

中国科学技术大学

专业硕士学位论文



基于光照语义引导的煤矿井下皮带 机异物检测方法研究

作者姓名： 罗里恒

学科专业： 计算机技术

校内导师： 赵云波 教授

实践导师： 李泽瑞 副研究员

完成时间： 二〇二六年五月二十二日

University of Science and Technology of China

A dissertation for master's degree



**Research on Foreign Object
Detection Method for
Underground Coal Mine Belt
Conveyors Based on
Illumination-Semantic Guidance**

Author: Luo Liheng

Speciality: Computer Technology

Supervisor: Prof. Yun-Bo Zhao

Practice supervisor: Assoc. Res. Fellow Zerui Li

Completion date: May 22, 2026

中国科学技术大学学位论文原创性声明

本人声明所呈交的学位论文，是本人在导师指导下进行研究工作所取得的成果。除已特别加以标注和致谢的地方外，论文中不包含任何他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的贡献均已在论文中作了明确的说明。

作者签名： 罗里恒

签字日期： 2026年5月25日

中国科学技术大学学位论文授权使用声明

作为申请学位的条件之一，学位论文著作权拥有者授权中国科学技术大学拥有学位论文的部分使用权，即：学校有权按有关规定向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅，可以将学位论文编入《中国学位论文全文数据库》等有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。本人提交的电子文档的内容和纸质论文的内容相一致。

控阅的学位论文在解除后也遵守此规定。

☒公开 ☐控阅（ 年）

作者签名： 罗里恒

导师签名： 袁波

签字日期： 2026年5月25日

签字日期： 2026年5月25日

摘要

煤矿井下皮带运输系统承担着原煤从采掘工作面输送至地面的核心任务，其安全稳定运行对于保障矿井高效生产和人员安全具有重要意义。皮带表面混入的矸石、锚杆等非煤异物极易导致皮带撕裂等恶性事故，有效检测这些异物并及时预警是保障皮带运输安全的核心任务。基于深度学习的视觉检测技术可以利用部署在皮带上方的工业摄像头实现对异物的实时监测，因此在煤矿智能化建设中逐渐得到应用。然而，该技术在井下实际应用中面临双重挑战：一方面，井下光照条件恶劣导致图像质量低下，严重影响目标特征的有效提取；另一方面，边缘计算资源受限与目标形态复杂之间存在矛盾，传统检测方法难以兼顾实时性与检测精度。因此，迫切需要分别研究侧重图像增强与轻量化检测的两种算法，并开发配套的智能监测系统。

针对上述问题，本文开展了面向煤矿井下皮带机的视觉检测技术研究，提出了图像增强与轻量化检测两种算法改进方法，并基于两者协同优化策略开发了融合上述改进方法的智能监测系统。具体研究工作包括：

(1) 针对井下暗区与强光点共存的复杂光照条件，提出了一种区域自适应低光照图像增强方法。该方法基于 Retinex 理论框架，通过区域分割模块实现暗区与亮区的像素级软划分，结合自适应光照增强器提升暗区亮度，同时利用光效抑制器削弱强光点引起的光斑和光晕伪影。通过设计由区域自适应重建损失、结构与高频细节保持损失、平滑性与正则化损失以及初始去噪一致性损失四部分构成的复合损失函数，实现了亮度增强、结构保持与伪影抑制的多目标平衡。在自建煤矿数据集和公开数据集上的实验表明，该方法有效改善了图像视觉质量，在 BRISQUE、NIQE、PIQE 等指标上均优于现有方法。

(2) 针对边缘计算平台资源受限及细长异物几何建模困难的问题，提出了一种光照语义引导的轻量化异物检测方法。该方法采用 GhostNetV2 构建高效主干网络，大幅降低参数量和计算量；引入动态蛇形卷积增强对锚杆等细长异物的几何建模能力；设计坐标注意力与光照区域语义联合引导的特征重标定机制，将增强模块输出的光照图和区域权重图作为语义先验融入检测网络。实验表明，该方法在保持较高检测精度的同时，参数量和计算量显著降低，检测精度较基线方法有明显提升。

(3) 开发了煤矿井下皮带机异物在线检测系统。系统采用“边缘端+服务端+前端”的三层分布式架构，集成图像增强与异物检测两种核心算法，实现了视频流接入、实时增强检测、告警联动、数据存储与可视化展示等功能。系统采用

时间窗投票的多帧一致性告警策略有效降低了误报率，通过 TensorRT 加速实现了边缘端实时推理。功能测试、性能测试和现场验证表明，系统在复杂井下环境中表现稳定，具有良好的工程应用价值。

关键词：低光照图像增强；目标检测；煤矿安全监测；轻量化网络；注意力机制

ABSTRACT

The belt conveyor system in underground coal mines undertakes the core task of transporting raw coal from the mining face to the surface, and its safe and stable operation is of critical importance for ensuring efficient mine production and personnel safety. Foreign objects such as gangue and rock bolts mixed onto the belt surface can readily cause catastrophic accidents including belt tearing; therefore, the effective detection of such objects together with timely early warning constitutes a central task in safeguarding belt transportation safety. Vision-based detection technology driven by deep learning, which leverages industrial cameras mounted above the belt to achieve real-time monitoring of foreign objects, has gradually been adopted in the intelligent construction of coal mines. However, this technology faces twofold challenges in practical underground deployment: on the one hand, the harsh underground illumination conditions lead to severely degraded image quality, which substantially hinders the effective extraction of target features; on the other hand, an inherent conflict exists between limited edge computing resources and the geometric complexity of target objects, rendering it difficult for conventional detection methods to simultaneously satisfy real-time and accuracy requirements. Consequently, there is a pressing need to investigate two dedicated algorithms focusing respectively on image enhancement and lightweight detection, and to develop a corresponding intelligent monitoring system.

To address the aforementioned issues, this thesis carries out research on vision-based detection technology for belt conveyors in underground coal mines. Two algorithmic improvement methods, namely image enhancement and lightweight detection, are proposed, and an intelligent monitoring system that integrates the two improvements through a collaborative optimization strategy is developed. The specific research contributions are as follows.

(1) To address the complex illumination conditions characterized by the coexistence of dark regions and intense point light sources in underground environments, a region-adaptive low-light image enhancement method is proposed. Built upon the Retinex theoretical framework, the method employs a region segmentation module to achieve pixel-level soft partitioning of dark and bright regions, incorporates an adaptive illumination enhancer to boost the brightness of dark areas, and exploits a light-effect suppressor to attenuate the glare and halo artifacts induced by intense point light sources. By designing a composite loss function consisting of four components—

namely a region-adaptive reconstruction loss, a structural and high-frequency detail preservation loss, a smoothness and regularization loss, and an initial denoising consistency loss—a multi-objective balance among brightness enhancement, structural preservation, and artifact suppression is achieved. Experiments conducted on a self-built coal mine dataset and public benchmarks demonstrate that the proposed method effectively improves the visual quality of images and outperforms existing approaches in terms of BRISQUE, NIQE, and PIQE metrics.

(2) To address the limited resources of edge computing platforms and the difficulty of geometrically modeling slender foreign objects, an illumination-semantics-guided lightweight foreign object detection method is proposed. The method adopts Ghost-NetV2 to construct an efficient backbone network, substantially reducing the number of parameters and computational cost; introduces dynamic snake convolution to enhance the geometric modeling capability for slender targets such as rock bolts; and devises a feature recalibration mechanism jointly guided by coordinate attention and illumination region semantics, which incorporates the illumination map and regional weight maps output by the enhancement module as semantic priors into the detection network. Experimental results demonstrate that the proposed method significantly reduces the number of parameters and computational cost while maintaining high detection accuracy, and yields a noticeable improvement in detection accuracy over baseline methods.

(3) An online foreign object detection system for underground coal mine belt conveyors is developed. The system adopts a three-tier distributed architecture comprising the “edge-side, server-side, and front-end” tiers, and integrates the two core algorithms of image enhancement and foreign object detection, thereby realizing functions such as video stream ingestion, real-time enhancement and detection, alarm linkage, data storage, and visualization. A multi-frame consistency alarm strategy based on temporal-window voting is employed to effectively reduce the false alarm rate, while TensorRT acceleration is leveraged to enable real-time inference on the edge side. Functional testing, performance testing, and on-site verification demonstrate that the system performs stably in complex underground environments and exhibits favorable engineering application value.

KEY WORDS: Low-light image enhancement, Object detection, Coal mine safety monitoring, Lightweight network, Attention mechanism

目录

第 1 章 绪论	1
1.1 研究背景与研究意义	1
1.1.1 煤炭能源的战略地位与运输系统关键性	1
1.1.2 非煤异物对皮带运输安全的危害分析	2
1.1.3 视觉感知面临的技术挑战与本课题研究意义	2
1.2 国内外研究现状	3
1.2.1 低光照图像增强技术研究现状	3
1.2.2 目标检测技术研究现状	5
1.2.3 煤矿井下视觉监测研究现状	6
1.2.4 研究现状总结	7
1.3 本文研究内容与结构安排	8
第 2 章 相关基础知识	10
2.1 煤矿井下视觉成像与光照特性分析	10
2.1.1 煤矿井下成像环境特点	10
2.1.2 煤矿井下光照分布与典型退化问题	11
2.1.3 低光照与强点光对视觉任务的影响	12
2.2 低光照图像增强基础理论	14
2.2.1 低光照图像退化模型	14
2.2.2 Retinex 理论及其数学表达	15
2.2.3 图像质量评价指标	16
2.3 深度学习目标检测基础理论	17
2.3.1 目标检测问题定义	17
2.3.2 YOLO 系列模型基本结构	18
2.3.3 注意力机制与轻量化网络简介	20
2.4 本章小结	23
第 3 章 区域自适应低光照图像增强方法与实验验证	24
3.1 引言	24
3.2 方法总体框架与关键模块设计	24
3.2.1 Retinex 分解与增强策略	24
3.2.2 总体框架概述	25
3.2.3 区域分割模块	26

3.2.4	自适应光照增强器	28
3.2.5	光效抑制器	28
3.3	损失函数与实现细节	30
3.3.1	复合损失函数设计	30
3.3.2	实现细节	33
3.4	实验与结果分析	33
3.4.1	数据集与评价指标	33
3.4.2	对比方法与实验设置	35
3.4.3	定量结果分析	35
3.4.4	可视化分析	36
3.4.5	消融实验	38
3.5	本章小结	41
第 4 章	面向煤矿皮带机的光照语义引导异物检测方法 与实验验证	42
4.1	引言	42
4.2	光照语义的定义与形式化表示	43
4.3	方法总体框架与关键模块设计	44
4.3.1	总体框架设计	44
4.3.2	轻量化主干网络：GhostNetV2	45
4.3.3	细长异物的局部形状建模：动态蛇形卷积	47
4.3.4	空间位置信息与光照区域语义的联合引导	49
4.4	损失函数与训练策略	52
4.4.1	检测损失函数	52
4.4.2	分阶段预训练与联合微调策略	53
4.5	实验设置与结果分析	54
4.5.1	数据集与评价指标	54
4.5.2	对比方法与实现细节	56
4.5.3	定量结果分析	57
4.5.4	消融实验	58
4.5.5	可视化分析	60
4.6	本章小结	60
第 5 章	煤矿井下皮带机异物在线检测系统设计 与实现	62
5.1	系统需求分析	62
5.1.1	系统功能性需求	62
5.1.2	系统非功能性需求	65

5.2 系统设计	65
5.2.1 系统总体设计	65
5.2.2 系统功能模块设计	66
5.3 系统实现	68
5.3.1 系统总体架构实现	68
5.3.2 检测模块实现	71
5.3.3 数据传输模块实现	72
5.3.4 数据存储模块实现	73
5.3.5 系统界面展示	74
5.4 系统测试	75
5.4.1 系统功能性测试	76
5.4.2 系统非功能性测试	77
5.5 本章小结	78
第 6 章 总结与展望	79
6.1 总结	79
6.2 展望	80
参考文献	82
致谢	88
在读期间取得的科研成果	89

插图和附表清单

图 1.1	2024 年中国一次能源消费结构占比统计 ^[1]	1
图 1.2	本文研究内容框架图	9
图 2.1	煤矿井下皮带运输典型成像场景	11
图 2.2	煤矿井下图像典型退化类型: (a) 整体亮度不足; (b) 局部过曝与光晕; (c) 噪声放大; (d) 细节丢失	12
图 2.3	不同光照条件下特征提取效果对比	13
图 2.4	低光照图像退化模型示意图	14
图 2.5	Retinex 分解示意图	16
图 2.6	目标检测任务定义示意图	18
图 2.7	YOLO 系列模型典型架构	19
图 2.8	坐标注意力模块结构图	21
图 2.9	典型轻量化模块结构对比: (a) 标准卷积; (b) 深度可分离卷积; (c) Ghost 模块	22
图 3.1	所提出的区域自适应低光照图像增强框架总体结构, 集成区域分割、自适应光照增强与光效抑制模块	25
图 3.2	区域分割模块 (RSM) 结构	27
图 3.3	自适应光照增强器 (AIE) 结构	28
图 3.4	光效抑制器 (LES) 结构	29
图 3.5	私有井下煤矿数据集典型样例: (a) 整体低光照场景; (b) 强点光场景; (c) 混合光照场景	34
图 3.6	LOL-v2 数据集上的视觉对比	37
图 3.7	LSRW 数据集上的视觉对比	37
图 3.8	井下煤矿数据集上的视觉对比。从左至右依次为: 输入图像、RUAS、SCI、ZeroDCE++、EnlightenGAN、NeRCO、Zero-IG、本章方法的结果, 展示亮度增强与伪影抑制效果	38
图 3.9	井下煤矿数据集上的视觉对比 (补充样本)	38
图 3.10	子网络消融实验。私有数据集上不同模块组合的视觉对比结果	39
图 3.11	损失函数消融实验。私有数据集上不同损失分量的视觉对比结果 ...	40
图 4.1	面向煤矿皮带机异物检测的整体网络架构	44
图 4.2	GhostNetV2 Bottleneck 结构图	46

图 4.3	动态蛇形卷积 (DS Conv) 结构示意图	47
图 4.4	细长异物目标上 DS Conv 采样轨迹可视化: (a) 原始低光照图像; (b) 增强后图像; (c) 标准卷积采样点; (d)DS Conv 自适应采样轨迹	48
图 4.5	坐标注意力 (CA) 模块结构图	50
图 4.6	光照语义重标定 (LSR) 模块结构示意图	51
图 4.7	自建数据集四类异物目标典型样例: (a) 矸石; (b) 锚杆; (c) 槽钢; (d) 木材	55
图 4.8	模块消融实验 mAP 变化柱状图	59
图 4.9	不同光照条件下检测可视化: (a)(b)(c) 低光照场景; (d)(e)(f) 强点光场景。每组包括原图、增强图、检测结果	60
图 5.1	异物检测需求用例图	63
图 5.2	数据传输管理需求用例图	64
图 5.3	异常数据回溯需求用例图	64
图 5.4	系统检测流程图	66
图 5.5	异物检测算法结构图	67
图 5.6	帧传递示意图	68
图 5.7	异物检测系统结构图	71
图 5.8	系统框架图	71
图 5.9	帧队列算法流程图	73
图 5.10	异物检测系统监控界面	74
图 5.11	监控设备管理界面	75
图 5.12	检测结果管理界面	75
表 2.1	不同光照条件对异物检测任务的影响分析	13
表 2.2	图像质量评价指标总结	17
表 2.3	YOLO 系列主要版本演进	20
表 2.4	轻量化主干网络对比	23
表 3.1	多数据集上的图像质量指标定量对比 (私有数据集、LOL-v2-Real、LOL-v2-Syn、LSRW-Huawei、LSRW-Nikon)。无监督方法中最优结果以 粗体 标注, 次优结果以下划线标注	35
表 3.2	私有煤矿数据集上各方法的运行时间、参数量与计算量对比	36
表 3.3	子网络消融实验。私有数据集与公开数据集上不同模块组合的图像质量指标定量对比	39

表 3.4	损失函数消融实验。私有数据集与公开数据集上不同损失分量的图像质量指标定量对比	40
表 4.1	不同主干网络参数量与计算量对比	47
表 4.2	训练超参数配置	54
表 4.3	自建煤矿皮带机异物数据集统计	55
表 4.4	测试集光照工况划分	55
表 4.5	实验实现细节	57
表 4.6	自建煤矿皮带机异物数据集主结果对比	57
表 4.7	不同光照工况下 mAP@0.5 对比	58
表 4.8	模块消融实验	58
表 4.9	训练策略消融实验	59
表 5.1	系统硬件环境配置	69
表 5.2	系统软件环境与技术栈	70
表 5.3	检测模块功能测试	76
表 5.4	数据传输模块功能测试	77
表 5.5	数据存储模块功能测试	77

第1章 绪论

1.1 研究背景与研究意义

1.1.1 煤炭能源的战略地位与运输系统关键性

能源是现代工业社会的血液，是国家经济发展和国防安全的基石。在当前及未来较长的一段时期内，煤炭作为我国主体能源的地位不会发生根本性改变。尽管近年来国家大力推动风能、太阳能等清洁能源的发展，构建多元化的能源供应体系，但受限于“富煤、贫油、少气”的资源禀赋特征，煤炭依然是我国能源安全的“压舱石”和“稳定器”。

根据国家统计局发布的《2024年国民经济和社会发展统计公报》数据显示，2024年我国全年能源消费总量比上年增长4.3%。在我国一次能源消费结构中，虽然清洁能源占比稳步提升，但煤炭消费量占比依然高达53.2%。如图1.1所示，煤炭在能源消费总量中占据主导地位，远超石油（18.0%）、天然气（9.2%）及非化石能源（19.6%）。值得注意的是，2024年非化石能源消费占比首次超过石油，标志着我国能源结构绿色转型取得重要进展。这一数据充分表明，保障煤炭工业的安全、高效、稳定生产，对于维持国民经济的平稳运行、保障国家能源安全战略具有举足轻重的意义^[1]。

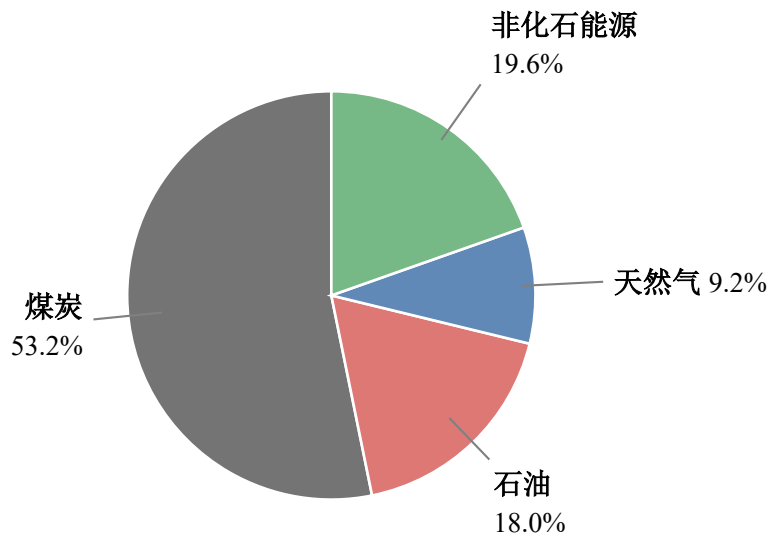


图 1.1 2024 年中国一次能源消费结构占比统计^[1]

在现代化的大型煤矿生产系统中，皮带运输系统被誉为矿井的“咽喉”与“大动脉”，其运行效率直接决定了矿井的生产能力。皮带运输机凭借其长距离、大运量、连续性强、易于实现自动化控制等显著优势，成为了井下原煤从采掘工作面输送至地面的核心装备^[2]。随着煤矿开采深度的不断增加和集约化生产规

模的扩大，皮带运输系统正朝着高带速、大功率、智能化的方向快速发展。然而，高强度的运输作业也使得皮带机长期处于满负荷运行状态，任何微小的故障若不能被及时发现和处理，都可能引发连锁反应，导致全矿停产，甚至造成不可估量的经济损失和人员伤亡。因此，提升皮带运输系统的安全监测水平，保障这条“黑色动脉”的畅通无阻，是实现煤矿智能化建设的关键一环^[3]。

1.1.2 非煤异物对皮带运输安全的危害分析

尽管煤矿生产机械化程度不断提高，但在实际采掘与运输过程中，受井下复杂地质条件、支护工艺以及设备老化等因素影响，各类非煤异物不可避免地会混入煤流中，并随皮带高速运行。这些异物主要包括：综采工作面脱落的大块矸石、断裂的金属锚杆、废弃的木材、铁丝网以及槽钢等。

这些异物对皮带输送系统的安全运行构成了极大的隐患。特别是细长型的金属锚杆或尖锐的矸石，一旦在溜槽、导料槽或托辊架等狭窄空间发生卡阻，极易刺穿高速运行的输送带。随着皮带的持续牵引运动，异物如同利刃般将皮带纵向划开，造成长距离的纵向撕裂。此类事故不仅会导致昂贵的皮带报废，更换皮带所需的漫长停机时间更会造成巨大的间接经济损失。此外，异物堆积还极易引发皮带跑偏、打滑，剧烈的机械摩擦产生的高温甚至可能引燃皮带或周围煤尘，诱发火灾、瓦斯爆炸等特别重大事故，严重威胁井下作业人员的生命安全。

在皮带运输系统引发的各类严重安全事故中，由非煤异物导致的皮带纵向撕裂、卡阻等机械故障是最为常见且破坏性最强的故障形式之一，已成为威胁运输系统安全稳定运行的首要因素^[4]。传统的依靠人工巡检的方式劳动强度大、效率低，且井下环境恶劣，人工巡检存在诸多安全盲区，已难以满足现代化煤矿对于“减人增效、少人则安”的迫切需求。因此，研究一种能够实时、准确、自动地检测皮带表面非煤异物的技术手段，对于预防恶性事故发生、保障煤矿安全生产具有重要的现实意义和巨大的工程应用价值。

1.1.3 视觉感知面临的技术挑战与本课题研究意义

随着计算机视觉与人工智能的飞速发展，以机器视觉为核心的非接触检测方案凭借其丰富的信息获取量、快速的处理响应、可全天候工作等优势，逐渐成为煤矿井下智能监测的首选方案。然而，将通用的视觉检测算法直接应用于煤矿井下这一特殊场景，面临着极为严峻的技术挑战，主要体现在以下三个方面：

第一，极端的低光照成像环境。煤矿井下是完全封闭的无光环境，仅依靠矿灯和局部照明设备提供光源，导致监控图像普遍存在照度极低、对比度严重不足的问题。大量纹理细节淹没在黑暗背景中，使得异物（尤其是深色的矸石或被煤尘覆盖的金属件）难以被有效区分^[5]。这种恶劣的光照条件直接导致了现有检

测算法特征提取能力的急剧下降。

第二，强点光源引起的非均匀光照干扰。为了满足照明需求，井下巷道通常每隔一定距离安装大功率防爆灯。这种点光源照明方式会在图像中形成强烈的非均匀光照分布。强光源直射区域会产生严重的光斑和光晕效应，造成局部过曝，与周围的暗区形成极大的动态范围反差。这种“亮处过曝、暗处死黑”的现象，严重破坏了图像的特征分布，极易导致检测算法出现漏检或误检。

第三，复杂的环境介质干扰与边缘计算资源受限。井下空气中长期漂浮着高浓度的煤尘颗粒和水雾，这些介质会对成像光线产生散射和吸收作用，进一步降低图像的清晰度和信噪比，呈现出类似雾霾天气的退化特征。同时，考虑到数据传输带宽和实时性要求，检测算法通常需部署在井下的边缘计算终端上。这些设备的算力、显存和功耗预算非常有限，难以支撑大规模、高复杂度的深度神经网络模型。

针对上述关键问题，本文围绕“煤矿井下皮带机异物检测”这一核心任务，突破了传统视觉任务中图像增强与目标检测相互割裂的研究范式，从两者协同优化的角度进行系统性探索。本文的研究意义主要体现在以下三个方面：

(1) 理论创新意义。本文提出了一种基于语义引导的协同优化框架，探索了如何在复杂光照条件下实现图像质量的自适应增强。通过在增强网络中引入光效抑制机制，解决了 Retinex 理论在强点光源下易产生伪影的难题；同时，将增强过程中提取的光照分布与空间位置信息作为语义先验融入检测网络，为解决恶劣环境下视觉特征失配问题提供了新的理论视角。

(2) 技术突破意义。针对边缘计算资源受限的约束，本文研究了轻量化检测网络的设计方法。通过引入高效的特征提取模块和动态卷积机制，在大幅降低模型参数量和计算量的同时，显著提升了对锚杆等细长、不规则异物的检测精度，平衡了算法的实时性与准确性，突破了高性能模型难以在井下边缘端落地的技术瓶颈。

(3) 工程应用价值。基于本文研究成果构建的皮带机异物在线检测系统，能够实现对关键异物的全天候实时监测与报警。这不仅能有效预防皮带撕裂等恶性事故，减少非计划停机时间，降低设备维护成本，更能从源头上消除安全隐患，推动煤矿运输系统向无人化、智能化迈进，助力我国煤炭工业的高质量发展。

1.2 国内外研究现状

1.2.1 低光照图像增强技术研究现状

低光照图像增强是计算机视觉领域的基础性研究问题，旨在改善光照不足条件下拍摄图像的可见性、对比度和整体视觉质量^[6]。此类图像广泛存在于夜

间摄影、视频监控、自动驾驶等实际应用场景中，光照不足导致的低对比度、高噪声和细节丢失严重制约目标检测、场景理解等下游任务的性能表现^[7]。

基于 Retinex 理论的图像增强方法为解决低光照问题提供了重要的理论框架。Retinex 理论认为观测图像可以分解为反射分量和光照分量的逐像素乘积，通过估计和调整光照分量实现图像增强^[8]。LIME 方法将光照估计作为独立优化目标，采用边缘保持平滑生成初始光照图，显著提升了局部对比度^[9]。RUAS 方法通过展开优化和协作先验搜索构建轻量化网络，在保持性能的同时实现了实时推理^[10]。URetinex-Net 进一步将 Retinex 分解过程展开为可学习深度网络，通过隐式先验正则化建模显著缓解了传统手工先验适应性不足的问题^[11]。

近年来，深度学习驱动的低光照图像增强技术实现了质的飞跃。在监督学习范式下，Retinex-Net 将经典 Retinex 模型与神经网络相融合，设计了分解子网络与增强子网络分别处理光照估计和图像增强任务^[12]。KinD 方法引入光照引导的反射分量恢复机制，有效抑制了噪声放大问题^[13]。SNR-Aware 方法基于像素级信噪比先验设计空间自适应增强策略，通过 Transformer 与卷积模块的动态融合实现了对极低信噪比区域的鲁棒增强^[14]。IAT 方法仅以约 90K 参数构建了轻量级照度自适应 Transformer，将 ISP 流程解耦为局部与全局分量，在保持极快推理速度的同时显著提升了低光照下的目标检测与语义分割性能^[7]。LLFormer 方法面向超高清低光照场景，提出基于轴向多头自注意力与跨层注意力融合的 Transformer 架构，将自注意力的计算复杂度降至线性^[15]。Retinexformer 方法首次将 Transformer 架构引入 Retinex 框架，通过照度引导的多头自注意力机制建模光照与反射分量之间的跨区域依赖关系^[16]。

在无监督与零参考学习范式下，Zero-DCE 方法开创性地提出了零参考增强框架，通过自适应曲线调整优化图像动态范围，避免了对成对训练数据的依赖^[17]。EnlightenGAN 方法通过全局-局部判别器和自正则化感知损失显著提升了无监督增强质量^[18]。SCI 方法采用权重共享的自校准光照学习模块，在提升曝光稳定性的同时降低了计算开销^[19]。PairLIE 方法仅利用同场景的成对低光照样本即可学习自适应先验，摆脱了对正常光照参考图像的强依赖^[20]。NeRCO 方法通过隐式神经表示统一多源退化因子，并引入视觉-语言预训练模型提供语义级监督，实现了感知友好的协同增强^[21]。Zero-IG 方法提出了光照引导的联合去噪与增强框架，实现了端到端的无监督优化^[22]。

生成式扩散模型为低光照图像增强带来了新的研究范式。Diff-Retinex 方法利用条件扩散模型重新定义增强任务，通过生成式扩散补充 Retinex 分解后丢失的高频信息^[23]。GSAD 方法通过对扩散过程 ODE 轨迹的曲率正则化和不确定性引导约束，在抑制扩散噪声的同时保留了图像的全局结构^[24]。HVI-CIDNet 方法构建了可学习的 Horizontal/Vertical-Intensity 色彩空间，并设计颜色与强度解耦的

双分支网络,从根源上缓解了 sRGB 空间增强过程中的颜色偏移与亮度伪影^[25]。

针对煤矿井下这一特殊工业场景,相关学者也开展了针对性研究。Yang 等人针对井下人员安全监测任务,构建了局部增强模块与基于 Transformer 的全局调整模块相级联的两阶段框架,并引入特征相似性损失约束目标特征的一致性,缓解了增强对下游检测的负面影响^[26]。

尽管上述方法在公开数据集上展现出良好的增强效果,但在煤矿井下等复杂工业场景中仍面临诸多挑战。现有方法普遍难以同时处理暗区增强与强光抑制两个相互冲突的需求,容易产生光晕扩散、过曝伪影和颜色失真等问题;此外,多数深度学习方法计算复杂度较高,难以满足边缘设备的实时部署需求。针对这些局限性,本文提出了一种区域自适应的低光照图像增强框架,通过区域分割模块实现暗区与亮区的差异化处理,结合光效抑制模块有效削弱强光伪影。

1.2.2 目标检测技术研究现状

目标检测是计算机视觉研究体系中的关键问题,其目标在于从静态图像或动态视频序列中实现感兴趣物体的空间定位与类别判定。受益于深度学习技术的蓬勃发展,检测算法在精度与速度两个维度上均实现了大幅提升^[27]。

从算法架构角度,深度学习目标检测方法可分为两阶段检测器和单阶段检测器两大类。以 Faster R-CNN 为代表的两阶段检测器首先通过区域建议网络生成候选框,再对候选区域进行分类和边界框回归,具有较高的检测精度但推理速度相对较慢^[28]。以 YOLO 系列为代表的单阶段检测器将目标检测视为端到端的回归问题,直接从图像预测目标类别和位置,在保持较高精度的同时实现了实时检测^[29]。RetinaNet 方法提出 Focal Loss 解决类别不平衡问题,使单阶段检测器的精度首次超越两阶段方法^[30]。

YOLO 系列算法自 2016 年首次提出以来,经历了从 YOLOv1 到 YOLOv11 的持续演进^[27]。YOLOv4 整合了大量训练技巧和网络改进,在保持高速的同时大幅提升了检测精度^[31]。YOLOv5 在工程实践中得到了广泛应用,成为目标检测领域的重要基线模型。YOLOv7 提出了 E-ELAN 高效聚合网络和复合缩放策略^[32]。YOLOv8 引入了无锚框检测头和改进的 C2f 模块。最新的 YOLOv9 引入了广义高效层聚合网络和可编程梯度信息等创新机制,进一步推动了单阶段检测器的性能边界^[33]。近年来,基于 Transformer 的 DETR 方法通过集合预测方式消除了对非极大值抑制(Non-Maximum Suppression, NMS)的依赖^[34],RT-DETR 方法在保持端到端特性的同时实现了与 YOLO 系列相当的推理速度^[35]。

针对边缘设备部署需求,轻量化目标检测网络成为研究热点。MobileNet 系列通过深度可分离卷积显著降低了计算复杂度^[36]。GhostNet 通过少量真实卷积生成大量冗余特征的方式大幅减少参数量和计算量^[37]。GhostNetV2 引入了解耦

全连接注意力机制,进一步增强了长距离依赖建模能力^[38]。ShuffleNet 利用通道混洗操作实现特征交互^[39]。近期研究中,LE-YOLO 在遥感图像目标检测中展现出良好的轻量化效果^[40],LEAF-YOLO 专门针对无人机图像中的小目标检测进行优化^[41]。

针对细长目标和形变目标的检测,可变形卷积通过学习采样点偏移量自适应调整感受野形状^[42]。动态蛇形卷积将采样网格展平为一维序列并引入可学习偏移,使卷积核能够沿目标轮廓自适应滑动,特别适合于管状、杆状等细长结构的检测^[43]。

尽管目标检测技术发展迅速,但在煤矿井下等低光照、高噪声、光照不均匀的复杂工业场景中,现有方法仍面临检测精度下降、误检漏检率升高等问题。如何在资源受限的边缘计算平台上实现高精度、实时化的异物检测,仍是亟待解决的关键技术难题。

1.2.3 煤矿井下视觉监测研究现状

煤矿井下视觉监测技术是矿山智能化建设的重要组成部分,涵盖皮带运输监测、人员行为识别、设备状态检测等多个应用方向。王国法等^[44]系统阐述了煤矿智能化建设的技术体系,指出基于视频图像识别的皮带机煤量检测、异物识别等技术已趋于成熟,正逐步实现皮带机的无人值守运行。近年来,随着深度学习技术与煤矿场景的深度融合,基于机器视觉的井下安全监测系统逐渐从实验室走向工程应用^[45]。

在皮带机异物检测方面,基于深度学习的方法已成为研究主流。Luo 等^[46]提出了一种基于改进 YOLOv5 的轻量化异物检测方法,通过引入暗通道去雾预处理和无参数注意力机制,在自建数据集上检测精度达 94.9%,模型参数量较 YOLOv5s 降低 43.1%。Fan 等^[47]针对低光照和粉尘雾环境中遮挡目标检测精度低的问题,采用 Soft-NMS 和 SimOTA 标签分配策略优化 YOLOv5s,结合 Group-Taylor 剪枝将模型压缩至 1.20 MB, $mAP_{0.5}$ 达到 91.30%。近期,基于 Transformer 的端到端检测方案也被引入该领域,Zhao 等^[48]提出了 MSC-DETR 算法,在保持实时推理速度的前提下,有效提升了对小目标和细长异物的检测精度。此外,Hong 等^[49]设计了基于改进 DLEA 增强和轻量化 SARC-DETR 模型的异物检测方法,参数量减少 47.01%,推理速度达 97.3 FPS。

在图像增强与检测协同方面,井下低光照环境对检测性能的影响已引起广泛关注。Wang 等^[50]提出了一种面向煤矿视频物联网的多模态图像增强方法,利用 CLIP 进行无监督优化并融合传统 ISP 与 CNN 框架,在 PSNR 指标上相较主流方法提升 2.9%~4.9%。Zhao 等^[51]采用改进 CycleGAN 进行无监督低光照增强,通过多尺度卷积特征增强模块提取不同尺度的特征,有效改善了井下图像质量

并提升了后续异物检测精度。Du 等^[52]提出了一种混合零参考与去雾网络，实现了对井下图像低光照和雾化退化的联合增强处理。然而，现有方法大多将增强与检测视为独立的串联环节，缺乏对两者协同机制的深入设计。Wu 等^[53]尝试利用语义感知知识引导低光照图像增强，为增强与检测的深度融合提供了新的思路。

在轻量化部署方面，考虑到井下边缘计算平台的算力和功耗限制，研究者提出了多种轻量化改进方案。Zhou 等^[45]提出了 UCM-Net 检测网络，并引入自监督预训练策略为模型提供更丰富的特征先验，在算力受限的矿用设备上取得了良好的检测效果。Ling 等^[54]以 YOLOv8n 为基线，采用 StarNet 轻量化主干和 DEConv 检测头，实现参数量降低 80%、计算量减少 60.49% 的同时精度提升 1.9%。

综上，现有煤矿井下视觉监测技术在异物检测、图像增强和轻量化部署等方面已取得重要进展，但仍面临以下挑战：一是图像增强与检测网络之间缺乏有效的信息传导机制，增强结果中蕴含的光照分布、区域语义等先验信息未能被检测网络充分利用；二是对锚杆、钢筋等细长异物的检测召回率仍有待提高；三是在满足边缘端实时性需求的同时保持较高检测精度仍然是一个开放性课题。

1.2.4 研究现状总结

综合上述分析可以看出，低光照图像增强、目标检测和煤矿井下视觉监测三个领域均已取得显著进展，但仍存在若干有待解决的关键问题。

在低光照图像增强方面，现有方法在处理暗区增强与强点光抑制的冲突需求时表现欠佳，难以同时实现亮度提升、伪影抑制和细节保持的多目标优化。特别是在煤矿井下等存在强点光源的特殊场景中，光晕扩散和过曝问题尤为突出。此外，多数深度学习方法计算复杂度较高，难以满足边缘设备的实时部署需求。

在目标检测方面，尽管 YOLO 系列等单阶段检测器已实现较好的精度-速度平衡，但在低光照、高噪声的复杂工业环境中性能仍有较大下降空间。现有轻量化方案在压缩网络规模的同时往往伴随检测精度的损失，如何在边缘计算资源约束下实现高精度检测仍是挑战。此外，针对锚杆等细长异物目标的几何建模能力有待加强。

在煤矿井下视觉监测方面，现有系统大多将图像增强与目标检测视为串联的独立模块，缺乏对两者协同优化机制的深入研究。增强模块输出的中间信息（如光照分布、区域权重等）未能有效利用，检测网络也未能针对增强图像的特性进行针对性设计。

针对上述研究现状与存在问题，本文拟从图像增强与目标检测协同优化的角度，研究面向煤矿井下皮带机的语义引导异物检测方法。通过构建区域自适应低光照图像增强框架解决复杂光照条件下的图像质量提升问题，通过设计轻量

化语义引导检测网络实现增强信息与检测特征的深度融合，为煤矿井下皮带机安全监测提供更加可靠的技术解决方案。

1.3 本文研究内容与结构安排

本文以煤矿井下皮带机异物检测为应用背景，围绕低光照图像增强与语义引导目标检测的协同优化问题展开研究。主要研究内容包括以下几个方面：

第一，区域自适应低光照图像增强方法研究。针对煤矿井下暗区与强光点共存的复杂光照条件，提出基于 Retinex 理论的区域自适应增强框架。该框架通过区域分割模块（RSM）实现暗区与亮区的软划分，通过自适应光照增强器（AIE）提升暗区亮度与对比度，通过光效抑制器（LES）削弱强光点引起的光斑和光晕伪影，并结合轻量化去噪网络（LDN）进行预处理，设计复合损失函数实现亮度增强、结构保持和伪影抑制的多目标平衡。增强模块输出增强图像 I_{out} 、光照图 S_2 以及区域权重图 W_{dark} 和 W_{bright} ，为后续检测提供高质量输入和语义引导信息。

第二，语义引导的轻量化异物检测方法研究。构建融合图像增强信息的轻量化检测网络，在主干网络中采用 GhostNetV2 等高效结构降低计算复杂度，引入动态蛇形卷积（DS Conv）增强对锚杆等细长异物的几何建模能力。在特征融合阶段，将增强模块输出的光照图 S_2 和区域权重图 W_{dark} 、 W_{bright} 映射为语义引导权重，结合坐标注意力（CA）机制对中间特征进行重标定，通过光照语义重标定（LSR）模块实现光照区域语义与空间位置语义的联合引导，提升检测网络在复杂光照条件下的鲁棒性。

第三，煤矿井下皮带机异物在线检测系统设计与实现。基于上述算法研究成果，开发面向实际应用的在线检测系统，采用“边缘端+服务端+前端”的三层分布式架构，实现视频流接入、实时图像增强与检测、告警联动、数据存储与回放、Web 可视化展示以及与煤矿综合自动化平台的接口对接等功能，为煤矿皮带运输安全监测提供完整的技术解决方案。

本文的总体结构安排如下：

第一章为绪论，阐述研究背景与意义，综述低光照图像增强、目标检测技术和煤矿井下视觉监测的国内外研究现状，分析现有研究的不足与本文的研究切入点，介绍本文的研究内容与结构安排。

第二章为相关理论基础，介绍煤矿井下成像特性与光照模型、Retinex 理论与低光照增强基本框架、图像质量评价指标，以及轻量化检测网络、注意力机制、语义引导等相关概念，为后续算法设计提供理论支撑。

第三章为区域自适应低光照图像增强算法研究，详细阐述所提出的增强框

架,包括轻量化去噪预处理(LDN)、区域分割与权重生成(RSM)、自适应光照增强(AIE)、光效抑制(LES)以及复合损失函数设计,并给出在自建煤矿数据集和公开数据集上的实验结果与分析。

第四章为语义引导的轻量化异物检测方法研究,介绍检测网络的整体架构设计,包括 GhostNetV2 轻量化主干、动态蛇形卷积(DS Conv)模块、坐标注意力(CA)与光照语义重标定(LSR)机制,以及检测头与训练策略设计,并通过对比实验和消融实验验证所提方法的有效性。

第五章为煤矿井下皮带机异物在线检测系统设计与实现,从系统需求分析、总体架构设计、关键功能模块实现等方面介绍系统开发过程,包括视频采集与预处理、增强推理、检测推理、告警管理、数据存储与回放、Web 可视化以及平台接口等模块,并结合实验室环境和实际井下工况对系统进行测试验证。

第六章为总结与展望,对全文研究工作进行总结,分析现阶段研究的不足,并对后续研究方向进行展望。

图 1.2展示了本文研究内容的总体框架,直观呈现了图像增强模块与检测模块的协同工作流程及语义引导信息的传递路径。

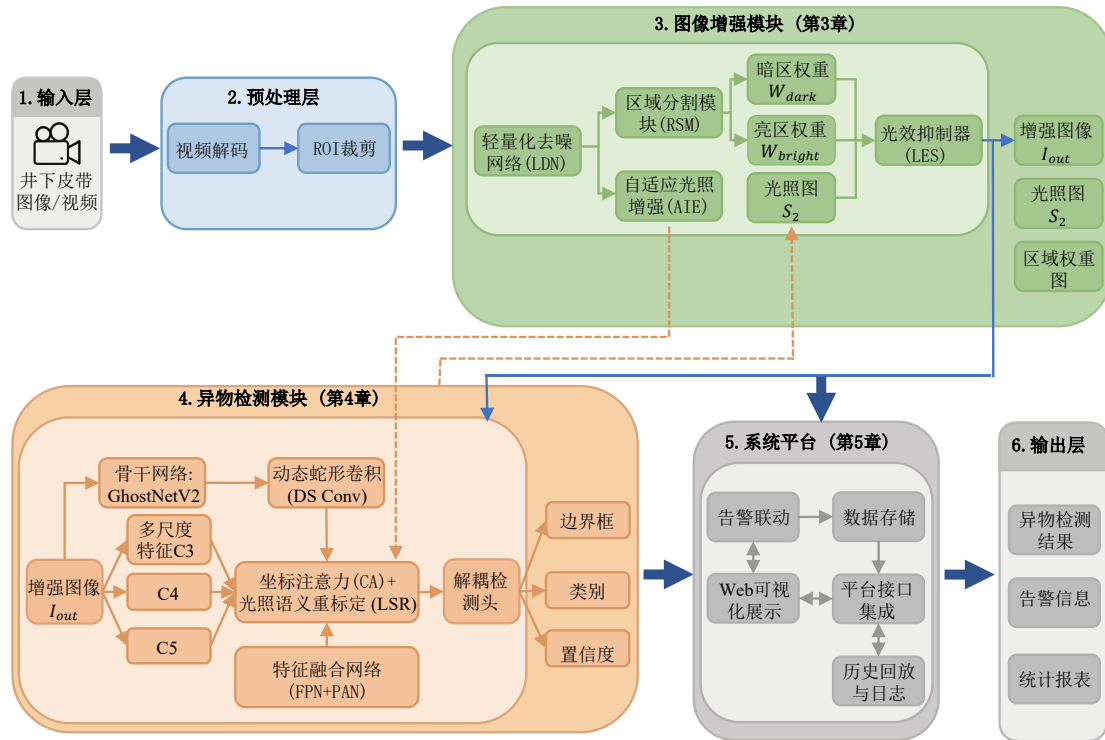


图 1.2 本文研究内容框架图

第2章 相关基础知识

本章系统介绍与煤矿井下皮带机异物检测研究相关的基础理论知识。首先分析煤矿井下特殊的视觉成像环境及其光照特性，揭示低光照与强点光并存条件下图像退化的内在机理；其次阐述低光照图像增强的基础理论，包括退化模型、Retinex 理论框架以及主流的深度学习增强方法；最后介绍深度学习目标检测的相关基础，涵盖检测问题定义、YOLO 系列模型以及注意力机制与轻量化网络设计。上述内容为后续章节的算法设计与系统实现奠定理论基础。

2.1 煤矿井下视觉成像与光照特性分析

2.1.1 煤矿井下成像环境特点

煤矿井下作为一种典型的受限空间工业环境，其成像条件与地面场景存在显著差异。从物理环境角度分析，井下巷道空间狭窄封闭，自然光线无法到达，照明完全依赖人工光源。常见的井下照明设备包括矿用防爆灯、LED 照明灯以及监控摄像机自带的红外补光灯等，这些光源在功率、光谱特性和空间分布上具有明显的局限性^[55]。

井下成像环境的复杂性主要体现在以下几个方面。第一，光照强度普遍不足。受能耗限制和防爆要求约束，井下照明设备的功率通常较低，导致整体环境亮度远低于正常成像所需的照度水平。第二，光照分布极不均匀。照明设备在巷道中呈离散点状布置，形成明显的光照梯度分布，部分区域存在强点光源，而远离光源的区域则几乎完全处于黑暗状态。第三，空气介质干扰严重。井下空气中悬浮着大量煤尘、岩粉等微粒，这些颗粒物对光线产生散射和吸收作用，进一步降低了成像质量。第四，设备工作环境恶劣。高温、高湿、振动等因素导致镜头污染、成像器件性能下降，加剧了图像退化程度。

图 2.1 展示了典型的煤矿井下皮带运输场景。从图中可以观察到，皮带机工作区域呈现出明显的明暗交替特征：照明灯附近区域亮度较高甚至出现局部过曝，而巷道深处及设备阴影区域则几乎不可见。这种极端的光照对比给视觉监测带来了严峻挑战。

从成像设备角度分析，井下监控系统普遍采用工业级防爆摄像机。受成本和防爆等级限制，这些设备的感光元件性能往往低于地面同类产品，在低照度条件下信噪比较低。此外，为满足宽动态范围要求，部分设备采用自动曝光策略，但在光照剧烈变化的场景中，自动曝光算法难以兼顾明暗区域的细节保留，往往导致暗区欠曝或亮区过曝的问题。

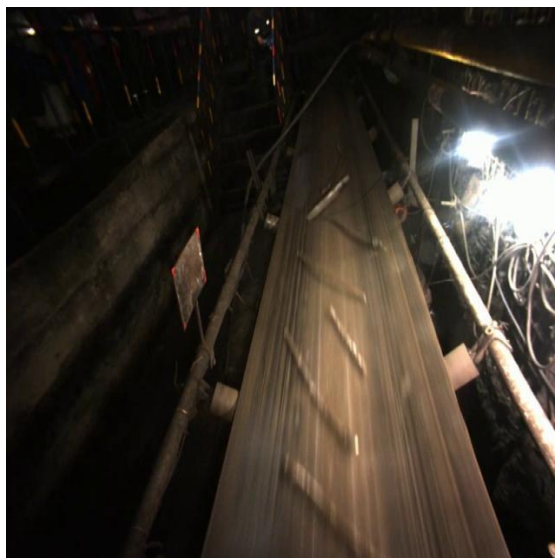


图 2.1 煤矿井下皮带运输典型成像场景

2.1.2 煤矿井下光照分布与典型退化问题

煤矿井下的光照分布呈现出典型的非均匀特征,这种非均匀性源于照明光源的点状分布特性和巷道空间的几何约束。设井下某一点 $\mathbf{p}$ 处的光照强度为 $E(\mathbf{p})$,根据点光源照明模型,该点接收到的光照强度可表示为:

$$E(\mathbf{p}) = \sum_{i=1}^N \frac{I_i \cos \theta_i}{r_i^2} \cdot \eta(\mathbf{p}) \quad (2.1)$$

其中, N 为光源数量, I_i 为第 i 个光源的发光强度, r_i 为该光源到点 $\mathbf{p}$ 的距离, θ_i 为光线入射角, $\eta(\mathbf{p})$ 为考虑空气介质散射衰减的透射系数。

由式 (2.1) 可知,光照强度随距离的平方衰减,这导致井下场景中光照分布呈现显著的空间梯度变化。在光源正下方或近处,照度可达数百勒克斯;而在远离光源的区域,照度可能低至几勒克斯甚至更低。这种量级的差异对成像传感器的动态范围提出了极高要求。

针对井下图像的典型退化问题,可从以下几个维度进行分析:

(1) 整体亮度不足

由于井下光照强度普遍偏低,采集到的图像整体呈现暗淡特征。设原始场景的反射率分布为 $R(\mathbf{p})$,则成像过程中获取的像素值 $I(\mathbf{p})$ 满足:

$$I(\mathbf{p}) = R(\mathbf{p}) \cdot E(\mathbf{p}) + n(\mathbf{p}) \quad (2.2)$$

其中 $n(\mathbf{p})$ 表示成像噪声。当 $E(\mathbf{p})$ 较小时,有效信号被噪声淹没,图像的信噪比显著降低。

(2) 局部过曝与光效伪影

在强点光源附近,过高的入射光强可能超出传感器的线性响应范围,导致像素值饱和,形成过曝区域。此外,光线在灰尘颗粒和镜头表面的散射会产生光晕

(Halo) 效应, 表现为光源周围出现的模糊亮斑。如图 2.2 所示, 光晕效应不仅降低了图像对比度, 还可能掩盖该区域内的目标信息。

(3) 噪声放大与细节丢失

低光照条件下, 为获得可辨识的图像, 相机通常会提高增益或延长曝光时间。增益放大在提升信号强度的同时也同比例放大了噪声, 导致图像呈现明显的颗粒感。长曝光则可能引入运动模糊, 对于高速运行的皮带机场景尤为不利。这些因素共同作用, 导致图像纹理细节严重退化^[56]。

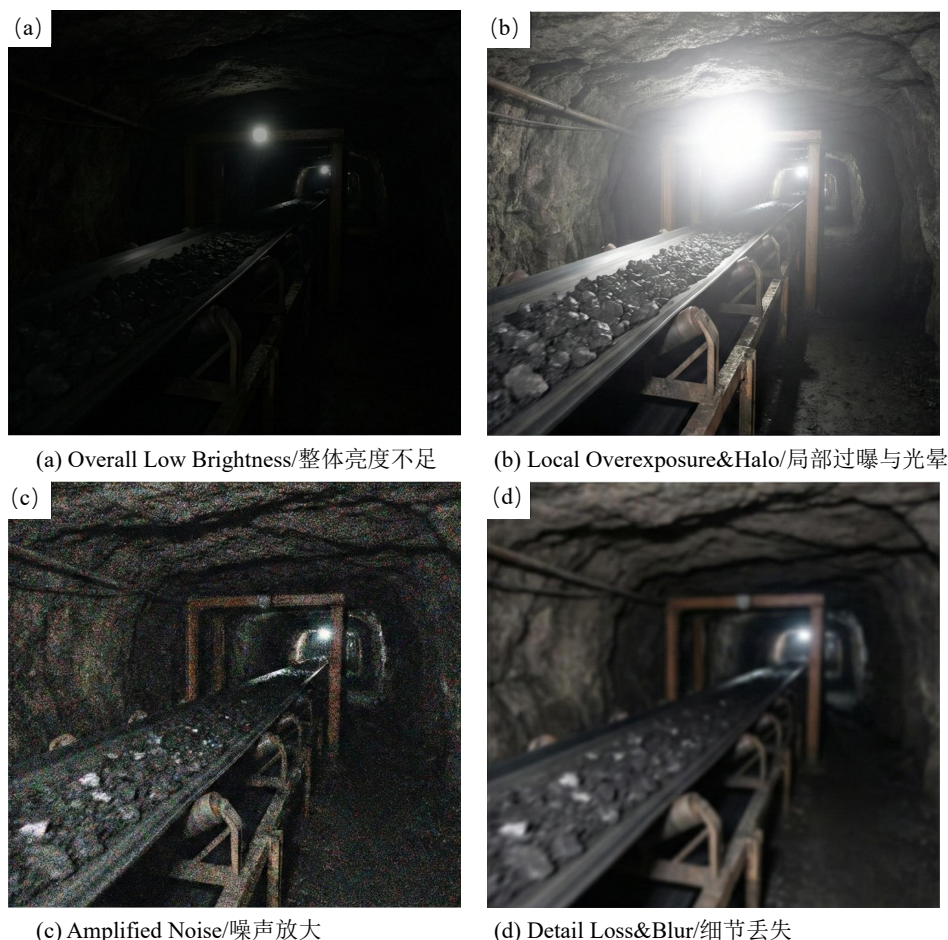


图 2.2 煤矿井下图像典型退化类型: (a) 整体亮度不足; (b) 局部过曝与光晕; (c) 噪声放大; (d) 细节丢失

2.1.3 低光照与强点光对视觉任务的影响

图像质量的退化对后续视觉任务产生直接而显著的影响。对于皮带机异物检测任务而言, 低光照与强点光并存的成像条件带来了多方面的挑战。

首先, 从特征提取角度分析, 深度神经网络通常依赖图像的边缘、纹理等底层特征进行目标识别。低光照导致图像对比度降低、边缘响应减弱, 使得卷积层难以提取到足够区分度的特征表示。图 2.3 对比了正常光照与低光照条件下卷积特征图的差异, 可见低光照图像的特征响应明显弱于正常光照图像, 这直接影响

了检测网络的判别能力。

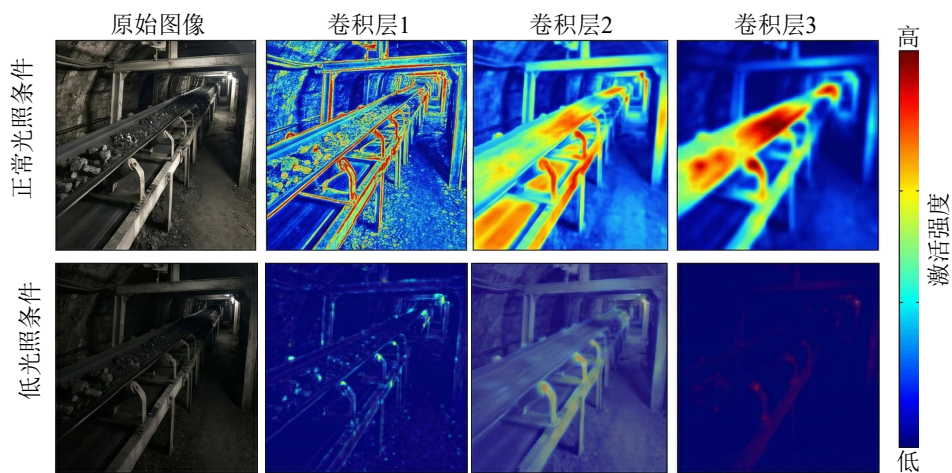


图 2.3 不同光照条件下特征提取效果对比

其次，从目标定位角度分析，异物目标在低光照区域往往与背景融为一体，边界模糊不清，增加了边界框回归的难度。而在强点光区域，过曝导致的像素饱和使得目标轮廓信息丢失，同样不利于精确定位。此外，光晕效应产生的虚假高亮区域可能被检测器误判为目标，增加了误检风险^[57]。

再次，从检测稳定性角度分析，井下皮带机监控场景中，光照条件可能随设备运行状态、人员走动等因素发生动态变化。这种光照波动导致同一目标在不同帧中的表观特征差异显著，影响检测结果的时序一致性和稳定性，对于需要实时响应的安全监测系统而言是不可忽视的问题。

表 2.1总结了不同光照条件对异物检测任务各环节的影响程度。从表中可以看出，低光照与强点光对检测任务的各个环节均产生负面影响，其中特征提取和目标定位受影响最为显著。这一分析表明，在进行异物检测之前对图像进行有效的增强处理具有必要性和迫切性。

表 2.1 不同光照条件对异物检测任务的影响分析

检测环节	正常光照	低光照	强点光
特征提取	正常	严重退化	局部失效
目标定位	准确	偏差增大	边界模糊
类别识别	可靠	易混淆	受干扰
检测稳定性	稳定	波动大	不稳定

2.2 低光照图像增强基础理论

2.2.1 低光照图像退化模型

建立准确的图像退化模型是设计有效增强算法的理论前提。低光照图像的形成过程涉及场景辐射、大气传输、光电转换等多个物理环节，可通过数学模型进行抽象描述。

基于物理成像原理，低光照图像的退化过程可表示为：

$$I_{\text{low}}(x, y) = f(R(x, y) \cdot S(x, y)) + n(x, y) \quad (2.3)$$

其中， $I_{\text{low}}(x, y)$ 为观测到的低光照图像， $R(x, y)$ 为场景反射率分量， $S(x, y)$ 为光照分量， $f(\cdot)$ 为相机响应函数， $n(x, y)$ 为成像噪声。在理想线性响应条件下， $f(\cdot)$ 可近似为恒等映射，退化模型简化为：

$$I_{\text{low}}(x, y) = R(x, y) \cdot S(x, y) + n(x, y) \quad (2.4)$$

从式 (2.4) 可以看出，低光照图像退化的本质是光照分量 $S(x, y)$ 取值过低。当 $S(x, y) \ll 1$ 时，乘积项 $R(x, y) \cdot S(x, y)$ 的有效信号强度显著降低，而加性噪声 $n(x, y)$ 的相对影响则被放大，导致信噪比下降。

在煤矿井下场景中，光照分量 $S(x, y)$ 的空间分布极不均匀。根据前文分析，光照强度随距光源距离的增加而快速衰减，这导致 $S(x, y)$ 在图像不同区域的取值可能相差数个数量级。设 S_{max} 和 S_{min} 分别为图像中光照分量的最大值和最小值，则动态范围比 γ 可定义为：

$$\gamma = \frac{S_{\text{max}}}{S_{\text{min}}} \quad (2.5)$$

对于井下强点光场景， γ 值可达 10^2 甚至更高，远超一般成像传感器 8 位量化 ($2^8 = 256$ 级) 的表示能力。这种高动态范围特性使得单一的全局增强策略难以同时照顾暗区提亮与亮区细节保持的需求^[58]。

图 2.4 以图示方式说明了低光照图像退化模型各组成部分及其相互关系。

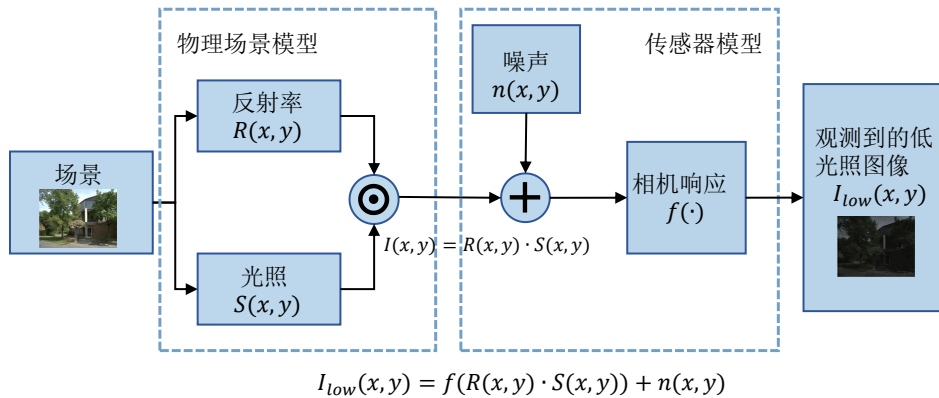


图 2.4 低光照图像退化模型示意图

2.2.2 Retinex 理论及其数学表达

Retinex 理论旨在解释人类视觉系统的颜色恒常性现象^[8]。该理论的核心假设是：人眼感知到的图像可以分解为物体固有的反射特性和环境光照条件两个独立分量的乘积。基于这一思想，Retinex 理论为低光照图像增强提供了重要的理论框架。

Retinex 成像模型的数学表达式为：

$$I(x, y) = R(x, y) \odot S(x, y) \quad (2.6)$$

其中， $I(x, y)$ 为输入低光照图像， $R(x, y)$ 为反射分量（Reflectance），表征物体表面的固有反射属性，与光照条件无关； $S(x, y)$ 为光照分量（Illumination），描述环境光照的空间分布； $\odot$ 表示逐像素乘法操作。

在对数域中，式 (2.6) 可转化为加性形式：

$$\log I(x, y) = \log R(x, y) + \log S(x, y) \quad (2.7)$$

这一变换将乘性分解问题转化为加性分离问题，便于采用线性滤波等方法进行处理。

基于 Retinex 理论的图像增强方法的核心思想是：从观测图像中估计光照分量 $S(x, y)$ ，进而恢复反射分量 $R(x, y)$ 。由于反射分量代表物体的本征特性，理论上不受光照变化影响，因此可作为增强结果输出。增强后的图像 I_{out} 可表示为：

$$I_{\text{out}}(x, y) = \frac{I(x, y)}{S(x, y)} = R(x, y) \quad (2.8)$$

然而，从单幅观测图像 $I(x, y)$ 中同时估计 $R(x, y)$ 和 $S(x, y)$ 是一个病态问题（Ill-posed Problem），因为存在无穷多组 (R, S) 的组合可以产生相同的观测图像。为求解这一问题，需要引入额外的先验约束条件。

常用的先验约束包括：

（1）光照平滑性约束

自然场景中的光照分布通常呈现平滑变化的特性，剧烈的光照突变较为罕见。基于这一观察，可对光照分量施加平滑性正则化：

$$\mathcal{L}_{\text{smooth}} = \sum_{x, y} (|\nabla_x S(x, y)| + |\nabla_y S(x, y)|) \quad (2.9)$$

其中， ∇_x 和 ∇_y 分别表示水平和垂直方向的梯度算子。

（2）反射分量结构保持约束

反射分量应保留原图像的结构和纹理信息，因此其边缘结构应与输入图像保持一致：

$$\mathcal{L}_{\text{struct}} = \sum_{x, y} w(x, y) \cdot |\nabla R(x, y) - \nabla I(x, y)| \quad (2.10)$$

其中, $w(x, y)$ 为边缘感知权重函数。

图 2.5 展示了 Retinex 分解的示意过程, 直观说明了从输入图像分离反射分量和光照分量的基本原理。

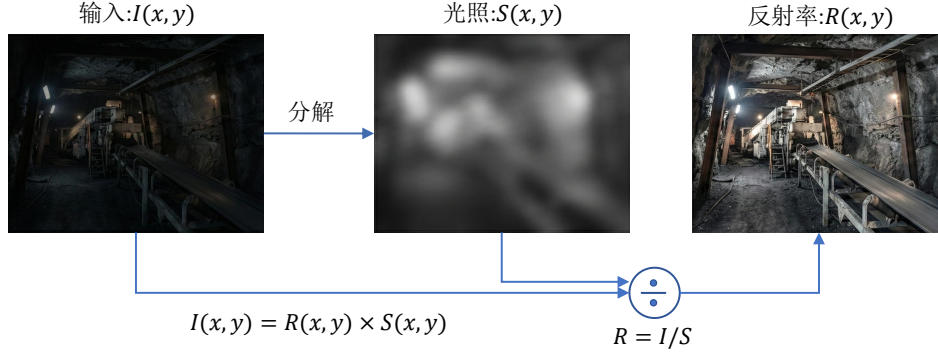


图 2.5 Retinex 分解示意图

经典的单尺度 Retinex (Single-Scale Retinex, SSR) 算法采用高斯滤波估计光照分量^[8]:

$$S(x, y) = I(x, y) * G_{\sigma}(x, y) \quad (2.11)$$

其中, G_{σ} 为标准差为 σ 的高斯核, $*$ 表示卷积操作。MSR 则通过融合多个尺度的估计结果来平衡局部对比度增强与全局色彩保真之间的关系。

近年来, 研究者将深度学习方法引入 Retinex 框架, 利用神经网络的强大表示能力学习更为准确的光照估计和反射分量恢复映射。这类方法通过数据驱动的方式隐式学习先验约束, 在复杂场景中展现出优于传统方法的性能表现^[11]。

2.2.3 图像质量评价指标

图像质量评价是衡量增强算法性能的重要环节。根据是否需要参考图像, 评价指标可分为全参考指标和无参考指标两类。

(1) 全参考图像质量评价指标

全参考指标通过比较增强图像与正常光照参考图像之间的差异来评估质量。常用指标包括:

峰值信噪比 (*Peak Signal-to-Noise Ratio, PSNR*): 衡量图像与参考图像之间的像素级误差^[59], 定义为:

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{L^2}{\text{MSE}} \right) \quad (2.12)$$

其中, L 为像素值的最大取值 (8 位图像为 255), MSE 为均方误差。PSNR 值越高表示增强图像与参考图像越接近。

结构相似性指数 (*Structural Similarity Index Measure, SSIM*): 从亮度、对比度和结构三个维度评估图像相似性^[60]:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (2.13)$$

其中, μ_x 、 μ_y 为局部均值, σ_x^2 、 σ_y^2 为局部方差, σ_{xy} 为协方差, C_1 、 C_2 为稳定常数。SSIM 取值范围为 $[0, 1]$, 值越大表示结构相似度越高。

(2) 无参考图像质量评价指标

在缺乏参考图像的实际应用场景中, 无参考指标具有更广泛的适用性。常用的无参考指标包括:

自然图像质量评估器 (*Natural Image Quality Evaluator, NIQE*): 基于自然场景统计特性, 通过比较待评图像与自然图像模型之间的偏差来评估质量^[61]。NIQE 值越低表示图像越接近自然图像的统计特性。

盲图像空间质量评估器 (*Blind/Referenceless Image Spatial Quality Evaluator, BRISQUE*): 利用归一化亮度系数的统计分布特征, 训练支持向量回归模型预测图像质量分数^[62]。与 NIQE 类似, BRISQUE 值越低表示质量越好。

感知图像质量评估器 (*Perception-based Image Quality Evaluator, PIQE*): 通过分析图像局部块的失真特性和空间活动性来评估质量^[63]。PIQE 适用于评估噪声、模糊等多种失真类型。

表 2.2 总结了上述评价指标的特点及适用场景。

表 2.2 图像质量评价指标总结

指标	类型	最优方向	评价维度
PSNR	全参考	越高越好	像素级误差
SSIM	全参考	越高越好	结构相似性
NIQE	无参考	越低越好	自然性统计
BRISQUE	无参考	越低越好	空间失真特征
PIQE	无参考	越低越好	感知质量

在煤矿井下图像增强评估中, 由于正常光照参考图像难以获取, 无参考指标具有更强的实用价值。综合采用 BRISQUE、NIQE 和 PIQE 三项无参考指标, 可以从不同角度全面评估增强效果。

2.3 深度学习目标检测基础理论

2.3.1 目标检测问题定义

目标检测是计算机视觉领域的核心研究任务之一, 旨在从图像或视频中自动定位感兴趣目标的空间位置并识别其类别。与图像分类任务仅需判断图像整体类别不同, 目标检测需要同时完成“在哪里”和“是什么”两个子任务^[64]。

形式化地, 给定输入图像 $\mathbf{I} \in \mathbb{R}^{H \times W \times 3}$, 目标检测任务的输出为一组检测结果 $\mathcal{D} = \{(\mathbf{b}_i, c_i, s_i)\}_{i=1}^N$, 其中:

- $\mathbf{b}_i = (x_i, y_i, w_i, h_i)$ 为第 i 个目标的边界框 (Bounding Box), (x_i, y_i) 表示边界框中心坐标, (w_i, h_i) 表示宽度和高度;
- $c_i \in \{1, 2, \dots, K\}$ 为目标的类别标签, K 为预定义类别总数;
- $s_i \in [0, 1]$ 为检测置信度分数, 反映检测结果的可靠程度。

目标检测算法的性能通常采用以下指标进行评估:

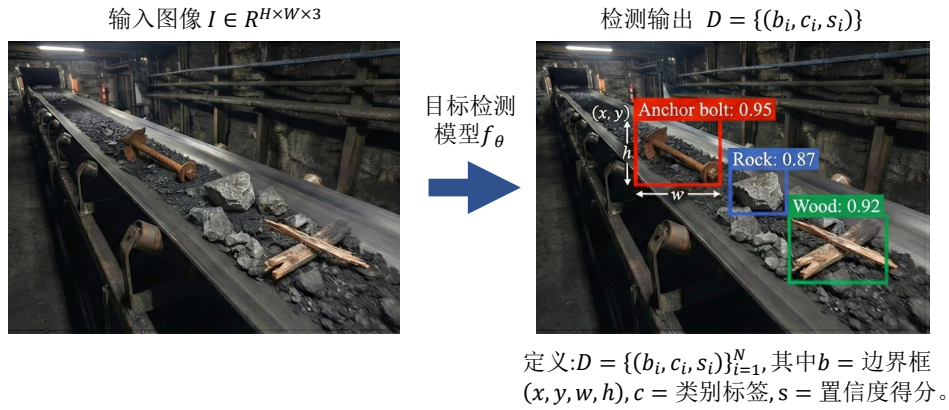
交并比 (Intersection over Union, IoU): 衡量预测边界框与真实边界框的重叠程度:

$$\text{IoU} = \frac{\text{Area}(\mathbf{b}_{\text{pred}} \cap \mathbf{b}_{\text{gt}})}{\text{Area}(\mathbf{b}_{\text{pred}} \cup \mathbf{b}_{\text{gt}})} \quad (2.14)$$

平均精度 (Average Precision, AP): 在不同召回率水平下精度的平均值, 综合反映算法的检测性能。

平均精度均值 (mean Average Precision, mAP): 所有类别 AP 值的平均, 是目标检测任务最常用的综合评价指标。

图 2.6 以图示方式说明了目标检测任务的输入输出关系及 IoU 的计算方式。



IOU (交并比) 示意图

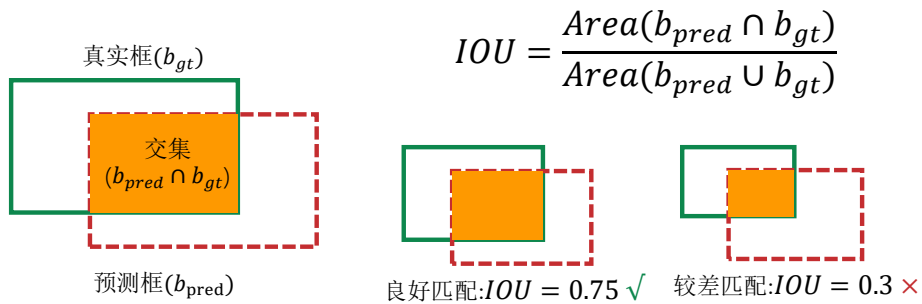


图 2.6 目标检测任务定义示意图

2.3.2 YOLO 系列模型基本结构

YOLO 系列算法在单阶段检测领域拥有最为深远的影响力, 自 2016 年首次提出以来, 经过多轮版本迭代与架构革新, 在检测准确率与推理速度方面持续突破^[65]。

(1) YOLO 整体架构

YOLO 系列模型通常包含三个核心组件：

主干网络 (Backbone)：承担输入图像多尺度特征编码的职责。早期版本使用 Darknet 架构，后续版本引入 CSPNet、EfficientNet 等更先进的骨干结构，以平衡特征表达能力和计算效率。

颈部网络 (Neck)：用于融合主干网络输出的多尺度特征，增强对不同尺寸目标的检测能力。常用结构包括特征金字塔网络 (Feature Pyramid Network, FPN)^[66] 和路径聚合网络 (Path Aggregation Network, PAN)^[67] 等。

检测头 (Head)：基于融合后的特征进行目标分类和边界框回归预测。根据预测方式的不同，可分为基于锚框 (Anchor-based) 和无锚框 (Anchor-free) 两种类型。

图 2.7 展示了 YOLO 系列模型的典型架构。

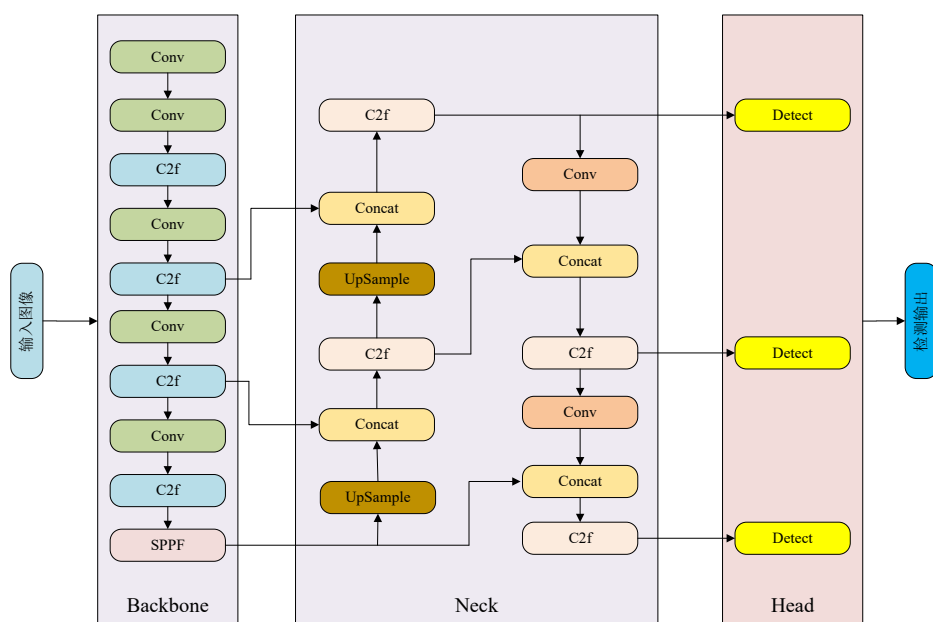


图 2.7 YOLO 系列模型典型架构

(2) YOLO 版本演进

YOLO 系列经历了从 v1 到 v11 的持续发展，各版本的主要改进如下：

YOLOv1-v3：奠定了 YOLO 系列的基础架构。v1 首次提出将检测问题转化为回归问题；v2 引入批归一化和锚框机制；v3 引入多尺度输出层与残差连接机制，大幅改善了小尺度目标的识别效果。

YOLOv4-v5：引入了大量训练技巧和结构优化。v4 整合了 CSPDarknet、SPP、PANet 等模块；v5 在工程实现上进行了优化，提供了多种规模的模型变体 (n/s/m/l/x)，成为应用最广泛的版本之一。

YOLOv6-v8：进一步提升了检测效率和精度。v6 针对工业部署进行优化；v7 引入 E-ELAN 和辅助训练头等创新；v8 采用无锚框检测头和改进的 C2f 模块，实

现了精度-速度的更优平衡。

YOLOv9-v11：代表了最新的技术进展。v9 引入可编程梯度信息和广义高效层聚合网络；v10 实现了端到端无 NMS 检测；v11 在多任务学习和部署效率方面进一步优化。

表 2.3总结了 YOLO 系列主要版本的关键改进。

表 2.3 YOLO 系列主要版本演进

版本	发布时间	关键改进
YOLOv1	2016	首次实现端到端实时检测
YOLOv3	2018	多尺度预测、残差连接
YOLOv5	2020	工程优化、多规模模型
YOLOv8	2023	无锚框检测头、C2f 模块
YOLOv10	2024	端到端 NMS-free 检测

(3) YOLOv8 关键模块

YOLOv8 作为当前应用较为广泛的版本，其关键模块设计具有代表性：

C2f 模块：融合了 CSP（Cross Stage Partial）结构和 ELAN（Efficient Layer Aggregation Network）的思想，通过密集连接和跨阶段特征融合增强了特征表达能力，同时保持了计算效率。其结构可表示为：

$$\mathbf{F}_{\text{out}} = \text{Conv}(\text{Concat}(\mathbf{F}_{\text{in}}, \text{Bottleneck}_1(\mathbf{F}_{\text{in}}), \dots, \text{Bottleneck}_n(\mathbf{F}_{\text{in}}))) \quad (2.15)$$

解耦检测头：将分类任务和回归任务分离到不同的分支进行处理，减少了任务间的梯度干扰，提升了收敛速度和最终性能。

无锚框预测：摒弃预定义锚框，直接预测目标中心点相对于网格的偏移量和目标尺寸，简化了模型设计并提高了对不规则形状目标的适应性。

2.3.3 注意力机制与轻量化网络简介

(1) 注意力机制

注意力机制（Attention Mechanism）是深度学习领域的重要技术，能够引导网络有选择性地关注输入中的重要信息，提升特征表示的判别性^[68]。在目标检测任务中，注意力机制的引入有助于增强对目标区域的关注，抑制背景干扰。

根据作用维度的不同，注意力机制可分为：

通道注意力：学习各通道特征的重要性权重。SENet^[69]提出的 Squeeze-and-Excitation 模块通过全局平均池化和全连接层建模通道间依赖关系：

$$\mathbf{s} = \sigma(\mathbf{W}_2 \cdot \delta(\mathbf{W}_1 \cdot \text{GAP}(\mathbf{F}))) \quad (2.16)$$

其中，GAP 表示全局平均池化， $\mathbf{W}_1$ 、 $\mathbf{W}_2$ 为全连接层权重， δ 和 σ 分别为 ReLU

和 Sigmoid 激活函数。

空间注意力：学习特征图各空间位置的重要性。通过聚合通道维度的信息，生成空间注意力图以突出关键区域。

混合注意力：同时考虑通道和空间两个维度。CBAM^[70]依次应用通道注意力和空间注意力模块，实现了更全面的特征增强。

坐标注意力：Coordinate Attention^[71]通过沿水平和垂直方向分别进行全局平均池化，将位置信息显式编码到通道注意力中。其计算过程如图 2.8所示，具体步骤包括：

$$\mathbf{z}_c^h = \frac{1}{W} \sum_{i=1}^W x_c(h, i) \quad (2.17)$$

$$\mathbf{z}_c^w = \frac{1}{H} \sum_{j=1}^H x_c(j, w) \quad (2.18)$$

其中， $\mathbf{z}_c^h$ 和 $\mathbf{z}_c^w$ 分别为沿宽度和高度方向聚合得到的特征向量。后续通过共享卷积和分支卷积生成两个方向的注意力权重，并与原特征逐元素相乘完成特征重标定。

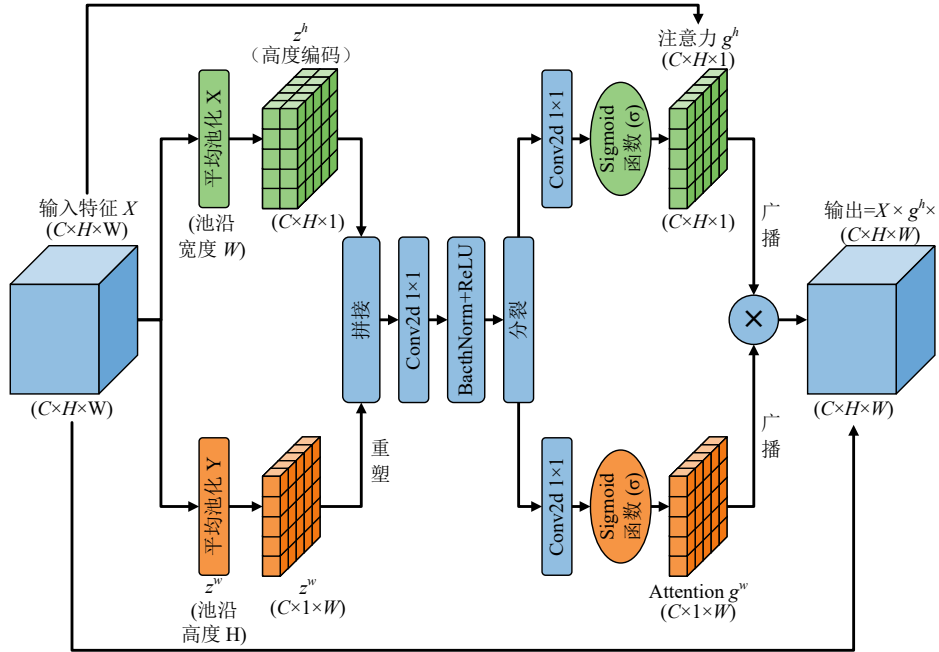


图 2.8 坐标注意力模块结构图

坐标注意力机制在保持较低计算开销的同时，能够捕获长距离的位置依赖信息，特别适合于需要感知目标空间分布的检测任务。在煤矿井下皮带机异物检测场景中，皮带表面在图像中的位置相对固定，坐标注意力可引导网络关注皮带区域及其周边的关键位置。

(2) 轻量化网络设计

在边缘计算和嵌入式设备上部署目标检测模型，需要在有限的算力和内存资源约束下实现高效推理。轻量化网络设计旨在降低模型的参数量和计算量，同时尽可能保持检测性能^[72]。

常用的轻量化技术包括：

深度可分离卷积：MobileNet^[36]提出将标准卷积分解为深度卷积（Depthwise Convolution）和逐点卷积（Pointwise Convolution）两步操作。设输入特征图通道数为 C_{in} ，输出通道数为 C_{out} ，卷积核大小为 $K \times K$ ，则参数量从 $K^2 \cdot C_{in} \cdot C_{out}$ 降低为 $K^2 \cdot C_{in} + C_{in} \cdot C_{out}$ ，压缩比约为 $1/C_{out} + 1/K^2$ 。

Ghost 模块：GhostNet^[37]观察到卷积神经网络中存在大量相似的冗余特征，提出通过少量真实卷积生成部分特征，再利用廉价的线性变换（如深度卷积）生成其余“幻影”特征。该设计在保持特征多样性的同时显著降低了计算量。GhostNetV2^[38]进一步引入 DFC（Decoupled Fully Connected）注意力机制，增强了跨空间位置的特征交互能力。

通道混洗：ShuffleNet^[39]通过分组卷积降低计算量，并引入通道混洗操作促进不同组之间的信息交换，解决了分组卷积导致的特征隔离问题。

图 2.9 对比展示了几种典型轻量化模块的结构。

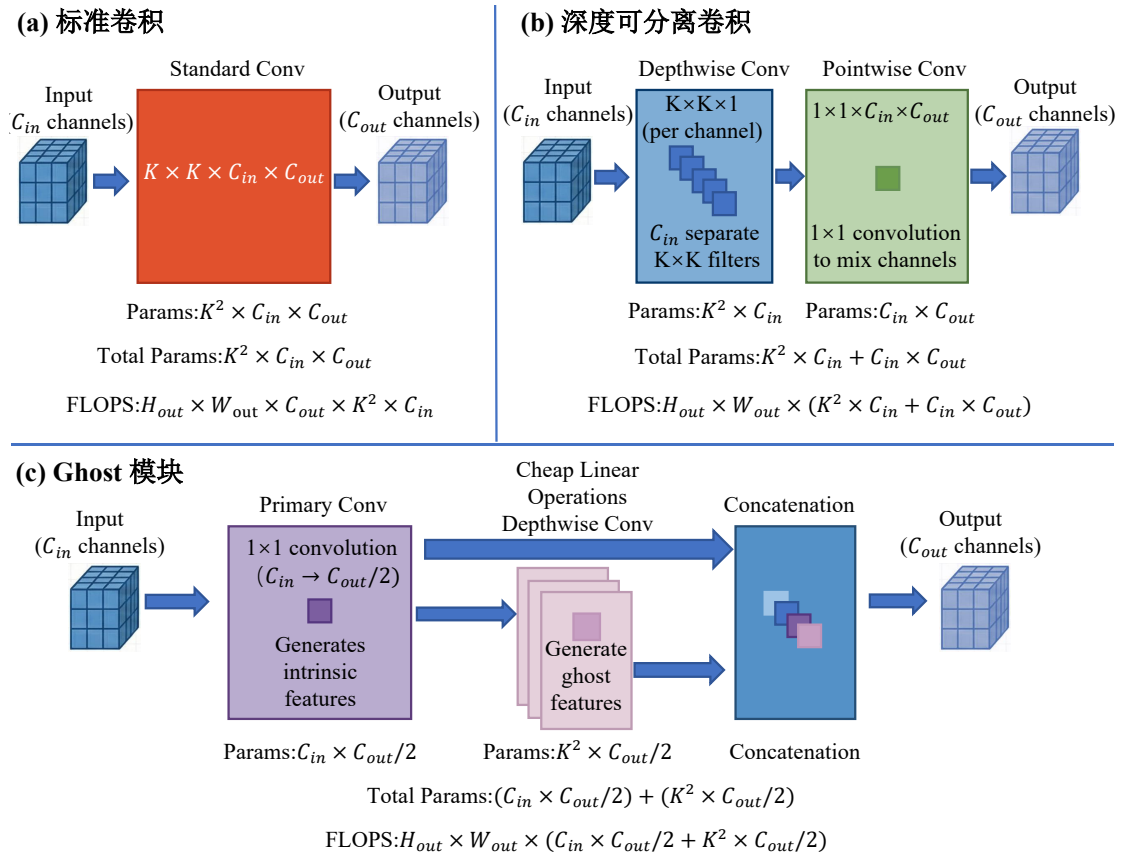


图 2.9 典型轻量化模块结构对比：(a) 标准卷积；(b) 深度可分离卷积；(c) Ghost 模块

表 2.4 从参数量、计算量和性能三个维度对比了几种轻量化主干网络。

表 2.4 轻量化主干网络对比

网络	参数量 (M)	FLOPs(G)	ImageNet Top-1(%)
MobileNetV2	3.4	0.30	72.0
ShuffleNetV2	2.3	0.15	69.4
GhostNet	5.2	0.14	73.9
GhostNetV2	6.1	0.17	75.3

在煤矿井下异物检测应用中，边缘计算设备的算力资源有限，轻量化网络设计对于实现实时检测至关重要。通过采用 GhostNetV2 等高效主干网络替代标准检测网络的主干部分，可以在保持检测精度的同时显著降低推理延迟和能耗。

2.4 本章小结

本章系统介绍了与煤矿井下皮带机异物检测研究相关的基础理论知识，主要包括以下几个方面：

首先，分析了煤矿井下特殊的视觉成像环境。井下空间封闭、光照依赖人工光源、粉尘干扰严重等因素共同导致了图像质量的显著退化。光照分布的非均匀性尤为突出，强点光源附近存在过曝和光晕效应，而远离光源的区域则严重欠曝。这种复杂的光照条件对视觉监测任务提出了严峻挑战。

其次，阐述了低光照图像增强的理论基础。建立了基于乘性模型的图像退化数学描述，介绍了 Retinex 理论的核心思想及其数学表达，分析了从观测图像中分离反射分量和光照分量的病态性质及常用的先验约束。在此基础上，综述了基于深度学习的低光照增强方法，包括监督学习和无监督学习两类代表性工作。同时，介绍了图像质量评价的全参考和无参考指标，为后续增强算法的评估提供了依据。

最后，介绍了深度学习目标检测的相关基础。明确了目标检测的问题定义和性能评价指标，对比分析了两阶段检测器和单阶段检测器的原理与特点，重点介绍了 YOLO 系列模型的架构演进和关键模块设计。此外，阐述了注意力机制和轻量化网络设计的基本原理，特别是坐标注意力机制和 Ghost 模块等适用于边缘部署的高效结构。

上述理论知识为后续章节中区域自适应低光照图像增强算法和语义引导的轻量化异物检测方法的设计奠定了坚实的理论基础。

第3章 区域自适应低光照图像增强方法与实验验证

3.1 引言

低光照图像增强是计算机视觉领域的基础性研究问题，旨在改善光照不足条件下拍摄图像的可见性、对比度和整体视觉质量。第二章已系统阐述了 Retinex 理论的数学基础、低光照图像退化模型以及主流增强方法的技术原理，本章在此基础上聚焦于算法设计与实验验证。

在煤矿井下等真实复杂场景中，低光照图像增强面临若干特殊困难。首先，光照分布呈现极端的不均匀性：照明设备呈离散点状分布，在光源附近形成高亮度区域，而远离光源的区域几乎完全处于黑暗状态，这种明暗共存的特征使得单一的全局增强策略难以同时兼顾暗区提亮与亮区保护^[73]。其次，强点光源引发的光效伪影问题突出：光线在灰尘颗粒和镜头表面的散射产生明显的光晕（Halo）效应，表现为光源周围的模糊亮斑，不仅降低图像对比度，还可能掩盖该区域内的目标信息。再次，暗区增强过程中噪声被同步放大：低光照条件下信噪比本就较低，直接提升亮度会导致噪声显著恶化，严重影响图像的可用性^[74]。

现有方法在处理上述问题时存在明显局限。LIME^[9]等传统 Retinex 方法缺乏区域差异化处理能力，在高动态范围场景中易导致局部过曝。SICE^[75]虽采用自适应增强策略，但未针对光效伪影进行抑制，导致光晕问题依然存在。RUAS^[10]侧重于计算效率优化，但在暗区增强与亮区伪影抑制的协同处理方面表现欠佳。基于深度学习的方法如 Zero-DCE^[17]和 SCI^[19]虽取得显著进展，但在强点光源场景中仍易产生不自然的伪影。

针对上述问题，本章提出一种区域自适应低光照图像增强框架。该框架在 Retinex 理论指导下，通过区域分割模块实现暗区与亮区的像素级软划分，结合自适应光照增强器提升暗区亮度，同时利用光效抑制器削弱亮区的光晕伪影。通过精心设计的复合损失函数，实现亮度增强、结构保持与伪影抑制的多目标平衡，为后续下游任务提供高质量的增强图像。

3.2 方法总体框架与关键模块设计

3.2.1 Retinex 分解与增强策略

本章方法基于第二章所述的 Retinex 理论框架。根据该理论，任意观测图像可分解为反射分量与光照分量的逐像素乘积：

$$\mathbf{I}_{in}(x, y) = \mathbf{R}(x, y) \odot \mathbf{S}(x, y) \quad (3.1)$$

其中, I_{in} 表示输入低光照图像, R 为反射分量, 表征物体表面的固有反射属性, S 为光照分量, 描述环境光照的空间分布, $\odot$ 表示逐元素乘法。

在低光照场景中, 光照分量 S 取值过低导致图像整体偏暗。传统 Retinex 方法通过直接估计原始光照分量进行增强, 但这种策略在非均匀光照条件下往往失效。本章采用不同的思路: 学习一个增强后的光照图 S_2 来补偿不良光照条件, 同时保持物体固有的反射特性。增强后的输出图像为:

$$I_{out} = R = \frac{I_{in}}{S_2} \quad (3.2)$$

其中, I_{out} 表示最终增强输出图像, S_2 为本章框架估计的增强光照图。与直接估计原始光照不同, S_2 通过区域自适应框架生成, 该框架能够区分需要增强的暗区与需要抑制伪影的亮区, 从而实现更鲁棒的图像分解。

3.2.2 总体框架概述

给定低光照输入图像 $I_{in} \in \mathbb{R}^{H \times W \times 3}$, 本章目标是生成具有优良视觉质量的增强图像 I_{out} 。如图 3.1 所示, 所提出的框架集成了去噪预处理、区域分割、自适应光照增强与光效抑制等模块, 基于 Retinex 理论进行图像分解并实现端到端优化。算法 3.1 给出了方法的完整处理流程。

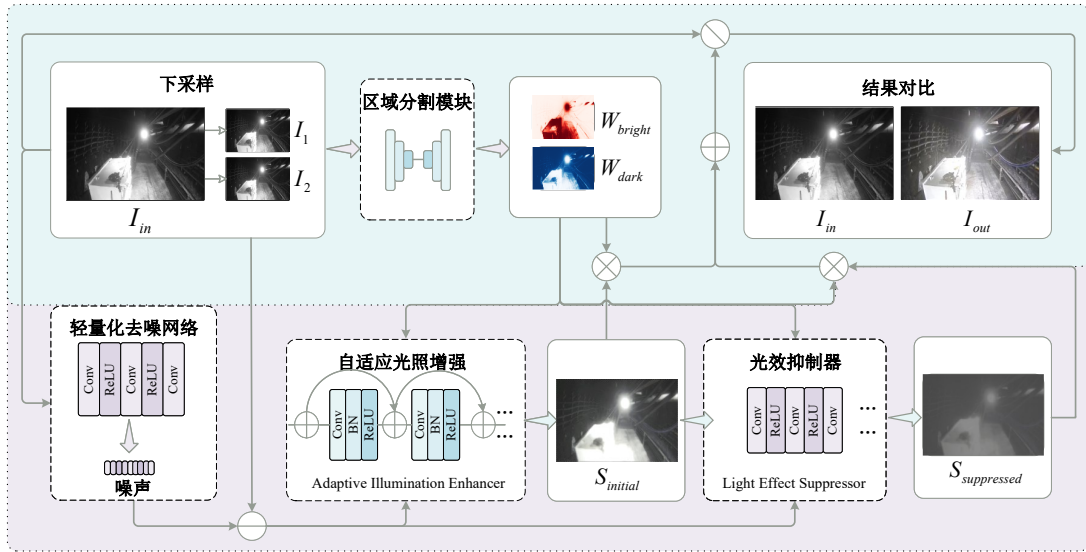


图 3.1 所提出的区域自适应低光照图像增强框架总体结构, 集成区域分割、自适应光照增强与光效抑制模块

在介绍各模块功能之前, 首先对图 3.1 中下采样模块的作用作必要说明。该模块对应于配对下采样操作 $G = (G_1, G_2)$, 通过对原始图像 I_{in} 在 2×2 邻域内按对角线规则进行互补抽取, 生成两幅低分辨率子图 $I_1 = G_1(I_{in})$ 与 $I_2 = G_2(I_{in})$ 。这两幅子图共享相同的场景内容, 但其上每个像素来自原图中不同空间位置的采样点, 因此各像素位置的噪声实现彼此独立。该性质是 LDN 实现自监督去噪的理论基础——网络可通过最小化两幅子图间的去噪一致性损失式 (3.15)、式

(3.16) 学习去噪映射，而无需任何干净参考图像。需要强调的是， I_1 和 I_2 仅在训练阶段为 LDN 提供监督信号，在推理阶段并不参与正向传播，最终增强结果由原始输入 I_{in} 经完整框架直接生成。

该框架包含四个主要阶段。首先，输入图像 I_{in} 经轻量化去噪网络 $\mathcal{F}_{LDN}$ 进行预处理，得到降噪后的图像 $L_2 = I_{in} - \mathcal{F}_{LDN}(I_{in})$ 。该步骤减轻噪声干扰，为后续光照调整提供更洁净的输入。同时，区域分割模块分析 I_{in} ，生成像素级权重图：用于标识需要增强的暗区权重 W_{dark} 和需要抑制过曝的亮区权重 W_{bright} 。

基于 L_2 和区域权重图，自适应光照增强器估计初始光照层 $S_{initial}$ ，优先增强暗区亮度。随后，光效抑制器针对亮区对 $S_{initial}$ 进行精调，生成抑制后的光照层 $S_{suppressed}$ ，以削弱光晕和过曝伪影。这两个光照分量通过区域权重融合形成最终光照层：

$$S_2 = S_{initial} \odot W_{dark} + S_{suppressed} \odot W_{bright} \quad (3.3)$$

依据 Retinex 理论，最终增强图像通过 $I_{out} = I_{in}/S_2$ 计算得到，确保暗区亮度提升、亮区自然抑制、结构细节保持。整个框架采用无监督方式训练，通过精心设计的复合损失函数实现区域自适应重建、结构一致性、高频细节保持与平滑性约束的综合平衡。

算法 3.1 区域自适应低光照图像增强总体框架

Require: 低光照输入图像 $I_{in} \in \mathbb{R}^{H \times W \times 3}$

Ensure: 增强输出图像 I_{out}

- 1: **初始去噪:** 应用轻量化去噪网络 $\mathcal{F}_{LDN}$
 - 2: $L_2 = I_{in} - \mathcal{F}_{LDN}(I_{in})$
 - 3: **区域分割:** 分析 I_{in} 生成像素级区域信息
 - 4: 提取暗区权重图 W_{dark} 与亮区权重图 W_{bright}
 - 5: **自适应光照估计:**
 - 6: 使用自适应增强器估计初始光照层
 - 7: $S_{initial} = \text{AdaptiveEnhancer}(L_2, W_{dark}, W_{bright})$
 - 8: **光效抑制:**
 - 9: 应用光效抑制器调整亮区
 - 10: $S_{suppressed} = \text{LightEffectSuppressor}(L_2, S_{initial}, W_{bright})$
 - 11: **光照融合:** 融合增强与抑制后的光照图
 - 12: $S_2 = S_{initial} \odot W_{dark} + S_{suppressed} \odot W_{bright}$
 - 13: **反射分量恢复:** 应用 Retinex 分解获取最终结果
 - 14: $I_{out} = I_{in}/S_2$
 - 15: **return** I_{out}
-

3.2.3 区域分割模块

如图 3.2 所示，区域分割模块（Region Segmentation Module, RSM）旨在将输入图像划分为暗区和亮区，并为每个像素生成平滑的归属权重。该模块通过三个主要步骤实现。

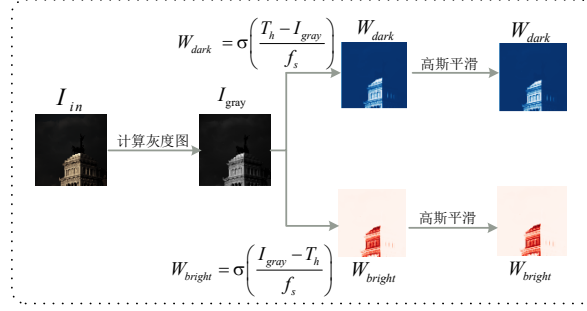


图 3.2 区域分割模块 (RSM) 结构

首先，将输入 RGB 图像 I_{in} 转换为灰度强度图 I_{gray} 。然后设定阈值 T_h （默认值为 0.5），基于每个像素强度值与 T_h 的偏差，通过 Sigmoid 函数计算暗区权重 W_{dark} 和亮区权重 W_{bright} ，以确保权重的平滑过渡：

$$W_{dark} = \sigma \left(\frac{T_h - I_{gray}}{f_s} \right) \quad (3.4)$$

$$W_{bright} = \sigma \left(\frac{I_{gray} - T_h}{f_s} \right) \quad (3.5)$$

其中， $\sigma(\cdot)$ 为 Sigmoid 函数， f_s 为平滑因子，控制权重过渡的陡峭程度。为进一步避免区域边界处可能产生的伪影，对生成的权重图施加高斯平滑处理。最终，该模块输出平滑的权重图 W_{dark} 和 W_{bright} 供后续处理使用。

与传统的硬阈值分割方法不同，本模块采用软分割策略，使得每个像素同时具有一定程度的暗区和亮区归属度。这种设计避免了硬边界带来的不连续性问题，使得后续的光照增强与抑制过程能够在区域之间平滑过渡。

需要进一步说明的是，本模块以原始输入图像 I_{in} 而非经过 LDN 处理后的图像 L_2 作为输入，主要基于以下三方面考虑。其一，从频域特性看，区域分割本质上依赖图像的全局光照亮度分布这一低频信息，而井下低光照场景中占主导的高斯噪声与泊松噪声均为零均值高频扰动。本模块内部已构成三级隐式低通滤波链路：RGB 至灰度的通道平均运算（可使独立同分布噪声的幅值降至 $1/\sqrt{3}$ 量级）、Sigmoid 函数的非线性平滑映射、以及输出端施加的 5×5 高斯平滑（标准差 1.0），三者协同可有效抑制噪声对权重图 W_{dark} 、 W_{bright} 的扰动。其二，从训练动力学看，若以 L_2 作为分割输入，则 RSM 的输出依赖于 LDN 的实时输出，而 LDN 在训练初期参数尚未稳定，会引发权重图的剧烈波动，进而反向干扰 AIE 与 LES 的优化过程，形成“分割—去噪”循环耦合。以 I_{in} 为输入则可保证 RSM 输出在整个训练过程中具有确定性，避免此类耦合震荡。其三，本章在私有煤矿数据集上对两种输入方案进行了对比验证，以 L_2 作为 RSM 输入时 BRISQUE、NIQE 和 PIQE 三项指标分别为 17.92、4.25、22.83，与采用 I_{in} 方案 (17.78、4.22、22.41) 的差异均在 1% 以内，而后者的训练损失收敛曲线更加平稳。综上，选择 I_{in} 作为 RSM 输入在数学合理性与工程稳定性之间达到了较好平衡。

3.2.4 自适应光照增强器

如图 3.3 所示，自适应光照增强器（Adaptive Illumination Enhancer, AIE）基于区域信息对预处理后的图像 L_2 进行光照估计与自适应调整。算法 3.2 给出了具体实现细节。

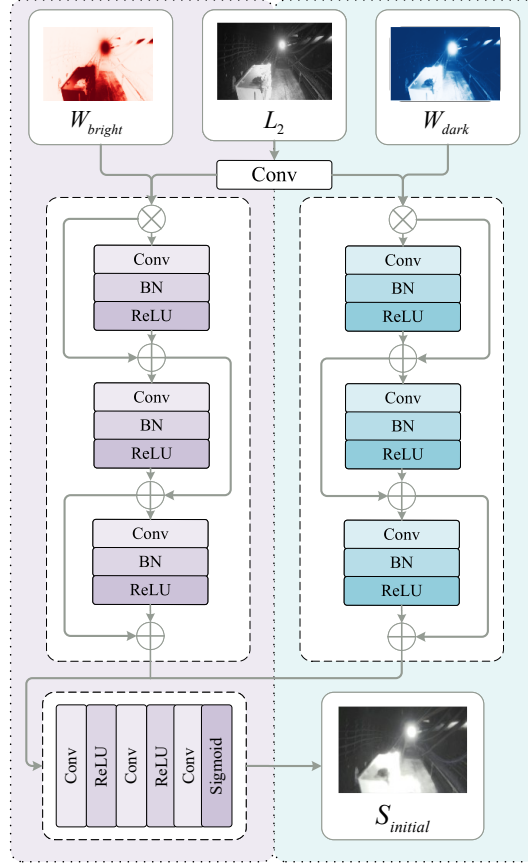


图 3.3 自适应光照增强器（AIE）结构

该增强器通过四阶段流水线实现。首先，通过卷积层进行基础特征提取，得到 F_{base} 。然后，利用区域分割模块输出的权重对 F_{base} 进行区域感知处理，分离出暗区特征 F_{dark} 和亮区特征 F_{bright} 。这些特征通过共享权重的残差卷积块并行处理，使得不同区域能够获得最优的特征变换。

处理后的特征进行拼接，送入自适应融合模块（由卷积层组成）学习其最优组合方式。最终，通过卷积层和 Sigmoid 激活函数生成初始光照图 $S_{initial}$ 。

3.2.5 光效抑制器

如图 3.4 所示，光效抑制器（Light Effect Suppressor, LES）专门针对过亮区域和光晕效应，通过对亮区光照估计进行特定调整来解决这些问题。算法 3.3 给出了实现细节。

该抑制器通过多阶段处理实现。首先，将去噪图像 L_2 与初始光照估计 $S_{initial}$

算法 3.2 自适应光照增强器

- Require:** 去噪图像 L_2 , 区域权重 W_{dark} , W_{bright}
- Ensure:** 初始光照图 S_{initial}
- 1: **基础特征提取:** 通过卷积层提取特征
 - 2: $F_{\text{base}} = \text{Conv}(L_2)$
 - 3: **区域感知处理:** 按区域权重调制特征
 - 4: $F_{\text{dark}} = F_{\text{base}} \odot W_{\text{dark}}$
 - 5: $F_{\text{bright}} = F_{\text{base}} \odot W_{\text{bright}}$
 - 6: **并行区域增强:** 通过共享权重残差块处理各区域特征
 - 7: $F_{\text{dark}}^{\text{enhanced}} = \text{ResidualBlocks}(F_{\text{dark}})$
 - 8: $F_{\text{bright}}^{\text{enhanced}} = \text{ResidualBlocks}(F_{\text{bright}})$
 - 9: **自适应融合:** 拼接并融合处理后的特征
 - 10: $F_{\text{fused}} = \text{AdaptiveFusion}(\text{Concat}(F_{\text{dark}}^{\text{enhanced}}, F_{\text{bright}}^{\text{enhanced}}))$
 - 11: **光照生成:** 生成初始光照图
 - 12: $S_{\text{initial}} = \text{Sigmoid}(\text{Conv}(F_{\text{fused}}))$
 - 13: 将数值钳制到 $[\epsilon, 1]$ 范围, 其中 ϵ 为小正常数
 - 14: **return** S_{initial}

沿通道维度拼接, 送入由多层卷积与 LeakyReLU 激活函数组成的编码器网络, 学习光效模式。编码器输出通过注意力模块 (卷积层后接 Sigmoid 函数) 生成注意力图 A_{light} , 用于突出潜在的光效区域。该注意力机制采用卷积操作聚焦于空间光效模式, 与传统自注意力方法有所不同。

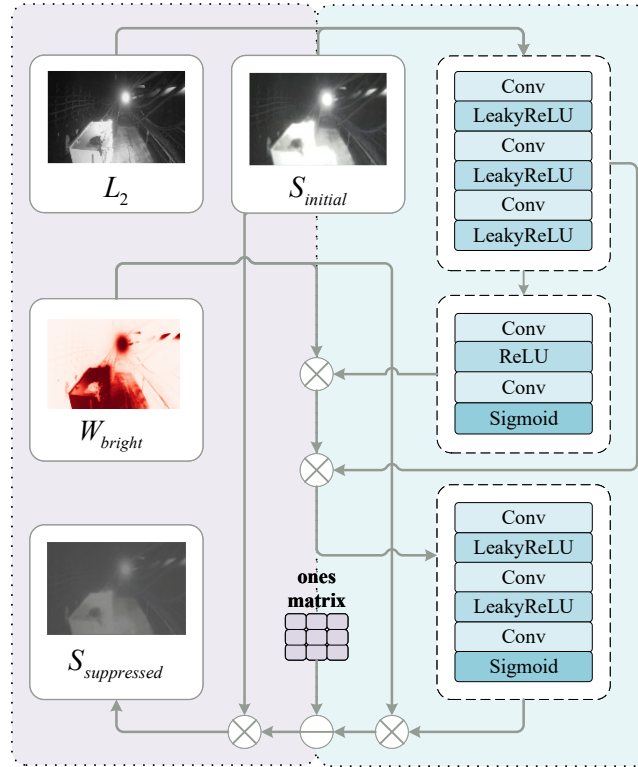


图 3.4 光效抑制器 (LES) 结构

注意力图 A_{light} 与区域权重 W_{bright} 的作用不同: W_{bright} 基于强度阈值识别

潜在亮区，而 $\mathbf{A}_{\text{light}}$ 通过特征分析学习识别复杂的光效模式。注意力图与输入的亮区权重 $\mathbf{W}_{\text{bright}}$ 逐元素相乘得到组合权重 $\mathbf{W}_{\text{combined}}$ ，确保抑制仅在同时满足强度条件和特征条件的区域生效。

编码器特征经 $\mathbf{W}_{\text{combined}}$ 加权后送入抑制器网络（卷积层加 LeakyReLU 激活后接 Sigmoid），生成逐像素抑制因子 $\mathbf{F}_{\text{suppress}} \in [0, 1]$ 。最后，通过以下方式自适应调整初始光照图：

$$\mathbf{S}_{\text{suppressed}} = \mathbf{S}_{\text{initial}} \odot (1 - \mathbf{F}_{\text{suppress}} \odot \mathbf{W}_{\text{bright}}) \quad (3.6)$$

通过该框架，获得针对暗区优化的 $\mathbf{S}_{\text{initial}}$ 和针对亮区优化的 $\mathbf{S}_{\text{suppressed}}$ ，两者通过区域权重融合形成最终光照图 $\mathbf{S}_2$ ，实现不同光照条件下的平衡增强。

算法 3.3 光效抑制器

Require: 去噪图像 $\mathbf{L}_2$ ，初始光照 $\mathbf{S}_{\text{initial}}$ ，亮区权重 $\mathbf{W}_{\text{bright}}$

Ensure: 抑制后光照 $\mathbf{S}_{\text{suppressed}}$

- 1: **输入构建:** 拼接输入特征
 - 2: $\mathbf{F}_{\text{input}} = \text{Concat}(\mathbf{L}_2, \mathbf{S}_{\text{initial}})$
 - 3: **特征编码:** 通过编码器网络学习光效模式
 - 4: $\mathbf{F}_{\text{encoded}} = \text{Encoder}(\mathbf{F}_{\text{input}})$ {多层 Conv + LeakyReLU}
 - 5: **注意力引导:** 生成光效模式注意力图
 - 6: $\mathbf{A}_{\text{light}} = \text{Sigmoid}(\text{Conv}(\mathbf{F}_{\text{encoded}}))$
 - 7: **区域加权注意力:** 将注意力与亮区权重组合
 - 8: $\mathbf{W}_{\text{combined}} = \mathbf{A}_{\text{light}} \odot \mathbf{W}_{\text{bright}}$
 - 9: **抑制因子生成:** 生成逐像素抑制因子
 - 10: $\mathbf{F}_{\text{weighted}} = \mathbf{F}_{\text{encoded}} \odot \mathbf{W}_{\text{combined}}$
 - 11: $\mathbf{F}_{\text{suppress}} = \text{Sigmoid}(\text{Suppressor}(\mathbf{F}_{\text{weighted}}))$ {Conv + LeakyReLU + Sigmoid}
 - 12: **自适应光照抑制:** 对初始光照施加抑制
 - 13: $\mathbf{S}_{\text{suppressed}} = \mathbf{S}_{\text{initial}} \odot (1 - \mathbf{F}_{\text{suppress}} \odot \mathbf{W}_{\text{bright}})$
 - 14: **return** $\mathbf{S}_{\text{suppressed}}$
-

3.3 损失函数与实现细节

3.3.1 复合损失函数设计

为有效训练所提出的网络并平衡不同增强目标，本章设计了一种新颖的自适应复合损失函数，包含多个分量从不同维度约束网络行为。该多分量设计解决以下关键挑战：（1）同时增强欠曝区域并抑制过曝区域；（2）处理含强光源的极端光照条件；（3）平衡亮度增强、伪影抑制与细节保持。

3.3.1.1 区域自适应重建损失

该损失项的核心思想是基于区域信息自适应调整光照图 $\mathbf{S}_2$ 的重建目标。

（1）暗区增强损失

对于暗区, 期望 $\mathbf{S}_2$ 能够显著增强图像亮度。目标亮度 $\mathbf{L}_{\text{dark}}^*$ 动态计算如下:

$$\mathbf{L}_{\text{dark}}^* = \mathbf{L}_2 \cdot \alpha \cdot \mathbf{W}_{\text{dark}} \quad (3.7)$$

其中, $\alpha = 0.5/(\mu_Y + \epsilon)$ 为增强因子, 由每幅输入图像的平均亮度 μ_Y 计算得到, $\epsilon = 10^{-9}$ 防止除零。暗区重建损失为:

$$\mathcal{L}_{\text{dark_rec}} = \|(\mathbf{S}_2 \odot \mathbf{W}_{\text{dark}}) - (\mathbf{L}_{\text{dark}}^* \odot \mathbf{W}_{\text{dark}})\|_2^2 \quad (3.8)$$

该损失项约束暗区的光照估计向目标亮度逼近, 增强因子 α 随输入图像亮度自适应调整, 避免对已较亮的图像过度增强。

(2) 亮区保持损失

对于亮区, 倾向于保持原始亮度而非过度增强。目标是使 $\mathbf{S}_2$ 在亮区接近去噪图像 $\mathbf{L}_2$ 的光照水平:

$$\mathcal{L}_{\text{bright_rec}} = \|(\mathbf{S}_2 \odot \mathbf{W}_{\text{bright}}) - (\mathbf{L}_2 \odot \mathbf{W}_{\text{bright}})\|_2^2 \quad (3.9)$$

该损失项确保亮区不被过度增强, 从而避免过曝和光效伪影的产生。

3.3.1.2 结构与高频细节保持损失

为确保增强图像的结构信息与高频细节得到保持, 引入以下损失项。

(1) 结构一致性损失

该损失采用输入图像 $\mathbf{I}_{\text{in}}$ 与增强输出 $\mathbf{I}_{\text{out}}$ 之间的 SSIM 计算:

$$\text{SSIM}(\mathbf{I}_{\text{in}}, \mathbf{I}_{\text{out}}) = \frac{(2\mu_1\mu_2 + C_1)(2\sigma_{12} + C_2)}{(\mu_1^2 + \mu_2^2 + C_1)(\sigma_1^2 + \sigma_2^2 + C_2)} \quad (3.10)$$

其中, μ_1 、 μ_2 为局部均值, σ_1^2 、 σ_2^2 为局部方差, σ_{12} 为局部协方差, C_1 、 C_2 为稳定常数。区域加权的结构损失为:

$$\mathcal{L}_{\text{struct}} = \mathbb{E} [(1 - \text{SSIM}(\mathbf{I}_{\text{in}}, \mathbf{I}_{\text{out}})) \odot (\mathbf{W}_{\text{dark}} + \lambda_s \mathbf{W}_{\text{bright}})] \quad (3.11)$$

其中 λ_s 为平衡因子。

该损失项确保增强过程中图像的整体结构信息不被破坏, 第二章已详细介绍 SSIM 指标的物理含义与计算方法, 此处不再赘述。

(2) 高频一致性损失

该损失利用 Laplacian 算子 ∇^2 提取输入图像 $\mathbf{I}_{\text{in}}$ 与增强输出 $\mathbf{I}_{\text{out}}$ 的高频分量, 并最小化其 L1 差异。区域权重侧重于暗区细节保持:

$$\mathcal{L}_{\text{hf}} = \mathbb{E} [|\nabla^2 \mathbf{I}_{\text{in}} - \nabla^2 \mathbf{I}_{\text{out}}| \odot (\lambda_{hf_d} \mathbf{W}_{\text{dark}} + \lambda_{hf_b} \mathbf{W}_{\text{bright}})] \quad (3.12)$$

其中 λ_{hf_d} 和 λ_{hf_b} 分别为暗区与亮区的权重因子。

高频分量包含图像的边缘、纹理等细节信息, 该损失项约束增强过程中这些细节不被模糊或丢失。

3.3.1.3 平滑性与正则化损失

为保持增强光照图 $\mathbf{S}_2$ 的视觉连贯性并避免伪影，引入以下损失项。

(1) 光照平滑性损失

借鉴文献^[76]的平滑损失和全变分损失 (Total Variation Loss)^[77]，约束生成的光照图 $\mathbf{S}_2$ 的平滑性，避免亮度突变：

$$\begin{aligned}\mathcal{L}_{\mathbf{S_smooth}} &= \text{SmoothLoss}(\mathbf{L}_2, \mathbf{S}_2) \\ \mathcal{L}_{\mathbf{S_tv}} &= \text{TV}(\mathbf{S}_2)\end{aligned}\quad (3.13)$$

第二章已述，光照分量在自然场景中通常呈平滑变化特性，该先验约束有助于生成符合物理规律的光照估计。

(2) 区域过渡平滑损失

这是一种新颖的损失项，专门确保光照图 $\mathbf{S}_2$ 在暗区与亮区过渡带保持平滑。首先通过比较 $\mathbf{W}_{\text{dark}}$ 与 $\mathbf{W}_{\text{bright}}$ 的差异识别区域边界，然后计算 $\mathbf{S}_2$ 在边界区域的梯度并最小化：

$$\mathcal{L}_{\text{trans_smooth}} = \mathbb{E}_{\text{boundary}}[|\nabla_x \mathbf{S}_2| + |\nabla_y \mathbf{S}_2|] \quad (3.14)$$

该损失项解决了区域自适应处理方法中常见的边界不连续问题，确保暗区增强与亮区抑制之间的自然过渡。

3.3.1.4 初始去噪一致性损失

为确保轻量化去噪网络 $\mathcal{F}_{\text{LDN}}$ 的有效性，其训练借鉴了 Zero-IG^[22] 中 LD-Net 的损失函数设计。

输入低光照图像 $\mathbf{I}_{\text{in}}$ 首先通过配对下采样操作 $G = (G_1, G_2)$ 分解为两幅独立的低分辨率图像 $I_1 = G_1(\mathbf{I}_{\text{in}})$ 和 $I_2 = G_2(\mathbf{I}_{\text{in}})$ (即图 3.1 中下采样模块的两个输出)。网络 $\mathcal{F}_{\text{LDN}}(\cdot)$ 旨在预测输入图像中的噪声。其损失函数 $\mathcal{L}_{\text{denoise}}$ 包含以下两部分：

(1) 残差损失

该损失鼓励网络学习能够解释两幅下采样图像差异的噪声：

$$\mathcal{L}_{\text{LD_res}} = \|I_1 - \mathcal{F}_{\text{LDN}}(I_1) - I_2\|_2^2 + \|I_2 - \mathcal{F}_{\text{LDN}}(I_2) - I_1\|_2^2 \quad (3.15)$$

其中 $\mathcal{F}_{\text{LDN}}(I_1)$ 和 $\mathcal{F}_{\text{LDN}}(I_2)$ 分别为网络对下采样图像 I_1 和 I_2 预测的噪声。

(2) 一致性损失

该损失确保不同尺度下去噪的一致性：

$$\begin{aligned}\mathcal{L}_{\text{LD_cons}} &= \|(I_1 - \mathcal{F}_{\text{LDN}}(I_1)) - G_1(\mathbf{I}_{\text{in}} - \mathcal{F}_{\text{LDN}}(\mathbf{I}_{\text{in}}))\|_2^2 \\ &\quad + \|(I_2 - \mathcal{F}_{\text{LDN}}(I_2)) - G_2(\mathbf{I}_{\text{in}} - \mathcal{F}_{\text{LDN}}(\mathbf{I}_{\text{in}}))\|_2^2\end{aligned}\quad (3.16)$$

其中 $\mathcal{F}_{\text{LDN}}(\mathbf{I}_{\text{in}})$ 为网络对原始输入图像 $\mathbf{I}_{\text{in}}$ 预测的噪声。

初始去噪网络的总损失为：

$$\mathcal{L}_{\text{denoise}} = \mathcal{L}_{\text{LD_res}} + \mathcal{L}_{\text{LD_cons}} \quad (3.17)$$

通过最小化该损失， $\mathcal{F}_{\text{LDN}}$ 网络学习去除噪声，输出低噪声水平的图像 $\mathbf{I}_2 = \mathbf{I}_{\text{in}} - \mathcal{F}_{\text{LDN}}(\mathbf{I}_{\text{in}})$ 供后续处理使用。

3.3.1.5 总损失函数

最终的总损失函数综合所有分量：

$$\begin{aligned} \mathcal{L}_{\text{total}} = & \lambda_1 \mathcal{L}_{\text{dark_rec}} + \lambda_2 \mathcal{L}_{\text{bright_rec}} + \lambda_3 \mathcal{L}_{\text{struct}} + \lambda_4 \mathcal{L}_{\text{hf}} \\ & + \lambda_5 \mathcal{L}_{\text{S_smooth}} + \lambda_6 \mathcal{L}_{\text{S_tv}} \\ & + \lambda_7 \mathcal{L}_{\text{trans_smooth}} + \lambda_8 \mathcal{L}_{\text{LD_res}} + \lambda_9 \mathcal{L}_{\text{LD_cons}} \end{aligned} \quad (3.18)$$

各损失权重设置为： $\lambda_1 = 700$ ， $\lambda_2 = 500$ ， $\lambda_3 = 200$ ， $\lambda_4 = 100$ ， $\lambda_5 = 5$ ， $\lambda_6 = 1600$ ， $\lambda_7 = 50$ ， $\lambda_8 = \lambda_9 = 1000$ 。这些权重通过系统实验确定，以在所有增强目标间达到最优平衡。

3.3.2 实现细节

在网络实现中，为确保数值稳定性，光照图 $\mathbf{S}_2$ 的取值被钳制到 $[\epsilon, 1]$ 范围内，其中 $\epsilon = 10^{-6}$ 为防止除零的小正常数。输入图像在送入网络前归一化到 $[0, 1]$ 范围。对于大尺寸图像，采用随机裁剪策略进行数据增强，裁剪块大小为 256×256 。

区域分割模块中的阈值 T_h 默认设为 0.5，平滑因子 f_s 设为 0.1。高斯平滑核大小为 5×5 ，标准差为 1.0。自适应光照增强器中的残差块数量设为 3，每个残差块包含两层 3×3 卷积。光效抑制器的编码器包含 4 层卷积，通道数分别为 32、64、64、32。

3.4 实验与结果分析

3.4.1 数据集与评价指标

3.4.1.1 实验数据集

为验证所提方法的泛化能力，在两个广泛使用的公开数据集（LOL-v2^[78]和LSRW^[79]）以及一个来自真实井下煤矿环境的私有数据集上进行实验。

LOL-v2 数据集于 2021 年提出，是低光照图像增强领域广泛使用的基准数据集 LOL 的扩展版本，包含两个子集：LOL-v2-Real 子集收录了 689 对真实场景低光照/正常光照配对图像（测试集 100 对），均通过调节相机曝光时间和 ISO 参数在室内家居、办公、走廊等场景中采集，分辨率为 400×600 ；LOL-v2-Synthetic 子集包含 900 对训练图像和 100 对测试图像，通过对高质量图像施加光照衰减

与噪声注入等操作合成生成，模拟了真实低光照条件下的亮度降低和传感器噪声特性。

LSRW 数据集于 2023 年提出，旨在提供更贴近实际使用场景的评测基准。该数据集使用消费级移动设备在自然环境中拍摄，包含两个设备子集：LSRW-Huawei 子集使用华为手机拍摄，包含 2480 对低光照/正常光照图像，涵盖城市街景、公园、建筑等场景；LSRW-Nikon 子集使用尼康相机拍摄，包含 3170 对图像，场景更加多样，包含室内外混合光照条件。与 LOL-v2 相比，LSRW 数据集的图像退化模式更加复杂，包含运动模糊、色彩偏移和噪声分布不均匀等真实场景中常见的退化特征。

私有井下煤矿数据集：从真实井下煤矿监控系统中采集了 3500 幅低光照图像（分辨率 600×400 ）。这些图像呈现强点光源、非均匀光照和复杂工业场景等独特挑战，与公开数据集中常见的室内/室外低光照场景存在显著差异。图 3.5 展示了私有数据集的典型样例，涵盖整体低光照、强点光和混合光照三种典型工况。

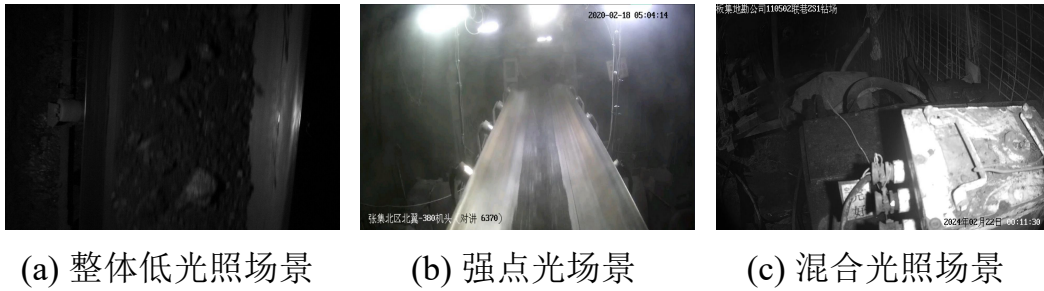


图 3.5 私有井下煤矿数据集典型样例：(a) 整体低光照场景；(b) 强点光场景；(c) 混合光照场景

3.4.1.2 评价指标

为全面评估增强效果，采用三种无参考图像质量评价指标：BRISQUE^[62]、NIQE^[61]和 PIQE^[63]。选择这些指标而非传统的 PSNR 或 SSIM 等全参考指标，原因如下：第一，私有煤矿数据集缺乏正常光照参考图像，需要无参考指标在无配对数据条件下评估图像质量；第二，本章采用的无监督训练框架关注感知质量而非像素级精度，感知质量指标更适合评估。第二章已详细介绍这些指标的定义与计算方法，此处仅需说明：三个指标均呈负相关特性，即数值越低表示图像质量越好。

3.4.2 对比方法与实验设置

3.4.2.1 对比基准方法

为全面评估所提方法的性能,选择六种代表性低光照图像增强方法作为对比基准,包括:RUAS^[10]、SCI^[19]、ZeroDCE++^[80]、EnlightenGAN^[18]、NeRCO^[21]和Zero-IG^[22]。所有对比实验均采用这些方法的官方开源代码实现,严格遵循其原始论文推荐的参数配置。实验在统一的硬件环境下进行,以确保实验结果的公平性和可重复性。

3.4.2.2 实验环境设置

实验采用 NVIDIA RTX 4090 GPU 作为计算平台,在 PyTorch 2.0.1 框架下实现算法。优化器配置为 Adam,学习率设为 0.0002,批次大小设为 1,训练轮数设为 2000。

3.4.3 定量结果分析

表 3.1总结了各方法在多个数据集上的定量对比结果。从表中可以看出,本章方法在大多数数据集和指标上均取得了优异性能。

表 3.1 多数据集上的图像质量指标定量对比(私有数据集、LOL-v2-Real、LOL-v2-Syn、LSRW-Huawei、LSRW-Nikon)。无监督方法中最优结果以粗体标注,次优结果以下划线标注

数据集	指标	RUAS CVPR2021	SCI CVPR2022	ZeroDCE++ TPAMI2021	EnlightenGAN TIP2022	NeRCO ICCV2023	Zero-IG CVPR2024	本章方法
私有数据集	BRISQUE↓	26.59	25.44	<u>18.38</u>	19.94	22.63	32.06	17.78
	NIQE↓	4.86	4.96	4.41	4.68	<u>4.23</u>	5.07	4.22
	PIQE↓	28.22	26.48	<u>25.24</u>	30.45	36.32	37.51	22.41
LOL-v2-Real	BRISQUE↓	20.85	32.51	38.35	17.42	18.68	28.15	<u>18.64</u>
	NIQE↓	7.47	8.05	8.18	7.12	<u>7.11</u>	7.96	7.09
	PIQE↓	24.62	40.00	47.48	28.15	29.42	<u>23.42</u>	23.17
LOL-v2-Syn	BRISQUE↓	38.60	22.81	22.89	18.62	<u>18.54</u>	38.04	18.14
	NIQE↓	6.51	4.87	4.83	<u>4.49</u>	4.50	5.08	4.47
	PIQE↓	50.04	46.49	43.67	40.42	36.35	48.12	<u>40.02</u>
LSRW-Huawei	BRISQUE↓	42.39	24.36	27.69	21.42	<u>21.36</u>	34.76	21.04
	NIQE↓	5.60	4.25	4.65	<u>4.13</u>	4.22	4.85	4.07
	PIQE↓	55.55	33.58	25.64	27.58	39.93	<u>27.55</u>	32.07
LSRW-Nikon	BRISQUE↓	45.37	28.70	30.27	24.94	<u>24.91</u>	35.39	21.11
	NIQE↓	5.13	3.94	3.61	4.53	<u>3.73</u>	4.18	4.26
	PIQE↓	59.75	33.84	34.12	33.94	47.89	<u>32.66</u>	32.17

在私有煤矿数据集上,本章方法在全部三项指标上均展现出卓越性能,相比次优方法分别取得了 BRISQUE 3.26%、NIQE 0.24%、PIQE 11.22% 的提升。这

些改进表明本章方法在处理含强点光源和非均匀光照的复杂真实场景时具有显著优势。

在公开基准数据集上,本章方法同样展现出竞争优势。在 LOL-v2-Real 数据集上,本章方法在 NIQE 和 PIQE 指标上取得最优,分别比次优方法提升 0.28% 和 1.07%,BRISQUE 指标位列第二。在 LOL-v2-Syn 数据集上,本章方法在 BRISQUE 和 NIQE 指标上取得最优,分别提升 2.16% 和 0.45%,PIQE 指标位列第二。在 LSRW-Huawei 数据集上,本章方法在 BRISQUE 和 NIQE 指标上取得最优,分别提升 1.50% 和 1.45%。在 LSRW-Nikon 数据集上,本章方法在 BRISQUE 和 PIQE 指标上取得最优,分别提升 15.25% 和 1.50%。

关于计算效率,表 3.2显示本章方法以 17.8ms 运行时间、0.156M 参数量、1.95G FLOPs 取得了具有竞争力的性能。虽然 RUAS 展现出最高效率(16.2ms 运行时间、0.004M 参数量、0.89G FLOPs),但如表 3.1所示,其图像质量指标明显较差。此外,尽管 SCI 展现较低的计算复杂度,但在其训练过程中普遍观察到训练崩溃现象,故出于公平性考虑将其从效率对比中排除。本章方法在增强质量与计算效率之间取得了最优平衡,仅比 RUAS 多 1.6ms 即可实现大幅度的质量提升,这一效率特性使其非常适合在安全关键应用中进行实时部署。

表 3.2 私有煤矿数据集上各方法的运行时间、参数量与计算量对比

方法	运行时间 (ms)	参数量 (M)	FLOPs (G)
RUAS	16.2	0.004	0.89
SCI	-	-	-
ZeroDCE++	18.4	0.167	2.05
EnlightenGAN	48.3	7.921	45.32
NeRCO	35.2	2.843	12.86
Zero-IG	31.5	0.327	4.82
本章方法	<u>17.8</u>	<u>0.156</u>	<u>1.95</u>

3.4.4 可视化分析

图 3.6和图 3.7展示了 LOL-v2 和 LSRW 数据集上代表性样本的定性对比结果。现有方法难以在不同光照条件下平衡亮度增强与细节保持。相比之下,本章方法在全局亮度提升与伪影抑制之间取得了更好的平衡,有效保持了明暗区域的纹理细节,同时维持了自然的视觉质量。例如,在 LOL-v2 数据集的图像上,本章方法在增强暗区墙壁可见度的同时提升了亮区书籍的可读性——无过曝现象、文字纹理保持良好、无颜色失真。这些结果表明本章方法在复杂光照场景中具有良好的泛化能力。

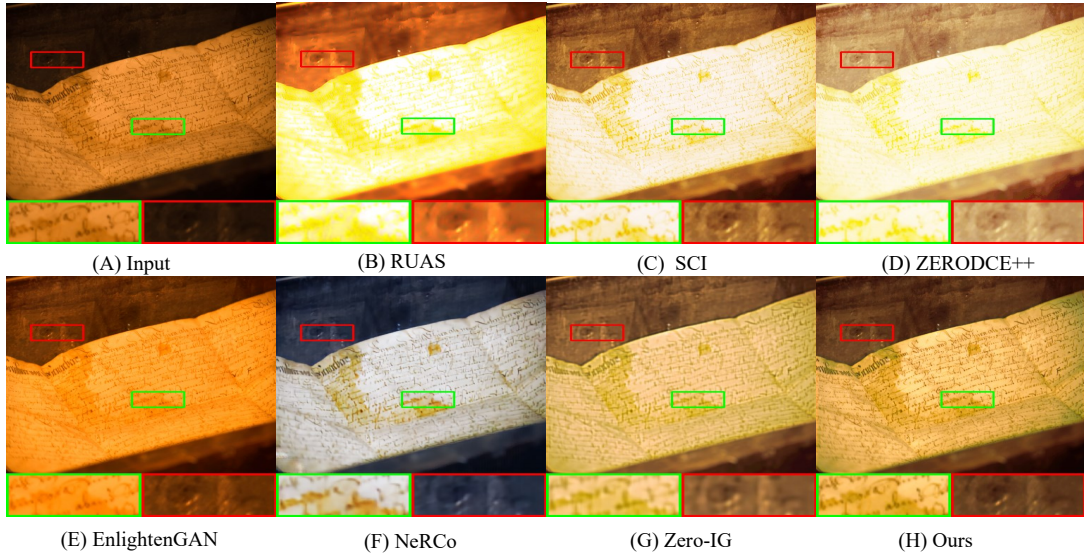


图 3.6 LOL-v2 数据集上的视觉对比

图 3.8和图 3.9展示了本研究构建的井下煤矿低光照数据集上代表性样本的定性对比结果。输入图像采集自真实采矿环境，具有强点光源和非均匀光照的特点。现有方法在此场景下面临诸多挑战：RUAS、SCI 和 ZeroDCE++ 在强光源周围产生明显的光晕效应并过度增强亮区；EnlightenGAN 生成不自然的颜色增强；NeRCO 和 Zero-IG 未能有效抑制强光源的扩散伪影，导致高亮度区域出现不自然的失真。而本章方法成功抑制了光晕效应，增强了暗区可见度，同时保持了整体自然的视觉质量，展现出优越的性能表现。

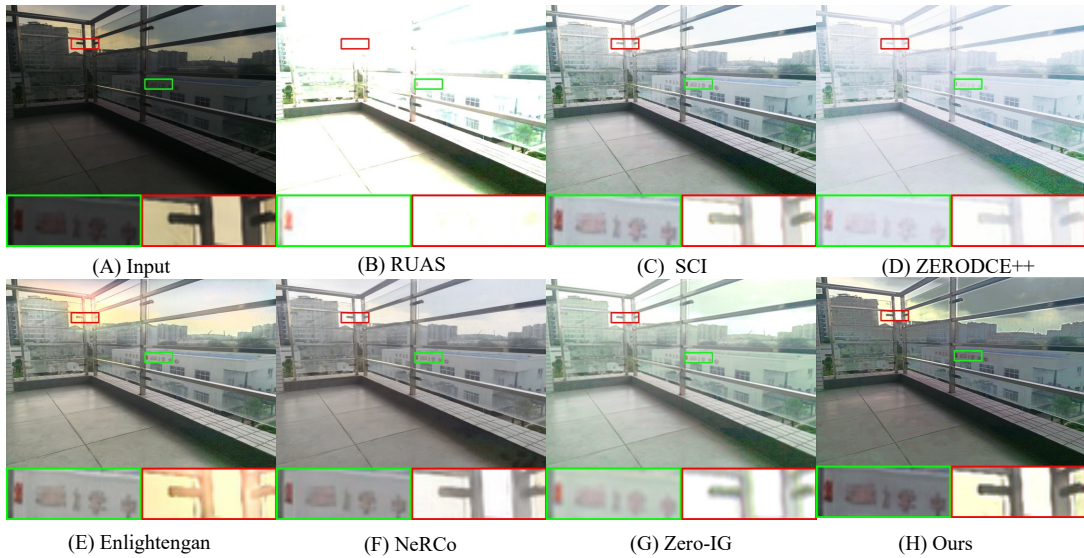


图 3.7 LSRW 数据集上的视觉对比

从视觉对比结果可以总结以下关键发现：第一，本章方法在暗区增强方面表现出色，能够有效提升原本不可见的暗区细节，同时避免噪声放大问题；第二，光效抑制模块有效削弱了强点光源引发的光晕效应，避免了过曝伪影的产生；第三，区域自适应处理策略使得暗区与亮区的过渡自然平滑，无明显边界伪影；第

四，增强后图像的颜色保持自然，未出现色偏或色彩失真现象。

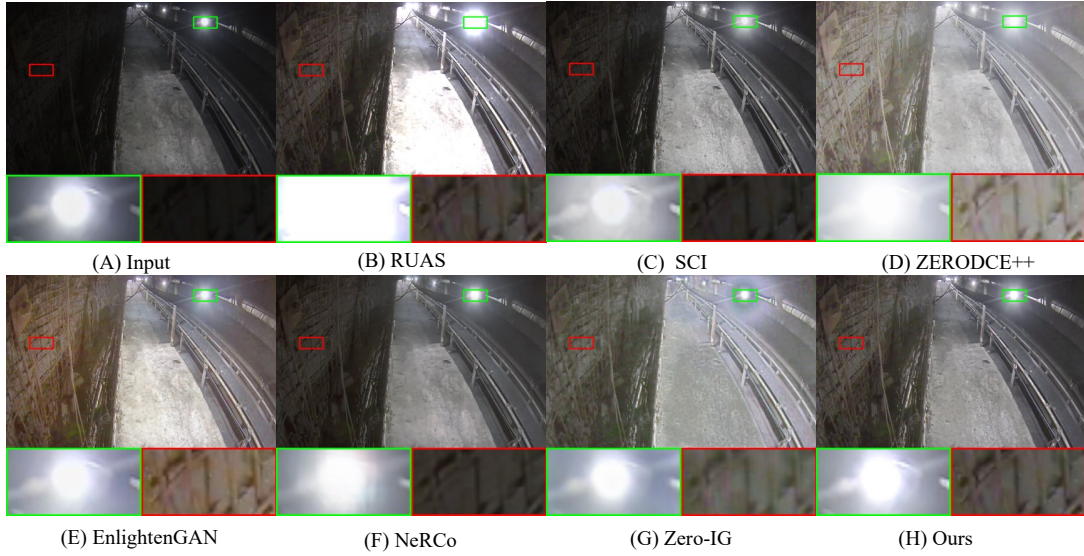


图 3.8 井下煤矿数据集上的视觉对比。从左至右依次为：输入图像、RUAS、SCI、ZeroDCE++、EnlightenGAN、NeRCO、Zero-IG、本章方法的结果，展示亮度增强与伪影抑制效果

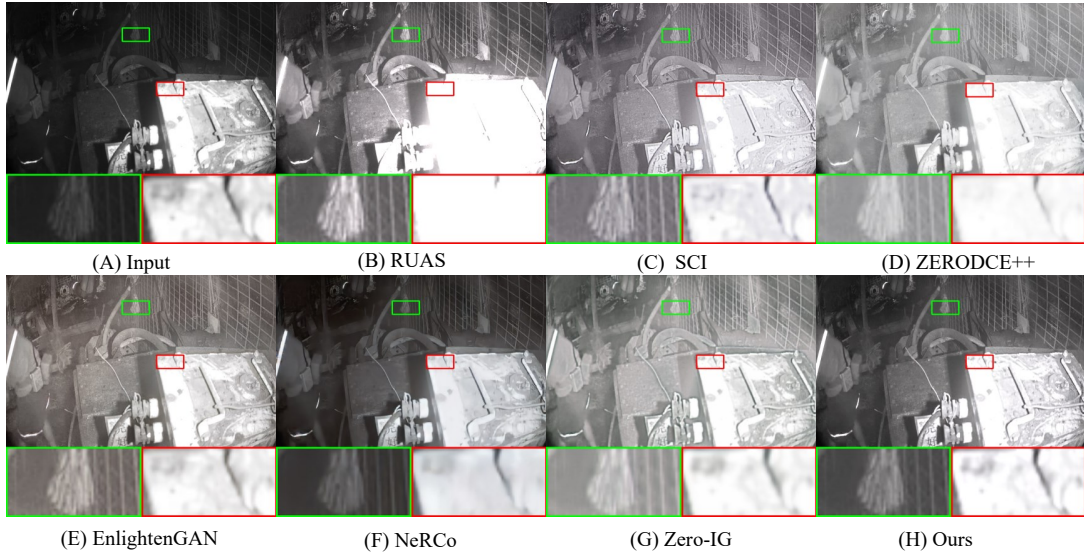


图 3.9 井下煤矿数据集上的视觉对比（补充样本）

3.4.5 消融实验

3.4.5.1 子网络消融实验

为评估各子网络的重要性，逐步添加自适应光照增强器（AIE）、区域分割模块（RSM）、轻量化去噪网络（LDN）和光效抑制器（LES）进行实验。表 3.3 给出了私有数据集和公开数据集上的定量结果，图 3.10 展示了定性对比。

首先评估在私有数据集上逐步添加子网络的效果。如表 3.3 所示，仅使用 AIE 时 BRISQUE 得分为 25.50、NIQE 为 4.43、PIQE 为 39.12，表明图像质量较差。

表 3.3 子网络消融实验。私有数据集与公开数据集上不同模块组合的图像质量指标定量对比

子网络				私有数据集			公开数据集		
AIE	RSM	LDN	LES	BRISQUE	NIQE	PIQE	BRISQUE	NIQE	PIQE
✓	×	×	×	25.50	4.43	39.12	22.94	4.52	41.92
✓	✓	×	×	20.97	4.65	32.37	20.55	4.56	41.80
✓	✓	✓	×	20.71	4.61	32.33	20.42	4.54	40.70
✓	✓	✓	✓	17.78	4.22	22.41	18.14	4.47	40.02

引入 RSM 后 BRISQUE 改善至 20.97、PIQE 改善至 32.37，图像质量提升，尽管 NIQE 略增至 4.65，说明 RSM 有助于区分暗区与亮区。添加 LDN 后 BRISQUE 略降至 20.71、NIQE 降至 4.61、PIQE 降至 32.33，展现出去噪效果。当所有模块集成时，BRISQUE 显著降至 17.78、NIQE 降至 4.22、PIQE 降至 22.41，图像质量大幅提升。

如图 3.10 所示，仅使用 AIE 导致严重的颜色偏差和失真；添加 RSM 后亮区出现过度增强和细节丢失；引入 LDN 减少了噪声但亮区过度增强问题仍存在；而完整模型有效增强了暗区、抑制了亮区过度增强、保持了细节与纹理。

在公开数据集上，消融实验同样验证了各子网络的作用。表 3.3 显示，仅使用 AIE 时 BRISQUE 为 22.94、NIQE 为 4.52、PIQE 为 41.92，图像质量欠佳。添加 RSM 后 BRISQUE 改善至 20.55、PIQE 改善至 41.80，尽管 NIQE 略增至 4.56。引入 LDN 后 BRISQUE 进一步降至 20.42、NIQE 降至 4.54、PIQE 降至 40.70，证明了去噪效果。所有模块集成时，BRISQUE 显著改善至 18.14、NIQE 改善至 4.47、PIQE 改善至 40.02，展现出大幅提升。这些发现凸显了 LES 在抑制亮区过度增强和保持细节方面的关键作用，而 LDN 则有效减少了噪声。

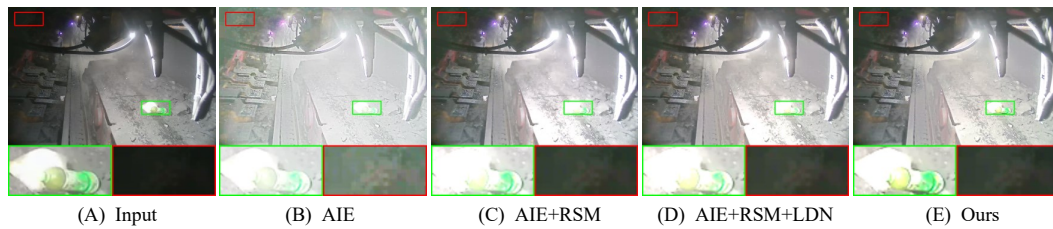


图 3.10 子网络消融实验。私有数据集上不同模块组合的视觉对比结果

3.4.5.2 损失函数消融实验

通过逐步添加区域自适应重建损失（RARL）、平滑性与正则化损失（SRL）、结构与高频细节保持损失（SHDPL）和初始去噪一致性损失（IDCL）来评估各损失分量的重要性。结果如表 3.4 和图 3.11 所示。

在私有数据集上，仅使用 RARL 导致较差的质量（BRISQUE: 27.32、NIQE:

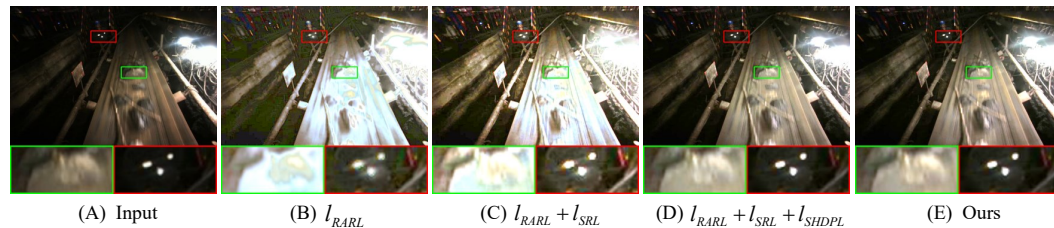
表 3.4 损失函数消融实验。私有数据集与公开数据集上不同损失分量的图像质量指标定量对比

损失函数				私有数据集			公开数据集		
RARL	SRL	SHDPL	IDCL	BRISQUE	NIQE	PIQE	BRISQUE	NIQE	PIQE
✓	×	×	×	27.32	4.62	38.72	63.95	9.88	67.92
✓	✓	×	×	26.93	4.56	35.82	52.54	8.52	63.47
✓	✓	✓	×	<u>19.83</u>	<u>4.31</u>	<u>26.93</u>	<u>20.21</u>	<u>4.63</u>	<u>41.65</u>
✓	✓	✓	✓	17.78	4.22	22.41	18.14	4.47	40.02

4.62、PIQE: 38.72)。添加 SRL 后各指标分别改善至 26.93、4.56、35.82。引入 SHDPL 后质量显著提升（19.83、4.31、26.93），而完整损失函数达到最优结果（17.78、4.22、22.41）。

如图 3.11 所示，仅使用 RARL 导致严重过曝、细节与纹理丢失、明显失真；添加 SRL 减少了失真并部分恢复纹理细节，但过曝问题仍存在；引入 SHDPL 恢复了纹理细节并抑制了过曝；而完整损失函数不仅解决了纹理细节和过曝问题，还有效降低了噪声。

在公开数据集上，仅使用 RARL 时 BRISQUE 为 63.95、NIQE 为 9.88、PIQE 为 67.92，图像质量极差。添加 SRL 后 BRISQUE 改善至 52.54、NIQE 改善至 8.52、PIQE 改善至 63.47。引入 SHDPL 后 BRISQUE 显著降至 20.21、NIQE 降至 4.63、PIQE 降至 41.65，展现出进一步改善。所有损失集成时，BRISQUE 降至 18.14、NIQE 降至 4.47、PIQE 降至 40.02，质量大幅提升。这些结果凸显了 SHDPL 在保持细节和抑制过曝方面的关键作用，IDCL 对降噪有显著贡献。

**图 3.11 损失函数消融实验。私有数据集上不同损失分量的视觉对比结果**

消融实验证实，所提框架中的每个子网络和损失函数分量对于实现高质量低光照图像增强都是不可或缺的。特别值得注意的是，LES 和 SHDPL 在抑制亮区过度增强和保持细节方面发挥着关键作用。这种协同机制确保了暗区的有效增强、亮区的自然抑制以及整体图像细节的保持。

3.5 本章小结

本章提出了一种新颖的区域自适应低光照图像增强框架，有效解决了复杂光照条件下的不自然伪影、过曝和细节丢失等问题。通过将基于 Retinex 理论的图像分解方法与区域差异化处理策略相结合，包括区域分割模块、自适应光照增强器、光效抑制器和复合损失函数，本章方法实现了亮度增强、伪影抑制和纹理保持的综合平衡。

在 LOL-v2、LSRW 和自建井下煤矿数据集上的广泛实验表明，本章方法在视觉质量方面展现出优越性能。与现有方法相比，在 BRISQUE、NIQE 和 PIQE 等指标上均取得了改进。消融实验证实了各组件在达成这些结果中的关键作用。

本章所提出的区域自适应增强框架不仅能够独立用于图像质量提升，其输出的光照图和区域权重图还可作为语义引导信息，为后续下游任务（如异物检测）提供有价值的先验知识。下一章将在此基础上，研究如何将增强模块的中间表示有效融入轻量化检测网络，实现图像增强与目标检测的协同优化。

第4章 面向煤矿皮带机的光照语义引导异物检测方法与实验验证

4.1 引言

煤矿井下皮带运输系统是矿井生产的核心环节，其安全稳定运行直接关系到煤矿的生产效率和人员安全。如第一章所述，大块矸石、金属锚杆、废弃木材及槽钢等非煤异物一旦混入煤流并随皮带高速运行，极易在溜槽、导料槽或托辊架等狭窄空间发生卡阻，进而刺穿并纵向撕裂输送带，造成严重的设备损失和安全事故^[4]。因此，研究一种能够实时、准确、自动检测皮带表面异物的视觉监测方法，对于保障煤矿安全生产具有重要意义。

然而，将通用目标检测算法直接应用于煤矿井下皮带机异物检测面临诸多特殊困难。首先，井下光照条件极为恶劣：整体照度极低导致图像对比度严重不足，强点光源产生的光斑和光晕效应造成局部过曝，使得暗区目标与背景难以区分、亮区目标轮廓信息丢失^[5]。其次，煤尘、岩粉等悬浮颗粒对光线的散射作用进一步降低了图像信噪比，呈现类似雾霾天气的退化特征。再次，异物目标本身具有特殊的形态特征：锚杆、钢筋等金属件呈细长杆状甚至弯曲形态，传统规则网格卷积难以对其几何结构进行有效建模。最后，井下边缘计算设备的算力、显存和功耗预算有限，难以支撑大规模复杂的深度神经网络模型。

第三章已详细阐述了区域自适应低光照图像增强框架的设计原理与实现方法。该框架通过区域分割模块输出暗区权重图 $\mathbf{W}_{\text{dark}}$ 和亮区权重图 $\mathbf{W}_{\text{bright}}$ ，通过自适应光照增强器和光效抑制器估计最终光照图 $\mathbf{S}_2$ ，最终输出亮度自然、细节清晰、伪影较少的增强图像 $\mathbf{I}_{\text{out}}$ 。本章的核心思想是：将增强过程中产生的中间表示（即光照图 $\mathbf{S}_2$ 和区域权重图 $\mathbf{W}_{\text{dark}}$ 、 $\mathbf{W}_{\text{bright}}$ ）视为一种光照区域语义编码，将皮带在图像中相对固定的位置分布及其由坐标注意力机制编码的空间位置信息视为空间位置信息语义，并将上述语义先验以引导权重的形式注入检测网络，对中间特征进行语义引导和重标定，从而在资源受限条件下实现检测精度、鲁棒性与实时性的综合平衡。

本章检测任务可形式化定义如下。给定输入图像 $\mathbf{I} \in \mathbb{R}^{H \times W \times 3}$ ，目标是预测图像中所有异物目标的位置和类别。设检测网络为 $\mathcal{F}_{\text{det}}$ ，其输出为一组检测结果 $\mathcal{D} = \{(\mathbf{b}_i, c_i, s_i)\}_{i=1}^N$ ，其中 $\mathbf{b}_i = (x_i, y_i, w_i, h_i)$ 为第 i 个目标的边界框坐标（中心点坐标及宽高）， $c_i \in \{1, 2, \dots, K\}$ 为类别标签， $s_i \in [0, 1]$ 为置信度分数， N 为检测到的目标数量， K 为异物类别总数。本章主要检测的异物类别包括矸石、锚杆、槽钢、木材等典型井下非煤异物。

本章的主要贡献包括：（1）构建融合图像增强模块的轻量化异物检测网络，在主干中采用 GhostNetV2 结构实现高效特征提取；（2）引入动态蛇形卷积增强对细长异物目标的几何建模能力；（3）设计坐标注意力与光照区域语义联合引导的特征重标定机制；（4）提出分阶段预训练与联合微调的训练策略，实现增强模块与检测模块的协同优化。

4.2 光照语义的定义与形式化表示

光照语义（Illumination Semantics）是指从图像增强过程中提取的、能够编码场景光照分布特性和区域属性的中间表示信息。与传统语义分割任务中基于物体类别的语义不同，光照语义所描述的并非目标的类别归属，而是图像中不同空间区域的光照状态属性——包括该区域属于暗区还是亮区、光照强度的空间分布梯度以及是否受到强光光源引起的光效伪影影响等。

具体地，给定输入低光照图像 $\mathbf{I}_{in} \in \mathbb{R}^{H \times W \times 3}$ ，经第三章区域自适应低光照增强模块 $\mathcal{F}_{\text{enhance}}$ 处理后，本文将光照语义定义为如下三元组：

$$\mathcal{S}_{\text{illu}} = (\mathbf{S}_2, \mathbf{W}_{\text{dark}}, \mathbf{W}_{\text{bright}}) \quad (4.1)$$

其中， $\mathbf{S}_2 \in \mathbb{R}^{H \times W \times 1}$ 为最终光照图，编码了场景中每个像素位置处的光照强度估计值，反映了光照分布的全局空间结构； $\mathbf{W}_{\text{dark}} \in [0, 1]^{H \times W}$ 为暗区权重图，其值越大表示该像素属于暗区的可能性越高，即该区域光照不足、特征提取越困难； $\mathbf{W}_{\text{bright}} \in [0, 1]^{H \times W}$ 为亮区权重图，其值越大表示该区域越可能存在过曝或光效伪影。由第三章区域分割模块的 Sigmoid 归一化设计可知，二者满足约束 $\mathbf{W}_{\text{dark}}(x, y) + \mathbf{W}_{\text{bright}}(x, y) = 1, \forall (x, y)$ 。

基于上述定义，本章所提出的光照语义引导是指将三元组 $\mathcal{S}_{\text{illu}}$ 作为先验条件注入检测网络的特征融合阶段，使检测网络在提取和融合特征时能够感知不同区域的光照状态，从而对暗区、亮区和正常光照区域的特征响应进行差异化调整。该引导过程可形式化为：

$$\hat{\mathbf{F}}_s = \text{LSR}(\mathbf{F}_s, \text{Resize}(\mathcal{S}_{\text{illu}}, H_s, W_s)), \quad s \in \{3, 4, 5\} \quad (4.2)$$

其中 $\mathbf{F}_s$ 为第 s 尺度的检测特征图， $\text{LSR}(\cdot)$ 为光照语义重标定模块（详见第 4.3.4 节）， H_s 、 W_s 为对应尺度的空间分辨率。

需要指出的是，光照语义引导机制与传统先增强、再检测的简单级联方式存在本质区别。在级联方式中，增强模块仅输出增强后的图像，检测网络对增强过程中的中间信息一无所知。而在本章的设计中，增强模块在估计光照分量的过程中隐式地学习了场景的光照分布结构，这些结构化信息通过 $\mathcal{S}_{\text{illu}}$ 显式传递给检测网络，使后者无需重新学习光照适应能力，而是直接利用增强模块的中间知识对检测特征进行重标定，从而在资源受限条件下实现更为高效的特征利用。

4.3 方法总体框架与关键模块设计

4.3.1 总体框架设计

本章构建的面向煤矿皮带机异物检测的整体算法框架如图 4.1 所示。该框架采用图像增强与轻量化目标检测的一体化设计思路，由图像增强模块、轻量化特征提取主干（Backbone）、多尺度特征融合网络（Neck）以及解耦检测头（Head）四个核心部分组成。

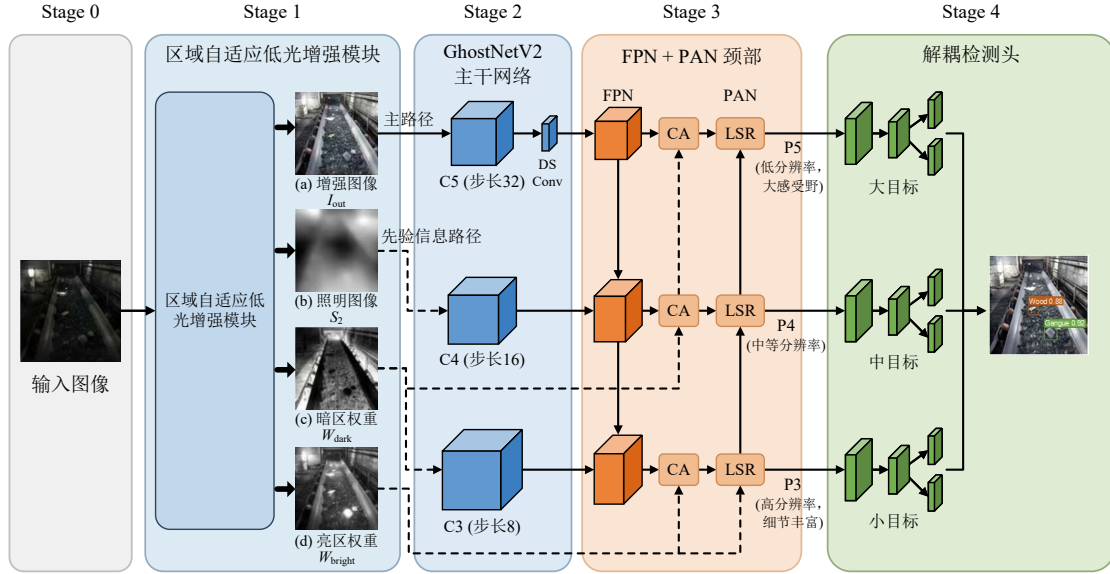


图 4.1 面向煤矿皮带机异物检测的整体网络架构

首先，煤矿井下采集的原始低光照图像 I_{in} 经第三章所述的区域自适应低光照图像增强模块处理，得到增强图像 I_{out} 以及光照图 S_2 、暗区权重图 W_{dark} 和亮区权重图 W_{bright} 等中间表示。这些中间表示蕴含了丰富的光照区域语义信息，将作为引导信号参与后续检测特征的重标定。

随后，增强图像 I_{out} 作为检测网络的输入，送入以 GhostNetV2 为主体构建的轻量化 Backbone，提取不同尺度的特征图 $\{C_3, C_4, C_5\}$ 。在 Backbone 深层引入动态蛇形卷积（DS Conv）模块，增强对锚杆等细长异物的几何建模能力。在 Neck 部分，采用 FPN 与 PAN 结构对多尺度特征进行融合与增强。在关键尺度上嵌入坐标注意力（CA）模块与光照语义重标定（LSR）模块，将光照图 S_2 和区域权重图下采样到各尺度特征尺寸后，与 CA 输出的位置注意力权重联合作用于中间特征，引导网络关注皮带表面及易发生异物堆积的关键区域。最后，Head 部分采用 YOLOv8 式解耦检测头，对各尺度特征图分别执行分类与边界框回归，实现对异物目标的定位与识别。

算法 4.1 给出了本文增强—检测协同框架的完整前向推理流程。给定一帧低光照输入图像，框架首先通过区域自适应低光照增强模块得到增强图像 I_{out} 与三类先验信息（照明图 S_2 、暗区权重 W_{dark} 、亮区权重 W_{bright} ）；随后以 I_{out} 作为

主输入，通过 GhostNetV2 主干提取多尺度特征 C_3 、 C_4 、 C_5 ，并将三类先验信息拼接重采样为多尺度引导图 G_3 、 G_4 、 G_5 ；颈部网络采用 FPN+PAN 双向结构，并在每个尺度上嵌入由 G_s 引导的坐标注意力（CA）与光照语义重标定（LSR）模块，最终输出 P_3 、 P_4 、 P_5 三个尺度的检测特征；解耦检测头分别对其进行小、中、大目标的预测，得到最终检测结果。与传统“先增强、再检测”的简单串联级联方式不同，本章将图像增强模块设计为可插拔的前端组件，并预留参数共享与联合微调接口。在训练阶段可选择以下两种模式：（1）固定增强模块参数，仅优化检测网络；（2）在总体损失稳定后对增强与检测两个子网络进行联合微调，使增强结果更贴合异物检测任务的需求。通过这种一体化设计，可在统一框架内同时兼顾图像质量提升与目标检测性能，减少信息在模块间传递时的损失。

算法 4.1 增强—检测协同框架前向推理流程

Require: 低光照输入图像 $I_{in} \in \mathbb{R}^{H \times W \times 3}$
Ensure: 检测结果 $\mathcal{D} = \{(b_i, c_i, s_i)\}$

- 1: 区域自适应低光照增强阶段:
- 2: $I_{out}, S_2, W_{dark}, W_{bright} = \mathcal{F}_{enhance}(I_{in})$ {调用第三章增强模块}
- 3: 多尺度特征提取阶段:
- 4: $C_3, C_4, C_5 = \text{GhostNetV2-Backbone}(I_{out})$ {以增强图像作为主输入}
- 5: $C_5 = \text{DSConv}(C_5)$ {动态蛇形卷积增强细长目标建模}
- 6: 多尺度光照-区域引导图构造:
- 7: **for** $s \in \{3, 4, 5\}$ **do**
- 8: $G_s = \text{Resize}(\text{Concat}(S_2, W_{dark}, W_{bright}), H_s, W_s)$ {将光照图与区域权重重采样至各尺度}
- 9: **end for**
- 10: 语义引导特征融合阶段:
- 11: $\tilde{C}_3, \tilde{C}_4, \tilde{C}_5 = \text{FPN}(C_3, C_4, C_5)$ {自顶向下传递语义信息}
- 12: **for** $s \in \{3, 4, 5\}$ **do**
- 13: $\tilde{C}_s = \text{CA}(\tilde{C}_s | G_s)$ {坐标注意力}
- 14: $\tilde{C}_s = \text{LSR}(\tilde{C}_s, G_s)$ {光照语义重标定}
- 15: **end for**
- 16: $P_3, P_4, P_5 = \text{PAN}(\tilde{C}_3, \tilde{C}_4, \tilde{C}_5)$ {自底向上传递定位信息}
- 17: 解耦检测输出阶段:
- 18: $\mathcal{D} = \text{DecoupledHead}(P_3, P_4, P_5)$ { $P_3/P_4/P_5$ 分别检测小/中/大目标}
- 19: **return** $\mathcal{D}$

4.3.2 轻量化主干网络：GhostNetV2

煤矿井下检测设备多部署于边缘端，受限于算力与功耗条件，检测网络必须在保证检测精度的前提下降低参数量与计算量。为此，本章在 YOLOv8 框架基础上采用 GhostNetV2^[38] 替代原始 CSP-C2f 主干结构，构建轻量化 Backbone。

GhostNetV2 的核心思想源于对卷积神经网络特征冗余性的观察：网络中存在大量相似的冗余特征图，这些“幻影”特征可通过廉价的线性变换从少量“真

实”特征中生成，而无需进行完整的卷积计算。如图4.2所示，Ghost模块首先通过常规卷积生成少量本征特征图，然后利用深度卷积等廉价操作生成幻影特征，最后将两者拼接得到完整输出。

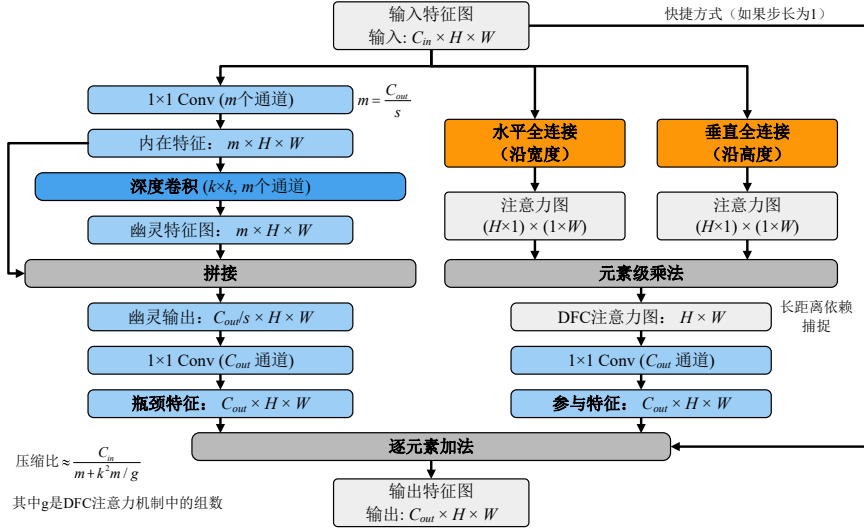


图 4.2 GhostNetV2 Bottleneck 结构图

设输入特征图通道数为 C_{in} ，期望输出通道数为 C_{out} ，Ghost 模块的计算过程可表示为：

$$\mathbf{Y}_{intrinsic} = \text{Conv}_{1 \times 1}(\mathbf{X}) \in \mathbb{R}^{H \times W \times m} \quad (4.3)$$

$$\mathbf{Y}_{ghost} = \text{DWConv}_{k \times k}(\mathbf{Y}_{intrinsic}) \in \mathbb{R}^{H \times W \times (C_{out} - m)} \quad (4.4)$$

$$\mathbf{Y} = \text{Concat}(\mathbf{Y}_{intrinsic}, \mathbf{Y}_{ghost}) \in \mathbb{R}^{H \times W \times C_{out}} \quad (4.5)$$

其中， m 为本征特征通道数，通常取 $m = C_{out}/s$ ， s 为压缩比（默认 $s = 2$ ），DWConv 表示深度可分离卷积。

与标准卷积相比，Ghost 模块的参数量压缩比约为：

$$r_{params} = \frac{C_{in} \times m + k^2 \times (C_{out} - m)}{k^2 \times C_{in} \times C_{out}} \approx \frac{1}{s} \quad (4.6)$$

GhostNetV2 在 Ghost 模块基础上引入了解耦全连接注意力（Decoupled Fully Connected Attention, DFC Attention）机制，增强了跨空间位置的特征交互能力。DFC 注意力通过沿水平和垂直方向分别进行全连接操作，捕获长距离依赖关系，其计算开销远低于标准自注意力机制。

在具体实现上，本章将 YOLOv8 Backbone 中的标准卷积模块替换为 Ghost-ConvV2，将原有的 C2f 模块替换为 C2f-GhostV2 结构。由于前端区域自适应低光照增强模块已提升了图像对比度和细节可见性，主干网络不再依赖过深、过宽的结构即可提取到具有较强判别能力的特征，形成前端增强与轻量特征提取的协同设计。表 4.1 对比了不同主干网络的参数量和计算量。

表 4.1 不同主干网络参数量与计算量对比

主干网络	参数量 (M)	FLOPs(G)	ImageNet Top-1(%)
YOLOv8-n Backbone	3.2	8.7	-
MobileNetV2	3.4	0.30	72.0
GhostNet	5.2	0.14	73.9
GhostNetV2	6.1	0.17	75.3
本章 GhostNetV2-Backbone	2.1	4.2	-

4.3.3 细长异物的局部形状建模：动态蛇形卷积

在煤矿皮带机异物中，锚杆、钢筋等金属物体往往呈细长、杆状甚至弯曲形态。传统规则网格卷积在固定感受野约束下，卷积采样点布置规则且刚性，容易出现采样位置与目标真实几何结构错位的问题，从而对细长目标的响应不够充分，影响小目标和细目标的召回率^[43]。

为增强网络对这类异物几何结构的感知能力，本章在 Backbone 深层引入动态蛇形卷积（Dynamic Snake Convolution, DS Conv）。如图 4.3所示，DS Conv 的核心思想是将标准规则卷积核的二维采样网格展平为一维序列，并在此基础上引入可学习的偏移量，对每个采样点的位置进行蛇形变形，使卷积核在空间上能够沿着潜在的细长结构自适应滑动。

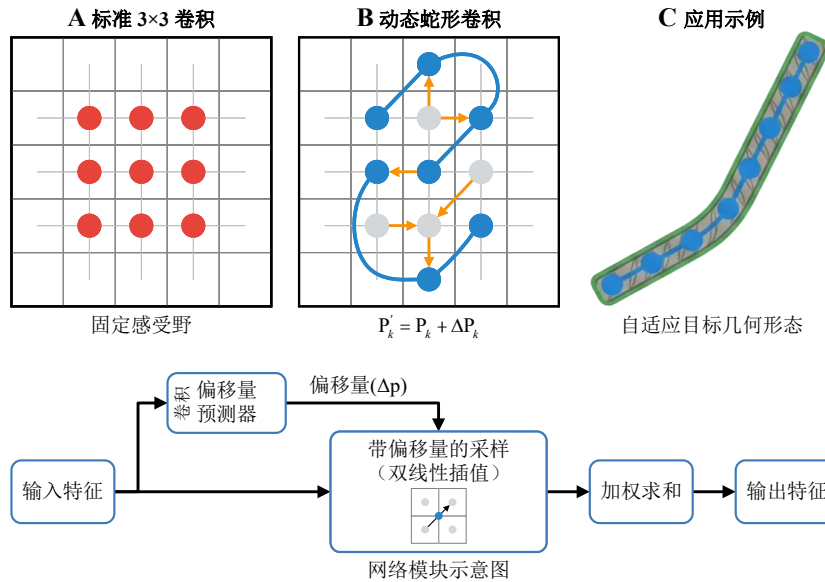


图 4.3 动态蛇形卷积（DS Conv）结构示意图

设标准 $K \times K$ 卷积在中心位置 $\mathbf{p}_0$ 处的采样点集合为 $\{\mathbf{p}_k\}_{k=1}^{K^2}$ ，DS Conv 首先将其展平为一维序列。对于沿 x 方向的蛇形变形，第 k 个采样点的位置更新为：

$$\mathbf{p}'_k = \mathbf{p}_k + \Delta \mathbf{p}_k \quad (4.7)$$

其中，偏移量 $\Delta \mathbf{p}_k$ 通过网络从输入特征中学习得到，并采用累积形式保证采样

点的连续性:

$$\Delta \mathbf{p}_k = \sum_{j=1}^k \delta_j \quad (4.8)$$

式中 δ_j 为相邻采样点之间的相对偏移。

基于更新后的采样坐标，DS Conv 的输出计算为:

$$\mathbf{y}(\mathbf{p}_0) = \sum_{k=1}^{K^2} w_k \cdot \mathbf{x}(\mathbf{p}'_k) \quad (4.9)$$

其中， w_k 为对应的卷积权重， $\mathbf{x}(\mathbf{p}'_k)$ 通过双线性插值从特征图中采样得到。

通过这种动态位置调整，DS Conv 可在局部区域内沿着目标的最长边方向持续聚焦，形成更符合细长目标轮廓的有效感受野覆盖。图 4.4 展示了 DS Conv 在细长异物目标上的采样轨迹可视化效果。

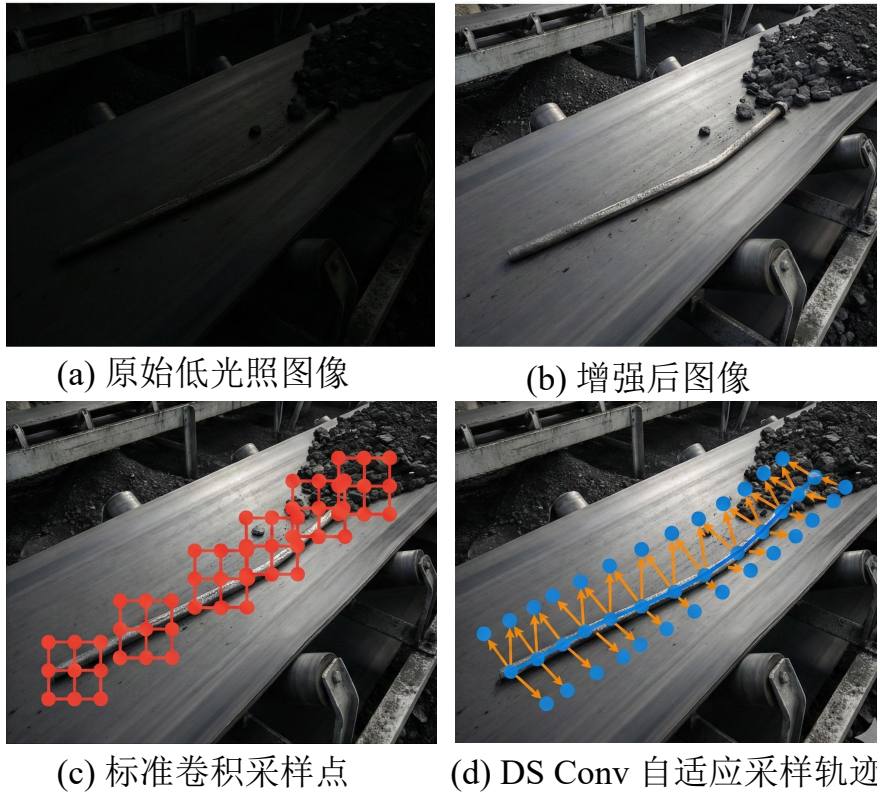


图 4.4 细长异物目标上 DS Conv 采样轨迹可视化: (a) 原始低光照图像; (b) 增强后图像; (c) 标准卷积采样点; (d) DS Conv 自适应采样轨迹

在网络集成方面，本章将 DS Conv 布置在负责高层语义信息的 C5 特征图之后，使其在分辨率相对较低但语义信息较丰富的特征图上对细杆状异物进行强化表征。一方面，该设计在较低分辨率下即可获得足够大的有效感受野，适合捕获长条形目标的整体形状；另一方面，结合 GhostNetV2 轻量化主干结构，DS Conv 引入的计算开销可控制在可接受范围内。

DS Conv 的优势与局限性分析 与标准卷积和可变形卷积相比，DS Conv 在细长目标建模方面的优势主要体现在三个方面。第一，几何自适应性强。DS Conv 通过式 (4.8) 中的累积偏移量机制保证了采样点的空间连续性，使卷积核能够沿目标轮廓形成连续的蛇形采样轨迹，相比可变形卷积中各采样点独立偏移可能导致的采样离散化问题，更适合捕获锚杆等具有连续细长结构的目标特征。第二，沿主轴方向的长距离依赖建模能力。通过将二维卷积核展平为一维序列并沿目标方向延伸，DS Conv 的有效感受野可沿细长目标的主轴方向显著扩展，在较低分辨率的 C5 特征图上即可覆盖完整的细长目标。第三，参数效率较高。DS Conv 的可学习参数仅来自偏移量预测网络（单层 3×3 卷积），相比可变形卷积需要预测 x 、 y 两个方向偏移量的开销更低。

同时，DS Conv 也存在以下局限性。第一，对紧凑型目标的增益有限。DS Conv 的蛇形采样策略主要针对高长宽比的细长目标设计，对于矸石等长宽比接近 1:1 的紧凑型目标，蛇形感受野相比标准网格卷积并无显著优势。第二，双线性插值带来的特征模糊。由于偏移后的采样坐标通常不落在整数网格点上，需要通过双线性插值获取特征值，这在一定程度上会模糊局部特征，对于边缘锐利度要求较高的场景可能产生影响。第三，训练收敛速度有所下降。偏移量的学习需要额外的训练迭代来建立有效的几何对应关系，实验中观察到引入 DS Conv 后模型达到收敛所需的训练轮数增加约 15%。

综合以上分析，DS Conv 特别适用于目标具有高长宽比细长形态（如管道、电缆、杆状结构）、目标可能呈弯曲或不规则蜿蜒形状、以及目标尺度较小但长度方向跨度大的检测场景。在本章的煤矿井下异物检测任务中，锚杆等细长金属异物恰好满足上述条件，因此 DS Conv 的引入带来了显著的检测性能提升。而对于矸石等紧凑型目标，其检测能力主要依赖 GhostNetV2 主干网络的基础特征提取以及 CA 与 LSR 语义引导机制的协同作用。

4.3.4 空间位置信息与光照区域语义的联合引导

4.3.4.1 坐标注意力机制

煤矿井下皮带表面区域在图像中的位置相对固定，异物往往出现在皮带中心带或边缘带，因此位置信息本身具有重要的参考价值。为充分利用这一空间位置先验并提高网络对目标分布的敏感度，本章在 Neck 部分引入坐标注意力（Coordinate Attention, CA）机制^[71]。

如图 4.5 所示，CA 模块首先对输入特征图 $\mathbf{F} \in \mathbb{R}^{C \times H \times W}$ 分别沿水平方向和

垂直方向进行全局平均池化，得到两条一维方向感知特征向量：

$$\mathbf{z}_c^h(h) = \frac{1}{W} \sum_{i=1}^W \mathbf{F}_c(h, i) \in \mathbb{R}^{C \times H \times 1} \quad (4.10)$$

$$\mathbf{z}_c^w(w) = \frac{1}{H} \sum_{j=1}^H \mathbf{F}_c(j, w) \in \mathbb{R}^{C \times 1 \times W} \quad (4.11)$$

这两条向量分别编码了特征在垂直和水平方向上的分布特性。

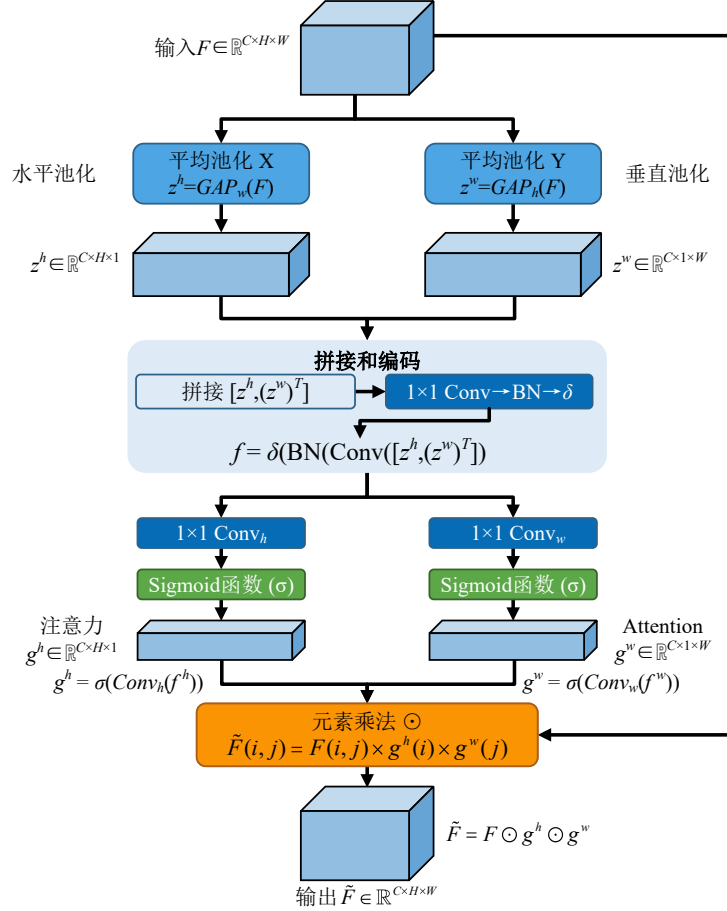


图 4.5 坐标注意力 (CA) 模块结构图

随后，两条方向特征拼接后通过共享的 1×1 卷积和非线性激活进行压缩与编码：

$$\mathbf{f} = \delta(\text{BN}(\text{Conv}_{1 \times 1}(\text{Concat}(\mathbf{z}^h, \mathbf{z}^w)))) \quad (4.12)$$

其中 δ 为非线性激活函数。编码特征 $\mathbf{f}$ 沿空间维度分割后，分别经过对应的 1×1 卷积和 Sigmoid 激活生成两个方向的注意力权重：

$$\mathbf{g}^h = \sigma(\text{Conv}_{1 \times 1}(\mathbf{f}^h)) \in \mathbb{R}^{C \times H \times 1} \quad (4.13)$$

$$\mathbf{g}^w = \sigma(\text{Conv}_{1 \times 1}(\mathbf{f}^w)) \in \mathbb{R}^{C \times 1 \times W} \quad (4.14)$$

最终，输入特征通过位置注意力权重进行重标定：

$$\tilde{\mathbf{F}} = \mathbf{F} \odot \mathbf{g}^h \odot \mathbf{g}^w \quad (4.15)$$

与传统仅在通道或空间维度上建模的注意力机制不同，CA 在保留较低参数量和计算开销的前提下，将空间位置信息显式嵌入到通道注意力中，特别适合于需要感知目标空间分布的检测任务。

4.3.4.2 光照语义引导特征重标定

为进一步利用增强模块提供的光照区域语义信息，本章设计了光照语义重标定（Light-Semantic Reweighting, LSR）模块。该模块将第三章输出的光照图 S_2 和区域权重图 W_{dark} 、 W_{bright} 作为引导信号，对检测网络的中间特征进行自适应调制。

如图 4.6 所示，LSR 模块的处理流程包括以下步骤。首先，构造多尺度引导图。将三通道引导信息拼接后下采样到与 Neck 各尺度特征图一致的空间尺寸：

$$\mathbf{G}_s = \text{Resize}(\text{Concat}(\mathbf{S}_2, \mathbf{W}_{\text{dark}}, \mathbf{W}_{\text{bright}}), H_s, W_s) \in \mathbb{R}^{3 \times H_s \times W_s} \quad (4.16)$$

其中下标 $s \in \{3, 4, 5\}$ 表示不同的特征尺度。

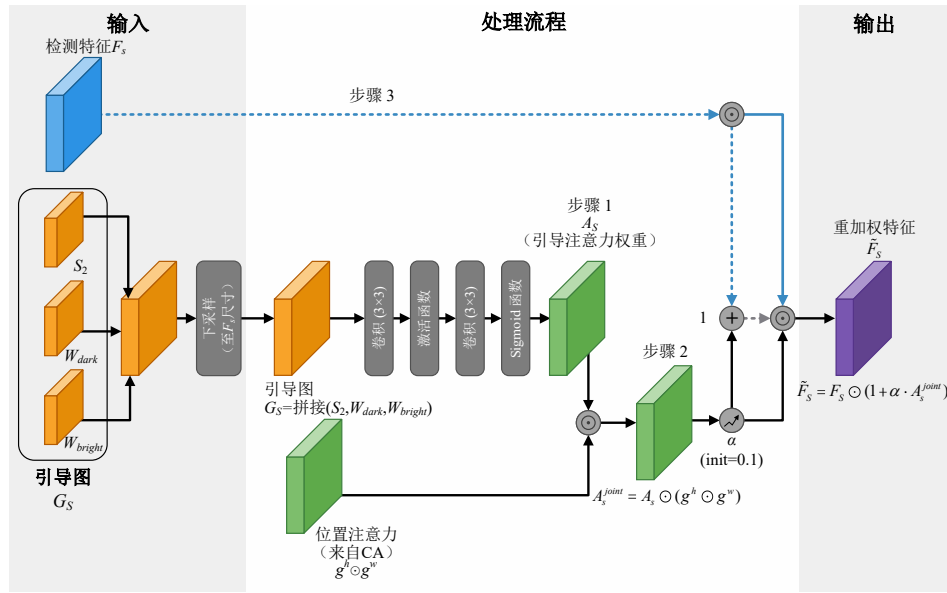


图 4.6 光照语义重标定（LSR）模块结构示意图

然后，通过轻量级卷积网络从引导图中生成引导权重：

$$\mathbf{A}_s = \sigma(\text{Conv}(\text{ReLU}(\text{Conv}(\mathbf{G}_s)))) \quad (4.17)$$

其中 σ 为 Sigmoid 激活函数，确保权重值落在 $[0, 1]$ 区间。

为实现坐标注意力与光照语义的联合引导，将 CA 模块输出的位置注意力权重与 LSR 生成的引导权重进行融合：

$$\mathbf{A}_s^{\text{joint}} = \mathbf{A}_s \odot (\mathbf{g}^h \odot \mathbf{g}^w) \quad (4.18)$$

最终，对输入特征进行自适应重标定：

$$\tilde{\mathbf{F}}_s = \mathbf{F}_s \odot (1 + \alpha \cdot \mathbf{A}_s^{\text{joint}}) \quad (4.19)$$

其中 α 为可学习的缩放因子，初始化为 0.1，控制引导强度。式 (4.19) 采用残差形式，确保在引导权重为零时特征保持不变，增强训练稳定性。

通过上述设计，暗区与亮区权重能够抑制远离皮带带面、光照极不均匀或噪声较重的背景区域响应，突出皮带表面及邻近区域；光照分布信息有助于网络区分真实高反射异物和由强点光产生的高亮伪影，在强光照区域抑制过度响应，降低误检概率。

4.4 损失函数与训练策略

4.4.1 检测损失函数

本章采用 YOLOv8 的标准检测损失函数设计，包含分类损失、置信度损失和边界框回归损失三个分量，总体损失函数为：

$$\mathcal{L}_{\text{det}} = \lambda_{\text{cls}}\mathcal{L}_{\text{cls}} + \lambda_{\text{obj}}\mathcal{L}_{\text{obj}} + \lambda_{\text{box}}\mathcal{L}_{\text{box}} \quad (4.20)$$

其中 λ_{cls} 、 λ_{obj} 、 λ_{box} 为各损失分量的平衡权重，默认分别取 0.5、1.0、7.5。

(1) 分类损失

采用二元交叉熵 (Binary Cross Entropy, BCE) 损失进行多标签分类：

$$\mathcal{L}_{\text{cls}} = -\frac{1}{N_{\text{pos}}} \sum_i \sum_c [y_{i,c} \log(\hat{y}_{i,c}) + (1 - y_{i,c}) \log(1 - \hat{y}_{i,c})] \quad (4.21)$$

其中 $y_{i,c}$ 为第 i 个目标的第 c 类真实标签， $\hat{y}_{i,c}$ 为预测概率， N_{pos} 为正样本数量。

针对井下异物检测中正负样本不平衡问题，引入 Focal Loss^[30] 进行缓解：

$$\mathcal{L}_{\text{cls}}^{\text{focal}} = -\frac{1}{N_{\text{pos}}} \sum_i \sum_c \alpha_c (1 - \hat{p}_{i,c})^\gamma \log(\hat{p}_{i,c}) \quad (4.22)$$

其中 $\hat{p}_{i,c} = y_{i,c}\hat{y}_{i,c} + (1 - y_{i,c})(1 - \hat{y}_{i,c})$ 为分类正确的概率， α_c 为类别平衡因子， γ 为聚焦参数（默认取 2.0）。

(2) 置信度损失

采用 BCE 损失约束目标存在性预测：

$$\mathcal{L}_{\text{obj}} = -\frac{1}{N} \sum_i [o_i \log(\hat{o}_i) + (1 - o_i) \log(1 - \hat{o}_i)] \quad (4.23)$$

其中 o_i 为真实目标置信度（正样本为 1，负样本为 0）， $\hat{o}_i$ 为预测置信度。

(3) 边界框回归损失

采用 CIoU 损失^[81] 进行边界框回归，该损失同时考虑重叠区域、中心点距离和长宽比一致性：

$$\mathcal{L}_{\text{box}} = 1 - \text{IoU} + \frac{\rho^2(\mathbf{b}, \mathbf{b}^{\text{gt}})}{c^2} + \alpha v \quad (4.24)$$

其中 IoU 为交并比， $\rho(\cdot)$ 为欧氏距离， $\mathbf{b}$ 和 $\mathbf{b}^{\text{gt}}$ 分别为预测框和真实框的中心点， c 为最小包围框的对角线长度， v 衡量长宽比一致性， α 为平衡系数。

4.4.2 分阶段预训练与联合微调策略

为充分利用增强模块与检测模块的协同效应，本章采用分阶段预训练与联合微调的训练策略，如算法 4.2所示。

算法 4.2 分阶段预训练与联合微调训练策略

Require: 训练数据集 $\mathcal{D}_{\text{train}}$ ，预训练增强模块参数 θ_{enh}^0
Ensure: 优化后的增强模块参数 θ_{enh}^* ，检测模块参数 θ_{det}^*

- 1: // 第一阶段：固定增强模块，预训练检测网络
- 2: 初始化检测网络参数 θ_{det} （可使用 COCO 预训练权重）
- 3: 冻结增强模块参数 $\theta_{\text{enh}} \leftarrow \theta_{\text{enh}}^0$
- 4: **for** epoch = 1 to E_1 **do**
- 5: **for** each batch ($\mathbf{I}, \mathbf{Y}$) in $\mathcal{D}_{\text{train}}$ **do**
- 6: $\mathbf{I}_{\text{out}}, \mathbf{S}_2, \mathbf{W}_{\text{dark}}, \mathbf{W}_{\text{bright}} = \mathcal{F}_{\text{enh}}(\mathbf{I}; \theta_{\text{enh}})$
- 7: $\mathcal{D} = \mathcal{F}_{\text{det}}(\mathbf{I}_{\text{out}}, \mathbf{S}_2, \mathbf{W}_{\text{dark}}, \mathbf{W}_{\text{bright}}; \theta_{\text{det}})$
- 8: $\mathcal{L} = \mathcal{L}_{\text{det}}(\mathcal{D}, \mathbf{Y})$
- 9: 更新 θ_{det} （仅优化检测网络）
- 10: **end for**
- 11: **end for**
- 12: // 第二阶段：联合微调增强与检测模块
- 13: 解冻增强模块参数
- 14: 降低学习率: $\text{lr} \leftarrow \text{lr} \times 0.1$
- 15: **for** epoch = 1 to E_2 **do**
- 16: **for** each batch ($\mathbf{I}, \mathbf{Y}$) in $\mathcal{D}_{\text{train}}$ **do**
- 17: $\mathbf{I}_{\text{out}}, \mathbf{S}_2, \mathbf{W}_{\text{dark}}, \mathbf{W}_{\text{bright}} = \mathcal{F}_{\text{enh}}(\mathbf{I}; \theta_{\text{enh}})$
- 18: $\mathcal{D} = \mathcal{F}_{\text{det}}(\mathbf{I}_{\text{out}}, \mathbf{S}_2, \mathbf{W}_{\text{dark}}, \mathbf{W}_{\text{bright}}; \theta_{\text{det}})$
- 19: $\mathcal{L} = \mathcal{L}_{\text{det}}(\mathcal{D}, \mathbf{Y}) + \lambda_{\text{enh}} \mathcal{L}_{\text{enh}}(\mathbf{I}_{\text{out}}, \mathbf{I})$
- 20: 联合更新 θ_{enh} 和 θ_{det}
- 21: **end for**
- 22: **end for**
- 23: **return** $\theta_{\text{enh}}^* \leftarrow \theta_{\text{enh}}, \theta_{\text{det}}^* \leftarrow \theta_{\text{det}}$

该策略分为两个阶段：

第一阶段（固定增强，预训练检测）：使用第三章预训练好的增强模块参数，冻结其权重，仅优化检测网络。训练轮数 $E_1 = 100$ ，学习率采用余弦退火策略，初始值为 0.01。该阶段使检测网络适应增强模块的输出特性，学习从增强图像和语义引导信息中提取有效的检测特征。

第二阶段（联合微调）：解冻增强模块参数，以较低的学习率（原学习率的 0.1 倍）对增强与检测两个子网络进行联合优化。训练轮数 $E_2 = 50$ ，总损失为检测损失与增强损失的加权和，权重 $\lambda_{\text{enh}} = 0.1$ 。该阶段使增强模块根据检测任务的反馈进行适应性调整，生成更有利于检测的增强结果。

表 4.2总结了各阶段的训练超参数配置。

表 4.2 训练超参数配置

超参数	第一阶段	第二阶段
训练轮数	100	50
初始学习率	0.01	0.001
学习率策略	余弦退火	余弦退火
批次大小	16	8
优化器	SGD	SGD
动量	0.937	0.937
权重衰减	0.0005	0.0005
增强模块状态	冻结	解冻
增强损失权重 λ_{enh}	-	0.1

4.5 实验设置与结果分析

4.5.1 数据集与评价指标

4.5.1.1 自建煤矿皮带机异物数据集

本章实验基于自建的煤矿皮带机异物数据集。该数据集采集自真实煤矿井下环境，涵盖不同光照条件、不同工况以及多种异物类型。数据集包含 4 类典型异物目标：矽石（Gangue）、锚杆（Anchor）、槽钢（Channel）和木材（Wood）。

该数据集的构建过程如下。在数据采集阶段，于某煤矿井下主运输巷道中部署 4 台工业防爆摄像头（海康威视 DS-2XE6222F-IS，200 万像素，IP68 防护等级），分别安装在皮带机头部、中段和尾部位置，采集 2024 年 3 月至 8 月期间的连续视频数据。在帧提取与筛选阶段，以 2 帧/秒的频率从原始视频流中抽帧，由人工筛选包含异物目标的有效帧，剔除大量纯煤流帧和画面模糊帧。在标注阶段，采用 LabelImg 标注工具，由 3 名具有井下工作经验的标注人员按照 PASCAL VOC 格式对矽石、锚杆、槽钢和木材四类异物目标进行边界框标注，标注完成后由第 4 名人员进行交叉审核以确保标注一致性。在数据划分阶段，按 7:2:1 的比例将数据集划分为训练集（5803 幅）、验证集（1658 幅）和测试集（829 幅），划分时确保同一摄像头采集的连续帧不跨越训练集和测试集，以避免数据泄露。

图 4.7 展示了数据集中四类异物目标典型样例。

表 4.3 给出了数据集的详细统计信息。数据集按照 7:2:1 的比例划分为训练集、验证集和测试集。为评估算法在不同光照条件下的鲁棒性，测试集进一步按照光照工况细分为正常光照、低光照和强点光三个子集。

表 4.4 展示了测试集按光照工况的划分情况。

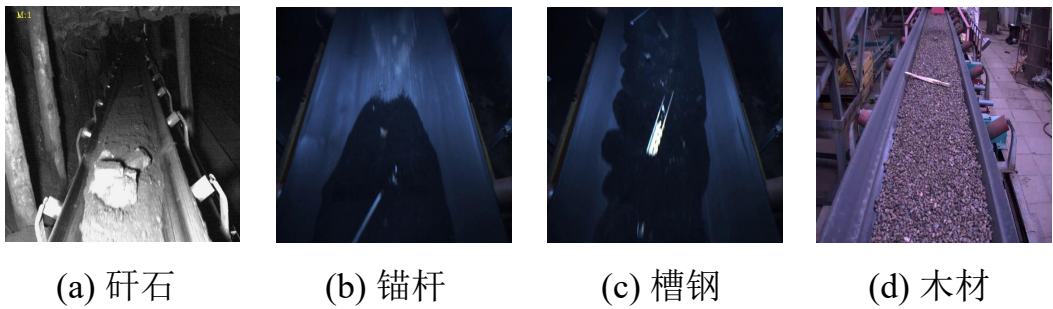


图 4.7 自建数据集四类异物目标典型样例：(a) 矸石；(b) 锚杆；(c) 槽钢；(d) 木材

表 4.3 自建煤矿皮带机异物数据集统计

类别	训练集	验证集	测试集	总计
矸石（Gangue）	2763	771	388	3922
锚杆（Anchor）	1916	535	269	2720
槽钢（Channel）	1286	358	180	1824
木材（Wood）	1463	408	206	2077
标注框总计	7428	2072	1043	10543
图像数量	5803	1658	829	8290

表 4.4 测试集光照工况划分

光照工况	图像数量	目标数量	特征描述
正常光照	312	398	照度适中，对比度良好
低光照	285	356	整体偏暗，噪声明显
强点光	232	289	局部过曝，光晕效应
合计	829	1043	-

4.5.1.2 评价指标

本章采用目标检测领域的标准评价指标评估算法性能，具体包括：

(1) 精确率 (Precision) 与召回率 (Recall)

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4.25)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4.26)$$

其中 TP、FP、FN 分别表示真正例、假正例和假负例数量。

(2) 平均精度 (Average Precision, AP) 与平均精度均值 (mAP)

$$\text{AP} = \int_0^1 p(r) dr \quad (4.27)$$

其中 $p(r)$ 为精确率-召回率曲线。mAP 为所有类别 AP 的均值。本章报告 mAP@0.5 (IoU 阈值 0.5) 和 mAP@0.5:0.95 (IoU 阈值从 0.5 到 0.95，步长 0.05 取均值) 两个指标。

(3) F1 分数

$$\text{F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4.28)$$

(4) 效率指标

包括推理帧率 (FPS)、模型参数量 (Params) 和浮点运算量 (FLOPs)，用于评估算法的实时性和部署友好性。

4.5.2 对比方法与实现细节

4.5.2.1 对比方法

为全面评估本章方法的有效性，选取以下基线方法进行对比：

(1) **原始 YOLO 系列**：YOLOv8-n、YOLOv8-s^[82]，代表当前主流的单阶段检测器。

(2) **轻量化检测器**：YOLOv8-n + MobileNetV2^[36]、YOLOv8-n + ShuffleNetV2^[39]，评估不同轻量化主干的效果。

(3) **增强 + 检测级联**：先使用第三章增强方法处理图像，再送入原始 YOLOv8-n 检测，评估简单级联的效果。

(4) **本章方法**：完整的语义引导检测框架，包括 GhostNetV2 主干、DS Conv、CA 和 LSR 模块。

4.5.2.2 实现细节

所有实验在配备 NVIDIA RTX 4090 GPU 的服务器上进行，使用 PyTorch 2.0.1 框架实现。输入图像统一缩放至 640×640 分辨率。数据增强策略包括：随机亮

度/对比度抖动（模拟井下光照变化）、随机高斯噪声添加（模拟煤尘干扰）、随机运动模糊（模拟皮带高速运行）、Mosaic 增强和随机水平翻转。具体配置如表 4.5所示。

表 4.5 实验实现细节

配置项	设置
输入分辨率	640 × 640
批次大小	16（第一阶段）/ 8（第二阶段）
优化器	SGD（momentum=0.937，weight_decay=0.0005）
学习率策略	余弦退火，warmup 3 epochs
训练轮数	100（第一阶段）+ 50（第二阶段）
数据增强	亮度抖动、噪声添加、运动模糊、Mosaic、水平翻转
NMS 阈值	IoU=0.45，置信度 =0.25

4.5.3 定量结果分析

4.5.3.1 主结果对比

表 4.6展示了各方法在自建煤矿皮带机异物数据集上的主要对比结果。

表 4.6 自建煤矿皮带机异物数据集主结果对比

方法	mAP@0.5	mAP@0.5:0.95	Precision	Recall	Params(M)	FPS
YOLOv8-n	78.3	52.1	82.6	74.5	3.2	215
YOLOv8-s	81.7	55.8	85.2	77.3	11.2	142
YOLOv8-n + MobileNetV2	76.8	50.4	81.3	72.9	2.8	198
YOLOv8-n + ShuffleNetV2	75.2	48.9	79.8	71.6	2.1	226
增强 + YOLOv8-n（级联）	82.5	56.3	85.8	78.1	3.4	168
本章方法	87.6	61.2	89.3	84.7	2.9	186

从表 4.6可以得出以下结论：（1）本章方法在 mAP@0.5 指标上达到 87.6%，相比基线 YOLOv8-n 提升 9.3 个百分点，相比更大规模的 YOLOv8-s 提升 5.9 个百分点，验证了所提出框架的有效性；（2）本章方法在召回率上显著优于其他方法，达到 84.7%，表明语义引导机制有效降低了漏检率；（3）在参数量和计算效率方面，本章方法仅有 2.9M 参数，推理速度达到 186 FPS，在精度与效率之间取得了良好平衡；（4）简单的增强 + 检测级联方案相比原始 YOLOv8-n 有一定提升，但仍低于本章方法，说明一体化设计和语义引导机制的必要性。

4.5.3.2 不同光照工况下的结果

为验证光照语义引导机制的有效性，表 4.7对比了各方法在不同光照工况子集上的检测性能。

表 4.7 不同光照工况下 mAP@0.5 对比

方法	正常光照	低光照	强点光	平均
YOLOv8-n	85.6	68.2	71.4	78.3
YOLOv8-s	87.9	73.5	75.8	81.7
增强 + YOLOv8-n（级联）	88.1	76.3	78.9	82.5
本章方法	90.2	84.1	85.6	87.6

从表 4.7可以观察到：（1）在正常光照条件下，各方法性能差异相对较小，本章方法达到 90.2%，仅比 YOLOv8-s 高 2.3 个百分点，说明光照良好时基线方法即可取得较好效果；（2）在低光照条件下，原始 YOLOv8-n 性能显著下降至 68.2%，而本章方法仍保持 84.1%，相比基线提升 15.9 个百分点，充分体现了增强模块与语义引导机制的协同作用；（3）在强点光条件下，本章方法达到 85.6%，相比 YOLOv8-n 提升 14.2 个百分点，验证了光照区域语义对于抑制过曝区域误检的有效性；（4）本章方法在三种光照工况下的性能波动最小（标准差约 3.2%），表现出最强的鲁棒性。

4.5.4 消融实验

为验证各关键模块的贡献，本章进行了系统的消融实验。

4.5.4.1 模块消融

表 4.8展示了逐步添加各模块对检测性能的影响。

表 4.8 模块消融实验

Baseline	GhostNetV2	DS Conv	CA	LSR	mAP@0.5	Params(M)	FLOPs(G)
✓					78.3	3.2	8.7
✓	✓				78.9	2.1	4.2
✓	✓	✓			82.4	2.3	4.8
✓	✓	✓	✓		84.8	2.5	5.2
✓	✓	✓	✓	✓	87.6	2.9	5.8

图 4.8以柱状图形式展示了消融实验的 mAP 变化趋势。

从消融实验结果可以分析各模块的作用：（1）引入 GhostNetV2 主干后，参数量从 3.2M 降至 2.1M，FLOPs 从 8.7G 降至 4.2G，降幅分别达到 34.4% 和 51.7%，

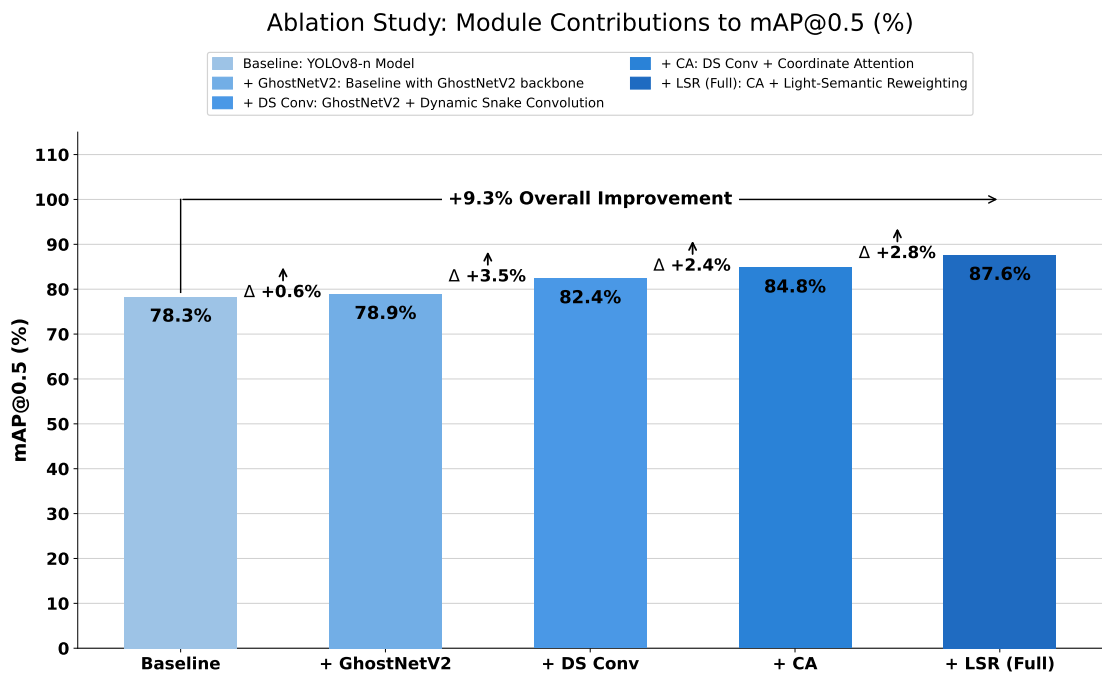


图 4.8 模块消融实验 mAP 变化柱状图

而 mAP 从 78.3% 提升至 78.9%，验证了轻量化设计的有效性；（2）DS Conv 模块带来 3.5 个百分点的 mAP 提升（从 78.9% 到 82.4%），对细长异物（如锚杆）的检测提升尤为明显，其 AP 从 71.2% 提升至 78.6%，验证了动态蛇形卷积对几何建模的有效性；（3）CA 模块通过编码位置信息使 mAP 从 82.4% 提升至 84.8%，增幅 2.4 个百分点，表明空间位置先验对于皮带区域关注的重要性；（4）LSR 模块通过引入光照语义引导使 mAP 从 84.8% 提升至最终的 87.6%，增幅 2.8 个百分点，尤其在极端光照条件下表现突出，低光照子集 mAP 提升 4.1 个百分点，强光光子集提升 3.8 个百分点。

4.5.4.2 训练策略消融

表 4.9对比了不同训练策略的效果。

表 4.9 训练策略消融实验

训练策略	mAP@0.5	训练时间 (h)
仅检测网络单独训练	82.1	6.5
增强模块冻结 + 检测训练	85.3	8.2
增强 + 检测联合训练（从头）	84.6	14.8
分阶段预训练 + 联合微调（本章）	87.6	12.3

结果表明，分阶段预训练与联合微调策略能够有效平衡训练效率和最终性能，先让检测网络适应增强模块的输出特性，再通过联合微调实现协同优化，获

得最优的检测效果。具体而言：（1）仅训练检测网络而不使用增强模块时，mAP 为 82.1%，作为参照基线；（2）冻结增强模块参数后训练检测网络可获得 85.3% 的 mAP，说明增强模块对检测性能有显著贡献；（3）从头联合训练反而导致性能下降至 84.6%，可能是由于两个子网络在训练初期相互干扰，难以找到全局最优解；（4）本章采用的分阶段策略达到最高的 87.6% mAP，同时训练时间相比从头联合训练减少约 17%。

4.5.5 可视化分析

4.5.5.1 不同光照条件下的检测可视化

图 4.9 展示了不同光照条件下的检测结果可视化。

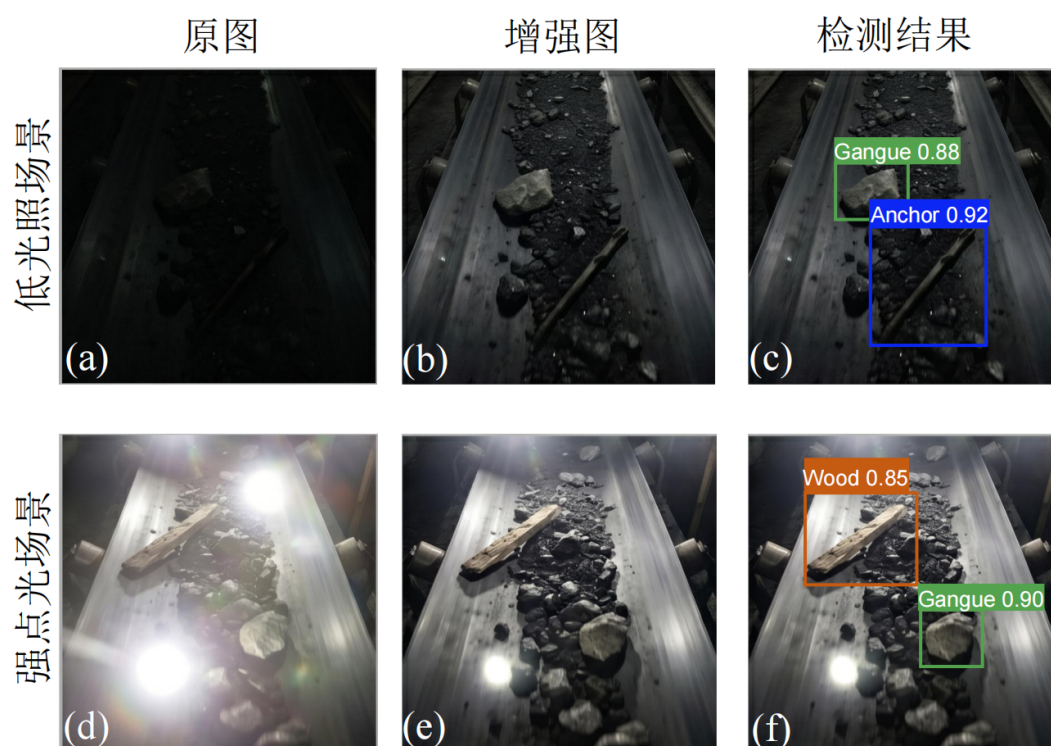


图 4.9 不同光照条件下检测可视化：(a)(b)(c) 低光照场景；(d)(e)(f) 强点光场景。每组包括原图、增强图、检测结果

从图 4.9 可以观察到：（1）在低光照场景中，原始图像几乎无法辨识异物目标，经增强后异物轮廓清晰显现，本章方法能够准确检出；（2）在强点光场景中，原图存在严重的光晕效应，增强模块有效抑制了光斑伪影，避免了因过曝导致的误检。

4.6 本章小结

本章提出了一种面向煤矿皮带机的光照语义引导异物检测方法，主要贡献包括以下几个方面：

(1) 构建了增强—检测一体化框架。将第三章的区域自适应低光照图像增强模块与轻量化检测网络有机融合,实现图像质量提升与目标检测的协同优化。增强模块可作为可插拔前端组件,支持固定或联合微调两种训练模式。

(2) 设计了轻量化检测网络。采用 GhostNetV2 构建高效 Backbone,在大幅降低参数量和计算量的同时保持检测精度,满足边缘设备部署需求。

(3) 引入动态蛇形卷积增强细长目标建模。通过可学习的采样点偏移使卷积核能够沿锚杆等细长异物的轮廓自适应滑动,显著提升了此类难检目标的召回率。

(4) 提出光照语义引导特征重标定机制。将增强模块输出的光照图和区域权重图作为语义引导信号,结合坐标注意力机制对检测特征进行自适应重标定,增强网络对皮带表面关键区域的关注,有效缓解复杂光照条件下的漏检误检问题。

(5) 设计了分阶段预训练与联合微调的训练策略。先固定增强模块训练检测网络,再以低学习率联合优化,实现两个子网络的协同提升。

实验结果表明,本章方法在自建煤矿皮带机异物数据集上达到 87.6% 的 mAP@0.5,相比基线 YOLOv8-n 提升 9.3 个百分点,尤其在低光照和强光等极端工况下表现出显著优势,相比基线分别提升 15.9 和 14.2 个百分点。消融实验验证了各关键模块的有效性,其中 DS Conv 对细长异物检测贡献最大,LSR 模块在极端光照条件下效果最为显著。

下一章将在本章研究成果基础上,设计并实现面向实际应用的煤矿井下皮带机异物在线检测系统,实现从算法研究到工程落地的完整闭环。

第5章 煤矿井下皮带机异物在线检测系统设计与实现

本文在第三章和第四章中分别研究了煤矿井下皮带机图像增强方法及异物检测方法,并且经过实验验证了方法的有效性。为推动上述算法成果在实际生产中的工程化应用,本章面向煤矿井下皮带运输场景,针对皮带表面异物的定位和识别任务设计煤矿井下皮带机异物在线检测系统。目前井下异物检测仍以人工巡检为主要手段,然而矿井内部照明不均、粉尘浓度高、空间狭窄等不利因素严重制约了巡检人员的视觉观测效果,微小或低对比度异物极易被遗漏。同时,告警事件的人工登记、视频回溯及归档统计等后续工作也占用了大量人力成本,且该成本随监控点位数量的增长呈近似线性上升,难以满足大规模矿区的部署需求。为克服上述人工模式的固有缺陷,本章构建的在线检测系统打通了从视频流采集、实时异物识别、告警触发到历史数据归档管理的完整处理链路,实现了监测全流程的自动化与智能化,可显著降低对人工巡检的依赖。

5.1 系统需求分析

本系统以煤矿井下皮带运输线的异物智能监测为核心应用目标。系统对接井下部署的工业防爆摄像设备获取实时视频数据后,依次调用第三章所提区域自适应低光照增强算法与第四章所提光照语义引导检测算法,完成对视频帧中矸石、锚杆等异物的增强、定位与分类,并将分析结果以可视化方式实时呈现给操作人员。围绕上述系统建设目标,下文分别从功能性需求与非功能性需求两个维度展开具体分析。

5.1.1 系统功能性需求

功能性需求分析的核心目的是在开发启动前厘清系统需要承载的各项业务能力。通过深入调研终端用户的操作习惯、梳理关键业务场景的工作流程并结合系统的总体建设目标,提炼出系统需实现的全部功能清单,为后续的模块划分和接口设计提供明确的输入依据。

综合井下生产现场的实际运行特点,本系统在功能层面需支撑三方面核心业务:异物的实时检测与可视化预警、多路视频数据流的稳定接入与传输控制、以及历史异常事件的持久化存储与追溯回查。以下分别对这三类功能需求进行阐述。

(1) 异物实时检测需求分析

工作人员需要利用系统实时掌握监控区域皮带表面的异物情况,这要求系

统能够通过接收处理工业防爆摄像头传输的视频流，利用第三章提出的区域自适应低光照增强算法对低光照图像进行增强处理，输出增强图像 I_{out} 、光照图 S_2 以及区域权重图 W_{dark} 和 W_{bright} ，然后利用第四章提出的光照语义引导检测算法定位分析图像中的矸石、锚杆、槽钢、木材等异物目标，在图像中绘制检测框标记异物区域辅助用户分析图像中是否存在异物。并且系统中要以实时流的方式推送图像，异物检测模块在保证精度的同时要具备较好的处理性能。异物检测需求的用例图如图 5.1 所示，用户可以通过检测模块训练模型，并调用模型接口检测视频流中是否存在异物。

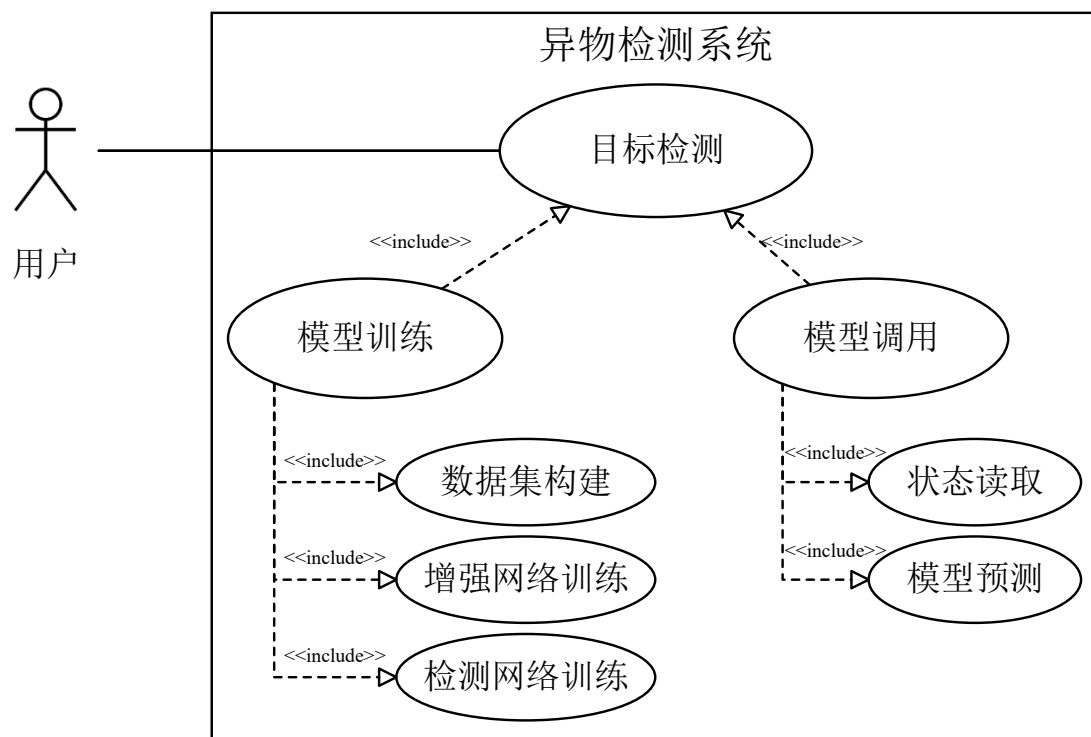


图 5.1 异物检测需求用例图

(2) 数据传输管理需求分析

井下皮带运输线通常分布于多条巷道，单一监控点位无法覆盖全部区域，因此系统需要同时接入多路视频数据源。每台工业防爆摄像头将采集到的视频帧经编码后通过矿用工业以太网推送至系统的流接收节点，检测模块再从对应节点拉取数据进行分析处理。在多设备并行工作的场景下，由于各摄像头的码率以及检测模块的算力存在差异，系统须具备灵活的设备关联机制和速率适配能力，确保上下游吞吐不匹配时仍能平稳运行。数据传输管理需求用例图如图 5.2 所示，除流节点的基本增删改查管理外，系统还应允许运维人员针对每个流节点分别设定帧缓冲区容量，在检测端短时处理速度跟不上数据产生速度时起到削峰缓冲作用，避免数据丢帧。同时提供可配置的缓冲区溢出应对策略，当长时间负载失衡导致缓冲区趋于饱和时，系统按照预设规则释放积压帧，防止处理管道因内存耗尽而中断服务。

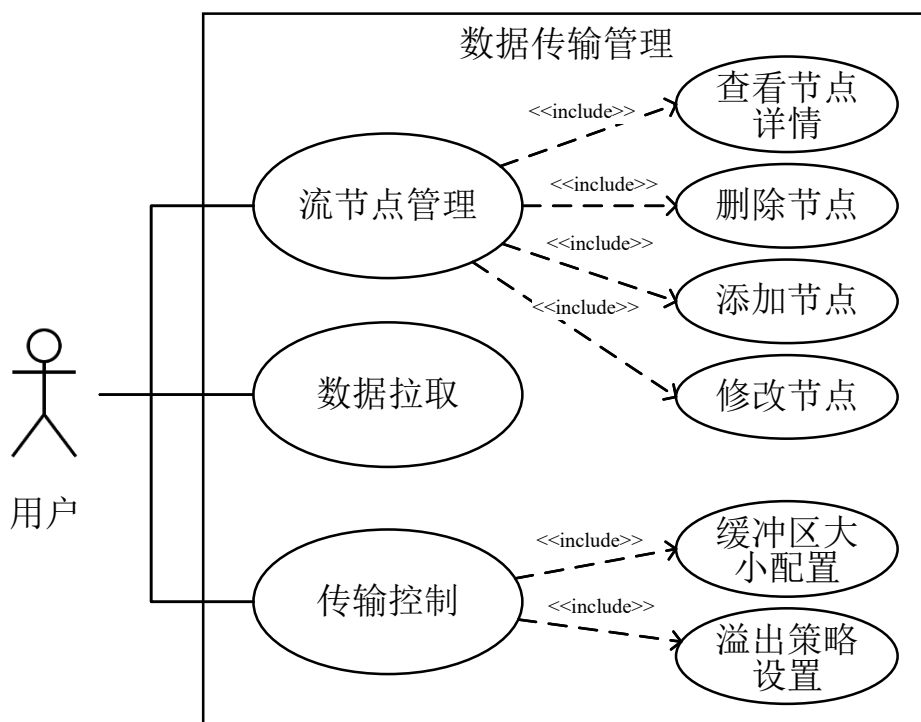


图 5.2 数据传输管理需求用例图

(3) 异常数据回溯需求分析

自动检测的结果需经人工二次确认后方可作为正式告警记录归档。因此，系统应将每一次疑似异物事件所涉及的视频片段完整保留。具体而言，当检测算法输出告警信号后，系统自动截取从告警触发帧至告警解除帧之间的连续图像序列，将其编码为标准视频格式并上传至对象存储服务。操作人员可随时在系统界面上检索、在线回放或下载对应的告警视频以完成人工复核。异常数据回溯需求的用例图如图 5.3 所示，包括视频存储查看、下载、删除和上传的功能需求。

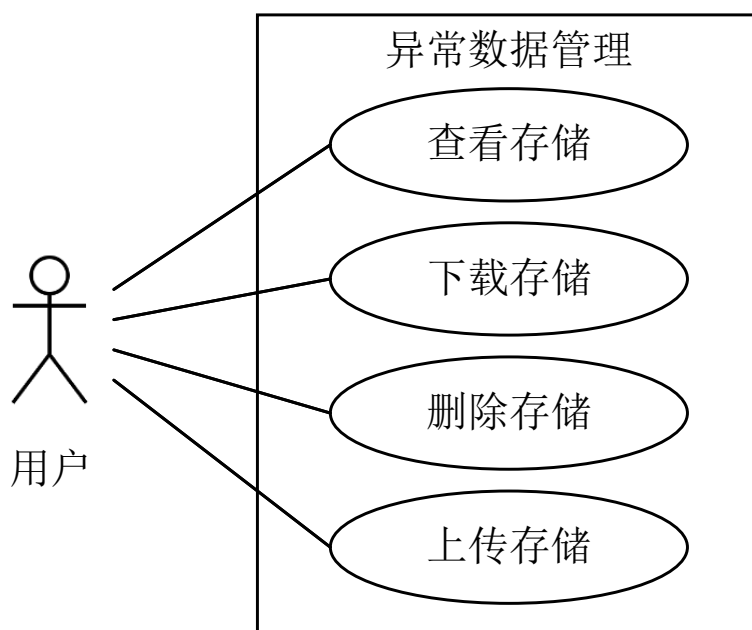


图 5.3 异常数据回溯需求用例图

5.1.2 系统非功能性需求

非功能性需求从宏观层面约束系统运行时的质量属性,涵盖响应速度、并发承载、持续可用性和信息安全等方面,直接决定了系统能否在实际工况中可靠运行以及后续的运维难度。在需求阶段充分明确这些约束,既能降低开发过程中的技术返工风险,又可为后续的技术选型和部署方案提供定量化参考。

结合皮带运行速度快、异物滞留窗口短等场景特征,对系统提出如下量化性能指标:(1)交互响应方面,任意单次业务操作请求的端到端往返时间须控制在 500 ms 以内,确保操作人员获得流畅的使用体验;(2)并发能力方面,系统应支持不少于 50 个终端的并行访问,以覆盖多个调度室和巡检岗位的同时在线使用场景;(3)实时性方面,在网络链路正常的前提下,端到端帧处理速率应达到 25 帧/秒以上,不低于监控设备的最低采集帧率;(4)时延方面,考虑到皮带典型运行速度为 2~4m/s,系统从图像采集到告警输出的全链路时延须控制在 200 ms 以内,确保异物在到达溜槽等危险部位前完成预警。

在可靠性层面,鉴于煤矿生产的连续性特征,系统须具备 7×24 小时不间断运行的能力,连续工作 168 小时期间不出现因内存泄漏或资源累积导致的性能退化或服务中断。在可维护性层面,整体架构遵循高内聚、松耦合的模块化原则,各功能组件可独立进行版本更新和性能调优而不影响其他模块的正常运作,降低系统后续演进的工程难度。在信息安全层面,视频流传输通道应采用加密机制防止数据在链路中被截获或篡改,持久化存储的告警记录和配置信息需设置访问权限控制,全面保障企业生产数据的完整性与保密性。

5.2 系统设计

系统设计阶段的任务是将上一节明确的需求规格转化为可指导编码实现的技术方案。该过程需要确定系统的分层架构、划定各功能模块的职责边界、规划模块间的数据交换流程并定义标准化的调用接口。本节在前述需求框架的约束下,依次对检测系统的总体流程和各子模块的功能逻辑进行详细设计。

5.2.1 系统总体设计

根据皮带机异物检测的需求,对系统的整体流程进行设计,如图 5.4所示:首先配置井下视频流拉流节点地址,然后系统从节点获取实时视频流;接着将视频流数据通过第三章提出的区域自适应低光照增强算法进行图像增强,该算法通过区域分割模块(RSM)实现暗区与亮区的像素级软划分,通过自适应光照增强器(AIE)提升暗区亮度,通过光效抑制器(LES)削弱强光引起的光斑和光晕伪影,输出增强图像 I_{out} 、光照图 S_2 以及区域权重图 W_{dark} 和 W_{bright} ; 然后

将增强结果输入第四章提出的光照语义引导检测算法检测异物区域，该算法采用 GhostNetV2 构建轻量化主干网络，引入动态蛇形卷积（DS Conv）增强对锚杆等细长异物的几何建模能力，通过坐标注意力（CA）与光照语义重标定（LSR）模块融合光照区域语义信息；如果存在异物目标，则根据检测结果在原始视频帧中标出异物区域和类别信息；最后输出结果，同时保存告警信息。

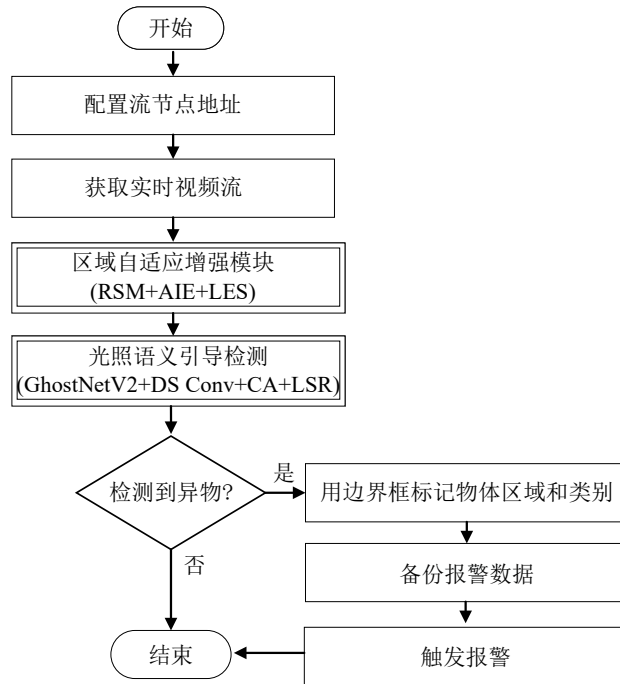


图 5.4 系统检测流程图

5.2.2 系统功能模块设计

按照系统的功能性需求分析，将系统功能模块设计为异物检测模块、数据传输模块和数据存储模块。

（1）异物检测模块设计

异物检测模块的功能是构建检测网络并为用户提供检测接口。异物检测算法由第三章和第四章设计的图像增强和异物检测算法组合而成，如图 5.5 所示。首先从井下监控服务器中拉取实时视频数据；然后将序列化的数据输入第三章设计的区域自适应低光照增强算法，该算法首先通过轻量化去噪网络（LDN）进行预处理，接着通过区域分割模块（RSM）生成暗区权重图 W_{dark} 和亮区权重图 W_{bright} ，然后分别通过自适应光照增强器（AIE）提升暗区亮度和光效抑制器（LES）削弱强光伪影，最终输出增强图像 I_{out} 和光照图 S_2 ；接着将增强结果和语义引导信息输入第四章设计的光照语义引导检测算法，该算法采用 GhostNetV2 作为轻量化主干网络提取多尺度特征，在特征提取阶段引入动态蛇形卷积（DS Conv）增强对锚杆等细长异物的几何建模能力，在特征融合阶段通

过坐标注意力 (CA) 机制编码空间位置信息, 并通过光照语义重标定 (LSR) 模块将光照图 S_2 和区域权重图 W_{dark} 、 W_{bright} 作为语义先验融入检测网络, 最终输出矸石、锚杆、槽钢、木材等异物目标的类别、位置和置信度; 最后根据检测结果使用检测框在当前帧上展示异物区域。

增强和检测模型由训练脚本进行训练, 训练得到的最佳权重会自动保存到系统模型权重位置, 后续检测任务会使用该位置的权重。为了便于使用, 组合后的模型会被封装为检测模块, 以接口的形式提供模型调用服务。

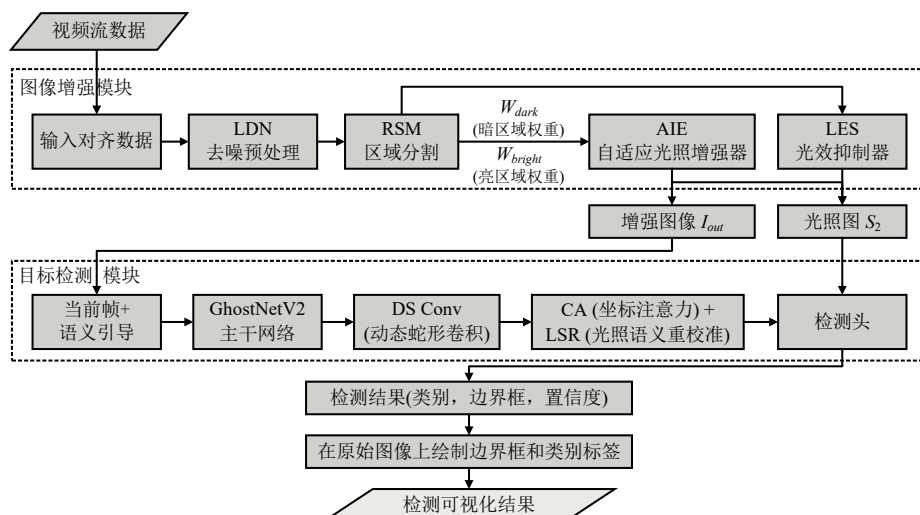


图 5.5 异物检测算法结构图

(2) 数据传输模块设计

数据传输模块的核心职能是保障多路视频数据在网络传输过程中的稳定性与连续性。在流节点的管理方面, 创建新节点时用户需填入节点名称、描述及拉流网络地址等必要信息, 模块随即对地址执行格式验证与连通性测试, 通过验证后将节点元数据持久化至数据库并标记为“可用”。已有节点的信息变更遵循与创建类似的校验流程, 区别仅在于以更新操作替代插入。节点列表查询通过在数据库中按指定条件进行筛选检索来实现, 而删除操作则会同步移除数据库中对应的节点记录。该模块的另一项关键能力是速率适配——即当算法处理吞吐量暂时低于前端推流速率时, 有效避免因数据积压而引发的内存溢出和服务崩溃。为此, 模块内部引入了生产者/消费者架构的帧缓冲区 (如图 5.6 所示): 拉流线程持续从流节点获取视频帧并写入缓冲区, 检测线程从缓冲区头部逐帧取出进行异物识别。缓冲区在短期速率波动时充当弹性蓄水池, 平滑上下游的速率差异。当系统长时间处于超负荷状态使缓冲区水位持续攀升时, 模块将根据实时监控到的队列深度启动预配置的淘汰策略, 可选先入先出丢弃或等间隔抽样丢弃, 以将缓冲区占用维持在安全阈值内。

(3) 数据存储模块设计

数据存储模块面向事后复查与告警统计分析提供数据底座, 支持告警记录

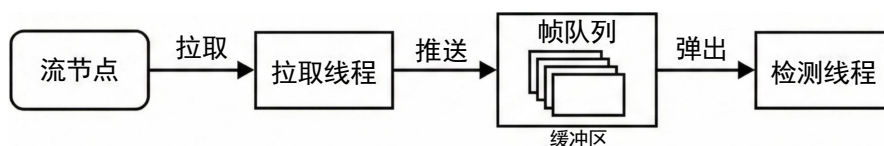


图 5.6 帧传递示意图

的创建、检索、下载与清除四项基本操作。其工作流程如下：检测模块判定出现疑似异物后触发归档流程——模块首先读取该检测任务在数据库中登记的关联信息（含设备编号、触发时刻与结束时刻），随后以帧为单位将告警时段内的检测结果逐帧写入临时文件，待告警事件完结后将该文件编码并上传至 MinIO 对象存储，同时在 MySQL 中插入一条包含文件存储路径与事件摘要的告警记录。用户执行查询操作时，系统按筛选条件在告警记录表中进行检索并返回符合条件的结果集。用户下载告警视频时，系统根据记录中的存储路径从 MinIO 拉取文件并还原为可播放的视频格式。删除操作则同步清理数据库记录与对象存储中的对应文件。

5.3 系统实现

系统实现阶段负责将设计方案落地为可运行的软件产品，包括各模块的编码开发、跨模块的接口集成、目标环境上的部署配置以及初步的功能验证与性能调测。本节首先介绍系统的总体技术架构与部署方案，随后逐一阐述检测模块、数据传输模块和数据存储模块的具体实现细节。

5.3.1 系统总体架构实现

表 5.1 和表 5.2 分别给出了系统开发与部署所采用的硬件环境配置和软件技术栈。

系统后端采用 Java 语言开发，基于 Spring Boot 框架构建 Web 服务，通过 MyBatis-Plus 实现数据持久化操作。检测模块采用 Python 语言实现，基于 PyTorch 框架完成模型训练，通过 TensorRT 进行推理加速。前端采用 Vue.js 框架实现响应式 Web 界面，结合 Element Plus 组件库构建交互页面，使用 flv.js 实现浏览器端的低延迟视频流播放。后端 Web 服务通过 HTTP RESTful 接口与前端通信，通过远程过程调用方式与 GPU 服务器上的检测模块进行交互。各模块通过 Docker 容器化部署，使用 Docker Compose 进行编排管理，边缘端检测服务、Web 后端服务和前端静态资源分别打包为独立容器，便于系统的部署、升级和运维。

本系统整体遵循 B/S（Browser/Server）体系架构，按照关注点分离原则自底向上划分为六个逻辑层次：客户端层、传输层、业务层、操作层、数据库层、运行环境层（如图 5.7 所示）。客户端层负责在 PC、平板及移动终端的浏览器上呈现

表 5.1 系统硬件环境配置

设备类型	配置项	具体参数
GPU 服务器	CPU	Intel Xeon Gold 6330 × 2
	GPU	NVIDIA RTX 4090 (24GB) × 1
	内存	128 GB DDR4 ECC
	存储	2 TB NVMe SSD
边缘计算设备	型号	NVIDIA Jetson AGX Orin 64GB
	GPU	Ampere 架构, 2048 CUDA Cores
	内存	64 GB LPDDR5
	功耗	15W–60W 可调
Web 服务器	CPU	Intel Xeon E-2388G
	内存	32 GB DDR4
	存储	1 TB SSD + 4 TB HDD

系统的交互页面。传输层采用 HTTP 协议承载前后端业务请求交互，采用 RTSP 协议完成视频数据的实时拉流。业务层中，数据传输与数据存储两个模块基于 Spring Boot 框架构建，异物检测模块则基于 PyTorch 框架训练并借助 TensorRT 进行推理加速，三者协同完成实时视频分析与告警管理的核心业务。操作层通过 MyBatis-Plus 封装数据库读写操作，以统一接口供上层调用。数据库层选用 MySQL 承担结构化业务记录的存储、Redis 提供高频数据的内存级缓存、MinIO 负责告警视频等大文件的对象化存储。运行环境层基于 Docker 容器化方案将前端服务、后端服务和检测推理服务分别封装为独立镜像，通过 Docker Compose 实现一键编排部署。

图 5.8描绘了系统的最终实现框架，整体由地下煤矿端、服务器端、存储端和客户端四个部分组成。在地下煤矿端，部署于各监控点位的工业防爆摄像头采集皮带机现场视频，经矿用工业以太网以 RTSP 协议将编码后的视频流推送至服务器端的流媒体服务器。在服务器端，流媒体服务器作为视频接入口，负责实时视频流的接收、解码与分发；GPU 推理服务器承载本文提出的区域自适应低光照增强与语义引导异物检测算法，通过检测 API 向流媒体服务器与 Web 应用服务器提供推理能力；Web 应用服务器作为系统调度中枢，负责请求路由、业务编排与状态管理；路由器统一承担服务器内部组件与外部网络之间的通信转发，对外通过 HTTP/WebSocket 协议经由互联网与远程客户端进行交互。在存储端，MySQL 关系型数据库存储任务配置、检测记录与告警元数据等结构化信息，MinIO 分布式对象存储管理告警视频片段及异物图片证据等非结构化文件，二者

表 5.2 系统软件环境与技术栈

类别	技术/工具	版本
操作系统	Ubuntu	22.04 LTS
容器化	Docker / Docker Compose	24.0.7 / 2.21
后端开发		
编程语言	Java	17 (OpenJDK)
Web 框架	Spring Boot	3.1.2
ORM 框架	MyBatis-Plus	3.5.3
流媒体处理	JavaCV (FFmpeg 封装)	1.5.9
检测模块		
编程语言	Python	3.10.12
深度学习框架	PyTorch	2.0.1 (CUDA 11.8)
模型部署	TensorRT	8.6.1
模型转换	ONNX / ONNX Runtime	1.14.0 / 1.15.0
视频处理	OpenCV	4.8.0
CUDA 工具包	CUDA / cuDNN	11.8 / 8.9.2
前端开发		
编程语言	JavaScript (ES6+) / HTML5 / CSS3	—
前端框架	Vue.js	3.3.4
UI 组件库	Element Plus	2.3.8
视频播放	flv.js	1.6.2
可视化图表	ECharts	5.4.3
数据库与存储		
关系型数据库	MySQL	8.0.33
缓存	Redis	7.0.12
对象存储	MinIO	RELEASE.2023-07-21
流媒体与通信		
流媒体协议	RTSP / RTMP	—
流媒体服务器	Nginx-RTMP	1.24.0
硬件解码	NVIDIA Video Codec SDK	12.1

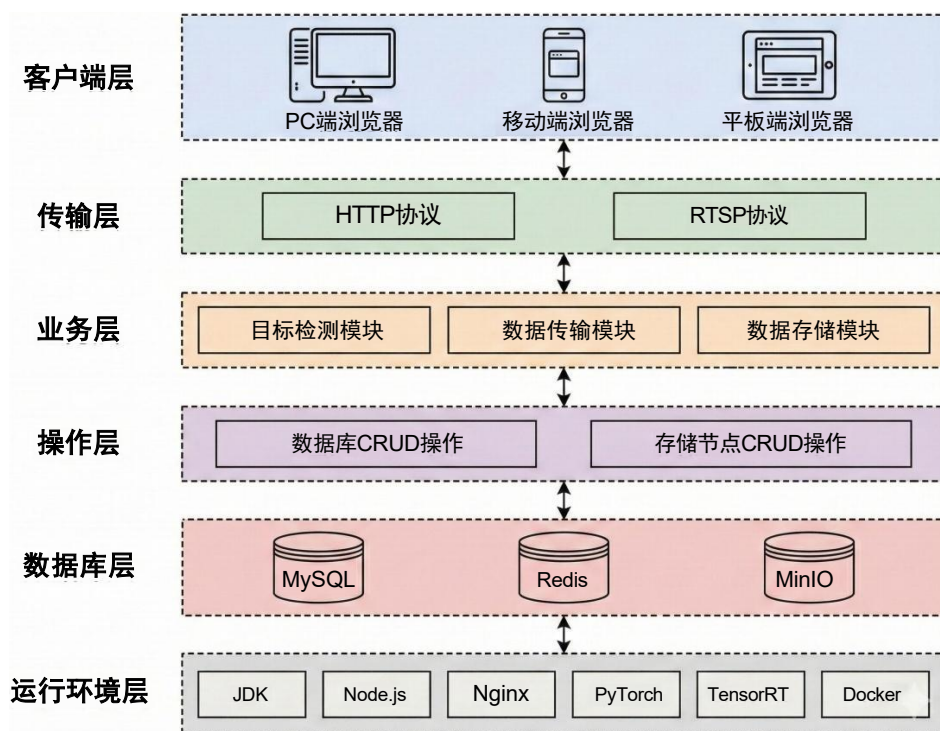


图 5.7 异物检测系统结构图

协同支撑系统数据的高效读写与持久化。在客户端，操作人员可通过 PC、平板电脑或手机等多种终端设备，借助浏览器经由互联网访问 Web 应用服务器，实时查看检测结果并接收告警推送，从而实现对皮带机异物状态的远程在线监测。

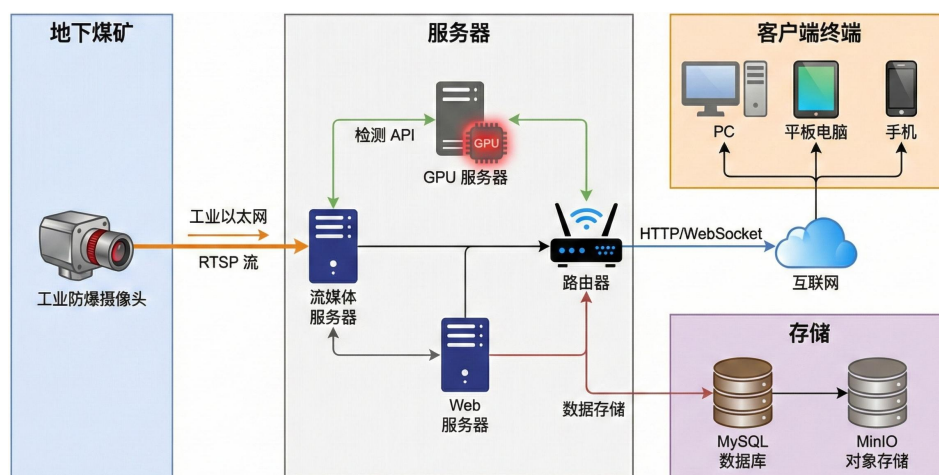


图 5.8 系统框架图

5.3.2 检测模块实现

检测模块向用户开放检测任务的全生命周期管理能力，每个检测任务记录包含摄像头拉流地址、任务名称及描述信息三个核心字段。该模块的后端控制入口为 `DetectorController` 类。在新建任务场景下，前端引导用户填写上述字段后将请求提交至 `addDetectorTask()` 方法。该方法首先对任务名称执行合法性校验，随后

通过 `getServiceStatus()` 向目标拉流地址发起连通性探测以确认数据源可达。上述校验均通过后,新任务记录被持久化至 `DetectorTaskTable` 表,标志着任务创建成功。任务信息变更由 `modDetectorTask()` 方法处理,其执行流程与新建任务基本一致,区别在于最终执行的是数据库 `UPDATE` 操作而非 `INSERT`。任务删除操作通过 `delDetectorTask()` 方法实现:该方法在移除 `DetectorTaskTable` 中目标记录的同时,级联查询 `AlertTable` 以定位与该任务相关联的全部告警条目,对这些条目执行批量删除,并通过异步线程清理 `MinIO` 中对应的告警视频文件,避免产生孤立的存储对象。启动检测任务时, `startDetectTask()` 方法首先通过 `getServiceStatus()` 确认流媒体服务器状态正常,接着经由 `remoteDetector()` 远程调用将目标拉流地址下发至部署在 GPU 服务器上的 `DetectorService` 推理服务,并触发其进入持续检测状态。`DetectorService` 内部通过 `objectDetector()` 方法循环拉取视频帧并执行增强一检测流水线,每帧的推理结果由 `resPush()` 实时推送至前端播放窗口。一旦算法判定当前帧序列中存在异物目标, `resUpload()` 负责将对应的告警视频片段异步上传至 `MinIO`,同时回调 Web 服务端的 `saveAlertRecords()` 接口完成告警元数据的入库。

为实现边缘端的高效推理,需要将 `PyTorch` 训练得到的模型转换为适合部署的格式。本系统采用“`PyTorch` → `ONNX` → `TensorRT`”的部署链路,通过以下技术实现推理加速:(1)层融合:将多个连续的小算子(如 `Conv+BN+ReLU`)融合为单个高效 `kernel`;(2)精度量化:支持 `FP16` 和 `INT8` 量化,以少量精度损失换取显著的速度提升和显存节省;(3)内存优化:自动优化中间张量的内存分配,减少显存占用;(4) `Kernel` 自动调优:针对特定硬件平台选择最优的 `kernel` 实现。

5.3.3 数据传输模块实现

在线检测流程涵盖视频拉流、算法推理与结果推流三个环节,三者分别运行在不同的硬件节点上且各自的数据吞吐能力存在差异。为协调这种速率不平衡,系统在拉流与推理环节之间引入帧缓冲队列作为弹性调节层。本节设计了一种自适应帧队列管理算法,每次消费者线程从队列取帧时触发执行,以动态维持队列水位在安全范围内,防止帧积压引发内存溢出。

帧队列控制算法的原理是动态监测和分级处理,如图 5.9 所示,算法在每次消费者线程取帧时触发执行,首先采集当前队列深度 `current_queue_num`,并计算其与上一次采集值之间的变化量 `delta_queue_num` 作为增长趋势指标。将二者相加得到队列深度的短期预测值 `estimate_num`,用于评估队列在下一个处理周期可能达到的水位。基于 `estimate_num` 与预设的队列容量上限 `queue_max` 之间的关系,算法将队列状态划分为三个等级并执行差异化的处理策略:当 `estimate_num` 超过

queue_max 时,判定队列即将溢出,立即执行全量清空操作以紧急释放负载;当 estimate_num 介于 queue_max 的 80% 与 queue_max 之间时,视为队列进入预警区间,按比例丢弃最旧的 20% 帧数据以将水位回退至安全阈值;当 estimate_num 低于 queue_max 的 80% 时,队列运行在健康区间,算法不执行任何干预。

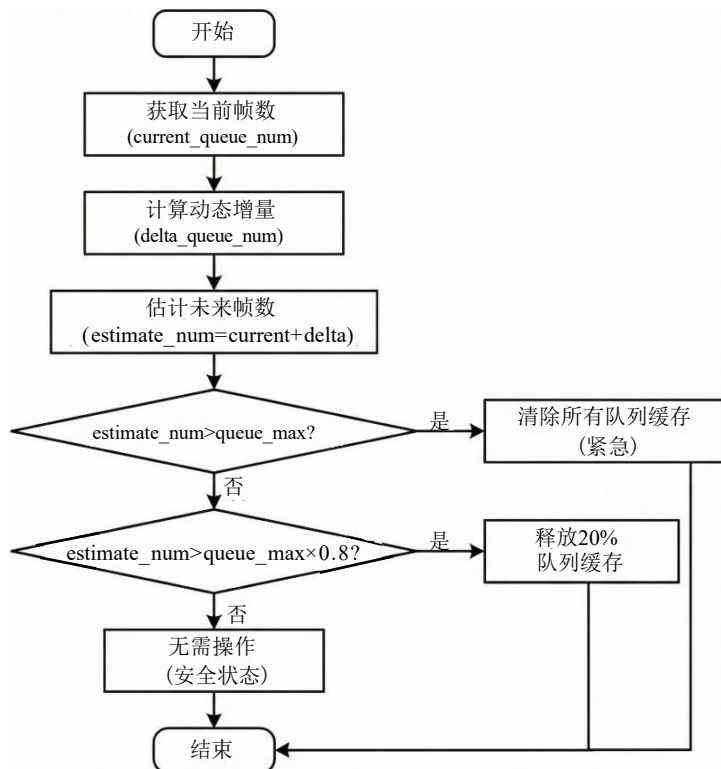


图 5.9 帧队列算法流程图

为降低解码延迟,系统优先使用 NVIDIA Video Codec SDK 提供的硬件解码能力(NVDEC),将 H.264/H.265 码流直接解码至 GPU 显存,避免 CPU 解码的性能瓶颈和 GPU-CPU 间的数据拷贝开销。当硬件解码不可用时,回退至 FFmpeg 软解码方案。同时采用 CUDA Stream 实现异步并行,使用独立的 CUDA Stream 进行视频解码、数据传输和模型推理,使三者能够并行执行;采用页锁定内存(Pinned Memory)加速 CPU-GPU 间的数据拷贝;通过双缓冲机制实现帧间流水线重叠。

5.3.4 数据存储模块实现

本文对于结构化数据例如检测任务数据的存储采用 MySQL 数据库,对于对象数据例如告警视频采用 MinIO 作为存储服务器。数据存储模块的后端控制器 AlertRecorderController 统一暴露告警数据的增删改查接口。在删除操作方面,提供两个粒度的方法:delAlertRecordsByDetector() 支持按检测任务维度进行级联删除,即同时清理 AlertTable 中该任务下的所有告警记录并通过异步任务移除 MinIO 中关联的视频文件;delAlertRecordsById() 则按单条告警记录的

主键进行定点删除。在写入与更新方面，`saveAlertRecords()` 负责将新产生的告警事件元数据（包括触发时间、结束时间、关联设备、视频存储路径等）插入 `AlertTable`，`modAlertRecords()` 则对已有记录的可编辑字段进行更新。在查询方面，`getAlertRecords()` 根据任务编号或时间范围等筛选条件返回匹配的告警记录列表，`getAlertData()` 则依据记录中的 MinIO 对象路径拉取完整的告警视频并以流式方式推送至前端播放器。

为降低因瞬时误检导致的误报，系统采用基于时间窗的多帧一致性确认策略。具体而言，设定一个滑动时间窗口 T_w （默认 1 秒），在窗口内统计同一目标（ $\text{IoU} > 0.5$ 视为同一目标）的出现次数 N_{appear} 。仅当 $N_{\text{appear}} \geq K$ （默认 $K = 3$ ）时才触发告警。此外，为避免对同一异物的重复告警，引入告警冷却机制：对于已触发告警的目标，在冷却时间 T_{cool} （默认 30 秒）内不再重复告警。

5.3.5 系统界面展示

图 5.10 展示了系统监控界面，页面采用左侧信息栏加右侧视频区的双栏布局。左侧信息栏上半部分为设备选择区，以列表形式展示已接入系统的全部监控摄像头，支持按名称关键字搜索以快速定位目标设备；选中某一设备后，下半部分的告警记录区将自动加载该设备的历史告警事件列表，每条记录提供在线查看、本地下载和删除三种操作入口，便于运维人员进行事后复核。右侧主体区域放置四个视频播放窗口：第一个窗口展示摄像头采集的原始低光照画面，第二个和第三个窗口展示经区域自适应增强算法处理后的增强画面，第四个窗口叠加显示检测框与类别标签的最终检测结果。

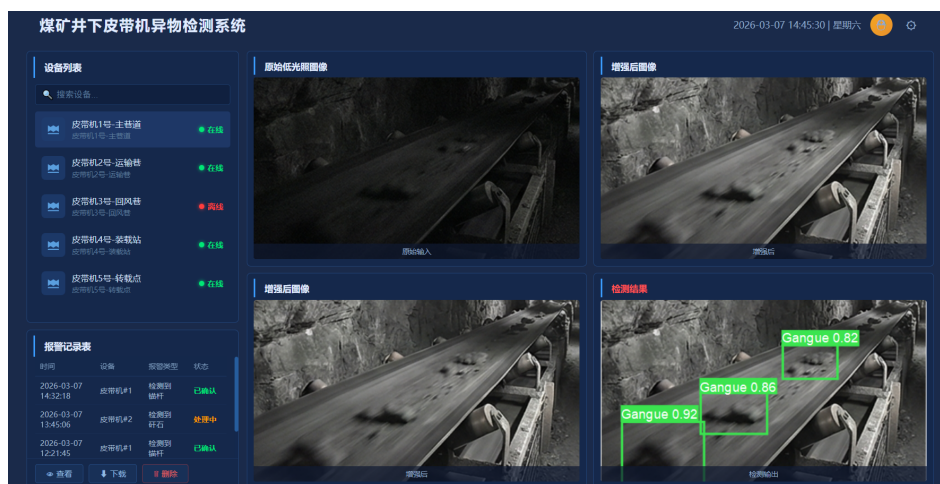


图 5.10 异物检测系统监控界面

图 5.11 展示了监控设备管理界面，该页面以数据表格的形式呈现系统中已登记的摄像设备清单，每行记录涵盖设备名称、所属监控区域、视频流推流地址、当前在线状态等关键属性。页面顶部提供新增和删除两种表单弹窗，运维人员可

据此完成设备信息的录入与变更。

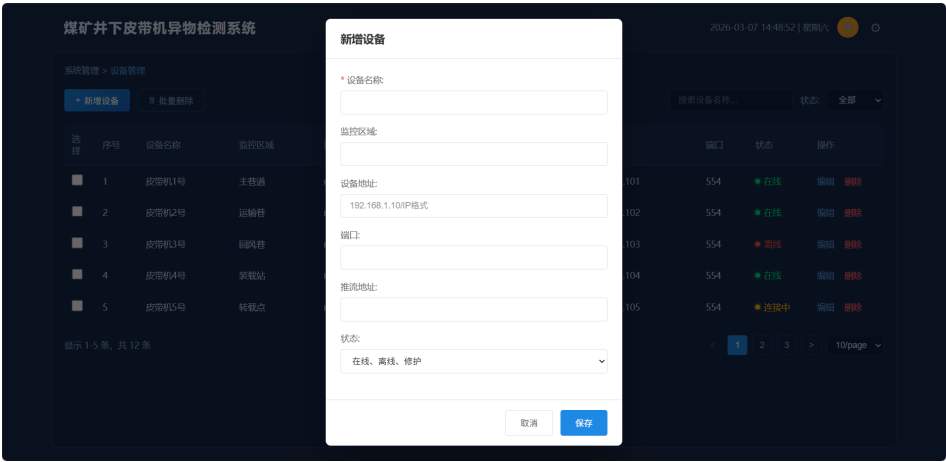


图 5.11 监控设备管理界面

图 5.12展示了检测结果管理界面，集中展示系统历史上检出异物的全部告警事件。每条告警记录包括关联设备的相关信息、告警视频的存储路径、事件起止时间戳等字段。管理员可通过该页面对告警记录执行浏览、下载和删除操作。



图 5.12 检测结果管理界面

5.4 系统测试

系统测试是在所有模块完成集成之后，对系统整体是否达到既定需求规格进行的端到端验证，重点考察功能正确性、性能指标达标情况以及长时间运行下的可靠性表现。本节分别针对异物检测模块、数据传输模块和数据存储模块设计了黑盒功能测试用例，并通过响应时间测试、并发压力测试、稳定性测试和浏览器兼容性测试对系统的非功能性表现进行了全面评估。

5.4.1 系统功能性测试

(1) 异物检测模块功能测试

异物检测模块的功能测试覆盖了检测任务的完整生命周期操作，包括任务列表加载、合法/非法参数下的任务创建、任务信息修改、任务删除、任务启动以及在线切换检测任务等场景。测试用例和测试结果如表 5.3所示，全部用例均按预期通过验证。

表 5.3 检测模块功能测试

测试场景	测试步骤	预期结果	测试结果
加载检测任务列表	进入检测任务管理界面	正确显示所有检测任务信息	通过
增加检测任务-参数异常	在检测任务管理界面添加无效拉流地址的检测任务	无法添加，显示拉流地址异常	通过
增加检测任务-参数正常	在检测任务管理界面添加参数正确的检测任务	正确添加，显示添加成功	通过
修改检测任务-参数异常	在检测任务管理界面修改已有检测任务的拉流地址为无效地址	无法修改，显示拉流地址异常	通过
修改检测任务-参数正常	在检测任务管理界面修改已有检测任务的拉流地址为其他有效地址	正确修改，显示修改成功	通过
删除检测任务	在检测任务管理界面删除已有检测任务	正确删除，显示删除成功	通过
开始一个检测任务	在系统监控界面选择一个检测任务	视频播放窗口正常显示实时监测结果	通过
更换检测任务	在系统监控界面更换检测任务	视频播放窗口更换实时监测结果	通过

(2) 数据传输模块功能测试

数据传输模块的测试重点在于验证帧缓冲队列与溢出管理策略在不同上下游速率比条件下的表现。测试方法为：分别以 15 fps、20 fps 和 25 fps 三种帧率向系统推送测试视频流，观察前端监控页面能否持续、无卡顿地渲染检测结果。测试用例和测试结果如表 5.4所示，三种速率条件下系统均保持了稳定输出，未出现丢帧黑屏或服务中断现象。

(3) 数据存储模块功能测试

在对数据存储模块功能测试时，会对查看、增加和删除异常告警数据进行功能测试。测试用例和测试结果如表 5.5所示。

表 5.4 数据传输模块功能测试

测试场景	测试步骤	预期结果	测试结果
不同任务的数据推送速率存在差异	在系统监控界面选择检测任务，更改检测任务的推送速率为15fps、20fps、25fps	视频播放窗口正常显示检测结果	通过

表 5.5 数据存储模块功能测试

测试场景	测试步骤	预期结果	测试结果
查看异常告警记录	进入检测结果界面	正确显示所有的异常告警记录	通过
查看异常告警视频	在系统监控界面选择异常数据查看	视频播放窗口正常显示异常告警视频	通过
检测任务发现并保存异常数据	检测任务推送存在异物的数据流	检测结果界面新增异常数据记录且 MinIO 正确保存异常视频	通过
删除检测任务时删除对应异常数据	检测任务管理界面删除包含异常数据的记录	检测结果界面对应检测任务的异常记录被删除	通过
删除异常告警数据	在检测结果删除特定异常告警记录	正确删除，显示删除成功	通过

5.4.2 系统非功能性测试

（1）算法性能测试

本文在第 3.4 节对区域自适应低光照增强算法进行了定量分析，在第 4.5 节对光照语义引导检测算法进行了定量分析。测试结果表明，在 RTX 4090 服务器上采用 FP16 精度时，组合后的端到端检测算法可达到较高的推理速度，单帧延迟较低；在 Jetson AGX Orin 边缘设备上采用 INT8 量化后，推理速度和延迟均能满足系统实时检测的性能需求。此外，FP16 量化相比 FP32 可获得显著加速，模型大小明显减少，精度损失较小；INT8 量化可获得更高加速比，精度损失在可接受范围内。

（2）系统响应性能测试

响应性能的评估通过接口级自动化压测完成：借助自动化工具对检测任务管理、设备管理和告警记录管理三组核心接口逐一发起 10 轮连续请求，统计每个接口的平均响应耗时。测试结果表明，所有被测接口的平均响应时间均在 500 ms 以下，满足系统实时响应的需求。

（3）系统并发性能测试

并发承载能力通过模拟 50 个客户端同时向检测任务查询接口发起并行请求来验证。测试期间监测了请求成功率、响应时间分布和服务器资源占用情况。结果显示, 50 路并发条件下的所有请求均成功返回结果, 且平均接口请求耗时低于 500ms, 系统能够满足实际业务并发需求。

(4) 浏览器兼容性测试

前端兼容性验证在 Chrome 浏览器、Firefox 浏览器、Edge 浏览器和 360 浏览器上分别展开。每款浏览器上均完整执行了以下操作序列: 检测任务的增删改查、在线检测启动与切换、告警记录的检索浏览、告警视频的在线回放与本地下载。各浏览器的页面渲染效果与交互行为均保持一致, 未出现布局错位、视频播放失败或功能按钮无响应等兼容性问题, 表明前端代码具备良好的跨浏览器适配性。

5.5 本章小结

本章针对传统煤矿皮带机异物检测存在的技术瓶颈, 基于第三章和第四章提出的区域自适应低光照增强算法与光照语义引导检测算法, 构建了面向煤矿井下皮带运输场景的智能化监测系统。在需求分析阶段, 结合井下作业环境的特殊性与传统人工巡检模式在响应速度、数据留存和安全保障等维度的不足, 提炼出系统在实时检测、稳定传输与数据持久化三个方面的核心能力要求。在系统设计与实现阶段, 异物检测模块集成了区域分割模块 (RSM)、自适应光照增强器 (AIE)、光效抑制器 (LES)、GhostNetV2 主干网络、动态蛇形卷积 (DS Conv) 和光照语义重标定 (LSR) 模块, 在保证检测功能的同时满足可靠、稳定的系统非功能性需求; 最后展示了系统的操作界面, 演示了操作细节, 并且设计了对系统功能和性能的测试方案, 验证了系统在煤矿井下环境监测场景中的工程适用性。通过本章实现的异物检测系统, 可有效实现异物事件的自动识别与预警, 相比传统人工检测方式, 显著提升了异物响应时效并降低了人力资源投入, 为煤矿皮带运输安全监测提供了可扩展的技术解决方案。

第6章 总结与展望

6.1 总结

煤矿井下皮带运输系统是矿井生产的核心装备，其安全稳定运行对于保障矿井高效生产和人员安全具有重要意义。皮带表面混入的矸石、锚杆、槽钢等非煤异物极易导致皮带撕裂等恶性事故，如何有效检测这些异物并及时预警是煤矿安全监测的主要难点。本文针对现阶段井下异物检测中存在光照条件恶劣、目标形态复杂、边缘计算资源受限等问题开展基于语义引导的煤矿井下皮带机异物检测方法研究。

本文对于煤矿井下皮带机异物检测的算法研究分为图像增强和轻量化检测两阶段。在低光照图像增强部分，通过引入区域分割机制，将图像中暗区与亮区进行像素级软划分；且利用井下强光点与暗区共存的特点，分别采用自适应光照增强器和光效抑制器进行差异化处理，改善图像视觉质量。在检测阶段中，利用图像增强模块输出的光照图和区域权重图作为语义先验，本文设计了坐标注意力与光照区域语义联合引导的特征重标定机制，且在检测网络中引入针对细长异物的动态蛇形卷积模块提高对锚杆等目标的几何建模能力。在本文的实验部分对两阶段工作都分别做了分析，结果表明本文方法对比已有方法在煤矿井下低光照条件下的异物检测具有更好的检测结果。

本文的主要工作内容和贡献包括：

(1) 针对煤矿井下暗区与强光点共存的复杂光照条件导致图像可见性差、对比度低，且强光点区域光斑和光晕伪影严重的问题，提出一种区域自适应低光照图像增强方法，融合区域分割与差异化增强策略的改进增强网络。首先，基于 Retinex 理论框架设计区域分割模块，通过可学习的软阈值机制动态划分暗区与亮区，生成区域权重图；其次提出自适应光照增强器，利用暗区权重引导光照分量调整，提升暗区亮度和对比度；最后设计光效抑制器，有效削弱强光点引起的光斑扩散和光晕效应，并引入复合损失函数实现亮度增强、结构保持与伪影抑制的多目标平衡。在消融实验中证明了各个模块的有效性，且通过对比经典的增强方法发现，本文算法能在提高增强质量的同时保持较低的计算复杂度。

(2) 针对边缘计算平台资源受限以及锚杆等细长异物几何建模困难的问题，本文提出一种光照语义引导的轻量化异物检测方法。首先采用 GhostNetV2 构建高效主干网络，通过廉价操作生成冗余特征的方式大幅降低参数量和计算量，在该网络中利用解耦全连接注意力增强长距离依赖建模能力。进一步引入动态蛇形卷积模块，使卷积核能够沿目标轮廓自适应滑动，增强对锚杆、钢筋等细长异

物的几何感知。构建坐标注意力与光照区域语义联合引导的光照语义重标定模块，将增强网络输出的光照图和区域权重图映射为语义引导权重，建立光照分布信息与检测特征的深度关联。实验结果表明，相对已有方法本文算法对细长异物的检测能力更强，在保持较高检测精度的同时显著降低了模型参数量和计算开销，能够满足边缘端实时部署需求。

(3) 针对传统煤矿皮带机监测系统存在的检测精度受光照影响大、实时性难以满足高速皮带监测需求等问题，本文结合提出的图像增强和轻量化检测算法设计了煤矿井下皮带机异物在线检测系统。首先分析了煤矿井下场景对于异物检测的基本要求和当前系统存在的问题，从系统需求到系统实现详细阐述了本文系统的解决方案，并针对各个模块给出了详细的流程设计图和实现框架，最后展示了系统的实际操作流程，且设计功能测试和性能测试方案进行了系统验证，验证了检测系统的有效性。

6.2 展望

本文围绕煤矿井下皮带机异物检测任务开展了研究，在低光照图像增强和轻量化目标检测方面取得了一定的进展，并整合设计了异物在线检测系统。但受限于实验条件等因素，未来可在如下方面继续深入探索：

(1) 在设计区域自适应低光照图像增强方法时，本文采用了区域分割与差异化增强策略以应对井下复杂光照条件带来的挑战，虽然提升了图像增强质量，但同时增加了网络结构复杂度，导致增强速度仍有所不足。未来可进一步研究轻量化的图像分解与增强结构，探索端到端的生成式模型以替代传统 Retinex 框架，降低建模复杂光照信息所需的网络深度，进而提高模型推理速度、减少模型计算量。

(2) 本文在语义引导检测网络设计中将光照图和区域权重图作为静态语义先验融入检测网络，然而在实际应用场景中，井下光照条件随设备运行状态和环境变化而动态改变，固定的语义引导机制可能难以完全适应光照的实时变化，进而对算法性能造成一定影响。未来可研究动态自适应的语义引导策略，使检测网络能够根据输入图像的光照特性实时调整注意力分布，从而提升本文方法在实际应用中的鲁棒性。

(3) 构建更加完备的煤矿井下异物检测数据集。虽然本文工作中采集了多个井下场景的图像数据并进行了数据增强，但目前的数据集仍无法涵盖所有典型工况，不同矿井的巷道环境、照明布置与皮带运行状态仍会对检测结果造成影响。未来可通过扩充数据集的矿井数量和场景类型、丰富不同异物类别与光照条件下的样本，进一步增强检测系统的通用性和鲁棒性。

(4) 开展增强方法对多类下游任务的独立验证。本文在第四章中通过端到端联合训练验证了所提区域自适应低光照增强方法与异物检测网络协同优化后的整体性能，但尚未将增强方法作为独立预处理模块在更广泛的下游任务上进行解耦式评估。未来可构建图像增强对下游任务的串联实验范式，从三个维度系统考察本文方法的通用性与工程价值：一是下游任务维度，覆盖目标检测（皮带跑偏识别、井下人员越界监测、设备状态识别）、图像分类（煤矸石分选、煤质等级评估）和语义分割（皮带边缘提取、堆煤区域分割）三类典型井下视觉任务；二是下游模型维度，分别选用 YOLOv8/YOLOv11、ResNet/Swin Transformer、DeepLabV3+/SegFormer 等主流网络作为下游基线；三是对比方法维度，与无增强基线、传统方法（CLAHE、LIME）以及代表性深度增强方法（URetinex-Net、Retinexformer、Zero-IG 等）进行横向比较。通过上述实验，进一步明确低光照增强对煤矿井下智能感知系统的贡献边界，为面向工业场景的“增强—感知”一体化设计提供更系统的实证依据。

参考文献

- [1] 国家统计局. 中华人民共和国 2024 年国民经济和社会发展统计公报[EB/OL]. 2025 [2025-04-01]. https://www.stats.gov.cn/sj/zxfb/202502/t20250228_1958817.html.
- [2] Guo X, Liu X, Zhou H, et al. Belt tear detection for coal mining conveyors[J]. *Micromachines*, 2022, 13(3): 449.
- [3] 王国法, 庞义辉, 任怀伟, 等. 智慧矿山系统工程及关键技术研究与实践[J]. *煤炭学报*, 2024, 49(01): 181-202.
- [4] Wang Y, Miao C, Miao D, et al. Hazard source detection of longitudinal tearing of conveyor belt based on deep learning[J]. *PLoS one*, 2023, 18(4): e0283878.
- [5] 张立亚, 郝博南, 马征, 等. 煤矿井下图像增强关键技术研究综述[J]. *煤炭科学技术*, 2025, 53(11): 101-116.
- [6] Li C, Guo C, Han L, et al. Low-light image and video enhancement using deep learning: A survey[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2021, 44(12): 9396-9416.
- [7] Cui Z, Li K, Gu L, et al. You only need 90K parameters to adapt light: A light weight transformer for image enhancement and exposure correction[C]//*Proceedings of the 33rd British Machine Vision Conference (BMVC)*. BMVA Press, 2022.
- [8] Jobson D J, Rahman Z u, Woodell G A, et al. Properties and performance of a center/surround retinex[J]. *IEEE transactions on image processing*, 1997, 6(3): 451-462.
- [9] Guo X. LIME: A method for low-light image enhancement[C]//*Proceedings of the 24th ACM International Conference on Multimedia*. 2016: 87-91.
- [10] Liu R, Ma L, Zhang J, et al. Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021: 10561-10570.
- [11] Wu W, Weng J, Zhang P, et al. URetinex-Net: Retinex-based deep unfolding network for low-light image enhancement[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2022: 5901-5910.
- [12] Wei C, Wang W, Yang W, et al. Deep retinex decomposition for low-light enhancement[C]//*Proceedings of the British Machine Vision Conference (BMVC)*. 2018: 155.
- [13] Zhang Y, Zhang J, Guo X. Kindling the darkness: A practical low-light image enhancer[C]//*Proceedings of the 27th ACM International Conference on Multimedia*. 2019: 1632-1640.
- [14] Xu X, Wang R, Fu C W, et al. SNR-aware low-light image enhancement[C]//*Proceedings*

- of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022: 17714-17724.
- [15] Wang T, Zhang K, Shen T, et al. Ultra-high-definition low-light image enhancement: A benchmark and transformer-based method[C]//Proceedings of the AAAI Conference on Artificial Intelligence: Vol. 37. 2023: 2654-2662.
- [16] Cai Y, Bian H, Lin J, et al. Retinexformer: One-stage retinex-based transformer for low-light image enhancement[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023: 12504-12513.
- [17] Guo C, Li C, Guo J, et al. Zero-reference deep curve estimation for low-light image enhancement[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 1780-1789.
- [18] Jiang Y, Gong X, Liu D, et al. Enlightengan: Deep light enhancement without paired supervision[J]. IEEE Transactions on Image Processing, 2021, 30: 2340-2349.
- [19] Ma L, Ma T, Liu R, et al. Toward fast, flexible, and robust low-light image enhancement [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 5637-5646.
- [20] Fu Z, Yang Y, Tu X, et al. Learning a simple low-light image enhancer from paired low-light instances[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2023: 22252-22261.
- [21] Yang S, Ding M, Wu Y, et al. Implicit neural representation for cooperative low-light image enhancement[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). 2023: 12918-12927.
- [22] Shi Y, Liu D, Zhang L, et al. Zero-IG: Zero-shot illumination-guided joint denoising and adaptive enhancement for low-light images[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024: 3015-3024.
- [23] Yi X, Xu H, Zhang H, et al. Diff-retinex: Rethinking low-light image enhancement with a generative diffusion model[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023: 12302-12311.
- [24] Hou J, Zhu Z, Hou J, et al. Global structure-aware diffusion process for low-light image enhancement[J]. Advances in Neural Information Processing Systems, 2023, 36: 79734-79747.
- [25] Yan Q, Feng Y, Zhang C, et al. HVI: A new color space for low-light image enhancement [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2025: 5678-5687.
- [26] Yang W, Wang S, Wu J, et al. A low-light image enhancement method for personnel safety

- monitoring in underground coal mines[J]. *Complex & Intelligent Systems*, 2024, 10(3): 4019-4032.
- [27] Murat A A, Kiran M S. A comprehensive review on yolo versions for object detection[J]. *Engineering Science and Technology, an International Journal*, 2025, 70: 102161.
- [28] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 39(6): 1137-1149.
- [29] 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*. 2016: 779-788.
- [30] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//*Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. 2017: 2980-2988.
- [31] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: Optimal speed and accuracy of object detection[EB/OL]. 2020[2026-04-01]. <https://arxiv.org/abs/2004.10934>.
- [32] Wang C Y, Bochkovskiy A, Liao H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2023: 7464-7475.
- [33] Wang C Y, Yeh I H, Mark Liao H Y. YOLOv9: Learning what you want to learn using programmable gradient information[C]//*Proceedings of the European Conference on Computer Vision (ECCV)*. 2024: 1-21.
- [34] Carion N, Massa F, Synnaeve G, et al. End-to-end object detection with transformers[C]//*Proceedings of the European Conference on Computer Vision (ECCV)*. 2020: 213-229.
- [35] Zhao Y, Lv W, Xu S, et al. DETRs beat YOLOs on real-time object detection[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2024: 16965-16974.
- [36] Howard A G, Zhu M, Chen B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications[EB/OL]. 2017[2026-04-01]. <https://arxiv.org/abs/1704.04861>.
- [37] Han K, Wang Y, Tian Q, et al. GhostNet: More features from cheap operations[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2020: 1580-1589.
- [38] Tang Y, Han K, Guo J, et al. GhostNetV2: Enhance cheap operation with long-range attention [C]//*Advances in Neural Information Processing Systems (NeurIPS)*: Vol. 35. 2022: 9969-9982.
- [39] Zhang X, Zhou X, Lin M, et al. ShuffleNet: An extremely computation-efficient CNN for mobile devices[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*

- dition (CVPR). 2018: 6848-6856.
- [40] Wang J, Huang Z, Dong Y, et al. LE-YOLO: a lightweight and effective yolo model for remote sensing image object detection[J]. *Journal of Real-Time Image Processing*, 2026, 23(1): 27.
 - [41] Nguyen H H, Hoang M S, et al. LEAF-YOLO: Lightweight edge-real-time small object detection on aerial imagery[J]. *Intelligent Systems with Applications*, 2025, 25: 200484.
 - [42] Dai J, Qi H, Xiong Y, et al. Deformable convolutional networks[C]//*Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. 2017: 764-773.
 - [43] Qi Y, He Y, Qi X, et al. Dynamic snake convolution based on topological geometric constraints for tubular structure segmentation[C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2023: 6070-6079.
 - [44] 王国法, 任怀伟, 富佳兴. 煤矿智能化建设高质量发展难题与路径[J]. *煤炭科学技术*, 2025, 53(1): 1-18.
 - [45] Zhou F, Zou J, Xue R, et al. Enhancing object detection in underground mines: UCM-Net and self-supervised pre-training[J]. *Sensors*, 2025, 25(7): 2103.
 - [46] Luo B, Kou Z, Han C, et al. A faster and lighter detection method for foreign objects in coal mine belt conveyors[J]. *Sensors*, 2023, 23(14): 6276.
 - [47] Fan H, Liu J, Yan X, et al. A fast and high-accuracy foreign object detection method for belt conveyor coal flow images with target occlusion[J]. *Sensors*, 2024, 24(16): 5251.
 - [48] Zhao A, Zheng Q, Zhao Y, et al. MSC-DETR: A lightweight real-time foreign object detection algorithm for underground belt conveyors in coal mines[J]. *Journal of King Saud University – Computer and Information Sciences*, 2026.
 - [49] Hong Y, Wang L, Su J, et al. Enhanced foreign body detection on coal mine conveyor belts using improved DLEA and lightweight SARC-DETR model[J]. *Signal, Image and Video Processing*, 2025, 19(5): 349.
 - [50] Wang S, Zhang S, Wu J, et al. Clip-optimized multimodal image enhancement via isp-cnn fusion for coal mine iovt under uneven illumination[J]. *IEEE Internet of Things Journal*, 2025.
 - [51] Zhao A, Zheng Q, Li L. Application of CycleGAN-based low-light image enhancement algorithm in foreign object detection on belt conveyors in underground mines[J]. *Scientific Reports*, 2025, 15(1): 27897.
 - [52] Du Q, Zhang S, Wang Z, et al. A hybrid zero-reference and dehazing network for joint low-light underground image enhancement[J]. *Scientific Reports*, 2025, 15(1): 10135.
 - [53] Wu Y, Pan C, Wang G, et al. Learning semantic-aware knowledge guidance for low-light image enhancement[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023: 1662-1671.
 - [54] Ling J, Fu Z, Yuan X. Lightweight coal mine conveyor belt foreign object detection based on

- improved YOLOv8n[J]. Scientific Reports, 2025, 15(1): 10361.
- [55] 孔二伟, 张亚邦, 李佳悦, 等. 面向煤矿井下低光照图像的增强方法[J]. 工矿自动化, 2023, 49(04): 62-69.
- [56] 孙林, 陈圣, 姚旭龙, 等. 矿井智能监控目标识别的图像增强方法与应用[J]. 煤炭学报, 2024, 49(S1): 495-504.
- [57] 王宏伟, 刘学刚, 王浩然, 等. 基于图像增强和改进 YOLOv8 的煤矿低光照目标检测[J]. 工矿自动化, 2025, 51(09): 33-42.
- [58] Han X, Khan I R, Rahardja S. High dynamic range image tone mapping: Literature review and performance benchmark[J]. Digital Signal Processing, 2023, 137: 104015.
- [59] Hore A, Ziou D. Image quality metrics: PSNR vs. SSIM[C]//2010 20th International Conference on Pattern Recognition (ICPR). IEEE, 2010: 2366-2369.
- [60] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: From error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.
- [61] Mittal A, Soundararajan R, Bovik A C. Making a “completely blind” image quality analyzer [J]. IEEE Signal Processing Letters, 2013, 20(3): 209-212.
- [62] Mittal A, Moorthy A K, Bovik A C. No-reference image quality assessment in the spatial domain[J]. IEEE Transactions on Image Processing, 2012, 21(12): 4695-4708.
- [63] Venkatanath N, Praneeth D, Sumohana S C, et al. Blind image quality evaluation using perception based features[C]//2015 Twenty First National Conference on Communications (NCC). IEEE, 2015: 1-6.
- [64] Zou Z, Chen K, Shi Z, et al. Object detection in 20 years: A survey[J]. Proceedings of the IEEE, 2023, 111(3): 257-276.
- [65] Ali M L, Zhang Z. The YOLO framework: A comprehensive review of evolution, applications, and benchmarks in object detection[J]. Computers, 2024, 13(12): 336.
- [66] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017: 2117-2125.
- [67] Liu S, Qi L, Qin H, et al. Path aggregation network for instance segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2018: 8759-8768.
- [68] Hassanin M, Anwar S, Radwan I, et al. Visual attention methods in deep learning: An in-depth survey[J]. Information Fusion, 2024, 108: 102417.
- [69] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2018: 7132-7141.
- [70] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional block attention module[C]//Proceedings

- of the European Conference on Computer Vision (ECCV). 2018: 3-19.
- [71] Hou Q, Zhou D, Feng J. Coordinate attention for efficient mobile network design[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 13713-13722.
- [72] Raza S M, Abidi S M H, Masduzzaman M, et al. Lightweight deep learning for visual perception: A survey of models, compression strategies, and edge deployment challenges[J]. Neurocomputing, 2025: 131357.
- [73] Joshua A S, Balasubramaniam P. An adaptive low-light image enhancement method via fusion of a new intuitionistic fuzzy generator and fractal-fractional derivative[J]. Signal, Image and Video Processing, 2025, 19(3): 233.
- [74] Xu L, Hu C, Zhu W, et al. Degrade for upgrade: Learning degradation representations for real-world low-light image enhancement[J]. Computers and Electrical Engineering, 2024, 119: 109622.
- [75] Cai J, Gu S, Zhang L. Learning a deep single image contrast enhancer from multi-exposure images[J]. IEEE Transactions on Image Processing, 2018, 27(4): 2049-2062.
- [76] Fu X, Zeng D, Huang Y, et al. A weighted variational model for simultaneous reflectance and illumination estimation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016: 2782-2790.
- [77] Rudin L I, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms[J]. Physica D: Nonlinear Phenomena, 1992, 60(1-4): 259-268.
- [78] Yang W, Wang W, Huang H, et al. Sparse gradient regularized deep Retinex network for robust low-light image enhancement[J]. IEEE Transactions on Image Processing, 2021, 30: 2072-2086.
- [79] Hai J, Xuan Z, Yang R, et al. R2RNet: Low-light image enhancement via real-low to real-normal network[J]. Journal of Visual Communication and Image Representation, 2023, 90: 103712.
- [80] Li C, Guo C, Loy C C. Learning to enhance low-light image via zero-reference deep curve estimation[J]. IEEE transactions on pattern analysis and machine intelligence, 2021, 44(8): 4225-4238.
- [81] Zheng Z, Wang P, Liu W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[C]//Proceedings of the AAAI Conference on Artificial Intelligence: Vol. 34. 2020: 12993-13000.
- [82] Ultralytics. YOLOv8: State-of-the-art YOLO models[EB/OL]. 2023[2026-04-01]. <https://github.com/ultralytics/ultralytics>.

致谢

在研究学习期间，我有幸得到了多位老师的悉心教导与热情帮助。首先，衷心感谢我的导师赵云波教授。赵云波老师以其深厚的学术功底、严谨的治学态度和敏锐的科学洞察力，为我的研究方向指引了清晰的道路。老师在学业上的悉心指导与生活上的温暖关怀，使我在科研道路上受益匪浅，铭感五内。

感谢吕文君老师在深度学习方面给予我的宝贵指导。仍记得刚接触科研时，吕老师耐心地帮助我打下基础，从模型原理到实验细节，均给予了深刻而透彻的讲解。这些知识积累为我后续的研究工作奠定了坚实的基础，令我受益至今。

感谢夏秀山老师和李泽瑞老师在目标检测方面的悉心指导。夏老师和李老师在专业实践过程中给予了诸多宝贵支持，使我在该领域的研究思路得到了极大的拓展与深化，在此深表谢意。

感谢我的父母及家人在我求学期间给予的默默支持与无条件的理解。他们的鼓励是我面对困难时最坚实的后盾，使我得以心无旁骛地投入到科学研究中。

感谢室友谢万砦、刘文豪、罗春晖在日常生活、学习和工作中给予的帮助与陪伴。三年的朝夕相处，留下了诸多美好的回忆，感恩有你们相伴。同时，也感谢所有曾在求学路上给予我帮助与关心的同学和朋友们，是你们的温暖让这段岁月更加充实而难忘。

谨以此文，向所有给予我帮助与支持的人，致以最诚挚的谢意。

在读期间取得的科研成果

已发表论文:

- [1] Luo L H, Xie W T, Xia X S, et al. Region-adaptive low-light image enhancement with light effect suppression and detail preservation[C]//Pacific Graphics Conference Papers, Posters, and Demos. The Eurographics Association, 2025.

发明专利:

- [1] 吕文君; 罗里恒; 康宇; 李鲲. 一种图像超分辨率方法. 授权号: CN 117274064 B, 2024-04-02.