

基于YOLOv7-IMN的钢表面缺陷检测算法

关怡馨, 于春和

(沈阳航空航天大学 电子信息工程学院, 辽宁 沈阳 110136)

摘要: 钢材是工业生产中的重要材料, 表面缺陷会严重影响其性能。因此, 设计快速、准确的缺陷检测算法具有重要意义。针对YOLOv7算法存在模型复杂的问题, 提出一种改进的YOLOv7-IMN算法用于检测钢材表面缺陷。首先构建Inception-ELAN模块以减少网络的参数量, 使用大核卷积扩大感受野并提高模型的检测性能; 其次, 在主干网络中使用多尺度特征提取策略(MFE), 以获得更多特征信息和细节信息, 进一步增强主干网络的特征提取能力, 并引入EMA注意力机制, 增强对缺陷位置信息的提取能力; 最后, 针对损失函数进行优化, 引入NWD降低模型对微小缺陷位置变化的敏感性, 提高检测能力。为验证改进算法的有效性, 使用NEU-DET数据集进行验证, 结果表明, 改进后的网络模型检测精度达到81.3%, 较YOLOv7提升了7.6%, 且参数量降低了24.3%, 计算量降低了31.8%, 检测速度提高了19.3%, 具有更高的检测精度和更快的检测速度。

关键词: 缺陷检测; YOLOv7; 特征提取; EMA; NWD

DOI: 10.11907/rjtk.241349

中图分类号: TP391.4

文献标识码: A

开放科学(资源服务)标识码(OSID):

文章编号: 1672-7800(2025)006-0192-07



Steel Surface Defect Detection Algorithm Based on YOLOv7-IMN

GUAN Yixin, YU Chunhe

(College of Electronic and Information Engineering, Shenyang Aerospace University, Shenyang 110136, China)

Abstract: Steel is a crucial material in industrial manufacturing, and any defects on its surface can significantly impact its functionality. Hence, it is crucial to conduct prompt and precise detection of faults on the surface of steel. This work presents an enhanced YOLOv7-IMN algorithm designed to accurately identify steel surface defects. Firstly, the parameter of network is decreased by developing the Inception-ELAN module, which utilizes large-kernel convolution to extend the receptive field and enhance the model's detection capability. Secondly, the multi-scale feature extraction (MFE) strategy is used in the backbone to acquire additional feature and detail information. The efficient multi-scale attention (EMA) mechanism is introduced to enhance the extraction ability of defect position information. Finally, the normalized wasserstein distance (NWD) is introduced to improve the ability to detect small defects. The enhanced algorithm on the NEU-DET dataset achieves a detection accuracy of 81.3%, surpassing the YOLOv7 by 7.6%. Additionally, the number of parameters is reduced by 24.3%, amount of computation is reduced by 31.8%, and inference speed is improved by 19.3%. These improvements make the model highly suitable for industrial applications, as it offers superior detection accuracy and faster speed.

Key Words: defect detection; YOLOv7; feature extraction; EMA; NWD

0 引言

钢材被广泛应用于制造业、建筑业等重要领域, 其表面缺陷会影响钢材的强度和耐久性, 造成经济损失和安全隐患^[1]。传统的缺陷检测方法效率低、实时性差, 难以满足现代工业生产线对效率和质量的要求。随着自动化技

术的快速发展, 基于机器视觉的钢表面缺陷检测技术在工业生产中得到了广泛关注^[2]。基于深度学习的检测方法通过对大量钢材图像数据进行训练和学习, 实现了对钢材缺陷的自动化检测与识别^[3]。例如, Wang等^[4]提出一种改进的Mask R-CNN算法用于检测钢轨表面缺陷, 该算法设计了新的特征金字塔用于多尺度融合, 使用CIoU以获得更好的检测效果。Liu等^[5]提出一种基于CNN和LSTM

收稿日期: 2024-04-08

扫描二维码阅读全文:



作者简介: 关怡馨(1992-), 女, 沈阳航空航天大学电子信息工程学院硕士研究生, 研究方向为图像处理与传输技术; 于春和(1976-), 男, 博士, 沈阳航空航天大学电子信息工程学院副教授, 研究方向为图像处理与传输技术。本文通讯作者: 于春和。

的周期性缺陷检测方法,使用CNN网络提取缺陷图像特征,然后将特征向量输入到LSTM网络中进行钢材缺陷识别。Kou等^[6]设计密集卷积块对YOLO网络进行改进,在NEU-DET和GC10-DET数据集上均取得了较好效果。

YOLOv7作为YOLO系列中较为先进的算法之一,因其具有快速、准确、稳定的特性,被广泛应用于多种检测场景^[7-10]。但对于钢表面缺陷检测任务,该方法仍需改进。首先,YOLOv7算法计算量和参数量较大,增加了硬件成本。其次,YOLOv7对于小缺陷目标的检测能力有待提高。因此,本文以YOLOv7算法为基线加以改进,并在NEU-DET数据集上进行实验以验证算法的有效性。本文主要贡献如下:

(1)对YOLOv7中的ELAN(ELAN-H)模块进行优化,降低模型的计算量和参数量,提高检测精度和速度。

(2)采用多尺度特征提取策略,从主干网络的不同阶段提取不同语义信息,以更有效地学习多尺度目标特征。

(3)引入EMA(Efficient Multi-scale Attention)注意力机制,使算法更加关注有用信息。

(4)对损失函数进行优化,引入NWD(Normalized Wasserstein Distance)代替IoU,降低边界框位置偏差对于小缺陷的影响,从而提升对小缺陷目标的缺陷检测能力。

1 相关工作

1.1 基于深度学习的缺陷检测算法

基于深度学习的钢材表面缺陷检测算法分为双阶段检测算法和单阶段检测算法两类。双阶段检测算法需要先提取候选区域再进行分类定位,因此通常需要较长的训练时间。Zhou等^[11]提出一种改进版本的Faster-R-CNN模型,使用K-Means聚类算法生成符合要求的锚框,将特征图与不同的感受野相结合用于缺陷检测。Wei等^[12]提出一种改进的Faster-R-CNN,设计了Weighted-RoI以消除位置偏差对小缺陷的影响,并提出融合多尺度特征的特征金字塔以提升检测能力。

相比之下,单阶段检测算法具有更快的检测速度。Xue等^[13]对SSD网络的主干网络进行改进,以获得更好的特征融合效果,并构建改进的EfficientNet对金属表面缺陷进行分类。Bai等^[14]使用轻量化的MobileNetv3替换YOLOv4的主干网络,并在PANet层应用深度可分离卷积实现对铁路表面的实时检测。Li等^[15]提出一种改进的YOLOv4算法,将注意力机制嵌入主干网络,设计了RFB模块替换SPP模块,获得了良好的检测效果。Zhao等^[16]提出一种基于YOLOv5的钢材表面缺陷检测算法,该算法主干网络中的Res2Net模块能够有效地扩大感受野,同时双特征金字塔网络结构和解耦的头提高了检测精度。

1.2 YOLOv7算法

YOLOv7在YOLO系列前作的基础上进行改进,在不

影响推理速度的情况下提高了准确率^[17]。YOLOv7算法的网络结构如图1所示。主干网络由多个高效的ELAN模块和MP模块组成,ELAN模块能够在不改变原有梯度路径的条件下使网络学习到更多特征,且具有更强的鲁棒性。MP模块通过下采样操作扩展感受野,从而获得更加丰富的特征表达。颈部网络中SPPCSPC模块能够通过不同尺度的感受野来区分物体大小。ELAN-H模块与主干网络中的ELAN模块结构类似,使模型在更深的结构中进行更高效的学习。检测头部分基于anchor机制使用RepConv模块调整通道数和预测大中小3种尺寸的特征图,再通过CIoU损失函数和NMS处理得到最终的预测结果。

2 改进的YOLOv7-IMN算法

2.1 Inception-ELAN模块

YOLOv7主干网络使用的ELAN结构由多个CBS(Conv-BN-SiLU)模块堆叠而成,导致网络参数量和计算量增大。为了节约计算成本,本文基于Inception网络结构,构建能提取丰富特征的轻量化模块,称为Inception-block。具体结构如图2所示。

对于通道数为 C 的输入 X ,本文沿着通道维度将其分为4组:

$$X_1, X_2, X_3, X_{id} = \text{Split}(X) = X_{:,1:g}, X_{:,g:2g}, X_{:,2g:3g}, X_{:,3g:} \quad (1)$$

式中, g 为卷积支路通道数。分支通道数由分割比率 r_g 确定:

$$g = r_g C \quad (2)$$

式中, C 为输入通道数。接下来使用CBS模块对各组进行特征提取,为获得多尺度的感受野,本文使用 k_s 和 k_l 两种大小不同的卷积核。同时,为了避免使用大卷积核而导致的计算成本增加,将 $k_l \times k_l$ 卷积拆分为 $1 \times k_l$ 和 $k_l \times 1$,其中卷积核 k_s 设置为3, k_l 设置为11。最后将各支路的输出进行拼接,得到该模块的输出。通过Inception-block能够捕捉到更多细节和不同尺度的特征,从而提高模型检测效果。

将Inception-block引入ELAN结构,构建特征提取模块Inception-ELAN。其结构如图3所示。其中,两个 1×1 卷积负责进行通道数调整,得到通道数减半的两个分支,其中一个分支经过两个Inception-block,并分别保留其输出,再将得到的4个分支进行拼接。Neck中的Inception-ELAN-H采取类似操作。由于只对部分通道进行了卷积操作,在减少计算量、提高计算效率的同时,保留了ELAN结构丰富的梯度组合,增强了特征提取能力,并保持较快的推理速度。与原结构相比,Inception-ELAN在保持特征提取能力的同时,实现了对网络的轻量化改进。

2.2 多尺度特征提取策略

引入大核卷积能够提升模型性能,如RepLKNet使用了 31×31 的超大卷积核。但考虑到钢材表面缺陷大小不一,且具有随机性,仅引入大卷积核可能导致模型无法捕

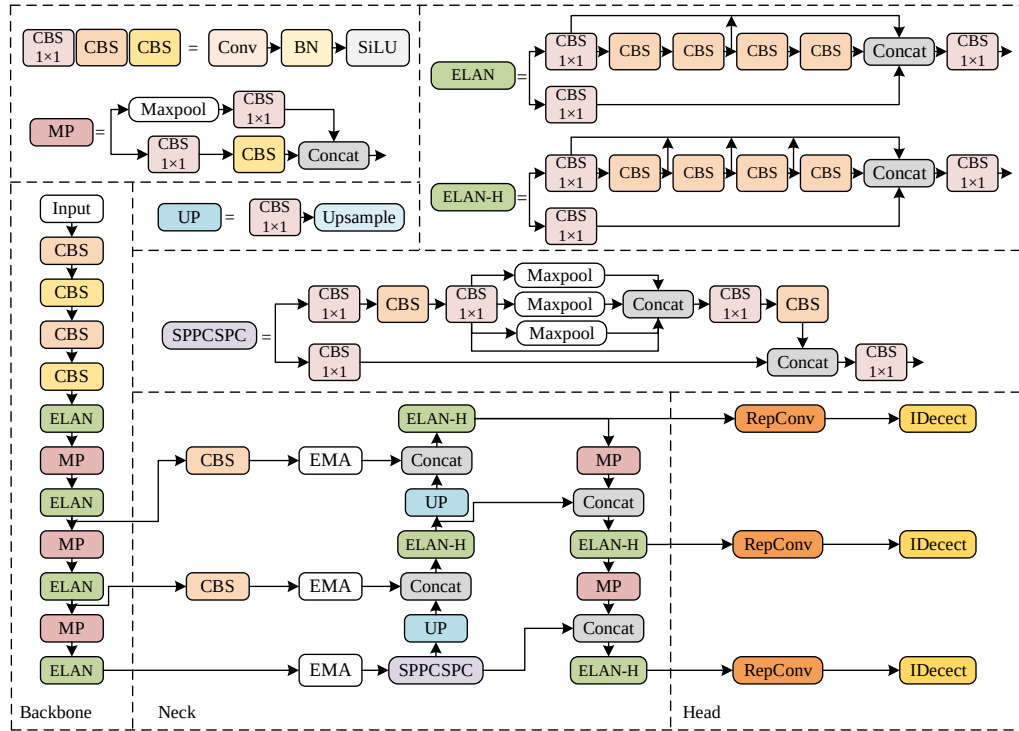


Fig. 1 Architecture of YOLOv7

图1 YOLOv7 结构

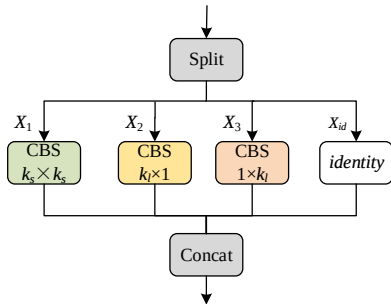


Fig. 2 Architecture of Inception-block

图2 Inception-block 结构

提到小缺陷的细节和边缘信息。此外,背景的干扰会导致缺陷与周围环境融合,使得算法对于缺陷边缘等关键位置信息的关注度降低,进而影响检测效果。

本文提出在不同阶段利用差异化卷积来捕获更丰富

特征的方法,称为多尺度特征提取策略(Multi-scale Feature Extraction, MFE)。基于Inception-block的多分支结构灵活地选择卷积核大小,在主干网络的不同阶段使用不同大小的卷积核,以便于同时提取细粒度和粗粒度的语义信息,使模型既可以捕捉到更细节的图像特征,又可以保留全局信息。本文根据网络深度和输入图像尺寸,从浅层到深层将 k_l 分别赋值为5、7、9和11。其结构如图4所示。

多尺度卷积可以在浅层阶段提取高分辨率特征以捕获细粒度的语义,过滤背景中包含的污染信息用于检测较小缺陷。在深层阶段可增加网络的感受野,提取低分辨率特征以捕获粗粒度语义,用于检测大型缺陷。多尺度特征提取策略在差异化卷积的帮助下可有效地学习并表达多尺度特征,提升检测精度。

2.3 EMA 注意力机制

EMA 注意力机制是一种高效且无需降维的多尺度注

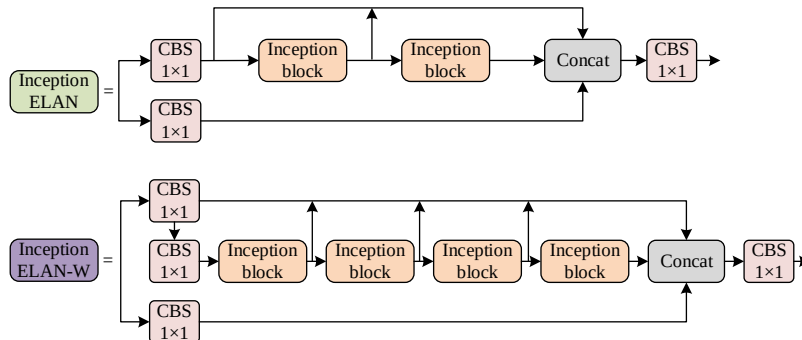


Fig. 3 Architecture of Inception-ELAN/Inception-ELAN-H

图3 Inception-ELAN/Inception-ELAN-H 结构

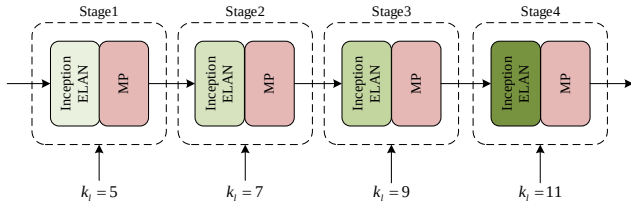


Fig. 4 Schematic diagram of multi-scale feature extraction strategy

图4 多尺度特征提取策略示意图

意力机制。通过对输入特征进行分组操作,将通道维度分组为多个子特征,并通过并行分支提取分组特征图的注意力权重。其跨空间学习结构高效提取了各个通道之间的依赖关系,并保留了更多空间信息。EMA中的并行子结构降低了网络深度,在不降低通道维度的情况下保留每个通道的信息。主干网络输出的特征图首先送入EMA模块,再送入FPN结构进行特征融合,兼顾了通道信息和位置信息。因此,EMA注意力机制帮助网络将更多有价值的信息融合到特征中,从而细化特征。

2.4 损失函数优化

由于钢材表面缺陷尺寸大小不一,且边界框中往往包含一些背景像素,对于较小的缺陷,IoU对位置偏差更为敏感,影响检测的准确性。因此,本文使用NWD判别预测边界框和真实边界框的相似度。NWD将边界框建模为二维高斯分布,即令边界框的中心像素权重最高,从中心到边界权重递减。对于水平边界框 $R = (c_x, c_y, w, h)$,其中 c_x, c_y 为中心坐标, w, h 分别为宽度和高度,可以建模为二维高斯分布 $N(\mu, \Sigma)$ 。

$$\mu = \begin{bmatrix} c_x \\ c_y \end{bmatrix}, \Sigma = \begin{bmatrix} \frac{w^2}{4} & 0 \\ 0 & \frac{h^2}{4} \end{bmatrix} \quad (3)$$

对于任意两个边界框 $A = (c_{xa}, c_{ya}, w_a, h_a)$ 和 $B = (c_{xb}, c_{yb}, w_b, h_b)$,其高斯分布距离为:

$$W_2^2(N_a, N_b) = \left\| \begin{bmatrix} c_{xa}, c_{ya}, \frac{w_a}{2}, \frac{h_a}{2} \end{bmatrix}^T, \begin{bmatrix} c_{xb}, c_{yb}, \frac{w_b}{2}, \frac{h_b}{2} \end{bmatrix}^T \right\|_2^2 \quad (4)$$

将距离度量 $W_2^2(N_a, N_b)$ 的指数形式归一化,获得归一化 Wasserstein 距离(NWD):

$$NWD(N_a, N_b) = \exp \left(- \frac{\sqrt{W_2^2(N_a, N_b)}}{C} \right) \quad (5)$$

其中, C 是与数据集有关的常数。在本文的 NEU-DET 数据集中,设置 $C=12.8$ 。

对于微小的缺陷而言,轻微的位置偏差会对 IoU 产生明显影响,影响标签分配的准确性,进而导致网络收敛困难和监督信息不足等问题。NWD 具有位置偏差平滑性,更加关注目标中心点的信息,使用 NWD 度量能够更全面地评估两个边界框之间的相似度。无论目标的边界框是

否发生重叠,都可以实现预测框与真实框之间高效、准确的相似性度量,且在检测较小缺陷时表现出良好的稳定性。引入 NWD 可克服传统 IoU 损失函数的局限性,缓解由于 IoU 对小目标的敏感性和正负样本不平衡导致的训练困难问题,更准确评估目标检测结果之间的相似性,使模型更快地达到收敛状态。

2.5 改进的 YOLOv7-IMN 网络结构

改进后的 YOLOv7-IMN 网络结构如图 5 所示。其使用轻量化的 Inception-ELAN 和 Inception-ELAN-H 模块替换原有结构,以降低网络计算复杂度,并提升算法的检测速度;采用 MFE 策略对主干网络进行优化,提升网络的特征提取能力;在 YOLOv7 的 Neck 中加入 EMA 模块,增强了对关键信息的提取能力;使用 NWD 度量综合考虑缺陷的尺寸、位置与置信度等因素,降低模型对于小缺陷的敏感度。此外,引入 NWD 度量加速模型收敛,提高了网络模型的缺陷定位检测能力。

3 实验与分析

3.1 数据集与实验环境

本文使用东北大学公开数据集 NEU-DET 钢材缺陷数据集进行训练和验证。该数据集包含 1 800 幅灰度图像,图像尺寸均为 200×200 像素,按常见的钢材表面缺陷分为裂缝、夹杂物、斑块、点蚀表面、氧化皮和划痕 6 类。将数据集按照 7:1:2 的比例随机进行划分,获得训练集图像 1 260 幅,验证集图像 180 幅,测试集图像 360 幅。所有实验均在相同环境下进行,具体实验环境及参数设置如表 1 所示。

3.2 评价指标

为了验证改进算法的有效性,实验采用参数量(Parameters)、每秒浮点运算次数(Floating-point Operations Per Second, FLOPs)、平均精度均值 mAP (mean Average Precision)、单张图片耗时(Frames Per Second, FPS)等指标进行评估。其中,准确率 P (Precision)、召回率 R (Recall)、平均精度 AP (Average Precision)、mAP 的定义如下:

$$\begin{aligned} P &= \frac{TP}{TP + FP} \\ R &= \frac{TP}{TP + FN} \\ AP &= \int_0^1 P(R) dR \\ mAP &= \frac{1}{c} \sum_{i=1}^c AP_i \end{aligned} \quad (6)$$

其中, TP 表示预测正确的正样本, TN 表示预测正确的负样本, FP 表示预测错误的正样本, FN 表示预测错误的负样本, c 表示类别数。

3.3 实验结果分析

3.3.1 消融实验

为了验证本文每个改进点的有效性,以 YOLOv7 算法

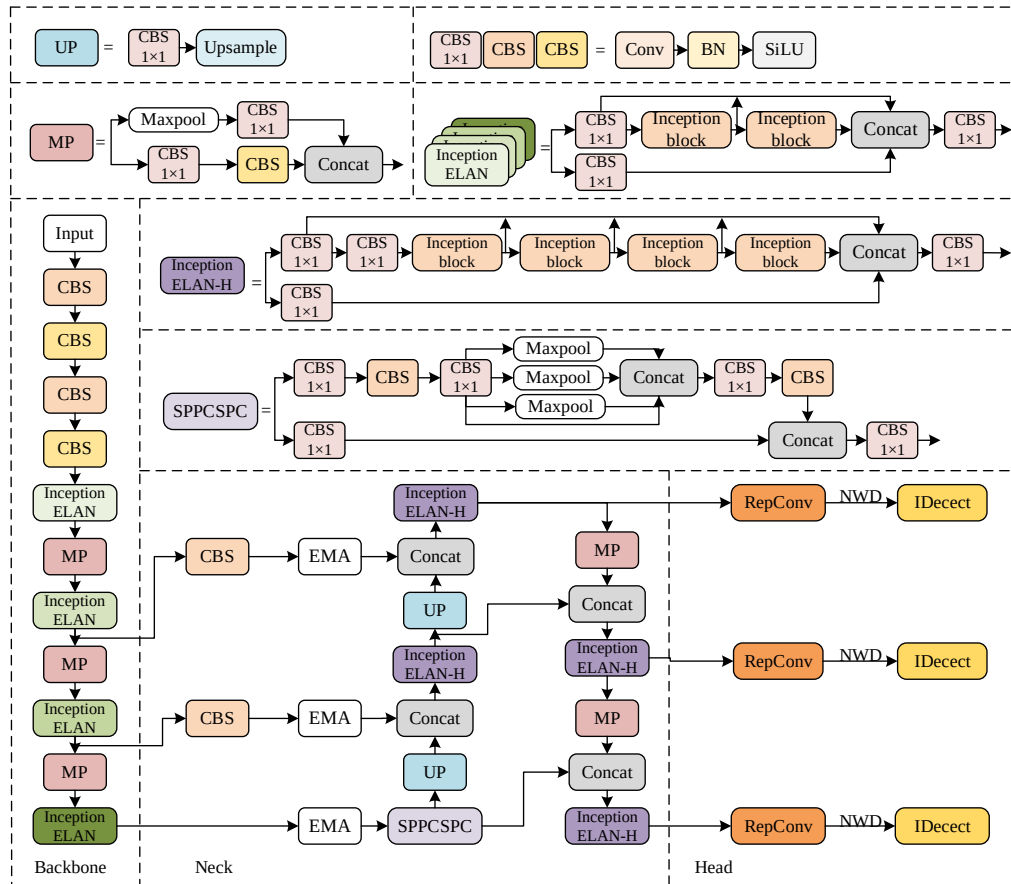


Fig. 5 Architecture of improved YOLOv7-IMN

图5 改进的YOLOv7-IMN结构

Table1 Experimental environment and parameter configuration

表 1 实验环境及参数设置

参数	配置
CPU	AMD Ryzen 7 5800H with Radeon Graphics 3.20 GHz
GPU	NVIDIA GeForce RTX 3060
操作系统	Windows 11
Pytorch	Pytorch 1.11.0+CUDA11.7
Python	Python 3.9.0
初始学习率	0.01
Batch size	8
Epoch	300
动量	0.937

为基线,在 NEU-DET 数据集上进行消融实验。实验结果如表 2 所示,“√”表示使用相应方法。

Table 2 Results of ablation experiment

表 2 消融实验结果

序号	YOLOv7	Inception-ELAN	EMA	NWD	MFE	FPS	Params/M	Flops/G	mAP@0.5/%
1	√					64.2	35.49	105.2	73.7
2	√	√				79.8	26.03	71.6	77.8
3	√		√			63.9	35.50	105.3	78.6
4	√			√		64.4	35.49	105.2	76.2
5	√	√			√	78.5	26.11	71.6	79.3
6	√	√	√		√	76.4	26.85	71.7	80.6
7	√	√		√	√	78.7	26.14	71.6	80.2
8	√	√	√	√	√	76.6	26.88	71.7	81.3

实验 1 为 YOLOv7 算法实验结果,作为其他实验的对比基准,其中参数量为 35.49 M,计算量为 105.2 G,mAP 为 73.7%,FPS 为 64.2。实验 2 引入 Inception-ELAN 和 Inception-ELAN-H 模块,由实验结果可知参数量和计算量显著下降,分别减少了 9.46 M 和 33.6 G,mAP 提高了 4.1%,FPS 提高了 15.6。实验 5 在实验 2 的基础上使用 MFE 策略,此时模型参数量增加 0.8 M,计算量保持不变,FPS 下降了 1.3,mAP 达到 79.3%。

实验 3 和实验 4 分别在实验 1 的基础上加入 EMA 与 NWD,实验 6 和实验 7 分别在实验 5 的基础上加入 EMA 与 NWD。由实验结果可知,加入 EMA 和 NWD 对模型的参数量和计算量影响并不大,但能够提升检测性能。图 6 展示了使用 IoU 和 NWD 度量的损失函数曲线图,引入 NWD 度量后模型的损失值下降显著,能够更快地达到较低的收敛值,模型的收敛速度更快。

实验 8 使用改进的 YOLOv7-IMN 算法,与实验 1 相比,参数量和计算量显著减少,分别为 26.88 M 和 71.7 G,分别降低了 24.3% 和 31.8%,mAP 值提升至 81.3%,升高了 7.6%,推理速度提升了 12.4 fps,证实了 YOLOv7-IMN 算法中每个改进点的有效性,算法性能在各方面也更占优势。

3.3.2 INCEPTION-ELAN 有效性实验

为验证 Inception-block 中各分支的有效性,本文通过

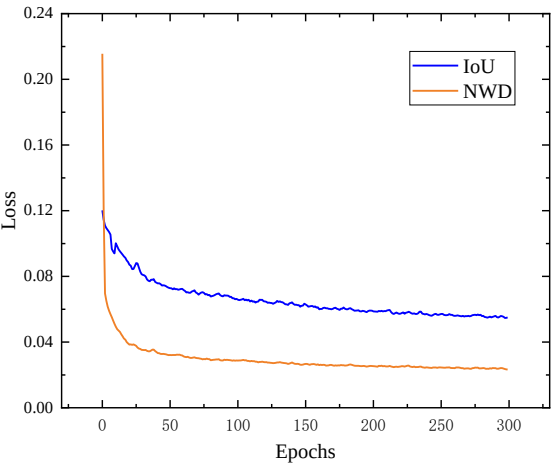


Fig. 6 Comparison of loss function curves between YOLOv7 and YOLOv7-IMN

图6 YOLOv7与YOLOv7-IMN损失函数曲线对比

去除其中任意一个分支的方式进行实验。其中, k_i 和 k_l 均设置为 3, g 为分支通道数, 0 表示去除该通道。实验结果见表 3。

Table 3 Results of branching validity experiments

表 3 分支有效性实验结果

$[k_s, 1 \times k_l, k_l \times 1, id]$	Parmas/M	Flops/G	mAP@0.5/%
$[g, g, g, 5g]$	26.09	71.6	77.8
$[0, g, g, 6g]$	26.08	71.6	77.5
$[g, 0, g, 6g]$	26.09	71.6	76.6
$[g, g, 0, 6g]$	26.09	71.6	76.2

由实验结果可以看出,当去除 $1 \times k_l$ 或 $k_l \times 1$ 分支时,性能显著下降,从 77.8% 下降到 76.2%,说明这两个分支的重要性。去除 $k_s \times k_s$ 分支时 mAP 略微下降,在对于检测精度的要求并不十分严格时,可以去除该分支。

3.3.3 多尺度特征提取策略有效性实验

为验证多尺度特征提取策略的有效性,本文通过设置不同卷积核大小进行对比实验,其中使用 $[s_1, s_2, s_3, s_4]$ 表示 s_i 阶段 k_i 的大小, k_s 设置为 3。实验结果见表 4。

Table 4 Comparative experiment results of MFE strategies

表 4 多尺度特征提取策略对比实验结果

序号	$[s_1, s_2, s_3, s_4]$	Parmas/M	Flops/G	mAP@0.5/%
1	$[3, 3, 3, 3]$	26.08	71.6	76.0
2	$[7, 7, 7, 7]$	26.09	71.7	76.9
3	$[11, 11, 11, 11]$	26.10	71.7	77.0
4	$[3, 5, 7, 9]$	26.09	71.6	78.2
5	$[5, 7, 9, 11]$	26.09	71.6	78.6
6	$[11, 9, 7, 5]$	26.09	71.6	75.8

由实验结果可知,增加卷积核大小并不会显著提升检测能力。然而,当使用 MFE 策略时,算法检测能力得到较大提升, mAP 相较于改进前提升了 2.6%, 优于具有相同卷积核大小的设置。由实验 6 可知,卷积核在各阶段的排列顺序也起着至关重要的作用,当浅层阶段使用大卷积核,深层阶段使用小卷积核时, mAP 下降了 2.8%, 表明深层阶段需要更大的感受野才能有效捕获粗粒度信息。

3.3.4 不同注意力机制比较

为验证 EMA 的有效性,将其与其它主流注意力机制进行比较,实验结果如表 5 所示。

Table 5 Comparative experiment results of different attention mechanisms

表 5 不同注意力机制对比实验结果

注意力机制	Parmas/M	Flops/G	mAP@0.5/%
YOLOv7	35.49	105.2	73.7
+SE	35.50	105.2	76.5
+CBAM	35.50	105.2	76.8
+CA	35.60	105.3	75.7
+GAM	35.63	105.5	77.1
+EMA	35.50	105.3	78.6
+ECA	35.50	105.7	78.2

从表 5 可以看出,引入 EMA 机制后,模型的参数量和计算量略微增加,但在检测精度上优于其他注意力机制, mAP 为 78.6%。与目前主流的注意力机制相比,EMA 机制通过跨维度的交互作用保留了每个通道上的信息,并行子网络保留了多尺度空间信息,可实现更丰富的特征聚合,有效提升了模型的检测能力。

3.3.5 不同算法性能比较

为验证改进算法的性能及优势,将 YOLOv7-IMN 算法与其他 YOLO 系列算法进行对比实验,结果如表 6 所示。

Table 6 Comparative experiment results of different algorithms

表 6 不同算法对比实验结果

方法	FPS	Parmas/M	Flops/G	mAP@0.5/%
YOLOv5l	59.8	47.89	147.7	72.9
YOLOv7-tiny	135.5	5.75	13.2	70.7
YOLOv7	64.2	35.49	105.2	73.7
YOLOv8l	65.5	41.66	165.7	79.2
YOLOv7-IMN	76.6	26.88	71.7	81.3

根据实验结果可知,YOLOv7-IMN 算法在检测精度上优于其他几种算法,其 mAP 为 81.3%;YOLOv7-tiny 算法虽然在推理速度上优于本文算法,但检测精度不佳; mAP 与 YOLOv7-IMN 算法相差较大。图 7 展示了使用 YOLO 系列算法与 YOLOv7-IMN 算法的检测效果,从图中可以看出, YOLOv7-IMN 算法对于漏检现象有明显改善,对于较小的划痕和夹杂物(Inclusion)缺陷也有较好的检出效果。

4 结语

本文针对目前钢材表面缺陷检测算法中网络结构参数量大、计算复杂度高的问题,首先,提出 Inception-ELAN 模块来减少网络复杂度,使模型更加轻量化。其次,提出 MFE 策略帮助模型更有效地学习并表达多尺度特征,并引入 EMA 注意力机制,加强网络对缺陷区域特征的关注度。最后,采用 NWD 替换 IoU,在提高网络收敛速度的同时,提升对小目标的定位和检测能力。通过一系列对比实验证明了改进算法的有效性,改进后的 YOLOv7-IMN 算法参数量和计算量显著降低,检测速度和 mAP 均有明显提升。

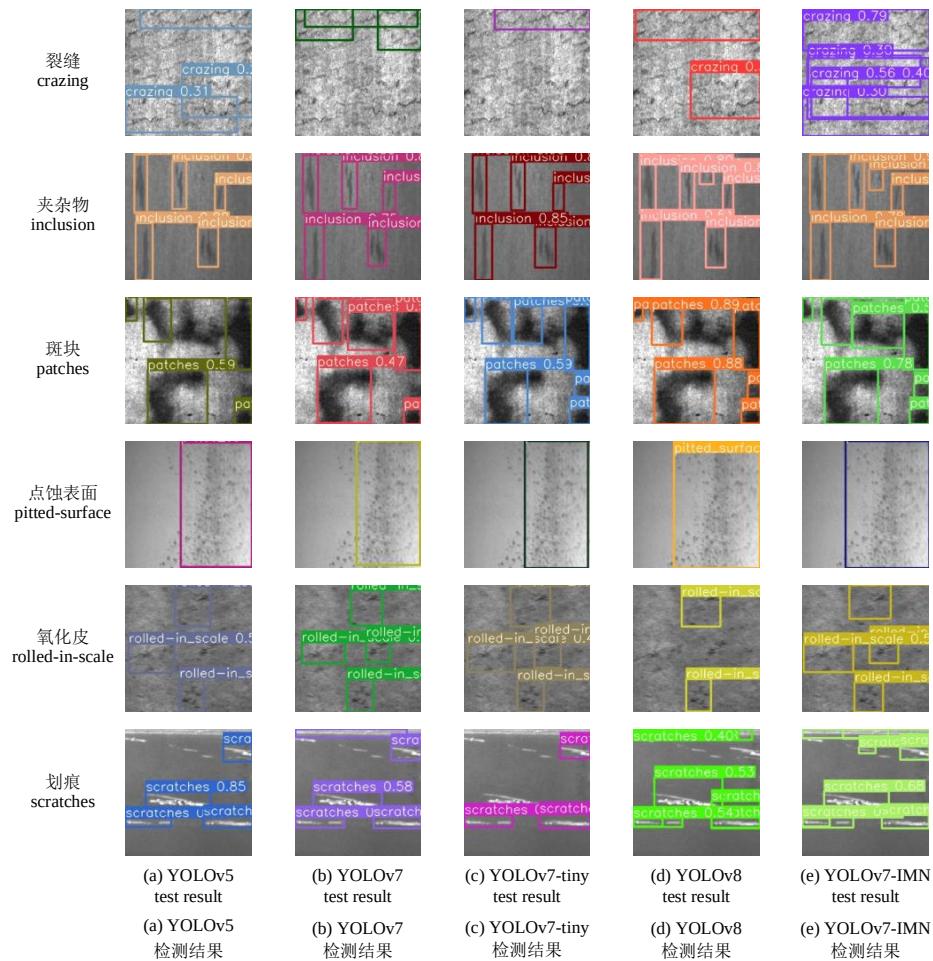


Fig. 7 Detection results comparison of different YOLO series algorithms

图7 YOLO 系列不同算法检测结果比较

参考文献:

- [1] LUO Q W, FANG X X, LIU L, et al. Automated visual defect detection for flat steel surface: a survey[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(3): 626-644.
- [2] LIU Q, LEI J S. Steel surface defect detection based on improved deep networks[J]. Computer Engineering and Design, 2022(9): 2654-2661. 刘琪, 雷景生. 基于改进深度网络的钢材表面缺陷检测[J]. 计算机工程与设计, 2022(9): 2654-2661.
- [3] WEIMER D, SCHOLZ-REITER B, SHPITALNI M. Design of deep convolutional neural network architectures for automated feature extraction in industrial inspection[J]. CIRP Annals, 2016, 65(1): 417-420.
- [4] WANG H, LI M, WAN Z. Rail surface defect detection based on improved Mask R-CNN[J]. Computers and Electrical Engineering, 2022, 102: 108269.
- [5] LIU Y, XU K, XU J. Periodic surface defect detection in steel plates based on deep learning[J]. Applied Sciences, 2019, 9(15): 3127.
- [6] KOU X, LIU S, CHENG K, et al. Development of a YOLO-V3-based model for detecting defects on steel strip surface[J]. Measurement, 2021, 182: 109454.
- [7] WANG C, BOCHKOVSKIY A, LIAO H. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[DB/OL]. <https://arxiv.org/abs/2207.02696>.
- [8] LIU S, WANG Y, YU Q, et al. CEAM-YOLOv7: improved YOLOv7 based on channel expansion and attention mechanism for driver distraction behavior detection[J]. IEEE Access, 2022, 10: 129116-129124.
- [9] NGUYEN H, BAE J, LEE Y, et al. Comparison of pre-trained YOLO models on steel surface defects detector based on transfer learning with GPU-based embedded devices[J]. Sensors, 2022, 22(24): 9926.
- [10] RAJ G, PRABADEVI B. Steel strip quality assurance with YOLOV7-CSF: a coordinate attention and SiLU fusion approach[J]. IEEE Access, 2023, 11: 129493-129506.
- [11] ZHOU Z, LU Q, WANG Z, et al. Detection of micro-defects on irregular reflective surfaces based on improved faster R-CNN [J]. Sensors, 2019, 19(22): 5000.
- [12] WEI R, SONG Y, ZHANG Y. Enhanced faster region convolutional neural networks for steel surface defect detection [J]. ISIJ International, 2020, 60(3): 539-545.
- [13] XUE P, JIANG C, PANG H. Detection of various types of metal surface defects based on image processing[J]. Traitement Du Signal, 2021, 38 (4): 1071-1078.
- [14] BAI T, GAO J, YANG J, et al. A study on railway surface defects detection based on machine vision[J]. Entropy, 2021, 23(11): 1437.
- [15] LI Z, TIAN X, LIU X, et al. A two-stage industrial defect detection framework based on improved-YOLOv5 and optimized-inception-resnetV2 models[J]. Applied Sciences, 2022, 12(2): 834.
- [16] ZHAO C, SHU X, YAN X, et al. RDD-YOLO: a modified YOLO for detection of steel surface defects[J]. Measurement, 2023, 214: 112776.
- [17] TERVEN J, CORDOVA-ESPARZA D. A comprehensive review of YOLO: from YOLOv1 to YOLOv8 and beyond[DB/OL]. <https://arxiv.org/pdf/2304.00501v1>.

(责任编辑: 黄 健)