

改进 YOLOv8 的金属表面缺陷检测模型^{*}

龙 阳 肖小玲

(长江大学计算机科学学院, 湖北荆州 434100)

摘要: 针对工业制造中金属表面缺陷检测的检测精度低、漏检率和误检率高等问题, 文章提出了一种改进 YOLOv8 的金属表面缺陷检测模型。首先采用 CG 模块替换掉 Backbone 的下采样卷积模块, 增强模型的上下文信息理解能力; 其次采用 RepGFPN 模块替换原始的特征金字塔网络, 提高模型的多尺度特征提取能力; 最后通过对检测头重设计, 提出 GDetect 模块, 提升模型的整体性能。实验结果表明, 改进后的模型在 GC10-DET 数据集中准确率、召回率和 mAP@0.5 达到了 71.2%、72.4% 和 74.5%, 分别提高了 2.8%、8.1% 和 6.0%, 参数量和计算量分别减少了 6% 和 22%。同时在 PASCAL VOC 和 NEU-DET 数据集验证了模型在不同数据集下的鲁棒性和泛化能力, 提高了对目标的检测精度。所提出的改进模型在金属缺陷检测领域取得了显著进展, 提高了检测精度, 解决了常见问题, 并在保持轻量级的同时实现了较高的性能, 为金属表面缺陷检测提供了一种高效且可行的解决方案。

关键词: 金属表面缺陷检测; YOLOv8; CG; RepGFPN; GDetect

中图分类号: TP391.4 文献标识码: A

DOI: [10.19287/j.mtmt.1005-2402.2024.08.027](https://doi.org/10.19287/j.mtmt.1005-2402.2024.08.027)

Improved YOLOv8 metal surface defect detection model

LONG Yang, XIAO Xiaoling

(School of Computer Science, Yangtze University, Jingzhou 434100, CHN)

Abstract: In order to solve the problems of low detection accuracy, high miss detection rate and high false detection rate of metal surface defect detection in industrial manufacturing, an improved YOLOv8 metal surface defect detection model is proposed. Firstly, the CG module is used to replace Backbone's downsampling convolution module to enhance the model's contextual information understanding ability; secondly, the RepGFPN module is used to replace the original feature pyramid network to improve the model's multi-scale feature extraction ability; finally, by redesigning the detection head, The GDetect module is proposed to improve the overall performance of the model. Experimental results show that the accuracy, recall and mAP@0.5 of the improved model reached 71.2%, 72.4% and 74.5% in the GC10-DET data set, which increased by 2.8%, 8.1% and 6.0% respectively. The number of parameters and calculation The volumes were reduced by 6% and 22% respectively. At the same time, the robustness and generalization ability of the model under different data sets were verified in the PASCAL VOC and NEU-DET data sets, and the detection accuracy of the target was improved. The proposed improved model has made significant progress in the field of metal defect detection, improved detection accuracy, solved common problems, and achieved high performance while maintaining lightweight, providing an efficient and feasible solution for metal surface defect detection.

Keywords: metal surface defect detection; YOLOv8; CG; RepGFPN; GDetect

金属表面缺陷检测一直是工业领域关注的重要问题之一。随着现代工业的发展, 金属制品在各行

各业中占据着重要地位, 而其质量直接关系到产品的性能和安全。然而, 金属制品在生产过程中难免

* 国家自然科学基金项目 (61771354)

会出现各种缺陷，如裂纹、气孔和凹坑等，这些缺陷不仅影响产品的外观美观，更可能导致产品在使用过程中出现故障，甚至危及安全。因此，发展高效准确的金属表面缺陷检测方法对于确保产品质量、提高生产效率和降低成本具有重要意义。

针对金属表面缺陷的检测一直是工业质量控制中的关键环节。传统的金属表面缺陷检测方法往往依赖于人工目视检查或使用基于规则的图像处理技术。人工目视检查虽然可靠，但存在主观性、疲劳度和效率低下的问题，尤其是对于大规模生产的金属制品而言。同时，基于规则的图像处理方法需要针对特定类型的缺陷设计专门的算法，难以应对不同类型和尺寸的缺陷，且对噪声和光照变化较为敏感，容易出现误判和漏检的情况。

随着计算机视觉和深度学习技术的不断发展，自动化缺陷检测方法逐渐取代了传统的人工检测方式，成为金属表面缺陷检测的主流手段之一。近年来，基于深度学习的目标检测算法在计算机视觉领域取得了显著的成果，为自动化缺陷检测提供了新的可能性。Tang M 等^[1]基于 Faster-RCNN 提出了一种基于注意力机制和多尺度最大池化的缺陷检测方法，能够有效提高检测精度。Fang J T 等^[2]基于 Cascade R-CNN 提出了一种新的缺陷检测方法，引入一种注意力块和新的 Mix-NMS 方法，可以提高有效网络的特征提取能力。Wang C L 等^[3]提出了金属表面缺陷检测的 MeDERT 模型，使用新的 STF 模块结构，能够侧重于多头注意模块，恢复丢失的细节，提高检测速度。Wang C L 等^[4]提出了一种用于金属表面缺陷检测的改进型 YOLOv4 模型，通过集成 SAF 模块、CRMA 方案和 PA 标签平滑方法到原始模型中，在特定的数据集上 mAP 得到了显著的提升。Xu Y T 等^[5]基于改进 YOLOv5，提出了一种多目标预测头的缺陷检测方法，能够快速准确地识别金属表面缺陷的类型。Zhao Y L 等^[6]在 YOLOv5 中引入轻量级的联邦融合细颈模块和 PSA 注意力机制，在经过数据增强的金属表面缺陷数据集中的平均检测精度达到 96.7%。然而，现有的模型在金属表面缺陷检测中仍面临一些挑战，金属表面缺陷的多样性和复杂性可能使得现有模型难以充分捕捉到所有类型的缺陷，导致检测精度不够高或者存在误检、漏检的情况。

为了解决上述问题，本文通过对 YOLOv8 模型的优化和改进，针对金属表面缺陷的特定和挑战，

提出了一种基于改进 YOLOv8 的金属表面缺陷检测模型，以提高检测精度、降低误判率以及降低模型参数量和计算量。本文的贡献具体如下：

(1) 使用 CG (context guided) 模块^[7] 替换 Backbone 的下采样卷积，使得主干网络能够同时处理局部特征、周围上下文和全局上下文，并将这些信息进行有效的融合，提高模型对特征的定位精度以及其周围环境的理解和识别。

(2) 使用 RepGFPN^[8] 替换原始的 Neck 网络，使得 Neck 网络能够高效地融合多尺度特征，提高模型的泛化能力和鲁棒性，从而提高检测精度。

(3) 改进 Head 部分的 Detect 模块：通过结合参数共享思想和分组卷积，设计了新的 GDetect 模块，有效地减少参数量和计算量，提升模型的性能。

1 算法模型

1.1 改进的 YOLOv8 模型

YOLOv8 模型是 Ultralytics 开发的一种基于深度学习的目标检测算法，相较于之前的 YOLO 系列^[9-13] 模型，YOLOv8 在保持较高准确率的同时，大大提高了检测速度，这对于需要实时响应的应用，如自动驾驶、无人机监控等具有重要意义。在整体设计上，YOLOv8 模型包括输入端、Backbone、Neck 和 Head 四部分。对于输入端，会通过数据预处理和增强的操作被转换为模型可以处理的格式。Backbone 部分是 YOLOv8 的主干网络，主要负责提取图像的特征。YOLOv8 的 Backbone 参考了 CSPDarkNet-53 网络，使用了 C2f 代替了 C3 模块，这样可以在保持高精度的同时降低计算量。Neck 是连接 Backbone 和 Head 的部分，主要作用是特征融合和处理，以提高检测的准确性和效率。Head 部分是模型的最后一层，其结构会根据不同的任务而有所不同。例如，在图像分类任务中，通常会使用 Softmax 分类器作为 Head；在目标检测任务中，Head 会输出预测的目标类别和位置信息。

虽然原始 YOLOv8 模型在性能方面具有较高的准确性和速度，但也存在一些缺点。首先，YOLOv8 的参数量相较于其前身 YOLOv5 有所增加，这可能会导致更大的存储空间需求和计算负担。其次，尽管 YOLOv8 在处理小物体检测任务时表现出色，但在处理一些特定场景或类别的目标时，可能会出现漏检或误检的情况。此外，由于其基于深度学习的

特性，对硬件设备的需求也相对较高，这可能限制了其在一些资源受限的环境中的应用。

在本文中，通过以上分析和比较，为满足金属表面缺陷检测场景的需求，提高缺陷检测的准确率和性能，故选用 YOLOv8 模型做相关改进，如图 1 所示。首先，在 Backbone 中将原始的下采样卷积模块替换为 CG 模块；然后，在 Neck 中将原始特征金字塔网络替换为 RepGFPN 模块；最后，将 Head 部分的 Detect 替换为本文改进的 GDetect 模块。

1.2 CG 模块

传统的采样操作存在信息丢失、分辨率降低以及感受野受限等问题。这些缺陷可能影响模型对图像细节和全局结构的理解，尤其在金属表面缺陷检测等任务中可能限制模型性能。而 CG 模块可以帮助模型更好地理解缺陷与周围环境的关系，通过整合全局和局部上下文信息，有助于提高模型对缺陷的识别准确性。CG 模块的结构如图 2 所示，主要由局部特征提取器、周围上下文提取器、联合特征提取器和全局上下文提取器组成。

CG 模块包含两个主要步骤，在第一步中，局部特征提取器相当于一个 3×3 的标准卷积层，从周围 8 个相邻的特征向量中学习局部特征；周围上下文提取器相当于一个 3×3 的空洞卷积，因为空洞卷

积具有相对较大的感受野，可以有效地学习周围的上下文信息；随后联合特征提取器将局部特征提取器和周围上下文提取器的输出特征连接起来，并通过一个批量归一化和参数化 ReLU 操作符。在第二步中，先利用一个全局平均池化层来聚合相对于的全局上下文，然后接着用一个多层感知器来进一步提取全局上下文，最后进行加权操作得到输出的特征图。

本文采用 CG 模块替代 YOLOv8 主干网络的传统下采样模块，从而提高模型对不同尺度和场景的适应能力，增强模型的特征表达能力，提高模型在金属表面缺陷检测任务中的性能和准确性，在后面的实验中也得到了证实。

1.3 RepGFPN 模块

特征金字塔网络是一种用于处理多尺度特征的架构，旨在解决目标检测中不同尺度物体检测的挑战。它通过在不同层级上构建特征金字塔，使网络能够同时利用来自不同分辨率特征图的信息，从而提高模型对物体的检测鲁棒性和准确性。Jiang Y Q 等^[14]提出的 GFPN 以相同优先级处理高层语义信息和低层空间信息的方式，有益于多尺度特征融合互补，提高了模型的多尺度检测能力。但是，特征复用和更多连接虽然提升了检测精度，

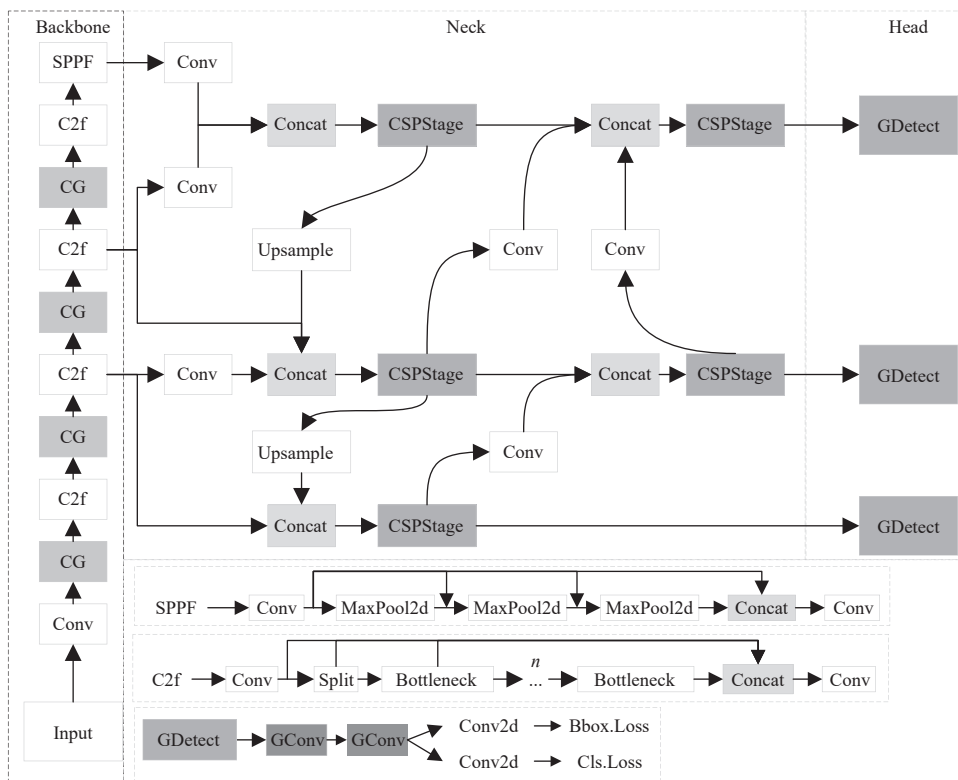


图 1 改进的 YOLOv8 模型结构图

但也存在很多的特征冗余和推理时间增加等问题。所以在此基础上，又提出了一种新的 RepGFPN 结构，如图 3 所示。

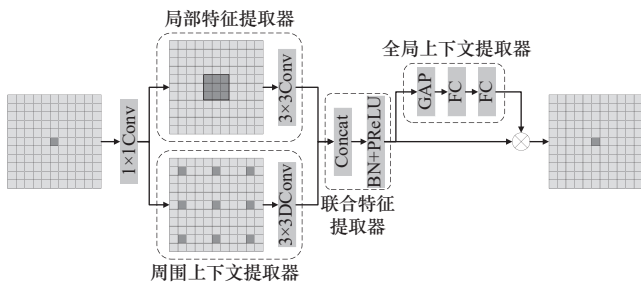


图 2 CG 结构图

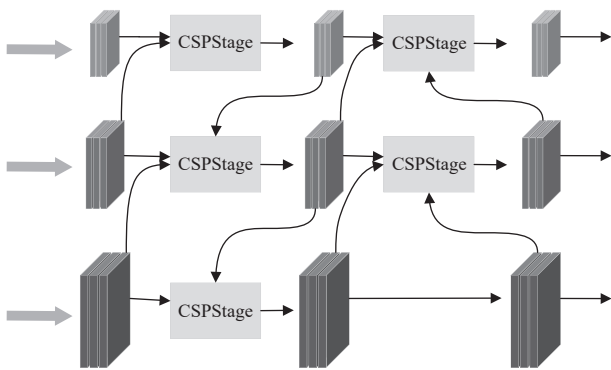


图 3 RepGFPN 结构图

相较于 GFPN，RepGFPN 主要有以下不同之处：首先，RepGFPN 在特征融合过程中，采用了不同通道维度的不同尺度特征图的设置，从而在轻量级计算量的约束下，能够灵活地控制高层特征和低层特征的表达能力；然后，去掉了 queen-fusion 中额外的上采样操作，提高 RepGFPN 实时性的性能；最后，使用 CSPStage^[15] 替换原始的 3×3 卷积的特征融合，CSPStage 结构如图 4 所示，输入为 2 个或 3 个层，经过 concat 之后，分为两条分支，一条分支经过 1×1 卷积降通道，另一条分支为 ELAN 特征聚合模块，由 N 个 Rep3×3 卷积和 3×3 卷积组成，两个分支再经过 concat 得到最终的输出。RepGFPN 在不带来额外巨大的计算负担下，实现了更高的精度。

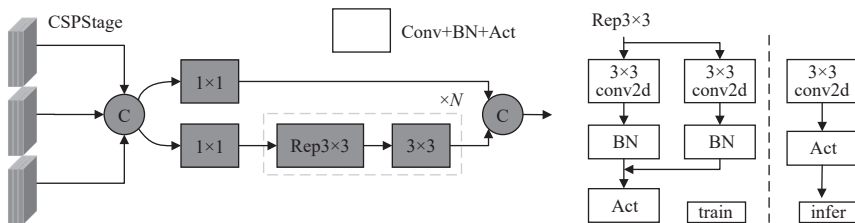


图 4 CSPStage 结构图

1.4 GDetect

原始 YOLOv8 的解耦头中的两个分支中都有两个 3×3 和一个 1×1 的卷积操作，而大量的 3×3 卷积会带来额外的计算量的负担。根据参数共享的原理，本文设计了图 5 所示的 GDetect 模块。特征图先经过两个参数共享的卷积操作，再通过解耦，进行 Cls 类别分支和 box 回归分支操作。在参数共享的卷积操作中，本文采用分组卷积替换原始的 3×3 卷积，GConv^[16] 可以减少模型的参数数量，从而降低计算成本，并在一定程度上提高模型的计算效率。

如图 6 所示，在分组卷积中，将输入通道均匀地分成多个组，每个组内地通道进行卷积操作，然后将各组的输出合并在一起形成最终的输出。

普通卷积层的参数数量和 FLOPs 计算公式为

$$Params = C_1 \times h_1 \times w_1 \times C_2 \quad (1)$$

$$FLOPs = C_1 \times C_2 \times W_2 \times H_2 \times w_1 \times h_1 \quad (2)$$

分组卷积的参数数量和 GLOPs 计算公式为

$$Params = \frac{1}{g} \times C_1 \times h_1 \times w_1 \times C_2 \quad (3)$$

$$GLOPs = \frac{1}{g} \times C_1 \times C_2 \times W_2 \times H_2 \times w_1 \times h_1 \quad (4)$$

可以看出这种分组结构可以减少每个卷积核的参数数量，从而降低模型的计算复杂度，其有效性在实验结果中也得到了证实。

2 实验结果与分析

2.1 数据集介绍

本文数据集使用的是 GC10-DET 数据集、PASCAL VOC 数据集和 NEU-DET 数据集。其中，GC10-DET 为真实工业钢材表面缺陷数据集，包含 10 种缺陷类型，即冲孔、焊缝、月牙弯、水斑、油斑、丝斑、异物、压痕、折痕和腰折。数据集首先经过人工筛选和预处理，得到 2 280 张尺寸均为 2 048×1 000 的缺陷图片，然后按照 8 : 1 : 1 的比例随机划分得到训练集 1 824 张、验证集 228 张和测试集 228 张。数据集部分示例图如图 7 所示。

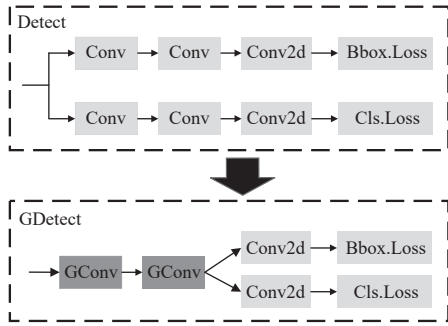


图5 GDetect 结构图

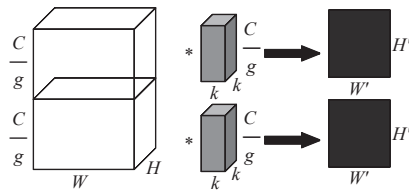


图6 GConv 结构图

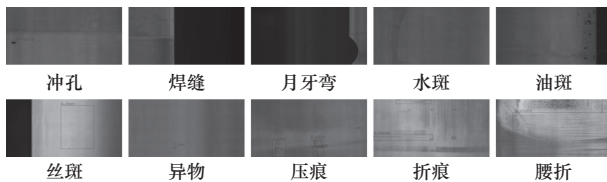


图7 数据集部分示例图

PASCAL VOC 为公开基准数据集，提供了用于检测算法和学习性能的标准评估系统。共有 20 个类别图像，即人类、船舶、汽车、狗和沙发等。本文采用的是 VOC2007+VOC2012 数据集的训练集和验证集一起来训练，VOC2007 数据集的测试集用来验证和测试，共计 16 551 张训练集，4 942 张测试集。

NEU-DET 为钢材表面缺陷数据集，包含了 6 种常见的钢材缺陷，即裂缝 (crazing, Cr)、斑块 (patches, Pa)、点蚀表面 (pitted-surface, Ps)、氧化皮 (rolled-in-scale, Rs)、夹杂物 (inclusion, In) 和划痕 (scratches, Sc)，共计 1 800 张尺寸为 200×200 的灰度图像，数据集按照 8 : 2 比例随机划分，得到训练集 1 440 张，验证集 360 张。

2.2 实验环境

本文的实验硬件环境配置为 NVIDIA GeForce RTX 3080，显存 12G，编译环境为 Python3.8.16+torch2.0.0+CUDA11.8。在 PyTorch 网络框架中进行训练，优化器采用的是 SGD，超参数设置批大小为 32，训练周期为 150，初始学习率为 0.01，输入图片的大小尺寸为 640×640。

2.3 评价指标

本文采用的评价指标有精确率 P 、召回率 R 、 $F1$ 因子、平均精确率 mAP 、参数量 $Params$ ，浮点运算次数 $GFLOPs$ 。相关公式如下：

$$P = \frac{TP}{TP + FP} \quad (5)$$

$$R = \frac{TP}{TP + FN} \quad (6)$$

$$F1 = \frac{2(P \times R)}{P + R} \quad (7)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n \int_0^1 P(R) d(R) \quad (8)$$

式中： TP 表示真正例的数量； FP 表示假正例的数量； FN 表示假负例的数量。

精确率衡量了模型在正类别预测方面的准确性；召回率衡量了模型在正类别样本中的覆盖率； $F1$ 因子用于综合衡量模型的性能；平均准确率计算了在不同类别上的平均精确率值，然后取平均值；参数量指模型中的可训练参数数量，通常，参数量越多，模型越复杂； $GFLOPs$ 表示模型在每秒进行的十亿次浮点运算，与模型的速度和性能有关。

2.4 消融实验

为了验证本研究改进点的有效性，在 GC10-DET 数据集上设计了消融实验，见表 1。其中，模型 1 将颈部网络替换为 RepGFPN 模块，模型 2 将主干网络的下采样模块替换为 CG 模块，两组实验结果虽然参数量和计算量有略微的增加，但是精确率、召回率和 $mAP_{0.5}$ 均有明显的提升，证明了 RepGFPN 模块能够提高模型的多尺度特征提取能力和 CG 模块能够提高模型的上下文信息提取能力。模型 3 将检测头替换为改进的 GDetect 模块，可以看出在不影响检测精度的情况下，GDetect 模块能够有效地减少模型的参数量和计算量，提高模型的性能。最后是将 3 个改进点同时添加到 YOLOv8n 网络中。实验结果表明，本文模型的 $Precision$ 为 71.2%， $Recall$ 为 72.4%， mAP 为 74.5%，相较于原始模型分别提高了 2.8%、8.1% 和 6.0%，参数量和计算量分别减少了 6% 和 22%。

图 8 所示为原始模型和本文模型在相同环境下 $mAP@0.5$ 的变化趋势。可以看出，与原始模型相比，本文模型在稳定性上表现更加一致，波动幅度更小，意味着本文模型对数据的变化或噪声具有更好的鲁棒性，能够更可靠地保持性能水平，不易受到外部因素的影响。这样的稳定趋势可以为模型的

预测结果提供更可靠的依据，提高模型在实际应用中的可信度和稳定性。

表 1 消融实验

模型	RepGFPN	CG	GDetect	P/	R/	mAP@0.5/	Params/	
				(%)	(%)	(%)	M	G
YOLOv8n				68.4	64.3	68.5	3.0	8.1
1	√			74.0	66.3	71.3	3.3	8.3
2		√		82.4	63.9	71.2	3.2	8.6
3			√	78.5	58.8	68.7	2.4	5.6
4	√		√	58.6	78.0	73.0	2.7	5.8
5		√	√	71.3	72.1	73.4	2.6	6.1
6	√	√		58.7	69.6	66.5	3.5	8.8
本文模型	√	√	√	71.2	72.4	74.5	2.8	6.3

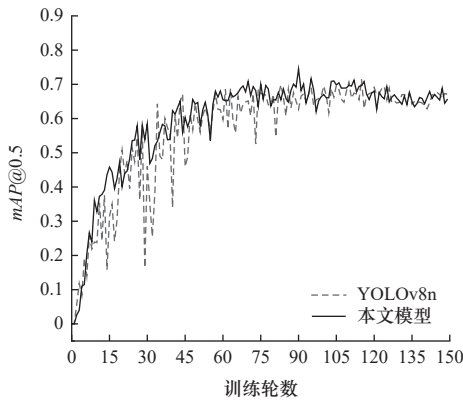


图 8 mAP@0.5 对比图

2.5 对比实验

为了验证本文模型与其他模型的性能对比，使用不同模型在 GC10-DET 数据集上进行实验，得到的实验结果见表 2。本文模型在金属表面缺陷中呈现出了卓越的性能优势。相较于其他经典模型，如 Faster R-CNN^[17]、SSD^[18]，以及不同版本的 YOLO 模型，本文模型不仅在 mAP@0.5 指标上达到了 74.5%，展现出更高的检测准确率和召回率，同时还具备较低的模型复杂度，仅有 2.8M 参数量和 6.3 G 计算量，体积更小、运行更高效。这种综合优势使得本文模型成为一个在精度和效率上均突出的选择，特别适用于资源有限的场景，如嵌入式设备或移动端应用。

为了进一步验证本文模型的有效性，在 Pascal VOC 数据集上进行了 3 组不同的实验。第 1 组实验只使用 Train2007 作为训练集，第 2 组实验只使用 Train2012 作为训练集，第 3 组实验综合 Train2007

和 Train2012 作为训练集，实验结果见表 3，可以看出，在不同的训练集情况下，本文模型相对于 YOLOv8n 在不同的评估指标上有着一定的优势。特别是在 mAP@0.5-0.95 这个严格度更高的指标上，本文模型表现更为出色，显示出更好的目标检测能力和泛化能力。这表明本文模型在更大规模和更丰富多样的数据集上训练时，能够更好地学习并捕捉目标的特征，从而提高了在验证集上的检测准确性和鲁棒性。

表 2 对比实验

模型	P/	R/	F1/	mAP@0.5/	Params/	GFLOPs/	Weight/
	(%)	(%)	(%)	(%)	M	G	M
Faster R-CNN	19.7	44.2	27.3	32.6	136.8	369.9	108.0
SSD	62.8	15.5	24.8	43.9	24.8	61.6	95.1
YOLOv3-tiny	68.7	55.8	61.6	64.0	12.1	18.9	24.4
YOLOv5s	68.2	63.8	65.9	68.7	7.0	15.8	14.4
YOLOv6n	70.4	68.3	69.3	70.3	4.2	11.8	8.7
YOLOv7-tiny	64.5	59.0	61.6	66.4	6.0	13.3	12.3
YOLOv8n	68.4	64.3	66.3	68.5	3.0	8.1	6.2
本文模型	71.2	72.4	71.8	74.5	2.8	6.3	6.1

表 3 PASCAL VOC 对比实验

实验	数据集	模型	P/	R/	F1/	mAP@0.5/	mAP@0.5-
			(%)	(%)	(%)	(%)	0.95/ (%)
1	Train2007	YOLOv8n	68.2	59.5	63.4	64.7	43.1
		本文模型	70.4	57.9	63.5	65.0	43.4
2	Train2012	YOLOv8n	74.4	66.1	70.0	72.9	51.9
		本文模型	73.5	66.7	69.9	73.7	52.7
3	Train2007+ Train2012	YOLOv8n	79.5	69.7	74.2	78.0	57.0
		本文模型	79.2	69.9	74.2	78.7	57.7

除此之外，在 NEU-DET 数据集上更进一步地验证了本文模型对金属表面缺陷检测的泛化能力和识别能力，实验结果见表 4。实验结果表明，相较于其他模型，本文模型在金属表面缺陷检测任务上表现出明显的优势，Precision 提升至 72.3%，Recall 达到 70.6%，而 mAP@0.5 整体也有所增加，达到了 76.8%。意味着本文模型可应用于不同的金属表面缺陷类型，并具有较高的可靠性和稳定性。

2.6 结果可视化

为了直观地看到本文模型与原始模型的对比效果，利用改进前后的模型在 GC10-DET 的测试集上进行了实验，10 种表面缺陷类型的检测精度对比如图 9 所示。可以看出本文模型在 10 种表面缺陷类型的检测精度上呈现出明显的优势，特别是在丝斑和异物等微小缺陷类型的检测中具有突出效果，为实际工业检测提供了更可靠的解决方案。

表 4 NEU-DET 对比实验

模型	P/ (%)	R/ (%)	mAP@0.5/ (%)	mAP@0.5/ (%)					
				Cr	Pa	Ps	Rs	In	Sc
Faster R-CNN	32.9	81.2	63.0	29.1	85.4	76.5	44.0	62.5	80.4
SSD	83.7	25.7	60.1	31.5	88.9	61.0	42.7	69.3	67.4
YOLOv3-tiny	59.6	67.1	66.6	34.1	89.0	66.6	60.6	80.2	69.4
YOLOv5s	67.3	69.0	71.8	34.8	89.4	80.8	58.5	80.4	87.1
YOLOv6n	67.3	72.9	76.5	41.4	92.3	88.0	61.8	83.4	91.9
YOLOv7-tiny	63.4	67.1	68.4	32.4	88.2	81.6	48.9	77.4	82.1
YOLOv8n	67.3	73.9	75.6	38.9	91.4	87.8	66.3	80.5	88.8
本文模型	72.3	70.6	76.8	42.3	90.3	87.1	67.0	85.6	88.6

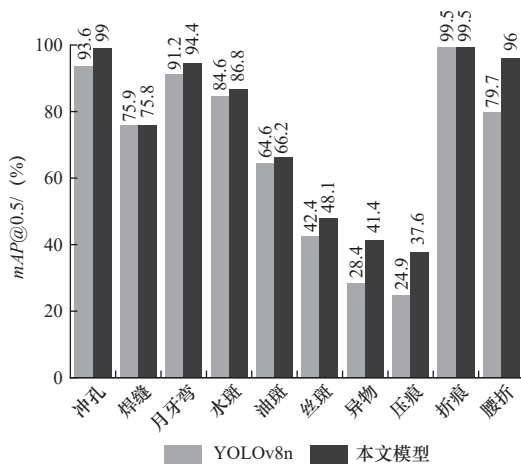


图 9 10 种缺陷类型检测精度对比图

部分检测结果如图 10 所示。首先，第 1 组和第 2 组结果图显示本文模型在漏检率方面有所降低，意味着对于较小或更细微的缺陷特征有更好的捕捉能力，从而提高了检测的全面性。其次，第 3 组和第 4 组结果图显示本文模型在误检率方面也有所降低，意味着能够更好地理解缺陷目标的上下文信息，减少误判，提高检测的准确性。最后，第 4 组和第

5 组结果图显示原始模型存在着预测边界框重合的问题，而本文模型能够更准确地理解缺陷目标的位置和大小，减少了边界框之间的重叠或者重复框的情况。综合而言，本文模型在漏检率、误检率和性能方面均取得了显著的改善，使其成为更可靠和高效的缺陷检测工具。

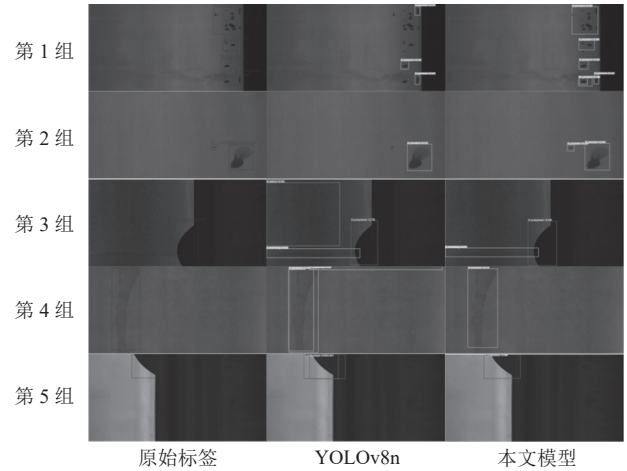


图 10 部分检测结果图

3 结语

本文提出的改进的 YOLOv8 模型在金属表面缺陷检测任务中表现出了卓越的性能和广泛的适用性。通过针对不同组件的改进，RepGFPN 模块、CG 模块和 GDetect 模块的引入分别提高了多尺度特征提取、上下文信息获取和模型参数效率。消融实验和对比分析结果验证了这些改进点的有效性，显示出本文模型在 mAP@0.5 等指标上的显著提升，同时在模型复杂度上保持着出色的表现。

在不同数据集的实验验证中，本文模型在 PASCAL VOC 和 NEU-DET 数据集上展现了出色的泛化能力和鲁棒性。并通过可视化分析，本文模型在漏检和误检方面都表现出明显改善，呈现最佳的缺陷检测和定位能力。

综上所述，改进的 YOLOv8 模型在金属表面缺陷检测任务中取得了显著的性能提升，不仅具备优异的精度和召回率，还在模型复杂度上表现出优越性，适用于资源有限的各类场景。未来将通过相关技术将进一步提高金属表面缺陷检测的精测精度，同时研究将模型在边缘设备和嵌入式设备的实际场景中部署和应用。

参 考 文 献

[1] Tang M, Li Y Y, Yao W, et al. A strip steel surface defect detection method

- based on attention mechanism and multi-scale maxpooling[J]. *Measurement Science and Technology*, 2021, 32(11): 115401.
- [2] Fang J T, Tan X Y, Wang Y H. ACRM: attention cascade R-CNN with mix-NMS for metallic surface defect detection[C]. 2020 25th International Conference on Pattern Recognition, 2021: 423-430.
- [3] Wang C L, Xie H. MeDERT: a metal surface defect detection model[J]. *IEEE Access*, 2023, 11: 35469-35478.
- [4] Wang C L, Zhou Z R, Chen Z M. An enhanced YOLOv4 model with self-dependent attentive fusion and component randomized mosaic augmentation for metal surface defect detection[J]. *IEEE Access*, 2022, 10: 97758-97766.
- [5] Xu Y T, Jiao P G, Liu J Q. CFM-YOLOv5: CFPNet module and multi-target prediction head incorporating YOLOv5 for metal surface defect detection[J]. *Plos One*, 2023, 18(12): e0289179.
- [6] Zhao Y L, Wang H, Xie X M, et al. An enhanced YOLOv5-based algorithm for metal surface defect detection[J]. *Applied Sciences*, 2023, 13(20): 11473.
- [7] Wu T Y, Tang S, Zhang R, et al. Cgnet: a light-weight context guided network for semantic segmentation[J]. *IEEE Transactions on Image Processing*, 2020, 30: 1169-1179.
- [8] Xu X Z, Jiang Y Q, Chen W H, et al. Damo-yolo: a report on real-time object detection design[J]. arXiv: 2211.15444, 2022.
- [9] Redmon J, Farhadi A. Yolov3: an incremental improvement[J]. arXiv: 1804.02767, 2018.
- [10] Bochkovskiy A, Wang C Y, Liao H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. arXiv: 2004.10934, 2020.
- [11] Ge Z, Liu S T, Wang F, et al. Yolox: exceeding yolo series in 2021[J]. arXiv: 2107.08430, 2021.
- [12] Li C Y, Li L, Jiang H L, et al. YOLOv6: a single-stage object detection framework for industrial applications[J]. arXiv: 2209.02976, 2022.
- [13] 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, 2023: 7464-7475.
- [14] Jiang Y Q, Tan Z Y, Wang J Y, et al. GiraffeDet: a heavy-neck paradigm for object detection[J]. arXiv: 2202.04256, 2022.
- [15] Wang C Y, Liao H Y M, Wu Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020: 390-391.
- [16] Xie S N, Girshick R, Dollár P, et al. Aggregated residual transformations for deep neural networks[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1492-1500.
- [17] Ren S Q, He K M, Girshick R, et al. Faster r-cnn: towards real-time object detection with region proposal networks [J]. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [18] Liu W, Anguelov D, Erhan D, et al. Ssd: single shot multibox detector[C]. *Computer Vision-ECCV 2016*, 2016: 21-37.
- 第一作者: 龙阳, 男, 1999 年生, 硕士研究生, 主要研究方向为计算机视觉与目标检测。E-mail: 2022710657@yangtzeu.edu.cn
- 通信作者: 肖小玲, 女, 1973 年生, 博士, 教授, 主要研究方向为智能信息处理与网络安全, 发表有关学术论文共 50 余篇, 被三大检索收录论文 20 余篇。E-mail: xxl@yangtzeu.edu.cn
- (编辑 曲书瑶)
(收稿日期: 2024-01-12)
- 文章编号: 20240828
如果您想发表对本文的看法, 请将文章编号填入读者意见调查表中的相应位置。

(上接第 186 页)

- [10] Yang Z J, Guo J Y, Tian H L, et al. Weakness ranking method for subsystems of heavy-duty machine tools based on FMECA information[J]. *Chinese Journal of Mechanical Engineering*, 2021, 34: 17.
- [11] 汪伦焰, 马莹, 李慧敏, 等. 基于模糊 VIKOR-FMEA 的南水北调运行管理安全风险评估 [J]. *中国农村水利水电*, 2020, 10: 194-202.
- [12] Ilangkuma R M, Shanmugam P, Sakthivel G, et al. Failure mode and effect analysis using fuzzy analytic hierarchy process[J]. *International Journal of Productivity and Quality Management*, 2014, 14(3): 296.
- [13] Yoon K P, Kim W K. The behavioral[J]. *Expert Systems with Applications*, 2017, 89: 266-272.
- [14] Khan M J, Ali M I, Kumam P, et al. Improved generalized dissimilarity measure-based VIKOR method for Pythagorean fuzzy sets[J]. *International Journal of Intelligent Systems*, 2021, 37(3): 1807-1845.
- [15] 荆树伟, 温志芳, 阎俊爱. 基于 FMEA 和模糊 VIKOR 的煤炭开采企业风险识别 [J]. *工业工程*, 2017, 20(2): 91-98.
- 第一作者: 崔英杰, 男, 1995 年生, 硕士研究生, 助理工程师, 从事数控机床可靠性技术研究。E-mail: cyingjie123@163.com
- 通信作者: 王军见, 男, 1988 年生, 博士, 高级工程师, 从事数控机床性能与可靠性测评技术研究。E-mail: wangjunjian@jcs.gt.cn
- (编辑 高扬)
(收修改稿日期: 2024-03-28)
- 文章编号: 20240827
如果您想发表对本文的看法, 请将文章编号填入读者意见调查表中的相应位置。