

戴志新, 王刚, 孙立辉. 基于改进的 YOLOv5s 钢材表面缺陷检测研究[J]. 智能计算机与应用, 2024, 14(10): 79-86. DOI: 10.20169/j.issn.2095-2163.241010

# 基于改进的 YOLOv5s 钢材表面缺陷检测研究

戴志新<sup>1</sup>, 王刚<sup>2</sup>, 孙立辉<sup>1</sup>

(1 吉林化工学院 信息与控制工程学院, 吉林 吉林 132022; 2 吉林交通职业技术学院 机械工程学院, 长春 130015)

**摘要:** 针对在钢材表面缺陷检测任务中对缺陷特征提取能力不足以及特征融合不充分的问题, 本文提出一种基于 YOLOv5s 改进的钢材表面缺陷检测算法 YOLOv5-TBC。首先, 在 Backbone 的核心特征提取模块中引入三重注意力机制 (Triplet Attention), 使 Backbone 更具有适应性和表征能力。其次, 研究引入了加权双向特征金字塔网络 (BiFPN), 提升网络的特征融合, 并引入了 CBAM 注意力机制优化模型对小尺度目标的检测能力。最后, 添加了轻量级上采样算子 CARAFE 用来扩大模型的感受野, 进一步提高对不同大小目标的检测效果。实验结果表明, 改进后的 YOLOv5s 模型在 NEU-DET 数据集上的精确率 (Precision,  $P$ ) 和平均准确率 (mAP) 分别为 74.0% 和 76.6%, 较 YOLOv5s 中  $P$  提升了 5.8%,  $mAP$  提升了 3.0%, 较 YOLOv7 中  $P$  提升了 3.8%,  $mAP$  提升了 2.1%, 证明该网络模型具有更良好的检测性能。

**关键词:** 缺陷检测; YOLOv5; 三重注意力机制; CBAM; 特征融合; CARAFE

中图分类号: TP391.41 文献标志码: A 文章编号: 2095-2163(2024)10-0079-08

## Research on steel surface defect detection based on improved YOLOv5s

DAI Zhixin<sup>1</sup>, WANG Gang<sup>2</sup>, SUN Lihui<sup>1</sup>

(1 School of Information and Control Engineering, Jilin Institute of Chemical Technology, Jilin 132022, Jilin, China;

2 School of Mechanical Engineering, Jilin Communications Polytechnic, Changchun 130015, China)

**Abstract:** This paper proposes an improved steel surface defect detection algorithm YOLOv5-TBC based on YOLOv5s to address the issues of insufficient defect feature extraction capability and inadequate feature fusion in steel surface defect detection tasks. Firstly, a triplet attention mechanism is introduced into the core feature extraction module of the Backbone to enhance its adaptability and representational ability. Secondly, a weighted Bi-directional Feature Pyramid Network (BiFPN) is introduced to improve feature fusion, and a CBAM attention mechanism is introduced to optimize the model's detection capability for small-scale targets. Finally, a lightweight upsampling operator CARAFE is added to enlarge the model's receptive field, further improving the detection performance for targets of different sizes. Experimental results demonstrate that the improved YOLOv5s model achieves a precision ( $P$ ) of 74.0% and a mean Average Precision ( $mAP$ ) of 76.6% on the NEU-DET dataset, which is an increase of 5.8% in  $P$  and 3.0% in  $mAP$  compared to YOLOv5s, and an increase of 3.8% in  $P$  and 2.1% in  $mAP$  compared to YOLOv7, proving that the network model has better detection performance.

**Key words:** defect detection; YOLOv5; triplet attention mechanism; CBAM; feature fusion; CARAFE

## 0 引言

钢材作为制造业的基础材料, 具备极高的实用价值, 并在各个工业领域广泛应用, 包括航空航天、深海探测工程和车辆制造等。钢材凭借其卓越的特性在这些领域中发挥着至关重要的作用。最初, 对钢材表面缺陷的检测主要依赖人工目检。然而, 这

种方法存在着昂贵的成本, 更为重要的是容易出现大量误检和漏检, 在实际的工业环境中并不切实可行。近年来, 深度学习的应用和改进在制造业中逐渐赢得青睐, 与机器视觉的结合展现出强大的优势<sup>[1-4]</sup>, 采用深度学习缺陷检测的方法, 可以及早地发现工艺流程上的缺陷, 避免不必要的成本损失, 因此将深度学习应用到钢材的表面缺陷检测具有重大

**基金项目:** 吉林省自然科学基金(20220101138JC, YDZJ202301ZYTS420); 吉林省教育厅科研项目(JJKH20241142KJ)。

**作者简介:** 戴志新(1999-), 男, 硕士研究生, 主要研究方向: 计算机视觉; 孙立辉(1976-), 男, 教授, 主要研究方向: 多智能体编队控制, 智能控制系统。

**通讯作者:** 王刚(1981-), 男, 博士, 教授, 主要研究方向: 多智能体编队控制, 图像与视频处理, 机器人与系统控制。Email: Wgzh198081@163.com

收稿日期: 2024-04-23

哈尔滨工业大学主办 ◆ 学术研究与应用

意义。

在钢材表面缺陷检测领域, YOLO<sup>[5]</sup> 作为主要的研究算法之一, 研究者们已经展开了广泛研究工作。Li 等学者<sup>[6]</sup> 针对缺陷漏检问题, 使用 ResNet-50 结构进行特征提取, 成功提高了缺陷的检出率。徐镭等学者<sup>[7]</sup> 基于 YOLOv3 网络进行改进, 将特征提取网络替换为轻量级 MobileNet, 并引入 Inception 结构以减少模型参数量, 从而提高了带钢表面缺陷整体的检测精度。李维刚等学者<sup>[8]</sup> 提出了一种创新的 YOLOv3 算法模型, 该模型通过加权聚类算法对先验框的参数进行优化, 从而增强了先验与特征图层的匹配程度。接着, 该团队还对 YOLOv3 算法的网络架构进行了调整, 以增强网络在检测带钢表面缺陷方面的准确性。Dong 等学者<sup>[9]</sup> 设计了一个并行混合注意力机制模块, 并将该模块应用到 YOLOv5 的骨干网络中, 以提高网络对缺陷特征的关注。Yang 等学者<sup>[10]</sup> 率先采用同态滤波预处理钢材缺陷图像, 随后利用引入注意力机制的改良 YOLOv5 模型对缺陷进行检测, 实现了精度的显著提升。Wang 等学者<sup>[11]</sup> 优化了 YOLOv5 算法, 加入了多尺度特征融合模块, 从而强化了模型对跨尺度语义特征的捕捉能力。Guo 等学者<sup>[12]</sup> 提出的改进 Msft-YOLO 模型, 基于 YOLOv5 模型整合了 Transformer 模块并实施数据增强, 有效提升了钢材

缺陷检测的精确性。

然而, 学者们的上述研究对于钢材表面缺陷检测任务仍存在很多问题, 如缺乏对特征图的特征提取能力, 检测精度有待提高等。针对这些问题, 本文提出了一种基于 YOLOv5s 改进的钢材表面缺陷检测算法 YOLOv5s-TBC。为解决特征提取能力不足的问题, 首先在 backbone 通过引入三重注意力机制 (Triple Attention, TA), 加强对小目标的特征提取能力。为解决特征融合不充分的问题, 引入加权双向特征金字塔网络 (Bi-directional Feature Pyramid Network, BiFPN) 增强图像的高层语义信息和低层特征信息融合性能, 同时添加了通道注意力与空间注意力融合的注意力模块 (Channel Block Attention Module, CBAM), 解决融合后细节信息丢失问题。另外, 引入轻量级上采样算子 CARAFE 用来扩大模型的感受野, 提高对小目标的检测精度, 实现对钢材表面缺陷的精确检测。

## 1 YOLOv5s 模型结构

目标检测是计算机视觉中的重要任务, 而 YOLOv5s (You Only Look Once version 5 small) 是 YOLOv5 系列中的一款轻量级模型, 专注于在保持高准确性的同时提供更高的实时性能。YOLOv5s 的网络结构如图 1 所示。

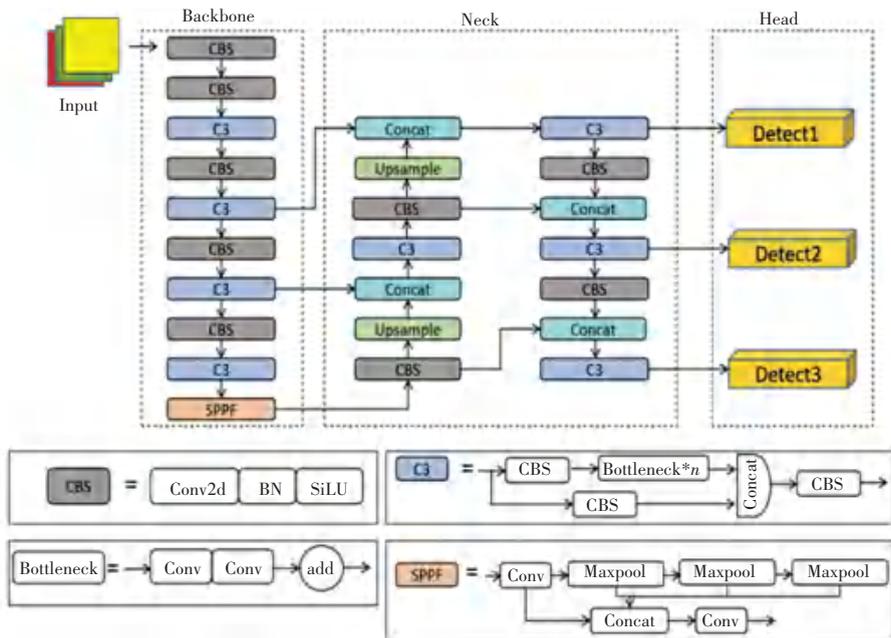


图 1 YOLOv5s 的网络结构图

Fig. 1 Network diagram of YOLOv5s

输入层部分进行图像预处理, 将输入图像的大小调整为  $640 \times 640$ , 并应用了 Mosaic 数据增强方

法<sup>[9]</sup> 等技术来处理输入数据。骨干网络层部分采用了 CSPDarknet53 作为骨干网络。CSPDarknet53

是一个基于 Darknet 框架的深度卷积神经网络, 具有较大的感受野和强大的特征提取能力, 同时快速空间金字塔池化 (Spatial Pyramid Pooling - Fast, SPPF) 模块能够快速地提取不同细粒度的图像特征。特征融合层采用特征金字塔网络 (FPN) 与金字塔注意力网络 (PAN) 结合的结构, FPN 通过将来自不同层次的特征图融合在一起, 有效地捕捉了图像中不同尺度的信息。这种多尺度的特征表示有助于模型更全面地理解图像内容, 提高目标检测的精度。而 PAN 则是将底层特征图的信息传递给深层并进行融合, 增强了不同尺度特征图的定位能力, 使得算法检测的鲁棒性和准确性大大提升。检测层是 YOLOv5s 的输出层, 总共有 3 个输出层, 输出图像大小为  $80 \times 80$ 、 $40 \times 40$  和  $20 \times 20$  三个大小的尺寸, 该模块主要负责生成目标的边界框和类别概率。YOLOv5s 采用了轻量级的检测头部设计, 以保证模

型的高效性能。检测头部输出目标的位置坐标、类别概率等信息, 为后续的目标框筛选和后处理提供了基础。

## 2 YOLOv5s 模型改进

### 2.1 特征提取网络的改进

#### 2.1.1 Triplet 注意力机制

注意力机制<sup>[13]</sup>在近年来已经被很多专家证明其能够在通道或空间位置之间建立相互依赖的关系, 因此在各种计算机视觉任务中受到广泛关注, 并得到了成功应用。三重注意力机制 (Triplet Attention)<sup>[14]</sup>是一种通过使用在 3 个不同通道结构上进行跨维度的交互来计算注意力权重的方法, 其结构设计如图 2 所示。该注意力机制既轻量化又高效, 可以作为附加模块插入到经典的主干网络中。

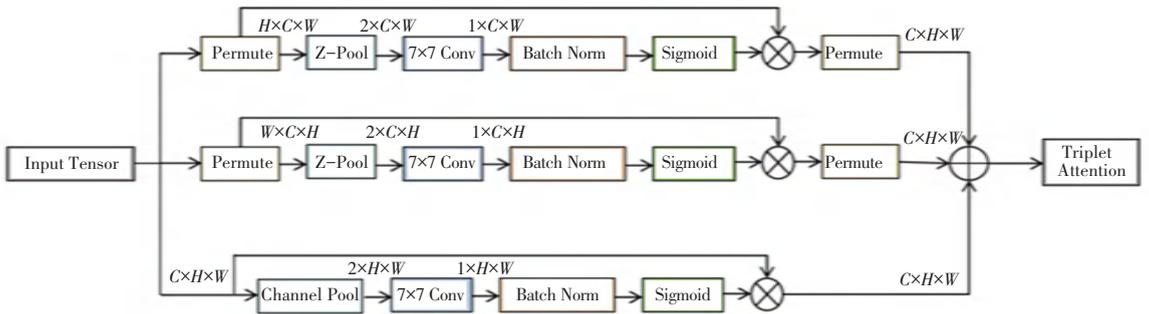


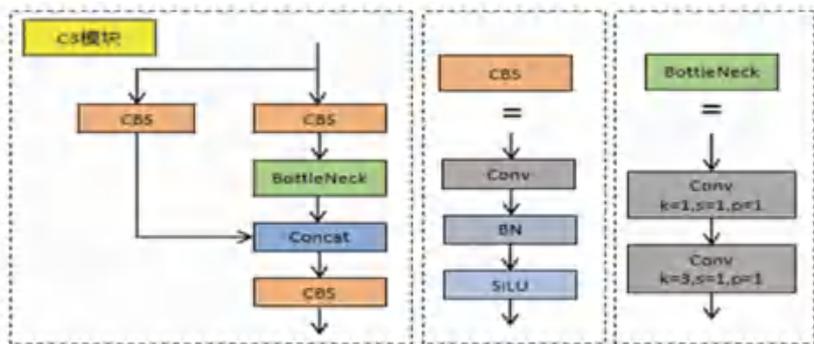
图 2 三重注意力机制结构图

Fig. 2 Structure diagram of the triple attention mechanism

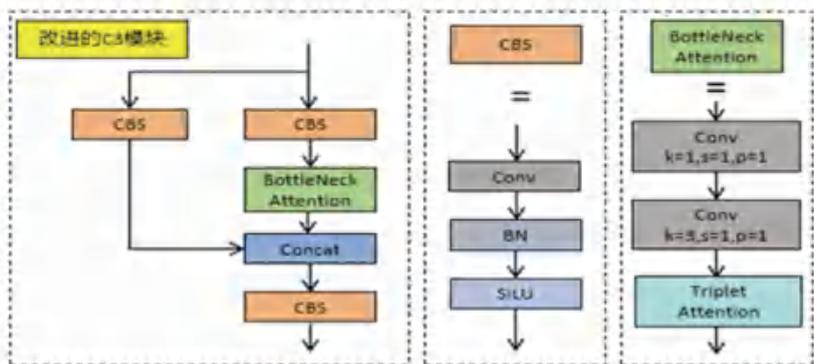
图 2 中, 三重注意力机制结构共有 3 个分支。最顶部的分支利用交互通道  $C$  和宽度  $W$  之间的空间注意力来构建。首先, 输入特征经过 Permute 操作以调整维度, 然后进行 Z-Pool 池化, 使得特征转变为  $2 \times C \times W$ 。接着, 经过一个  $7 \times 7$  卷积层、批量归一化层和 Sigmoid 激活函数处理, 最终再次通过 Permute 调整维度, 将特征大小重新转变为  $C \times H \times W$ 。中间的分支则利用了交互通道  $C$  和高度  $H$  之间的注意力权重。输入特征同样经过 Permute 操作后, 特征变为  $W \times H \times C$ , 然后进行与顶层分支同样的卷积池化操作, 最后再次通过 Permute 调整维度, 将特征大小重新转变为  $C \times H \times W$ 。底部的最后一个分支关注于交互宽度  $W$  和高度  $H$  之间的空间依赖关系。与前 2 层不同的是, 该层输入首先经过 Channel Pool 层, 然后再经过  $7 \times 7$  卷积层和批量归一化层处理, 最后再通过使用 Sigmoid 激活函数生成一个空间注意力权重。最终, 3 个分支生成的权重通过简单平均聚合。

#### 2.1.2 C3 模块的改进

在 YOLOv5s 中, C3 模块通常位于模型的 backbone 中。主干网络是负责从输入图像中提取特征的部分, C3 模块的加入有助于提高特征的代表能力。首先将特征图分为左右两部分, 其中左半部分经过一个卷积模块后等待右半部分处理完成, 然后进行融合。右半部分先经过一个相同的卷积模块, 然后通过  $n$  个瓶颈 (bottleneck) 模块, 最终与左半部分融合。整个过程可概括分为 2 路: 左路保持相对简单, 右路通过多个瓶颈模块进行降维、升维和特征提取, 最后再合并为一路。融合后, 经过一个卷积模块输出特征图。总体而言, C3 模块的特征图经过特征提取, 其尺寸和通道数都保持不变。虽然 C3 模块参数量小, 推理速度也更快, 但是其在小目标检测效果上的表现不尽如人意, 这可能是由于模块内的结构简单导致的结果, 因此本文对图 3 中 (a) 所示的 C3 模块进行了改进, 改进后的 C3\_TA 模块结构图如图 3 (b) 所示。



(a) C3模块结构图



(b) C3\_TA 模块结构图

图3 C3模块改进前后结构图

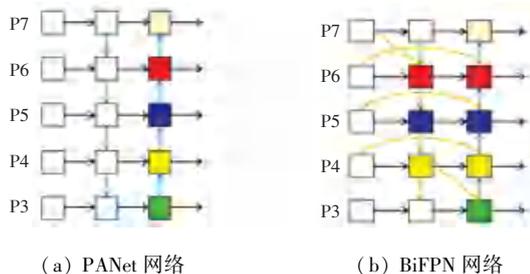
Fig. 3 Structure of C3 module before and after improvement

## 2.2 特征融合网络的改进

### 2.2.1 BiFPN 特征融合网络

随着特征提取网络层数的不断加深,很多重要的特征信息会丢失,因此在缺陷检测任务中使用特征融合的主要目的是提高模型对不同尺度目标的感知能力,以提高模型的检测性能。缺陷检测任务中的物体尺度各异,有些目标可能很小,而另一些可能很大。为了有效地检测和识别这些不同尺度的目标,特征融合成为一种关键的技术。常见的特征融合网络有:特征金字塔网络(FPN)<sup>[15]</sup>、路径聚合网络(PANet)<sup>[16]</sup>、高分辨率特征金字塔网络(HRFPN)<sup>[17]</sup>以及双向特征金字塔网络(BiFPN)<sup>[18]</sup>等。其中,FPN是一种经典的特征金字塔网络,通过自顶向下的信息传递,从不同层级的特征金字塔中提取语义信息。特征融合网络如图4所示。图4(a)中,PANet通过在FPN的基础上引入自底向上的信息传递,实现了更全面的特征融合,传统的YOLOv5模型在特征融合部分使用了PANet,但该融合方式较为简单。而HRFPN专注于解决目标检测中对小目标的挑战,通过保持高分辨率的特征图,提高了对小目标的感知能力。BiFPN旨在解决传统的FPN中存在的信息流单一、信息损失等问题,通过

引入双向信息流,改进了特征金字塔的设计,为了使模型能够对特征进行更好的融合,本文引用了图4(b)中的BiFPN特征融合网络代替了YOLOv5s中原有的PANet特征融合网络。



(a) PANet 网络

(b) BiFPN 网络

图4 特征融合网络结构

Fig. 4 Feature fusion network structure

BiFPN网络是一种用于目标检测任务的特征融合网络,在EfficientDet等一些目标检测模型中取得了很好的性能。与传统的单向特征金字塔不同,BiFPN引入了双向信息流,使得特征可以在不同层之间自由流动。这样可以更好地捕捉不同尺度的信息,有助于提高模型对小目标和大目标的检测性能。

### 2.2.2 CBAM 注意力机制

对于特征融合后的结果再进行特征提取时,网络采用相同的方式对不同重要程度的特征进行加

权,这就给小尺度的缺陷检测目标的提取带来不小的难度,因此本文在 Neck 层引入 CBAM (Channel Block Attention Module) 注意力机制<sup>[19]</sup>,用来解决网络对小目标特征信息的提取较难的问题。

CBAM 注意力机制如图 5 所示。CBAM 注意力机制首先通过通道注意力模块 (Channel Attention Module, CAM) 对每个通道来计算全局平均池化。然后,通过全连接层学习通道间的权重,得到每个通道的重要性权重。最后,对每个通道进行加权,得到经过通道注意力调整后的特征图。再将 CAM 模块生成的特征图送入到空间注意力模块 (Spatial Attention Module, SAM),针对每个空间位置,计算每个通道的最大值,得到每个空间位置的最大值。接着,通过全连接层学习空间位置间的权重,得到每个空间位置的重要性权重,再对每个空间位置进行加权,得到经过空间注意力调整后的特征图。CBAM 最终的输出是将 2 个模块的输出按元素相乘得到的最终特征图。

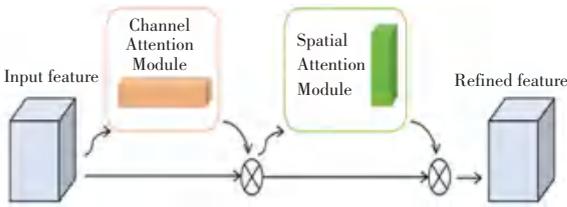


图 5 CBAM 注意力机制

Fig. 5 Attention mechanism of CBAM

### 2.2.3 CARAFE 上采样

在深度学习中,上采样经常用于恢复特征图的

分辨率,以利于更好地理解 and 定位目标。YOLOv5s 常用的是最邻近插值法或者双线性插值法作为上采样方式,这 2 种上采样方式的感受野都很小,并没有利用到特征图中丰富的语义信息。本文引入对的 CARAFE (Channel Attention ReLU Upsampling) 是由 Wang 等学者<sup>[20]</sup>提出的一种用于上采样的神经网络操作,旨在提高上采样的性能。CARAFE 的优势在于能够更好地捕获上采样过程中的细节信息,并且通过通道注意力机制能够更加灵活地调整各个通道的重要性,从而提高网络的性能,这对于需要高质量上采样的任务,如语义分割和目标检测等,是非常有益的。

CARAFE 由上采样核预测模块和特征重组模块两部分组成。其中,上采样核预测模块的作用是动态地学习并进行预测,上采样核是在空间上对特征进行插值的关键部分。CARAFE 通过训练可学习的参数,使模型能够自适应地生成不同位置、不同分辨率上采样核,这种动态生成的上采样核可以更好地适应不同任务和不同输入情境,提高上采样的灵活性。特征重组模块的目的是通过内容感知的方式对特征进行重组。CARAFE 引入了通道注意力机制,允许模型在上采样的过程中关注和调整不同通道的重要性,这有助于更好地捕捉特征图中的有用信息。

### 2.3 YOLOv5-TBC 算法

本文提出的 YOLOv5-TBC 算法是在 YOLOv5-s 的基础上进行模型改进,改进模型主要由 Backbone 和 Neck 两个部分构成。具体改进后的整体结构框架如图 6 所示。

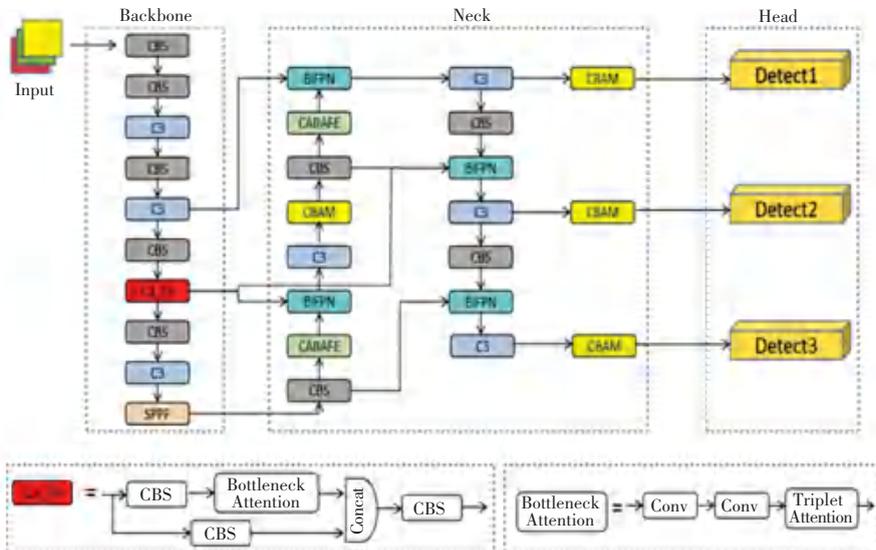


图 6 YOLOv5-TBC 算法整体结构框图

Fig. 6 Block diagram of the overall structure of the YOLOv5-TBC algorithm

首先图像通过输入端进入到 Backbone 中,在 Backbone 的第七层的 C3 模块中引入三重注意力机制形成新的 C3\_TA 模块,使网络更集中地关注到当前任务重要的特征,这有助于提高网络对输入数据的表达能力,使其更好地捕捉关键信息。然后,将原来 Neck 网络中的 PANet 替换成 BiFPN 进行多尺度融合,同时在 Neck 网络中的每个 C3 模块后添加一个 CBAM 注意力机制,使模型对小目标的特征在通道和空间上引入注意力机制,提高模型的鲁棒性。最后,将 CARAFE 引入到 Neck 网络中进行上采样,获得更广阔的感受野信息,提高模型的计算效率,进一步提升算法的检测性能。

### 3 实验结果与分析

#### 3.1 数据集

本文实验所采用的是东北大学的公开数据集 NEU-DET<sup>[21]</sup>,该数据集共包含裂纹(Crazing, Cr)、内含物(Inclusion, In)、氧化轧皮(Rolled-in\_Scale, RS)、划痕(Scratches, Sc)、点蚀面(Pitted\_Surface, PS)和斑块(Patches, Pa)共6个种类的钢材表面缺陷,每一类各有300张灰度图像,共计1800张图片。该数据集提供了缺陷目标的位置和类别信息,并将这些标注信息保存在txt文件中。将数据集内的1800张灰度图片按照7.0:1.5:1.5的比例划分为训练集、验证集和测试集,即随机抽取训练集1260张,验证集和测试集各270张。对模型进行训练并验证YOLOv5-TBC算法的有效性。NEU-DET数据集部分图片如图7所示。

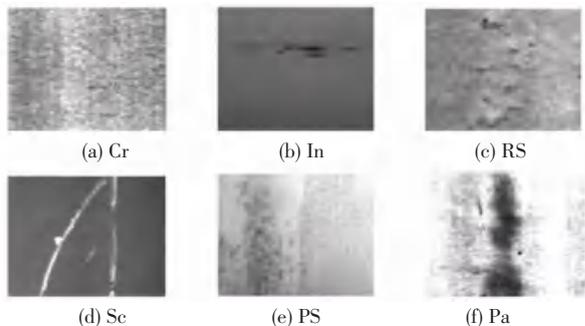


图7 NEU-DET数据集部分图片

Fig. 7 Part of the images of the NEU-DET dataset

#### 3.2 配置环境及评价标准

本文在Windows操作系统上,显卡为NVIDIA GEFORCE RTX3050,用PyCharm软件进行实验分析,软件环境基于PyTorch 1.12.1搭建,Python版本为3.9,CUDA版本为11.3。

本文主要使用精确率(Precision,  $P$ )、召回率

(Recall,  $R$ )和平均精度( $mAP$ )三个指标对模型进行评估。其中,精确率表示被模型正确检测为目标的数量与所有模型判定为目标的本样本数量的比例。召回率表示模型成功检测为目标的数量与所有实际目标的样本数量的比例。而平均精度则结合了精确率-召回率曲线(Precision-Recall Curve)下的面积,用于评估模型在多个类别上的平均表现, $mAP$ 经常被用作评价目标检测模型的一个重要标准,是当前主流的评价指标。 $P, R, AP, mAP$ 的数学计算公式具体如下:

$$P = \frac{TP}{TP + FP} \quad (1)$$

$$R = \frac{TP}{TP + FN} \quad (2)$$

$$AP = \sum_{i=1}^n P(i) \Delta R(i) = \int_0^1 P(R) dR \quad (3)$$

$$mAP = \frac{\sum_{i=1}^n AP(i)}{n} \quad (4)$$

其中, $TP$ 表示真正样本的数量; $FP$ 表示假正样本的数量; $FN$ 表示假负样本的数量; $AP(i)$ 表示第*i*类样本的准确率。

#### 3.3 消融实验结果分析

基于本文提出的4种改进方向,得到本研究的改进模型YOLOv5-TBC,为了验证改进后的网络检测性能是否有所提升,本文一共设置了5组消融实验,每组都是在相同网络环境和参数设置下对NEU-DET数据集进行的消融实验。实验结果见表1。

表1 消融实验结果

Table 1 Ablation test results				%
网络模型	$P$	$R$	$mAP @ 0.5$	
YOLOv5s	68.5	74.0	73.6	
YOLOv5s+A	71.6	69.9	73.5	
YOLOv5s+B	69.8	73.2	73.7	
YOLOv5s+C	69.0	<b>75.0</b>	74.1	
YOLOv5s+D	70.2	74.3	75.3	
YOLOv5s+A+B+C+D	<b>74.3</b>	74.1	<b>76.6</b>	

表1中,A表示在Backbone中的第七层的C3模块中引入三重注意力机制TA,B表示在Neck层的C3模块后加入CBAM注意力模块,C表示使用BiFPN作为特征融合网络,D表示引入CARAFE上采样模块,而本文提出的YOLOv5-TBC算法则为YOLOv5s+A+B+C+D。实验结果显示各模块的改进

都分别对网络起到不同程度的提升效果,证明了各模块的有效性。特别是模型在替换特征融合网络为 BiFPN 和引入 CARAFE 上采样时表现最为突出,无论是在检测精度,还是召回率上都有提升,说明 BiFPN 能够快速进行多尺度融合, CARAFE 上采样能够更好地捕获上采样过程中的细节信息,提高了模型的检测精度;在增加 TA 注意力机制和 CBAM 注意力机制后原始模型在精确率方面有小幅度的上升。本文提出的模型 YOLOv5s-TBC,在  $mAP$  值上较原始 YOLOv5s 模型提升了 3.0%,精确率提升了 5.8%,能够更好地完成钢材表面缺陷检测任务。

### 3.4 对比实验结果分析

为了进一步验证本文算法的可行性,在相同实验环境和数据集下,与其他主流的单阶段检测模型进行比较,对比结果见表 2。首先对比的模型有 YOLOv3、YOLOv4、YOLOv5s、YOLOv7 和 YOLOv7-tiny 五个模型,其中本文提出的模型的  $mAP@0.5$  分别提升了 3.1%、6.6%、3.0%、2.1% 和 5.5%;召回率  $R$  分别提升了 3.0%、1.7%、0.1%、2.7% 和 4.6%;精确率  $P$  分别提升了 5.5%、5.3%、5.8%、4.1% 和 6.9%。从对比结果来看,本文提出的 YOLOv5s-TBC 模型相较于其他模型在检测性能上

有较大的优势,证实了本文提出的改进措施可以有效地提高模型的检测能力。

表 2 对比实验结果

网络模型	$P$	$R$	$mAP@0.5$
YOLOv3	68.8	71.1	73.5
YOLOv4	69.0	72.4	70.3
YOLOv5s	68.5	74.0	73.6
YOLOv7	70.2	71.4	74.5
YOLOv7-tiny	67.6	69.5	71.1
Ours	<b>74.3</b>	<b>74.1</b>	<b>76.6</b>

### 3.5 缺陷检测效果

为了更直观有效地对比模型改进前后的检测效果,将本文改进后模型的部分检测结果和 YOLOv5s 模型检测结果进行比较,部分检测结果如图 8 所示。由图 8 可知,原 YOLOv5s 的检测精度不高,并对一些小目标缺陷、如氧化轧皮类 (Rolled-in Scale) 存在较为明显的漏检问题。而改进后的模型改善了原始 YOLOv5s 模型检测时存在的漏检问题,得到的整体检测精度明显优于原始 YOLOv5s 模型,验证了本文提出的 YOLOv5s-TBC 模型的有效性。

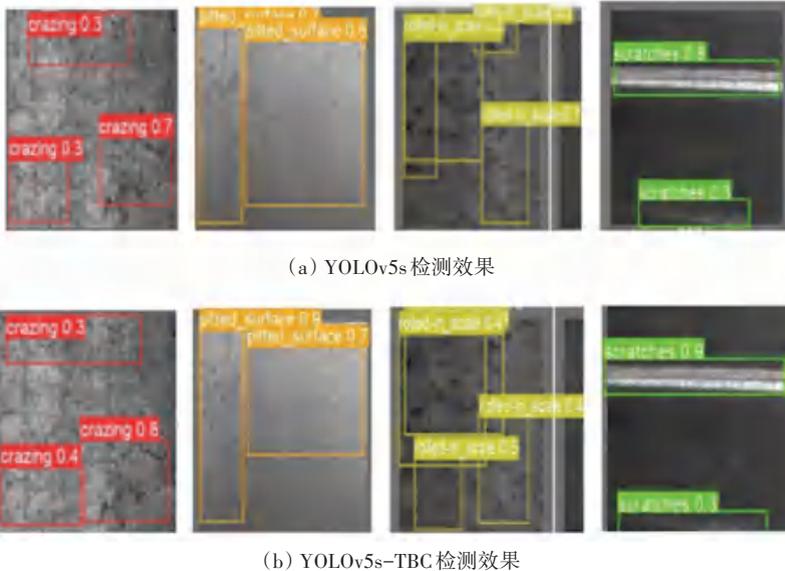


图 8 YOLOv5s 模型改进前后部分检测结果对比图

Fig. 8 Comparison of the detection results of the YOLOv5s model before and after the improvement

## 4 结束语

本文通过改进 YOLOv5s 算法提出一种钢材表面缺陷检测模型。通过引入三重注意力机制,增强模型对小目标特征的提取能力;引入 BiFPN 特征融合网络、CARAFE 上采样模块和 CBAM 注意力机

制,增强模型的特征融合能力,使得模型具有更好的提取目标特征的能力。使用公开的 NEU-DET 数据集进行实验验证,与其他主流算法进行了比较,证明了本文提出的算法的有效先进性。

本文提出的算法仍存在一些问题,比如网络计算参数量大,计算耗时较长。未来,将继续研究模型

的轻量化改进,在提升模型的检测速度及检测精度方面做更深入的探索讨论。

## 参考文献

- [1] 李少波,杨静,王铮,等. 缺陷检测技术的发展与应用研究综述[J]. 自动化学报,2020,46(11):2319-2336.
- [2] 袁洪强,杜国锋,余泽禹,等. 基于轻量神经网络的钢材表面缺陷快速识别[J]. 科学技术与工程,2021,21(34):14651-14656.
- [3] 李翰超,古阳,于亚婷. 基于涡流无损检测技术的钢材表面缺陷三维模拟分析[J]. 电子科技大学学报,2020,49(6):942-948.
- [4] 颜学坤,楚建安. 基于YOLOv5改进算法的印花图案疵点检测[J]. 电子测量技术,2022,45(4):59-65.
- [5] JIANG Peiyuan, ERGU D, LIU Fangyao, et al. A Review of Yolo algorithm developments [J]. Procedia Computer Science, 2022, 199: 1066-1073.
- [6] LI Jiangyun, SU Zhenfeng, GENG Jiahui, et al. Real-time detection of steel strip surface defects based on improved Yolo detection network[J]. IFAC-PapersOnLine, 2018, 51(21): 76-81.
- [7] 徐镭,朱洪锦,范洪辉,等. 改进的YOLOv3网络在钢板表面缺陷检测研究[J]. 计算机工程与应用, 2020, 56(16): 265-272.
- [8] 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.
- [9] DONG Hao, YUAN Mu, WANG Shu, et al. PHAM-YOLO: A parallel hybrid attention mechanism network for defect detection of meter in substation[J]. Sensors, 2023, 23(13): 6052.
- [10] YANG N, GUOW. 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.
- [11] WANG Han, YANG Xiuding, ZHOU Beishi, et al. Strip surface defect detection algorithm based on YOLOV5 [J]. Materials, 2023, 16(7): 2811.
- [12] GUO Zexuan, WANG Chensheng, YANG Guang, et al. Msft-Yolo: Improved YOLOv5 based on transformer for detecting defects of steel surface[J]. Sensors,2022,22(9):3467.
- [13] QIN Xu, WANG Zhilin, BAI Yuanchao, et al. FFA-Net: Feature fusion attention network for single image dehazing[C]//Proceedings of the AAAI Conference on Artificial Intelligence (AAAI). AAAI,2020:11908-11915.
- [14] MISRA D, NALAMADA T, ARASANIPALAI A U, et al. Rotate to attend: Convolutional triplet attention module[J]. arXiv preprint arXiv:2010.03045v1, 2020.
- [15] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York, USA: IEEE,2017:936-944.
- [16] 张倩,刘紫燕,陈运雷,等. 融合Transformer和改进PANet的YOLOv5s交通标志检测[J]. 传感技术学报,2023,36(2):232-241.
- [17] 孙海蓉,李莉,周映杰,等. 基于特征金字塔融合高分辨率网络的光伏热斑识别[J]. 太阳能学报,2023,44(9):109-116.
- [18] 韩岩江,王伟,王峰萍. 融合坐标注意力和BiFPN的YOLO v5s交通标志检测方法[J]. 国外电子测量技术, 2022, 41(11): 170-179.
- [19] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [M]//FERRARI V, HEBERT M, SMINCHISECCU C. Computer Vision (ECCV), Lecture Notes in Computer Science. Cham:Springer, 2018: 3-19.
- [20] 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, Republic of Korea :IEEE,2019: 3007-3016.
- [21] HE Yu, SONG Kechen, MENG Qinggang. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features [J]. IEEE Transactions on Instrumentation and Measurement , 2020,69(4):1493-1504.