

KThin-YOLOV7:轻量级的焊接件表面缺陷检测^{*}

卢开喜 段先华 陶宇诚 倪东海

(江苏科技大学计算机学院 镇江 212100)

摘要: 针对目前市面主流焊接件表面缺陷的模型检测精度不高,模型复杂和不满足实时监测等问题,提出了一种基于 YOLOV7-tiny 改进得到的焊接件表面缺陷新型检测模型 KThin-YOLOV7。首先,设计了基于模拟人类视觉感受野的 EMA-BasicRFBC 模块,更换 YOLOV7-tiny 模型的空间特征金字塔 SPP 模块,从而加强模型特征表达的性能。其次,以 SlimNeck 设计范式结构为基础设计了 ThinNeck 结构,并用其更换 YOLOV7-tiny 的 NECK 特征融合部分,减少模型的参数量和计算量的同时提高了模型的平均检测精度。最后,引入 K-means++ 算法找出合适的锚框,并用 FEIOU 损失函数更换原模型的 LOSS,进一步帮助模型优化目标框的位置和大小。KThin-YOLOV7 相对原始 YOLOV7-tiny 模型的 mAP 提升了 7.11%,达到 87.64%,同时模型的参数量和计算量分别下降了 11.14% 和 15.26%。实验结果表明,KThin-YOLOV7 能够高效且准确地定位检测焊接件表面的缺陷。

关键词: 焊接件表面缺陷;目标检测;YOLOv7-tiny;RFB;ThinNeck;K-means++

中图分类号: TP183;TN249.2;TP391.41 **文献标识码:** A **国家标准学科分类代码:** 520.6040

KThin-YOLOV7: Lightweight inspection of surface defects on welded parts

Lu Kaixi Duan Xianhua Tao Yucheng Ni Donghai

(School of Computer Science, Jiangsu University of Science and Technology, Zhenjiang 212100, China)

Abstract: Aiming at the current market mainstream weldment surface defects model detection accuracy is not high, the model is complex and does not meet the real-time monitoring and other issues, a new detection model of weldment surface defects based on the improvement of YOLOV7-tiny obtained KThin-YOLOV7 is proposed. Firstly, the EMA-BasicRFBC module, which is based on simulating the human visual sensory field, was designed to replace the spatial feature pyramid SPP module of the YOLOV7-tiny model, so as to enhance the performance of the model feature expression. Secondly, the ThinNeck structure is designed based on the SlimNeck design paradigm structure, and it is used to replace the NECK feature fusion part of YOLOV7-tiny, which reduces the number of parameters and computation of the model and improves the average detection accuracy of the model at the same time. Finally, the K-means++ algorithm is introduced to find out the appropriate anchor frame and replace the LOSS of the original model with the FEIOU loss function, which further helps the model to optimize the position and size of the target frame. The mAP of the KThin-YOLOV7 is improved by 7.11% to 87.64% compared to the original YOLOV7-tiny model, while the number of parameters and computation of the model are decreased by 11.14% and 15.5%, respectively. Experimental results show that KThin-YOLOV7 can efficiently and accurately locate and detect defects on the surface of welded parts.

Keywords: weldment surface defects; target detection; YOLOv7-tiny; RFB; ThinNeck; K-means++

0 引言

焊接件表面缺陷检测是焊接质量控制中至关重要的一环。焊接件表面缺陷的存在会对产品的外观和质量产生负面影响,因此在焊接件的生产过程中,表面缺陷检测显得尤

为重要。常见的焊接件表面缺陷包括冲孔、焊缝、月牙弯、油斑和水斑等。然而,在真实而复杂的工业环境中,表面缺陷检测面临着多重挑战,如缺陷成像与背景之间的差异较小、对比度较低以及缺陷尺度变化较大等。传统的基于机器视觉的方法往往难以达到理想的检测效果^[1-3]。

在焊接件缺陷检测阶段,传统的工业检测方法效率低,用工成本也较高,已经逐渐不满足焊接件缺陷检测的发展现状,而计算机视觉中人工智能的深度学习具有出色的学习能力和对大量数据进行特征提取的能力,因此许多学者将其应用于产品缺陷的识别,其可以以更准确且快速地帮助检测人员对目标缺陷进行定位和分类^[4-6]。

目前,缺陷检测主要采用计算机视觉中基于深度学习的目标检测算法。这些算法可以分为双阶段和单阶段两种。双阶段目标检测算法将整个检测过程分为两个阶段,首先生成候选框,然后对候选框进行分类和定位。代表性的算法有,Mask-R-CNN^[7]算法以及 Faster-R-CNN^[8]算法等。单阶段目标检测算法则将整个检测过程放在一起进行,不需要生成候选框。常见的算法有 YOLO^[9-12]系列算法的以及 SSD^[13]算法等。

在上述算法中,单阶段的 YOLO 算法更为适用,其检测速度超过双阶段算法,但在精度上略有不足。考虑到其在检测速度和模型复杂性方面的优化潜力,学者们相应地提出了多种改进策略。

Zhang 等^[14]采用最新 YOLOv3 版本应用于桥梁表面缺陷定位,引入了参数调节,进一步提高了缺陷检测性能。霍超等^[15]改进的 YOLOv4 模型对绝缘子缺陷的检测精度高,模型的尺寸较小。杜娟等^[16]提出一种改进 YOLOv7 的复杂道路场景目标检测算法。用 K-means++ 重聚类先验框,使得先验框更贴合目标,采用 WIoU 损失函数与 P-ELAN 模块,提高了网络对目标的定位能力。赵春华等^[17]针对钢材表面缺陷提出一种基于 YOLOv7-tiny 网络改进的轻量级算法模型 PC-YOLOv7。PC-YOLOv7 算法在 NEU-DET 数据集上平均精度均值(mAP)达到了 78.5%,其不仅提高了检测精度并且在一定程度上减少了模型的尺寸。

虽然,国内外学者已经开始着手解决优化检测精度和模型尺寸的问题,但其仍然有改进空间。为了进一步平衡焊接件表面缺陷检测精度与模型尺寸的问题,本文提出了一种基于 YOLOV7-tiny 网络改进的焊接件表面缺陷检测模型 KThin-YOLOV7,主要工作如下:

1)针对上述焊接件表面缺陷数据集,通过翻转、随机变形缩放、亮度变化和 prewitt 算子锐化去噪等操作来扩张、增强原始数据。这样不仅增加了图片的多样性,而且易于增强模型的鲁棒性。

2)设计了特征提取模块(efficient multi-scale attention-basic receptive field block cheaper, EMA-BasicRFBC),更换 YOLOV7-tiny 模型特征金字塔模块(spatial pyramid pooling, SPP),在保持简单和轻量级的同时,能够增强模型的特征表达能力,从而提高区分焊接件表面缺陷特征的能力。

3)设计了纤细颈部(thinner neck, ThinNeck)来更换 yolov7-tiny 的 NECK 部位减少模型的参数量和计算量,提

高了模型的平均检测精度。

4)引入 K-means++ 算法与 FEIOU 损失函数,可以优化锚框,从而准确地定位缺陷位置并且减少网络收敛速度。

本文 KThin-YOLOV7 检测模型能够对焊接件表面缺陷进行快速且准确的识别,并且模型尺寸小,在工业缺陷检测方面具有广阔的应用前景。

1 KThin-YOLOV7 模型

KThin-YOLOV7 在不改变基准模型 YOLOV7-tiny 算法的整体框架下,主要通过增强头部特征提取、特征融合、多尺度处理、上下文信息的融合和参数优化等方面的能力,从而去更好地定位缺陷,减少 YOLOV7-tiny 参数数量的同时,进一步提升了模型的精度。其拥有更灵活的输入特征应对策略,更优越的检测性能,可以适用于焊接件表面缺陷检测,其结构如图 1 所示。

图 1 为本文提出的 KThin-YOLOV7 的模型结构。其改进措施如下:对于特征金字塔 SPP 模块,设计了 EMA-BasicRFBC 模块。此外,设计了 ThinNeck 结构来代替原 YOLOV7-tiny 中的 Neck 特征融合部分,其中 MCB 模块替换成改进的轻量级模块(efficient multi-scale attention-vanilla object vocabulary global semantic contextual spatial pyramid cheaper, EMA-VoVGSCSPC),协助采用了上采样算子(content-aware reassembly of features, CARAFE)与分组空间卷积(grouped spatial convolution, GSConv),并将 Head 部分的 CBL 卷积全部更换为 GSconv 卷积。

1.1 EMA-BasicRFBC 模块

基于原来 YOLOV7-tiny 虽然能够降低计算量和参数量,但是在现实检测状况下,不同缺陷的检测难度不同,油斑与水斑的缺陷尺度、样貌接近,故其表现特征接近。原有模型无法达到最优检测性能。为了解决上述问题,本文对特征金字塔 SPP 模块进行改进,提出了基于人类视觉感受野的 EMA-BasicRFBC 模块。BasicRFP^[18]它是一个提高感受野的模块,它能够提取不同缺陷的细节,从而增强特征的表达能力和语义信息。而高效的多尺度注意力模块(efficient multi-scale attention, EMA),它以保留每个通道上的信息和降低计算开销为目标,将空间语义特征在每个特征组中均匀分布。其二维全局池化操作如式(1)所示。

$$z_c = \frac{1}{H \times W} \sum_j^H \sum_i^W x_c(i, j) \quad (1)$$

它是为编码全局信息和模拟长距离的依赖关系而设计的。其中 z_c 表示二维全局池化输出, x_c 表示第 c 个通道的输入特征。H 和 W 分别表示输入特征的空间尺寸。这样设计,着重于保留每个通道的信息并减少计算开销,其跨空间学习的能力,也可以实现更丰富的特征聚合。其 EMA^[19]文献中说明,EMA 较于 CBAM 注意力以及 CA 注意力等具有更少的计算开销和更好的特征提取效果。将以上两模块结合,能够进一步加强模型特征提取不同缺陷的

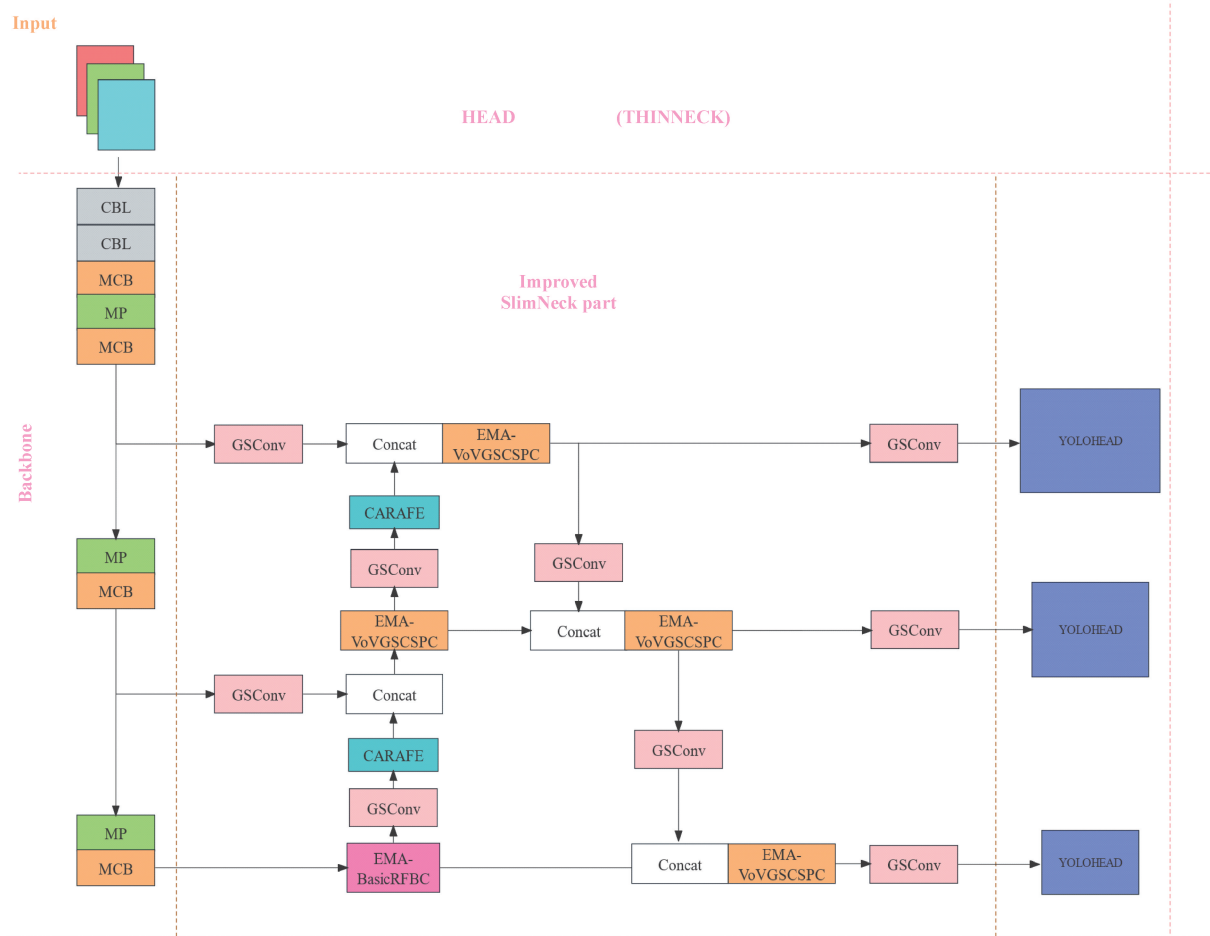


图 1 KThin-YOLOV7 网络结构图

能力。改进后的 EMA-BasicRFBC 如图 2 下部分所示。

具体的实验操作步骤如下:首先,使用 BasicRFP 模块来替换 SPP 模块,如图 2 上部分所示。其 BasicRFP 结构是将主干的 MCB 的输出当做其输入,然后进行 3 个分支 branch 的特征堆叠,并最后经过非线性函数 Relu 来增加模型的非线性,完成了多尺度的特征提取。相对于 SPP 模块来说,BasicRFB 使用多个小卷积核代替大卷积核,减少模型的参数数量,增加模型的非线性,又使用空洞卷积增加感受野,提高对目标的识别能力。文献[20]中特征提取模块(basic receptive field block-slim,BasicRFB-S)的结构是将 branch 调节为 4 个分支,而 5×5 的卷积被分摊在其他两个 branch 支路上,这样做,虽然有效地提升了特征提取的能力,但是由于增加了分支,增加了网络的深度与模型复杂度。于是,接着对原始的 BasicRFP 模块进行了结构优化,对于原先的 branch3 中 5×5 的 BasicConv 将其换成了 branch1 中的 3×3 的 BasicConv 并更改 branch3 中的卷积为 3×3 的 BasicConv,这样得到改进版本 BasicRFBC 模块。这样在减少参数量的同时可以保持模型的深度,并可以提高模型的表达能力和性能。最后,为了进一步提升模型提取不同缺陷特征的能力,在基本保持计算量不变的

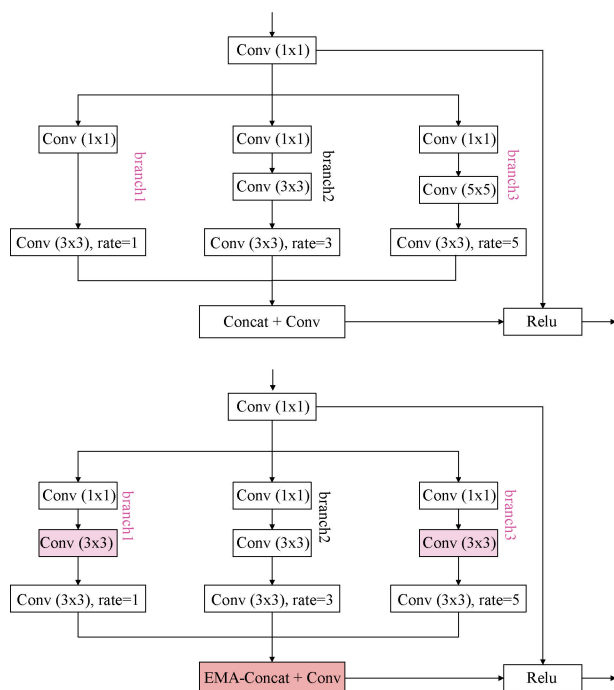


图 2 改进的 RFP 模块的结构参照图

情况下,使用 EMA 注意力机制添加在 BasicRFBC 模块的特征堆叠模块后面,从而实现更丰富的特征聚合。

为了验证 EMA-BasicRFBC 模块的效果,将图 2 中两个模块和文献[20]中的复现出 BasicRFP-S 模块分别加入到原始模型,并在本文构建的焊接件表面缺陷数据集下进行实验,每组实验的数据为 3 次实验获得的结果取平均值。由于图 2 中的 5 个模块的计算量同等级,主要观察 mAP 指标来选择最好的模块。实验指标如表 1 所示。

表 1 优化 SPP 模块

SPP 模块改进	mAP/ %	参数量/ M	计算量 (GFLOPs)	检测 速度/fps
Baseline	81.82	6.01	13.1	119.04
BasicRFP	82.54	6.35	13.5	108.34
BasicRFP-S	82.98	6.49	13.6	122.26
BasicRFPC	82.97	6.11	13.2	122.35
EMA-BasicRFPC	83.67	6.12	13.2	123.58

实验数据表明,EMA-BasicRFBC 模块可以在几乎不改变模型大小的同时明显提高了模型的检测性能。

1.2 ThinNeck 结构

1)改进 VoVGSCSP 模块

SlimNeck 是文献[21]提出的一个轻量级颈部结构,它是由分组空间瓶颈 (grouped spatial bottleneck, GSBottleneck) 组成的,其设计结构完全可以应用于 YOLOV7-tiny 的 HEAD 前部位 (Neck 特征融合部分)。运用 SlimNeck 结构可以帮助信息在网络中更快地传播,减少参数数量的同时提高网络的训练效果。

原始 ThinNeck 结构是基于 SlimNeck 网络结构进行的改进。首先,对于 SLimNeck 网络,在 GSBottleneck 上进行了改进,设计了更轻量的分组空间瓶颈结构 (grouped spatial bottleneck cheaper, GSBottleneckc),如图 3(c)所示,GSBottleneckc 结构是将 shortcut 卷积层改用全局上下文感知卷积(global context-aware convolution, GWConv),其可降低参数量的大小,这样改进后,VoVGSCSP 模块通过调用 GSBottleneckc 模块即可改进提升为 VoVGSCSPC 结构,在保持性能的同时,可实现更轻量化的操作。VoVGSCSP 模块的结构如图 3(b)所示。

同时,为了避免特征融合阶段的特征丢失以及网络过深导致模型检测速度变慢问题,在最后一层卷积操作后再加上 EMA 注意力模块,来实现跨空间学习。改进的 EMA-VoVGSCSPC 的结构如图 3(d)所示。

为了验证改进版 EMA-VoVGSCSPC 的效果,本文展开了消融实验,其实验指标如表 2 所示。

由表 2 可知,EMA-VoVGSCSPC 模块的改进对于整个模型的参数数量的调节减少效果是最佳的,比原先的 MCB 模块降低了 8.4%。其精度较于比原先 MCB 模块提高了

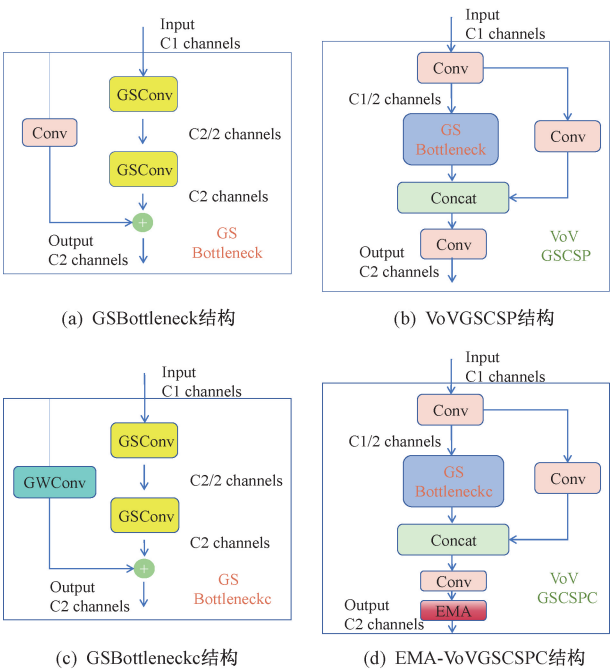


图 3 模块改进对比图

表 2 优化 SlimNeck 模块

SlimNeck 模块改进	mAP/ %	参数量/ M	计算量 (GFLOPs)	检测 速度/fps
Baseline(MCB)	81.82	6.01	13.1	119.04
VoVGSCSP	82.41	5.59	12.2	118.56
VoVGSCSPC	82.23	5.50	12.0	118.56
EMA-VoVGSCSPC	83.37	5.50	12.0	123.77

1.9%。可见 EMA-VoVGSCSPC 模块可以较高地提升模型的性能。

在 VoVGSCSP 模块中存在着轻量级卷积 GSConv。通常情况下,用普通的卷积神经网络对输入数据进行训练,在训练完成后,会产生很多冗余的特征图。为了在输出相同大小的特征图的情况下减少计算量和参数数量,而 GSConv 模块采用了一种新的卷积方式。这种方式不仅有效地减少了计算量和参数数量,而且还能保持原始特征的语义信息,从而确保准确性。GSConv 原理如图 4 所示。

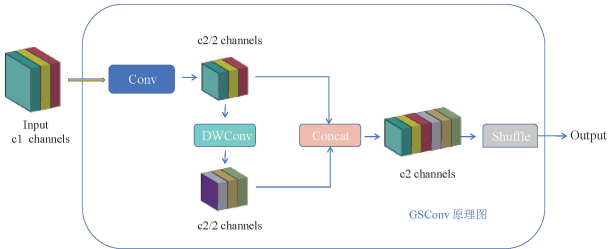


图 4 GSConv 原理图

鉴于上述 GSConv 的优势主要体现在其轻量级的检测

器,这些检测器通过插入深度可分离卷积层(depthwise separable convolution, DSC)和 Shuffle 来提升非线性表达能力。然而,如果在模型的全部阶段都采用 GSConv,模型的网络层会因此变得更深,这将增大数据流的阻力并且明显增长推理时间。当这些特征图抵达瓶颈阶段(此模型 Head 的前部),他们已经足够精细(通道维度已经达到最大,而宽度和高度维度已经缩小到最小),此时的转换已经变得适中。因此,本文在 Neck 阶段使用 GSConv 更为合适,相比于常规的 CBL 卷积,它在保持准确性的同时具有更少的参数,并且可以有效优化模型的深度和推理速度。本文已在表 3、4 中给出该实验, SlimNeck 模块和 GSConv 的组合效果。

2)上采样算子 CARAFE

此外,本文中在结合改进 SlimNeck 与 GSConv 的基础上,使用了 CARAFE 上采样算子对网络的 Neck 部位的上采样进行了优化,CARAFE 上采样算子是文献[22]提出来的,CARAFE 通过学习像素级的重组方式,在上采样过程中能够保持更高的缺陷细节保真度,并减少参数的数量。本文通过调节 CARAFE 的 k_encoder、k_up 来提升性能,并在参数变化可接收的范围内选择以下协同系数来进行对比实验。CARAFE 参数调节后的指标如表 3 所示。

表 3 展示了在引入改进版 SlimNeck 与 GSConv 结合之后,再进行 CARAFE 调整实验的一个示例表格。(S+G)代表 SlimNeck 与 GSConv 结合。实验中,选择了 CARAFE 算子以优化上采样过程,其主要目的在于提升模型的准确性,使模型能更精准地识别不同缺陷之间的细微

表 3 CARAFE 参数调节表

组别	CARAFE k_encoder	CARAFE k_up	mAP/ %	参数量/ M	计算量 (GFLOPs)
1	1	3	80.76	4.54	10.3
2	1	5	81.73	4.54	10.3
3	1	7	81.44	4.56	10.4
4	3	3	82.61	4.57	10.4
5	3	5	84.78	4.64	10.5
6	3	7	83.58	4.76	10.8
(S+G)	—	—	82.48	4.54	10.4
Baseline	—	—	81.82	6.01	13.1

差别。在追求精度提升的同时,也要确保模型的参数保持在可接受的范围内。因此,实验中只选择了 6 对协同系数。实验结果显示 CARAFE 运算符与优化版 SlimNeck 结构以及 GSConv 模块具有协同效应。将三模块叠加最好的调参后的 mAP 效果较于 SlimNeck 与 GSConv 结合的 mAP 提升了 2.8%,而参数量仅仅只提升了 0.1 M。实验表明,CARAFE 算子能进一步增强模型的准确性。

最终,通过以上模块的组合得到最后设计的 ThinNeck 结构(改进的 Slimneck+CARAFE+GSConv)。为了验证最终 ThinNeck 结构的效果,分别将 3 个结构(ThinNeck 的结构)加入到该模型在本文构建的焊接件表面数据集下进行消融实验。实验结果如表 4 所示(此时 CARAFE 已选出最优的)。

表 4 ThinNeck 结构消融实验

组别	改进 Slimneck	CARAFE	GSConv	mAP/ %	参数量/ M	计算量 (GFLOPs)
Baseline				81.82	6.01	13.1
1	✓			83.37	5.50	12.0
2		✓		82.45	5.90	13.3
3			✓	82.96	4.78	11.3
4	✓	✓		84.54	5.63	12.3
5	✓		✓	82.48	4.54	10.4
6		✓	✓	83.58	4.92	11.5
7	✓	✓	✓	84.78	4.64	10.5

表 4 中行 2 中的 CARAFE 选取的是基于原 YOLOV7-tiny 网络时调参最好的测试效果,行 4~7 的 CARAFE 是基于模块组合的调参实验最好的测试效果。行 5 中改进的 Slimneck 与 GSConv 模块的结合较于行 1 中仅改进的 Slimneck,在 mAP 仅损失 1.0%的情况下,它的参数量降低了 16%,计算量降低了 13%。而行 7 的 ThinNeck 结构比仅改进的 Slimneck,不仅 mAP 提升了 1.7%,其参数量也降低了 15%,计算量降低了 13%。综

上,应用 ThinNeck 模块可以同时减少模型的参数量并提升其性能。并且 ThinNeck 内部的模块之间不存在互斥的现象,有良好的协同作用。

1.3 锚框的优化

考虑到文中的焊接件表面缺陷数据集具有不同的尺度缺陷,而原有的 YOLOV7 的预设锚框,不适用于文中数据集。而 K-means++ 算法^[23]可以很好地去根据数据集的尺度特征,将其 anchors 分类来达到一个不错的检测效

果。K-means++算法的主要思想是改进 K-means 算法的初始聚类中心的选择方法。具体而言,K-means++算法首先随机选择一个样本点作为第一个聚类中心,然后计算每个样本点与最近的聚类中心之间的距离,并将每个样本点归类到距离最近的聚类中心所对应的类别中。接下来,K-means++算法对每个样本点进行概率分析,选择具有较大概率值的样本点作为下一个聚类中心。通过这种方式,K-means++算法能够选择更优的初始聚类中心,使得聚类结果更加准确和稳定。其概率计算公式为:

$$p = \frac{D(x)^2}{\sum_{i=1}^n D(x_i)^2} \quad (2)$$

其中, $D(x)$ 表示每个采样点与当前中心之间的最小距离, p 表示数据点 x 被选为新中心的概率。在每次迭代中,选择一个目标,然后根据已存在的聚类目标重新计算,直到没有目标被重新分配到其他聚类中。最终,筛选出 K 个最佳的聚类中心。根据数据集,本文使用 K-means++聚类算法得到了一组先验锚框的宽度和高度,它们的尺寸为(27,60)、(33,37)、(52,97)、(70,176)、(80,353)、(113,606)、(123,248)、(313,50)、(635,46)。与 YOLOV7 默认的先验框相比,这些先验锚框之间的差异更大。因此,使用这些锚框进行网络训练时,其收敛速度更快,并且能够更好地定位和分类缺陷。

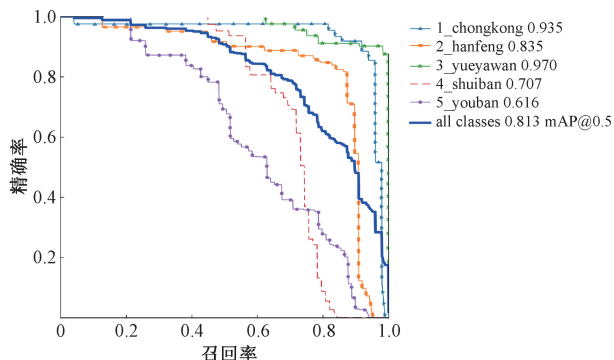
在此基础上,使用了 FEIOU^[24] 对锚点进行了进一步的优化。FEIOU 损失函数综合考虑了重叠损失、中心距离损失和宽高损失。其具有更快的收敛速度和模型的训练速度。同时,它引入了 Focal Loss 来解决目标检测中的样本不平衡问题,使模型更加专注于高质量的锚框。其 $LoSS_{Focal-EIOU}$ 的损失公式如式(3)所示。

$$\begin{cases} Loss_{EIOU} = 1 - IOU + \frac{\rho^2(b, b^{gt})}{C^2} + \frac{\rho^2(w, w^{gt})}{C_w^2} + \frac{\rho^2(h, h^{gt})}{C_h^2} \\ Loss_{Focal-EIOU} = IOU^\gamma \times Loss_{EIOU} \end{cases} \quad (3)$$

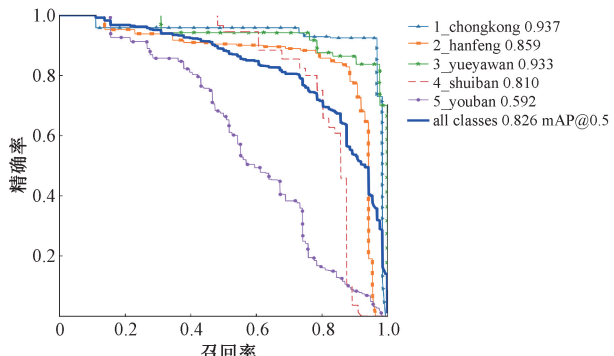
式中: C_w 与 C_h 分别代表最小外接矩形框的宽度和高度, IOU 代表交并比, $\rho^2(b, b^{gt})$, $\rho^2(w, w^{gt})$, $\rho^2(h, h^{gt})$ 分别代表中心点 b 与 b^{gt} 的欧式距离的平方,预测边界框的宽度 w, w^{gt} 以及高度 h, h^{gt} 的差的平方和。其中 gt 代表真实标注, γ 为控制异常值抑制程度的参数。

为了验证锚框优化的效果,分别对初始 YOLOV7-tiny 模型与锚框优化后的 YOLOV7-tiny 模型进行训练。训练后的模型性能对比图如图 5 所示。

由图 5 可以看出,锚框优化后,整体的 mAP 的值进一步上升,其中部分缺陷指标也稳固提升。同时锚框优化后,模型 PR 曲线下的面积也变得更大,面积越大,模型的性能越好。实验表明,同时使用 K-means++和 FEIOU 损失函数对 anchors 优化可以提高模型的性能。



(a) 锚框优化前



(b) 锚框优化后

图 5 锚框优化对比图

2 改进模型的消融实验分析

为验证本文所提出的 3 种改进算法的有效性,本文将使用上述改进方法在扩充增强后的焊接件表面缺陷数据集进行训练以及测试,均基于以上实验环境进行训练,每个方法进行 3 次重复实验,严格控制变量。最终数据为 3 次实验获得的数据的均值。实验结果如表 5 所示。

由表 5 第 2 组实验可知,引入 EMA-BasicRFB 结构来代替原先的 SPP 模块,将模型的 mAP 从 81.82% 提升到了 83.67%,提升了 2.26%。由表 5 第 3 组实验可知,在 YOLOV7-tiny 模型上将原先的 NECK 模块改为基于改进版本的 SlimNeck 设计的 ThinNeck 结构,模型的 mAP 从 81.82% 提升到了 84.78%,提升了 3.61%,并且模型的参数量由原来的 6.01 MB 降低到了 4.64 MB,减少了 22.79%的参数量。计算量由原来的 13.1 GFLOPs 减少到 10.5 GFLOPs,减少了 19.84%的计算量。表明本文提出的 ThinNeck 结构相对模型中原有 NECK 模块可以在显著降低模型的参数量以及计算量的同时有效提高模型检测精度。表 5 的第 4 组实验结合锚框实验可知引入锚框的优化(K-means++和 FEIOU)来进行先验框的聚类加快模型收敛速度,使得模型 FPS 检测速率提升了 18.13%,同时模型的 mAP 从 81.82% 提升到了 82.73%,提升了 1.11%。以上 3 组实验表明本文提出的 3 种改进方法在保持模型参数在可接受的范围内,或者在降低模型复杂性的

表 5 消融实验结果对比

组别	EMA- BasicRFB	ThinNeck	Anchors 优化	mAP/ %	参数量/ M	计算量 (GFLOPs)
1				81.82	6.01	13.1
2	✓			83.67	6.12	13.2
3		✓		84.78	4.64	10.5
4			✓	82.73	6.02	13.1
5	✓	✓		86.59	5.34	11.2
6	✓		✓	84.31	6.46	13.6
7		✓	✓	85.23	4.65	10.5
8	✓	✓	✓	87.64	5.34	11.1

同时,能有效地提高模型的平均精度。

由表 5 第 5、6、7 组实验,可以得出结论,这 3 个改进方法并不会相互抑制,而是可以相互补充,共同提升模型的检测性能。此外,由表 5 第 8 组实验可知,3 个改进方法均采用之后,既本文提出的 KThin-YOLOV7 模型的 mAP 最高,相对 YOLOV7-tiny 模型提升了 7.11%,达到 87.64%,同时模型的参数量和计算量分别下降了 11.14% 和 15.26%。这说明,本文提出的 KThin-YOLOV7 相对原始 YOLOV7-tiny 模型能够以更低的成本获得更高的检测精度。

3 实验分析

3.1 数据增强

本文搜集来自 AI Studio 网络平台上公开的焊接件表面缺陷数据集,其数据集中某些缺陷种类的数量过于稀少,而且人工标注困难,故只选择常见缺陷包括油斑、水斑、月牙弯、焊缝和冲孔等 5 类,其图像分辨率为 $2\,048 \times 1\,000$,图像格式为 jpg,用 python 代码进一步去除空标签后,初步的缺陷图片一共为 1 254 张。由于样本数量太小会造成训练过拟合,并考虑到其数据集的图片的亮度、对比度不太明显。于是,本文使用 python 的 Pillow(PIL)库,对标注好的数据集图片进行翻转、对比度调节、多尺度 Top-hat 亮度调节,以及使用 MATLAB 代码的 Prewitt 算子来增强扩张数据集,如图 6 所示。Prewitt^[25]算子是一种常用于边缘检测的一阶微分算子,它通过比较像素点的上下、左右邻点的灰度差异,能够检测出边缘的位置并实现极值检测。使用 Prewitt 算子能够抑制噪声的同时,可以更好地提升图像的锐化效果。

最终,缺陷图片扩张到了 4 376 张,这样增加了图片的多样性,并且有助于提升模型的鲁棒性,实验效果也得到了显著的改善。

3.2 实验环境配置

本环境选择的深度学习框架为 1.8.0 版本的 Pytorch,编译语言为 Python 其版本为 3.8.13,cuda 版本为 11.1,Cudnn 版本为 8.0.5,操作系统为 Linux,GPU 为英伟达

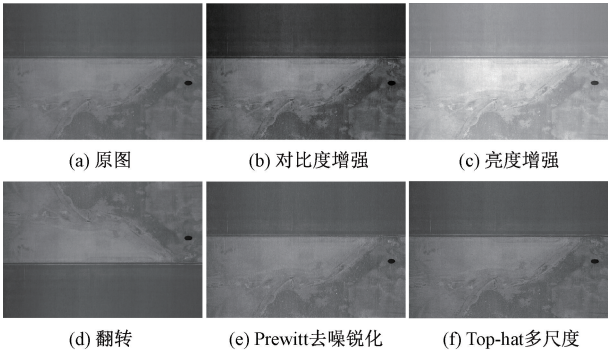


图 6 数据集扩增效果

(NVIDIA) Tesla A100-PCIE-40GB,处理器为 Intel(R) Xeon(R) Gold 6240R@2.40 GHZ,内存为 16 GB。模型训练采用了随机梯度下降(SGD)优化算法,Batch size 设置为 8。初始学习率为 0.01,Momentum 为 0.937,Weight decay 为 0.000 5。训练过程中引入了 mosaic 增强,Epochs 设定为 300,输入图片尺寸设置为 640×640 。

3.3 模型评价指标

本文采用模型参数量(Params)以及计算量(FLOPs)来评价 KThin-YOLOV7 模型大小以及复杂度。使用平均精度均值(mean average precision, mAP)以及每秒传输帧数(frames per second, FPS)来评价模型的性能。其评价指标公式分别如下:

$$Params = C_o \times (k^2 \times C_i + 1) \tag{4}$$

$$FLOPs = C_i \times k^2 \times C_o \times W \times H \tag{5}$$

$$P = \frac{TP}{TP + FP} \tag{6}$$

$$AP = \int_0^1 P(R) dR \tag{7}$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \tag{8}$$

式(4)、(5)中 C_o 表示输出通道数, C_i 表示输入通道数, k 表示卷积核大小, W 和 H 表示特征图的宽和高。

式(6)~(8)中 P 是精确率, $P(R)$ 是由精确率和召回率构成的曲线; TP 是正确预测为正的实例; FP 是误判为

正 的 负 实 例。 AP 表 示 某 个 类 别 的 准 确 率， N 是 类 别 个 数。

3.4 实验结果与分析

为了直观展示本文改进方法对模型优化的效果，以及分析三种方法对不同缺陷检测的影响，如图 7 所示，展现了它们对各类缺陷检测效果的对比。

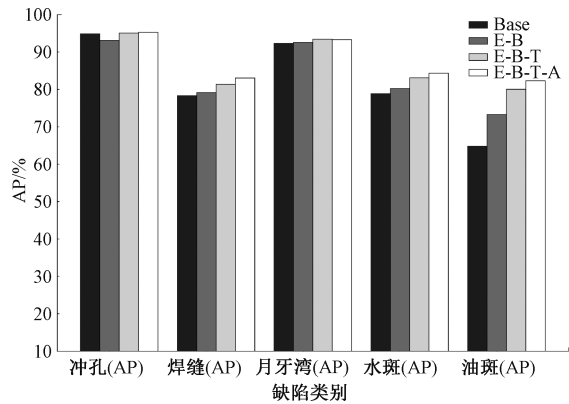


图 7 缺陷的检测效果

如图 7 可知，相对原始模型 YOLOv7-tiny 中的 SPP 模块，当引入 EMA-BasicRFB 模块后，模型检测对水斑与油斑这两种特征相似的小目标缺陷的检测精度提升最明显。在引入 EMA-BasicRFB 模块的基础上引入设计的 ThinNeck 结构之后模型对各类缺陷检测的精度均有提升，最后通过优化 anchors 使得再次优化了模型对缺陷的检测性能，FPS 检测速度也有显著的提高。

同时，为了验证本文提出的 KThin-YOLOV7 模型的

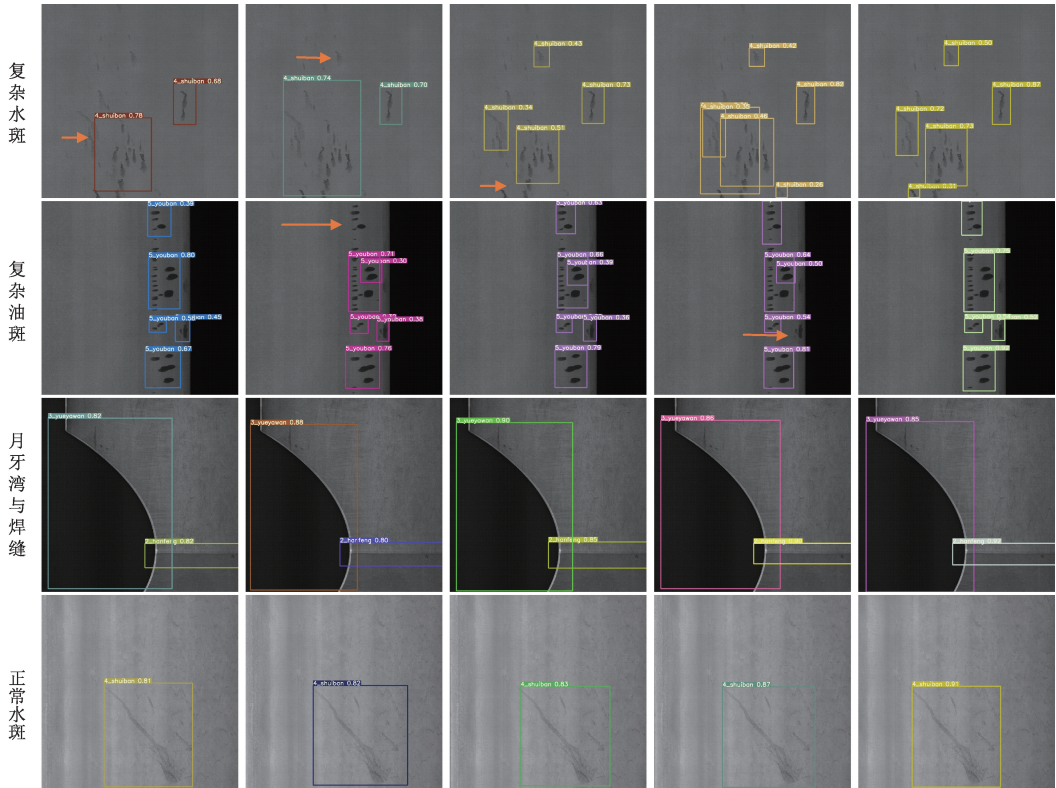
性能以及实用性，选取了 YOLOv5-S、YOLOv7-tiny、文献[17]中的 PC-YOLOV7，以及文献[21]中的 SlimNeck-tiny，5 种模型在保持同一数据集，固定超参数在本文 3.2 节实验环境下进行训练，采用 mAP(平均精度均值)、FPS (检测速度)、参数量(Params)作为评价指标来评价模型的性能以及实用性。实验指标如表 6 所示。

表 6 不同改进方法各缺陷检测精度

实验模型	mAP/%	参数量/M	检测速度/fps
YOLOv5-S	80.65	7.02	88.13
YOLOv7-tiny	81.82	6.01	119.04
PC-YOLOV7	85.93	5.97	113.65
SlimNeck-tiny	86.14	5.59	131.64
KThin-YOLOV7	87.64	5.34	138.56

表中的数据是 3 次试验结果的平均值。如表 6 所示，本研究提出的算法在检测精度和检测速度上均表现出色，同时模型复杂度亦不高。因此，可以看出，本研究提出的 KThin-YOLOV7 在一定程度上具有优越性。

为了能更直观地看出主流算法的检测效果。对主要模型 a YOLOv5-S、模型 b YOLOv7-tiny、模型 c PC-YOLOV7、模型 d SlimNeck-tiny，以及本文的模型 e KThin-YOLOV7 的检测效果进行可视化分析。如图 8 所示，主要选取了缺陷较为繁杂的复合图以及常见的典型缺陷图，用来分别看各个模型的效果。



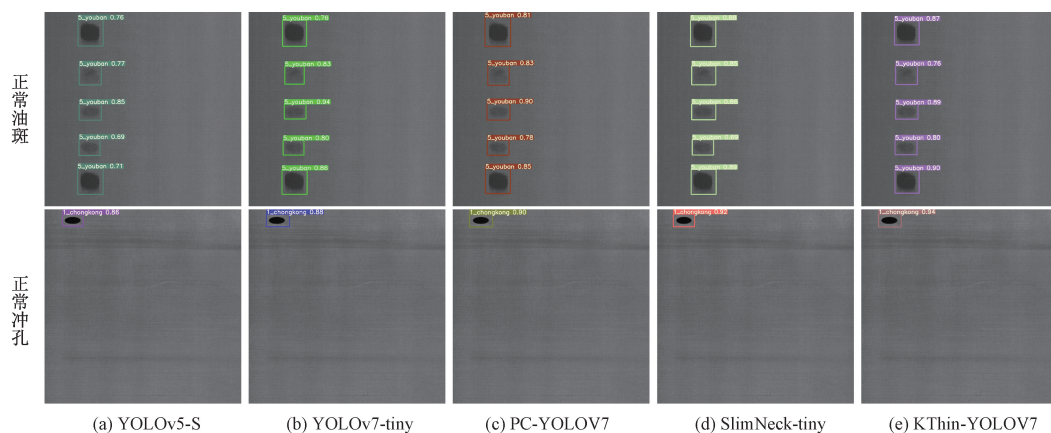


图 8 5 种主要检测算法模型对比图

由图 8 可以看出 KThin-YOLOV7 相比于几个轻量级网络,模型的检测精度更高,在微小缺陷的检测效果上都有很好的提升。在复杂油斑的模型表现上,图 8(b)、(d)的模型在红色箭头标出漏检情况,而图 8(e)的模型并未出现漏检。同样在复杂水斑上,图 8(a)~(c)的模型也出现了漏检的情况,图 8(e)模型不仅没漏检还具有更高的检测精度。同样,图中在正常缺陷情况下,图 8(e)模型的精度检测效果明显优于其他主流算法模型。可见最终 3 个改进方法都使用之后,即本文提出的 KThin-YOLOV7 模型在工业检测场景下对焊接件表面缺陷检测性能拥有最佳的检测性能。

4 结 论

文中提出了一种基于改进 YOLOV7-tiny 的焊接件表面缺陷检测方法,设计了基于改进的 SlimNeck 轻量级网络替换 YOLOV7-tiny 的 HEAD 前部分(Neck 结构),能够在提升检测精度的同时有效减小模型尺寸;通过替换原有的 SPP 模块为改进的 EMA-BasicRFBC 模块,能够提高模型的感受野,从而加强了网络的特征提取的能力,而且避免了灰度背景的干扰;通过应用 K-means++ 算法,对焊接件表面缺陷数据集进行了有力的聚类分析,从而确定了适合的宽度和高度的先验框,这种方法有效地提升了模型的检测准确性。在这个基础上,加入了 FEIOU 损失函数,这个函数进一步优化了锚点,其宽高损失直接将预测框和真实框之间的宽度和高度差异降至最低,从而实现了更快的收敛速度。

实验表明,改进后的 KThin-YOLOV7 模型比 YOLOV7-tiny 模型平均精度提升了 7.11%,模型大小缩减了 11.14%,检测速度提高了 138.56 fps。相较于主流的目标检测算法,文中提出的方法在缩小模型体积的同时,也提高了精度和检测效率,为实际工业环境中的焊接部件表面缺陷检测带来了新的可能性。未来的工作将涉及收集更多种类的焊接部件缺陷(如压痕、凹坑等),并通过数据预处理和数据增强来训练和优化模型,以进一步提升模

型的泛化性能。这将对实际工业环境中的焊接部件生产过程具有重要的实际价值。

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.
- [2] 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.
- [3] CHU M, LIU X, GONG R, et al. Multi-class classification method using twin support vectormachines with multi-information for steel surface defects [J]. Chemometrics and Intelligent Laboratory Systems, 2018, 176: 108-118.
- [4] 段岳飞,马嵩华,胡天亮.基于全卷积神经网络的焊缝识别方法[J]. 制造技术与机床,2023(4): 44-49.
- [5] SHANG L, YANG Q, WANG J, et al. Detection of rail surface defects based on CNN image recognition and classification[C]. 2018 20th International Conference on Advanced Communication Technology (ICACT), IEEE, 2018.
- [6] CHEN J, LIU Z, WANG H, et al. Automatic defectdetection of fasteners on the catenary supportdevice using deep convolutional neural network [J]. IEEE Transactions on Instrumentation and Measurement, 2017, 67(2): 257-269.
- [7] HE K, GKIOXARI G, DOLLÁR P, et al. Mask r-cnn [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961-2969.
- [8] HE Y, SONG K, MENG Q, et al. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features[J]. IEEE Transactions on Instrumentation and Measurement, 2019, 69(4): 1493-1504.
- [9] 卢进南,刘扬,王连捷,等.基于改进 YOLOX 的电铲铲齿断裂检测方法[J]. 电子测量与仪器学报,2023,

- 37(5):46-57.
- [10] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[J]. ArXiv Preprint, 2020, ArXiv:2004.10934.
- [11] 肖粲俊, 潘睿志, 李超, 等. 基于改进 YOLOv5s 绝缘子缺陷检测技术研究[J]. 电子测量技术, 2022, 45(24): 137-144.
- [12] HAN L, LI F, YU H, et al. BiRPN-YOLOvX: A weighted bidirectional recursive feature pyramid algorithm for lung nodule detection[J]. Journal of X-Ray Science and Technology, 2023, 31(2): 1-17.
- [13] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Singshot multibox detector[C]. Computer Vision-ECCV 2016; 14th European Conference, Amsterdam, The Netherlands, 2016.
- [14] ZHANG C, CHANG C, JAMSHIDI M. Concrete bridge surface damage detection using a single-stage detector[J]. Computer-Aided Civil and Infrastructure Engineering, 2020, 35(4): 389-409.
- [15] 霍超, 谷晓钢, 黄玲琴, 等. 基于 YOLOv4 优化的航拍绝缘子缺陷图像检测模型[J]. 电子测量技术, 2023, 46(9): 175-181.
- [16] 杜娟, 崔少华, 晋美娟, 等. 改进 YOLOv7 的复杂道路场景目标检测算法[J]. 计算机工程与应用, 2024, 60(1): 96-103.
- [17] 赵春华, 罗顺, 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.
- [18] BODIN H A B, KRAKOWSKI R A, ORTOLANI S. The reversed-field pinch: From experiment to reactor [J]. Fusion Technology, 1986, 10(31): 307-353.
- [19] OUYANG D, HE S, ZHANG G, et al. Efficient multi-scale attention module with cross-spatial learning[C]. ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2023.
- [20] 史先鹏, 王宏勃. 基于 YOLOv4 改进的轻量级水下目标检测网络[J]. 哈尔滨工程大学学报, 2023, 44(1): 154-160.
- [21] XIE T H, QI X Y, HE P, et al. BaDExpert: Extracting backdoor functionality for accurate backdoor input detection[J]. ArXiv Preprint, 2023, ArXiv:2308.12439.
- [22] BHOWMIK N, GAUS Y F A, BRECKON T P. On the impact of using x-ray energy response imagery for object detection via convolutional neural networks[C]. 2021 IEEE International Conference on Image Processing (ICIP), IEEE, 2021: 1224-1228.
- [23] 程婧怡, 段先华, 朱伟. 改进 YOLOv3 的金属表面缺陷检测研究 [J]. 计算机工程与应用, 2021, 57(19): 252-258.
- [24] 王昱, 姚兴智, 李斌, 等. 基于改进 YOLOv7-tiny 的甜椒畸形果识别算法[J]. 农业机械学报, 2023, 54(11): 236-246.
- [25] AHMAD Z, KHAN N. Inertial sensor data to image encoding for human action recognition[J]. IEEE Sensors Journal, 2021, 21(9): 10978-10988.

作者简介

卢开喜, 硕士, 主要研究方向为目标检测与图像分割。

E-mail: 1632875540@qq.com

段先华, 博士, 教授, 主要研究方向为计算机视觉、模式识别、数据融合等。

陶宇诚, 硕士, 主要研究方向为计算机视觉、目标检测。

倪东海, 硕士, 主要研究方向为目标检测与图像识别。