

改进 YOLOv10 的钢材表面缺陷检测方法^{*}

吴 葛 朱宇凡 叶天成

(华中科技大学电气与电子工程学院 武汉 430074)

摘要: 针对工业系统现有钢材表面缺陷检测模型较老,准确率有限,存在错判漏判的问题,提出了一种改进 YOLOv10 的钢材表面缺陷检测方法 FAA-YOLO。该方法通过引入轻量化网络 FasterNet 和多尺度注意力机制 EMA,设计了 C2f_Faster_EMA 模块,实现 Backbone 网络轻量化与特征提取能力增强的平衡;在 Backbone 网络末端添加自适应细粒度通道注意力机制 AFGC,增强模型骨干网络对输入特征的初步提取能力,进而提高模型的检测精度;将 Neck 部分换为注意力尺度序列融合框架 ASF,提升模型整合多尺度特征信息的能力。在 NEU-DET 钢材表面缺陷数据集上进行对比实验和消融实验,实验结果表明,提出的 FAA-YOLO 模型较基线 YOLOv10n 模型参数量降低 11.01%,计算量降低 7.69%,检测精度提高 2.9 个点,达到 83.6% 的检测准确率,在降低模型复杂度的同时实现了较高的检测准确率,在工业系统中具有高可用性与高实时性。

关键词: 钢材表面缺陷检测;YOLOv10;FasterNet;EMA;AFGC;ASF-YOLO

中图分类号: TP391.4; TN791 文献标识码: A 国家标准学科分类代码: 520.6040

Improving the steel surface defect detection method of YOLOv10

Wu Ge Zhu Yufan Ye Tiancheng

(School of Electrical & Electronic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China)

Abstract: Aiming at the issues of outdated existing models for steel surface defect detection, limited accuracy, and the existence of misjudgment and omission, an improved YOLOv10 method for steel surface defect detection, named FAA-YOLO, is proposed. This method introduces the lightweight network FasterNet and multi-scale attention mechanism EMA, designing the C2f_Faster_EMA module to balance the lightweight of the backbone network with enhanced feature extraction capabilities. An adaptive fine-grained channel attention mechanism AFGC is added at the end of the backbone network to enhance the preliminary feature extraction ability of the model's backbone network, thereby improving the model's detection accuracy. The Neck part is replaced with an attention scale sequence fusion framework ASF to enhance the model's ability to integrate multi-scale feature information. Comparative experiments and ablation experiments on the NEU-DET steel surface defect dataset show that the proposed FAA-YOLO model reduces the number of parameters by 11.01%, the computational load by 7.69%, and increases the detection accuracy by 2.9 percentage points, achieving a detection accuracy of 83.6%. This method reduces the complexity of the model while achieving high detection accuracy, demonstrating high usability and real-time performance in industrial systems.

Keywords: steel surface defect detection; YOLOv10; FasterNet; EMA; AFGC; ASF-YOLO

0 引言

在现代工业生产中,钢材质量控制的准确性直接关系到企业的竞争力。钢材表面的缺陷,如划痕、污染和凹陷等,是决定产品质量的关键因素。因此,对钢材表面进行严格的质量检测显得尤为关键^[1]。早期的钢材表面缺陷检测技术,如频闪法、红外检测法、漏磁检测法、涡流检测法和激

光扫描检测法等,虽然在实践中有所应用,但它们在自动化程度和精确性方面存在一定的局限性^[2]。传统的计算机视觉方法,例如基于图像处理的检测技术,通过图像采集设备获取钢材表面的图像,并运用边缘检测、阈值分割、形态学操作等图像处理算法来识别和定位缺陷。这些方法虽然能够提供直观的缺陷检测结果,但在复杂环境中的适应性和对微小缺陷的识别能力仍有限。

收稿日期:2024-11-13

* 基金项目:教育部 2020 年第二批新工科研究与实践项目(E-NYDQHGC20202219)资助

随着深度学习技术的兴起,一种新的缺陷检测方法应运而生。这种方法通过深度神经网络自动学习和识别产品表面的缺陷特征,相较于传统方法,它在自动化程度、实时性和检测精度方面都有了显著提升^[3]。目前,基于深度学习的缺陷检测方法已经成为钢材表面缺陷检测领域的新趋势,为提高钢材产品质量提供了强有力的技术支持^[4]。

当前深度学习目标检测算法主要分为两大类:一阶段算法和二阶段算法^[5]。二阶段算法首先生成候选区域,然后对这些区域进行分类和定位,典型的二阶段算法包括 R-CNN^[6],Fast R-CNN^[7] 和 Faster R-CNN^[8] 等。虽然二阶段算法在精度上通常表现较好,但它们需要更多的计算资源,并且检测速度相对较慢,这使得它们不太适合部署在计算资源受限的平台以及实时性要求较高的应用场景,如钢材表面缺陷检测。而以 YOLO^[9-10],SSD^[11-12] 为代表的一阶段算法在目标检测中表现出更快的响应速度和较高的准确率。这些算法直接在图像上预测类别和边界框,无需先生成候选区域,因此在处理速度上具有优势。由于这些特点,一阶段算法更加适合应用于钢材表面缺陷检测的场景,特别是在需要快速检测和实时反馈的工业生产线上^[13]。

针对钢材表面缺陷检测问题,陶友凤等^[14]提出了一种基于改进 YOLOv8 的钢材表面缺陷检测方法,采用 R_HGNet 替换 YOLOv8 骨干网络,提升模型轻量化,并采用 C2f-DWR 模块融合浅层与深层特征,最后引入 Inner-CIOU 损失函数提高对小目标的定位准确性,获得较好的检测效果。然而 R_HGNet 轻量化特征提取网络在处理更复杂的缺陷模式时会大幅限制模型的表示能力,因此不适用于对检测准确率要求高的钢材表面缺陷检测场景。敖思铭等^[15]提出了一种基于改进 YOLOv5 的钢材表面缺陷检测方法,首先使用空洞空间金字塔池化模块获取更大感受野,然后采用 K-means 算法聚类得到更匹配的锚框,并采用 SIOU 损失函数进一步提升模型的定位能力,获得了较为准确的检测结果。但是,空洞空间金字塔池化和 SIOU 损失函数的引入增加了模型的计算开销,而 K-means 的敏感性可能导致局部最优解,这些因素使得该方法不适用于实时性要求高的场景。Wang 等^[16] 采用多尺度探索块(multi-scale explore block)和空间注意力机制(spatial attention mechanism)改进 YOLOv5,在降低计算开销的同时取得了较为准确的目标检测率。然而,以 YOLOv5、YOLOv8 作为改进的基线模型相对较旧,目前已有一系列更新、检测效果更好的 YOLO 系列基线模型。针对 YOLOv10 改进,黄毅等^[17]引入多通道逆残差模块(multi-channel inverted residual module, MCIR)、CBAM 注意力机制和 Focal-EIoU 损失函数改进 YOLOv10 模型,虽然增强了特征的表达能力与利用率,但是 MCIR 会大幅提升计算冗余,特别是在模型的深层部分,并且 Focal-EIoU 损失函数直接将边长作为惩罚项,在目标尺寸泛化强的钢材表面缺陷检测中会错误的放大边长的影响,使得模型准

确率下降。杨海涛等^[18]采用膨胀卷积(dilated convolution2f, DC2f)拓展感受野、双向特征金字塔(bidirectional feature pyramid network, BiFPN)实现多尺度特征融合以及 BAM 注意力机制改进 YOLOv10s,模型的大目标检测能力和目标区域聚焦能力得到提升,但是 DC2f 增加模型计算量的同时会捕捉过于细节的特征造成过拟合,BiFPN 则使模型更难学习到泛化的特征表示,这在钢材表面缺陷检测中难以接受。李军等^[19]通过部分卷积、Repulsion-IoU 损失函数和三层特征提取结构 AFPN-3 改进 YOLOv10,在天台光伏涉电区域行为人预警的特定场景下表现良好,但是 Repulsion-IoU 损失函数针对遮挡目标设计,缺乏不同应用场景的泛化能力,而部分卷积和三层特征提取结构 AFPN-3 增加了模型的复杂性与对融合策略的依赖性。此外,上述基线及改进模型计算开销仍然较大,且检测准确率有限,无法满足钢材表面缺陷检测对准确性与实时性的高需求。

综上所述,本文针对钢材表面缺陷检测中缺陷尺寸分化强、模型参数大、泛化能力弱以及实时性要求高等挑战。选取 YOLOv10 作为基线模型,针对其本身计算冗余、网络架构效率有限、特征权重分配不够精细和小目标识别能力不足和已有改进泛化能力弱、模型复杂度高和过拟合风险大等问题,提出了一种改进 YOLOv10 的检测方法。本文贡献如下:1)网络架构优化:采用轻量化网络 FasterNet 中的 FasterBlock 替换 C2f 中的 Bottleneck,并引入 EMA 注意力机制设计 C2f-Faster-EMA 模块,以替换 Backbone 网络中的 C2f 模块,有效提高模型的每秒浮点运算(floating point operations per second,FLOPs)效率;2)引入自适应细粒度通道注意力机制(adaptive fine-grained channel attention, AFGC):这种机制通过捕获全局和局部信息之间的相关性,并在不同粒度级别上促进它们的交互,以实现更有效的特征权重分配;3)ASF 注意力尺度序列融合框架:在 Neck 网络引入 ASF 注意力尺度序列融合框架,结合空间和尺度特征进行目标识别,提高模型在处理小目标和密集目标时的准确性。实验结果证明,本文方法在降低模型计算开销的同时提高了钢材表面缺陷的检测准确率,不仅有助于提高工业生产中钢材质量监控的效率,也为确保工业系统的安全稳定运行提供了有力支持。

1 YOLOv10 模型

YOLOv10 模型,由敖思铭等^[15]于 2024 年提出,它保留了 YOLO 系列单阶段检测算法的核心特点,能够在一次前向传递中预测图像中所有目标的类别和边界框。YOLOv10 在继承前代版本优势的基础上,通过模型架构和训练策略的优化,实现了速度、精度和模型大小之间更优的平衡。

YOLOv10 的关键创新点是其无需非极大值抑制(non-maximum suppression,NMS)的训练策略。传统目标检测

模型在推理时通常依赖 NMS 来去除冗余检测框, 这不仅增加了计算负担, 还可能导致错误地抑制有效检测框被。YOLOv10 采用了新颖的双标签分配机制, 结合了一对多和一对一的标签分配策略, 这不仅丰富了监督信号, 还避免了在推理过程中使用 NMS, 显著提升了模型的推理速度和效率。此外, YOLOv10 还引入了一致性匹配公式, 确保了一对一标签分配策略能够提供更高质量的训练样本, 从而进一步提高模型性能。在模型设计方面, YOLOv10 进行了精心优化。包括采用轻量化的分类头和空间与通道解耦的下采样方式, 旨在降低计算成本并提升运行效率。YOLOv10 还引入了紧凑逆残差模块 (compact inverted block, CIB) 模块, 这是一种高效的卷积块, 有助于进一步降低模型的计算负担。同时, 通过大核卷积和部分自注意力 (partial self-attention, PSA) 机制的应用, YOLOv10 在保持低计算成本的同时, 显著提高了检测精度。

凭借其创新的训练策略和优化的模型设计, YOLOv10 为实时目标检测任务提供了一个高效、准确的解决方案, 特别是在对小目标检测准确率和实时性要求较高的场景中, YOLOv10 展现出了卓越的性能。

但是, YOLOv10 基线模型存在以下缺点需要改进: 1) 层的设计不够优化, 如 Backbone 部分在特征提取中存在重复计算, 导致明显的计算冗余; 2) Backbone 部分对不同特征的学习能力不同, 导致特征权重分配不均; 3) Neck 部分难以捕捉到图像中小目标区域足够的特征信息, 造成小目标识别能力弱。

基于上述优缺点, 本文采用 YOLOv10 作为改进的基线模型, 其网络结构图如图 1 所示。

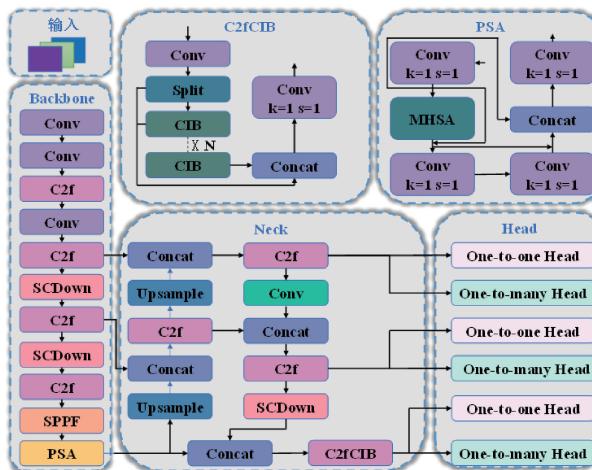


图 1 YOLOv10 网络结构图

Fig. 1 Architecture of YOLOv10

2 改进 YOLOv10 模型

本文针对 YOLOv10 的网络架构效率低、特征权重分配不均和小目标识别能力不足这 3 个重要缺点, 结合目前

已有的 YOLOv10 改进泛化能力弱、模型复杂度高和过拟合风险大的问题做出如下改进。

2.1 C2f-Faster-EMA 模块

针对钢材表面缺陷检测中检测精度不足以及模型参数量、计算量过大难以部署和 YOLOv10 网络架构效率低的问题, 本节中对 YOLOv10 骨干网络的 C2f 进行优化设计, 使其具有更高的检测精度及更轻量的网络结构。

1) FasterNet 轻量化网络

现有的 YOLO 轻量化改进模型通常致力于减少 FLOPs 量, 以提升模型的实时性。然而, 这种减少并不总是能带来预期的延迟降低, 因为操作符频繁的内存访问, 尤其是深度卷积, 可能会导致 FLOPs 的效率不高。超轻量级网络 FasterNet 引入了创新的部分卷积 (partial convolution, PConv) 技术, 这一技术不仅保持了模型的低复杂度, 而且还通过减少冗余计算和内存访问, 增强了模型对空间特征的提取能力^[20]。

PConv 利用特征图中的冗余信息, 通过仅对输入特征图一部分通道应用卷积核, 从而有效减少计算量和内存访问。对于一个特征输入图, PConv 的计算公式如下:

$$\left\{ \begin{array}{l} \mathbf{I} \in R^{c \times h \times w} \\ \gamma = \frac{c_p}{c} \\ FLOPs = h \times w \times k^2 \times c_p \\ MemoryAccess = h \times w \times 2c_p \end{array} \right. \quad (1)$$

式中: \mathbf{I} 为输入特征图, c 为通道数, h 为高度, w 为宽度; c_p 为参与卷积计算的通道数, γ 为部分卷积率; k 为卷积核的大小, $FLOPs$ 为浮点运算量; $MemoryAccess$ 为内存访问量。

由式(1)可知, 对于同样的特征输入图 I , 相比全卷积, 使用 FasterNet 的浮点运算量和内存访问量显著减小, 实现高速运行和高准确性。

FasterNet 由 4 个层次化的阶段组成, 每个阶段都由一系列 FasterNet 块堆叠而成, 以有效提取特征, 其整体架构如图 2 所示。每个 FasterNet 块包含一个 PConv 层和两个逐点卷积 (pointwise convolution, PWConv) 层, 形成了类似于倒置残差结构的设计。PConv 层主要负责提取空间特征, 而 PWConv 层则负责融合特征, 这使得网络能够在减少计算量的同时保持强大的特征表达能力。4 个层次化的阶段, 每个阶段之前都设有一个嵌入层或合并层, 它们负责进行空间下采样和通道数扩展, 帮助网络在不同尺度上捕获特征, 同时降低计算复杂度。此外, FasterNet 在每个中间 PWConv 层之后采用批量归一化和激活函数, 其中批量归一化促进了网络的训练过程, 而激活函数引入了非线性, 增强了模型的表达能力。得益于这种架构, FasterNet 能在运行速度和准确性之间取得优秀的平衡性。

2) EMA 多尺度注意力机制

在各种计算机视觉任务中, 通道或空间注意力机制对

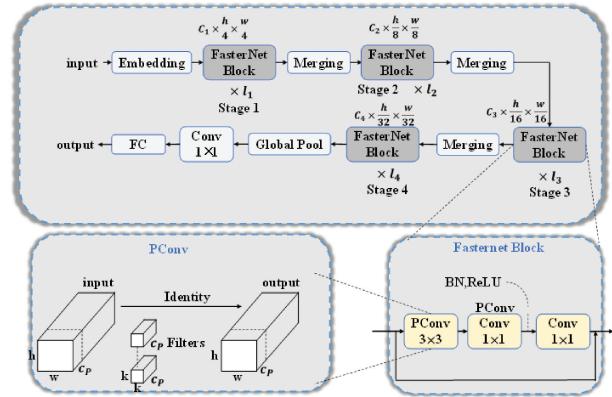


图 2 FasterNet 架构图

Fig. 2 Architecture of FasterNet

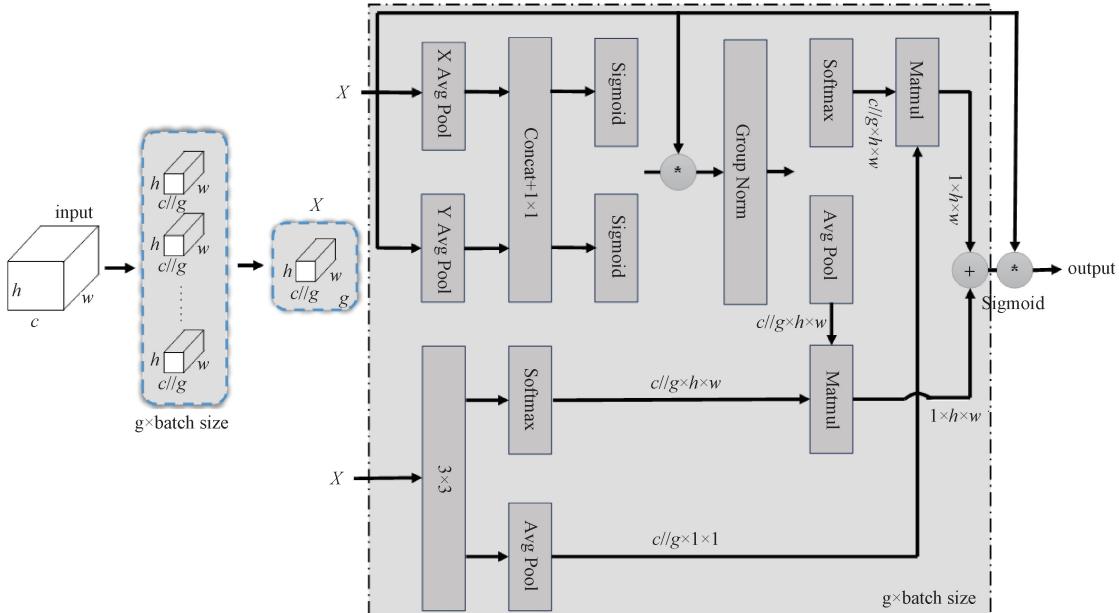


图 3 EMA 架构图

Fig. 3 Architecture of EMA

在 1×1 卷积分支中, 输入特征图 X 分别沿水平和垂直方向进行 1D 全局平均池化, 得到两个平均特征向量后, 将两个平均特征向量拼接后通过 1×1 卷积进行处理, 来提取紧凑的特征表示; 3×3 卷积分支直接应用于输入特征图, 捕获更广泛的上下文信息, 进而保留更多的空间细节。跨空间学习 Matmul 对 1×1 卷积分支的输出进行 2D 全局平均池化, 然后与 3×3 卷积分支的输出进行矩阵点积操作, 以捕获像素级的成对关系, 增强特征之间的关联性。接着对每个组的输出特征图应用 Sigmoid 函数得到最终的空间注意力图, Sigmoid 函数能够将注意力权重归一化到 $[0, 1]$ 区间, 使得模型更加灵活地调整不同区域的特征重要性。最后将空间注意力图与原始输入特征图进行逐个元素乘法, 得到加权后的特征图, 且最终输出的特征图与原始输入特征图尺寸相同, 可以直接堆叠到现代架构中。EMA

于捕获更多可识别的特征具有显著的有效性, 然而, 通过通道降维来建模跨通道关系可能会对提取深度视觉表征带来副作用。EMA 是一种高效的多尺度注意力机制, 通过将部分通道重构为批处理维度, 并将通道维度分组为多个子特征, 可以保留每个通道上的信息并减少计算开销, 使得空间语义特征在每个特征组内均匀分布^[21]。除了编码全局信息以重新校准每个并行分支中的通道权重外, 还通过跨维交互进一步聚合两个并行分支的输出特征, 以捕获像素级成对关系。EMA 的网络结构如图 3 所示。

对于任意给定的输入特征图 X , EMA 将 X 分为 G 个子特征组, 使得每组能够专注于学习不同的语义特征, 这种分组方式不仅增强了模型对关键语义区域特征的学习, 还有效压缩了噪声信息。EMA 包含两个并行的卷积分支:

机制的流程可以用公式表示为:

$$\left\{ \begin{array}{l} X \in R^{C \times H \times W} \\ z_X = \frac{1}{H} \sum_{i=1}^H x_{c,i,j} \\ z_Y = \frac{1}{W} \sum_{j=1}^W x_{c,i,j} \\ a_X = \sigma[W_{1 \times 1} * concat(z_X, z_Y)] \\ A = softmax(\frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W a_{X,i,j} a_{Y,i,j}) \\ F = A \odot X \end{array} \right. \quad (2)$$

式中: X 为输入特征图, C 、 H 和 W 分别为输入特征图的通道数、高度和宽度; $x_{c,i,j}$ 为输入特征图在第 c 个通道, 第 i 行第 j 列的特征值, z_X 和 z_Y 分别为沿水平方向和垂直方向的全局平均池化结果; a_X 为经过 Sigmoid 函数处理后

的通道注意力图, σ 为 Sigmoid 函数; $W_{1 \times 1}$ 为 1×1 卷积分支的权重, $concat(z_X, z_Y)$ 为 z_X 和 z_Y 沿通道拼接的结果; A 为跨空间学习得到的空间注意力图, $softmax$ 为将输入转换为概率分布的函数, $a_{X,i,j}$ 为来自 1×1 卷积分支的注意力图在第 i 行, 第 j 列的值, $a_{Y,i,j}$ 为来自 3×3 卷积分支的注意力图在第 i 行, 第 j 列的值; F 为最终输出的特征图, \odot 为逐元素乘法算子。

本文引入 FasterNet 网络对原 YOLOv10 模型 Backbone 部分进行优化, 但直接引用会导致模型准确性下降, 不利于目标场景使用。因此, 本文仅引入 FasterBlock 结构, 将骨干网络 C2f 模块中的 Bottleneck 结构替换为 FasterBlock 得到 C2f_Faster 模块, 在此基础上添加 EMA 注意力机制设计得到 C2f_Faster_EMA 模块。原 C2f 结构如图 4 所示, 改进后的 C2f_Faster_EMA 结构如图 5 所示。

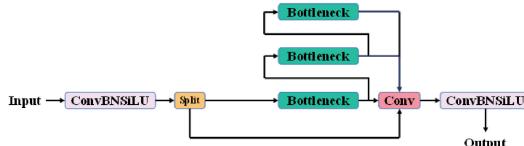


图 4 C2f 架构图

Fig. 4 Architecture of C2f

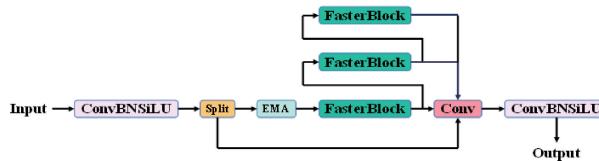


图 5 C2f_Faster_EMA 架构图

Fig. 5 Architecture of C2f_Faster_EMA

将 YOLOv10 骨干网络中的 C2f 模块全部替换为 C2f_Faster_EMA 模块, 在轻量化网络结构、降低模型复杂度的同时, 增强了骨干网络对输入钢材表面图象的特征提取能力, 有助于提升模型对钢材表面缺陷的检测准确性和实时性。

2.2 AFGC 注意力机制

在钢材表面缺陷检测中, 由于钢材表面的纹理、反光特性以及生产过程中不可预测的各种瑕疵, 背景噪声通常较大, 这对缺陷的准确检测带来了挑战。基线 YOLOv10 模型在处理这种复杂背景下的钢材表面缺陷时, 往往难以准确识别有效的缺陷信息, 造成特征权重分配不均, 进而影响了检测的准确性。为了解决这个问题, 现有研究尝试引入了多种注意力机制如 SENet (squeeze-and-excitation networks)、GAM (global attention mechanism)、CBAM (convolutional block attention module)。这些机制在一定程度上增强了模型对关键信息的关注能力。然而, 它们在提取特定异物局部特征方面存在局限性, 并且可能会增加模型的计算量, 不利于模型的轻量化。针对这些问题, 本文引入一种新的注意力机制: AFGC^[22]。该机制的核心优势在于能够有效整合全局和局部信息的权重向量, 通过自适应地分配特征通道权重, 能让模型更加灵活地关注图像

中的关键区域, 从而在不增加过多计算负担的情况下, 提升特征提取和权重分配能力, 并保持模型的轻量化。AFGC 的网络结构如图 6 所示。

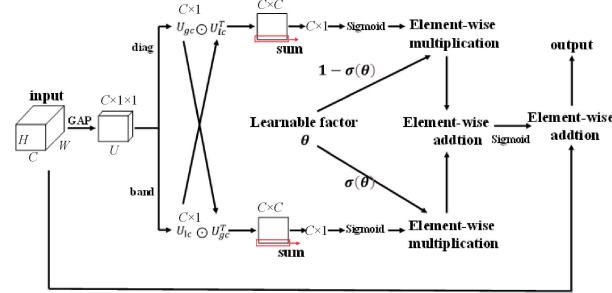


图 6 AFGC 注意力机制结构图

Fig. 6 Architecture of AFGC attention

AFGC 的计算原理可以分为以下几步。首先, 特征图通过全局平均池化转换为通道描述符, 将每个通道的空间信息压缩为单一数值, 进而提取出每个通道的全局特征。该步骤可以用公式表示为:

$$U_n = GAP(F_n) = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W F_n(i, j) \quad (3)$$

式中: $F_n(i, j)$ 为第 n 个通道特征图在位置 (i, j) 的值, H 和 W 分别代表特征图的长度和宽度, F_n 为输入特征图第 n 个通道, U_n 为通道描述符。然后进行局部通道信息、全局通道信息以及全局与局部通道信息之间的交互: 在局部通道信息交互中, 使用带状矩阵保持模型参数量较小的同时获取局部通道信息; 为了捕获所有通道之间的依赖关系作为全局信息, 使用对角矩阵; 将对角矩阵获得的全局信息与带状矩阵获得的局部信息相结合, 通过交叉相关操作捕获它们在不同粒度上的关联, 获得相关矩阵, 该步骤可以用公式表示为:

$$\begin{cases} U_{lc} = \sum_{i=1}^k U b_i \\ U_{gc} = \sum_{i=1}^c U d_i \\ M_1 = U_{gc} U_{lc}^T \\ M_2 = U_{lc} U_{gc}^T \end{cases} \quad (4)$$

式中: U_{lc} 为局部信息, b_i 为带状矩阵 B 第 i 个局部值, k 为临接通道数; U_{gc} 为全局信息, d_i 为对角矩阵 D 第 i 个全局值, c 为通道数; M_1 和 M_2 为相关性矩阵。接着从相关性矩阵及其转置矩阵中提取行和列信息作为全局和局部信息的权重向量, 并通过学习因子动态融合, 该步骤可以用公式表示为:

$$\begin{cases} U_{wg} = \sum_{j=1}^c (M_1)_{ij, ji} \\ U_{wl} = \sum_{j=1}^c (M_2)_{ij, ji} \\ W = \sigma(\theta) \sigma(U_{wg}) + [1 - \sigma(\theta)] \sigma(U_{wl}) \end{cases} \quad (5)$$

式中: U_{wg} 和 U_{wl} 分别为全局和局部通道权重融合后的权

重, σ 为 sigmoid 激活函数。最后将融合权重与输入特征图相乘得到输出,该步骤可以用公式表示为:

$$F^* = W \otimes F \quad (6)$$

式中: F^* 为输出特征图, F 为输入特征图。

本文在原 YOLOv10 骨干网络的末端 PSA 模块后添加 AFGC,进一步提取并融合 PSA 模块输出特征图中的局部和全局特征,进而提升骨干网络对输入的钢材表面图像的初步特征提取能力,有助于提高模型对钢材表面缺陷的检测准确率。

2.3 ASF 注意力尺度序列融合框架

基线 YOLOv10 模型的多尺度信息提取能力和小目标检测能力有限,针对此问题,引入特征金字塔网络(feature pyramid network,FPN)、融合注意力机制和新增小目标检测头只能在一定程度上改善此缺陷。为进一步提升 Neck 网络的信息提取能力和小目标检测能力,本文引入 ASF-YOLO 注意力尺度序列融合框架^[23]。该框架采用尺度序列特征融合(scale-sequential feature fusion,SSFF)模块来增强网络的多尺度信息提取能力,并采用三重特征编码器(triple feature encoder,TFE)模块来融合不同尺度的特征图,以增加细节信息。在此基础上,进一步引入了通道和位置注意力机制(channel and position attention mechanism,CPAM),将 SSFF 和 TFE 模块整合在一起,重点关注信息通道和与空间位置相关的小目标,以提高检测和分割性能。ASF-YOLO 框架的网络结构如图 7 所示。

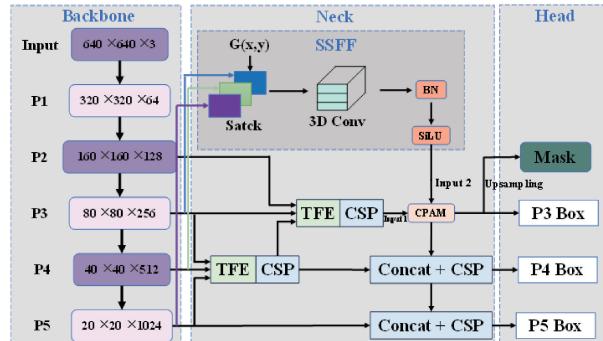


图 7 ASF-YOLO 结构图

Fig. 7 Architecture of ASF-YOLO

ASF-YOLO 框架可以结合空间和多尺度特征。在 SSFF 模块中,将从骨干网络提取的 P3,P4,P5 特征图被归一化为相同大小,经过上采样叠加在一起作为三维卷积的输入,以捕捉并结合不同空间尺度的特征。作为 SSFF 输入的缩放图像计算可以用公式表示为:

$$\begin{cases} F_\sigma(w, h) = G_\sigma(w, h) f(w, h) \\ G_\sigma(w, h) = \frac{1}{2\pi\sigma^2} e^{-\frac{w^2+h^2}{2\sigma^2}} \end{cases} \quad (7)$$

式中: $f(w, h)$ 为宽度为 w 高度为 h 的二维输入图像, $F_\sigma(w, h)$ 由使用二维高斯滤波器 $G_\sigma(w, h)$ 的一系列平滑卷积生成, σ 为用于卷积的二维高斯滤波器标准偏差的缩

放参数。

传统特征金字塔网络 FPN 的融合机制只是对小尺寸特征图进行上采样,接着将其分割或添加到上一层的特征中,这种做法忽略了大尺寸特征层丰富的细节信息。ASF-YOLO 中使用的 TFE 通过分割大、中、小尺寸特征,添加大尺寸特征图并进行特征放大,可以得到完整的特征信息。TFE 模块的结构如图 8 所示。

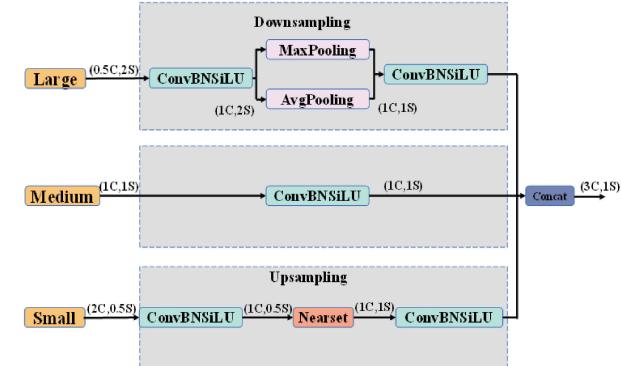


图 8 TFE 结构图

Fig. 8 Architecture of TFE

在进行特征编码前,先调整特征通道的数量,保证其与主要尺度特征通道数量相同。对于大尺寸特征图,在利用卷积模块将通道数调整为 1C 后,采用最大池化和平均池化的混合结构进行降采样,保留了高分辨率特征和缺陷图像的有效性和丰富性。针对小尺寸特征图,先使用卷积模块调整通道数为 1C,再采用最近邻插值法进行上采样,这种操作可以保留低分辨率图像特征的多样性和小目标特征信息。将尺寸相同的大、中、小特征图进行一次卷积后,在通道维度上进行拼接。该步骤可以用公式表示为:

$$F_{TFE} = \text{Concat}(F_L, F_M, F_S) \quad (8)$$

式中: F_{TFE} 为 TFE 模块输出的特征图, F_L, F_M, F_S 分别表示大、中、小尺寸的特征图, F_{TFE} 的分辨率与 F_M 相同,通道数则是 F_M 的 3 倍。

为了提取不同通道中的代表性特征信息,ASF-YOLO 利用 CPAM 来整合细节特征信息和多尺度特征信息,CPAM 模块的网络结构如图 9 所示。

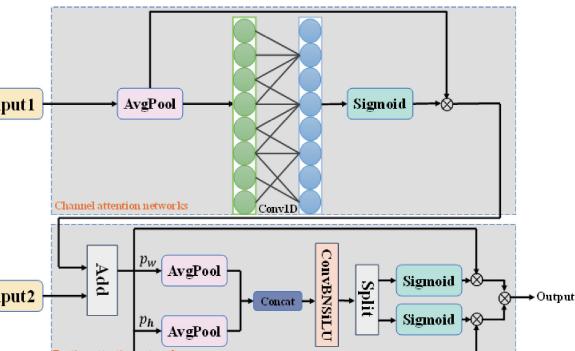


图 9 CPAM 结构图

Fig. 9 Architecture of CPAM

输入 1 是 PANet 之后的特征图, 其中保留了 TFE 的细节特征。对于输入 1, 先进行全局平均池化, 利用每个通道及其 k 个近邻来捕捉局部跨通道交互, 即使用大小为 k 的一维卷积, k 为局部跨通道交互的覆盖率, 且 k 与信道维度 C 存在映射关系, 一般情况下映射关系可以用公式表示为:

$$\begin{cases} C = \psi(k) = 2^{\gamma k - b} \\ k = \phi(C) = \left\lceil \frac{\log_2(c) + b}{\gamma} \right\rceil_{odd} \end{cases} \quad (9)$$

式中: k 为卷积核大小, C 为信道维度, $\lceil \cdot \rceil_{odd}$ 为最近邻奇数, γ 通常设为 2, b 通常设为 1。显然, 高值通道的交换时间较长, 而低值通道的交换时间较短, 因此通道注意力机制可以捕获多个通道特征。

通道注意力网络输出和 SSFF 模块输出叠加后作为输入 2, 可以提供互补信息。输入 2 在水平轴和垂直轴上分别进行全局平均池化, 减小特征图维度的同时保留重要的特征信息, 再进行合并。该步骤可以用公式表示为:

$$\begin{cases} p_w = \frac{1}{H} \sum_{0 \leq j \leq H} E(w, j) \\ p_h = \frac{1}{W} \sum_{0 \leq i \leq W} E(i, h) \end{cases} \quad (10)$$

式中: w 为输入特征图的宽度, h 为输入特征图的高度, $E(w, j)$ 和 $E(i, h)$ 是输入特征图中位置 (i, j) 的值。

对水平轴和垂直轴进行连接和卷积操作, 捕捉特征图的全局上下文信息, 生成位置注意力坐标, 该步骤可以用公式表示为:

$$P(a_w, a_h) = \text{Conv}[\text{Concat}(p_w, p_h)] \quad (11)$$

式中: $P(a_w, a_h)$ 为位置注意力坐标的输出。

同时, 在对注意力特征进行分割时, 会生成一对位置相关特征图, 进一步细化了特征图, 该步骤可以用公式表示为:

$$\begin{cases} s_w = \text{Split}(a_w) \\ s_h = \text{Split}(a_h) \end{cases} \quad (12)$$

式中: s_w 和 s_h 为分割输出的宽度和高度。

最后 CPAM 模块的输出定义可以用公式表示为:

$$F_{CPAM} = E \times s_w \times s_h \quad (13)$$

式中: E 代表通道和位置注意力的权重。

原 ASF-YOLO 框架基于 YOLOv5 进行改进并用于细胞图像的实例分割, 在 Neck 部分进行了创新的网络设计。本文借助 ASF-YOLO 框架的灵活性和可扩展性, 在基线 YOLOv10 模型的 Neck 部分引入 ASF 框架, 并针对钢材表面缺陷的检测任务做了特定优化, SSFF 和 TFE 模块提高了模型对多尺度和小对象目标识别的性能, 通道和位置注意力机制则进一步挖掘这两个模块提供的特征信息, 增强了模型对多尺度、多空间和多目标的钢材表面缺陷的适应能力, 进而提高了模型对钢材表面缺陷的检测准确率。

2.4 FAA-YOLO 模型

基于上述分析, 本文基于 YOLOv10n 提出了改进模型: FAA-YOLO 模型, 该模型结构如图 10 所示。

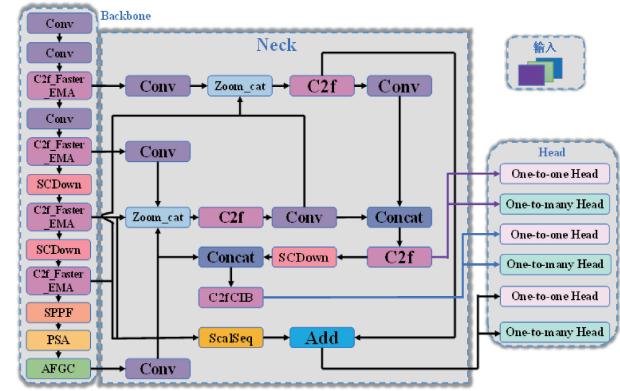


图 10 FAA-YOLO 结构图
Fig. 10 Architecture of FAA-YOLO

首先在 YOLOv10n 的 Backbone 网络中, 将 C2f 中的 Bottleneck 结构替换为 FasterBlock, 并在 C2f_Faster 模块中添加 EMA 注意力机制, 替换原 C2f 模块, 在轻量化网络结构的同时, 增强了骨干网络对输入钢材表面图象的特征提取能力。其次, 在 YOLOv10n 的 Backbone 网络末端 PSA 模块后引入 AFGC 注意力机制, 进一步提取并融合 PSA 模块输出特征图中的局部和全局特征, 有助于提高模型对钢材表面缺陷的检测准确率。最后, 在 Neck 网络中引入 ASF-YOLO 框架, 增强了模型对不同尺度特征信息的提取能力。

具体而言, FasterBlock 相较 Bottleneck 降低了模型参数量和计算量, 同时保持原有 Bottleneck 几乎相同的精度, 引入 EMA 注意力机制则可以保留每个通道上的信息并减少计算开销, 使得空间语义特征在每个特征组内均匀分布。AFGC 注意力机制有效的整合了全局和局部信息的权重向量, 并实现了对特征通道权重的自适应分配。ASF-YOLO 注意力尺度序列融合框架则重点关注信息通道和与空间位置相关的小目标, 在多尺度、多空间以及多目标的场景能够提取更多特征信息, 使模型在小目标、多噪声背景条件下的性能进一步提升。

3 实验结果与分析

3.1 实验环境

本文实验基于 Linux20.04 操作系统, 使用基于 Python3.8.0 的 Pytorch2.0.0 搭建深度学习模型。在模型训练方面, 使用 NVIDIA GeForce RTX 4090D GPU 作为模型训练平台, 并采用 CUDA11.7 对 GPU 进行加速。本文深度学习模型参数设置如表 1 所示。

3.2 实验数据集

本文实验采用东北大学采集的钢材表面缺陷(NEUDET)数据集^[24]。该数据集标注包含裂纹(crazing)、斑块

表 1 模型参数

Table 1 Model parameters

| 参数 | 设置 | 参数 | 设置 |
|------|----------|--------|------|
| 类别数量 | 6 | 学习阈值 | 0.01 |
| 输入宽 | 640 | WARMUP | True |
| 输入高 | 640 | 批大小 | 16 |
| 学习率 | 0.001 25 | 训练轮数 | 300 |

(patches)、内含物(inclusion)、点蚀表面(pitted-surface)、轧制氧化皮(rolled-in_scale)和划痕 (scratches) 共 6 类缺陷。每类缺陷 300 张共计 1 800 张,本文实验按照 8:1:1 划分为训练集、验证集和测试集。

3.3 评价指标

本文选取准确率(precision, P)、召回率(recall, R)和 IoU 为 0.5 时的平均精度均值(mean average precision, mAP)作为模型对钢材表面缺陷检测准确率的评价指标。其中, mAP 是本文检测准确率的关键评价指标。P、R、mAP 越接近 1 说明模型检测准确率越高,三者的计算公式

如下:

$$P = \frac{TP}{TP + FP} \quad (14)$$

$$R = \frac{TP}{TP + FN} \quad (15)$$

$$mAP = \frac{1}{n} \sum_{i=1}^N AP_i = \frac{1}{n} \sum_{i=1}^N \int P dR \quad (16)$$

式中: TP 为图像中检测缺陷正确的目标数量; FP 为图像中检测错误缺陷的目标数量; FN 为图像中漏检缺陷的目标数量; AP 为平均精度,即 PR 曲线下的面积; N 为目标检测类别的数量。

同时,本文选取参数量与计算量作为模型效率的评价指标,参数量与计算量越小说明模型越高效。

3.4 消融实验

为了验证本文提出的改进模型所做的各项改进对表面缺陷检测的影响,对各项改进进行消融实验。实验之间严格控制模型参数一致,消融实验结果如表 2 所示,“√”代表添加此改进。

表 2 消融实验结果

Table 2 Results of ablation experiment

| 实验组别 | C2f_Faster_EMA | AFGC | ASF | Params/ 10^6 | GFLOPs | P/% | R/% | mAP/% |
|------|----------------|------|-----|----------------|--------|------|------|-------|
| 1 | — | — | — | 2.27 | 6.5 | 84.3 | 73.2 | 80.7 |
| 2 | √ | — | — | 1.91 | 5.6 | 77.2 | 77.3 | 80.9 |
| 3 | — | √ | — | 2.33 | 6.5 | 81.5 | 77.2 | 81.5 |
| 4 | — | — | √ | 2.31 | 6.9 | 77.3 | 75.3 | 81.2 |
| 5 | √ | √ | — | 1.97 | 5.6 | 76.5 | 75.3 | 80.2 |
| 6 | √ | — | √ | 1.94 | 6.0 | 82.7 | 73.4 | 80.9 |
| 7 | — | √ | √ | 2.37 | 6.9 | 75.3 | 73.6 | 79.5 |
| 8 | √ | √ | √ | 2.02 | 6.0 | 77.1 | 78.7 | 83.6 |

由表 2 进行分析可知:在 Backbone 网络中引入 C2f_Faster_EMA 模块后,模型参数量和计算量分别下降 15.56% 和 13.85%,同时精度也有小幅度提升,这是因为该模块减少了不必要的计算和参数,使模型能更加专注于有效特征信息,表明改进后的骨干网络能够实现有效提取特征信息和模型轻量化的平衡;在 Backbone 网络末端引入 AFGC,检测精度提升了 0.8%,且模型计算量不变,参数量仅增加 2.64%,说明该模块能有效的整合特征信息,提高了特征的表达能力,从而提升了检测精度,而参数量的增加可能是由于引入了额外的权重向量来整合特征信息;在 Neck 网络中引入 ASF-YOLO 框架后,模型参数量和计算量分别增加了 1.76% 和 6.15%,检测精度提升 0.5%,表明该模块采用的额外的上下文信息和特征融合机制,能够有效提升模型的检测精度,参数量和计算量的增加可能是额外的特征融合操作导致的。注意到,同时添加 ASF-YOLO 框架和 AFGC 会导致模型检测精度降低 1.2% 的同时参数量和计算量分别增加 4.41% 和 6.15%,

这是因为 AFGC 的注意力机制与 ASF-YOLO 框架的通道注意力分配不一致,使得模型在不同尺度的特征图上分配注意力不均衡,导致有效的特征信息被抑制或噪声信息被放大。

同时引入 C2f_Faster_EMA、AFGC 和 ASF 的最终改进网络在参数量降低 11.01%,计算量降低 7.69% 的同时, mAP 提升了 2.9 个点,达到 83.6% 的准确率。实验结果说明 FAA-YOLO 模型能够捕获更丰富的特征信息,在轻量化的同时有效增强模型性能。

3.5 模型对比实验

为了证明本文提出的 FAA-YOLO 模型的整体效果,将本文提出的模型与其他目标检测算法进行对比分析,实验之间严格控制参数一致,实验结果如表 3 所示。

由表 3 可知,本模型相较二阶段架构的 Faster-RCNN 模型 mAP 提升 11.5%,并且模型参数量和计算量显著降低,相对 FasterRCNN 分别降低 95.24% 和 95.66%。相较其他一阶段目标识别算法,本文模型在参数量仅高于

表 3 不同模型实验结果对比

Table 3 Comparison of different models

| 模型 | Params/ 10^6 | GFLOPs | mAP/% | P/% | R/% |
|-------------------------------|----------------|--------|-------|------|------|
| FasterRCNN | 42.5 | 138.1 | 72.1 | 71.9 | 64.5 |
| SSD | 26.1 | 82.3 | 71.6 | 71.3 | 63.8 |
| YOLOv3-tiny | 9.52 | 14.3 | 74.5 | 66.8 | 73.2 |
| YOLOv3 | 61.5 | 156.3 | 65.3 | 69.8 | 61.2 |
| YOLOv5s | 7.0 | 17.6 | 76.6 | 74.1 | 72.4 |
| YOLOv6 | 4.16 | 11.5 | 79.2 | 74.2 | 73.4 |
| YOLOv7-tiny | 6.0 | 13.2 | 65.9 | 62.3 | 63.3 |
| YOLOv7 | 37.1 | 105.3 | 70.4 | 66.6 | 65.2 |
| YOLOv8n | 3.0 | 8.2 | 78.6 | 77.7 | 71.8 |
| ASF-YOLO ^[23] | 2.73 | 7.2 | 80.9 | 78.6 | 75.6 |
| YOLOv9-tiny | 1.96 | 6.4 | 79.8 | 77.3 | 72.6 |
| YOLOv10 | 2.27 | 6.5 | 80.7 | 84.3 | 73.2 |
| YOLOv10-AD ^[25] | 2.53 | 6.7 | 81.1 | 78.7 | 73.4 |
| REIW-YOLOv10n ^[26] | 3.18 | 8.2 | 81.0 | 75.9 | 77.6 |
| YOLO11 | 2.58 | 6.3 | 82.6 | 76.0 | 80.0 |
| 本文模型 | 2.02 | 6.0 | 83.6 | 77.1 | 78.7 |

YOLOv9-tiny, 计算量则实现一阶段模型中最低, mAP 为所有模型最高, 达到 83.6% 的检测精度。值得注意的是, 相比 YOLO11、YOLOv10-AD 和 REIW-YOLOv10n 等最新 YOLO 模型及其改进, 本文模型相比 YOLO11 参数量降低 21.70%, 计算量降低 4.76%, mAP 提升 1%; 相比 YOLOv10-AD 参数量降低 20.16%, 计算量降低 4.76%, mAP 提升 2.5%; 相比 REIW-YOLOv10n 参数量降低 36.48%, 计算量降低 36.67%, mAP 提升 2.6%, 说明本文模型在模型复杂度控制和检测能力方面均优于 YOLOv10 及其现有改进。结果表明, 本文模型在保证轻量化的同时, 显著提升了钢材表面缺陷检测的准确率, 较好地平衡了模型轻量化和性能, 在应用场景中具有低计算量、高准确性和实时性等优点。

3.6 实验结果可视化分析

为了更直观的对比本文提出的 FAA-YOLO 模型与基线 YOLOv10n 模型的检测效果, 从测试集中随机抽取部分图像进行测试, 测试结果如图 11 所示。

由图 11 可知, 本文提出的 FAA-YOLO 模型能检测出原 YOLOv10n 模型无法检测到的目标, 捕捉更低细节的缺陷。

3.7 公开数据集对比

为了验证本文提出的 FAA-YOLO 模型的泛化能力, 本文使用北京大学智能机器人开放实验室发布的印刷电路板瑕疵数据集(PKU-Market PCB)进行对比试验。选取其中适用于检测任务的 693 张裸板缺陷图片, 其中包含漏孔、鼠咬、开路、短路、杂散、杂铜共 6 类缺陷。实验设置与 3.1 节保持一致, 实验结果如表 4 所示。

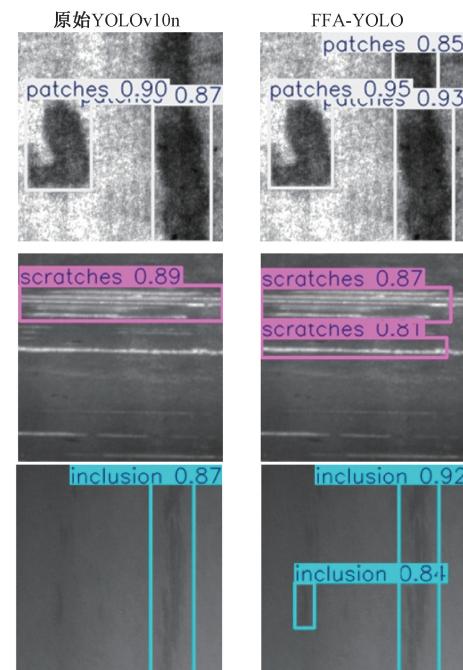


图 11 YOLOv10n 与 FAA-YOLO 检测结果对比

Fig. 11 Comparison of detection results between YOLOv10N and FAA-YOLO

由表 4 可知, 本文提出的 FAA_YOLO 模型相比 YOLOv10 及其改进, 不仅适用于钢材表面缺陷检测, 而且在 PCB 缺陷检测这类小目标数据集同样实现了检测效果增强, 说明本文模型在泛化能力增强和过拟合风险控制方面优于 YOLOv10 及其现有改进。提出的 FAA_YOLO 模

表 4 PCB 数据集实验结果对比

Table 4 Comparison of PCB Dataset Experimental Results

| 模型 | Params/ 10^6 | GFLOPs | mAP/% | P/% | R/% |
|-------------------------------|----------------|--------|-------|------|------|
| YOLOv10n | 2.27 | 6.5 | 91.0 | 91.1 | 84.7 |
| YOLOv10-AD ^[25] | 2.53 | 6.7 | 90.2 | 90.0 | 83.5 |
| REIW-YOLOv10n ^[26] | 3.18 | 8.2 | 91.3 | 90.3 | 86.6 |
| FFA_YOLO | 2.02 | 6.0 | 93.2 | 92.7 | 88.5 |

型相比基线 YOLOv10n 在模型参数量、计算量分别降低 11.01%、7.69% 的同时, mAP 增长了 2.2%。

4 结 论

针对钢材表面缺陷检测准确率有限,无法充分平衡模型检测准确性与实时性,难以满足工业系统稳定运行要求的问题,提出了一种改进 YOLOv10 的钢材表面缺陷检测方法。主要结论为:本文引入 FasterNet 和 EMA 改进 Backbone 部分,在降低模型参数量与计算量的同时提升了的检测准确率;本文在 Backbone 网络末端引入 AFGC,增强模型骨干网络对输入特征的初步提取能力,进而提高模型的检测精度;本文引入 ASF-YOLO 框架改进 Neck 部分,通过多尺度特征融合有效提升了模型的检测性能。综上所述,本文提出的改进 YOLOv10 的钢材表面缺陷检测方法相较于原始 YOLOv10n 模型及其现有改进具有更高的检测准确性与实时性,有助于钢材表面缺陷检测智能化的实现。

参考文献

- [1] 赵佰亭,张晨,贾晓芬. ECC-YOLO:一种改进的钢材表面缺陷检测方法[J]. 电子测量与仪器学报,2024,38(4):108-116.
ZHAO B T, ZHANG CH, JIA X F. ECC-YOLO: An improved method for detecting surface defects in steel [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(4): 108-116.
- [2] 周彦,孟江南,王冬丽,等. 基于多尺度轻量化注意力的钢材缺陷检测[J]. 控制与决策,2024,39(3): 901-909.
ZHOU Y, MENG J N, WANG D L, et al. Steel defect detection based on multi-scale light-weight attention [J]. Control and Decision, 2024, 39 (3): 901-909.
- [3] 黄志海,邓耀华,吴光栋. IC 器件表面缺陷多光谱图像特征融合检测方法[J]. 仪器仪表学报,2024,45(9): 24-33.
HUANG ZH H, DENG Y H, WU G D. Multispectral image feature fusion method for detecting surface defects in IC devices [J]. Chinese Journal of Scientific Instrument, 2024, 45(9): 24-33.
- [4] 程锦锋,方贵盛,高惠芳. 表面缺陷检测的机器视觉技术研究进展[J]. 计算机应用研究,2023,40 (4): 967-977.
CHENG J F, FANG G SH, GAO H F. Research progress of surface defect detection based on machine vision technology [J]. Application Research of Computers, 2023,40(4):967-977.
- [5] 张福豹,吴婷,赵春峰,等. 基于弱光增强与 YOLO 算法的锯链缺陷检测方法[J]. 电子测量技术,2024,47(6): 100-108.
ZHANG F B, WU T, ZHAO CH F, et al. Saw chain defect detection system based on low-light enhancement and YOLO algorithm [J]. Elelctronic Measurement Technology, 2024,47(6):100-108.
- [6] YASIR M, LIU S W, XU M M, et al. ShipGeoNet: SAR image-based geometric feature extraction of ships using convolutional neural networks [J]. IEEE Transactions on Geoscience and Remote Sensing, 2024,62(1):1-13.
- [7] GIRSHICK R. FastR-CNN [C]. IEEE International Conferenceon Computer Vision, 2015:1440-1448.
- [8] REN SH Q, HE K M, ROSS G, et al. FasterR-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactionson Pattern Analysis and Machine Intelligence, 2016, 39 (6): 1137-1149.
- [9] 余浪,苗鸿宾,苏赫朋,等. 基于改进 YOLOv3 的轴承端面缺陷检测算法[J]. 机床与液压,2024,52(9): 209-214.
YU L, MIAO H B, SU H P, et al. Defect detection algorithm of bearing end face based on improved YOLO v3 [J]. Machine Tool & Hydraulics, 2024, 52(9): 209-214.
- [10] WANG J X, LIU M, DU Y R, et al. PG-YOLO: An efficient detection algorithm for pomegranate before fruit thinning [J]. Engineering Applications of Artificial Intelligence, 2024,134(1):108700.
- [11] 李波,李志康,周钰彬. 结合特征融合和注意力机制的 SAR 舰船检测[J]. 电子测量技术,2024, 47 (10): 134-140.
LI B, LI ZH K, ZHOU Y B. SAR ship detection algorithm combining feature fusion and attention mechanism[J]. Elelctronic Measurement Technology,

- 2024, 47(10): 134-140.
- [12] LI X S, WANG C L, ZENG Z. WS-SSD: Achieving faster 3D object detection for autonomous driving via weighted point cloud sampling [J]. Expert Systems with Applications, 2024, 249(1):123805.
- [13] 易磊, 黄哲玮, 易雅雯. 改进 YOLOv8 的输电线路异物检测方法 [J]. 电子测量技术, 2024, 47(15): 125-134. YI L, HUANG ZH W, YI Y W. Improved YOLOv8 foreign object detection method for transmission lines [J]. Elelctronic Measurement Technology, 2024, 47(15): 125-134.
- [14] 陶友凤, 姜麟, 大妹. 基于 RCD-YOLO 的钢材表面缺陷检测算法 [J]. 控制工程, 2024, DOI: 10.14107/j.cnki.kzgc.20240426. TAO Y F, JIANG L, DA M. Algorithm for steel surface defect detection based on RCD-YOLO [J]. Control Engineering of China, 2024, DOI: 10.14107/j.cnki.kzgc.20240426.
- [15] 敖思铭, 周诗洋, 杨智颖, 等. 基于 KAS-YOLO 的钢板表面缺陷检测 [J]. 组合机床与自动化加工技术, 2024(8): 168-174. AO S M, ZHOU SH Y, YANG ZH Y, et al. Surface defect detection of steel plate based on KAS-YOLO [J]. Modular Machine Tool & Automatic Manufacturing Technique, 2024(8): 168-174.
- [16] WANG A, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection [J]. ArXiv preprint arXiv:2405.14458, 2024.
- [17] 黄毅, 周纯, 刘欣军, 等. 基于 YOLOv10 的无人机复杂背景下多尺度检测模型 [J/OL]. 光通信研究, 1-8 [2024-12-30]. <http://kns.cnki.net/kcms/detail/42.1266.TN.20240822.1307.002.html>. HUANG Y, ZHOU CH, LIU X J, et al. Multiscale detection model for complex backgrounds in UAV images based on YOLOv10 [J/OL]. Study on Optical Communications, 1-8 [2024-12-30]. <http://kns.cnki.net/kcms/detail/42.1266.TN.20240822.1307.002.html>.
- [18] 杨海涛, 赵俊羽, 王瑞, 等. 基于 DBB-YOLOv10s 的无人机小目标检测 [J/OL]. 红外技术, 1-8 [2024-12-30]. <http://kns.cnki.net/kcms/detail/53.1053.TN.20241126.1129.002.html>. YANG H T, ZHAO J Y, WANG R, et al. Small object detection for UAVs based on DBB-YOLOv10s [J/OL]. Infrared Technology, 1-8 [2024-12-30]. <http://kns.cnki.net/kcms/detail/53.1053.TN.20241126.1129.002.html>.
- [19] 李军, 房志远, 周昊星. 改进 YOLOv10 的天台光伏涉电区域行为人预警算法 [J/OL]. 计算机工程与应用, 1-13 [2024-12-30]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20241031.1517.010.html>.
- [20] CHEN J, KAO S H, HE H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2023.
- [21] OUYANG D, HE S, ZHANG G, et al. Efficient multi-scale attention module with cross-spatial learning [C]. 2023 IEEE International Conference on Acoustics, Speech and Signal Processing(ICASSP), 2023.
- [22] SUN H, WEN Y, FENG H, et al. Unsupervised bidirectional contrastive reconstruction and adaptive fine-grained channel attention networks for image dehazing [J]. Neural Networks, 2024, 176: 106314.
- [23] KANG M, TING C M, TING F F, et al. ASF-YOLO: A novel YOLO model with attentional scale sequence fusion for cell instance segmentation [J]. Image and Vision Computing, 2024, 147: 105057.
- [24] BAO Y, SONG K, LIU J, et al. Triplet-graph reasoning network for few-shot metal generic surface defect segmentation [J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-11.
- [25] 靳学萌, 梁西银, 邓鹏飞. 基于改进 YOLOv10 的轻量级黄花菜分级检测模型 [J]. 智慧农业 (中英文), 2024, 6(5): 108-118. JIN X M, LIANG X Y, DENG P F. Lightweight daylily grading and detection model based on improved YOLOv10 [J]. Smart Agriculture, 2024, 6 (5): 108-118.
- [26] 高立鹏, 周孟然, 胡锋, 等. 基于 REIW-YOLOv10n 的井下安全帽小目标检测算法 [J]. 煤炭科学技术, 2024, DOI: 10.12438/cst.2024-1031. GAO L P, ZHOU M R, HU F, et al. Small target detection algorithm for underground helmet based on REIW-YOLOv10n [J]. Coal Science and Technology, 2024, DOI: 10.12438/cst.2024-1031.

作者简介

吴葛(通信作者),硕士,工程师,主要研究方向为自动控制原理、电力电子学、信号与控制教学与科研。

E-mail:gew@hust.edu.cn

朱宇凡,本科,主要研究方向为深度学习、机器视觉。

E-mail:zhuyufan@hust.edu.cn

叶天成,本科,主要研究方向为深度学习、机器视觉。

E-mail:yetiancheng@hust.edu.cn