

## 改进GBS-YOLOv7t的钢材表面缺陷检测

梁礼明,龙鹏威,卢宝贺,李仁杰

#### 引用本文:

梁礼明,龙鹏威,卢宝贺,等. 改进GBS-YOLOv7t的钢材表面缺陷检测[J]. 光电工程, 2024, **51**(5): 240044. Liang L M, Long P W, Lu B H, et al. Improvement of GBS-YOLOv7t for steel surface defect detection[J]. *Opto-Electron Eng*, 2024, **51**(5): 240044.

https://doi.org/10.12086/oee.2024.240044

收稿日期: 2024-02-29; 修改日期: 2024-03-20; 录用日期: 2024-03-22

# 相关论文

#### 基于改进YOLOv5s网络的斜拉桥拉索表面缺陷检测

王鹏峰,李运堂,黄永勇,朱文凯,林婕,王斌锐 **光电工程 2024, 51(5): 240028** doi: 10.12086/oee.2024.240028

#### PAW-YOLOv7: 河道微小漂浮物检测算法

栾庆磊,常昕昱,吴叶,邓从龙,史艳琼,陈梓华 光电工程 2024, **51**(4): 240025 doi: 10.12086/oee.2024.240025

局部和全局特征融合的太阳能电池片表面缺陷检测 陶志勇,何燕,林森,易廷军,张尧晟 光电工程 2024, **51**(1): 230292 doi: 10.12086/oee.2024.230292

基于改进YOLOv5s的无人机图像实时目标检测

陈旭,彭冬亮,谷雨 光电工程 2022, **49**(3): 210372 doi: 10.12086/oee.2022.210372

更多相关论文见光电期刊集群网站



http://cn.oejournal.org/oee





DOI: 10.12086/oee.2024.240044

# 改进 GBS-YOLOv7t 的钢材 表面缺陷检测

梁礼明,龙鹏威\*,卢宝贺,李仁杰 江西理工大学电气工程与自动化学院,江西赣州 341000



摘要:针对钢材表面缺陷区域小目标居多,现有大部分方法无法均衡检测精度和速度的问题,提出一种基于YOLOv7tiny 的钢材表面缺陷检测算法 (GBS-YOLOv7t)。该方法一是设计 GAC-FPN 网络,采用渐进和跨层的方式充分融合 目标语义信息,以改善传统特征金字塔中存在限制信息流问题;二是嵌入双层路由注意力模块,使模型具备动态查询 和感知稀疏性能力,以提高对小目标的检测精度;三是引入 SloU 损失函数,提升模型训练和推理能力,增强网络鲁 棒性。最后在公共数据集 NEU-DET 进行实验验证, mAP 和精确度分别为 72.9% 和 69.9%,相较于 YOLOv7-tiny 原 模型分别提升 4.2% 和 8.5%; FPS 达到 104.1 帧,具有较强实时性;与其他检测算法相比,GBS-YOLOv7t 算法对 钢材表面区域小目标的检测更有效,实验表明改进后的算法能够更好地均衡检测精度和速度。

 关键词:
 缺陷检测; YOLOv7-tiny; GAC-FPN 网络; 双层路由注意力; SloU

 中图分类号:
 TP391.4

 文献标志码:
 A

梁礼明,龙鹏威,卢宝贺,等.改进 GBS-YOLOv7t 的钢材表面缺陷检测 [J]. 光电工程,2024,**51**(5): 240044 Liang L M, Long P W, Lu B H, et al. Improvement of GBS-YOLOv7t for steel surface defect detection[J]. *Opto-Electron Eng*, 2024, **51**(5): 240044

# Improvement of GBS-YOLOv7t for steel surface defect detection

# Liang Liming, Long Pengwei<sup>\*</sup>, Lu Baohe, Li Renjie

School of Electrical Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou, Jiangxi 341000, China

Abstract: Given that small targets are predominant in the steel surface defect areas, most existing methods cannot balance the trade-off between detection accuracy and speed. In this paper, we propose a steel surface defect detection algorithm based on YOLOv7-tiny (GBS-YOLOv7t). Firstly, we design the GAC-FPN network to fully integrate the target semantic information progressively and across layers, aiming to address the limited information flow issue in traditional feature pyramids. Secondly, we embed a bi-level routing attention (BRA) module to endow the model with dynamic query and sparse perception capabilities, thus enhancing the detection accuracy of small targets. Thirdly, we introduce the SIoU loss function to improve the training and inference capabilities of the model, and to enhance the network robustness. Experimental validation on the public dataset NEU-DET demonstrates an mAP of 72.9% and a precision of 69.9% for GBS-YOLOv7t, achieving improvements of 4.2% and 8.5%, respectively, over the original YOLOv7-tiny model. The FPS reaches 104.1 frames, indicating strong real-time performance. Compared to other detection algorithms, GBS-YOLOv7t is more effective in detecting small targets in

收稿日期: 2024-02-29; 修回日期: 2024-03-20; 录用日期: 2024-03-22

基金项目: 国家自然科学基金资助项目 (51365017, 61463018); 江西省自然科学基金面上项目 (20192BAB205084); 江西省教育厅科 学技术研究重点项目 (GJJ170491); 江西省研究生创新专项资金项目 (YC2022-S676)

\*通信作者:龙鹏威,2637018663@qq.com。 版权所有©2024 中国科学院光电技术研究所 steel surface areas, with experimental results showing that the improved algorithm better balances the detection accuracy and speed.

Keywords: defect detection; YOLOv7-tiny; GAC-FPN network; bi-level routing attention; SIoU

# 1 引 言

随着工业化的快速发展,钢材在日常生活中随处 可见,比如日常的交通工具、建筑设施以及制造业的 工具等都需要钢材作为原材料。然而,在生产制造过 程中因生产设备、人为因素和环境因素等影响,钢材 表面会出现氧化、开裂和划痕等缺陷<sup>[1]</sup>,严重影响美 观和使用寿命,若应用在工业领域,将带来潜在的风 险。因此,亟待研究一种精准且高效的钢材表面缺陷 检测算法。

传统缺陷检测法主要包括人工检查法和光电检测 法,人工检测方法耗费人力,而且效率低,会出现漏 检的情况。光电检测方法虽然速度快、灵敏度高,但 却难以检测轻微划痕和区分缺陷种类,故两者均难以 胜任钢材表面缺陷检测的任务。随着深度学习[2]的快 速发展,基于卷积神经网络的目标检测算法得到广泛 应用,并逐渐取代传统缺陷检测方法。赵朗月等<sup>[3]</sup>研 究了基于机器视觉的表面缺陷检测方法进展,阐述近 十年来使用图像处理、机器学习和深度学习实现缺陷 检测的原理,分析相应优缺点,并对未来工作进行展 望。基于深度学习的目标检测算法主要包括一阶段和 二阶段两类算法。二阶段检测代表算法诸如文献 [4-6] 提及的 R-CNN、Fast R-CNN 和 Faster R-CNN 算法 等,有较高检测精度,但速度慢,无法满足实时性需 求。针对其不足,一阶段检测算法相应问世,诸如文 献 [7] 提及的 SSD 算法和文献 [8] 涉及的 YOLO 算法 等,通过删除候选区域,直接对目标的位置和类别进 行检测,有效提高了实时性,但精度有所欠缺。由于 钢材表面缺陷微小且分辨率低,属于小目标范畴,为 提升 YOLO 系列算法对小目标的检测精度, Lin 等<sup>[9]</sup> 提出特征金字塔网络(FPN),利用自上而下的结构处 理上下文信息, 增加对小目标特征映射的分辨率, 以 获得更多有用信息,但单向传递结构存在限制信息流 问题。Liu 等<sup>[10]</sup>构建了一种路径聚合网络 (PANet), 在 FPN 的基础上增加一条自下而上的路径, 缩短信 息路径、增强小目标语义特征的流动,但自下而上的 路径反而会使浅层特征信息在传递和交互过程中丢失。 方钧婷等[11]构建一种注意力级联网络的金属表面缺陷 检测算法,将两个 IOU 阈值递增的检测头部网络级

联,提升检测精度和模型鲁棒性,但模块化的网络部 署部署能力较差。熊聪等<sup>[12]</sup>提出改进 YOLOX 的钢 材表面缺陷检测算法,通过引入 Swin Transformer 等 方法,有效提升检测精度,但检测速度大幅度下降。 赵春华等<sup>[13]</sup>提出基于 PC-YOLOv7 的钢材缺陷检测算 法,采用 PC-ELAN 结构替换 ELAN 结构等方法改善 模型,却以牺牲检测速度为代价换来模型参数量的优 化。因此, YOLO 系列在小目标上的应用还有待改进, Lian 等<sup>[14]</sup>设计一种改进 YOLOv4的缺陷检测算法,应 用知识蒸馏和预测尺度的方式提高检测精度;马燕婷 等<sup>[15]</sup> 提出改进 YOLOv5 网络的带钢表面缺陷检测方 法,采用 Transformer 与 BiFPN 相结合等方法增强图 像浅层特征信息与深层特征信息的融合,检测精度得 到一定提升。上述文献虽然在小目标领域带来一定的 进展,但无法同时均衡参数量,检测精度和实时性。 YOLOv7-tiny 在参数量和检测速度上具备一定优势, 但其融合网络和损失函数依旧沿用 PANet 网络和 CloU,导致对小目标语义信息融合不充分,模型训 练和推理能力差,故精度有待提升。

针对上述钢材表面缺陷检测和 YOLOv7-tiny 网络存在的问题,本文设计一种改进 GBS-YOLOv7t (GAC-FPN+BRA+SIoU, GBS)的钢材表面缺陷检测算法,主要研究工作如下:①设计 GAC-FPN 融合网络,改善传统金字塔网络存在的不足,提升网络对小目标特征的融合能力;②添加双层路由注意力(bi-level routing attention, BRA)模块,在参数量增加不明显的 情况下,使网络具备动态查询和感知能力,均衡检测 速度的同时提升对小目标的检测精度;③使用 SIoU 损失函数,提高模型训练推理和鲁棒性能力。

# 2 YOLOv7-tiny 模型

Wang 等<sup>[16]</sup>提出 YOLOv7-tiny 算法,该算法在参数量和检测速度上都超越以往的 YOLO 系列算法,该网络结构主要由输入端 (Input)、特征提取网络骨干 (Backbone)、特征融合网络 (Neck)和检测头 (Head) 四部分组成。

① 在 Input 中,对输入图像进行裁剪、缩放等操作,使图片像素满足主干提取网络的要求。采用 Mosaic 技术提升检测速度,减少内存消耗。 ② 在 Backbone 中,对输入信息进行提取,包含 卷积层、长程聚合网络层和池化层。

③ 在 Neck 中,由 SPP 层提供特征提取信息,再 通过特征融合金字塔 (PAFPN) 和聚合网络 (PANet) 对 提取信息进行融合。

④在 Head 中,使用标准卷积替代 RepConv 卷积 对输出通道数进行调整,采用 Idetect 作为检测头。

## 3 本文模型

#### 3.1 GBS-YOLOv7t 网络模型

本文以 YOLOv7-tiny 为基础模型,提出基于 GBS-YOLOv7t 的检测算法,以改善钢材表面小目标 缺陷居多和原模型融合网络存在限制信息流、语义信 息丢失、训练和推理能力差的问题,GBS-YOLOv7t 网络结构图如图 1 所示。该算法首先设计GSC-FPN 网络作为融合网络,改善原 PANet 网络在融合过程 中存在语义信息丢失的问题,以提高模型的特征融合 能力;其次在主干网络添加 BRA 注意力模块,加强 特征提取的同时使模型具备稀疏采样能力,提高对小 目标检测效果;最后通过 SIoU 损失函数,提升模型 鲁棒性以及训练和推理能力。

#### 3.2 鬼影渐进跨层融合网络

原模型在 Neck 部分沿用文献 [10] 提出的 PANet

作为融合网络,其在 FPN 的基础上增加一条自下而 上的路径, 使深层特征中包含更多浅层特征, 但自下 而上的路径会使浅层特征信息在传播和交互过程中丢 失。为提升对小目标的检测精度, Yang 等<sup>[17]</sup>提出渐 进特征金字塔网络 (AFPN),通过将两个不同分辨率 的浅层特征进行组合,再将深层特征纳入融合,该过 程极大缩小了非相邻层次之间的语义差距,减少信息 在传递过程中的丢失。但该网络由大量的标准卷积 (SC)组成,随着层数的加深和卷积的深入,冗余信息 会逐渐增多,导致网络参数量大且无法完全避免语义 信息的丢失。Li 等[18] 结合深度可分离 (DSC) 和 SC 的 特点提出一种轻量化的鬼影混洗卷积 (GSConv), 在 降低参数量和计算复杂度的同时维持与 SC 相同的输 出贡献。本文受文献 [17-18] 启发,设计鬼影渐进跨 层特征金字塔 (ghost asymptotic cross-layer feature pyramid network, GAC-FPN)网络, 其网络结构如 图 2 所示, GAC-FPN 网络首先采用 GSConv 替换 AFPN 网络中的标准卷积,降低模型复杂度;然后将 主干每层的特征提取信息与网络中的 ELAN 结构进行 跨层连接,充分减少信息在传递过程中的丢失。

GAC-FPN 网络通过在第一阶段将两个不同分辨 率的浅层特征进行组合,第二阶段将深层特征纳入融 合,最终融合主干的顶级特征。其中深层特征与来自 浅层特征的详细信息融合、浅层特征与来自深层特征



图 1 GBS-YOLOv7t 网络结构 Fig. 1 Network structure of GBS-YOLOv7t

#### 梁礼明, 等. 光电工程, 2024, 51(5): 240044

的语义信息融合。与此同时,为避免该网络在元素求和时因不同层次之间可能存在信息矛盾,本文通过引入自适应空间融合 (adaptively spatial fusion, ASF)<sup>[19]</sup>操作来过滤多级融合过程的特征,ASF 结构如图 2 中所示,其融合过程可表示为

$$y_{ij}^{l} = \alpha_{ij}^{l} \cdot x_{ij}^{1 \to l} + \beta_{ij}^{l} \cdot x_{ij}^{2 \to l} + \gamma_{ij}^{l} \cdot x_{ij}^{3 \to l}, \qquad (1)$$

其中, $y_{ij}^{l}$ 表示输出特征的第 (*i*,*j*)个向量在通道之间 映射 $y^{l}$ , $x_{ij}^{n \to l}$ 表示从第 *n* 层调整到第 *l* 层的特征映射 上 (*i*,*j*) 位置的特征向量, $\alpha_{ij}^{l}$ 、 $\beta_{ij}^{l}$ 、 $\gamma_{ij}^{l}$ 分别为网络自 适应学习到第 *l* 层的三个不同层次特征映射的空间重 要度权重, 且 $\alpha_{ij}^{l}$ + $\beta_{ij}^{l}$ + $\gamma_{ij}^{l}$ =1。三个参数可以被网络 自定义的学习,并定义为

$$\alpha_{ij}^{l} = \frac{e^{\lambda_{a_{ij}}}}{e^{\lambda_{a_{ij}}^{l}} + e^{\lambda_{p_{ij}}^{l}} + e^{\lambda_{\gamma_{ij}}^{l}}}.$$
 (2)

其中: $\lambda_{\alpha_{ij}}^{l}$ 、 $\lambda_{\beta_{ij}}^{l}$ 和 $\lambda_{\gamma_{ij}}^{l}$ 分别是 $\alpha_{ij}^{l}$ 、 $\beta_{ij}^{l}$ 、 $\gamma_{ij}^{l}$ 的控制参数。



Fig. 2 GAC-FPN network structure

GSConv 是一种轻量级卷积方法,其结构如图 3 所示。相比于 SC,GSConv 结合普通卷积高效的特征 提取融合能力和 DSC 轻量且速度快的特点,通过压

#### https://doi.org/10.12086/oee.2024.240044

缩内部通道来减少参数量和计算量。首先输入信息 经 SC 后得到信息 A,此时通道数压缩为原来的一半, 信息 A 再经过 DSC 操作充分渗透得到信息 B;然后 将信息 A 与 B 进行拼接融合;最后再经过 Shuffle 模 块实现信息交互并恢复通道数。该过程不仅保留 SC 的高效处理,同时还减少参数量和计算量且平衡检测 速度。



Fig. 3 GSConv network structure

#### 3.3 双层路由注意力

为捕获重要语义信息,注意力机制得到广泛关注, 传统注意力机制大部分通过增加头数来提升表现,但 对模型的计算和内存的需求也会增加。为提升对小目 标的检测效果且不过多增加模型复杂度,本文引入一 种具备动态稀疏的双层路由注意力机制<sup>[20]</sup>(bi-level routing attention, BRA),该注意力机制是基于稀疏采 样操作,一是可以减少网络模型计算量;二是保留细 粒度的空间细节信息,从而提升对小目标的检测效果。

双层路由注意力的核心思想是过滤掉特征图中大部分不相关的键值对,只保留小范围的路由区域,从而去除掉冗余信息。BRA由三部分组成:①区域划分和输入投影;②有向图区域到区域路由;③token-to-token注意力,其结构如图4所示。



图 4 BRA 网络结构 Fig. 4 BRA network structure

① 区域划分和输入投影:先将输入图片  $X \in \mathbb{R}^{H \times W \times C}$ 划分为 $S \times S$ 个不同区域,区域中包含  $HW/S^2$ 个特征向量。再通过reshape变换,把X变成  $X^r \in \mathbb{R}^{S^2 \times HW/S^2 \times C}$ 。最后经线性映射得到Q、K、V三个特征向量,其计算式为

$$\boldsymbol{Q} = \boldsymbol{X}^{\mathrm{r}} \boldsymbol{W}^{\mathrm{q}}, \quad \boldsymbol{K} = \boldsymbol{X}^{\mathrm{r}} \boldsymbol{W}^{\mathrm{k}}, \quad \boldsymbol{V} = \boldsymbol{X}^{\mathrm{r}} \boldsymbol{W}^{\mathrm{v}}, \quad (3)$$

其中:  $W^q$ ,  $W^k$ ,  $W^{v} \in \mathbb{R}^{C \times C}$ 分别是 query, key, value 的投影权重。

② 有向图区域到区域路由:为过滤掉不相关的 区域,减少 GPU 运算。分别在 Q 和 K 中应用每个区 域平均值推导出区域级Q<sup>r</sup>, K<sup>r</sup>∈ ℝ<sup>s<sup>2</sup>×C</sup>,再将Q<sup>r</sup>与K<sup>r</sup>的 转置进行矩阵乘法得到区域到区域相关性的邻接矩阵 A<sup>r</sup>∈ ℝ<sup>s<sup>2</sup>×s<sup>2</sup></sup>,其计算式为

$$\boldsymbol{A}^{\mathrm{r}} = \boldsymbol{Q}^{\mathrm{r}} (\boldsymbol{K}^{\mathrm{r}})^{\mathrm{T}} .$$
 (4)

通过邻接矩阵找到一个与已知特征图区域语义相关的另一个区域,在通过取 top-k 保存特征区域所在的地址,推导出路由索引矩阵**F**,其计算式为

$$I^{\rm r} = {\rm topkIndex}(A^{\rm r}) , \qquad (5)$$

其中,topkIndex 表示将邻接矩阵中当前每一行的特征区域和每一列的其他 k 个特征区域的列索引存在向量里。

③有向图区域到区域路由:通过矩阵ⅠT计算细粒 度的区域到区域注意力,对于 k 个路由区域并集中的 所有键值对,这些区域分散在整个特征图中,而 GPU 依赖内存合并操作。因此,要先将键值对的 tensor 值进行聚集 (gather)得到 K<sup>9</sup>和V<sup>9</sup> ∈ ℝ<sup>s<sup>2</sup>×HW/S<sup>2</sup>×C</sup>, 然后将注意力应用在聚集后的键值对上得到 O。其计 算公式分别为

$$\boldsymbol{K}^{9} = \text{gather}(\boldsymbol{K}, \boldsymbol{I}^{\mathrm{r}}) , \qquad (6)$$

$$V^9 = \text{gather}(V, I^r) , \qquad (7)$$

$$\boldsymbol{O} = \text{Attention}\left(\boldsymbol{Q}, \boldsymbol{K}^9, \boldsymbol{V}^9\right) + \text{LCE}\left(\boldsymbol{V}\right) \,, \tag{8}$$

其中,LCE(•)是基于深度卷积的上下文增强函数。

#### 3.4 损失函数

定位损失函数在目标检测中至关重要,合理的定 义将给网络性能带来显著的提升,原检测网络沿用 CloU 作为损失函数,虽然考虑了重叠面积和归一化 中心点等因素,但存在自由度高、训练收敛速度慢且 鲁棒性不强等问题。为改善网络性能,本文引入 SloU 损失函数<sup>[21]</sup>,SloU 作为一种新的边界损失函数, 增加角度成本,重新对距离进行定义,减少了惩罚项 为0出现的概率,使得收敛更加平稳,进一步提升回 归进度,从而降低预测误差。SIoU由角度成本、距 离成本和形状成本组成,其计算所涉及的参数如图 5 所示。



图 5 SIOU 硕大函数计并方式 Fig. 5 Calculation method of SIoU loss function

图中  $C_x$  和  $C_y$  表示坐标轴的宽度和高度,  $C_w$  和  $C_h$  为真实框和预测框中心点的宽度差和高度差,  $\sigma$  真 实框和预测框中心点的距离,  $b_{gt}$  为真实框中心坐标 的高度。

#### 3.4.1 角度成本

判断角度与 45°的关系,当小于 45°时用φ作为最 小化,反之用φ作为最小化。增加角度感知能够减少 与距离之间相关的参数。其计算公式为

$$\Lambda = 1 - 2 \times \sin^2 \left( \arcsin(x) - \frac{\pi}{4} \right), \tag{9}$$

其中:

$$x = \frac{C_{\rm h}}{\sigma} = \sin\varphi,$$
  
$$\sigma = \sqrt{\left(b_{C_{\rm y,gt}} - b_{C_{\rm x}}\right)^2 + \left(b_{C_{\rm y,gt}} - b_{C_{\rm y}}\right)^2},$$
  
$$C_{\rm h} = \max\left(b_{C_{\rm y,gt}}, b_{C_{\rm y}}\right) - \min\left(b_{C_{\rm y,gt}}, b_{C_{\rm y}}\right),$$

#### 3.4.2 距离成本

距离成本表示真实框与预测框的中心点距离,结 合对角度成本的定义,SIoU对距离成本重新进行定 义,其计算公式为

$$\Delta = \sum_{t=x,y} (1 - e^{-\gamma \rho t}), \qquad (10)$$

其中:

$$\gamma = 2 - \Lambda ,$$
  

$$\rho_x = \left(\frac{b_{C_x,\text{gt}} - b_{C_x}}{C_w}\right)^2, \quad \rho_y = \left(\frac{b_{C_y,\text{gt}} - b_{C_y}}{C_h}\right)^2.$$

结合式 (9) 可知, 当 $\varphi \rightarrow 0$ 时, 对距离成本的贡献 最低,随着 $\varphi$ 在(0, $\pi$ /4)区间内逐渐增大, 对距离成本 和 $\Delta$ 值的贡献也逐渐增加。当角度继续增加,  $\gamma$ 将随 着距离值的变化赋予时间优先权。

#### 3.4.3 形状成本

形状成本的定义式为

$$\Omega = \sum_{t=w,h} (1 - e^{-w_t})^{\theta}, \qquad (11)$$

其中:  $\omega_{w} = \frac{\left|w - w_{gt}\right|}{\max(w, w_{gt})}, \quad \omega_{h} = \frac{\left|h - h_{gt}\right|}{\max(h, h_{gt})}$ 

式(11)中θ值控制形状成本的关注程度,θ值可根据不同数据集自定义。为防止过于关注形状损失而忽略对预测框的移动,本文设置为4。

$$L_{\text{SIOU}} = 1 - I_{\text{IOU}} + \frac{\Delta + \Omega}{2}, \qquad (12)$$

$$| \pm \psi, \ I_{\text{IOU}} = \frac{|b \cap b_{\text{gt}}|}{|b \cup b_{\text{gt}}|} \circ$$

# 4 实验结果及分析

#### 4.1 实验环境与参数设置

实验环境为 Windows11 操作系统,深度学习框架为 pytorch1.11.0, CPU为 Intel Core i7-12700H, GPU为 RTX4060,显卡为 8 GB。在训练阶段,采用 Adam 作为优化器,设置输入图像尺寸为 640×640、 初始学习率为 0.01、动量为 0.973、权重衰减因子为 0.0005。其训练 batchsize 为 16、epoch 为 300 轮。

#### 4.2 数据集

本文采用东北大学 NEU-DET 钢材缺陷数据集<sup>[1]</sup> 共 1800 张图片,训练集、验证集和测试集按 6:2:2 随 机划分,其中训练集 1040 张、验证集 和测试集 360 张。共包含裂缝 (Crazing)、夹杂物 (Inclusion)、 斑块 (Patches)、点蚀表面 (Pitted-surface)、氧化皮 (Rolled-in-scale) 和划痕 (Scratches) 六种缺陷,样本如 图 6 所示。



Fig. 6 Images of various defects on steel surface

#### 4.3 评价指标

本实验采用平均精度 (average precision, AP)、平 均精度均值 (mean AP, mAP)、每秒检测帧数 (frames per second, FPS)、模型参数大小 (Params) 和计算量 (floating-point operations, FLOPs) 作为评价指标, 衡量 上述指标的计算式分别为

$$AP = \int_0^1 P dR , \qquad (13)$$

$$mAP = \frac{\sum_{i=1}^{N} AP_i}{N}, \qquad (14)$$

$$Precision = \frac{TP}{TP + FP}, \qquad (15)$$

$$Recall = \frac{TP}{TP + FN} , \qquad (16)$$

$$FPS = \frac{n_{\rm F}}{T_{\rm Elopsed}} , \qquad (17)$$

其中: AP 表示 P-R 曲线下方与坐标轴围成的面积; mAP 是对所有类别的 AP 求平均值所得; N 表示检测 的类别数,本实验 N=6; TP、FP 和 FN 分别表示正确 检测到目标框、误检框和漏检框的数量; n<sub>F</sub> 是需要检 测图片的总数, T<sub>elap</sub>则表示检测所需花费的总时间。

#### 4.4 改进 GAC-FPN 实验

本文为验证所提出 GAC-FPN 融合网络的有效性, 在 NEU-DET 数据集上,通过逐一消融的方式进行验 证。其中,A 表示添加跨层连接,B 表示用 GSConv 替换 AFPN 中的标准卷积。"√"表示加入该模块,结 果如表1所示,其中加粗部分表示最优值。

从表 1 可以看出,在 AFPN 网络的基础上加入跨 层连接时,mAP 值提高 1.5%,参数量和计算量略微 增加,检测速度减少 12.04 帧,表明跨层连接能有效 减少信息在传递过程中的丢失,提高检测精度。在此 基础上用 GSConv 替换融合网络中的标准卷积 mAP 值提高 1.2%、参数量减少 0.88 M、计算量减少 1.4 G、 检测速度提升 16.77 帧,表明 GSConv 的输出贡献明 显优于标准卷积且具备轻量化,能更好地平衡检测速 度和精度。

表	1	GAS-FPN 消融实验结果
Table 1	Res	ults of GAS-FPN ablation experiment

A	В	mAP/%	Params/M	FLOPs/G	FPS
		69.2	7.11	14.1	106.38
$\checkmark$		70.7	7.44	14.6	94.34
$\checkmark$	$\checkmark$	71.9	6.56	13.2	111.11

为验证设计的 GAC-FPN 网络具备优越性,在 NEU-DET 数据集和相同实验条件下,与 PANet 和 AFPN 网络进行对比分析,结果如表 2 所示,其中加 粗部分表示最优值。

表 2 GAS-FPN 对比实验 Table 2 Comparison experiment of GAS-FPN

Model	mAP/%	Params/M	FPS
PANet	68.7	6.02	108.12
AFPN	69.2	7.11	106.38
GAC-FPN	71.9	6.56	111.11

从表 2 可知, GAC-FPN 网络除参数量相较于 PANet 增加 0.54 M 外,其余指标均具有明显优势, 其中 mAP 为最佳达到 71.9%,比原模型提升 3.2%。 GAC-FPN 网络相比于 AFPN 网络, mAP、参数量和 FPS 均得到了优化,能更好地平衡检测精度、速度和 轻量化。实验结果表明,本文设计的 GAC-FPN 融合 网络具有优越性。

GAC-FPN 网络、PANet 和 AFPN 网络对检测各 类缺陷的平均精度值如图 7 所示,可以看出本文提出 的 GAC-FPN 网络相比于 PANet 和 AFPN 网络,对各 类缺陷的平均检测精度值均有所提升,表明 GAC-FPN 网络能有效提高对小目标的检测精度。



缺陷 AP 值对比

Fig. 7 Comparison of AP values of GAC-FPN, PANet and AFPN networks for detecting various types of defects

#### 4.5 BRA 模块对比实验

本文为确定 BRA 模块加在 YOLOv7-tiny 网络主 干的所处位置对各指标的综合提升效果最好,针对以 下 3 种情况进行实验,其中,Stage3、Stage4 和 Stage5 分别表示只在主干的第三、第四、第五个特征 提取后加入 BRA 模块,实验结果如表 3 所示,其中 加粗部分表示最优值。

表3 BRA位置实验结果

|--|

Location	mAP/%	Params/M	FPS
Baseline	68.7	6.02	108.69
Stage3	69.0	6.09	79.37
Stage4	69.9	6.20	104.17
Stage5	69.3	7.08	111.11

从表 3 可知,对三个指标综合考虑,当 BRA 模型处于 Stage3 时带来的参数量最少,但速度减少 29.32 帧,mAP 值提升较少;处于 Stage4 时mAP 值为 69.9%,相比于基础模型提升 1.2% 为最优,在牺牲较少 Params 的情况下,保持 FPS 基本不变;虽然 Stage5 的检测速度更优,但检测精度和参数量均不如 Stage4。综上分析,将 BRA 模块加在主干网络第四个特征提取层后面效果最佳。

同时,为验证 BRA 模块的优越性,在 NEU-DET 数据集和同一实验条件下,将 BRA 与 SE 注意 力、坐标注意力 (coordinate attention, CA) 和三重注意 力 (triplet attention, TA) 进行对比,实验结果如表 4 所 示,其中加粗部分为最优值。

从表 4 中实验结果可知,对三个指标综合考虑, SE 的 mAP 值达到 69.3%,相比于原模型提升 0.6%, 但以牺牲巨大参数量和计算量为代价。TA 和 CA 所 带来的参数量较少,但 TA 相比于原模型 mAP 值只 提升 0.2%,检测性能不够优越,CA 相比于原模型降 低 0.4%。反观 BRA 模块,mAP 为 69.9%,达到最优 值,同时兼顾了参数量和计算量。原因可能是 BRA 基于稀疏采样操作,在减少网络模型计算量的同时保 留了细粒度的空间细节信息,从而提升检测能力。

表 4 BRA 对比实验 Table 4 BRA comparison experiments

			-
Model	mAP/%	Params/M	FLOPs/G
Baseline	68.7	6.02	13.1
SE	69.3	11.57	30.8
TA	68.9	6.02	13.2
CA	68.3	6.03	13.5
BRA (Ours)	69.9	6.20	13.2

#### 4.6 消融实验

本文为验证所提模块的有效性,在 NEU-DET 数

据集下,通过逐一消融方式进行验证,实验得到 mAP、Params、P和R。其中,M1表示采用GAC-FPN网络作为融合网络,M2表示用SIoU 替换CIoU, M3表示添加BRA模块,"√"表示加入该模块,结果 如表5所示,加粗部分表示最优值。

从表 5 可知,首先用改进后的特征金字塔网络替换 PANet 网络,mAP 值提高 3.2%,精确度增加 2.4%,表明 GAC-FPN 网络能有效提高模型特征融合能力和检测能力。然后把 CloU 替换成 SloU 损失函数,mAP 值提高 3.8%,精确度提高 3.9%,说明改进后的损失函数能有效够提高模型鲁棒性。最后添加具有动态查询感知稀疏性的 BRA 模块,即本文所提GBS-YOLOv7t算法,mAP 值提高 4.2%,精确度提高 8.5%,说明具备稀疏采样的 BRA 模块能提升对小目标的检测效果。虽然参数量与原模型相比增加 0.81 M、召回率略微降低 2.2%,但对于硬件和部署的影响较小。消融实验结果表明,本文对基础模型做出的改进均能提高其检测精度,说明 GBS-YOLOv7t算法的合理性和有效性。

为直观展示本文算法和原始模型对各类缺陷检测 的平均精度 AP 值,采用柱状图进行展示,两种算法 的检测结果如图 8 所示。从图 8 中可以看出,本文算 法相比于原模型除斑块类缺陷平均精度值降低 0.7%, 其他类别缺陷平均精度值均有所提升,其中裂缝、氧

			08.7	6.02	01.4	12.1
$\checkmark$			71.9	6.56	63.8	70.2
$\checkmark$	$\checkmark$		72.5	6.56	65.3	73.5
$\checkmark$	$\checkmark$	$\checkmark$	72.9	6.83	69.9	70.5
	100 90 80 70 60 50 40 30 30	* YOLO • Ours 32.7 25.3	v7-tiny <sup>93.1</sup> 92 <sub>18.2</sub> 69.6	.4 95.0 <sup>96.5</sup> 49.	88. 81.4	6

图 8 改进算法检测结果对比

;	表 5	消融实验结果
Table 5	Resu	Its of ablation experiment

Params/M

P/%

**R**/%

mAP/%

M2

M1

M3

化皮和划痕的平均精度提升较大,分别为7.4%、 8.4%和7.2%。

#### 4.7 对比实验分析

为客观评价本文算法的优势,在 NEU-DET 数据集和相同实验条件下,与 Faster R-CNN、SSD、FCOS、RetinaNet 和 YOLO 系列算法以及文献 [13]、[22] 和 [23] 进行对比复现,结果如表 6 所示,其中加粗表示最优值。

从表 6 可知, Faster R-CNN 算法的 mAP为 65.7% 与 SSD 相比提高 4.7%, 但参数量和检测速度 都不如 SSD。YOLOv3 在 SSD 的基础上 mAP 提高 6%, 但参数量大、检测速度慢,同样不利于部署。 YOLOv4 的 mAP 最低为 51%, 检测效果最差。 YOLOv5s 算法的 mAP 为 70.1%、参数量为 7.07 M、FPS 为 102 帧,明显优于前三种检测算法,但相比本 文算法各项指标均不够理想。YOLOX-s 算法的 mAP 在 YOLOv5s 算法的基础上有所提升,但 FPS 下降 56 帧。YOLOv7 算法与 Faster R-CNN、YOLOv3 和 YOLOv4 存在同样的问题,对硬件设备要求高,且不利于实时检测。YOLOv7-tiny 算法的参数量较低,检测速度也较快,但 mAP 值较低,检测精度较差。 FCOS 和 RetinaNet 的检测精度和实时性较低且参数 量也不具备优势。文献 [13] 算法在参数量上具备优势。

表6 对比实验结果

Table 6	Comparison of the experimental results		
Model	mAP/%	Params/M	FPS
Faster R-CNN	65.7	72.0	17.8
SSD	61.0	24.4	41.0
YOLOv3	67.0	61.5	31.5
YOLOv4	51.0	52.5	45.0
YOLOv5s	70.1	7.07	102.0
YOLOX-s	71.8	8.0	46.0
YOLOv7	70.0	37.2	36.1
YOLOv7-tiny	68.7	6.02	108.1
FCOS	68.8	43.2	12.0
RetinaNet	69.5	18.3	15.1
PC-YOLOv7 <sup>[13]</sup>	71.2	5.97	61.0
文献[22]	73.0	9.54	_
文献[23]	74.1	23.9	75
GBS-YOLOv7t	72.9	6.83	104.1

				-	
Fig. 8	Comparison of	detection r	esults of th	e improved	algorithm

#### 梁礼明,等.光电工程,2024,51(5):240044

但无法同时均衡检测速度,可能是因为 PC-YOLOv7 算法采用模型压缩等技术。尽管参数量减少,但对于 大规模数据的处理和复杂情况的识别需要更多的计算 资源和时间,从而影响检测速度。文献 [22] 和 [23] 算法虽然有较高的检测精度,但以牺牲参数量和检测 速度为代价。反观本文算法,参数量仅为 3.83 M, mAP 和 FPS 值达到较优,分别为 72.9% 和 104.1 帧。 在 mAP、参数量和 FPS 上相比于其他对比算法取得 较好的平衡,表明 GBS-YOLOv7t 算法相比于其他对 比算法具备一定优越性,能更好地平衡检测精度和 速度。

#### 4.8 可视化对比

为直观展示本文算法的检测效果,使用 GBS-YOLOv7t 模型和其他检测算法在 NEU-DET 钢材表面 缺陷图片上进行检测,检测结果如图 9 所示。

对于不同的缺陷类别分别用不同颜色的锚定框来 表示。从图 9 中可知, YOLOv4 算法的检测效果最差, 对于裂缝、氧化皮和划痕均未识别,其余算法也存在 不同程度漏检情况。而本文算法能检测出更多有效缺 陷区域,结合图 8 中本文算法对各类缺陷的 mAP 结 果,本文算法对裂缝、氧化皮和划痕三类缺陷的检测 指标提升较大,从图 9 中可视化结果图也可直观看出。 特别是对于难以检测的裂缝缺陷,其他算法漏检率较 高,而本文算法能检测出多个标识框,说明改进后的 算法有更强的浅层特征提取与融合能力,一定程度上 能够避免漏检情况,对钢材表面小目标缺陷的检测效 果更佳。

# 5 结 论

针对钢材表面缺陷小目标居多,现有大部分方法 无法均衡检测精度和速度等问题,本文提出 GBS-YOLOv7t 的检测算法。

 1)设计鬼影渐进跨层征融合网络,改善传统特征 金字塔对语义信息融合不充分和限制信息流问题。

2)利用具备动态查询和感知稀疏性的 BRA 注意 力机制,使网络保留更多细粒度语义信息,以提高对 小目标的检测精度。

3) 联合 SIoU 损失函数改善模型训练、推理和鲁 棒性能力,更好地平衡检测精度和速度。

在 NEU-DET 数据集上进行消融实验和对比实验, 在参数量不明显增加的情况下,检测精度达到 72.9%, 速度达到 104.1 帧,表明 GBS-YOLOv7t 算法在钢材



图 9 本文算法与其他算法检测效果对比 Fig. 9 Comparison of detection effect between the proposed algorithm and other algorithms

#### 梁礼明,等.光电工程,2024,51(5):240044

表面缺陷检测的有效性和优越性,对工业钢材表面缺陷检测有一定的应用价值。未来的研究针对部分缺陷 在特征图中占比较大的问题,可通过添加 P6 检测层 提升钢材表面部分大目标的检测精度;为进一步优化 模型参数量,可以通过替换轻量级主干、模型剪枝等 操作;为提升整体检测精度,可对数据集做预处理。

#### 利益冲突:所有作者声明无利益冲突

# 参考文献

- [1] Chen H Y, Zhao P, Yan H W. Crack detection based on multi-scale Faster RCNN with attention[J]. *Opto-Electron Eng*, 2021, 48(1): 200112.
  陈海永, 赵鹏, 闫皓炜. 融合注意力的多尺度Faster RCNN的裂纹 检测[J]. 光电工程, 2021, 48(1): 200112.
- [2] Liang L M, Lu B H, Long P W, et al. Adaptive feature fusion cascade transformer retinal vessel segmentation algorithm[J]. *Opto-Electron Eng*, 2023, **50**(10): 230161.
   梁礼明, 卢宝贺, 龙鹏威, 等. 自适应特征融合级联Transformer视 网膜血管分割算法[J]. 光电工程, 2023, **50**(10): 230161.
- Zhao L Y, Wu Y Q. Research progress of surface defect detection methods based on machine vision[J]. *Chin J Sci Instrum*, 2023, **43**(1): 198-219.
   赵朗月, 吴一全. 基于机器视觉的表面缺陷检测方法研究进展[J]. 仪器仪表学报, 2023, **43**(1): 198-219.
- [4] Liu Z, Yeoh J K W, Gu X Y, et al. Automatic pixel-level detection of vertical cracks in asphalt pavement based on GPR investigation and improved mask R-CNN[J]. *Autom Constr*, 2023, **146**: 104689.
- [5] Zhang W M, Zhu Q K, Li Y Q, et al. MAM Faster R-CNN: improved Faster R-CNN based on malformed attention module for object detection on X-ray security inspection[J]. *Digital Signal Process*, 2023, **139**: 104072.
- [6] Hu B, Wang J H. Detection of PCB surface defects with improved faster-RCNN and feature pyramid network[J]. *IEEE* Access, 2020, 8: 108335–108345.
- [7] Qian H M, Wang H L, Feng S, et al. FESSD: SSD target detection based on feature fusion and feature enhancement[J]. *J Real-Time Image Process*, 2023, 20(1): 2.
- [8] Hussain M. YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection[J]. *Machines*, 2023, **11**(7): 677.
- [9] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, 2017: 936–944. https://doi.org/10.1109/CVPR.2017. 106.
- [10] Liu S, Qi L, Qin H F, et al. Path aggregation network for instance segmentation[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, 2018: 8759–8768. https://doi.org/10.1109/CVPR. 2018.00913.
- [11] Fang J T, Tan X Y. Defect detection of metal surface based on

attention cascade R-CNN[J]. *J Front Comput Sci Technol*, 2021, **15**(7): 1245–1254.

方钧婷, 谭晓阳. 注意力级联网络的金属表面缺陷检测算法[J]. 计 算机科学与探索, 2021, **15**(7): 1245-1254.

- [12] Xiong C, Yu A N, Gao X H, et al. Steel surface defect detection algorithm based on improved YOLOX[J]. *Electron Meas Technol*, 2023, 46(9): 151-157. 熊聪,于安宁,高兴华,等. 基于改进YOLOX的钢材表面缺陷检测 算法[J]. 电子测量技术, 2023, 46(9): 151-157.
- [13] Zhao C H, Luo S, Tan J L, et al. Detection of steel surface defects based on PC-YOLOv7 algorithm[J]. Foreign Electron Meas Technol, 2023, 42(9): 137-145. 赵春华, 罗顾, 谭金铃, 等. 基于PC-YOLOv7算法钢材表面缺陷检 测[J]. 国外电子测量技术, 2023, 42(9): 137-145.
- [14] Lian J W, He J H, Niu Y, et al. Fast and accurate detection of surface defect based on improved YOLOv4[J]. Assem Autom, 2022, 42(1): 134–146.
- [15] Ma Y T, Zhao H D, Yan C, et al. Strip steel surface defect detection method by improved YOLOv5 network[J]. *J Electron Meas Instrum*, 2022, **36**(8): 150–157.
  马燕婷,赵红东,阎超,等. 改进YOLOv5网络的带钢表面缺陷检测方法[J]. 电子测量与仪器学报, 2022, **36**(8): 150–157.
- [16] Wang C Y, Bochkovskiy A, Liao H Y M. YOLOv7: Trainable bagof-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, 2023: 7464–7475. https://doi.org/10.1109/CVPR52729.2023.00721.
- [17] Yang G Y, Lei J, Zhu Z K, et al. AFPN: Asymptotic feature pyramid network for object detection[C]//2023 IEEE International Conference on Systems, Man, and Cybernetics, Honolulu, 2023: 2184–2189. https://doi.org/10.1109/ SMC53992.2023.10394415.
- [18] Li H L, Li J, Wei H B, et al. Slim-neck by GSConv: A better design paradigm of detector architectures for autonomous vehicles[Z]. arXiv: 2206.02424, 2022. https://doi.org/10.48550/ arXiv.2206.02424.
- [19] Liu S T, Huang D, Wang Y H. Learning spatial fusion for singleshot object detection[Z]. arXiv: 1911.09516, 2019. https://doi. org/10.48550/arXiv.1911.09516.
- [20] Zhu L, Wang X J, Ke Z H, et al. BiFormer: Vision transformer with bi-level routing attention[C]//Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, 2023: 10323–10333. https://doi.org/ 10.1109/CVPR52729.2023.00995.
- [21] Gevorgyan Z. SloU loss: More powerful learning for bounding box regression[Z]. arXiv: 2205.12740, 2022. https://doi.org/10. 48550/arXiv.2205.12740.
- [22] Li Z G, Wei X M, Hassaballah M, et al. A deep learning model for steel surface defect detection[J]. *Complex Intell Syst*, 2024, 10(1): 885–897.
- [23] Cao Y Q, Wu M L, Xu L. Steel surface defect detection based on improved YOLOv5 algorithm[J]. J Graphics, 2023, 44(2): 335–345.

曹义亲, 伍铭林, 徐露. 基于改进YOLOv5算法的钢材表面缺陷检测[J]. 图学学报, 2023, 44(2): 335-345.

### 梁礼明, 等. 光电工程, 2024, **51**(5): 240044

#### https://doi.org/10.12086/oee.2024.240044

#### 作者简介



梁礼明(1967-),男,教授,硕士生导师,主要 研究方向为机器学习、模式识别与图像处理等。 公开发表学术论文百余篇,其中被 SCI、EI、 ISTP 收录论文二十余篇。获得中国发明专利 六项(排名第一)、出版研究生教材一部。 E-mail: 9119890012@jxust.edu.cn



【通信作者】龙鹏威(1998-),男,硕士研究生, 主要研究方向为机器学习、模式识别与图像处 理等。

E-mail: 2637018663@qq.com



# Improvement of GBS-YOLOv7t for steel surface defect detection

# Liang Liming, Long Pengwei<sup>\*</sup>, Lu Baohe, Li Renjie



Network structure of GBS-YOLOv7t

**Overview:** Aiming at the problem that most of the existing methods are unable to equalize the detection accuracy and the speed because of the predominance of small targets in the defective region of the steel surface, this paper proposes a steel surface defect detection algorithm based on YOLOv7-tiny (GBS-YOLOv7t). The method, firstly, takes into account that the feature fusion network of the original YOLOv7-tiny model adopts the traditional path aggregation network (PANet), which is designed with a bottom-up structure, but the bottom-up structure will have the problem of limiting the information flow. To address this problem, this paper compresses the model complexity and further preserves the semantic information of small targets by introducing the asymptotic feature pyramid (AFPN), and on its basis, by introducing the ghost shuffle mixing convolution (GSConv) and cross-layer connectivity. Based on the above improvements, the Ghost Asymptotic Cross-layer Fusion Network (GAC-FPN) is designed and replaces the original YOLOv7-tiny path aggregation network. The GAC-FPN network adopts an asymptotic and cross-layer approach to fully fuse the semantic information of the target features, which effectively improves the problem of restricting the flow of information in the top-down structure in the traditional feature pyramid. Secondly, to increase the model's accuracy in detecting the small targets. To improve the detection accuracy of the model for small targets, a Bi-Level Routing Attention module is embedded in the backbone network, and the optimal location of the module in the backbone network is verified through experiments, and the results show that the module makes the model possess the ability of dynamic querying and sparsity perception while taking into account the number of network parameters and the computational complexity, which effectively improves the detection accuracy of the model for small targets; thirdly, a SIOU loss function is introduced to replace the CIoU loss function of the original network, effectively improving the model training and reasoning ability, which improves the model training and inference ability, and enhances the network robustness. Finally, experimental validation is carried out on the publicly available Northeastern University Steel Surface Defect Dataset (NEU-DET), and the experimental results show that the mAP and accuracy of the GBS-YOLOv7t algorithm reach 72.9% and 69.9%, respectively, which are improved by 4.2% and 8.5%, respectively, compared with the original model of YOLOv7-tiny; the FPS reaches 104.1 frames, which is strong real-time performance. Compared with other classical detection models and current mainstream algorithms, the GBS-YOLOv7t algorithm has better performance and is more effective in detecting small targets on the surface area of steel, and the experiments show that the improved algorithm better balances lightweight, detection accuracy and speed.

Liang L M, Long P W, Lu B H, et al. Improvement of GBS-YOLOv7t for steel surface defect detection[J]. *Opto-Electron Eng*, 2024, **51**(5): 240044; DOI: 10.12086/oee.2024.240044

School of Electrical Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou, Jiangxi 341000, China

\* E-mail: 2637018663@qq.com

Foundation item: Project supported by National Natural Science Foundation of China (51365017, 61463018), Jiangxi Provincial Natural Science Foundation (20192BAB205084), Jiangxi Provincial Department of Education Scientific and Technological Research Key Project (GJJ170491), and Jiangxi Provincial Postgraduate Innovation Special Funds Project (YC2022-S676)