

DOI: 10.13382/j.jemi.B2307142

面向煤矸识别的目标检测算法*

叶志宇¹ 贾晓芬^{1,2} 王天奇¹

(1. 安徽理工大学电气与信息工程学院 淮南 232001; 2. 安徽理工大学人工智能学院 淮南 232001)

摘要:煤与矸石具有目标密集、特征差异小等特征,基于图像处理的识别方法普遍存在检测速度慢、精度低等问题。为了提高煤矸检测的速度和精度,提出一种 GE-YOLOv5s 煤矸检测模型。首先在 YOLOv5s 的基础上引入 Ghost Conv 代替卷积操作,并设计新模块 GhostCSP,在提升模型检测速度的同时实现网络的轻量化;其次在预测层中加入 GC 自注意力机制,融合 SENet 的轻量化和 NLNet 长距离信息全局捕获的优势,使网络记忆、放大煤矸图像间的细微差异特征,提升模型的表现力;然后在 Neck 部分采用双向特征金字塔网络(BiFPN)结构,利用 BiFPN 融合 3 个不同维度的特征,通过加权特征融合机制提高模型计算效率,进一步提升煤矸检测速度;最后设计一种新型激活函数 Eswish 替代 SiLU 激活函数,提高参数利用率的同时加快模型收敛速度并提升鲁棒性。实验数据表明:相较于 YOLOv5s 模型减少了 34.1% 的参数数量和 38.6% 的浮点运算量,并且在 mAP 0.5:0.95 指标上提升了 1.9%。对比实验显示,相较于 YOLOv3、SSD、FasterR-CNN 和 YOLOv5-scSE 的 mAP 0.5:0.95 指标分别提高了 16.6%、4.8%、13.6% 和 3.8%。将 GE-YOLOv5s 模型应用于煤矸石目标检测过程中,具有更优的识别性能、鲁棒性、网络泛化能力,可有效避免漏检、误检和重叠现象。

关键词:煤矸检测;目标识别;Eswish 激活函数;注意力机制;BiFPN 结构

中图分类号: TP391.4; TN98 **文献标识码:** A **国家标准学科分类代码:** 520.2060

Target detection algorithm for coal and gangue identification

Ye Zhiyu¹ Jia Xiaofen^{1,2} Wang Tianqi¹

(1. School of Electrical and Information Engineering, Anhui University of Science and Technology, Huainan 232001, China;

2. School of Artificial Intelligence, Anhui University of Science and Technology, Huainan 232001, China)

Abstract: Coal and gangue have the characteristics of dense targets and small feature differences, and the recognition methods based on image processing generally have the problems of slow detection speed and low accuracy. To further improve the speed and accuracy of coal gangue detection, a GE-YOLOv5s coal gangue detection model is proposed. Firstly, Ghost Conv is introduced based on YOLOv5s instead of convolution operation, and a new module GhostCSP is designed to improve the detection speed of the model while realizing the lightweight of the network; secondly, the GC self-attention mechanism is added in the prediction layer, which integrates the lightweight of SENet and the advantage of global capture of long-distance information of NLNet, to enable the network to memorize and magnify the Then in the Neck part, a bidirectional feature pyramid network (BiFPN) structure is adopted, and BiFPN is used to fuse the features of three different dimensions to improve the computational efficiency of the model through the weighted feature fusion mechanism to further enhance the speed of coal gangue detection; finally, a new type of activation function is designed to replace the activation function of SiLU, which can improve the utilization rate and accelerate the convergence of the model. Finally, a new activation function Eswish is designed to replace the SiLU activation function, which improves the parameter utilization rate, accelerates the convergence speed of the model and improves the robustness. The experimental data show that compared with the YOLOv5s model, the number of parameters is reduced by 34.1% the amount of floating-point operations is reduced by 38.6%, and the mAP 0.5:0.95 index is improved by 1.9%.

收稿日期: 2023-12-19 Received Date: 2023-12-19

* 基金项目:国家自然科学基金面上项目(52174141)、安徽省自然科学基金面上项目(2108085ME158)、安徽高校协同创新项目(GXXT-2020-54)资助

Comparison experiments show that the mAP 0.5:0.95 metric is improved by 16.6%, 4.8%, 13.6% and 3.8% compared to YOLOv3, SSD, FasterR-CNN, and YOLOv5-scSE, respectively. Applying the GE-YOLOv5s model to the gangue target detection process has better recognition performance, robustness, and network generalization ability, and can effectively avoid the phenomena of leakage, misdetection and overlapping.

Keywords: coal gangue detection; target recognition; Eswish activation function; attention mechanism; BiFPN structure

0 引言

在煤炭开采过程中,通常会把在成煤过程中与煤层伴生的一种比煤坚硬、含碳量较低、不易燃烧的黑灰色岩石也一起采上来,使运到地面上的煤炭中含有矸石。因此在采上来的矿物中煤与矸石往往密集混杂在一起。传统依靠人工来进行煤矸石分选的方法成本高、安全性差,容易造成资源浪费且危害身体健康。为了响应国家自动化、智能化和绿色节能的能源开发号召,研究出新型的高效煤矸检测方法迫在眉睫。近年来,随着人工智能的快速发展,其各种方法已经广泛应用于国防、制造、通信、医疗、教育、安全等领域^[1-2]。

目标检测作为计算机视觉和图像处理的基石,同时也是目标识别、缺陷检测、体积分割、图像描述等更复杂更高层次的视觉任务的关键。以 HOG、DPM、Viola-Jones 为代表的传统目标检测算法在鲁棒性、泛化能力和效率上都有所欠缺,不利于目标检测技术的进一步发展。卷积神经网络(convolutional neural networks, CNN)的出现给图像处理领域带来了重大突破,随之而来的各种神经网络的推陈出新也为研究学者进行更深入的研究带来了启发^[3]。利用目标检测原理,通过机器视觉技术实现煤与矸石混合物中煤和矸石的智能检测,为煤矸自动分选环节的整体系统提供算法支撑^[4]。

目标检测方法主要分为两类:一类是以 YOLO 系列^[5-7]、SSD^[8]、RetinaNet 为代表的单阶段检测算法,另一类是以 FasterR-CNN 系列^[9]为代表的双阶段算法。在对煤与矸石分类识别研究领域,郜亚松等^[10]提出了一种改进型轻量级深度识别网络模型的煤矸识别方法,以 MobileNetV3-large 模块结构为基础,在保证模型参数体积及复杂度较少增加的前提下,使识别精度和召回率相对原模型有明显的提升;徐志强等^[11]利用 K-means++ 算法判断不同卷积核内特征的相似度,实现基于多种 CNN 模型结构的煤矸识别,并且对模型进行了优化和压缩;李德永等^[12]采用 ConvNeXt V2 特征提取模块替换主干网络末端的 2 个 C3 模块,同时采用 Focal-EIOU 损失函数替换原 CIoU 损失函数,有效缓解特征崩溃问题以及保持特征在网络传递过程中的多样性,能较好的提取尺度变化较大或低分辨率煤矸目标关键特征;杨政等^[13]基于煤矸垮落冲击液压支架尾梁的声音信号,提出了一种融

合低级听觉特征 Mel 频谱和高级听觉特征听觉神经递质发放率的放顶煤煤矸识别方法,用来进行煤矸识别;李鹤群等^[14]提出基于灰度区域大小矩阵的激光散斑图像纹理特征提取方法,利用激光散斑成像表征矿物类别差异实现煤矸识别;Yan 等^[15]将多分支结构与跨尺度融合策略相结合,利用膨胀卷积扩展模型感受野,提出一种新型深度学习模型 GAF-DCNN,实现对细粒度局部光谱特征和空间光谱特征的有效学习,用于识别煤矸石的光谱图像;Alfarzaei 等^[16]以卷积神经网络(CNN)为基础,以热图像作为煤矸石识别的标准图像,建立了一种新的模型(CGR-CNN),CGR-CNN 在识别准确率上相对于 SVM-YCbCr 有明显提升;Lei 等^[17]提出一种基于 CNN 的视觉分类方法,构建视觉深度神经网络 FCCN,实现可视化煤炭分类检测算法,与同类算法 Mask R-CNN 相比,提高了分类精度,减少了模型的训练时间,具有良好的应用价值。

由于煤与矸石具有目标密集、特征差异小等特征,现有基于图像处理的上述识别方法普遍存在检测速度慢、精度低等问题,YOLOv5 使用 Pytorch 框架,能够方便地训练自己的数据集,模型训练也非常快速,且模型参数更小,精度更高。鉴于此,在 YOLOv5 基础上,引入幽灵卷积来提升检测速度,融合全局注意力机制(global context, GC)和双向特征金字塔网络(bidirectional feature pyramid network, BiFPN)来记忆、放大煤矸图像间的细微差异特征,同时改进特征提取、特征融合方法,设计新的激活函数,最终建立了 GE-YOLOv5s 煤矸检测模型,通过对比、消融实验验证模型的有效性。

1 GE-YOLOv5s 网络模型的结构

YOLOv5 依据网络宽度、深度分为 YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x 4 类模型^[18]。煤矸检测的实时性要求很高,为了实现对煤与矸石的高效率、高精度检测,因此选择特征映射图宽度最小,深度最浅,检测速度最快的 YOLOv5s 网络作为基础网络进行改进,得到如图 1 所示的 GE-YOLOv5s。

图 1 由输入层、骨干网络、颈部网络和预测层 4 部分组成。首先为提升检测速度,骨干网络中除首层的二维卷积外,其余的普通卷积全部替换为 GhostConv 幽灵卷积,并将原网络中的两种 CSP 模块替换为新设计的

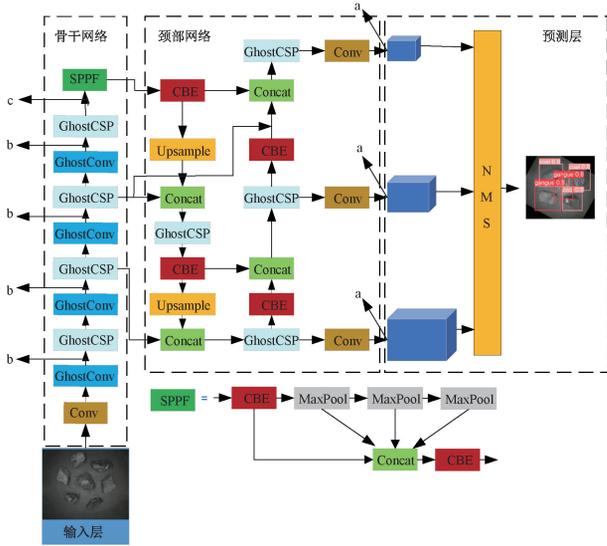


图 1 GE-YOLOv5s 网络结构图

Fig. 1 GE-YOLOv5s network structure diagram

GhostCSP 模块。其次,在网络的预测层引入了 GC 注意力机制,然后用 BiFPN 结构替换颈部网络中的 FPN + PAN 结构。最后,设计新型激活函数 Eswish,并构建 CBE 模块代替原网络中的 CBS 模块。

2 GE-YOLOv5s 的设计

2.1 GhostCSP 模块设计

1) 幽灵卷积

研究发现在进行煤矸石检测的过程中,卷积运算时会使网络产生大量的计算,因此从卷积层方面改进可以更有效地对模型进行轻量化。Han 等^[19]提出的仅通过少量卷积计算就能生成大量特征图的卷积操作——幽灵卷积(ghost conv),可以很好地应用于模型轻量化的工作。

煤矸石分选对细节特征要求严格,文献[20]提到了传统卷积提取图像特征会导致细节信息的遗漏,而幽灵卷积内部线性运算在强化输入数据的区分能力的同时削弱不相关因素。借助幽灵卷积和煤与矸石相似性之间的冗余信息,线性地生成 Ghost 映射,可以实现对煤与矸石图像特征的全面、快速学习。在每个独立的残差模块中,利用不同大小的卷积核逐步提取特征信息,同时利用可变速长卷积实现图像通道信息的维数增减,前后层的信息流通过残差连接进行监督和加强。

2) GhostCSP 模块设计

2020 年华为诺亚方舟实验室提出了轻量级模型 GhostNet^[21-22],如图 2 所示的网络中步长为 1 的 GhostBottleneck 的结构,它的第 1 个 Ghostconv 模块作为扩展层,增加了输出通道数,第 2 个 Ghostconv 模块将通

道数恢复目标输出通道数,使得输入特征图大小不发生改变。然后经过 shortcut 连接,最后将所有输入特征图逐点相加进行特征融合。

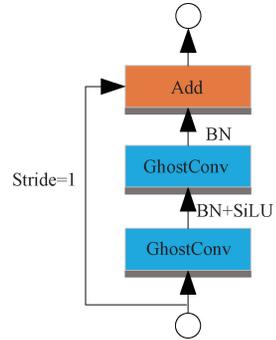


图 2 GhostBottleneck 结构

Fig. 2 GhostBottleneck structure

由于 YOLOv5s 中的 CSP 模块包含的 Bottleneck 结构使用了较多 3×3 普通卷积,会引起较高额参数数量和浮点计算量,因此用 GhostBottleneck 结构替换 CSP 模块中所有的 Bottleneck 模块,构建网络轻量化的基础,减少模型的计算量,形成如图 3 所示的 GhostCSP 模块。

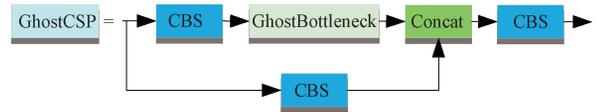


图 3 GhostCSP 模块

Fig. 3 GhostCSP module

2.2 GC 注意力机制的引入

注意力机制能够从大量繁杂信息中快速获取有用信息,可以在不给网络增加额外的负荷的情况下有效提升网络的效率和准确度,进而实现更精准的煤矸检测效果。CBAM^[23]、SE^[24]、ECA^[25]等常见的注意力机制,主要利用图像特征在空间和通道方面的邻近区域的特征信息,此特性无法有效检测煤矸的差异特征。GC^[26]自注意机制通过结合 SENet^[24]的轻量化和 NLNet^[27]捕获长距离信息全局依赖关系,实现全局上下文信息与通道的融合。它能够对煤与矸石之间的细小差异进行精准捕捉,并在特征提取时重点学习此类特征。

2.3 BiFPN 结构

文献[28]提出了一种双向特征金字塔网络 BiFPN。该网络构建了一个双向跨尺度连接的结构,增强了信息融合,并引入了类似于 attention 的权重,使不同尺度的特征信息得到充分利用。

煤与矸石都是固体物质,并且可以有各种不同的颜色和质地。一些煤和矸石都可以呈现类似黑色、棕色、灰色等外观,图像特征相似度较高且具有丰富的不同维度

的纹理信息,因此获得具有更高语义的特征,在颈部网络中将原来的 FPN+PAN 结构替换为图 4 的 BiFPN 结构,来对 3 个维度的深层特征和浅层特征进行跨尺度融合,最终从 3 个维度进行预测。

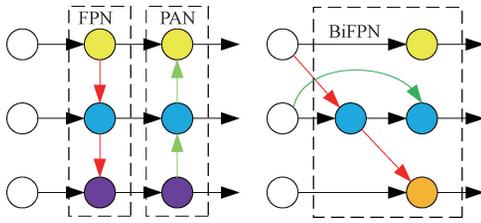


图 4 BiFPN 结构图

Fig. 4 BiFPN structure diagram

相较于 PANet, BiFPN 在 3 个方面进行了改进。首先移除无贡献的节点,增加额外边,叠加特征网络层,以实现更高级的特征融合;其次,设计与 ResNet 类似特征图融合方法,将自上而下的下采样直接与自下而上的残差融合,可以分散梯度并防止退化,增强特征表示能力;最后,针对融合的各个尺度特征增加一个权重,调节每个尺度的贡献度,提出了快速归一化融合,提高目标检测速度。

2.4 Eswish 激活函数的设计

Google 提出的激活函数 Swish^[29],在大数据集上和部分中小数据集上都表现出了不错的效果。Swish 有下界而无上界,不会出现梯度饱和,同时其自身和一阶、二阶导数均平滑,处处连续且可导训练简便,且作为一种非线性激活函数,Swish 可以更好地捕捉数据中的复杂特征,在煤矸识别中,这种非线性特性有助于更好地识别煤矸的复杂结构和纹理。为应对煤矸检测的实时性需求,在 Swish 激活函数的基础上进行了改进,设计了一个新型激活函数 Eswish (efficient swish),其数学模型为:

$$\sigma(x) = \max(x, xe^{-|x|}) \tag{1}$$

数利用分段函数表示为:

$$\sigma(x) = \begin{cases} x, & x > 0 \\ xe^x, & x < 0 \end{cases} \tag{2}$$

其导函数为:

$$\sigma'(x) = \begin{cases} 1, & x > 0 \\ (1+x)e^x, & x < 0 \end{cases} \tag{3}$$

Eswish 函数图像如图 5 所示。

由图 5 可知,函数图像在 0 附近时,随着 x 的增长,先缓慢下降再逐渐增加,即函数在包含负半轴的原点附近是饱和的。在训练模型时一般采用正态分布初始化参数或者均匀初始化参数,对于 ReLU 系列函数,饱和和区域的参数和不饱和区域的参数各有一半,即有一半的参数未被得到利用,而 Eswish 函数在负值区间也有一定的

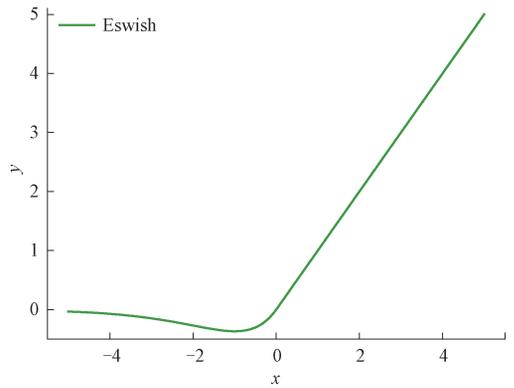


图 5 Eswish 图像

Fig. 5 Eswish image

饱和区域,这使得一部分处于负半轴的参数也可以被利用到,并且正值区间使用线性函数 x 有利于提高模型稳定性与简洁性,使函数直接进行线性输出,易于使用梯度下降法进行优化。平滑的函数图像,也符合生物神经元特征,有助于模型在训练过程中快速稳定收敛,函数的导数特性有助于保持训练中的梯度信息,避免梯度弥散和神经元死亡,负值区域的指数函数增强了激活函数的多样性和复杂性,在处理煤与矸石之间的复杂数据时比其他激活函数更具优势,收敛速度更快。

使用新型激活函数 Eswish 替换原网络中的 SiLU 激活函数,生成如图 6 所示的 CBE 模块,用以代替原来的 CBS 模块。



图 6 CBE 模块

Fig. 6 CBE structure diagram

替换后的 Eswish 激活函数结合了线性和指数特性的优点,具有连续性、光滑性、非单调性、软饱和等特点,既确保了神经网络在训练过程中不会因负值区域的梯度消失问题导致训练速度显著下降,又在一定程度上降低了过拟合的风险。在正值区域提供简单的线性激活,在负值区域通过指数函数增强了激活的表现力。这种设计使得函数不仅能处理复杂的数据特征,同时也降低了计算上的复杂性。与激活函数 Swish 和 Mish 相比,Eswish 在处理负值时的非线性变换,既保留了必要的梯度信息,又避免了计算复杂度的显著增加。

3 实验及结果分析

3.1 煤矸数据集建立

为了更好地开展煤与矸石识别的目标检测实验,采集了大量煤和矸石实际场景图像,大小均为 2 448 ×

2 048, 并制作成数据集, 其中每张图像约包含 4~8 个煤或矸石样本, 以便更好的突出煤与矸石目标密度与关键特征相似等特点。使用 LabelImg 软件对图像进行人工标注, 并通过程序代码生成相对应的 xml 和 txt 标签文件。

针对煤与矸石混合样本图像进行水平翻转、随机裁剪、高斯模糊、亮度变化、仿射变换等相关数据增强操作, 以提升模型的泛化能力, 每张影像增强数量为 130, 共得到增强后的图像 7 205 张。图像扩充后的结果如图 7 的示例。将所有图像按照 4 : 1 的比例划分为训练集和测试集两部分, 其中训练集 5 764 张, 测试集 1 441 张。

由于 YOLOv5 的架构和预训练权重通常基于较小尺寸的图像进行训练和调优。将图像尺寸缩小到 640×640 有助于利用预训练权重的优势, 从而提升模型性能, 并且较大的图像在进行训练时会占用大量的计算资源, 将图像尺寸缩小到 640×640 也可以显著降低内存和计算需求, 加快训练速度, 所以在实验中将输入图像分辨率设置为 640×640。

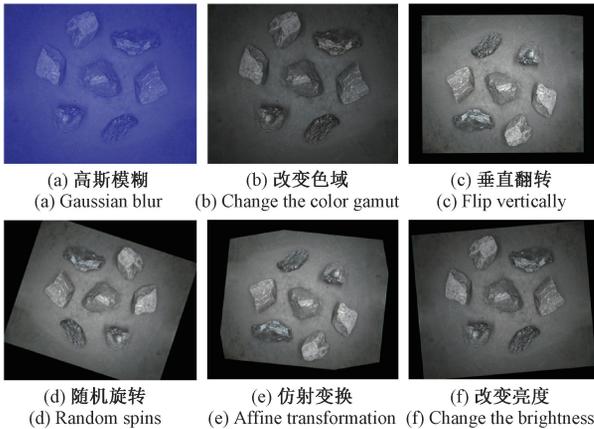


图 7 图像增强结果

Fig. 7 Image enhancement results

3.2 实验环境及评价指标

实验在 GPU 型号为 NVIDIA RTX 3080Ti 的 PC 上进行, Pytorch 版本为 1.9.0, torchvision 版本为 0.10.0, CUDA 版本为 11.1, 训练网络时, 采用随机梯度下降法 (SGD) 训练, Epoch 设置为 120。权重衰减系数设置为 0.000 5, 初始学习速率设置为 0.01, 动量设置为 0.937, 批次设置为 12, 输入图像大小设置为 640×640。

实验对象使用如图 7 所示的采集了大量煤和矸石实际场景图像制作的煤矸数据集。为了评估模型整体性能, 实验中使用平均精确率 (mean average precision, mAP)、每秒浮点运算次数 (floating-point operations, FLOPs) 和参数量 (params) 作为模型的评估指标。mAP0.5 指 IoU 阈值设为 0.5 时, 计算每一个类别的 AP

所得的平均值; mAP0.5 : 0.95 指在不同 IoU 阈值 (0.5、0.55、0.6、0.65、0.7、0.75、0.8、0.85、0.9、0.95) 下的平均 mAP。mAP0.5 : 0.95 综合评估了模型在不同 IoU 阈值下的性能, 更加全面和严格, 更能体现模型的精确度和抗干扰能力。mAP 指标的值越高越好, FLOPs 和 Params 两个指标分别用来衡量计算量和模型大小, 值越低越好。每项指标最优值用粗体标出。

3.3 GC 模块替换

为了放大煤与矸石之间的细微差异, 将 GC 模块引入到 YOLOv5s 网络。因为 YOLOv5 的架构设计是基于层次化特征提取, 从低级到高级逐步提取特征。在每个重要层次引入注意力机制可以最大化每一层次特征提取的效果, 所以一般只在输出层, 输入层, 特征提取层中引入注意力机制。故选择添加位置为 a、b、c 处, 具体如图 1 所示, 其中 a 处为输出层, 在该处添加注意力机制可以在预测阶段增强重要特征的权重; b 处为输入层, 在该处添加注意力机制可以帮助模型在早期阶段提取更加有用的特征; c 处为特征提取层, 在该处添加注意力机制可以提高特征提取的效果。经过多次实验后, 显示 GC 模块加在输出层 a 处特提取效果最佳。

表 1 注意力机制的添加位置

Table 1 The add location of attention mechanism

位置	参数量/M	Map0.5	mAP0.5 : 0.95
a	7.36	94.5	70.2
b	7.37	93.4	66.6
c	7.27	93.5	67.9

表 1 为添加后的实验数据, 可得: GC 模块在少量增加参数的情况下, 能实现全局通道信息间的自我关注, 并且实现煤矸的高精度、高效率检测。

3.4 消融实验

为了验证改进策略的有效性, 将 Ghost、GC、BiFPN 和 Eswish 分别引入 YOLOv5s, 在自制煤矸数据集上进行消融实验, 在相同参数配置下, 开展了表 2 中的 6 组实验。

表 2 消融仿真实验结果

Table 2 Ablation simulation experimental results

模型	参数量/M	浮点数/G	Map0.5	mAP0.5 : 0.95
YOLOv5s	7.02	15.8	95.0	69.0
+GC	7.36	16.0	94.5	70.2
+Ghost	3.74	8.2	92.9	68.0
+BiFPN	7.16	16.4	95.8	63.2
+Eswish	7.02	15.8	94.9	70.8
GE-YOLOv5s	4.62	9.7	96.8	70.9

由表 2 中数据可知, 网络进行了不同改进后, mAP0.5 均保持在较高的水平, 保证了该网络检测的高

精度水平。将原函数中的 SiLU 函数替换为 Eswish 函数,生成新型 CBE 模块,使 mAP 0.5 : 0.95 值提升了 1.8%。加快了收敛速度,提升了模型鲁棒性,很好的提升了模型的准确性和泛化能力。

引入的 Ghost Conv,通过线性变换基于原始特征挖掘出所需信息的特征图,GhostBottleNeck 的设计可大大降低 FLOPs 和 Params,在少量降低精确度的情况下,FLOPs 减少了 48.1%,Params 减少了 46.7%。引入 GhostConv 后,采用线性变换,先生成 channel 较小的特征图,再通过廉价线性运算(cheap linear operations)大大减少由卷积过程产生的计算量,从而达到网络轻量化效果,与此同时还加快了网络训练速度,提升了检测效率。

在网络预测层中引入 GC 注意力模块,相比于原网络,mAP 0.5 : 0.95 值提升了 1.2%,证明 GC 注意力机制的全局信息提取和信息高效传输的特性被充分发挥,提高了被检测对象的显著性,将提取到的煤与矸石之间的差异信息进行放大,赋予不同的权重,可以得到不同权重通道的特征图,重点学习重要特征,极大提高了网络训练效率。

使用 BiFPN 结构替换颈部网络中的 FPN+PAN 结构后,通过有效的双向特征融合和加权特征融合,增强了模型对煤与矸石所处的复杂背景和不同目标尺寸的处理能力。由于引入了可学习的权重机制,BiFPN 结构可以对不同尺度特征的有效性进行评估,并调整特征的贡献,保留融合过程中重要的特征,显著提高了特征表达的有效性和丰富性。该部分使 mAP 0.5 值提升了 0.8%。

将 YOLOv5s 进行上述 4 处改进后,构成新的网络 GE-YOLOv5s,分别降低了 34.1% 的 Params 和 38.6% 的 FLOPs,并且在 mAP 0.5 : 0.95 指标上提升了 1.9%,说明本模型在不同的阈值下的检测结果均非常准确,抗干扰能力较强,覆盖面广,可以适应不同的场景和应用需求。

3.5 对比实验

在保证使用相同的数据集和实验环境配置的情况下,将 GE-YOLOv5s 网络与 YOLOv5-scSE^[5]、YOLOv3^[6]、SSD^[8]、FasterR-CNN^[9] 几种主流网络进行对比实验,结果如表 3 所示。

表 3 对比仿真实验结果

Table 3 Comparison of the simulation results

模型	参数量/M	浮点数/G	Map0.5	mAP0.5 : 0.95
YOLOv5-scSE	8.80	17.9	96.8	67.1
YOLOv3	61.50	154.7	93.5	54.3
SSD	23.89	129.8	95.8	66.1
FasterR-CNN	41.08	82.5	93.3	57.3
GE-YOLOv5s	4.62	9.7	96.8	70.9

低值,mAP 0.5 : 0.95 高达 70.9%,各项指标参数均是最优。相较于 YOLOv5-scSE、YOLOv3、SSD 和 FasterR-CNN 的 mAP 0.5 : 0.95,分别提高了 3.8%、16.6%、4.8% 和 13.6%。除原模型外,YOLOv5-scSE 的 Params 指标和 FLOPs 指标排名第 2,但 mAP 0.5 : 0.95 仅为 67.1%。GE-YOLOv5s 算法的 Params 为 4.62 M,低于原始算法模型的 7.02 M。说明改进算法在提高模型精度的同时减小了模型规模,具体检测结果如图 8 所示。

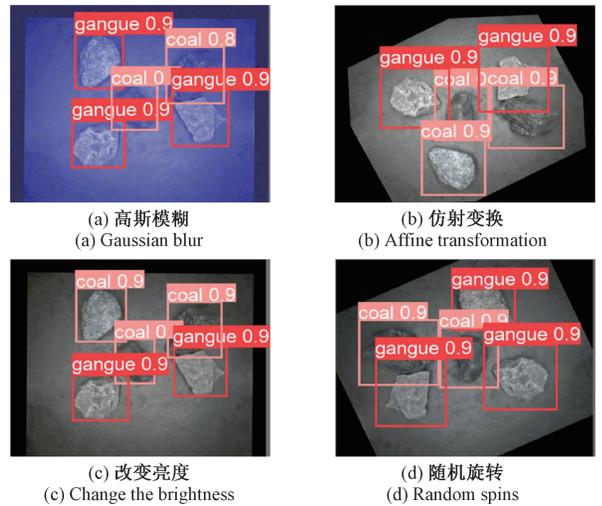


图 8 目标检测结果

Fig. 8 Target detection result

GE-YOLOv5s 更加充分地利用了低层次的特征信息,将煤与矸石之间的细节特征进行放大记忆并加强学习,从而使得模型对煤与矸石等小目标检测具有更大的优势;同时缩减无用特征对模型的干扰和影响,增强模型稳定性,提高网络精度的同时完成了模型的轻量化。

4 结论

文中设计的 GE-YOLOv5s 能解决煤与矸石的错检漏检问题,它利用 GhostConv 在骨干网络有效地生成冗余特征图,从而在保持模型性能的前提下,大幅度降低了计算成本和参数数量。在预测层引入 GC 注意机制,可以通过全局特征聚合获取全局上下文信息,增强对有效特征信息的捕获能力,提高网络的训练速度。结合线性和指数函数特点设计的新型激活函数 Eswish,能提高参数利用率,从而让更有效的特征信息流入神经网络中,确保网络的准确性和泛化能力。利用 BiFPN 结构修改颈部网络,可以学习不同输入特征的重要性,同时有区分的融合不同输入特征,可以更好的增强网络的稳定性和收敛速度。在煤矸自建数据集上的仿真实验表明,相较于 YOLOv5s,GE-YOLOv5s 在减少了 34.1% 参数数量的同

改进后模型的 Params 和 FLOPs 在所有模型中为最

时,mAP 0.5 : 0.95 提升了 1.9%。后续将在此基础上继续优化网络结构,提高 map0.5 精度,并且将工作重心转换为对模型部署方面,使网络可以更好地应用于实际工作中。

参考文献

- [1] 常丽,张雪,蒋辉,等.融合YOLOv5s与SRGAN的实时隧道火灾检测[J].电子测量与仪器学报,2022,36(8):223-230.
CHANG L, ZHANG X, JIANG H, et al. Real-time tunnel fire detection by fusion of YOLOv5s and SRGAN [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(8): 223-230.
- [2] 董兆苒,董明利,何彦霖,等.血管介入手术导丝末端检测方法研究[J].仪器仪表学报,2023,44(2):221-229.
DONG ZH R, DONG M L, HE Y L, et al. Study on the detection method of putting guide wire end in vascular interventional surgery [J]. Chinese Journal of Scientific Instrument, 2023, 44(2): 221-229.
- [3] 雷春丽,夏奔锋,薛林林,等.小样本下自校正卷积神经网络的滚动轴承故障识别方法[J].仪器仪表学报,2022,43(9):122-130.
LEI CH L, XIA B F, XUE L L, et al. Fault identification for rolling bearing by self-calibrated convolutional neural network under small samples conditions[J]. Chinese Journal of Scientific Instrument, 2022, 43(9): 122-130.
- [4] 赵学军,李建.一种基于深度学习的煤矸石检测方法[J].矿业科学学报,2021,6(6):730-736.
ZHAO X J, LI J. A detection method of gangue based on deep learning [J]. Journal of Mining Sciences, 2021, 6(6): 730-736.
- [5] LI Y, FAN Y, WANG S, et al. Application of YOLOv5 based on attention mechanism and receptive field in identifying defects of thangka images[J]. IEEE Access, 2022, 10: 81597-81611.
- [6] REDMON J, FARHADI A. YOLOv3: An incremental improvement [J]. ArXiv preprint arXiv: 1804.02767, 2018.
- [7] 丁伟利,张志鹏,雷子琦,等.深度学习陶瓷表面缺陷检测算法研究[J].电子测量与仪器学报,2023,37(11):161-169.
DING W L, ZHANG ZH P, LEI Z Q, et al. Research on deep learning ceramic surface defect detection algorithm[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(11): 161-169.
- [8] ZHANG B, ZHANG H B. Coal gangue detection method based on improved SSD algorithm[C]. 2021 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), Xi'an, China, 2021: 634-637.
- [9] 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 P F, SUN H Q, LAI X P, 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 Coal Science, 2022, 47(3): 1382-1394.
- [10] 郜亚松,张步勤,郎利影.基于深度学习的煤矸石识别技术与实现[J].煤炭科学技术,2021,49(12):202-208.
GAO Y S, ZHANG B Q, LANG L Y. Coal gangue identification technology and implementation based on deep learning [J]. Coal Science and Technology, 2021, 49(12): 202-208.
- [11] 徐志强,吕子奇,王卫东,等.煤矸智能分选的机器视觉识别方法与优化[J].煤炭学报,2020,45(6):2207-2216.
XU ZH Q, LYU Z Q, WANG W D, et al. Machine vision recognition method and optimization of intelligent sorting of coal gande [J]. Journal of China Coal Society, 2020, 45(6): 2207-2216.
- [12] 李德永,王国法,郭永存,等.基于CFS-YOLO算法的复杂工况环境下煤矸图像识别方法[J].煤炭科学技术,2024,52(6):226-237.
LI D Y, WANG G F, GUO Y C, et al. Image recognition method of coal waste in complex working environment based on CFS-YOLO algorithm [J]. Coal Science and Technology, 2024, 52(6): 226-237.
- [13] 杨政,王世博,饶柱石,等.基于听觉特征融合的煤矸识别方法研究[J].振动与冲击,2024,43(8):136-144.
YANG ZH, WANG SH B, RAO ZH SH, et al. Coal gangue identification method based on auditory feature fusion research [J]. Journal of Vibration and Shock, 2024, 43(8): 136-144.
- [14] 李鹤群,郑予菲,杨涵夕,等.基于激光散斑成像的煤矸特征提取与识别方法研究[J/OL].光学学报,1-20[2024-07-02]. <http://kns.cnki.net/kcms/detail/31.1252.O4.20240625.0902.002.html>.
LI H Q, ZHENG Y F, YANG H X, et al. Based on the coal waste of laser speckle imaging feature extraction and recognition method research [J/OL]. Journal of optics, 1-20 [2024-07-02]. <http://kns.cnki.net/kcms/detail/31.1252.O4.20240625.0902.002.html>.
- [15] YAN Z, XIAO D. Identification of coal, gangue, and

- surrounding rock based on LIBS and deep learning[C]. IEEE Transactions on Instrumentation and Measurement, 2024;1-9.
- [16] ALFARZAEI M S, NIU Q, ZHAO J, et al. Coal/Gangue recognition using convolutional neural networks and thermal images [C]. IEEE Access, 2020, 8; 76780-76789.
- [17] LEI S, XIAO X, ZHANG M, et al. Visual classification method based on CNN for coal-gangue sorting robots[C]. 2020 5th International Conference on Automation, Control and Robotics Engineering (CACRE), Dalian, China, 2020;543-547.
- [18] CHEN Y, LIU W, ZHANG J. An enhanced YOLOv5 model with attention module for vehicle-pedestrian detection[C]. 2022 IEEE 31st International Symposium on Industrial Electronics (ISIE), Anchorage, AK, USA, 2022; 1035-1040.
- [19] HAN K, WANG Y, TIAN Q, et al. GhostNet: More features from cheap operations [C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020;1577-1586.
- [20] 王天奇,贾晓芬,杜圣杰,等. 煤矸的轻量级智能分选网络[J]. 光电子·激光,2023,34(1):19-25.
WANG T Q, JIA X F, DU S J, et al. Coal gangue lightweight intelligent network [J]. Journal of electronics, laser, 2023,34(1) : 19-25.
- [21] LIU C, SUI H, WANG J, et al. Real-time ground-level building damage detection based on lightweight and accurate YOLOv5 using terrestrial images [J]. Remote Sensing, 2022, 14(12) : 2763.
- [22] KONG L, WANG J, ZHAO P. YOLO-G: A lightweight network model for improving the performance of military targets detection[J]. IEEE Access, 2022.
- [23] WOO S, PARK J, LEE J Y. CBAM: Convolutional block attention module [J]. Springer Cham, 2018; 3-19, DOI: 10.1007/978-3-030-01234-2_1.
- [24] JIE H, LI S, GANG S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017; 2011-2023, DOI: 10.1109/TPAMI.2019.2913372.
- [25] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020; 11531-11539.
- [26] CAO Y, XU J, LIN S, et al. Genet: Non-local networks meet squeeze-excitation networks and beyond [C]. 2019IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). IEEE,2020.
- [27] WANG X, GIRSHICK R, GUPTA A, et al. Non-local neural networks[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. 2018; 7794-7803.
- [28] 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.
- [29] RAMACHANDRAN P, ZOPH B, LE Q V. Searching for activation functions [J]. ArXiv preprint arXiv: 1710.05941, 2017.

作者简介



叶志宇,现为安徽理工大学硕士研究生,主要研究方向为深度学习、图像处理。

E-mail:y15855436325@163.com

Ye Zhiyu is a M. Sc. candidate at Anhui University of Science and Technology now. His main research interests include deep learning and image processing.



贾晓芬(通信作者),2009年于哈尔滨工业大学获得硕士学位,2019年于安徽理工大学获得博士学位。现为安徽理工大学人工智能学院教授,博士生导师,主要研究方向为深度学习、机器学习和图像处理。

E-mail:jxfzbt2008@163.com

Jia Xiaofen (Corresponding author) received her M. Sc. degree from Harbin Institute of Technology in 2009 and her Ph. D. degree from Anhui University of Science and Technology in 2019. She is now a professor and doctoral supervisor at the School of Artificial Intelligence, Anhui University of Science and Technology. Her main research interests include deep learning, machine learning and image processing.