

文章编号: 2095-2163(2023)11-0069-08

中图分类号: TP391.4

文献标志码: A

基于改进 YOLOv7-Tiny 的工业缺陷检测研究

刘凌峰¹, 陈洪刚¹, 卿颀波¹, 孙承行²

(1 四川大学 电子信息学院, 成都 610065; 2 四川省隆鑫科技包装有限公司, 四川 遂宁 629200)

摘要: 针对现代工业生产对工业产品缺陷的高效检测需求,以钢材缺陷为切入点,提出了一种基于改进 YOLOv7-Tiny 的缺陷检测方法。首先,将 DCNv3 网络与 YOLOv7-Tiny 的主干网络层进行融合,提高模型对不同尺度、形状缺陷的灵敏度;其次,引入了 CABM 注意力机制,加强网络对特征图中重要信息的感知能力;然后,在网络特征融合层使用了 CARAFE 算子,改善原始网络语义信息利用不充分的问题;最后,采用基于归一化 Wasserstein 距离的损失函数来改善在小目标检测中网络对预测框与真实框位置偏移敏感的问题。实验结论表明,本文改进后的 YOLOv7-Tiny 算法在开源钢材缺陷数据集 NEU-DET 上平均准确率达到 77.5%,较原始算法提升了 4.3%,同时模型参数量相较于原始 YOLOv7-Tiny 算法降低了 11.3%,达到了检测精度与模型复杂度的平衡。

关键词: 钢材缺陷检测; YOLOv7-Tiny; 可变形卷积; 注意力机制; 归一化 Wasserstein 距离

Research on industrial defect detection based on improved YOLOv7-Tiny

LIU Lingfeng¹, CHEN Honggang¹, QING Linbo¹, SUN Chenghang²

(1 College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China;

2 Sichuan Longxin Technology Packaging Co., Ltd., Suining Sichuan 629200, China)

Abstract: Aiming at the requirement of high efficiency detection of industrial product defects in modern industrial production, a defect detection method based on improved YOLOv7-Tiny is proposed, taking steel defects as an entry point. Firstly, the DCNv3 network is fused with the backbone network layer of YOLOv7-Tiny to improve the sensitivity of the model to defects of different sizes and shapes. Secondly, CABM attention mechanism is introduced to strengthen the network's ability to perceive important information in feature maps. Then, CARAFE operator is used in the network feature fusion layer to improve the inadequate use of the original network semantic information. Finally, the loss function based on normalized Wasserstein distance is used to improve the sensitivity of the network to the position deviation of the predicted frame and the real frame in small target detection. Experimental results show that the improved YOLOv7-Tiny algorithm in this paper has an average accuracy of 77.5% on the open source steel defect dataset NEU-DET, which is 4.3% higher than that of the original algorithm, and the number of model parameters is 11.3% lower than that of the original YOLOv7-Tiny algorithm. The balance between detection accuracy and model complexity is achieved.

Key words: steel surface inspection; YOLOv7-Tiny; deformable convolution; attention mechanism; normalized Wasserstein distance

0 引言

随着工业 4.0 的到来以及中国制造 2025 的提出,国家大力支持工业的发展。钢材作为一种经济高效的工业材料在各个领域得到广泛应用,如建筑、化工、船舶、冶金、航空航天等行业。拥有高质量的钢材对我国工业的发展至关重要。在钢材的生产过程中,由于生产原材料、生产工艺和外部环境等不可控因素的影响,钢表面可能会出现各种不同类型的

缺陷^[1]。因此,对钢材缺陷进行高效的检测具有极其重要的意义。

基于传统机器视觉的目标检测算法使用滑动窗口的方法来获取候选区域的 HOG 特征^[2]和 SIFT 特征^[3],并利用 SVM 分类器^[4]进行目标分类。这种检测方法主要适用于待测目标特征明显、背景简单的场景,并且在检测精度以及速度上存在一定的局限。在钢材缺陷图像中,背景与目标的区分度较低且缺陷目标类型复杂多变,使用传统的抽象特征很难完

作者简介: 刘凌峰(1999-),男,硕士研究生,主要研究方向:计算机视觉、人工智能;陈洪刚(1991-),男,博士,副研究员,主要研究方向:计算机视觉、人工智能;孙承行(1982-),男,工程师,主要研究方向:机械设计。

通讯作者: 卿颀波(1982-),男,博士,教授,主要研究方向:计算机视觉、人工智能。Email: qing_lb@scu.edu.cn

收稿日期: 2023-04-12

成对缺陷的准确检出。

目前,基于深度学习的目标检测算法根据其是否采用多阶段的处理,主要分为2类:Two-Stage^[5]和One-Stage^[6]。Two-stage典型算法代表为Faster R-CNN^[7]和Mask R-CNN^[8],其使用区域建议网络生成候选框,而后进行分类和回归;One-Stage算法典型代表为SSD^[9]、RetinaNet^[10]、EfficientDet^[11]、YOLO系列等,是利用处理回归问题的方法来有效解决目标边框定位的问题,因此不需要事先产生候选框。在Two-Stage目标检测算法领域,LUO等学者^[12]基于CNN构建了一种解耦的两阶段的工业目标检测网络,并在缺陷定位任务中提出了多级聚合块作为定位特征增强模块,用于PCB板的缺陷检测。Shi X等学者^[13]将ConvNeXt架构作为特征提取主干,提出了一种基于Faster R-CNN的改进网络用于钢表面缺陷检测。在One-Stage目标检测算法领域,KOU等学者^[14]基于YOLOv3进行改进,使用无锚框(Anchor-free)的方式缩短了模型的计算时间,设计密集卷积块(Dense block)用于提取更丰富的特征信息,从而提升模型的准确率与鲁棒性。Yang等学者^[15]首先将钢材缺陷图像经过同态滤波预处理,然后将图片送入融合了注意力机制的改进YOLOv5算法中进行检测,达到了更高的检测精度。

YOLO系列作为One-Stage算法典型的代表网络,不仅在速度上优势明显,而且模型体积小,同时维持较高的识别精度。在2023年,Wang等学者提

出了YOLOv7算法,相比于领域内其他算法,YOLOv7在目标识别和定位方面具有更高的准确性和更快的速度。但是YOLOv7网络使用了大量E-ELAN模块,导致模型的参数数量和复杂度大大增加,同时钢材表面缺陷通常是小目标,在图像中所占的像素点很少,且变形类型多样、缺陷类型多样,很容易被忽略或误判,导致漏检的问题。

为了提高钢材表面缺陷检测的效率,本文基于YOLOv7的轻量化版本YOLOv7-Tiny,进一步优化网络对目标的特征提取能力,引入混合注意力机制,增强网络的特征感知;将可变形卷积融合进网络骨干层,构建ELAN-DCNv3模块,增强模型对不同尺寸和形状的特征提取能力;在网络特征融合层加入了基于内容感知的特征重组模块,更加充分地利用上下文语义信息;最后,采用基于归一化Wasserstein距离的损失函数实现对小目标缺陷更加精准的检测。

1 YOLOv7-Tiny 网络结构

YOLOv7-Tiny保留了YOLOv7基于级联的模型缩放策略,并改进了高效长程聚合网络(ELAN),是一种轻量级的目标检测算法,在保证检测精度没有大幅降低的基础上参数数量更少、检测速度更快,更适合应用于钢材表面缺陷实时检测的需求。YOLOv7-Tiny网络结构如图1所示。

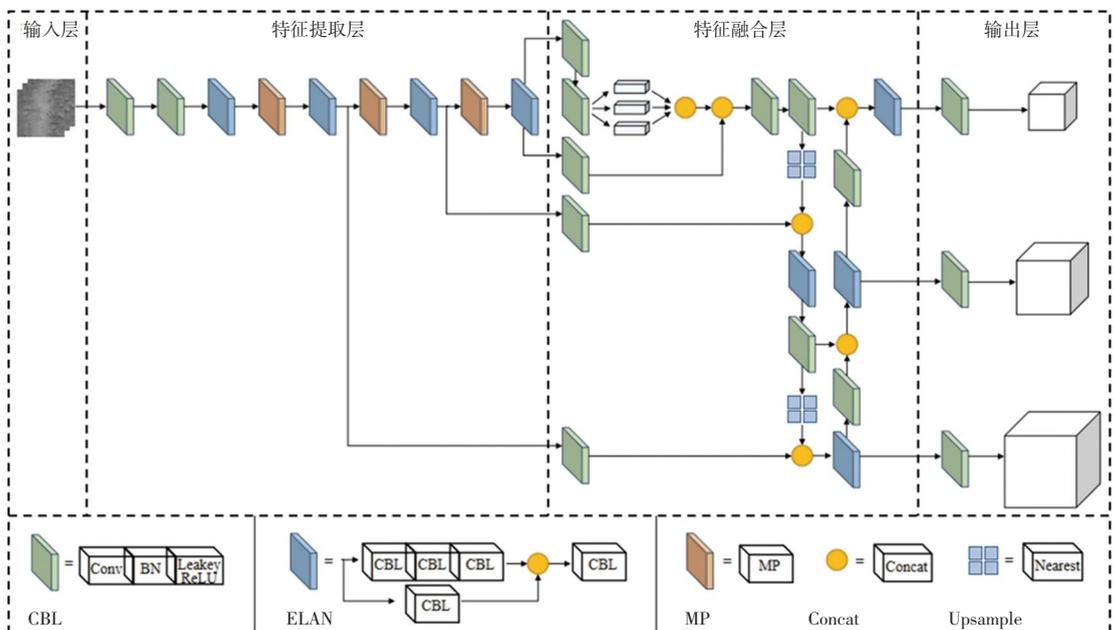


图1 YOLOv7-Tiny网络

Fig. 1 YOLOv7-Tiny network

Yolov7-Tiny 算法由输入层、特征提取层、特征融合层及输出层四个部分构成。将固定尺寸的图片作为输入, 首先将其送入由普通卷积层、MP 与 ELAN 卷积层构成的特征提取层。然后, 将提取的特征图送入在 SPP 结构基础上优化后的 SPPCSPC 模块, 进行处理后再送入特征融合层, 采用聚合特征金字塔结构对整个金字塔特征图进行增强, 将下层的显著性特征信息传递到上层, 保留更丰富的特征信息。最后, 将这些语义丰富的特征图送入输出层, 采用卷积对不同尺度的特征图进行预测, 在目标特征图上应用锚框生成带有类别概率和目标框的最终输出向量。

尽管 YOLOv7-Tiny 具有较少的参数和计算量, 能保持较高的精度, 但是针对钢材表面缺陷这类小目标, 存在特征提取能力不足的情况。因此, 本文提出改进方法, 在不增加网络参数的前提下加强特征感知, 以满足钢材缺陷这类特征较弱的目标在实际应用场合的检测需求。

2 改进型 YOLOv7-Tiny 钢材表面缺陷检测算法

本文基于 YOLOv7-Tiny, 将可变形卷积融合进网络骨干层, 增强模型对不同尺寸和形状的特征提取能力; 引入混合注意力机制, 增强网络的特征感知; 在网络特征融合层加入了基于内容感知的特征重组模块, 更加充分的利用上下文语义信息; 最后采用基于归一化 Wasserstein 距离的损失函数实现对小目标缺陷更加精准的检测。

2.1 构建 ELAN-DCNv3 模块

钢材表面缺陷具有目标小、变形类型多样、缺陷类型多样的特点, 普通卷积层对此类目标的特征提取能力较差。本文中使用的 DCNv3^[17] (Deformable Convolution Network v3) 是在 DCNv2^[18] (Deformable Convolution Network v2) 的基础上进行改进的。DCNv3 首先借用了可分离卷积的思想, 将原始卷积权重分离成深度部分和点部分, 实现了卷积神经元之间的权重共享; 其次, 引入了多组卷积的机制, 将空间聚合过程分成多组, 每组具有单独的采样偏移和调制尺度, 从而为下游任务带来更强的特征; 最后, 通过用 k 个调制因子的 *softmax* 归一化代替调制标量 *sigmoid*, 整个训练过程变得更加稳定。完整的 DCNv3 运算符如式(1)所示:

$$y(p_0) = \sum_{g=1}^G \sum_{k=1}^K w_g m_{gk} x_g(p_0 + p_k + \Delta p_{gk}) \quad (1)$$

其中, G 表示卷积组的数量; W_g 表示每组内的共享投影权重; m_{gk} 表示第 g 组中第 k 个采样点的归一化调制因子; x_g 表示切片输入的特征图; Δp_{gk} 表示第 g 组中的网格采样位置 (g, k) 相对应的偏移量。

本文将可变形卷积 DCNv3 与特征提取层的 ELAN 结构融合, 重建 ELAN-DCNv3 结构, 结构如图 2 所示。

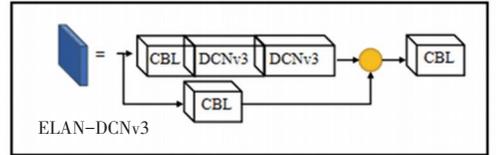


图 2 ELAN-DCNv3 结构

Fig. 2 Structure of ELAN-DCNv3

使用 DCNv3 算子可以弥补传统卷积在长期依赖性和自适应空间聚合方面的不足, 并可使卷积算子更适合于变化多样缺陷的检测, 同时通过实现稀疏全局建模并适当地保留 CNN 的归纳偏差, 更好地实现了检测复杂性和准确性之间的平衡。

2.2 构建 ELAN-CBAM 模块

注意力机制的原理是选择性地关注更感兴趣的领域, 而忽略信息的其他部分。在钢材缺陷检测中引入注意力机制可以提高缺陷目标区域的权重。神经网络可以更完整地地区分待检测的缺陷目标和背景信息, 有效地解决由于网络层次加深导致的钢材缺陷目标信息丢失的问题, 使网络学习能够朝着更有利的方向学习。

在真实的钢材缺陷数据中, 缺陷目标较多且种类繁多, 为了解决小目标漏检、误检的问题, 需要让网络更加关注裂纹、麻点等小目标缺陷。本文选用 CBAM 注意力机制 (如图 3 所示), 依次应用通道和空间注意力模块, 强化了通道和空间两个维度上的缺陷特征, 具体计算过程如下。

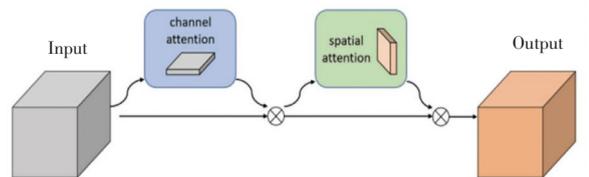


图 3 CBAM 注意力机制结构

Fig. 3 Structure of CBAM attention mechanism

给出输入特征图为 $X \in R^{C \times H \times W}$, CBAM 模块首先通过通道注意力模块 ($F_{CH} \in R^{C \times 1 \times 1}$) 生成通道注意力特征图, 然后将其与输入图像相乘生成中间变

量 X' , 计算过程可由式(2)来表示:

$$X' = F_{CH}(X) \otimes X \quad (2)$$

其中,“ \otimes ”表示加权乘法。

此后通过二维空间注意力模块 ($F_s \in R^{1 \times W \times H}$) 在特征图 X' 的基础上生成空间注意力特征图,再将二维空间注意力模块 F_s 与通道自适应化后的特征图 X' 相乘,得到最终特征加强后的特征图 F' , 计算过程可用式(3)来描述:

$$X'' = F_s(X') \otimes X' \quad (3)$$

在特征提取层中,随着网络的逐渐加深,特征图的尺度会变小,并且会丢失一些信息。因此,本文将 CBAM 注意力机制与卷积层融合,嵌入到特征提取层提取了特征后的特征图输出处,构建 ELAN-CBAM 模块(如图4所示),以强化随着网络深度而逐渐削弱的特征,增强网络的表达能力,提高最终检测的精度。

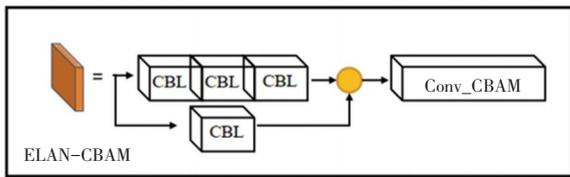


图4 ELAN-CBAM 结构

Fig. 4 Structure of ELAN-CBAM

2.3 引入轻量级特征重组模块

在 YOLOv7 的特征融合层中,使用了最近邻上采样算法进行特征图的上采样,是通过像素之间的空间距离来指导上采样的过程,利用采样点四周相邻最近的像素点的灰度值,但却忽略了其他像素点的影响,无法充分利用特征图中的空间特征,可能会造成输出图像灰度值上的不连续,从而使得图像灰度变化明显区域出现噪声,这会对钢材缺陷这类小目标的检测造成极大的困扰,导致检测精度降低。基于此,本文引入了基于内容感知的轻量级特征重组算子 CARAFE^[19](如图5所示)来代替 YOLOv7-Tiny 特征融合层中的最近邻上采样。CARAFE 算子能够在更大的感受野中聚合上下文信息,摒弃了最近邻上采样算法中使用单一内核采样的方式,采用了基于自适应内容感知的采样方式,其具体计算流程如图5所示。

在经过 CARAFE 模块重建后的特征图中,每个像素都是根据邻域的特征内容进行上采样的,充分利用了上下文信息,让局部区域的相关点信息得到了更多的关注,使其语义信息更加丰富。

2.4 基于归一化 Wasserstein 距离的损失函数

本文所检测的钢材表面缺陷中存在许多小目标缺陷,这对目标检测来说是一个难点,因为其中往往只包含几个像素,因此,提高检测小目标的能力对于提升整个钢材缺陷检测的效果十分重要。

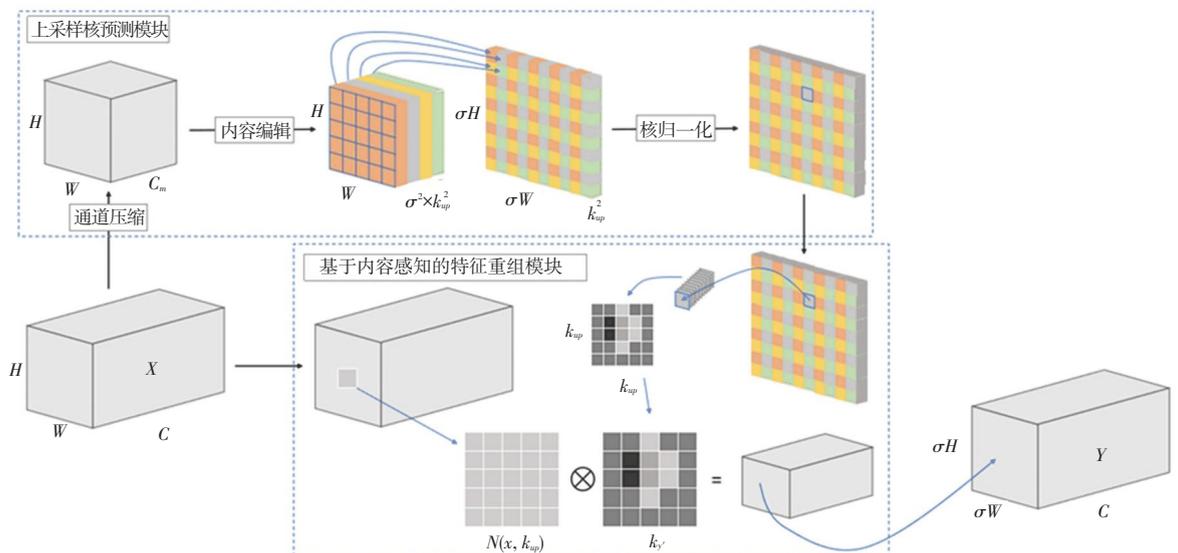


图5 CARAFE 计算流程图

Fig. 5 Calculation flowchart of CARAFE

目前, 各类目标检测算法通常使用基于 IoU (Intersection over Union) 的各类损失函数作为模型优化的指引, 例如 $CIoU$ 、 $SIoU$ 等。但是这些度量标准对小目标的位置偏差敏感度较高, 很容易受小像素分布的影响。预测框 P 中几个像素的偏差就会导致 P 与标注框 GT 之间的重叠部分变化特别大, 导致基于 IoU 的损失函数并不适用于小目标物体的检测。

针对这个问题, 本文采用了一种基于归一化 Wasserstein 距离^[20] (NWD, Normalized Wasserstein Distance) 的损失函数。

首先将建模为二维高斯分布, 然后使用归一化

Wasserstein 距离 (NWD) 的计算标准来计算其相应高斯分布之间的相似性; 接着将 NWD 度量标准嵌入到模型的损失函数中, 取代 YOLOv7-Tiny 原始的 $CIoU$ 损失函数。基于 NWD 的损失函数计算过程如下:

$$L_{NWD} = 1 - NWD(N_p, N_g) \quad (4)$$

其中, N_p 是预测框 P 的高斯分布模型, N_g 是标注框 GT 的高斯分布模型。基于 NWD 的损失函数即使在小目标缺陷检测时像素点偏差的情况下也能够拥有梯度, 进一步提高钢板缺陷目标检测的精度。

2.5 改进后的 YOLOv7-Tiny 模型

本文改进后的 YOLOv7-Tiny 网络结构图如图 6 所示。

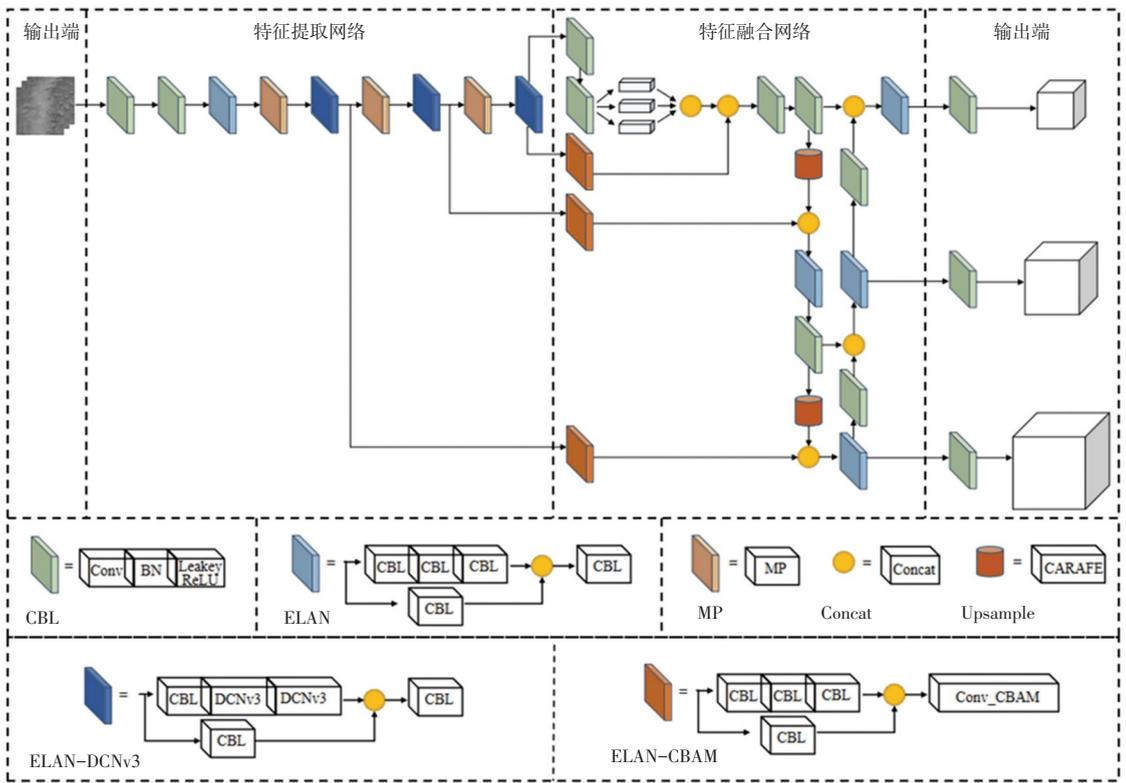


图 6 改进后的 YOLOv7-Tiny 网络结构图

Fig. 6 Structure of improved YOLOv7-Tiny network

3 实验结果与分析

3.1 实验数据集

本文采用了东北大学发布的钢材表面缺陷数据集来验证改进的 YOLOv7-Tiny 算法的有效性。NEU-DET 包含 6 种钢材表面缺陷, 包括裂纹 (crazing)、夹杂 (inclusion)、斑块 (patches)、麻点 (pitted surface)、轧入氧化皮 (rolled in scale) 以及划痕 (scratches)。每种缺陷图片的数量为 300 张, 共 1 800 张, 图像大小均为 $200 * 200$ 。6 类缺陷样例如图 7 所示。

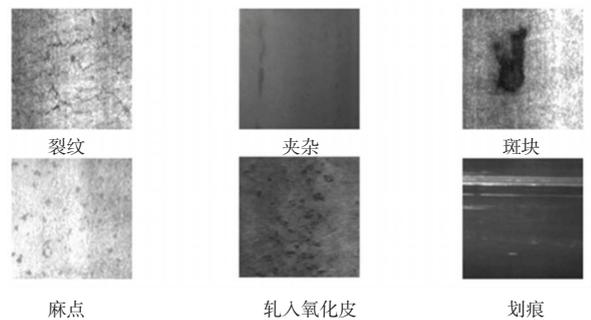


图 7 钢材表面缺陷样例

Fig. 7 Examples of steel surface defects

实验以每类 8 : 1 : 1 的比例随机划分, 选取 1 260 张图片作为训练集, 260 张图片作为验证集, 260 张图片作为测试集。

3.2 实验环境及训练参数

本文提出的钢材表面缺陷检测算法的实验硬件环境以及软件环境见表 1, 训练参数见表 2。

表 1 实验采用的硬件环境与软件环境

Tab. 1 The hardware environment and software environment adopted in the experiment

名称	参数
操作系统	Ubuntu18.04
CPU	双路 NVIDIA 2080Ti
GPU	Intel(R) Core(TM) i7-9700
深度学习环境	Pytorch 1.9.0 + cuda10.2
编译软件	VS Code

表 2 实验采用的训练参数

Tab. 2 The training parameters used in the experiment

参数名称	参数值
epochs	300
batch-size	64
learning-rate	0.01
momentum	0.937
weight-decay	0.000 5
img-size	640 * 640
learning-rate-strategy	cosine annealing

3.3 模型评价指标

本文采用平均准确率 ($mAP@0.5$)、模型大小 (Model Size) 作为算法性能的衡量指标。其中, mAP 是各类缺陷准确率 AP 的均值, AP 则由各类的精准度 (Precision, 下文简称 P) 以及召回 (Recall, 下文简称 R) 计算得到, 其值为 $P-R$ 曲线与坐标轴围成的面积; 研究推得的数学定义公式具体如下:

$$P = \frac{TP}{TP + FP} \quad (5)$$

$$R = \frac{TP}{TP + FN} \quad (6)$$

$$mAP = \frac{\sum_{n=1}^N \int_0^1 P(R) dR}{N} \quad (7)$$

其中, TP 为正确检测到缺陷的样本数量; FP 是错误地将非缺陷目标检测为缺陷目标的数量; FN 是错误地将缺陷目标认作非缺陷的数量。 $mAP@0.5$ 表示 IoU 阈值为 0.5 的 mAP 。

3.4 对比实验设计

为了验证本文提出方案的效果, 在 NEU-DET 数据集中测试了多种目前主流的目标检测算法, 包括 SSD、RetinaNet、YOLOv5s、YOLOv7 以及文献[21-22]。选取平均精度 $mAP@0.5$ 以及网络模型大小作为评价指标。对比实验结果见表 3。通过表 3 可以看出, 本文提出的改进算法平均准确率 $mAP@0.5$ 比 YOLOv7-Tiny 高 4.3%, 模型大小压缩了 11.3%, 实际检测效果对比如图 8 所示。图 8(a)~(c) 中, 从左到右分别是裂纹、夹杂、斑块、裂纹、轧入氧化、划痕。可以看到, 经过改进之后的网络对小目标的缺陷识别能力有着明显的提升, 识别精度也有所提高。相较于 Faster-RCNN、RetinaNet、YOLOv3-Tiny、YOLOv7 算法, 评价指标有着不同程度的领先。相较于文献[21-22], 在检测精度上也有着明显的优势。由此证明, 本文提出的改进算法的性能优于目前领域内的主流算法。

表 3 对比实验结果

Tab. 3 Comparative experimental results

算法	$mAP@0.5/\%$	Model Size/MB
SSD	52.7	102.0
RetinaNet	63.4	139.0
YOLOv5s	72.9	14.5
YOLOv7	75.4	72.0
YOLOv7-Tiny	73.2	12.3
文献[21]	75.1	
文献[22]	74.1	
Improved-YOLOv7-Tiny	77.5	10.9

3.5 消融实验设计

为了进一步验证本文提出的改进方法对模型的有效性, 对改进算法进行消融实验。

消融实验结果分析见表 4。从表 4 的结果可知, 引入 DCNv3 构建 ELAN-DCNv3 模块后, 模型的识别准确率有着明显的提升, $mAP@0.5$ 提升了 1.1%, 同时, 由于 DCNv3 采用了可分离卷积以及分组卷积的思想, 模型的大小也有了一定程度的压缩; 在 ELAN-DCNv3 的基础上进一步引入 CBAM 注意力机制, 加强网络对钢材表面缺陷目标的特征感知, 学习通道与空间两个维度间的相关性, 使 $mAP@0.5$ 提升了 0.6%, 但是参数有了小幅度的增加; 将网络特征融合层原始的最近邻上采样改进为基于内容感知的轻量级特征重建模块 CARAFE 后, 更加丰富了上采样后特征图的语义信息, 进一步增强了利于网络的识别能力, 使得 $mAP@0.5$ 提升了 0.9%; 最后将

网络原始的 $CIoU$ 损失函数改进为更适用于小目标检测的基于 NWD 的损失函数, 将 $mAP@0.5$ 在之前

的基础上再次提升了 1.7%, 达到 77.5%。

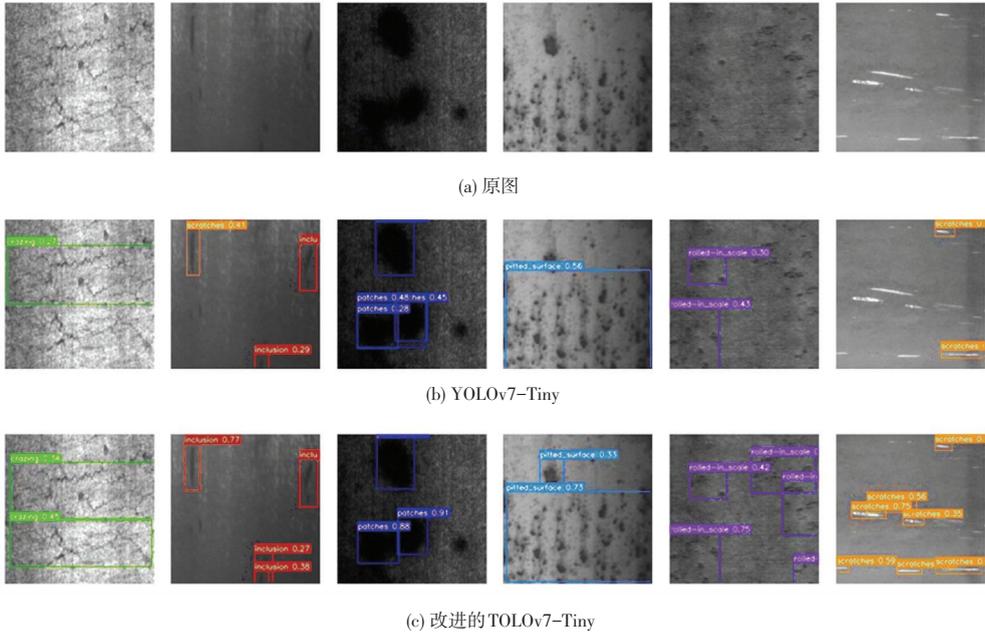


图 8 实际检测效果图

Fig. 8 Actual test effect drawing

表 4 消融实验结果分析

Tab. 4 Analysis of ablation experimental results

方法	ELAN-DCNv3	ELAN-CBAM	CARAfe	NWD	$mAP@0.5/\%$	Model Size/MB
YOLOv7-Tiny	×	×	×	×	73.2	12.3
改进一	√	×	×	×	74.3	10.5
改进二	√	√	×	×	74.9	10.8
改进三	√	√	√	×	75.8	10.9
改进四	√	√	√	√	77.5	10.9

经过消融实验结果分析, 本文确定了最终的改进方案为构建 ELAN-DCNv3、构建 ELAN-CBAM 模块、引入 CARAFE 和采取基于 NWD 的损失函数。YOLOv7-Tiny 改进前后训练效果对比如图 9 所示。由图 9 可以看出, 在经过 300 个 $Epochs$ 的训练迭代

后, 改进后的 YOLOv7-Tiny 算法在 6 类钢材缺陷上的检测精度均有提升, 平均准确率 $mAP@0.5$ 相较于原始 YOLOv7-Tiny 提高 4.3%, 同时模型大小压缩了 11.3%。

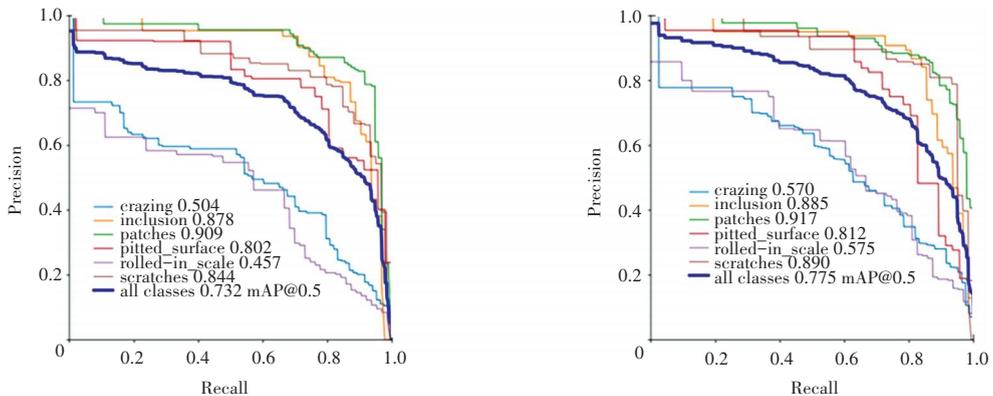


图 9 YOLOv7-Tiny 改进前后训练效果对比图

Fig. 9 Comparison of YOLOv7-Tiny training effect before and after improvement

4 结束语

本文针对工业生产环境需要对钢材表面缺陷进行高效检测的应用背景,提出了一种基于改进YOLOv7-Tiny的钢材表面缺陷识别方法。该方法以YOLOv7-Tiny为基础,改进特征提取层的ELAN结构,引入DCNv3构建ELAN-DCNv3模块,在降低一定模型参数量的同时增强了网络对多种类型、形状缺陷的特征提取能力;增加CBAM注意力机制,提升了网络对钢材表面缺陷的特征感知能力;采用基于内容感知的特征重组算子CARAFE进行特征图的上采样,让局部区域的相关点信息得到了更多的关注,更加丰富了输出特征图的语义信息;此外,针对IoU损失函数对小目标缺陷位置偏移敏感的问题,采取基于NWD的损失函数,使得网络有了更好的小目标检测性能。实验结果表明,本文提出的方法对于钢材表面缺陷有着很好的识别效果,识别准确率达到77.5%,同时将模型压缩了11.3%。本文方法在不增加模型复杂度的前提下,识别准确率方面较YOLOv7-Tiny有着较为明显的提升,在实际应用中展示出了巨大的潜力。未来将继续优化算法,提升算法的准确率,对模型进行剪枝、知识蒸馏,进一步降低模型的推理复杂度,提升模型在边缘计算端的性能,更适应实际生产应用。同时考虑通过数据增强的方式扩充数据集,使得模型的泛化性能进一步提高。

参考文献

- [1] 徐紫琪. 基于机器视觉的螺纹钢表面检测算法研究[D]. 西安: 长安大学, 2022.
- [2] ZHANG Jinhui, WU Bin, SHAO Yanhua. Method human detection based on HOG and CSS characteristics[J]. Computer Knowledge and Technology, 2018, 14(2): 146-148.
- [3] CHANG Henghua, CHAN W C. Automatic registration of remote sensing images based on revised SIFT with trilateral computation and homogeneity enforcement [J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, PP(99): 1-16.
- [4] ZHENG Kai, CHEN Yabo, JIANG Yi, et al. A SVM based ship collision risk assessment algorithm[J]. Ocean Engineering, 2020, 202: 107062.
- [5] GUAN Jian, LIU Jiabei, FENG Pengming, et al. Multiscale deep neural network with two-stage loss for SAR target recognition with small training set [J]. IEEE Geoscience and Remote Sensing Letters, 2021, PP(99): 1-5.
- [6] FENG Pengming, LIN Youlian, GUAN Jian, et al. TOSO: Student's-T distribution aided one-stage orientation target detection in remote sensing images[C]//Proceedings of the 45th International Conference on Acoustics, Speech, and Signal Processing. Barcelona, Spain: IEEE, 2020.
- [7] REN Shaoqing, HE Kaiming, 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.
- [8] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. Hongkong: IEEE, 2017: 2961-2969.
- [9] LI Zouxin, ZHOU Fuqiang. FSSD: Feature fusion single shot multibox detector [EB/OL]. (2017-12-04) [2021-07-19]. <https://arxiv.org/abs/1712.00960>.
- [10] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. arXiv preprint arXiv:1708.02002, 2017.
- [11] TAN Mingxing, PANG Ruoming, LE Q V. EfficientDet: Scalable and efficient object detection[C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE, 2020: 10778-10787.
- [12] LUO Jiaxiang, YANG Zhiyu, LI Shipeng. FPCB surface defect detection: A decoupled two-stage object detection framework[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-11.
- [13] SHI X, ZHOU S, TAI Y, et al. An Improved Faster R-CNN for steel surface defect detection [C]//2022 IEEE 24th International Workshop on Multimedia Signal Processing (MMSP). IEEE, 2022: 1-5.
- [14] KOU Xupeng, LIU Shuaijun, CHENG Kaiqiang, et al. Development of a YOLO-V3-based model for detecting defects on steel strip surface[J]. Measurement, 2021, 182: 109454.
- [15] YANG N, GUO W. Application of improved YOLOv5 model for strip surface defect detection [C]//2022 Global Reliability and Prognostics and Health Management (PHM - Yantai). Yantai, China: IEEE, 2022: 1-5.
- [16] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023: 7464-7475.
- [17] WANG Wenhai, DAI Jifeng, CHEN Zhe, et al. InternImage: Exploring large-scale vision foundation models with deformable convolutions[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023: 14408-14419.
- [18] ZHU Xizhou, HU Han, LIN S, et al. Deformable convnets v2: More deformable, better results[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 9308-9316.
- [19] WANG Jiaqi, CHEN Kai, XU Rui, et al. Carafe: Content-aware reassembly of features [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 3007-3016.
- [20] WANG Jinwang, XU Chang, YANG Wen, et al. A normalized Gaussian Wasserstein distance for tiny object detection[J]. arXiv preprint arXiv:2110.13389, 2021.
- [21] 李刚, 邵瑞, 周鸣乐, 等. 基于注意力的轻量级工业产品缺陷检测网络[J]. 计算机工程, 2023, 49(11): 275-283.
- [22] 曹义亲, 伍铭林, 徐露. 基于改进YOLOv5算法的钢材表面缺陷检测[J]. 图学学报, 2023, 44(2): 335-345.