

DOI:10.19651/j.cnki.emt.2211841

# 基于 CBE-YOLOv5 的钢材表面缺陷检测方法\*

赵林榕 甄国涌 储成群 单彦虎  
(中北大学仪器与电子学院 太原 030051)

**摘要:** 针对钢材表面缺陷种类多,背景干扰强且尺度变化多样导致的检测效率低、精度差的问题,提出了一种钢材表面缺陷检测算法 CBE-YOLOv5。在 YOLOv5 算法基础上进行改进,通过主干采用坐标注意力机制,加强对目标的关注,提高特征提取能力;用 BiFPN 作为特征提取网络,给出有效的特征对应权重,以充分融合不同尺度的特征,并通过 EIOU 来计算模型损失,使模型能更精确的回归。在公开数据集 NEU-DET 上的实验结果表明,CBE-YOLOv5 算法 mAP 为 75.5%,较 YOLOv5 提高了 3.8%,检测速度也高于一些常见的目标检测算法,能够更准确、更快速地检测到钢材表面的缺陷。

**关键词:** 钢材表面缺陷;YOLOv5;注意力机制;特征金字塔;EIOU

**中图分类号:** TP391.4 **文献标识码:** A **国家标准学科分类代码:** 520.6040

## Detection method of steel surface defects based on CBE-YOLOv5

Zhao Linrong Zhen Guoyong Chu Chengqun Shan Yanhu  
(School of Instrument and Electronics, North University of China, Taiyuan 030051, China)

**Abstract:** Aiming at the problems of low detection efficiency and poor accuracy caused by the variety of steel surface defects, strong background interference and various scale changes, this paper proposes a steel surface defect detection algorithm: CBE-YOLOv5. On the basis of YOLOv5 algorithm, improve the ability of feature extraction by using coordinate attention mechanism in the trunk to strengthen the focus on the target; BiFPN is used as a special extraction network, and effective feature corresponding weights are given to fully integrate features of different scales. The model loss is calculated through EIOU, so that the model can be regressed more accurately. The experimental results on the public dataset NEU-DET show that the CBE-YOLOv5 algorithm mAP is 75.5%, 3.8% higher than YOLOv5, and the detection speed is also higher than some common target detection algorithms, it can detect the defects on the steel surface more accurately and quickly.

**Keywords:** steel surface defects;YOLOv5;attention mechanism;feature pyramid;EIOU

## 0 引言

钢铁作为国防领域和建筑行业不可或缺的原材料,对生产质量有着极高的要求。传统的钢材表面缺陷检测主要依靠肉眼观察,严重依赖质检人员的主观判断,存在工作量大,效率低,漏检误检率高等问题,开发出检测准确率高、符合自动化生产的钢材表面缺陷检测方法非常有现实需要。

随着人工智能理论以及深度学习技术的发展,钢材表面的缺陷检测算法逐步得到研究和应用。如以 R-CNN<sup>[1]</sup> 系列为主的基于候选区域的两阶段目标检测算法,以及以 YOLO<sup>[2]</sup> 系列为主的基于回归的单阶段目标检测算法。韩强等<sup>[3]</sup> 在 Faster R-CNN 算法的基础上使用 VGG-16 进行特征提取,级联两个 IOU 阈值不同的检测网络,提出了 FF R-CNN 钢材表

面缺陷检测算法。吴越等<sup>[4]</sup> 用 ROI Align 改进 Faster R-CNN 算法的特征池化模块,解决了 Faster R-CNN 中池化操作产生的像素偏差和空间位置偏差问题,提出的钢板检测模型对小目标的检测精度显著提高。杨莉等<sup>[5]</sup> 引入 GA-RPN 代替 RPN 预测锚点的形状和位置,采用多任务特征金字塔网络进行特征融合,改进后的 Faster R-CNN 提高了检测的准确率和召回率。虽然深度学习的引入使检测准确性有了较大的提升,但是存在网络模型大,识别时间长,缺乏实时性的缺点,无法满足工业现场实时轻量化部署需求。

本文针对这些问题,提出了一种基于 YOLOv5 的改进算法,主要思路如下:1)在主干部分采用坐标注意力机制(coordinate attention, CA),增加模型对钢材表面缺陷的敏

收稿日期:2022-10-25

\* 基金项目:国家自然科学基金重点项目(62131018)、山西省基础研究计划项目(202103021222012)资助

感程度,提高特征提取能力。CA 插入位置对模型的精度有不同程度的影响,文献[6]提出的船舶检测模型讨论了检测器与 CA 融合的 9 种情况,将 CA 插入到 9 个不同位置,9 种模型精度各不相同。由于应用场景不同,具体嵌入网络结构中的哪一部分并没有固定的说法。为了保证模型的轻便性和高效性,本文使用一个 CA 模块,在实验部分评估其位置对模型性能的影响;2)在特征融合阶段使用加权双向特征金字塔结构 BiFPN,实现多尺度特征融合,降低随着网络深度的加深目标信息丢失的概率,充分利用浅层特征

图能够加强对小目标的检测;3)用 EIOU 损失函数对模型进行优化,以提升网络的回归精度以及稳定性。

改进的算法基于 CA、BiFPN、EIOU 模块,以下简称 CBE-YOLOv5。

### 1 YOLOv5 算法介绍

本文研究的算法是基于 YOLOv5S 的 v6.0 版本,它可以分为 4 个部分,分别为 Input、Backbone、Neck 和 Prediction,整体结构如图 1 所示。

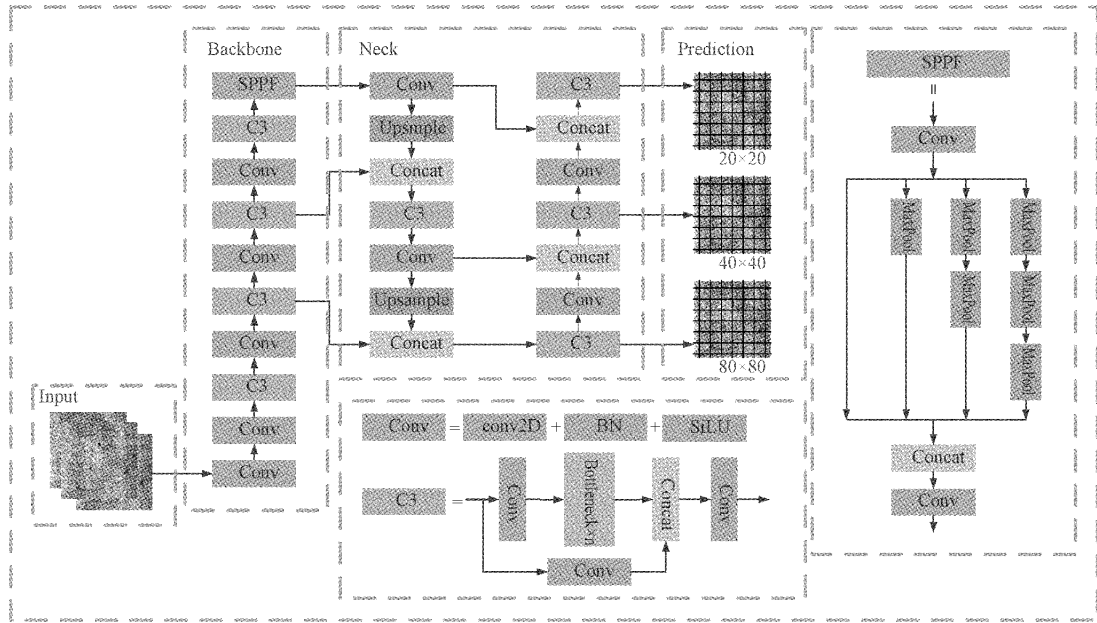


图 1 YOLOv5S 网络结构

输入端对图片进行马赛克数据增强丰富数据集,提高网络的鲁棒性。再通过自适应图片缩放技术将输入的图片缩放至统一大小送进模型。

主干网络主要包括卷积模块、C3 模块和 SPPF 模块,用来对图片进行特征提取,通过多重卷积和池化获得不同下采样倍数的特征图便于缺陷的检测<sup>[7]</sup>。

颈部采用了特征金字塔网络<sup>[8]</sup>(feature pyramid networks, FPN)和路径聚合网络<sup>[9]</sup>(path aggregation networks, PAN)进行特征融合,FPN 自顶向下传达顶层利于分类的强语义特征,PAN 自底向上传达利于定位的强位置信息。

YOLOv5 的预测层采用 CIoU\_Loss 作为边框损失函数,考虑了边框高宽比,使预测框的精度更高,NMS 对检测后处理阶段目标框进行筛选,得到最优的目标检测框。

## 2 模型的优化与改进

### 2.1 CA 注意力机制

注意力机制作为一种仿生视觉机制,已经被证明有助于提高深度神经网络模型的性能,其本质是帮助模型对输入

图像的每个部分赋予不同的权重,对感兴趣的目标信息赋予更大的权重,让模型重点学习,同时抑制无用的背景信息。

坐标注意力机制<sup>[10]</sup>将位置信息嵌入到通道注意中,它不仅考虑跨通道信息,还可以捕获方向感知和位置敏感信息,能够帮助模型更准确的定位和识别感兴趣区域,有效解决了广泛使用的 SE-Net<sup>[11]</sup>、CBAM<sup>[12]</sup>等注意力机制对位置信息不关注导致的模型精度低等问题。CA 主要通过坐标信息嵌入和坐标注意生成两个步骤,获得精确的位置信息来对通道关系和远程依赖进行编码,工作原理流程图 2 所示。

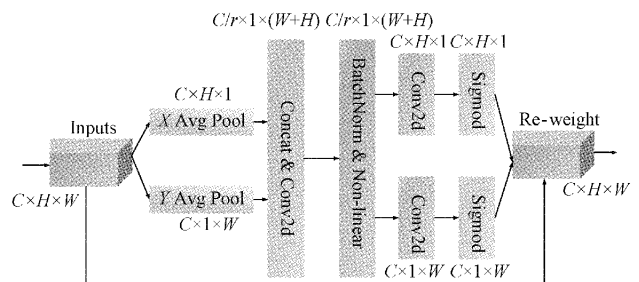


图 2 CA 模块原理

步骤1)坐标信息嵌入:CA将全局池化分解为一对1D特征编码操作,对输入维度为 $C \times H \times W$ 的特征图所有通道进行池化,使用大小为 $(H, 1)$ 的池化核沿宽度方向进行特征提取,使用大小为 $(1, W)$ 的池化核沿高度方向进行特征提取,分别获得维度为 $C \times H \times 1$ 和 $C \times 1 \times W$ 的特征图。在高度为 $h$ 的第 $c$ 个通道的输出可以表示为:

$$Z_c^h(h) = \frac{1}{W} \sum_{0 \leq i \leq W} x_c(h, i) \quad (1)$$

同理,在宽度为 $w$ 的第 $c$ 个通道的输出可以表示为:

$$Z_c^w(w) = \frac{1}{H} \sum_{0 \leq i \leq H} x_c(i, w) \quad (2)$$

以上两个变换,其中一个沿空间方向捕获远程依赖关系,第2个沿另一个空间方向保留精确的位置信息,获得到的两个方向感知特征图能够帮助模型准确定位缺陷目标所在位置。

步骤2)坐标注意生成:将获取到的一对特征图进行拼接,得到维度为 $C \times 1 \times (H+W)$ 的特征图,然后经过共享的 $1 \times 1$ 卷积变换函数 $F_1$ ,得到中间特征图 $f$ ,它的变换公式如下:

$$f = \delta(F_1([Z^h, Z^w])) \quad (3)$$

其中, $f \in R^{\frac{C}{r} \times (H+W)}$ , $[\cdot, \cdot]$ 代表沿空间维度进行拼接操作, $\delta$ 表示非线性激活函数, $r$ 表示缩减因子,卷积操作以大小为 $r$ 的缩减倍数,将通道维从 $C$ 维压缩至 $C/r$ 维。然后,对 $f$ 进行拆分操作,沿空间维度将其拆分为 $f^h \in R^{\frac{C}{r} \times H}$ 和 $f^w \in R^{\frac{C}{r} \times W}$ 两个单独的张量。另外,再用两个 $1 \times 1$ 卷积变换 $F_h$ 和 $F_w$ 分别作用于特征图 $f^h$ 和特征图 $f^w$ ,由于广播机制的存在,能够得到两个与输入 $X$ 具有相同尺寸的张量 $g^h$ 、 $g^w$ ,具体变换过程如下:

$$g^h = \sigma(F_h(f^h)) \quad (4)$$

$$g^w = \sigma(F_w(f^w)) \quad (5)$$

其中, $\sigma$ 为激活函数sigmoid,最后CA注意力机制输出 $Y$ 可以用式(6)进行表示:

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (6)$$

CA模块作为一个灵活轻便、即插即用的模块,可以轻松插入到YOLOv5模型中。

## 2.2 改进的特征金字塔结构

颈部的作用是对主干网络输出的信息进行特征融合。YOLOv5特征融合网络是PANet,融合过程中的信息流是基于临近特征结果,自底向上通路的每一处输入信息都经过FPN网络处理,如图3(b)所示。主干网络提取到的原始图片信息无法到达PAN,缺少原始信息容易导致模型在训练过程中出现偏差。同时,使用上采样或下采样将不同尺寸的特征图变成相同的尺寸再进行简单的相加或拼接,无法充分利用不同尺度间的特征,影响模型的检测精度。加权双向特征金字塔<sup>[13]</sup>BiFPN结构引入了跨尺度连接和加权特征融合环节,解决了PANet存在的不足,其结构如图3(c)所示。

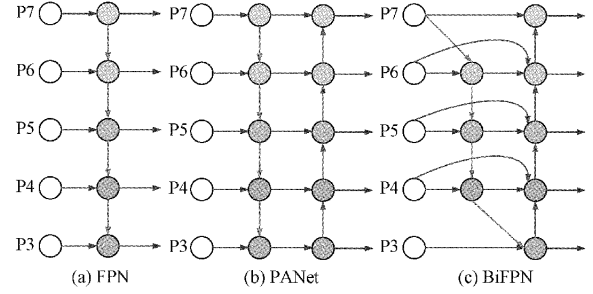


图3 特征融合网络图

1)跨尺度连接:只有一个输入的节点对特征融合网络的贡献很小,删除这种节点对模型性能影响不大,且能够简化网络结构,例如图3(c)中的P3、P7。针对PANet忽略主干网络中的原始信息,导致位置信息等流失的问题,BiFPN在处于同一等级的原始输入节点和输出节点之间添加一条额外边,如图3(c)中紫色线条标出,在没有增加太多计算量的同时融合了更多尺度的特征。

2)加权特征融合:无论是FPN还是PANet都将所有输入对输出的影响视为同等重要,但是不同分辨率的特征图含有的语义信息和位置信息对输出特征的重要程度不同。BiFPN的加权特征融合层对每个输入都赋予一个额外的权重,模型在训练过程中会自主学习权重值。为了避免训练过程不稳定,或占用过多计算成本提出了基于softmax的快速归一化融合方法,公式如下:

$$O = \sum_i \frac{w_i}{\epsilon + \sum_j w_j} \cdot I_i \quad (7)$$

其中, $O$ 是节点输出值, $I_i$ 为节点 $i$ 输入值, $w_i$ 是输入端可学习的权重, $j$ 为输入节点个数总和, $\epsilon=0.0001$ 用来避免数值不稳定。

以其中节点 $P_5$ 为例,BiFPN整个过程可用如下两个公式表示:

$$P_5^{td} = Conv\left(\frac{w_1 \cdot P_5^{in} + w_2 \cdot R(P_6^{td})}{w_1 + w_2 + \epsilon}\right) \quad (8)$$

$$P_5^{out} = Conv\left(\frac{w'_1 \cdot P_5^{in} + w'_2 \cdot P_5^{td} + R(P_4^{out})}{w'_1 + w'_2 + w'_3 + \epsilon}\right) \quad (9)$$

式中: $P_5^{in}$ 表示输入节点, $P_5^{td}$ 表示自顶向下通路中的中间节点, $P_5^{out}$ 表示自底向上通路中的输出节点, $R$ 表示上采样或者下采样操作, $Conv$ 表示卷积操作。

一张待测图中可能包含多个尺度的目标,在检测过程中漏检误检的情况时常发生,用BiFPN模块作为特征融合网络,跨尺度连接能够使网络学习到更多各阶段特征,自主学习权重策略在此基础上对不同重要程度的信息进行有效的抑制或增强,这样一来既有助于小目标的检测,还能降低误检率。

## 2.3 损失函数改进

损失函数为分类损失、置信度损失和边界框回归损失的加权和,YOLOv5选用CIoU作为边界框损失函数,分

类损失和置信度损失用二元交叉熵损失函数计算。

CIOU 将中心点距离和高宽比作为惩罚因子<sup>[14]</sup>, 计算公式如下:

$$L_{CIOU} = 1 - \left( IOU - \frac{\rho^2(b, b^{gt})}{c^2} - \alpha\gamma \right) \quad (10)$$

$$IOU = \frac{A \cap B}{A \cup B} \quad (11)$$

$$\alpha = \frac{\gamma}{1 - IOU + \gamma} \quad (12)$$

$$\gamma = \frac{4}{\pi^2} \left( \arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (13)$$

其中, A 为预测框, B 为真实框;  $b$  和  $b^{gt}$  分别为 A、B 框的中心点,  $\rho$  为中心点之间的距离;  $c$  为 AB 最小外接矩形的对角线长度,  $w$ 、 $h$  和  $w^{gt}$ 、 $h^{gt}$  分别表示 A、B 两个框的宽和高。

虽然 CIOU 考虑了真实框与预测框之间的重叠面积、中心点距离和高宽比, 但是其高宽比是一个相对值, 整体考虑高宽比忽略了预测高宽值各自与其真实值的差异。我们选用 EIOU 作为改进后模型的边界框损失函数, EIOU 在 CIOU 的基础上将高宽比影响因子拆开, 分别计算高度和宽度的差异, 直接使用真实框与预测框高宽之差的最小值代替高宽损失, 能够加快模型收敛速度, 使模型回归精度更高。其公式如下所示, 包含重叠损失、中心距离损失、高宽损失 3 个部分。

$$L_{EIOU} = L_{IOU} + L_{dis} + L_{asp} = 1 - IOU + \frac{\rho^2(b, b^{gt})}{c^2} + \frac{\rho^2(w, w^{gt})}{c_w^2} + \frac{\rho^2(h, h^{gt})}{c_h^2} \quad (14)$$

式中:  $c_w$  和  $c_h$  为 AB 最小外接矩形的宽度和高度。

### 3 实验部分

#### 3.1 实验环境

本文实验环境基于 Ubuntu18.04, CPU 型号为 Intel (R) Xeon (R) Platinum 8350C CPU@2.60 GHz, 运行内存 45 GB, GPU 型号为 RTX3080Ti, 显存 12 GB。该模型基于深度学习框架 PyTorch 1.9.0, 运行环境为 Python 3.8, 集成开发环境为 PyCharm, 使用 CUDA 11.1 进行加速。

#### 3.2 模型训练

模型训练参数设置如表 1 所示。

表 1 模型参数设置

参数名	参数值
初始学习率	0.01
学习率动量	0.973
权重衰减系数	0.000 5
批量大小	8
训练轮数	300

#### 3.3 数据集介绍

本实验使用来自东北大学的公开数据集 NEU-DET<sup>[15]</sup>, 该数据集包含 6 种分辨率  $200 \times 200$  的钢材表面缺陷灰度图, 每种 300 张样本图片。6 种缺陷分别为: 裂纹 (crazing, Cr)、夹杂物 (inclusion, In)、斑块 (patches, Pa)、点蚀面 (pitted surface, PS)、氧化铁皮压入 (rolled-in scale, RS)、划痕 (scratches, Sc), 数据集部分示例如图 4 所示。

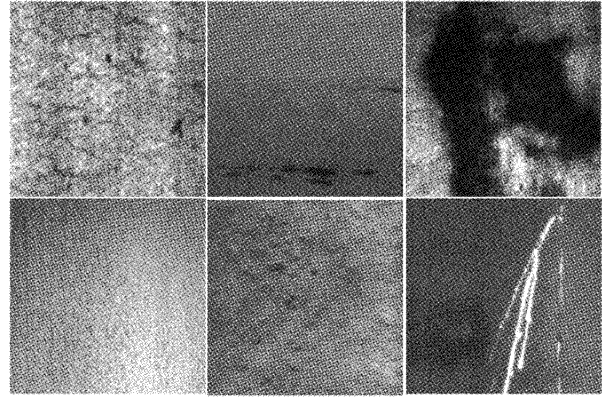
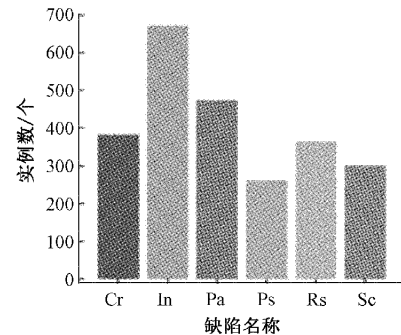
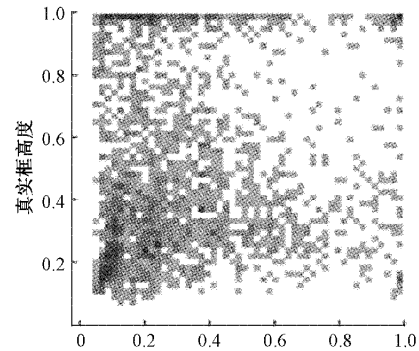


图 4 数据集部分示例

将数据集以 6:2:2 的比例划分为训练集、验证集和测试集, 对 6 种缺陷的分布情况进行可视化分析, 结果如图 5 所示。其中图 5(a) 为每种缺陷类别的实例个数分布情况, 横坐标表示缺陷种类, 纵坐标表示数量; 图 5(b) 为缺



(a) 数据集类别分布



(b) 目标大小分布

图 5 数据集可视化分析

陷实例大小分布图,横坐标表示物体的宽度,纵坐标表示高度。从图4、5中可以发现待测缺陷与背景颜色相近,尺度多样,且小目标占大多数,原始YOLOv5模型辨别起来具有一定困难。

3.4 评价指标

本文从模型精度、检测速度和模型复杂度3个层面评估。模型精度指标:精度(Precision)、召回率(Recall)、平均精度(AP)、平均精度均值(mAP);运行速度指标:每张图片处理所需时间*t*;复杂度指标:权重文件大小(Weight)、参数量(Parameters)和网络层数(Layers)。其中精度、召回率和平均精度均值的计算公式如下:

$$Precision = \frac{TP}{TP + FP} \tag{15}$$

$$Recall = \frac{TP}{TP + FN} \tag{16}$$

$$mAP = \frac{1}{m} \sum_{i=1}^m AP_i \tag{17}$$

式中:TP(true positives)、FP(false positives)和FN(false negatives)分别表示正确检出、漏检、误检,m表示缺陷类别数。

3.5 YOLOv5与CA模块融合位置的对比试验

为了获得性能最佳的CA融合模型,将CA模块插入不同位置做了5组对比实验,YOLOv5作为baseline,实验结果如表2所示,表中CA融合的位置和图6中一一对应。AP1~AP6分别表示裂纹、夹杂物、斑块、点蚀面、氧化铁皮压入、划痕6种缺陷的平均精度。

表2 CA模块在不同位置下的实验结果

序号	CA融合的位置	mAP	AP1	AP2	AP3	AP4	AP5	AP6
1	baseline	0.717	0.386	0.818	0.925	0.786	0.504	0.882
2	backbone(1)	0.719	0.373	0.822	0.921	0.799	0.506	0.893
3	backbone(2)	0.723	0.378	0.827	0.915	0.8	0.517	0.902
4	backbone(3)	0.726	0.392	0.836	0.93	0.797	0.514	0.89
5	backbone(4)	0.738	0.387	0.856	0.923	0.803	0.564	0.897
6	Neck(5)	0.725	0.412	0.839	0.922	0.79	0.516	0.872

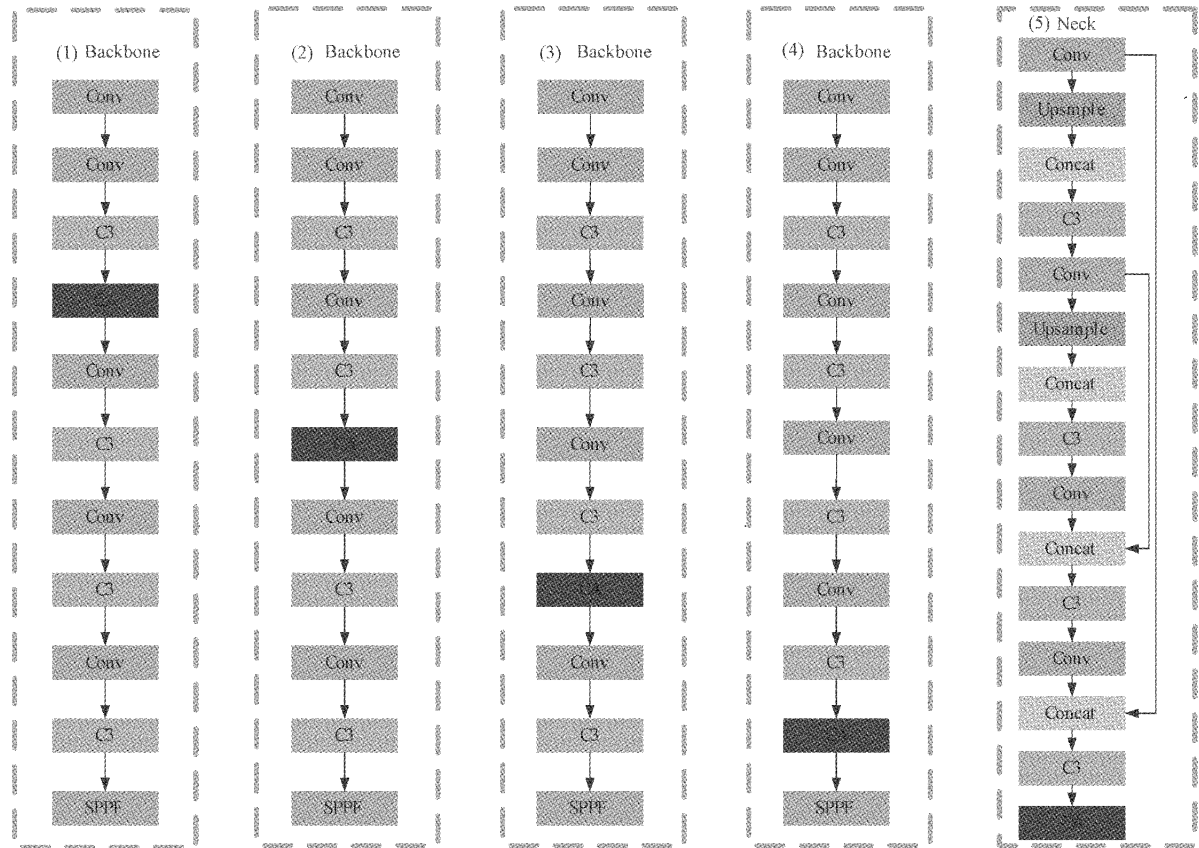


图6 CA融合的位置分布图

由表 2 可以看出,实验 2~6 的 mAP 均比实验 1 的 mAP 高,表明 CA 模块有助于模型检测效果。不同位置对模型的精度有不同的影响,实验 2~实验 5 为分别在主干 4 个 C3 模块之后添加 CA 模块,实验 5 为在 Neck 的最后一个 C3 模块之后添加 CA 模块,最佳添加位置是实验 5, mAP 为 73.8%,相较于 baseline 提高了 2.1%。原因可能是浅层网络的特征图尺寸太大,通道数太少,浅层的 CA 模块学习到的特征信息不够充分,得到的通道权重不能概括具体特征,提取到的权重也不够泛化,导致模型精度没有明显提升。随着网络层数的加深,通道数变多,但是过多的通道数易导致模型过拟合,更重要的是,越靠近分类层注意力机制

对模型分类的结果越敏感,导致检测精度有所下降。所以我们选择实验 5 的融合方式,在整个网络的中间层添加 CA 模块,位于主干网络的后端,输入图片的特征信息经过前面骨干网络的特征提取,已经被模型充分学习,此时的通道权重能够较好的概括整体情况,加入 CA 让模型重点学习感兴趣的区域,最大程度发挥模型对钢材表面缺陷的检测效果。

### 3.6 局部对比实验

将改进特征金字塔结构的模型命名为 B-YOLOv5,改

进损失函数的模型命名为 E-YOLOv5,通过局部对比实验,验证两种改进方法的有效性,具体实验结果如表 3 所示。

表 3 特征提取网络和损失函数改进前后局部对比

序号	实验	特征提取网络	损失函数	mAP
1	YOLOv5	PANet	CIoU	0.717
2	B-YOLOv5	BiFPN	CIoU	0.724
3	E-YOLOv5	PANet	EIoU	0.726

表 3 中实验 1 未进行特征金字塔结构和损失函数的改进,特征提取网络为 PANet,边界框损失函数为 CIoU,此时钢材表面缺陷检测的 mAP 值为 71.7%。实验 2 仅改进特征金字塔结构,使用 BiFPN 替换 v5 的 PANet,mAP 值达到 72.4%。实验 3 仅改进损失函数,使用 EIoU 替换 CIoU,mAP 值达到 72.6%。通过局部对比实验发现,改进的特征金字塔结构和损失函数均能有效提高模型对钢材表面缺陷的检测精度,mAP 值别提高了 0.7%、0.9%。

### 3.7 消融实验

为验证各个新增模块单独的优化效果,依次做了消融实验,×表示未添加该模块,√表示已添加该模块,实验结果如表 4、5 所示。

表 4 消融实验模型复杂度对比

实验	CA	BiFPN	EIoU	Layers	Weight	Parameter
1	×	×	×	213	13.81	7 026 307
2	√	×	×	223	13.87	7 051 955
3	×	√	×	229	15.87	8 092 823
4	×	×	√	213	13.81	7 026 307
5	√	√	×	239	15.95	8 128 503
6	√	√	√	239	15.95	8 128 503

表 5 消融实验模型精度对比

	实验 1	实验 2	实验 3	实验 4	实验 5	实验 6
AP1(Cr)	0.386	0.387	0.367	0.366	0.408	0.447
AP2(In)	0.818	0.856	0.822	0.826	0.854	0.863
AP3(Pa)	0.925	0.923	0.918	0.917	0.94	0.936
AP4(PS)	0.786	0.803	0.795	0.807	0.798	0.79
AP5(RS)	0.504	0.564	0.551	0.544	0.572	0.577
AP6(Sc)	0.882	0.897	0.888	0.896	0.92	0.917
mAP	0.717	0.738	0.724	0.726	0.749	0.755

在表 4 和 5 中,实验 1 用 YOLOv5 作为对照实验,mAP 为 71.7%,对比实验 2、3、4,分别添加 CA、BiFPN、EIoU 模块均使 mAP 提高。实验 2,CA 模块的融合将模型 mAP 提高了 2.1%,难以分辨的夹杂物(In)和氧化铁皮压入(RS)两种缺陷 AP 值分别提高了 3.8%和 6.0%,但是网络层数和参数量只有少量增加,说明 CA 注意力机制能

够在不带来过多计算成本的前提下很好的区分目标与背景,显著提升模型对钢材表面缺陷的关注度。实验 3,BiFPN 模块由于增加了跨尺度链接,避免不了模型复杂度的增加,模型权重增加了 2 MB,但 mAP 提高到 72.4%。存在尺寸差异的夹杂物(In)、点蚀面(PS)、氧化铁皮压入(RS)3 种缺陷的 AP 值得到提升,这表明跨尺度连接有助

于提高模型检测精度。实验4使用EIOU代替原有的CIoU,并没有带来任何模型复杂度的增加,但是mAP值提高到72.6%,说明原有的CIoU损失函数确实存在不足,EIOU将高宽比影响因子拆开,分别计算高度和宽度的差异,能够提高回归精度,进而增强缺陷检测效果。实验6,CBE-YOLOv5模型综合考虑3个模块的影响,mAP从71.7%提高到75.5%,尽管网络层数增加了26层,模型权重增加2MB左右,但是mAP值提高了3.8%,对裂纹、夹杂物、斑块、点蚀面、氧化铁皮压入、划痕6种缺陷AP值分别提高了6.1%、4.5%、1.1%、0.4%、7.3%、3.5%。

### 3.8 纵向对比实验

为验证本文所提模型的有效性,使用目前比较主流的目标检测算法YOLOv3和YOLOv5l作为对比模型,结果如表6所示。

分析表6,综合考虑模型的检测效率和实时性应用的需求,本文所提模型性价比最高。CBE-YOLOv5的mAP值高达75.5%,相较于YOLOv5S、YOLOv5l和YOLOv3分别提高了3.8%、1.8%和3.4%,说明我们的模型具备更强的检测能力。YOLOv5l和YOLOv3的检测速度远不

表6 纵向对比实验结果

模型	mAP	时间/ ms	Weight/ MB	Parameters
YOLOv5S	0.717	3.2	13.8	7 026 307
YOLOv5l	0.737	9.4	88.6	46 135 203
YOLOv3	0.721	11.6	62.2	61 524 355
CBE-YOLOv5	0.755	3.4	15.95	8 128 503

上YOLOv5,它们权重文件极大,模型复杂度与v5也差一个量级,不仅会占用大量的计算资源,而且部署到移动端比较困难,在实际工业应用中很难推广。本文提出的CBE-YOLOv5模型,相比于YOLOv5S模型权重大小和模型复杂度只有小幅提升,不会带来过多的计算量,不仅精度提高了3.8%,处理每张图片花费的时间只比YOLOv5多了0.2ms,在钢铁工业生产过程中依旧满足实时性的需求。

### 3.9 横向对比实验

与现有的钢材表面缺陷检测算法相比,本文提出的算法具有一定的优越性,检测效果对比实验如表7所示。

表7 横向对比试验结果

实验方法	baseline	改进方法	mAP	Parameters/ $10^6$
CBE-YOLOv5(本文)	YOLOv5S	+CA+BiFPN+EIOU	75.5%	8.1
文献[16]	YOLOv5m	+SPP_Res+C3_MHSA+多层融合	74.1%	23.9
文献[17]	YOLOv5l	+BiFPN+TRANS	75.2%	90.8

文献[16]在YOLOv5m的主干引入SPP\_Res和C3\_MHSA结构,颈部进行多层特征融合,改进后的YOLOv5算法在NEU-DET上mAP为74.1%;文献[17]在YOLOv5l的基础上添加Transformer和BiFPN模块,提出的MSPT-YOLO算法在NEU-DET上mAP为75.2%。本文的CBE-YOLOv5基于v5S,v5S与文献[16]的m、文献[17]的l相比,

模型更简单,但mAP分别提高了1.4%和0.3%,参数量也大大减少,因此,本文所提模型能更好的完成钢材表面缺陷检测任务,同时模型更轻便,更具有实际应用价值。

### 3.10 检测效果对比

为了更加直观的比较改进前后模型的检测效果,图7展示了部分效果对比图。

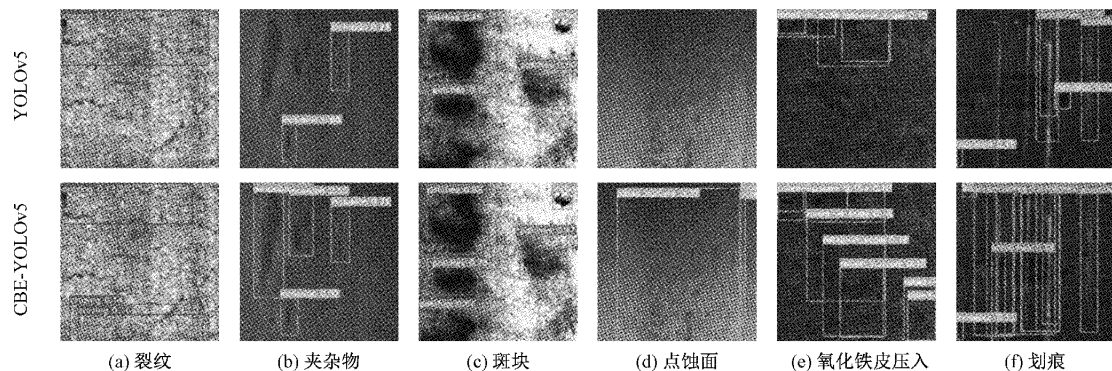


图7 模型检测效果对比

对比发现原模型存在严重的漏检情况,且置信度普遍更低。裂纹、夹杂物、点蚀面和氧化铁皮压入4种缺陷背景干扰强,原模型漏检严重,甚至检测不到点蚀面缺陷,

CBE-YOLOv5模型加入了坐标注意力机制,模型对缺陷的敏感程度增强,漏检情况明显改善。夹杂物和划痕缺陷尺寸变化多样,加入BiFPN模块后能够检测出尺度较小的夹

杂物,证明 BiFPN 模块能够适应缺陷尺寸变化,解决了缺陷尺度变化多样的问题。整体观察图 7,可以直观地发现改进后的模型检测效果明显提升。

#### 4 结 论

钢材缺陷检测在钢材生产过程中具有重要意义,本文基于 YOLOv5 提出了 CBE-YOLOv5 模型,在主干网络的最后一个 C3 模块之后加入 CA 模块,解决了钢材表面缺陷与背景难以区分、缺陷之间易混淆等导致模型检测精度不高的问题。在颈部使用加权双向特征金字塔网络 BiFPN 进行特征融合,使用自主学习权重策略对不同尺度的特征进行融合,既能够增强模型对不同尺度物体的调节能力,还能降低误检率,提高模型检测精度。最后,EIOU 改进原有 CIOU 损失函数中边界框损失函数的不足,提高了模型的回归精度和稳定性。

通过消融实验对 CBE-YOLOv5 添加的模块进行验证,CA、BiFPN、EIOU 模块单独添加时 mAP 分别提高 2.1%、0.7%、0.9%,3 个模块同时加入,模型的检测精度提高 3.8%。

本文将常见的目标检测算法与所提方法进行对比试验,CBE-YOLOv5 检测精度和检测速度综合表现更优,mAP 值为 75.5%,比 YOLOv5s、YOLOv3、YOLOv5l 分别高了 1.8%、3.4%、3.8%,每张图片的处理速度为 3.4 ms,比 YOLOv3、YOLOv5l 分别快了 8.3、6 ms。在实际应用过程中,虽然本文所提算法满足实时性检测要求,但是进一步提高模型的检测精度是未来研究中应该加强的地方。

#### 参考文献

- [1] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [2] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [3] 韩强,张喆,续欣莹,等.基于 FF R-CNN 钢材表面缺陷检测算法[J]. 太原理工大学学报,2021,52(5):754-763.
- [4] 吴越,杨延竹,苏雪龙,等.基于 Faster R-CNN 的钢板表面缺陷检测方法[J]. 东华大学学报(自然科学版),2021,47(3):84-89.
- [5] 杨莉,张亚楠,王婷婷,等.基于改进 Faster R-CNN 的钢材表面缺陷检测方法[J]. 吉林大学学报(信息科学版),2021,39(4):409-415.
- [6] XIE F, LIN B, LIU Y. Research on the coordinate attention mechanism fuse in a YOLOv5 deep learning

detector for the SAR ship detection task[J]. Sensors, 2022, 22(9): 3370.

- [7] 李鑫,汪诚,李彬,等.改进 YOLOv5 的钢材表面缺陷检测算法[J]. 空军工程大学学报(自然科学版),2022, 23(2):26-33.
- [8] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 5936-5944.
- [9] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [10] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2021:13713-13722.
- [11] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018:7132-7141.
- [12] WOO S, PARK J, LEE J Y, et al. CBAM: Convolution block attention module [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018:3-19.
- [13] TAN M X, PANG R M, LE Q V. EfficientDet: Scalable and efficient object detection [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2020: 10778-10787.
- [14] 李衍照,于镭,田金文.基于改进 YOLOv5 的金属焊缝缺陷检测[J]. 电子测量技术,2022,45(19):70-75.
- [15] 宋克臣. NEU surface defect database[DB/OL]. (2019-07-22) [2022-10-25]. [http://faculty.neu.edu.cn/songkechen/zh\\_CN/zdylm/263270/list/index.html](http://faculty.neu.edu.cn/songkechen/zh_CN/zdylm/263270/list/index.html).
- [16] 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.
- [17] GUO Z, WANG C, YANG G, et al. MSFT-YOLO: Improved YOLOv5 based on transformer for detecting defects of steel surface[J]. Sensors, 2022, 22(9): 3467.

#### 作者简介

**赵林熔**, 硕士研究生, 主要研究方向为机器视觉。

E-mail: zzzlr99@163.com

**甄国涌**(通信作者), 教授, 主要研究方向为动态测试技术、机器视觉。

E-mail: zengguoyong@nuc.edu.cn

**储成群**, 讲师, 主要研究方向为动态测试技术、机器视觉。

E-mail: chuchengqun@nuc.edu.cn

**单彦虎**, 讲师, 硕士生导师, 主要研究方向为动态测试技术、机器视觉。

E-mail: shanyanhu@nuc.edu.cn