

基于YOLOv5的表面缺陷检测优化算法

渠逸, 汪诚*, 余嘉博, 孔亚康, 陈贤聪

(空军工程大学基础部, 西安, 710051)

摘要 快速、准确地检测材料表面缺陷已成为各领域研究的重要目标,为增加检测效率,实现设备轻量化,提出了一种基于YOLOv5的目标检测优化算法,添加DyHead检测头,融合多个注意力机制,增强模型的检测精度;更换aLRPLoss损失函数,减少超参数调节工作,优化训练过程;基于FasterNet提出C3-Faster,代替网络中的C3模块,以PConv的思想提升模型检测性能,减少模型体积;最后添加轻量级上采样算子CARAFE,扩大模型感受野,提升对不同大小目标的检测效果。实验结果表明,改进后的YOLOv5模型相比于原版模型,在钢材表面缺陷数据集上总体平均精度提高了4.174%,参数量减少了11.25%,计算复杂度减少了13.75%,权重体积减少了10.72%,检测性能高于SSD、RetinaNet、FCOS、YOLOv3、YOLOv4等主流目标检测算法,在工业检测中具有较高的应用价值。

关键词 目标检测;YOLOv5;钢材表面缺陷

DOI 10.3969/j.issn.2097-1915.2023.05.010

中图分类号 TP391.41 **文献标志码** A **文章编号** 2097-1915(2023)05-0080-08

Optimized Algorithm for Surface Defect Detection Based on YOLOv5

QU Yi, WANG Cheng*, YU Jiabo, KONG Yakang, CHEN Xiancong

(Fundamentals Department, Air Force Engineering University, Xi'an 710051, China)

Abstract The rapid and accurate detection of surface defects in materials has become an important objective across various research domains. To enhance detection efficiency and realize lightweight equipment, this paper proposes a target detection optimization algorithm based on YOLOv5, adding DyHead detection head to enhance the detection accuracy of the model by fusing multiple attention mechanisms; replacing the aLRPLoss loss function to reduce the hyperparameter adjustment work and optimize the training process; propose C3-Faster based on FasterNet to replace the C3 module in the network to improve the model detection performance and reduce the model size with the idea of PConv; finally add the lightweight upsampling operator CARAFE to expand the model perceptual field and improve the detection effect on targets of different sizes. The experimental results show that the improved YOLOv5 model improves the overall average accuracy by 4.174%, reduces the parameter volume by 11.25%, reduces the computational complexity by 13.75%, and reduces the weight volume by 10.72% on the steel surface defect dataset compared with the original model, and the detection performance is also higher than that of SSD, RetinaNet, FCOS, YOLOv3, and YOLOv4 and other mainstream target detection algorithms, which have high application

收稿日期: 2023-05-26

基金项目: 陕西省自然科学基金(2023-JC-QN-0696)

作者简介: 渠逸(2000-),男,内蒙古赤峰人,硕士生,研究方向为表面微缺陷智能检测技术。E-mail: strsky778@163.com

通信作者: 汪诚(1974-),男,江苏扬州人,教授,博士生导师,研究方向为航空测试技术与表面强化。E-mail: valid_01@163.com

引用格式: 渠逸,汪诚,余嘉博,等.基于YOLOv5的表面缺陷检测优化算法[J].空军工程大学学报,2023,24(5):80-87. QU Yi, WANG Cheng, YU Jiabo, et al. Optimized Algorithm for Surface Defect Detection Based on YOLOv5[J]. Journal of Air Force Engineering University, 2023, 24(5): 80-87.

value in industrial detection.

Key words object detection; YOLOv5; steel surface defects

随着现代工业的发展,各行业对金属材料性能的把控能力愈加重要,钢材作为目前产量最高、应用最广的金属材,在交通运输、机械制造、航空航天和国防工业等领域其他金属材料是无法替代的。随着我国工业水平的迅速发展,市场需求不断变化,各领域对钢材的质量要求也在不断提升。钢材在生产制造的过程中,受到生产设备、工艺水平等因素制约,钢材表面会产生如裂纹、划痕、凹坑、斑块等无规律的缺陷^[1],导致在后续应用中,材料的力学性能、抗腐蚀性、抗磨损性能都会有不同程度地下降,严重的缺陷甚至会引发安全事故^[2]。因此,探索一种准确、高效率的表面缺陷检测方法成为了当前工业发展的迫切需求。

近年来,随着深度学习的发展,计算机视觉广泛应用在各个领域,对材料表面的目标检测是计算机视觉中重要的一环,该项技术大幅降低了人工成本,相较于传统人工检测手段具有稳定、高效、准确度高等优势。

当前基于深度学习的目标检测算法主要分为 one-stage 和 two-stage 2 种。one-stage 目标检测算法的检测速度快,但精度相对较低,代表算法有 SSD^[3]、YOLO^[4]等。two-stage 目标检测算法检测精度较高,但检测速度慢,以 Fast R-CNN^[5]、Faster R-CNN^[6]、Mask R-CNN^[7]等为代表。对于一般的物体表面缺陷检测,one-stage 目标检测算法的精度已经可以满足检测需求,又因为其具有更高的速度,所以被广泛应用^[8]。YOLO 网络作为 one-stage 检测算法里最受欢迎的系列,近年来被大量学者使用和优化。Fang 等^[9]通过改进的 YOLOv3 算法实现了对金属表面的缺陷检测,通过对输入端的图像进行处理和损失函数优化提升了检测效果。李维刚等^[10]使用加权 K-means 聚类算法,并通过融合浅层特征与深层特征的方式改进了 YOLOv3 网络,在带钢表面缺陷数据集上的检测精度比原始网络提高了 11%。Li 等^[11]通过构建全卷积 YOLO 检测网络检测带钢表面缺陷,为此类检测提供了端到端的检测方案,并取得了 97.55 的 mAP。Zhang 等^[12]提出了改进的 YOLOv5 算法,该算法在原有算法的基础上增加了微尺度检测层,并加入了 CBAM 关注机制,以控制小目标缺陷等缺陷的特征信息丢失。

大量研究已经初步使目标检测的精度和检测速度得到提升,但各类材料的应用场景变化较大,为满

足各类环境下个人检测任务需求,检测设备需尽可能轻量化。在此基础上,本文针对工业应用场景,以提升检测精度与速度、减小模型体积为目标,提出了改进的 YOLOv5 算法,并在 NEU-DET 数据集上进行测试。

1 实验基础

1.1 数据集

本文所涉及的实验数据集均采用东北大学宋克臣团队制作的钢材表面缺陷数据集 (NEU-DET),数据集内包含 6 个缺陷类别: crazing, inclusion, patches, pitted surface, rolled in scale, scratches, 每类缺陷 300 张识别图像,各类别的缺陷标注信息通过 xml 文件保存,共计 1 800 张灰度图像,4 189 个检测缺陷边界框。

将数据集内 1 800 张图像按照 8 : 1 : 1 的比例划分为训练集、验证集和测试集,得到训练图片 1 440 张,验证图片和测试图片各 180 张,数据集内各类缺陷图片如图 1 所示,其中缺陷目标所在位置位于红框标注处。

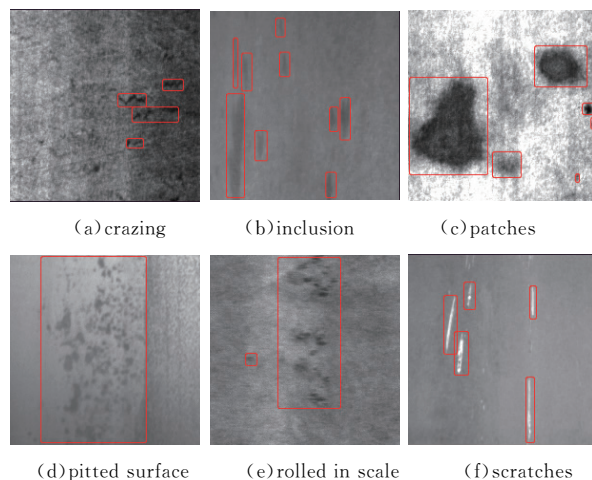


图 1 数据集部分图片

1.2 YOLOv5

YOLO 是一种通用的 one-stage 目标检测算法,侧重于推理速度,被广泛应用在图像或视频的目标分类、定位工作中。迄今为止,学者通过不断改进,发布了各种版本的 YOLO 模型,其中 YOLOv5 便是在 YOLOv4 的基础上改进而来,其在输入端的模型训练阶段加入了 Mosaic 数据增强、自适应图片缩放和锚框计算功能;融合了 Focus 结构与 CSP 结构改进了基准网络;在 Neck 部

分添加了 FPN+PAN 结构;改进了预测筛选框的 DIoU NMS 和训练时的 GIOU Loss,对比前代在精度和速度上都得到了很大提升。YOLOv5 目前已经更新到 7.0 版本,根据不同的使用场景,有 n/s/m/l/x 5 个模型可以使用,由 n 至 l 版本,模型大小和运算量逐步增加,精度也同时提高。根据本文的实验对象,经过对比试验后,采用体积相对较小、速度较快的 s 版本模型。

2 算法改进

2.1 添加目标检测头

DyHead 是由 Dai 等^[13]提出的一种基于注意力机制的目标检测头,通过将多头 self-attention 注意力机制结合在尺度感知的特征层、空间感知的空间位置、任务感知的输出通道处,并将其统一在一个框架内,以此在没有多余算力开销的情况下有效提升目标检测头的检测性能。其结构如图 2 所示。

从图 2 可以看出,DyHead 将尺度注意力、空间

注意力和任务注意力叠加在了一个模块内,将 Head 部分的输入,也就是 General View 视为一个三维的 tensor,其中 L 为不同尺度(即不同层次和阶段)的特征图; S 为空间位置信息,也就是特征图的宽高乘积; C 为通道信息。DyHead 使用分离式的注意力机制,使以上每个维度都能独立通过注意力机制进行特征感知,3 个注意力机制如下:

1)Scale-aware Attention 为添加在 level 维度的尺度感知注意力,不同 level 的特征图所对应的目标尺度也不同,该机制的引入可以增强模型的尺度感知能力。

2)Spatial-aware Attention 为添加在 spatial 维度的空间感知注意力,当空间位置发生变化,所对应的便是检测目标的几何变换,该机制的引入可以增强模型的空间位置感知能力。

3)Task-aware Attention 为添加在 channel 维度的任务感知注意力,不同的 channel 所对应的检测任务也不同,该机制的引入可以增强模型对不同任务的感知能力。

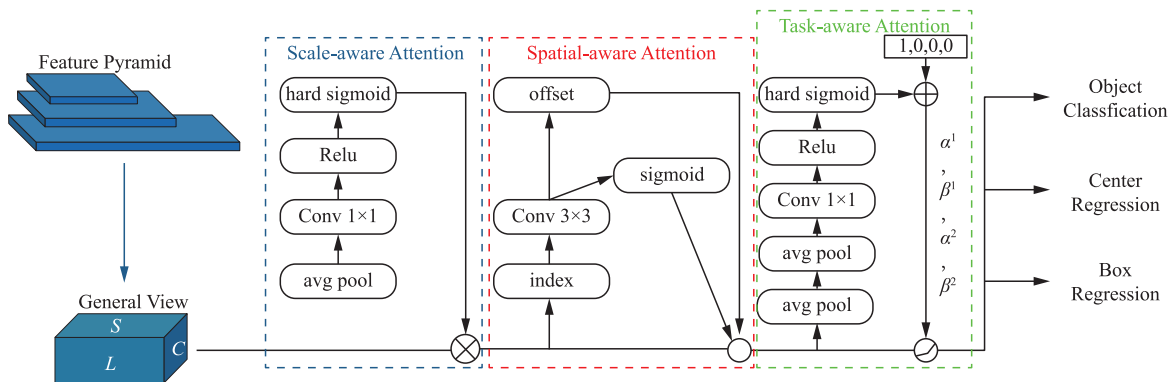


图 2 DyHead 结构

2.2 改用 aLRP Loss

一般的目标检测任务需要通过结合平衡超参数,来联合优化分类目标与定位目标,但这种方法在定位任务和分类任务之间缺乏关联,无法同时兼顾定位精度和分类精度;在训练过程中,对超参数的调节较为困难,导致在需要较长时间训练的网络内收敛变慢;容易受到正负样本不均匀或者正负样本异常值变动的影 响,在参数调节上需花费更多精力^[14]。为解决以上问题,Oksuz 等^[15]提出了 aLRP Loss。这是一种基于 ranking 的损失函数,具有统一、有界、平衡等特点,通过对 AP Loss 的扩展,建立一个能够优化不可导排序损失的通用框架,能有效用于目标检测中的分类和定位任务。与精确度和 AP 损失之间的关系类似,aLRP Loss 为正样本 LRP 值($l^{\text{LRP}}(i)$)的平均值,如式(1)所示:

$$L^{\text{aLRP}} = \frac{1}{|P|} \sum_{i \in P} l^{\text{LRP}}(i) \quad (1)$$

式中: P 为精确率。LRP 值定义为:

$$l^{\text{LRP}}(i) = \frac{1}{\text{rank}(i)} \cdot$$

$$\left[N_{\text{FP}}(i) + \epsilon_{\text{loc}}(i) + \sum_{k \in P, k \neq i} \epsilon_{\text{loc}}(k) H(x_{ik}) \right] \quad (2)$$

$$\text{rank}(i) = N_{\text{FP}} + N_{\text{TP}} \quad (3)$$

$$\epsilon_{\text{loc}}(k) = \frac{[1 - \text{IoU}(k)]}{1 - \tau} \quad (4)$$

式中: N_{FP} 为错误预测的数量; N_{TP} 为正确预测的数量;当 $k \geq 0$ 时, $H(k) = 1$,否则 $H(k) = 0$; τ 为 TP 标记阈值,一般取 0.5。

相对于传统损失函数,aLRP Loss 具有以下优点:

1)在分类和定位任务中首次采用排名策略,将独立工作的 2 个分支合为统一整体,两者同时调用检测器的全部输出,从而同时兼顾检测精度和定位质量。

2)相对于目前 SOTA 方法中的 6 个超参数, aLRP Loss 仅有一个不需要调试的超参数,能够减少参数调节的工作量,并使模型更容易训练。

3)aLRP Loss 由 AP Loss 拓展而来,能够替换 AP Loss、SmoothL1 Loss 等损失函数,并以单个损失函数参与训练,具有出色的正负样本策略,以提高网络性能。

2.3 基于 FasterNet 轻量化模型提出 C3-Faster

为提高算法的检测速度,Chen 等人^[16]提出了一种新的部分卷积,其基本思想是在卷积操作中只对输入中非零的部分进行卷积,对输入中全零的部分不进行操作,可以同时减少计算冗余和内存的访问,更有效地提取空间特征,并在其基础上提出了 FasterNet,在不影响检测精度的同时,大幅提升了检测速度。

在此基础上,本文提出了 C3-Faster,替换了 YOLOv5 中用于提取特征的 C3 模块。在 C3-Faster 中调用了 C3 模块的构造函数,初始化部分基本参数,并引入了 Partial Convolution 部分卷积层和 Faster_Block 模块,二者都使用了类似 PConv^[17]的思想,其通过使用一个额外的掩蔽矩阵实现,该矩阵与输入图像大小相同,其中缺失或遮挡的像素位置用 0 填充,而正常的像素位置用 1 填充。随后卷积核仅对被预设为 1 的像素进行卷积运算,其中在掩蔽矩阵中对应为 0 的像素位置,卷积核中相应的权重也被设为 0。这意味着在卷积计算过程中,这些被掩蔽的权重不会对输出特征图产生影响,不但降低了运算量,在处理不完整的输入数据时也能提供更好的结果,从而提高了模型的鲁棒性和泛化能力,可以增强模型的非线性表达能力,提高检测精度;Faster_Block 通过部分卷积来处理输入特征图,从而减少计算量和模型参数。

2.4 添加轻量级上采样算子 CARAFE

上采样在目标识别算法中可以将低分辨率的物体特征图扩展到更高分辨率,在卷积神经网络中,高层特征图中的目标特征比较抽象,分辨率较低;低层特征图中物体特征更为精细,分辨率更高。通过将低层特征图上采样到高层尺寸,可以将低层精细的特征和高层抽象的特征结合起来,从而更好地捕获物体的特征,提高识别的准确性。

为了深度神经网络中实现特征图的高效上采样,Wang 等^[18]提出了轻量级通用上采样算子(content-aware reassembly of features, CARAFE),这是一种通用性强、轻量化、高效率的模块,相比于传统上采样模块,CARAFE 在更轻、更快的条件下,感

受野进一步扩大,不依靠亚像素邻域运行,而是在更大的接受域内融合信息,并且可以针对特定内容动态生成自适应内核,能够更优秀地进行内容感知。

CARAFE 运行的具体步骤是在每个位置为中心的预定义区域内,使用加权组合来重新组装特征,并通过内容感知的方式生成权重,并且上述的每个位置都会存在多组生成的上采样权重,最后将生成的权重重新排列为一个完整的空间快,以此进行特征上采样。具体来说,给定一个大小为 $C \times H \times W$ 的特征图 X 和一个上采样率 σ (假设 σ 为一个整数),将生成一个大小为 $C \times \sigma H \times \sigma W$ 的新特征图 X' ,其原理如式(5)、(6)所示:

$$W_{l'} = \psi(N(X_l, K_{\text{encoder}})) \quad (5)$$

$$X'_{l'} = \varphi(N(X_l, K_{\text{up}}), W_{l'}) \quad (6)$$

对于输出 X' 的任一个目标位置 $l' = (i', j')$,在输入 X 中都能找到一个相应的源位置 $l = (i, j)$,其中 $N(X_l, k)$ 为以 l 为中心的特征图 X 的 $k \times k$ 子区域, K_{up} 为重组核的大小,同时考虑到性能和效率,使用核大小的卷积层 K_{encoder} 一般取值为 $K_{\text{up}} - 2$ 。

式(5)中,内核预测模块 ψ 根据 X_l 的子区域,为每个 l' 预测该位置上的内核 $W_{l'}$,即输出大小为 $C_{\text{up}} \times H \times W$ 的重组核, C_{up} 为:

$$C_{\text{up}} = \sigma^2 k_{\text{up}}^2 \quad (7)$$

式(6)中 φ 是内容感知重组模块,可以将 X_l 的子区域与内核 $W_{l'}$ 重新组合,如式(8)所示:

$$\varphi(X'_{l'}) = \sum_{n=-r}^r \sum_{m=-r}^r W_{l'(n,m)} X_{(i+n, j+m)} \quad (8)$$

式中: $r = K_{\text{up}} / 2$

2.5 改进的 5 种模型

基于以上 4 点改进方向,本文共提出 5 种改进的 YOLOv5 模型如下:

1)YOLOv5-D:为 YOLOv5 网络添加基于分离式注意力机制的 DyHead 目标检测头;

2)YOLOv5-A:改用由 AP_Loss 拓展而来的 aLPRLoss 损失函数;

3)YOLOv5-F:将 YOLOv5 算法中的 C3 模块替换为基于 FasterNet 网络提出的 C3-Faster;

4)YOLOv5-C:为 YOLOv5 网络添加 CARAFE 轻量级上采样算子;

5)YOLOv5-DAFC:将 Dyhead 目标检测头、aLPRLoss 损失函数、C3-Faster 模块、CARAFE 上采样算子同时部署在 YOLOv5 网络中。

改进后的 YOLOv5-DAFC 网络结构如图 3 所示:

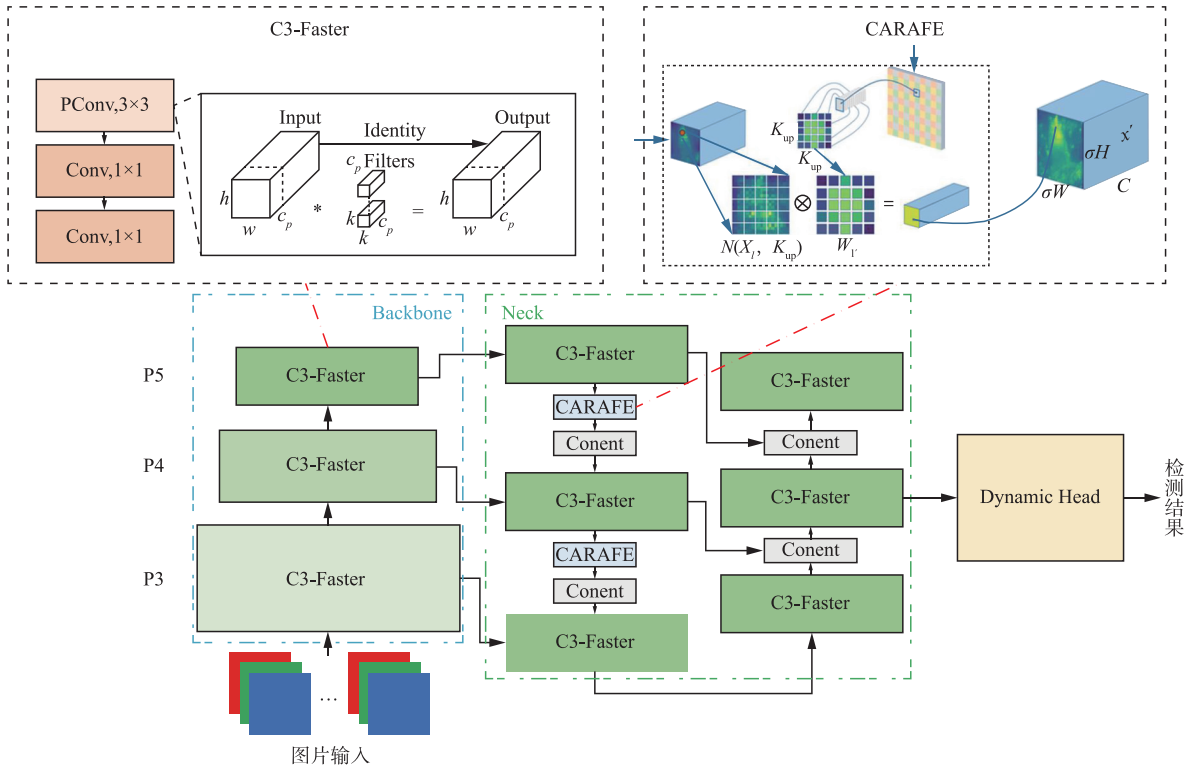


图3 YOLOv5-DAFC 结构

3 实验

本文实验环境基于 Windows 操作系统,运行内存 128 GB,使用 Pytorch 1.11.0 作为深度学习框架,Python 版本为 3.8,CUDA 版本为 11.2。CPU 为 Intel(R) Xeon Gold 6330 CPU @ 2.00 GHz。GPU 为 NVIDIA GeForce RTX 2080Ti 11G。

3.1 实验评价标准

本文从模型自身参数、模型检测参数两方面进行效果对比。

其中模型自身参数包括以下几类:模型参数量,即网络结构中所有可训练参数的数量,通常是衡量模型复杂度和计算所需资源的指标;FLOPs,指模型在推理期间进行的浮点运算次数,一般采用 GFLOPs,通常被用来评估模型的计算复杂度;权重体积,指模型训练后所生成权重文件的大小。

模型的检测参数包括以下几类:精确率,指模型正确预测的目标数量与总预测目标数量之比;召回率,即在所有真实目标数量中,模型准确预测的目标数量占比。 $AP^{[19]}@0.5$ (IoU 阈值为 0.5 时各个类别的平均 AP 值)而 mAP 则是在不同召回率下的精确率的均值,这一指标考虑了准确率和召回率的综合表现,是当前主流的目标检测评价指标。相关计算公式为:

$$P = \frac{T_p}{(T_p + F_p)} \quad (9)$$

$$R = \frac{T_p}{(T_p + F_N)} \quad (10)$$

$$A_p = \sum_{i=1}^n P(i) \Delta R(i) = \int_0^1 P(R) dR \quad (11)$$

$$A_{MP} = \frac{\sum_{i=1}^N P_{A_i}}{N} \quad (12)$$

式中: P 、 R 分别为精确率和召回率; N 表示目标类别数; T_p 为 $IoU > 0.5$ 时的检测框数量(同一个 Ground Truth 值计算一次); F_p 为 $IoU \leq 0.5$ 时的检测框数量(或者是检测到同一个 Ground Truth 的多余检测框的数量); F_N 为指没有检测到的 Ground Truth 数量; A_{MP} 为均值平均精度,即 mAP。

3.2 训练参数设置

YOLOv5 使用了 K-means 聚类算法来生成锚框(anchor box),通过统计训练集中不同大小的目标,自动生成合适的 K 值(锚框的数量),以便网络能够检测不同大小和比例的目标。本文针对目标数据集与服务器性能,初始学习率设为 0.01,使用随机梯度下降法(stochastic gradient descent,SGD),warmup_epoch 设为 5,学习率动量为 0.98,权重衰减系数 weight_decay 设为 $5e-5$,work-number 设为 16,原始 YOLOv5 模型 batch-size 设为 32,后续训练根据模型大小与显存冗余调整至最优。依次对原 YOLOv5 模型的 s 版本与 5 个改进模型进行训练

与测试,对比分析各模型的参数和性能。

3.3 训练过程损失

在训练过程中,本文提取了 YOLOv5 原模型与 YOLOv5-DAFC 模型的损失函数值,如图 4 所示,

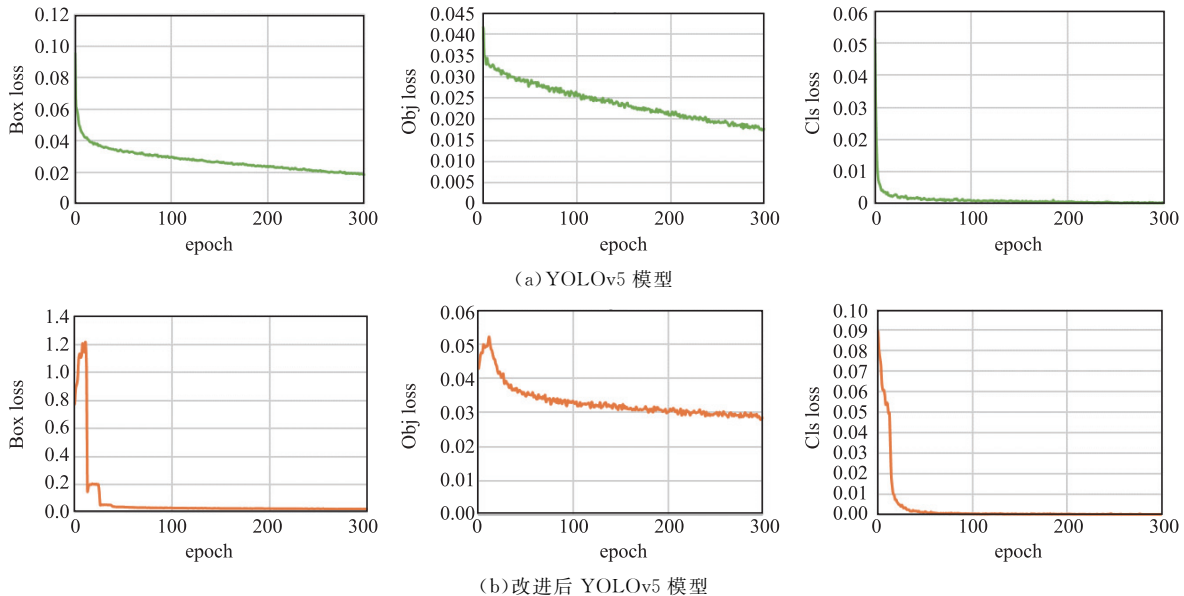


图 4 训练损失曲线

图 4 中 Box loss 是定位损失,通过计算预测边界框与实际边界框差异所得,用于衡量模型对物体位置的预测能力。Obj loss 是置信损失,通过 Focal Loss 计算所得,用于衡量模型对预测物体与背景的置信度误差。Cls loss 是分类损失,通过计算分类预测结果与实际标签的交叉熵损失函数所得,用于衡量模型对每个类别的分类能力。

可以看出,在训练 epoch 达到 300 时,2 个模型均已收敛,损失值趋于平稳。添加了 aLRP Loss 的改进模型损失曲线与原版区别较大,在 Box loss 曲线中,DAFC 模型在训练初期出现了过拟合^[20]现象,相比原版的平滑下降,呈阶段性下降趋势,并且收敛速度明显快于原版,仅在 50 epoch 左右就已经进入收敛状态,且收敛后曲线平稳度高于

原版;在 Obj loss 曲线中,DAFC 模型训练初期仍有过拟合现象存在,但模型收敛后曲线平稳度仍高于原版。可以看出,引入了 aLRP Loss 的 DAFC 模型在收敛速度和收敛后曲线稳定性方面均高于原版模型。

4 实验结果对比

4.1 模型性能对比

为确定每个优化结构对网络性能的影响,本文进行消融实验,分别对 YOLOv5、YOLOv5-D、YOLOv5-A、YOLOv5-F、YOLOv5-C、YOLOv5-DAFC 6 个模型进行训练,得到不同模型自身参数与测试性能如表 1 所示。

表 1 不同模型的参数与检测效果对比

检测模型	Parameters	GFLOPs	P/%	R/%	AP@0.5/%	Weight/KB
YOLOv5	7035811	16.0	73.1	64.8	72.817	14 137
YOLOv5-D	7598355	17.2	69.7	71.5	73.918	14 746
YOLOv5-A	7035811	16.0	64.5	73.4	72.307	14 137
YOLOv5-F	5795779	12.7	72.4	72.4	75.642	11 727
YOLOv5-C	7035811	16.0	74.5	68.9	72.888	14 137
YOLOv5-DAFC	6244235	13.8	71.2	72.6	76.991	12 622

通过分析表 1 实验结果可知,在网络中增加 DyHead 目标检测头、用 C3-Faster 替换 C3 模块后,网络整体的训练参数量和计算复杂度会有较为明显的变化,前者增加了 8% 的参数量和 7.5% 的浮点运算量,将 AP@0.5 提升了 1.101%,同时权重体积

增加了 4.3%;后者则减少了 17.62% 的参数量和 20.63% 的浮点运算量,同时将 AP@0.5 提升了 2.825%,提升效果显著。

而在改用 ALRPLoss 损失函数与增加 CARAFE 上采样算子的 A、C 模型中,相应的改进主要

作用在检测精度的提升上,A、C 两模型相比于原版模型,AP@0.5 分别提升了 0.49%和 1.071%。

DAFC 模型则是 4 个模型优点的整合,保留 4 个改进点对检测精度提升的同时,综合减少模型参数量与计算复杂度。相比于原模型,改进的 DAFC 模型在参数量减少了 11.25%,计算复杂度减少了 13.75%,权重体积减少了 10.72%的同时,AP@

0.5 增加了 4.174%。可以看出,改进的 DAFC 模型同时发挥了 4 个改进点的优势,在尽可能节约内存占用与计算资源的同时,做到了检测效果的最佳优化。

4.2 不同缺陷检测性能对比

为研究本文所提出的模型对于不同类别缺陷的检测效果,摘取实验数据如图 5 所示。

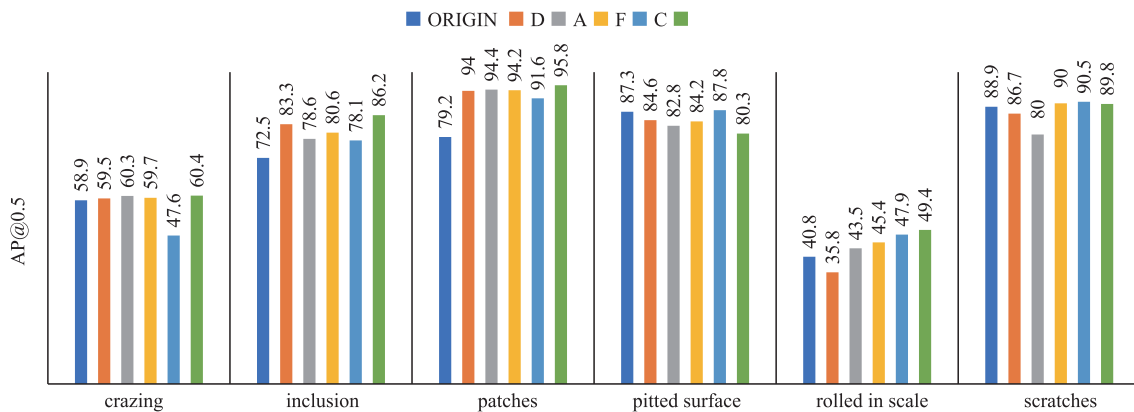
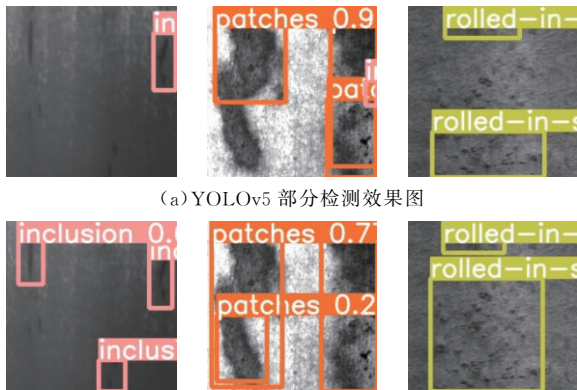


图 5 各类缺陷检测效果

从图 5 的分析结果可以看出,对于本数据集内的 6 种缺陷类别,最优检测效果均分布在 5 个改进模型上,其中 C 模型对 pitted surface 和 scratches 类别缺陷检测效果最佳,检测精准度分别提升了 0.5%、1.6%;对于 crazing、inclusion、patches、rolled in scale 4 种缺陷,均为 DAFC 模型检测效果最佳,相比于原版模型,改进后的 DAFC 模型对 crazing 的识别准确率提升了 1.5%,inclusion 提升了 13.7%,patches 提升了 16.6%,rolled in scale 提升了 8.6%,scratches 提升了 0.9%。改进后的模型对本数据集整体检测效果大幅提升,并对部分难检目标的检测效果提升明显,对实际应用场景的可用性有实质性的加强。

4.3 检测效果对比

为了直观了解改进后的模型对目标缺陷的检测效果提升,使用原模型与改进后的 5 种模型对测试集 180 张图片进行了检测,YOLOv5s 与改进后的 YOLOv5-DAFC 模型对部分测试图片识别效果如图 6 所示。



(a) YOLOv5 部分检测效果图

(b) YOLOv5-DAFC 部分检测效果图

图 6 部分检测效果图

通过对比可以看出,原版模型对较小目标容易发生漏检,对较大目标存在定位错误,并且存在重复描框、同一目标检测为多个等问题,改进后的 DAFC 模型相对于原版 YOLOv5,感受野增大后对较大目标识别效果提升,对较难检测的目标更为敏感,可以有效避免目标漏检的情况发生,并且对于紧邻的多个目标、叠加目标的检测策略更加智能,显著减少了对同一个目标重复描框、分段描框的情况和将多个目标识别为同一对象的情况,有效避免原版模型的检测错误问题。

4.4 主流检测算法性能对比

为了进一步评估本文提出的模型的性能,在相同实验条件下,在同一数据集上选择了 SSD、RetinaNet、FCOS、YOLOv3、YOLOv4 和 YOLOv7 6 种主流的一阶目标检测算法,与本文提出的 YOLOv5-DAFC 模型进行实验对比分析,通过比较实验结果,可以更加客观地评估本文所提出改进模型的有效性和优越性。具体实验结果如表 2 所示。

表 2 与主流目标检测算法的性能对比

算法	AP@0.5/%	FPS/s	Weight/MB
SSD	27.71	19.4	96.2
RetinaNet	69.58	13.1	35.68
FCOS	68.82	9.0	87.25
YOLOv3	69.75	20.7	29.51
YOLOv4	67.29	19.4	32.63
YOLOv7	71.23	43.1	71.40
YOLOv5-DAFC	76.99	54.3	12.32

通过以上对比可以明显看出,本文所提出的YOLOv5-DAFC模型相比于几种主流的 one-stage 目标检测算法,在整体检测精度、检测速度、模型体积3个方面都有较大提升,并且在各方面均超过了更新的YOLOv7模型,基本可以实现高精度、高速度、轻量级需求的工业检测。

5 结语

为实现精准、快速的材料表面缺陷检测,本文提出了改进的YOLOv5优化算法,添加DyHead目标检测头,通过分离式注意力机制将多种注意力同时应用在同一模块中,有效提升目标检测头的检测性能;更换ALRPLoss损失函数,简化训练参数、优化训练效果的同时兼顾分类和定位精度;基于FasterNet网络提出C3-Faster模块,用于替换原版C3模块,提升模型鲁棒性同时减少权重体积;添加CARAFE上采样算子,扩大感受野,提升内容感知性能。实验证明,改进后的YOLOv5-DAFC模型相比原版YOLOv5模型,检测精度提升的同时,模型参数量、计算复杂度、权重体积都有明显优化,对于轻量化工业检测领域有较高应用价值。

参考文献

- [1] 彭向前. 产品表面缺陷在线检测方法研究及系统实现[D]. 武汉:华中科技大学,2008:13-24.
- [2] 许锬俊. 缺陷、损伤、微裂纹对航空发动机构件服役总寿命及可靠性的影响[J]. 航空发动机,2003(2):5.
- [3] BERG A C, FU C Y, SZEGEDY C, et al. SSD: Single Shot MultiBox Detector: 10.1007/978-3-319-46448-02[P]. 2015.
- [4] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]// Computer Vision & Pattern Recognition. Piscataway: IEEE Press, 2016.
- [5] GIRSHICK R. Fast R-CNN [C]//IEEE International Conference on Computer Vision. Piscataway: IEEE Press 2016:1440-1448.
- [6] 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 on Pattern Analysis and Machine Intelligence, 2017, 39(6):1137-1149.
- [7] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN [C]// Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017:2961-2969.
- [8] 周俊宇, 赵艳明. 卷积神经网络在图像分类和目标检测应用综述[J]. 计算机工程与应用, 2017, 53(13):8.
- [9] 方叶祥, 甘平, 陈俐. 金属表面缺陷检测的改进YOLOv3算法研究[J]. 机械科学与技术, 2020, 39(9):5.
- [10] 李维刚, 叶欣, 赵云涛, 等. 基于改进YOLOv3算法的带钢表面缺陷检测[J]. 电子学报, 2020, 48(7):9.
- [11] LI J, SU Z, GENG J, et al. Real-Time Detection of Steel Strip Surface Defects Based on Improved YOLO Detection Network-Science Direct[J]. IFAC PapersOn Line 2018, 51(21):76-81.
- [12] ZHANG R, WEN C. SOD-YOLO: A Small Target Defect Detection Algorithm for Wind Turbine Blades Based on Improved YOLOv5 [J]. Advanced Theory and Simulations, 2022, 7(5):2100631.
- [13] DAI X, CHEN Y, XIAO B, et al. Dynamic Head: Unifying Object Detection Heads with Attentions [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, TN: IEEE Press, 2021.
- [14] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation [C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH: IEEE Press, 2014.
- [15] OKSUZ K, CAM B C, AKBAS E, et al. A Ranking-based, Balanced Loss Function Unifying Classification and Localisation in Object [DB/OL]. arXiv Preprint: 2609.12592, 2020.
- [16] CHEN J, KAO S, HE H, et al. Run, Don't Walk: Chasing Higher FLOPS for Faster Neural Networks [DB/OL]. ArXiv Preprint :2303.03667, 2023.
- [17] LIU G, REDA F A, SHIH K J, et al. Image Inpainting for Irregular Holes Using Partial Convolutions [DB/OL]. ArXiv Preprint:1804.07723, 2018.
- [18] WANG J, CHEN K, XU R, et al. CARAFE: Content-Aware ReAssembly of Features [C]//2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea: IEEE Press, 2019.
- [19] EVERINGHAM M, GOOL L V, WILLIAMS C, et al. The Pascal Visual Object Classes (VOC) Challenge [J]. International Journal of Computer Vision, 2010, 88(2):303-308.
- [20] 陈大伟, 闫昭, 刘昊岩. SVD系列算法在评分预测中的过拟合现象[J]. 山东大学学报(工学版), 2013, 44(3):15-21.

(编辑:徐楠楠)