

改进 YOLOv5s 的钢板表面缺陷检测算法

周彦^{1*}, 孟江南^{1**}, 吴佳¹, 罗智², 王冬丽¹¹湘潭大学自动化与电子信息学院, 湖南 湘潭 411105;²湖南华菱湘潭钢铁有限公司, 湖南 湘潭 411105

摘要 针对传统方式检测钢板表面缺陷存在检测精度低、检测速度慢的问题,提出一种改进 YOLOv5s 算法。首先,使用基于交并比(IoU)度量距离的 K-means 算法对钢铁数据集进行重新聚类,获得多组锚框,通过遗传算法对其进行变异运算,得到与全体标注框更匹配的多组锚框;其次,在 Mosaic 数据增强上融合 MixUp,抑制过拟合,提升模型的泛化能力;然后,对网络结构进行改进,融入注意力模块,进一步提高了网络的特征提取能力;最后,针对难识别样本,在损失函数中融入 Focal loss,提高网络的收敛速度与检测精度。实验结果表明,改进后的 YOLOv5s 算法在测试集上的平均精度均值(mAP)可达 78.4%,比原始的 YOLOv5s 算法提高了 3.0 个百分点,速度上与原始 YOLOv5s 基本持平。所提算法在保持高检测速度的基础上,检测性能也优于 DDN、Faster R-CNN 和 YOLOv3。

关键词 YOLOv5s; 钢板表面缺陷检测; 注意力机制; Focal loss

中图分类号 TP183 文献标志码 A

DOI: 10.3788/LOP213302

Steel-Plate Surface-Defect Detection Algorithm Based on Improved YOLOv5s

Zhou Yan^{1*}, Meng Jiangnan^{1**}, Wu Jia¹, Luo Zhi², Wang Dongli¹¹School of Automation and Electronic Information, Xiangtan University, Xiangtan 411105, Hunan, China;²Hunan Valin Xiangtan Iron and Steel Co., Ltd., Xiangtan 411105, Hunan, China

Abstract To solve the problem of the low accuracy and slow speed of traditional methods for detecting surface defects in steel plates, we propose an improved YOLOv5s algorithm. First, the steel datasets were re-clustered using K-means algorithm based on the intersection-over-union (IoU) metric distance, to obtain multiple groups of anchor boxes; a genetic algorithm was used to perform mutation operations and obtain multiple groups of anchor boxes that match the entire ground truth box better. Second, MixUp was fused with the Mosaic data enhancement to avoid over-fitting and improve the generalizability of the model. Then, the network structure was improved, and an attention module was incorporated to improve feature extraction capability of the network further. Finally, Focal loss was incorporated into the loss function to improve the convergence speed and detection accuracy of the network for hard-to-identify samples. Our experimental results show that the mean average precision (mAP) of the improved YOLOv5s algorithm on a test set is 78.4%, which is 3.0 percentage points higher than that of the original algorithm, and the speed is same as the original YOLOv5s. The detection performance of the improved YOLOv5s algorithm is better than that of DDN, Faster R-CNN, and YOLOv3, and it maintains a high detection speed.

Key words YOLOv5s; steel plate surface defect detection; attention mechanism; Focal loss

1 引言

在钢板生产加工过程中,限于生产工艺技术水平与实际生产环境不确定性,钢板表面容易出现裂纹、划

痕、斑块等各种缺陷^[1],这些缺陷的存在会对钢板性能造成较大影响,因此缺陷检测是保证钢板质量的重要环节。目前,部分企业的这一重要环节主要由人工目检完成,检测准确率极大地依赖工人的工作状态和经

收稿日期: 2021-12-21; 修回日期: 2022-02-17; 录用日期: 2022-03-29; 网络首发日期: 2022-04-10

基金项目: 国家自然科学基金(61773330)、湖南省国家应用数学中心项目(2020YFA0712503)、湖南省教育厅科研项目(19C1740)、湖南省科技计划(2020GK2036)、上海市科委项目(19511120900)

通信作者: *yanzhou@xtu.edu.cn; **mjnshizhu@163.com

验水平,且长时间工作下,准确率下降明显,存在产品质量不可靠不稳定以及生产效率低问题。因此,研究一种智能且高效的钢板表面缺陷检测算法具有十分重要的意义。

对于缺陷检测,早期的计算机视觉技术主要通过人工提取特征,再进行缺陷分类。例如:Arjun等^[2]利用脉冲涡流检测钢铁表面缺陷;Joung等^[3]利用红外成像技术对管道缺陷进行检测;You等^[4]提出一种利用以太赫兹成像系统提取缺陷轮廓的方法;范宏等^[5]针对扣件视觉检测在背景复杂和光照多变的条件下检测精度低的问题,提出一种基于局部区域特征的缺陷扣件检测算法;Tsukada等^[6]利用漏磁进行缺陷检测;苗玲等^[7]提出了新型传感结构下的多物理电涡流热成像技术,该技术用于对钢轨滚动接触疲劳裂纹的动态检测;Jeon等^[8]利用光照法进行钢铁表面缺陷检测。这些传统缺陷检测方法存在精度低、步骤繁琐以及检测速度慢等问题,在实际生产中局限性太大,难以满足生产需求。

近年来,基于深度学习的缺陷检测算法具有稳定高效、鲁棒性强的特点,因此得到了迅速的发展和广泛的应用,在很多应用场景中都取得了瞩目的成绩,逐渐取代传统表面缺陷检测方法。Yi等^[9]提出了一种基于深度学习的端到端的钢铁表面缺陷检测方法。王宸等^[10]通过使用 Mish 激活函数和 CIoU 损失函数改进 YOLOv3 算法,实现对轮毂焊缝缺陷的智能化检测。胡勃等^[11]引入可形变卷积改进 Faster R-CNN,提出一种木板表面刮痕缺陷检测模型,该模型能够识别和定位不同纹理背景下的木板表面刮痕缺陷。赵振兵等^[12]针对输电线路中典型金具尺度变化大且部分金具为小尺度目标的问题,提出一种基于改进 Cascade R-CNN 的典型金具及其部分缺陷检测方法,大幅度提高了检测精度。汪权等^[13]提出基于 Gaussian YOLOv3 的绝缘子检测算法,该算法可以快速精确地识别航拍图像中的绝缘子缺陷。邢俊杰等^[14]使用 Cutmix 方法扩充小缺陷数量,并改进 YOLOv3 网络结构,提出 CSYOLOv3 方法,该方法能够有效识别工件表面小缺陷。李彬等^[15]为解决航空发动机部件表面缺陷检测精度低、检测速度慢的问题,使用改进的 YOLOv4 算法进行智能检测,检测精度提高了 2.17%。He 等^[16]基于 Faster R-CNN 提出了 defect detection network (DDN),将多个分层特征组合成一个特征,在 NEU-DET 数据集上采用 ResNet-50 作为 Backbone,实现了 82.3% 的平均精度均值 (mAP),但作为一个 two-stage 目标检测算法,DDN 存在检测速度不足的问题。

YOLOv5 作为一种 one-stage 目标检测算法,在保证较快检测速度的同时,相比 YOLOv3,精度上有较大提升,相比 YOLOv4,模型体积小,速度快,能很好地满足工业生产中精度与实时性的要求^[17-18]。综上所

述,本研究综合考虑精度与速度,在 YOLOv5s 的基础上,根据钢板表面缺陷的特点进行针对性改进。首先,改进 YOLOv5s 的自适应锚框算法,使用交并比 (IoU) 度量距离,并对数据集进行重新聚类,获得匹配度更高的锚框参数;然后,在 Mosaic 上融合 MixUp 方法进行数据增强,对原始图像模糊处理结果进行拼接;其次,在原网络的特征提取部分增加 convolutional block attention 模块 (CBAM),提高模型特征提取能力;最后,针对难识别样本,使用 Focal loss 改进损失函数,提升困难样本的权重。在 NEU-DET 数据集上进行实验验证,评估不同改进策略对模型检测性能的影响,同时与目前主流的检测算法进行对比,验证了改进的 YOLOv5s 算法的有效性。

2 YOLOv5s 算法

2.1 网络结构

YOLOv5s 是 YOLOv5 (第五版) 的轻量化版本,是 one-stage 检测算法,检测时把特征图分割成网格,每个网格中心通过预先设定长宽比的锚框 (anchor box) 生成预测框,对检测目标的标注框 (ground truth box) 进行预测。其网络结构主要分为 4 个部分,包括 Input 端、Backbone、Neck、Head 检测端,如图 1 所示,其中 k 为卷积核尺寸, s 为步长, n 表示模块的重复次数。Input 端主要完成数据预处理工作,包括 Mosaic 数据增强、缩放和色彩空间调整,其中 Mosaic 增强对小目标检测有重要作用,其次还有自适应锚框功能,能根据输入的数据集匹配较合适的锚框和自适应图像填充,能减少信息冗余,提高推理速度。Backbone 作为主干网络,主要用来提取特征,YOLOv5s 的 Backbone 包括 Conv、Focus、C3 和空间金字塔池化 (SPP) 模块,其中 Conv 是卷积单元,用来完成卷积、归一化和激活运算。与之前版本不同的是,YOLOv5s 第五代版本采用 SiLU 激活函数。SiLU 激活函数表达式为

$$\text{SiLU}(\mathbf{x}) = \mathbf{x} \times \frac{1}{1 + e^{-\mathbf{x}}}, \quad (1)$$

式中 \mathbf{x} 为输入矩阵。Focus 模块用来减少计算量,提升速度。C3 模块源自 BottleneckCSP 模块,由 Bottleneck 模块和 CSP 结构组成,相比 BottleneckCSP 模块,C3 模块在拥有相同性能的基础上更快、更轻量化。SPP 模块则聚合不同池化尺度信息,为网络获取更多信息。Neck 作为具有特征融合功能的网络,包含特征金字塔网络 (FPN) 和路径聚合网络 (PAN),FPN 能在网络中自上而下传递语义信息,PAN 则自下而上传递定位信息,实现不同深度特征图的信息融合,有效提高网络表达能力。Head 检测端用于最终目标检测,3 个 Detect 头用于在不同尺度特征图上预测不同尺寸的目标,通过生成的锚框同时预测类别和边界框偏移量。

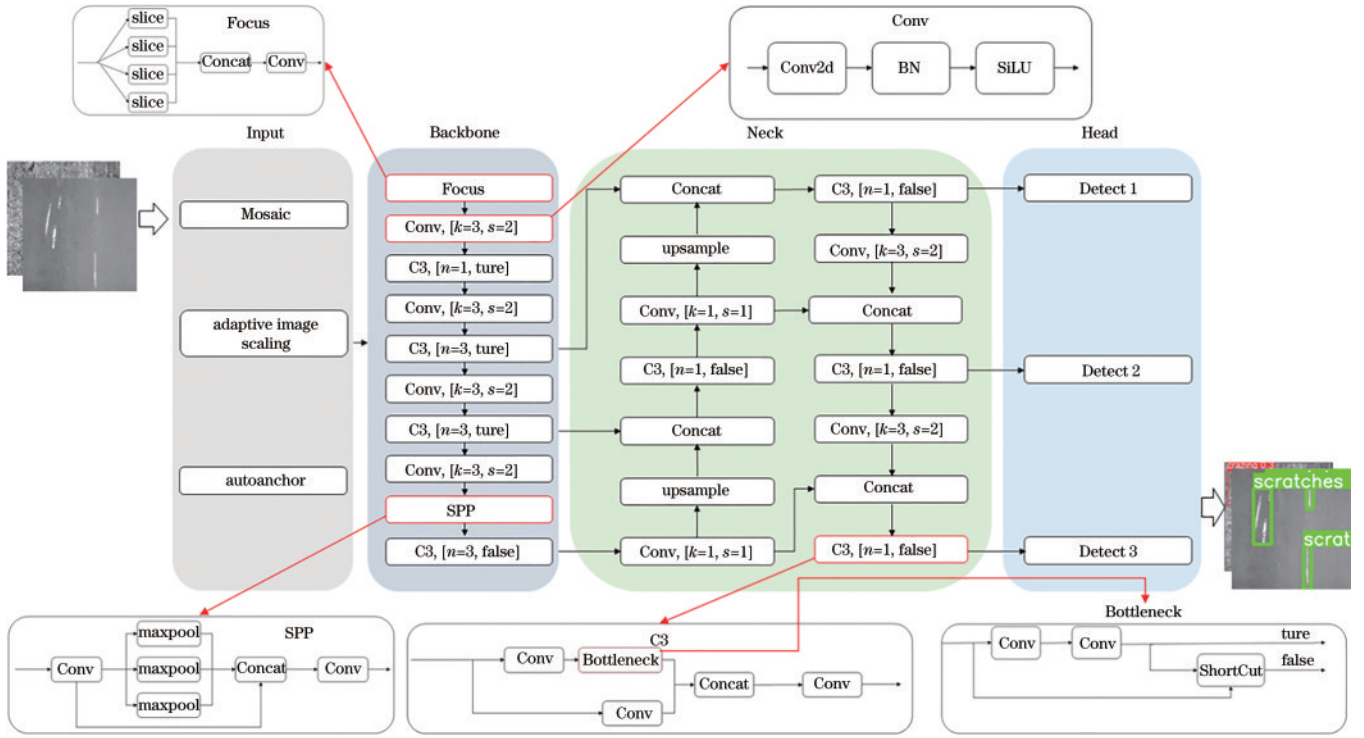


图 1 YOLOv5s 网络结构

Fig. 1 YOLOv5s network structure

2.2 损失函数

YOLOv5 损失函数由置信度损失 L_{conf} 、分类损失 L_{cls} 、回归框定位损失 L_{loc} 三部分组成, 表达式为

$$L = L_{\text{conf}} + L_{\text{cls}} + L_{\text{loc}} \quad (2)$$

其中, 置信度损失 L_{conf} 为

$$L_{\text{conf}} = - \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} \left[\bar{C}_i^j \log C_i^j + (1 - \bar{C}_i^j) \log (1 - C_i^j) \right] - \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{noobj}} \left[\bar{C}_i^j \log C_i^j + (1 - \bar{C}_i^j) \log (1 - C_i^j) \right], \quad (3)$$

式中: S^2 为网格数; B 为每个网格中锚框个数; I_{ij}^{noobj} 为预测框不含目标的指示值, 取值为 1 表示不包含, 0 表示包含; C_i^j 为预测置信度, \bar{C}_i^j 为真实置信度; λ_{noobj} 为自行设定的参数值。

分类损失 L_{cls} 为

$$L_{\text{cls}} = - \sum_{i=0}^{S^2} I_{ij}^{\text{obj}} \left\{ \bar{P}_i^j(c) \log [P_i^j(c)] + [1 - \bar{P}_i^j(c)] \log [1 - P_i^j(c)] \right\}, \quad (4)$$

式中: c 为检测目标所属的种类; $P_i^j(c)$ 为预测概率; $\bar{P}_i^j(c)$ 指网格中目标属于类别 c 的实际概率。

不同于 YOLOv3^[19] 使用均方误差 (MSE), YOLOv5 默认使用 CIoU loss 作为回归定位损失函数, 表达式为

$$L_{\text{loc}} = 1 - R_{\text{IoU}}(M, N) + \frac{\rho^2(M_{\text{ctr}}, N_{\text{ctr}})}{m^2} + \alpha v, \quad (5)$$

式中: $R_{\text{IoU}}(M, N)$ 表示预测框和标注框的 IoU 值; $\rho^2(M_{\text{ctr}}, N_{\text{ctr}})$ 表示预测框和检测目标标注框的中心点的欧氏距离; m 表示包含预测框和标注框的最小外接

矩形的对角线距离; α 为权重系数; v 为长宽比一致性参数。它们的计算公式为

$$\alpha = \frac{v}{[1 - R_{\text{IoU}}(M, N)] + v}, \quad (6)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w_{\text{gt}}}{h_{\text{gt}}} - \arctan \frac{w}{h} \right)^2, \quad (7)$$

式中: w_{gt} 和 h_{gt} 为真实框的宽和高; w 和 h 为预测框的宽和高。

3 改进的 YOLOv5s 算法

3.1 改进的锚框生成算法

在原始 YOLO 算法中, 预设的锚框 (anchor box) 是 K-means 算法在 COCO 数据集上进行聚类得到的。合理的锚框应该与标注框 (ground truth box) 的长宽比相接近, 因此需要对当前数据集进行重聚类。YOLOv5s 项目中有自适应锚框机制, 能根据输入的数据集聚类生成与之匹配度更高的锚框, 但是自适应锚框是通过误差平方和, 即欧氏距离来定义样本之间的距离的。经实验测试, 用 IoU 定义样本之间的距离生成的锚框与样本标注框更匹配。如果检测目标的标注框与聚类中心框 IoU 越大, 则距离越近。使用适应度 (fitness) 作为指标来评估通过不同算法生成的锚框与样本标注框的匹配程度, fitness 表示样本标注框与锚框 IoU 大于阈值 0.25 的数量占全部标注框数量的比值。表 1 为锚框统计表。

从表 1 可以看出: 基于 COCO 数据集生成的锚框对本实验数据集并不十分合适, fitness 值只有

表 1 锚框统计表

Table 1 Anchor box statistics table

Method	Anchor box									Fitness / %
Raw	(10,13)	(16,30)	(33,23)	(30,61)	(62,45)	(59,119)	(116,90)	(156,198)	(373,326)	68.46
Euclidean	(31,42)	(36,86)	(80,63)	(38,206)	(64,136)	(135,79)	(199,60)	(109,200)	(201,207)	72.87
IoU	(20,39)	(25,82)	(50,45)	(62,82)	(192,31)	(34,190)	(146,74)	(78,166)	(185,188)	75.56
Genetic	(20,42)	(25,82)	(51,46)	(176,29)	(64,85)	(34,190)	(146,74)	(78,166)	(185,188)	75.91

68.46%；基于欧氏距离重新聚类的锚框有一定程度改善；使用IoU进行样本之间的聚类，效果得到进一步提升，fitness值达75.56%；最后通过遗传算法(genetic algorithm)在此聚类结果上进行变异，fitness值达75.91%。可以看到，改进后的锚框生成算法能获得与样本标注框更匹配的锚框。

3.2 改进的Mosaic:融合MixUp

在Mosaic^[20]图像融合的过程中加入了MixUp^[21]背景虚化，来提升模型的性能。Mosaic是一种有效的数据增强策略，源自Cutmix^[22]。Cutmix的表达式为

$$\begin{cases} \tilde{x} = M \odot x_a + (1 - M) \odot x_b \\ \tilde{y} = \lambda y_a + (1 - \lambda) y_b \end{cases}, \quad (8)$$

式中：通过融合两个样本 (x_a, y_a) 和 (x_b, y_b) ，生成新的样本； $M \in \{0, 1\}^{W \times H}$ 表示一个二元掩模； $\mathbf{1}$ 为元素全为1的二元掩模； \odot 表示逐元素相乘运算；

$\lambda \sim \text{Beta}(\alpha, \alpha)$, $\alpha \in (0, \infty)$ 。Mosaic在Cutmix的基础上，将融合的样本数量由2张扩展为4张。

MixUp对两张样本图按比例进行插值来生成一张混合样本，操作表达式为

$$\begin{cases} \tilde{x} = \lambda x_a + (1 - \lambda) x_b \\ \tilde{y} = \lambda y_a + (1 - \lambda) y_b \end{cases}. \quad (9)$$

对样本 (x_a, y_a) 和样本 (x_b, y_b) 进行线性叠加，生成新样本 (\tilde{x}, \tilde{y}) ，通过这种方式可以让样本间分布更加平滑。将Mosaic与MixUp结合起来，改进后的表达式为

$$\begin{cases} \tilde{x} = M_1 \odot x_a + M_2 x_b + M_3 \odot x_c + \mu x_d \\ \tilde{y} = \lambda_1 y_a + \lambda_2 \odot y_b + \lambda_3 \odot y_c + \lambda_4 \odot y_d \end{cases}, \quad (10)$$

式中： $M_i \in \{0, 1\}^{W_i \times H_i}$ 表示第*i*个二元掩模，其中 $\sum W_i = W$, $\sum H_i = H$ ； μ 和 λ 均为透明度调整系数。融合MixUp，改进后的Mosaic处理前后的效果如图2所示。

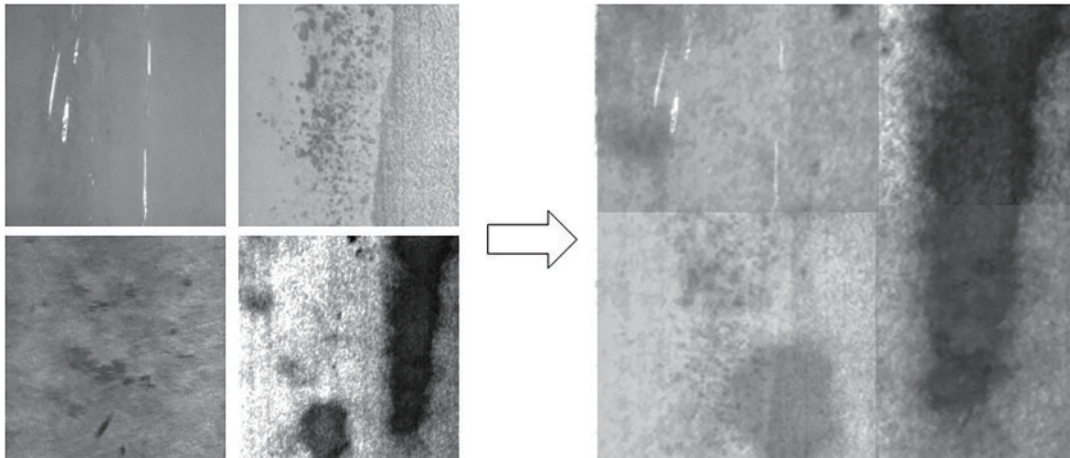


图2 融合MixUp的Mosaic数据扩充
Fig. 2 Mosaic data expansion with MixUp

3.3 改进的网络结构:融合注意力机制

在计算机视觉领域，注意力机制的有效性得到了广泛的证明。为了提高YOLOv5s网络的表达能力，将CBAM^[23]融入网络Backbone结构中。CBAM由通道注意力模块(CAM)和空间注意力模块(SAM)两个子模块组成，结构如图3所示。

CBAM由CAM和SAM串行组合在一起，对于输入特征图 $F \in \mathbb{R}^{C \times H \times W}$ ，CBAM在其一维通道上和二维空间上分别进行重构，得到特征图 F_c 和图 F_s 。

F_c 表示经过通道注意力处理后的特征图，处理的

过程可表示为

$$F_c = M_c(F) \otimes F, \quad (11)$$

$$M_c = \text{Sigmoid}\left\{ \text{MLP}\left[\text{GAP}(F)\right] + \text{MLP}\left[\text{GMP}(F)\right] \right\}, \quad (12)$$

式中： M_c 表示通道注意力模块。输入特征图 F 的各个通道并行经过全局最大池化(GMP)和全局平均池化(GAP)，得到两个 $1 \times 1 \times C$ 的特征图，再将其分别送入多层感知机(MLP)，对输出特征进行逐元素加和运算和Sigmoid激活函数处理，生成通道注意力特征 M_c 。 M_c 是一个 C 维向量，表达式为

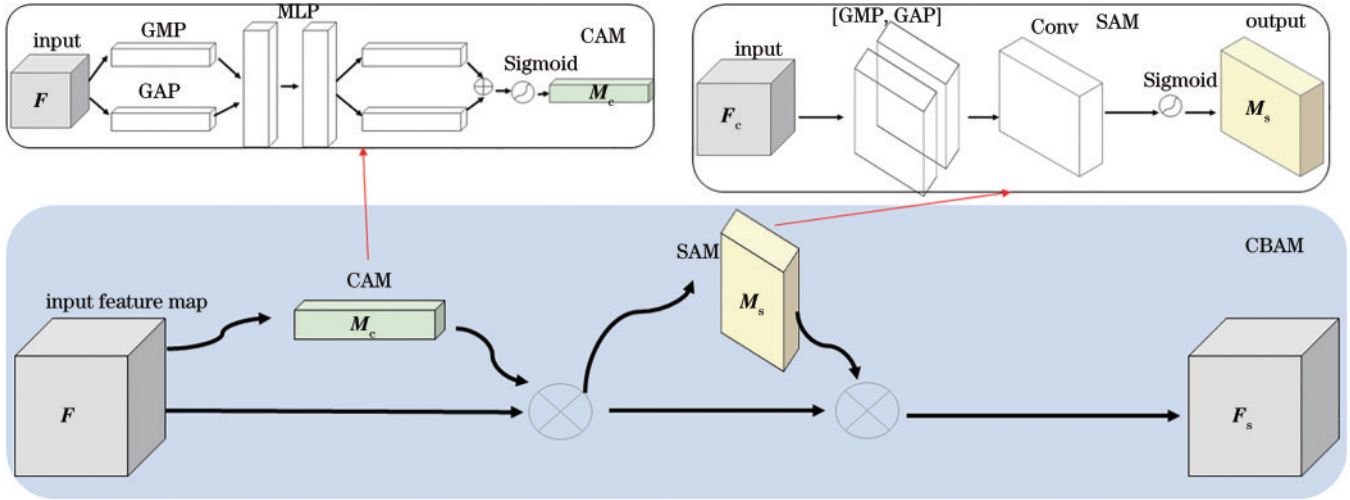


图3 CBAM结构

Fig.3 CBAM structure

$$M_c = (\alpha_1, \alpha_2, \dots, \alpha_C)^T. \quad (13)$$

M_c 将与 F 进行逐元素乘法运算,生成带通道注意力的特征图 F_c ,该模块输出能够有效关注重要通道,抑制非重要通道的影响。 F_s 表示经过空间注意力处理后的特征图,处理过程可表示为

$$F_s = M_s(F_c) \otimes F_c, \quad (14)$$

$$M_s = \text{Sigmoid}\{f^{7 \times 7}[\text{GAP}(F_c), \text{GMP}(F_c)]\}, \quad (15)$$

式中: M_s 表示空间注意力模块; $f^{7 \times 7}$ 表示卷积核为 7×7 的卷积运算。 F_c 首先分别进行最大池化和平均池化,得到两个 $H \times W \times 1$ 的矩阵;然后进行基于通道的Concat,即在通道维度上进行拼接,再通过 7×7 的卷积降维成 $H \times W \times 1$ 的矩阵,送入激活函数Sigmoid生成空间注意力特征 M_s ;最后与 F_c 逐元素相乘,得到带空间注意力的特征图 $F_s \in \mathbb{R}^{C \times H \times W}$ 。该模块能在空间维度上对有价值的信息进行提取。

在一个检测模型中,Backbone的主要作用是特征提取,Neck的主要作用是特征融合,而Head则基于融合后的特征进行分类和回归预测。所以本文实验中把CBAM嵌入到YOLOv5s的Backbone部分中,利用通道注意力和空间注意力来挖掘特征表征潜能,加强模型的特征提取能力,这对整个网络表达力的提升有重要作用。融合CBAM后的Backbone结构如图4所示。

3.4 改进的损失函数:困难样本挖掘

在使用NEU-DET数据集训练模型的过程中,通过对失败案例进行分析,能够发现部分类别的平均精度(AP)在整个训练过程中明显低于其他类别,某次训练结果如图5所示。

对缺陷scratches的AP能达90%,而对缺陷crazing的AP不超过50%,存在样本难易不均衡问题。因此,引入Focal loss^[24],改进YOLOv5s损失函数,对困难样本进行挖掘。二分类交叉熵损失函数为

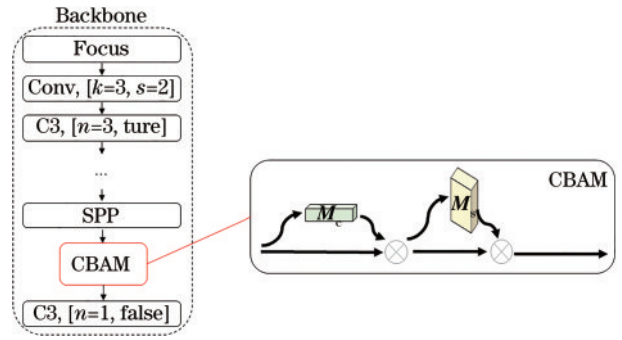


图4 融合注意力模块

Fig. 4 Fusion attention module

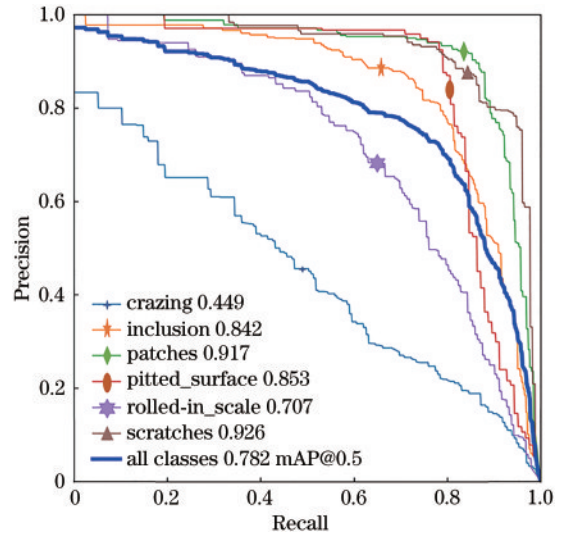


图5 对各类别缺陷的AP统计

Fig. 5 AP statistics of various types of defects

$$L_{ce} = \begin{cases} -\log p_i, & y = 1 \\ -\log(1 - p_i), & y = 0 \end{cases}, \quad (16)$$

式中: y 是真实样本标签; p_i 为正样本预测概率。损失函数容易淹没在大量的负样本和简单样本中,影响迭

代速度。Focal loss 通过引入两个参数,降低负样本、简单样本对损失值的影响,提高正样本与困难样本的权重,表达式为

$$L_{fl} = \begin{cases} -\alpha_t(1-p_t)^\gamma \log p_t, & y=1 \\ -(1-\alpha_t)p_t^\gamma \log(1-p_t), & y=0 \end{cases}, \quad (17)$$

式中: α_t 是权重系数, $\alpha_t \in [0, 1]$, 用来改善正负样本不平衡问题; γ 是调制系数, γ 取值越大, 简单样本的损失衰减越快, 困难样本对损失值的影响权重提高, 模型将更专注于对困难样本的学习。当 $\gamma=1, \alpha_t=0$ 时, Focal loss 即退化为交叉熵损失函数。

改进后的类别损失函数为

$$L_{lcls} = -\sum_{i=0}^{S^2} I_{ij}^{obj} \left\{ \alpha_t [1 - P_i^j(c)]^\gamma \bar{P}_i^j(c) \log [P_i^j(c)] + (1 - \alpha_t) [P_i^j(c)]^\gamma [1 - \bar{P}_i^j(c)] \log [1 - P_i^j(c)] \right\}. \quad (18)$$

图 6 是引入 Focal loss 后网络检测效果的对比。YOLOv5-A 是 YOLOv5 的组合改进版本 (锚框 + MixUp + CBAM), YOLOv5-B 是在此基础之上还引入 Focal loss。对于难识别缺陷样本 crazing, YOLOv5-A 的 AP 为 44.9%, YOLOv5-B 的 AP 为 48.8%。在使用 Focal loss 改进损失函数后, 对难识别样本 crazing 的 AP 大幅提升 3.9 个百分点, YOLOv5-B 对困难样本类别识别有很好的改善。

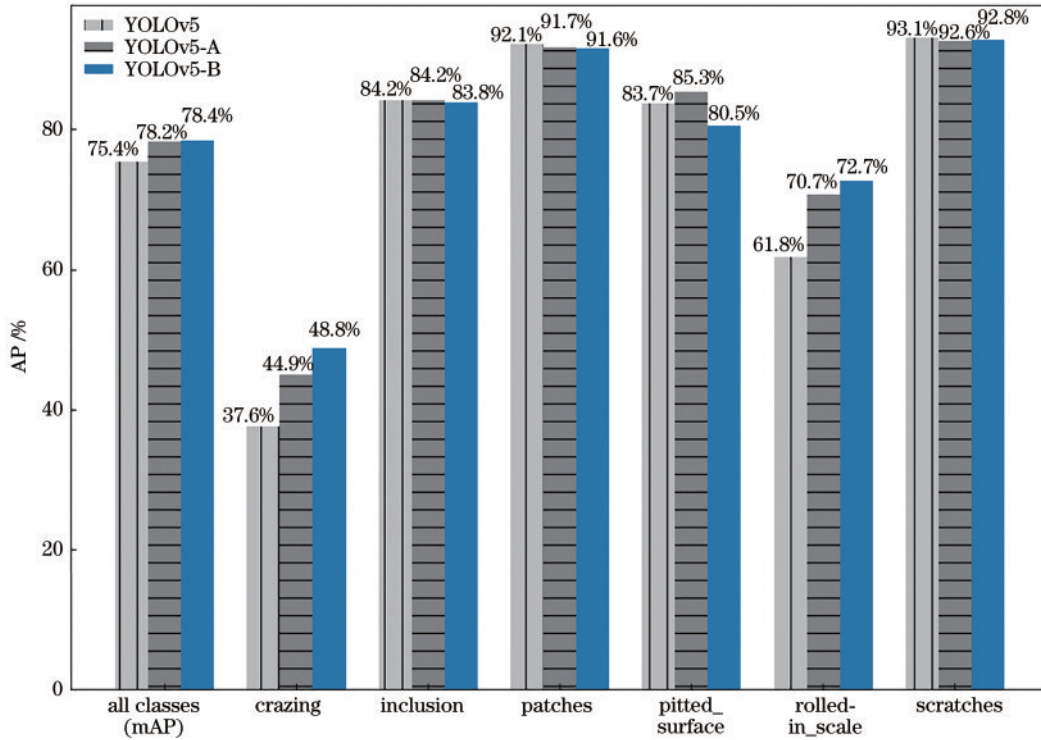


图 6 改进损失函数前后各类别缺陷 AP 对比

Fig. 6 AP comparison of various types of defects before and after improvement of loss function

4 实验结果与分析

4.1 实验环境及参数

在 Windows10 系统中采用 PyTorch 深度学习框架搭建网络。处理器为 Inter(R) Core(TM) i5-9300H CPU @ 2.40 GHz; 显卡为 NVIDIA GeForce GTX 1660 Ti; PyTorch 版本为 1.7.0; CUDA 版本为 10.1。

本实验使用 COCO 预训练模型进行参数初始化, 整个训练过程有 500 epoch, 优化器为 SGD, 初始学习率为 0.01, 动量为 0.937, 权重衰减因子为 0.005, batchsize 为 16。

4.2 数据集

实验使用 NEU-DET 数据集, NEU-DET 是东北大学发布的钢材表面缺陷样本数据集, 数据集包含 6 类

钢材表面缺陷各 300 张, 图片大小为 200×200 , 部分图片存在不止一种缺陷。如图 7(a) 所示, 6 类钢材表面缺陷分别为压入氧化铁 (rolled-in_scale)、斑块 (patches)、裂纹 (crazing)、麻点 (pitted_surface)、夹杂 (inclusion) 及划痕 (scratches)。图 7(b) 为 6 类缺陷样本数统计。将原始数据集以 7:3 的比例随机划分为训练集和测试集, 即训练集 1260 张, 测试集 540 张。原始数据集标注信息为 xml 文件, 通过格式转换, 脚本转换为包含缺陷类别和位置信息的 txt 格式文件, 以满足 YOLOv5s 网络模型程序能接收的标注文件类型要求。

4.3 评价指标

采用目标检测模型常用评估指标, 平均精度 (AP)、平均精度均值 (mAP) 以及检测速度三项指标进行评估。AP 指查准率 (precision, P) - 查全率 (recall,

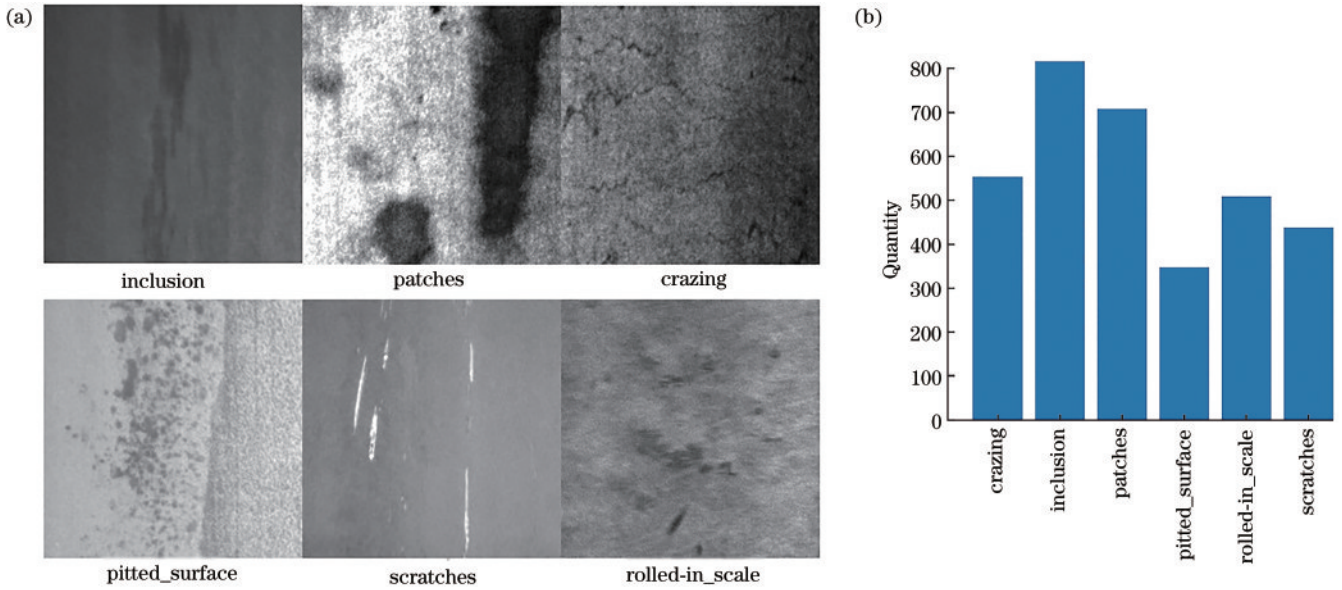


图 7 数据集及各缺陷样本数统计
Fig. 7 Data set and statistic of samples of each defect

R) 曲线下的面积, 用来评估模型在单个检测类别上的精度。mAP 是所有类别 AP 的均值, 用来评估整体检测精度。查准率和查全率的公式为

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}}, \quad (19)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad (20)$$

式中: 真阳性 (TP) 表示检测正确的目标数量; 假阳性 (FP) 表示检测错误的目标数量; 假阴性 (FN) 表示漏检的目标数量。

4.4 消融实验及对比实验

4.4.1 改进方法对模型性能的影响

为验证对 YOLOv5s 的多种改进策略的有效性, 在 NEU-DET 数据集上进行消融实验, 实验结果如表 2 所示。对表 2 进行分析可知: 第 1 组为原 YOLOv5s 算法, mAP 值为 75.4%; 第 2 组, 使用基于 IoU 定义样

本之间距离的 K-means 算法重新生成锚框, 并利用遗传算法对其进行变异, mAP 和 recall 分别提升 0.8 个百分点和 1.3 个百分点, 说明生成的锚框对本实验数据集更合适; 第 3 组, 在 Mosaic 图像融合过程中加入基于 MixUp 的背景虚化, mAP 提升了 0.7 个百分点; 第 4 组, 在前组基础上引入了 CBAM, 由于通道信息和空间信息的加入, 仅增加少量参数, 网络性能提升明显, mAP 达 78.2%; 第 5 组, 在损失函数中融入 Focal loss, 提高模型对困难样本的关注, 检测精度获得进一步提升, 相较于改进前, mAP 值提高了 3.0 个百分点, 效果提升显著。改进后的 YOLOv5s 网络的 mAP 值为 78.4%。图 8 为改进前后效果图对比, 可以看到, 改进后一方面识别准确率得到提升, 另一方面小目标检测效果也有所改善, 如图片右下角椭圆中的缺陷在原始 YOLOv5s 中未被检测到, 但是在改进后的网络中被准确地识别出来。

表 2 YOLOv5s 的消融实验
Table 2 Ablation experiments of YOLOv5s

Method	Anchor	MixUp	CBAM	Focal loss	Precision / %	Recall / %	Number of parameters	mAP / %
YOLOv5s					77.1	72.5	7035811	75.4
Improved 1	✓				75.9	73.8	7035811	76.2
Improved 2	✓	✓			77.0	72.6	7035811	76.9
Improved 3	✓	✓	✓		77.9	72.3	7068677	78.2
Improved 4	✓	✓	✓	✓	78.0	73.2	7068677	78.4

4.4.2 与主流目标检测模型性能对比

为了验证所提改进的 YOLOv5s 的检测性能, 与主流目标检测模型 YOLOv3、DDN、Faster R-CNN 等进行对比, 采用 mAP 和检测速度两项指标进行评估,

实验结果如表 3 所示。由表 3 分析可知: 改进的 YOLOv5s 算法的 mAP 达 78.4%, 比原始 YOLOv5s 算法提高 3.0 个百分点, 相比于其他主流目标检测网络模型, 能保持较快的检测速度和较好的检测性能。

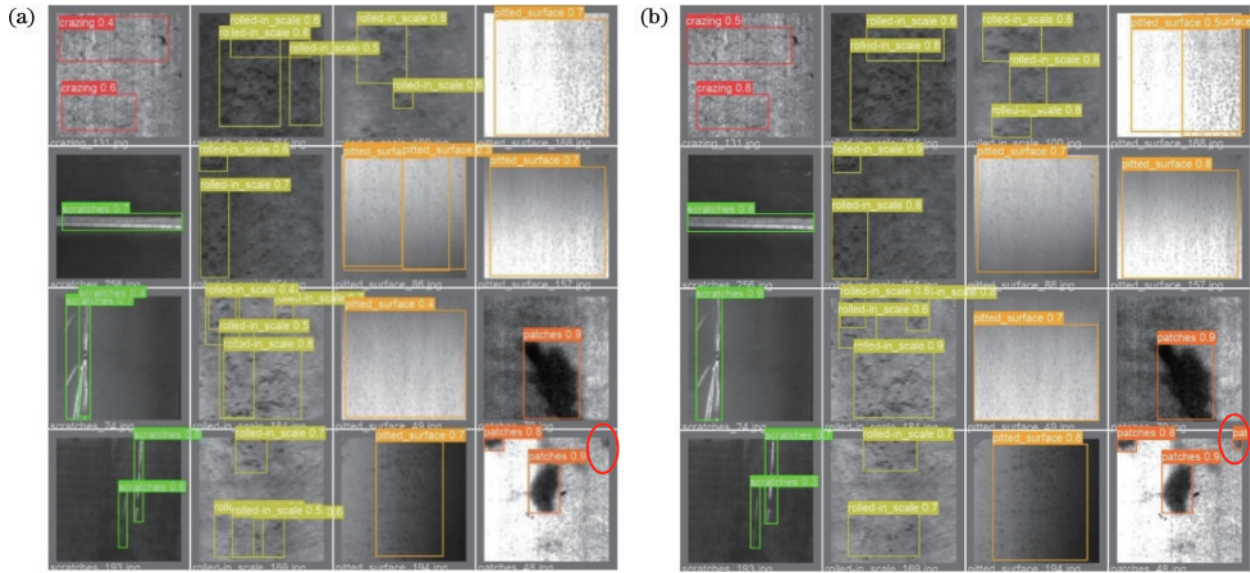


图 8 改进前后检测效果对比。(a)原始 YOLOv5s; (b)改进的 YOLOv5s

Fig. 8 Comparison of detection results before and after improvement. (a) Original YOLOv5s; (b) improved YOLOv5s

表 3 不同模型的性能对比

Table 3 Performance comparison of different models

Network	Speed / (frame · s ⁻¹)	mAP / %	AP / %					
			crazing	inclusion	patches	pitted_surface	rolled-in_scale	scratches
YOLOv3	29	69.1	44.7	60.8	84.4	74.5	61.1	87.2
DDN(ResNet34) ^[16]	<20	74.8	48	75.9	87.4	78.3	68.4	90.8
DDN(ResNet50) ^[16]	<10	82.3	62.4	84.7	90.3	89.7	76.3	90.1
Faster R-CNN(ResNet34) ^[16]	<20	70.2	46.7	61.3	82.8	76.5	70.7	83.4
YOLOv5s	41	75.4	37.6	84.2	92.1	83.7	61.8	93.1
Improved YOLOv5s	40	78.4	48.8	83.8	91.6	80.5	72.7	92.8

5 结 论

提出一种改进的 YOLOv5s 算法并将其应用于钢铁表面缺陷检测。首先使用基于 IoU 度量距离的 K-means 算法对钢铁表面缺陷数据集进行聚类,通过遗传算法进行变异运算,得到更合适的锚框;然后在 Mosaic 上加入 MixUp,并且在网络主干部分引入注意力模块,加强网络的特征提取能力;最后对于困难样本,引入 Focal loss 函数来改进类别损失函数,提高难样本权重。实验结果表明:相比于原始 YOLOv5s 算法,改进后的 YOLOv5s 算法可在保持较快检测速度的情况下有效提高对钢铁表面缺陷的检测精度,实现对钢铁表面缺陷的智能高效检测。下一步,将尝试融合 Transformer 来改进目标检测模型,进一步提高模型的表达能力。

参 考 文 献

[1] 李少波, 杨静, 王铮, 等. 缺陷检测技术的发展与应用研究综述[J]. 自动化学报, 2020, 46(11): 2319-2336.
Li S B, Yang J, Wang Z, et al. Review of development and application of defect detection technology[J]. Acta

Automatica Sinica, 2020, 46(11): 2319-2336.
[2] Arjun V, Sasi B, Rao B P C, et al. Optimisation of pulsed eddy current probe for detection of sub-surface defects in stainless steel plates[J]. Sensors and Actuators A: Physical, 2015, 226: 69-75.
[3] Joung O J, Kim Y H. Application of an IR thermographic device for the detection of a simulated defect in a pipe[J]. Sensors, 2006, 6(10): 1199-1208.
[4] You C W, Lu C, Wang T Y, et al. Method for defect contour extraction in terahertz non-destructive testing conducted with a raster-scan THz imaging system[J]. Applied Optics, 2018, 57(17): 4884-4889.
[5] 范宏, 侯云, 李柏林, 等. 基于区域特征的缺陷扣件视觉检测[J]. 铁道学报, 2021, 43(8): 132-138.
Fan H, Hou Y, Li B L, et al. Visual inspection of defective fasteners based on regional features[J]. Journal of the China Railway Society, 2021, 43(8): 132-138.
[6] Tsukada K, Majima Y, Nakamura Y, et al. Detection of inner cracks in thick steel plates using unsaturated AC magnetic flux leakage testing with a magnetic resistance gradiometer[J]. IEEE Transactions on Magnetics, 2017, 53(11): 1-5.
[7] 苗玲, 高斌, 石永生, 等. 基于电涡流热成像的钢轨滚动接触疲劳裂纹动态检测研究[J]. 机械工程学报,

- 2021, 57(18): 86-97.
- Miao L, Gao B, Shi Y S, et al. Research on dynamic detection of rail rolling contact fatigue crack based on eddy current thermography[J]. *Journal of Mechanical Engineering*, 2021, 57(18): 86-97.
- [8] Jeon Y J, Choi D C, Lee S J, et al. Steel-surface defect detection using a switching-lighting scheme[J]. *Applied Optics*, 2016, 55(1): 47-57.
- [9] Yi L, Li G Y, Jiang M M. An end-to-end steel strip surface defects recognition system based on convolutional neural networks[J]. *Steel Research International*, 2017, 88(2): 1600068.
- [10] 王宸, 张秀峰, 刘超, 等. 改进 YOLOv3 的轮毂焊缝缺陷检测[J]. *光学精密工程*, 2021, 29(8): 1942-1954.
- Wang C, Zhang X F, Liu C, et al. Detection method of wheel hub weld defects based on the improved YOLOv3[J]. *Optics and Precision Engineering*, 2021, 29(8): 1942-1954.
- [11] 胡勃, 秦威, 刘成良, 等. 具有复杂纹理的木板表面刮痕缺陷检测模型[J/OL]. *计算机集成制造系统*: 1-20 [2021-11-28]. <http://kns.cnki.net/kcms/detail/11.5946.TP.20211109.1542.008.html>.
- Hu Q, Qin W, Liu C L, et al. Scratch defect detection model on wooden board surface with complex texture[J/OL]. *Computer Integrated Manufacturing Systems*: 1-20 [2021-11-28]. <http://kns.cnki.net/kcms/detail/11.5946.TP.20211109.1542.008.html>.
- [12] 赵振兵, 熊静, 李冰, 等. 基于改进 Cascade R-CNN 的典型金具及其部分缺陷检测方法[J]. *高电压技术*, 2022, 48(3): 1060-1067.
- Zhao Z B, Xiong J, Li B, et al. Typical fittings and its partial defect detection method based on improved Cascade R-CNN[J]. *High Voltage Engineering*, 2022, 48(3): 1060-1067.
- [13] 汪权, 易本顺. 基于 Gaussian YOLOv3 的航拍图像绝缘子缺陷识别[J]. *激光与光电子学进展*, 2021, 58(12): 1210022.
- Wang Q, Yi B S. Insulator defect recognition in aerial images based on Gaussian YOLOv3[J]. *Laser & Optoelectronics Progress*, 2021, 58(12): 1210022.
- [14] Xing J J, Jia M P, Xu F Y, et al. A method for workpiece surface small-defect detection based on CutMix and YOLOv3[J]. *Journal of Southeast University (English Edition)*, 2021, 37(2): 128-136.
- [15] 李彬, 汪诚, 丁相玉, 等. 改进 YOLOv4 的表面缺陷检测算法[J/OL]. *北京航空航天大学学报*: 1-10 [2021-11-28]. DOI:10.13700/j.bh.1001-5965.2021.0301.
- Li B, Wang C, Ding X Y, et al. Surface defect detection algorithm based on improved YOLOv4 [J/OL]. *Journal of Beijing University of Aeronautics and Astronautics*: 1-10 [2021-11-28]. DOI:10.13700/j.bh.1001-5965.2021.0301.
- [16] He Y, Song K C, Meng Q G, et al. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features[J]. *IEEE Transactions on Instrumentation and Measurement*, 2020, 69(4): 1493-1504.
- [17] 来文豪, 周孟然, 王锦国, 等. 多光谱波段筛选的煤矸石快速定位[J]. *中国激光*, 2021, 48(16): 1611001.
- Lai W H, Zhou M R, Wang J G, et al. Fast location of coal gangue based on multispectral band selection[J]. *Chinese Journal of Lasers*, 2021, 48(16): 1611001.
- [18] 张力, 黄丹平, 廖世鹏, 等. 基于目标检测网络的轮对踏面缺陷检测方法[J]. *激光与光电子学进展*, 2021, 58(4): 0410020.
- Zhang L, Huang D P, Liao S P, et al. Wheelset tread defect detection method based on target detection network [J]. *Laser & Optoelectronics Progress*, 2021, 58(4): 0410020.
- [19] Farhadi A, Redmon J. Yolov3: an incremental improvement[EB/OL]. (2018-04-08)[2021-02-03]. <https://arxiv.org/abs/1804.02767>.
- [20] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23)[2021-05-04]. <https://arxiv.org/abs/2004.10934>.
- [21] Zhang H Y, Cisse M, Dauphin Y N, et al. Mixup: beyond empirical risk minimization[EB/OL]. (2017-10-25)[2021-02-05]. <https://arxiv.org/abs/1710.09412>.
- [22] Yun S, Han D, Chun S, et al. CutMix: regularization strategy to train strong classifiers with localizable features [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 6022-6031.
- [23] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018*. Lecture notes in computer science. Cham: Springer, 2018, 11211: 3-19.
- [24] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2999-3007.
- [25] Zhu X K, Lü S C, Wang X, et al. TPH-YOLOv5: improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios[C]//2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), October 11-17, 2021, Montreal, BC, Canada. New York: IEEE Press, 2021: 2778-2788.
- [26] 王红梅, 王晓鸽, 王晓燕. 基于深度学习的复杂背景下目标检测研究[J/OL]. *控制与决策*: 1-8 [2021-11-29]. DOI:10.13195/j.kzyjc.2021.0686.
- Wang H M, Wang X G, Wang X Y. Research on target detection under complex background based on deep learning[J/OL]. *Control and Decision*: 1-8 [2021-11-29]. DOI:10.13195/j.kzyjc.2021.0686.