

基于改进 yolov8 的金属瑕疵检测算法

魏锦程 肖璇 石涛

湖南工商大学, 湖南长沙, 410000;

摘要: 金属表面瑕疵检测是工业质量控制中至关重要的一部分, 尤其是在自动化生产线上。然而, 现有的目标检测算法在小目标检测、复杂纹理干扰等方面存在较大挑战。为了提高检测精度并减少漏检与误检, 本研究提出了一种改进的 YOLOv8 模型, 结合了 CBAM (Convolutional Block Attention Module) 注意力机制。通过增强通道与空间特征表达, 以及优化多尺度特征融合策略, 模型在 NEU-DET 金属瑕疵数据集上的检测精度显著提高。实验结果表明, 改进后的模型在 mAP 和误检率 (FPR) 等指标上优于基准 YOLOv8 模型。

关键词: YOLOv8; CBAM; 金属瑕疵检测; 目标检测; 注意力机制; 多尺度特征融合

DOI:10.69979/3041-0673.25.02.041

引言

金属表面瑕疵检测是现代工业质检中的一个关键环节, 广泛应用于汽车、航空、电子等多个行业。在这些行业中, 金属材料的表面瑕疵可能导致产品性能下降、使用寿命缩短, 甚至引发安全事故。因此, 快速、准确地识别和修复金属表面瑕疵, 对于确保产品质量和生产安全至关重要。

随着深度学习技术的进步, 自动化的金属瑕疵检测方法逐渐取代了传统的人工检查和基于规则的机器视觉检测。特别是目标检测算法, 如 YOLO 系列, 因其在实时性和准确性上的优势, 已成为金属瑕疵检测的主流方法。然而, 现有的目标检测方法仍然面临一些挑战, 主要包括(1) 金属表面瑕疵通常具有较小的尺寸, 尤其是微裂纹、细小划痕等。这些小目标在传统目标检测模型中容易被忽略, 导致漏检现象较为严重。(2) 金属表面具有强烈的反光、纹理变化和背景干扰, 尤其是在光照条件不佳时, 瑕疵区域的特征可能与背景高度相似, 传统算法难以准确区分。(3) 金属瑕疵的尺度差异较大, 从微小裂纹到较大的划痕或凹陷等, 传统检测方法难以统一处理不同尺度的目标, 容易导致检测效果的不稳定。

为了克服现有方法的局限性, 本研究提出了一种改进的 YOLOv8 模型, 通过引入 CBAM 注意力机制, 来增强网络对金属表面瑕疵的特征表达能力。具体创新点如下:

(1) 集成 CBAM 注意力机制, CBAM 模块通过通道注意力机制和空间注意力机制双重强化特征的选择性和聚焦性。在通道维度上, CBAM 能够通过全局池化操作为

每个通道分配权重, 从而抑制无关特征, 突出对金属瑕疵区域的敏感度。在空间维度上, CBAM 通过加权特征图上的空间位置, 使得模型能够聚焦于重要的瑕疵区域, 从而提升检测精度。(2) 多尺度特征融合优化, 改进后的模型通过加强不同尺度特征图之间的信息融合(如上采样、特征拼接), 提高对小目标瑕疵的检测能力, 尤其是在复杂背景和光照变化条件下。

1 改进 yolov8 算法

1.1 算法原理

YOLOv8 是 YOLO 系列算法之一, 继承并优化了 YOLO 系列在实时目标检测中的高效性。与之前的版本相比, YOLOv8 在精度和速度上都有显著提高, 特别是在处理小目标、复杂背景和多尺度目标时展现出了更强的能力。其核心思想仍然是利用卷积神经网络 (CNN) 进行端到端的目标检测, 即同时进行目标的分类和定位。在保留实时检测速度的优势的同时, YOLOv8 对网络结构和算法进行了多项改进, 使其在多个视觉任务中表现得更为精准。

YOLOv8 的网络结构主要包括主干网络 (Backbone)、颈部网络 (Neck) 和检测头 (Head)。其中, 主干网络负责从输入图像中提取特征, 通常使用一些先进的卷积模块如 C2f 和 SPPF, 以帮助网络捕捉不同层次的特征信息。这些模块有助于从图像中提取细节丰富且具有高辨识度的特征。颈部网络则承担了跨尺度特征融合的任务, 通过特征金字塔 (FPN) 或 PANet 等技术, 增强了对多尺度目标的检测能力。通过多层次的特征融合, 网络能在不同的尺度上同时进行目标检测, 确保检测到各

种大小的目标。检测头则是将网络输出的特征图转化为实际的目标检测结果，包括目标的位置、类别以及置信度。YOLOv8 的检测头不仅可以处理单个尺度的目标，还能应对多个尺度的目标，特别是在小目标和密集目标的检测上表现得尤为出色。

1.2 算法改进

1.2.1 CBAM 注意力机制

近年来，注意力机制已成为深度学习中非常重要的研究方向，特别是在计算机视觉任务中。注意力机制模仿人类视觉注意力的机制，使得网络能够更加专注于图像中的关键区域，从而提高特征表达的有效性。注意力机制大致分为两类：通道注意力机制和空间注意力机制。CBAM (Convolutional Block Attention Module) 是一种典型的结合通道和空间注意力机制的模块，通过对输入特征图的通道和空间两个维度进行注意力加权，提升了网络的特征选择能力。

(1) 通道注意力机制：通道注意力通过对每个特征通道进行加权，从而提升对重要通道的关注，抑制冗余或无关的通道信息。在计算机视觉任务中，不同的通道代表了不同的特征信息，通过对通道进行注意力加权，能够使得网络更加关注那些对任务重要的特征，提升模型对目标的识别能力。

(2) 空间注意力机制：空间注意力机制则是基于空间维度加权特征图中的每个像素点，通过空间位置上的注意力加权，网络能够聚焦于图像中的目标区域，从而避免背景干扰。空间注意力机制对于定位和检测目标尤其有效，可以使得模型更精确地识别目标的位置。

为此我们主要在主干网络与检测头上进行了改进。主干网络用于提取输入图像的基础特征。在 YOLOv8 的默认配置中，使用了经典的卷积层与 C2f (CSPDarknet) 模块组合进行特征提取。本模型在此基础上进行增强，加入了 CBAM 模块来加强特征选择。而检测头负责对多尺度特征进行检测，输出最终的检测结果。通过在每个尺度的特征图中插入 CBAM 模块，我们能够让网络更加聚焦于目标区域，尤其是小目标和边缘区域。

1.2.2 特征融合网络改进

特征金字塔网络用于生成多尺度的特征图。我们在 YOLOv8 的默认特征融合方式上进行了一些调整，采用跨尺度拼接 (Concat) 来融合不同尺度的特征，增强多尺度瑕疵的检测能力。在 YOLOv8 的默认结构中，使用了

特征金字塔网络 (FPN) 来实现多尺度特征融合。而我们进一步优化了这一过程，通过跨尺度拼接 (Concat) 与上采样技术来增强不同尺度之间的信息互补性。首先在网络的上采样部分，我们采用了改进的双线性插值方法来恢复特征图的分辨率。传统的上采样方法可能会在恢复过程中丢失信息，特别是在小目标的检测中。通过优化的双线性插值方法，我们能够更好地保留特征信息，减少上采样带来的信息损失，从而增强了小目标的检测能力。其次，通过跨尺度拼接技术，我们将浅层的细节信息与深层的语义信息进行融合。在金属瑕疵检测中，瑕疵可能在浅层特征中呈现出较为细微的局部结构，而在深层特征中则可能有更强的全局语义信息。通过跨尺度拼接，模型能够同时关注细节和语义，增强对不同尺度瑕疵的检测能力。

2 仿真实验与结果分析

2.1 数据集来源

本实验采用的 NEU-DET (Northeastern University Defect Dataset) 是一个金属表面瑕疵检测的标准数据集，包含 6 种不同类型的金属表面瑕疵。该数据集共有 1,800 张灰度图像，其中每个类别有 300 张图像。瑕疵种类包括：Crack (裂纹)、Pitted Surface (坑洼表面)、Scratches (划痕)、Spots (斑点)、Rust (锈蚀) 和 Weld Line (焊接线)。

2.2 实验设备及评价标准

实验采用 Pytorch 深度学习框架进行网络模型部署，Python 语言环境为 3.8.18 采用 cuda11.3 对训练进行加速。本文模型的硬件测试环境 CPU 选用 Intel (R) Core (TM) i7-11800H @ 2.30GHz，GPU 选用 NVIDIA 的 RTX30 60, 16GB 显存。训练时针对数据集与服务器的性能，设置输入图像为 640×640 ，初始学习速率为 0.08，模型训练周期 (epoch) 为 100 轮，超参数批量大小 (batch_size) 为 2。

本文使用的评价指标平均精度 mAP、帧率 FPS、参数量 Params。mAP@0.5：这是常用的检测精度指标，表示模型在 IoU 阈值为 0.5 时的平均精度。mAP 越高，表示模型在检测任务中的表现越好。推理速度 (FPS)：推理速度表示模型每秒可以处理多少帧图像，越高表示模型的实时性越好。模型参数量 (Params)：模型的参数量影响计算开销和存储需求。较小的模型参数量可以

降低模型的存储需求和推理成本。

2.3 缺陷检测结果对比

我们对比了基准 YOLOv8、yolov8+CBAM、yolov8+CBAM+多尺度融合模型在 NEU-DET 数据集上的性能。以下是结果概述：

模型	mAP@0.5	FPS	Params
YOLOv8n	80.05%	49.75	12.0M
YOLOv8n + CBAM	80.10%	60.24	11.7M
YOLOv8n + CBAM + 多尺度融合	80.24%	64.10	11.9M

根据实验数据对比分析，原始 YOLOv8n 模型在 mAP@0.5 指标上达到 80.05%，引入 CBAM 注意力机制后检测精度微升至 80.10%，结合多尺度融合技术后进一步提升至 80.24%，表明注意力机制与特征融合策略对目标检测任务具有协同优化作用。在计算效率方面，基准模型的 FPS 为 49.75 帧/秒，添加 CBAM 模块后显著提升至 60.24 帧/秒，多尺度融合技术的应用使实时性能进一步增强至 64.10 帧/秒，验证了改进算法在保持精度的同时有效优化了计算效率。值得注意的是，模型参数量从原始版本的 12.0M 缩减至 11.7M 后，最终改进型略增至 11.9M，反映了模型在结构优化过程中通过轻量化设计平衡了精度与复杂度。该实验结果表明，基于注意力机制与多尺度特征融合的改进策略能够实现检测精度、推理速度及模型轻量化三者的协同提升，为将来金属瑕疵轻量级实时目标检测系统的工程部署提供了新的解决方案。

3 结论

本研究提出了一种基于 YOLOv8 的改进金属瑕疵检测方法，集成了 CBAM (Convolutional Block Attention

Module) 注意力机制，以提高模型在金属表面瑕疵检测中的精度和鲁棒性。通过对 NEU-DET 数据集的实验验证，我们得出了以下主要结论：1) 在金属瑕疵检测中，尤其是针对小尺寸瑕疵，改进后的模型展示了更好的检测精度。传统的 YOLO 模型在小目标检测上存在较大的误差，而集成 CBAM 后的模型能够有效抑制背景干扰，聚焦在关键瑕疵区域，提高了检测的准确性。2) 金属表面的复杂纹理、反光及遮挡问题，往往使得传统检测方法在复杂场景中的表现不佳。然而，通过引入 CBAM 的通道和空间注意力机制，模型能够有效地提取与瑕疵相关的特征，忽略无关背景，从而提高了在复杂背景下的鲁棒性。3) 在金属瑕疵检测任务中，不同规模的瑕疵需要从不同层级的特征中获取信息。

参考文献

- [1]涂福泉,戚晏奇,刘建,等.基于改进 YOLOv8s 的金属齿轮表面瑕疵检测算法[J].计算机与现代化,2025,(01):100-106.
- [2]吴祖旺.基于改进 YOLO 算法的金属缺陷检测研究[D].江西理工大学,2024. DOI: 10.27176/d.cnki.gnfy.c.2024.000279.
- [3]王涵,刘海明,邵雨虹.一种改进 YOLOv5 算法的金属表面缺陷检测研究[J/OL].机械科学与技术,1-6[2025-02-24]. <https://doi.org/10.13433/j.cnki.1003-8728.20230315>.
- [4]虞佳佳,李玉,周延锁,等.基于改进 YOLOv5n 的田间铺盘装置设计与试验[J/OL].农业机械学报,1-10[2025-02-24]. <http://kns.cnki.net/kcms/detail/11.1964.s.20240624.1344.015.html>.
- [5]彭佑菊.基于深度学习的野生菌视觉分类算法的研究[D].贵州大学,2024. DOI: 10.27047/d.cnki.ggudu.2024.000662.