

基于 AGD-YOLO 的钢材表面缺陷检测算法

李思思, 葛华勇, 孟亚群, 吉宇凯

(东华大学 信息科学与技术学院, 上海 201620)

摘要: 钢材在制造业中应用广泛, 其表面缺陷严重影响其质量; 缺陷形态的多样性以及检测背景的干扰给现有检测模型带来挑战, 为此提出钢材表面缺陷检测算法 AGD-YOLO; 设计了自适应多尺度下采样 AMSD 模块, 该模块利用膨胀卷积与多池化操作捕捉并融合多尺度特征信息, 并集成到主干网络中以提升缺陷识别能力; 结合膨胀卷积和空间金字塔池化提出了增强型全局语义空间池化金字塔 EGC-SPP 模块以整合全局背景与边缘特征; 设计了双流融合网络 DSFN 增强特征表示, 提升算法上下文互补性和细节特征的识别能力; 实验结果表明, 相较于基线算法, 改进后算法在 NEU-DET 数据集上的检测精度提升了 7.1% mAP, 有效解决了钢表面缺陷检测的难点。

关键词: YOLO; 缺陷检测; 钢材; 膨胀卷积; 下采样

Algorithm for Steel Surface Defect Detection Based on AGD-YOLO

LI Sisi, GE Huayong, MENG Yaqun, JI Yukai

(College of Information Science and Technology, Donghua University, Shanghai 201620, China)

Abstract: Steel is widely used in the manufacturing industry, and its surface defects significantly affect its quality. The diversity of defect patterns and background interference in the detection pose a challenge to existing detection models. To address this, a surface defect detection algorithm for steel, AGD-YOLO, is proposed. An adaptive multi-scale downsampling (AMSD) module is designed, which uses dilated convolutions and multi-pooling operations to capture and integrate multi-scale feature information, and is incorporated into the backbone network to enhance defect recognition capability. An enhanced global context-space pooling pyramid (EGC-SPP) module is proposed by combining dilated convolutions and spatial pyramid pooling to integrate global background and edge features. A dual-stream fusion network (DSFN) is designed to enhance feature representation, improving the algorithm's context complementarity and the ability to recognize detailed features. Experimental results show that, compared to the baseline algorithm, the improved model achieves a 7.1% mAP increase in detection accuracy on the NEU-DET dataset, effectively solving the difficulties in detecting surface defects on steel.

Keywords: YOLO; defect detection; steel; dilated convolution; downsampling

0 引言

钢材作为一种重要的结构材料, 因其高耐磨性和低成本等优异特性, 在汽车制造、建筑施工等多个行业中发挥着不可或缺的作用, 钢材生产质量对于经济发展具有重要意义。然而, 由于受生产技术和环境条件的影响, 钢材表面的划伤、夹杂等缺陷不可避免, 严重影响了带钢的生产效率和产品质量^[1]。检测钢材表面缺陷是保证生产高质量钢材的关键步骤, 是工业生产领域中必不可少的环节。由于照明条件、背景环境和缺陷类型的

多样性和复杂性, 钢材表面缺陷检测通常需要耗费大量的人力成本。人工缺陷检测过程不仅耗时, 还严重依赖于操作者的主观经验。因此, 自动化钢材缺陷检测技术至关重要, 但其仍然是一项具有挑战性的任务。

近年来, 人工智能算法已广泛应用于自动化缺陷检测, 通常可以分为两大类: 传统的基于机器学习的目标检测^[2]和基于深度学习的目标检测^[3]。基于机器学习的目标检测容易受到外部环境的影响, 例如照明条件和背景噪声, 并且识别的准确性严重依赖于特征工程, 导致鲁棒性和模型泛化能力差。而深度学习具有强大的学习

收稿日期: 2024-10-30; 修回日期: 2024-12-09。

基金项目: 国家自然科学基金(62372100)

作者简介: 李思思(2001-), 女, 硕士。

通讯作者: 葛华勇(1976-), 女, 博士, 副教授。

引用格式: 李思思, 葛华勇, 孟亚群, 等. 基于 AGD-YOLO 的钢材表面缺陷检测算法[J]. 计算机测量与控制, 2025, 33(11): 142

- 148.

能力, 且消除了机器学习的冗长过程, 被视为机器学习的一种进化形式^[4]。

钢铁表面缺陷检测算法作为深度学习的重要下游应用之一, 已得到广泛研究。根据算法的框架结构, 可分为单阶段算法和双阶段算法, 前者主要包括 YOLO 系列^[5-7]、CenterNet^[9]和 SSD^[8], 后者主要包括 Mask R-CNN^[10]和 Fast R-CNN^[11]。文献 [12] 针对 Faster R-CNN 在缺陷检测中的性能不足问题, 提出了一种基于改进 Faster R-CNN 的新方法, 其引入了引导锚点区域提议网络和多任务特征金字塔网络结构, 改进了锚点位置和形状预测, 增强了特征融合, 但检测速度较慢。由于工业级钢表面缺陷检测对实时性有较高要求, 近几年研究更侧重于单阶段算法。文献 [13] 在 YOLOv5s 的基础上引入了渐进细化再分布金字塔网络, 帮助网络从复杂背景中区分缺陷。文献 [14] 通过构建融合边界框中心位置的 CIoU 损失函数, 提升了 YOLOX 算法在目标框定位中的精度。文献 [15] 使用对比度受限的自适应直方图均衡化作为数据增强方法, 提高缺陷图像的对比度, 突出带钢表面图像的缺陷特征, 进一步在 YOLOX 算法基础上引入协调注意模块来取代主干网络的空间金字塔池结构, 有效提高了网络定位缺陷的能力。文献 [16] 构建了一个精细残差双向聚合网络, 通过优化和调整提取的特征, 增强了双向传播能力, 并有效复用高、低层特征, 提升了特征聚合能力。

尽管在钢表面缺陷检测算法的改进上已有大量研究并取得了显著进展, 但在实际应用中依旧面临挑战, 主要原因有两方面: 1) 钢铁表面缺陷形状多样且不规则, 相邻缺陷之间也没有规律性。2) 生产环境中的光线不均匀、光反射、噪声以及运动模糊等因素也对钢表面缺陷检测造成了困难。在不稳定的成像环境下, 缺陷图像可能出现低对比度或高对比度的问题, 甚至在相同缺陷的情况下, 成像条件不佳会导致类内变化大的缺陷被误检。

YOLOv7 算法^[17]在多个应用场景中取得了显著成果, 但针对上述检测过程中存在的漏检和误检问题解决能力仍显不足。本文提出了一种改进的 YOLOv7 算法, 以提高钢表面缺陷检测的准确性, 主要研究内容包括以下几个方面:

1) 设计了增强型全局语义空间池化金字塔 (EGC-SPP, enhanced global context-space pooling pyramid) 提升算法对高级特征图处理能力。该模块增强空间池化金字塔的上下文理解能力的同时扩大特征图的感受野, 从而提高对类内变化较大的缺陷检测能力;

2) 设计了自适应多尺度下采样模块 (AMSD, adaptive multi-scale downsampling block) 增强主干网络特征表达能力, 在保留全局信息的同时突出显著特征;

3) 提出了双流融合网络 (DSFN, dual-stream fusion network) 优化颈部网络特征融合机制, 通过促进高层语义特征与低层空间细节交互, 提升特征表达的丰富性。

1 YOLOv7 算法原理

YOLOv7 是一种实时目标检测算法, 适用于中对工业生产线上的产品缺陷进行实时检测和监控。该算法设计了扩展高效层聚合网络 (E-ELAN, extended efficient layer aggregation networks) 作为核心模块, 通过群卷积增加特征维度, 提升特征图的学习能力, 并控制梯度的最短与最长路径, 使网络能够学习更丰富的特征。YOLOv7 分为 4 个部分, 分别是输入端、主干网络、颈部网络和检测头。

输入层主要负责调整输入图像的尺寸, 将其标准化为 640×640 像素的三通道格式。主干网络对经输入端处理后的特征图进行多尺度特征提取, 该网络由多个不同尺寸的卷积模块、最大池化卷积 (MPConv, max-pooling convolution) 和 E-ELAN 组成。其中 MPConv 通过池化和卷积的双重处理实现了下采样, 减少了特征损失。颈部网络对来自主干网络的 3 个不同层级特征图进行深度特征融合, 该网络设计了 SPPCSPC 模块, 通过引入大规模的残差分支来优化特征提取, 降低计算量, 并有效扩展感受野。检测头引入了重新参数化的卷积, 针对颈部网络输出的 3 种不同尺度的特征图进行通道匹配优化, 该模块能够自适应选择最优的先验框, 最终, 网络输出目标的类别预测、置信度评分及对应的锚框位置。

2 改进的 YOLOv7 算法

2.1 增强型全局语义空间池化金字塔

在工业环境中, 由于光照不足和噪声干扰, 钢材表面缺陷常常表现为低对比度特征, 这显著增加了目标特征提取的难度。因此, 在低对比度场景下提升算法的缺陷识别能力已成为钢材表面缺陷检测领域的核心挑战之一。

YOLOv7 算法的主干网络负责提取 3 个关键特征层用于对象检测, 分别是中间层、中低层和底层。底层特征图能够捕捉丰富细节, 但这些细节在经过逐层处理后会逐渐衰减。因此, 在特征融合过程中充分利用底层特征图的信息变得尤为重要。SPPCSPC 模块作为处理底层特征图的首要模块, 发挥着至关重要的作用。受文献 [18] 中对 SPP 模块改进的启发, 本文设计了 EGC-SPP 模块替代原 YOLOv7 算法中的 SPPCSPC 模块, 以增强全局显著信息的提取能力。

原 SPPCSPC 模块主路径通过三层最大池化操作, 提取出包含局部显著信息的特征图。但在检测低对比度

缺陷时, 仅关注边缘信息而忽视背景信息, 往往会导致部分关键语义信息的缺失。因此新增全局平均池化层和全局最大池化层以同时整合全局背景和边缘信息, 从而辅助网络从复杂背景中更精确的识别边缘信息。如图 1 所示, 新增了两条路径, 分别执行全局最大池化和全局平均池化操作。两条路径的输出分别记为 u 和 v , 经过 $\exp$ 函数处理后被扩展至与输入特征相同的尺度。

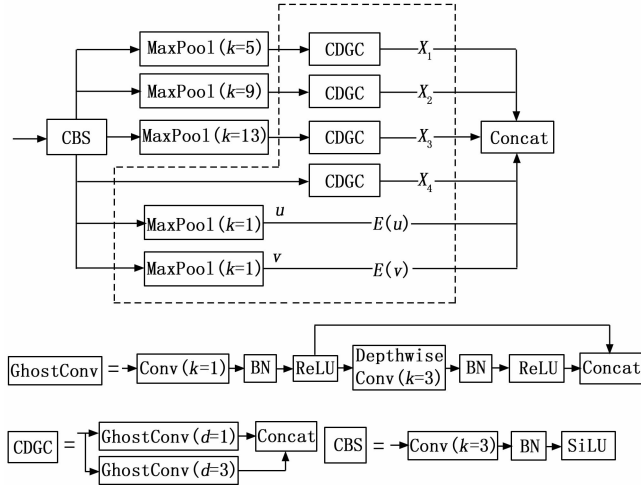


图 1 EGC-SPP 模块结构图

为进一步提升算法对类内变化大的缺陷的识别能力, 本文引入空洞卷积^[19]来处理不同池化程度的特征图。空洞卷积可以在不降低分辨率的情况下扩大感受野, 但在深层网络中, 它会增加计算量和资源消耗。为了减轻这一负担, 本文在 GhostConv 模块的 3×3 卷积中引入了膨胀系数 d , 通过调整 d 的大小来控制感受野的扩展程度。为了充分发挥空洞卷积的优势, 本文设计了联合空洞幽灵卷积模块 (CDGC, combined dilated ghost convolution)。如图 1 所示, 该模块将膨胀系数分别设置为 1 和 3 的两个 GhostConv 生成的特征图在通道维度上进行拼接。通过这种设计, CDGC 模块能够有效整合多尺度特征, 充分利用现有特征图中的显式信息, 提升算法鲁棒性。最终 EGC-SPP 模块的输出公式如式 (1) 所示:

$$Y = C[X_1, X_2, X_3, X_4, E(u), E(v)] \quad (1)$$

式中, X_1 、 X_2 和 X_3 分别表示经过最大池化层 (核大小分别为 5、9 和 13) 以及 CDGC 模块处理所生成的特征图, X_4 则跳过了最大池化操作, 从而保留更多的空间信息, 这对于检测较小缺陷尤为重要。 $E(u)$ 和 $E(v)$ 分别表示经过 $\exp$ 函数处理后的全局最大池化和全局平均池化输出的结果。最终, 将 6 个特征层 X_1 、 X_2 、 X_3 、 X_4 、 $E(u)$ 和 $E(v)$ 沿通道维度拼接, 生成的多尺度特征图 Y 包含全局背景和显著信息, 这有助于网络更有效地识别低对比度的缺陷特征。

2.2 自适应多尺度下采样模块

在 YOLOv7 算法的主干网络中, MPConv 模块包括两个下采样分支。第一个分支通过一个 1×1 卷积层调整通道数量, 在经过一个 3×3 卷积层实现进一步的下采样; 第二个分支则先通过一个池化核尺寸 k 为 2 的最大池化层进行初步下采样, 随后使用一个 1×1 卷积层调整通道数量。然而, 仅依赖最大池化和 3×3 卷积进行下采样可能导致在特征图提取过程中部分细节信息的丢失。

为解决这一问题, 本文提出了一种自适应多尺度下采样模块 AMSD, 旨在提升算法提取特征的多样性和细节保留能力, 其结构图如图 2 所示。AMSD 包含两个下采样分支: 第一个分支通过一个膨胀系数为 3 的 3×3 膨胀卷积进行下采样并减少通道数, 在不增加计算成本的情况下扩大感受野, 从而捕获更多全局信息, 提升对不同尺度目标的检测能力; 第二个分支首先通过一个 1×1 卷积调整通道数, 接着采用 chunk 操作将特征图分为两部分, 分别使用核大小为 3 的最大池化和平均池化操作。这种双重池化设计有效地融合了局部和全局信息, 最大池化能够保留显著特征, 而平均池化则有助于平滑噪声并保持细节。最后, 将这两部分的特征图进行融合, 以提高特征表达的多样性与完整性。相较于 MPConv, AMSD 模块通过引入膨胀卷积和多池化操作, 提升了算法的适应性, 有助于在低对比度场景下提高检测效果。

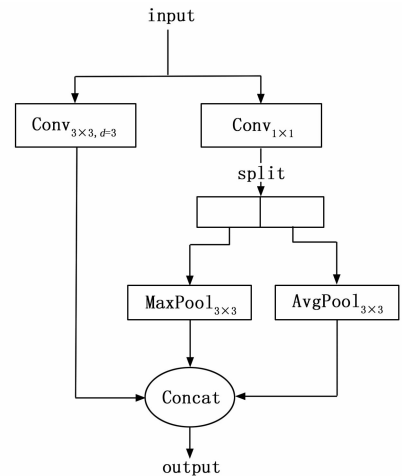


图 2 AMSD 模块结构图

本文在主干网络的多个部分引入了自适应多尺度下采样模块 AMSD, 如图 3 所示。通过将主干网络中的所有 MPConv 模块替换为 AMSD, 并且将步长为 2 的 3×3 卷积一并替代, 增强了主干网络提取不同尺度的细节信息能力。

2.3 双流融合网络

特征金字塔网络 (FPN, feature pyramid network)

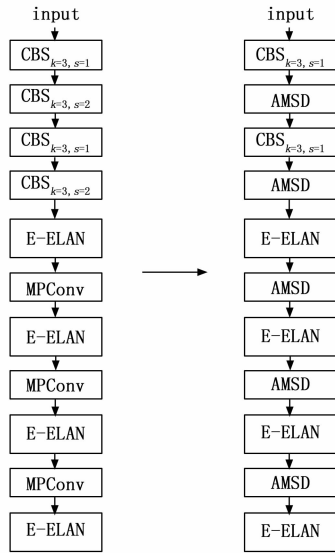


图 3 Backbone 改进对比图

通过横向连接与自上而下的路径结构, 在网络内部构建了一个特征金字塔, 将不同层次的特征进行有效融合, 使各个尺度的特征具备更加丰富的语义信息与空间细节^[22]。FPN 的引入对目标检测领域产生了深远的影响, 充分展现了多尺度特征的关键作用。YOLOv7 算法引入 PANet^[21] 通过自底向上融合特征以实现可靠的信息传递, 其架构如图 4 所示。

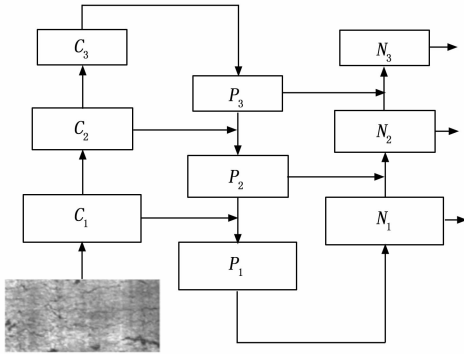


图 4 YOLOv7 算法架构图

图中, C_1 、 C_2 和 C_3 分别表示从主干网络提取的不同尺度特征图; P_1 、 P_2 和 P_3 则是对这些特征图进行融合后的结果; N_1 、 N_2 和 N_3 是经过路径增强和特征聚合后的特征图, 用于输入到检测头中进行目标检测。YOLOv7 算法直接将特征图 P_i ($i=1, 2, 3$) 与 N_{i+1} 的下采样图 N_i 进行拼接, 由于没有考虑跨尺度特征之间的未对准, 可能会导致不同尺度特征间的混淆, 进而影响模型性能及细粒度识别能力, 随着骨干网跨尺度融合的深入, 混叠效应愈发明显, 进一步加剧算法信息冗余和特征表达的模糊性。为应对特征融合过程中的混叠效应与信息冗余问题, 本文提出了双流融合网络 DS-

FN, 该网络引入了双线性融合机制。双线性融合机制最早由在文献 [22] 提出用于细粒度视觉识别领域, 旨在通过模块化分离提升特征的表达能力。通过双线性池化生成的紧凑表示, 双线性卷积神经网络不仅实现了局部与全局信息的高效融合, 还能减少特征的高维冗余, 使网络在多样化的视觉任务中表现优异。DSFN 引入双线性融合机制后能更有效地整合多尺度特征, 在提升算法判别力的同时减少信息冗余, 从而增强了特征表达的丰富性和精确度, 其结构如图 5 所示。

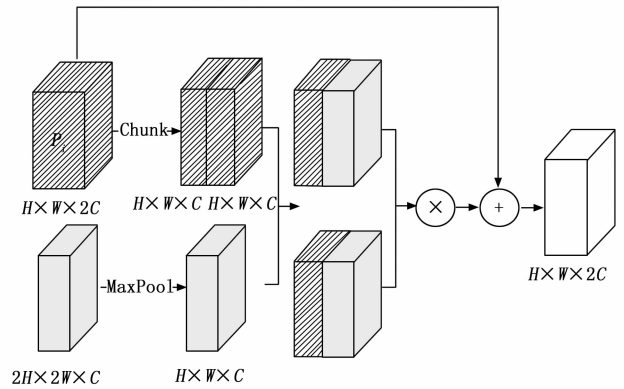


图 5 DSFN 架构图

首先, 将特征图 P_i 沿通道维度均分为两个子特征图 x_0 , x_1 , 并将其与下采样后的特征图 N_i 拼接, 生成两个新的融合特征图。接着, 对这两个融合特征图使用 1×1 卷积操作压缩通道维度后, 再通过逐元素乘积实现双线性信道融合, 生成新的下采样图。该融合过程生成的下采样特征图 y 与原始特征图 P_i 进行拼接, 最终得到 DSFN 模块的输出。上述过程的数学表达如式 (2) 所示:

$$\begin{aligned} x_0, x_1 &= \text{Chunk}(P_i) \\ y &= \text{Concat}(N_i, x_0) \cdot \text{Concat}(N_i, x_1) \\ \text{DSFN} &= \text{Concat}(P_i, y) \end{aligned} \quad (2)$$

通过对特征图 P_i 进行通道维度的切分、拼接及与特征图 N_i 的双线性融合, 本模块实现了有效的特征交互与多尺度信息融合。尽管池化后的特征图 N_i 能够保留局部区域的显著特征, 但其在一定程度上忽略了特征的具体空间位置信息。因此, 需结合特征图 P_i 恢复原始的空间信息, 尤其是细粒度的位置信息和边缘细节特征。通过双线性融合机制, 有效解决了特征融合过程中可能出现的冗余信息问题, 利用逐元素的点对点运算进一步增强了有用特征的表达, 同时也加强 P_i 内各通道之间的信息交互, 提升了特征图的多样性和辨别能力, 优化了下采样特征图 N_i 的表达效果, 最后将优化后的特征图 N_i 与 P_i 进行拼接, 用于为 E-ELAN 模块的提供更丰富的特征输入。最终本文提出的 AGD-YOLO 算法结构图如图 6 所示。

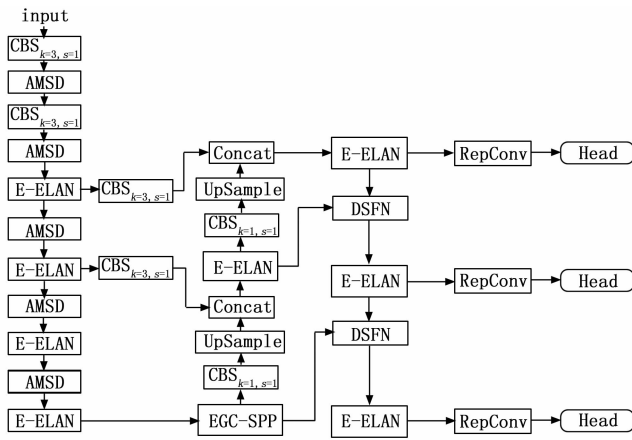


图 6 AGD-YOLO 算法架构图

3 实验结果与分析

3.1 实验数据集

本文实验使用的是 NEU-DET 钢材缺陷数据集^[23]，共包含 1800 张灰度图像，涵盖 6 种典型的表面缺陷，每类缺陷均有 300 张样本。这 6 类缺陷分别为裂纹 (Cracking)、夹杂物 (Inclusion)、斑块 (Patches)、点蚀 (Pitted Surface)、氧化皮 (Rolled-in Scale) 以及划痕 (Scratches)。本实验将数据集按照 8 : 1 : 1 的比例随机划分为训练集、验证集和测试集。为了验证 YOLOv7 模型中各改进模块的效果，分别在该数据集上进行了消融实验和对比实验。实验过程中严格控制所有其他变量，以确保结果的公平性和实验的有效性。

3.2 实验评估指标

实验性能评价指标有以下 5 个，分别是：召回率 (R, recall)、平均精度 (AP, average precision)、平均精度均值 (mAP, mean average precision)、每秒传输帧数 (FPS, frames per second) 和用于衡量模型的计算复杂度的计算量 (FLOPs)。精度 P 、召回率 R 、平均精度 AP 、平均精度均值 mAP 和每秒传输帧数 FPS 的计算公式分别如式 (3) ~ (6) 所示：

$$P = \frac{TP}{TP + FP} \times 100\% \quad (3)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (4)$$

$$AP = \int_0^1 P(R) dR \times 100\% \quad (5)$$

$$mAP = \frac{\sum_{i=1}^n AP_i}{n} \times 100\% \quad (6)$$

$$FPS = \text{FrameNum} / \text{ElapsedTime} \quad (7)$$

式中， FP (False Positive) 表示模型将负样本错误识别为正样本的数量， FN (False Negative) 表示模型将正样本错误识别为负样本的数量， TP (True Positive)

则表示模型正确检测到的正样本数量， n 为数据集类别数；FrameNum 为图片总数，ElapsedTime 为检测图片所用总时间。

3.3 实验环境配置

实验使用的硬件设备为 NVIDIA GeForce 3090 GPU，操作系统为 Linux。实验环境搭建在 Python 3.11 版本下，采用了 PyTorch 2.0.1 作为深度学习框架，并结合 CUDA 11.8 的加速计算架构。在模型训练过程中，训练轮次 (epoch) 设定为 300，批样本大小 (batch size) 设置为 16，动量参数设定为 0.937。

3.4 改进点消融实验

为验证 EGC-SPP 模块、AMSD 模块和 DSFN 在提升模型检测性能方面的作用，本文在 NEU-DET 数据集上进行了消融实验，结果如表 1 所示。该实验旨在评估各模块的贡献，以进一步分析其对模型性能的改进效果。

表 1 改进点消融实验

算法	$R/\%$	$mAP/\%$	FPS
YOLOv7	72.7	76.5	132
YOLOv7+EGC-SPP	71.3	79.4	125
YOLOv7+AMSD	77.6	78.5	121
YOLOv7+EGC-SPP+AMSD	72.4	80.5	119
YOLOv7+EGC-SPP+DSFN	78.1	81.1	124
本文算法	79.3	83.6	99

首组为原 YOLOv7 算法检测结果，作为后续 5 组实验的对照基准，用以评估各改进模块的性能提升效果。通过改进 YOLOv7 算法金字塔池化层为 EGC-SPP 模块后， mAP 由 76.5% 提升至 79.4%，这验证了该模块提升了对底层特征图的利用率，以及整合全局背景信息对于检测低对比度缺陷的关键作用，但 EGC-SPP 模块的引入增加了计算复杂度，导致 FPS 下降了 7。改进主干网络的下采样模块为 AMSD 对提升算法对小目标缺陷的检测能力有着重要作用，召回率 R 和平均精度均值 mAP 分别提升了 4.9 个百分点和 2.0 百分点，这是由于该模块弥补了原算法在下采样过程中丢失的不同尺度的细节信息，增强了网络对不同大小目标的感知能力。联合引入 EGC-SPP 和 AMSD 模块后，模型的 mAP 提升了 4.0 个百分点，相较于单独应用任一模块的效果均有提升，且 FPS 仅下降了 13，低于分别应用 EGC-SPP 和 AMSD 模块时 FPS 下降之和，表明两模块的共同作用在提升检测精度的同时，减小了计算开销，但召回率 R 相比于只改进主干网络的下采样模块时暂时有所降低。为此进一步引入双流融合网络 DSFN， R 和 mAP 分别提升了 6.9 个百分点和 3.1 百分点，其中召回率 R 有着显著提升，主要归因于该网络采用双线性融合操作，增强了有用特征的表达能

从而提升了算法的特征提取能力。通过第二组、第四组、第五组和第六组实验结果可以得出, AMSD 模块与 DSFN 模块之间存在显著的协同效应, 这两者的联合应用能够有效提升模型的特征提取和融合能力。其中, AMSD 模块通过自适应多尺度下采样增强了对不同尺寸目标的感知能力, 而 DSFN 模块则进一步优化了高层语义信息与低层细节信息的结合, 显著提升了模型的整体检测性能。

综上所述, 本文提出的 3 种改进方法均有效提升算法检测缺陷能力, 相较于原 YOLOv7 算法, 改进后算法召回率提高了 6.6 百分点, mAP 提高了 7.1 百分点, 尽管 FPS 下降了 33, 但该幅度仍处于可接受范围内, 能够满足工业环境实时检测需求, 同时检测精度和召回率得到显著提升, 确保了改进算法在精度和实时性之间的良好平衡。

3.5 改进效果对比实验

为直观展示改进算法在钢材表面缺陷检测中的效果, 本文对比了改进前后模型在测试集上的检测表现, 部分检测结果见图 7。

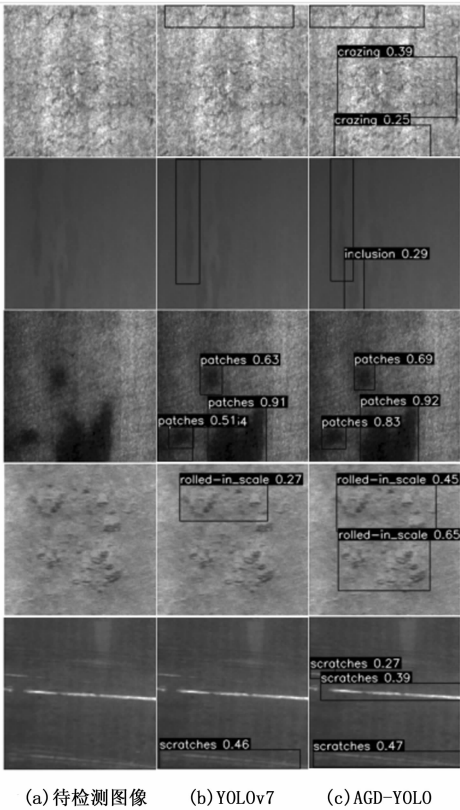


图 7 YOLOv7 算法和 AGD-YOLO 算法检测效果对比

从图中可以看出, 在检测裂纹、夹杂物、点蚀、氧化皮和划痕等缺陷时, 原 YOLOv7 算法存在不同程度的漏检问题, 而通过本文提出的改进模型, 漏检情况得到了有效改善, 检测的召回率显著提升。此外,

本文提出的算法在多组实验中均表现出更高的检测置信分数。例如, 在第四组实验中, YOLOv7 算法对氧化皮缺陷的检测置信度为 0.27, 而本文算法提升至 0.45, 明显优于原算法的检测性能。总体实验结果表明, 本文改进后的算法在缺陷检测的召回率和置信度方面均有显著提升, 有效提高了对钢材表面缺陷的检测精度与可靠性。

3.6 不同模型对比实验

为进一步验证所提出算法的有效性, 本研究在 NEU-DET 数据集上与其他目标检测模型进行了对比实验。参与对比的模型包括 SSD、Faster-RCNN、YOLOX 和 YOLOv8, 实验结果如表 2 所示。

表 2 不同算法对比实验

算法	$mAP/\%$	$R/\%$
Faster-RCNN	69.1	59.1
SSD	62.1	56.0
YOLOX	74.9	70.4
YOLOv7	76.5	72.7
YOLOv8	79.4	74.9
文献[24]	80.3	72.2
文献[25]	81.8	75.4
本文算法	83.6	79.3

根据实验结果可以看出, 本文提出的算法在 mAP 指标和 R 指标上表现最佳, 分别达到了 83.6% 和 79.3%, 显著高于双阶段目标检测算法 Faster-RCNN 和单阶段目标检测算法 SSD、YOLOX、YOLOv7 和 YOLOv8。与较新的 YOLOv8 算法相比, 本文算法在 mAP 和 R 指标上分别提升了 4.2 个百分点和 4.4 个百分点, 表现出更优的性能。进一步与其他 YOLOv7 改进算法进行比较, 本文算法在 mAP 指标上分别高于文献 [24] 和文献 [25] 3.3 个百分点和 1.8 个百分点。这表明, 本文提出的改进算法不仅显著提升了缺陷检测精度, 还有效降低了漏检率。综上所述, 本文提出的改进算法在多个评价指标上均优于对比算法, 展示了较强的实际应用潜力。

4 结束语

本文结合钢材表面缺陷检测的需求, 提出了基于 YOLOv7 的改进模型。该模型通过构建自适应多尺度下采样 AMSD 模块, 实现了不同尺度特征的高效捕捉与融合, 从而提升了模型对小目标缺陷的敏感度。引入的双流融合网络 DSFN 优化了高层语义信息与低层空间特征的交互, 增强了模型对复杂背景干扰的鲁棒性。此外, EGC-SPP 模块通过膨胀卷积与空间金字塔池化相结合, 有效提升了对多尺度上下文信息的提取能力, 使得缺陷检测更加精准。实验表明, 在针对算法下采样信

息丢失问题进行优化后,改进后的模型在召回率和检测精度方面显著提升。未来,模型将进一步结合去噪与亮度增强模块,以提高模型在多样化工业环境中的适应性与应用价值。

参考文献:

- [1] WANG H, ZHANG J, TIAN Y, et al. A simple guidance template-based defect detection method for strip steel surfaces [J]. *IEEE Transactions on Industrial Informatics*, 2018, 15 (5): 2798–2809.
- [2] WANG A, SHA M, LIU L, et al. A new process industry fault diagnosis algorithm based on ensemble improved binary-tree SVM [J]. *Chinese Journal of Electronics*, 2015, 24 (2): 258–262.
- [3] HUSSAIN N, ATTIQUE KHAN M, TARIQ U, et al. Multiclass cucumber leaf diseases recognition using best feature selection [J]. *Computers, Materials & Continua*, 2022, 70 (2): 3281–3294.
- [4] ALZUBAIDI L, BAI J, AL-SABA AWI A, et al. A survey on deep learning tools dealing with data scarcity: definitions, challenges, solutions, tips, and applications [J]. *Journal of Big Data*, 2023, 10 (1): 1–82.
- [5] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 779–788.
- [6] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection [EB/OL]. (2024-12-04) [2020-04-23]. <https://arxiv.org/abs/2004.10934>.
- [7] GE Z, LIU S, WANG F, et al. YOLOX: exceeding yolo series in 2021 [EB/OL]. (2024-12-04) [2021-07-18]. <https://arxiv.org/abs/2107.08430>.
- [8] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [J]. *Lecture Notes in Computer Science*, 2016: 21–37.
- [9] DUAN K, BAI S, XIE L, et al. CenterNet: keypoint triplets for object detection [C] // *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019: 6569–6578.
- [10] HE K, GKIOXARI G, DOLLAR P, et al. Mask R-CNN [C] // *Proceedings of the IEEE International Conference on Computer Vision*, 2017: 2961–2969.
- [11] GIRSHICK R. Fast R-CNN [C] // *IEEE International Conference on Computer Vision*, 2015: 1440–1448.
- [12] 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.
- [13] YU X, LYU W, WANG C, et al. Progressive refined re-distribution pyramid network for defect detection in complex scenarios [J]. *Knowledge-Based Systems*, 2022, 260: 110176.
- [14] 熊聪, 于安宁, 高兴华, 等. 基于改进 YOLOX 的钢材表面缺陷检测算法 [J]. *电子测量术*, 2023, 46 (9): 151–157.
- [15] CHEN H, DU Y, FU Y, et al. DCAM-Net: a rapid detection network for strip steel surface defects based on deformable convolution and attention mechanism [J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 1–12.
- [16] LIANG Y, LI J, ZHU J, et al. A lightweight network for defect detection in nickel-plated punched steel strip images [J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 1–15.
- [17] 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*, 2023: 7464–7475.
- [18] SU P, HAN H, LIU M, et al. MOD-YOLO: Rethinking the YOLO architecture at the level of feature information and applying it to crack detection [J]. *Expert Systems with Applications*, 2024, 237: 121346.
- [19] YU F, KOLTUN V. Multi-Scale Context Aggregation by Dilated Convolutions [EB/OL]. (2024-12-04) [2015-11-23]. <https://arxiv.org/abs/1511.07122>.
- [20] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature Pyramid Networks for Object Detection [C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2017: 2117–2125.
- [21] LIU S, QI L, QIN H, et al. Path Aggregation Network for Instance Segmentation [C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018: 8759–8768.
- [22] LIN T Y, ROYCHOWDHURY A, MAJI S. Bilinear CNN models for fine-grained visual recognition [C] // *Proceedings of the IEEE International Conference on Computer Vision*, 2015: 1449–1457.
- [23] HE Y, SONG K, MENG Q, et al. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features [J]. *IEEE Transactions on Instrumentation and Measurement*, 2019, 69 (4): 1493–1504.
- [24] 高春艳, 秦 臻, 李满宏, 等. 改进 YOLOv7 算法的钢材表面缺陷检测研究 [J]. *计算机工程与应用*, 2024, 60 (7): 282–291.
- [25] 孙卫波, 丁 卫. 改进 YOLOv7 的带钢表面缺陷检测算法 [J]. *工业控制计算机*, 2024, 37 (8): 94–96.