

基于多尺度改进的YOLOv5电解槽设备及样品检测方法

吴姣姣, 刘铮

引用本文

吴姣姣, 刘铮. 基于多尺度改进的YOLOv5电解槽设备及样品检测方法[J]. 计算机科学, 2023, 50(11A): 230200163-6.

WU Jiaojiao, LIU Zheng. Electrolyzer Equipment and Sample Detection Method Based on Multi-scale Improved YOLOv5 [J]. Computer Science, 2023, 50(11A): 230200163-6.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[一种融合CNN和Swin Transformer的医学显微图像分割模型](#)

Medical Microscopic Image Segmentation Model Based on CNN Structure and Swin Transformer
计算机科学, 2023, 50(11A): 230200119-8. <https://doi.org/10.11896/jsjcx.230200119>

[基于注意力机制和ConvLSTM的船舶交通流量预测算法](#)

Ship Traffic Flow Prediction Algorithm Based on Attention Mechanism and ConvLSTM
计算机科学, 2023, 50(11A): 230800067-7. <https://doi.org/10.11896/jsjcx.230800067>

[基于图卷积网络和注意力机制的诊断预测](#)

Diagnosis Prediction Based on Graph Convolutional Network and Attention Mechanism
计算机科学, 2023, 50(11A): 221100232-6. <https://doi.org/10.11896/jsjcx.221100232>

[基于知识蒸馏和高效通道注意力的异常检测](#)

Novelty Detection Method Based on Knowledge Distillation and Efficient Channel Attention
计算机科学, 2023, 50(11A): 220900034-10. <https://doi.org/10.11896/jsjcx.220900034>

[融合物品关系的图神经网络推荐算法](#)

Graph Neural Network Recommendation Algorithm Based on Item Relations
计算机科学, 2023, 50(11A): 230100019-9. <https://doi.org/10.11896/jsjcx.230100019>

基于多尺度改进的 YOLOv5 电解槽设备及样品检测方法

吴姣姣 刘 铮

长沙理工大学电气与信息工程学院 长沙 410114

(1228815302@qq.com)

摘 要 针对电解铝车间电解槽转运机器人实时识别问题,电解槽设备和铝锭样品的目标检测存在识别物体尺寸差异过大的问题,通常使用的目标检测算法参数较大,部署在电解槽转运机器人上难以达到实时检测的要求。因此,提出一种解决目标尺寸差异过大的轻量化多尺度的 YOLOv5 网络模型,替换主干特征提取网络为轻量化 ShuffleNetV2 网络;添加 SE 注意力机制提高小目标识别准确率;在加强特征提取网络中增加一层浅层检测层作为更小目标的检测层,实现对多尺度以及尺寸变化大的目标的识别准确率。实验结果表明,改进后的 YOLOv5 算法在电解槽转运机器人的电解槽设备和样品识别中物体总类别的平均检测精度为 93.5%,相比 YOLOv5 算法平均检测精度提升了 1.5%,模型参数量降低了约 39.4%,每张图片平均检测速度提升了 2.5 ms,有利于部署到电解槽转运机器人中。

关键词: YOLOv5;轻量化;注意力机制;多尺度目标检测;目标尺寸差异

中图法分类号 TP391

Electrolyzer Equipment and Sample Detection Method Based on Multi-scale Improved YOLOv5

WU Jiaojiao and LIU Zheng

School of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha 410114, China

Abstract Aiming at the real-time recognition problem of electrolytic cell transfer robot in electrolytic aluminum workshop, there is a problem that the size difference of the recognition object is too large for the target detection of electrolytic cell equipment and aluminum ingot samples. Generally, the parameters of the target detection algorithm are large, and it is difficult to meet the requirements of real-time detection when deployed on electrolyzer transfer robots. Therefore, a lightweight multi-scale YOLOv5 network model that solves the excessive difference in target size is proposed, and the backbone feature extraction network is replaced by a lightweight ShuffleNet V2 network. Add SE attention mechanism to improve the accuracy of small target recognition. In the enhanced feature extraction network, a shallow detection layer is added as the detection layer for smaller targets to achieve the recognition accuracy of multi-scale and large size changes. Experimental results show that the average detection accuracy of the improved YOLOv5 algorithm in the electrolyzer equipment and sample identification of the electrolyzer transfer robot is 93.5%, which is 1.5% higher than the average detection accuracy of the YOLOv5 algorithm, the number of model parameters reduces by about 39.4%, and the average detection speed of each picture increases by 2.5 milliseconds, which is conducive to deployment to the Electrolyzer transfer robot.

Keywords YOLOv5, Lightweight, Attention mechanism, Multi-scale object detection, Target size difference

1 引言

电解槽是电解铝车间的主要设备,环境存在磁场强、温度高、噪声大等因素,导致人工对电解铝车间的铝锭样品转运存在效率低、危险系数大等问题,因此有必要研制电解槽转运机器人,代替人工对铝锭样品的运输。

电解槽转运机器人的转运过程中需要实现对电解槽设备和铝锭样品的实时检测,完成铝锭样品的抓取和转运。目前,电解槽设备和样品铝锭的检测存在的问题有:1)检测目标尺寸差异过大,导致目标检测精准度过低;2)模型网络复杂度高,难以达到转运机器人对电解槽设备的实时检测。

近年来,计算机视觉中出现了很多基于深度学习的目标检测算法,同时也有很多针对识别物体尺寸差异过大的多尺度识别的目标检测算法。多尺度特征融合的检测网络的代表

是 2016 年 Liu 等提出的 SSD^[1],其利用各个不同阶段的特征图分别负责不同大小的检测物体。目前应用最为广泛的是自上而下单融合的 FPN^[2],其在很多经典目标检测网络中都有所应用,如双阶段目标检测网络 Faster RCNN^[3]和单阶段目标检测网络 YOLO^[4-5]系列。在 FPN 提出后,PA-Net 提出自顶向下、自底向上的双向融合模型,用于缩短层之间的路径^[6]。2020 年, Tan 等提出 BiFPN 加权的特征金字塔网络,可以简单、快速地进行多尺度特征融合^[7]。多尺度的特征融合一方面增加了目标识别的精度,一方面也增加了设备的计算量和计算的存储空间。鉴于轻量级网络的发展,如 SqueezeNet^[8]是公认轻量级模型的最早设计之一,通过堆叠包含了 squeeze 部分和 expand 部分在内的 fire module 大大减少了参数。ShuffleNet^[9]利用 group convolution 和 channel shuffle 两个操作设计卷积神经网络模型,使得模型参数得到

较大程度的下降,加速了模型的推理。Xception^[10]与 MobileNetv2^[11]同样采用了深度可分离卷积,主要是在 Inception-Netv3^[12]的基础上引入了 depthwise separable convolution。因此,本文在轻量化网络的基础上实现多尺度特征融合。

本文提出采用轻量化网络 ShuffleNetV2 代替 YOLOv5 的主干特征提取网络 CSPDarknet53,保证更好地在嵌入式设备上部署;根据电解槽设备和铝锭样品的尺寸识别差异,对于铝锭较小目标的识别,加入 maxpooling 层和 SPPF 模块,同时降低计算成本;引入 SE 注意力机制^[13],进一步提高对小目标铝锭的识别精度。对于电解槽设备大目标和铝锭样品的尺寸差异,在加强特征提取部分,在三尺度的基础上,再增加一层浅层检测层,当输入图片尺寸大小是 640×640 时,输出 4 个特征层,分别是 20×20,40×40,80×80,160×160,20×20 用于预测电解槽主体,160×160 用于预测样品铝锭,提高改进 YOLOv5 网络对电解槽设备和样品识别的准确率。

2 改进的 YOLOv5 模型原理

本文提出的改进的 YOLOv5 目标检测算法结构图如图 1 所示,整体结构主要分为 3 个部分:主干特征提取网络(Backbone)、加强特征提取网络(Neck)、预测结果(Predict)。

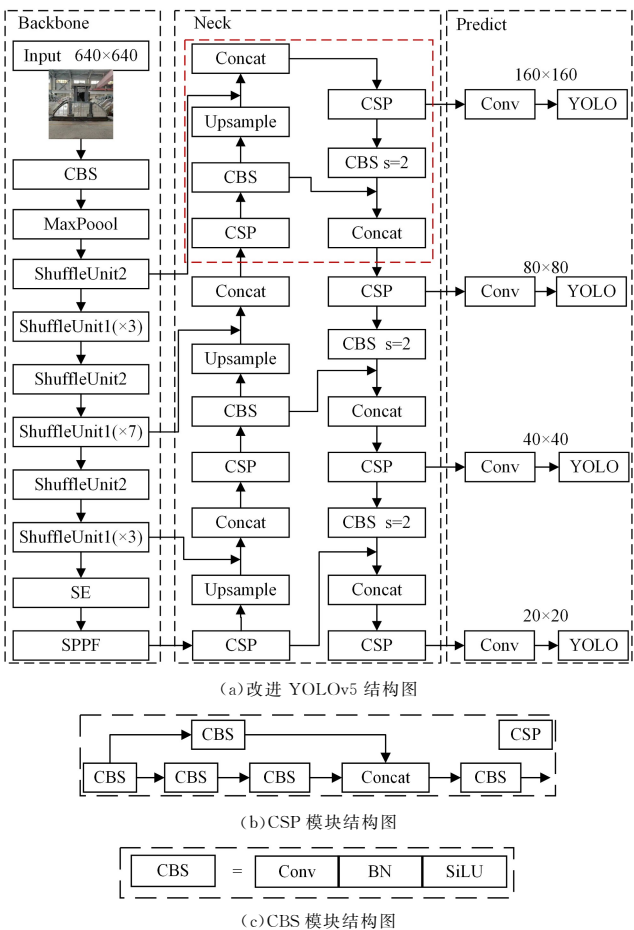


图 1 改进 YOLOv5 结构图(电子版为彩图)

Fig. 1 Improved YOLOv5 structure diagram

主干特征提取网络(Backbone): YOLOv5 算法使用的 Backbone 是 CSPDarknet53,本文主要组件采用 ShuffleNetV2。首先对输入图像进行 3×3 卷积操作,其中步长(s)为 2;同时删除原始 YOLOv5 主干网络中对输入图像进行的切片(Fous)操作,主要原因是 ShuffleNetV2 采用了逐点分组

卷积和通道重组两种技术,能够有效满足特征提取的通道交换。Fous 模块增加了过多的计算量,不利于轻量化,因此删除了 Fous 模块。其次加入 maxpooling 层对输入图像进行 2 倍的下采样,maxpooling 层主要是打乱输入图像至少一半的信息,有利于提高铝锭样品小目标的识别准确率,同时增加的计算量也比 Fous 模块少得多。在 maxpooling 之后,可以看作执行 3 个阶段,包含来自 ShuffleNetV2 的多个单元。阶段 2 和阶段 4 包含 3 个 ShuffleNetV2 单元,第三阶段有 7 个 ShuffleNetV2 单元。主干网络的末端分别是 SE 通道注意力机制和 SPPF 模块,其中 SE 注意力机制主要是关注上面通道交换后传过来的特征图中并不容易发现的铝锭小目标,增加对铝锭小目标识别的准确率。空间金字塔池化使用的是 SPPF 模块,该模块能在保证图片不失真的情况下,使全连接层的输入的尺寸保持统一,同时能减小模型计算量。

加强特征提取网络(Neck):本文在 YOLOv5 三尺度检测的基础上,增加一层小目标检测层,特征融合部分改为四尺度特征输出,输出部分为四尺度特征层。如图 1(a)所示,Neck 部分的红色虚线内是相应增加的第四尺度的特征融合部分, Predict 部分的 160×160 是增加的第四尺度的检测。

预测层(Predict):获得预测框与得分,对得分进行筛选,得到得分满足置信度的预测框,同时对预测框进行非极大抑制,即筛选出一定区域内属于同一种类得分最大的框。计算损失部分,损失函数采用原 YOLOv5 中的 GIOU_Loss 函数^[14]。

2.1 ShuffleUnit 模块

ShuffleNetV2 网络包含两个基本组件,其中基本单元组件为 ShuffleUnit1,如图 2(a)所示,下采样单元组件为 ShuffleUnit2,如图 2(b)所示。ShuffleNetV2 主要采用了 pointwise-groupconvolution(逐点分组卷积)和 channelshuffle(通道重组)两种技术。分组卷积就是对特征层分组后分别进行卷积。分组卷积可以减少网络的容量,使网络更加轻便。逐点分组卷积就是卷积核为 1×1 的分组卷积,但若只使用单独的逐点分组卷积,分组会导致信息的丢失,使特征提取效果不佳。ShuffleNetV2 加入通道重组,进行通道交换,能够使特征信息更加丰富,提取更多的有效特征,而且目标检测结果更加准确。图 2 中 CBR 中的激活函数是 ReLU 激活函数。

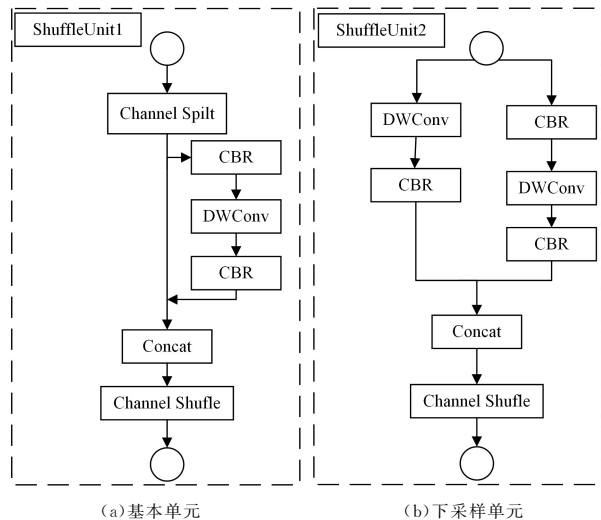


图 2 ShuffleNetV2 网络结构

Fig. 2 ShuffleNetV2 network structure

2.2 SE 注意力机制

SE 注意力机制在通道维度增加注意力机制,可以衡量各特征通道权重。主要操作是压缩(squeeze)和激励(excitation),结构如图3所示。

图3中, X, U, \tilde{X} 表示特征图像, W', H' 和 C' 表示 X 的宽、高和特征通道数; W, H 和 C 表示 U 和 \tilde{X} 的宽、高和特征通道数; F_{tr} 表示将 X 变换到 U 的映射算子; $F_{sq}(\cdot)$, $F_{ex}(\cdot)$, $F_{scale}(\cdot)$ 分别表示压缩函数、激励函数、比例函数。

压缩函数对每个通道特征图进行全局平均池化操作,通道权重 Z_c 的计算公式如式(1)所示,其中 i, j 分别表示每个特征图上 H 和 W 的坐标点。

$$Z_c = F_{tr}(u_c) = \frac{\sum_{i=1}^H \sum_{j=1}^W u_c(i, j)}{H \times W} \quad (1)$$

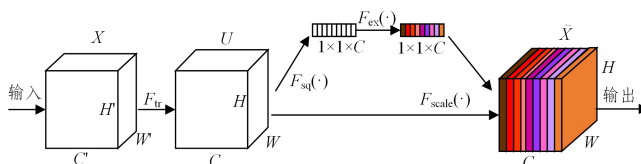


图3 SE 注意力机制结构

Fig. 3 SE attention mechanism structure

2.3 SPPF 空间金字塔池化

SPPF 是 SPP 模块的升级版,二者在精度方面旗鼓相当的情况下,SPPF 的速度有大幅度的提升,两者的结构上有一些细小的差别,主要表现在最大池化层上,SPPF 是将输入的特征图输入3个相同的 5×5 的最大池化层中,再将每一层的结构进行拼接,它在主干网络的末端,在主干网络特征提取向量图进入颈部网络之前,对特征图尺度进行统一,确保图像中的目标位置和图像大小保持相同,从而保证图像的准确性。本文选择在主干特征提取网络的末端加入 SPPF 模块,目的是为了在进入下一步的全连接层的时候,解决全连接输入尺寸不一致的问题。SPPF 模块如图4所示。

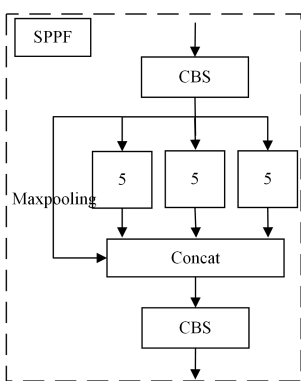


图4 SPPF 结构

Fig. 4 SPP structure

2.4 多尺度改进

YOLOv5 算法采用了 FPN + PAN 网络的多尺度检测^[15],在输入图像尺寸为 640×640 时,算法使用 20×20 的特征层检测大尺寸的目标,使用 40×40 的特征层检测中等尺寸的目标, 80×80 的特征层用于检测小尺寸目标。原YOLOv5通过3个不同尺度可以实现对大、中、小3种尺寸目标的检测,有效解决对于单尺度检测目标识别差异小的问题,大幅度提高检测精度^[16]。

在得到 squeeze 的通道表示后,激励函数采用2层全连接层实现对每个通道重要性的预测,激励函数公式如式(2)所示,其中 W, W_1, W_2 分别为全局平均池化,使用了 ReLU 激活函数、sigmoid 激活函数。

$$S_c = F_{ex}(Z_c, W) = \text{sigmoid}(W_2 \times \text{ReLU}(W_1, Z_c)) \quad (2)$$

最后将上一步生成的权重向量 S_c 对特征图 U 进行权重赋值,得到想要的特征图 \tilde{X} 。加权表达式为式(3),其中 \otimes 表示逐元素相乘, \tilde{X} 表示 SE 模块输出的特征图像。

$$\tilde{X} = F_{scale}(u_c, S_c) \otimes S_c \quad (3)$$

本文在主干网络中引入 SE 通道注意力机制模块,提升主干网络的目标检测精度,同时也是为了确保能提高对铝锭样品小目标识别准确度,在带来极小的参数量的同时较大幅度地提高网络精度。

本文数据集中,电解槽主体和铝锭样品尺寸如表1所列。由表1可知,电解槽主体尺寸远大于铝锭样品,仅有YOLOv5的3个尺度特征,对浅层信息融合不充分,易使部分小目标信息丢失。所以本文将三尺度检测扩展为四尺度检测,即增加 160×160 检测层用来检测本文数据集中的最小目标铝锭。将原算法的特征融合部分也改为四尺度特征融合,增加一层特征层,与FPN网络中上采样后的特征层进行融合输出,得到 160×160 的目标检测层,用于检测更小的目标。通过增加检测的尺度,增大算法对物体检测尺度的范围,算法更加适用于本文的目标物体尺寸差异大的数据集,并且有效地提高了检测精度。本文改进的多尺度特征融合模型如图5所示。

表1 电解槽主体和铝锭样品尺寸

Table 1 Electrolyzer body and aluminium ingot sample size

(单位:cm)		
设备	高	宽
电解槽主体	364	482
铝锭样品	3.4	3.0

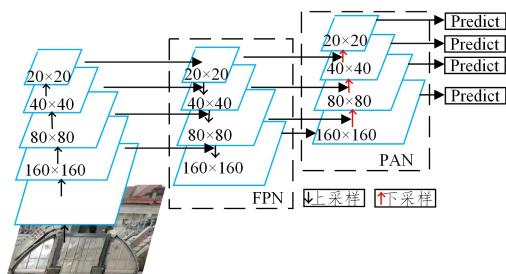


图5 多尺度特征融合模型

Fig. 5 Multi-scale feature fusion model

3 实验评估

3.1 实验数据集制作和实验环境配置

本实验采用的数据集是模拟机器人在巡检过程中采集到

的图片,一共包含 16 种数据集,分别是电解槽主体(electro-bath)、电解槽的母线(母线槽)(mid)、电解槽标牌(num)、电解槽外壳罩板(s_coverplate、coverplate)和样品铝锭(al),其中电解槽标牌中还存在数字 0-9 的标签,电解槽外壳罩板存在两侧和中间部分的区分。

数据集总共包含 2631 张图片,将图片数量按照 9:1 进行训练集和测试集的划分。使用 labelingm 工具对图片目标进行标记,同时生成相对应的 xml 目标框信息位置文件,编写 python 程序将 xml 文件转化为包含文件信息的 txt 文件,作为电解槽设备和样品铝锭的数据集。

实验采用 Win10 的 64 位系统、Pytorch1.6.0 和 Python3.7 环境以及 NVIDIA GeForce RTX 2080Ti 显卡。具体实验环境配置如表 2 所列。

表 2 实验环境配置

Table 2 Experiment environment configuration

名称	配置
显卡	GeForce RTX 2080Ti
GPU 加速库	CUDA10.1,CUDNN 7.6
深度学习框架	Pytorch1.6.0
操作系统	Windows10(64)

3.2 模型训练和评价指标

本文模型训练的 Batch-size 设置为 16,总迭代次数(epoch)是 300 次,选择 Adam 优化算法^[17],前 60 个 epoch 设置学习率为 0.001,后 140 个 epoch 的学习率设置为 0.0001,学习率的调整采用余弦退火的方式^[18]。训练中的损失变化如图 6 所示,其中 train loss 为训练集损失, val loss 为测试集损失。

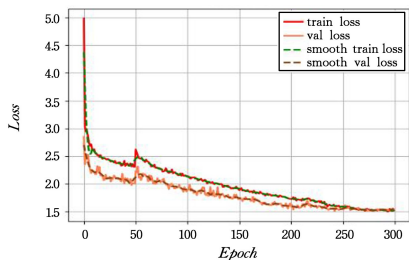


图 6 损失变化过程图

Fig. 6 Loss change process diagram

本文从各类的平均精确率 AP、召回率 R、F1 分数(F1-score)、模型平均检测速度和模型参数量 5 个方面对模型进行评价,同时计算出总数据集的平均精度 mAP。精确率 P 和召回率 R 的计算公式如式(4)所示,AP 和 mAP 的计算公式如式(5)所示,F1-score 的计算公式如式(6)所示。

$$\begin{cases} P = \frac{T_p}{T_p + F_p} \\ R = \frac{T_p}{T_p + F_N} \end{cases} \quad (4)$$

$$\begin{cases} AP = \int_0^1 P(R) dR \\ mAP = \frac{\sum_{i=1}^K AP_i}{K} \end{cases} \quad (5)$$

$$F_1 = 2 \cdot \frac{P \cdot R}{P + R} \quad (6)$$

其中, T_p 为正类预测为正类的数量; F_p 为负类预测为正类的

数量; F_N 为正类预测为负类的数量。F1-score 为精确率和召回率的调和平均数。AP 表示某一类的平均精确度, k 表示待测目标的类别数。mAP 表示所有类 AP 值总和的平均值。图 7 给出了训练过程中的 mAP 值的变化曲线图,可以看出,整个类别的平均精度在迭代 200 次后接近 1,且无明显波动。

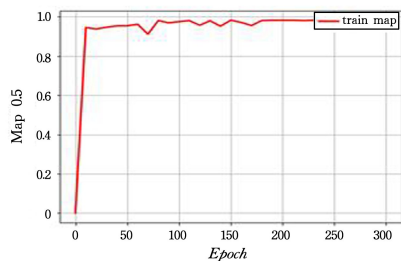


图 7 mAP 曲线图

Fig. 7 mAP graph

3.3 实验结果评估

本文改进的 YOLOv5 模型检测结果如图 8 所示。



(a) 电解槽正面检测结果图

(b) 电解槽侧面检测结果图



(c) 母线槽检测结果图 (d) 电解槽标牌识别结果图 (e) 样品铝锭检测结果图

图 8 改进 YOLOv5 检测结果图

Fig. 8 Improved YOLOv5 test result graph

图 8(a) 电解槽的正面检测结果图包括整个需要识别的所有类别,其中电解槽下半部分的罩壳分为侧边的扇形罩板和中间矩形罩板。图中检测框上的数值表示置信度,即物体属于该类目标的概率,最高为 1。在查准的情况下,置信度越高,代表检测效果越可靠;在误检的情况下,置信度通常较低,表面检测为无目标或者为其他目标。由图 8 可见,改进的 YOLOv5 检测后设备检测置信度基本都在 90% 以上,表示检测精度可靠。

本文算法模型对电解槽设备和样品的检测置信度基本都在 90% 以上,为了进一步评价改进 YOLOv5 的网络性能,采用每一类数据集的平均精确度(AP)、召回率和 F1-score 对检测结果进行分析。检测结果如表 3 所列。由表 3 可知,改进的 YOLOv5 网络在 7 种设备上的平均精确度、召回率和 F1-score 分别为 93.5%, 94.8%, 94.7%。针对大目标电解槽主体和小目标铝锭,电解槽主体(electro-bath)类别的平均检测

精度比原 YOLOv5 的检测精度高 1.0%,样品铝锭(al)类别的平均检测精度提高 1.9%,证明了多尺度改进对尺寸差异大

的目标检测的优越性。改进后的网络模型在降低了模型的参数量的同时,精准度有所提高,检测速度也大幅提高。

表 3 电解槽设备和样品识别结果分析

Table 3 Analysis of identification results of electrolyzer equipment and samples

网络名称	分类名称	精确率/%	召回率/%	F1-score/%	平均检测速度/ (ms/张)	模型参数量/ ($\times 10^6$)
YOLOv5 网络	电解槽主体	91.6	93.4	94.3	8.7	7.1
	母线槽	96.3	94.7	96.5		
	电解槽标牌	93.2	93.7	92.6		
	数字 0-9	89.1	90.4	91.6		
	电解槽外壳罩板(正)	90.6	91.3	93.7		
	电解槽外壳罩板(侧)	92.3	93.3	92.9		
	样品铝锭	90.5	92.8	91.2		
	平均值	92.0	92.8	93.3		
改进的 YOLOv5 网络	电解槽主体	92.6	94.3	94.8	6.2	4.3
	母线槽	96.8	95.2	97.7		
	电解槽标牌	95.6	94.9	94.4		
	数字 0-9	90.3	91.4	92.8		
	电解槽外壳罩板(正)	92.5	93.7	95.5		
	电解槽外壳罩板(侧)	94.5	96.2	95.0		
	样品铝锭	92.4	93.6	92.5		
	平均值	93.5	94.8	94.7		

为了进一步证实改进的 YOLOv5 网络算法的可行性和各个模块对检测结果的不同影响,我们对各个模块进行了消融实验,结果如表 4 所列。

表 4 消融实验分析

Table 4 Results of ablation studies

YOLOv5	ShufflenetV2	SE 注意力 机制	四尺度	mAP (IoU=0.5)/%	平均检测 速度/ (ms/张)
✓				92.0	8.7
✓	✓			89.2	5.3
✓	✓	✓		92.7	6.1
✓	✓	✓	✓	93.5	6.2

由表 4 可以看出,替换主干网络为 ShuffleNetV2 后,虽然检测精度下降了 3.8%,但是检测速度提高了 3.4ms;嵌入 SE 注意力机制后,检测精度提高了 0.7%;改进 Neck 部分,加入第四尺度后,检测精度在原基础上提高了 1.5%。结果表明,引入轻量化主干网络可以提高网络的检测速度,加入注意力机制和改进为四尺度输出后能够提高网络的检测精度。

本文提出的 YOLOv5 多尺度的轻量化改进,提高了对尺寸差异大的目标的检测精度,同时满足轻量化要求在电解槽转运机器人上部署,实现对电解槽设备和样品的实时检测。

结束语 针对炼铝厂的电解槽转运机器人下的视觉识别研究,本文提出轻量的 YOLOv5 目标识别方法,首先改进 YOLOv5 的主干特征提取网络进行轻量化,其次在特征融合部分和预测部分将三尺度检测模型改为四尺度检测模型,最后通过改进的 YOLOv5 进行特征拼接从而实现对设备的检测,得到以下结论:

1)用轻量级的 ShuffleNetV2 网络替换 YOLOv5 的主干特征提取网络 CSPDarknet53,保证网络的轻量化,便于部署到嵌入式设备中;主干网络引入 SE 注意力机制,SPPF 提高较小目标铝锭样品的识别准确率。

2)为了实现多尺度融合,添加第四尺度,增加预测特征层的尺寸差异。与原 YOLOv5 相比,在电解槽转运机器人的实时识别中具有更高的检测精度和识别速度,同时还保证了模型的轻量化。

该算法部署到智能转运机器人上,可以实现对铝锭的实时转运,取代人工运输,提高了铝电厂的智能化。本文针对电解铝车间中环境较好的地方的设备进行识别,但是该方法对于一些存在灰尘或者污渍的情况是否可行以及高分辨率的图片对轻量级网络精度的影响等问题还有待继续研究。

参 考 文 献

- [1] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot Multi Box Detector[C]// European Conference on Computer Vision. Cham; Springer, 2015: 21-37.
- [2] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]// Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition. 2017: 936-944.
- [3] GIRSHICK R. Fast r-cnn[C]// Proceedings of the IEEE International Conference on Computer Vision. 2015: 1440-1448.
- [4] ZHANG Y, ZHANG M L, LV X L, et al. Review of deep learning small target detection algorithm[J]. Computer Engineering and Application, 2022, 58(15): 1-17.
- [5] FAN L L, ZHAO H W, ZHAO H Y, et al. Review of object detection based on deep convolutional neural network[J]. Optics and Precision Engineering, 2020, 28(5): 1152-1164.
- [6] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 8759-8768.
- [7] TAN M, PANG R, LE Q V. Efficientdet: Scalable and efficient object detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 10781-10790.
- [8] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size[J]. arXiv: 1602. 07360, 2016.
- [9] ZHANG X, ZHOU X, LIN M, et al. Shufflenet: An extremely efficient convolutional neural network for mobile devices[C]//

- Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018:6848-6856.
- [10] CHOLLET F. Xception: Deep learning with depthwise separable convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017:1251-1258.
- [11] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018:4510-4520.
- [12] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016:2818-2826.
- [13] XIONG K F, FAN S S, WU J. Identification method of power equipment in foggy substation based on improved YOLOv4[J]. Radio Engineering, 2022, 52(8):1504-1512.
- [14] PI J, LIU Y H, LI J H. Research on lightweight forest fire detection algorithm based on YOLOv5s[J]. Journal of Graphics, 2023, 44(1):26-32.
- [15] WANG P F, HUANG H M, WANG M Q. Improved complex road target detection algorithm for YOLOv5[J]. Computer Engineering and Applications, 2022, 58(17):81-92.
- [16] WANG G W. Multi-scale feature fusion to improve pedestrian and vehicle detection in YOLOv3 network[D]. Nanjing University of Posts and Telecommunications, 2019.
- [17] DUBEY S R, CHAKRABORTY S, ROY S K, et al. DiffGrad: an optimization method for convolutional neural networks [J]. IEEE transactions on neural networks and learning systems, 2019, 31(11):4500-4511.
- [18] LOSHCHILOV I, HUTTER F. Sgdr: Stochastic gradient descent with warm restarts[J]. arXiv:1608.03983, 2016.



WU Jiaojiao, born in 1998, postgraduate, is a member of China Computer Federation. Her main research interest is visual image processing.



LIU Zheng, born in 1977, Ph. D. His main research interests include power electronics and power robots.