

基于轻量化 PAM-M-YOLO 模型的煤矸石图像检测*

郭栋梁, 张延军

(太原科技大学 机械工程学院, 山西 太原 030024)

摘要:针对传统煤矸石检测算法中人为提取煤矸石图像特征过程复杂、检测精度低等问题,提出了一种轻量化的 PAM-M-YOLO 煤矸石检测模型。首先,使用 MobileNetv3 特征提取网络替换原模型主干网络,采用深度可分离卷积替换传统卷积进行煤矸石图像的特征提取;其次,设计 PAM 并联注意力模块提升目标检测网络层拼接后特征图通道和空间信息关注度;最后,基于 CAM 激活限制分支给模型添加先验信息,以降低模型在非关键特征上的局部坍塌。试验结果表明,轻量化 PAM-M-YOLO 煤矸石检测模型准确率、召回率、mAP 值分别为 98.7%、97.5%、98.8%,较原 M-YOLO 模型分别提升了 3.6、2.3、2.0 个百分点;参数量为 3.8 MB,比 YOLOv5 模型降低了近 1/2。热力图可视化效果表明,轻量化 PAM-M-YOLO 模型在检测过程中所关注的信息更集中于煤矸石区域,有效解决了模型在煤矸石区域的局部坍塌问题。

关键词:煤矸石图像检测;YOLOv5 模型;轻量化 PAM-M-YOLO 模型;深度学习;注意力机制;损失函数

中图分类号:TD849+.5;TP391.41 **文献标识码:**A

文章编号:1005-2763(2024)05-0220-08

0 引言

煤炭作为我国的主要能源之一,广泛应用于发电、冶金、化工等领域^[1]。开采和选煤过程中所产生的煤矸石固体废弃物占煤炭总产量的 10% 以上,当煤炭中杂质含量较高时会导致燃烧过程中煤炭利用率低,并对环境造成严重危害^[2]。因此,煤矸石的精确分选成为选煤领域的关键环节。

相较于传统选煤过程中凭借人工经验筛选和利用煤炭、矸石物理性质差异进行机器混选的方法,基于算法的煤矸石识别成为选煤行业的主流方法。传统煤矸石图像识别中依靠人为提取和挑选利于算法识别的图像特征信息,例如灰度直方图和矸石纹理特征等。DOU 等^[3]用支持向量机(Support Vector Machine, SVM)从煤炭和煤矸石图像提取的 19 个

特征中识别最优特征,并以此构建最优分类器。杨慧刚等^[4]用 X 射线扫描煤炭和矸石的图像厚度信息,综合厚度信息和灰度信息确定煤炭和矸石的分离阈值。吴景涛^[5]以支持向量机和随机森林为分类模型,提取灰度共生矩阵特征和 HOG(Histogram of Oriented Gradient)特征为依据进行煤矸石分类。余乐等^[6]通过对煤矸石 0~255 级灰度图像的前部分灰度信息进行灰度级压缩和灰度矩阵扩阶,计算共生矩阵,提取新特征。

随着深度学习在图像识别领域取得突破性成功,深度学习算法也被用于煤炭和矸石的图像识别,即利用卷积神经网络自主提取图像特征和网络参数来实现煤矸石的快速识别。陈佳鑫等^[7]以智能图像识别为核心,设计了用于井下的煤矸石分选装置。孙涛等^[8]基于卷积神经网络提出了一种煤岩信息识别边缘计算系统,实现了煤岩界面的分割识别。蔡秀凡等^[9]基于改进的 YOLOv4 模型,克服了小块煤矸石分选过程中漏检的缺陷问题。单鹏飞等^[10]提出改进的 CBAM Faster R-CNN 模型,实现煤矸石放落瞬间的识别定位,为放煤量的智能监测提供依据。郭永存等^[11]提出通过迁移权重和简化神经元进行模型优化的方法,使煤矸石识别准确率达到 97.461%。尽管上述煤矸石检测算法的检测效果有所提升,但是研究中仍存在人为提取矸石特征过程复杂、算法参数量大、不适合部署在嵌入式平台、识别精度不足等问题。

针对上述情况,本文提出了一种轻量化的 PAM-M-YOLO(Parallel Attention Module-MobileNet-YOLO, PAM-M-YOLO)模型。首先,采用深度可分离卷积代替传统卷积进行煤矸石图像特征提取,以此降低模型参数量来实现模型的轻量化;其次,添加融合通道和空间注意力的并联注意力模块(Parallel Attention Module, PAM);再次,基于热力

* 收稿日期:2023-07-12

作者简介:郭栋梁(1999—),男,山西临汾人,硕士研究生,主要研究方向为计算机视觉与目标检测,E-mail:2482669786@qq.com。

通信作者:张延军(1982—),男,河北承德人,博士,教授,主要研究方向为机电液复合传动及智能控制,E-mail:hebei0735@126.com。

图(Class Activation Mapping, CAM)激活限制分支,给模型添加先验信息,降低模型对非关键特征的关注度,提升模型检测精度;最后,通过试验验证该模型算法的可行性。

1 PAM-M-YOLO 模型

本文所提出的煤矸石检测模型是在 YOLO(You Only Look Once)目标检测模型基础上改进得到的。相较于 Faster-RCNN (Faster Region-Convolutional Neural Network)^[12]、SSD (Single Shot MultiBox Detector)^[13]等目标检测模型,YOLOv5 模型的检测速度和检测精度较高,更适用于实时检测任务。但是,为了模型方便部署在嵌入式平台且具备更高的检测精度,需要在原模型的基础上进行轻量化操作来实现 PAM-M-YOLO 煤矸石检测模型的搭建。

1.1 主干网络轻量化

YOLOv5 网络模型由输入端、Backbone 主干网络、Neck 颈部网络层、Head 预测层构成。

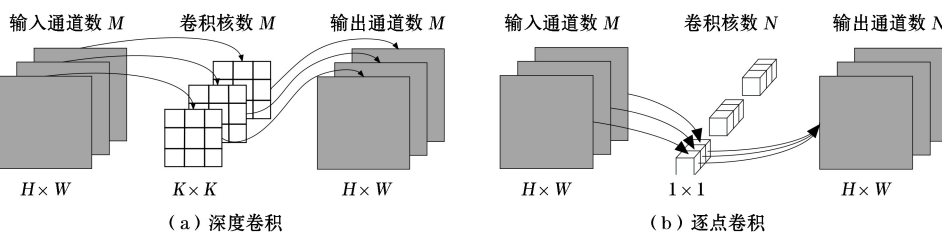


图 1 深度可分离卷积

Fig.1 Deep separable convolution

对于一个输入通道数为 M 、输出通道数为 N 、大小为 $H \times W$ 的特征图,采用尺寸为 $K \times K$ 的卷积核对该特征图进行卷积操作。传统卷积中,先采用单个卷积核对所有输入通道进行卷积并求和,得到输出特征图的一个通道;再利用 N 个同等尺寸的卷积核重复上述操作,得到通道数为 N 的输出特征图。传统卷积过程中的运算量为 K^2MNWH 。而深度可分离卷积中,首先使用 M 组卷积核对输入特征图进行独立卷积,获得与输入特征图通道相同的新特征图;再使用 N 组通道数为 M 的 1×1 卷积核调整该特征图通道数,以此得到最终的输出特征图。深度可分离卷积过程中,深度卷积和逐点卷积的运算量分别为 K^2MWH 与 $MNWH$,深度可分离卷积与传统卷积的运算量之比为:

$$\frac{K^2MWH + MNWH}{K^2MNWH} = \frac{1}{N} + \frac{1}{K^2} \quad (1)$$

理论上讲,针对尺寸为 3×3 的卷积核,模型使

BOCHKOVSKIY 等^[14]在 YOLOv4 网络模型的基础上进行改进,提出跨阶段局部结构(Cross Stage Partial, CSP),并用 Leaky ReLU 和 Sigmoid 激活函数替换了 YOLOv4 原有的 Mish 激活函数。主干网络采用 CSP Darknet 网络对原始图像进行特征提取,模型检测精度相较于之前系列有较好的提升。但考虑到 YOLOv5 模型参数量大,其模型大小不满足部署在嵌入式平台的要求,本文使用一种轻量化的卷积操作提取图像特征,保证模型在有较小参数量的同时保持较高精度。

为了在 MobileNet 系列模型中选取最优特征提取网络作为本文模型的主干网络,设计了一组替换主干网络的对比试验,并对替换后不同模型的准确率、平均精度(mean Average Precision, mAP)和参数量进行比较,选择最优的特征提取主干网络。

1.1.1 深度可分离卷积

深度可分离卷积不同于传统卷积,可在特征提取过程中与逐点卷积独立进行,如图 1 所示。

用深度可分离卷积的运算量将达到传统卷积的 $1/8$ 至 $1/9$ 。

1.1.2 煤矸石检测模型轻量化

文献[15]中 MobileNetv3 模型被应用于图像分类任务,采用多层全连接网络。为在该模型的基础上实现轻量化的煤矸石图像特征提取,需要在主干网络最后层修改输出通道数并添加 YOLO 模型的目标检测网络。目标检测网络结合特征金字塔网络(Feature Pyramid Network, FPN)和路径聚合网络(Path Aggregation Network, PAN)。首先,将主干网络所提取特征图的语义信息和位置信息融合;其次,对主干网络和目标检测网络的特征信息进行融合,使模型获得多层次特征信息;最后,通过损失函数和非极大值抑制来预测特征图中是否具有煤矸石图像。轻量化 M-YOLO (MobileNet-YOLO, M-YOLO)煤矸石检测模型结构如图 2 所示。

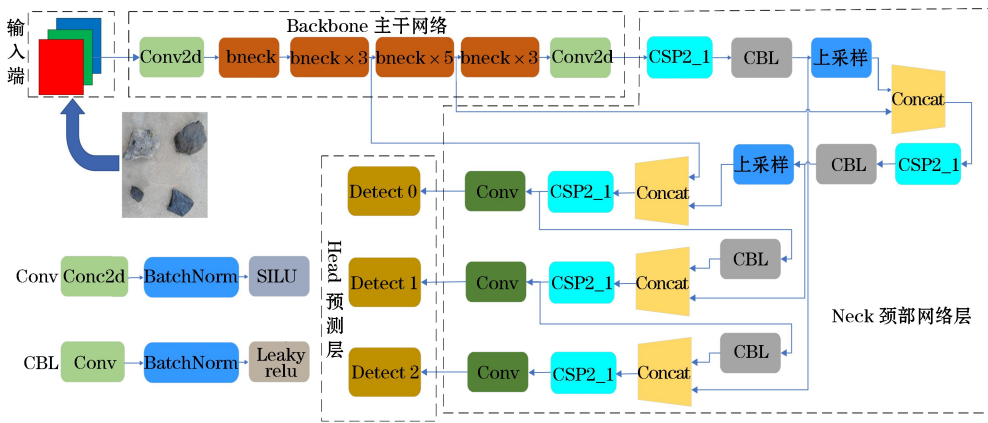


图2 轻量化 M-YOLO 煤矸石检测模型结构

Fig.2 Lightweighted M-YOLO coal gangue detection model structure

从图2可知,对于像素大小为 640×640 的彩色输入图像,M-YOLO模型主干网络对输入图像进行特征提取,目标检测网络将所提取的特征信息进行多尺度融合,并分别对大、中、小3种目标进行检测。相比于传统YOLO模型,轻量化M-YOLO煤矸石检测模型能有效降低模型参数量及运算量,且对模型的检测精度影响较小。

1.2 添加PAM注意力模块

由于SE(Squeeze and excited block)^[16]注意力模块只考虑了特征图的通道信息,忽略了位置信息的重要性。而CBAM(Convolutional Block Attention Module)^[17]注意力模块虽然对特征图先后进行了通道和空间信息的特征融合,但是在对通道信息融合过程中的两次全连接操作造成了部分通道信息的损失,且两者串联的连接方式影响了空间注意力模块对输入特征位置信息的提取。因此,为了加强改进模型对目标检测网络层拼接后特征图通道和空间信息的关注,从而更好地提升轻量化M-YOLO模型对煤矸石图像的检测精度,本文设计了PAM注意力模块。PAM注意力模块由通道注意力模块和空间注意力模块并联组成,具体结构如图3所示。

由图3(b)可知,PAM注意力模块上支路首先对输入特征 F 分别进行全局平均池化和全局最大池化,并对池化结果进行1D卷积。1D卷积过程中选择卷积核的大小为3、填充为1,以保证卷积后的特征图维度与池化结果一致;其次,将两个1D卷积结果进行相加来获得输入特征 F 不同通道上的权重值;最后,对该权重值进行Sigmoid激活,得到带通道权重的通道特征 M_c 。

由图3(c)可知,PAM注意力模块下支路对输

入特征 F 在通道维度进行全局平均池化和全局最大池化,两种池化后的特征图在通道维度进行拼接,得到一个尺寸为 $2 \times H \times W$ 的特征图,其中 H 、 W 分别为特征图的高和宽。对该特征图进行2D卷积,并进行Sigmoid激活得到一个包含输入特征 F 空间信息的空间特征 M_s 。2D卷积过程中选取卷积核大小为3、填充为1。最后将通道特征 M_c 和空间特征 M_s 分别与输入特征 F 相乘得到带通道和空间权重参数的输出特征 F' ,如图3(a)所示。

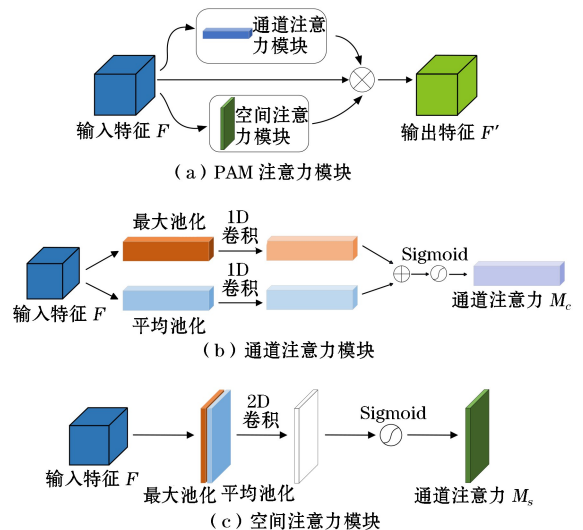


图3 PAM注意力模块结构

Fig.3 PAM attention module structure

1.3 基于CAM激活限制分支添加先验信息

目标检测任务中,YOLO模型虽能够自主学习到一些有辨识度的图像特征,但也会不可避免地拟合到一些非关键特征,造成模型对图像关键特征的拟合不足,影响模型检测精度。因此,本文通过在模型的目标检测网络Head预测层添加CAM激活限

制分支,以解决模型在非关键特征信息上的局部坍塌问题,模型改进前后对比如图 4 所示。

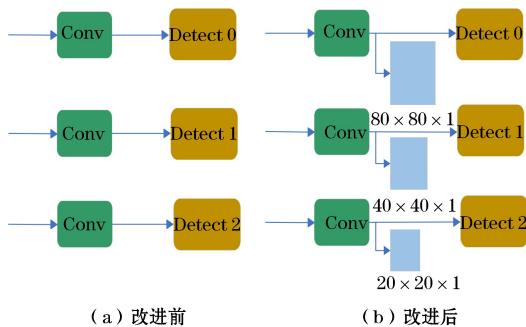


图 4 模型改进前后对比

Fig.4 Comparison of model before and after improvement

1.3.1 添加 CAM 激活限制分支

首先,在模型的 3 个预测层前各添加一个独立分支,每个分支在通道维度对输入特征进行全局平均池化,得到第 k 个通道的权重值 W^k ;然后,将该权重值 W^k 与对应通道的输入特征相乘并在通道维度求和,得到煤矸石图像的激活 M ;最后,使用 Leaky ReLU 激活函数对 M 进行激活,生成一个通道数为 1,且大小与输入特征图一致的 CAM 激活限制分支。其中,权重值 W^k 和激活 M 的计算公式为:

$$W^k = \frac{\sum_{x,y} F_k(x,y)}{h \times w} \quad (2)$$

$$M(x,y) = \sum_k W^k F_k(x,y) \quad (3)$$

式中, W^k 表示特征图在第 k 个通道上的权重值; $F_k(x,y)$ 表示输入特征图中第 k 个通道上坐标为 (x,y) 的值; $M(x,y)$ 表示特征图在空间 (x,y) 点处的激活值; h 为输入特征图的高; w 为输入特征图的宽。

由于经过 CAM 激活限制分支处理后的特征图与输入特征图在空间维度保持相同的网格信息。因此,该分支能够有效展示煤炭和煤矸石在特征图中的空间位置信息。

1.3.2 定义损失函数

煤矸石图像数据集标注过程中,人为将煤矸石的标签标为 0,煤炭的标签标为 1。因此,经过 CAM 激活限制分支处理后的特征图上出现负值的激活区域是煤矸石的激活,即标签为 0 的激活;出现正值的激活区域是煤炭的激活,也就是标签为 1 的激活。为了通过给模型添加先验信息实现模型只关注煤矸石激活区域的目的,本文通过定义损失函数来对标

签为 1 的正激活值进行限制。

首先,设定 $n(x,y)$ 为 CAM 激活限制分支输出特征图中坐标为 (x,y) 的值;其次,将 $n(x,y)$ 作为函数 $F(n)$ 的输入来求得 CAM 激活限制分支的目标特征 $N(x,y)$;最后,对输出特征 $n(x,y)$ 和目标特征 $N(x,y)$ 对应值相乘并求和,得到基于 CAM 激活限制的损失。其中,函数 $F(n)$ 和损失函数 $Loss$ 的计算公式为:

$$F(n) = \begin{cases} 0, & \text{if } n \leq 0 \\ 1, & \text{if } n > 0 \end{cases} \quad (4)$$

$$Loss = \frac{\sum_{x,y} N(x,y) \cdot n(x,y)}{h' \times w'} \quad (5)$$

式中, h' 为经过 CAM 激活限制分支处理后特征图的高; w' 为经过 CAM 激活限制分支处理后特征图的宽。

2 基于改进 PAM-M-YOLO 模型的煤矸石检测试验

2.1 试验数据采集

当前深度学习领域所使用的煤矸石图像主要利用工业相机对输送带上的煤矸石和煤炭进行图像采集。以山西长治某选煤厂的煤矸石图像为例,从中挑选多种角度、不同距离的 3300 张煤炭和煤矸石图像样本进行试验。其中,煤炭图像为 950 张,煤矸石图像为 1300 张,煤炭和煤矸石混合图像为 1050 张。用 LabelImg 标注工具对所采集的数据集图像进行标注,并按照 8 : 1 : 1 的比例将其分为训练集、验证集、测试集 3 部分。

2.2 改进的 PAM-M-YOLO 模型训练

训练过程中选用的试验环境为 Windows 10、Python3.10.10 和 Pytorch,相关硬件配置和模型参数见表 1。单个训练周期有 413 个训练批次,共训练 200 个周期。

表 1 试验相关硬件配置和模型参数

Table 1 Test related hardware configuration and model parameters

硬件名称	配置	参数	参数值
CPU	Core™ i7-9700	优化器	SGD
GPU	GTX 1660	学习率	0.01
CUDA	SUPER 23.3.1	图片像素 批量大小	640×640 8

本文选用准确率 P 、召回率 R 、 mAP 和参数量作为模型的评价指标。其中 P 、 R 和 mAP 的计算公式如下:

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

$$R = \frac{TP}{TP + FN} \quad (7)$$

$$AP = \int_0^1 P(R) dR \quad (8)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP \quad (9)$$

式中, TP 表示判定为正类的正样本数量; FP 表示判定为正类的负样本数量; FN 表示判定为负类的正样本数量; N 表示检测类别数; AP 表示模型在每个类别的检测效果; mAP 表示模型在所有类别上的检测效果。

对 PAM-M-YOLO 模型进行多次训练并选取最优结果用于评估, 训练过程中的损失函数曲线如图 5 所示, 损失值越小表示模型预测结果越精准。由图 5 可知, 当模型训练至 140 轮时, 损失曲线趋于平稳, 损失函数收敛, 说明该模型的训练效果良好。

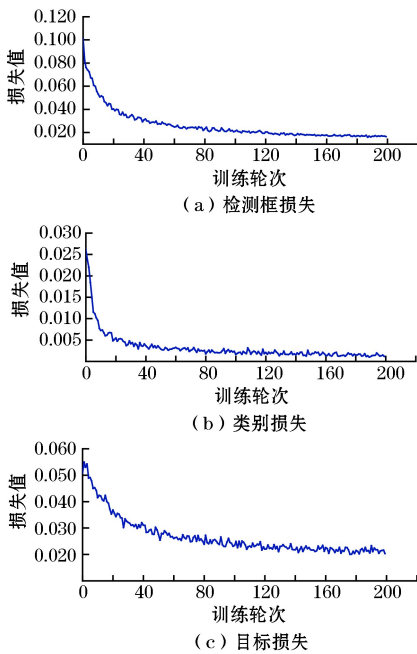


图 5 训练损失

Fig.5 Training loss

3 试验结果分析

3.1 模型主干网络试验对比分析

为在 MobileNet 系列模型中选择最合适的特征提取网络作为本文模型的主干网络, 设计一组对比试验。试验中分别用 MobileNetv1、MobileNetv2 和 MobileNetv3 的特征提取网络对原模型 CSP Darknet

主干进行替换, 对替换后的网络进行训练, 得到各模型准确率 P 、 mAP 和参数量, 试验结果见表 2。由表 2 可知, 替换 v1、v2 特征提取网络后的模型准确率和 mAP 明显低于原模型, 而替换 v3 特征提取网络后, 虽然模型准确率略低于原模型, 但 mAP 相较于之前有较小提升, 且模型参数量减少近一半。因此, 本文选择 MobileNetv3 特征提取网络作为本文模型的主干网络。

表 2 同系列特征提取网络试验对比

Table 2 Comparison of the same series of feature extraction network tests

特征提取主干网络	准确率/%	$mAP@0.5/\%$	参数量/MB
YOLOv5s	96.2	96.5	7.0
MobileNetv1	88.0	89.1	4.3
MobileNetv2	89.5	87.3	3.5
MobileNetv3	95.1	96.8	3.7

3.2 消融试验

为合理分析本文提出的改进模型对原模型的影响, 验证各改进点的可行性, 分别对不同改进策略进行消融试验, 试验数据见表 3。

表 3 消融试验

Table 3 Ablation test

模型	准确率/%	召回率/%	$mAP@0.5/\%$
M-YOLO	95.1	95.2	96.8
M-YOLO+PAM	97.1	96.5	97.9
M-YOLO+CAM	96.3	95.8	97.3
轻量化 PAM-M-YOLO	98.7	97.5	98.8

由表 3 可知, 在 M-YOLO 模型中添加 PAM 注意力模块后, 模型检测准确率、召回率、 mAP 分别为 97.1%、96.5%、97.9%, 相比 M-YOLO 模型分别提高了 2.0, 1.3, 1.1 个百分点。这是由于模型中添加 PAM 注意力模块能够更好地提取拼接后特征层的通道和空间信息。在 M-YOLO 模型中基于 CAM 激活限制分支添加损失函数后, 模型检测准确率、召回率、 mAP 相较于 M-YOLO 模型分别提高了 1.2, 0.6, 0.5 个百分点。在 M-YOLO 模型中同时添加 PAM 注意力模块和基于 CAM 激活限制分支添加损失函数后, 模型检测准确率为 98.7%、召回率为 97.5%、 mAP 为 98.8%, 均高于 M-YOLO 模型和添加单个改进点的值。因此, 在模型中添加 PAM 注意力模块和在预测层添加基于 CAM 激活限制分支的先验信息均能够有效提高模型检测精度。而二者的联合使用可以使模型的检测性能得到进一步提升。

3.3 改进的轻量化 PAM-M-YOLO 模型与其他网络模型对比分析

为验证本文提出的轻量化 PAM-M-YOLO 煤矸石检测模型的识别精度,使用相同的煤矸石数据集,在相同试验条件下对 YOLOv3、YOLOv4-Tiny、YOLOv5 以及本文模型进行训练。不同模型的准确率、 mAP 和参数量见表 4。

表 4 煤矸石数据集下不同模型效果对比

Table 4 Effects comparison of different models under coal gangue data set

模型	准确率/%	$mAP@0.5$ /%	参数量/MB
YOLOv3	95.4	94.8	61.8
YOLOv4-Tiny	93.7	92.5	6.10
YOLOv5	96.2	96.5	7.0
轻量化 PAM-M-YOLO	98.7	98.8	3.8

由表 4 可知,相比于 YOLOv5 原模型及其他 YOLO 系列模型,轻量化 PAM-M-YOLO 模型的参

数量最少,同时检测准确率和 mAP 均高于对比模型。通过对比试验的数据结果可知,改进后的模型在降低模型参数数量的同时有效提高了模型检测精度,满足部署在嵌入式平台的轻量化和高识别度要求。

为了直观了解模型在煤矸石检测过程中所重点关注的图像信息,本文采用 Grad-CAM^[18] 热力图可视化策略对提出的模型进行可解释性分析。Grad-CAM 热力图可视化策略通过计算模型最后卷积层中煤矸石的类别权重,加权求和得到粗粒度的煤矸石定位图,将此定位图映射至输入图像中,突出显示模型检测煤矸石过程中关注的区域。

根据热力图可视化策略对改进前后的模型所关注的煤矸石图像信息进行可视化,可视化结果如图 6 所示。从图 6 中可看出,改进后的轻量化 PAM-M-YOLO 模型相较于原模型,检测过程中所关注的信息更集中于煤矸石区域上,有效解决了模型在非煤矸石区域的局部坍塌问题。

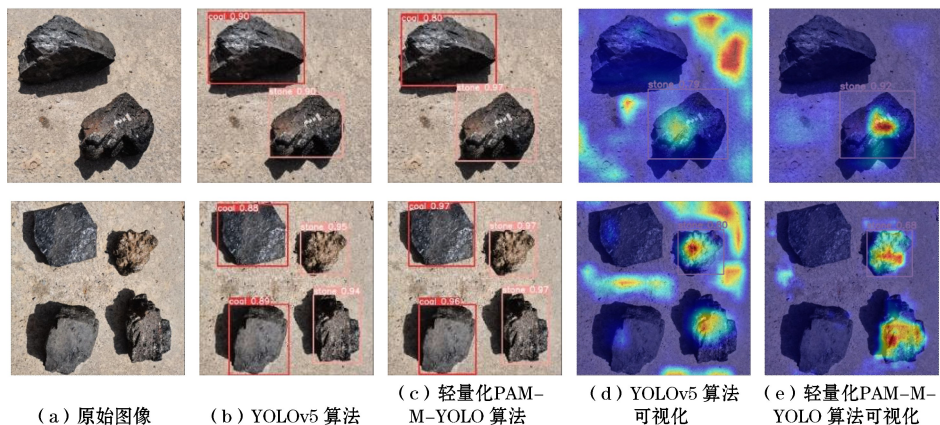


图 6 不同算法结果对比

Fig.6 Comparison of different algorithm results

4 结论

(1) 针对煤矸石图像检测模型算力要求较高的问题,本文提出了一种轻量化的 PAM-M-YOLO 煤矸石检测模型。首先,采用深度卷积与逐点卷积相结合策略,使用深度可分离卷积代替传统卷积,对煤矸石图像的特征信息进行提取。其次,将通道注意力机制与空间注意力机制融合。最后,在模型目标检测 Head 预测层添加 CAM 激活限制分支并定义损失函数,减少模型对非煤矸石图像区域信息的关注。

(2) 在 M-YOLO 模型中添加 PAM 注意力模块和在预测层添加基于 CAM 激活限制分支的先验信息均能提高模型检测精度,而二者的联合使用使模型

检测性能再进一步提升,较 M-YOLO 模型,轻量化 PAM-M-YOLO 模型的准确率、召回率和 mAP 分别提升了 3.6, 2.3 和 2.0 个百分点。

(3) 相较于 YOLO 系列模型,轻量化 PAM-M-YOLO 模型具有更高的检测精度,模型参数量相较于原系列模型有明显降低,满足选煤行业的实际应用需求。

(4) 热力图可视化效果表明,相比于原模型,轻量化的 PAM-M-YOLO 模型对煤矸石特征的关注效果更加明显。

参考文献(References):

- [1] 袁亮,张农,阚甲广,等.我国绿色煤炭资源量概念、模型及预测[J].中国矿业大学学报,2018,47(1):1-8.

- YUAN Liang, ZHANG Nong, KAN Jianguang, et al. The concept, model and reserve forecast of green coal resources in China [J]. *Journal of China University of Mining & Technology*, 2018, 47(1):1-8.
- [2] LI J Y, WANG J M. Comprehensive utilization and environmental risks of coal gangue: A review [J]. *Journal of Cleaner Production*, 2019, 239:117946.
- [3] DOU Dongyang, WU Wenze, YANG Jianguo, et al. Classification of coal and gangue under multiple surface conditions via machine vision and relief-SVM [J]. *Powder Technology*, 2019, 356:1024-1028.
- [4] 杨慧刚, 乔志敏. 基于 X 射线和机器视觉的煤与矸石分选系统设计 [J]. *工矿自动化*, 2017, 43(3):85-89.
- YANG Huigang, QIAO Zhimin. Design of separation system of coal and gangue based on X-ray and machine vision [J]. *Journal of Mine Automation*, 2017, 43(3):85-89.
- [5] 吴景涛. 机器视觉技术在煤矸石智能识别中的应用研究 [D]. 徐州: 中国矿业大学, 2019.
- WU Jingtao. Application of machine vision technology in coal gangue recognition [D]. Xuzhou: China University of Mining & Technology, 2019.
- [6] 余乐, 郑力新, 杜永兆, 等. 采用部分灰度压缩扩阶共生矩阵的煤和煤矸石图像识别 [J]. *华侨大学学报 (自然科学版)*, 2018, 39(6):906-912.
- YU Le, ZHENG Lixin, DU Yongzhao, et al. Image recognition method of coal and coal gangue based on partial grayscale compression extended coexistence matrix [J]. *Journal of Huaqiao University (Natural Science)*, 2018, 39(6):906-912.
- [7] 陈佳鑫, 赵国贞, 程伟, 等. 基于动态称量和图像识别技术的井下煤矸石智能分选装置研发 [J]. *矿业研究与开发*, 2023, 43(2):178-183.
- CHEN Jiaxin, ZHAO Guozhen, CHENG Wei, et al. Development of underground coal gangue intelligent sorting device based on dynamic weighing and image recognition technology [J]. *Mining Research and Development*, 2023, 43(2):178-183.
- [8] 孙涛, 王宏伟, 闫志蕊. 基于卷积神经网络和语义分割混合模型的煤岩识别系统研究 [J]. *矿业研究与开发*, 2022, 42(11):179-187.
- SUN Tao, WANG Hongwei, YAN Zhirui. Research on coal-rock recognition system based on convolutional neural network and semantic segmentation hybrid model [J]. *Mining Research and Development*, 2022, 42(11):179-187.
- [9] 蔡秀凡, 谢金辰. YOLOv4 煤矸石检测方法研究 [J]. *煤炭工程*, 2022, 54(8):157-162.
- CAI Xiufan, XIE Jinchun. YOLOv4-based detection method of coal and gangue [J]. *Coal Engineering*, 2022, 54(8):157-162.
- [10] 单鹏飞, 孙浩强, 来兴平, 等. 基于改进 Faster R-CNN 的综放煤矸混合放出状态识别方法 [J]. *煤炭学报*, 2022, 47(3):1382-1394.
- SHAN Pengfei, SUN Haoqiang, LAI Xingping, et al. Identification method on mixed and release state of coal-gangue masses of fully mechanized caving based on improved Faster R-CNN [J]. *Journal of China Coal Society*, 2022, 47(3):1382-1394.
- [11] 郭永存, 王希, 何磊, 等. 基于 TW-RN 优化 CNN 的煤矸识别方法研究 [J]. *煤炭科学技术*, 2022, 50(1):228-236.
- GUO Yongcun, WANG Xi, HE Lei, et al. Research on coal and gangue recognition method based on TW-RN optimized CNN [J]. *Coal Science and Technology*, 2022, 50(1):228-236.
- [12] GIRSHICK R. Fast R-CNN [C] // *International Conference on Computer Vision*. Piscataway, NJ, USA: IEEE, 2015:1440-1448.
- [13] LIU Wei, ANGELOV D, ERHAN D, et al. SSD: single shot multibox detector [C] // *European Conference on Computer Vision*, 2016:21-37.
- [14] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection [J]. 2020. <https://doi.org/10.48550/arXiv.2004.10934>.
- [15] HOWARD A, SANDLER M, CHU G, et al. Searching for MobileNetV3 [C] // *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019:1314-1324.
- [16] HU J, SHEN L, SUN G, et al. Squeeze-and-excitation networks [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Los Alamitos: IEEE Computer Society Press, 2018:7132-7141.
- [17] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional Block Attention Module [C] // *Computer Vision - ECCV 2018: 15th European Conference Germany*; Springer, 2018:3-19.
- [18] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization [C] // *IEEE International Conference on Computer Vision*. Venice Italy: IEEE, 2017:618-626.

Coal Gangue Image Detection Based on Light-Weighted PAM-M-YOLO Model

GUO Dongliang, ZHANG Yanjun

(School of Mechanical Engineering, Taiyuan University of Science and Technology, Taiyuan, Shanxi 030024, China)

Abstract: Aiming at the problems of complex process and low detection accuracy of artificially extracting coal gangue image features in traditional coal gangue detection algorithms, a light-weighted PAM-M-YOLO coal gangue detection model was

proposed. Firstly, the MobileNetv3 feature extraction network was used to replace the original model backbone network, and the depth separable convolution was used to replace the traditional convolution to extract the features of coal gangue images. Secondly, PAM parallel attention module was designed to improve the attention of feature map channel and spatial information after the splicing of target detection network layer. Finally, a priori information is added to the model based on the CAM activation restriction branch to reduce the local collapse of the model on non-key features. The experimental results show that the accuracy, recall rate and mAP value of the light-weighted PAM-M-YOLO coal gangue detection model are 98.7%, 97.5% and 98.8%, respectively, which are 3.6, 2.3 and 2.0 percentage points higher than those of the original M-YOLO model. The number of parameters is 3.8 MB, which is nearly 1/2 lower than the YOLOv5 model. The visualization effect of the heat map shows that the information concerned by the light-weighted PAM-M-YOLO model in the detection process is more concentrated in the coal gangue area, which can effectively solve the local collapse problem of the model in the coal gangue area.

Key words: Coal gangue image detection, YOLOv5 model, Light-weighted PAM-M-YOLO model, Deep learning, Attention mechanism, Loss function