

文章编号 1004-924X(2023)15-2260-13

基于深度学习的金刚线光斑点检测

丰宗强, 应一鹏*, 章甫君, 于勇波, 刘毅
(燕山大学 机械工程学院, 河北 秦皇岛 066004)

摘要: 金刚线断线检测是金刚线生产过程中的重要环节。针对现有接触式检测敏感度低、断线反馈滞后等问题, 提出了一种基于机器视觉检测强光下金刚线反射的光斑点的非接触式断线检测方法。在金刚线光斑点检测的嵌入式平台上, 针对传统图像处理的光斑点检测操作复杂、易受外部光照影响的局限性, 研究了基于深度学习的光斑点目标检测, 对多种 Yolo 系列模型进行了训练部署, 针对原有模型网络层次较深、模型体积较大, 在嵌入式设备中存在检测实时性较差的问题, 提出了一种基于 YoloX 改进的轻量化目标光斑点检测模型 MCA-YoloX, 利用 MobileNetV3 轻量化特征提取网络替换 YoloX 模型的主干特征提取网络, 对模型进行轻量化改进, 然后利用深度可分离卷积和倒残差结构对加强特征提取网络进行了轻量化改进。结合 CA 注意力机制提高了轻量化模型的检测精度。最后, 将改进后的模型部署于嵌入式平台。实验结果表明, 改进后模型 MCA-YoloX 的大小和运算量减小到 YoloX 模型的 1/3 以下, 与同样规模的 YoloX-Tiny 和 YoloV4-Tiny 相比具有更高的检测精度, 模型的 mAP 提升了 1% 以上, 加速优化后检测速度可达 30 frame/s, 提供了一种基于深度学习检测金刚线断线的完整工业检测方案。

关键词: 机器视觉; 金刚线; 断线检测; 光斑点; 深度学习; 嵌入式

中图分类号: TP391.4 **文献标识码:** A **doi:** 10.37188/OPE.20233115.2260

Light spot detection of diamond wire based on deep learning

FENG Zongqiang, YING Yipeng*, ZHANG Fujun, YU Yongbo, LIU Yi

(School of Mechanical Engineering, Yanshan University, Qinghuangdao 066004, China)

* Corresponding author, E-mail: yyp2022111@163.com

Abstract: Break detection is an important part of the diamond wire production process. To address the problems of low sensitivity and lag feedback of existing contact detection, a non-contact wire break detection method is proposed based on machine vision detection of light spots reflected by diamond lines under strong light. Here, the study addresses the limitations of complex spot detection operation and ease of influence from external illumination of traditional image processing by investigating spot target detection on the embedded platform of diamond line spot detection using deep learning. A variety of Yolo-type models were trained and deployed. The problem of poor real-time detection in embedded devices due to the deep network level and large model volume of the original model was also addressed through a lightweight target spot detection model MCA-YoloX based on YoloX. The MobileNetV3 lightweight feature extraction network was used to replace the backbone feature extraction network of the YoloX model, and the model was lightweighted. Then, the enhanced feature extraction network was lightweight using the deep separable convolution and inverted residual structures. Then, combined with the CA attention mechanism, the

收稿日期: 2022-08-27; 修订日期: 2022-10-19.

基金项目: 国家自然科学基金资助项目 (No. U2037202)

detection accuracy of the lightweight model was improved. Finally, the improved model was deployed on the embedded platform. The experimental results show that the size and computation amount of the improved model, MCA-Yolox, are reduced to less than 1/3 those of the Yolox model, and compared with Yolox-Tiny and YOLOv4-Tiny of the same scale, it has higher detection accuracy. The mAP of the model increased by more than 1%, and the detection speed can reach 30 frames/s after accelerated optimization. In summary, this paper presents a complete industrial detection scheme based on deep learning to detect diamond wire breaks.

Key words: machine vision; diamond wire; detection of broken wire; light spot; deep learning; embedded type

1 引 言

金刚线^[1]是硬脆材料的切割工具,广泛应用于宝石、硅片^[2]、太阳能板^[3]的切割作业中。在金刚线生产线^[4]上,任何额外的接触带来的振动或者其他不稳定因素都有可能导致线受力不均而断裂。所以金刚线断裂的快速检测和断线位置的准确捕捉是金刚线生产过程中的重要课题。目前,金刚线断线普遍在生产线线头和线尾采用拉力传感器进行检测,通过测量线的张力来判定金刚石线是否断裂,此方案具有可靠且稳定的优点,但是此方法只有断线位置接近线头线尾时才能触发报警,故受限于金刚线 1.5 m/s 的线速,存在断线反馈滞后导致增加损失且设备成本较高的缺点,原有方案的实际滞后时间为 50~90 s。

由于在激光照射下金刚线网呈现明显的光斑点,故将金刚线是否断线的状态表征为光斑点有无。通过视觉检测金刚线反射的亮斑点达到检测是否断线的目的,将断线检测转化为视觉光斑点目标检测的问题。斑点检测是机器视觉研究的重要内容之一,随着图像处理技术的日趋成熟,斑点检测已经应用于产品缺陷检测^[5]、医学图像检测^[6]、遥感识别^[7]等诸多领域,众多应用方案已经在工业生产与实际生活中实施。针对传统图像处理方法斑点检测的缺点,研究人员将深度学习技术应用到斑点检测任务中,并取得了不错的效果。Xu^[8]等提出 UH-DOG 斑点检测模型,联合 DOG 法和 U-Net 深度学习模型,在细胞斑点检测中取得了优异的检测结果。张轩宇^[9]等则利用 VGGNet 和 ResNe-t34 卷积神经网络实现了遥感图像中船舶、养殖箱等不同形状的斑点物体的检测及分类。Cao^[10]等也利用 YOLOv3 卷积

神经网络结构实现了遥感图像中船舶斑点的识别和检测。

目标检测是相对斑点检测更高的范畴,目标检测的主要任务是确定目标在图像中的位置、尺寸和边缘等信息,这依旧是当前计算机视觉领域充满挑战性的问题^[11]。Redmon^[12]等提出了 YOLO 系列目标检测算法,马立^[13]等针对小目标漏检率高的问题,提出了改进的 Tiny YOLO3 模型,对算法的特征提取网络、预测网络和损失函数进行改进。Ahmad^[14]等将低级特征的特征图与常规 SSD 的反卷积层相结合提出 EDF-SSD 模型。王宸^[15]等利用 YOLOv3 对轮毂焊缝缺陷进行检测。陈仁祥^[16]结合注意力机制提出多注意力的 Faster RCNN 结构对电路板进行缺陷检测。Kumar^[16]等结合空间金字塔操作,对 YOLOv4 模型进行改进,提高了口罩检测的准确率。

本文以生产线上金刚线的反射光斑点为研究对象,在前期算法开发中经过灰度处理、空间滤波和图像增强的方法对图像进行预处理,并分别结合基于阈值的亮斑点分割方法、图像形态学操作对目标亮斑点进行了有效检测,得到了较理想的光斑点分割与检测效果。但由于经典图像处理方法主要利用单通道图像进行检测,且在阈值分割过程中会丢失大量的图像特征,同时依靠人为提取特征进行目标分割,步骤多、程序复杂度高。传统图像检测方法易受环境光照条件的影响,为解决传统图像处理的问题,研究了基于深度学习的光斑点目标检测算法,并对算法存在的网络层次较深、模型体积较大、检测实时性较差的问题,进行轻量化改进和优化,将深度学习技术应用于工业领域的亮斑点检测任务中,简化检测流程,提高断线检测的准确率,并将改进后

的模型部署于嵌入式平台,提出了一套工业落地的深度学习嵌入式金刚线断线检测方案。

2 材料与方 法

2.1 检测原理与智能检测平台

光斑点视觉检测系统原理如图 1 所示。以线扫激光光源向生产线上的金刚石线投射激光,激光照射在金刚石线网上形成一排光斑点,将识别金刚石线的运动状态表征为金刚石线上漫反射形成的亮斑状态,并通过 CCD 摄像头进行图像采集及视觉检测,从而判定金刚石线的工作状态。

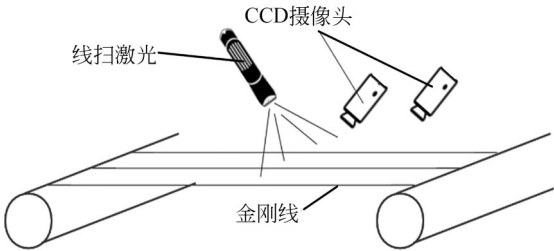


图 1 光斑点视觉检测系统原理
Fig.1 Schematic diagram of light spot vision detection system

检测系统分为图像采集端和控制器端两部分。图像采集端由图像采集和辅助功能组成,辅助功能包括光源以及 LED 交互灯组的控制,图像采集和传输主要依靠 3 个 720P 工业相机模组和有源 USB H-UB 模组。前期研究发现,在绿光下斑点的成像效果最优,故光源选择一字线绿色激光灯,相机使用 LRCP7650_720P 高清面阵工业相机,其最大分辨率为 1 280×720,帧频为 30 frame/s,数据接口为 USB2.0。USB HUB 模组可输出 4 个 USB2.0 接口。

图像采集端以 STM32 作为控制器接收控制信号并控制舵机、LED 灯组和激光笔等从设备,从控制器接受控制器指令调整光源角度和人机交互模组提供良好的图像识别环境。控制器端由 Jetson Nano 嵌入式模组和外围电路组成,控制器端的 Nano 模组接收相机模组的视频流数据并进行目标检测,硬件框架组成如图 2 所示。

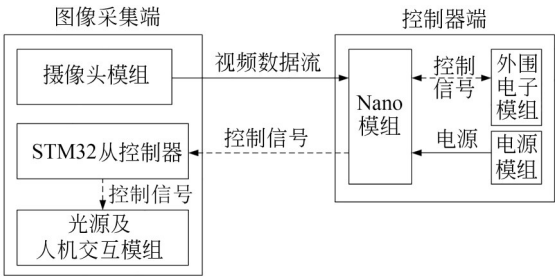


图 2 光斑点视觉检测硬件组成
Fig.2 Hardware composition of light spot vision detection system

如图 3 所示,在金刚线生产线上,系统对金刚线网实时扫描,捕捉实时图像并进行光斑点目标检测,实现对产线上线网状态的过程监测。金刚线网由 63 条金刚线组成,为了满足实际检测工况、成像效果、相机视野覆盖金刚线网等条件,需将金刚线网均分为左、中、右 3 个部分检测,图像采集端左、中、右 3 个相机分别捕获金刚线网的左、中、右区域的图像,即分别对各自区域的 21 个光斑进行检测,并使用黑色光面亚克力板隔除复杂背景,创造良好的成像环境。系统控制端循环获取 3 个相机的图像帧进行目标检测,完成一次金刚线网的断线检测,图 4 为检测系统左、中、右相机采集的原始图像。

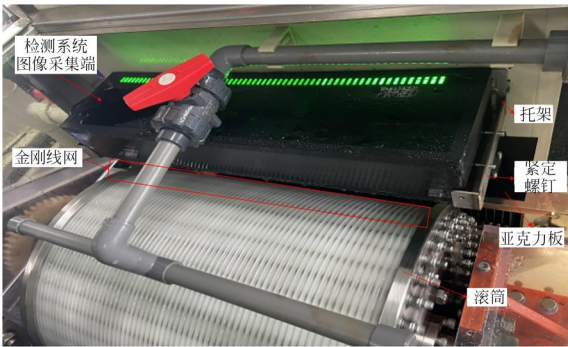


图 3 金刚线断线智能检测平台
Fig.3 Prototype of light spot vision inspection system

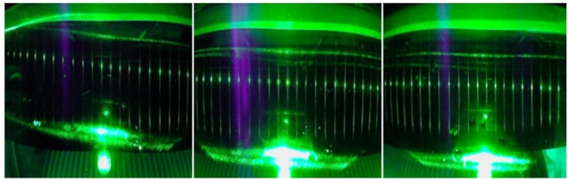


图 4 检测系统左、中、右相机采集的原始图像
Fig.4 Raw images captured by left, center, and right cameras of inspection system

2.2 数据采集与制作

对处于生产状态下的不同产线进行光斑点图像采集,将图像分为多段,针对不同段,选取不同图像阈值,便于标注时斑点的识别同时缩小目标数量,使目标标注的区域选择更加准确。数据增广是深度学习模型训练的常用方法之一,能够有效增加训练数据,避免模型过拟合,同时提升模型的鲁棒性,并采用镜像变换、旋转、透视变换、亮度调整、添加噪声等方法对数据集进行增广,再对数据集进行标注。图 5 为采集图像处理后的部分数据集。

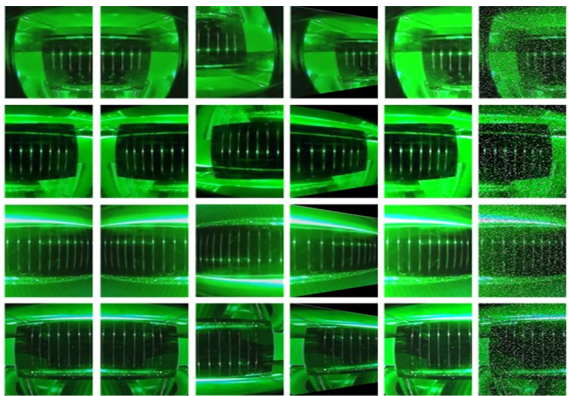


图 5 部分光斑点检测数据集展示
Fig. 5 Partial presentation of blob detection dataset

对实际生产线上采集得到的原始图像进行筛选,滤除斑点代表性较差、图像过度模糊、图像相似度高的图像,并利用数据增广方法进一步扩充数据集,最终挑选出 15 000 张较为理想的图片作为模型训练的数据集,各阶段数据集的参数如表 1 所示。

表 1 数据集参数表			
Tab. 1 Dataset parameter			
原始(张)	初步挑选后 (张)	图像增广后 (张)	目标斑点 (个)
51 563	9 000	15 000	约 105 000

将得到的训练数据集按照 8:1:1 的数量比例随机划分为训练集、验证集和测试集 3 个部分供后续模型训练。其中,训练集 12 000 张,验证集 1 500 张。实验使用图像标注工具 LabelImg,标注图像数据中光斑点的位置,共标注 13 500 张图

像。为进一步验证光斑点目标检测的准确率,在测试数据集中增加了 500 张非金刚线反射形成的光斑杂点作为干扰项,模拟实际产线的成像环境。

2.4 模型训练与部署

本文模型训练及相关实验均在相同的设备上完成,采用的硬件配置和软件版本的具体参数如表 2 所示。模型训练采用迁移学习方法 Pytorch 官方提供的预训练模型进行模型参数初始化,同时采用冻结模型的训练方法,模型总共进行 100 个 Epoch 的训练。在前 50 个 Epoch 训练过程中,先冻结模型的主干特征提取网络的权重系数,使之不进行反向梯度更新,冻结阶段网络训练的 batch_size 为 8,初始学习率为 0.001。在后 50 个 Epoch 的训练过程中,再将主干特征提取网络解冻,此时对整个网络模型参数进行反向梯度更新。由于需要更新的参数量增加,为避免超出内存,减小输入模型的图片 batch_size 为 4,且由于解冻阶段模型更深,故减小初始学习率为 0.000 1。在两个阶段的训练过程中,模型的学习率变化均采用固定步长衰减的策略,学习率每隔一定步数(1 个 Epoch)就衰减为原来的 0.92 倍。

表 2 实验的硬件配置和软件版本参数
Tab. 2 Experiment hardware configuration and software version parameter

硬件或软件版本信息	参 数
CPU	AMD Ryzen 7 3700x
内存	16G
GPU	NVIDIA GeForce RTX 3060Ti
操作系统	Ubuntu18. 04LTS
CUDA	11. 3
CUDNN	8. 0
Pytorch	1. 10. 0
Opencv	4. 5. 4

在相同的软硬件配置下,本文分别对 Yolo v4, Yolo v5 和与前两者模型大小相近的 Yolo x-m 3 个模型进行训练。检测速度和检测平均精度(mAP)的关键指标统计如表 3 所示。

表 3 Yolo 系列光斑点的检测精度与速度

Tab. 3 Spot detection accuracy and speed of Yolo series

Model	mAP/%	Speed/(frame·s ⁻¹)
Yolov4	96.97	2
Yolov5	96.12	3
Yolox-m	97.28	3

由表 3 可见,3 种模型均具有较好的检测效果,mAP 值均在 96% 以上。其中 Yolox-m 的检测精度最高,达到了 97.28%,但三者的检测速度均较低,只有 2~3 frame/s,主要是由于三者模型层次较深且体积较大,实时性较差,需进行轻量化改进。在检测系统的实际应用中,从检测速度和平均精度考虑,选择对 Yolox-m 进行轻量化改进。图 6 为 Yolox-m 模型的光斑点检测结果。

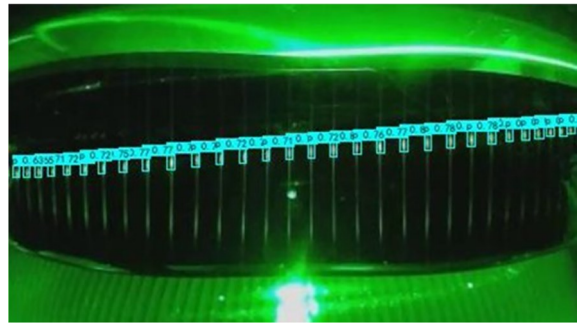


图 6 Yolox-m 模型的光斑点检测结果

Fig. 6 Detection results of light spots in Yolox-m model

3 关键算法

3.1 Yolox 原理

Yolox 目标检测算法是新一代高性能目标检测算法,其算法框架可分为主干特征提取网络、加强特征提取网络和分类器三部分,如图 7 所示。算法框架中,主干特征提取网络为 CSPDarkNet 网络,它主要由 5 个局部跨阶段(Cross Stage Partial,CSP)残差网络块构成。加强特征提取网络 (Feature Pyramid Network, FPN) 主要由空间金字塔操作 (Spatial Pyramid Pooling Network, SPP) 和路径聚合网络 (Path Aggregation Network, PANet) 组成。输入图片经过 CSPDarkNet 和 FPN 得到 3 个特征图完成图像特征提取,最后在 Yolo Head 中完成目标检测。与 Yolov4 最大的不同在于:分类器中 Yolox 将回归预测和置信度预测进行解耦,并采用无锚框(anchor-free)的边框预测方法代替原有的基于锚框(anchor-based)的边框预测方法。

Yolox 采用解耦检测器,模型在检测器中增加预测分支,将边界框的位置及宽高的回归任务和种类置信度预测的二分类任务相分离,去除检测器在预测过程中回归任务和二分类任务可能存在的不利耦合作用,最后再通过通道堆叠操作将输出的 3 个预测三维张量进行通道堆叠。解耦检测器虽然增加了少量的模型计算量,但有效地提高了模型训练的收敛速度和预测精度。解耦检测器结构如图 7 中右侧框检测器所示,其中通道堆叠前的卷积层不执行批归一化和激活函数操作。

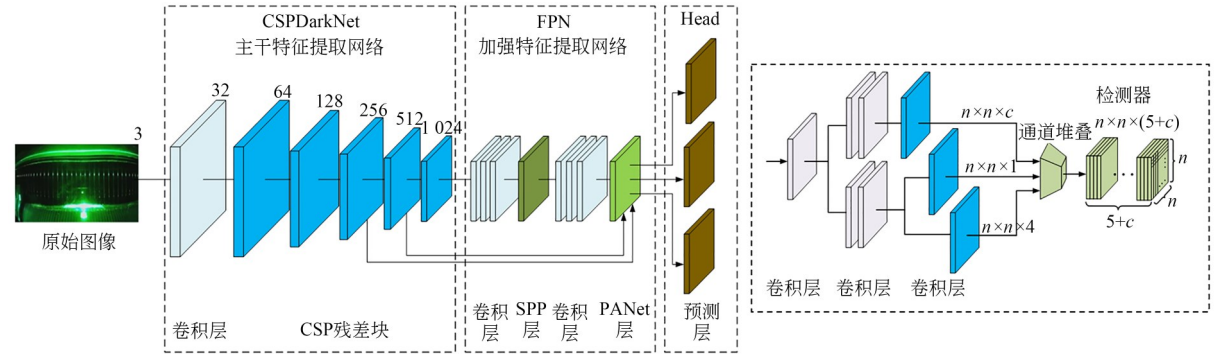


图 7 Yolox 目标检测算法框架

Fig. 7 Schematic diagram Yolox target detection algorithm

相较于早先 Yolo 系列模型的基于 anchor-based 的边界框预测方法, YoloX 采用了目前流行的 anchor-free 的预测框回归策略。在进行正负样本选择时, 基于锚框的边界框预测方法需要进行大量的真实框与锚框的交并比计算。

无锚框回归采用点回归的形式, 在预测目标物体边缘框的过程中不再借助锚框尺寸。特征图中每个栅格只有一个预测点, 根据经过解耦检测器得到的预测参数 t_x, t_y, t_w, t_h 可以确定该预测点的预测框形状。如图 8 所示, 与 YOLOv4 不同, 经过网络预测得到的宽高的预测参数通过指数函数后直接得到预测框的宽高, 而不再是预测框相对于锚框的比例系数。

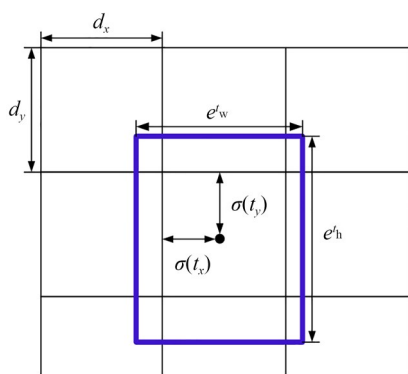


图 8 预测框位置及宽高示意图

Fig. 8 Schematic diagram of position and width and height of prediction box

根据网络所得的预测参数可得预测框的中心点坐标及其宽高, 计算公式如下:

$$\begin{cases} b_x = \sigma(t_x) + d_x \\ b_y = \sigma(t_y) + d_y \\ b_w = e^{t_w} \\ b_h = e^{t_h} \end{cases} \quad (1)$$

其中: b_x, b_y 为预测框中心坐标相对于图像左上角的偏移量, b_w, b_h 为预测框宽高, t_x, t_y 为预测框中心点坐标的预测参数, t_w, t_h 为预测框宽高的预测参数, d_x 和 d_y 为预测点所在栅格左上角点与图像左上角点的距离。

通过解耦检测器可得到 3 个包含预测参数的三维张量, 分别对应加强特征提取网络输出的 3 张特征图, 根据式 (1) 即可利用三维张量中的预测参数得到相应特征图中每个栅格的预测框。若特征图尺寸为 $n \times n$, 则每张特征图可得

到 $n \times n \times 1$ 个待分类样本, 与基于锚框预测的模型相比减少了 2/3 的待分类样本数目。但由于负样本数量相对正样本数量的比例依旧较大, 为了进一步减小正负样本的数量差距, 利用 Simple Optimal Transport Assignment (SimOTA) 算法进行正负样本分类, 增加正样本的数量。在 SimOTA 算法中, 先将真实框内及真实框向外扩展 2.5 个栅格的区域划分为候选区域, 位于候选区域的预测点作为候选点, 候选点对应的预测框为候选框。再计算候选框与其对应的真实框的 IOU 值, 并计算候选框与真实框的代价函数, 代价函数为:

$$C = L_{cls} + x \times L_{reg}, \quad (2)$$

其中: L_{cls} 为候选框的分类误差, L_{reg} 为候选框的回归误差, x 为系数取 0.5。

再由图像中真实框的数量确定参数 y , y 的计算公式如下:

$$y = \max(\text{真实框数量}, 10). \quad (3)$$

取出每个真实框 IOU 最大的前 y 个候选框, 将其 IOU 值相加得到 y' 。将每个真实框的候选框中代价函数最小的前 y' 个候选框作为正样本, y' 最小值取 1, 其余样本为负样本。若一个正样本候选框对应多个真实框, 则该候选框归为代价函数值小的真实框的正样本。

YoloX 中 SimOTA 算法在限制的候选区域内进行正样本分类, 保证了正样本质量, 同时根据算法确定自适应的选取样本数, 选取多个预测框为正样本, 提高正样本的数量。再将其余预测框归为负样本, 由于预测框总数减少了 2/3, 所以有效控制了负样本数量, 进而提高了正负样本比例, 使用于模型训练的正负样本数量更加均衡。同时由于放弃了锚框, 模型不再需要提前指定参考锚框的尺寸, 使模型训练更加简便。

3.2 主干特征提取网络替换

结合深度可分离卷积的模型轻量化方法和倒残差块的结构设计对 YoloX 模型的网络结构进行轻量化改进, 将 YoloX 中的 CSPDarknet 主干特征提取网络替换为采用倒残差块结构的轻量化网络 Mobile-NetV3-Large, 得到初步的轻量化网络模型 YoloX-MobileNetV3。在主干特征提取网络替换前, 删去 MobileNetV3-Large 模型最后的全连接层。YoloX 模型中默认的输入图像尺寸为

640×640,故主干特征提取网络输出的3个特征图的宽高分别为20×20,40×40,80×80,MobileNetv3-Large网络的3个输出特征图的尺寸应与之相同。

3个特征图的输出位置及裁剪后的MobileNetv3-Large网络结构如表4所示。图中,3×3,5×5表示残差块中卷积操作中卷积核尺寸, s 表示卷积操作步长。Scale通道为输出特征图通道,3个输出的特征图尺寸和通道数分别为80×80×40,40×40×112和20×20×160。

表4 裁剪后的MobileNetv3-Large网络结构

Tab.4 Cropped MobileNetv3-Large network structure

网络结构	输出
卷积层,3×3, $s=2$	320×320×16
倒残差块,3×3, $s=1$	320×320×16
倒残差块,3×3, $s=2$	160×160×24
倒残差块,3×3, $s=1$	160×160×24
倒残差块,5×5, $s=1$	80×80×40
倒残差块,5×5, $s=1$	80×80×40
倒残差块,5×5, $s=1$	80×80×40
倒残差块,3×3, $s=2$	40×80×80
倒残差块,3×3, $s=1$	40×80×80
倒残差块,3×3, $s=1$	40×80×80
倒残差块,3×3, $s=1$	40×80×80
倒残差块,3×3, $s=1$	40×40×112
倒残差块,3×3, $s=1$	40×40×112
倒残差块,5×5, $s=2$	20×20×160
倒残差块,5×5, $s=1$	20×20×160
倒残差块,5×5, $s=1$	20×20×160

3.2 加强特征提取网络的改进

在对主干特征提取网络进行轻量化修改后,对加强特征提取网络也进行轻量化修改。由于Yolox模型的加强FPN中也含有大量的卷积网络层和传统残差结构块,为了进一步减小模型大小,还可以对FPN进行轻量化改进。FPN主要由上采样、下采样、普通卷积层和CSP层组成,CSP层为CSP残差块结构。原加强特征提取网络结构如图9所示。

加强特征提取网络的轻量化即将FPN的卷积层中的普通卷积替换成深度可分离卷积,并对CSP层中的卷积结构和残差结构进行改进。

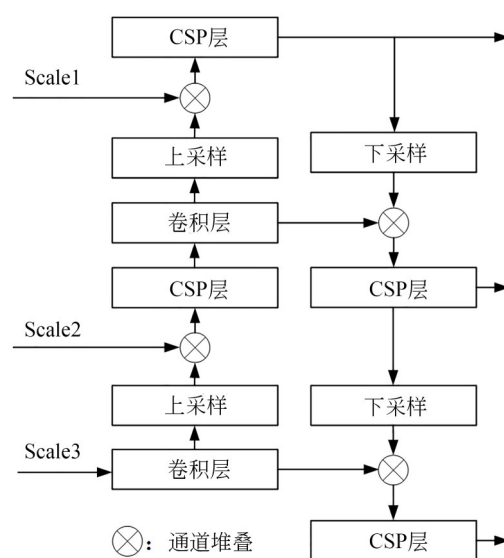


图9 原加强特征提取网络结构

Fig.9 The original enhanced feature extraction network structure

Yolox的FPN中CSP层的轻量化改造如图10所示。为了清晰描述网络结构,省略了批归一化层和激活函数,图中左部分所示网络结构为原始CSP残差块结构;右部分所示为轻量化后的CSP残差块结构,其中卷积层中执行普通卷积操作,DP卷积中执行深度可分离卷积操作, m 表示残差块的堆叠次数。

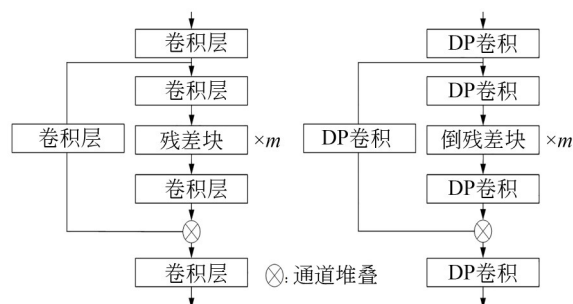


图10 CSP层结构轻量化

Fig.10 Lightweight CSP layer structure

将改进后的加强特征提取网络FPN和改进后的主干特征提取网络进行组合,最终轻量化改进模型结构Yolox-MobileNetV3如图11所示。在最终轻量化改进的模型结构中,将主干特征提取网络和加强特征提取网络均进行轻量化处理,同时对于Yolox模型的Yolo Head检测器,保持其原有的解耦检测和无锚框的检测机制。图中,

Scale 通路表示从修改的主干特征提取网络中输出的 3 个特征图,Scale 通路左部分为采用经过裁剪的 MobileNet 网络替换后的主干特征提取网络,Scale 通路右侧紧接着的是轻量化后的加强特征提取网络,模型的 Yolo Head 检测器依旧保持原有结构。DP 卷积表示深度可分离卷积层,InvCSP 层表示轻量化改进的采用倒残差结构的 CSP 层。

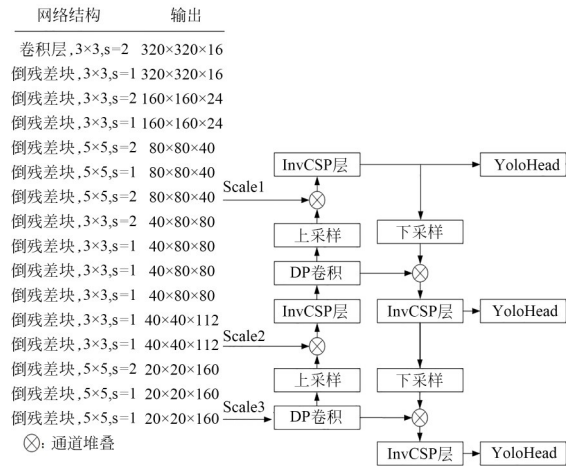


图 11 轻量化改进后的 YOLOX-MobileNetV3 网络模型结构
Fig. 11 Lightweight and improved YOLOX-MobileNetV3 network model structure

3.3 基于 CA 的注意力机制改进

添加注意力机制是提高模型性能的有效手段。由于斑点目标存在较强的空间分布特征且斑点之间存在线性的空间分布规律,故增强网络的空间特征信息的提取能力有助于目标斑点的识别。

在 CA 机制中,为了有效地捕捉位置信息和通道关系。通过将 SE 中全局池化的 $H \times W$ 池化核更换为 $H \times 1$ 和 $1 \times W$ 的两个池化核,将二维全局池化操作分解为两个一维池化编码,沿宽度方向和高度方向的编码方式如下:

$$\begin{cases} z_s^h(h) = \frac{1}{W} \sum_{0 \leq i \leq W} x_s(h, i) \\ z_s^w(w) = \frac{1}{H} \sum_{0 \leq i \leq H} x_s(i, w) \end{cases}, \quad (4)$$

其中: H, W 分别表示特征张量的高和宽, h 表示水平池化所在行高度, w 表示竖直池化所在列的宽度, z_s^h 表示沿高度方向的编码值, z_s^w 表示沿宽度方向的编码值。

CA 模块的添加位置如图 12 所示。为了避免破坏 MobileNet 原有的网络结构导致预训练权重失效,故在主干特征提取网络的 3 个特征图输出通道后添加 CA 模块。同时在加强特征提取网络 FPN 中,在特征图上采样之后也添加 CA 模块以增强网络采样在空间方向上的灵敏度。将经过轻量化及添加 CA 机制的 YOLOX 模型记为 MCA-YOLOX 模型。

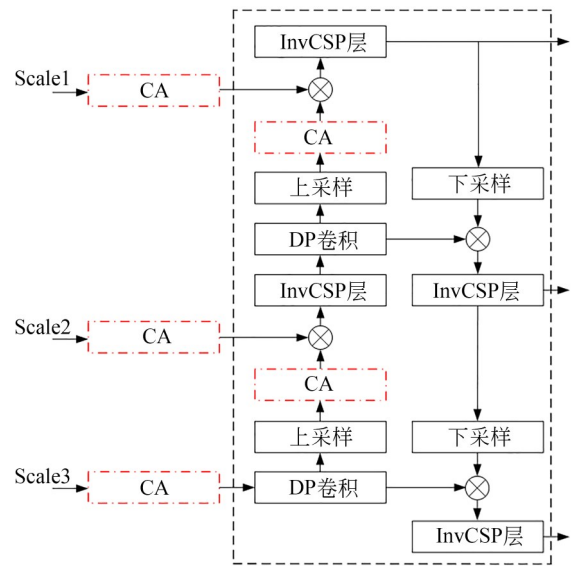


图 12 注意力机制改进网络
Fig. 12 Improved network of attention mechanism

4 实验与结果

4.1 模型评价指标

深度学习中通常用模型的平均精度 (Average Precision, AP) 作为衡量模型目标检测效果的指标。而模型的平均精度受模型的精确度 (Precision) 和召回率 (Recall) 影响。准确率表示预测为真的样本中正样本数的比例,召回率表示在所有真实情况为正的样本中预测为真的样本数的比例。精确度和召回率的计算公式分别为:

$$R = \frac{TP}{TP + FN}, \quad (5)$$

$$P = \frac{TP}{TP + FP}, \quad (6)$$

其中: TP 表示预测为真的正样本 (True Positive), FN 表示预测为假的负样本 (False negative), FP 表示预测为假的正样本 (False positive)。

tive)。AP 是不同召回率下的准确率的平均值, mAP 表示不同类别的 AP 值的平均值。mAP 的计算公式为:

$$mAP = \frac{\sum_{i=1}^n AP_i}{C} = \frac{\sum_{i=1}^n \int_0^1 P_i(R) dR}{C}, \quad (7)$$

其中: P 表示精确度, R 表示召回率, C 表示目标类别数。

F1 分数(F1-score)是 Precision 和 Recall 的调和平均数, 最大为 1, 最小为 0, 分数越高, 说明模型越稳健, 计算公式为:

$$F1 = 2 \times \frac{P \times R}{P + R} = \frac{2TP}{2TP + FN + FP}. \quad (8)$$

由于受到嵌入式平台算力成本限制, 在实际应用场景中除了模型的检测能力, 人们往往更关注模型的复杂度, 通常用模型前向传播的每秒浮点运算次数(Floating-Point Operations Per second, FLOPs)和模型参数个数(Parameters)来描述模型的复杂度。对模型进行实验比较时, 需要结合模型的准确度和复杂度来选择。

4.2 网络轻量化对比

除了 MobileNet 模型, 常见的轻量化网络模型还有 ShuffleNet, GhostNet 和 EfficientNet 等系列。为了验证 YoloX-MobileNetV3 模型的性能, 选取 YoloX-ShuffleNetV2, YoloX-GhostNet 和 YoloX-EfficientNet 3 种轻量化模型, 采取与 YoloX-MobileNet 相同的方式替换 YoloX 模型的主干特征提取网络, 在相同的数据集上进行对比实验。Mobilenet 网络中为了灵活控制模型大小, 设置了两个超参数, 通道因子 α 和分辨率因子 β , 分别用来控制模型中特征张量的通道数和尺寸, 通过因子参数可以控制模型大小。实验中默认通道因子和分辨率因子均为 1.0。

模型训练的学习率依旧采用固定步长衰减的策略, 学习率每隔一定步数(1 个 Epoch)就衰减为原来的 0.92 倍。训练采用先冻结主干特征提取网络, 再解冻的训练步骤, 模型训练 100 个 Epoch, 冻结的训练阶段为前 50 个 Epoch。各模型的 mAP 值默认取 threshold 为 0.5 时的 mAP 值。

如表 5 所示, 与 YoloX-m 模型相比, 4 个轻量化改进模型的参数量、大小和运算量均大幅地减小。YoloX-MobileNetV3 模型的检测精度在 4 个模型中最高, mAP 值达到了 95.02%, 模型的大

小也仅有 20.23 Mb, 在模型的检测准确率和模型存储上都比较优秀。

表 5 不同主干特征提取网络替换模型参数

Tab. 5 Model parameter for different backbone feature extraction network replacement

Model	Parameters/M	Size/Mb	FLOPs /G	mAP/%
YoloX-m	25.28	96.44	36.75	97.28
YoloX-MobileNetV3	5.30	20.23	7.51	95.02
YoloX-ShuffleNetV2	11.83	45.15	9.65	94.43
YoloX-EfficientNet	6.84	26.10	6.13	94.04
YoloX-GhostNet	5.03	19.20	6.87	92.86

训练过程中, 4 种轻量化模型的 mAP 值变化曲线如图 13 所示。图 14 为模型训练后期的 mAP 变化曲线。在模型解冻即训练 50 个 Epoch 后, 模型得到进一步的训练, 模型的 mAP 值有较大的提高。4 种轻量化改进模型在训练 70 个 Epoch 后趋向收敛, mAP 值趋向于稳定。与其他模型相比, 利用 MobilenetV3 轻量化改进的模型在模型训练后期能够达到更高的 mAP 值, 体现出 MobileNetV3 的改进模型相比于其他几个模型在该斑点数据集上具有更高的检测准确率。

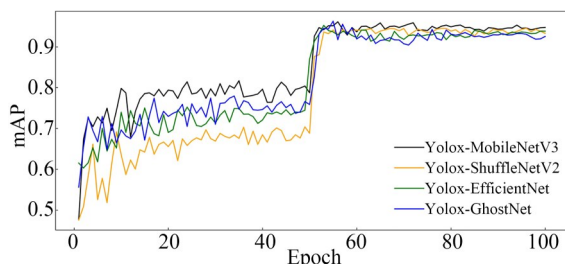


图 13 模型训练过程中 mAP 值

Fig. 13 mAP in target detection of light spot

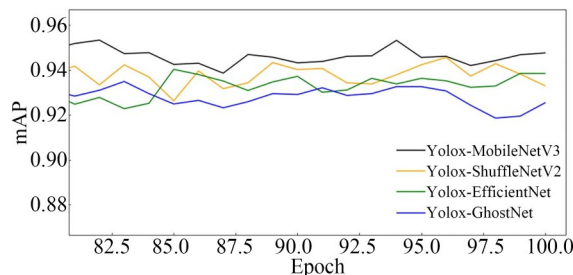


图 14 训练后期模型 mAP 值

Fig. 14 mAP value of post-training model

4.3 改进模型性能验证

为了验证改进模型的检测精度,同时为了进一步验证添加CA机制对提高模型检测精度的有效性,对 YoloX-MobileNetV3, MCA-YoloX, YoloX-Tiny 和 YOLOv4-Tiny 的检测精度进行比较。

4个模型训练过程中的 mAP 曲线、F1 值曲线分别如图 15 和图 16 所示。4个模型在模型解冻后,模型的 mAP 值和 F1 值均有较大的提升,说明模型得到进一步收敛,模型的检测精度进一步得到提升。与其他模型相比,在训练过程中改进模型 MCA-YoloX 均保持较高的 mAP 值,说明改进模型具有较高的检测精确度。由实验数据得到, MCA-YoloX 模型的 mAP 值达到 95.43% 左右, YoloX-MobileNetV3, YoloX-Tiny 和 YOLOv4-Tiny 的 mAP 值分别在 94.52%, 94.24% 和 92.52% 左右。在模型 F1 值上,改进模型比其他 3 种模型更高,说明模型同时具备更高的 Precision 和 Recall 值,模型检测性能更加稳定。MCA-YoloX 模型的 F1 值达到 0.952 左右, YoloX-MobileNetV3, YoloX-Tiny 和 YOLOv4-Tiny 的 F1 值分别在 0.943, 0.922 和 0.922 左右。

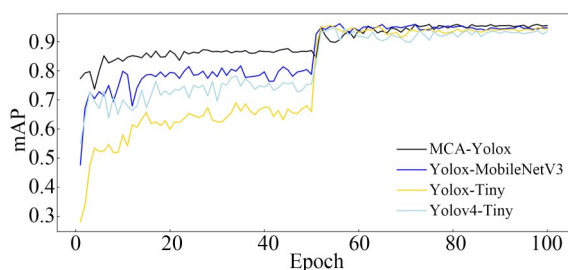


图 15 模型训练的 mAP 曲线

Fig. 15 mAP curve of model training

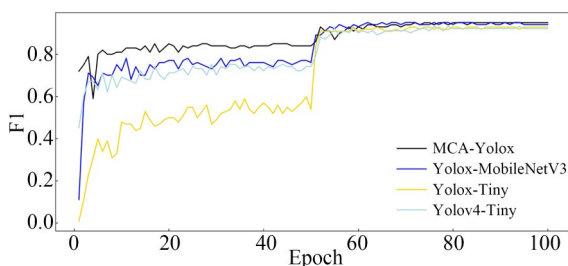


图 16 模型训练的 F1 曲线

Fig. 16 F1 curve of model training

模型的 Precision-Recall (PR) 曲线如图 17 所示,可验证模型的 mAP 和 F1 曲线的特征。改进

模型 MCA-YoloX 的 PR 曲线基本将其他模型的曲线包围,曲线与横纵坐标轴包围的面积比另外 3 个模型的面积都要大,说明改进模型具有较高的 mAP 值,曲线的右上角点同时满足更高的 Precision 和 Recall 坐标值,说明模型具有更高的 F1 值。综上,从模型的 mAP 值和 F1 值参数上比较,改进模型 MCA-YoloX 模型具有最好的检测精度和准确率。同时, MCA-YoloX 模型比 YoloX-MobileNetV3 模型在 mAP 值上提升了约 1%,验证了添加 CA 机制对模型性能的提升。

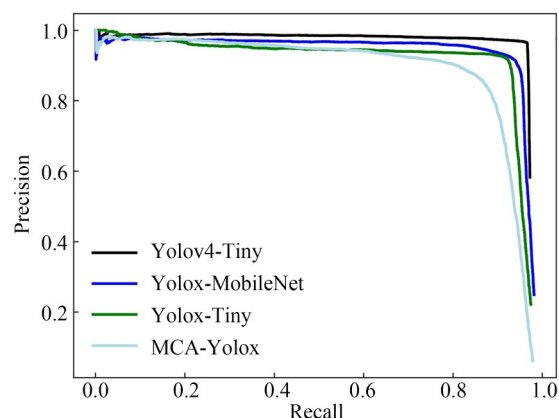


图 17 模型训练的 PR 曲线

Fig. 17 PR curve of model training

各模型在 Jetson Nano 平台上的实验参数如表 6 所示。经过模型的轻量化,改进后的 MCA-YoloX 模型的参数量仅占到 YOLOv4 模型的 1/13 左右。在模型大小和运算量方面, MCA-YoloX 也只有 YOLOv4 模型的 1/10 左右。与 YOLOv4-Tiny 和 YoloX-Tiny 模型相比, MCA-YoloX 在模型大小和运算量相差不多的情况下, mAP 提升了 1% 以上,体现出更佳的性能。

YOLOv4-Tiny, YoloX-Tiny 和改进模型 MCA-YoloX 在 Jetson Nano 平台上的检测结果如图 18 所示,可以看到 YoloX-Tiny 和 YOLOv4-Tiny 模型对于金刚线网边缘较暗的光斑点都有出现漏检的现象, YOLOv4-Tiny 对于金刚线网中间形状较不规则的亮斑点也出现了漏检,而改进的 MCA-YoloX 模型则成功地检测出图像中的所有亮斑点。同时 YoloX-Tiny 和 YOLOv4-Tiny 模型检测出的亮斑点的置信度都在 0.60~0.75 之间,而 MCA-YoloX 模型检测出的亮斑点的置信度都高

表 5 模型实验参数
Tab. 5 Parameters of model experiment

Model	Parameters/M	Size/Mb	FLOPs/G	mAP/%	F1	Speed/(frame·s ⁻¹)
Yolov4-Tiny	5.87	22.42	6.82	92.52	0.922	15
Yolox-Tiny	5.03	19.20	7.57	94.24	0.922	13
Yolox-MobileNetV3	5.30	20.23	7.51	94.52	0.943	13
MCA-Yolox	5.65	21.58	7.64	95.43	0.952	12
Yolov4	64.36	245.53	71.40	97.28	0.972	1
Yolox	25.28	96.44	36.75	96.97	0.963	2

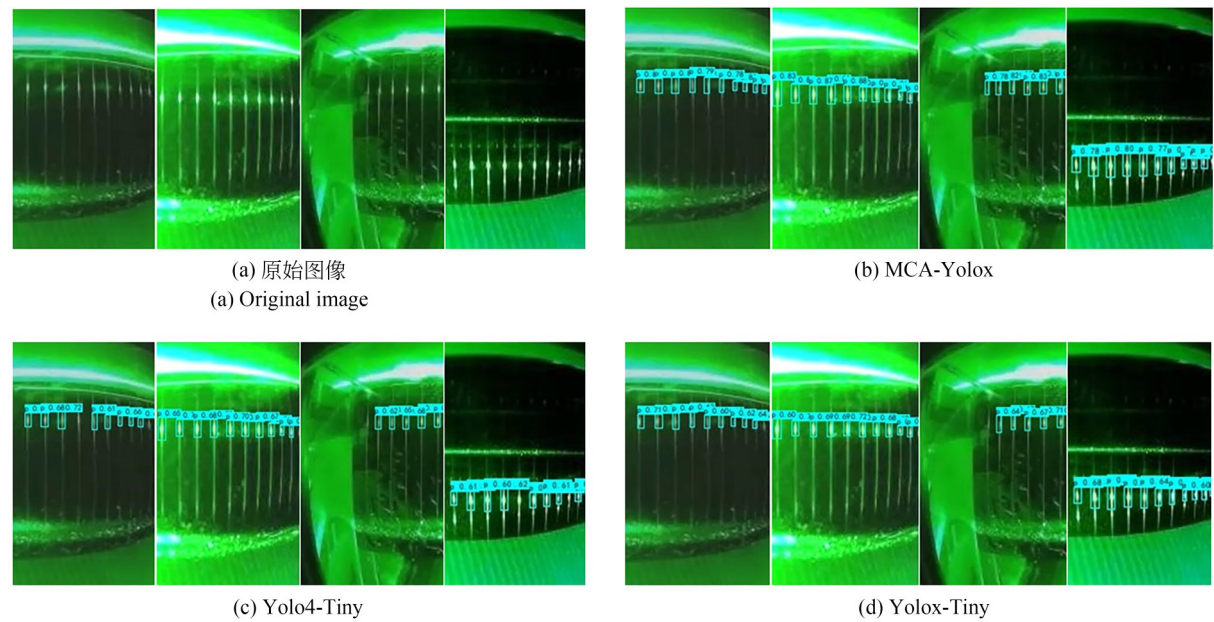


图 18 各模型检测效果对比
Fig. 18 Comparison of detection effect of each model

于 0.75,体现出 MCA-Yolox 模型更高的检测精度和稳定性。轻量化检测模型中 Yolov4-Tiny 模型的检测速度最快达到 15 frame/s,MCA-Yolox 模型也达到 12 frame/s。

最后,利用 TensorRT 库对改进后的模型进行加速优化,随后将优化后的模型部署于智能检测嵌入式平台,经过 TensorRT 加速后的 MCA-Yolox 模型的检测速度能够达到 30 frame/s 左右,若不考虑系统图像传输时间,系统可在 0.1 s 左右内完成 3 个相机捕获的图像帧斑点检测,并进行断线异常报警。从模型存储空间、检测速度和检测准确率等方面考虑,改进的 MCA-Yolox 模型在满足检测实时性的要求上,

检测准确率较高,达到了理想的光斑点检测效果。

5 结 论

本文针对传统图像处理算法易受外部环境干扰,实时性较差等问题,引入深度学习算法对光斑进行检测,并提出了一种改进的轻量化目标检测模型 MCA-Yolox,对 Yolox 模型提出了 3 点改进策略:利用 MobileNetV3 轻量化特征提取网络替换 Yolox 模型的主干特征提取网络对模型进行了轻量化改进;利用深度可分离卷积和倒残差结构对加强特征提取网络进行了轻量化改进;对模型添加 CA 机制,提高了模型的检测精度。

实验结果表明,改进后的MCA-Yolox模型的大小和运算量减小到Yolov4和Yolox模型的1/3以下,相比同样规模的Yolov4-Tiny和Yolox-Tiny轻量化模型具有更高的检测精度,其mAP值高于Yolov4-Tiny约2.9%,高于Yolox-Tiny约

1.2%。在Jetson Nano设备上的检测实验中,MCA-Yolox模型体现出最好的亮斑点检测效果,并达到12 frame/s的检测速度。最后,将加速优化后的模型部署于智能检测平台,检测速度可达30 frame/s。

参考文献:

- [1] 王宝玉,宋为,张太超. 电镀金刚线技术探讨[J]. 金属制品, 2018, 44(3): 10-13.
WANG B Y, SONG W, ZHANG T CH. Discussion on electroplating diamond wire technology[J]. *Metal Products*, 2018, 44(3): 10-13. (in Chinese)
- [2] 孟雪,李和胜. 太阳能硅片切割用金刚线发展评述[J]. 超硬材料工程, 2019, 31(1): 46-51.
MENG X, LI H SH. Review on diamond wires for slicing solar silicon wafers[J]. *Superhard Material Engineering*, 2019, 31(1): 46-51. (in Chinese)
- [3] 赵雷,李欢,胡孝伟. 金刚线在硅晶体切割领域的应用研究[J]. 电子工业专用设备, 2019, 48(3): 33-36.
ZHAO L, LI H, HU X W. Research on the application of diamond wire in single polysilicon cutting[J]. *Equipment for Electronic Products Manufacturing*, 2019, 48(3): 33-36. (in Chinese)
- [4] ZHANG M, LIU Z D, PAN H W, *et al.* Experimental research on combined processing of diamond wire saw and molybdenum twisted wire[J]. *The International Journal of Advanced Manufacturing Technology*, 2019, 101(9/10/11/12): 2751-2759.
- [5] 梁志敏. 视频图像质量检测技术研究与实践[D]. 北京: 北京邮电大学, 2013: 5-80.
LIANG ZH M. *Research and Practice of Video Image Quality Detection Technology* [D]. Beijing: Beijing University of Posts and Telecommunications, 2013: 5-80. (in Chinese)
- [6] 葛乃红. 二维斑点追踪超声心动图与心电图对心肌梗死的诊断价值对比[J]. 影像研究与医学应用, 2021, 5(23): 154-155.
GE N H. Comparison of diagnostic value between two-dimensional speckle tracking echocardiography and electrocardiogram in myocardial infarction[J]. *Journal of Imaging Research and Medical Applications*, 2021, 5(23): 154-155. (in Chinese)
- [7] 唐沐恩,林挺强,文贡坚. 遥感图像中舰船检测方法综述[J]. 计算机应用研究, 2011, 28(1): 29-36.
TANG M E, LIN T Q, WEN G J. Overview of ship detection methods in remote sensing image[J]. *Application Research of Computers*, 2011, 28(1): 29-36. (in Chinese)
- [8] XU Y Z, WU T, GAO F, *et al.* Improved small blob detection in 3D images using jointly constrained deep learning and Hessian analysis[J]. *Scientific Reports*, 2020, 10: 326.
- [9] 张轩宇. 遥感影像海上船舶检测技术研究[D]. 大连: 大连理工大学, 2021: 5-15.
ZHANG X Y. *Research on Detection Technology of Remote Sensing Images for Ships at Sea* [D]. Dalian: Dalian University of Technology, 2021: 5-15. (in Chinese)
- [10] CAO C Q, WU J, ZENG X D, *et al.* Research on airplane and ship detection of aerial remote sensing images based on convolutional neural network[J]. *Sensors*, 2020, 20(17): 4696.
- [11] 范丽丽,赵宏伟,赵浩宇,等. 基于深度卷积神经网络的目标检测研究综述[J]. 光学精密工程, 2020, 28(5): 1152-1164.
FAN L L, ZHAO H W, ZHAO H Y, *et al.* Survey of target detection based on deep convolutional neural networks[J]. *Opt. Precision Eng.*, 2020, 28(5): 1152-1164. (in Chinese)
- [12] REDMON J, DIVVALA S, GIRSHICK R, *et al.* You only look once: unified, real-time object detection[C]. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 779-788.
- [13] 马立,巩笑天,欧阳航空. Tiny YOLOV3目标检测改进[J]. 光学精密工程, 2020, 28(4): 988-995.
MA L, GONG X T, OUYANG H K. Improvement of Tiny YOLOV3 target detection[J]. *Opt. Precision Eng.*, 2020, 28(4): 988-995. (in Chinese)
- [14] AHMAD T, CHEN X N, SAQLAIN A S, *et al.* EDF-SSD: an improved feature fused SSD for ob-

- ject detection [C]. 2021 *IEEE 6th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA)*. April 24-26, 2021, Chengdu, China. IEEE, 2021: 469-473.
- [15] 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 CH, ZHANG XF, LIU CH, *et al.* Detection method of wheel hub weld defects based on the improved YOLOv3[J]. *Optics and Precision Engineering*, 2021, 29(8): 1942-1954. (in Chinese)
- [16] KUMAR A, KALIA A, SHARMA A, *et al.* A hybrid tiny YOLO v4-SPP module based improved face mask detection vision system [J]. *Journal of Ambient Intelligence and Humanized Computing*, 2023, 14(6): 6783-6796.

作者简介:



丰宗强(1990—),男,山西朔州人,硕士,实验师,2017年于燕山大学获得硕士学位,主要从事并联机器人及机电一体化系统的研究。E-mail:zqf@ysu.edu.cn

通讯作者:



应一鹏(1998—),男,浙江金华人,硕士,2022年于燕山大学获得硕士学位,主要从事智能制造及机器视觉的研究。E-mail:yyp2022111@163.com