

文章编号 1004-924X(2026)18-2936-23

面向铁路扣件的多尺度结构感知低照度 图像增强方法

马思维¹, 胡贺南¹, 李荣华^{2*}

(1. 大连交通大学 机械工程学院, 辽宁 大连 116028;

2. 大连交通大学 电气工程学院, 辽宁 大连 116045)

摘要:针对车载铁路扣件检测中隧道内光照不足、补光不均和局部阴影造成的图像亮度不足、边缘模糊及关键结构难以辨识问题,提出一种多尺度结构感知低照度图像增强网络(Multi-Scale Structure-Aware Network, MSA-Net)。该网络以自校准光照网络(Self-Calibrated Illumination, SCI)为基础,通过多尺度特征提取模块(Multi-Scale Feature Extraction Module, MSFE)获取不同感受野下的光照、纹理和边缘信息,利用低照度结构感知注意力模块(Structure-Sensing Attention Module, SSA)强化螺栓轮廓、弹条纹理和扣件边界响应,并采用自适应残差融合模块(Adaptive Residual Fusion Module, ARF)根据区域退化程度动态调节增强强度。在福州滨海快线低照度铁路扣件数据集上开展消融、算法对比及下游检测实验。结果表明,MSA-Net的标准差、平均梯度、信息熵、自然图像质量评价和无参考图像空间质量评价分别为46.39, 30.49, 4.88, 21.36和26.34;模型参数量为0.016 M,计算量为9.782 GFLOPs,增强速度达到302.39 FPS。采用正常光照图像训练的YOLOv8检测器时,MSA-Net增强输入的精确率、召回率、mAP@0.5和mAP@0.5:0.95分别达到86.10%, 83.60%, 85.80%和59.00%,端到端处理速度达到99.02 FPS;采用低照度原图训练的YOLOv8检测器时,四项指标分别达到88.10%, 85.40%, 88.60%和61.80%。在两个公开低照度图像数据集上的实验进一步验证了该方法的场景适应性。所提方法能够在保持实时处理效率的同时改善低照度扣件图像质量和结构可辨识性,可为车载铁路扣件检测提供稳定的图像输入。

关键词:低照度图像增强;铁路扣件;多尺度特征提取;结构感知注意力;自适应残差融合

中图分类号:TP391.41 **文献标识码:**A

doi:10.37188/OPE.20263418.2936 **CSTR:**32169.14.OPE.20263418.2936

Multi-scale structure-aware low-light image enhancement method for railway fasteners

MA Siwei¹, HU Henan¹, LI Ronghua^{2*}

(1. School of Mechanical Engineering, Dalian Jiaotong University, Dalian 116028, China;

2. School of Electrical Engineering, Dalian Jiaotong University, Dalian 116045, China)

* Corresponding author, E-mail: lironghua705@163.com

Abstract: To address insufficient image brightness, blurred edges, and poor distinguishability of key structures caused by inadequate tunnel illumination, uneven supplementary lighting, and local shadows in vehi-

收稿日期:2026-06-25;修订日期:2026-07-21.

基金项目:国家自然科学基金项目(No. U24B20159);辽宁省教育厅科学研究重点项目(No. LJKZ0475);大连市高层次人才创新支持计划(No. 2022RJ03);辽宁省交通科技基金项目(No. 202416)

cle-mounted railway fastener detection, a Multi-Scale Structure-Aware Network (MSA-Net) was proposed for low-light image enhancement. Built upon the Self-Calibrated Illumination (SCI) network, MSA-Net employed a Multi-Scale Feature Extraction Module (MSFE) to extract illumination, texture, and edge information under different receptive fields, a Structure-Sensing Attention Module (SSA) to strengthen the responses of bolt contours, clip textures, and fastener boundaries, and an Adaptive Residual Fusion Module (ARF) to dynamically adjust the enhancement intensity according to regional degradation levels. Ablation studies, algorithm comparisons, and downstream detection experiments were conducted on a low-light railway fastener image dataset collected from the Fuzhou Binhai Express Line. The results show that MSA-Net achieves a standard deviation, average gradient, information entropy, Natural Image Quality Evaluator, and Blind/Referenceless Image Spatial Quality Evaluator of 46.39, 30.49, 4.88, 21.36, and 26.34, respectively. The model contains 0.016 M parameters, requires 9.782 GFLOPs, and achieves an enhancement speed of 302.39 FPS. For the YOLOv8 detector trained on normal-light images, the Precision, Recall, mAP@0.5, and mAP@0.5:0.95 obtained using MSA-Net-enhanced images reach 86.10%, 83.60%, 85.80%, and 59.00%, respectively, with an end-to-end processing speed of 99.02 FPS. For the detector trained on original low-light images, the corresponding four metrics reach 88.10%, 85.40%, 88.60%, and 61.80%, respectively. Experiments on two public low-light image datasets further demonstrate the adaptability of the proposed method to different scenes. MSA-Net improves low-light railway fastener image quality and structural distinguishability while maintaining real-time processing efficiency, thereby providing stable image inputs for vehicle-mounted railway fastener detection.

Key words: low-light image enhancement; railway fasteners; multi-scale feature extraction; structure-sensing attention; adaptive residual fusion

1 引 言

基于机器视觉的铁路扣件检测已成为铁路智能运维的重要技术手段。车载视觉检测系统能够在列车运行过程中连续采集扣件图像,并为螺栓缺失、弹条断裂、扣件松动和结构缺陷等异常状态识别提供图像依据^[1]。与人工巡检相比,车载视觉检测具有检测效率高、连续性强和部署灵活等优势。然而,在夜间运行、隧道检测和车底车载采集等实际场景中,铁路扣件图像常受到光照不足、补光不均、局部阴影、车辆遮挡和复杂轨旁背景等因素影响,导致图像亮度不足、对比度降低、纹理细节弱化和结构边缘模糊等低照度退化问题^[2]。

对于铁路扣件检测而言,低照度退化不仅影响图像视觉质量,还会影响实时检测、结构定位和状态识别。螺栓轮廓、弹条边缘和扣件边界等关键结构通常占据区域较小,且包含细粒度结构特征。当图像亮度和对比度不足时,这些结构难

以从背景中有效区分,容易削弱特征提取和目标定位能力。尤其是边缘模糊和结构轮廓弱化,会降低扣件定位及后续视觉测量的稳定性^[3]。因此,研究具有结构保持能力的低照度图像增强方法,对于提高铁路扣件图像的可见性、可测性和检测稳定性具有重要意义。

现有低照度图像增强方法主要包括传统图像增强方法和基于深度学习的方法。传统方法通常基于图像先验、灰度变换或 Retinex 分解实现亮度和对比度提升,例如直方图均衡化^[4]、伽马校正^[5]和 Retinex 类算法^[6]。郑爽爽等提出了融合全变分与 Gamma 的低照度图像增强算法,通过 Retinex 变分模型分解光照图像和反射图像,并结合自适应 Gamma 校正改善低照度图像亮度和细节表现^[7];Gao 等将 Retinex 风格分解与去噪深度图像先验相结合,以缓解噪声干扰和亮度退化问题^[8];张箴等提出了基于多通道 Retinex 模型的低照度图像增强网络,通过初始化模块获得光照和反射分量,并采用分通道增强策略解决低照

度增强后的色偏问题^[9];Wang 等将 Retinex 分解与 Swin Transformer 及类 U-Net 结构相结合,实现光照和反射分量增强^[10]。上述方法能够在一定程度上改善图像质量,但在复杂工业检测场景中仍容易出现噪声放大、局部过增强、边缘模糊和结构保持不足等问题。

近年来,基于深度学习的低照度图像增强方法通过学习低照度图像到增强图像之间的非线性映射,在亮度恢复和视觉质量改善方面取得了较好效果。零参考深度曲线估计方法(Zero-Reference Deep Curve Estimation, Zero-DCE)^[11]和自校准光照网络(Self-Calibrated Illumination, SCI)^[12]等方法为无参考或轻量化低照度增强提供了有效思路。陈清江等提出多通道融合注意力网络,通过多层级特征提取和跨尺度注意机制改善低照度图像的亮度、对比度和纹理细节^[13];Yuan 等结合 Transformer 和卷积神经网络,协同建模全局亮度关系与局部纹理信息^[14];江泽涛等提出基于全局一致性的非配对低照度图像增强方法,通过全局一致性保持、局部平滑调制和深度特征强化,提高增强图像的曝光一致性、噪声抑制能力和细节保持效果^[15];Liu 等通过融合空间信息和通道上下文信息构建双信息低照度增强网络^[16]。上述研究推动了低照度增强方法在亮度恢复、细节保持和视觉自然性方面的发展。然而,现有方法主要面向自然场景图像,通常侧重于全局亮度恢复、颜色保真和主观视觉质量改善,对工业检测场景中的局部边缘、细粒度纹理和可测结构信息缺少显式约束。对于铁路扣件检测任务,增强结果不仅需要改善图像亮度和对比度,还应保持螺栓轮廓、弹条纹理和扣件边界等关键结构线索的清晰性与稳定性,以满足后续目标检测、结构定位和视觉测量的需求。

作为一种具有代表性的轻量化低照度增强方法,SCI 具有无需配对正常光照图像监督、网络结构紧凑和便于部署等优势。但将其直接应用于低照度铁路扣件图像时仍存在一定局限:一是单尺度卷积对弱纹理、局部阴影和非均匀光照分布的表征能力有限;二是缺少对螺栓轮廓、弹条纹理和扣件边界等关键可测结构的显式关注;三是固定的残差相加方式难以根据局部退化程度自适应调节增强强度,容易造成欠增强、过增强

或结构不一致,从而影响扣件结构的可辨识性和后续检测模型的稳定性。

针对上述问题,本文提出一种面向铁路扣件的多尺度结构感知低照度图像增强方法,并构建多尺度结构感知网络 MSA-Net。该网络以 SCI 为基础,引入多尺度特征提取模块(Multi-Scale Feature Extraction Module, MSFE)、低照度结构感知注意力模块(Structure-Sensing Attention Module, SSA)和自适应残差融合模块(Adaptive Residual Fusion Module, ARF)。通过三个模块的协同作用,MSA-Net 在保持轻量化设计的同时,提高低照度铁路扣件图像的亮度恢复、关键结构保持和增强稳定性。本文的主要工作如下:

(1)提出多尺度特征提取模块,通过 3×3 卷积、 5×5 卷积和空洞卷积构建不同感受野分支,提取低照度铁路扣件图像中的纹理、边缘和光照分布信息,提高网络对弱光区域、局部阴影和非均匀光照的建模能力。

(2)设计低照度结构感知注意力模块,通过融合特征显著性、暗区感知和边缘结构响应,引导校准网络关注螺栓轮廓、弹条纹理和扣件边界等关键区域,增强可测结构的保持能力和边缘细节表达。

(3)构建自适应残差融合模块,根据输入图像与增强映射的联合特征生成空间自适应融合权重,动态调节不同位置的增强强度,从而缓解固定残差相加造成的局部欠增强、过增强和结构失真问题。

(4)在自建铁路扣件数据集上开展消融、算法对比、复杂度分析及不同训练策略下的下游检测实验,并在多个公开低照度图像数据集上进行扩展验证,从增强质量、计算效率、检测性能和场景适应性等方面综合评价所提方法的有效性。

2 多尺度结构感知增强方法

2.1 MSA-Net 结构

为提高低照度铁路扣件图像的可见性和结构可辨识性,本文在自校准增强网络 SCI 的基础上,构建多尺度结构感知增强网络 MSA-Net,其整体结构如图 1 所示。该网络采用多阶段自校准增强策略,主要包括光照估计、反射分量恢复和输入校准三个过程。与原始 SCI 相比,MSA-

Net引入多尺度特征提取模块、低照度结构感知注意力模块和自适应残差融合模块,以提高非均匀光照建模、关键结构保持和区域自适应增强能力。

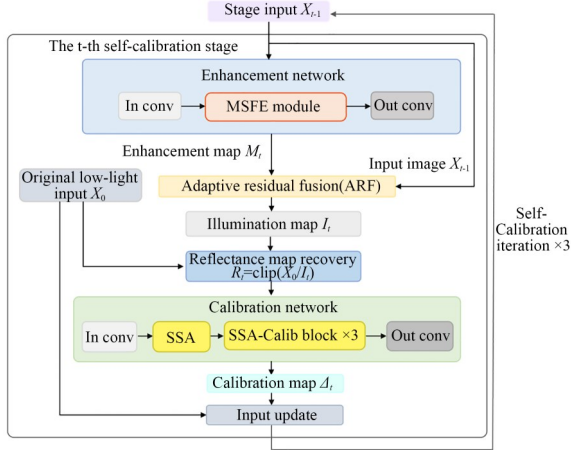


图1 MSA-Net的整体架构

Fig. 1 Overall architecture of MSA-Net

设原始低照度铁路扣件图像为 X_0 , 在训练阶段, 网络进行 T 次自校准迭代, 本文设置 $T=3$ 。在第 t 阶段, 当前输入图像 X_{t-1} 首先输入增强网络。增强网络由输入卷积层、多尺度特征提取模块和输出卷积层组成, 用于从不同感受野中提取光照分布、边缘纹理和局部结构特征, 并生成当前阶段的增强图 M_t 。

随后, 将增强图 M_t 与当前输入图像 X_{t-1} 输入自适应残差融合模块, 得到当前阶段的光照图 I_t 。与 SCI 中固定残差相加方式不同, 自适应残差融合模块能够根据局部退化程度动态调节增强强度, 使暗光区域得到充分增强, 同时抑制局部过增强和结构失真。

获得光照图 I_t 后, 根据 Retinex 成像模型, 以原始低照度图像 X_0 恢复第 t 阶段的反射图 R_t , 如式(1)所示:

$$R_t = \text{clip}\left(\frac{X_0}{I_t}, 0, 1\right), \quad (1)$$

其中: 光照图 I_t 的数值范围限制为 $[0.000\ 1, 1]$, 以避免反射恢复过程中出现除零和数值不稳定; $\text{clip}(\cdot)$ 表示截断操作, 用于将反射图像素值限制在 $[0, 1]$ 范围内。恢复得到的 R_t 作为当前阶段的增强结果, 并用于后续输入校准。

随后, 将 R_t 输入校准网络生成校准图 Δ_t 。校准网络由输入卷积层、低照度结构感知注意力模块、三个结构感知校准块和输出卷积层组成。低照度结构感知注意力模块用于引导网络关注螺栓轮廓、弹条纹理和扣件边界等关键结构区域, 结构感知校准块则进一步细化结构特征并修正残余光照偏差。生成的校准图 Δ_t 用于更新下一阶段输入, 如式(2)所示:

$$X_t = U(X_{t-1}, \Delta_t) = \text{clip}(X_0 + \Delta_t, 0, 1), \quad (2)$$

其中: $U(\cdot)$ 表示阶段输入更新函数, X_{t-1} , Δ_t 和 X_t 分别表示当前阶段输入、校准图和更新后的下一阶段输入。该函数将原始输入 X_0 与 Δ_t 逐像素相加, 并通过 $\text{clip}(\cdot)$ 将结果限制在 $[0, 1]$ 范围内, 即小于 0 的值置为 0, 大于 1 的值置为 1。 X_{t-1} 的信息已通过当前阶段的校准过程包含在 Δ_t 中, 不与其直接相加。更新后的 X_t 作为第 $t+1$ 阶段的输入。

MSA-Net 沿用 SCI 的无监督训练方式, 不使用配对正常光照图像作为监督标签。训练损失由光照保真损失和加权光照平滑损失组成, 3 个阶段的总损失如式(3)所示:

$$L = \sum_{t=1}^3 (1.5L_{\text{fid}}^t + L_{\text{smooth}}^t), \quad (3)$$

其中: T 表示自校准增强阶段数; L_{fid}^t 和 L_{smooth}^t 分别表示第 t 阶段的光照保真损失和加权光照平滑损失。总损失由 3 个阶段的损失累加得到, 其中光照保真损失的权重设置为 1.5。

光照保真损失采用均方误差约束第 t 阶段光照图 I_t 与该阶段输入图像 X_{t-1} 之间的整体差异, 其计算如式(4)所示:

$$L_{\text{fid}}^{(t)} = \frac{1}{3HW} \sum_{c=1}^3 \sum_{i=1}^H \sum_{j=1}^W [I_t^c(i, j) - X_{t-1}^c(i, j)]^2, \quad (4)$$

其中: H 和 W 分别表示图像的高度和宽度, c 表示 RGB 通道。 $I_t^c(i, j)$ 和 $X_{t-1}^c(i, j)$ 分别表示光照图与当前阶段输入在通道 c 、位置 (i, j) 处的像素值。

该损失用于限制光照估计结果相对于当前阶段输入的整体偏差, 避免光照图发生过度变化, 并提高多阶段光照估计的稳定性。

为约束非均匀光照的空间连续性, 同时降低跨颜色边界过度平滑的影响, 加权光照平滑损失根据当前阶段输入图像相邻像素在 YCbCr 颜色

空间中的差异构造平滑权重,其计算如式(5) 所示:

$$\begin{cases} L_{\text{smooth}}^{(t)} = \sum_{\delta \in D} \frac{1}{|\Omega_{\delta}|} \sum_{p \in \Omega_{\delta}} w_{p,p+\delta}^{(t)} \|I_t(p) - I_t(p+\delta)\|_1 \\ w_{p,p+\delta}^{(t)} = \exp \left[-\frac{\|C(X_{t-1})(p) - C(X_{t-1})(p+\delta)\|_2^2}{2\sigma^2} \right] \end{cases} \quad (5)$$

其中: p 表示像素位置, δ 表示邻域方向, D 表示由一阶和二阶邻域组成的24组方向偏移, Ω_{δ} 表示方向 δ 下的有效像素集合; $C(\cdot)$ 表示由RGB颜色空间至YCbCr颜色空间的转换; $w_{p,p+\delta}^{(t)}$ 表示相邻像素之间的平滑权重; σ 为颜色敏感系数,沿用SCI的设置取10。

输入图像及转换后的颜色分量均处于归一化数值域,因此该权重主要用于稳定光照图的空间平滑过程;扣件轮廓、弹条纹理和结构边界的保持主要由MSFE和SSA实现。

总体而言,MSA-Net在保留SCI无需成对数据训练和轻量化优势的基础上,通过MSFE提取多尺度光照与结构特征,利用SSA增强扣件关键结构响应,并采用ARF自适应调节区域增强强度,提高了铁路扣件图像的亮度恢复、结构保持和增强稳定性。

2.2 MSFE 模块

低照度铁路扣件图像通常存在亮度不足、局部阴影、纹理弱化和光照不均等问题。原始SCI增强网络主要采用单尺度卷积,固定的特征采样范围限制了网络对局部纹理、结构边缘和区域光照变化的协同表征能力,容易导致增强后弱纹理恢复不足、结构边缘不清晰和亮度分布不均。

为解决上述问题,本文在增强网络中引入多尺度特征提取模块MSFE,其结构如图2所示。MSFE采用3个具有不同卷积核尺寸和采样方式的并行分支: 3×3 卷积分支用于提取局部纹理和细粒度结构; 5×5 卷积分支通过连续密集采样捕获较大范围的结构与光照信息;空洞卷积分支采用膨胀率为2的 3×3 卷积,在等效 5×5 覆盖范围内进行稀疏采样,以较少参数补充不同空间位置的上下文信息。

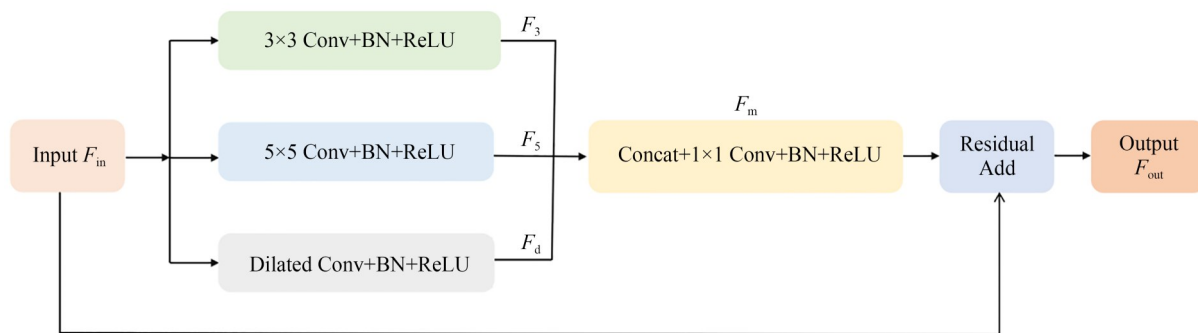


图2 MSFE模块结构

Fig. 2 Structure of the MSFE module

设MSFE的输入特征为 F_{in} ,3个分支的特征提取过程如式(6)~式(8)所示:

$$F_3 = \sigma(BN_3(Conv_{3 \times 3}(F_{in}))), \quad (6)$$

$$F_5 = \sigma(BN_5(Conv_{5 \times 5}(F_{in}))), \quad (7)$$

$$F_d = \sigma(BN_d(DConv_{3 \times 3}(F_{in}))), \quad (8)$$

其中: $Conv_{3 \times 3}(\cdot)$ 和 $Conv_{5 \times 5}(\cdot)$ 分别表示 3×3 和 5×5 卷积操作, $DConv(\cdot)$ 表示空洞卷积操作, $BN_3(\cdot)$, $BN_5(\cdot)$, $BN_d(\cdot)$ 分别表示各分支相互独

立的批归一化(Batch Normalization, BN)操作, $\sigma(\cdot)$ 表示ReLU激活函数。

在网络实现中,MSFE输入特征 F_{in} 的通道数设置为 $C=3$,空间尺寸为 $H \times W$ 。3个并行分支均保持输入、输出通道数一致。其中, 3×3 卷积分支的步长为1、padding为1、空洞率为1; 5×5 卷积分支的步长为1、padding为2、空洞率为1;空洞卷积分支采用 3×3 卷积核,步长为1、空洞率

为 2, padding 为 2。各分支的卷积层后均依次连接批归一化层和 ReLU 激活函数,输出特征的空间尺寸均保持为 $H \times W$ 。

随后,将 3 个分支的输出特征沿通道维度进行拼接,得到通道数为 $3C$ 的多尺度特征,再通过 1×1 卷积将通道数压缩至 C ,并依次进行批归一化和 ReLU 激活,如式(9)所示:

$$F_m = \text{ReLU} \left[\text{BN}_m \left(\text{Conv}_{1 \times 1} \left(\text{Cat}(F_3, F_5, F_d) \right) \right) \right], \quad (9)$$

其中: $\text{Cat}(\cdot)$ 表示通道维度上的特征拼接。 1×1 卷积用于降低通道冗余,并促进不同尺度特征之间的信息交互。 $\text{BN}_m(\cdot)$ 表示融合层的批归一化操作。

最后,引入残差连接以保留输入结构信息,输出特征如式(10)所示:

$$F_{\text{out}} = F_m + F_{\text{in}}. \quad (10)$$

通过上述多分支特征提取、通道融合和残差连接,MSFE 能够在保持较低参数量和和计算开销的同时,增强网络对局部纹理、结构边缘和非均匀光照信息的表征能力,为后续光照估计和反射恢复提供多尺度特征支撑。

2.3 SSA 模块

在 MSA-Net 中,校准网络根据当前阶段恢复得到的反射图 R_t 生成校准图 Δt ,用于修正残余光照偏差并更新下一阶段输入。由于低照度铁路扣件图像中的螺栓轮廓、弹条纹理和扣件边界容易在校准过程中被弱化,因此,校准网络不仅需要修正光照退化,还应保持关键结构信息。原始 SCI 校准网络主要采用标准卷积,对暗区、弱纹理和结构边缘缺乏显式区分能力。

针对上述问题,本文引入低照度结构感知注意力模块 SSA。如图 3(a)所示,SSA 由简单无参数注意力模块(A Simple, Parameter-Free Attention Module, SimAM)分支^[17]、低照度感知分支和边缘结构感知分支组成。给定输入特征 $F \in \mathbb{R}^{C \times H \times W}$, SimAM 分支生成特征显著性响应 A_s ,低照度感知分支根据通道平均亮度生成暗区权重 A_l ,边缘结构感知分支在亮度响应图上利用 Sobel 算子生成边缘权重 A_e 。随后融合 3 类响应,得到结构感知注意力图 A_{SSA} ,用于对输入特征进行重标定。

SimAM 分支通过衡量神经元相对于同通道

特征分布的统计差异生成三维注意力权重,无需引入额外可学习参数。对于铁路扣件特征,轮廓、纹理和边界位置通常具有较明显的局部响应差异, SimAM 分支生成的特征显著性响应如式(11)所示:

$$A_s(c, x, y) = \text{Sigmoid} \left[\frac{(F(c, x, y) - \mu_c)^2}{4(v_c + \lambda)} + \frac{1}{2} \right], \quad (11)$$

其中: μ_c 和 v_c 分别表示输入特征第 c 个通道的均值和方差, λ 为防止分母为零的稳定参数,设置为 10^{-4} 。

低照度感知分支用于描述图像退化程度。首先对输入特征 F 沿通道维度求平均,得到二维亮度响应图 B ,并进行归一化得到 $\hat{B}$,并利用其反向响应生成暗区权重 A_l ,如式(12)~式(14)所示:

$$B(x, y) = \frac{1}{C} \sum_{c=1}^C F(c, x, y), \quad (12)$$

$$\hat{B}(x, y) = \frac{B(x, y) - B_{\min}}{B_{\max} - B_{\min} + \epsilon}, \quad (13)$$

$$A_l(x, y) = 1 - \hat{B}(x, y), \quad (14)$$

其中: $B_{\min}$ 和 $B_{\max}$ 分别表示亮度响应图 B 的最小值和最大值, ϵ 为防止分母为零的极小常数。

边缘结构感知分支在亮度响应图 B 上采用 Sobel 算子提取水平方向和垂直方向的梯度,并计算梯度幅值 G ,如式(15)所示:

$$\begin{cases} G_x = S_x * B \\ G_y = S_y * B \\ G = \sqrt{G_x^2 + G_y^2} \end{cases}, \quad (15)$$

其中: S_x 和 S_y 分别表示水平和垂直方向的 Sobel 卷积核, $*$ 表示卷积操作。

根据梯度响应计算边缘结构权重 A_e ,如式(16)所示:

$$A_e(x, y) = \frac{G(x, y) - G_{\min}}{G_{\max} - G_{\min} + \epsilon}, \quad (16)$$

其中: $G_{\min}$ 和 $G_{\max}$ 分别表示梯度幅值图 G 的最小值和最大值, ϵ 为数值稳定项。

随后,将 A_s , A_l 和 A_e 进行拼接与融合,生成最终的结构感知注意力图 A_{SSA} ,如式(17)所示:

$$A_{\text{SSA}} = \text{Sigmoid}(\text{Conv}_{1 \times 1}(\text{Cat}(A_s, A_l, A_e))), \quad (17)$$

其中: $\text{Cat}(\cdot)$ 表示通道拼接, $\text{Conv}_{1 \times 1}(\cdot)$ 用于融合

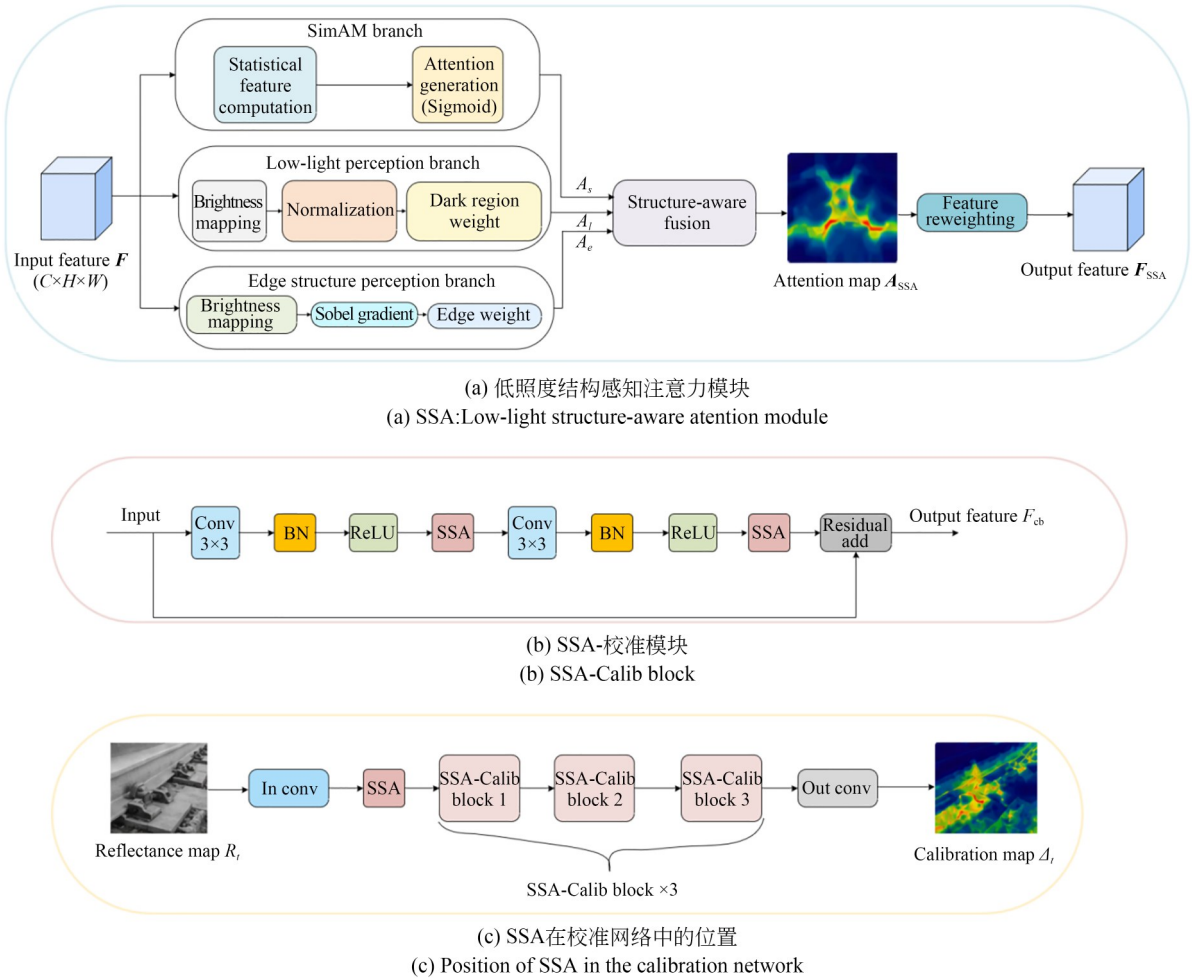


图3 SSA模块结构

Fig. 3 Structure of the SSA module

不同分支的注意力信息。

生成的注意力图用于对输入特征进行重标定,如式(18)所示:

$$F_{SSA} = F \odot A_{SSA}, \quad (18)$$

其中, $\odot$ 表示逐元素相乘。

经过重标定后,输出特征能够同时增强暗区响应、边缘结构响应和扣件关键区域响应。

为进一步提高校准网络的结构恢复能力,本文将SSA嵌入SSA-Calib Block中,如图3(b)所示。该模块由卷积、批归一化、ReLU和SSA操作组成,并采用残差连接,其输出如式(19)所示:

$$F_{cb} = F_{in} + H(F_{in}), \quad (19)$$

其中: F_{cb} 表示SSA-Calib Block的输出特征, F_{in} 表示输入特征, $H(\cdot)$ 表示非线性校准映射。

SSA在校准网络中的位置如图3(c)所示。

在第 t 个自校准阶段,反射图 R_t 首先经过输入卷积层进行特征映射,然后依次经过SSA和3个级联的SSA-Calib Block,以逐步细化校准特征。最后,输出卷积层生成用于下一阶段输入更新的校准图 Δ_t ,如式(20)所示:

$$\Delta_t = C_{out}(B_3(B_2(B_1(SSA(C_{in}(R_t)))))), \quad (20)$$

其中: $C_{in}(\cdot)$ 和 $C_{out}(\cdot)$ 分别表示输入卷积和输出卷积操作, B_1 、 B_2 和 B_3 表示3个连续的SSA-Calib Block。

通过上述设计,SSA将特征显著性、暗区感知和边缘结构信息引入校准网络,使其在修正残余光照偏差的同时保持扣件关键结构,并提高多阶段校准的恢复稳定性。

2.4 ARF模块

低照度铁路扣件图像中不同位置的退化程度通常存在差异。暗光和阴影区域需要较强的

亮度补偿,而亮度较充分或结构边缘清晰的区域应避免过度增强。原始SCI采用固定残差相加方式生成光照图,对所有位置使用相同的融合形式,难以根据局部退化差异调节增强映射的参与程度,容易造成局部欠增强、过增强或结构失真。

为解决上述问题,本文在增强网络之后引入自适应残差融合模块ARF。如图4所示,ARF将

当前阶段输入 X_{t-1} 和增强映射 M_t 沿通道维度拼接,然后通过由卷积层和Sigmoid激活函数组成的门控分支,得到自适应融合权重 α ,如式(21)所示:

$$\alpha = \text{Gate}([X_{t-1}, M_t]), \quad (21)$$

其中: $[X_{t-1}, M_t]$ 表示通道拼接操作, $\text{Gate}(\cdot)$ 表示门控分支, α 表示ARF预测的像素级融合权重,其取值范围为 $[0, 1]$ 。

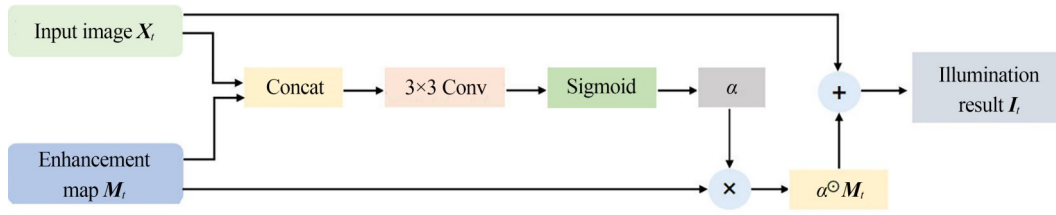


图4 ARF模块结构

Fig. 4 Structure of the ARF module

随后,ARF利用 α 对增强图 M_t 进行加权,并将加权后的增强图与当前输入图像 X_{t-1} 相加,得到当前阶段的光照图 I_t ,如式(22)所示:

$$I_t = X_{t-1} + \alpha \odot M_t, \quad (22)$$

其中: $\odot$ 表示逐元素相乘。自适应融合权重 α 用于调节增强映射 M_t 对光照估计的参与程度。

ARF根据当前输入与增强映射的联合特征生成空间变化的融合权重,使不同位置采用不同的增强映射参与程度。对于暗光、阴影和亮度较充分区域,网络可通过联合调节 M_t 与 α 改变光照估计结果,从而在反射恢复过程中平衡暗区可见性提升与局部过增强抑制。

通过上述自适应残差融合策略,ARF将SCI的固定残差相加改进为空间自适应融合,根据当前输入与增强映射的联合特征调节不同位置的光照估计,并通过反射恢复兼顾暗区增强、过增强抑制与扣件关键结构保持,从而提高增强稳定性。

3 实验设计与分析

3.1 实验设置

为保证实验结果的公平性和可复现性,本文所有实验均在相同的硬件和软件环境下完成,具体配置见表1。实验操作系统为Ubuntu 20.04,

表1 实验环境配置

Tab. 1 Experimental environment configuration

Configuration item	Parameter
CPU	16 vCPU Intel(R) Xeon(R) Gold 6430
GPU	NVIDIA GeForce RTX 4090
GPU memory	24 GB
Deep learning framework	PyTorch 1.11.0
CUDA	11.3

开发环境为Python 3.8。铁路扣件图像尺寸为 640×320 pixels(宽 $\times$ 高),训练过程中不进行裁剪或缩放,仅将RGB像素值归一化至 $[0, 1]$ 。MSA-Net采用Adam优化器进行训练,初始学习率为0.0003,动量参数 β_1 和 β_2 分别设置为0.9和0.999,权重衰减系数为 3×10^{-4} ,训练轮数为100,批量大小为1,训练阶段的自校准阶段数设置为 $T=3$ 。

3.2 评价指标

为全面评价不同方法的图像增强性能,本文采用标准差(Standard Deviation, SD)、平均梯度(Average Gradient, AG)、信息熵(Information Entropy, IE)、自然图像质量评价指标(Natural Image Quality Evaluator, NIQE)和盲/无参考图

像空间质量评价指标 (Blind/Referenceless Image Spatial Quality Evaluator, BRISQUE)。上述指标分别从图像对比度、边缘清晰度、信息丰富度、自然性和失真程度等方面评价增强结果。

标准差 SD 用于反映图像灰度分布的离散程度,其值越大,通常表示图像灰度分布越分散、对比度越高。设待评价灰度图像为 I , 图像尺寸为 $H \times W$ 。其计算方法如式 (23) 所示:

$$AG = \frac{1}{(H-1)(W-1)} \sum_{i=1}^{H-1} \sum_{j=1}^{W-1} \sqrt{\frac{[I(i+1,j)-I(i,j)]^2 + [I(i,j+1)-I(i,j)]^2}{2}}, \quad (24)$$

其中: H, W 和 $I(i, j)$ 的含义同式 (23); i 和 j 分别表示像素的行、列索引。AG 值越大, 通常表示图像的边缘和纹理变化越明显, 结构细节越清晰。

信息熵 IE 用于描述图像灰度信息的丰富程度, 其值越大通常表示图像包含的信息越丰富, 但噪声、背景干扰和过增强伪影也可能造成信息熵升高, 因此需要结合其他指标和视觉结果进行综合分析。其计算方法如式 (25) 所示:

$$IE = - \sum_{k=0}^{L-1} p_k \log_2 p_k, \quad (25)$$

其中: L 表示图像灰度级数, p_k 表示灰度级 k 在图像中出现的概率, 当 $p_k=0$ 时对应项取 0。

NIQE 通过计算待评价图像与自然图像模型之间的统计分布距离评价图像自然性, 其值越小, 表示待评价图像越符合自然图像的统计规律。其计算方法如式 (26) 所示:

$$NIQE = \sqrt{(\mu_n - \mu_t)^T \left(\frac{\sum_n + \sum_t}{2} \right)^{-1} (\mu_n - \mu_t)}, \quad (26)$$

其中: μ_n 和 $\sum_n$ 分别表示自然图像统计特征的均值向量和协方差矩阵, μ_t 和 $\sum_t$ 分别表示待评价图像统计特征的均值向量和协方差矩阵。

BRISQUE 首先提取图像局部归一化亮度的自然场景统计特征, 再利用支持向量回归模型预测图像质量分数, 其值越小, 通常表示图像失真程度越低、感知质量越好。其计算方法如式 (27) 所示:

$$\begin{cases} BRISQUE(x) = \sum_{q=1}^{N_s} a_q k(x, s_q) + b \\ k(x, s_q) = \exp[-\gamma \|x - s_q\|_2^2] \end{cases}, \quad (27)$$

其中: x 表示待评价图像的自然场景统计特征向

$$\begin{cases} SD = \sqrt{\frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W [I(i, j) - \mu]^2} \\ \mu = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W I(i, j) \end{cases}, \quad (23)$$

其中: $I(i, j)$ 表示位置 (i, j) 处的灰度值, μ 表示图像平均灰度。

平均梯度 AG 用于衡量相邻像素之间的平均灰度变化, 其计算方法如式 (24) 所示:

量, s_q 表示第 q 个支持向量, a_q 和 b 分别表示回归系数和偏置项, $k(\cdot)$ 表示径向基核函数。

为评价图像增强对下游铁路扣件检测性能的影响, 本文采用精确率 (Precision)、召回率 (Recall)、交并比 (Intersection over Union, IoU) 阈值为 0.5 时的平均精度均值 (mean Average Precision, mAP@0.5) 以及 IoU 阈值为 0.5~0.95 时的平均精度均值 (mAP@0.5:0.95) 进行评价。

Precision 和 Recall 越大, 通常表示误检和漏检越少。其计算分别如式 (28) 和式 (29) 所示:

$$Precision = \frac{TP}{TP + FP}, \quad (28)$$

$$Recall = \frac{TP}{TP + FN}, \quad (29)$$

其中: TP, FP 和 FN 分别表示正确检测、误检和漏检的目标数量。

指标 mAP@0.5 和 mAP@0.5:0.95 均为数值越大, 检测性能越好。设检测类别总数为 C , 其计算分别如式 (30) 和式 (31) 所示:

$$mAP@0.5 = \frac{1}{C} \sum_{c=1}^C AP_c(0.5), \quad (30)$$

$$mAP@0.5:0.95 = \frac{1}{10C} \sum_{c=1}^C \sum_{k=0}^9 AP_c(0.50 + 0.05k), \quad (31)$$

其中: $AP_c(0.5)$ 表示第 c 类目标在 IoU 阈值为 0.5 时的平均精度。 $k=0, 1, \dots, 9$, 对应 0.50~0.95、步长为 0.05 的 10 个 IoU 阈值。

在模型计算效率方面, 采用参数量 (Parameters)、十亿次浮点运算量 (Giga Floating-Point Operations, GFLOPs)、单张图像推理时间和每秒处理帧数 (Frames Per Second, FPS) 进行评价。参数量以百万为单位, 记为 Parameters/M, 主要反映模型的存储开销; GFLOPs 表示模型处

理单幅图像、完成一次前向推理所需的十亿次浮点运算量,用于衡量模型的理论计算复杂度。单张图像推理时间表示模型处理一幅图像所需的平均时间,FPS表示模型每秒能够处理的图像数量。Parameters/M,GFLOPs和推理时间通常越小越好,FPS越大表示实时处理能力越强。

3.3 数据集

本文实验数据采集于中国福州滨海快线某区段,主要用于研究实际车载检测场景下的低照度铁路扣件图像增强问题。图像采集装置如图5所示,车载成像设备安装于列车底部轨检梁,成像距离约为500 mm。系统采用DALSA LA-GM-02K08A工业线扫相机,相机分辨率为 2048×1 ,并配合车载补充光源,在列车运行过程中连续采集铁路扣件区域的线图像。



图5 铁路扣件图像采集装置

Fig. 5 Railway fastener image acquisition device

在线扫成像过程中,相机沿列车运行方向连续采集一维线数据,并按照采集顺序进行累积,形成二维轨道图像。随后提取铁路扣件区域,得到尺寸为 640×320 pixels(宽 $\times$ 高)的低照度扣件图像,用于模型训练、图像增强评价和下游检测实验。采集图像包含亮度不足、补光不均和局部阴影等典型低照度退化情况。

构建的低照度铁路扣件数据集共包含2 000张图像,其中1 800张用于增强模型训练,200张作为独立测试集,用于评价各方法在未见图像上的增强性能。MSA-Net训练仅使用低照度图像,

不使用配对正常光照图像或目标框标注作为监督信息,独立测试集不参与模型训练。

为开展下游铁路扣件检测实验,本文另采集1 100张正常光照图像,其中1 000张用于检测器训练,100张用于模型验证。为保持不同光照条件下检测器的训练和验证样本数量一致,从上述1 800张低照度训练图像中选取1 100张,其中1 000张用于低照度检测器训练,100张用于模型验证。此外,对200张独立低照度测试图像进行处理,并将其作为不同检测与增强策略的统一测试集。检测训练集、验证集与统一测试集之间不存在图像重叠。

3.4 模块有效性与内部响应分析

3.4.1 模块消融与计算效率分析

为验证MSA-Net各组成模块的作用,本文以SCI为基线,分别引入MSFE,SSA和ARF进行消融实验。所有配置采用相同的数据划分、训练参数和测试条件,定量结果见表2,指标对比见图6。

由表2可知,3个单模块均能改善SCI的增强性能。其中,MSFE的综合提升最为明显,使SD,AG和IE分别由19.04,11.55和3.84提高至32.18,21.36和4.29,并将NIQE降低至23.74,表明多尺度特征提取有助于改善图像灰度分布、梯度响应和信息表达。SSA使SD和AG分别提高至28.64和18.75,并取得单模块中最低的BRISQUE值26.85,说明结构感知注意力有助于增强结构响应并改善图像自然性。ARF的各项指标均优于基线,表明空间自适应融合能够优化增强映射在光照估计中的参与程度。

双模块组合进一步体现了各模块之间的互补性,其中MSFE+SSA取得最佳双模块结果。完整MSA-Net的SD,AG和IE分别达到46.39,30.49和4.88,NIQE和BRISQUE分别降低至21.36和26.34,5项指标均为最优,说明3个模块能够协同提升图像的亮度恢复、信息表达和增强自然性。

在计算效率方面,按测试阶段的实际增强路径统计,MSA-Net的参数量由SCI的0.006 M增加至0.016 M,GFLOPs由9.191增加至9.782,增幅约为6.43%,处理速度仍达到302.39 FPS。由于SSA仅参与训练阶段的自校准,不增加测试

阶段的计算开销,因此所提方法在增强性能、模型复杂度和实时处理效率之间取得了较好的平

衡。图 6 的质量指标变化与表 2 一致,进一步验证了各模块的有效性和互补性。

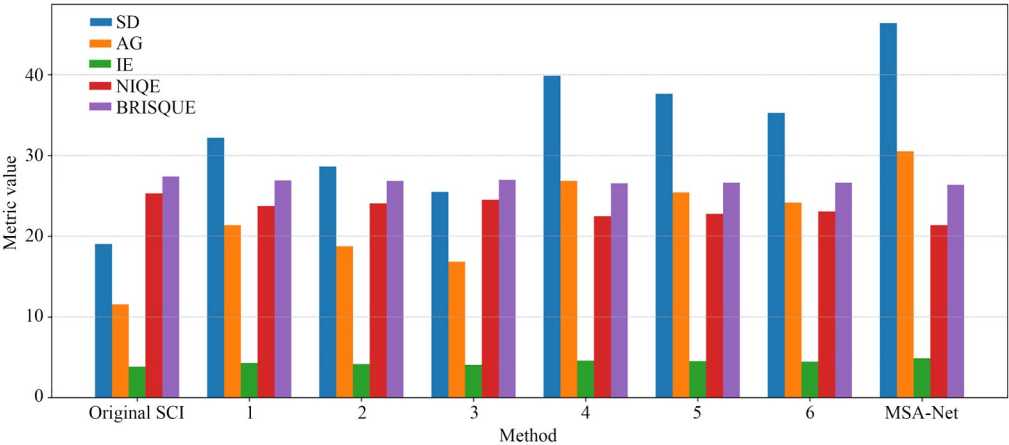


图 6 图像质量评价指标上的模块消融结果

Fig. 6 Module ablation results on image quality assessment metrics

表 2 不同模块的消融实验结果

Tab. 2 Ablation results of different modules

Method	MSFE	SSA	ARF	SD	AG	IE	NIQE	BRISQUE	Params/M	GFLOPs	FPS
Original SCI				19.04	11.55	3.84	25.28	27.38	0.006	9.191	326.90
1	✓			32.18	21.36	4.29	23.74	26.92	0.012	9.636	312.42
2		✓		28.64	18.75	4.17	24.06	26.85	0.006	9.191	326.90
3			✓	25.47	16.82	4.05	24.52	26.98	0.008	9.303	320.45
4	✓	✓		39.86	26.83	4.58	22.48	26.56	0.012	9.636	312.42
5	✓		✓	37.62	25.41	4.50	22.76	26.61	0.016	9.782	302.39
6		✓	✓	35.27	24.16	4.43	23.05	26.63	0.008	9.303	320.45
MSA-Net	✓	✓	✓	46.39	30.49	4.88	21.36	26.34	0.016	9.782	302.39

3.4.2 模块特征响应与关键结构增强分析

为直观分析各改进模块在低照度铁路扣件图像增强中的作用,本文对多尺度特征提取模块(Multi-Scale Feature Extraction, MSFE)、低照度结构感知注意力模块(Structure-Sensing Attention, SSA)和自适应残差融合模块(Adaptive Residual Fusion, ARF)的中间响应进行可视化。对于多通道特征,采用通道绝对响应均值生成二维响应图,计算方法如式(32)所示:

$$M(x,y)=\frac{1}{C}\sum_{c=1}^C|F_c(x,y)|, \tag{32}$$

其中: $F_c(x,y)$ 为第 c 个通道在位置 (x,y) 处的响应, C 为通道数。

图 7 给出了 MSFE 三个分支、融合特征及输

出结果的响应。为突出各分支响应的空间分布,图中响应图分别归一化至 $[0,1]$,颜色表示同一响应图内部的相对强弱。 3×3 卷积分支主要响应螺栓、弹条等区域的局部纹理和边缘; 5×5 卷积分支通过连续密集采样,形成较完整的扣件轮廓与光照响应;膨胀率为 2 的 3×3 空洞卷积分支在等效 5×5 覆盖范围内进行稀疏采样,以较少参数补充不同空间位置的上下文信息。经过多尺度融合,各分支响应得到互补,为后续光照估计和结构保持提供同时包含局部细节与空间上下文的特征表示。

为分析训练阶段 SSA 不同分支对铁路扣件关键结构的响应特性,选取螺栓轮廓、弹条纹理和扣件边界 3 类代表性区域,基于校准网络输入

特征 F , 分别生成特征显著性响应 A_s 、暗区权重 A_l 和边缘结构权重 A_e 。随后融合得到结构感知注意力图 A_{SSA} , 并对输入特征进行重标定, 得到输出特征 F_{out} 。为便于比较, 各响应图均归一化至 $[0, 1]$, 结果如图 8 所示。

图 8 中第 1~3 行分别对应螺栓轮廓、弹条纹理和扣件边界区域。特征显著性响应 A_s 主要表征局部高响应结构, 暗区权重 A_l 补充亮度退化信息,

边缘结构权重 A_e 则响应轮廓和边界变化。三个分支融合后, 注意力图 A_{SSA} 综合不同响应信息, 使重标定特征 F_{out} 中的螺栓轮廓、弹条纹理和扣件边界得到相对突出, 同时抑制部分均匀背景响应。

上述结果表明, SSA 能够综合利用特征显著性、暗区退化和边缘结构信息, 引导训练阶段的校准网络关注低照度区域及扣件关键结构, 从而为增强网络参数优化供给结构约束。

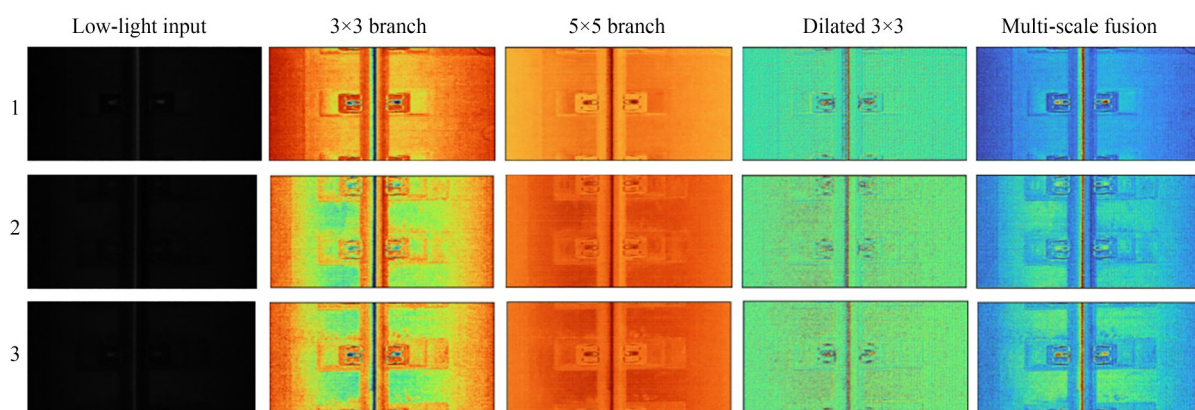


图 7 MSFE 多尺度特征响应可视化

Fig. 7 Visualization of multi-scale feature responses of MSFE

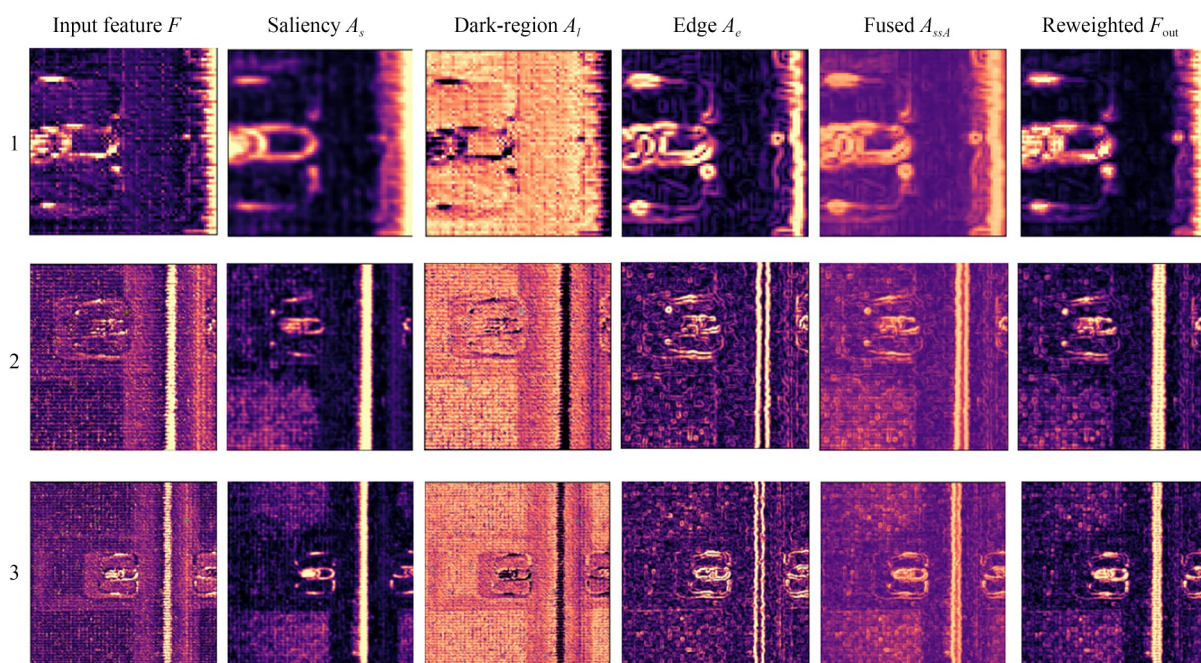


图 8 SSA 训练阶段三支响应、融合注意力及特征重标定结果

Fig. 8 Three-branch responses, fused attention, and feature recalibration results of SSA during training

为分析 ARF 融合权重与图像亮度分布之间的关系, 首先采用 BT. 601 亮度转换关系计算输

入图像的亮度分量 Y , 如式 (33) 所示:

$$Y = 0.299R + 0.587G + 0.114B, \quad (33)$$

其中: R, G 和 B 分别表示归一化后的红、绿、蓝通道值。

根据每幅输入图像的亮度分布,将亮度不高于第 30 百分位数的像素划分为暗区,处于第 30~70 百分位数的像素划分为中等亮度区,高于第 70 百分位数的像素划分为相对亮区。上述区域划分仅用于模型推理后的分组统计,不参与 ARF 权重预测。

在各区域内分别统计输入亮度、增强映射均值和 ARF 平均融合权重。将融合权重 α_1 固定为 1 时得到的反射恢复结果作为固定融合基线,并与采用 ARF 预测空间权重得到的反射恢复结果进行比较,结果见图 9 和表 3。需要说明的是,图 9 和表 3 中的固定融合输出与 ARF 输出均指反射恢复后的增强结果,而不是光照图 I_1 。

图 9 中第 1~3 行分别对应 3 幅低照度铁路扣

件图像,各列依次为输入图像及退化区域划分、增强映射 M_1 、ARF 融合权重 α_1 、固定融合结果和 ARF 输出。第 1 列中的蓝色、橙色和红色分别表示暗区(Dark region, D)、中等亮度区(Moderate-luminance region, M)和相对亮区(Bright region, B),其后数值为对应区域内 ARF 权重 α_1 的平均值。为突出局部差异,3 幅权重图统一采用 0.40~0.48 的显示范围(彩图见期刊电子版)。

由图 9 可见,3 幅图像的区域平均权重均呈现暗区最高、相对亮区最低的趋势,其中 D, M 和 B 区域的平均权重分别处于 0.434 9~0.435 0, 0.433 4~0.433 8 和 0.432 0~0.432 4。尽管区域间差异较小,权重图在螺栓和扣件主体附近仍表现出局部变化,说明 ARF 能够根据输入图像与增强映射的联合特征调节不同位置的增强映射参与程度。

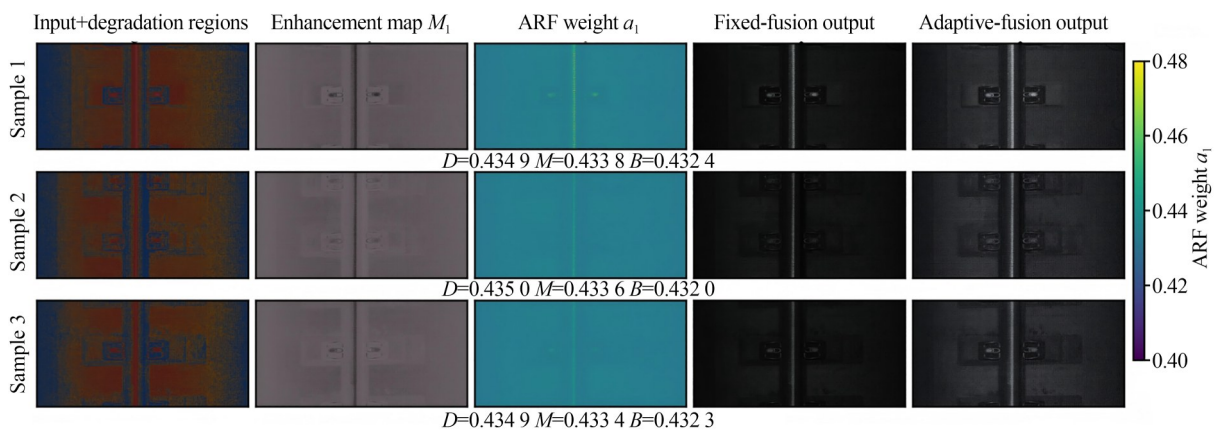


图 9 不同退化区域的 ARF 权重及增强结果

Fig. 9 ARF weights and enhancement results in regions with different degradation levels

表 3 汇总了 3 幅测试图像在不同亮度区域内的统计结果。输入亮度、增强映射及恢复结果亮度均归一化至 $[0, 1]$,相对亮度增益定义为 ARF 恢复结果平均亮度与输入平均亮度之比。随着

区域由暗区变化至相对亮区,增强映射均值由 0.386 82 降低至 0.327 58, ARF 平均权重由 0.434 93 降低至 0.432 23, 相对亮度增益由 5.262 倍降低至 5.198 倍,整体呈轻微递减趋势,

表 3 不同退化区域的 ARF 权重及增强结果统计

Tab. 3 Statistics of ARF weights and enhancement results in regions with different degradation levels

Region	Mean input luminance	Mean enhancement-map value	ARF weight α_1	Fixed-fusion luminance	Adaptive-fusion luminance	Adaptive luminance gain/ $\times$
Dark region	0.018 95	0.386 82	0.434 93	0.046 90	0.099 71	5.262
Medium-luminance region	0.030 96	0.362 38	0.433 60	0.078 88	0.162 23	5.240
Relatively bright region	0.047 50	0.327 58	0.432 23	0.128 84	0.246 90	5.198

反映出 ARF 对不同区域增强映射参与程度的空间调节。

暗区、中等亮度区和相对亮区经 ARF 恢复后的平均亮度分别为 0.099 71, 0.162 23 和 0.246 90, 较固定融合结果分别提高 0.052 81, 0.083 35 和 0.118 06。绝对亮度增量随区域初始亮度升高而增大, 而相对亮度增益在暗区最高、相对亮区最低, 表明 ARF 兼顾了暗区可见性恢复与不同亮度区域的差异化调节。

3.4.3 MSA-Net 中间变量与多阶段反射恢复分析

为分析 MSA-Net 训练阶段增强映射、光照估计、反射恢复和校准更新之间的关系, 本文对第 1 阶段中间变量及 3 个自校准阶段的反射恢复过程进行可视化, 结果如图 10 所示。该分析用于展示多阶段自校准对增强网络参数的优化过程, 不代表测试阶段执行 3 阶段推理。为突出光照分

布差异, 光照图 I_1 单独归一化显示; 校准图采用统一的正负值颜色映射, 其中蓝色和红色分别表示负向与正向校正(彩图见期刊电子版)。

图 10(a)展示了第 1 训练阶段的主要中间计算过程。增强网络首先根据原始低照度输入 X_0 生成增强映射 M_1 , ARF 进一步得到光照图 I_1 , 并通过反射恢复获得 R_1 。随后, 校准网络根据 R_1 生成校准图 Δ_1 , 其响应主要分布在轨缝、螺栓轮廓和扣件边界等亮度或结构变化区域。校准图用于更新下一阶段的训练输入; 图中由 Δ_1 指向 R_3 的箭头表示 Stage 2 和 Stage 3 的后续自校准过程, 并非 Δ_1 直接生成 R_3 。

图 10(b)展示了同一样本在 3 个训练阶段中的反射恢复结果。Stage 1 完成暗区的初步恢复, Stage 2 继续调整非均匀光照和扣件结构响应, Stage 3 进一步减小残余校正, 得到训练阶段第 3 阶段的反射恢复结果 R_3 。

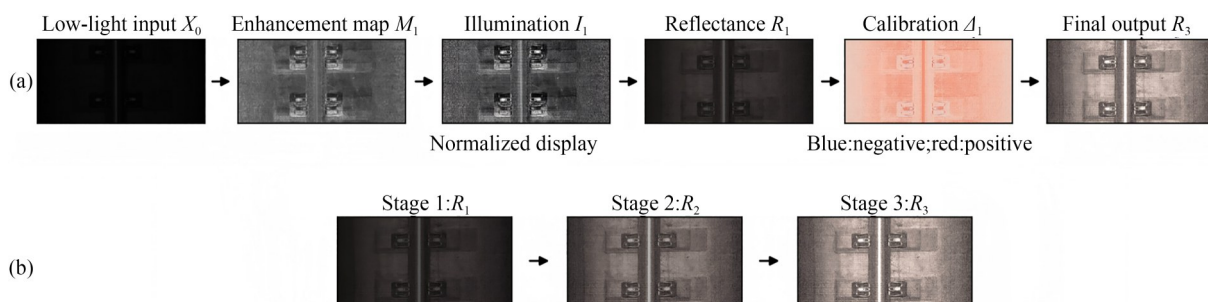


图 10 MSA-Net 训练阶段中间变量及多阶段反射恢复结果

Fig. 10 Intermediate variables and multi-stage reflectance recovery results of MSA-Net during training

为观察训练阶段中间变量的空间响应, 本文选取直接处理原始低照度输入且响应较明显的 Stage 1 进行三维可视化。图 11 分别给出了增强映射 M_1 、光照图 I_1 、反射图 R_1 和校准幅度 $|\Delta_1|$ 的三维响应曲面, 其中横轴和纵轴表示归一化后的图

像空间位置, 竖轴表示响应值。各曲面采用独立的纵轴尺度, 以突出对应变量的空间响应分布。多通道变量采用通道平均值绘制, 校准图以绝对值表示局部校正幅度, 其正、负方向已在图 10 中通过颜色区分(彩图见期刊电子版)。

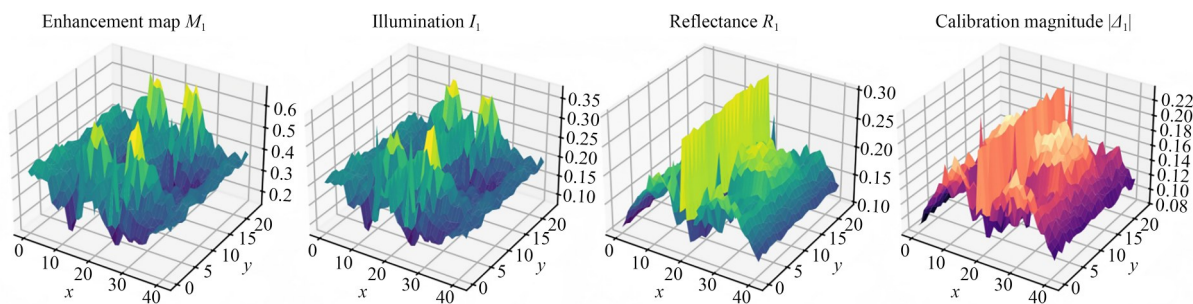


图 11 MSA-Net 训练阶段 Stage 1 中间变量三维响应曲面结果

Fig. 11 Three-dimensional response surfaces of intermediate variables at Stage 1 of MSA-Net during training

由图 11 可见,增强映射 M_i 在低照度背景与扣件主体附近形成非均匀响应;光照图 I_i 表征场景的空间照度变化;反射图 R_i 在轨缝、螺栓轮廓和扣件边界附近呈现较明显的起伏;校准幅度 $|\Delta_i|$ 则在亮度变化和结构边缘附近出现局部峰值。

为减少二维图和三维曲面分析的主观性,本文进一步统计图 10 所示样本在各训练阶段的增强映射 M_i 、光照图 I_i 和反射图 R_i 的均值与标准差,同时计算校准幅度 $|\Delta_i|$ 的均值与最大值,结果见表 4。其中, M_i 、 I_i 和 R_i 的取值范围均为 $[0,1]$ 。

由表 4 可知,随着自校准阶段推进,增强映射 M_i 的均值由 0.374 69 降低至 0.071 76,光照图 I_i 的均值由 0.194 39 降低至 0.062 31,说明后续阶

段所需补充的增强信息和光照修正逐渐减少。与此同时,反射图 R_i 的均值由 0.164 00 依次提高至 0.367 00 和 0.537 97,标准差由 0.054 18 增加至 0.186 32,表明图像亮度和不同区域之间的灰度差异逐步恢复。

校准幅度 $|\Delta_i|$ 的均值由 0.135 37 降低至 0.001 09,最大值由 0.464 86 降低至 0.007 16,说明 Stage 1 承担了较大幅度的初始校正,后续阶段主要修正残余偏差。至 Stage 3,校准响应已接近于 0,表明该样本的多阶段自校准过程趋于稳定。结合图 10 的中间变量、图 11 的空间响应和表 4 的数值统计可知,MSA-Net 通过“增强映射-光照估计-反射恢复-校准更新”的训练过程,逐步优化增强网络的亮度恢复与结构保持能力。

表 4 MSA-Net 训练阶段各阶段中间变量统计

Tab. 4 Numerical statistics of intermediate variables at different training stages of MSA-Net

Stage	M_i (Mean±Std.)	I_i (Mean±Std.)	R_i (Mean±Std.)	Mean of $ \Delta_i $	Maximum of $ \Delta_i $
1	0.374 69±0.087 26	0.194 39±0.076 43	0.164 00±0.054 18	0.135 37	0.464 86
2	0.135 71±0.050 77	0.090 12±0.047 64	0.367 00±0.125 61	0.062 06	0.215 93
3	0.071 76±0.038 83	0.062 31±0.038 32	0.537 97±0.186 32	0.001 09	0.007 16

3.5 对比实验

3.5.1 不同算法的铁路扣件图像增强效果对比

为验证 MSA-Net 在低照度铁路扣件图像增强任务中的有效性,选取 Retinex 启发的展开架构搜索方法 (Retinex-inspired Unrolling with Architecture Search, RUAS)^[18], Zero-DCE^[11]、鲁棒 Retinex 分解网络 (Robust Retinex Decomposition Network, RRDNet)^[19] 和原始 SCI^[12] 进行对比。各方法均在相同的独立测试集上进行评价,视觉

对比结果如图 12 所示,定量结果见表 5。

如图 12 所示,原始低照度图像亮度和对比度较低,螺栓轮廓、弹条纹理及扣件边界难以辨识。RUAS 能够提升暗区亮度,但部分结构边缘仍较模糊;Zero-DCE 的亮度提升较明显,但存在背景同步增亮和灰度层次减弱现象;RRDNet 能够恢复部分灰度信息,但局部区域仍然偏暗;SCI 改善了图像可见性,但对弱纹理和细小结构的恢复仍显不足。

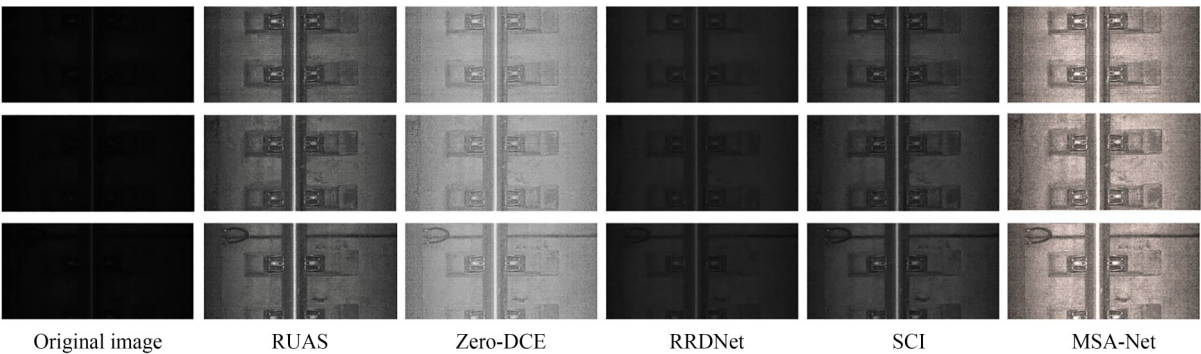


图 12 不同方法在低照度铁路扣件图像上的视觉对比

Fig. 12 Visual comparison of different methods on low-light railway fastener images

表 5 低照度铁路扣件图像增强结果对比

Tab. 5 Comparison of enhancement results on low-light railway fastener images

Method	SD	AG	IE	NIQE	BRISQUE
RUAS	28.24	17.14	3.54	28.86	40.61
Zero-DCE	26.14	19.22	6.70	24.38	37.20
RRDNet	8.65	7.05	5.13	23.58	42.03
SCI	19.04	11.55	3.84	25.28	27.38
MSA-Net	46.39	30.49	4.88	21.36	26.34

相比之下,MSA-Net在提高整体亮度的同时较好地保持了扣件结构边界,使螺栓轮廓、弹条纹理和扣件边缘更加清晰。在所示样本中,其未出现明显的局部过增强和边缘模糊,在亮度恢复、局部对比度和结构保持之间取得了较好的平衡。

由表 5 可知,MSA-Net 的 SD 和 AG 分别达到 46.39 和 30.49,均为对比方法中的最高值,表明其能够扩大图像灰度分布并增强边缘与纹理响应。其 NIQE 和 BRISQUE 分别为 21.36 和 26.34,均取得最低值,说明增强结果在自然度和整体质量方面表现较好。

Zero-DCE 和 RRDNet 的 IE 分别为 6.70 和 5.13,高于 MSA-Net 的 4.88,但 IE 仅反映灰度信息分布的丰富程度,背景噪声和过增强产生的灰度波动同样可能使其升高。结合图 12 可见,MSA-Net 的 IE 较 SCI 的 3.84 有所提升,同时取得最高的 SD,AG 和最低的 NIQE,BRISQUE,说明其在亮度恢复、细节增强和失真抑制之间取得了较好的综合平衡。

为评价不同增强方法的工程应用能力,本文进一步比较各方法的参数量、计算量和推理效率,结果见表 6。对于 RUAS,Zero-DCE,SCI 和 MSA-Net 等前馈式方法,在相同硬件环境、输入尺寸和计时方式下统计模型推理时间。RRDNet 需要针对每幅输入图像进行迭代优化,其完整处理过程与前馈式方法存在差异,因此仅报告参数量和计算量,不参与推理时间及 FPS 的直接比较。

由表 6 可知,MSA-Net 的参数量为 0.016 M,低于 Zero-DCE 和 RRDNet,仅略高于 RUAS 和 SCI。其计算量为 9.782 GFLOPs,低于 Zero-DCE 和 RRDNet,但略高于 SCI。MSA-Net 的单

幅图像推理时间为 3.307 ms,处理速度达到 302.39 FPS,表明所提方法在提升图像增强质量的同时仍保持了较低的计算开销和较高的推理效率,能够满足铁路扣件图像实时增强需求。

综合图 12 及表 5 和表 6 可知,MSA-Net 在恢复图像亮度的同时,能够较好地保持螺栓轮廓、弹条纹理和扣件边界等关键结构,并在无参考质量指标、模型复杂度和推理效率之间取得较好的平衡。其处理速度达到 302.39 FPS,可为后续铁路扣件检测提供清晰、稳定的低照度增强图像。

表 6 不同增强方法的计算复杂度与推理效率对比

Tab. 6 Comparison of computational complexity and inference efficiency of different enhancement methods

Method	Params/M	GFLOPs	Time/ms	FPS
RUAS	0.003	0.669	2.525	396.04
Zero-DCE	0.079	16.220	2.237	447.03
RRDNet	0.128	26.159	—	—
SCI	0.006	9.191	3.059	326.90
MSA-Net	0.016	9.782	3.307	302.39

3.5.2 不同增强与训练策略下的铁路扣件检测性能

为评价低照度图像增强对下游铁路扣件检测性能的影响,并比较增强预处理与直接使用低照度图像训练检测器两种技术路线,本文分别构建正常光照图像训练检测器和低照度原图训练检测器。两个检测器采用相同的 YOLOv8(You Only Look Once version 8)^[20]模型结构、输入尺寸、训练轮数、优化器及其他训练参数。

首先,固定正常光照图像训练得到的检测器权重,分别输入原始低照度图像以及 RUAS,Zero-DCE,RRDNet,SCI 和 MSA-Net 的增强结果,用于评价不同增强方法在不重新训练检测器条件下对检测性能的影响。

测试时,YOLOv8 输入尺寸统一设置为 640×640 ,批量大小为 1,NMS 交并比阈值为 0.45。置信度阈值设置为 0.25。

如图 13 所示,3 幅原始低照度图像均未产生有效检测框。经过不同方法增强后,YOLOv8 能够恢复对扣件目标的检测,但 RUAS,Zero-DCE,RRDNet 和 SCI 仍存在检测框置信度偏低

的问题。相比之下,MSA-Net在所样本中获得了较完整的检测响应,且检测框置信度相对较

高,说明其增强结果有助于提高低照度扣件结构的可辨识性。

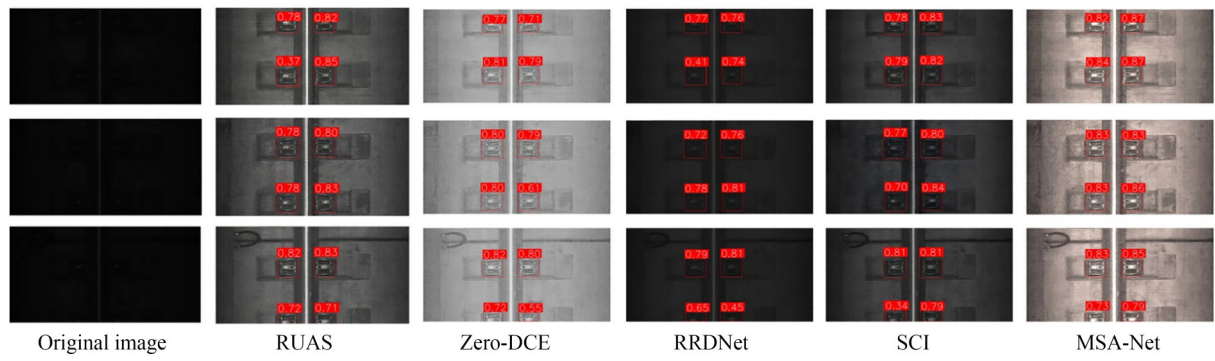


图 13 不同输入图像下铁路扣件检测结果对比

Fig. 13 Detection results of railway fasteners under different input images

由于图 13 仅展示典型样本,进一步采用完整标注测试集上的 Precision, Recall 和 mAP 进行定量评价,结果见表 7。

由表 7 可知,在正常光照图像训练检测器下,各增强方法均能不同程度地改善原始低照度图像的检测性能。其中,MSA-Net 的 Precision, Recall, mAP@0.5 和 mAP@0.5:0.95 分别达到 86.10%, 83.60%, 85.80% 和 59.00%, 较原始低照度输入分别提高 14.70%, 22.80%, 19.10% 和 15.90%。与 RUAS 相比,MSA-Net 的四项指标分别提高 1.60%, 3.40%, 2.40% 和 2.70%, 表明其增强结果能够更有效地改善低照度条件下的目标检出与定位性能。

表 7 正常光照图像检测器在不同输入下的检测性能
Tab. 7 Detection performance with different inputs using the detector trained on normally illuminated images (%)

Test input	Precision	Recall	mAP@	mAP@0.5
			0.5	:0.95
Original image	71.40	60.80	66.70	43.10
RUAS	84.50	80.20	83.40	56.30
Zero-DCE	83.80	78.20	81.50	54.10
RRDNet	78.90	73.10	75.70	49.60
SCI	82.60	79.40	82.90	55.80
MSA-Net	86.10	83.60	85.80	59.00

随后,固定低照度原图训练得到的 YOLOv8 检测器权重,在同一测试集上分别输入原始低照度图像和不同方法的增强结果,用于评价图像增

强对低照度专项训练检测器的影响,结果见表 8。

由表 8 可知, RUAS, Zero-DCE 和 SCI 使检测器的多数指标有所提升,而 RRDNet 的四项指标均略有下降,可能与增强图像和低照度训练图像之间的特征分布差异有关。MSA-Net 的 Precision, Recall, mAP@0.5 和 mAP@0.5:0.95 分别达到 88.10%, 85.40%, 88.60% 和 61.80%, 较原始低照度输入分别提高 1.80%, 2.20%, 2.20% 和 2.50%, 并取得各增强方法中的最佳结果。该结果表明,在本实验条件下,MSA-Net 增强不仅能够适配正常光照图像训练的检测器,也能进一步改善低照度原图训练检测器的目标检出与定位性能。

为评价不同增强方法接入 YOLOv8 后的整体处理效率,固定正常光照图像训练得到的检测器权重,并在同一测试集上分别输入原始低照度

表 8 低照度原图检测器在不同输入图像下的检测性能
Tab. 8 Detection performance of the detector trained on original low-light images under different image inputs (%)

Test input	Precision	Recall	mAP@	mAP@
			0.5	0.5:0.95
Original image	86.30	83.20	86.40	59.30
RUAS	86.74	83.86	87.03	60.08
Zero-DCE	86.96	84.11	87.39	60.47
RRDNet	84.82	82.73	85.95	58.84
SCI	86.28	84.56	87.77	60.92
MSA-Net	88.10	85.40	88.60	61.80

图像和各方法的增强结果。YOLO 检测时间包括图像预处理、网络前向推理和 NMS 后处理,端到端时间为图像增强时间与 YOLO 检测时间之和;原始低照度图像的增强时间记为 0。

由表 9 可知,MSA-Net 的端到端处理时间为 10.099 ms,对应处理速度为 99.02 FPS,虽然略低于 RUAS 和 Zero-DCE,但高于 SCI 及原始低照度输入。RRDNet 在测试阶段需要针对每幅输入图像进行迭代优化,其完整处理流程与前馈式增强方法不同,因此不参与端到端效率的直接比较。结合表 7 的检测结果,MSA-Net 在取得最佳 Precision,Recall 和 mAP 的同时仍保持接近 100 FPS 的端到端处理速度,表明其在检测性能与实时处理效率之间取得了较好的平衡。

表 9 不同增强输入下 YOLO 检测及端到端处理效率
Tab.9 YOLO detection and end-to-end processing efficiency under different enhanced inputs

Test input	YOLO detection time/ ms	End-to-end processing time/ms	End-to-end FPS
Original image	12.700	12.700	78.74
RUAS	6.987	9.512	105.13
Zero-DCE	7.327	9.564	104.56
RRDNet	7.544	-	-
SCI	7.201	10.260	97.47
MSA-Net	6.792	10.099	99.02

3.5.3 不同注意力机制的增强性能对比

为验证 SSA 在训练阶段对铁路扣件细粒度结构的保持作用,本文开展不同注意力机制的替换对比实验。在保持 MSFE,ARF 及网络其他结构不变的条件下,将校准网络中的 SSA 分别替换

为卷积块注意力模块(Convolutional Block Attention Module,CBAM)^[21]、纹理注意力模块(Texture Attention Module,TAM)^[22]和照明引导多头自注意力(Illumination-Guided Multi-head Self-Attention,IG-MSA)^[23]。各组实验采用相同的数据划分、损失函数、训练参数和测试条件,仅替换注意力模块,增强性能结果见表 10。

由表 10 可知,不同注意力机制对增强网络的训练结果存在差异。TAM 取得最低的 NIQE 值 19.41,但其 SD,AG 和 IE 相对较低,BRISQUE 为 32.71;CBAM 在对比度和细节增强方面表现较好,但各项指标仍低于 SSA;IG-MSA 的 IE 达到 4.63,但 BRISQUE 升至 41.48。采用 SSA 时,SD,AG 和 IE 分别达到 46.39,30.49 和 4.88,均为最高值,BRISQUE 降至最低的 26.34。虽然 SSA 的 NIQE 并非最优,但在灰度分布、边缘细节、信息表达和整体失真控制方面取得了更均衡的结果,表明 SSA 能够在训练阶段为增强网络提供有效的结构感知约束。

表 10 不同注意力机制的增强性能对比
Tab.10 Comparison of enhancement performance using different attention mechanisms

Attention	SD	AG	IE	NIQE	BRISQUE
CBAM	39.37	28.70	4.54	21.40	27.05
TAM	34.45	18.99	4.19	19.41	32.71
IG-MSA	36.41	20.86	4.63	20.37	41.48
SSA	46.39	30.49	4.88	21.36	26.34

由图 14 可见,CBAM 能够增强扣件区域的局部响应,但背景颗粒较为明显;TAM 对纹理和噪声具有一定调节作用,但部分区域的结构对比度仍然偏低;IG-MSA 改善了整体光照均匀性,

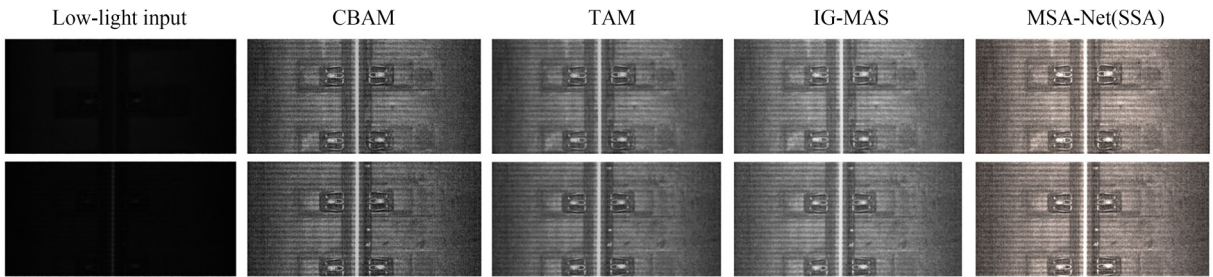


图 14 不同注意力机制下铁路扣件图像增强结果对比
Fig. 14 Comparison of railway fastener image enhancement results using different attention mechanisms

但部分细小边缘略有平滑。相比之下,采用 SSA 训练得到的增强结果具有较清晰、连续的螺栓轮廓、弹条纹理和扣件边界,同时保持了较均衡的亮度分布。结合表 10 的定量结果可知,SSA 能够在训练阶段引导校准网络关注低照度区域及扣件关键结构,从而提高增强网络的结构保持能力和增强稳定性。

3. 5. 4 LOL 数据集上的增强效果验证

为验证 MSA-Net 在通用低照度场景中的适用性,本文选取公开低照度图像数据集 (Low-Light dataset, LOL)^[24] 中的 485 张低照度图像,并重新划分为训练集和测试集,其中 400 张用于训练,85 张用于测试。MSA-Net, RUAS, Zero-DCE 和 SCI 均采用相同的数据划分重新训练,并在统一测试集上进行评价;RRDNet 按照原始方法对每幅测试图像进行自适应优化。

由表 11 可知,MSA-Net 的 SD 和 AG 分别达到 47.38 和 16.74,均为对比方法中的最高值,表明其能够扩大图像灰度分布并增强边缘与纹理响应。其 IE 为 6.94,仅低于 SCI 的 7.06;NIQE 取得最低值 8.18, BRISQUE 为 27.93,仅高于 Zero-DCE。综合各项指标与图 15 的视觉结果,MSA-Net 在亮度恢复、细节增强和视觉自然性之间取得了较好的平衡,表现出一定的通用低照度

场景适应能力。

表 11 LOL 测试集上不同增强方法的定量对比结果

Tab. 11 Quantitative comparison results of different enhancement methods on the LOL test set

Method	SD	AG	IE	NIQE	BRISQUE
RUAS	35.38	11.93	6.86	8.73	29.62
Zero-DCE	29.24	12.41	6.36	8.93	24.89
RRDNet	20.60	3.40	6.03	17.75	28.97
SCI	41.61	14.35	7.06	8.23	30.49
MSA-Net	47.38	16.74	6.94	8.18	27.93

如图 15 所示,原始低照度图像整体亮度不足,暗区结构和纹理难以辨识,图中最后一列为对应的正常光照参考图,可用于比较不同方法在亮度、色彩和细节恢复方面的差异。RUAS 和 Zero-DCE 能够提高图像亮度,但部分区域存在过曝、发白或色彩偏移;RRDNet 的增强结果较为保守,暗区细节恢复有限;SCI 改善了整体可见性,但局部亮度仍不均匀。相比之下,MSA-Net 在恢复暗区信息的同时较好地保持了场景色彩、边缘轮廓和纹理细节,且未出现明显的大面积过曝,在亮度恢复与视觉自然性之间取得了较好的平衡。

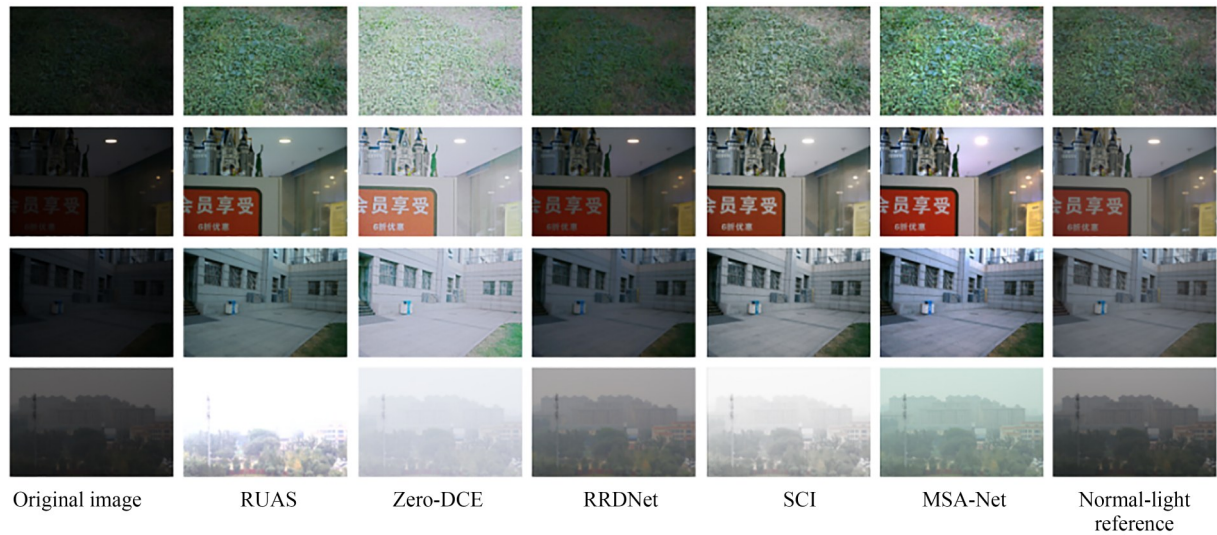


图 15 LOL 数据集上低照度图像增强结果的视觉对比

Fig. 15 Visual comparison of low-light image enhancement results on the LOL dataset

3. 5. 5 LSRW 数据集上的增强效果验证

为进一步验证 MSA-Net 在真实低照度场景

中的适应能力,本文在大规模真实场景低照度/正常光照配对图像数据集 (Large-Scale Real-

World paired low/normal-light images dataset, LSRW)^[25] 的华为子集上开展对比实验。本文选取其中 2 450 张低照度图像并重新划分,其中 2 000 张用于训练,450 张用于测试,训练集与测试集之间不存在图像重叠。

无监督增强融合网络(Unsupervised Boosted Fusion Network,UBF-Net)^[26]、零参考相机响应网络(Zero-Reference Camera Response Network,ZRCRN)^[27],RUAS,Zero-DCE,SCI 和 MSA-Net 均采用相同的数据划分,并按照各自的训练策略重新训练;非正则化自监督 Retinex 方法(Non-Regularization Self-Supervised Retinex Approach, NeurBR)^[28] 和 RRDNet 无需数据集级预训练,按照原始方法对测试图像逐幅进行自适应优化。所有方法均不使用正常光照图像作为监督信息,并在相同的 450 张低照度图像上进行评价。

由表 12 可知,MSA-Net 的 SD 达到 57.87,为所有方法中的最高值,表明增强图像具有较充分的灰度分布和明暗对比。其 AG 和 IE 分别为 11.36 和 7.17,均居次优,其中 AG 仅低于 Zero-DCE,IE 仅低于 NeurBR,说明该方法能够较好地保留边缘与信息内容。MSA-Net 的 NIQE 为 7.39,优于 RUAS,Zero-DCE,RRDNet 和 SCI;BRISQUE 为 21.66,低于 NeurBR,UBF-Net,

RUAS,Zero-DCE 和 RRDNet,且仅略高于 ZRCRN。综合各项指标,MSA-Net 在对比度提升、细节保持和感知质量之间取得了较好的平衡。

表 12 LSRW 测试集上不同增强方法的定量对比结果
Tab. 12 Quantitative comparison results of different enhancement methods on the LSRW dataset

Method	SD	AG	IE	NIQE	BRISQUE
NeurBR	51.66	7.12	7.23	6.83	24.91
UBF-Net	42.04	6.97	7.01	7.04	22.56
ZRCRN	37.25	5.70	6.79	7.16	21.45
RUAS	56.33	11.27	5.75	8.00	30.59
Zero-DCE	45.38	11.79	6.93	8.14	29.23
RRDNet	38.19	4.44	6.74	8.63	22.06
SCI	47.17	6.81	6.72	7.45	18.41
MSA-Net	57.87	11.36	7.17	7.39	21.66

由图 16 可见,NeurBR,UBF-Net 和 ZRCRN 能够改善暗区可见性,但部分场景仍存在亮度不足或色彩偏移;RUAS 和 Zero-DCE 的亮度提升较明显,但易造成高亮区域过曝和局部细节丢失;RRDNet 与 SCI 增强较为保守,部分暗区信息仍不够清晰。相比之下,MSA-Net 在提高整体亮度的同时较好地抑制了高亮区域过增强,表现出较好的视觉自然性和场景适应能力。

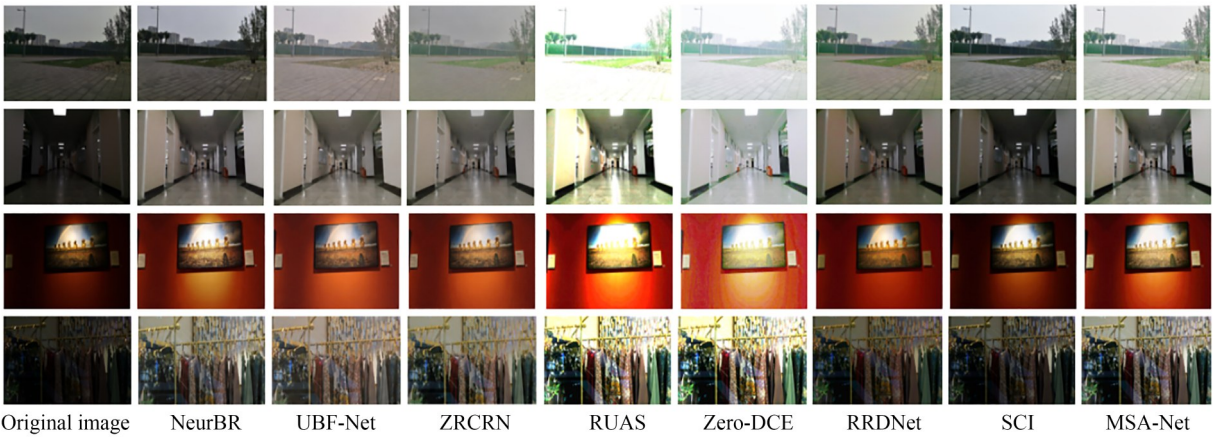


图 16 LSRW 数据集上低照度图像增强结果的视觉对比

Fig. 16 Visual comparison of low-light image enhancement results on the LSRW dataset

4 结 论

针对车载铁路巡检中低照度、补光不均和局

部阴影导致扣件图像亮度不足、细节模糊及关键结构难以辨识的问题,本文提出多尺度结构感知低照度图像增强网络 MSA-Net。在 SCI 自校准

增强框架基础上,引入多尺度特征提取模块、低照度结构感知注意力模块和自适应残差融合模块,分别利用不同感受野提取光照与结构特征、强化螺栓轮廓、弹条纹理和扣件边界等关键结构响应,并依据区域退化程度自适应调节增强强度,使网络在恢复图像亮度的同时兼顾细节保持与过增强抑制。

在福州滨海快线低照度铁路扣件数据集上的实验结果表明,MSA-Net的SD,AG,IE,NIQE和BRISQUE分别为46.39,30.49,4.88,21.36和26.34。消融实验、模块响应分析及注意力机制替换实验进一步验证了MSFE,SSA和ARF在亮度恢复、结构保持和增强稳定性方面的有效性。该模型参数量为0.016 M,计算量为9.782 GFLOPs,单幅图像推理时间为3.307 ms,处理速度达到302.39 FPS,在提高质量、模型复杂度和推理效率之间取得了较好的平衡。

在下游铁路扣件检测实验中,采用正常光照图像训练的YOLOv8检测器时,MSA-Net增强输入的Precision,Recall,mAP@0.5和mAP@

0.5:0.95分别达到86.10%,83.60%,85.80%和59.00%;采用低照度原图训练的检测器时,上述指标分别达到88.10%,85.40%,88.60%和61.80%,均为各对比输入中的最佳结果。在采用正常光照图像训练检测器的端到端效率实验中,MSA-Net与YOLOv8串联后的端到端处理速度为99.02 FPS,表明该方法能够在保持实时处理能力的同时改善低照度条件下扣件目标的检出与定位性能。在LOL和LSRW数据集上的实验结果进一步表明,MSA-Net对通用低照度场景具有一定的适应能力。

后续研究将扩充不同线路、光照条件和运行环境下的现场数据,并结合定位误差、异常识别精度及长期运行稳定性,进一步评估增强方法在实际铁路巡检系统中的应用效果。

作者贡献声明:

马思维:设计实验及初稿写作;
胡贺南:算法提出,论文构思;
李荣华:论文审核与编辑写作。

参考文献:

- [1] 沙伯男,白堂博,许贵阳,等.基于改进YOLOv9的铁路扣件状态检测方法[J].北京交通大学学报,2025,49(6):75-84.
SHA B N, BAI T B, XU G Y, et al. Railway fastener condition detection method based on improved YOLOv9[J]. *Journal of Beijing Jiaotong University*, 2025, 49(6): 75-84. (in Chinese)
- [2] 袁小翠,王咏涛,刘宝玲,等.基于3D线激光传感器的轨道弹条扣件结构缺陷检测方法[J].红外与激光工程,2024,53(7):20240176.
YUAN X C, WANG Y T, LIU B L, et al. Detection method for structural defects of railway clip fastener based on 3D line laser sensor[J]. *Infrared and Laser Engineering*, 2024, 53(7): 20240176. (in Chinese)
- [3] 蒋占军,吴佰靖,马龙,等.基于金字塔边缘增强的自矫正低光照目标检测[J].光学精密工程,2024,32(20):3099-3111.
JIANG Z J, WU B J, MA L, et al. A self correcting low-light object detection method based on pyramid edge enhancement[J]. *Opt. Precision Eng.*, 2024, 32(20): 3099-3111. (in Chinese)

- [4] HAN Y, CHEN X Y, ZHONG Y, et al. Low-illumination road image enhancement by fusing retinex theory and histogram equalization[J]. *Electronics*, 2023, 12(4): 990.
- [5] WANG Y L, LIU Z, LIU J Z, et al. Low-light image enhancement with illumination-aware gamma correction and complete image modelling network[C]. 2023 IEEE/CVF International Conference on Computer Vision (ICCV). October 1-6, 2023, Paris, France. IEEE, 2024: 13082-13091.
- [6] LI X F, WANG W W, FENG X C, et al. Deep parametric Retinex decomposition model for low-light image enhancement[J]. *Computer Vision and Image Understanding*, 2024, 241: 103948.
- [7] 郑爽爽,卫文学,徐聪.融合全变分与Gamma的低照度图像增强算法[J].激光与光电子学进展,2023,60(12):1210022.
ZHENG S S, WEI W X, XU C. Low illumination image enhancement algorithm combining total variation and gamma[J]. *Laser & Optoelectronics Progress*, 2023, 60(12): 1210022. (in Chinese)
- [8] GAO X, ZHANG M, LUO J. Low-light image enhancement via Retinex-style decomposition of de-

- noised deep image prior [J]. *Sensors*, 2022, 22 (15): 5593.
- [9] 张箴, 鹿阳, 苏奕铭, 等. 基于多通道 Retinex 模型的低照度图像增强网络[J]. *信息与控制*, 2024, 53 (5): 652-661, 672.
ZHANG Z, LU Y, SU Y M, *et al.* Low-light image enhancement network based on multichannel retinex model[J]. *Information and Control*, 2024, 53 (5): 652-661, 672. (in Chinese)
- [10] WANG Z X, QINGGE L, PAN Q Y, *et al.* Retinex decomposition based low-light image enhancement by integrating Swin transformer and U-Net-like architecture[J]. *IET Image Processing*, 2024, 18(11): 3028-3041.
- [11] GUO C L, LI C Y, GUO J C, *et al.* Zero-reference deep curve estimation for low-light image enhancement [C]. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 13-19, 2020, Seattle, WA, USA. IEEE, 2020: 1777-1786.
- [12] MA L, MA T Y, LIU R S, *et al.* Toward fast, flexible, and robust low-light image enhancement [C]. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 18-24, 2022, New Orleans, LA, USA. IEEE, 2022: 5627-5636.
- [13] 陈清江, 顾媛. 多通道融合注意力网络的低照度图像增强[J]. *光学精密工程*, 2023, 31(14): 2111-2122.
CHEN Q J, GU Y. Low-light image enhancement algorithm based on multi-channel fusion attention network[J]. *Opt. Precision Eng.*, 2023, 31(14): 2111-2122. (in Chinese)
- [14] YUAN N Z, ZHAO X Y, SUN B Y, *et al.* Low-light image enhancement by combining transformer and convolutional neural network [J]. *Mathematics*, 2023, 11(7): 1657.
- [15] 江泽涛, 黄钦阳, 张惠鹏, 等. 基于全局一致的非配对低照度图像增强方法[J]. *计算机研究与发展*, 2025, 62(4): 876-887.
JIANG Z T, HUANG Q Y, ZHANG H J, *et al.* Unpaired low-light image enhancement method based on global consistency [J]. *Journal of Computer Research and Development*, 2025, 62 (4) : 876-887. (in Chinese)
- [16] LIU M L, LI X S, FANG Y. Low-light image enhancement *via* dual information-based networks [J]. *Electronics*, 2024, 13(18): 3713.
- [17] YANG L, ZHANG R Y, LI L, *et al.* SimAM: a simple, parameter-free attention module for convolutional neural networks [C]. *Proceedings of the 38th International Conference on Machine Learning. Proceedings of Machine Learning Research*, 2021, 139: 11863-11874.
- [18] LIU R S, MA L, ZHANG J A, *et al.* Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement [C]. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 20-25, 2021, Nashville, TN, USA. IEEE, 2021: 10556-10565.
- [19] ZHU A Q, ZHANG L, SHEN Y, *et al.* Zero-shot restoration of underexposed images *via* robust retinex decomposition [C]. 2020 *IEEE International Conference on Multimedia and Expo (ICME)*. July 6-10, 2020, London, UK. IEEE, 2020: 1-6.
- [20] 范华琦, 杨柳. 改进 YOLOv8 的轨道扣件状态检测方法 [J]. *国外电子测量技术*, 2024, 43(4): 19-29.
FAN H Q, YANG L. Improved YOLOv8 track fastener state detection method [J]. *Foreign Electronic Measurement Technology*, 2024, 43(4): 19-29. (in Chinese)
- [21] WOO S, PARK J, LEE J Y, *et al.* CBAM: Convolutional Block Attention Module [M]. *Computer Vision-ECCV 2018*. Cham: Springer International Publishing, 2018: 3-19.
- [22] WANG Y, DING Z, SONG J Y, *et al.* Peanut defect identification based on multispectral image and deep learning [J]. *Agronomy*, 2023, 13(4): 1158.
- [23] CAI Y H, BIAN H, LIN J, *et al.* Retinexformer: one-stage retinex-based transformer for low-light image enhancement [C]. 2023 *IEEE/CVF International Conference on Computer Vision (ICCV)*. October 1-6, 2023, Paris, France. IEEE, 2024: 12470-12479.
- [24] WEI C, WANG W J, YANG W H, *et al.* Deep Retinex Decomposition for Low-light Enhancement [EB/OL]. 2018: arXiv: 1808.04560. <https://arxiv.org/abs/1808.04560>
- [25] HAI J, XUAN Z, YANG R, *et al.* R2RNet: low-light image enhancement *via* real-low to real-normal network [J]. *Journal of Visual Communica-*

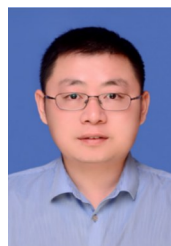
- tion and Image Representation, 2023, 90: 103712.
- [26] ZHANG J F, LI H X, HUO Z Q. Unsupervised boosted fusion network for single low-light image enhancement [J]. *IEEE Access*, 2024, 12: 179252-179264.
- [27] WANG X F, HUANG L, LI M X, *et al.* Fast, zero-reference low-light image enhancement with camera response model [J]. *Sensors*, 2024, 24 (15): 5019.
- [28] ZHAO Z J, LIN H X, SHI D M, *et al.* A non-regularization self-supervised Retinex approach to low-light image enhancement with parameterized illumination estimation [J]. *Pattern Recognition*, 2024, 146: 110025.

作者简介:



马思维(1998—),男,辽宁沈阳人,博士研究生,主要从事方向为机器视觉与深度学习, E-mail: ma_siwei2021@163.com

通讯作者:



李荣华(1983—),男,河北唐山人,博士,教授,博士生导师,2005年于大连交通大学获得学士学位,2011年于大连理工大学获得博士学位,主要从事方向为遥感图像处理、航天器姿态视觉测量, E-mail: lironghua705@163.com