

李强华. 基于轻量化 YOLOv7 的焊缝表面缺陷检测模型 [J]. 智能计算机与应用, 2025, 15(11): 178–184. DOI: 10.20169/j. issn. 2095-2163. 251129

基于轻量化 YOLOv7 的焊缝表面缺陷检测模型

李强华

(景德镇陶瓷大学 机械电子工程学院, 江西 景德镇 333403)

摘要: 针对焊缝表面缺陷检测模型在计算资源有限的情况下, 难以平衡计算耗时、精度和速度的问题, 本文在 YOLOv7 模型的基础上, 提出了一种融合 GhostNet V2 和 GSConv 模块的轻量化模型。首先, 在主干特征提取网络中用 C3GhostNet V2 模块替换 ELANB 模块, 在不改变输出特征图和通道大小的前提下, 减少了卷积层的参数量、计算量; 其次, 在检测头中, 用 GSConv 卷积设计的 GSELANN 模块部分替换原始 ELANN 模块, 减少冗余和重复的信息, 从而降低计算成本, 实现轻量化。实验结果表明, 本文模型在检测精度提升 2.9 个百分点的基础上, 参数量下降 14.4%、计算量下降 16.5%, 为焊缝表面缺陷检测系统实时检测提供了有力保障。

关键词: 缺陷检测; 焊缝表面; YOLOv7; 轻量化

中图分类号: TP391.41; TG441.7

文献标志码: A

文章编号: 2095-2163(2025)11-0178-07

Surface defect detection model for welds based on lightweight YOLOv7

LI Qianghua

(School of Mechanical and Electronic Engineering, Jingdezhen Ceramic University, Jingdezhen 333403, Jiangxi, China)

Abstract: In response to the problem of balancing computation time, accuracy, and speed in the detection model of weld surface defects with limited computing resources, this paper proposes a lightweight model that integrates GhostNet V2 and GSConv modules based on the YOLOv7 model. Firstly, replacing the ELANB module with the C3GhostNet V2 module in the backbone feature extraction network reduces the parameter and computational complexity of the convolutional layer without changing the output feature map and channel size; Secondly, in the detection head, the original ELANN module is partially replaced with the GSELANN module designed by GSConv convolution, reducing redundant and repetitive information, thereby reducing computational costs and achieving lightweighting. The experimental results show that, on the basis of improving detection accuracy by 2.9 percentage points, the model in this paper reduces parameter count by 14.4% and computational complexity by 16.5%, providing strong support for real-time detection of surface defects in welds.

Key words: defect detection; weld surface; YOLOv7; lightweight

0 引言

机器视觉和深度学习在表面质量检测中的应用研究越来越广泛和深入, 焊缝表面质量的检测是其中的一个重要研究热点。焊接工艺在零件生产制造中广泛应用, 然而由于设计和操作中存在各种人为因素和随机因素, 在焊接部位可能形成不同类型的缺陷, 如气孔、裂纹、焊瘤、咬边等缺陷^[1-2]。现有的人工检测方法主观性较强、效率有限, 所以亟需一种检测算法能够对焊缝表面缺陷快速、准确的检测, 以满足企业的生产需求。

随着人工智能的迅速发展, 基于深度学习的视觉识别方法不断的应用在工业领域, 也在焊缝表面缺陷分类识别领域有了新的研究方向^[3]。随着研究的深入, 将改进的深度学习单阶段模型如 SSD (Single Shot MultiBox Detector)、YOLO (You Only Look Once) 等和两阶段模型 Faster R-CNN 等用于焊缝缺陷检测成为新趋势^[4]。Yang 等^[5]将目标检测算法 YOLOv5 应用于钢管焊缝缺陷检测, 提高了焊接质量的检测精度; Ding 等^[6]提出一种基于改进 MobileNet V2 的焊缝表面缺陷轻量化识别方法, 通过嵌入注意力模块细化图像特征信息, 在识别精度

基金项目: 江西省研究生创新基金(JYC202229)。

作者简介: 李强华(2000—), 男, 硕士研究生, 主要研究方向: 机械视觉, 深度学习。Email: 1353770880@qq.com。

收稿日期: 2024-02-25

和计算效率方面均得到提升;Pan 等^[7]设计了一种新的基于 MobileNet 的迁移学习模型作为焊接缺陷的特征提取工具,预测准确率为 97. 69%;Knaak 等^[8]提出了一种融合卷积神经网络、门控循环单元与高性能分类器的集成深度学习架构,并应用于焊缝缺陷检测;Chen 等^[9]在 Faster R-CNN 基础上,融合特征金字塔网络,并引入空间-位置注意机制和几何变换的数据增强方法,对不同的焊缝缺陷进行了多种情况的模拟;王宸等^[10]使用 Mish 激活函数替换 YOLOv3 主干网络中的激活函数,并将改进的 YOLOv3 算法应用于焊缝缺陷的检测;李衍照等^[11]提出改进的 YOLOv5 焊缝缺陷检测方法,引入 GhostNet 并采用完全交并比损失函数作为坐标位置损失,实现了焊缝缺陷的检测;唐茂俊等^[12]提出改进 Faster R-CNN 检测模型,利用 K-means 算法对锚框进行优化,引入特征金字塔网络并加入解耦分类细化结构,在焊缝缺陷检测中检测效果良好。

以上研究已经把深度学习成功应用到焊缝的缺陷检测，并验证了其可行性。但在计算资源有限的情况下，难以平衡计算耗时、精度和速度的问题。本

文基于 YOLOv7 模型,设计并提出了轻量化改进方案 YOLOv7-GG,通过引入轻量级特征提取结构 C3GhostNet V2 和改进的 GSELANN 模块,有效压缩模型参数和计算量,同时提升检测精度,实现了焊缝表面缺陷的高效检测。

1 基于轻量化 YOLOv7 的焊缝表面缺陷检测模型

1.1 YOLOv7 模型

Wang 等^[13]提出 YOLOv7 模型,与 Jocher 等^[14]提出的 YOLOv5 模型结构相似。YOLOv7 模型采用高效层聚合网络(ELAN)、基于级联模型的模型缩放、模型重新参数化等策略^[15-16]。YOLOv7 模型在不增加计算和推理成本的情况下,提供高度增强和准确的目标检测性能,在检测效率与精度之间取得了非常好的平衡,是当前最流行和成功的目标检测算法之一。YOLOv7 模型由输入(Input)、主干特征提取网络(Backbone)、特征融合网络(Neck)和检测头(Head)4 部分组成,网络结构如图 1 所示。

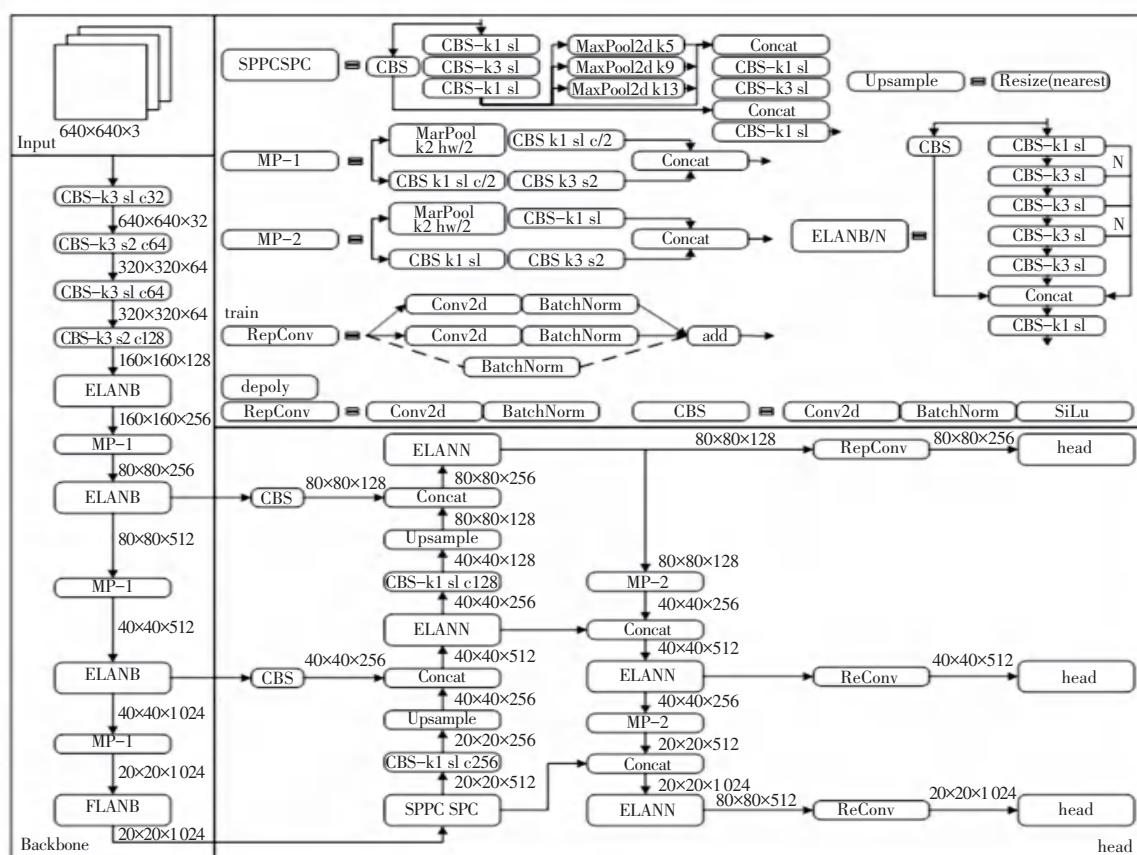


图 1 YOLOv7 模型的网络结构

Fig. 1 Network structure of the YOLOv7 mode

1.2 本文轻量化模型的结构设计

YOLOv7 模型具有精度高、易于训练和部署、检测速度快等特点,但在有限计算资源下,存在较难平衡检测精度与训练耗时的问题^[17]。本文在 YOLOv7 模型的基础上,构建了一种 C3GhostNet V2 轻量化模

块、设计了一种 GSELANN 轻量化模块,并提出融合上述两个模块的轻量化模型 YOLOv7-GG, YOLOv7-GG 模型的网络结构如图 2 所示。YOLOv7-GG 模型能够显著减少计算量和参数量,从而提高检测系统的实时性,确保检测系统能够迅速、及时的识别出缺陷。

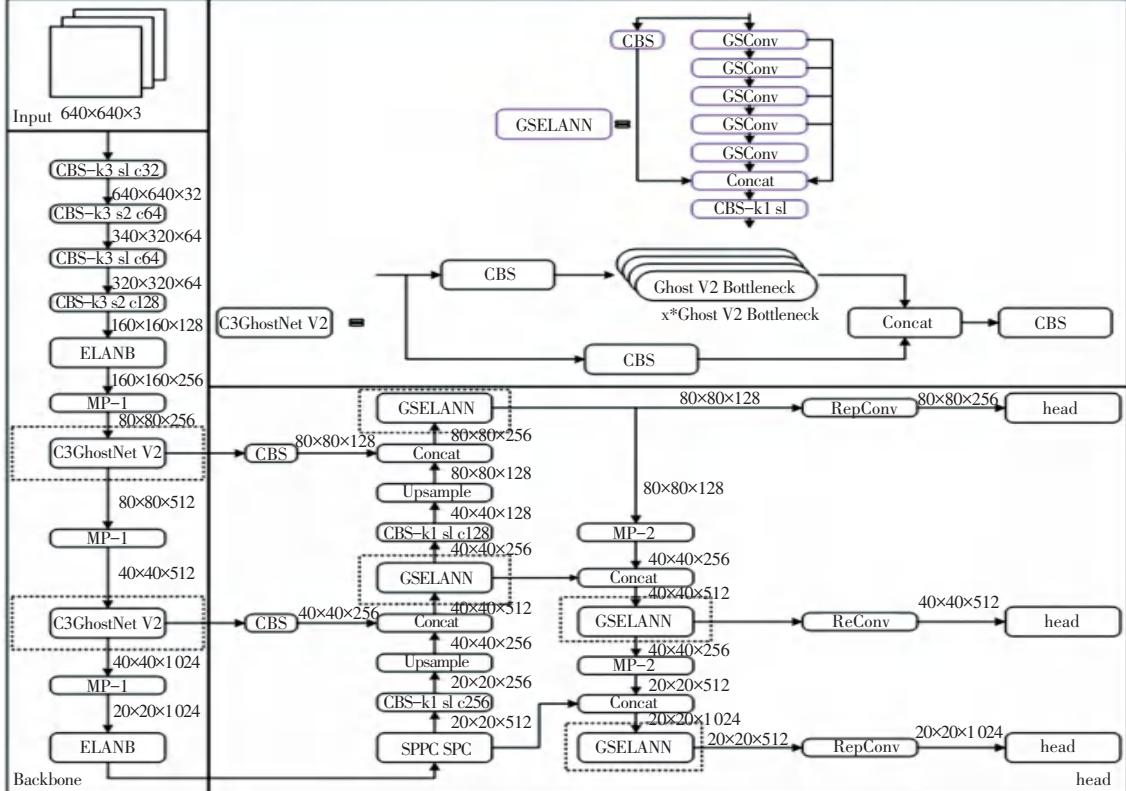


图 2 YOLOv7-GG 模型的网络结构

Fig. 2 Network structure of the YOLOv7-GG model

1) GhostNet V2 轻量化改进主干特征提取网络

卷积神经网络中,大量的卷积计算可能导致中间特征图中存在相当大的冗余,进而导致大量的计算资源消耗和不必要的成本。为了将 YOLOv7 模型更好的应用于实时焊缝表面缺陷检测场景,有必要对模型进行轻量化设计,以进一步减少计算资源的消耗。GhostNet 网络是一种轻量级的卷积神经网络,专为在移动端上进行高效推理而设计,其核心组件是 Ghost 模块^[18]。Ghost 模块作为一种新型卷积模块,相较于传统卷积神经网络(CNN)更轻量、高效。Ghost 模块通过廉价线性运算生成更多的特征图,在保持输出特征图尺寸和通道数不变的前提下,有效替代部分标准卷积操作,从而显著减少参数量与计算量,实现更高的效率和轻量化。尽管 Ghost 模块取得了显著效果,但卷积操作仅能捕获窗口区域内的局部信息,制约了性能的进一步提升。为了更好地捕获全局信息,C3GhostNet V2 使用动态特征卷积(DFC)注意机制来增

强 Ghost 模块的输出特征 Y ,以捕捉不同空间像素之间的长程依赖关系,通过 DFC 注意机制提升了轻量级模型的表征能力。基于 GhostNet V2 的优势,本文设计了一种轻量级的特征提取结构 C3GhostNet V2,具体结构如图 3 所示。

2) 轻量级 GSELANN 模块设计

分组混洗卷积(GSConv)模块的结构示意图如图 4 所示,GSConv 模块结合标准卷积(SConv)和深度可分离卷积(DWConv)的优点,在处理焊缝表面缺陷图像时展现出显著优势^[19]。与 DWConv 完全切断通道之间的联系不同,GSConv 尽可能地保留通道之间的联系,以保证模型的高精度。GSConv 将卷积结果组合和混洗,以增强特征非线性表示。针对焊接环境变化导致的焊缝表面缺陷目标,GSConv 提取的非线性特征能更好地表征焊缝表面缺陷图像的变形和膨胀过程,进而提升了模型的泛化能力和鲁棒性,更好地适应多样化的焊接缺陷检测场景。

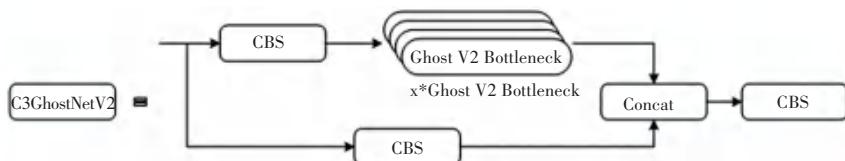


图 3 C3GhostNet V2 模块结构示意图

Fig. 3 C3GhostNet V2 module structure diagram

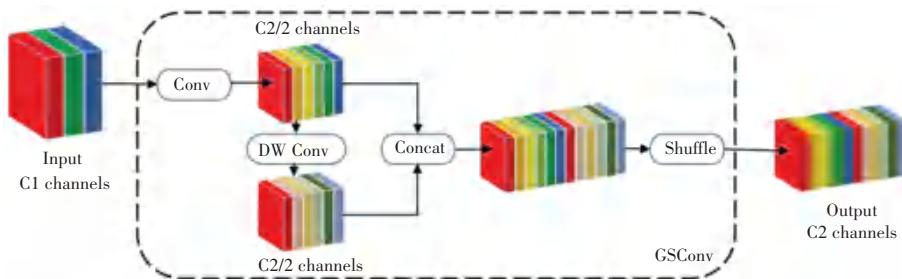


图 4 GSConv 模块结构示意图

Fig. 4 Schematic diagram of GSConv module structure

在 YOLOv7 模型中都使用 GSConv, 模型的深度将显著增加, 从而增加检测焊缝表面缺陷的时间。本文选择仅在部分阶段引入 GSConv 以实现轻量化优化。在 YOLOv7 模型中, 主干特征提取网络需通过多层卷积提取丰富特征, 且其标准卷积具有较强的通道相关性, 因而主干特征提取网络仍保留标准卷积结构。而在检测头中引入 GSConv 替代部分卷积, 能够减少冗余和重复的信息, 从而降低计算成本, 实现轻

量化。为了进一步发挥 GSConv 的作用, 本文对 ELANN 进行结构改进, 提出 GSELANN 模块, 并使用改进后的 GSELANN 模块替换检测头中原始的 ELANN 模块。改进前后的 ELANN 结构如图 5 所示。

输入的焊缝表面缺陷图像经过一系列的 GSConv 卷积, 每个通道洗牌的卷积操作能够有效地融合不同通道的焊缝表面缺陷特征图, 同时减少了模型的计算量和参数量, 最终实现了与标准卷积相近的性能。

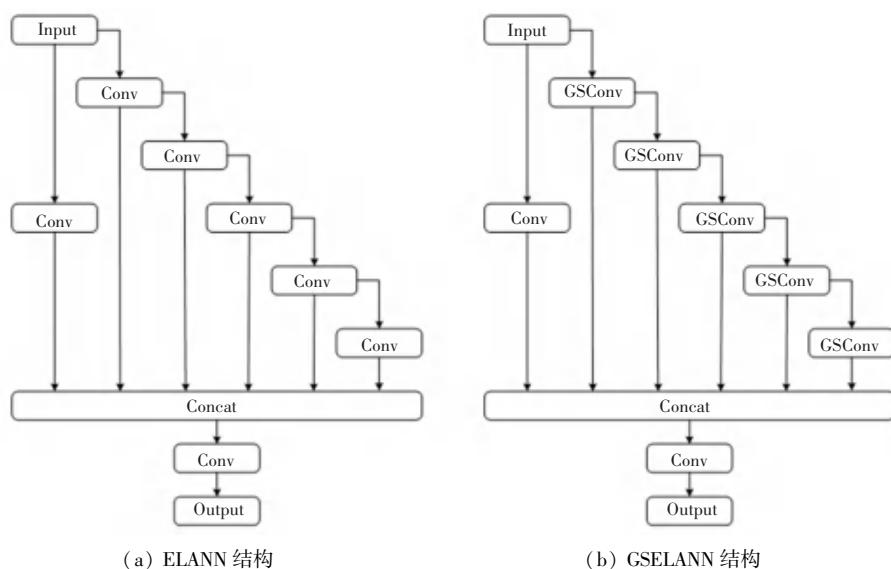


图 5 改进前后的 ELANN 结构

Fig. 5 ELANN structure before and after improvement

2 实验与分析

2.1 焊缝表面缺陷及数据集制备

目前没有焊缝表面缺陷的公开数据集, 故在某企

业实际热丝氩弧焊(TIG)焊接环境下, 对大型结构件的多层次长焊缝进行图像采集, 共采集焊缝表面缺陷原始图像 1 529 张, 缺陷以气孔、裂纹、焊瘤和咬边为主, 如图 6 所示。

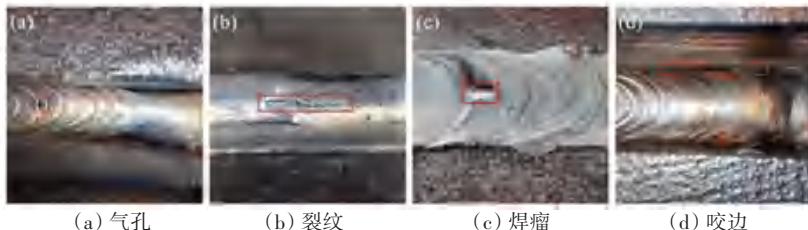


图 6 焊缝表面缺陷图片
Fig. 6 Pictures of weld surface defects

焊缝表面缺陷较小,为提高缺陷像素面积占比,对图像进行裁剪,并保留缺陷区域;采用加噪声、平移、旋转、改变亮度、随机遮挡(Cutout)和镜像方法对裁剪图像进行数据增强如图 7 所示,提升模型的鲁棒

性和泛化性。数据增强后数据集扩充为 8 154 张,分辨率 320×320,使用 LabelImg 软件对数据集进行缺陷标注,在标注完成后,将数据集以 8 : 1 : 1 的比例划分为训练集、测试集和验证集。



图 7 图像数据增强
Fig. 7 Image data enhancement

2.2 模型的评价指标

为了准确客观的对模型的检测性能进行评价,本文选用平均精度均值(mAP@0.5)、参数量(Params)和计算量(FLOPs)为评价指标。

2.3 训练环境

模型的训练环境为 Windows11、Python3.10.9 和 PyTorch1.12.1 等,相关软硬件配置见表 1。训练 epoch 为 300,输入图像尺寸为 640×640,批样本大小(batch size)为 12、参数学习率为 0.01。

表 1 实验相关软硬件配置

Table 1 Hardware and software configurations related to the experiment

名称	配置
处理器	Intel(R) Core(TM) i5-12400
Ram	DDR4 32 G
显卡	NVIDIA GeForce RTX 4070
软件	CUDA 12.1+cuDNN 8.9.0+OpenCV 4.7.0+pycharm

2.4 消融实验

为了合理评估所提轻量化模型在焊缝表面缺陷检测中的性能提升效果,本文在保持参数和环境设置一致的前提下,进行了 4 组消融实验。消融实验的结果见表 2,√ 表示选择了相应模块。

(1) 第一组为 YOLOv7 模型的实验结果,以此作为后续消融实验的基准;

表 2 消融实验的实验结果

Table 2 Experimental results of ablation experiments

Number	GhostNet V2	GSConv	Params/	FLOPs/	mAP@0.5 /
			M	G	%
N1			37.21	105.2	96.0
N2	√		34.52	93.2	97.6
N3		√	34.54	99.7	98.7
N4	√	√	31.85	87.8	98.9

(2) 第二组为主干特征提取网络中用轻量级的特征提取结构 C3GhostNet V2 替换 ELANB 模块的实验结果,训练结果较对比基准 mAP 值提升 1.6 个百分点,参数量下降 7.2%,计算量下降 11.4%;

(3) 第三组为在检测头中将改进后的 GSELANN 模块替换原始的 ELANN 模块的实验结果,训练结果较对比基准 mAP 值提升 2.7 个百分点,参数量下降 7.2%,计算量下降 5.2%;

(4) 第四组为轻量化 YOLOv7 的实验结果,训练结果较对比基准 mAP 值提升 2.9 个百分点,参数量下降 14.3%,计算量下降 16.4%。

2.5 对比实验

为了保证对比实验的公平性,在实验环境和参数设置不变的情况下,分别用 YOLOv5、YOLOv7 和 YOLOv7-GG 模型进行训练,训练结果如图 8~图 10 所示。

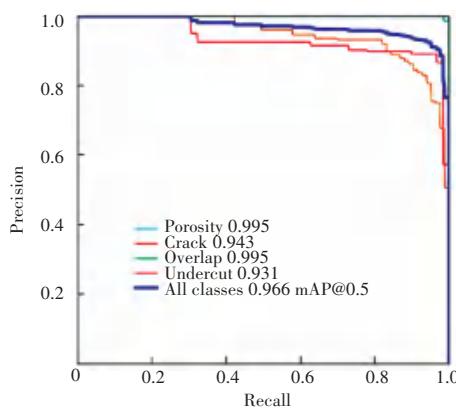


图 8 YOLOv5 训练结果

Fig. 8 YOLOv5 training results

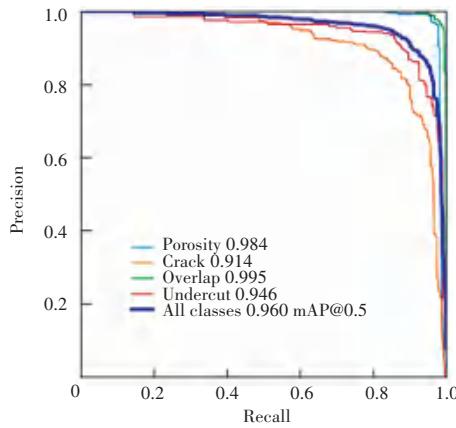


图 9 YOLOv7 训练结果

Fig. 9 YOLOv7 training results

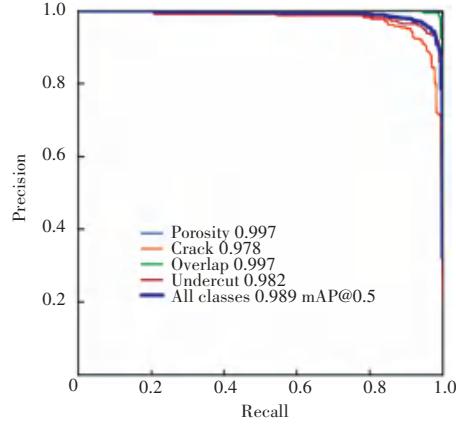


图 10 YOLOv7-GG 训练结果

Fig. 10 YOLOv7-GG training results

实验结果表明, YOLOv7-GG 模型与 YOLOv5、YOLOv7 模型相比具有更好的焊缝表面缺陷检测性能。YOLOv7-GG 算法训练结果的平均精度均值相较于 YOLOv7 算法由 96.0% 提高到 98.9%, 提升 2.9 个百分点;气孔缺陷平均精度均值提升 1.3 个百分点;裂纹缺陷平均精度均值提升 6.4 个百分点;焊瘤缺陷平均精度均值提升 0.2 个百分点;咬边缺陷值平均精度均值提升 3.6 个百分点,即 YOLOv7-

GG 模型相较于 YOLOv5 模型,4 种缺陷的检测精度均有一定提升。

2.6 本文模型与其他网络模型的对比

为验证所提轻量化模型的有效性,本文在训练参数与实验环境保持一致的条件下,将其与 Faster R-CNN、SSD、YOLOv5 及 YOLOv7 等经典目标检测模型进行对比实验,实验结果见表 3。实验结果表明,在相同输入分辨率和图像尺寸下,所提轻量化模型在 mAP 指标上均优于上述主流模型,具有更高的检测精度与综合性能。

表 3 不同模型检测结果

Table 3 Results of different models

模型	mAP/%	AP/%			
		气孔	裂纹	焊瘤	咬边
Faster R-CNN	92.7	95.5	91.1	94.3	89.9
SSD	93.6	96.1	92.0	95.3	91.0
YOLOv5	96.6	99.5	94.3	99.5	93.1
YOLOv7	96.0	98.4	91.4	99.5	94.6
本文模型	98.9	99.7	97.8	99.7	98.2

2.7 本文模型检测效果

针对焊缝表面缺陷目标的检测,本文利用 YOLOv7 和 YOLOv7-GG 算法训练的模型对焊缝表面缺陷图像进行检测,检测结果如图 11 所示。



图 11 检测效果

Fig. 11 Detect the effect

YOLOv7-GG 算法的检测效果均优于 YOLOv7 算法。图 11 使用不同的标识框来区别各算法的检测效果,并在标识框上方注明了检测类别的置信度。综上,YOLOv7-GG 算法与 YOLOv7 算法相比, YOLOv7-GG 算法能够更好的检测焊缝表面缺陷。

3 结束语

针对 YOLOv7 模型在焊缝表面缺陷检测中存在体积大和参数冗余等问题,本文提出了一种轻量化改进模型 YOLOv7-GG。在主干网络中引入重新设计的轻量级特征提取结构 C3GhostNet V2 替换原 ELANB 模块;将改进后的 GSELANN 模块替换原 ELANN 模块。实验结果表明,YOLOv7-GG 相较于

YOLOv7 模型精度提升 2.9 个百分点,参数量下降 14.4%,计算量下降 16.5%。在保证精度的前提下,本文模型大幅压缩了参数量和计算量,提高了模型的精简度和可部署性,更适用于焊缝表面缺陷检测的场景中,对于工业生产中焊缝表面缺陷的检测具有一定的参考价值。

参 考 文 献

- [1] 宋寿鹏, 乔梦丽. 基于 NLFM Barker 编码的板材焊缝缺陷超声检测方法研究[J]. 仪器仪表学报, 2020, 41(4): 246–254.
- [2] 王飞, 朱建江. 机器视觉在复杂环境下焊接缺陷检测的应用研究[J]. 焊接技术, 2017, 46(5): 127–133.
- [3] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436–444.
- [4] 陶显, 侯伟, 徐德. 基于深度学习的表面缺陷检测方法综述[J]. 自动化学报, 2021, 47(5): 1017–1034.
- [5] YANG D, CUI Y, YU Z, et al. Deep learning based steel pipe weld defect detection [J]. Applied Artificial Intelligence, 2021, 35(15): 1237–1249.
- [6] DING Kai, NIU Zhangqi, HUI Jizhuang, et al. A weld surface defect recognition method based on improved mobilenetv2 algorithm[J]. Mathematics, 2022, 10(19): 3678.
- [7] PAN H, PANG Z, WANG Y, et al. A new image recognition and classification method combining transfer learning algorithm and mobilenet model for welding defects[J]. IEEE Access, 2020, 8 (1): 119951–119960.
- [8] KNAAK C, VONEBEN J, KRÖGER M, et al. A spatio-temporal ensemble deep learning architecture for real-time defect detection during laser welding on low power embedded computing boards [J]. Sensors, 2021, 21(12): 4205.
- [9] CHEN J, WANG H, LI H. Defects detection in weld joints based on visual attention and deep learning[J]. Nondestructive Test and Evaluation International, 2022, 133: 102764.
- [10] 王宸, 张秀峰, 刘超, 等. 改进 YOLOv3 的轮廓焊缝缺陷检测 [J]. 光学精密工程, 2021, 29(8): 1942–1954.
- [11] 李衍照, 于镭, 田金文. 基于改进 YOLOv5 的金属焊缝缺陷检测[J]. 电子测量技术, 2022, 45(19): 70–75.
- [12] 唐茂俊, 黄海松, 张松松, 等. 改进的 Faster-RCNN 在焊缝缺陷检测中的应用[J]. 组合机床与自动化加工技术, 2021(12): 83–86.
- [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. Piscataway, NJ: IEEE, 2023: 7464–7475.
- [14] JOCHER G, CHANGYU L, HOGAN A, et al. Ultralytics/YOLOv5: Initial Release [J]. Zenodo, 2020. DOI: 10.5281/ZENODO.3908560
- [15] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. Scaled-yolov4: Scaling cross stage partial network[C]//Proceedings of the IEEE/cvf Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2021: 13029–13038.
- [16] DING X, ZHANG X, MA N, et al. Repvgg: Making vgg-style convnets great again [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2021: 13733–13742.
- [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. Piscataway, NJ: IEEE, 2023: 7464–7475.
- [18] HAN K, WANG Y, TIAN Q, et al. Ghostnet: More features from cheap operations [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2020: 1580–1589.
- [19] LI H, LI J, WEI H, et al. Slim-neck by GSConv: A better design paradigm of detector architectures for autonomous vehicles [J]. arXiv preprint arXiv:2206.02424, 2022.