

# 改进 YOLOv5 的钢材表面缺陷检测算法

李鑫, 汪诚\*, 李彬, 郭振平, 李秋良, 李卓越

(空军工程大学基础部, 西安, 710038)

**摘要** 针对传统钢材表面缺陷检测方法存在检测效率低、检测精度差等问题, 提出一种基于改进 YOLOv5 的钢材表面缺陷检测算法。首先使用 GhostBottleneck 结构替换原 YOLOv5 网络中的 C3 模块和部分卷积结构, 实现网络模型轻量化; 其次在 Backbone 部分引入 SE 注意力机制, 对重要的特征通道进行强化; 最后针对数据集特点在网络中增加一个检测层, 强化特征提取能力, 并在 Neck 部分增加特征融合结构, 使用 DW 卷积替换部分标准卷积以减少运算量。实验表明, 改进的 YOLOv5s-GSD 算法, 模型体积减少了 10.4%, 在测试集上的 mAP 值为 76.8%, 相比原 YOLOv5s 网络提高了 3.3%, 检测精度和速度也明显高于一些主流算法。相比传统的钢材表面缺陷检测方法, 提出的算法能够更加准确、快速地检测出钢材表面缺陷的种类和位置, 并且具有较小的模型体积, 便于在移动端的部署。

**关键词** 钢材表面缺陷; YOLOv5; 检测算法; 注意力机制

**DOI** 10.3969/j.issn.1009-3516.2022.02.005

**中图分类号** TP391.4 **文献标志码** A **文章编号** 1009-3516(2022)02-0026-08

## Steel Surface Defect Detection Algorithm Based on Improved YOLOv5

LI Xin, WANG Cheng\*, LI Bin, GUO Zhenping, LI Qiuliang, LI Zhuoyue

(Fundamentals Department, Air Force Engineering University, Xi'an 710038, China)

**Abstract** Aiming at the problems of low efficiency and poor accuracy of traditional steel surface defect detection methods, a steel surface defect detection algorithm based on improved YOLOv5 is proposed in this paper. Firstly, the C3 module and part of convolutional structure in YOLOv5 network are replaced by GhostBottleneck structure to realize the lightweight of network model. Secondly, SE attention mechanism is introduced in Backbone to strengthen the important feature channels. Finally, according to the characteristics of the data set, a detection layer is added to the network to strengthen the feature extraction ability, and a feature fusion structure is added in the Neck part. DW convolution is used to replace part of the standard convolution to reduce the computation. Experimental results show that the improved YOLOv5s-GSD algorithm reduces the model volume by 10.4%, and the mAP value on the test set is 76.8%. Compared with the original YOLOv5s network, the detection accuracy and speed are obviously higher than some mainstream algorithms. Compared with traditional steel surface defect detection methods, the algorithm proposed in this paper can detect the type and location of steel surface defects more accurately and quickly, and has a smaller model volume, which is convenient for deployment in mobile terminals.

**Key words** steel surface defect; YOLOv5; detection algorithm; attentional mechanism

收稿日期: 2021-12-09

基金项目: 陕西省重点研发计划(2020GY-307)

作者简介: 李鑫(1997—), 男, 河南许昌人, 硕士生, 研究方向为目标检测算法。E-mail: lx202006t10@163.com

通信作者: 汪诚(1974—), 男, 江苏扬州人, 教授, 研究方向为航空发动机高温材料及表面检测。E-mail: warrant\_74@163.com

**引用格式:** 李鑫, 汪诚, 李彬, 等. 改进 YOLOv5 的钢材表面缺陷检测算法[J]. 空军工程大学学报(自然科学版), 2022, 23(2): 26-33. LI Xin, WANG Cheng, LI Bin, et al. Steel Surface Defect Detection Algorithm Based on Improved YOLOv5[J]. Journal of Air Force Engineering University (Natural Science Edition), 2022, 23(2): 26-33.

钢材作为一种重要的工业产品,是航空航天、机械制造和国防工业等领域不可或缺的原材料。随着我国工业水平的不断发展,钢材生产技术取得了巨大进步<sup>[1]</sup>。而在钢材的生产制造过程中,由于受到原料质量、制造设备和生产环境等因素影响,产品表面会形成裂纹、孔洞、划痕等不同类型的缺陷<sup>[2]</sup>。这些缺陷会使钢材的强度、抗疲劳和耐磨性能下降,影响正常使用,甚至可能造成严重后果。因此,为保证钢材的安全性能,必须对其表面进行缺陷检测。

传统的钢材表面缺陷检测方法有人工检测方法和频闪光检测方法<sup>[3]</sup>,均为非自动化的检测方法。检测人员需要进行大量重复的工作,容易产生视觉疲劳,造成漏检或误检的情况。

近年来,机器视觉技术迅速发展,在图像分类、人脸识别和目标检测等领域得到了应用。机器视觉技术具有稳定、高效、准确等优势<sup>[4]</sup>,有学者将该技术应用于缺陷检测,提高了检测效率和精度。韩强等利用检测网络对 RPN 网络生成的区域建议框进行分类回归,提出了一种特征融合和级联检测网络的 Faster R-CNN 钢材表面缺陷检测算法<sup>[5]</sup>。叶欣提出了一种使用自适应空间特征融合结构和 EIoU 损失函数的改进 YOLOv4 算法,提高了对热轧带钢

表面缺陷的检测精度<sup>[6]</sup>。杨莉等通过引入导向锚点候选区域和多任务 FPN 结构解决了传统 Faster R-CNN 网络在检测钢材表面小目标缺陷时特征融合不充分的问题<sup>[7]</sup>。

本文基于钢材表面缺陷检测任务,提出一种改进的 YOLOv5 算法,旨在提高算法对钢材表面缺陷的识别准确率和检测速度,并在 NEU-DET 数据集上验证了算法的有效性。

### 1 YOLOv5 算法介绍

YOLO 系列算法<sup>[8]</sup>是一种基于深度学习和卷积神经网络的目标检测算法,具有推理速度快、检测精度高、可实时检测等优点。YOLOv5 是该系列算法的最新一代,相比前几代具有更高的精度、更快的速度和更小的体积。该算法共有 4 种模型:YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x,网络的深度和宽度依次增加。本文选择 YOLOv5s 作为预训练模型,在其基础上进行迁移学习,建立钢材表面缺陷检测算法。YOLOv5s 的网络结构如图 1 所示。

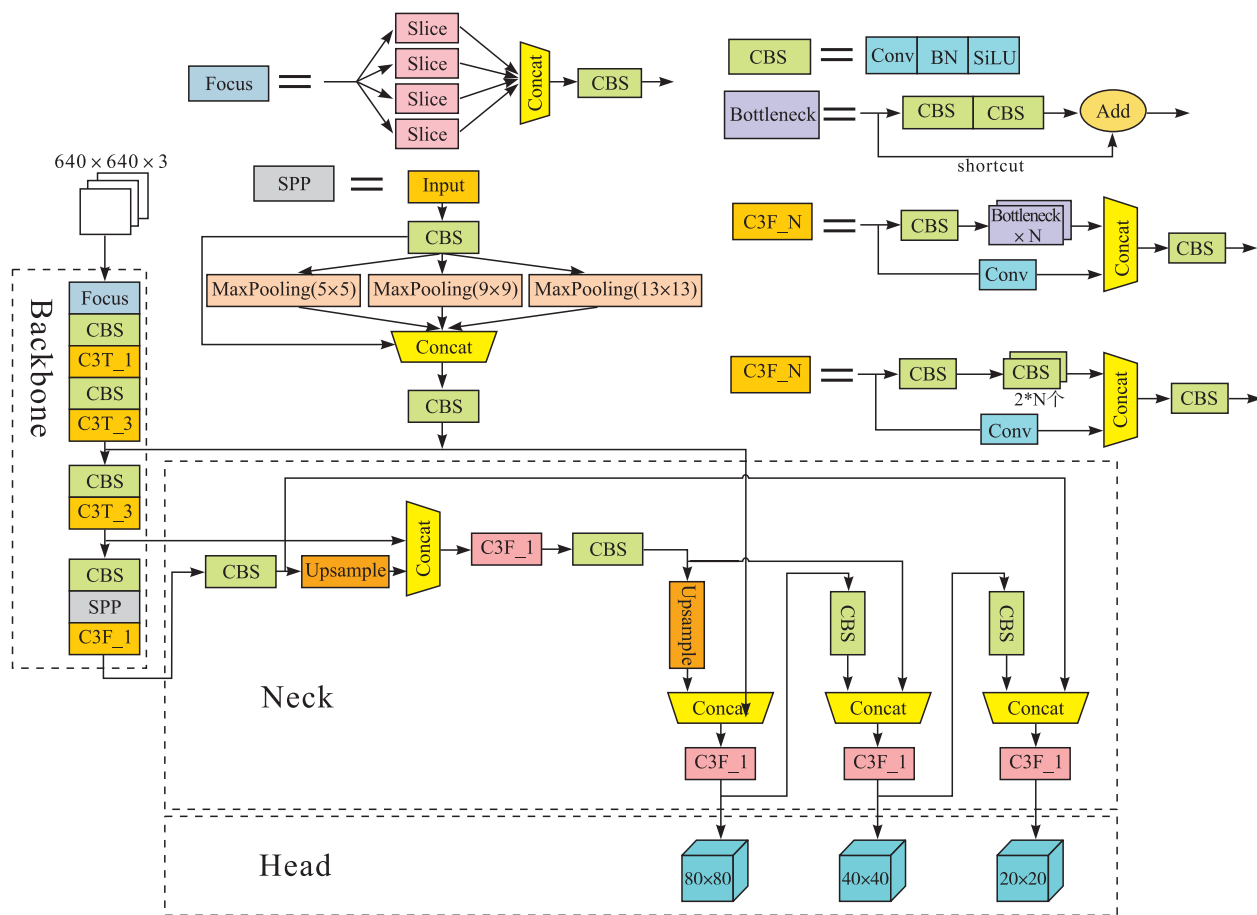


图 1 YOLOv5s 网络结构图

YOLOv5s 模型主要由 4 部分组成: Input、Backbone、Neck 和 Head。Input 部分采用了 Mosaic 数据增强,每次选取 4 张图片进行随机组合,丰富了数据集的多样性<sup>[9]</sup>;Backbone 部分用于提取图像特征,主要由 Focus、C3、SPP 等模块组成。Focus 用来对输入网络的图像进行切片操作,可以增加输出深度,同时减少计算量;C3 主要由若干个 Bottleneck 结构和卷积层组合而成,有利于加快推理速度;SPP<sup>[10]</sup> 为空间金字塔池化层,分别使用尺寸为 5、9、13 的 3 种池化核对输入图像进行最大池化操作,有效提升了网络的感受野;Neck 是网络的融合部分,采用了特征金字塔网络(feature pyramid networks,FPN)<sup>[11]</sup> + 金字塔注意力网络(pyramid attention networks,PAN)<sup>[12]</sup> 的结构,二者结合增强了对不同尺度特征的融合效果;Head 为网络的预测部分,输出 3 组包含预测框类别、置信度和坐标位置的向量。

## 2 改进的 YOLOv5 算法

### 2.1 使用 GhostBottleneck 结构

GhostNet<sup>[13]</sup> 是一种新颖的计算模块,可以通过使用较少的参数来生成更多的特征,原理如图 2 所示。使用该模块可以减少参数量、消除冗余特征、获得更加轻量化的模型。

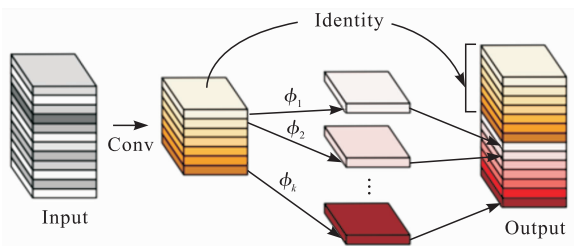


图 2 Ghost 模块原理

GhostNet 的具体过程为:

$$Y' = X * f' \quad (1)$$

$$y_{ij} = \Phi_{i,j}(y'_i), i \in 1, 2, \dots, m, j \in 1, 2, \dots, s \quad (2)$$

第 1 步为卷积操作,  $X$  为输入特征图,  $Y'$  为输出

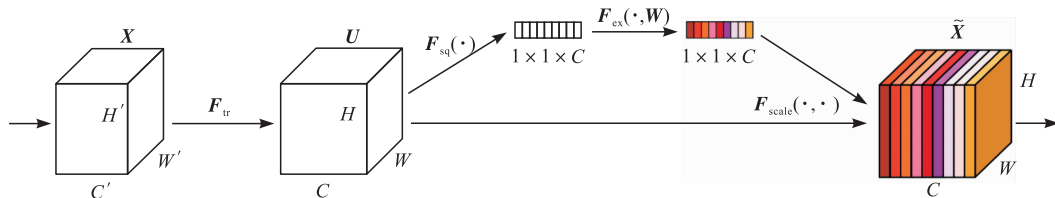


图 4 SE 模块

SE 模块主要包括 squeeze、excitation 和 scale 等操作:squeeze 将特征图的每个通道进行压缩,对  $H \times W$  的二维向量进行全局平均池化,输出  $1 \times 1 \times C$  的向量。

的  $m$  个特征图,  $f'$  为尺寸为  $k \times k$  的卷积核;第 2 步操作是对  $Y'$  中的每个特征图  $y'_i$  进行线性运算  $\Phi_{i,j}$ , 最终得到  $n = ms$  个输出特征图  $Y = [Y_{11}, Y_{12}, \dots, Y_{ms}]$ , 经计算可以得出普通卷积运算量约为 Ghost 模块的  $s$  倍。与普通的卷积神经网络相比, Ghost 模块在不改变输出特征图大小的情况下,降低需要的参数量和计算复杂度,实现模型轻量化。

GhostBottleneck 模块分为 2 种结构,如图 3 所示。步长为 1 的 GhostBottleneck 主要由 2 个 Ghost 模块堆叠而成;步长为 2 的 GhostBottleneck 的 2 个 Ghost 模块之间通过一个步长为 2 的 DW 卷积层连接。

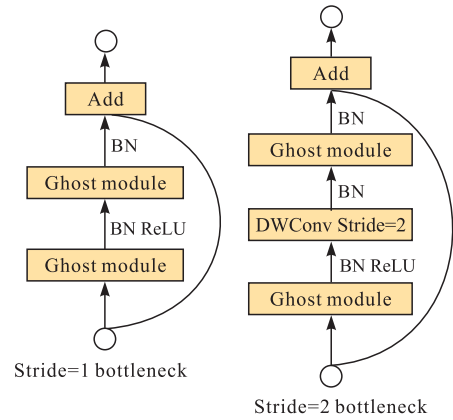


图 3 GhostBottleneck 的两种结构

YOLOv5s 网络中的 Bottleneck 结构对输入特征图先后使用 32 个  $1 \times 1$  卷积核和 64 个  $3 \times 3$  卷积核进行卷积操作,因此本文采用步长为 1 的 GhostBottleneck 替换网络中的 C3 模块,使用步长为 2 的 GhostBottleneck 替换网络中的 CBS 模块,以减少模型计算量。

### 2.2 添加注意力机制 SE 模块

SENet(squeeze and excitation network)<sup>[14]</sup> 是一种通道注意力机制,获得了 ImageNet 2017 竞赛分类任务的冠军。该网络的原理如图 4 所示,通过学习的方式自动获取每个特征通道的重要程度,并利用得到的重要程度来增强有用的特征,抑制不重要的特征,从而实现对特征权重的调整。

$$z_k = \frac{1}{H \times W} \sum_i^H \sum_j^W u_k(i, j) \quad (3)$$

excitation 在得到  $1 \times 1 \times C$  向量后,利用全连接层,对每个通道生成一个权重,用来代表特征通道的

重要程度。

$$s = \sigma(W_2 \delta(W_1 z)) \quad (4)$$

最后 scale 操作是将上一步输出的权重看作每个特征通道的重要性,再加权到之前的特征上,完成在通道维度上对原始特征的重新标定。

$$\tilde{X}_C = s c u_C \quad (5)$$

### 2.3 增加检测层

YOLOv5s 网络的 head 部分共输出 3 个尺度的检测层,输入尺寸为  $640 \times 640$  的图像,输出的检测层尺度为  $80 \times 80$ 、 $40 \times 40$  和  $20 \times 20$ ,分别用于检测小、中、大目标。考虑到本文所用数据集中部分缺陷占整张图片的比例较大,因此在原 YOLOv5s 网络的 head 部分再增加一个尺度为  $10 \times 10$  的检测层,相应地在 neck 部分添加特征融合结构,以输出新的特征图。这样在改进后的网络中,head 部分共输出 4 个尺度的检测层。

### 2.4 使用 DW 卷积

DWConv(depthwise convolution,深度卷积)<sup>[16]</sup>的原理是将卷积核拆分成单通道形式,在不改变输入特征图像深度的情况下,对每个通道进行独立的卷积运算,得到与输入特征图通道数一致的输出特征图。相比标准卷积,DW 卷积能够降低运算量,节约计算成本。本文在引入注意力机制并增加检测层后运算量增加较多,因此在 neck 部分使用 DW 卷积替代部分标准卷积,从而达到减少运算量的目的。

## 3 实验结果与分析

### 3.1 实验环境及数据集

本文的实验环境基于 Windows10 操作系统,内存为 64 GiB,使用 PyTorch 1.9.1 作为深度学习框架,Python 版本为 3.7.11,CUDA 版本为 10.2。CPU 为 Intel(R) Xeon(R) Gold 5218 CPU @2.30 GHz。GPU 为 NVIDIA GeForce RTX 2080Ti。

实验所用数据集为东北大学宋克臣团队制作的钢材表面缺陷数据集(NEU-DET),包含 6 种缺陷类型:裂纹(crazing)、内含物(inclusion)、斑块(patches)、点蚀面(pitted surface)、氧化轧皮(rolled-in scale)、划痕(scratches),共计 1 800 张图片。图片的原始尺寸为  $200 \times 200$ ,类别为灰度图。采集钢材表面缺陷的一种方法为:将 2 个 LED 光源对称倾斜安装于待测钢材上方,在两光源的中心轴线上放置一台工业相机,对钢材表面的缺陷进行照片采集,如图 5 所示。对相机采集到的图像进行预处理,消除高频噪声,进行灰度变换<sup>[17]</sup>。

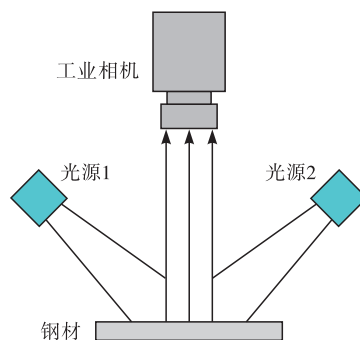


图 5 采集缺陷示意图

数据集图片均使用 LabelImg 软件依次对缺陷标注真实框,得到格式为 xml 的标签文件。将数据集的 1 800 张图片按照 6 : 2 : 2 划分为训练集、验证集和测试集,得到训练集 1 080 张,验证集 360 张,测试集 360 张。数据集部分图片如图 6 所示。

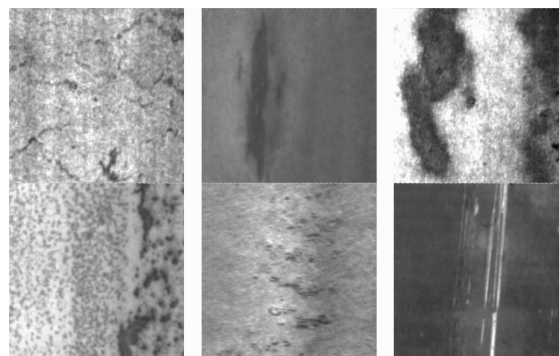


图 6 数据集部分图片

### 3.2 训练策略与评估指标

YOLOv5 网络的训练策略为:使用 K-means 聚类算法自动根据数据集的标注框计算三组初始锚框的大小;初始学习率 lr0 为 0.01,使用余弦退火算法动态调整学习率;学习率动量 momentum 为 0.937,权重衰减系数 weight decay 为  $5 \times 10^{-4}$ ;训练过程中开启 Mosaic 数据增强,提高模型的泛化能力;batch-size 设置为 16,训练轮次为 300 个 epoch。依次对原 YOLOv5s 模型和改进的模型进行训练和测试,对比分析各模型的参数和性能。

本文从模型参数量、运算量、模型体积、平均推理时间、精确率 P、召回率 R、mAP<sup>[18]</sup>@0.5(IoU 阈值为 0.5 时各个类别的平均 AP 值)等方面对模型进行比较。参数量表示训练过程中模型的总参数量;运算量 FLOPs 表示浮点运算数;模型体积表示训练完成后生成的模型权重大小;平均推理时间是指 GPU 推理一张图片所需的时间;精确率指检测出的正样本中实际仍为正样本的概率,召回率指在实际的正样本中检测出正样本的概率;平均精度 AP 指的是在不同召回率下精确率的均值,即 P-R 曲线与坐标轴围成的面积,反映了准确率和召回率的综合性能,是目前主流目标检测模型的评价指标。

对各个类别的 AP(式中简记为  $P_A$ ) 求均值即可得到 mAP(式中简记为  $P_{mA}$ ), 即所有类别的平均精度。它们的具体计算公式为:

$$P = \frac{P_T}{P_T + P_F} \quad (6)$$

$$R = \frac{P_T}{P_T + N_F} \quad (7)$$

$$P_A = \sum_{i=1}^n P(i) \Delta R(i) = \int_0^1 P(R) dR \quad (8)$$

$$P_{mA} = \frac{\sum_{i=1}^N P_{Ai}}{N} \quad (9)$$

式中:  $P$ 、 $R$  分别指精确率和召回率;  $N$  表示总的类别数;  $P_T$  表示模型预测为正样本且实际也为正样本的数量;  $P_F$  表示模型预测为正样本但实际为负样本的数量;  $N_F$  表示模型预测为负样本但实际为正样本的数量。

### 3.3 改进的 4 种模型

本文共设计了 4 种模型与原 YOLOv5s 模型进

行对比, 分别为:

1) YOLOv5s-G: 使用 GhostBottleneck 模块替换 Backbone 中的 CBS 和 C3 模块。

2) YOLOv5s-S: 在 Backbone 中增加 3 个注意力机制 SE 模块。

3) YOLOv5s-D: 在 head 部分增加 1 个尺度为  $10 \times 10$  的检测层, 同时在 neck 部分增加特征融合结构。

4) YOLOv5s-GSD: 在原 YOLOv5s 网络中同时使用 GhostBottleneck 替换 CBS 和 C3, 添加 3 个注意力机制 SE 模块, 在网络中增加 1 个  $10 \times 10$  的检测层和对应的特征融合结构, 使用 DW 卷积替换 neck 中的部分标准卷积。该模型即为本文所提出的改进 YOLOv5s 模型, 其网络结构如图 7 所示, 其中 GB-1 和 GB-2 分别表示步长为 1 和步长为 2 的 GhostBottleneck 模块。

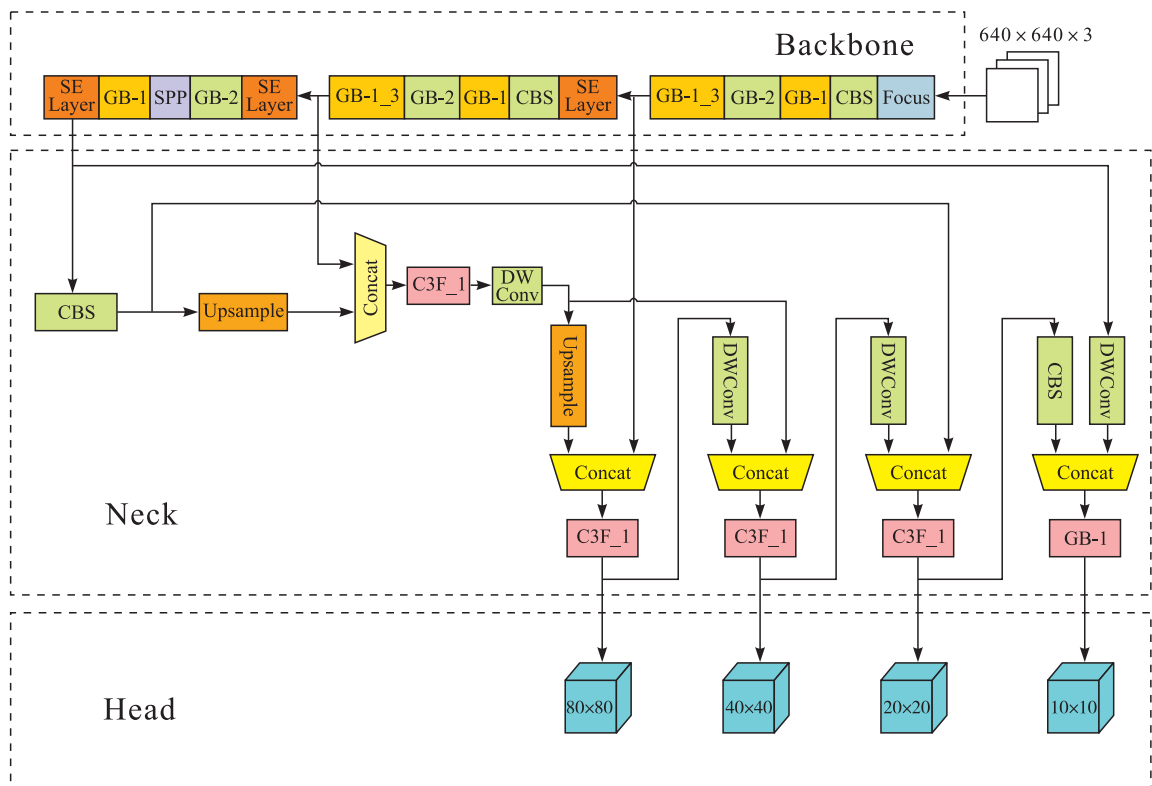


图 7 改进的 YOLOv5s-GSD 模型结构

### 3.4 实验对比

#### 3.4.1 模型参数与性能对比

为了验证单个结构改变对网络性能的影响, 本文进行了消融实验, 对 YOLOv5s、YOLOv5s-G、YOLOv5s-S、YOLOv5s-D、YOLOv5s-GSD 进行训练和测试, 得到 5 种模型的复杂度和性能对比如表 1 和表 2 所示。

表 1 各模型复杂度对比

| Model       | Parameters | GFLOPs | Weight/MB |
|-------------|------------|--------|-----------|
| YOLOv5s     | 7 077 027  | 16.4   | 14.4      |
| YOLOv5s-G   | 4 169 107  | 9.7    | 8.7       |
| YOLOv5s-S   | 7 120 035  | 16.5   | 14.5      |
| YOLOv5s-D   | 10 949 956 | 16.2   | 22.2      |
| YOLOv5s-GSD | 6 261 556  | 9.1    | 12.9      |

表 2 各模型性能对比

| Model       | P    | R    | mAP@0.5 | Speed |
|-------------|------|------|---------|-------|
|             | /%   | /%   | /%      | /ms   |
| YOLOv5s     | 69.7 | 75.1 | 73.5    | 2.6   |
| YOLOv5s-G   | 68.4 | 71.0 | 74.5    | 2.2   |
| YOLOv5s-S   | 76.5 | 74.5 | 75.1    | 2.8   |
| YOLOv5s-D   | 76.3 | 69.6 | 74.1    | 2.9   |
| YOLOv5s-GSD | 68.9 | 75.6 | 76.8    | 2.5   |

分析实验结果可知, YOLOv5s-G 模型相比原 YOLOv5s 的参数量减少了 41.1%, 运算量减少了 40.9%, 模型大小减少了 39.9%, 平均推理时间减少了 0.4 ms。这得益于 GhostNet 使用了线性运算替代部分卷积过程, 从而在得到近似于原始特征图的同时减少计算量和特征图冗余<sup>[19]</sup>。YOLOv5s-S 相比于原 YOLOv5s 模型, 由于 SE 注意力机制的引入, 虽然增加了模型复杂度, 使得模型的参数量、运算量和模型体积略有增大, 导致平均推理时间有 0.2 ms 的增加, 但 mAP 值提升了 1.6%, 这证明了注意力机制对重要特征的强化作用。YOLOv5s-D 由于增加一个检测层和部分卷积结构, 相比原 YOLOv5s 网络参数量增加较多, 平均推理时间增加 0.3 ms, 但 mAP 值提升了 0.6%。本文提出的 YOLOv5s-GSD 模型, 相比原 YOLOv5s 模型, 参数量减少了 11.5%, 运算量减少了 44.5%, 模型体积减少了 10.4%, mAP 值提升了 3.3%, 平均推理时间减少了 0.1 ms。

YOLOv5s-GSD 和 YOLOv5s 检测各类缺陷的平均精度 AP 值如图 8 所示。从图中可以看出, 改进后的 YOLOv5s-GSD 模型相比原 YOLOv5s 模型, 除 scratches 类缺陷的 AP 值降低 2.2% 外, 其他类别均有不同程度的提高, 其中 crazing、inclusion、patches、pitted surface 和 rolled-in scale 的平均精度分别提高 5.9%、2.0%、4.6%、2.6% 和 7.2%, 对于原模型检测效果较差的 crazing 和 rolled-in scale 这两类缺陷的检测精度有了明显提升。

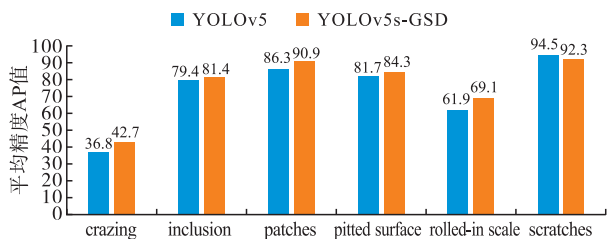


图 8 YOLOv5s-GSD 和 YOLOv5s

各类别平均精度 AP 值对比

综合以上分析, 可以得出: Ghost 模块和 DW 卷积对于模型轻量化和减少运算量有显著作用; SE 注

意力机制对于精确率有明显的提升, 很好地改善了误检的情况; 在 head 部分新增的检测层对于较大目标的检测精度有所提高。本文提出的 YOLOv5s-GSD 模型相比原 YOLOv5s 模型在参数量、运算量、模型体积、mAP 值以及推理速度等方面均有明显优势。

### 3.4.2 训练过程损失

YOLOv5 网络的损失函数共包含 3 个部分:  $L_{\text{box}}$  为定位损失函数, 用于衡量预测框对坐标定位错误的误差;  $L_{\text{obj}}$  为置信度损失函数, 反映了预测框的置信度误差;  $L_{\text{cls}}$  为分类损失函数, 反映了预测框对目标类别预测错误引起的误差。其中  $L_{\text{box}}$  采用 CIoU\_Loss<sup>[20]</sup> 函数,  $L_{\text{obj}}$  和  $L_{\text{cls}}$  采用交叉熵损失函数。本文提出的 YOLOv5s-GSD 模型在训练过程中的损失变化如图 9 所示。

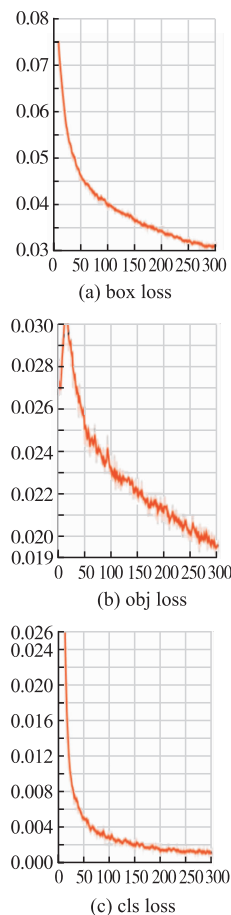


图 9 YOLOv5s-GSD 模型训练过程损失

训练设置迭代次数为 300 次, 从图中可以看出, 在 0 到 50 次迭代的过程中,  $L_{\text{box}}$  和  $L_{\text{cls}}$  急剧下降, 之后的训练过程中缓慢下降, 而  $L_{\text{obj}}$  在训练初期出现了局部的过拟合现象; 经过 300 次迭代后, 损失值逐渐趋于平稳。

### 3.5 检测效果对比

为更加直观地比较改进后的 YOLOv5s-GSD 模型与原 YOLOv5s 模型的实际检测效果, 本文分

别用 2 种模型对测试集的缺陷图片进行了检测,部分检测结果对比如图 10 所示。



图 10 改进前和改进后模型的检测效果对比

图 10(a)为原 YOLOv5s 模型的检测效果,(b)为改进的 YOLOv5s-GSD 模型检测效果。对比两者可以看出,原 YOLOv5s 模型对于一些目标存在漏检的情况,且置信度普遍较低。改进的 YOLOv5s-GSD 模型检测缺陷目标的置信度相比原 YOLOv5s 模型有了明显提高,尤其是 crazing 类缺陷,在使用原 YOLOv5s 模型检测时效果较差,且在较多漏检目标,改进后的模型明显提升了检测效

果,而且对于缺陷的定位也更加准确。

### 3.6 不同算法对比

为了进一步评估 YOLOv5s-GSD 模型的性能,本文选取了 3 种主流的目标检测算法 SSD<sup>[21]</sup>、YOLOv3 和 YOLOv4,在同样的数据集上与本文改进的 YOLOv5 模型进行实验对比,结果如图 11 和表 3 所示。

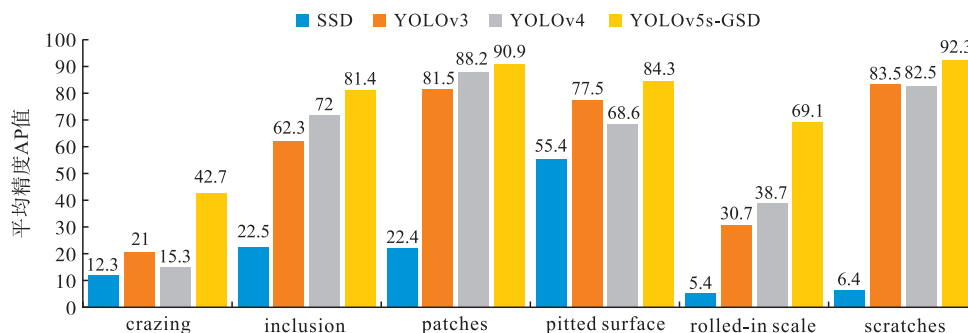


图 11 不同算法检测的各类别平均精度 AP 值对比

表 3 不同算法检测性能对比

| Model       | Parameters | mAP@0.5/% | Weight /MB | FPS  |
|-------------|------------|-----------|------------|------|
| SSD         | 24 013 232 | 20.7      | 91.6       | 18.1 |
| YOLOv3      | 61 539 889 | 59.4      | 234.8      | 34.9 |
| YOLOv4      | 63 953 841 | 60.9      | 244.0      | 25.7 |
| YOLOv5s-GSD | 6 261 556  | 76.8      | 12.9       | 43.3 |

根据图 11 和表 3 的实验结果可得:本文所改进的 YOLOv5s-GSD 模型相比几种主流目标检测算法,有效提升了对各种类别缺陷的检测精度,尤其是主流算法检测效果较差的 crazing 和 rolled-in scale 类缺陷,AP 值有大幅提高,mAP 值达 76.8%;参数量和模型体积大幅减小,更加便于移动端的部署;同

时具有较高的检测速度,达到了 43.3 FPS,实现了对缺陷目标的实时检测。因此可以得出,本文改进的 YOLOv5 模型在模型体积、检测精度和检测速度等方面均优于主流目标检测算法,能更好地完成钢材表面缺陷检测的任务。

### 3.7 实验结果总结与展望

通过上述的实验结果可以得出,本文改进的 YOLOv5s-GSD 模型对比原 YOLOv5s 模型和几种主流的目标检测算法,在模型复杂度和检测性能等方面都具有明显优势。相比原 YOLOv5s 模型,参数量减少了 11.5%,运算量减少了 44.5%,模型体积减少了 10.4%,实现了模型的轻量化;同时 mAP 值提升了 3.3%,几乎对各类缺陷的检测精度均有不同程度的提升。其中 crazing 类缺陷的检测精度

相对较低,有较大提升空间,本文推测可能是该类别缺陷的标注存在一定的误差。可使用标注软件对图片的标注信息进行调整,提高图片标注精度,从而进一步提升 mAP 值。

## 4 结论

本文提出了一种基于改进 YOLOv5 算法的钢材表面缺陷检测模型,利用 GhostBottleneck 结构和 DW 卷积实现了网络模型的轻量化;引入 SE 注意力机制,通过强化重要的特征来增强检测效果;改进了特征融合层结构,增加了一个尺度为  $10 \times 10$  的检测层,提升了对部分类别缺陷的检测精度。下一步工作将验证模型在不同数据集上的训练和检测效果,同时进一步优化网络结构,提高网络的检测速度和精度。研究如何将模型部署在移动端,在实际的缺陷检测应用中对模型加以完善和改进。

### 参考文献

- [1] 胡贵超. 基于计算机视觉的钢材表面缺陷检测系统的研究[D]. 哈尔滨:哈尔滨理工大学,2013:1-3.
- [2] 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.
- [3] 冯永毅. 基于人工免疫方法的带钢表面缺陷分割[D]. 成都:电子科技大学,2007:2-4.
- [4] KONG X, LI J. Image Registration-Based Bolt Loose-ning Detection of Steel Joints [J]. *Sensors*, 2018, 18 (4):1000.
- [5] 韩强,张喆,续欣莹,等. 特征融合和级联检测网络的 Faster R-CNN 钢材表面缺陷检测算法[J/OL]. 太原理工大学学报, 2021: 1-10. [2021-11-29]. <http://kns.cnki.net/kcms/detail/14.1220.N.20210616.1818.012.html>.
- [6] 叶欣. 基于深度学习的热轧带钢表面缺陷检测算法研究[D]. 武汉:武汉科技大学,2021:40-43.
- [7] 杨莉,张亚楠,王婷婷,等. 基于改进 Faster R-CNN 的钢材表面缺陷检测方法[J]. 吉林大学学报(信息科学版),2021,39(4):409-415.
- [8] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection [C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE,2016:779-788.
- [9] 于娟,罗舜. 基于 YOLOv5 的违章建筑检测方法[J]. 计算机工程与应用,2021,57(20):236-244.
- [10] HE K, ZHANG X, REN S, et al. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*,2014,37(9):1904-1920.
- [11] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature Pyramid Networks for Object Detection [C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu:IEEE,2017:936-944.
- [12] LIU S, QI L, QIN H F, et al. Path Aggregation Network for Instance Segmentation [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City:IEEE,2018:8759-8768.
- [13] HAN K, WANG Y, TIAN Q, et al. GhostNet: More Features From Cheap Operations [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). [S. l.]:IEEE,2020:1580-1589.
- [14] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-Excitation Networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42 (8): 2011-2023.
- [15] 李成. 基于改进 YOLOv5 的小目标检测算法研究[J]. 长江信息通信,2021,34(9):30-33.
- [16] CHOLLET F. Xception: Deep Learning with Depth-wise Separable Convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). [S. l.]:IEEE,2017:1800.
- [17] 刘亚楠. 基于机器视觉的钢材表面缺陷识别方法分析[J]. 计量与测试技术,2018,45(12):97-98.
- [18] 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):303-338.
- [19] 蒋明,陈雨,周青华,等. 适用于非合作目标捕获的轻量级位姿估计算法[J/OL]. 计算机工程,2021:1-12. [2021-11-29]. <https://doi.org/10.19678/j.issn.1000-3428.0060882>.
- [20] ZHENG Z, WANG P, LIU W, et al. Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*,2019,34:12993-13000.
- [21] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot Multibox Detector[C]//Proceedings of European Conference on Computer Vision. [S. l.]: Springer, 2016:21-37.

(编辑:姚树峰)