

多尺度融合的轻量级钢材表面缺陷检测^①



杨本臣¹, 李世熙¹, 金海波¹, 康洁²

¹(辽宁工程技术大学 软件学院, 葫芦岛 125105)

²(北京工业大学 信息学部, 北京 100124)

通信作者: 李世熙, E-mail: 865026279@qq.com

摘要: 钢材表面缺陷的检测质量直接影响工业生产安全性和机器性能质量, 现实工厂钢材质量把控受限于设备条件, 在实现高精度强实时的检测效果仍面临着挑战. 为了解决这一问题, 提出一种多尺度融合的轻量级 YOLOv8n 检测算法. 首先引入一种结合 HGnetv2 与 RepConv 的轻量级多尺度融合主干网络 (RepHGnetv2), 提高 Backbone 的特征提取能力与泛化能力同时降低了模型的复杂度; 在 Head 部分, 利用 ADown 下采样模块替换原算法的普通卷积 (Conv), 降低计算量并提高语义保留能力; 最后将原算法的 Loss 函数替换为 SlideLoss, 改善样本之间不平衡的问题. 在 NEU-DET 数据集上进行消融与对比实验, 改进算法与原算法相比, $mAP@0.5$ 提升 6.7%, Precision 提升 9.3%, 模型大小下降 25.5%, 计算量下降了 17.2%, FPS 也有一定的提升; 并在 VOC2012 数据集上进行了通用性对比实验, 实验结果表明改进算法可以有效提高缺陷检测精度与效率, 同时具有较好的通用性.

关键词: 钢材表面缺陷检测; 轻量级 YOLOv8n; RepHGnetv2; ADown; SlideLoss

引用格式: 杨本臣, 李世熙, 金海波, 康洁. 多尺度融合的轻量级钢材表面缺陷检测. 计算机系统应用, 2024, 33(11): 58-67. <http://www.c-s-a.org.cn/1003-3254/9678.html>

Lightweight Steel Surface Defect Detection with Multiscale Fusion

YANG Ben-Chen¹, LI Shi-Xi¹, JIN Hai-Bo¹, KANG Jie²

¹(Software College, Liaoning University of Technology, Huludao 125105, China)

²(Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China)

Abstract: The quality of steel surface defect inspection directly affects industrial production safety and machine performance. However, in real factories, steel quality control is limited by equipment conditions, making it challenging to achieve high-precision and real-time inspection. To solve this problem, a lightweight YOLOv8n detection algorithm with multi-scale fusion is proposed. Firstly, a lightweight multi-scale fusion backbone network (RepHGnetv2) is introduced, combining HGnetv2 and RepConv to improve the feature extraction and generalization capabilities of Backbone and reduce the complexity of the model. In the Head part, the ordinary convolution of the original algorithm is replaced with the ADown downsampling module, which reduces computational complexity and improves semantic retention. Finally, the loss function of the original algorithm is replaced by SlideLoss to address sample imbalance. Ablation and comparison experiments are conducted on the NEU-DET dataset. Compared with the original algorithm, the improved algorithm increases precision by 9.3%, reduces the model size by 25.5%, decreases computational complexity by 17.2%, and improves FPS to a certain extent. Comparative experiments are conducted on the VOC2012 dataset to evaluate the generalizability of the improved algorithm, and the results show that the improved algorithm exhibits strong generalizability and effectively improves the accuracy and efficiency of defect detection.

Key words: steel surface defect detection; lightweight YOLOv8n; RepHGnetv2; ADown; SlideLoss

① 基金项目: 国家自然科学基金 (62173171)

收稿时间: 2024-04-22; 修改时间: 2024-05-29; 采用时间: 2024-06-04; csa 在线出版时间: 2024-09-29

CNKI 网络首发时间: 2024-09-30

随着科技的进步,钢铁在各种工业领域得到广泛应用,在生产和日常应用的过程中受生产设备,恶劣环境等因素影响容易出现划痕,氧化等缺陷,严重影响钢材的物理性能,使得企业生产成本增加,造成资源浪费和经济损失,甚至影响人们的生命.为落实我国高质量发展战略,探索高度可靠的钢材表面缺陷检测方法十分重要.目前传统的缺陷检测方法有磁粉检测法、渗透检测法、涡流检测法、超声波检测法^[1]、人工检测法^[2],但是这些方法大多只适用于特定类型的材料,同时检测效率过低,容易受到人工操作因素影响,从而影响检测结果的准确性.

伴随机器学习出现,目标检测算法逐渐应用在缺陷检测领域^[3],传统的目标检测算法^[4,5]主要是基于滑动窗遍历选择区域,进行特征提取.但由于计算复杂度较高,检测效果不佳,逐渐淡出视野.近年来,深度神经网络(deep neural network, DNN)快速发展,基于深度学习的目标检测算法^[6]在精度与效率方面均优于传统的检测方法.主流算法分为两类:以 Faster-RCNN 算法^[7]为代表的两阶段算法与以 SSD 算法^[8]、RetinaNet 算法^[9]与 YOLO 算法^[10]为代表的一阶段算法.两阶段的检测算法又称基于候选区域的目标检测,首先获取输入图像,生成包含图像的候选框,再将候选框的图像输入到卷积神经网络(convolutional neural network, CNN)中进行分类识别.向宽等^[11]提出了一种基于 Faster-RCNN 的铝材表面缺陷检测方法,在主干网络增加特征金字塔网络(FPN),运用 ROI Align 算法去解决小缺陷定位不准确等问题.陈仁祥等^[12]基于 Faster-RCNN 引入多注意力机制,获得具有抗干扰能力的特征表示,降低噪声干扰,提升检测效果.两阶段算法虽然具有较高的检测精度但计算复杂度较高,仍然无法达到实际工业生产要求.一阶段算法不同于前者,运用预设锚框代替生成候选框,直接得出物体类别概率与位置,可以更好地适应缺陷检测任务.Li 等^[13]提出了一种将 SSD 主干网络替换为 MobileNet 的灌装生产线容器密封表面缺陷检测方法,减少了模型的计算复杂度.吴鸿杨等^[14]提出了一种基于 RetinaNet 的缺陷检测方法,通过注意力模块对特征金字塔网络的输出特征图进行处理,让模型更加关注缺陷特征提高缺陷检测性能.

YOLO 系列算法是典型的一阶段目标检测算法.2016年由 Redmon 等人^[10]提出,核心思想是将目标检测转变为单次前向回归问题,采用多尺度训练与预测

策略,保证精度的同时实现了实时检测.YOLOv2-YOLOv4^[15,16]相继提出.Kou 等^[17]提出了一种改进的 YOLOv3 端到端的缺陷检测算法,采用 Anchor-Free 的结构替换原结构,减少了计算时间.Yu 等^[18]提出了一种基于 YOLOv4 的高效尺度感知的检测算法,通过强化浅层特征以减少目标损失,并提出了一种具有动态感受野的检测头,解决感受野与目标尺度不匹配的问题.王磊等^[19]提出了一种改进 YOLOv5s 缺陷检测算法,引入 CBAM 注意力机制并在池化模块中加入上下文卷积模块,增强网络对信息的提取效率,增强对特征表达能力.齐向明等^[20]提出了一种改进的 YOLOv7-tiny 钢材表面缺陷检测算法,将原网络激活函数更换为 SiLU,引入轻量级上采样模块 CARAFE,提高特征提取与融合能力.张曙文等^[21]提出了基于改进 YOLOX 算法的缺陷检测方法,采用 ASFF 自适应多尺度特征融合方法,修改损失函数为 Varifocal 函数,有效减少样本噪声.

目前钢材表面缺陷检测算法计算复杂度高,参数量较大,虽然可以很好的提升检测精度,但检测速度底下难以适用实际生产.本文针对上述一系列问题提出一种轻量级 YOLOv8n 钢材表面缺陷检测算法.主要工作包括如下内容.

(1) 在 Backbone 部分,提出一种轻量级多尺度融合主干网络 RepHGnetv2,将 RepConv 融入到 HGnetv2 网络中,替换原网络 LightConv,并引入到 YOLOv8n 算法中.提高模型的泛化能力和性能,减少计算量与内存消耗.

(2) 在 Head 部分,使用 ADown 下采样结构替换原算法中的部分普通卷积(Conv),减少计算复杂度并增强语义保留能力.

(3) 将原算法的损失函数更换为 SlideLoss 函数,以应对样本不平衡问题.

1 YOLOv8 算法

YOLOv8 是一种 SOTA 模型,在 YOLOv5 模型结构的基础进行改进,保留了 YOLOv5 算法高效快速的检测能力,同时为训练模型提供统一对象检测、实例分割与图像分类框架.基于缩放系数提供 n/s/m/l/x 尺度的不同大小模型,以适用于不同规模应用需求.网络结构分为 Input、Backbone、Head,具体结构如图 1 所示.

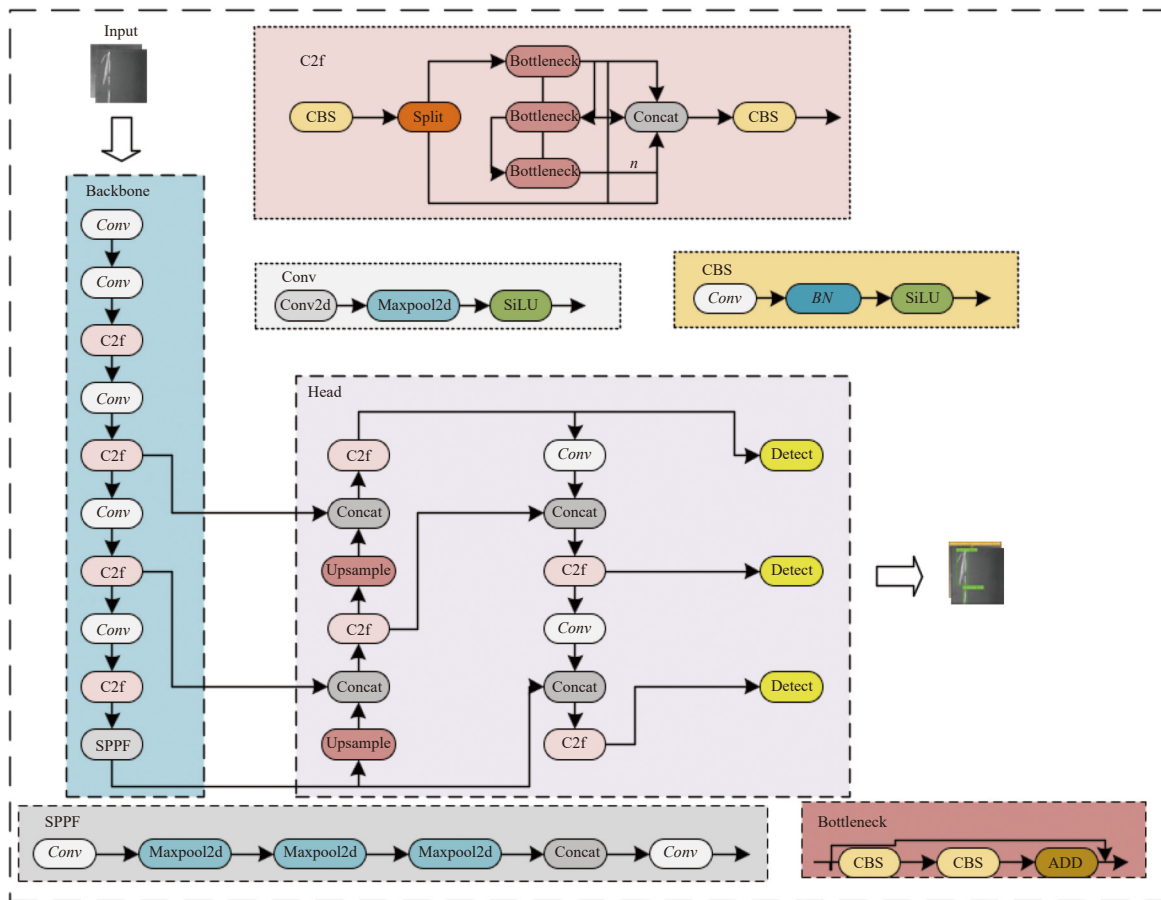


图1 YOLOv8 算法结构图

1.1 Input

Input 部分使用 Mosaic 数据增强技术, 将图片输入到 Input 进行裁剪, 缩放等一些列预处理操作, 增加了数据集的多样性, 提升训练速度, 减少内存消耗, 使 Backbone 更好地提取特征。

1.2 Backbone

Backbone 部分相比 YOLOv5 将第 1 个卷积层的 Kernel 由 6×6 转为 3×3, 引入 CBS、C2f 与 SPPF 等模块, 相比与原 YOLOv5 的 C3 模块, C2f 梯度更加丰富且轻量化。

1.3 Head

在 Head 部分, 将原解耦头中的分类和检测头分离, 使用 Anchor-Free 代替原有 Anchor-Based 方法, 使得模型更加灵活。在 Loss 计算方面 YOLOv8 使用 Task-AlignedAssigner 正样本分配策略并引入了 Distribution Focal Loss, 可以更好地分配正样本, 提高模型的检测性能。

综上所述, YOLOv8 作为一种广泛应用于目标检

测领域的深度学习模型, 具有高速高精度的检测性能。本文选用 YOLO8n 为基础框架。

2 改进的 YOLOv8n 轻量级钢材表面缺陷检测算法

2.1 RepHGnetv2 轻量级网络

2.1.1 HGnetv2 网络

以往的目标检测算法通过加宽加深主干网络的方式, 虽然可以大幅度的提升检测精度, 但同时也增加了计算复杂度, 使得基于深度学习的缺陷检测算法在实际工业生产中难以落实。针对这个问题, 2019 年 Sun 等^[2]提出了一种高性能高分辨率主干网络 HGnet。该网络通过并行连接高分辨率到分辨率卷积保持高分辨率表示, 并重复跨并行卷积执行多尺度融合来增强高分辨率表示。

HGnet 整体结构如图 2 所示。通过串联多个分辨率递减的子网络来获取多层特征, 实现分辨率金字塔式特征提取。该方法不再依靠单一的、由低到高的上

采样操作,而是在整个过程中不断融合不同尺度特征.具体来说, HGnet 通过使用减缓单元,在不同子网络之间进行特征交互和融合,从各层网络提取多视角信息.这种结构方式既保留各级特征,也实现了跨层结合,有效促进特征学习.并联网络结构如图 3 所示.

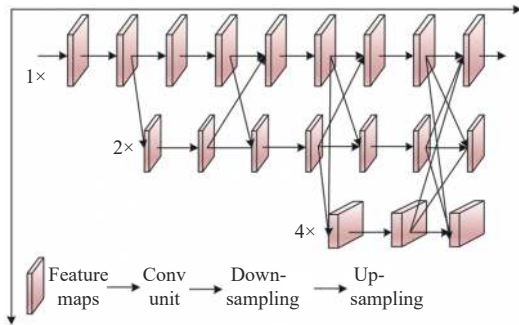


图 2 HGnet 网络结构

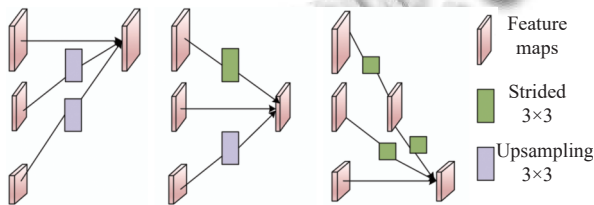


图 3 HGnet 并联网络结构

HGnetv2^[23]于 2023 年由 PaddleDetection 团队提出,在原网络基础上进行改进.首先在激活函数部分添加了 use_lab 结构,用来进行分支残差计算.其次使用 LightConv 模块,替换原网络的 ESE 模块,并且使用了 DWConv^[24].相比原网络,有效解决了梯度消失问题,显著减少了模型的参数量,提高模型检测准确性与泛化能力. HGnetv2 结构如图 4 所示.

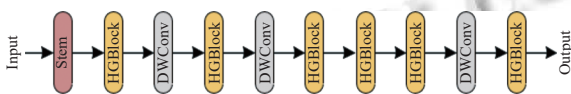


图 4 HGnetv2 网络结构

Stem 层为网络的预处理层,从输入的数据中挖掘深度特征.核心网络层为 HGBlock 层,即并联的卷积模块与通道压缩解压模块,不同 HGBlock 块处理不同的层次,从而对输入变换具有更强的建模能力,保持了高分辨率的特征表示,能够更好地处理多尺度目标,提高了模型的检测能力.

2.1.2 RepConv

RepConv 最初应用于 VGG 网络^[25],是一种重参数化技术,在推理过程中将多个计算模块合,使模型效率

与性能提升,其主要思想是在训练过程中采用多分支卷积层,在推理时将分支的参数量重参数化到主分支上,实现减少计算量与内存消耗. RepConv 首先将 3×3 卷积与 BN 层融合,公式如下:

$$Conv(x) = W(x) + b \tag{1}$$

$$BN(x) = \gamma \times \frac{x - mean}{\sqrt{var}} + \beta \tag{2}$$

其中,式(1)和式(2)分别为卷积层公式与 BN 层公式, γ 与 β 是可学习参数 weight 与 bias 值.将式(1)的卷积结果代入式(2)得出:

$$BN(Conv(x)) = \frac{\gamma \times W(x)}{\sqrt{var}} + \left(\frac{\gamma \times (b - mean)}{\sqrt{var}} + \beta \right) \tag{3}$$

其中, $\frac{\gamma \times W(x)}{\sqrt{var}}$ 为融合后的 weight, $\frac{\gamma \times (b - mean)}{\sqrt{var}} + \beta$ 为融合后卷积的 bias.然后,将 1×1 卷积的权重 padding 成 3×3 卷积,再与 3×3 卷积的权重相加,串联 1×1 卷积与 3×3 卷积,将 1×1 卷积的第 0 维与第 1 维调换位置,再将 3×3 卷积核与转置后的 1×1 卷积核进行卷积.公式如下:

$$output = Conv3 \times 3(Conv1 \times 1(x)) \tag{4}$$

最后将 1×1 卷积, 1×3 卷积与 3×1 卷积并联,卷积的 weight 全部被 padding 成 3×3 卷积的 weight,再将其相加.结构如图 5 所示.

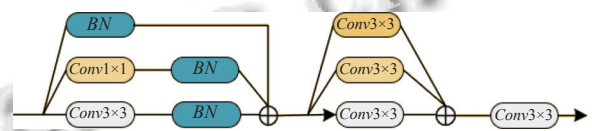


图 5 RepConv 结构

2.1.3 RepHGnetv2

在改进算法中,提出一种新型检测算法主干网络 RepHGnetv2.运用 RepConv 的重参数思想,与 HGnetv2 中的 HGBlock 进行融合,设计一个 RepHGBlock 结构,兼顾检测速度和算法性能,减少模型的参数量,实现实时的钢材表面缺陷检测. RepHGBlock 结构如图 6 所示.

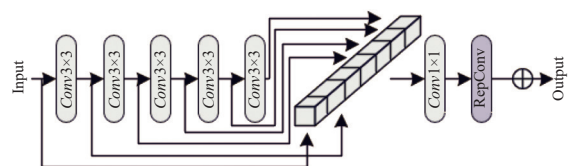


图 6 RepHGBlock 结构

2.2 ADown 下采样模块

由于钢材表面缺陷种类多且大小不一,为了改善特征融合网络性能,在 Neck 部分引入 YOLOv9^[26]中提出的 ADown 下采样模块。

ADown 是一种自适应下采样模块,可以根据不同的数据场景自动调整下采样的比例和方式,传统的下采样方法,如 MaxPooling 和 AvgPooling,在不同场景下效果可能不尽如人意,ADown 可以自适应地选择合适的下采样方法,以更好地保留关键特征。ADown 通过学习一组权重因子来决定使用哪种下采样操作,从而使网络能够更好地适应不同的数据分布和特征。结构如图 7 所示,其中“/”表示整除。

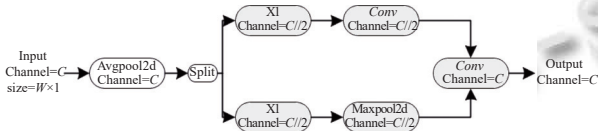


图 7 ADown 模块结构

首先输入特征图: 假设输入特征图的形状为 (C, H, W) , 其中 C 为 Channel 数, H 和 W 分别为高度和宽度。

权重生成: ADown 模块会生成 3 组权重因子 α , β 每组都是 $(1, 1, 1)$ 的形状。这 3 组权重因子分别对应 3 种下采样方式:

$$\begin{cases} \alpha : \text{MaxPooling} \\ \beta : \text{AvgPooling} \end{cases}$$

加权下采样: 对输入特征图进行 3 种下采样操作, 得到 3 个下采样特征图:

$$\begin{cases} \text{MaxPooling特征图: } F_{max} = \text{MaxPool}(input) \\ \text{AvgPooling特征图: } F_{avg} = \text{AvgPool}(input) \end{cases}$$

加权融合: 将 3 个下采样特征图按照学习到的权重因子 α, β 进行加权融合, 得到最终的下采样输出:

$$output = \alpha \cdot F_{max} + \beta \cdot F_{avg} \quad (5)$$

这个加权融合过程可以自适应地选择最佳的下采样方式, 以更好地保留关键特征, 从而提高网络在不同数据场景下的性能。

ADown 模块的权重因子 α, β 是通过端到端的训练学习出来的, 网络会根据输入特征自动调整这些权重, 实现自适应下采样。

2.3 SlideLoss 损失函数

钢材表面缺陷具有差异性, 容易造成样本不平衡, 针对这个问题, 引入 SlideLoss 损失函数^[27], 其主要作

用是让模型更加关注难例, 进而解决简单样本和困难样本之间的不平衡问题, 改善模型在难例上的效果, 有效地处理数据集中类别之间的样本数量差异, 平衡不同类别的损失权重。

SlideLoss 又称滑动损失, 可以自适应学习样本阈值参数与负样本阈值参数 μ , 引导模型提高对难样本的关注度。通过参数 μ 将样本分为正样本与负样本, 在使用加权函数 Slide 对边界处样本强调。在 μ 附近设置较高的权重会增加难分类示例的相对损失, 从而将更多注意力集中在难分类的错误示例上。公式如下:

$$f(x) = \begin{cases} 1, & x \leq \mu - 0.1 \\ e^{1-\mu}, & \mu \leq x \leq \mu + 0.1 \\ e^{1-x}, & x \geq \mu + 0.1 \end{cases} \quad (6)$$

由于分类不明确, 希望模型能够学习优化边界附近的样本往往遭受较大的损失。此类样本数量比较少, 因此, 需要为困难样本分配更高的权重。首先使用参数 μ 将样本分为正样本和负样本, 再使用加权函数 Slide 对边界样本进行强调, 如图 8 所示。

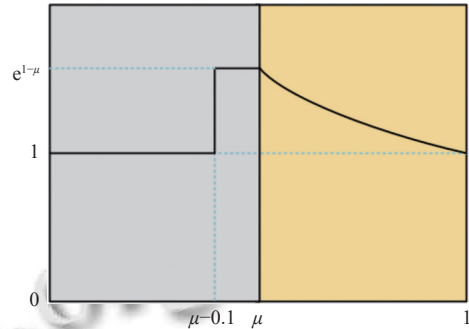


图 8 SlideLoss 函数

改进后的 YOLOv8n 算法结构如图 9 所示。

3 实验分析

本文使用东北大学开源的 NEU-DET 钢材表面缺陷数据集对算法进行训练测试, 验证算法的有效性。

3.1 实验准备

3.1.1 数据集

NEU-DET 数据集共由 1800 张图片组成, 缺陷种类分为裂纹 (Crazing, Cr)、夹杂 (Inclusion, In)、斑块 (Patches, Pa)、麻点 (Pitted surface, Ps)、压入氧化铁皮 (Rolled-in scale, Rs)、划痕 (Scratches, Sc), 如图 10 所示。每种缺陷类别 300 张图片, 本次实验分训练集 1440 张, 测试集与验证集各 180 张。

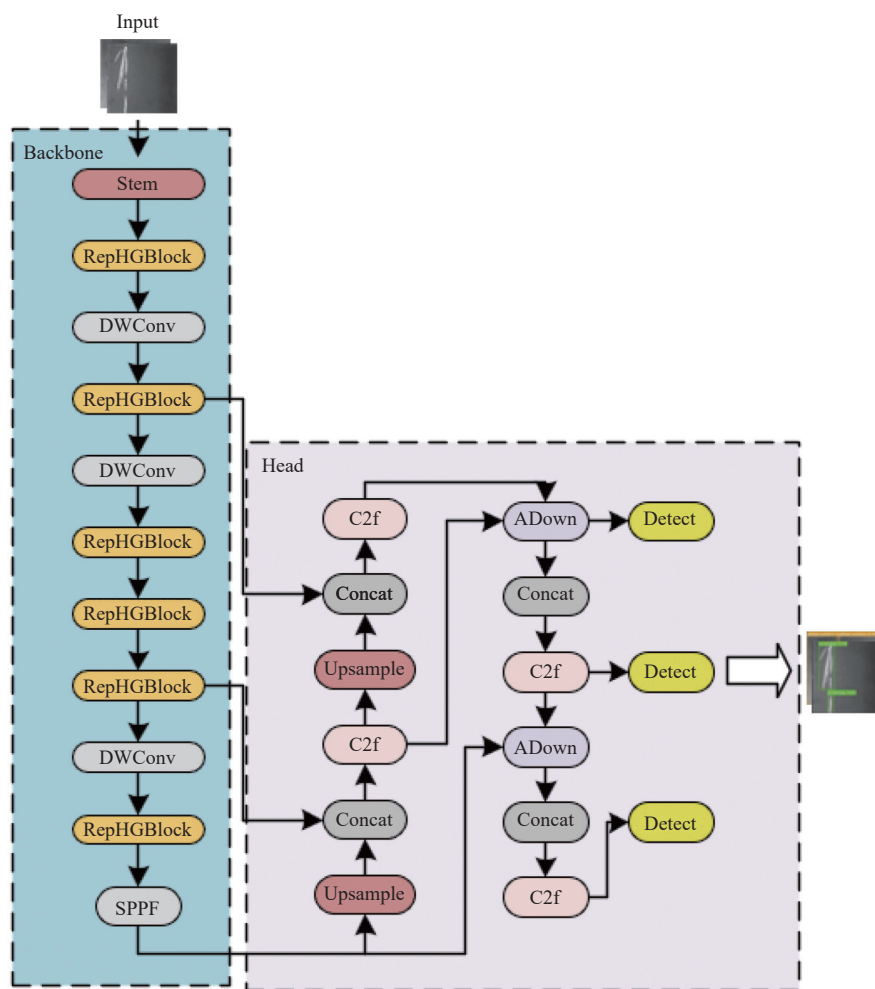


图9 改进后的YOLOv8n网络结构图

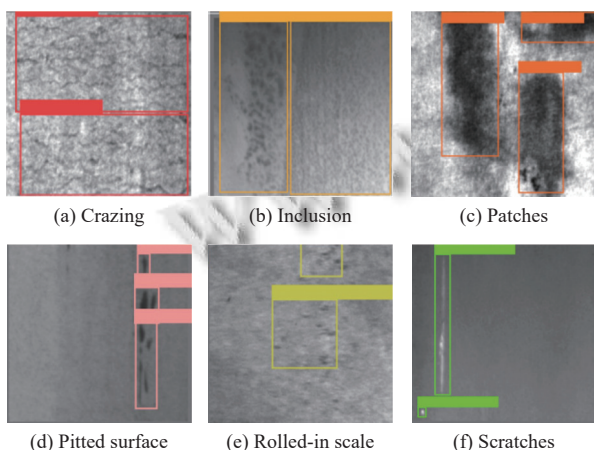


图10 NEU-DET数据集缺陷类型

3.1.2 实验环境及参数配置

实验环境及配置参数分别如表1和表2所示. 为确保实验的公平性, 不设置预训练权重.

表1 实验环境

配置	设置
CPU	i9-13900HX 2.20 GHz
GPU	NVIDIA RTX 4060
RAM	DDR4 8 GB
操作系统	Windows 11
部署环境	Python 3.10.11
深度学习框架	PyTorch 2.0.0
加速计算架构	CUDA 11.7

表2 实验参数配置

参数	设置
输入图像尺寸	640×640
参数学习率	0.01
动量	0.937
权重衰减系数	0.0005
IoU阈值	0.5
batch-size	16
epoch	200

3.1.3 实验评测标准

本文使用 5 种指标对算法进行评估, 分别为平均精度 (mAP)、精确率 (Precision, P)、计算量 (FLOPs)、刷新率 (FPS)、参数量 (Parameters, PA)。

在缺陷检测中常用 mAP 对模型性能进行评价, 公式如下:

$$mAP = \frac{\sum_{i=0}^n AP(i)}{n} \quad (7)$$

其中, n 为图像的种类, i 为检测次数, AP 为单一类别的识别平均准确率。本次实验使用 $mAP@0.5$, 计算每类图片的 AP , 再对所有种类取平均值。

Precision 表示模型检测的准确程度, 指在所有预测的正样本中, 预测正确的比例。公式如下:

$$P = \frac{TP}{TP + FP} \quad (8)$$

其中, TP (true positive) 表示模型预测的正例, FP (false positive) 表示模型预测为正例的反例。

FLOPs 表示模型复杂程度计算量, 在模型运行时所涉及的浮点运算量, 单位为 G。

FPS 表示模型检测图片的数量, 公式如下:

$$FPS = \text{Framenum} / \text{ElapsedTime} \quad (9)$$

其中, Framenum 表示图片的数量, ElapsedTime 表示模型检测时间。

Parameters 表示模型的参数量, 单位为 M, 公式如下:

$$PA = (K_h \times K_w \times C_m) \times C_{out} + C_{out} \quad (10)$$

其中, C_m 与 C_{out} 为输入和输出模型特征图的通道数, K_h 与 K_w 为卷积核的宽和高。

3.2 消融实验

为了更加直观展现改进算法在钢材表面缺陷方面的性能, 以标准的 YOLOv8n 算法的结果为基准方法与依次加入改进点后的算法模型结果进行比较, 结果如表 3 所示。

表 3 消融实验结果

组别	YOLOv8	RepHGnetv2	ADown	Slide Loss	mAP (%)	P (%)	FLOPs (G)	FPS	PA (M)
1	√	—	—	—	78.1	67.1	8.1	222	6.3
2	√	√	—	—	81.6	64.2	6.9	250	5
3	√	√	√	—	82.1	64.7	6.7	222	4.7
4	√	√	√	√	84.8	76.4	6.7	227	4.7

在表 3 中, 根据第 1 组与第 2 组实验结果显示, 原网络在使用 RepHGnetv2 重建网络后, mAP 由 78.1% 提升至 81.6%, 虽然 Precision 有所降低但 FLOPs 由 8.1G 降低至 6.9G, Parameters 下降 0.7M。结果表明重建后的主干网络可以提取更丰富的特征信息, 同时有效减少模型的参数量与计算量。第 2 组与第 3 组实验结果对比显示, 在把 Head 部分的卷积操作替换为 ADown 模块后, mAP 从 81.6% 提升至 82.1%, Precision 提升 0.5%, FLOPs 下降 0.2, Parameters 下降 0.3M。结果表明, ADown 模块可以有效降级参数量, 减少计算冗余, 微量提升检测精度。根据第 3 组与第 4 组实验, 将原算法的损失函数替换为 SlideLoss 函数后, mAP 由 82.1% 提升至 84.8%, Precision 从第 1 组的 67.1% 提升至 76.4%, 结果表明在添加 SlideLoss 函数后, 模型整体性能得到有效提升。

本文提出的改进的 YOLOv8n 算法在 NEU-DET 数据集上 mAP 达到了 84.8%, 计算量下降 17.3%, 参数量下降 25.3%, FPS 具有一定提升。图 11 为改进算法检测中缺陷的效果对比图, 其中第 2 行为原算法的检测效果, 第 3 行为改进算法的检测效果。根据图中效果对比可知, 改进算法在所有缺陷类别数据集检测中均优于原检测算法。其中在 Cr 类别与 Pa 类别中, 改进算法均检测出原算法未检测到的区域。因此本文提出的改进算法可以有效检测钢材表面缺陷, 并且适用于在设备受限的条件检测。

3.3 不同轻量级主干网络对比实验

为进一步证明本文提出的改进算法的有效性, 在 NEU-DET 数据集上与经典 GNN 网络 EfficientViT^[28]、GhostNet^[29]、MobileNetV3^[30]、2023 年提出新型主干网络 FasterNet^[31], 并于提出的主干网络 RepHGnetv2 的原型 HGnetv2 进行性能对比。结果如表 4 所示。

在表面缺陷检测中, 希望精确度与 FPS 具有较高的值, 计算量与参数量也要尽可能小。根据表 4 中结果, RepHGnetv2 网络在所有指标中均优于 YOLOv8n 原算法, RepHGnetv2 网络在精确度方面高于 EfficientViT 20.6%, 然而 FPS 比 GhostNet 高 12.6%, 参数量相对于 FasterNet 减少了 35%。根据 HGnetv2 与 RepHGnetv2 的结果显示在添加 RepConv 后 mAP 与 FPS 得到提升, 其他指标均没有改变, 证明 RepConv 可以有效提高模型检测精度与检测速度。

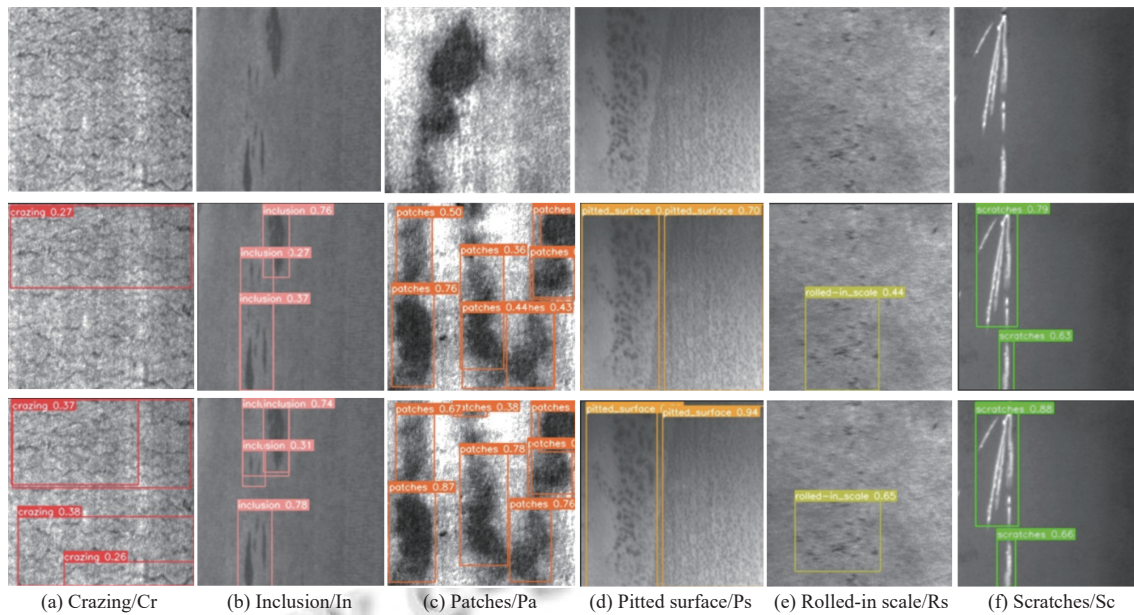


图 11 检测缺陷效果对比图

表 4 与其他网络模型比较结果

主干网络	<i>mAP</i> (%)	FLOPs (G)	<i>FPS</i>	<i>PA</i> (M)
YOLOv8n	78.1	8.1	222	6.3
EfficientViT	61.0	6.5	70	5.2
GhostNet	78	5.5	222	4.4
MobileNetV3	78.7	5.4	243	4.7
FasterNet	77.3	10.7	125	8.6
HGnetv2	80.2	6.9	243	5.0
RepHGnetv2	81.6	6.9	250	5.0

3.4 主流目标检测算法对比实验

为了评估改进算法的综合性能,与主流的目标检测算法在 NEU-DET 数据集上进行对比实验,包括 Faster-RCNN、SSD、RetinaNet、YOLOv3、YOLOv4、YOLOv5s、YOLOv7-tiny 与 YOLOXs。对比结果如表 5 所示。

根据实验结果,在 Cr 缺陷检测方面,本文算法相比其他算法具有巨大提升,由于 Cr 缺陷的图像对比度较低且具有复杂特征与无关噪声,对检测网络造成困难,属于难例,所以其他检测算法容易漏检。改进算法在使用 SlideLoss 损失函数,可以增加对难例的关注,有效增强了对 Cr 特征的检测。Faster-RCNN 由于是两阶段网络,过程中冗余信息与无效信息过多,在参数量与计算量都未具有良好的效果,不适合实时检测任务需求。SSD 相比 Faster-RCNN 具有更小的参数量与更快的速度,但检测精度不如前者。RetinaNet 在检测精度和速度方面较为均衡,在 Cr 缺陷检测方面尚未达到所需的标准。YOLOv3 虽然在精度方面优于 SSD,然而参

数量指标明显高于 SSD,难以应用到实际检测任务。YOLOv4 相比 YOLOv3 参数量减少并检测速度提升,但检测精度有所降低。YOLOv5s 达到了 73.6% 的精确度,同时拥有 4.5M 的参数量与 75.1 的检测速度,但是在 Cr 与 Rs 缺陷检测方面仍然无法达到要求。YOLOv7-tiny 在检测速度方面要优于之前所有算法,具有良好的参数量,但是检测精度却大打折扣。YOLOX 的整体效果比较均衡,特别是在 Sc 缺陷检测方面在所有算法中最高,在检测速度方面还有提升的空间。对比 YOLOv8n 算法,本文改进算法在 Pa、Ps、Rs 和 Sc 缺陷种类检测方面取得可观的准确率,在 Cr 与 In 缺陷类型检测效果提高 26.8%、5.7%,参数量与检测速度均优于所有算法,因此,本文提出的方法具有显著优势。

3.5 通用性对比实验

为了验证改进泛化效果,可以适用不同的环境,在公开的 VOC2012 数据集进行通用性对比实验。结果如表 6 所示。

根据表 6 可知,改进算法与原算法比较,*mAP* 值从 80.8% 提升至 84.4%,*FPS* 有一定提升,虽然 Precision 有小幅度降低,但参数量减少了 24%,计算量下降了 15%。由实验结果可知,改进算法不仅可以应用于特定的缺陷检测数据集,具有通用性。

4 结论与展望

本文针对在工业实际生产中缺陷检测任务,提出

了一种多尺度融合的轻量级 YOLOv8n 钢材表面缺陷检测算法。该算法在 Backbone 部分, 引入 HGnetv2 网络架构, 有效提取了更丰富的特征信息, 将原有的 Light-Conv 卷积替换为 RepConv, 通过重参数化操作提高了算法的检测速度和效率, 同时减少了计算量和内存消耗, RepConv 的引入能够更有效地捕获和提取图像特征, 从而改善了目标检测任务的准确性和效率。在 Head

部分, 采用 ADown 下采样模块, 使网络能够更好地适应不同的数据场景; 利用 SlideLoss 损失函数, 增强了网络对难例的学习能力, 实现了更加准确的检测和定位。实验结果表明, 相比 YOLOv8n 及其他检测算法, 改进算法在精确度, 检测速度, 参数量和计算量等方面都有显著提升并有通用性。改进后的 YOLOv8n 算法仍有提升空间, 需要进一步改进研究。

表5 主流检测算法对比实验

算法	mAP (%)	Parameters (M)	FPS	AP (%)					
				Cr	In	Pa	Ps	Rs	Sc
Faster-RCNN ^[7]	76.4	108	27.2	47.2	70.7	76.2	85.6	66.4	82.5
SSD ^[8]	68.1	24.4	63.7	12.5	64.1	29.6	74	45	57.1
RetinaNet ^[9]	74.7	32.1	34.6	39.4	82.6	92.6	82.7	61.8	89.1
YOLOv3 ^[15]	75.3	61.5	26.4	59.8	64	89.7	93.2	60.9	84.3
YOLOv4 ^[16]	67.8	35.1	42	34.7	76.8	87.4	76.5	50.3	81.6
YOLOv5s	73.6	4.5	75.1	53.5	61.4	87.7	93.1	63.5	83.3
YOLOv7-tiny ^[32]	66.4	6	165	39	86.6	79.2	66.8	55.9	71.3
YOLOXs ^[33]	75.3	9	42.1	36.7	82.2	92.8	88.3	57.6	94.3
YOLOv8n	78.1	4.7	222	46.3	69.6	97.7	99.5	71.4	91.6
本文算法	84.8	4.7	227	74.1	75.3	97.2	99.5	71.4	91.6

表6 VOC2012 数据集实验对比结果

算法	mAP (%)	Precision (%)	Parameters (M)	FPS	FLOPs (G)
YOLOv8n	80.8	57.8	6.3	263	8.1
改进算法	84.4	55.9	4.8	270	6.9

参考文献

- 李少波, 杨静, 王铮, 等. 缺陷检测技术的发展与应用研究综述. 自动化学报, 2020, 46(11): 2319–2336. [doi: 10.16383/j.aas.c180538]
- Wang Y, Wang HY, Xin ZH. Efficient detection model of steel strip surface defects based on YOLO-V7. IEEE Access, 2022, 10: 133936–133944. [doi: 10.1109/ACCESS.2022.3230894]
- Tabernik D, Šela S, Skvarč J, et al. Segmentation-based deep-learning approach for surface-defect detection. Journal of Intelligent Manufacturing, 2020, 31(3): 759–776. [doi: 10.1007/s10845-019-01476-x]
- Jiao LC, Zhang F, Liu F, et al. A survey of deep learning-based object detection. IEEE Access, 2019, 7: 128837–128868. [doi: 10.1109/ACCESS.2019.2939201]
- Padilla R, Netto SL, da Silva EAB. A survey on performance metrics for object-detection algorithms. Proceedings of the 2020 International Conference on Systems, Signals and Image Processing (IWSSIP). Niteroi: IEEE, 2020. 237–242.
- Xiao YZ, Tian ZQ, Yu JC, et al. A review of object detection

based on deep learning. Multimedia Tools and Applications, 2020, 79(33–34): 23729–23791. [doi: 10.1007/s11042-020-08976-6]

- Ren SQ, He KM, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137–1149. [doi: 10.1109/TPAMI.2016.2577031]
- Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector. Proceedings of the 14th European Conference on Computer Vision. Amsterdam: Springer, 2016. 21–37.
- Lin TY, Goyal P, Girshick R, et al. Focal loss for dense object detection. Proceeding of the 2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017. 2999–3007.
- Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016. 779–788.
- 向宽, 李松松, 栾明慧, 等. 基于改进 Faster RCNN 的铝材表面缺陷检测方法. 仪器仪表学报, 2021, 42(1): 191–198. [doi: 10.19650/j.cnki.cjsi.J2007109]
- 陈仁祥, 詹赞, 胡小林, 等. 基于多注意力 Faster RCNN 的噪声干扰下印刷电路板缺陷检测. 仪器仪表学报, 2021,

- 42(12): 167–174. [doi: [10.19650/j.cnki.cjsi.J2108477](https://doi.org/10.19650/j.cnki.cjsi.J2108477)]
- 13 Li YT, Huang HS, Xie QS, *et al.* Research on a surface defect detection algorithm based on MobileNet-SSD. *Applied Sciences*, 2018, 8(9): 1678. [doi: [10.3390/app8091678](https://doi.org/10.3390/app8091678)]
- 14 吴鸿杨, 康维新. 基于改进 RetinaNet 的多类焊接缺陷 X 射线图像检测模型. *计算机应用*, 2022, 42(S1): 100–105.
- 15 Redmon J, Farhadi A. YOLOv3: An incremental improvement. *arXiv:1804.02767*, 2018.
- 16 Bochkovskiy A, Wang CY, Liao HYM. YOLOv4: Optimal speed and accuracy of object detection. *arXiv:2004.10934*, 2020.
- 17 Kou XP, Liu SJ, Cheng KQ, *et al.* Development of a YOLO-V3-based model for detecting defects on steel strip surface. *Measurement*, 2021, 182: 109454. [doi: [10.1016/j.measurement.2021.109454](https://doi.org/10.1016/j.measurement.2021.109454)]
- 18 Yu XY, Lyu WT, Zhou D, *et al.* ES-Net: Efficient scale-aware network for tiny defect detection. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71: 3511314.
- 19 王磊, 郝涌汀, 潘明然, 等. 电力巡检中改进 YOLOv5s 的缺陷检测算法研究. *计算机工程与应用*, 2024, 60(10): 256–265.
- 20 齐向明, 董旭. 改进 YOLOv7-tiny 的钢材表面缺陷检测算法. *计算机工程与应用*, 2023, 59(12): 176–183.
- 21 张曙文, 钟振宇, 朱大虎. 基于改进 YOLOx 网络的金属齿轮表面缺陷检测方法. *激光与光电子学进展*, 2023, 60(22): 2215005.
- 22 Sun K, Xiao B, Liu D, *et al.* Deep high-resolution representation learning for human pose estimation. *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019. 5686–5696.
- 23 Zhao Y, Lv WY, Xu SL, *et al.* DETRs beat YOLOs on real-time object detection. *arXiv:2304.08069*, 2023.
- 24 Howard AG, Zhu ML, Chen B, *et al.* MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv:1704.04861*, 2017.
- 25 Ding XH, Zhang XY, Ma NN, *et al.* RepVGG: Making VGG-style ConvNets great again. *Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE, 2021. 13728–13737.
- 26 Wang CY, Yeh IH, Liao HYM. YOLOv9: Learning what you want to learn using programmable gradient information. *arXiv:2402.13616*, 2024.
- 27 Yu ZP, Huang HB, Chen WJ, *et al.* YOLO-FaceV2: A scale and occlusion aware face detector. *arXiv:2208.02019*, 2022.
- 28 Liu XY, Peng HW, Zheng NX, *et al.* EfficientViT: Memory efficient vision Transformer with cascaded group attention. *Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver: IEEE, 2023. 14420–14430.
- 29 Han K, Wang YH, Tian Q, *et al.* GhostNet: More features from cheap operations. *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020. 1577–1586.
- 30 Howard A, Sandler M, Chen B, *et al.* Searching for MobileNetV3. *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision*. Seoul: IEEE, 2019. 1314–1324.
- 31 Chen JR, Kao SH, He H, *et al.* Run, don't walk: Chasing higher FLOPS for faster neural networks. *Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver: IEEE, 2023. 12021–12031.
- 32 Wang CY, Bochkovskiy A, Liao HYM. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver: IEEE, 2023. 7464–7475.
- 33 Wang GH, Zheng H, Zhang X. A robust checkerboard corner detection method for camera calibration based on improved YOLOX. *Frontiers in Physics*, 2022, 9: 819019. [doi: [10.3389/fphy.2021.819019](https://doi.org/10.3389/fphy.2021.819019)]

(校对责编: 张重毅)