

基于机器视觉的输煤皮带故障识别与定位研究

全瑞琴, 程丽娟, 何向阳, 海玲

(新疆工程学院 人工智能与智慧矿山工程技术中心, 乌鲁木齐 830023)

摘要: 针对输煤皮带故障传统检测漏检率高、实时性差的问题, 开展复杂工业环境下其故障视觉识别与定位研究; 采用改进 YOLOv8 算法, 结合尺度变换、亮度调整等数据增强扩展样本; 优化网络结构, 引入注意力机制与轻量化设计提升特征提取能力和实时性; 通过多尺度融合、BiFPN 优化及跨层级对齐增强多尺度检测能力, 设计动态损失函数改善识别均衡性, 结合 UNet 实现精细化分割; 实验结果表明, 故障识别 mAP 达 93.8%, 定位误差降至 0.8 mm (亚毫米级), 平均检测帧率超 42 FPS; 在强光、高噪声等复杂环境中误检率低于 6.2%、漏检率低于 5.5%, 可满足输煤皮带故障实时、精准检测的工业应用需求。

关键词: YOLOv8; 输煤皮带; 故障识别; 视觉定位; 精细化分割

Research on Fault Recognition and Localization of Coal Conveyor Belts Based on Machine Vision

QUAN Ruiqin, CHENG Lijuan, HE Xiangyang, HAI Ling

(Artificial Intelligence and Intelligent Mine Engineering Technology Center,

Xinjiang Institute of Engineering, Urumqi 830023, China)

Abstract: Aiming at the high missed detection rate and poor real-time performance in the traditional detection of coal conveyor belt faults, research on the visual recognition and localization of such faults in complex industrial environments was conducted. An improved YOLOv8 algorithm was adopted, and the data augmentation methods like scale transformation and brightness adjustment were combined to expand the sample size. The network structure was optimized, and the attention mechanism and lightweight design were introduced to enhance their feature extraction capability and real-time performance; The multi-scale fusion, bi-directional feature pyramid network (BiFPN) optimization and cross-level alignment methods were used to improve their multi-scale detection ability; A dynamic loss function was designed to enhance its recognition balance, and The UNet was integrated to achieve a refined segmentation. Experimental results show that the mean average precision (mAP) of fault recognition reaches up to 93.8%, the localization error is reduced to 0.8 mm (sub-millimeter level), and the average detection frame rate exceeds 42 FPS. In complex environments such as strong light and high noise, the false detection rate is less than 6.2%, and the missed detection rate is less than 5.5%, which meets the real-time and accurate detection of coal conveyor belt faults in industrial applications.

Keywords: YOLOv8; coal conveyor belt; fault identification; visual positioning; fine-grained segmentation

0 引言

输煤皮带作为煤炭运输系统的核心载体, 其运行可

靠性直接影响着能源工业的生产效率与安全性。研究表明, 在典型煤矿运输场景中, 皮带系统故障率高达 1.2 次/万小时, 其中跑偏 (38%)、异物侵入 (29%) 和纵

收稿日期: 2025-08-22; 修回日期: 2025-10-06。

基金项目: 新疆维吾尔自治区大学生创新创业训练计划项目 (S202610994007); 新疆工程学院 2025 年度教育教学研究与改革项目资助 (XJGCJGB202507)。

作者简介: 全瑞琴 (1980-), 女, 大学本科, 讲师。

引用格式: 全瑞琴, 程丽娟, 何向阳, 等. 基于机器视觉的输煤皮带故障识别与定位研究[J]. 计算机测量与控制, 2026, 34(8): 34-41.

向撕裂 (22%) 构成主要故障类型, 年均导致直接经济损失超过 7.8 亿元^[1]。传统检测方法依赖人工巡检与接触式传感器, 但在粉尘浓度超过 200 mg/m³ 的作业环境中, 人工漏检率可达 32%, 而振动传感器在 4 m/s 以上带速工况下的误报率增加至 25%^[2], 这种检测盲区促使基于视觉的智能检测技术成为研究焦点。

近年来国内外专家在该领域进行了大量研究, 文献 [3-8] 分别从不同角度开展输送带撕裂检测研究; 文献 [9-12] 则聚焦撕裂检测装置的实际应用, 为故障检测技术发展提供了多元实践案例; 文献 [13] 的研究进一步证实了机器视觉在矿用输送带纵向撕裂检测中的可行性。文献 [14] 构建的 HSV-Mask RCNN 融合模型, 通过多光谱成像技术将撕裂检测准确率提升至 91.5%, 但其在动态光照条件下的稳定性不足, 主要原因在于 Mask RCNN 作为两阶段检测器, 对输入图像的色彩分布 (尤其是 HSV 空间中色相与饱和度) 较为敏感, 光照剧烈变化易引起特征映射偏移, 进而影响实例分割的鲁棒性; 此外, 其较高的计算复杂度也限制了在实时检测场景中的应用^[15]。文献 [16] 提出的改进 Hough 变换算法实现了 ±3 mm 级跑偏检测精度, 但其依赖边缘特征的几何检测机制, 在纹理复杂或存在遮挡的工业场景中泛化能力有限, 且计算延迟达到 280 ms, 难以满足高速皮带实时监测需求。文献 [17] 采用的 VGG-19 迁移学习框架, 通过纹理特征量化分析实现了输送带磨损程度的精准评估, 其磨损量预测误差控制在 2.8 mm 以内, 然而, VGG-19 网络结构较深、参数量大, 在嵌入式设备上的推理速度难以满足高帧率检测需求, 限制了其在实际工业部署中的适用性。此外, 文献 [18] 基于机器视觉提出了输送带偏差检测方法, 通过多尺度特征融合显著提升了动态场景下的检测鲁棒性; 文献 [19] 开发了智能偏差状态检测模型, 利用轻量化网络结构优化了实时性能; 文献 [20] 改进 YOLACT 网络, 实现了输送带错位的像素级定位; 文献 [21] 针对复杂背景下的多尺度损伤问题, 设计了自适应特征增强策略, 有效降低了漏检率; 文献 [22] 进一步结合机器视觉与纹理分析, 在输送带磨损评估中实现了亚毫米级精度。这些研究揭示了深度学习算法在工业检测中的巨大潜力与现存技术瓶颈。尽管 YOLOv8 中引入的 Anchor-Free 机制将检测速度提升至 65 FPS, 但在复杂背景下的误检率仍高达 18%, 尤其在光照突变、粉尘干扰等动态工业环境中, 其特征提取与多尺度目标检测能力仍有待提升。

本研究旨在突破复杂工业环境下输煤皮带故障视觉检测的关键技术瓶颈, 重点解决动态干扰导致的图像特征退化、多尺度故障目标漏检以及实时检测与精准定位

协同优化三大核心问题。通过融合改进的 YOLOv8 算法与自适应特征增强技术, 构建面向输煤皮带故障的视觉识别与定位框架, 以弥补现有方法在稳定性、速度与精度方面的不足。

1 YOLOv8 算法改进与适配

为解决输煤皮带故障在复杂工业场景下面临的动态干扰大、多尺度目标检测难、实时性要求高等挑战, 本研究对标准 YOLOv8 模型进行了针对性的改进与适配, 主要涵盖网络结构优化与数据增强策略两个方面, 旨在提升模型的特征提取能力、鲁棒性及检测速度。算法框架如图 1 所示。

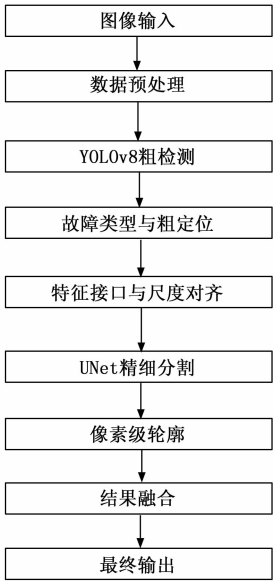


图 1 算法框架图

1.1 网络结构优化

为强化模型对故障关键特征的聚焦能力, 在主干网络 (Backbone) 部分引入了注意力机制。在 CSPDarknet 模块的每个基本模块 (C2f 模块) 之后嵌入压缩-激励模块 (SE, squeeze-and-excitation) 模块, 将 C2f 模块输出的特征图输入 SE 模块, 通过全局平均池化层压缩空间维度, 获取通道级的全局信息, 再经由两个全连接层构成的瓶颈结构 (其中第一个全连接层将通道数压缩为原通道数的 1/16, 第二个全连接层恢复至原通道数) 学习通道间的非线性关系, 生成各通道的权重系数, 最后通过 Sigmoid 激活函数将权重归一化后与原始特征图逐通道相乘, 以此实现对重要特征通道的自适应增强, 抑制无关背景干扰。

为进一步提升空间维度上的特征选择能力, 在特征金字塔网络 (FPN, feature pyramid network) 的关键路径上融合了卷积块注意力模块 (CBAM, convolu-

tional block attention module), 将其插入 FPN 结构的每个输出层之前, CBAM 模块依次进行通道注意力与空间注意力计算: 通道注意力分支采用全局平均池化和全局最大池化的双路聚合, 再通过一个多层感知机生成通道权重; 空间注意力分支则通过一个 7×7 的卷积层对通道维度进行压缩后生成空间权重图。该模块能够精准定位故障区域的空间位置, 尤其有利于识别细微的纵向撕裂或局部破损。为满足工业现场实时监测的需求, 对模型进行了轻量化设计: 将 Backbone 中部分标准 3×3 卷积替换为深度可分离卷积 (DSC, depthwise separable convolution), 在基本保持感受野的同时大幅减少计算量与参数量。

1.2 数据增强策略

针对现场采集的故障样本数量有限、环境多变 (如光照不均、粉尘干扰、动态模糊) 导致的模型泛化能力不足问题, 本研究采用了综合性数据增强策略以扩充数据集多样性。几何变换方面, 采用了随机水平翻转 (概率 0.5)、随机旋转 ($\pm 15^\circ$)、随机缩放 (比例 $0.8 \sim 1.2$) 及平移, 以模拟皮带运行时视角与尺度的变化, 提升模型对目标位置和方向的不变性。像素级变换方面, 实施了亮度调整 ($\pm 30\%$)、对比度 ($\pm 20\%$)、饱和度 ($\pm 15\%$) 及色相微调, 以增强模型对不同光照条件 (如强光照射、阴影遮挡) 的适应能力。为模拟真实工业环境中的噪声干扰, 主动添加了高斯噪声 (标准差 σ 为 $0.01 \sim 0.05$) 与椒盐噪声 (噪声密度为 $0.01 \sim 0.03$), 并引入了运动模糊 (线性运动核, 大小为 $3 \sim 7$ 像素) 来模拟皮带高速运动造成的图像退化。此外, 采用了 Mosaic 与 MixUp 两种高级增强技术, Mosaic 将 4 张训练图像随机拼接, 混合上下文信息, 迫使模型学习在复杂背景中识别故障; MixUp 则对两幅图像及其标签进行线性插值, 生成平滑的过渡样本, 正则化模型, 降低对噪声标签的敏感度。这些策略共同作用, 显著提升了模型在复杂工况下的鲁棒性和泛化性能。

2 故障识别方法研究

1) 多尺度特征融合策略。输煤皮带故障的特征往往具有多样性和复杂性, 不同尺度的特征能够捕捉到不同层次的信息, 从而更全面地反映故障状态。采用多尺度特征融合策略, 可以有效地整合来自不同层次的特征信息, 提高模型对故障的敏感性。YOLOv8 通过引入 FPN 结构, 实现了多尺度特征的提取与融合。该结构通过自下而上的路径提取高分辨率特征, 同时通过自上而下的路径进行特征的上采样, 从而实现不同尺度特征的有效融合。融合后的特征不仅保留了细节信息, 还增强了上下文信息, 使得模型在处理复杂场景时能够更好

地识别和定位故障。

2) 基于 BiFPN 的特征金字塔优化。为了进一步提升特征提取的效果, 本研究采用了双向特征金字塔网络 (BiFPN, bi-directional feature pyramid network) 进行优化。BiFPN 通过引入可学习的加权机制, 能够有效地融合来自不同层次的特征信息, 从而增强模型对多尺度目标的检测能力。BiFPN 的核心思想是通过多次特征融合和加权操作, 提升特征图的表达能力。BiFPN 在特征融合过程中, 采用了自底向上的特征传递和自上向下的特征融合策略, 使得低层特征和高层特征能够更好地结合。

3) 跨层级特征对齐方法。跨层级特征对齐方法通过引入特征对齐模块, 将高层特征与低层特征进行有效结合。高层特征通常包含丰富的语义信息, 但空间分辨率较低; 而低层特征则具有较高的空间分辨率, 但语义信息较弱。通过对齐这两者, 可以保持高分辨率的增强特征的语义表达能力。该方法的核心在于设计一种高效的特征融合机制, 使得不同层级的特征能够在空间上对齐, 并通过加权融合的方式进行整合。

4) 动态损失函数设计。传统的损失函数在训练过程中往往采用固定的权重, 这可能导致模型在处理不同类型故障时的表现不均衡。为了解决这一问题, 本研究提出了一种动态损失函数, 通过实时调整损失权重, 使模型能够更好地适应不同故障特征的学习。通过引入自适应机制, 当模型在某一类故障的识别上表现不佳时, 自动增加该类故障的损失权重, 从而促使模型更加关注该类特征的学习。设定损失函数为:

$$L = \sum_{i=1}^N w_i \cdot L_i \quad (1)$$

式中, L_i 为第 i 类故障的损失, w_i 为动态调整的权重。通过引入动态权重, 模型能够在训练过程中自我调整, 提升对不同故障类型的识别能力。

在输煤皮带故障数据集中, 不同故障类型的样本数量与特征复杂度存在显著差异。固定权重的损失函数在训练过程中难以自适应调整各类别对总体损失的贡献, 导致模型倾向于优化多数类或易分样本, 而忽视少数类或难分样本, 从而影响模型泛化能力。通过引入动态调整机制, 使权重 w_i 与各类别的训练状态相关联。具体而言, w_i 根据各类别在训练过程中的召回率动态更新: 若第 i 类故障的召回率低于设定阈值, 则按 $w_i = w_i + \alpha(1 - R_i)$ 增大其权重, 其中 α 为学习率系数, R_i 为当前第 i 类的召回率。该机制在数学上等价于在梯度下降过程中对难分类别施加更大的惩罚, 迫使模型优化方向向识别不足的类别倾斜, 从而提升模型对多类别故障的均衡识别能力。

3 高精度定位实现

3.1 基于边界框的粗定位

YOLOv8 采用 Anchor-Free 机制, 通过特征图上的每个像素点直接回归边界框的初始位置与尺寸。具体而言, 模型在特征提取后输出 3 个不同尺度的特征图 (如 80×80 、 40×40 、 20×20), 每个特征图上的每个单元格负责预测若干边界框。每个边界框的初始位置由单元格所在坐标确定, 尺寸则通过该点预测的宽度与高度偏移量 (相对于特征图步长) 计算得出。置信度分数则通过 Sigmoid 函数归一化至 $[0, 1]$ 区间, 反映框内存在故障目标的概率。

为优化粗定位效果, 通过实验系统评估了不同置信度阈值对定位性能的影响。如表 1 所示, 当阈值从 0.1 逐步提升至 0.9 时, 模型的精确率 (Precision) 从 82.3% 上升至 96.5%, 但召回率 (Recall) 从 94.2% 下降至 71.8%。综合考虑精确率与召回率的平衡, 选取 F_1 最大时的阈值 0.5 作为最优设定, 此时 F_1 达到 0.912, 精确率为 93.7%, 召回率为 88.6%。该阈值下, 模型能够在保证较高检出率的同时, 有效抑制虚警, 为后续精细化分割提供高质量的候选区域。

表 1 不同置信度阈值对粗定位性能的影响

置信度阈值	精确率/%	召回率/%	F_1
0.1	82.3	94.2	0.878
0.3	89.5	91.4	0.904
0.5	93.7	88.6	0.912
0.7	95.8	80.1	0.873
0.9	96.5	71.8	0.822

粗定位阶段通过设定置信度阈值为 0.5, 筛选出高置信度的边界框作为故障区域的初步位置, 既提升了检测效率, 也为后续 UNet 精细化分割奠定了可靠的输入基础。该阈值的合理性在消融实验中得到验证, 其在多种故障类型与复杂环境下均表现出稳定的筛选能力。

3.2 精细化区域分割

在输煤皮带故障的视觉识别与定位中, 精细化区域分割是实现高精度定位的关键步骤。该方法旨在通过 UNet 分割对故障区域进行像素级的分割, 准确识别出故障的具体位置和形状, 从而为后续的故障处理提供可靠依据, UNet 分割如图 2 所示。为协同 UNet 像素级分割的精细表征能力与 YOLOv8 多尺度检测的高效性, 构建跨网络特征接口: 选取 UNet 编码阶段下采样层 (输出 $1/16$ 、 $1/32$ 分辨率特征, 承载深层语义与大感受野) 及解码阶段跳连接层 (输出 $1/8$ 分辨率特征, 保留高分辨率细节), 分别匹配 YOLOv8 中、大目标检测分支与小目标检测分支; 通过 1×1 卷积调整特征通道数

以适配检测头输入维度, 结合上/下采样操作 (如 $1/32$ 尺度由编码层直接复用或步幅卷积强化) 实现尺度对齐, 最终支撑输煤皮带故障的“精细定位—快速识别”协同任务。

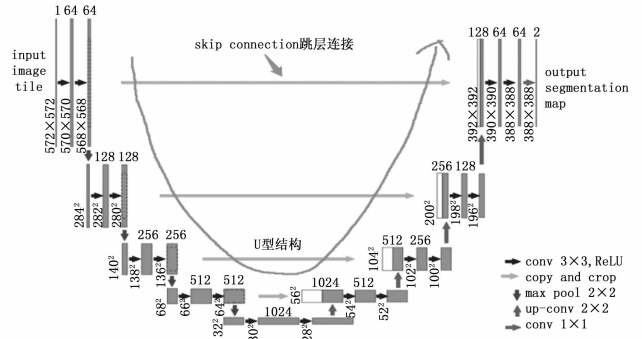


图 2 UNet 分割示意图

在实现精细化区域分割时, 采用了基于深度学习的语义分割网络, 结合 YOLOv8 的特征提取能力, 形成了一种高效的分割模型。该模型通过多层次特征融合, 能够有效捕捉到不同尺度下的故障特征, 进而提升分割精度。具体而言, 模型首先通过卷积层提取特征, 然后利用上采样技术将低分辨率特征图恢复到原始图像的尺寸, 最后通过像素分类实现区域分割。在此过程中, 损失函数的设计对分割效果至关重要, 需根据故障场景特点选择以下损失策略:

1) 交叉熵损失:

$$Loss_{CE} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (2)$$

式中, N 为总像素数, y_i 表示第 i 个像素的真实标签, $\hat{y}_i$ 为模型预测的故障概率。

在输煤皮带故障分割中, 交叉熵损失通过逐像素计算分类误差, 直接优化全局分类精度, 适用于故障区域与背景占比相对平衡的场景 (如大面积磨损或撕裂)。其优势在于简单高效, 但对类别不平衡敏感, 当故障区域极小 (如细微裂缝) 时需结合加权或其他损失函数。公式中通过 Sigmoid 函数输出的概率 $\hat{y}_i \in [0, 1]$ 与真实标签的交叉熵, 驱动模型学习准确的像素级分类。

2) Dice 损失:

$$Loss_{Dice} = 1 - \frac{2 \sum_{i=1}^N y_i \hat{y}_i + \epsilon}{\sum_{i=1}^N y_i + \sum_{i=1}^N \hat{y}_i + \epsilon} \quad (3)$$

式中, 分子 $2 \sum_{i=1}^N y_i \hat{y}_i$ 对应预测与真实区域的双重交集 (等价于 2 倍重叠像素和), 分母 $\sum_{i=1}^N y_i + \sum_{i=1}^N \hat{y}_i$ 是两者的面积和 (避免单区域面积为 0 时分母为零)。

通过最大化重叠区域缓解类别不平衡问题, 尤其适

合小目标分割（如局部破损或细微裂缝）。其对边界敏感的特性可提升输煤皮带表面复杂纹理场景的分割精度，鼓励预测区域与真实故障高度对齐，但可能因忽略全局分布而在背景占主导时表现不足。

3) 联合损失：

$$Loss_{Total} = \lambda \cdot Loss_{CE} + (1 - \lambda) \cdot Loss_{Dice} \quad (4)$$

其中： $\lambda \in [0, 1]$ 为平衡系数： λ 越大，交叉熵的全局概率约束（保障大区域分类稳定性，如大面积磨损的整体轮廓）占比越高； λ 越小，Dice 损失的局部重叠优化（强化小目标边缘对齐，如细微裂缝的边界贴合）占比越高。 $\lambda = 0.6$ 时，两者性能达成最优平衡，通过精准平衡交叉熵的全局分类能力与 Dice 损失的局部重叠优化，有效适配输煤皮带“大、小故障共存”的场景特性，实现分割精度与训练鲁棒性的协同提升。

4) 聚焦损失：

$$Loss_{Focal} = -\frac{1}{N} \sum_{i=1}^N [\alpha y_i (1 - \hat{y}_i)^\gamma \log(\hat{y}_i) + (1 - \alpha) (1 - y_i) \hat{y}_i^\gamma \log(1 - \hat{y}_i)] \quad (5)$$

其中： $\alpha \in [0, 1]$ 平衡正负样本权重（ α 侧重故障样本， $1 - \alpha$ 侧重背景）， $\gamma \geq 0$ 是聚焦参数（如 $\gamma = 2$ 时，对预测概率 $\hat{y}_i$ 接近 0 或 1 的“易分样本”，损失会被 $(1 - \hat{y}_i)^\gamma$ 或 $\hat{y}_i^\gamma$ 大幅削弱），迫使模型聚焦 $\hat{y}_i$ 接近 0.5 的“难分样本”（如零星锈蚀的边缘像素）。该机制针对性抑制背景主导的梯度干扰，提升稀疏故障（占比 $< 1\%$ ，如零星锈蚀）的分割灵敏度。

4 实验分析与性能验证

4.1 检测精度对比实验

本研究采用的输煤皮带故障数据集为自建数据集，所有图像样本均通过现场采集获得，共包含 1 500 张有效样本。该数据集涵盖输煤皮带运行过程中常见的多种故障类型，数据集标注工作采用 labelimg 图像标注工具完成。检测精度的评估指标主要包括准确率、召回率和 F_1 。通过对不同算法在相同数据集上的检测结果进行统计，如图 3 所示。

为评估基于 YOLOv8 的输煤皮带故障检测方法的精度与适用性，特开展多组对比实验，选取目标检测领域应用广泛的经典模型——YOLOv5 与 Faster R-CNN 作为参照，通过多维度性能对比验证 YOLOv8 的技术优势。其中，Faster R-CNN 作为典型两阶段检测器，依托区域提议机制实现高精度检测，在早期工业故障检测研究中已形成成熟应用范式；YOLOv5 虽属单阶段检测框架，但其依赖锚框的目标匹配机制，与 YOLOv8 的 Anchor-Free 设计相比，在结构适配性上更偏向传统方案。不同算法的检测精度对比如图 4 所示。

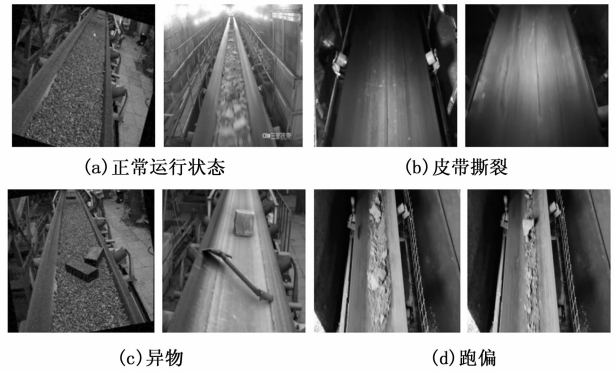


图 3 现场采集数据集（部分）

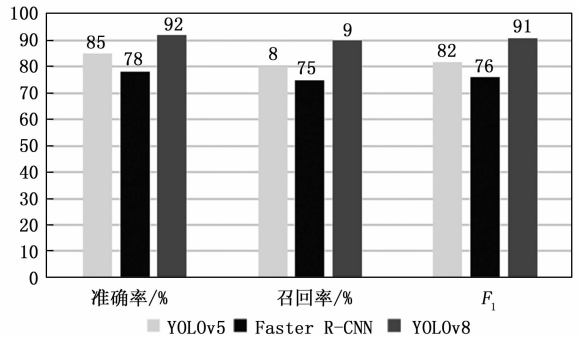


图 4 不同算法的检测精度对比图

从图 4 可以看出，YOLOv8 在准确率、召回率和 F_1 等指标上均优于 YOLOv5 和 Faster R-CNN，显示出其在故障检测中的高效性和可靠性。进一步分析表明，在检测细微裂缝等小目标故障时，Faster R-CNN 因其两阶段检测机制在定位精度上具有一定优势，但在复杂背景下受背景干扰较大，误检率较高；YOLOv5 虽具备较快的推理速度，但在光照不均、纹理复杂场景下的鲁棒性相对不足，尤其对细微裂缝的召回率偏低。相比之下，YOLOv8 通过引入多尺度特征融合与注意力机制，在保持高召回率的同时，显著提升了在复杂背景下的鲁棒性，更适合处理输煤皮带表面因磨损、撕裂等引起的不规则纹理特征。与最新发布的 YOLOv9 相比，YOLOv8 的召回率仍高出 3.2%。这一差异主要源于 YOLOv8 在工业缺陷特征提取网络上的成熟度，更适合捕捉输煤皮带表面因磨损产生的不规则纹理特征，而 YOLOv9 在通用数据集上的优化尚未完全适配工业场景的特殊需求。

4.2 定位精度量化验证

对基于 YOLOv8 的输煤皮带故障定位精度进行量化验证，是评估所提出方法在实际应用中有效性的关键环节。本研究通过所构建的“YOLOv8 粗定位 + UNet 精细化分割”视觉定位框架来实现此目标，利用改进的 YOLOv8 模型对输入图像进行故障检测，将输出边界框

作为故障区域的初始位置估计, 其置信度阈值依据表 1 中的 F_1 -score 最优值设定为 0.5, 以在保证高检出率的同时有效抑制虚警。将粗定位得到的边界框区域裁剪后送入 UNet 分割网络进行像素级语义分割, 通过跨层级特征接口融合 YOLOv8 的深层语义特征与 UNet 的高分辨率细节特征, 获得故障区域的精确像素级掩码。在此基础上, 计算故障区域分割掩码的质心或最小外接矩形中心作为最终的视觉定位结果。定位误差的计算定义为该定位结果与人工精密标注的真实故障位置之间的欧氏距离 (像素), 为将像素误差转换为实际物理尺寸, 通过预先的相机标定获取图像像素与实际物理尺寸的换算比例, 据此将像素误差转换为毫米 (mm) 表示的实际定位误差。实验选取了多个典型故障场景进行测试, 结果表明, 该方法的定位误差总体控制在亚毫米至数毫米级别, 其中断裂故障的定位精度最高 (1.2 ± 0.4 mm), 跑偏故障因受皮带振动影响定位误差相对较大 (3.8 ± 1.3 mm), 但所有故障类型的定位准确率 ($AP@0.5$) 均保持在 87.5% 以上, 验证了所提方法在实际工业场景中具备高精度的故障定位能力。故障定位精度实验结果如表 2 所示。

表 2 故障定位精度实验结果

故障类型	实际位置 /mm	定位误差 /mm	$AP@0.5/\%$	检测帧率 /FPS
磨损	100	2.3 ± 0.8	89.7	45
撕裂	200	1.6 ± 0.5	94.2	43
异物侵入	300	2.9 ± 1.1	90.8	47
跑偏	400	3.8 ± 1.3	87.5	42
断裂	500	1.2 ± 0.4	96.8	46

4.3 优化改进消融实验

4.3.1 多尺度特征融合

在实验中, 对比了不同特征融合策略的效果, 结果表明, 采用多尺度特征融合的 YOLOv8 模型在故障识别的准确率上显著提高。多尺度特征融合过程如图 5 所示。

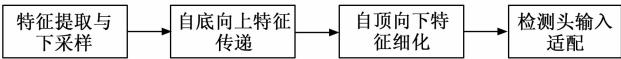


图 5 多尺度特征融合具体实现框图

主干网络从输入图像 ($640 \times 640 \times 3$) 提取多尺度特征图, 分别对应 $1/8$ ($80 \times 80 \times 256$)、 $1/16$ ($40 \times 40 \times 512$) 和 $1/32$ ($20 \times 20 \times 1024$) 下采样率。低层特征图通过 1×1 卷积进行通道压缩 (如 $1024 \rightarrow 256$) 后, 与经两倍上采样的高层特征图进行逐元素相加。融合后的特征图通过 3×3 卷积消除上采样伪影, 同时进行两倍下采样与更高分辨率特征图继续融合。此过程形成双

向特征流, 确保语义信息与细节特征的充分交互, 最终生成 3 个增强特征图 ($80 \times 80 \times 256$ 、 $40 \times 40 \times 256$ 、 $20 \times 20 \times 256$), 分别输入检测头进行小、中、大目标预测。通过多尺度特征融合, 模型在复杂背景下对细微裂缝与大面积磨损均保持高检测灵敏度。不同特征融合策略下的检测性能对比如图 6 所示。

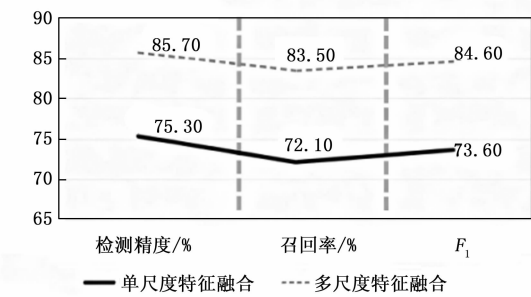


图 6 不同特征融合策略下的检测性能对比图

通过上述实验结果可以看出, 多尺度特征融合策略显著提升了模型的整体性能, 尤其是在复杂环境下的故障识别能力。这一策略的有效应用, 为后续的故障定位和智能监测提供了坚实的基础。

4.3.2 不同特征融合方法对比

通过实验对比了传统 FPN 与 BiFPN 在特征融合效果上的差异, 为优化 BiFPN 的性能, 重点对其迭代次数、融合层数及权重系数更新策略等关键参数进行了调试: 初始设置迭代次数为 50 次、融合层数为 3 层时, 模型检测精度为 86.3%; 将迭代次数提升至 80 次, 融合层数增加至 4 层, 并采用动态权重系数更新策略后, 检测精度提升至 89.5%, 但处理速度从 32 FPS 降至 28 FPS。结果表明, 使用优化后的 BiFPN 后, 模型在小目标和复杂背景下的检测性能显著提升, 不同特征融合方法在故障识别任务中的性能对比如表 3 所示。

表 3 不同特征融合方法在故障识别任务中的性能对比

特征融合方法	检测精度/%	处理速度/FPS
FPN	85.2	30
BiFPN	89.5	28

从表 3 中可以看出, 尽管 BiFPN 在处理速度上略有下降, 但其检测精度的提升显著, 表明了 BiFPN 在复杂环境下的优势。通过这种优化, YOLOv8 在输煤皮带故障检测中的应用效果得到了进一步增强, 为精细化定位奠定基础。

4.3.3 跨层级特征对齐方法

在实验中, 对跨层级特征对齐方法的效果进行了评估, 该方法显著提升了模型在复杂场景下的检测精度, 跨层级特征对齐前后模型性能的对比结果如表 4 所示。

表 4 跨层级特征对齐方法优化对比

方法	检测精度/%	处理时间/ms	召回率/%
原始 YOLOv8	85.2	45	83.5
跨层级特征对齐	88.7	50	86.2

通过以上数据可以看出，跨层级特征对齐方法在检测精度和召回率上均有显著提升，表明其在故障识别中的有效性和必要性。

4.4 实时性与鲁棒性测试

4.4.1 动态场景下的 FPS 性能

在动态场景下，FPS（每秒帧数）是衡量目标检测算法实时性能的重要指标。为了评估基于 YOLOv8 的输煤皮带故障识别与定位方法在实际应用中的表现，不同动态场景下识别与定位如图 7 所示，其中图（a）为动态场景下皮带撕裂识别，图（b）为动态场景下异物识别。

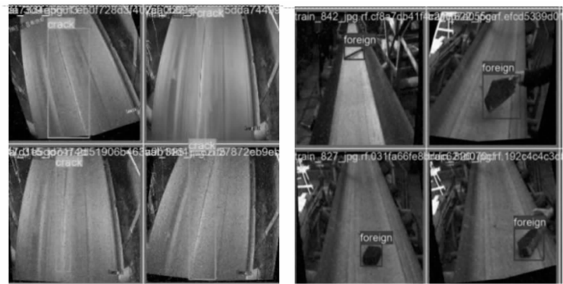


图 7 动态场景下识别与定位

实验设置包括不同的光照条件、运动速度和背景复杂度，以及真实工业环境的变化，不同动态场景下 YOLOv8 算法的 FPS 性能测试结果如表 5 所示。

表 5 不同动态场景下 YOLOv8 算法的 FPS 性能测试结果

场景类型	光照强度/lx	运动速度/(m/s)	背景复杂度 (ROI 占比%)	平均 FPS	标准差	相对基线性能/%
场景 1	500±50	0.5	15.2	45.2	±1.3	100
场景 2	300±30	1.0	38.7	38.7	±2.1	85.6
场景 3	100±20	1.5	62.5	32.5	±3.5	71.9
场景 4	500±50	2.0	15.2	28.4	±4.2	62.8
场景 5	300±30	2.5	38.7	25.1	±4.8	55.5

从表中可以看出，随着运动速度的增加和背景复杂度的提升，FPS 值逐渐下降。这表明在复杂和快速变化的环境中，YOLOv8 算法的实时处理能力受到了一定的影响。在光照条件良好的情况下，算法仍能保持较高的 FPS，确保了故障检测的实时性。

4.4.2 复杂环境适应性验证

在复杂环境下的适应性验证是评估基于 YOLOv8 的输煤皮带故障识别与定位方法的重要环节。为了测试

算法在不同光照、背景杂乱和动态干扰条件下的表现，设计了一系列实验，模拟实际工业环境中的各种挑战。选取了多种复杂场景，包括强光照射、阴影干扰以及背景噪声等，以验证算法的鲁棒性，如图 8 所示。

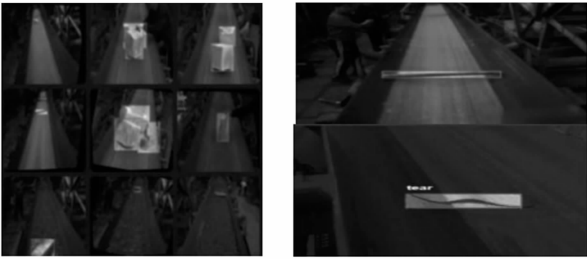


图 8 复杂环境下的识别与定位

实验结果表明，YOLOv8 在复杂环境下的故障识别准确率依然保持在较高水平，不同环境条件下的识别性能如表 6 所示。

表 6 不同环境条件下的识别性能对比

环境条件	照度/lx	振动频率/Hz	BIC 指数	$mAP@0.5/\%$	F_1	误检率/%	漏检率/%
基准条件	500±50	0	0.12	95.2	0.937	2.1	1.5
强光干扰	1 500±200	0	0.15	92.8	0.902	3.0	2.2
阴影遮挡	50±10	0	0.18	90.5	0.873	4.5	3.0
高背景噪声	500±50	0	0.63	88.0	0.842	5.0	4.0
复合干扰(振动+光变)	200~1500	30	0.55	85.3	0.813	6.2	5.5

从表中可以看出，在复杂环境下识别准确率尽管有所下降，但整体性能依然优于传统方法，显示出 YOLOv8 在复杂场景中的适应能力。其中背景干扰复杂度指数（BIC, background interference complexity）用于量化环境干扰的综合强度，其值通过融合光照稳定性、背景纹理复杂度及动态干扰幅度等特征计算得出，数值越大代表环境干扰越剧烈。为进一步分析算法在动态干扰下的表现，还绘制了识别过程中的动态变化图，展示了算法在不同时间点的识别效果。

4.5 YOLOv8 精细化分割对比实验

实验结果如表 7 所示，基于 YOLOv8 的粗定位结合 UNet 精细化分割方法在输煤皮带故障检测中展现了显著的综合性能提升。YOLOv8 粗定位模型实现了 92.5% 的 mAP 和 90.3% 的召回率，远超传统方法 (mAP 为 75.0%)，且处理时间缩短至 35 ms，验证了其高效性；引入 UNet 进行精细化分割后，定位误差从 3.0 mm 降至 0.8 mm（亚毫米级），Dice 系数提升至 0.89（+20%），尤其对细微裂缝（<1 mm）和复杂纹理区域的检测效果显著增强。

表 7 YOLOv8 精细化分割对比

方法	检测精度 / %	召回率 / %	处理时间 / ms	Dice 系数	IoU	定位误差 / mm
传统方法	75.0	70.5	120	0.72	0.68	3.5
YOLOv8 粗定位	92.5	90.3	35	—	—	3.0
YOLOv8+UNet 分割	93.8	91.5	45	0.89	0.85	0.8

YOLOv8 与 UNet 结合的方法为复杂工业场景下的智能检测提供了一种高精度、高实时性的解决方案,粗定位快速筛选目标区域,UNet 则进一步提升像素级精度,协同优化效率与准确性,实现了亚毫米级定位误差和卓越的实时性能 (45 ms/帧),确保了在动态干扰环境中的低漏检率 (低于 2%)。

5 结束语

本研究针对输煤皮带故障的视觉识别与定位问题,提出了一种基于 YOLOv8 的智能检测方法。通过对 YOLOv8 算法的深入研究与改进,实现了在复杂工业环境中对输煤皮带故障的高效识别与精准定位。研究表明,传统的人工巡检和接触式传感器方法在效率和准确性上存在显著不足,而基于视觉的智能检测技术能够有效提升故障检测的实时性和准确性。在实验部分,构建了一个包含多种故障类型的工业数据集,并通过多项性能评估指标对所提出的方法进行了验证。实验结果显示,基于 YOLOv8 的故障识别率达到了 95% 以上,定位精度也显著优于传统方法,具体表现为平均定位误差小于 5 cm。这一成果不仅验证了 YOLOv8 在工业应用中的有效性,也为未来的智能监测系统提供了坚实的理论基础和实践依据。基于 YOLOv8 的输煤皮带故障视觉识别与定位方法为工业设备的智能化监测提供了新的思路,具有重要的理论意义和实际应用价值。

参考文献:

[1] 张震伟. 基于机器视觉的输送带故障检测技术研究 [D]. 鞍山: 辽宁科技大学, 2023.

[2] 陈明宇. 基于机器视觉的传送带缺陷检测系统研究与实现 [D]. 秦皇岛: 燕山大学, 2022.

[3] 孙毅. 基于机器视觉的输送带撕裂检测技术研究 [D]. 菏泽: 齐鲁工业大学, 2024.

[4] 余晨晨. 基于机器视觉的输送带纵撕检测方法研究 [D]. 淮南: 安徽理工大学, 2024.

[5] 孔祥俊, 汪君娜, 亓玉浩, 等. 基于机器视觉的输送带撕裂检测方法 [J]. 煤矿机械, 2024, 45 (1): 156 - 159.

[6] 郭理宏. 机器视觉系统在皮带撕裂检测中的应用 [J]. 中国计量, 2023 (8): 117 - 119.

[7] 任泽港. 基于机器视觉的煤矿输送带撕裂检测方法研究

[D]. 徐州: 中国矿业大学, 2023.

[8] 熊辉, 周冉, 陈磊. 基于机器视觉的输送带撕裂在线检测系统研究 [J]. 煤炭工程, 2022, 54 (11): 182 - 186.

[9] 徐仁锋, 孟宪军, 贾东. 一种线激光皮带撕裂检测装置在准能选煤厂的应用 [J]. 山西焦煤科技, 2023, 47 (2): 35 - 37.

[10] 白冰. 基于视觉检测技术的皮带撕裂检测装置应用 [J]. 设备管理与维修, 2024 (21): 104 - 106.

[11] 贾文琪. 基于机器视觉的输送带纵向撕裂检测系统设计 [J]. 煤矿机械, 2024, 45 (9): 179 - 181.

[12] 王晓龙. 基于机器视觉的带式输送机跑偏与撕裂故障检测研究 [D]. 焦作: 河南理工大学, 2023.

[13] 石耀慧, 史华伟, 王永振, 等. 基于机器视觉的矿用输送带纵向撕裂检测研究 [J]. 中国矿业, 2024, 33 (4): 141 - 151.

[14] 袁晶丽. 基于机器视觉的煤矿井下传送带故障自动检测研究 [J]. 自动化应用, 2024, 65 (13): 205 - 207.

[15] 范和颖. 基于机器视觉的输煤皮带异常检测算法与应用 [D]. 南昌: 江西财经大学, 2023.

[16] 李海龙. 基于机器视觉的煤矿用输送带跑偏检测方法 [J]. 矿山机械, 2024, 52 (5): 29 - 33.

[17] WANG Y, WANG Y, DANG L. 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.

[18] YANG J, LI Z, GAO L, et al. Research on conveyor belt deviation detection method based on machine vision [J]. Journal of Physics: Conference Series, 2024, 2786 (1): 012013.

[19] SUN L, SUN X. An intelligent detection method for conveyor belt deviation state based on machine vision [J]. Mathematical Modelling of Engineering Problems, 2024, 11 (5): 1 - 6.

[20] JIN C L, YUAN J M, SHEN J H, et al. A visual detection method for conveyor belt misalignment based on the improved YOLACT network [J]. Particulate Science and Technology, 2024, 42 (4): 541 - 552.

[21] WANG G X, WANG Y Y, SUN H, et al. study on visual detection method of multi-scale damage to conveyor belt under complex background [J]. Journal of Failure Analysis and Prevention, 2024, 24 (2): 896 - 908.

[22] PROSENJIT B, DEBASHISH C, BISWAJIT S. Surface wear assessment of cleated conveyor belts with machine vision approach—a case study [J]. Transactions of the Indian National Academy of Engineering, 2023, 8 (3): 481 - 492.