

深度学习中单阶段金属表面缺陷检测算法优化综述

董甲东⁺, 郭庆虎, 陈琳, 桑飞虎

安庆师范大学 电子工程与智能制造学院, 安徽 安庆 246011

⁺ 通信作者 E-mail: djd@aqnu.edu.cn

摘要: 金属表面的划痕、凹坑、波纹等缺陷会直接影响产品的质量。传统的检测方法耗时耗力, 准确性受限于操作人员的经验和技能。近年来, 深度学习技术在图像识别领域的突破性进展为金属表面缺陷检测提供了新的解决方案, 基于深度学习的金属表面缺陷检测方法在检测精度和速度方面取得了显著成效。为了便于金属表面缺陷检测算法的研究, 综合分析了单阶段深度学习算法在金属表面缺陷检测中的优化方法及应用。介绍了目前常用的金属表面缺陷数据集和算法评价指标; 总结了目标检测算法的发展史以及单阶段目标检测算法的基本概念和典型模型; 从数据增强、特征的提取与融合、锚框优化三个方面, 对比总结了不同算法不同优化方式的优缺点, 并研究了金属表面缺陷检测算法的轻量化; 从多模态融合、大数据应用技术、现实与虚拟结合三个方面对金属表面缺陷检测算法的未来研究方向进行了展望。

关键词: 金属表面缺陷检测; 深度学习; 单阶段目标检测算法; 模型优化

文献标志码: A **中图分类号:** TP391.41 **doi:** 10.3778/j.issn.1002-8331.2408-0098

Review on Optimization Algorithms for One-Stage Metal Surface Defect Detection in Deep Learning

DONG Jiadong⁺, GUO Qinghu, CHEN Lin, SANG Feihu

School of Electronic Engineering and Intelligent Manufacturing, Anqing Normal University, Anqing, Anhui 246011, China

Abstract: Scratches, pits, ripples and other defects on the metal surface will directly affect the quality of the product. Traditional detection methods are time consuming, and the accuracy is limited by the operator's experience and skills. In recent years, breakthroughs of deep learning technology in the field of image recognition have provided new solutions for metal surface defect detection, and the deep learning-based metal surface defect detection method have achieved remarkable results in terms of detection accuracy and speed. In order to facilitate the research of metal surface defect detection algorithm, the optimization method and application of one-stage deep learning algorithm in metal surface defect detection are comprehensively analyzed. The commonly used metal surface defect datasets and algorithm evaluation indexes are introduced. The development history of object detection algorithms, the basic concepts and typical models of one-stage object detection algorithms are summarized. From three aspects of data enhancement, feature extraction and fusion, anchor frame optimization, the advantages and disadvantages of different algorithms and different optimization methods are compared and summarized, and the light weight of metal surface defect detection algorithm is also studied. The future research direction of metal surface defect detection algorithm is prospected from three aspects: multi-mode fusion, big data application technology, reality and virtual combination.

Key words: metal surface defect detection; deep learning; one-stage target detection algorithm; model optimization

在复杂的工业生产过程中, 受生产设备、加工工艺、生产原料等因素的影响, 可能会造成金属表面产生划痕、裂缝、脏点等缺陷, 如图1所示。金属表面存在缺陷不仅影响产品外观, 损害材料的机械性能, 缩短制成品

的使用寿命, 还会增加企业的生产成本, 造成不必要的资源浪费^[1]。因此, 检测金属表面缺陷是企业在不影响生产的前提下提高产品质量的核心能力。与人工检测相比, 自动检测缺陷技术具有明显的优势, 不仅可以适

基金项目: 国家自然科学基金(62205005); 安徽省高校科研计划重大项目(2024AH040174)。

作者简介: 董甲东(1968—), 男, 博士, 教授, 研究方向为智能制造、工业互联网; 郭庆虎(1999—), 男, 硕士研究生, 研究方向为机器视觉、缺陷检测; 陈琳(1999—), 女, 硕士研究生, 研究方向为工业互联网、边缘计算; 桑飞虎(1998—), 男, 硕士研究生, 研究方向为机器视觉、目标检测。

收稿日期: 2024-08-07 **修回日期:** 2024-09-22 **文章编号:** 1002-8331(2025)04-0072-18

应恶劣的生产环境,而且可以长期工作,检测精度和效率也比较高。对缺陷检测技术的研究可以降低生产成本,提高生产效率和产品质量,为制造业的智能化转型奠定坚实的基础。

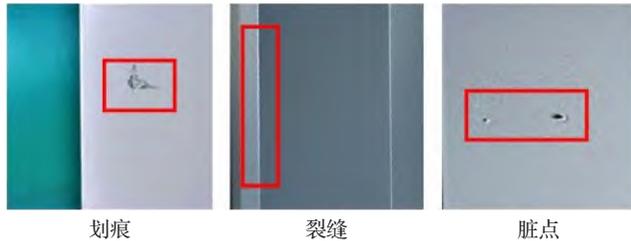


图1 金属表面缺陷

Fig.1 Metal surface defects

近年来,众多学者使用机器视觉和图像处理技术检测图片或视频中的目标,通过识别图像中目标的类型,确定目标在图像中的位置,从而实现不同目标的类别判断和定位。然而传统的目标检测方法容易受到检测环境的影响,在特定的检测背景下可以取得一定的效果。但在实际的生产环境中,光照、摆放角度等客观因素会影响检测的准确性和速度。另外,传统的目标检测方法依赖于人工设计的特征,需要大量的先验知识和繁琐的步骤执行检测,这一过程不仅复杂而且计算量大,不满足实时检测的要求。而金属表面缺陷通常在生产线上进行检测,便于处理不良品,实时性要求较高,传统的目标检测算法很难满足需求。

随着深度学习的快速发展,目标检测技术取得了显著进步。深度学习通过模拟人脑处理信息的方式,能够自主学习和提取图像中的特征,从而实现对目标的高效识别和定位。这种检测方法不仅提高了算法的准确性和鲁棒性,还简化了检测过程,减少了对先验知识的依赖。使用深度学习算法检测目标具有诸多优势。第一,深度学习算法通过在大规模数据集上进行训练,能够学习到更加复杂的特征,从而提高目标检测的准确率。如Ling等^[2]收集制作了一种大规模场景数据集DL3DV-10K,为深度学习探索三维空间提供了重要数据。第二,深度学习算法可通过调整网络结构和使用不同的数据集,以适应不同的场景,减少外部因素对检测结果的影响。Yang等^[3]就是通过调整RT-DETR(real-time transformer-based detector)网络结构,设计了一种具有双融合骨干网和加权多尺度双向特征混合融合金字塔网络(bidirectional feature hybrid fusion pyramid network, BdfPN)。然后将该算法应用到自制的航空电子焊点数据集上,平均正确率均值达到了92.7%,优于原算法。第三,利用深度学习进行特定目标检测时,具有很强的扩展性。这表明可以根据不同的需求和场景对算法网络结构进行改进和扩展,以应用于其他领域。Zheng等^[4-5]在对深度学习算法进行改进后,将算法应用到无人机通信领域,

利用不同的轻量级神经网络完成无人机的实时自动调制分类(automatic modulation classification, AMC),确保AMC准确性的同时提高推理效率。深度学习凭借这些优势在金属表面缺陷检测领域取得了显著成果。伍麟等^[6]从光学成像技术、图像预处理技术与缺陷检测器三个方面综述了金属表面缺陷检测现状以及未来的研究趋势。李宗祐等^[7]以监督方法为主线综述了近年来基于深度学习的金属材料表面缺陷的研究现状,总结了当下金属表面缺陷检测主要问题。另外几篇综述也总结了近几年学者使用深度学习算法检测金属表面缺陷的成果^[8-9]。

这些综述从学习方式、缺陷分类等方面总结了当前金属表面缺陷检测的研究趋势,并没有系统地归纳某一类算法的具体优化方法及应用效果。本文首先介绍了常用的金属表面缺陷数据集和算法的评价指标;随后概述了目标检测算法的发展史及典型模型;接着从数据增强、特征的提取与融合、锚框优化三个方面,对近年来国内外学者的研究进行归纳,对比总结了不同算法不同优化方法的优缺点,研究了金属表面缺陷检测算法的轻量化;最后从多模态融合、大数据应用技术、现实与虚拟结合三个方面对金属表面缺陷检测算法的未来研究方向进行了展望。

1 数据集和评价指标

1.1 数据集

自20世纪80年代初,研究者们开始初步探索如何通过机器学习算法识别和定位图像中的特定对象,目标检测数据集的构建与发展便成为了推动该领域进步的重要基石。

在早期阶段,针对目标检测的研究多依赖于小规模的数据集,通常包含简单的目标和有限的背景复杂度,如手工标注的小规模数据集。这些数据集的局限性制约了模型的应用和推广。随着深度学习的崛起,尤其是卷积神经网络(convolutional neural network, CNN)的广泛应用,目标检测模型开始展现出强大的能力,制作和丰富规模庞大的数据集则成为了训练高效和精确模型的必要条件。

进入2010年以来,像Pascal VOC(visual object classes)^[10]和ImageNet^[11]这样的数据集相继问世,为目标检测提供更为丰富的样本和多样化的目标类别。Pascal VOC数据集不仅提供了包含人物、动物、工具等20类检测对象,还开创了规范的评测标准,推动了目标检测算法的发展。此外,ImageNet数据集的庞大规模与多样性也为目标检测提供了宝贵的训练数据。

随着目标检测技术的不断进步和检测需求的日益增长,COCO(common objects in context)^[12]数据集于2014年推出。COCO数据集的特点在于其不仅涵盖了

多种类别的对象,还关注对象在不同背景中的复杂性,增强了对现实场景的反应能力。COCO的引入不仅提升了模型的泛化能力,也促进了诸如YOLO(you only look once)、SSD(single shot multibox detector)和Faster R-CNN等先进模型不断涌现。

此外,研究者们意识到,单一的数据集无法全面覆盖各种应用场景,因此针对特定领域的数据集应运而生。例如,针对行人检测的Caltech Pedestrian数据集^[13]、针对面部检测的Wider Face数据集^[14]等,这些数据集为目标检测的研究提供了新的机遇和挑战。

为了能够更好地研究金属表面缺陷检测技术,研究者们以钢材、铝材等一些具体的金属产品为对象,制作了一些数据集。目前常用的数据集有NEU-DET数据集(东北大学钢材表面缺陷检测数据集)^[15]和天池铝型材表面缺陷数据集^[16]。

NEU-DET数据集包含训练集、验证集和测试集三部分。这种划分有利于研究者在训练集上进行模型训练,在验证集上进行模型参数的调整 and 选择,最后在测试集上评估模型的性能。数据集包含夹杂、划痕、轧制氧化皮、裂纹、麻点和斑块6种缺陷,每种缺陷300张,图像尺寸为200×200。每个图像文件都包含了对应的xml标注文件,这些标注文件详细记录了图像中缺陷的位置和类别信息。

天池铝型材表面缺陷数据集专门用于铝型材表面瑕疵识别的研究。数据集包含1万张高分辨率铝型材彩色图像,共包含10类缺陷,如凹陷、不导电、擦花等,缺陷特征较为清晰。每张图像中包含一个或多个瑕疵类型,图片尺寸为2560×1920,每张图像附有对应的缺陷类型和标注信息,适合于监督学习。

其他金属表面缺陷数据集的信息及特点,如GC10-DET^[17]、RSDDs^[18]、DeepPCB^[19]等见表1。对于其他的特定研究对象如活塞环、黄铜螺母等金属产品,研究人员需要进行图像收集、筛选和整理,自行制作数据集,供模型检测训练。

1.2 评价指标

目标检测算法的评价指标不仅能够客观反映模型

的性能,在不同的应用场景中,这些指标还可以帮助研究者更好地理解和改进目标检测算法,确保其在实际应用中达到最佳效果。主要包括以下评价指标:

(1) 准确率和召回率

准确率(Precision),又称查准率(式(1)),指在所有实际为正的样本中,被预测为正样本的占比。召回率(Recall),又称查全率(式(2)),指在所有被预测为正的样本中,确实是正样本的占比。

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

式中,TP表示正样本被正确识别的数量;FP表示正样本被识别成负样本,即误报的负样本数量;TN表示负样本被正确识别的数量;FN表示负样本被识别成正样本,即漏报的正样本数量。准确率和查全率越高,模型性能越好。

(2) 交并比

交并比(intersection over union, IoU)表示预测框与真实框之间的重叠度,计算方法见式(3)。其中A、B为两个相互独立的预测检测框,其值越高,表示模型的检测结果与真实结果越接近。

$$IoU = \frac{A \cap B}{A \cup B} \quad (3)$$

(3) F1分数

F1分数(F1-Score)是精确率和召回率的调和平均值,能够在一定程度上平衡这两个指标,使得在关注精确率和召回率之间的权衡时,提供一个单一的评价标准。F1分数介于0-1之间,其值越大,代表模型既能很好地查全,也能很好地查准,计算方法见式(4)。

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

(4) 平均正确率和平均正确率均值

平均正确率(average precision, AP)表示单个类别的平均精确率,计算方法如式(5),其中R表示不同交并比下的召回率。

表1 相关数据集信息

Table 1 Related dataset information

名称	数量/张	对象	特点
GC10-DET	3 570	钢板	包含10类缺陷,缺陷种类多,对模型的分类能力具有挑战性;含有少量无标签图像和错误的标签名称,需提前修正
Severstal	18 074	带钢	包括凹坑、锈蚀、划痕等缺陷,样本数量多,有利于模型训练;由于图像数量多,分辨率较高,且一张图像中含多种类型的缺陷,因此对计算机硬件设备要求较高
RSDDs	195	铁轨	包含从快车轨道和普通/重型运输轨道拍摄的数据,但总体样本数量较少;每幅图像至少包含一个缺陷,这些缺陷已由专业人员进行标记,但图像背景复杂且噪声很大,使用前需对数据进行预处理
KolektorSDD	400	电子换向器	数据集中的图像是在不同背景和光照条件下拍摄的,有利于训练模型的泛化能力;每张图像配有缺陷和类别标签,适合用于监督学习
DeepPCB	1 386	PCB	包含缺失孔、老鼠咬坏、开路、短路、杂散、伪铜6种缺陷,每种瑕疵类型都有相应的标注,并且图像多为高分辨率的RGB格式,细节较为清晰,可用于检测、分类和配准任务

$$AP = \int_0^1 P(R) dR \quad (5)$$

平均正确率均值(mean average precision, mAP)表示多个物体类别的AP的平均值,通常将mAP作为评估模型性能的最终指标。平均正确率越高,表示模型检测缺陷的种类和位置越准确,计算方法如式(6),其中C为类别数目。

$$mAP = \frac{\sum_{i=1}^c AP_i}{C} \quad (6)$$

除了上述评价指标之外,常用评估模型的指标还有每秒传输帧数(frames per second, FPS)和每秒浮点运算次数(floating point operations per second, FLOPS),用于表示模型的检测速度。

2 目标检测算法

2.1 传统目标检测

传统的目标检测算法为后来基于深度学习的方法奠定了理论和实践的基础。传统目标检测算法主要依赖手工提取特征或机器学习进行目标的识别和定位,如支持向量机(support vector machine, SVM)^[20]、决策树^[21]、AdaBoost^[22]等。传统目标检测算法主要包括三个步骤:区域选择、特征提取和目标分类。

在区域选择阶段,传统算法常采用滑动窗口的方法来定位图片中的目标。这种方法通过在图像上以不同的尺度有序滑动固定大小的窗口,并把每个窗口内的图像送入后续的特征提取和分类环节。如图2所示,绿色部分为滑动窗口,大小为4,设定步长为1向右移动,直至检测到红色圆圈目标。由于一幅图像可能产生大量的窗口需要处理,使用这种方法计算量巨大。为了优化这一过程,一些改进的算法如级联分类器^[23]被提出,通过简单的特征快速排除背景区域,减少需要进一步处理的窗口数量。

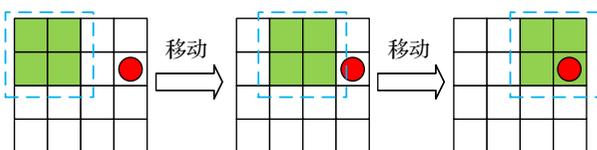


图2 滑动窗口法

Fig.2 Sliding window method

在特征提取阶段,传统的目标检测算法主要依赖于如尺度不变特征变换(scale invariant feature transform, SIFT)^[24]、方向梯度直方图(histogram of oriented gradient, HOG)^[25]、加速鲁棒特征(speeded up robust feature, SURF)^[26]等手工设计的特征描述符。这些特征描述符通过捕捉图像中的梯度、边缘、纹理等关键信息,算法能够对目标进行有效的表征。例如,HOG特征通过描述图像局部区域的梯度强度和方向分布用于区分行人与其他对象,其特征提取流程如图3所示。

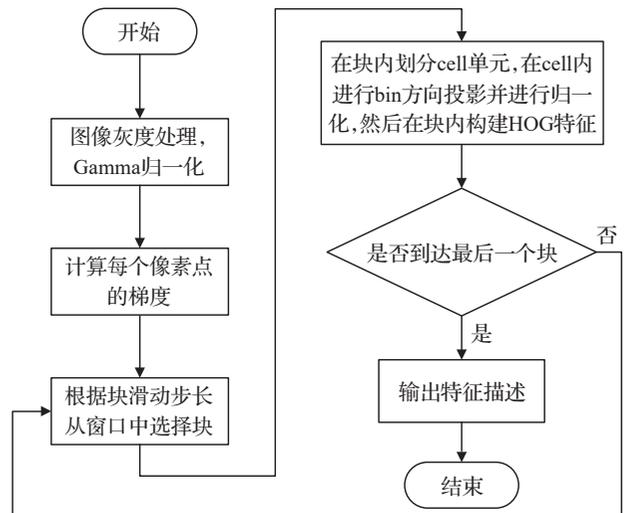


图3 HOG特征提取流程

Fig.3 HOG feature extraction process

在目标分类阶段,传统的目标检测算法通常使用诸如SVM之类的机器学习模型对提取的特征进行分类。SVM是一种强大的分类器,尤其当配合合适的核函数使用时,能够在高维空间中寻找最优的决策边界,有效区分各类目标与背景,其训练过程如图4所示。分类间隔宽度越宽(即最大化),训练集的局部干扰所引起的影

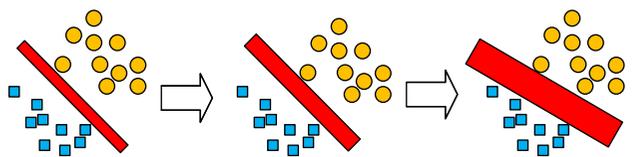


图4 SVM分类器训练示意图

Fig.4 SVM classifier training diagram

此外,集成学习方法如AdaBoost也被广泛应用,通过组合多个弱分类器的决策来提高整体的检测精度。尽管传统的目标检测算法在某些情况下表现出色,但由于其依赖于手工特征提取且泛化能力有限,它们在处理复杂、多变场景下的目标检测任务时,性能往往无法达到现代深度学习算法的水平。

2.2 两阶段目标检测算法

两阶段目标检测算法主要由区域提取和区域分类两个阶段组成。第一阶段是区分前景和背景,会生成有可能包含待检物体的预选框,称作感兴趣区域(region of interest, ROI)。第二阶段是在感兴趣区域的特征图上提取特征,然后再进行分类和预测,得出物体所在位置。

2014年, Girshick等^[27]提出R-CNN算法。该算法首先使用选择性搜索(selective search, SS)算法在图像中生成一系列候选区域,然后将这些区域送入CNN网络进行特征提取,最后使用SVM分类器对每个区域进行分类判断。R-CNN还引入了边框回归技术(bounding

box regression)以提高检测精度。目标检测算法发展历程如图5所示。

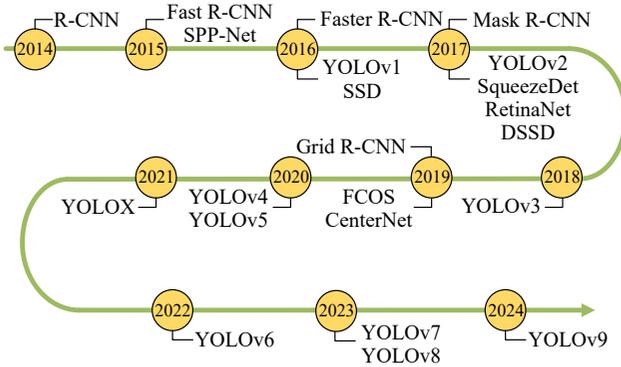


图5 目标检测算法发展时间轴

Fig.5 Object detection algorithm development timeline

2015年,何恺明等^[28]提出空间金字塔池化网络(spatial pyramid pooling network, SPP-Net)算法。该算法引入空间金字塔池化层,把特征图划分为多个不同大小的网格,从而使得网络能够处理任意尺寸的输入图像。相比R-CNN算法, SPP-Net只需对整个图像进行一次卷积计算,避免了多次重复计算,显著提高了目标检测的效率。

同年, Girshick等^[29]在R-CNN和SPP-Net的基础上进一步改进,提出了Fast R-CNN算法,可以实现在网络里同时预测目标的分类概率和位置偏移。随后他又提出Faster R-CNN^[30]算法,第一个端到端的算法。该算法在Fast R-CNN的基础上引入区域提议网络(region proposal network, RPN),该网络直接在卷积特征图上生成候选区域,避免了传统方法中耗时的区域选择步骤。Fast R-CNN网络结构如图6所示。

两阶段检测算法的优势在于高精度的目标检测能力。但也面临着计算量大、实时性差的挑战,尤其是在区域提议阶段的计算成本较高。尽管后来的学者在原有算法的基础上进行改进,提出了新算法,如Mask

R-CNN^[31]、Grid R-CNN^[32],但对于需要快速响应的应用场景,两阶段检测算法仍然不够理想。因此研究人员另辟蹊径,提出单阶段目标检测算法以提高检测速度。

2.3 单阶段目标检测算法

单阶段目标检测算法通常只需一次前向传播就可以完成目标检测。算法通过将图像分割成不同的锚点,对每个锚点进行分类和回归,以确定是否存在目标物体以及物体位置和大小,不需要生成目标候选框,直接输出目标类别和位置信息单阶段满足能够实时检测的任务。目前单阶段目标检测算法主要包括YOLO、SSD和RetinaNet等。

2.3.1 YOLO目标检测算法

2015年, Redmon等^[33]第一次提出YOLO算法。YOLOv1的网络结构主要包括24个卷积层和2个全连接层;检测过程中将图像分成 7×7 的网格,每个网格预测2个边界框及类别概率,输出为 $7 \times 7 \times 30$ 的张量。YOLOv1在检测速度上优于Faster R-CNN,每秒可处理45帧图片,但对于小目标和近似目标的检测效果较差。

次年, Redmon等^[34]引入锚框机制,同时增加多尺度训练和联合训练策略,使用Darknet-19网络结构,提出YOLOv2算法。该算法提升了检测精度,同时保持了较快的检测速度。随后Redmon在YOLOv2的基础上做了一些改进,提出YOLOv3^[35]。特征提取部分采用Darknet-53网络结构;利用特征金字塔网络(feature pyramid network, FPN)^[36]进行多尺度检测,在兼顾实用性的同时保证了目标检测的准确性。

2020年, YOLOv4由Bochkovskiy等^[37]提出。他们在YOLOv3的基础上进行了多项改进,包括使用CSPNet作为主干网络,引入Mish激活函数,使用CIoU_Loss等,提高了小目标的检测效果。YOLOv4发布一个月后, YOLOv5正式提出。YOLOv5优化了网络训练和推理过程,更加注重工业实用性,是研究者在金属表面缺陷检测领域研究较多的算法,其网络结构如图7所示。

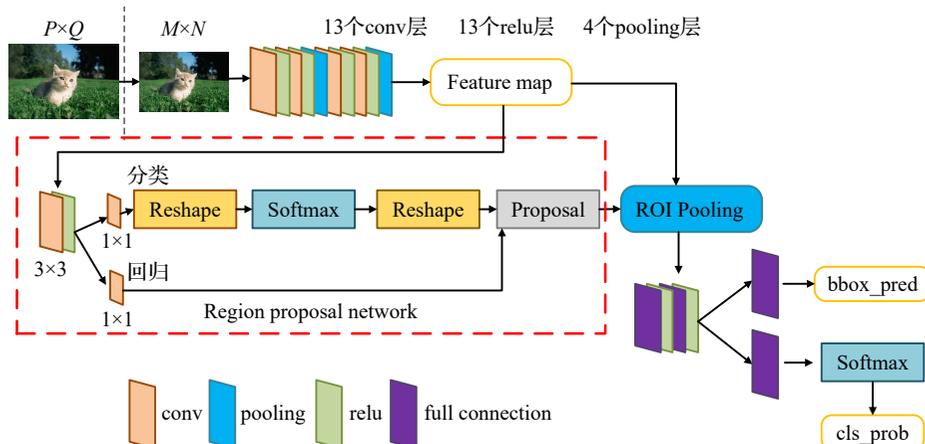


图6 Faster R-CNN网络结构

Fig.6 Faster R-CNN network structure

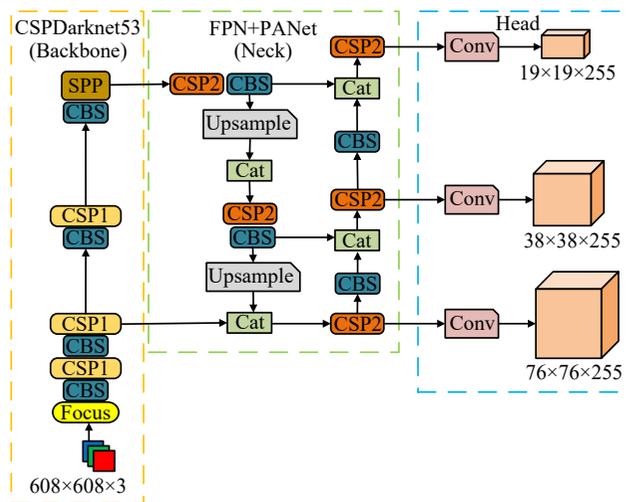


图7 YOLOv5 网络结构

Fig.7 YOLOv5 network structure

Ge等^[38]在YOLOv3的基础上,引入Decoupled Head、Data Aug、Anchor Free和SimOTA样本匹配的方法,构建了一种anchor-free的端到端目标检测框架,于2021年提出了YOLOX。YOLOv6由Li等^[39]提出,该算法取消了YOLOv1到YOLOv5一直沿用的锚框,优化了主干网络和颈部网络,并对检测头进行了解耦,分开了边框与类别的分类过程。

Wang等^[40]在2022年提出YOLOv7,他们对头部结构进行了重构,引入新的注意力机制和特征融合方法,在多个标准数据集上表现出色。YOLOv8由YOLOv5团队提出,网络结构与YOLOv5类似,在损失函数计算方面,采用了TAL(task aligned assigner)分配策略。由分类损失VFL(varifocal loss)、回归损失CIoU(complete-IoU)和DFL(distribution focal loss)三个损失函数加权组合而成。

2024年,YOLOv9^[41]由YOLOv7团队提出。YOLOv9引入了程序化梯度信息(programmable gradient information, PGI)技术,解决了深层网络中的信息丢失问题;并设计了兼顾轻量级、推理速度和准确性的广义高效层聚合网络(generalized efficient layer aggregation network,

GELAN),在多个标准数据集上展现了优异的性能。表2总结了YOLO系列算法的网络结构。

2.3.2 SSD和RetinaNet目标检测算法

除了YOLO系列算法,SSD和RetinaNet算法的目标检测性能也比较优秀。2016年,Liu等^[42]在欧洲计算机视觉国际会议(ECCV)上首次提出SSD算法。SSD采用了多尺度特征映射图和先验框(prior box)的方法,一次性完成目标的定位和检测,显著提高了运算速度,达到了与Faster R-CNN相当的准确率。针对SSD对小目标检测效果不佳的问题,该团队又提出DSSD(deconvolutional single shot detector)^[43]算法,通过增加网络深度和引入残差连接,提高了检测精度。

Lin等^[44]在2017年提出RetinaNet算法。该算法主要由ResNet主干网络、FPN、分类子网络和边框回归子网络组成。其中FPN用于融合不同层次的特征图,提高了对多尺度目标的检测能力;分类和回归子网络则负责在FPN输出的特征图上进行目标分类和位置回归。此外,该算法通过引入Focal Loss损失函数,有效解决了样本不平衡问题。

2.3.3 其他目标检测算法

SqueezeDet是由Wu等^[45]提出的一种单阶段目标检测算法。主干网络使用SqueezeNet,通过堆叠卷积滤波器提取低分辨率特征图,并利用卷积层计算大量物体边界框和类别预测,以减少内存访问次数,提高能效。该网络在实现高准确性的同时,具有实时推理速度、小模型尺寸和高能效等优点。

2019年,沈春华团队^[46]提出了FCOS(fully convolutional one-stage)算法。相较于传统的Anchor-Based方法,FCOS避免了复杂的锚框设计,减少了超参数调整的需求,同时缓解了正负样本不平衡的问题。通过引入中心度分支,能够更精确地定位目标中心,提高了检测精度。FCOS在后处理中使用非极大值抑制(non-maximum suppression, NMS)来过滤重叠的边界框,提高了检测结果的质量。

除上述经典算法之外,还有CenterNet^[47]、EfficientDet^[48]等目标检测算法也可在不同的场景用于目标检测。

表2 YOLO算法网络结构

Table 2 YOLO algorithm network structure

模型	骨干网络(Backbone)	颈部网络(Neck)	检测头(Head)
YOLOv1	GoogLeNet, VGG-16	2×fully connected layers	IoU Loss, NMS
YOLOv2	Darknet-19	fully connected layers	IoU Loss, NMS
YOLOv3	Darknet-53	FPN	IoU Loss, NMS
YOLOv4	CSPDarknet53	SPP, PANet	CIoU Loss, DIoU NMS
YOLOv5	CSPDarknet53	SPPF, CSP-PAN	GIoU Loss, DIoU NMS, 邻域正负样本分配策略
YOLOX	Darknet53	SPP, FPN	CIoU Loss, DIoU NMS, Decoupled Head, SimOTA 标签分配策略
YOLOv6	Efficient Rep	SPP, Rep-PAN	SIoU Loss, DIoU NMS, Efficient Decoupled Head, SimOTA 标签分配策略
YOLOv7	Extended-ELAN	SPPCSPC	CIoU Loss, DIoU NMS, SimOTA 标签分配策略
YOLOv8	Darknet-53	SPP, PAN	CIoU Loss, DFL Loss, DIoU NMS, TAL 标签分配策略, Decoupled Head
YOLOv9	GELAN	PGI	CIoU Loss, DFL Loss, BCE Loss, TAL 标签分配策略

3 单阶段金属表面缺陷检测算法优化措施

单阶段目标检测算法的不断更新为金属表面缺陷检测提供了更多的检测方法。国内外研究学者根据检测过程中遇到的问题,以原算法为基础,进行不同方面的优化,提供了更可靠的缺陷检测方案。本文将从数据增强、特征的提取与融合、锚框优化、模型轻量化四个方面进行总结,了解算法的优化措施以及应用在金属表面缺陷的检测效果。

3.1 数据增强

使用单阶段目标检测算法检测金属表面缺陷需要大量的训练数据,而获取高质量的标注数据往往成本高昂,现有的公开数据集图像种类和缺陷数量又比较有限,图像的质量参差不齐。因此,可通过在现有数据的基础上生成更多的训练样本,或者提高图像的清晰度再训练模型,从而提高模型的泛化能力。

Liang 等^[49]为解决数据不平衡的问题,采用数据增强方法对铝材数据集 APSDD(aluminum profile surface detection database)进行均衡操作。除了使用 YOLOv5 的随机裁剪、翻转、旋转等方法外,还应用镜像、高斯模糊、均值模糊、高斯噪声等操作对数据集进行进一步增强,将原有的 1 218 张图像扩充到 5 652 张图像,并通过 YOLOv5 进行测试。算法在数据增强后的测试集上 mAP 提高 33.9%,达到 91.3%,降低了模型对特定特征的依赖,增强了模型的鲁棒性。Xiao 等^[50]考虑到摄像机在拍摄镀锌层钢表面缺陷图像时,由于生产线设备的高速运行、曝光时间、光照条件等因素,会导致图像出现运动模糊,影响算法检测结果,遂采用维纳滤波(Wiener filter)对图像进行降噪。但对非高斯噪声的处理效果不佳,对应用场景有限制。

数据增强一方面可以增加数据多样性,提高模型的鲁棒性;另一方面针对数量较少的小目标,可通过数据增强增加数量,提升模型对小目标缺陷的识别能力。但随着样本数量的增加,模型的复杂度和训练时间也可能随之增加,造成训练过程中的资源需求增加,并且并非所有的数据增强方法对模型训练都有利。Chen 等^[51]仅对 NEU-DET 数据集中的小缺陷图像进行裁剪、复制、粘贴,虽然在数量上满足了模型训练的条件,但图像上小目标缺陷的清晰度依然较低,模型很难捕捉到缺陷的信息。实验证明,使用这种方法前后,AP 仅提升 0.1%,mAP 保持不变,对模型的检测精度影响较小。因此,需要根据实际的应用场景以及数据的质量选择是否进行数据增强。例如多数模型在检测 NEU-DET 数据集中的“裂纹”类缺陷时并不能达到很好的效果,此时,就可以考虑针对这一类缺陷的图像进行处理,然后再进行模型训练。

3.2 特征的提取与融合

原始的金属表面缺陷图像通常包含了许多冗余信息,使用原始数据训练检测模型效率较低且计算量较大。通过增强算法特征提取和融合能力,减少原始图像

的复杂性,提高模型捕获图像中不同尺寸缺陷信息的精准度,有利于金属表面缺陷的检测。

Cheng 等^[52]为解决因钢材表面缺陷种类繁多、存在大量小面积和模糊边缘损伤而导致检测精度不高的问题,使用 TBConv(tied block convolution)卷积模块对 YOLOv4 骨干网络中的标准卷积层进行改进。引入高效通道注意力机制(efficient channel attention, ECA)到检测层,并在空间特征金字塔池化模块(spatial pyramid pooling, SPP)之后增加级联瓶颈残差结构。模型检测钢铁表面缺陷准确率提高 6.2%,mAP 为 78.34%;模型大小缩小 27.46%,检测速度为 47 FPS。

Ren 等^[53]除了引入 ECA 机制,还引入了 SiLU(sigmoid linear unit)激活函数和 SIoU 损失函数优化 YOLOv5。模型检测钢材表面缺陷 mAP 为 78.8%,比原始模型提高 7.1%;召回率达到 76.4%,模型对于表征不明显的缺陷检测准确率较低。

ECA 机制使用一维卷积的方式实现跨通道间的信息交互,避免了复杂的降维和升维过程,因此能在不增加显著计算负担的情况下提升模型性能,计算效率较高。但该机制主要关注通道信息,相对忽略了空间信息,在检测任务中并不是最优的选择。考虑到该机制的局限性,Yu 等^[54]和 Jia 等^[55]引入其他的注意力机制优化 YOLOv5。前者引入 CBAM(convolutional block attention module)机制,不仅关注通道信息,还关注空间信息,实现从通道到空间的顺序注意力结构。在自制的包含 5 种缺陷(圆形、裂缝、损坏、划痕、空洞)的六角螺栓数据集上进行验证,mAP 比原 YOLOv5 提高 1.1%。后者引入 CA(coordinate attention)机制到 YOLOv5,将位置信息嵌入到通道注意力中,减少空间信息的损失。提出的模型 YOLOv5-CDG 在引入 CA 机制后检测轴承表面划痕 mAP 提高了 2.2%。

引入不同的注意力机制虽然可以增强模型的特征提取能力,提高检测精度,但也会带来较多的计算资源,降低模型检测速度。Zhang 等^[56]使用这种方法优化 YOLOv5,模型的精度虽有提升,但平均每张图片的检测时间增加了 3 ms。

为了在保证检测精度的前提下,提高检测速度,朱德平等^[57]在引入 CBAM 的基础上,在 Neck 中设计 F2C 特征融合模块,并利用 NWD(normalized Wasserstein distance)优化回归损失,降低了模型参数大小。在包含 3 530 张齿轮图像的数据集中 mAP 提高了 3.2%,减少了缺陷误检、漏检的情况;检测速度提升到 43 FPS,比 YOLOv7 高 8 FPS。Li 等^[58]则是在引入 SA(spatial attention)机制的基础上,将 CSPCrossLayer 模块融入到 YOLOX 骨干网络中,建立跨层连接;又引入 PSblock 卷积模块替代特征融合网络中的 CSPLayer 结构,减少冗余计算。实验测得 mAP 达到 77%,检测速度达到 100 FPS,提升了 34 FPS;在钢轨数据集上,该算法的

mAP 达到 88.8%,检测速度为 93 FPS,在保证检测精度的前提下,实现了检测速度的提升。但针对 NEU-DET 数据集中的裂纹类小目标缺陷识别效果不佳。

小目标缺陷检测是金属表面缺陷检测的一个难点,如果使用传统的多级卷积运算,可能导致像素占比少的小目标在该过程中丢失。为了提高小目标缺陷的检测精度,自 YOLOv3 开始在模型的 Neck 结构中使用 FPN。该网络能够在增加较少计算量的前提下融合低分辨率语义信息较强的特征图和高分辨率语义信息较弱但空间信息丰富的特征图。后续研究人员在此基础上进行优化、调整,并与基础算法融合,能够有效提高算法提取特征的能力。

李思雨^[59]在 2023 年为检测铝材表面的脏点、擦花等缺陷,提出一种 YOLOv5-CA-GFPN 模型。该模型采用全局特征金字塔网络(global feature pyramid network, GFPN)替换 PANet 模块,引入 CA 注意力机制;同时,为增强模型的训练速度和适应性,使用 IoU 距离和 Mosaic-9 分别代替欧氏距离和 Mosaic-4。在包含 766 张铝材表面缺陷图像的数据集中 mAP 为 92%,检测速度为 83.3 FPS,提升了铝材表面缺陷的检测精度。但模型使用的数据集数据量较少,与实际工业数据存在一定的偏差,模型的泛化能力较弱。

Zhao 等^[60]在 FPN 的基础上设计了一种新的特征金字塔结构:双特征金字塔网络(double feature pyramid network, DFPN)。DFPN 主要由两个 FPN 组成,一个与骨干网络连接,另一个与检测头连接,如图 8 所示。两者之间通过 3 个 SPP 模块建立连接关系。SPP 块的输出不仅与 FPN 连接,还与预测部分之前的特征映射连接,DFPN 不仅缓解了信息丢失问题,还增强了特征复用和融合的能力。在引入 Res2Net 模块的基础上,使用 DFPN 优化 YOLOv5 的 Neck 结构,所得模型 RDD-YOLO 具有

良好的鲁棒性和泛化能力,在 NEU-DET 和 GC10-DET 数据集上的平均检测精度和检测速度都有所提升。

特征金字塔网络比较侧重于多尺度特征的融合和利用,在训练和推理时通常需要更多的 GPU 内存和计算资源。2024 年,高春艳等^[61]为节约计算资源,放弃使用 FPN 优化 YOLOv7,提出一种高精度实时的缺陷检测算法 CDN-YOLOv7。通过加入 CARAFE 轻量化上采样算子来改善网络特征融合能力,相对于 FPN, CARAFE 更侧重于上采样过程中的内容感知和细节保留,减少了计算开销。此外,使用 NF-EIoU 替换 CIoU 损失并重新设计了解耦头。实验结果表明,模型检测钢材表面缺陷的 mAP 可达 80.3%,较原 YOLOv7 算法提高 6.0%,模型推理速度可达 60.8 FPS,满足实时性需求。但 CARAFE 比较适用于需要高质量上采样的场景以及资源受限的设备上运行。

赵泽民^[62]通过引入深度网络注意力模型 SNR(self notice ratio)与 YOLOv6 进行结合,用于检测轧钢表面缺陷。但针对图像受外部光照条件的影响,导致缺陷目标和背景的颜色混淆,算法识别准确率不佳的问题,提出 SNR-YOLOv6-U-Net 模型。通过将 YOLOv6 处理的图像输入 U-Net 卷积模型,以确认缺陷的存在并检测缺陷的精确位置,并以二进制掩码的方式输出,放在原图像上,达到检测目的。U-Net 的引入提升了模型的检测精度,识别面状、周期、单一等轧钢缺陷的准确率分别达到了 96.6%、93.6%和 94.5%。但模型的检测计算时间较长,后续需要做轻量化处理。

Lu 等^[63]在 YOLOv8 的基础上提出 WSS-YOLO 模型。通过在骨干网络中添加 C2f-DSC(dynamic snake convolution)模块,增强模型对复杂缺陷的特征提取能力,在 Neck 结构中引入 GSConv 和 VOV-GSCSP 模块,使用 WIoU 损失函数代替 CIoU 进行锚框分类。优化后

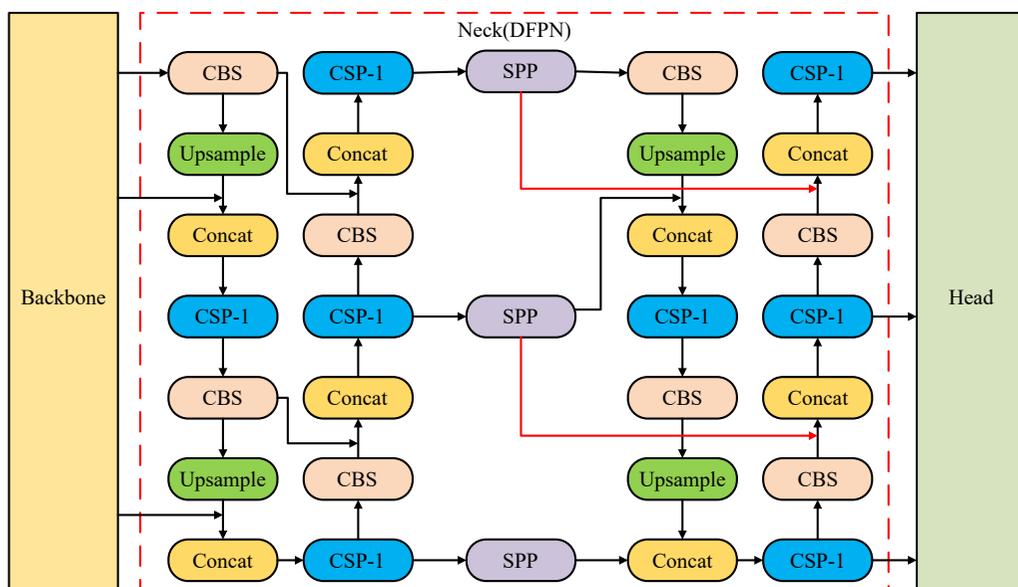


图8 DFPN 结构

Fig.8 DFPN structure

的模型mAP提高1.85%,达76%,精测速度低于原算法。上述研究是在原有算法的基础上重新设计卷积模块,增加了卷积层的数量以增强模型理解目标缺陷及其空间关系。但增加模型的计算复杂度,会影响检测速度,Chen等^[64]在引入深度感知卷积模块(depth aware convolution module, DAC)到YOLOv7-tiny的基础上,为了不影响检

测速度,利用快速归一化融合和BiFPN设计一种新的特征融合网络AWFP添加在concat模块后面,最后使用Focal-SIoU替代CIoU损失函数。所提模型CAF-YOLO在GC10-DET和NEU-DET两个数据集上的mAP分别提高11%和9.2%,检测速度都超过100 FPS。表3总结了部分YOLO算法的优化方法及实验结果。

表3 YOLO算法优化方法

Table 3 YOLO algorithm improvement methods

文献	优化方法	实验结果	准确率/%	FPS
文献[65]	提出YOLOv3-SA模型,引入空间注意力机制SA嵌入到残差模块内部。引入K-means生成新的预设锚框	模型的查全率和mAP有所提升,mAP提升约9%。对于缺陷的识别和定位不够精准,查准率较低	78.30 NEU-DET	—
文献[66]	提出YOT-Net模型,通过引入三重损失函数,构建LCCT(location-confidence-class probability-triplet)损失模块,利用图像相似性提高特征提取能力	提高了铜弯头表面缺陷检测精度,与原YOLOv3相比,mAP提高8.23%。但模型检测速度略慢	67.42 自建集	4.6
文献[67]	在YOLOv3主干网络Darknet-53的第2、3、4、5个残差块后加入CBAM混合注意力机制模块,并引入DYReLU动态激活函数。新增一个尺度为104×104的特征层,并进行跨层密集连接。使用K-means++算法聚类分析先验框	提升了金属表面缺陷的检测精度,比原算法提高6.81%,平均每张图片检测时间为30.96 ms,略慢于原算法	78.69 NEU-DET	32.3
文献[68]	使用改进的Retinex增强钢轨图像。引入Res2Net和CBAM模块,以改进感受野和小目标位置权重	与原YOLOv4相比,mAP提高4.98%,平均每张图片的检测时间增加了13 ms	92.68 自建集	14.7
文献[69]	使用Transformer取代YOLOv5的Darknet骨干网络,并将Swin注意力机制添加到骨干网络,增强特征提取能力,引入Wise-IoU损失函数代替原函数	模型在NEU-DET和GC10-DET数据集上的mAP分别提高了9.6%和2.04%,对图片质量要求较高,反光和拍摄角度不佳会降低模型检测性能	84.73 NEU-DET	—
文献[70]	提出CGS-YOLOv5,使用Ghost模块代替标准卷积层,采用坐标注意力机制模块CA,提高PCB小目标的检测能力,引入SIoU损失函数,提高模型的精度	改进后的YOLOv5在平均准确率提高了1.4%,参数大小缩小了47%,仅8.2 MB	95.40 自建集	57.8
文献[71]	提出CCFSC(cross-convolution feature strengthening connection)方法,缩短信息传递路径,提高特征表征能力。引入CBAM模块与concat模块融合,构成EFCM模块,增强特征之间的信息传播。使用K-means++算法聚类先验框	相较于原YOLOv5s算法,平均检测精度略有提升。模型识别白斑缺陷效果不佳	77.40 自建集	—
文献[72]	在YOLOv5骨干网络中设计多个MB(multi-scale block)块,每个块包含两个MSCConv(multi-scale convolutions)。引入SA机制,提高网络表征能力	模型以72%的mAP检测出钢铁表面缺陷,并满足实时速度要求,达192.3 FPS。不宜用于检测同一图像中出现多种缺陷	72.00 NEU-DET	192.3
文献[73]	在检测前加入图像增强模块,引入MHSA(multi-head self-attention)模块嵌入YOLOv5骨干网络末端。引入跨层加权级联,减少钢轨图像噪声对目标检测的影响	模型检测钢轨的裂纹、剥落、磨损三类缺陷的精度和速度有所提升。受限于数据集中缺陷种类较少,模型检测其他缺陷效果未知	98.20 自建集	77.0
文献[74]	利用多尺度特征融合(multi-scale feature fusion,MSF)策略融合浅层和深层特征,丰富目标缺陷的细节信息。引入CRA block模块(CSPLayer Res2Attention block),减少分层传输过程中的缺陷信息损失,增强特征提取能力	优化后的YOLOv5可以准确地识别缺陷种类,定位缺陷位置。但用在其他金属表面缺陷数据集,效果一般	78.50 NEU-DET	54.0
文献[75]	引入SiLU激活函数和BiFPN网络增强特征融合能力。使用上采样算子CARAFE和MHSA方法改善输出端的检测性能	与YOLOv7-tiny算法比较,mAP提升11.7%,检测速度满足实时性要求	93.20 NEU-DET	192.0
文献[76]	提出一种小样本驱动的训练样本生成方法。将SE(squeeze and excitation)模块与YOLOv7的ELAN模块融合,引入ACmix注意力模块和WiIoU损失函数	优化后的模型在自建钢板表面缺陷数据集上mAP达64.7%,后续可降低模型参数的规模,提升模型的检测速度	64.70 自建集	—
文献[77]	提出STFE-Net模型。在预处理阶段,采用CLAHE算法增强源图像的对比度;在特征融合阶段,引入SPA-Res模块和SFW(shallow feature weighted)分支来加权纹理特征,并添加特征层提取更丰富的语义特征;在预测阶段,添加检测头和ECIoU损失函数以精细化金属缺陷定位	相较于YOLOv8m,mAP提高6.9%,检测精度有所提升,检测速度一般	82.30 NEU-DET	54.0
文献[78]	引入SimAM注意力机制实现对感兴趣特征区域的快速提取,在SPFF模块中添加CBAM模块扩大感受野	与原始YOLOv8算法相比,改进后的算法mAP提高9.3%,召回率提高10%	97.90 NEU-DET	—

3.3 锚框优化

从数据增强和特征的提取与融合两个方面优化算法主要是针对图像上缺陷信息的处理,但是在实际应用中,缺陷的尺寸、形状各不相同,如果仅依靠优化算法的主干网络和颈部网络,难以充分覆盖不同缺陷的特征。因此为了进一步提升目标检测的准确性,可对锚框(anchor boxes)进行优化。锚框为模型提供一组预定义的框,这些框具有不同的大小和宽高比。模型通过锚框进行预测,然后根据实际目标缺陷的位置和形状进行调整,以便更准确地预测目标的真实边界框。锚框的优化主要涉及到对锚框的尺寸、形状和数量的调整,以更好地适应不同类型的缺陷。

刘浩翰等^[65]通过引入K-means算法对真实目标框进行聚类,将模型的检测精度提升了1%,但该算法的聚类结果比较依赖初始簇中心的选择,会导致聚类结果不稳定。为了降低K值对结果的影响,Fan等^[79]在K-means算法的基础上设计了一种新的锚点优化算法GA-K-means,提出ACD-YOLO新模型。通过GA(genetic algorithm)算法调整锚点边界,降低依赖,优化后锚点的BPR(best possible recall)可达1.00。此外还在YOLOv5骨干网络的末端和颈部结构分别添加一个上下文增强模块(context augmentation module,CAM)提取语境特征,并引入高效卷积算子DSConv降低模型复杂度。模型在实验中取得了79.3%的mAP和72 FPS。尽管已经降低了计算复杂度,但检测速度仍落后于YOLOv7和YOLOv8,需要进一步改进。

陈明文^[67]和胡欣等^[80]使用K-means++算法对锚框重聚类,通过改变初始聚类中心的生成方式,增大初始锚框之间的差距,使锚框更适应整体数据分布,得到匹配度更好的多尺度锚框,进而提升模型的检测性能。前者在引入该方法后mAP提升了0.6%,后者在包含4 500张螺纹钢表面缺陷图像数据集中的mAP提升了1.3%,说

明聚类后的锚框对检测精度有所提高。

Liu等^[81]提出YOLO-IMF模型,以解决铝板表面缺陷识别问题。使用EIoU损失函数代替YOLOv8的CIoU损失函数。EIoU是在CIoU惩罚项的基础上将预测框和真实框的纵横比的影响因子拆开,分别计算预测框和真实框的长和宽,可以更好地衡量小目标和不规则形状目标之间的相似性,增强边界框回归的有效性。实验结果表明模型的检测精度略有提升,mAP从98.1%提高到99.3%。模型仅对损失函数做了调整,且原基础算法的检测精度较高,并不能完全体现出EIoU损失函数的优化效果。Kong等^[82]在YOLOv8的基础上引入WIoU损失函数,更加重视普通质量的锚框,减少极端样本中产生的有害梯度。使用卷积残差模块来提高特征提取能力,模型的mAP比原YOLOv8提高了3.4%,达74.9%,但计算量增加了约20%。

通过表3及其他学者的研究不难发现,现阶段对于金属表面缺陷检测算法的优化方法主要体现在引入不同的机制^[62,66-67]、替换原模型的骨干网络^[68-70]、改进卷积层和连接方式^[72-74,77]以及在输出阶段对锚框进行优化^[66-67,71,82]四个方面。使用这些方法虽然可以提高模型检测金属表面缺陷的精度,但不同的机制对缺陷信息的关注程度不一样,对于不同的检测对象需要使用合适的机制,这会降低模型的泛化能力,导致模型的鲁棒性较差^[67-68]。此外,针对一些不明显的特征,特别是小目标缺陷,对卷积层数量和连接方式的改进确实有利于模型更好地提取和融合缺陷信息,但可能会带来计算量的增加,降低模型的检测速度,不适用于实时性要求较高的应用场景^[75-76]。引入不同的损失函数优化锚框的位置和大小,使其更贴近目标对象,提升定位精度特征,但需要进行实验和调优,可能会造成时间和资源的浪费^[82]。

3.4 网络模型轻量化

表4总结了SSD算法的优化方法及优缺点。从表3、

表4 SSD算法优化方法
Table 4 SSD algorithm improvement methods

文献	优化方法	优缺点	准确率/%	FPS
文献[83]	引入ResNet50网络作为基础网络。使用反卷积融合方法弥补浅层语义信息	mAP比原SSD算法提高了4.6%。加入卷积层导致计算量增大,实时性不高	94.13 NEU-DET	—
文献[84]	使用ResNet50网络作为特征提取网络。引入通道注意机制对特征图进行过滤和保留重要信息	相较于SSD算法,RAF-SSD模型mAP提升5.3%。小目标缺陷的识别效果不佳	75.10 NEU-DET	53.0
文献[85]	在特征提取阶段,使用ResNet50的前七层作为特征提取网络。在编码阶段,设计语义融合结构CFS(context fusion structure)。在解码阶段,使用特征重构模块FRM(feature refinement module),以过滤出不同尺度特征图融合带来的语义冲突和冗余	mAP提升7.5%,检测速度下降34 FPS。模型检测划痕、凹坑缺陷的精准度较低,不适合用于在低分辨率和低对比度的环境	79.50 NEU-DET	71.0
文献[86]	提出SDDNet模型。引入FRB模块将多个分辨率不同的金字塔特征图融合在一起,并插入到池化层之上,以在保留纹理信息。引入SDCM(skip densely connected module)模块将细节从低级特征图传播到高级特征图,以更好地预测缺陷	模型的检测精度和泛化能力大大提高。在DAGM和Magnetic-Tile的mAP分别为99.1%和93.4%。但模型计算较复杂,输入样本数的大小会影响检测速度的快慢	88.80 NEU-DET	—
文献[87]	使用Transformer多头注意力机制模块替代原SSD结构中的Conv5_1层,以提高对小目标缺陷的检测能力;将Conv7操作替换为Involution算子操作,减少运算的参数量	相比改进前mAP提高了4.5%,检测速度提高了13.6%。识别小目标缺陷存在漏检的现象	73.18 自建集	34.5

表4中可以看出,优化后的深度学习算法可以提高金属表面缺陷的检测精度,但模型的复杂度和参数量会影响检测速度。研究人员为解决这一问题,考虑在保证检测精度的前提下,降低模型的计算复杂度和空间复杂度,使其能够被安装在计算及存储资源受限的嵌入式边缘设备上,推动深度学习模型从理论走向实际。

金属表面缺陷检测模型主要通过使用轻量化网络代替原主干网络,或使用网络剪枝、参数量化、知识蒸馏等技术压缩模型实现网络轻量化的部署。常用的轻量化网络有 MobileNet、ShuffleNet、GhostNet 等。

MobileNet 系列网络自 2017 年推出,始终在深度学习领域扮演着重要的角色。MobileNetv1^[88]首次将深度可分离卷积(depthwise separable convolution)引入到主流的卷积神经网络设计中。这一创新有效减少了模型的参数量和计算复杂度,使得在资源受限的设备上能够实现实时的图像识别和处理,但在分类任务中的准确率较低。

随后的 MobileNetv2^[89]在架构上进行了进一步的优化,提出了逆残差(inverted residual)结构,这一设计允许信息和梯度在网络中更有效地传递。此外,MobileNetv2 引入了线性瓶颈(linear bottleneck)层,减少了非线性变换对计算的影响,提升了模型在移动设备上的推理速度和精度。但相对于 v1, v2 架构较为复杂,可能会增加实现和调优的难度。

MobileNet 系列网络的基本结构如图 9 所示。在 MobileNetv3^[90]中,Google 采用了混合搜索空间(mixed search space)和自动机器学习(auto machine learning)技术,结合经验设计与搜索策略进行优化,引入 SE (squeeze-and-excitation)注意力机制,优化输入数据的处理方式,使得模型在处理精度和处理时间之间达成了更好的平衡。MobileNet3 的参数量相较于 MobileNetv2 增

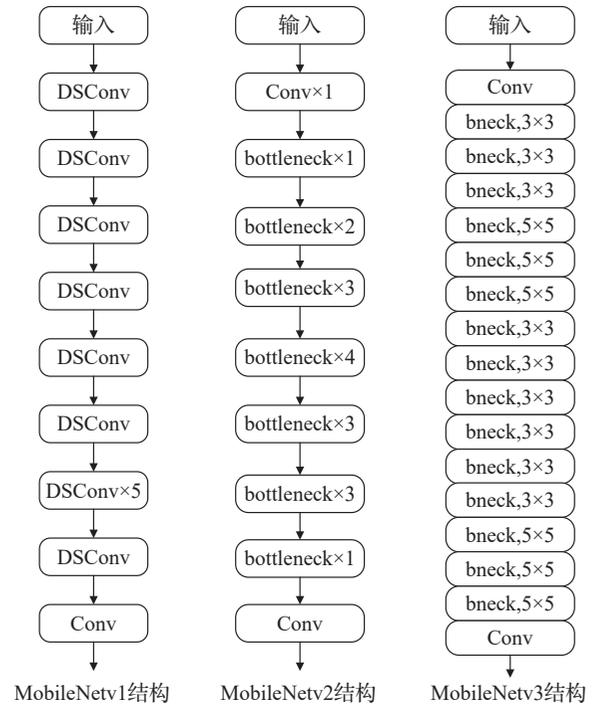


图9 MobileNet 网络结构

Fig.9 MobileNet network structure

加了 2 MB,为 5.4 MB;计算量减少了 81 MB,为 219 MB,适用于性能要求极高的场景。

同在 2017 年提出的还有 ShuffleNet^[91]轻量级卷积网络。研究者们通过将卷积操作拆分为多个小的部分,结合通道混洗的方式,设计出 ShuffleNetv1 网络,有效地减少了计算复杂度,但在某些应用场景中面临着网络深度和特征表达能力的挑战。ShuffleNet 网络单元结构如图 10 所示。

2018 年,研究人员根据设计轻量快速模型的四个准则,提出了 ShuffleNetv2^[92]。在 ShuffleNetv1 的基础上引入通道分离策略,以提高特征传播的效率。模型通过综

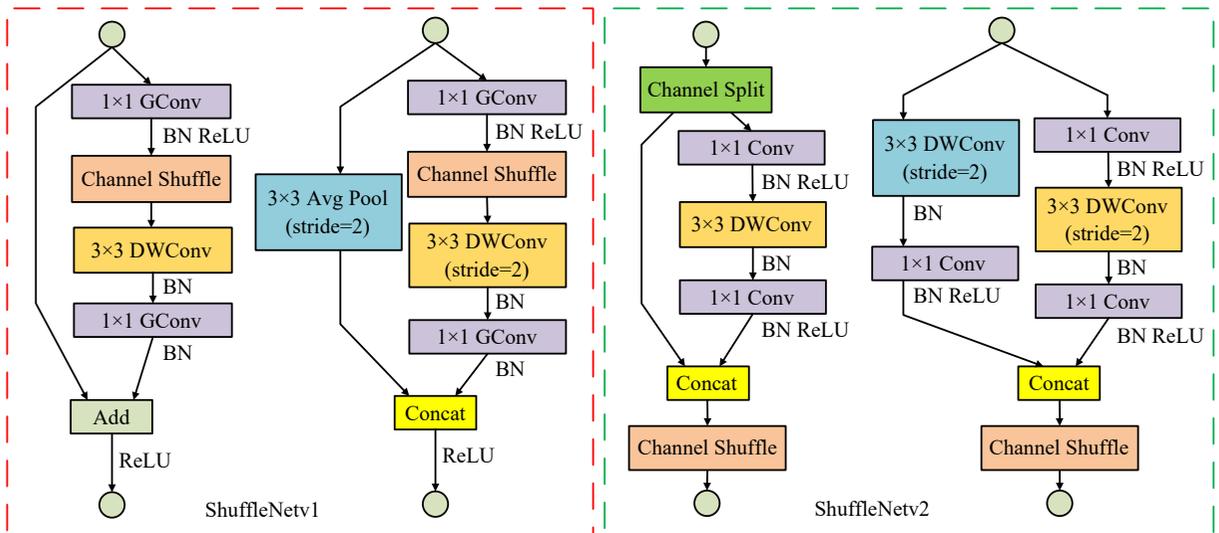


图10 ShuffleNet 网络单元结构

Fig.10 ShuffleNet network unit structure

合考虑输入特征通道的数量和计算复杂度,使得每个模块都能在保持高性能的同时更好地协调各个层之间的特征传递。shuffle操作不仅改善了特征的混合能力,还增强了模型的泛化能力。

尽管ShuffleNet和MobileNet在轻量级网络方面已经取得了一定的成果,但二者仍存在一些问題。ShuffleNet在执行阶段比较复杂并且传递信息效率不足。MobileNet在特征提取过程中可能会由于过于简化的结构而导致信息损失,影响模型的准确性。为了弥补上述缺点,GhostNet^[93]于2019年被正式提出。GhostNet借鉴深度可分离卷积的思想,最重要的创新在于“Ghost模块”。该模块通过简化的线性运算生成额外的特征图,从而有效降低模型的参数数量和计算量,其结构如图11所示。这种高效的特征生成方式使得GhostNet在执行速度和内存占用方面表现卓越,具备良好的边缘适应性。

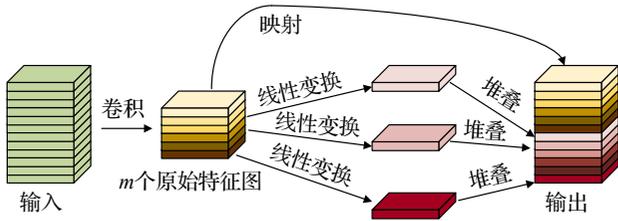


图11 Ghost Module 结构

Fig.11 Ghost Module structure

GhostNet虽然在性能上优于另外两种算法,但并不代表使用该网络优化其他算法一定能够提升模型的检测精度和速度。是否使用轻量级卷积网络或者使用哪一种网络,需要根据金属表面缺陷检测的需求及应用场景来选择。

比如Xie等^[94]就没有直接使用轻量级卷积网络,而是通过设计轻量级卷积模块DSConv对YOLOv4骨干网络进行轻量化处理,引入DFPN,并在CIoU和HIoU的基础上建立一种新的回归损失函数CHIoU加速模型收敛。计算方式如下:

$$L_{CHIoU} = \frac{\omega_1 L_{CIoU} + \omega_2 L_{HIoU}}{\omega_1 + \omega_2} \tag{7}$$

$$\omega_1 = \left(1 - \cos \frac{Steps}{Epochs} \pi\right) / 2 \tag{8}$$

$$\omega_2 = \left(1 + \cos \frac{Steps}{Epochs} \pi\right) / 2 \tag{9}$$

式中, $\omega_1 \in [0, 1]$ 是余弦递增因子, $\omega_2 \in [0, 1]$ 是余弦衰减因子, Steps是训练记录的周期数, Epochs是实验中的总周期数。提出的FE-YOLO轻量化近80%,参数量仅10.36 MB;精度提升3.9%,达83.9%;在NVIDIA GeForce RTX 2080Ti上运行速度为90 FPS。

李丰润^[95]在2023年提出一种新思路用于网络模型轻量化——网络双流,把单路网络改为一个深分支和一个浅分支网络,如图12所示。深分支层数多,负责从低分辨率图像中提取高级语义特征;浅分支层数少,负责从高分辨率图像中提取低级语义特征;两个分支间通过桥接层进行连接,以促进梯度传播和信息沟通。优化后的YOLOv4模型参数量为18.4 MB,远少于SSD、YOLOv4等目标检测算法,实现了网络模型的轻量化,但对图像的分辨率要求较高,图像分辨率较小的情况下,该模型可能会降低检测速度。

Wu等^[96]和Liu等^[97]则是借助轻量级网络GhostNet降低YOLOv5模型复杂度。前者通过设计卷积鬼影网络模块(convolution-GhostNet hybrid module, CGH)和多卷积特征融合模块(multi-convolution feature fusion block, MCFF),并采用CARAFE取代双线性插值上采样,引入BiFPN增强模型对不同尺度目标的适应性。模型的参数量和计算量降低至3.4 MB和6.4 GFLOPs,mAP提高了6.3%,达到76.1%;FPS也由原来的96提升到121。后者不同于前者的是卷积模块较为简单,在特征融合部分使用渐进式特征金字塔网络(asymptotic feature pyramid network, AFPN)增强特征的利用能力,并增加解耦头层的数量增大感受野,最终在保证检测精度的前提下,模型的参数量减少了一半,检测螺栓的mAP达到了98.6%,在同前者相同的设备NVIDIA RTX 3090上运行的FPS由原来的140提升到244。

上述两种方法都实现了网络模型的轻量化,但后者

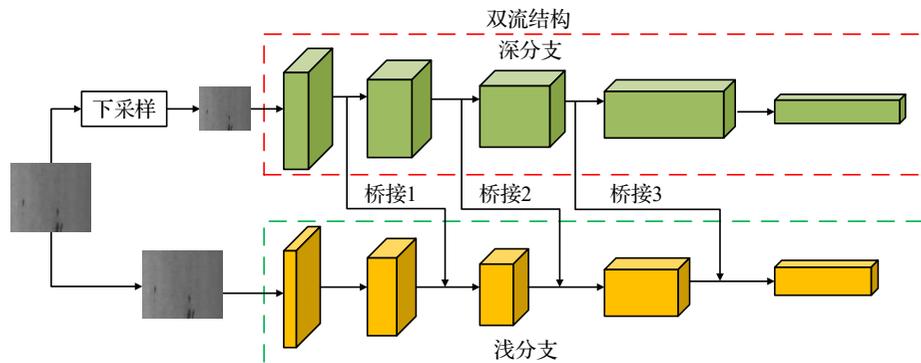


图12 双流结构示意图

Fig.12 Double flow structure

模型在精确度和检测速度方面的改变都优于前者。除了优化方法不同带来的影响之外,数据集的质量也起到了重要作用。前者使用的NEU-DET数据集,图像清晰度较差且未对数据集中的黑白图像进行处理,而后者使用的是自制的螺栓数据集,图像背景简单,清晰度较高,因此,后者的模型表现相对更佳。

Zhang等^[98]在MobileNetv2的基础上添加Swish-1激活函数和额外的拓展层,设计出轻量级网络RexNet用于特征提取;在Neck部分使用轻量级模块DSConv(depthwise separable convolution)和VoV-GSCSP来替代标准卷积,减少模型复杂度和计算复杂度;将动态检测头(dynamic detection head, DyHead)与解耦头相结合组成DdyHead检测头,以增强模型对细微缺陷的识别能力;最后通过剪枝算法进一步压缩模型。相较于原YOLOv7算法,所提模型YOLO-RDP参数量减少了40%,仅3.5 MB;计算量减少了25%,为9.95 GFLOPs。模型的mAP提升至79.8%。

梁礼明等^[99]为降低模型的参数量和计算量,在YOLOv7-tiny的基础上引入鬼影混洗卷积GhostConv压缩通道数来减少计算量。实验结果表明,模型的参数量减少了1.6 MB,计算量减少了3.8 GFLOPs。苏佳等^[7]将这种方法用到YOLOv8算法上,并在骨干网络引入卷积模块ELK(efficient large kernel)和多尺度上下文模块MCM(multi-scale context module);在Neck结构中引入VoVGSCSP模块替换C2f模块,得到的EML-YOLO,模型参数量仅2.7 MB,降低了12.9%,计算量降低14.6%,仅7.0 GFLOPs。模型在GC10-DET和DeepPCB两个数据集上的mAP均有所提升。表5总结了金属表面缺陷检模型轻量化改进的方法及结果。

Chen等^[107]注意到YOLOv8骨干网络中C2f卷积模块的计算较为复杂,于是对该模块进行优化,减少计算量。使用轻量级网络EfficientViT作为特征提取网络,设计轻量级卷积模块C2f-Faster-EffectiveSE集成到模型颈部,利用共享参数的概念重建了检测头。优化后的EFS-YOLO模型参数量和计算量分别减少了49.5%和62.7%,分别为5.6 MB和9.6 GFLOPs;在NEU-DET数据集上取得了82.7%的mAP,提高了2.3%;在NVIDIA GeForce RTX 3090上运行FPS达92.4。

2024年,Xie等^[108]在YOLOv8骨干网络中引入高效全局注意力机制(efficient global attention mechanism, EGAM),并设计了轻量级多尺度混合卷积模块(light weight multi-scale mixed convolution, LMSMC),与C2f融合得到C2f_LMSMC。采用调整通道的方法进行特征融合,使用NWD代替CIoU作为回归损失函数。与YOLOv8相比,LMS-YOLO mAP提高2.8%,为81.1%;浮点运算减少了17.4%,参数减少了34.2%,在NVIDIA GeForce RTX 3070上运行FPS达61.3。

王春梅等^[109]以YOLOv8为基准,提出轻量级带钢表面缺陷检测模型YOLOv8-VSC。该模型使用轻量级网络VanillaNet作为骨干特征提取网络,并引入SPD(space to depth)模块和CARAFE模块。模型的参数量与计算量仅为1.96 MB和6.0 GFLOPs,为原算法的65.1%和74.1%,实现了模型的轻量化。在NEU-DET钢材数据集和在铝材表面缺陷数据集APSPC上的mAP分别为80.8%和68.7%,但由于部分缺陷的背景噪声较大,不易区分背景与缺陷,会存在误检、漏检的概率。

Liu等^[110]基于SSD模型,将ResNet50作为学生模型,ResNet101作为教师模型进行知识蒸馏得到KD-

表5 金属表面缺陷检测模型轻量化改进

Table 5 Lightweight improvement of metal surface defect detection models

文献	网络结构&拓展模块	参数量/MB	计算量/GFLOPs	准确率/%	FPS	运行平台
文献[94]	DSConv,DFP,CHIou(YOLOv4)	10.36	—	83.90	90.0	NVIDIA GeForce RTX 2080
文献[95]	将单路网络改为双路网络(YOLOv4)	18.40	—	78.41	30.0	NVIDIA GeForce RTX 3060
文献[96]	CGH,CARAFE,BiFPN(YOLOv5s)	3.40	6.40	76.10	121.0	NVIDIA RTX 3090
文献[97]	Ghost Bottleneck,AFPN(YOLOv5)	6.80	—	98.60	244.0	NVIDIA RTX 3090
文献[98]	RexNet,GSCConv,DdyHead(YOLOv7)	3.50	9.95	79.80	—	NVIDIA RTX 3090
文献[99]	VoVGA-FPN,TA,GSCConv(YOLOv7)	5.41	8.90	74.40	87.0	RTX 4060
文献[100]	GhostNet, α -CIoU loss(YOLOv4)	11.25	—	81.99	26.1	NVIDIA TITANX
文献[101]	RepGhost,GSCConv(YOLOv5s)	5.30	12.70	77.60	187.0	NVIDIA GeForce RTX 3060
文献[102]	GAC-FPN,BRA,SIoU(YOLOv7)	6.83	13.20	72.90	104.1	RTX 4060
文献[103]	VanillaNet,C2fDSConv,CA,WIoU(YOLOv8)	1.90	5.40	79.80	109.0	NVIDIA GeForce RTX 3070 Ti
文献[104]	EfficientViT,H-C2f,HWD(YOLOv8s)	7.00	15.60	79.70	136.0	RTX 4060
文献[105]	MobileNetv3,DGSCConv,D-SimSPPF(YOLOv7)	29.50	21.80	97.00	—	NVIDIA RTX 3060
文献[106]	CA,Faster Block,Ghost Conv(YOLOv8n)	2.80	—	75.80	384.6	RTX 3090
文献[107]	EfficientViT,EffectiveSE(YOLOv8s)	5.60	9.60	82.70	92.4	NVIDIA GeForce RTX 3050
文献[108]	LMSMC,EGAM,NWD(YOLOv8)	28.30	134.20	81.10	61.3	NVIDIA GeForce RTX 3070
文献[109]	VanillaNet,CARAFE,SPD(YOLOv8)	1.96	6.00	80.80	263.0	GeForce RTX 3060
文献[110]	KD-ResNet50,注意力机制(SSD)	24.71	3.80	71.40	61.0	NVIDIA GeForce 1070

ResNet50网络。在此基础上,添加特征融合模块和注意力机制。优化后的模型参数量为24.71 MB,计算量为3.8 GFLOPs。模型的准确率高于SSD模型、Faster RCNN模型和YOLO-v3模型,检测速度比最快的SSD模型高12.9%,能有效满足钢带表面缺陷的实时检测要求。

使用轻量化网络以及改进网络中的卷积模块是降低模型计算量的重要手段,轻量化的模型可以加快推理速度,提高模型的实时应用性能。但并不是所有的轻量化方法都适用,一些轻量化方法,如知识蒸馏,实施起来较为复杂,需要额外的模型训练时间并且优化后的模型可能在特定硬件上表现较好,而在其他硬件上效果较差。此外,模型的参数量和计算量下降并不代表模型检测速度一定会提升,模型运行的设备以及初始缺陷图像资源的大小也会对模型的检测速度产生影响,使用较好的运行设备,模型的计算速度相对更快。

4 总结与展望

本文总结了目标检测算法的发展历程,介绍了常用的金属表面缺陷数据集和评价指标,详细说明了近年来单阶段的金属表面缺陷检测算法的优化方式,分析了算法优化的优缺点及实验效果,最后介绍了常用的轻量级网络以及应用。未来金属表面缺陷算法可围绕以下方向继续展开研究。

4.1 多模态融合

目前的金属表面缺陷检测主要依赖于数据集中的图像信息,但是在实际场景中,单一的图像信息并不能提供足够的缺陷特征来实现目标识别。未来可通过整合多模态数据,即将来自不同模态(如文本、音频等)的数据信息进行整合,以提取出更加全面、准确的缺陷信息。如刘凡凡等^[11]使用像素级多模态融合优化算法检测小目标,从不同模态中提取共享信息和特殊信息,对不同模态之间的信息进行互补,提高了模型的信息提取能力。多模态融合可以弥补单一模态数据在表达信息时的局限性,从而帮助模型更好地理解数据背后的含义和上下文关系,使模型能够更全面地捕捉缺陷的特征,提高检测精度。但如何将这些具有不同的特征表示和语义结构的模态数据融合是一个关键问题,以及如何评估融合效果的优劣,也是多模态数据融合需要解决的问题。

4.2 大数据应用技术

Sun等^[12]利用多无人机联合部署收集交通事故数据,并分析这些大数据的信息以提高交通系统的智能化。金属表面缺陷检测同样可以利用这种思路提高模型性能。通过在生产线上安装各种传感器,实时收集金属表面的数据,并通过网络传输到云端数据处理中心(如IBM Cloud),这样不仅可以实现实时监控和缺陷检测,而且拍摄的图片不会占用本地设备过多的内存,影响模型的检测速度。同时可以利用大数据分析技术从大量的传感数据中挖掘出潜在的缺陷模式,并结合算法

进行预测和决策,从而提高检测效率和准确性。例如,通过分析温度、压力、湿度等环境参数,可以预测哪些条件下更容易产生缺陷,并采取预防措施。

4.3 现实与虚拟结合

当前算法在检测到金属表面缺陷之后是以标注的形式在屏幕上呈现,想要知道具体的缺陷位置和类型需要调取检测结果。未来可利用增强现实技术,将检测到的缺陷直接在实际环境中的金属表面上标记出来,帮助操作人员及时发现和处理问题。而虚拟现实技术可以通过模拟不同的光照条件、视角和表面状态,创建一个仿真的环境,用于训练和测试新的检测算法。Park等^[13]在2024年将这种现实与虚拟结合的思想应用到医学领域,用于检测轻度认知障碍,为检测早期老年痴呆提供了重要参考。后续可尝试将这种技术应用到金属表面缺陷检测领域,提高检测算法的适应性和可靠性。

参考文献:

- [1] AMERI R, HSU C C, BAND S S. A systematic review of deep learning approaches for surface defect detection in industrial applications[J]. *Engineering Applications of Artificial Intelligence*, 2024, 130: 107717.
- [2] LING L, SHENG Y, TU Z, et al. DL3DV-10K: a large-scale scene dataset for deep learning-based 3D vision[C]//*Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Piscataway: IEEE, 2024: 22160-22169.
- [3] YANG X, WEN G, MEI S, et al. DBHF: a double fusion backbone and bidirectional feature hybrid fusion detector for high-precision inspection of avionics solder joint defects[J]. *Measurement*, 2024, 237: 115221.
- [4] ZHENG Q, TIAN X, YU Z, et al. MobileRaT: a lightweight radio transformer method for automatic modulation classification in drone communication systems[J]. *Drones*, 2023, 7(10): 596.
- [5] ZHENG Q, SAPONARA S, TIAN X, et al. A real-time constellation image classification method of wireless communication signals based on the lightweight network MobileViT[J]. *Cognitive Neurodynamics*, 2024, 18(2): 659-671.
- [6] 伍麟, 郝鸿宇, 宋友. 基于计算机视觉的工业金属表面缺陷检测综述[J]. *自动化学报*, 2024, 50(7): 1261-1283.
- [7] WU L, HAO H Y, SONG Y. A review of metal surface defect detection based on computer vision[J]. *Acta Automatica Sinica*, 2024, 50(7): 1261-1283.
- [8] 李宗祐, 高春艳, 吕晓玲, 等. 基于深度学习的金属材料表面缺陷检测综述[J]. *制造技术与机床*, 2023(6): 61-67.
- [9] LI Z Y, GAO C Y, LV X L, et al. A review of surface defect detection for metal materials based on deep learning[J]. *Manufacturing Technology & Machine Tool*, 2023(6): 61-67.
- [10] 王慧慧, 杨长辉, 吕庆. 基于机器视觉的金属表面缺陷检测方法综述[J]. *微纳电子与智能制造*, 2022, 4(4): 71-81.
- [11] WANG H J, YANG C H, LYU Q. Review of metal surface defect detection methods based on machine vision[J]. *Micro-*

- Nano Electronics and Intelligent Manufacturing, 2022, 4(4): 71-81.
- [9] 费佳杰, 李宏胜, 任飞, 等. 基于深度学习的钢表面缺陷检测方法综述[J]. 现代信息科技, 2023, 7(19): 107-112.
FEI J J, LI H S, REN F, et al. Overview of steel surface defect detection methods based on deep learning[J]. Modern Information Technology, 2023, 7(19): 107-112.
- [10] EVERINGHAM M, ALI ESLAMI S M, VAN GOOL L, et al. The pascal visual object classes challenge: a retrospective [J]. International Journal of Computer Vision, 2015, 111(1): 98-136.
- [11] DENG J, DONG W, SOCHER R, et al. ImageNet: a large-scale hierarchical image database[C]//Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2009: 248-255.
- [12] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: common objects in context[C]//Proceedings of the 13th European Conference on Computer Vision. Cham: Springer, 2014: 740-755.
- [13] DOLLAR P, WOJEK C, SCHIELE B, et al. Pedestrian detection: a benchmark[C]//Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2009: 304-311.
- [14] YANG S, LUO P, LOY C C, et al. Wider Face: a face detection benchmark[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 5525-5533.
- [15] 徐洪俊, 唐自强, 张锦东, 等. 钢材表面缺陷检测的YOLOv5s算法优化研究[J]. 计算机工程与应用, 2024, 60(7): 306-314.
XU H J, TANG Z Q, ZHANG J D, et al. Research on optimization of YOLOv5s detection algorithm for steel surface defect[J]. Computer Engineering and Applications, 2024, 60(7): 306-314.
- [16] XU J, HU X, ZHANG Y, et al. BHE-YOLO: effective small target detector for aluminum surface defect detection[J]. Advanced Theory and Simulations, 2024, 7(1): 2300563.
- [17] 苏佳, 贾泽, 秦一畅, 等. 面向工业表面缺陷检测的改进YOLOv8算法[J]. 计算机工程与应用, 2024, 60(14): 187-196.
SU J, JIA Z, QIN Y C, et al. Improved YOLOv8 algorithm for industrial surface defect detection[J]. Computer Engineering and Applications, 2024, 60(14): 187-196.
- [18] GAN J, LI Q, WANG J, et al. A hierarchical extractor-based visual rail surface inspection system[J]. IEEE Sensors Journal, 2017, 17(23): 7935-7944.
- [19] JI T, ZHAO Q, AN K, et al. A real-time PCB defect detection model based on enhanced semantic information fusion[J]. Signal, Image and Video Processing, 2024, 18(6): 4945-4959.
- [20] CORTES C, VAPNIK V. Support-vector networks[J]. Machine Learning, 1995, 20(3): 273-297.
- [21] BOMBARA G, VASILE C I, PENEDO F, et al. A decision tree approach to data classification using signal temporal logic[C]//Proceedings of the 19th International Conference on Hybrid Systems: Computation and Control. New York: ACM, 2016: 1-10.
- [22] FREUND Y, SCHAPIRE R E. A decision-theoretic generalization of on-line learning and an application to boosting[J]. Journal of Computer and System Sciences, 1997, 55(1): 119-139.
- [23] VIOLA P, JONES M. Rapid object detection using a boosted cascade of simple features[C]//Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2001.
- [24] LINDEBERG T. Scale invariant feature transform[J]. Scholarpedia, 2012, 7(5): 10491.
- [25] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]//Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2005: 886-893.
- [26] BAY H, TUYTELAARS T, VAN GOOL L. SURF: speeded up robust features[C]//Proceedings of the 9th European Conference on Computer Vision. Berlin, Heidelberg: Springer, 2006: 404-417.
- [27] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2014: 580-587.
- [28] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [29] GIRSHICK R. Fast R-CNN[C]//Proceedings of the 2015 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2015: 1440-1448.
- [30] 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.
- [31] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]//Proceedings of the 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2017: 2980-2988.
- [32] LU X, LI B Y, YUE Y X, et al. Grid R-CNN[C]//Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2019: 7355-7364.
- [33] 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.
- [34] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 6517-6525.
- [35] REDMON J, FARHADI A. YOLOv3: an incremental improvement[J]. arXiv:1804.02767, 2018.
- [36] 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. Piscataway: IEEE, 2017: 936-944.
- [37] BOCHKOVSKIY A, WANG C Y, LIAO H M. YOLOv4: optimal speed and accuracy of object detection[J]. arXiv: 2004.10934, 2020.
- [38] GE Z, LIU S, WANG F, et al. YOLOX: exceeding YOLO series in 2021[J]. arXiv:2107.08430, 2021.
- [39] LI C, LI L, JIANG H, et al. YOLOv6: a single-stage object detection framework for industrial applications[J]. arXiv: 2209.02976, 2022.
- [40] WANG C Y, BOCHKOVSKIY A, LIAO H M. YOLOv7: trainable bag-of-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. Piscataway: IEEE, 2023: 7464-7475.
- [41] WANG C Y, YE H I, LIAO H M. YOLOv9: learning what you want to learn using programmable gradient information [J]. arXiv:2402.13616, 2024.
- [42] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//Proceedings of the 14th European Conference on Computer Vision. Cham: Springer, 2016: 21-37.
- [43] FU C Y, LIU W, RANGA A, et al. DSSD: deconvolutional single shot detector[J]. arXiv:1701.06659, 2017.
- [44] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2017: 2999-3007.
- [45] WU B, WAN A, IANDOLA F, et al. SqueezeDet: unified, small, low power fully convolutional neural networks for real-time object detection for autonomous driving[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE, 2017: 446-454.
- [46] TIAN Z, SHEN C, CHEN H, et al. FCOS: fully convolutional one-stage object detection[J]. arXiv:1904.01355, 2019.
- [47] DUAN K, BAI S, XIE L, et al. CenterNet: keypoint triplets for object detection[C]//Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2019: 6568-6577.
- [48] TAN M, PANG R, LE Q V. EfficientDet: scalable and efficient object detection[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2020: 10778-10787.
- [49] LIANG J, KONG R, MA R, et al. Aluminum surface defect detection algorithm based on improved YOLOv5[J]. Advanced Theory and Simulations, 2024, 7(2): 2300695.
- [50] XIAO D, XIE F T, GAO Y, et al. A detection method of spangle defects on zinc-coated steel surfaces based on improved YOLO-v5[J]. The International Journal of Advanced Manufacturing Technology, 2023, 128(1/2): 937-951.
- [51] CHEN F, DENG M, GAO H, et al. NHD-YOLO: improved YOLOv8 using optimized neck and head for product surface defect detection with data augmentation[J]. IET Image Processing, 2024, 18(7): 1915-1926.
- [52] CHENG Y, WANG S. Improvements to YOLOv4 for steel surface defect detection[C]//Proceedings of the 2022 5th International Conference on Intelligent Autonomous Systems. Piscataway: IEEE, 2022: 48-53.
- [53] REN F, FEI J, LI H, et al. Steel surface defect detection using improved deep learning algorithm: ECA-SimSPPF-SIoU-Yolov5 [J]. IEEE Access, 2024, 12: 32545-32553.
- [54] YU J, SHI X, WANG W, et al. LCG-YOLO: a real-time surface defect detection method for metal components[J]. IEEE Access, 2024, 12: 41436-41451.
- [55] JIA H, ZHOU H, CHEN Z, et al. Research on bearing surface scratch detection based on improved YOLOv5[J]. Sensors, 2024, 24(10): 3002.
- [56] ZHANG H, MIAO Q, LI S, et al. An efficient and real-time steel surface defect detection method based on single-stage detection algorithm[J]. Multimedia Tools and Applications, 2024. DOI:10.1007/s11042-024-19565-2.
- [57] 朱德平, 程光, 姚景丽. 改进YOLOv5的汽车齿轮配件表面缺陷检测[J]. 计算机工程与应用, 2024, 60(5): 321-327.
- ZHU D P, CHENG G, YAO J L. Improved YOLOv5 model for surface defect detection of automotive gear components [J]. Computer Engineering and Applications, 2024, 60(5): 321-327.
- [58] LI C, XU A, ZHANG Q, et al. Steel surface defect detection method based on improved YOLOX[J]. IEEE Access, 2024, 12: 37643-37652.
- [59] 李思雨. 基于改进Yolov5的铝型材表面缺陷检测方法研究[D]. 西安: 西安理工大学, 2023.
- LI S Y. Research on surface defect detection method of aluminum profile based on improved Yolov5[D]. Xi'an: Xi'an University of Technology, 2023.
- [60] ZHAO C, SHU X, YAN X, et al. RDD-YOLO: a modified YOLO for detection of steel surface defects[J]. Measurement, 2023, 214: 112776.
- [61] 高春艳, 秦桑, 李满宏, 等. 改进YOLOv7算法的钢材表面缺陷检测研究[J]. 计算机工程与应用, 2024, 60(7): 282-291.
- GAO C Y, QIN S, LI M H, et al. Research on steel surface defect detection with improved YOLOv7 algorithm[J]. Computer Engineering and Applications, 2024, 60(7): 282-291.
- [62] 赵泽民. 基于改进YOLOv6的轧钢表面细小缺陷检测研究与应用[D]. 太原: 中北大学, 2023.
- ZHAO Z M. Research and application of small defects detection on steel rolling surface based on improved YOLOv6 [D]. Taiyuan: North University of China, 2023.
- [63] LU M, SHENG W, ZOU Y, et al. WSS-YOLO: an improved industrial defect detection network for steel surface defects [J]. Measurement, 2024, 236: 115060.
- [64] CHEN S, ZHOU F, GAO G, et al. Unleashing the power of AI in detecting metal surface defects: an optimized YOLOv7-tiny model approach[J]. PeerJ Computer Science, 2024, 10: e1727.

- [65] 刘浩翰, 孙铖, 贺怀清, 等. 基于改进YOLOv3的金属表面缺陷检测[J]. 计算机工程与科学, 2023, 45(7): 1226-1235.
LIU H H, SUN C, HE H Q, et al. Metal surface defect detection based on improved YOLOv3[J]. Computer Engineering & Science, 2023, 45(7): 1226-1235.
- [66] XIAN Y, YU Y, LIAN Y, et al. An EA-based pruning on improved YOLOv3 for rapid copper elbow surface defect detection[J]. Engineering Applications of Artificial Intelligence, 2023, 123: 106412.
- [67] 陈明文. 基于深度学习的金属表面缺陷检测方法研究与实现[D]. 大庆: 东北石油大学, 2023.
CHEN M W. Research and implementation of metal surface defect detection method based on deep learning[D]. Daqing: Northeast Petroleum University, 2023.
- [68] MI Z, CHEN R, ZHAO S. Research on steel rail surface defects detection based on improved YOLOv4 network[J]. Frontiers in Neurorobotics, 2023, 17: 1119896.
- [69] SUN S, DENG M, LUO J, et al. ST-YOLO: an improved metal defect detection model based on Yolov5[C]//Proceedings of the 2024 3rd Asia Conference on Algorithms, Computing and Machine Learning. New York: ACM, 2024: 158-164.
- [70] ZHU H, JIANG J, WANG Y. CGS-YOLOv5: a defect detection algorithm for PCB board based on YOLOv5 algorithm [C]//Proceedings of the 2nd International Conference on Physics, Photonics, and Optical Engineering. Bellingham: SPIE, 2024: 320-329.
- [71] ZHU X, LIU J, ZHOU X, et al. Enhanced feature fusion structure of YOLOv5 for detecting small defects on metal surfaces[J]. International Journal of Machine Learning and Cybernetics, 2023, 14(6): 2041-2051.
- [72] WANG L, LIU X, MA J, et al. Real-time steel surface defect detection with improved multi-scale YOLO-v5[J]. Processes, 2023, 11(5): 1357.
- [73] 杜少聪, 张红钢, 王小敏. 基于改进YOLOv5的钢轨表面缺陷检测[J]. 北京交通大学学报, 2023, 47(2): 129-136.
DU S C, ZHANG H G, WANG X M. Rail surface defect detection based on improved YOLOv5[J]. Journal of Beijing Jiaotong University, 2023, 47(2): 129-136.
- [74] ZHANG H, LI S, MIAO Q, et al. Surface defect detection of hot rolled steel based on multi-scale feature fusion and attention mechanism residual block[J]. Scientific Reports, 2024, 14(1): 7671.
- [75] 齐向明, 董旭. 改进Yolov7-tiny的钢材表面缺陷检测算法[J]. 计算机工程与应用, 2023, 59(12): 176-183.
QI X M, DONG X. Improved Yolov7-tiny algorithm for steel surface defect detection[J]. Computer Engineering and Applications, 2023, 59(12): 176-183.
- [76] 窦智, 胡晨光, 李庆华, 等. 改进YOLOv7的小样本钢板表面缺陷检测算法[J]. 计算机工程与应用, 2023, 59(23): 283-292.
DOU Z, HU C G, LI Q H, et al. Improved YOLOv7 algorithm for small sample steel plate surface defect detection [J]. Computer Engineering and Applications, 2023, 59(23): 283-292.
- [77] ZHONG H, FU D, XIAO L, et al. STFE-Net: a multi-stage approach to enhance statistical texture feature for defect detection on metal surfaces[J]. Advanced Engineering Informatics, 2024, 61: 102437.
- [78] YOU C, KONG H. Improved steel surface defect detection algorithm based on YOLOv8[J]. IEEE Access, 2024, 12: 99570-99577.
- [79] FAN J, WANG M, LI B, et al. ACD-YOLO: improved YOLOv5-based method for steel surface defects detection [J]. IET Image Processing, 2024, 18(3): 761-771.
- [80] 胡欣, 周运强, 肖剑, 等. 基于改进YOLOv5的螺纹钢表面缺陷检测[J]. 图学学报, 2023, 44(3): 427-437.
HU X, ZHOU Y Q, XIAO J, et al. Surface defect detection of threaded steel based on improved YOLOv5[J]. Journal of Graphics, 2023, 44(3): 427-437.
- [81] LIU Z, YE K. YOLO-IMF: an improved YOLOv8 algorithm for surface defect detection in industrial manufacturing field [C]//Proceedings of the 19th International Conference on Metaverse. Cham: Springer, 2023: 15-28.
- [82] KONG X, FAN W. Surface defect detection of steel strips based on an improved YOLOv8[C]//Proceedings of the 2024 5th International Seminar on Artificial Intelligence, Networking and Information Technology. Piscataway: IEEE, 2024: 2325-2328.
- [83] 刘艳菊, 王秋霖, 张惠玉, 等. 基于改进SSD的工件表面缺陷检测[J]. 热加工工艺, 2024, 53(2): 134-139.
LIU Y J, WANG Q J, ZHANG H Y, et al. Defect detection of workpiece surface based on improved SSD[J]. Hot Working Technology, 2024, 53(2): 134-139.
- [84] LIU X, GAO J. Surface defect detection method of hot-rolled steel strip based on improved SSD model[J]. EasyChair, 2020, 4681: 1-11.
- [85] LI Y, HE L, ZHANG M, et al. Improving the performance of the single shot multibox detector for steel surface defects with context fusion and feature refinement[J]. Electronics, 2023, 12(11): 2440.
- [86] CUI L, JIANG X, XU M, et al. SDDNet: a fast and accurate network for surface defect detection[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 2505713.
- [87] 阎馨, 杨月川, 屠乃威. 基于改进SSD的钢材表面缺陷检测[J]. 现代制造工程, 2023(5): 112-120.
YAN X, YANG Y C, TU N W. Steel surface defect detection based on improved SSD[J]. Modern Manufacturing Engineering, 2023(5): 112-120.
- [88] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[J]. arXiv:1704.04861, 2017.
- [89] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and

- Pattern Recognition. Piscataway: IEEE, 2018: 4510-4520.
- [90] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3[C]//Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2019: 1314-1324.
- [91] ZHANG X, ZHOU X, LIN M, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices [C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 6848-6856.
- [92] MA N, ZHANG X, ZHENG H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design[C]//Proceedings of the 15th European Conference on Computer Vision. Cham: Springer, 2018: 122-138.
- [93] HAN K, WANG Y, TIAN Q, et al. GhostNet: more features from cheap operations[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2020: 1577-1586.
- [94] XIE Y, HU W, XIE S, et al. Surface defect detection algorithm based on feature-enhanced YOLO[J]. Cognitive Computation, 2023, 15(2): 565-579.
- [95] 李丰润. 基于深度学习的金属表面缺陷检测的研究与实现[D]. 广州: 华南理工大学, 2023.
LI F R. Research and implementation of metal surface defect detection based on deep learning[D]. Guangzhou: South China University of Technology, 2023.
- [96] WU Y, CHEN R, LI Z, et al. SDD-YOLO: a lightweight, high-generalization methodology for real-time detection of strip surface defects[J]. Metals, 2024, 14(6): 650.
- [97] LIU Z, LV H. YOLO_Bolt: a lightweight network model for bolt detection[J]. Scientific Reports, 2024, 14(1): 656.
- [98] ZHANG G, LIU S, NIE S, et al. YOLO-RDP: lightweight steel defect detection through improved YOLOv7-tiny and model pruning[J]. Symmetry, 2024, 16(4): 458.
- [99] 梁礼明, 龙鹏威, 冯耀, 等. 改进轻量化 VTG-YOLOv7-tiny 的钢材表面缺陷检测[J]. 光学精密工程, 2024, 32(8): 1227-1240.
LIANG L M, LONG P W, FENG Y, et al. Improving the lightweight VTG-YOLOv7-tiny for steel surface defect detection[J]. Optics and Precision Engineering, 2024, 32(8): 1227-1240.
- [100] 叶卓勋. 基于深度学习网络的金属小目标外观缺陷检测[D]. 杭州: 浙江大学, 2023.
YE Z X. Appearance defect detection of small metal targets based on deep learning network[D]. Hangzhou: Zhejiang University, 2023.
- [101] 苏盈盈, 何亚平, 邓圆圆, 等. 基于改进 YOLOv5s 轻量化带钢表面缺陷检测方法[J]. 光电子·激光, 2024, 35(7): 723-730.
SU Y Y, HE Y P, DENG Y Y, et al. Lightweight strip surface defect detection method based on improved YOLOv5s [J]. Journal of Optoelectronics · Laser, 2024, 35(7): 723-730.
- [102] 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): 61-73.
LIANG L M, LONG P W, LU B H, et al. Improvement of GBS-YOLOv7t for steel surface defect detection[J]. Opto-Electronic Engineering, 2024, 51(5): 61-73.
- [103] 刘凤春, 张靖, 薛涛, 等. YOLO-VDCW: 一种新的轻量化带钢表面缺陷检测算法[J]. 中国冶金, 2024, 34(6): 125-135.
LIU F C, ZHANG J, XUE T, et al. YOLO-VDCW: a new lightweight surface defect detection algorithm for strip steel[J]. China Metallurgy, 2024, 34(6): 125-135.
- [104] 梁礼明, 龙鹏威, 卢宝贺, 等. EHH-YOLOv8s: 一种轻量级的带钢表面缺陷检测算法[J/OL]. 北京航空航天大学学报 (2024-08-08)[2024-09-15]. <https://link.cnki.net/urlid/11.2625.V.20240807.1403.003>.
LIANG L M, LONG P W, LU B H, et al. EHH-YOLOv8s: a lightweight algorithm for strip surface defect detection [J/OL]. Journal of Beijing University of Aeronautics and Astronautics (2024-08-08)[2024-09-15]. <https://link.cnki.net/urlid/11.2625.V.20240807.1403.003>.
- [105] LU J, YU M, LIU J. Lightweight strip steel defect detection algorithm based on improved YOLOv7[J]. Scientific Reports, 2024, 14(1): 13267.
- [106] YUAN Z, NING H, TANG X, et al. GDGP-YOLO: enhancing steel surface defect detection using lightweight machine learning approach[J]. Electronics, 2024, 13(7): 1388.
- [107] CHEN B, WEI M, LIU J, et al. EFS-YOLO: a lightweight network based on steel strip surface defect detection[J]. Measurement Science and Technology, 2024, 35(11): 116003.
- [108] XIE W, SUN X, MA W. A light weight multi-scale feature fusion steel surface defect detection model based on YOLOv8 [J]. Measurement Science and Technology, 2024, 35(5): 055017.
- [109] 王春梅, 刘欢. YOLOv8-VSC: 一种轻量级的带钢表面缺陷检测算法[J]. 计算机科学与探索, 2024, 18(1): 151-160.
WANG C M, LIU H. YOLOv8-VSC: lightweight algorithm for strip surface defect detection[J]. Journal of Frontiers of Computer Science and Technology, 2024, 18(1): 151-160.
- [110] LIU X, GAO J. Surface defect detection method of hot rolling strip based on improved SSD model[C]//Proceedings of the 2021 International Workshops on Database Systems for Advanced Applications. Cham: Springer, 2021: 209-222.
- [111] 刘凡凡, 朱成梅, 赵娜娜, 等. 基于多模态融合的遥感小目标检测[J]. 激光与光电子学进展, 2024, 61(24): 10.
LIU F F, ZHU C M, ZHAO N N, et al. Remote sensing small target detection based on multi-modal fusion[J]. Laser & Optoelectronics Progress, 2024, 61(24): 10.
- [112] SUN L, WANG J S, WANG J, et al. Efficient joint deployment of multi-UAVs for target tracking in traffic big data [J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(7): 7780-7791.
- [113] PARK B, KIM Y, PARK J, et al. Integrating biomarkers from virtual reality and magnetic resonance imaging for the early detection of mild cognitive impairment using a multimodal learning approach: validation study[J]. Journal of Medical Internet Research, 2024, 26: e54538.