

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

改进轻量化的FCM-YOLOv8n钢材表面缺陷检测

梁礼明, 陈康泉, 陈林俊, 龙鹏威

引用本文:

梁礼明, 陈康泉, 陈林俊, 等. 改进轻量化的FCM-YOLOv8n钢材表面缺陷检测[J]. 光电工程, 2025, 52(2): 240280.

Liang L M, Chen K Q, Chen L J, et al. Improving the lightweight FCM-YOLOv8n for steel surface defect detection[J]. *Opto-Electron Eng*, 2025, 52(2): 240280.

<https://doi.org/10.12086/oe.2025.240280>

收稿日期: 2024-11-30; 修改日期: 2025-01-04; 录用日期: 2025-01-06

相关论文

PIC2f-YOLO: 金属表面缺陷检测轻量化方法

胡依伦, 杨俊, 许聪源, 夏亚金, 邓文斌

光电工程 2025, 52(1): 240250 doi: 10.12086/oe.2025.240250

改进GBS-YOLOv7t的钢材表面缺陷检测

梁礼明, 龙鹏威, 卢宝贺, 李仁杰

光电工程 2024, 51(5): 240044 doi: 10.12086/oe.2024.240044

基于改进YOLOv5s网络的斜拉桥拉索表面缺陷检测

王鹏峰, 李运堂, 黄永勇, 朱文凯, 林婕, 王斌锐

光电工程 2024, 51(5): 240028 doi: 10.12086/oe.2024.240028

更多相关论文见光电期刊集群网站 



光电工程
Opto-Electronic Engineering

<http://cn.oeejournal.org/oe>



 OE_Journal

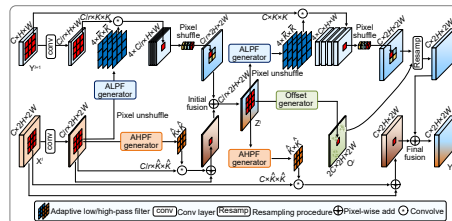


Website

改进轻量化的 FCM-YOLOv8n 钢材表面缺陷检测

梁礼明, 陈康泉*, 陈林俊, 龙鹏威

江西理工大学电气工程与自动化学院, 江西 赣州 341000



摘要: 针对现有钢材表面缺陷检测算法在资源消耗、检测精度和效率等方面存在的不足, 提出一种基于 YOLOv8n 的轻量级钢材缺陷检测算法 (FCM-YOLOv8n)。该方法一是采用频率感知特征融合网络, 高效提取并融合高频信息, 以降低计算成本并提升检测速度; 二是重构轻量化特征交互模块 (Cc-C2f), 有效保留空间和通道依赖关系, 减少特征冗余, 以降低模型参数量和计算复杂度; 三是利用多谱注意力机制, 从频域维度减少特征信息缺失, 以提升复杂缺陷的识别准确度。在 Severstal 和 NEU-DET 钢材缺陷数据集上的实验结果表明, 相较于 YOLOv8n 算法, FCM-YOLOv8n 算法的 mAP@0.5 分别提高 2.2% 和 1.5%; 参数量和复杂度分别降低 0.5 M 和 1.5 G; FPS 分别达到 143 f/s 和 154 f/s, 展示优异的实时性。该算法在检测精度、计算成本和效率之间实现良好的平衡, 为边缘终端设备应用提供有力的支持。在 GC10-DET 数据集上的进一步验证表明, FCM-YOLOv8n 相较于基线模型 mAP@0.5 提升 2.9%, 充分佐证其卓越的泛化能力。

关键词: 缺陷检测; YOLOv8n; 频率感知特征融合网络; Cc-C2f; 多谱注意力

中图分类号: TP391.41

文献标志码: A

梁礼明, 陈康泉, 陈林俊, 等. 改进轻量化的 FCM-YOLOv8n 钢材表面缺陷检测 [J]. 光电工程, 2025, 52(2): 240280

Liang L M, Chen K Q, Chen L J, et al. Improving the lightweight FCM-YOLOv8n for steel surface defect detection[J]. *Opto-Electron Eng*, 2025, 52(2): 240280

Improving the lightweight FCM-YOLOv8n for steel surface defect detection

Liang Liming, Chen Kangquan*, Chen Linjun, Long Pengwei

School of Electrical Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou, Jiangxi 341000, China

Abstract: In response to the deficiencies of existing steel surface defect detection algorithms in terms of resource consumption, detection accuracy, and efficiency, a lightweight steel defect detection algorithm based on YOLOv8n (FCM-YOLOv8n) is proposed. First, a frequency-aware feature fusion network is utilized to efficiently extract and integrate high-frequency information, reducing computational costs while enhancing detection speed. Second, a lightweight feature interaction module (Cc-C2f) is restructured to effectively preserve spatial and channel dependencies while reducing feature redundancy, thereby lowering model parameters and computational complexity. Finally, a multi-spectrum attention mechanism is applied to mitigate feature information loss in the frequency domain, improving the accuracy of detecting complex defects. Experimental results on the Severstal and

收稿日期: 2024-11-30; 修回日期: 2025-01-04; 录用日期: 2025-01-06

基金项目: 国家自然科学基金资助项目 (51365017, 61463018); 江西省自然科学基金资助项目 (20192BAB205084); 江西省教育厅科学技术研究青年项目 (GJJ2200848)

*通信作者: 陈康泉, 1136344152@qq.com。

版权所有©2025 中国科学院光电技术研究所

NEU-DET steel defect datasets show that, compared to YOLOv8n, the FCM-YOLOv8n algorithm achieves a 2.2% and 1.5% improvement in mAP@0.5, respectively, with a 0.5 M and 1.5 G reduction in parameters and computational complexity. The FPS reaches 143 f/s and 154 f/s, respectively, demonstrating excellent real-time performance. The algorithm achieves an optimal balance between detection accuracy, computational cost, and efficiency, providing robust support for edge device applications. Further validation on the GC10-DET dataset shows a 2.9% improvement in mAP@0.5 compared to the baseline model, fully demonstrating the algorithm's exceptional generalization ability.

Keywords: defect detection; YOLOv8n; frequency-aware feature fusion network; Cc-C2f; multi-spectral attention

1 引言

钢材表面缺陷检测算法主要分为传统方法和基于深度学习的方法。传统方法普遍存在准确性不足、系统复杂和实施困难等问题,限制了其实际应用。随着深度学习^[1]技术的飞速发展,基于深度学习的目标检测^[2]算法逐渐成为钢材表面缺陷检测的主流。开发高效且精准的检测技术对于保障生产安全和提升应用效率至关重要^[3]。

在深度学习的缺陷检测中,主流算法分为两类:一类是以 R-CNN^[4]为代表的两阶段方法,另一类是以 SSD^[5]和 YOLO^[6-8]系列为代表的单阶段方法。两阶段方法虽然精度较高,但计算量大、速度慢,难以在边缘终端上高效运行;而单阶段方法通过回归算法同时处理分类和定位任务,在检测速度上具有显著优势,但精度较低。由于深度学习模型的训练和推理过程消耗大量计算资源,终端设备的计算能力有限,这给实际部署带来挑战。在应用中,实时性也是钢材表面缺陷检测系统面临的关键问题,要求模型具备较高的性能。

为解决这些问题,众多研究者提出一系列钢材表面缺陷检测模型。例如, Li 等^[9]提出了一种轻量级的缺陷检测网络,采用轻量化的注意力机制,增强复杂缺陷的检测性能。Liu 等^[10]在 YOLOX 算法中引入 ASFF 模块,有效融合多尺度特征,从而提升了检测精度,但同时增加了模型体积。文献 [3] 构建了 VTG-YOLOv7-tiny 算法,通过轻量化设计降低计算消耗,但速度有所下降。Ma 等^[11]通过优化 Anchor 聚类锚框提升 YOLOv5 的检测精度,但导致信息冗余,同时精度提升仍需进一步优化。Xu 等^[12]在 YOLOv8 算法基础上使用可变形卷积提升检测精度,避免额外的计算成本。尽管上述方法在钢材缺陷检测中取得一定进展,但在检测精度、计算成本和速度的平衡上仍存

在不足。

鉴于现有研究的局限性,本文提出一种轻量化的钢材表面缺陷检测算法 FCM-YOLOv8n (FreqFusion+ Cc-C2f+MA, FCM),以 YOLOv8n 为基线模型,旨在平衡检测精度、资源消耗和效率。其主要工作如下: 1) 融合频率感知特征网络,强化高频细节边界信息,压缩模型体积并加速推理; 2) 设计轻量化特征交互模块 (Cc-C2f),减少冗余特征,降低参数量和计算复杂度; 3) 采用多谱注意力模块,从频域维度关注关键缺陷特征,增强对复杂缺陷的识别精度; 4) 在 Severstal 和 NEU-DET 数据集上验证 FCM-YOLOv8n 算法的杰出性能,并在 GC10-DET 数据集上进一步展示其在处理多样化数据和复杂场景时的稳定性和适应性。

2 YOLOv8n 算法

YOLOv8 作为经典的单阶段目标检测算法,以其快速且精确的性能在图像或视频的多目标检测中广受欢迎。其改进点包括:首先,通过不同的缩放系数提供多种尺寸模型,以适应多元化应用场景;其次,采用 C2f 模块实现轻量化;然后,采用主流的解耦头设计简化模型结构;最后,通过 Distribution Focal Loss 损失函数及 Task Aligned Assigner 正态样本分配策略,优化训练流程。YOLOv8n 网络框架如图 1 所示, Backbone、Neck 和 Head 网络分别用于特征提取、特征融合和预测输出。

3 本文算法

3.1 FCM-YOLOv8n 算法

提出一种轻量化的 FCM-YOLOv8n 检测模型,旨在优化现有钢材表面缺陷检测算法在检测精度、计算成本和效率方面的表现,特别适用于资源有限且对实时性要求较高的应用场景。FCM-YOLOv8n 模型首

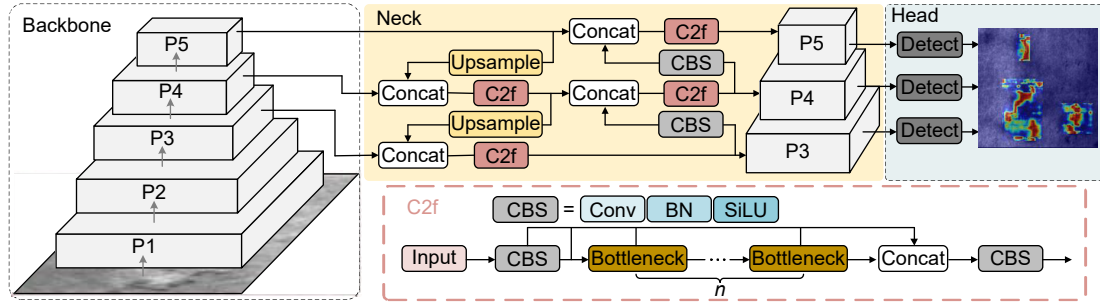


图 1 YOLOv8n 网络结构

Fig. 1 YOLOv8n network structure

先通过频率感知特征融合网络, 高效处理高频信息, 从而优化计算开销并加快检测速度。其次, 设计轻量化特征交互模块 (Cc-C2f), 替换原有的特征融合模块 (C2f), 有效保留空间和通道之间的关联性, 减少冗余特征。最后, 在骨干网络的底端添加多谱注意力模块, 从频域维度减少特征信息缺失, 增强对复杂缺陷的识别精度。其结构如图 2 所示。

3.2 频率感知特征融合模块

YOLOv8n 采用深层卷积神经网络结构来提取高层语义信息, 但随着网络层数的加深, 特征分辨率逐渐降低, 导致高频信息在传递过程中逐渐损失, 进而难以捕捉到精细的边界。为攻克这一难题, Chen 等^[13]提出一种频率感知特征融合网络 (frequency-aware feature fusion, FreqFusion), 其结构如图 3 所示。

该网络首先采用自适应低通滤波器 (adaptive low-pass filter, ALPF) 发生器减少上采样时的类内不一致

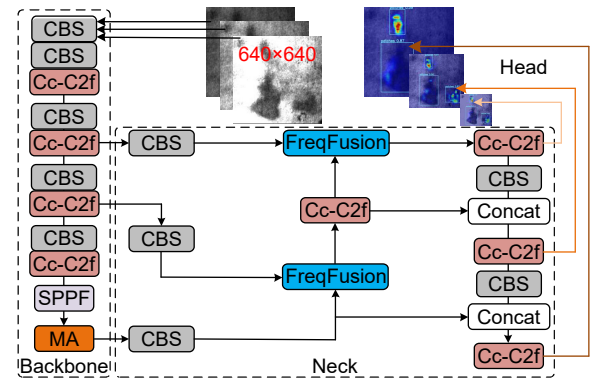


图 2 FCM-YOLOv8n 网络结构

Fig. 2 FCM-YOLOv8n network structure

性, 增强微小缺陷识别能力; 其次, 利用偏移生成器优化不一致特征和薄边界; 最后, 借助自适应高通滤波器 (adaptive high-pass filter, AHPF) 发生器增强缩减采样期间丢失的高频详细边界信息, 进一步增强缺陷边界识别能力。

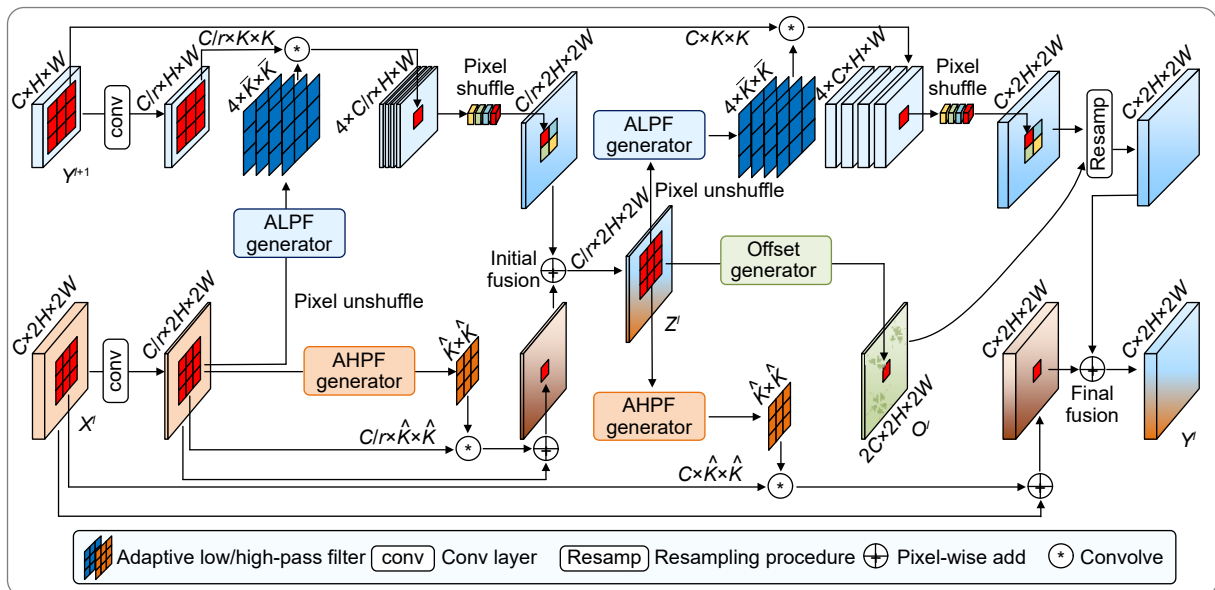


图 3 频率感知特征融合网络结构

Fig. 3 Frequency-aware feature fusion network structure

假设给定输入特征 Y^{l+1} 和 X^l , 频率感知特征融合模块对其进行压缩及融合, 在数学上可分别表示为

$$\begin{cases} Y_{i,j}^l = \tilde{Y}_{i+u,j+v}^{l+1} + \tilde{X}_{i,j}^l \\ \tilde{Y}^{l+1} = \Gamma^{\text{UP}}(\Gamma^{\text{LP}}(Y^{l+1})), \\ \tilde{X}^l = \Gamma^{\text{HP}}(X^l) + X^l \end{cases} \quad (1)$$

$$Z^l = \Gamma^{\text{UP}}(\text{Conv}_{1 \times 1}(Y^{l+1})) + \text{Conv}_{1 \times 1}(X^l), \quad (2)$$

式中: l 表示层级索引; (u, v) 表示由偏移生成器预测的坐标 (i, j) 处特征的偏移值; Γ^{UP} 表示由 AHPF 生成器预测的高通滤波器; Γ^{LP} 表示由 ALPF 生成器预测的低通滤波器; Γ^{HP} 表示由 AHPF 生成器预测的高通滤波器; $Z^l \in \mathbb{R}^{C/r \times 2H \times 2W}$ 表示融合的压缩特征, 其中 r 表示通道压缩率, C 、 W 、 H 分别表示特征的通道数、宽度和高度; $\text{Conv}(\cdot)$ 表示卷积操作。

ALPF 生成器通过 3×3 卷积层和 Softmax 层对初步融合特征 Z^l 进行预测, 得到空间变化的低通滤波器。上述过程可分别表示为

$$\begin{cases} \tilde{V}^l = \text{Conv}_{3 \times 3}(Z^l) \\ \tilde{W}_{i,j}^{l,p,q} = \text{Softmax}(\tilde{V}_{i,j}^l) = E^{p,q} - \frac{\exp(\tilde{V}_{i,j}^{l,p,q})}{\sum_{p,q \in \Omega} \exp(\tilde{V}_{i,j}^{l,p,q})} \end{cases} \quad (3)$$

式中: $\tilde{V}^l \in \mathbb{R}^{K^2 \times 2H \times 2W}$ 表示空间变化的滤波器权重; q 表示位置索引; E 表示基础权重; Ω 表示 $K \times K$ 的大小, 其中 K 表示低通滤波器的核大小。然后以像素重排的方式重塑 $\tilde{W}^l$, 将高度和宽度减半, 并将通道扩展 4 倍, 获取 4 组低通滤波器。最后使用子像素上采样技术对 $Y^{l+1} \in \mathbb{R}^{C \times H \times W}$ 采样得到上采样高级特征 $\tilde{Y}^{l+1} \in \mathbb{R}^{C \times 2H \times 2W}$ 。上述过程可分别表示为

$$\begin{cases} \tilde{Y}_{i,j}^{l+1,g} = \sum_{p,q \in \Omega} \tilde{W}_{i,j}^{l,g,p,q} \cdot Y_{i+p,j+q}^{l+1} \\ \tilde{Y}^{l+1} = \text{PS}(\tilde{Y}^{l+1,1}, \tilde{Y}^{l+1,2}, \tilde{Y}^{l+1,3}, \tilde{Y}^{l+1,4}) \end{cases} \quad (4)$$

式中: $\text{PS}(\cdot)$ 表示子像素上采样操作。

偏移生成器以 Z^l 和 S 作为输入, 通过 3×3 卷积层预测偏移方向和尺度, 并计算局部余弦相似性, 对具有高同类相似性的特征进行采样, 从而减少边界或同类不一致区域的模糊性。上述过程可分别表示为

$$S_{i,j}^{l,p,q} = \frac{\sum_{c=1}^C Z_{c,i,j}^l \cdot Z_{c,i+p,j+q}^l}{\sqrt{\sum_{c=1}^C (Z_{c,i,j}^l)^2} \sqrt{\sum_{c=1}^C (Z_{c,i+p,j+q}^l)^2}}, \quad (5)$$

$$\begin{cases} O^l = D^l \cdot A^l \\ D^l = \text{Conv}_{3 \times 3}(\text{Concat}(Z^l, S^l)) \\ A^l = \text{Sigmoid}(\text{Conv}_{3 \times 3}(\text{Concat}(Z^l, S^l))) \end{cases} \quad (6)$$

式中: $S \in \mathbb{R}^{8 \times H \times W}$ 包含每个像素与其 8 个相邻像素中间的余弦相似性; $D^l \in \mathbb{R}^{2G \times H \times W}$ 表示偏移方向; $A^l \in \mathbb{R}^{2G \times H \times W}$ 表示用于控制偏移的幅度; $O^l \in \mathbb{R}^{2G \times H \times W}$ 表示每个高级特征像素的最终预测偏移量; G 表示偏移组的数量。

AHPF 生成器通过 3×3 卷积层、Softmax 层和滤波器反转操作对初步融合特征 Z^l 进行预测, 得到空间变化的高通滤波器。其过程可分别表示为

$$\begin{cases} \hat{V}^l = \text{Conv}_{3 \times 3}(Z^l) \\ \hat{W}_{i,j}^{l,p,q} = E - \text{Softmax}(\hat{V}_{i,j}^l) = E^{p,q} - \frac{\exp(\hat{V}_{i,j}^{l,p,q})}{\sum_{p,q \in \Omega} \exp(\hat{V}_{i,j}^{l,p,q})} \end{cases} \quad (7)$$

式中: $\hat{V}^l \in \mathbb{R}^{K^2 \times H \times W}$ 表示高通滤波器。

3.3 Cc-C2f 网络

C2f 模块通过多个尺度提取特征来增强表达能力, 从而提升检测性能。然而, 在处理局部特征的同时保持上下文的平衡存在一定挑战, 这导致局部与全局信息不协调, 进而影响检测准确性和鲁棒性。此外, C2f 模块采用深层网络结构, 增加了参数量和计算复杂度。针对这些问题, Zhang 等^[14]提出了一种轻量化卷积加性自注意力机制 (conv additive self-attention, CDSA), 该模块综合考虑通道和空间信息, 采用快速线性变换减少模型的参数量和复杂度。文献^[15]提出了一种轻量级的卷积门控线性单元 (convolutional gated linear unit, CGLU), 该模块结合局部和全局信息以增强网络的表征能力。受此启发, 本文融合 C2f、CDSA 和 CGLU 模块, 设计 Cc-C2f(CDSA+CGLU, Cc) 模块, 其网络结构如图 4 所示, 其中 N 表示特征向量组。

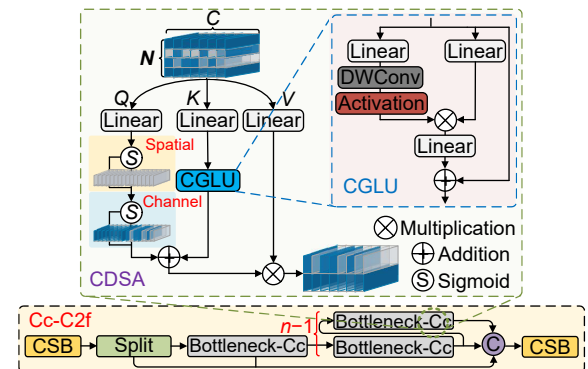


图 4 Cc-C2f 网络结构

Fig. 4 Cc-C2f network structure

Cc 模块通过三条独立的线性变换获取查询 (query, Q)、键 (key, K) 和值 (value, V), 这一过程不依

赖于手动设计的上下文, 同时能够通过卷积操作实现, 从而实现关键信息的交互, 在数学上可分别表示为

$$\begin{cases} Q = W_q x \\ K = W_k x \\ V = W_v x \end{cases}, \quad (8)$$

$$O = \Gamma(\Phi(Q) + \Phi(K)) \cdot V, \quad (9)$$

式中: q 、 k 、 v 分别表示单个标记的 Query、Key 和 Value 向量; $\Phi(\cdot)$ 表示基于 Sigmoid 的通道注意力 $C_A(\cdot) \in \mathbb{R}^{N \times d}$ 和空间注意力 $S_A(\cdot) \in \mathbb{R}^{N \times d}$ 的上下文映射函数, 其中 d 表示维度; $\Gamma(\cdot) \in \mathbb{R}^{N \times d}$ 表示用于集成上下文信息的线性变换; O 为输出特征。此外, 卷积门控线性单元结合卷积操作和门控机制, 增强局部特征与全局特征之间的协同交互能力。其数学上可表示为

$$C_A(X) = (XW_1 + B_1) \odot \sigma(\text{DWConv}(XW_2 + B_2)), \quad (10)$$

式中: $\sigma(\cdot)$ 为激活函数; $\text{DWConv}(\cdot)$ 表示深度可分离卷积操作; B_1 、 B_2 表示偏置项; W_1 、 W_2 表示权重矩阵。

3.4 多谱注意力机制

注意力机制在缺陷检测领域取得显著进展, 通过通道和空间注意力机制有效地捕捉输入信息。然而, 由于池化操作较为简单, 它在处理复杂输入特征时往往存在局限性。为了解决这一问题, Qin 等^[16]提出了频率域通道注意力网络, 通过在频域中增强对特征的关注, 从而增强目标特征的表达能力。为了减少与目标特征无关的图像信息对检测结果的干扰, 提高钢材缺陷检测的精度, 在优化模型的主干网络底部引入多谱注意力 (multi-spectral attention module, MA) 模块。其结构如图 5 所示。

该模块主要从频域角度增强注意力的关注度, 给定输入 $X^i \in \mathbb{R}^{C \times H \times W}$, 首先沿通道维度将输入 X^i 分割成 n 个部分, 记作 $X^i \in \mathbb{R}^{C \times H \times W}, i \in \{0, 1, \dots, n-1\}$, 其中

$C = C/n$ (C 能被 n 整除), 每个部分会被分配一个对应的离散余弦变换 (discrete cosine Transform, DCT) 频率分量, 这些 DCT 频率成分被压缩为信道注意力特征。其数学表示式为

$$\text{Freq}^i = \text{DCT}^{u_i, v_i}(X^i), i \in \{0, 1, \dots, n-1\}, \quad (11)$$

式中: $[u^i, v^i]$ 为 X^i 对应的频率分量; $\text{Freq}^i \in \mathbb{R}^C$ 为压缩后的信道注意力特征; $\text{DCT}(\cdot)$ 为离散余弦变换操作。然后, 将压缩后的信道注意力特征进行串联融合得到多谱特征矢量 Freq , 具体为

$$\text{Freq} = \text{cat}([\text{Freq}^0, \text{Freq}^1, \dots, \text{Freq}^{n-1}]), \quad (12)$$

式中: Freq 为多谱特征矢量; $\text{cat}(\cdot)$ 为特征融合操作。最后, 通过特征尺度变换将处理后的特征恢复到原始尺度, 从而有效剔除与目标特征无关的图像信息, 提高钢材缺陷检测的准确性。上述多谱注意力的整体过程可以表示为

$$\text{msatt} = \text{sigmoid}(\text{FC}(\text{Freq})). \quad (13)$$

4 实验

4.1 实验环境及参数设置

实验环境基于 Windows 系统, 采用 Python 编程语言和 Pytorch 框架进行开发。硬件配置为 AMD Ryzen 97945HX 处理器和 8 GB 显存的 NVIDIA GeForce RTX 4060 显卡。输入图像分辨率为 640 pixel × 640 pixel, 优化器为 SGD, 初始学习率为 0.01, 动量参数为 0.937, 权重衰减系数为 0.0005, 训练轮次为 300 轮, 批处理数量为 16 张。

4.2 实验数据集

本实验采用 NEU-DET^[17]、Severstal 和 GC10-

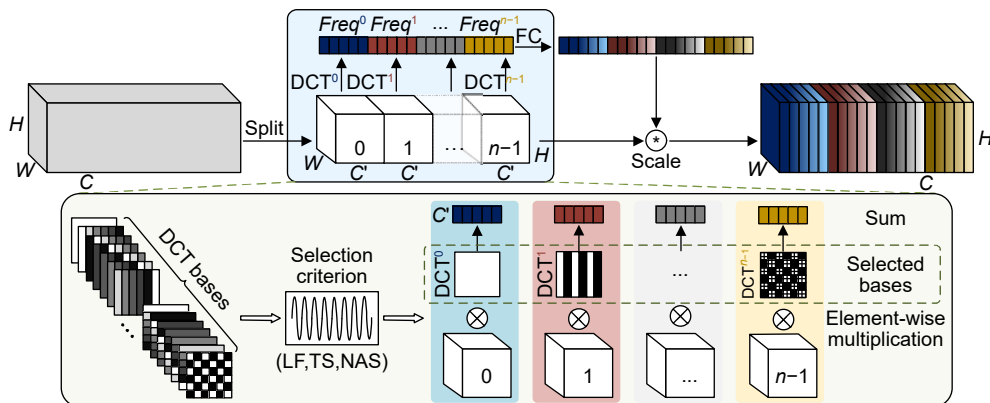


图 5 多谱注意力机制

Fig. 5 Multi-spectral attention mechanism

DET^[18] 数据集作为研究基础。NEU-DET 数据集经东北大学标注, 囊括 1800 张样本; Severstal 数据集则源自 Kaggle 平台竞赛, 收录 1120 张图像。这两个数据集均覆盖钢材的 6 种缺陷类型: 开裂 (crazing)、夹杂物 (inclusion)、斑块 (patches)、点蚀表面 (pitted-surface)、轧制氧化皮 (rolled-in-scale) 和划痕 (scratches)。而 GC10-DET 数据集来自工业场景, 包含冲孔 (Pu)、焊缝 (Wl)、新月形缝隙 (Cg)、水斑 (Ws)、油斑 (Os)、丝斑 (Ss)、夹杂物 (In)、轧坑 (Rp)、折痕 (Cr) 和腰部折痕 (Wf) 类别, 共计 2294 张图像。实验按照 8 : 2 的比例随机划分数据集为训练集和验证集。

4.3 数据增强

为提升模型的稳健性并降低过拟合的概率, 针对样本量偏少的 Severstal 数据集使用数据增强技术: 旋转、平移、镜像、裁剪、添加噪声、调整亮度和应用 Cutout, 将训练集的 896 张图像按照 3 : 1 的比例扩展到 2688 张图像, 此时训练集与验证集的图像总数为 2912 张。

4.4 评价指标

本文使用准确率 (precision, P)、召回率 (recall, R)、平均准确率 (average precision, AP)、平均准确率均值 (mean average precision, mAP)、帧率 (frame per second, FPS)、浮点运算量 (floating-point operations, FLOPs) 和参数量 (params, Par) 评估模型性能。其计算式分别为

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

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

$$AP = \int_0^1 P(R) dR, \quad (16)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i, \quad (17)$$

$$FPS = \frac{Frameum}{ElapsedTime}, \quad (18)$$

式中: TP 为模型正确地将正类样本识别出来的数量;

FP 为模型错误地将负类样本识别为正类的数量; FN 指模型错误地将正类样本识别为负类的数量; AP 为准确率和召回率曲线所围成的区域面积; mAP 为各类别 AP 值的算数平均; N 为类别数; $Frameum$ 为图像数; $ElapsedTime$ 为模型处理时间; FPS 为每秒处理的图像帧数。

4.5 Cc-C2f 模块实验

为验证 Cc-C2f 模块的优越性, 将 Cc-C2f 和 C2f 模块在 NEU-DET 数据集上进行对比。实验结果如表 1 所示, 其最优值加粗显示。

分析表 1 可知, 相较于基线模型, Cc-C2f 的 mAP@0.5 值为 75.9%, 提升了 0.7%。同时参数量和计算复杂度分别减少了 0.4 M 和 1.2 G, 检测速度达到 161 f/s, 表现出较高的实时性。Cc-C2f 的轻量化设计旨在减少参数量和计算复杂度, 同时提升检测精度。然而, 由于引入了额外的运算层级, 模型在处理复杂特征时, 可能会增加计算时间, 从而影响推理速度。总体而言, Cc-C2f 整体性能优异。

4.6 消融实验

为验证改进模块的有效性, 在 Severstal 和 NEU-DET 数据集上开展消融实验。其结果如表 2 所示, 其中最优值加粗显示, $\sqrt{}$ 表示添加该网络。

分析表 2 可知, 在 Severstal 和 NEU-DET 数据集上, 融合 FreqFusion 网络后, 有效降低了模型参数量和计算复杂度, 并提高了检测精度和速度。分别添加 Cc-C2f 和 MA 模块, 检测精度均有所提升。结合 FreqFusion 和 Cc-C2f 模块, 参数量和计算复杂度显著减少, 但速度有所降低, 在 Severstal 数据集上的精度略有下降, 而在 NEU-DET 数据集上精度有所提升, 其原因在于, Cc-C2f 的轻量化设计虽减少参数量和计算复杂度, 但也限制 FreqFusion 提升精度的潜力, 增加了推理时间。结合 FreqFusion、Cc-C2f 和 MA 模块, 模型实现较低的参数量和计算复杂度以及较高的检测精度, 同时效率较高。综上所述, 相较于基线模型, FCM-YOLOv8n 的检测精度分别提升 2.2% 和 1.5%, 参数量和计算复杂度分别降低 0.5 M 和 1.5 G,

表 1 Cc-C2f 与 C2f 对比实验

Table 1 Comparison experiment between Cc-C2f and C2f

Module	mAP@0.5/%	Par/M	FLOPs/G	FPS
C2f	75.2	3.0	8.1	181
Cc-C2f	75.9	2.6	6.9	161

速度分别达到 143 f/s 和 154 f/s, 具有较高的实时性。FreqFusion、Cc-C2f 和 MA 模块的协同作用显著推动性能的提升。

图 6 展示了本文算法与基线模型在经过 300 轮次训练后, 在 mAP@0.5 评估指标上的具体表现。观察训练曲线可以发现, FCM-YOLOv8n 模型在 mAP@0.5 指标上显著优于基线模型, 实现更为精确的缺陷分类和回归。在 Severstal 和 NEU-DET 数据集上, 经过 300 轮次的训练, FCM-YOLOv8n 模型的 mAP@0.5 训练曲线均呈现出平稳的趋势。与 YOLOv8n 相比, FCM-YOLOv8n 的训练曲线波动更小, 显示出更好的稳定性。

4.7 对比实验分析

为验证 FCM-YOLOv8n 的检测性能, 本文在 Severstal 和 NEU-DET 数据集上, 将其与主流目标检测算法 YOLOv3-tiny、YOLOv4-tiny、YOLOv5n、YOLOv5s、YOLOv6n、YOLOv7-tiny、YOLOv8n、

YOLOv8s、文献 [3]、文献 [10]、文献 [11]、文献 [12] 和文献 [13] 算法进行对比。实验结果如表 3 所示, 其中最优值加粗显示。

分析表 3 可知, FCM-YOLOv8n 算法在 Severstal 和 NEU-DET 数据集上表现优异。与 YOLOv3-tiny、YOLOv4-tiny、YOLOv5n、YOLOv5s、YOLOv6n、YOLOv7-tiny、YOLOv8s 以及基线模型 YOLOv8n 相比, FCM-YOLOv8n 算法在多个指标上表现出显著优势, 其 mAP@0.5 值分别达到 75.1% 和 76.7%; 参数量仅 2.5 M; 计算复杂度仅 6.6 G; 检测速度分别达到 143 f/s 和 154 f/s, 实时性较高。相比之下, 文献 [10] 的模型参数量最低, YOLOv5n 在参数量和复杂度上也较低, 检测速度分别达到 175 f/s 和 185 f/s, 虽速度最优, 但检测精度一般。文献 [3] 和文献 [10] 实现较好的轻量化, 但检测精度较低; 文献 [11]、文献 [12] 和文献 [13] 检测精度较高, 尤其文献 [13], 但仍需进一步提升。综合分析, FCM-YOLOv8n 算法在计算成本、检测精度和效率方面取得良好平衡, 整体性能优异。

表 2 消融实验数据

Table 2 Ablation experimental data

Dataset	FreqFusion	Cc-C2f	MA	mAP@0.5/%	Par/M	FLOPs/G	FPS
Severstal				72.9	3.0	8.1	153
	√			74.5	2.8	7.7	156
		√		73.0	2.6	6.9	141
			√	74.1	3.0	8.1	145
	√	√		74.0	2.5	6.6	145
	√	√	√	75.1	2.5	6.6	143
NEU-DET				75.2	3.0	8.1	181
	√			75.7	2.8	7.7	188
		√		75.9	2.6	6.9	161
			√	75.9	3.0	8.1	182
	√	√		76.0	2.5	6.6	154
	√	√	√	76.7	2.5	6.6	154

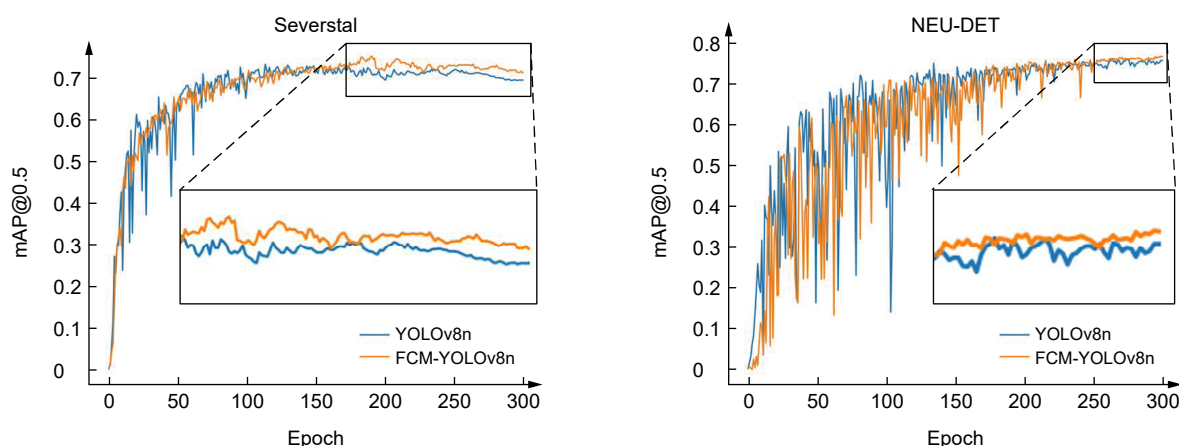


图 6 mAP@0.5 训练曲线

Fig. 6 mAP@0.5 training curves

4.8 可视化对比

为直观展示本文算法的检测效果, 将其与原始模型在 Severstal 和 NEU-DET 数据集上对比, 可视化结果见图 7。从图 7(a) 可见, YOLOv8n 在 Crazing、Inclusion、Patches、Pitted-surface 和 Rolled-in-scale 多个类别中

出现不同程度的漏检; 在 Scratches 类别中, 部分图像边缘被误判为划痕缺陷。如图 7(b) 显示, YOLOv8n 在 6 个类别中均存在严重漏检, 漏检率较高, 主要原因是 YOLOv8n 在设计时侧重提升推理速度和降低计算开销, 导致其特征提取能力相对较弱。相比之下,

表 3 不同算法检测数据对比

Table 3 Comparison of detection data from different algorithms

Dataset	Model	mAP@0.5/%	Par/M	FLOPs/G	FPS
Severstal	YOLOv3-tiny	60.2	12.1	18.9	151
	YOLOv4-tiny	55.5	5.9	16.1	97
	YOLOv5n	72.6	2.5	7.1	175
	YOLOv5s	72.1	9.1	23.8	120
	YOLOv6n	74.5	4.2	11.8	153
	YOLOv7-tiny	61.5	6.0	13.1	76
	YOLOv8n	72.9	3.0	8.1	153
	YOLOv8s	73.7	11.1	28.4	106
	Ours	75.1	2.5	6.6	143
NEU-DET	YOLOv3-tiny	64.4	12.1	18.9	156
	YOLOv4-tiny	64.0	5.9	16.1	120
	YOLOv5n	73.2	2.5	7.1	185
	YOLOv5s	75.8	9.1	23.8	106
	YOLOv6n	75.9	4.2	11.8	185
	YOLOv7-tiny	68.6	6.0	13.1	89
	Reference [3]	74.4	5.4	8.9	87
	Reference [10]	75.1	2.3	9.0	-
	Reference [11]	75.7	14.4	-	109
	Reference [12]	75.7	7.5	16.8	94
	Reference [13]	76.0	3.0	-	-
	YOLOv8n	75.2	3.0	8.1	181
	YOLOv8s	75.2	11.1	28.4	108
	Ours	76.7	2.5	6.6	154

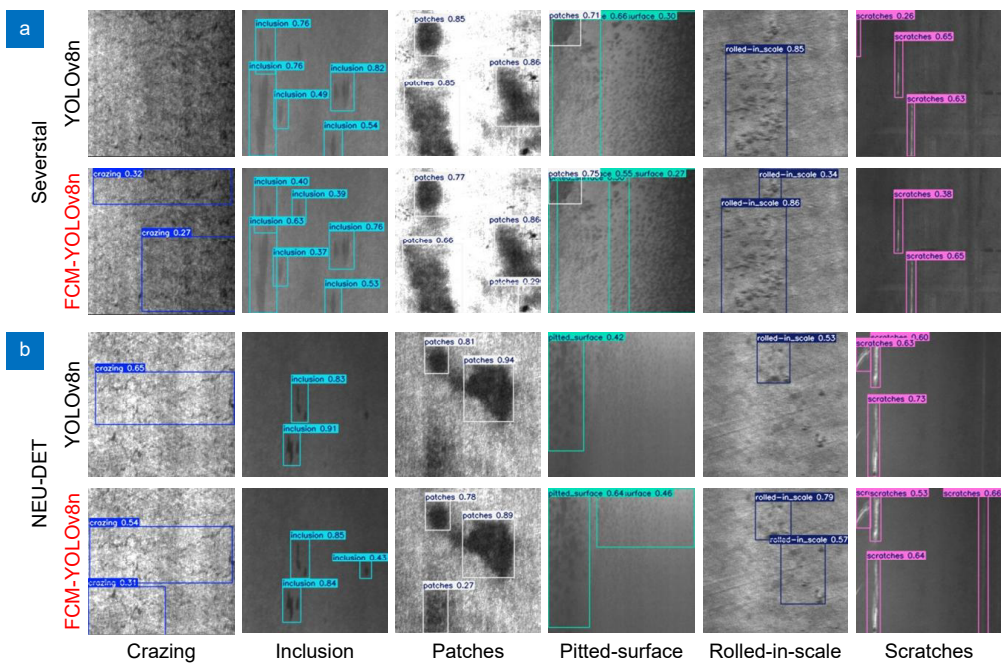


图 7 在 Severstal 和 NEU-DET 数据集上的检测效果对比

Fig. 7 Comparison of detection performance on the Severstal and NEU-DET datasets

FCM-YOLOv8n 算法在形态多样且结构复杂的缺陷中表现更优, 能够实现更精准的目标定位与识别, 显著减少误检和漏检情况。

此外, 为展示模型在 Severstal 数据增强后复杂场景中的表现, 评估 YOLOv3-tiny、YOLOv5n、YOLOv6n、YOLOv8n、YOLOv8s 和 FCM-YOLOv8n 算法在不同检测任务中的性能, 可视化结果如图 8 所示。

由图 8 可见, 对于噪声注入的裂缝类别, YOLOv3-tiny、YOLOv5n、YOLOv6n、YOLOv8n 和 YOLOv8s 均出现漏检, 主要由于裂缝缺陷与背景纹理和颜色相似, 难以区分。在随机遮挡、镜像和旋转等变换后的夹杂物类别中, YOLOv3-tiny、YOLOv5n、YOLOv6n、

YOLOv8n 和 YOLOv8s 均存在不同程度的漏检, 尤以 YOLOv5n、YOLOv8n 和 YOLOv8s 为甚。与之相比, FCM-YOLOv8n 能够准确检测所有夹杂物。在经过噪声注入和裁剪的斑块类别时, YOLOv6n、YOLOv8n 和 YOLOv8s 准确识别斑块, 而 YOLOv3-tiny 和 YOLOv5n 出现漏检。FCM-YOLOv8n 虽能识别斑块, 但发生误检, 将斑块误判为夹杂物, 表明该类别检测精度仍需改进。对于亮度调整和旋转的斑块和点蚀表面类别, YOLOv5n、YOLOv6n、YOLOv8n 与 YOLOv8s 均有漏检, 而 FCM-YOLOv8n 能准确识别出点蚀表面特征, 展现了出色的适应性。在面对噪声注入和随机遮挡的轧制氧化皮类别时, YOLOv8n

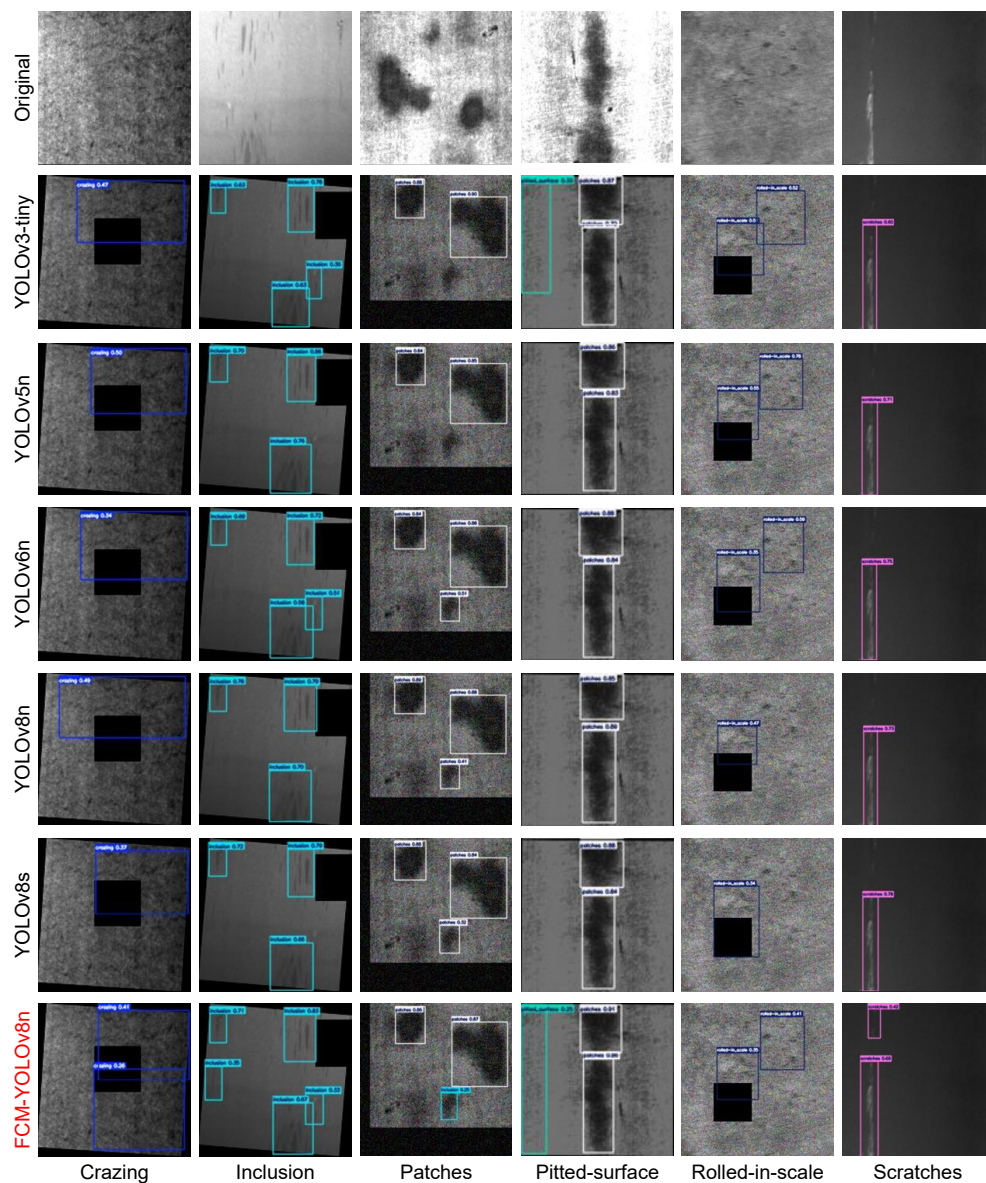


图 8 不同模型在数据增强后的 Severstal 数据集上的检测效果对比

Fig. 8 Comparison of detection performance of different models on the Severstal dataset after data augmentation

和 YOLOv8s 仅识别部分目标, 而 FCM-YOLOv8n 成功识别所有目标, 展示其强大的识别能力。对于经过亮度调整后的划痕类别, 仅 FCM-YOLOv8n 能够成功识别所有细微划痕, 进一步体现了其在挑战性图像中的可靠性。综上, FCM-YOLOv8n 在应对复杂背景干扰时, 展现出卓越的识别能力和强大的鲁棒性。

4.9 泛化性验证

在 GC10-DET 数据集上对 FCM-YOLOv8n 算法进行泛化性能验证, 将其与基线模型展开对比。其结果如表 4 所示, 其中最优值加粗显示。

分析表 4 可知, 本文算法 mAP@0.5 值达到 71.7%, 相比基线模型提升 2.9%。在 10 种缺陷类别中, 除夹杂物外, 其余 AP 值均显著提升。同时参数量和复杂度分别减少 0.5 M 和 1.5 G, 检测速度达到 270 f/s, 实时性较高。由此可见 FCM-YOLOv8n 算法

具备较强的泛化能力。

为更直观地比较本文算法与基线模型在 GC10-DET 数据集上的性能差异, 采用热力图进行定性分析, 结果如图 9 所示。观察图 9 可知, 热力图的可视化效果与锚定框图以及表 4 中数据基本一致, 即 FCM-YOLOv8n 算法在检测夹杂物时热力效果相对较差, 但在冲孔、焊缝、新月形缝隙、水斑、油斑、丝斑、轧坑、折痕和腰部折痕 9 种缺陷类别中, 热力效果显著, 且缺陷检测区域置信度更高, 漏检率较低。可视化结果表明, 本文算法表现优异。

5 结 语

提出一种轻量化的 FCM-YOLOv8n 钢材表面缺陷检测算法, 有效解决现有算法计算成本高、检测精度低和速度慢的问题。该算法在 Severstal 和 NEU-

表 4 GC10-DET 数据集检测结果对比
Table 4 Comparison of GC10-DET dataset detection results

Model	AP/%										mAP@0.5/%	Par/M	FLOPs/G	FPS
	Pu	WI	Cg	Ws	Os	Ss	In	Rp	Cr	Wf				
YOLOv8n	97.9	89.2	96.0	77.9	68.2	63.0	37.5	28.1	44.7	85.6	68.8	3.0	8.1	303
FCM-YOLOv8n	98.6	89.3	96.5	78.4	69.7	67.3	29.1	36.6	60.8	90.4	71.7	2.5	6.6	270

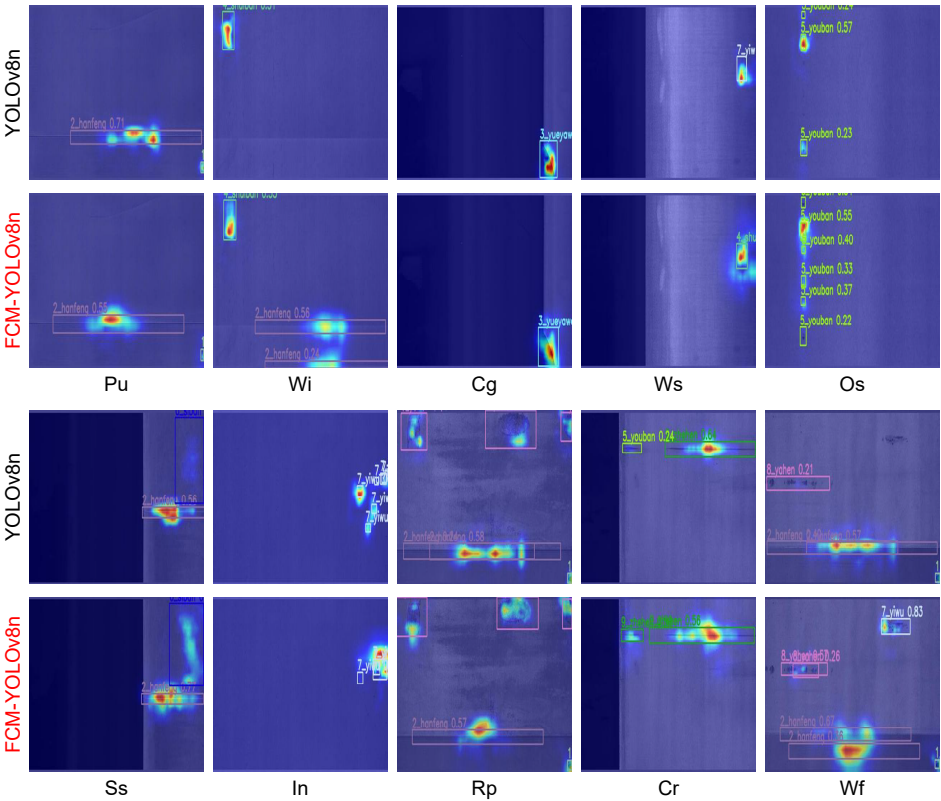


图 9 在 GC10-DET 数据集上的检测效果对比
Fig. 9 Comparison of detection performance on the GC10-DET dataset

DET 数据集上展现出卓越的性能, 其 mAP@0.5 分别提升 2.2% 和 1.5%, 参数量和计算量仅为原模型的 83.3% 和 81.5%。此外, 在 GC10-DET 数据集上的验证结果进一步彰显该算法的出色泛化能力, mAP@0.5 提升 2.9%。然而 FCM-YOLOv8n 的检测速度相较基线模型略有放缓, 未来将通过通道剪枝、层剪枝、量化技术、稀疏化策略和混合精度训练等方法, 优化检测精度、速度和资源消耗之间的平衡关系。当前, 该模型的部署仍处于实验阶段, 后续将继续改进和优化, 以期进一步提高其综合性能, 为边缘终端设备应用提供新视角。

参考文献

- [1] Huang S Q, Huang J G. Improved steel defect detection method based on enhanced fusion of RFB and YOLOv5 features[J]. *Comput Eng*, 2024.
黄硕清, 黄金贵. 基于 RFB 和 YOLOv5 特征增强融合改进的钢材缺陷检测方法[J]. *计算机工程*, 2024.
- [2] Liang L M, Long P W, Lu B H, et al. Improvement of GBS-YOLOv7t for steel surface defect detection[J]. *Opto-Electron Eng*, 2024, 51(5): 240044.
梁礼明, 龙鹏威, 卢宝贺, 等. 改进 GBS-YOLOv7t 的钢材表面缺陷检测[J]. *光电工程*, 2024, 51(5): 240044.
- [3] Liang L M, Long P W, Feng Y, et al. Improving the lightweight VTG-YOLOv7-tiny for steel surface defect detection[J]. *Opt Precis Eng*, 2024, 32(8): 1227-1240.
梁礼明, 龙鹏威, 冯耀, 等. 改进轻量化 VTG-YOLOv7-tiny 的钢材表面缺陷检测[J]. *光学精密工程*, 2024, 32(8): 1227-1240.
- [4] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//*Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014: 580-587.
<https://doi.org/10.1109/CVPR.2014.81>.
- [5] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector[C]//*Proceedings of the 14th European Conference on Computer Vision-ECCV 2016*, 2016: 21-37.
https://doi.org/10.1007/978-3-319-46448-0_2.
- [6] Reis D, Kupec J, Hong J, et al. Real-time flying object detection with YOLOv8[Z]. arXiv: 2305.09972, 2023.
<https://doi.org/10.48550/arXiv.2305.09972>.
- [7] Wang X Q, Gao H B, Jia Z M, et al. BL-YOLOv8: an improved road defect detection model based on YOLOv8[J]. *Sensors*, 2023, 23(20): 8361.
- [8] Zeng S, Yang W Z, Jiao Y Y, et al. SCA-YOLO: a new small object detection model for UAV images[J]. *Vis Comput*, 2024, 40(3): 1787-1803.
- [9] Li G, Shao R, Zhou M L, et al. Lightweight industrial products defect detection network based on attention[J]. *Comput Eng*, 2023, 49(11): 275-283.
李刚, 邵瑞, 周鸣乐, 等. 基于注意力的轻量级工业产品缺陷检测网络[J]. *计算机工程*, 2023, 49(11): 275-283.
- [10] Liu Y, Jiang S X. Steel surface defect detection algorithm based on improved YOLOX[J]. *Mod Electron Tech*, 2024, 47(9): 131-138.
刘毅, 蒋三新. 基于改进 YOLOX 的钢材表面缺陷检测研究[J]. *现代电子技术*, 2024, 47(9): 131-138.
- [11] Ma D M, Zhu J H. The optimization of YOLOv5 algorithm for detecting surface defects on hot rolled strips[J]. *Manuf Technol Mach Tool*, 2024, (6): 153-160.
马冬梅, 朱佳浩. 面向热轧带钢表面缺陷检测的 YOLOv5 算法优化分析[J]. *制造技术与机床*, 2024, (6): 153-160.
- [12] Xu X Y, Shen T, Lv J. Steel Surface Defect Detection Based On Improved YOLOv8 algorithm[J]. *Autom Appl*, 2024, 65(15): 6-10.
徐薪羽, 沈通, 吕佳. 基于改进 YOLOv8 算法的钢材表面缺陷检测[J]. *自动化应用*, 2024, 65(15): 6-10.
- [13] Chen L W, Fu Y, Gu L, et al. Frequency-aware feature fusion for dense image prediction[J]. *IEEE Trans Pattern Anal Mach Intell*, 2024, 46(12): 10763-10780.
- [14] Zhang T F, Li L, Zhou Y, et al. CAS-ViT: convolutional additive self-attention vision transformers for efficient mobile applications[Z]. arXiv: 2408.03703, 2024.
<https://doi.org/10.48550/arXiv.2408.03703>.
- [15] Shi D. TransNeXt: robust foveal visual perception for vision transformers[C]//*Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024: 17773-17783.
<https://doi.org/10.1109/CVPR52733.2024.01683>.
- [16] Qin Z Q, Zhang P Y, Wu F, et al. FcaNet: frequency channel attention networks[C]//*Proceedings of 2021 IEEE/CVF International Conference on Computer Vision*, 2021: 783-792.
<https://doi.org/10.1109/ICCV48922.2021.00082>.
- [17] Yeung C C, Lam K M. Efficient fused-attention model for steel surface defect detection[J]. *IEEE Trans Instrum Meas*, 2022, 71: 2510011.
- [18] Wang X, Zhuang K Y. An improved YOLOX method for surface defect detection of steel strips[C]//*Proceedings of 2023 IEEE 3rd International Conference on Power, Electronics and Computer Applications*, 2023: 152-157.
<https://doi.org/10.1109/ICPECA56706.2023.10075827>.

作者简介



梁礼明 (1967-), 男, 硕士, 教授, 主要研究方向为机器学习、模式识别与图像处理。

E-mail: lianglm67@163.com



【通信作者】陈康泉 (1995-), 男, 硕士研究生, 主要研究方向为机器学习、模式识别与图像处理。

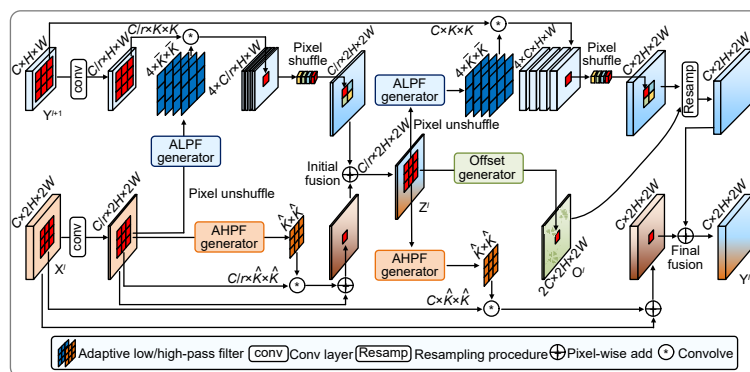
E-mail: 1136344152@qq.com



扫描二维码, 获取PDF全文

Improving the lightweight FCM-YOLOv8n for steel surface defect detection

Liang Liming, Chen Kangquan^{*}, Chen Linjun, Long Pengwei



Frequency-aware feature fusion network structure

Overview: In response to the deficiencies of existing steel surface defect detection algorithms in terms of resource consumption, detection accuracy, and efficiency, a lightweight steel defect detection algorithm based on YOLOv8n (FCM-YOLOv8n) is proposed. This algorithm incorporates three principal innovative elements. First, a frequency-aware feature fusion network is utilized to efficiently extract and integrate high-frequency information, reducing computational costs while enhancing detection speed. This network ingeniously integrates an adaptive low-pass filter generator (ALPF), an offset generator, and an adaptive high-pass filter generator (AHPF). The ALPF generator forecasts spatially-variant low-pass filters, which serve to attenuate high-frequency constituents within objects, thereby diminishing intra-class disparities during the up-sampling procedure. The offset generator plays a pivotal role in refining pronounced inconsistent features and tenuous boundaries. It achieves this by substituting inconsistent elements with more congruous ones via resampling. Meanwhile, the AHPF generator functions to augment the high-frequency detailed boundary information that is otherwise lost during down-sampling. Collectively, this fusion paradigm substantially augments feature consistency and sharpens object boundaries. Secondly, a lightweight feature interaction module (Cc-C2f) is restructured to effectively preserve spatial and channel dependencies while reducing feature redundancy, lowering model parameters and computational complexity. The Cc-C2f module integrates the lightweight convolutional additive self-attention mechanism (CDSA) and the lightweight convolutional gated linear unit (CGLU). The CDSA module takes into account both channel and spatial information, and employs fast linear transformation to reduce the number of model parameters and computational complexity. The CGLU module combines local and global information to enhance the network's representational ability. Finally, a multi-spectrum attention mechanism is applied to mitigate feature information loss in the frequency domain, improving the accuracy of detecting complex defects. Experimental results on the Severstal and NEU-DET steel defect datasets show that, compared to YOLOv8n, the FCM-YOLOv8n algorithm achieves a 2.2% and 1.5% improvement in mAP@0.5, respectively, with a 0.5 M and 1.5 G reduction in parameters and computational complexity. The FPS reaches 143 f/s and 154 f/s, respectively, demonstrating excellent real-time performance. The algorithm achieves an optimal balance between detection accuracy, computational cost, and efficiency, providing robust support for edge device applications. Further validation on the GC10-DET dataset shows a 2.9% improvement in mAP@0.5 compared to the baseline model, demonstrating the algorithm's exceptional generalization ability. Through comparative analysis with disparate algorithms, the superiority of the proposed algorithm's performance is further accentuated.

Liang L M, Chen K Q, Chen L J, et al. Improving the lightweight FCM-YOLOv8n for steel surface defect detection[J]. *Opto-Electron Eng*, 2025, 52(2): 240280; DOI: [10.12086/oe.2025.240280](https://doi.org/10.12086/oe.2025.240280)

Foundation item: National Natural Science Foundation of China (51365017, 61463018), Natural Science Foundation of Jiangxi Province (20192BAB205084), and Jiangxi Provincial Department of Education Science and Technology Research Youth Project (GJJ2200848)

School of Electrical Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou, Jiangxi 341000, China

^{*} E-mail: 1136344152@qq.com