类特征点进行缺陷检测则会影响模型的检测性能。

针对以上问题，本文提出基于关键点特征融合的方式进行检测，即将原本基于单个特征点的检测拓展为基于一组关键点进行检测，使用一组关键点而不是单个特征点描述缺陷特征，丰富输入网络的信息，提升检测网络的表现。

如图 4 所示，左图为使用单个特征点检测，输入检测网络的信息为此特征点对应的感受野区域的信息，右图为使用关键点集合检测，输入检测网络的信息为这组关键点对应的感受野区域的综合信息，相较于单个特征点，关键点集合能够提供更丰富的缺陷信息，从而更易于后续的检测。

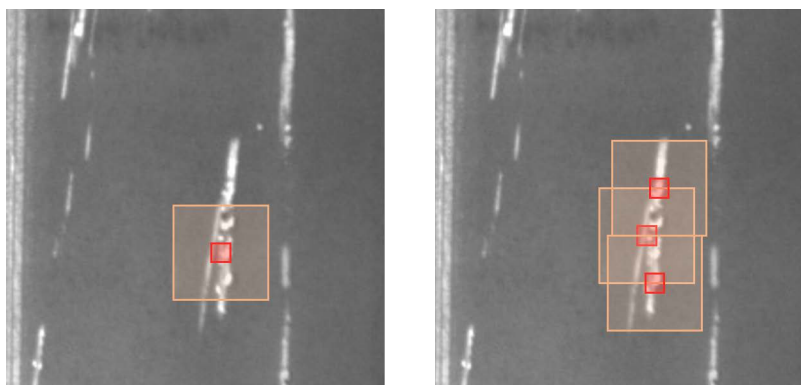


Figure 4. Sketch map of single feature point representation and key point set representation (red rectangle is the feature point, and yellow rectangle is the receptive field corresponding to the feature point)

图 4. 单个特征点表征与关键点集合表征示意图(红色矩形为特征点，黄色矩形为特征点对应的感受野)

关键点集合的生成与融合使用 DCN 实现，如图 5 所示，首先使用 1×1 卷积生成 $2k$ 个需要融合的关键点坐标，然后根据坐标使用双线性插值获得关键点集合对应的特征，然后使用 1×1 卷积对这些特征进行融合，并与原特征点特征合并，作为检测网络输入。

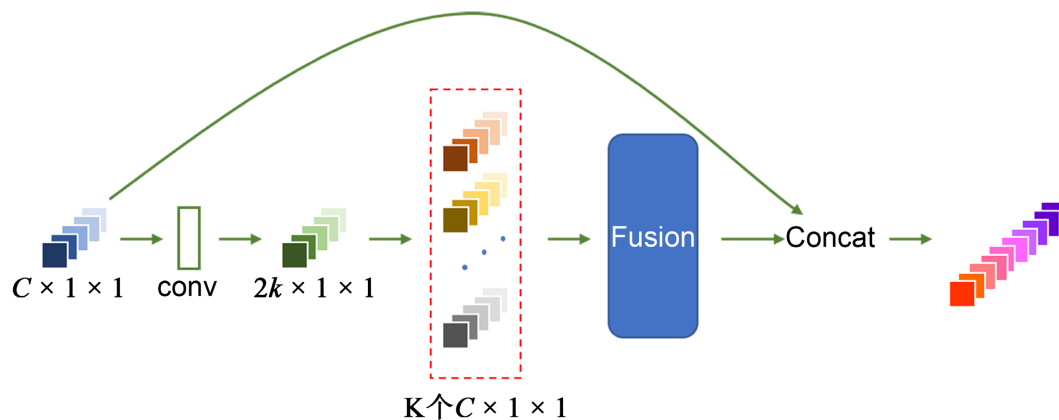


Figure 5. Generation and fusion of key point sets

图 5. 关键点集合生成与融合

3.3. 中心采样策略

在原始的 FCOS 实现中, 只要特征点在原图上的映射点处在标注框内部, 此特征点会被视作正样本。然而标注框内部靠近边缘的部分往往仍然是背景部分, 因此这部分正样本实际对应的点应该属于负样本的背景部分, 使用这些样本进行训练会对模型的学习造成困扰。

针对以上问题, 本文使用中心采样策略选取正样本(如图 6 所示), 即以标注框中心点为圆心, 取一个比标注框更小的圆形区域, 只有在原图上的映射点处于此圆形区域的特征点才会被视作正样本, 这样就将大多数处在标注框内边缘、实际落在背景部分的特征点正确识别为负样本, 有利于模型的正确学习。在本文提出的改进 FCOS 模型中, 使用标注框短边的 0.75 作为中心采样区域的直径。

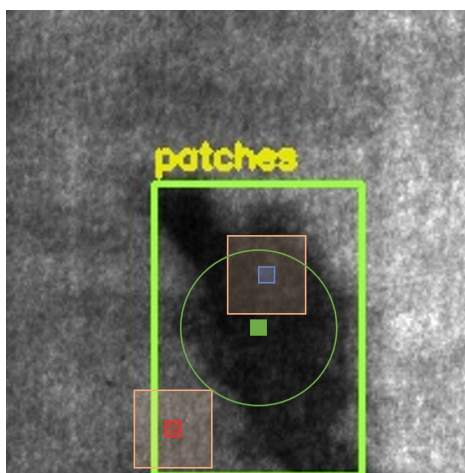


Figure 6. Center sampling (the feature points falling in the green ring are regarded as positive samples, and vice versa)

图 6. 中心采样(落在绿色圆环内的特征点被视作正样本, 反之为负样本)

4. 实验结果与分析

为了验证本模型的可行性和优势, 我们使用东北大学的 NEU-DET 数据集进行实验。

4.1. NEU-DET 数据集

NEU-DET 数据集是东北大学宋克臣团队[9]制作的钢带表面缺陷数据集, 包含 6 种缺陷类别, 分别

为裂纹(Crazing), 夹杂(Inclusion), 斑块(Patch), 点蚀表面(Pitted Surface), 轧制氧化皮(Rolled-in Scale), 划痕(Scratch), 每种缺陷含有 300 张分辨率为 200×200 的灰度图片, 总样本数为 1800。数据集同时提供每张图片所对应的缺陷标注信息的 XML 文件, 标注信息包括缺陷所属类别与边界框的信息(矩形框左上角和右下角的坐标信息), 共计 4189 个边界框。图 7 为 NEU-DET 数据集中 6 种表面缺陷图像的示例样本。

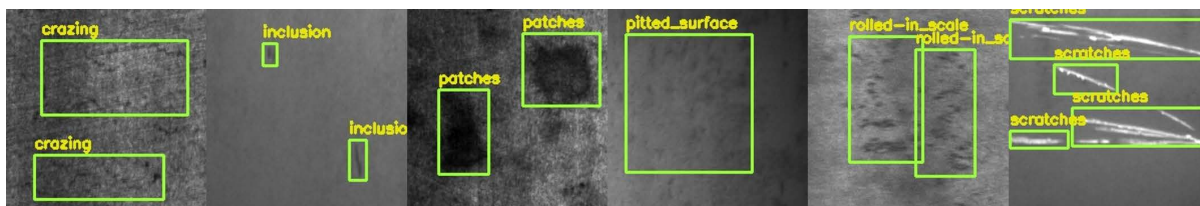


Figure 7. Examples of defect sample in NEU-DET dataset

图 7. NEU-DET 数据集缺陷样本示例

4.2. 评价指标

本文采用的评价指标为 COCO 数据集评价指标[10], 包括 mAP、AP₅₀、AP₇₅、AP_S、AP_M、AP_L。COCO 评价指标根据模型输出的目标类别和模型输出检测框与实标注框的交并比(Intersection over Union, IoU)判断是否正确检测到目标。其中 AP₅₀、AP₇₅ 指的是 IoU 阈值分别为 0.5 和 0.75 时的检测精度, IoU 阈值越高意味着对检测精度要求越高, mAP 则是 IoU 阈值从 0.5 到 0.95 的平均精度。AP_S、AP_M、AP_L 分别为对小目标、中目标、大目标检测精度, 在 COCO 数据集中, 面积小于 32^2 的为小目标, 面积大于 32^2 小于 96^2 为中目标, 面积大于 96^2 的为大目标。

4.3. 实验结果

实验采用的硬件配置为 GTX1080Ti 显卡, 软件环境为 CUDA10.1, Cudnn7.6.4。每个批次使用 8 张图片进行训练, 使用含动量的随机梯度下降优化参数, 学习率为 0.2, 动量为 0.9, 权重衰减速率为 0.0001, 每个模型共训练 36 轮, 实验结果如图 8 和表 1 和所示。

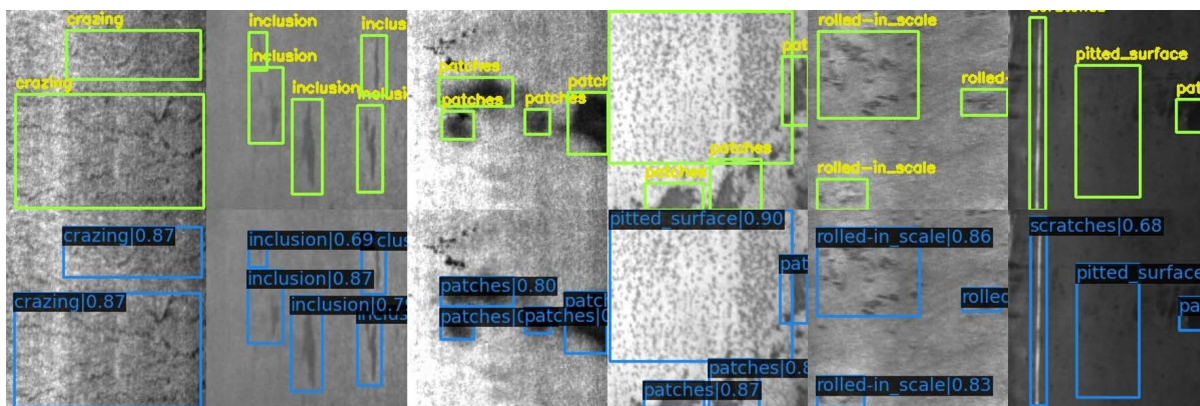


Figure 8. Detection results of improved FCOS (top row is ground truth, bottom row is the results of improved FCOS)

图 8. 改进 FCOS 模型检测结果(上图为真实标注框, 下图为改进 FCOS 输出结果)

本文在不同主流模型上进行实验对比, 实验结果如表 1 所示。从表中可以看出改进 FCOS 模型在 NEU-DET 数据集上的各项表现均优于其他主流模型, 尤其在 AP₇₅ 指标和 AP_S 指标上, 改进 FCOS 相较于主流模型提升巨大, 相较于 Faster R-CNN, AP₇₅ 提升 11.8%, AP_S 提升 16.1%, 相较于 YOLOv3, AP₇₅

提升 22.1%， AP_S 提升 25.8%。改进 FCOS 在不使用基于先验知识设计的预设检测框的情况下，不仅实现了高精度检测，而且大幅提升了缺陷检测算法对小尺度目标的检测精度，证明 Anchor-free 算法在缺陷检测领域的应用价值。

Table 1. Comparison of experimental results of various defect detection models

表 1. 各类缺陷检测模型实验结果对比

模型	mAP (%)	AP_{50} (%)	AP_{75} (%)	AP_S (%)	AP_M (%)	AP_L (%)
Faster R-CNN	63.0	94.3	72.5	49.8	61.7	72.7
YOLOv3	56.4	93.1	62.2	40.1	53.9	66.1
FCOS	63.6	94.1	71.3	57.5	60.6	75.1
改进 FCOS	74.0	96.6	84.3	65.9	71.9	82.9

5. 总结

针对钢带表面缺陷检测问题，本文提出一种基于改进 FCOS 的表面缺陷检测算法，使用直接回归检测框的思想实现缺陷检测，避免引入先验知识导致模型性能波动和泛化性降低。在特征提取部分使用形变卷积增强卷积网络对不同尺度、形状缺陷的特征提取能力。在检测部分使用关键点特征融合方法丰富输入网络的信息。使用中心采样策略定义样本正负，提高训练样本质量，改善模型训练效果。最终改进 FCOS 模型在 NEU-DET 数据集上 mAP 达到 74.0%，检测速度为 31.4 FPS，并且在检测精确度和小目标检测上均大幅领先主流算法，证明 Anchor-free 思想在缺陷检测领域的应用价值，为缺陷检测模型设计提供新的思路。

参考文献

- [1] 吴平川, 路同浚. 钢板表面缺陷的无损检测技术与应用[J]. 无损检测, 2000, 22(7): 312-315.
- [2] 顾佳晨, 高雷, 刘路路. 基于深度学习的目标检测算法在冷轧表面缺陷检测中的应用[J]. 冶金自动化, 2019, 43(6): 19-22.
- [3] 王海云, 王剑平, 罗付华. 融合多层次特征 Faster R-CNN 的金属板带材表面缺陷检测研究[J]. 机械科学与技术, 2021, 40(2): 262-269.
- [4] Dai, X., Chen, H. and Zhu, C. (2020) Research on Surface Defect Detection and Implementation of Metal Workpiece Based on Improved Faster R-CNN. *Surface Technology*, **49**, 362-371.
- [5] 李维刚, 叶欣, 赵云涛, 等. 基于改进 YOLOv3 算法的带钢表面缺陷检测[J]. 电子学报, 2020, 48(7): 1284-1292.
- [6] 刘艳菊, 王秋霖, 赵开峰, 等. 基于卷积神经网络的热轧钢条表面实时缺陷检测[J/OL]. 仪器仪表学报: 1-10. <http://kns.cnki.net/kcms/detail/11.2179.TH.20211230.1604.024.html>
- [7] Tian, Z., Shen, C., Chen, H., et al. (2019) FCOS: Fully Convolutional One-Stage Object Detection. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 9627-9636. <https://doi.org/10.1109/ICCV.2019.00972>
- [8] Dai, J., Qi, H., Xiong, Y., et al. (2017) Deformable Convolutional Networks. 2017 *IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 764-773. <https://doi.org/10.1109/ICCV.2017.89>
- [9] Song, K. and Yan, Y. (2013) A Noise Robust Method Based on Completed Local Binary Patterns for Hot-Rolled Steel Strip Surface Defects. *Applied Surface Science*, **285**, 858-864. <https://doi.org/10.1016/j.apsusc.2013.09.002>
- [10] Lin, T.-Y., Maire, M., Belongie, S., et al. (2014) Microsoft COCO: Common Objects in Context. *European Conference on Computer Vision*, Zurich, 6-12 September 2014, 740-755. https://doi.org/10.1007/978-3-319-10602-1_48