



煤炭科学技术 COAL SCIENCE AND TECHNOLOGY

# 煤炭科学技术

## 基于特征增强与Transformer的煤矿输送带异物检测

高涵 赵培培 于正 肖涛 李肖利 李良先

### 引用本文:

高涵, 赵培培, 于正, 等. 基于特征增强与Transformer的煤矿输送带异物检测[J]. 煤炭科学技术, 2024, 52(7): 199–208.  
GAO Han, ZHAO Peipei, YU Zheng. Coal mine conveyor belt foreign object detection based on feature enhancement and Transformer[J]. Coal Science and Technology, 2024, 52(7): 199–208.

在线阅读 View online: <https://doi.org/10.12438/cst.2023-1336>

### 您可能感兴趣的其他文章

#### Articles you may be interested in

#### 基于双流融合网络的输送带跑偏检测方法

Conveyor belt deviation detection method based on dual flow network

煤炭科学技术. 2023, 51(S2): 259–267 <https://doi.org/10.13199/j.cnki.cst.2023-0215>

#### 基于激光测距矿用带式输送机多参数检测方法研究

Research on multi-parameters detection method of mine belt conveyor based on laser ranging

煤炭科学技术. 2020, 48(8): 131–138 <http://www.mtkxjs.com.cn/article/id/77cc2e01-881c-4fff-990b-ac99b657cf4c>

#### 基于轻量化的输送带块煤实时监测方法

A lightweight-based method for real-time monitoring of lump coal on conveyor belts

煤炭科学技术. 2023, 51(S2): 285–293 <https://doi.org/10.12438/cst.2023-1217>

#### 机器视觉灰度化金字塔卷积模型的煤流异物识别

Recognition of unwanted objects in coal flow based on gray pyramid convolution model of machine vision

煤炭科学技术. 2022, 50(11): 194–201 <http://www.mtkxjs.com.cn/article/id/fb838c5c-a1b8-4ba5-98a1-343a9352252a>

#### 输送带输煤采样技术研究现状及发展趋势

Research status and development trend of belt coal sampling technology

煤炭科学技术. 2022, 50(9): 200–206 <http://www.mtkxjs.com.cn/article/id/662f110f-ee8b-4c7b-ad39-6b5d1d490945>

#### 煤矿带式输送机火灾光纤传感检测技术研究

Technology research on optical fiber sensing detection for belt conveyor fire in coal mine

煤炭科学技术. 2019(2) <http://www.mtkxjs.com.cn/article/id/3c985764-301b-40cb-b4c9-7b7afe3f8f2d>



关注微信公众号，获得更多资讯信息



移动扫码阅读

高 涵, 赵培培, 于 正, 等. 基于特征增强与 Transformer 的煤矿输送带异物检测[J]. 煤炭科学技术, 2024, 52(7): 199–208.

GAO Han, ZHAO Peipei, YU Zheng, *et al.* Coal mine conveyor belt foreign object detection based on feature enhancement and Transformer[J]. Coal Science and Technology, 2024, 52(7): 199–208.

## 基于特征增强与 Transformer 的煤矿输送带异物检测

高 涵<sup>1</sup>, 赵培培<sup>1</sup>, 于 正<sup>1</sup>, 肖 涛<sup>2</sup>, 李肖利<sup>2</sup>, 李良先<sup>3</sup>

(1. 中国矿业大学 信息与控制工程学院, 江苏 徐州 221116; 2. 常州海图信息科技股份有限公司, 江苏 常州 213000;

3. 焦作煤业(集团)新乡能源有限公司赵固二矿, 河南 新乡 453633)

**摘 要:** 输送带是煤矿井下最重要的运输设备之一, 在输送工作中会因锚杆、槽钢、大块矸石等异物混入造成输送带撕裂、落煤口阻塞等重大安全事故, 严重影响运输效率, 甚至威胁工人生命安全。为了解决现有输送带异物检测算法存在的对细长物体表征能力弱、弱语义特征提取能力差等问题, 设计了一种基于低层级特征增强与 Transformer 机制的异物检测算法 (Feature Enhancement and Transformer YOLO, FET-YOLO)。首先, 针对现有检测网络难以提取细长物体特征的问题, 引入可变形卷积以提升网络对细长异物的形状特性的适应性, 并使用 MobileViT 模块增加图像中异物与背景的区分度, 以提取出更符合细长异物的多样性特征, 削弱背景噪声的干扰; 其次, 构建低层级特征增强模块 (Low-Level Feature Enhancement Module, LFEM), 提升异物弱语义特征在检测网络中的表达能力, 以降低漏检、错检的概率; 最后, 引入鬼影混洗卷积 (GSConv) 减少因特征图尺寸变化造成的信息丢失, 保证网络高效提取特征的同时, 减少模型参数量。利用煤矿井下输送带工作视频制作训练集和验证集, 并将提出的算法与现有的 3 种输送带检测算法对比, 实验结果表明: 所提出的算法可以更好的解决输送带异物目标检测中细长物体检测效果差、弱语义特征提取困难的问题, 具有更高的检测精度、同时符合输送带检测场景对检测实时性的要求, 对于分辨率大小为 640×640 的图像 mAP@0.5 可达 0.875, mAP@0.5:0.95 可达 0.543, 检测速度为 75 fps。

**关键词:** 异物检测; YOLOv7-tiny; 鬼影混洗卷积; 输送带; 深度学习

中图分类号: TP391.4

文献标志码: A

文章编号: 0253-2336(2024)07-0199-10

## Coal mine conveyor belt foreign object detection based on feature enhancement and Transformer

GAO Han<sup>1</sup>, ZHAO Peipei<sup>1</sup>, YU Zheng<sup>1</sup>, XIAO Tao<sup>2</sup>, LI Xiaoli<sup>2</sup>, LI Liangxian<sup>3</sup>

(1. School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China; 2. Changzhou Haitu Information Technology Co., Changzhou 213000, China; 3. Zhaogu No.2 Mine, Jiaozuo Coal Industry (Group) Xinxiang Energy Co., Ltd., Xinxiang 453633, China)

**Abstract:** Conveyor belt is one of the most important transportation equipment in underground coal mines. Anchor rod, channel steel, large gangue, and other foreign objects that are mixed with the conveyor belt during conveying operations cause the belt to tear, clog the coal drop opening, and cause other serious safety incidents. They seriously affect the efficiency of transportation and even threaten the lives of workers. Aiming at the existing conveyor belt foreign object detection algorithm's problems such as weak characterization of slender objects and poor weak semantic feature extraction, a foreign object detection algorithm based on low-level feature enhancement with Transformer is designed, notated as FET-YOLO. Firstly, to address the problem that existing detection networks have difficulty in extracting features of elongated objects, deformable convolution is introduced to enhance the network's adaptability to the shape characteristics of elongated foreign objects, and the MobileViT module is used to increase the differentiation between the foreign objects and the background in the image, in order to extract features that are more consistent with the diversity of the elongated foreign objects and to

收稿日期: 2023-11-20

责任编辑: 常 琛

DOI: 10.12438/cst.2023-1336

基金项目: 国家重点研发计划资助项目(2022YFC3004700)

作者简介: 高 涵(1999—), 男, 黑龙江哈尔滨人, 硕士研究生。E-mail: gaohan2024@163.com

通讯作者: 赵培培(1979—), 女, 河北三河人, 副教授, 博士。E-mail: zppcumt@163.com

weaken the interference of background noise. Secondly, constructing a low-level feature enhancement module LFEM, to improve the representation of weak semantic features of foreign objects in the detection network in order to reduce the probability of wrong detection. Finally, the introduction of gsconv reduces the information loss due to changes in the size of the feature map and ensures that the network extracts features efficiently while reducing the number of model parameters. The training set and validation set are produced by using the video of conveyor belt work in an underground coal mine. The proposed algorithm is compared to three other conveyor belt detection algorithms, and the experimental results show that the proposed algorithm can better solve the problems of poor detection of elongated objects and difficulty in weak semantic feature extraction in conveyor belt foreign object target detection, with higher detection accuracy. For images with a resolution size of  $640 \times 640$ , the performance metrics mAP@0.5 can be up to 0.875, mAP@0.5:0.95 can be up to 0.543, and the detection speed is 75 fps.

**Key words:** foreign object detection; YOLOv7-tiny; ghost convolution; conveyor belt; deep learning

## 0 引言

中国是世界上最大的煤炭生产国, 2022 年全国煤炭开采量突破 45.6 亿 t, 占中国一年能源消耗的 56%, 煤炭是我国能源战略的压舱石。煤炭的安全开采、高效运输已成为煤炭生产的核心问题<sup>[1]</sup>。输送带是煤矿井下最重要的运输设备, 其运行状态直接影响煤炭开采运输量。在复杂井下开采作业中, 输送带上异物诸如大块矸石、锚杆、铁块等异物, 如不能及时发现并清除可能会引起输送带撕裂、堵塞落煤口等安全事故, 严重影响煤矿生产工作, 威胁煤矿工人生命安全<sup>[2]</sup>。因此对输送物料中混杂的异物进行检测识别, 以便提前预警处理, 对提高井下输送带运行的稳定性和安全性有重要意义。

针对井下复杂恶劣的环境, 国内外众多研究学者很早就尝试使用传统计算机图像处理的方法识别运输带异物。刘富强等<sup>[3]</sup>通过煤和矸石的灰度直方图曲线差异对其进行区分。程建等<sup>[4]</sup>应用高斯混合模型区分输送带输送物料中的矸石等异物。余乐<sup>[5]</sup>构造一种基于非线性灰度压缩-膨胀共生矩阵的煤与矸石识别方法。然而上述方法虽然有处理速度快、信息压缩潜力大等特点, 但仍存在人工提取目标特征困难、多目标检测效果差、受环境影响大的缺点, 使得基于传统图像处理方法的输送带异物检测算法精度较差<sup>[6]</sup>。

基于深度学习的目标检测算法拥有通过自学习发现图像中待检测目标多样性特征的能力, 对于多目标检测任务具有较好的检测能力。其可接单阶段检测算法和双阶段检测算法来分类<sup>[7]</sup>。双阶段检测算法首先需要在特征图上生成候选框, 然后在此基础上修正目标位置和候选区域范围并进行目标分类, 典型算法包括以 Fast R-CNN<sup>[8]</sup>、Faster R-CNN<sup>[9]</sup>等。吕志强<sup>[10]</sup>通过优化传统候选框生成方式, 设计了一种基于 Faster R-CNN 的输送带异物识别算法, 并通

过实验证明该算法具有较高的识别精度。吴守鹏<sup>[11]</sup>等将双向特征金字塔与 Faster R-CNN 结合, 解决输送带异物多尺度问题, 有效提升算法对小尺寸异物的检测能力。单阶段网络跳过了生成候选框阶段, 直接输出目标位置和目标类别, 主要算法包括 SSD<sup>[12]</sup>、YOLOv3<sup>[13]</sup>、YOLOv4<sup>[14]</sup>、YOLO9000<sup>[15]</sup>等。CHEN 等<sup>[16]</sup>使用了 KinD++ 微光图像增强算法, 通过特征处理提高捕获的低质量图像的质量, 并通过 GridMask 方法, 对训练图像中的异物进行部分遮盖, 实现了对输送带异物检测场景的高效检测。来文豪<sup>[17]</sup>等利用多光谱系统采集 3 个波段并构建了伪 RGB 图像数据集, 之后利用 YOLOv4 网络进行检测, 解决了井下图像不清晰的问题。ZHANG 等<sup>[18]</sup>通过引入深度可分离卷积改进 YOLOv4 的主干网络和颈部, 实现了基于 YOLOv4 的轻量化网络, 其检测精度在接受的范围内提升了 200% 以上的速度。XIAO 等<sup>[19]</sup>则通过剪枝算法简化 YOLOv3 网络, 实现了更小的计算、更快的处理和更小的尺寸。相较于传统的图像识别方法, 基于深度学习的目标检测方法更适用于多目标场景, 在保证实时性的同时, 拥有较高的准确率和鲁棒性。

但现有基于深度学习的异物检测算法仍存在不足之处: 首先, 输送带工作图片存在图像模糊、颜色信息单一等特点, 并且异物目标特征主要由形状、边缘、纹理、轮廓等弱语义特征构成, 而现有算法主干网络对异物弱语义信息表征能力弱; 其次, 部分异物诸如锚杆等异物是宽高比大的细长对象。算法主要使用  $3 \times 3$  标准卷积提取目标特征, 对细长物体特征不敏感; 最后, 某些异物的前景内容只占标注框中的很小部分, 所提取出的候选区域特征大部分都是图像的背景区域或其他物体, 使网络学习过程中容易受到相似背景噪声及非本目标特性引导, 导致检测效果差。为此, 本文设计了一种基于低层级特征增强和 Transformer 机制的 YOLOv7-tiny 井下输送带

异物检测算法 FET-YOLO,以期能够自动、快速、准确的发现输送带上的异物,其主要创新点为:①针对细长异物的几何属性,引入可变形卷积<sup>[20]</sup>改善网络的感受野,以更好提取出感兴趣物体的完整特征;②针对异物前景信息少的问题,构建基于 Transformer 机制的全局特征提取模块 (Spatial Pyramid Pooling with Transformer,SPPT),提升目标在标注框范围内的显著度,进而提升网络精度;③提出了一种有效的低层级特征图增强模块 LFEM,该模块通过融合不同尺度的主干网络输出特征图,增强特征图中属于异物目标的语义特征;④在特征融合阶段引入鬼影混洗卷积 GSConv<sup>[21]</sup>替换  $3\times 3$  大小的标准卷积,保证特征融合效率的同时减少特征图冗余。

## 1 YOLOv7-tiny 算法原理

2022 年 8 月 WANG 等提出了 YOLOv7 算法。YOLOv7 在 YOLOv5 的基础上升级改造,使其在检测速度和检测精度上均优于 YOLOv5。YOLOv7-tiny 是 YOLOv7 的轻量化版本,虽然精准度有一定损失,但是在运算速度和参数量上有优势,更适合部

署煤矿井下的边缘计算 GPU 设备,其结构由 Input, Backbone, Neck, Head 四个部分组成。YOLOv7-tiny 整体网络结构如图 1 所示。

Input 部分主要负责对输入特征图作预处理,包括调整输入图像为  $640\times 640$  大小、mosaic 数据增强、自适应锚框计算与随机裁剪等操作。Backbone 部分由高效的聚合网络 (Efficient Layer Aggregation Networks, ELAN) 和改进空间金字塔池化 (CSPSPP) 构成,其中高效的聚合网络通过控制最短和最长的梯度路径,提升网络的特征提取能力,增强系统鲁棒性。CSPSPP 为一个添加了卷积操作的多分支并行化特征融合模块,通过对特征图进行 3 次最大池化,增加网络感受野,有效解决了图像失真的问题,增强网络的线性表达能力;在 Neck 部分,YOLOv7-tiny 采用 PANet 结构保证主干网络不同层次的特征图信息充分融合,避免梯度衰减,增强网络推理信息的传输,Neck 部分接受主干网络 C3, C4, C5 层特征图,并将融合后的特征图传给 Head。在 Head 部分使用 GIOU\_Loss (Generalized Intersection over Union Loss) 函数来估算目标检测矩形框损失。

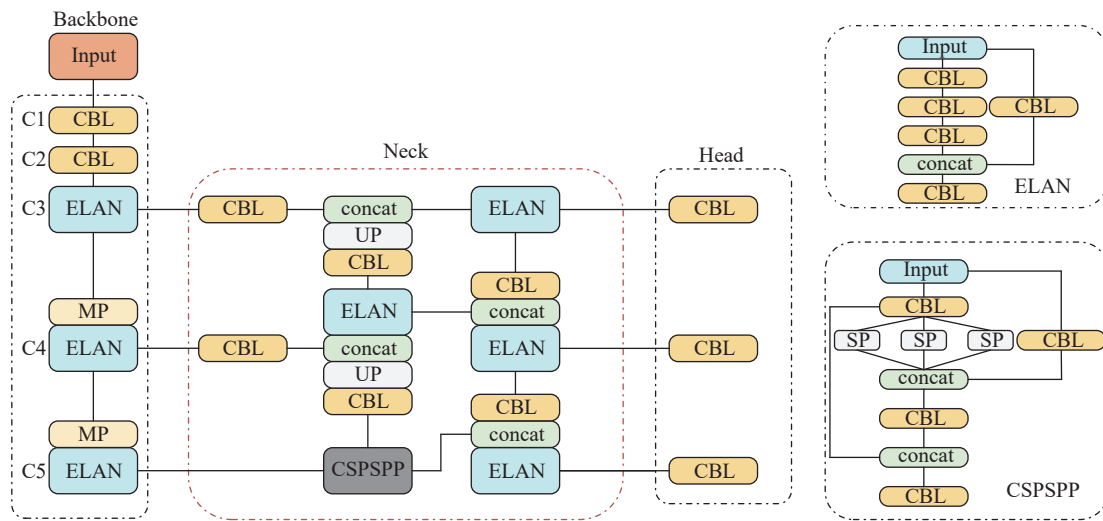


图 1 YOLOv7-tiny 网络结构  
Fig.1 YOLOv7-tiny network structure

YOLOv7-tiny 算法拥有简洁的网络结构,还同时兼顾检测速度和检测精度。不过相较于自然场景下的目标,井下异物目标存在特征弱、信息量少等特点,现有算法仍有优化空间。

## 2 FET-YOLO 网络

本文算法架构如图 2 所示,其计算流程为:首先,训练集图像经由添加可变形卷积的 Backbone 提取特征,再将 C3, C4 和 C5 层特征图输入给 Neck 部分;然

后在 Neck 部分对输入的 C3 和 C4 层特征图使用低层级特征增强模块 LFEM 增强特征图中弱语义特征,对 C5 层特征图使用基于 Transformer 机制的 SPPT 模块增强,提升网络对异物前景信息的偏好,并融合增强后的特征图传给 Head 部分检测,得到训练权重;最后将训练好的网络权重在准备好的测试数据集上得出的客观评价指标对比以证明优化后网络性能的优越性。

### 2.1 可变形卷积

标准的卷积操作仅计算卷积核中心矩形范围内



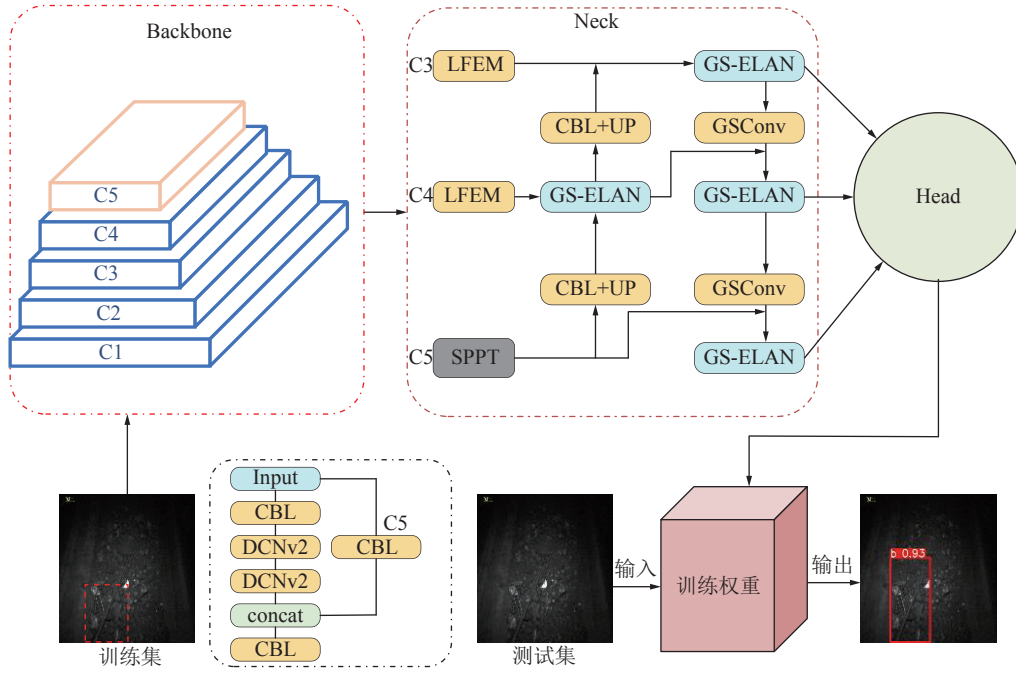


图 2 FET-YOLO 网络结构

Fig.2 FET-YOLO network structure

的像素,无论特征传播的深度和特征图大小如何变化,网络特定位置的感受野始终是固定的。但如图 3 所示中异物所示,异物目标具有宽高比大的几何属性,通过一些仿射变换的数据增强操作不足以弥补的这种场景。



图 3 细长异物

Fig.3 Slender foreign object

引入可变形卷积模块增加对图像中物体的特殊几何属性具备良好的适应性,通过章节 4 中消融实验证明,该改进对井下输送到异物检测任务具有良好的提升效果。

可变形卷积就是卷积核的位置的是可变形的,并非在传统的  $N \times N$  的网格上的卷积,可以通过自适应学习精确定位物体的形状。可变形卷积的操作一般在二维空间中进行的假设卷积和的采样网格点是  $R, R = \{(-1, 1), (-1, 0), \dots, (0, 1), (1, 1)\}$ 。则传统卷积输出的特征图中每个元素有  $p_0$ , 其公式为:

$$y(p_0) = \sum_{p_n \in R} w(p_n) x(p_0 + p_n) \quad (1)$$

式中,  $x$  为输入特征图;  $p_n$  为枚举  $R$  中的第  $n$  个位置  $w(p_n)$  表示  $p_n$  位置的权重。本文使用的可变形卷积模块 DCNv2 在传统卷积操作上增加了一个偏移量  $\{\Delta p_n | n = 1, 2, \dots, N\}$ , 其中  $N = |R|$ , 同时引入惩罚权重  $\Delta m_k$ , 来区分计算的区域是否为该兴趣的区域, 减少了可能引入的无用上下文信息, 增强建模能力。让卷积形变为不规则的卷积, 由式 (1) 可以扩展为

$$y(p_0) = \sum_{p_n \in R} w(p_n) x(p_0 + p_n + \Delta p_n) \Delta m_k \quad (2)$$

因为偏移量  $\Delta p_n$  是一个微小的浮点数, 对于  $x(p_0 + p_n + \Delta p_n)$ , 在计算时可以通过双线性插值得到式 (3):

$$x(p) = \sum_{q \in R} G(q, p) x(q) \quad (3)$$

式中,  $p = p_0 + p_n + \Delta p_n$  为任意位置;  $q$  为枚举输入特征图中的任意位置;  $G$  为二维的双线性插值卷积核, 可以被分为两个一维核, 如式 (4) 所示:

$$G(q, p) = g(q_x, p_x) g(q_y, p_y) \quad (4)$$

其中,  $g(a, b) = \max(0, 1 - |a - b|)$ 。

可变形卷积的运行流程与普通卷积类似, 首先需要额外的卷积层计算输入特征图的偏移量, 再通过双线性插值算法得到最终的输出特征图。在训练阶段, 可变形卷积网络需要同时学习用于输出特征的和计算偏移量的卷积核, 其采样点更能贴近

异物的形状,表征能力更强。模块算法流程图如图 4 所示。

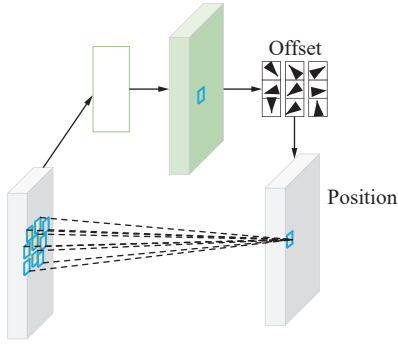


图 4 可变形卷积运行流程

Fig.4 Deformable convolution running process

考虑到可变形卷积可能会造成额外的参数量和计算量,为了保证检测网络的轻量化特性,本文仅在主干网络 C5 部分引入,替换其中  $3\times 3$  大小的标准卷积,以增强主干网络对细长异物的表征能力。

## 2.2 低层级特征增强模块 LFEM

网络实现异物检测需要依赖主干网络输出低层级特征图中包含的形状、纹理、位置等细节信息,这些特征图处于网络的起始阶段,未经过充分处理导致其包含的语义信息少、感受野有限,不能很好的适应输送带异物检测任务需求。

因此本文在金字塔池化模块 (Pyramid Pooling Module, PPM) 的基础上改进,构建低层级特征增强模块 LFEM,通过采用多种尺度不同的池化特征金字塔所构建的多分支结构,拼接形成多通道特征图,横向扩展网络宽度,提高网络对弱语义特征的表征能力的提取能力并增加感受野。LFEM 网络结构如图 5 所示。其中第 4 条分支通过  $1\times 1$  卷积将从前面网络中提取出的特征图调整为输出特征数。其他 3 条支路首先通过全局平均池化 (Global Average Pooling, GAP) 将特征图宽高缩放放到原图大小的  $1/2, 1/4$  和  $1/8$ , 来获取不同尺度的特征信息,然后通过  $1\times 1$  大小卷积降低通道数为输出通道数的  $1/16$ , 最后通过双线性插值算法上采样到输入图像尺寸。最后将 4 条分支的特征图级联,通过  $1\times 1$  卷积调整通道数为输出通道数,以不同尺度的池化操作提取更细腻的特征。特征增强模块 LFEM 的计算过程可以表达为

$$W_1 = f_{up}(f_{Conv}^{1\times 1}(f_{gp}^{1/2}(F))) \quad (5)$$

$$W_2 = f_{up}(f_{Conv}^{1\times 1}(f_{gp}^{1/4}(F))) \quad (6)$$

$$W_3 = f_{up}(f_{Conv}^{1\times 1}(f_{gp}^{1/8}(F))) \quad (7)$$

$$Y = f_{Conv}^{1\times 1}(\text{concat}(W_1, W_2, W_3, F)) \quad (8)$$

式中,  $F$  为输入特征图;  $f_{Conv}^{1\times 1}$  为卷积核大小为  $1\times 1$  的常规卷积操作,主要起到调整输出特征通道数作用;  $f_{gp}^{1/8}$  为全局平均池化,上角标表示池化的特征图缩放倍率;  $f_{up}$  为双线性插值填充上采样,将特征图尺寸调整与  $F$  相同;  $W_1, W_2, W_3$  是 3 条尺度不同的池化特征金字塔的输出;  $Y$  为 LFEM 得到的新特征图。

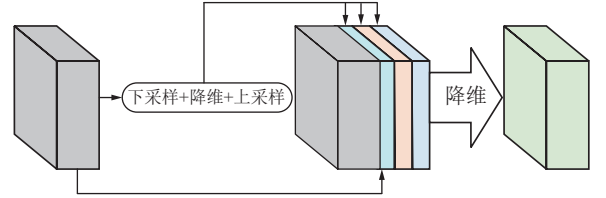


图 5 LFEM 模块

Fig.5 Low-level feature enhancement module

主干网络的低层级输出特征图经过 LFEM 处理后,在不损失分辨率的情况下,提升了网络对于输送带异物图像中颜色信息少,异物边缘不清晰等情况下获取目标特征的能力。

## 2.3 基于 Transformer 的特征融合模块

输送带异物普遍存在前景信息弱的问题,如图 6 所示,异物所占的像素不足其包围框的 10%,主干网络提取的特征受背景噪声主导,导致网络鲁棒性较差,识别效果欠佳。为此,本文重构特征融合模块 CSPSPPT,形成基于 Transformer 模块的特征融合模块 SPPT,通过在下采样后添加 MobileViT 模块<sup>[22]</sup>实现对特征图全局信息提取,以提升主干网络输出特征图对异物特征的偏好,使网络在检测过程中能更关注感兴趣的目标。其结构如图 7 所示。

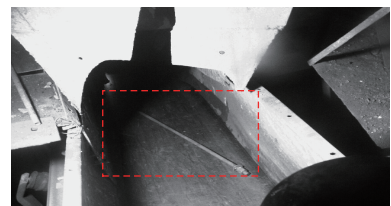


图 6 前景信息弱的异物

Fig.6 Foreign object with weak foreground information

MobileViT 模块通过深度可分离卷积和 Transformer 结合获取特征图的全局信息。相较于传统的 Transformer, MobileViT 模块运算速度更快。对于大小为  $H\times W\times C$  的特征图  $X$  ( $H$  为输入特征图高度,  $W$  为特征图宽度,  $C$  为特征图通道数), 首先通过一个卷积核大小为  $3\times 3$  大小的深度可分离卷积得到局部建模特征图  $X_M$ , 然后通过 unfold 操作将  $X_M$  划分为  $N$  个大小相同的图像块, 每个像素中包含  $P$  个像素 ( $P = w\times h$ ,  $N = (H\times W)/P$ ,  $w$  和  $h$  是预设好的特征块

的宽和高),再将其展开为 $P \times N \times C$ 大小的特征序列 $X_U$ , $X_U$ 中不同像素块中位置相同的像素通过 Transformer 模块处理得到大小为 $P \times N \times C$ 全局特征序列 $X_G$ ,如式 (9) 所示。

$$X_G(p) = \text{Transformer}(X_U(p)), 1 \leq p \leq P \quad (9)$$

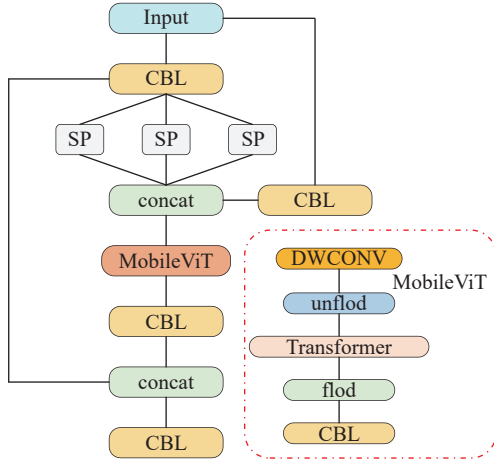


图 7 SPPT 网络结构

Fig.7 SPPT network structure

然后通过 Transpose 与 Reshape 函数组合成的 flod 操作折叠 $X_G$ 到特征图 $X_F$ , $X_F$ 的大小为 $H \times W \times C$ 。与传统的视觉 Transformer 不同, MobileViT 模块通过展平和折叠操作保留了图像块内部像素和图像块之间的位置信息,所以可以不用添加位置编码。最后通过 $1 \times 1$ 卷积作为 MobileViT 模块的输出。

MobileViT 模块对全局信息学习有效性如图 8 所示。图中红色像素点通过 Transformer 与蓝色像素点建立联系,而蓝色像素点已经通过卷积核大小为 $3 \times 3$ 的深度可分离卷积与邻近感受了邻近像素的特征。因此 $X_G$ 中的每一个像素都能对输入特征图 $X$ 中所有的像素进行编码。

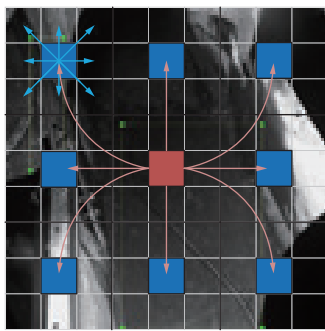


图 8 MobileViT 感受野

Fig.8 MobileViT receptive field

本文通过设计 $h = w = 2$ 来保证 MobileViT 模块输出特征图 $X_w$ 的感受野能够覆盖分辨率为 $H \times W$ 的

输入特征图。相较于直接应该用原版 Transformer 提取全局特征, MobileViT 仅消耗约 1/4 的计算量,符合井下输送带检测任务对实时性的要求。

## 2.4 GSConv 模块

标准卷积 (Standard Convolution, SConv) 通过对输入特征图中所有通道进行点积映射来提取特征,卷积核个数等于输出特征数。过多的卷积层堆叠会生成冗余的特征图,消耗过多的计算量和参数量。Ghost Conv 模块是由 HAN 等<sup>[23]</sup>提出的卷积模块,可以在降低计算量的同时保证模型性能,多用于轻量化网络研究中。其运算操作分为两步,首先利用少量的卷积核对输入特征图进行特征提取,然后对生成的特征图进行线性变化运算,最后将两步操作得到的特征图拼接作为输出特征图。

虽然 Ghost Conv 替代标准卷积可以有效降低网络的计算需求,但模块使用的线性变化运算会损失通道信息。为此, Li H 等提出了 GSConv 模块,其结构如图 9 所示。假设模块的输入通道数为 $C_1$ ,输出通道数为 $C_2$ 。首先通过标准卷积将通道数调整为 $C_2/2$ ,再通过卷积核大小为 $5 \times 5$ 的深度可分离卷积并保证通道数不变,最后将两个卷积操作得到的特征图拼接并混洗。混洗操作可以将输出的特征图均匀打乱,使标准卷积生成的信息渗透到深度可分离卷积生成的信息中以增强通道信息融合,提高网络的表征能力。

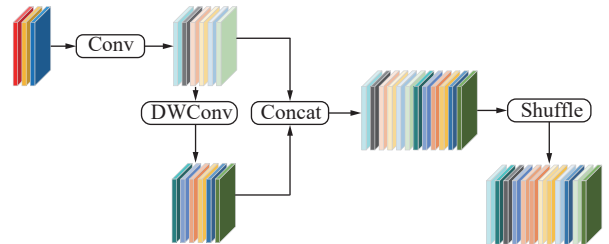


图 9 GSConv 结构图

Fig.9 GSConv network structure

YOLOv7-tiny 网络在 Neck 部分融合不同网络层级的特征图,防止网络退化。随着特征信息不断向深层次网络传递,频繁的对特征图大小和通道数压缩会损失部分语义信息,影响检测精度。为此,本文将 GSConv 引入 Neck 部分中,使用 GSConv 对特征图下采样,减少因采样带来的特征损失,并替换 ELAN 中卷积核大小为 $3 \times 3$ 大小的标准卷积。改进后的 ELAN 模块拥有更好的特征融合能力、更大的感受野,并且参数量和计算量更少。可以在提高网络精度的对网络计算速度。



### 3 模型训练及结构分析

本文的实验平台选择 ubuntu20.04 系统; CPU 型号为 Inter(R) Xeon(R) Platinum 8255 c@2.5 GHz, 共 12 个核心; GPU 型号为 RTX3080, 显存容量为 10 G; 系统内存为 40 G; CUDA 版本为 11.3 版本; Pytorch 框架为 1.11.0 版本。

#### 3.1 数据集制作

本文模型训练所采用的数据集取自某煤矿井下输送带, 输送带的运动速度为 5 m/s, 共取得含有大块矸石、锚杆、槽钢等异物的图像 3 000 张, 使用 Labelimg 工具标注。共标注目标总数 4 458, 其中标签为矸石的数量为 572, 标签为锚杆的数量为 1 598, 标签为铁器的数量为 2 288。通过分析发现数据集现有样本较少且类别不平衡, 难以满足需求, 有过拟合风险, 因此通过采用水平翻转、随机旋转、以及 HSV 变换操作对数据集进行扩充, 扩充后得到图像 6 332 张, 将自建数据集按照 7 : 3 的比例划分为训练集和验证集, 划分后数据集具体信息见表 1。

表 1 数据集详情  
Table 1 Dataset details

数据集	图片数	目标数	矸石	锚杆	铁器
训练集	4 432	7 620	2 412	2 230	2 978
测试集	1 900	3 584	1 020	966	1 598
总数	6 332	11 204	3 432	3 196	4 576

#### 3.2 网络模型训练

实验过程中训练批次 (batch size) 大小设置为 32, 初始学习率为 0.01, 共训练 300 个轮次, 优化器采用 SGD。Momentum 设置为 0.937, Weight\_decay 为 0.000 5, 输入图像尺寸为 640×640。

#### 3.3 评价指标

评价目标检测网络的指标主要有精度 (Precision, P), 召回率 (Recall, R), 平均精度 (Average Precesion, AP), 平均精度均值 (mean Average Precesion, (mAP))。

本文选取 mAP@0.5 和 mAP@0.5: 0.95 作为模型评价指标, mAP@0.5 指 IoU 阈值取 50% 时的 mAP, mAP@0.5:0.95 指 IoU 阈值从 50% 以 5% 为步长增长到 95% 下 mAP 的平均值。各种评价指标的计算公式如下:

$$P = \frac{tp}{tp + fp} \quad (10)$$

$$R = \frac{tp}{tp + fn} \quad (11)$$

$$AP = \int P dR \quad (12)$$

$$mAP = \frac{\sum_{k=1}^C AP(k)}{C} \quad (13)$$

式中, tp 为预测正确的正样本数量; fn 为预测错误样本的负样本数; fp 为预测错误的正样本数。

### 4 实验结果分析

为验证本文算法的有效性, 将本文算法与 YOLOv7-tiny 进行, 结果如图 10 所示。

为了验证本文算法做出的各个改进是否有效, 设计了一组消融实验进行对比分析, 其结果见表 2。表 2 第 1 行为原始 YOLOv7-tiny 网络的检测结果。通过第 2 行可以看出, 添加 GSConv 后, 网络参数量减少 6.6%、mAP@0.5 指标提升 2.3%、在 mAP@0.5: 0.95 提升 2.7%, 证明 GSConv 能够提高网络的特征融合能力, 并减少参数量, 且检测速度没有下降。从第 3 行可以看出添加 MobileViT 后 mAP@0.5 指标提升 1.0%、在 mAP@0.5:0.95 提升 0.4%, 检测速度几乎没有下降, 证明 MobileViT 确实可以增强网络对前景信息的提取能力, 并且对检测速度影响不大, 但其会多使用 0.98 M 的参数。添加 LFEM 后, mAP@0.5 指标提升 0.6%、在 mAP@0.5:0.95 提升 0.6%, 检测速度下降 1.5%; 添加 DCNv2 后网络对细长物体的表征能力有所提升, mAP@0.5 指标提升 2.5%、在 mAP@0.5:0.95 提升 1.8%, 参数量和速度几

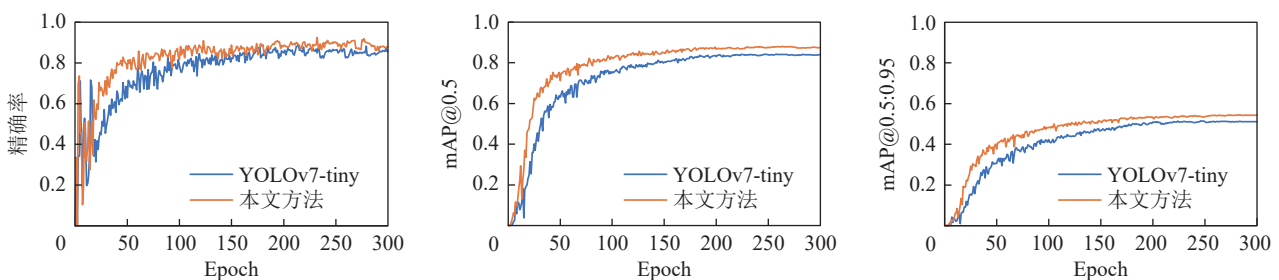


图 10 评价指标

Fig.10 Evaluating indicator



乎没有变化;将 4 种改进融合后 mAP@0.5 指标提升 4.4%、在 mAP@0.5:0.95 提升 5.6%、参数量增加 0.7 M、检测速度下降约 10%,本文算法很好的融合了各个模块的优点,并且有很好的实时性。

为了验证本文算法的性能,选取 3 种前沿的轻量化输送带异物检测算法 YOLOv5.1-scSE<sup>[24]</sup>、CBA-YOLO<sup>[25]</sup>、Fire-Dense-YOLO<sup>[26]</sup>。这 3 种算法

对煤矿井下环境有特殊优化,在实际场景中拥有很高的检测精度和实时性,与其对比可以保证实验的可信度和算法的有效性。选取含有细长形状异物、低层级信息弱、异物形状纹理信息不明显、含大块矸石 4 种典型场景进行测试验证,如图 11a 所示,图中目标已用红框标注,对比结果如图 11b~11e 所示。

表 2 消融实验结果  
Table 2 Ablation results

GSConv	MobileViT	LFEM	DCNv2	mAP@0.5	mAP@0.5: 0.95	参数量/M	检测速度/fps
				0.838	0.514	6.01	83.72
√				0.857	0.528	5.61	81.72
	√			0.846	0.516	6.99	76.32
		√		0.843	0.517	6.02	82.3
			√	0.859	0.523	6.13	81.46
√	√	√	√	0.875	0.543	6.73	75.26

注:空白为未添加;√为添加模版。

从第 1 组试验可以看出,YOLOv5.1-scSE, CBA-YOLO, Fire-Dense-YOLO 算法生成的包围框不能很好的涵盖异物,而且存在漏检和错检的情况,1 个异物会被识别成 2 个目标。而本文算法通过在主干网络添加可变形卷积操作提高了网络对细长物体的特征的提取能力,生成的检测框更贴合目标。

从第 2 组试验可以看出,因光照不足影响,异物混合在煤堆中,其特征诸如形状、纹理、边缘等非常不明显。Fire-Dense-YOLO 算法只确定出其位置,而 YOLOv5.1-scSE 和 CBA-YOLO 算法均漏检。而本文算法通过在低层级添加特征增强模块 LFEM,增强低层级特征图中特征信息,进而能有效将目标从相似背景中提取出来。

从第 3 组试验可以看出,原图中一个大异物的包围框包含很多小异物,且大异物在其包围框的占比很低。在 Fire-Dense-YOLO, YOLOv5.1-scSE, CBA-YOLO 生成的结果中,大异物干扰了小异物的识别,并且,大异物自身的识别效果,而本文算法通过引入 SPPT 模块,实现了对特征图的全局信息提取,增强异物的前景信息,并弱化不相关区域对本目标的影响,进一步提升了网络对目标的表征能力,较好的解决了漏检问题。

从第 4 组试验可以看出,矸石目标和背景中的煤块形状相似度很高,YOLOv5.1-scSE, CBA-YOLO, Fire-Dense-YOLO 均出现错检漏检的情况,而本文算法的通过 LFEM 模块增强矸石的纹理、颜色信息,有

效提高了网络对矸石的区分度,从而提高网络的检测精度。

为了更客观展示本文算法的优越性,将常用的轻量级目标检测网络诸如 YOLOX-tiny<sup>[27]</sup>、YOLOv4-tiny, YOLOv5 s, DETR<sup>[28]</sup> 等传统目标检测网络和上文提到的 3 种前沿输送带异物检测算法在相同数据集和训练环境下进行对比,对比实验结果见表 3。

表 3 对比实验结果  
Table 3 Comparative experimental results

模型	mAP@0.5	mAP@0.5:0.95	参数量/M	FPS
YOLOv5.1-scSE	0.87	0.528	7.09	70.14
CBA-YOLO	0.863	0.523	7.07	72.2
Fire-Dense-YOLO	0.831	0.472	4.83	117.06
SSD	0.645	0.371	40.1	70.22
YOLOv3-tiny	0.792	0.443	8.67	116.95
YOLOv4-tiny	0.818	0.464	6.27	83.6
YOLOX-tiny	0.837	0.489	5.03	114.16
YOLOv5 s	0.857	0.527	7.02	76.38
DETR	0.808	0.445	41.28	28.44
YOLOv7-tiny	0.838	0.514	6.01	83.72
本文方法	0.875	0.543	6.73	75.26

由表 3 结果可以看出,本文算法在 mAP@0.5 和 mAP@0.5:0.95 指标上远高于其他 9 种算法,虽然在检测速度和参数量上略低于 Fire-Dense-YOLO 和 YOLOv3-tiny 等算法,但其仍然满足实时性要求。与

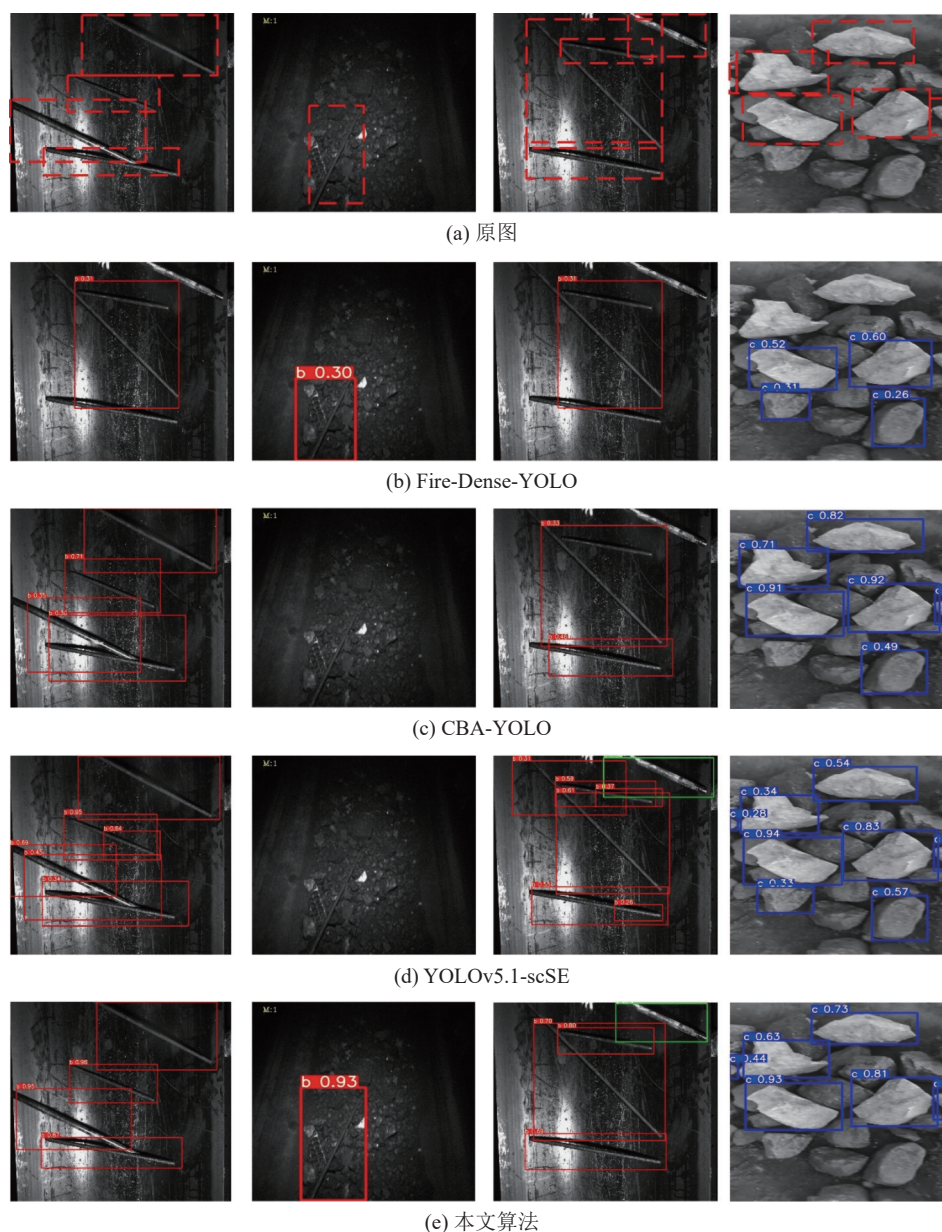


图 11 检测结果对比

Fig.11 Comparison of test results

同为 Transformer 架构的 DETR 网络相比,本文算法不仅在模型收敛速度上占优,而且在 4 项性能指标均有领先。因此,本文算法相较于其他算法具有明显优势,能够有效的检测出输送带上的异物,提高运输效率。

## 5 结 论

1) 基于低层级特征增强和 Transformer 机制设计了一种轻量级输送带异物检测算法 FET-YOLO。该算法通过可变形卷积克服现有网络模型对细长物体特征困难的问题;采用 GSConv 增强特征融合能力,减少因特征图大小变化造成的通道信息损失;引入 MobileViT 模块增强异物目标前景信息,减少背

景特征信息对异物检测的干扰;加入低层级特征增强模块 LFEM,增加网络对低层级特征的提取能力,进而解决目标难以检测的问题。

2) 通过大量试验证明本算法在  $640 \times 640$  分辨率的输送带异物图像上的平均精度可达 0.875,检测速度达到 75fps,性能远优于 CBA-YOLO, Fire-Dense-YOLO, YOLOv5.1-scSE 等轻量级输送带异物检测算法。

3) 本文模型在当前数据集下表现良好,但是对于泛化能力的提升仍需进一步研究。此外模型基于较少数据训练所得,存在一定程度的过拟合问题,未来将进一步扩展数据集的尺寸和场景的丰富程度,以满足矿井下复杂场景的应用要求。

4) 本文提出的网络能够有效提升井下输送带异物检测效果,降低安全隐患,保障运输效率,为无人化、智能化煤炭开采提供理论基础,促进计算机视觉技术在煤炭输送领域的发展与应用。

#### 参考文献(References):

- [1] 中矿煤炭产业景气指数研究课题组, 吴吟. 2022-2023 年中国煤炭产业经济形势研究报告[J]. 中国煤炭, 2023, 49(3): 2-10.  
China Mining Coal Industry Prosperity Index RESEARCH Group, WU Yin. Research report on the economic situation of China's coal industry in, 2022-2023[J]. China Coal, 2023, 49(3): 2-10.
- [2] 程德强, 徐进洋, 寇旗旗, 等. 融合残差信息轻量级网络的运煤皮带异物分类[J]. 煤炭学报, 2022, 47(3): 1361-1369.  
CHENG D Q, XU J Y, KOU Q Q, *et al.* Lightweight network for coal belt foreign object classification with fusion of residual information[J]. Journal of Coal, 2022, 47(3): 1361-1369.
- [3] 刘富强, 钱建生, 王新红, 等. 基于图像处理与识别技术的煤矿矸石自动分选[J]. 煤炭学报, 2000, 45(5): 534-537.  
LIU Fuqiang, QIAN Jiansheng, WANG Xinhong, *et al.* Automatic sorting of coal mine gangue based on image processing and recognition technology[J]. Journal of Coal, 2000, 45(5): 534-537.
- [4] 程健, 王东伟, 杨凌凯, 等. 一种改进的高斯混合模型煤矸石视频检测方法[J]. 中南大学学报(自然科学版), 2018, 49(1): 118-123.  
CHENG J, WANG D W, YANG L K, *et al.* An improved video detection method for coal gangue with Gaussian mixture model[J]. Journal of Central South University(Natural Science Edition), 2018, 49(1): 118-123.
- [5] 余乐. 一种煤和煤矸石图像识别的新方法[J]. 现代计算机(专业版), 2017(17): 66-70.  
YU Le. A new method for image recognition of coal and coal gangue[J]. Modern Computer(Professional Edition), 2017(17): 66-70.
- [6] 李明, 鹿朋, 朱美强, 等. 基于改进 YOLO-tiny 的闸板阀开度检测[J]. 煤炭学报, 2021, 46(S2): 1180-1190.  
LI M, LU P, ZHU M Q, *et al.* Gate valve opening detection based on improved YOLO-tiny[J]. Journal of Coal, 2021, 46(S2): 1180-90.
- [7] 赵永强, 饶元, 董世鹏, 等. 深度学习目标检测方法综述[J]. 中国图象图形学报, 2020, 25(4): 629-54.
- [8] He K, Gkioxari G, Dollár P, *et al.* Mask R-CNN[J]. Ieee Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 386-397.
- [9] REN S, HE K, GIRSHICK R, *et al.* Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Trans Pattern Anal Mach Intell, 2017, 39(6): 1137-1149.
- [10] 吕志强. 复杂环境下煤矸皮带运输异物图像识别研究[D]. 徐州: 中国矿业大学, 2020.  
LV Zhiqiang. Research on foreign body image recognition in coal mine belt transportation under complex environment [D]. Xuzhou: China University of Mining and Technology, 2020.
- [11] 吴守鹏, 丁恩杰, 俞啸. 基于改进 FPN 的输送带异物识别方法[J]. 煤矿安全, 2019, 50(12): 127-30.  
WU Shoupeng, DING Enjie, YU Xiao. Foreign body identification method in conveyor belt based on improved FPN[J]. Safety in Coal Mines, 2019, 50(12): 127-30.
- [12] LIU W, ANGUELOV D, ERHAN D, *et al.* SSD: Single Shot MultiBox Detector[C]// proceedings of the 14th European Conference on Computer Vision( ECCV ). Amsterdam, NETHERLANDS, 2016: 21-37.
- [13] REDMON J, FARHADI A J A E-P. YOLOv3: An Incremental Improvement [J]. Arxiv, 2018.
- [14] BOCHKOVSKIY A, WANG C-Y, MARK LIAO H-Y. YOLOv4: Optimal Speed and Accuracy of Object Detection [J]. Arxiv, 2020.
- [15] Redmon J, Farhadi A, Ieee. YOLO9000: Better, Faster, Stronger[C]. 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 6517-6525.
- [16] CHEN Y, SUN X, XU L, *et al.* Application of YOLOv4 Algorithm for Foreign Object Detection on a Belt Conveyor in a Low-Illumination Environment[J]. SENSORS, 2022, 22(18): 6851.
- [17] 来文豪, 周孟然, 王锦国, 等. 多光谱波段筛选的煤矸石快速定位[J]. 中国激光, 2021, 48(16): 190-200.  
LAI W H, ZHOU M R, WANG J G, *et al.* Rapid localization of coal gangue by multispectral waveband screening[J]. China Laser, 2021, 48(16): 190-200.
- [18] ZHANG M C, CAO Y S, JIANG K, *et al.* Proactive measures to prevent conveyor belt Failures: Deep Learning-based faster foreign object detection[J]. Engineering Failure Analysis, 2022, 141: 14.
- [19] XIAO D, LIU P P, WANG J C, *et al.* Mining belt foreign body detection method based on YOLOv4\_GECA model [J]. Scientific Reports, 2023, 13(1).
- [20] ZHU X, HU H, LIN S, *et al.* Deformable ConvNets v2: MORE Deformable, Better Results [C]//, 2019 IEEE/CVF conference on computer vision and pattern recognition(CVPR 2019). Shanghai, 2019: 9300-9308.
- [21] Li H, Li J, Wei H, *et al.* Slim-neck by GSConv: A better design paradigm of detector architectures for autonomous vehicles[J]. arXiv preprint arXiv: 2206.02424, 2022.
- [22] Mehta S, Rastegari M. MobileViT: Light-weight, General-purpose, and Mobile-friendly Vision Transformer[J]. arXiv preprint arXiv: 2110.02178, 2021.
- [23] Han K, Wang Y, Tian Q, *et al.* GhostNet: More Features from Cheap Operations[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) [C]. Seattle. IEEE, 2020: 1577-1586.
- [24] YAN P, SUN Q, YIN N, *et al.* Detection of coal and gangue based on improved YOLOv5.1 which embedded scSE module\* [J]. MEASUREMENT, 2022, 188.
- [25] 桂方俊, 李尧. 基于 CBA-YOLO 模型的煤矸石检测[J]. 工矿自动化, 2022, 48(6): 128-33.  
GUI Fangjun, LI Yao. Coal gangue detection based on CBA-YOLO model[J]. Automation of Industry and Mine, 2022, 48(6): 128-33.
- [26] 许鹏. 基于边缘计算的煤矿井下皮带异物检测关键技术研究[D]. 徐州: 中国矿业大学, 2021.  
XU P. Research on key technology of detecting foreign matter in coal mine belt based on edge computing [D]. Xuzhou: China University of Mining and Technology, 2021.
- [27] GE Z, LIU S, WANG F, *et al.* YOLOX: Exceeding YOLO Series in, 2021 [J]. Arxiv, 2021.
- [28] CARION N, MASSA F, SYNNAEVE G, *et al.* End-to-End Object Detection with Transformers [J], Arxiv 2020.