

基于 CA-YOLOv8 的输送带大块煤检测方法

油亚鹏¹⁾, 马波^{1,2,3)*}, 赵乐¹⁾, 王鹏琪¹⁾

¹⁾ (北京化工大学机电工程学院 北京 100029)

²⁾ (北京化工大学高端压缩机及系统技术全国重点实验室 北京 100029)

³⁾ (北京化工大学高端机械装备健康监控与自愈化北京市重点实验室 北京 100029)
(mabo@buct.edu.cn)

摘要: 针对煤矿环境复杂多变以及输送带图像模糊, 导致大块煤难以被准确检测及泛化性差的问题, 提出一种基于 CA-YOLOv8 的大块煤检测方法。首先针对输送带图像特点, 采用图像复原技术对模糊的输送带图像进行预处理, 使得图像恢复清晰, 为后续的大块煤检测提供了高质量的图像输入; 然后通过分析大块煤特征对 YOLOv8 算法进行改进, 包括在输入端对图像进行基于现实尺寸的自适应缩放、网络结构中嵌入坐标注意力机制以及移除多余的检测头, 实现对大块煤的准确、稳定和高效的检测。在由某煤矿 4 条不同的输送带的监控视频图像制作的数据集上进行实验, 结果表明, 所提方法在相同环境条件下精确率达 95.6%, 召回率达 89.0%; 与原 YOLOv8 算法相比, 在不同环境条件下精确率提升超 17.0%, 召回率提升超 15.3%。

关键词: 大块煤检测; 输送带; 图像复原; YOLOv8; 坐标注意力机制

中图分类号: TP391.41 **DOI:** 10.3724/SP.J.1089.2023-00471

Detection Method of Large Coal Blocks on Conveyor Belt Based on CA-YOLOv8

You Yapeng¹⁾, Ma Bo^{1,2,3)*}, Zhao Le¹⁾, and Wang Pengqi¹⁾

¹⁾ (College of Mechanical and Electrical Engineering, Beijing University of Chemical Technology, Beijing 100029)

²⁾ (National Key Laboratory of Advanced Compressor and System Technology, Beijing University of Chemical Technology, Beijing 100029)

³⁾ (Beijing Key Laboratory of Health Monitoring and Self-Recovery for High-End Mechanical Equipment, Beijing University of Chemical Technology, Beijing 100029)

Abstract: The complex and variable environment of coal mines, coupled with the blurriness of conveyor belt images, results in challenges for accurately detecting large coal blocks and poor model generalization. In this situation, a detection method for large coal blocks based on CA-YOLOv8 is proposed. Firstly, considering the characteristics of conveyor belt images, image restoration is used to preprocess the blurred images, thereby enhancing their clarity and providing high-quality inputs for detection of large coal blocks. Then, based on the characteristics of large coal blocks, the YOLOv8 algorithm is improved, including adaptive scaling of the images at the input based on real size, embedding a coordinate attention in the network structure, and removing unnecessary detection heads, which facilitate accurate, stable, and efficient detection of large coal blocks. Experiments conducted on a dataset created from surveillance video images of four different conveyor belts in a coal mine demonstrate that the proposed method achieve a precision of over 95.6% and a recall of over 89.0% under the same environmental conditions. Compared to the original YOLOv8 algorithm, the precision is improved by over 17.0% and the recall is improved by over 15.3% under different environmental conditions.

收稿日期: 2023-07-24; 修回日期: 2024-02-02. 基金项目: 国家自然科学基金(21978013). 油亚鹏(1998—), 男, 硕士研究生, 主要研究方向为计算机视觉、图像处理; 马波(1977—), 男, 博士, 教授, 硕士生导师, 论文通信作者, 主要研究方向为信号处理、模式识别、故障诊断; 赵乐(1999—), 女, 硕士研究生, 主要研究方向为计算机视觉、图像处理; 王鹏琪(2001—), 男, 硕士研究生, 主要研究方向为信号处理、模式识别。

Key words: detection of large coal blocks; conveyor belt; image restoration; YOLOv8; coordinate attention

输送带是煤炭开采和运输过程中的重要设备,在其运行过程中常常因混入未被充分破碎的大块煤引起转接处堵塞,造成煤料堆积,或导致煤料分布不均,引发输送带跑偏、加剧磨损甚至撕裂,严重威胁到煤炭的安全高效生产^[1]. 运用计算机视觉技术对输送带进行智能监控,自动准确地检测出大块煤,对保障安全生产和推动煤矿智能化建设至关重要.

目前,基于计算机视觉的大块煤检测方法包括传统的图像处理方法和基于深度学习的目标检测方法^[2]. 传统的图像处理方法通常依赖于图像处理技术提取边缘特征,使用阈值分割的方式识别煤块^[3-5],存在检测精度低和实时性差等问题. 相比之下,基于深度学习的目标检测方法在大块煤检测中表现出更高的识别准确率和更强的鲁棒性,并得到了广泛的推广和应用. 杨建辉等^[6]通过图像灰度化方法结合光照补偿处理,实现对输送带实时图像进行亮度校正和无关像素剔除,再通过构造双通道金字塔卷积模型完成输送带异物识别;郝帅等^[7]通过自适应直方图均衡算法增强煤矿井下输送带图像的对比度,在 YOLOv5 算法框架下引入卷积块注意力模型提升目标检测精度;Xiao 等^[8]采用中值滤波方法提高矿石边缘的清晰度,并改进 YOLOv3 算法以提升处理速度,有效地实现了矿石输送带上异物的实时识别;Song 等^[9]通过数据增强方法扩大数据集,并改进 YOLOv5 算法,以提升特征提取能力.

煤矿环境复杂多变,煤尘的干扰、昏暗不均的光照以及输送带的高速运动等因素,使得采集到的图像经常存在噪声、模糊不清、缺失细节信息及物体轮廓不明显等图像退化现象^[7],直接影响图像的质量,从而降低了目标检测的准确性. 尽管上述方法在处理这些问题上已取得了一定的进展,但未充分考虑在不同环境条件下模型的泛化能力,限制了其在实际应用中的效果.

针对煤矿环境中大块煤难以被准确检测及泛化性差的问题,本文结合坐标注意力机制(coordinate attention, CA)^[10]和 YOLO(you only look once)算法^[11]构建 CA-YOLOv8 模型,并基于 CA-YOLOv8 模型提出一种大块煤检测方法. CA-YOLOv8 模型在 YOLOv8 算法的基础上,通过分析大块煤特征,有针对性地改进其输入端和网络结构. 本文方法利用图像复原技术对模糊的输送带图像进行预处理,使得图像恢复清晰,获得高质量的图像;

并将其输入 CA-YOLOv8 模型进行目标识别与定位,完成对大块煤的检测. 最后,通过实际煤矿输送带监控视频,验证了本文方法的有效性.

1 相关工作

1.1 YOLO 算法

2016 年,Redmon 等^[11]提出首个单阶段目标检测算法 YOLO. 该算法以回归思想为基础,同步进行目标识别与目标定位,实现了端到端的目标检测,极大地提高了目标检测速度. 目前,YOLO 系列算法已经从 YOLOv1 发展到 YOLOv8,并衍生出多种改进算法,如 YOLOX 和 YOLOR 成为实时目标检测的主流算法^[12],已经在安防^[13]、工业^[14]、医疗^[15]、交通^[16]等领域取得了很好的应用效果.

通常,YOLO 算法的网络结构分为 4 个部分:输入端,Backbone,Neck 和 Head. 输入端对输入图片进行数据增强、自适应图片缩放等操作,数据增强包括色调、饱和度、亮度等颜色变换,旋转、平移、缩放、裁剪、翻转等几何变换,以及 Mosaic^[17], Mixup^[18]等特殊变换. Backbone 采用卷积神经网络提取图像特征,并引入跨阶段局部网络(cross stage partial network, CSPNet)^[19],在保证检测精度的同时提高检测速度. Neck 引入空间金字塔池化(spatial pyramid pooling, SPP)^[20],在多个尺度上进行池化操作以增大感受野,捕获不同尺度下的特征;同时采用特征金字塔网络(feature pyramid network, FPN)^[21]结合路径聚合网络(path aggregation network, PAN)^[22]的结构,汇集不同尺度的图像语义信息与图像位置信息,进行多尺度的特征融合. Head 负责预测目标类别和边界框,以交并比(intersection over union, IOU)^[23]及其各类改进方法作为边框回归的损失函数,通过非极大值抑制(non-maximum suppression, NMS)方法^[24]输出最优结果.

图 1 所示为 YOLOv8 算法的结构. 其中, CBS 层由 1 层卷积(convolution, Conv), 1 层批归一化(batch normalization, BN), 1 层 SiLU(Sigmoid Linear Unit)激活函数构成; C2f(CSP bottleneck with 2 convolutions fast)层在 CSP 模块的基础上增加了更多的跳层连接,取消了分支中的卷积操作,让特征信息更丰富的同时,减少了计算量; SPPF(spatial pyramid pooling fast)层采用多个小尺寸池化核级联代替 SPP 层中单个大尺寸池化核,在融合不同

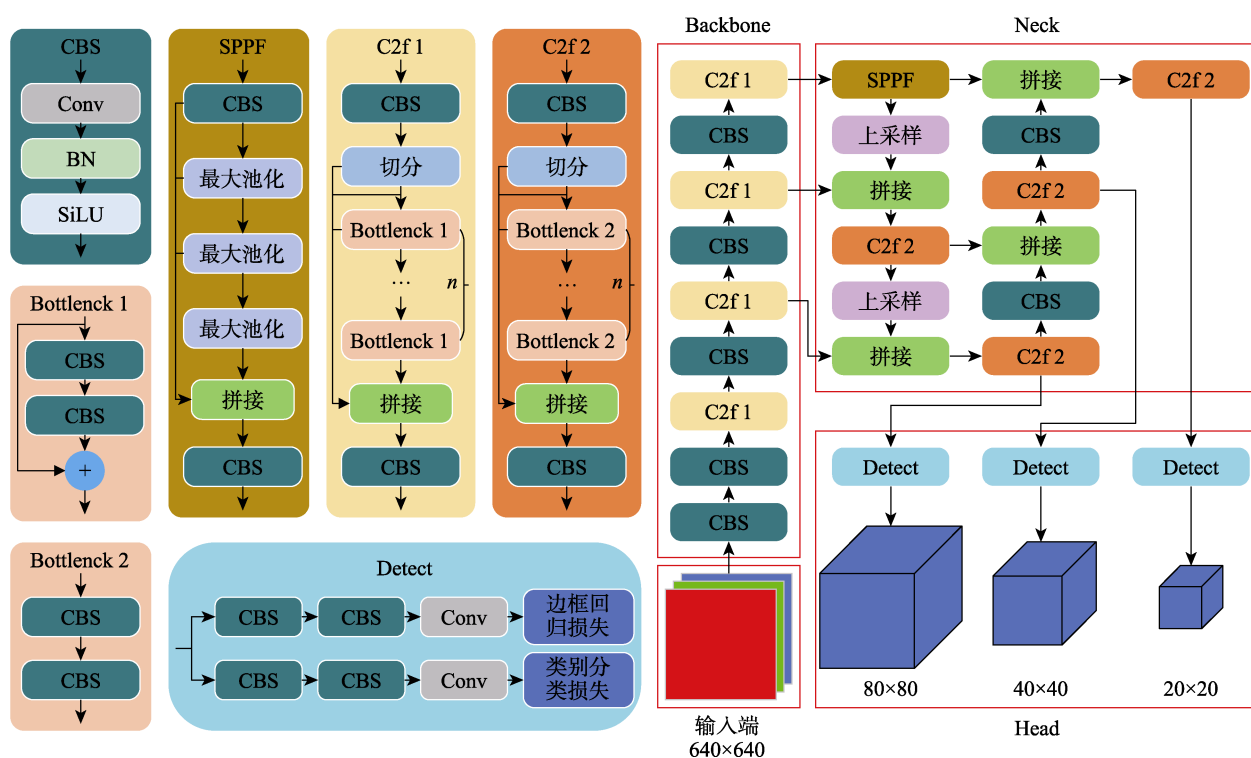


图1 YOLOv8 算法结构

感受野的特征图、丰富特征图的表达能力的情况下, 进一步提高了运行速度。

1.2 CA

注意力机制源于对人类视觉系统的研究, 其主要思想是在处理复杂的输入信息时, 神经网络应该能够集中注意力于最相关或最重要的部分, 而不是平等地对待所有信息^[25]。通过引入注意力机制, 可以有效地提高模型的性能和泛化能力^[26]。目前, 注意力机制被广泛用于图像分类、目标检测、语义分割等各种不同的计算机视觉任务中^[27]。

CA 通过将位置信息嵌入到通道注意力中, 使网络能够关注大区域, 同时避免产生显著的计算开销。与常规的通道注意力只关注编码通道间的信息而忽视位置信息的重要性不同, 坐标注意力将通道注意力分解为2个一维特征编码过程, 分别沿2个空间方向进行特征聚合。这种方式使得能够在一个空间方向上捕获到远程依赖关系, 同时在另一个空间方向上保留精确的位置信息; 然后将得到的特征图单独编码成一对方向感知和位置敏感的注意力图, 这些图以互补的方式应用于输入特征图, 增强感兴趣对象的表示。CA 结构如图2所示。

2 本文方法

本文对煤矿输送带图像特点以及大块煤特征,

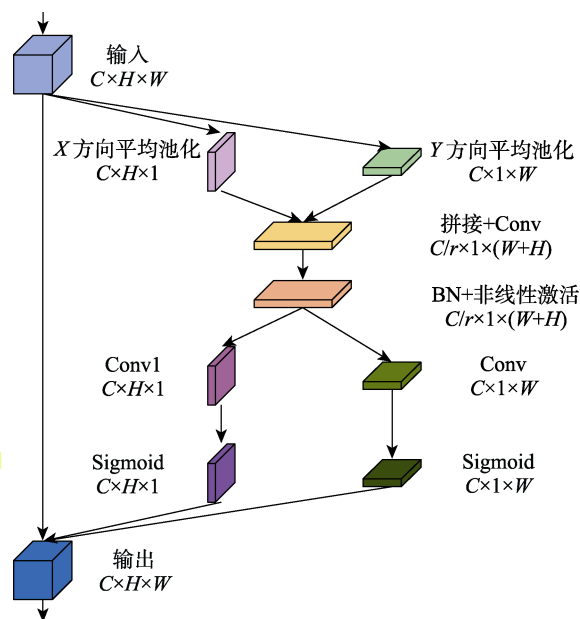


图2 CA 结构

提出一种专门针对输送带大块煤的检测方法, 其流程图如图3所示。首先针对输送带图像特点, 采用图像复原技术进行预处理, 获取高质量的图像; 然后分析大块煤特征, 在 YOLOv8 算法的基础上, 有针对性地对其输入端、网络结构进行改进, 构建 CA-YOLOv8 模型, 旨在增强模型对大块煤的识别能力和在不同环境下的泛化能力, 实现对煤矿输送带上大块煤的准确、稳定和高效的检测。

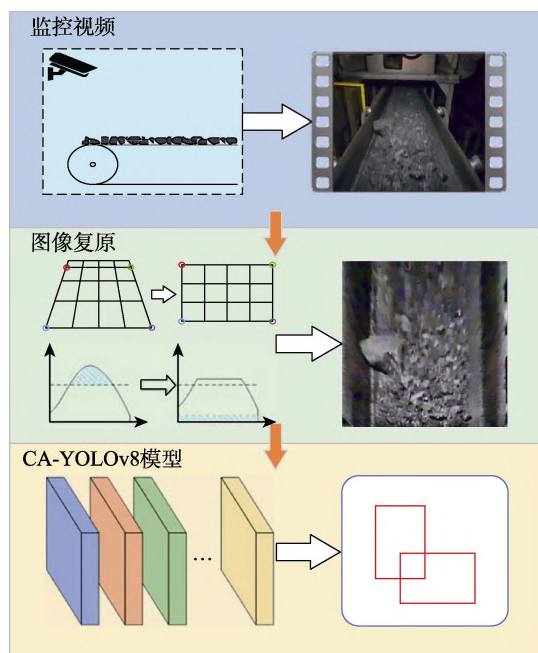


图3 本文方法流程

2.1 图像复原及步骤

2.1.1 输送带图像退化分析

以图像左上角点作为原点, 图像水平方向和垂直方向分别作为 X 轴与 Y 轴建立坐标系, 坐标示意图如图 4 所示.

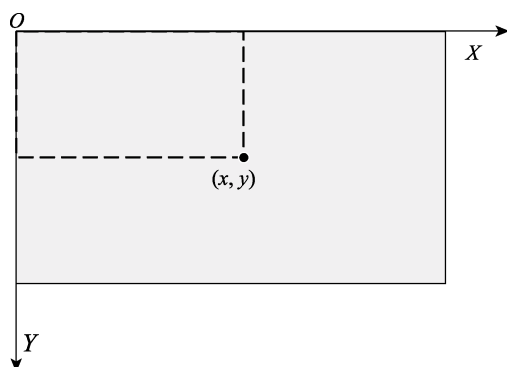


图4 输送带图像坐标示意图

设输送带原始图像为 $f(x, y)$, 输送带图像在形成、记录等过程中, 受光照 $l(x, y)$ 的影响产生灰度失真, 受输送带运动 $m(x, y)$ 的影响产生运动模糊, 受摄像头姿态 $k(x, y)$ 的影响产生几何失真, 受煤尘 $\eta(x, y)$ 等因素的影响产生噪声干扰, 导致摄像头采集获得的图像 $p(x, y)$ 质量有所下降. 经分析, 退化过程如图 5 所示.

根据输送带图像的退化过程建立输送带图像退化模型, 图像 $f(x, y)$ 逐步退化至图像 $p(x, y)$ 的

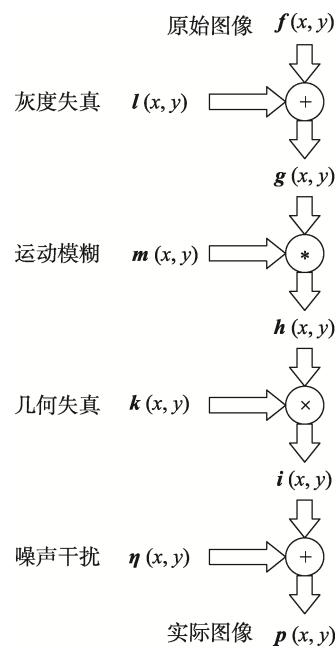


图5 输送带图像退化过程

过程为

$$p(x, y) = [f(x, y) + l(x, y)] * m(x, y) \times k(x, y) + \eta(x, y).$$

2.1.2 图像复原步骤

图像复原是在退化模型的基础上, 采用与退化过程相反的过程来复原图像, 将实际图像 $p(x, y)$ 重建成接近于原始图像 $f(x, y)$ 的过程. 由于图像复原时只有退化的实际图像, 而原始的理想图像是无法获知的, 因此图像复原实际上是对理想图像的估计过程. 复原后的输送带图像称为 $f'(x, y)$, 输送带图像复原过程为

$$f'(x, y) = I_r(p(x, y), \eta(x, y), k(x, y), m(x, y), l(x, y)).$$

其中, $I_r(\cdot)$ 表示图像复原模型. 图像复原过程的步骤如下.

Step1. 消除噪声干扰. 因为噪声干扰 $\eta(x, y)$ 难以通过有限样本直接估计, 所以采用双边滤波对图像 $p(x, y)$ 进行去噪, 在滤除噪声的同时保存边缘信息, 去噪后得到图像 $i'(x, y)$, 图像去噪过程为

$$i'(x, y) = \frac{\sum_{(u,v) \in S(x,y)} p(u,v) \times w(x, y, u, v)}{\sum_{(u,v) \in S(x,y)} w(x, y, u, v)},$$

$$w(x, y, u, v) = w_s(x, y, u, v) \times w_r(x, y, u, v) =$$

$$e^{-\frac{(x-u)^2 + (y-v)^2}{2\sigma_s^2}} \times e^{-\frac{\|p(x, y) - p(u, v)\|^2}{2\sigma_r^2}}.$$

其中, $S(x, y)$ 表示以 (x, y) 为中心的 $c \times c$ 的矩阵; σ_s 表

示空间邻近度标准差; σ_r 表示颜色相似度标准差。

Step2. 校正几何失真. 采用透视变换^[28]校正图像中的几何失真, 由于透视变换适用于平面图像矫正, 而传送带的表面并不是完全平面的, 因此使用透视变换时, 在离摄像头较远的位置大块煤的形状会出现失真. 为了解决这一问题, 在图 $i'(x, y)$ 中选取输送带离摄像头较近区域内的 4 个顶点 O_s , 对应于校正后图像 $h'(x, y)$ 中顶点 O_d , 如图 6 所示. 几何校正过程为

$$h'(x, y) = K' \times i'(x, y).$$

其中, K' 表示根据顶点 O_s 和 O_d 求得的透视矩阵。

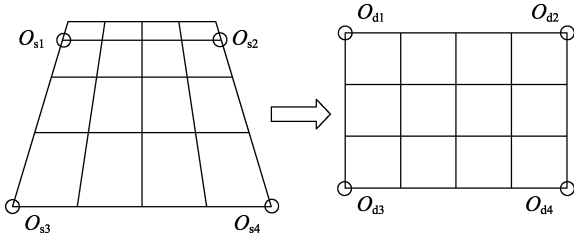


图 6 几何失真校正示意图

Step3. 消除因运动引起的模糊. 输送带与监控设备间作相对匀速直线运动, 且运动方向与图像 X 轴垂直, 则运动模糊退化函数为

$$m'(x, y) = \begin{cases} \frac{1}{d}, & \text{if } 0 \leq y \leq d \text{ and } x = 0 \\ 0, & \text{otherwise} \end{cases}$$

其中, d 表示退化函数的长度, 与输送带运动速度和监控设备曝光时间相关。

采用维纳滤波对图像 $h'(x, y)$ 进行复原, 得到图像 $g'(x, y)$, 图像去模糊过程为

$$G'(u, v) = \frac{|M'(u, v)|^2}{|M'(u, v)|^2 + 10^{-\frac{z}{10}}} \cdot \frac{H'(u, v)}{M'(u, v)}.$$

其中, $G'(u, v)$, $H'(u, v)$ 和 $M'(u, v)$ 分别表示 $g'(x, y)$, $h'(x, y)$ 和 $m'(x, y)$ 的频域数据; z 表示图像信噪比。

Step4. 亮度均衡化. 采用限制对比度自适应直方图均衡化(contrast limited adaptive histogram equalization, CLAHE)^[29] $h_{c,a}(\cdot)$ 对图像 $g'(x, y)$ 进行处理, 抑制过亮区域, 同时提亮昏暗区域, 得到图像 $f'(x, y)$, 亮度均衡化过程为

$$f'(x, y) = h_{c,a}(g'(x, y), s, t).$$

亮度均衡化过程中, 先将图像分成 $s \times s$ 块, 求得每块图像的直方图 $R_i = \{r_j | j = 0, 1, \dots, 255\}$, $i = 0, 1, \dots, s \times s$. 其中, r_j 表示图像中灰度级为 j 的像素数量; 然后根据阈值系数 t 对直方图进行裁切得到 R'_i , 如图 7 所示; 再将裁切后的直方图进行均衡化得到 R''_i ; 最后采用双线性插值重构, 得到图像 $f'(x, y)$ 。

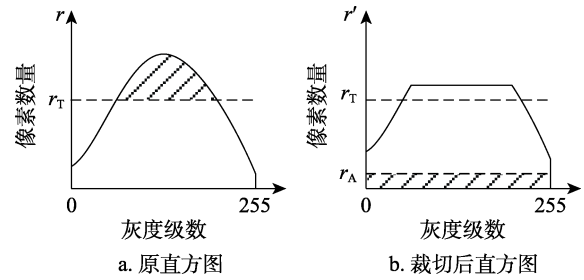


图 7 CLAHE 变换过程

图 7 中, 像素数量阈值 r_T , 像素数量分配均值 r_A 的计算公式为

$$r_T = \max\left(t \times \frac{w_s \times h_s}{256}, 1\right),$$

$$r_A = \sum_{j=0}^{255} \max(r_j - r_T, 0) / 256.$$

其中, w_s 和 h_s 分别表示每个子块在宽、高方向的像素个数。

2.2 检测模型结构

2.2.1 CA-YOLOv8

为了实现对煤矿输送带上大块煤的准确、稳定和高效的检测, 在 YOLOv8 算法基础上进行如下改进. 首先在输入端对输入图像进行基于现实尺寸的自适应图像缩放, 其次在 Neck 部分引入 CA, 最后移除小目标检测层与小目标检测头。

(1) 输入端改进

YOLO 算法在输入端采用自适应图片缩放的方式, 将任意分辨率的图像缩放至统一的分辨率大小, 通常是 640×640 像素; 在缩放过程中, 不考虑物体的实际尺寸造成对物体尺寸特征的破坏, 使得同一物体在不同场景中尺度不一致。

输送带工作环境相对单一, 且现实场景下输送带的宽度是确定的, 大块煤的尺寸通常依据与宽度的比例确定. 以输送带尺寸为基础对输送带图像进行缩放, 采取非固定分辨率图像作为输入图像, 可以保证缩放后图像中物体尺度一致. 将输送带宽度 a 映射为 640 像素, 考虑 YOLO 算法下采样倍数为 32, 即输入图像分辨率应为 32 的倍数, 图像高度可能不满足要求, 需在缩放后填充图像高度. 经输入端处理后, 图像分辨率为 $(640, h_{f'})$, 图像高度的计算公式为

$$h_{f'} = 32 \times \left\lceil \frac{b \times 640}{a \times 32} \right\rceil.$$

其中, b 表示图像中输送带长度(根据托辊间距对输送带长度进行估计); $\lceil \cdot \rceil$ 表示向上取整。

根据上述方法对图像 $f'(x, y)$ 进行图像缩放,

得到图像 $f''(x, y)$. 如图 8 所示, 采用双线性插值方法进行图像缩放, 公式为

$$f''(x', y') = f'(x, y) \approx \frac{f'(x_1, y_1)}{(x_2 - x_1)(y_2 - y_1)}(x_2 - x)(y_2 - y) + \frac{f'(x_2, y_1)}{(x_2 - x_1)(y_2 - y_1)}(x - x_1)(y_2 - y) + \frac{f'(x_1, y_2)}{(x_2 - x_1)(y_2 - y_1)}(x_2 - x)(y - y_1) + \frac{f'(x_2, y_2)}{(x_2 - x_1)(y_2 - y_1)}(x - x_1)(y - y_1).$$

其中, (x_1, y_1) , (x_1, y_2) , (x_2, y_1) 和 (x_2, y_2) 表示与点 (x, y) 相邻的 4 个点.

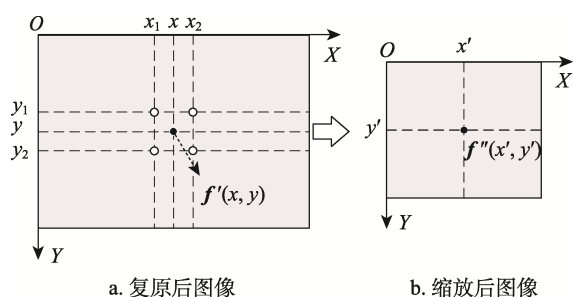


图 8 图像缩放示意图

(2) 嵌入 CA

不同的目标之间或目标与背景之间应存在明显的差异性特征, 这对于有效地识别和定位目标是非常重要的. 目标检测算法通过解析局部特征分布和空间语义关联, 在图像或视频中找出与这些特征匹配的区域.

输送带大块煤检测的情况相对复杂, 主要问题是大块煤与背景中的散煤具有较大的相似性, 在色彩、纹理等视觉特征上大块煤与散煤极为相似; 在大块煤与散煤的尺寸差异上, 虽然大块煤的尺寸通常较大, 但在运输过程中, 大块煤可能会与散煤混合、叠加, 形成复杂的空间结构.

为了使网络更好地将注意力集中到大块煤的特征上, 在 YOLOv8 算法的 Backbone 与 Neck 之间引入 CA, 使得在整合特征时 Neck 能够凸显大块煤的关键特征, 提升网络的检测能力.

(3) 移除多余检测层

在目标检测中, 物体的尺寸通常根据其在图像中的分辨率进行分类: 小于 32×32 像素的物体被视为小目标, 大于 96×96 像素的物体视为大目

标, 而介于两者之间的被视为中目标. 为了确保能有效地检测这 3 种不同尺度的目标, YOLO 算法在 Neck 部分通过 FPN+PAN 结构实现了多尺度特征的融合, 并在 Head 部分为不同尺度的目标设计了 3 个独立的检测头.

在实际应用中, 对于尺寸与输送带宽度比例大于 80% 的大块煤需要停机进行人工破碎. 然而在数据标注过程中, 对于比例大于 15% 且明显区别于散煤的煤块均进行标注, 并以此进行验证. 采用本文方法在输入端进行基于现实尺寸的自适应图像缩放后, 所有大块煤目标尺寸均满足大于 96 像素, 不存在小目标与中目标.

为了更有效地提取图像位置信息, 选择保留中目标检测层与大目标检测层, 移除小目标检测层与小目标检测头, 并通过 FPN+PAN 结构进行特征融合, 确保对大块煤的定位精度. 最后通过大目标检测头与中目标检测头进行预测输出.

2.2.2 检测模型参数及模型训练

CA-YOLOv8 结构如图 9 所示. 将复原后的图像作为检测模型的输入, 在输入端进行基于现实尺寸的自适应缩放; 经 Backbone 部分提取特征, CA 凸显大块煤的特征; Neck 部分进行多尺度特征融合; 最终在 Head 部分完成识别与定位, 将大块煤检测结果进行输出. CA-YOLOv8 的详细参数如表 1 所示.

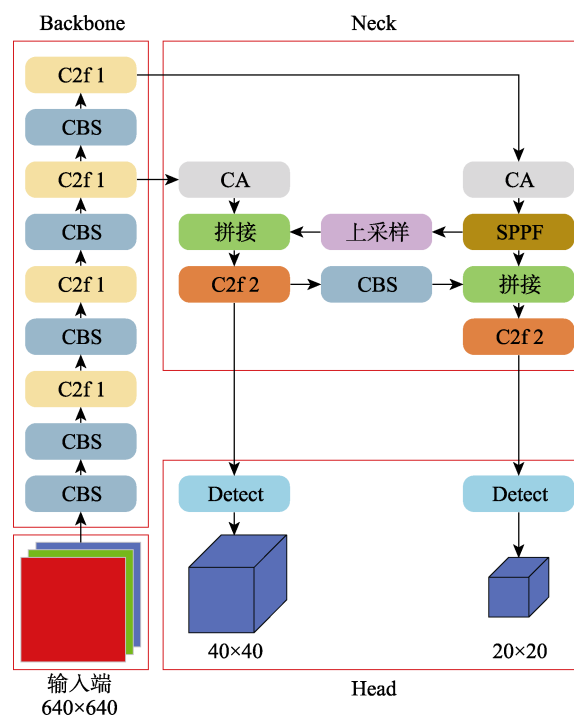


图 9 CA-YOLOv8 结构

表1 CA-YOLOv8的参数

层数	层名称	输入层	通道数	层数	层名称	输入层	通道数
0	CBS	输入端	64	10	CA	9	1 024
1	CBS	0	128	11	SPPF	10	1 024
2	C2f 1	1	128	12	上采样	11	
3	CBS	2	256	13	拼接	7, 12	
4	C2f 1	3	256	14	C2f 2	13	512
5	CBS	4	512	15	CBS	14	512
6	C2f 1	5	512	16	拼接	11, 15	
7	CA	6	512	17	C2f 2	16	1 024
8	CBS	6	1 024	18	Detect	14, 17	
9	C2f 1	8	1 024				

将历史监控视频逐帧进行标注得到数据集,其中,90%用于训练模型,10%用于模型验证。以交叉熵损失函数作为类别分类损失函数,CIOU(complete-IOU)与DFL(distribution focal loss)作为

边框回归损失函数,完成训练后即可得到大块煤检测模型。

3 实验与结果分析

使用某煤矿输送带监控视频对本文方法进行有效性验证。实验平台为 Windows 10 操作系统,11th Gen Intel Core i7-11800H@2.30GHz CPU, NVIDIA GeForce GTX 3060 GPU, PyTorch 深度学习框架。

3.1 实验数据

本文实验数据来自某煤矿 4 条不同的输送带的监控视频图像,其工作环境与图像特点各不相同,如表 2 所示。分别对这些图像逐帧进行标注,自行制作成 4 个数据集 $D_1 \sim D_4$,其图像特点和基本参数如表 3 所示。

表2 不同的输送带的监控视频图像复原前后对比







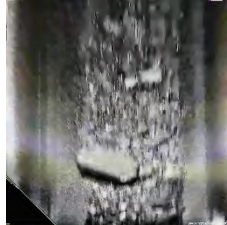

图像	数据集			
	D_1	D_2	D_3	D_4
原图				
复原图				

表3 实验数据集基本参数

数据集	特点	训练集/幅	验证集/幅
D_1	较昏暗、较模糊	1 137	125
D_2	光亮、模糊	1 359	148
D_3	光照不均、较模糊		126
D_4	光亮、较模糊		176

表4 各数据集图像复原相关参数

数据集	c	σ_s	σ_r	d	z	s	t
D_1	5	2.5	30	6	8	8	2.5
D_2	5	2.5	30	16	10	8	2.0
D_3	5	2.5	30	5	5	8	3.0
D_4	5	2.5	30	8	6	8	2.5

3.2 方法验证

3.2.1 图像复原效果分析

图像复原过程中,双边滤波参数参考文献[30],设置卷积核大小 $c=5$,空间邻近度标准差 $\sigma_s=c/2$,颜色相似度标准差 $\sigma_r=30$;维纳滤波参数通过逐步调节对比效果进行选择;对于限制对比度自适应直方图均衡化参数,分块数量选择经验值 $s=8$,逐步调节阈值系数 t 并通过对比效果选择,相关参数设置如表 4 所示。

输送带监控图像复原前后对比如表 2 所示。可以看出,与复原前的图像相比,经过复原处理后的图像清晰度明显提高,煤块的轮廓更加清楚可辨。

为了客观评价本文方法的有效性,选取 Brenner 函数、Tenengrad 函数和 Entropy 函数作为评价图像清晰度的客观指标,它们的值越大,表示图像的清晰度越高。对于分辨率为 (N_x, N_y) 的图像 $f(x, y)$, Brenner 函数表达式为

$$D_B(f) = \sum_{x=1}^{N_x-2} \sum_{y=1}^{N_y-2} (f(x+2, y) - f(x, y))^2.$$

Tenengrad 函数表达式为

$$D_T(f) = \sum_{x=2}^{N_x-1} \sum_{y=2}^{N_y-1} (G(x, y))^2.$$

其中, $G(x, y)$ 表示像素点 (x, y) 处的梯度, 公式为

$$G(x, y) = \sqrt{G_x^2(x, y) + G_y^2(x, y)} = \sqrt{(f(x, y) * g_x)^2 + (f(x, y) * g_y)^2}.$$

其中, $G_x(x, y)$ 和 $G_y(x, y)$ 表示像素点水平方向和垂直方向的梯度值; g_x 和 g_y 表示 Sobel 算子横向模板和纵向模板, 公式分别为

$$g_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad g_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}.$$

表 5 不同评价函数对图像清晰度评价结果

数据集	Brenner		Tenengrad		Entropy	
	原图	复原图	原图	复原图	原图	复原图
D_1	5 403 246	60 561 197	229 025 120	1 576 703 998	4.52	4.92
D_2	10 653 721	68 564 861	580 913 378	2 414 119 094	5.10	5.14
D_3	22 455 246	122 276 843	1 031 775 384	4 153 976 214	5.28	5.30
D_4	25 770 291	63 265 439	708 854 458	1 944 385 806	4.87	4.91

训练完成后, 使用相同环境下的验证集测试训练得到的检测模型. 选用精确率和召回率这 2 个指标衡量检测模型的性能. 实验结果如表 6 所示. 可以看出, 使用相同环境的训练集和验证集($D_1 \rightarrow D_1$, $D_2 \rightarrow D_2$)时, 本文方法表现出了良好的性能, 表明该方法能够实现高效且准确的大块煤检测.

表 6 检测结果对比

检测任务	精确率	召回率
$D_1 \rightarrow D_1$	95.6	89.0
$D_2 \rightarrow D_2$	95.8	90.4

进一步, 验证模型在不同环境条件时的泛化能力, 结果如表 7 所示. 使用 D_1 数据集训练得到的检测模型在 $D_2 \sim D_4$ 数据集上进行验证, 分别用 $D_1 \rightarrow D_2$, $D_1 \rightarrow D_3$ 和 $D_1 \rightarrow D_4$ 表示; 使用 D_2 数据集训练得到的检测模型在 D_1 , D_3 和 D_4 数据集上进行验证, 用 $D_2 \rightarrow D_1$, $D_2 \rightarrow D_3$ 和 $D_2 \rightarrow D_4$ 表示. 如图 10 所示为检测模型在不同环境条件下的检测效果. 可以看出, 无论环境条件是否相同, 图像中的大块煤都被

Entropy 函数表达式为

$$D_E(f) = -\sum_{i=0}^{L-1} p_i \ln(p_i).$$

其中, L 表示灰度级数(通常取为 $L = 256$), p_i 表示图像中灰度值为 i 的像素出现的概率.

图像复原前后输送带区域图像清晰度评价结果如表 5 所示. 可以看出, 3 个函数下, 复原后的图像的各项评价指标均超越原图, 说明本文方法能有效地提升图像的清晰度.

3.2.2 检测效果分析

为验证本文方法的有效性, 分别使用 D_1 和 D_2 的训练集训练检测模型. 在模型训练过程中, 参数设置如下: 初始学习率设置为 0.01; 权重衰减设置为 0.000 5, 以防止过拟合; 动量因子设置为 0.937, 避免陷入局部最优解或跳过最优解; 类别损失系数设置为 0.5; IOU 损失系数设置为 7.5; DFL 损失系数设置为 1.5; 批大小设置为 16; 训练轮数为 100.

表 7 泛化验证结果对比

检测任务	精确率	召回率
$D_1 \rightarrow D_2$	87.9	64.4
$D_1 \rightarrow D_3$	91.7	68.7
$D_1 \rightarrow D_4$	90.6	78.1
$D_2 \rightarrow D_1$	85.4	62.2
$D_2 \rightarrow D_3$	82.6	74.7
$D_2 \rightarrow D_4$	85.5	79.0

准确地标识和定位.

对比表 7 与表 6 可以看出, 在不同环境条件的验证过程中, 本文方法的检测精确率和召回率有所下降, 但整体上仍然保持较高的检测性能. 实验结果表明, 本文方法能够学习到有效的、可在不同环境条件下迁移的特征, 在不同环境条件下具备良好的泛化能力, 能够在多种场景下进行有效的大块煤检测.

对比 D_1 在其他环境条件时的检测性能(如 $D_1 \rightarrow D_3$), 与 D_2 在其他环境条件时的检测性能(如 $D_2 \rightarrow D_3$)可以看出, D_1 模型的精确率高于 D_2 模型的精

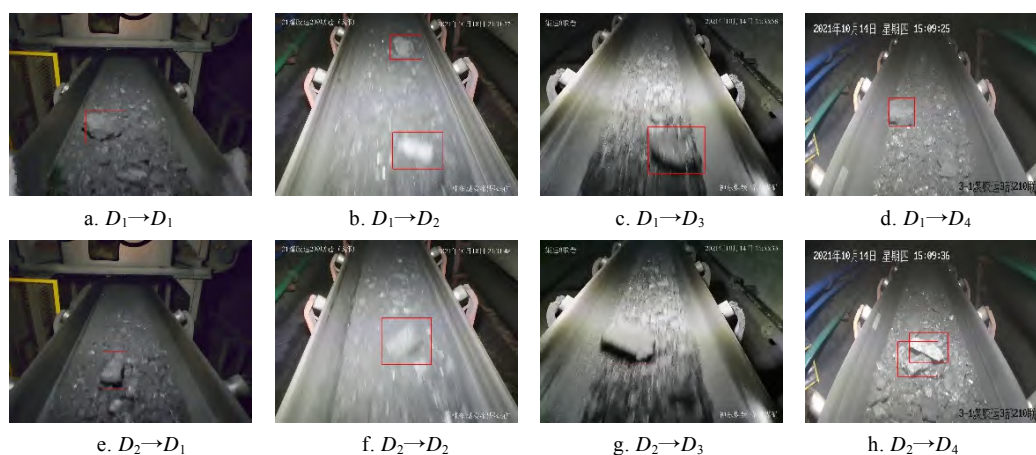


图 10 模型检测结果对比

确率, 这种差距是由于 D_1 和 D_2 在图像质量、背景噪声等特性上存在差异, 影响了模型的泛化性能。

3.3 消融实验

通过消融实验验证本文方法各部分的作用。因为本文方法的输入端改进是基于复原后图像才可以进行的, 所以分为图像复原+输入端改进、嵌入 CA 和移除多余检测层 3 部分进行消融实验, 共设计 5 组实验, 实验中采用相同的数据集和软硬件

设备, 结果如表 8 所示。

本文使用如表 1 所示不同尺度目标检测层的热力图(14 层为大目标检测层, 17 层为中目标检测层)分析模型, 验证嵌入 CA 与移除多余检测层对检测模型的影响。实验过程发现, 中目标检测层和小目标检测层无热力图输出, 如图 11 所示为消融实验组 1、实验组 3 和实验组 4 检测模型的大目标检测层热力图。

表 8 消融实验结果对比

实验组	图像复原+ 输入端改进	嵌入 CA	移除多余 检测层	% %							
				$D_1 \rightarrow D_1$		$D_1 \rightarrow (D_2, D_3, D_4)$		$D_2 \rightarrow D_2$		$D_2 \rightarrow (D_1, D_3, D_4)$	
				精确率	召回率	精确率	召回率	精确率	召回率	精确率	召回率
1				90.1	88.4	73.0	42.4	94.9	91.9	66.1	57.5
2	√			91.1	89.0	87.5	67.2	93.3	92.1	80.5	75.5
3		√		91.1	87.8	80.3	51.1	95.2	90.1	78.1	52.8
4			√	91.3	85.4	81.2	48.0	95.3	91.1	69.8	55.1
5	√	√	√	95.6	89.0	90.0	71.0	95.8	90.4	84.6	72.8

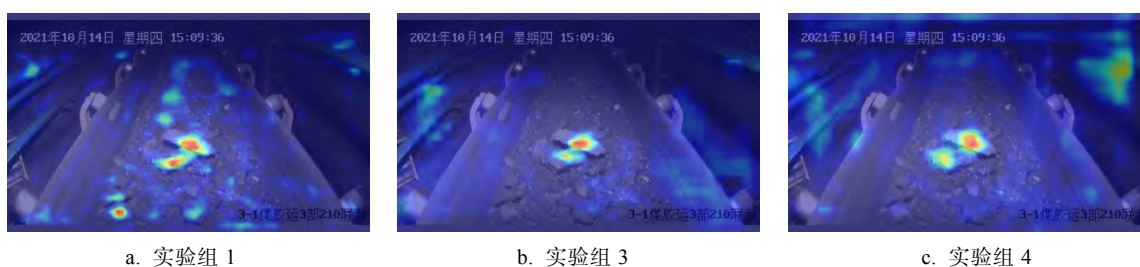


图 11 大目标检测层热力图

从表 8 可以看出, 图像复原与输入端改进统一了煤块的尺度, 相同环境条件验证时检测精度与原模型基本相当, 但不同环境条件验证时的精确率与召回率有大幅的提升。实验结果表明, 对图像进行预处理和尺度统一化, 有助于提高模型在复杂环境下的泛化能力。

结合表 8 和图 11 可以看出, 通过嵌入 CA 增强检测模型对大块煤特征的提取能力, 提升对大块煤检测的精确率, 证明了注意力机制在目标检测任务中的重要性, 尤其是在复杂和多变的环境中; 但同时也会丢失一些其他的特征, 进而导致召回率的降低, 出现漏检的情况。

从图 11 可以看出, 移除多余检测层后, 模型在图像中的关注区域变得更加集中和精确, 提升了大块煤检测的精确率; 但同时会忽略一些细节信息, 因此对召回率的提升效果不大, 甚至出现下降的情况。

3.4 对比分析

将本文方法与未作改进的 YOLOv8、文献[7]和文献[9]方法进行对比实验, 进一步验证本文方法的有效性。实验中使用相同的软件和硬件环境, 结果如表 9 所示。可以看出, 在相同环境条件下,

本文方法的召回率略低于其他方法。这是因为虽然其通过嵌入 CA、移除多余检测层使检测模型更关注于大块煤的关键特征, 从而在不同环境条件下都表现出较高的泛化能力; 但这种设计牺牲了对大块煤细节信息的特征提取能力, 导致召回率有所下降。尽管如此, 本文方法在相同环境条件下的精确率, 以及在不同环境条件下的精确率和召回率, 都明显高于其他 3 种方法, 证明了其在大块煤检测任务上的优越性。

表 9 不同方法对比实验结果对比

%

方法	$D_1 \rightarrow D_1$		$D_1 \rightarrow (D_2, D_3, D_4)$		$D_2 \rightarrow D_2$		$D_2 \rightarrow (D_1, D_3, D_4)$	
	精确率	召回率	精确率	召回率	精确率	召回率	精确率	召回率
YOLOv8	90.1	88.4	73.0	42.4	94.9	91.9	66.1	57.5
文献[7]	92.0	91.0	84.1	57.2	92.6	91.6	75.9	64.8
文献[9]	94.3	90.6	80.5	58.4	94.1	91.1	82.4	63.6
本文	95.6	89.0	90.0	71.0	95.8	90.4	84.6	72.8

实验结果表明, 本文通过图像复原技术显著地提升了图像清晰度; 构建的 CA-YOLOv8 模型不仅在相同的环境条件下展现出了出色的检测性能, 而且在不同的环境条件下也表现出了很好的泛化能力。但是, 本文方法在召回率上仍有提升空间, 未来将进一步研究如何在保持高精确率的同时提高召回率, 从而提升模型在实际应用中的综合性能。

4 结 语

本文针对煤矿输送带图像的特点, 应用图像复原技术以提升图像清晰度, 为后续的煤块检测提供了更清晰、更准确的图像输入。通过分析大块煤的特征构建 CA-YOLOv8 模型, 在 YOLOv8 算法的基础上进行改进, 在保证检测性能的同时, 有效地提升了检测模型在不同环境下的泛化能力。实验结果表明, 本文方法不仅适用于输送带大块煤检测, 对于煤矿场景及其他恶劣场景中的目标检测也具有一定的借鉴意义。然而本文方法对大块煤细节信息的特征提取能力不足, 存在较大提升空间。未来, 将进一步改进模型, 强化局部细节表征, 提升模型在复杂工业环境下检测结果的完整性与可靠性。

参考文献(References):

[1] Wang Y, Guo W, Zhao S F, *et al.* A big coal block alarm detection method for scraper conveyor based on YOLO-BS[J]. *Sensors*, 2022, 22(23): Article No.9052

[2] Wang Y B, Wang Y J, Dang L F. Video detection of foreign objects on the surface of belt conveyor underground coal mine based on improved SSD[J]. *Journal of Ambient Intelligence and Humanized Computing*, 2023, 14(5): 5507-5516

[3] Wang Yan, Guo Xiaoqiang, Liu Xinhua. Design of visual detection system for large foreign body in belt conveyor[J]. *Mechanical Science and Technology for Aerospace Engineering*, 2021, 40(12): 1939-1943(in Chinese)
(王燕, 郭潇楠, 刘新华. 带式输送机大块异物视觉检测系统设计[J]. *机械科学与技术*, 2021, 40(12): 1939-1943)

[4] Zhang Jianli, Sun Shenshen, Qin Shuqi. Ore image segmentation based on optimal threshold segmentation based on genetic algorithm[J]. *Science Technology and Engineering*, 2019, 19(7): 105-109(in Chinese)
(张建立, 孙深深, 秦书棋. 基于遗传算法最佳阈值分割的矿石图像分割[J]. *科学技术与工程*, 2019, 19(7): 105-109)

[5] Zhang Z L, Yang J G, Ding L H, *et al.* Estimation of coal particle size distribution by image segmentation[J]. *International Journal of Mining Science and Technology*, 2012, 22(5): 739-744

[6] Yang Jianhui, Huang Ziyang, Wang Mei, *et al.* Recognition of unwanted objects in coal flow based on gray pyramid convolution model of machine vision[J]. *Coal Science and Technology*, 2022, 50(11): 194-201(in Chinese)
(杨建辉, 黄子洋, 汪梅, 等. 机器视觉灰度化金字塔卷积模型的煤流异物识别[J]. *煤炭科学技术*, 2022, 50(11): 194-201)

[7] Hao Shuai, Zhang Xu, Ma Xu, *et al.* Foreign object detection in coal mine conveyor belt based on CBAM-YOLOv5[J]. *Journal of China Coal Society*, 2022, 47(11): 4147-4156(in Chinese)
(郝帅, 张旭, 马旭, 等. 基于 CBAM-YOLOv5 的煤矿输送带异物检测[J]. *煤炭学报*, 2022, 47(11): 4147-4156)

[8] Xiao D, Kang Z, Yu H, *et al.* Research on belt foreign body detection method based on deep learning[J]. *Transactions of the*

- Institute of Measurement and Control, 2022, 44(15): 2919-2927
- [9] Song Q J, Liu Z J, Jiang H Y. Coal gangue detection method based on improved YOLOv5[C] //Proceedings of the 3rd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering. Los Alamitos: IEEE Computer Society Press, 2022: 466-469
- [10] Hou Q B, Zhou D Q, Feng J S. Coordinate attention for efficient mobile network design[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2021: 13708-13717
- [11] Redmon J, Divvala S, Girshick R, *et al.* You only look once: unified, real-time object detection[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2016: 779-788
- [12] Jiang P Y, Ergu D, Liu F Y, *et al.* A review of Yolo algorithm developments[J]. *Procedia Computer Science*, 2022, 199: 1066-1073
- [13] Fan Jingjing, Xue Haowei, Wu Xinhong, *et al.* Gesture recognition algorithm introducing ghost feature mapping and channel attention mechanism[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2022, 34(3): 403-414(in Chinese)
(范晶晶, 薛皓玮, 吴欣鸿, 等. 引入重影特征映射和通道注意力机制的手势识别算法[J]. *计算机辅助设计与图形学学报*, 2022, 34(3): 403-414)
- [14] Jiang Xingqun, Liu Bo, Song Li, *et al.* Surface damage detection and recognition of wind turbine blade based on improved YOLO-v3[J]. *Acta Energaie Solaris Sinica*, 2023, 44(3): 212-217(in Chinese)
(蒋兴群, 刘波, 宋力, 等. 基于改进 YOLO-v3 的风力机叶片表面损伤检测识别[J]. *太阳能学报*, 2023, 44(3): 212-217)
- [15] Chen Jia, Ding Qi'an, Liu Longshen, *et al.* Early detection of broilers respiratory diseases based on YOLO v5 and short time tracking[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(1): 271-279(in Chinese)
(陈佳, 丁奇安, 刘龙申, 等. 基于 YOLO v5 与短时跟踪的鸡只呼吸道疾病早期检测[J]. *农业机械学报*, 2023, 54(1): 271-279)
- [16] Tang Xinyao, Song Huansheng, Wang Wei, *et al.* 3D vehicle information recognition algorithm of monocular camera based on self-calibration in traffic scene[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2020, 32(8): 1305-1314 (in Chinese)
(唐心瑶, 宋焕生, 王伟, 等. 单目交通场景下基于自标定的车辆三维信息识别算法[J]. *计算机辅助设计与图形学学报*, 2020, 32(8): 1305-1314)
- [17] Wang C Y, Bochkovskiy A, Liao H Y M. Scaled-YOLOv4: scaling cross stage partial network[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2021: 13024-13033
- [18] Zhang H Y, Cisse M, Dauphin Y N, *et al.* Mixup: beyond empirical risk minimization[OL]. [2023-07-24]. <https://arxiv.org/abs/1710.09412>
- [19] Wang C Y, Liao H Y M, Wu Y H, *et al.* CSPNet: a new backbone that can enhance learning capability of CNN[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Los Alamitos: IEEE Computer Society Press, 2020: 1571-1580
- [20] He K M, Zhang X Y, Ren S Q, *et al.* Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904-1916
- [21] Huang S Z, Li X Y, Jiang Z Q, *et al.* Hyper feature fusion pyramid network for object detection[C] //Proceedings of the IEEE International Conference on Multimedia & Expo Workshops. Los Alamitos: IEEE Computer Society Press, 2018: 1-6
- [22] Liu S, Qi L, Qin H F, *et al.* Path aggregation network for instance segmentation[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 8759-8768
- [23] Yu J H, Jiang Y N, Wang Z Y, *et al.* UnitBox: an advanced object detection network[C] //Proceedings of the 24th ACM International Conference on Multimedia. New York: ACM Press, 2016: 516-520
- [24] Neubeck A, van Gool L. Efficient non-maximum suppression[C] //Proceedings of the 18th International Conference on Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2006: 850-855
- [25] Jiang Zetao, Xiao Yun, Zhang Shaoqin, *et al.* Low-illumination object detection method based on Dark-YOLO[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2023, 35(3): 441-451(in Chinese)
(江泽涛, 肖芸, 张少钦, 等. 基于 Dark-YOLO 的低照度目标检测方法[J]. *计算机辅助设计与图形学学报*, 2023, 35(3): 441-451)
- [26] Ye X, He Z H, Heng W, *et al.* Toward understanding the effectiveness of attention mechanism[J]. *AIP Advances*, 2023, 13(3): 035019
- [27] Guo M H, Xu T X, Liu J J, *et al.* Attention mechanisms in computer vision: a survey[J]. *Computational Visual Media*, 2022, 8(3): 331-368
- [28] Zhou Dengke, Yang Ying, Zhu Jie, *et al.* Tilt correction method of pointer meter based on deep learning[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2020, 32(12): 1976-1984(in Chinese)
(周登科, 杨颖, 朱杰, 等. 基于深度学习的指针式仪表倾斜校正方法[J]. *计算机辅助设计与图形学学报*, 2020, 32(12): 1976-1984)
- [29] Feng Xiongbo, Chen Xi, Min Huina, *et al.* DR image enhancement of aero-engine guide vane based on improved CLAHE[J]. *Journal of Aerospace Power*, 2022, 37(7): 1425-1436(in Chinese)
(冯雄博, 陈曦, 闵慧娜, 等. 基于改进 CLAHE 的航空发动机导向叶片 DR 图像增强[J]. *航空动力学报*, 2022, 37(7): 1425-1436)
- [30] Ai Jiaqiu, Wang Feifan, Yang Xingming, *et al.* SAR image speckle noise suppression algorithm based on background homogeneity and bilateral filtering[J]. *National Remote Sensing Bulletin*, 2021, 25(5): 1071-1084(in Chinese)
(艾加秋, 王非凡, 杨兴明, 等. 基于背景均匀性双边滤波的 SAR 图像斑点噪声抑制算法[J]. *遥感学报*, 2021, 25(5): 1071-1084)