

文章编号: 1671-7449(2026)04-0413-10

J. Test Measur. Technol., 2026, 40(4): 413-422.

基于改进CA-ESRGAN的红外图像增强算法

韩龙*, 何辉煌, 赵雅婷, 陈楚, 马志远

(黑龙江科技大学 电气与控制工程学院, 黑龙江 哈尔滨 150022)

摘要: 针对电力设备在低照度、遮挡、小目标密集等复杂环境中实际巡检时, 红外热成像常存在清晰度差、边缘模糊等问题, 提出一种基于改进通道注意力的超分辨率生成对抗网络(Channel Attention Enhanced Super-Resolution Generative Adversarial Netural, CA-ESRGAN)的红外图像增强算法。首先, 采用线性可变形卷积替代标准卷积, 通过动态调整卷积核采样位置提高模型对复杂巡检环境的适应性。其次, 采用空间-通道协同注意力替换原通道注意力, 增强对局部细节和全局信息的捕捉能力。最后, 引入正则化相对性GAN损失函数优化对抗训练过程, 提高图像质量和训练稳定性。在FLIR公开数据集和变电所采集红外数据集进行实验, 结果表明, 该算法在两个数据集上均取得较好性能, 尤其在变电所采集数据集上, 峰值信噪比、结构相似性和学习感知图像块相似性3项评价指标分别较GTISR算法提升了1.84 dB、0.038 1和0.013 1, 进一步验证了其在复杂实际场景中能够有效提升红外图像质量。

关键词: 红外图像增强; 超分辨率重建; 线性可变形卷积; 协同注意力; 对抗损失

中图分类号: TP391.41

文献标识码: A

DOI: 10.62756/csjs.1671-7449.2026037

引用格式: 韩龙, 何辉煌, 赵雅婷, 等. 基于改进CA-ESRGAN的红外图像增强算法[J]. 测试技术学报, 2026, 40(4): 413-422.

Han Long, He Huihuang, Zhao Yating, et al. Infrared image enhancement algorithm based on improved CA-ESRGAN[J]. Journal of Test and Measurement Technology, 2026, 40(4): 413-422.

Infrared Image Enhancement Algorithm Based on Improved CA-ESRGAN

Han Long*, He Huihuang, Zhao Yating, Chen Chu, Ma Zhiyuan

(School of Electrical & Control Engineering, Heilongjiang University of Science & Technology, Harbin 150022, China)

Abstract: In complex environments commonly encountered during routine inspection of power equipment, such as low illumination, occlusion, and dense distributions of small targets, infrared thermal imaging often suffers from poor clarity and blurred edges. To address these issues, an improved CA-ESRGAN-based infrared image enhancement algorithm is proposed. Firstly, LDConv is used to replace the standard convolution, dynamically adjusting the convolution kernel sampling positions to enhance the model's adaptability to complex inspection conditions. Secondly, SCSA is adopted to replace the original channel attention mechanism, strengthening the capture of both local details and global contextual information. Finally, R-RaGAN is introduced to optimize the adversarial training process, improving image quality and training stability. Experiments conducted on the FLIR public dataset and an infrared dataset collected from a substation demonstrate that well performance on both datasets the is achieved by the proposed algorithm. In particular, on the substation dataset, PSNR, SSIM,

【收稿日期】2025-07-09

【修回日期】2025-08-18

【录用日期】2025-10-21

【基金项目】2024年度黑龙江省省属高等学校基本科研业务费项目(2024-KYYWF-1108); 黑龙江省自然科学基金资助项目(LH2021F051)

*【通信作者】韩龙(1978—), 男, 博士, 教授, 主要从事轴承故障诊断、图像处理研究。E-mail: hljkc_hl@usth.edu.cn。

and *LPIPS* are improved by 1.84 dB, 0.038 1, and 0.013 1, respectively, compared to the best-performing baseline, further confirming its effectiveness in enhancing infrared image quality under complex real-world conditions.

Key words: infrared image enhancement; super-resolution reconstruction; linear deformable convolution; collaborative attention; counteract losses

0 引言

红外热成像技术因其非接触性、直观性和易于分析的特点^[1-3],已广泛应用于工农业等领域^[4],特别是在电力设备巡检的复杂环境中,价值尤为突出。然而,由于受灰尘、辐射等外部环境因素,以及元件噪声等内部干扰,红外图像常表现出低对比度、低分辨率等缺陷^[5]。因此,如何有效改善复杂环境中红外图像的对比度和清晰度^[6],提升其图像质量,已成为该领域重要的研究课题。

针对此类问题,研究者们提出了多种解决方法。传统方法主要有直方图均衡化、多尺度Retinex等^[7],这些方法通过调整像素强度分布或融合多尺度亮度信息来改善图像对比度与细节表现。然而,其在增强整体亮度的同时,往往带来局部过增强、噪声放大等问题。随着机器学习的发展,基于数据引导的图像增强方法逐渐兴起。Wang等^[8]提出了结合语义分割的增强策略,能够在保持全局一致性的同时突出关键区域,从而有效提升图像质量。但此类方法通常依赖人工特征提取,泛化能力和适应性也存在一定不足。

随着深度学习的不断发展,Zhu等^[9]提出的生成式对抗网络(Generative Adversarial Network, GAN)为红外图像超分辨率重建奠定了基础,之后,Sun等^[10-11]将其引入红外图像增强领域,提出了基于GAN的超分辨率生成网络(Super-Resolution Generative Adversarial Network, SRGAN),并取得了显著进展。Rifai等^[12]提出的增强型SRGAN(Enhanced SRGAN, ESRGAN)通过引入残差密集块(Residual-in-Residual Dense Block, RRDB),显著改善了生成图像的真实感与视觉质量,成为领域内重要的基准方法。为提升模型对真实退化图像的适应性,李刚等^[13]提出了真实场景ESRGAN(Real-world ESRGAN, Real-ESRGAN),能够在实际场景中处理复杂退化模式。在此基础上,基于通道注意力的ESRGAN(Channel Attention ESRGAN, CA-ESRGAN)^[14]通过引入通道注意力机制以强化特征

表达能力,从而在细节保真与纹理重建方面进一步提升了生成图像的质量。此外,近年来Transformer结构被逐步引入到红外图像增强任务,其中,引导式热红外图像超分辨率(Guided Thermal Image Super-Resolution, GTISR)^[15]作为其中的代表性方法,通过全局注意力机制提升了对远程依赖信息的建模能力,展现出较强的图像重建性能。随着双鉴别器架构、混合感知损失等被提出,现有基于GAN的超分辨率重建方法逐渐暴露出一些问题,如难以同时建模空间和通道特征;模型训练过程中易出现模式崩溃、收敛不稳定等问题;损失函数无法有效约束生成器学习方向等。

针对上述问题,本文提出一种改进CA-ESRGAN的红外图像增强算法。在特征提取阶段中引入线性可变形卷积(Linear Deformable Convolution, LDConv),通过动态调整卷积核采样位置,提高模型对不同环境的适应性。在生成器中,将原通道注意力替换为空间-通道协同注意力(Spatial and Channel Squeeze-and-Excitation Co-Attention, SCSA),并通过同时建模空间特征和通道特征来提升模型对图像不同尺度特征的感知能力。此外,针对传统GAN模型在训练过程中常出现的模式崩溃和收敛不稳定的问题,本文算法引入正则化相对性生成式对抗网络损失(Regularized Relativistic Generative Adversarial Network Loss, R-RaGAN),以增强损失函数对生成器优化方向的约束力,从而优化对抗训练。

在FLIR公开红外图像数据集及变电站采集红外图像数据集上进行实验,结果表明,本文算法在主观评价和峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)、结构相似性(Structural Similarity Index, SSIM)和学习感知图像块相似性(Learned Perceptual Image Patch Similarity, LPIPS)等客观指标上均优于ESRGAN、Real-ESRGAN、CA-ESRGAN和GTISR算法,能够有效提升红外图像质量。

1 基本原理

本文算法的主要步骤如下:

1) 预处理: 输入低分辨率图像, 引入高阶退化模型对原始图像进行模糊、下采样等操作, 增强对复杂退化模式的适应性。

2) 特征提取: 通过一层 3×3 卷积提取浅层特征, 捕获图像的边缘和局部纹理信息。再利用多层引入LDConv的可快速重新配置的数据库(Rapidly Reconfigurable Date Base, RRDB)提取深层特征。

3) 非线性映射: 将低分辨率特征映射至高分辨率特征空间, 补偿丢失的高频细节。映射过程

中引入SCSA, 同时建模空间与通道维度的特征相关性, 增强模型对目标区域和细节信息的关注。

4) 高频重建: 将映射后的特征通过像素重排模块进行上采样, 恢复至目标分辨率, 生成细节丰富且清晰的高分辨率红外图像。

5) 对抗训练: 采用生成器与判别器的对抗训练进行优化。引入R-RaGAN损失函数, 通过相对性判别和正则化约束, 增强模型的对抗性与稳定性。

算法网络结构图如图 1 所示。

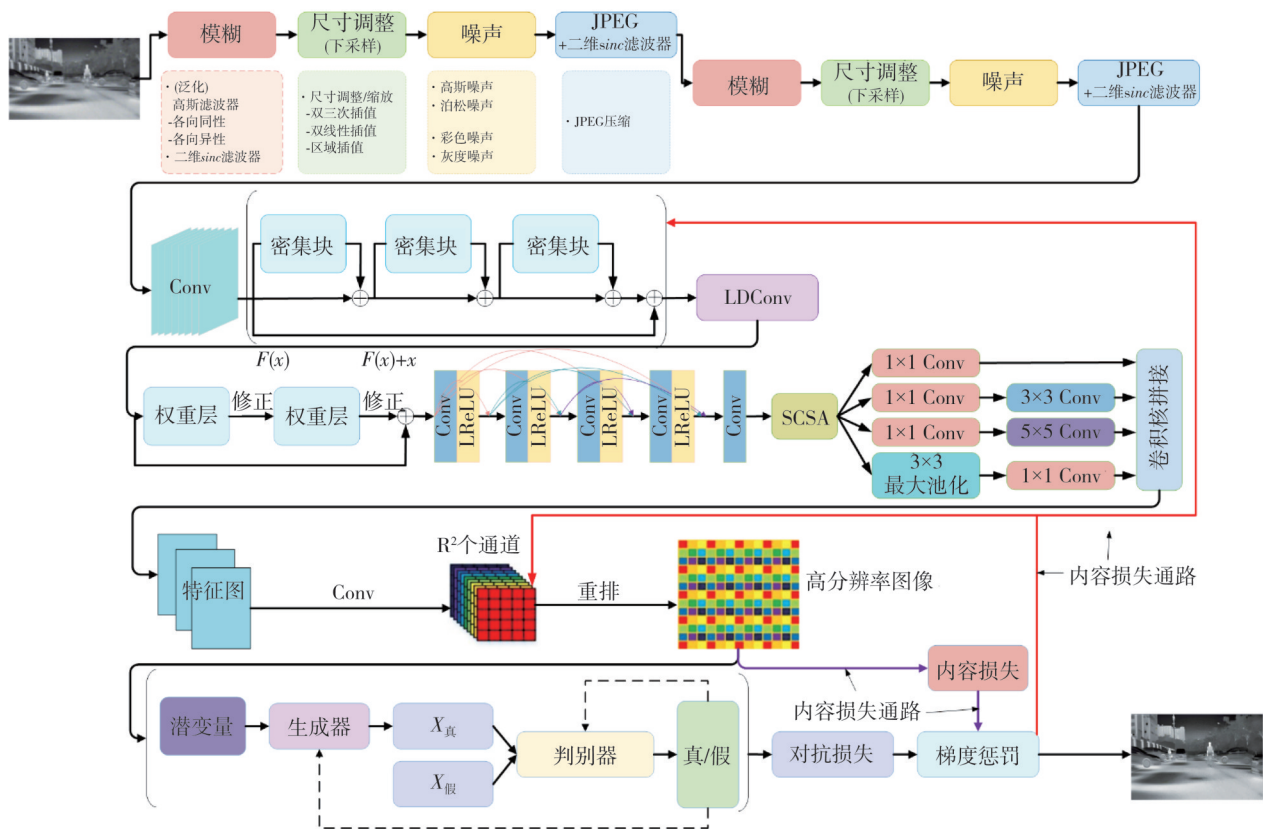


图 1 本文算法结构示意图

Fig. 1 Cchematic diagram of algorithm structure in this paper

1.1 LDConv

标准卷积存在两大局限: 一是固定窗口限制了信息捕捉范围, 难以适应复杂结构变化; 二是卷积核大小固定为 $k \times k$, 参数量随卷积核尺寸的平方增长, 计算开销过大, 影响模型效率。此外, 偏移量的学习过程易受梯度传播影响模型的稳定性。

为提升模型能力, 降低计算复杂度, 本文采用LDConv替代标准卷积, 结构如图 2 所示。LDConv通过动态调整卷积采样形状, 提高了网络对复杂纹理的建模能力。同时, LDConv采用线性增长方式调整卷积核的参数量, 避免了参数随卷积核尺寸平

方增长的情况, 显著降低了计算开销。

LDConv是在数字通信网(Digital Communication Network, DCN)基础上进行的改进, 使得卷积核参数和采样形状更加灵活, 其公式可表示为

$$y(p) = \sum_{p_n \in S} \omega(p_n) \cdot (p + p_n + \Delta p_n), \quad (1)$$

式中: S 为非固定形状采样, 不局限于 $k \times k$ 窗口; Δp_n 为学习得到的动态偏移量, 可进行优化; $\omega(p_n)$ 为卷积核权重, 能适应采样形状大小。

参数量方面, 相比于标准卷积(k^2 规模)和DCN(k^2 规模), LDConv选用更小的 $|S|$ 以减少计算量, 同时保持了良好的特征提取能力。

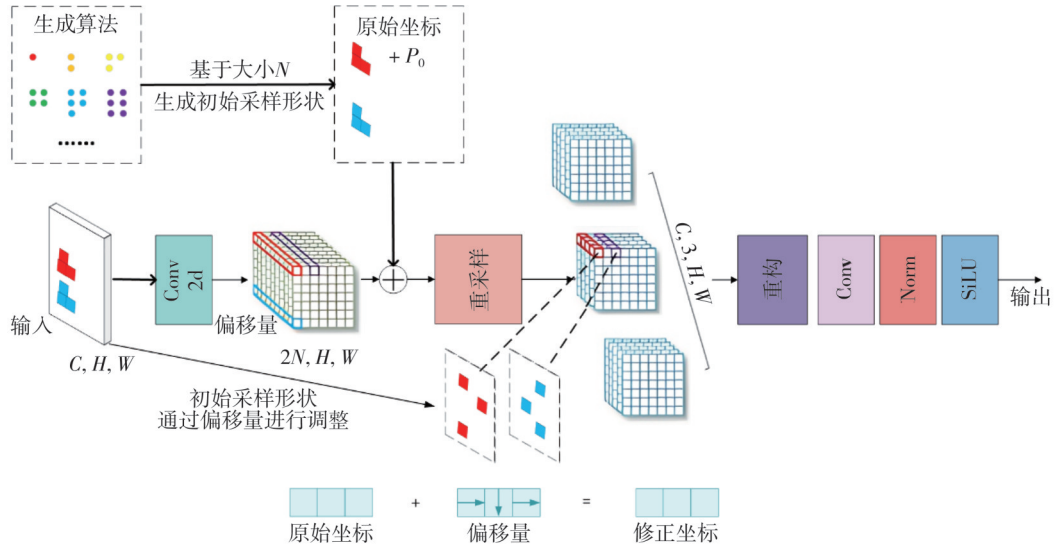


图2 LDConv结构图

Fig. 2 Structure diagram of LDConv

1.2 SCSA

本文将原通道注意力机制替换为SCSA，加

入了对空间特征的建模，提升了模型对复杂图像细节特征的表达能力。

SCSA原理如图3所示。

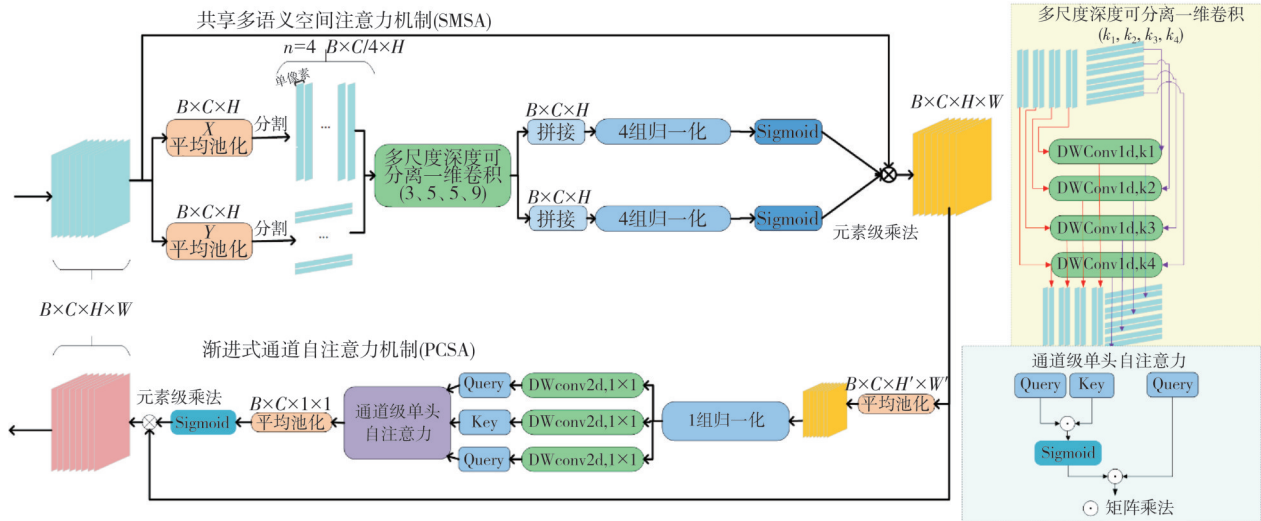


图3 SCSA结构图

Fig. 3 Structure diagram of SCSA

SCSA由两个主要部分组成：共享多语义空间注意力(Shared Multi-Semantic Space Attention, SMSA)和渐进通道自注意力(Progressive Channel-Wise Self-Attention, PCSA)。

SMSA整合多语义信息，通过渐进压缩策略将判别空间先注入到自注意力中，引导通道再校准。首先对输入特征进行多尺度处理，获取不同语义层次的空间特征并融合，继而生成空间注意力图

$$S = \sigma(f_{\text{conv}}(\text{Concat}(X_1, X_2, \dots, X_n))), \quad (2)$$

式中： X_i 为第*i*个尺度的特征图； $\text{Concat}(\cdot)$ 为通道

维度上的拼接； f_{conv} 为特征融合的 1×1 卷积； σ 为Sigmoid激活函数； $S \in \mathbb{R}^{1 \times H \times W}$ 为空间注意力图。

最后， S 通过逐元素惩罚作用域原始特征图 X ，得到空间增强后的特征 X_s 为

$$X_s = X \otimes S, \quad (3)$$

式中： $\otimes$ 为逐元素乘法。

而PCSA通过注意力机制对不同通道的贡献进行重新加权。对 X_s 进行通道注意力计算为

$$A = \text{softmax}(QK^T), \quad (4)$$

式中： $Q = X_s W_q$ 为通过线性变换得到的查询矩

阵; $K = X_s W_k$ 为通过线性变换得到的键矩阵; $W_q, W_k \in \mathbb{R}^{C \times d}$ 为科学系的参数矩阵; $A \in \mathbb{R}^{C \times C}$ 为通道间的子注意力权重。

再利用 A 对特征进行加权求和, 得

$$X' = AV, \quad (5)$$

式中: $V = X_s W_v$, W_v 为通道转换的权重矩阵; X' 为最终特征表示, 融合空间与通道信息。

SCSA 整体结构为先通过 SMSA 计算空间注意力图 S , 用于增强输入 ($X_s = X \otimes S$), 再通过 PCSA 计算权重 A 并应用于特征图 ($X' = AV$)。

1.3 R-RaGAN 损失函数

在 GAN 的训练过程中, 损失函数的选择会直接影响模型的稳定性和生成质量。传统模型采用最小化交叉熵损失(Binary Cross Entropy, BCE), 即

$$L_D^{BCD} = -E_{x \sim p_{data}} [\lg D(x)] - E_{z \sim p_z} [\lg(1 - D(G(z)))], \quad (6)$$

$$L_G^{BCE} = -E_{z \sim p_z} [\lg D(G(z))], \quad (7)$$

式中: L_D^{BCD} 为判别器交叉熵损失, 用于优化判别器区分真实样本与生成样本的能力; L_G^{BCE} 为生成器交叉熵损失, 用于优化生成器生成样本的真实性; $x \sim p_{data}$ 为真实样本; $z \sim p_z$ 为输入的噪声向量; $D(x)$ 为判别器对输入 x 在 BCE 损失中判定为真的概率; $G(z)$ 为生成器根据输入噪声 z 生成的样本。

然而, BCE 容易导致梯度消失, 生成器难以学习有效特征。此外, 对抗训练的不稳定性往往导致模式崩溃, 生成器只能产生少数固定模式样本, 无法覆盖数据的真实分布。

为解决这些问题, 研究者们提出了 WGAN-GP 损失函数, 通过优化 Wasserstein 距离, 使训练过程更稳定, 公式如下

$$L_D^{WGAN-GP} = E_{\hat{x} \sim p_{\hat{x}}} [D(\hat{x})] - E_{x \sim p_{data}} [D(x)] + \lambda E_{\hat{x} \sim p_{\hat{x}}} [(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2], \quad (8)$$

$$L_G^{WGAN-GP} = -E_{z \sim p_z} [D(G(z))], \quad (9)$$

式中: $L_D^{WGAN-GP}$ 为 WGAN-GP 的判别器损失; $L_G^{WGAN-GP}$ 为 WGAN-GP 的生成器损失; $\hat{x}$ 为在真实样本 $x \sim p_{data}$ 和生成样本 $x_f = G(z)$ 之间随机插值得到的样本; $D(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2$ 为梯度惩罚项, 用于限制梯度范数, 以满足 1-Lipschitz 约束; λ 为梯度惩罚项的权重系数。

然而 WGAN-GP 仍存在较大计算开销、生成器优化过程困难、梯度爆炸或收敛不稳定等问题。

本文采用 R-RaGAN 损失函数, 相较 WGAN-GP, 其引入了相对性损失

$$L_D^{R-RaGAN} = -E_{x_r \sim p_{data}, x_f \sim p_g} [\lg \sigma(D(x_r) - \overline{D(x_f)})] + R(D), \quad (10)$$

$$L_G^{R-RaGAN} = -E_{x_r \sim p_{data}, x_f \sim p_g} [\lg \sigma(D(x_f) - \overline{D(x_r)})], \quad (11)$$

式中: $L_D^{R-RaGAN}$ 为 R-RaGAN 的判别器损失; $L_G^{R-RaGAN}$ 为 R-RaGAN 的生成器损失; $x_r \sim p_{data}$ 为真实样本; $x_f \sim p_g$ 为生成样本; $D(x_r)$ 、 $D(x_f)$ 为当前正在更新的判别器对真实/生成样本的评分; $\overline{D(x_f)}$ 为同一批次中所有生成样本的判别器评分平均值; $\overline{D(x_r)}$ 为同一批次中所有真实样本的判别器评分平均值; $R(D)$ 为梯度惩罚项; σ 为 Sigmoid 函数。

与 WGAN-GP 相比, R-RaGAN 能够在不增加计算成本的情况下提供更平滑的梯度优化, 提高生成器的训练效率。由于其在损失函数中引入相对得分机制, 使得 GAN 模型在训练过程中能更有效地学习数据的真实分布, 减少模式崩溃问题。

2 实验

2.1 实验设置

为验证本算法的有效性, 选用 FLIR 官方红外图像数据集 D_1 和变电站采集红外图像数据集 D_2 进行实验。 D_1 包含 2 556 张图像, 分辨率为 640×512 , 涵盖多种环境条件和目标类型; D_2 由 WP-GS030/M 型红外相机采集, 包含 4 753 张图像, 分辨率为 640×480 , 包含低照度、遮挡、小目标密集等复杂环境。数据集经归一化、随机裁剪、水平翻转等处理后用于训练和测试。训练集与测试集按 8:2 比例划分。

本实验在 Ubuntu 23.10 64 位操作系统下进行, 计算平台包括 Intel Core i5-14400 处理器、NVIDIA RTX 4060 Ti GPU, 软件环境为 PyTorch 2.0.1、CUDA 11.8 和 cuDNN 8.4.1。

训练过程中, 批量大小设为 16, 采用混合精度训练提高效率。优化器采用 Adam, 初始学习率设为 1×10^{-4} , 并使用 Step Decay 学习率衰减策略, 每 10 轮衰减 50%。参数设置为 $\beta_1 = 0.9$, $\beta_2 = 0.999$, 权重衰减系数设为 1×10^{-5} 。损失函数由 L1、感知损失和 R-RaGAN 组成, 平衡系数设为 $\alpha = 0.5$, 调节感知损失与对抗损失的权重。总轮次设为 50。

本文选取 ESRGAN、Real-ESRGAN、CA-ESRGAN 和 GTISR 作为对比算法, 这些方法代

表了当前基于生成对抗网络的红外图像增强算法的研究主线。ESRGAN具备出色的高频细节重建能力,能够显著提升生成图像的视觉逼真度;Real-ESRGAN在真实退化图像的纹理还原与整体清晰度方面表现优异;CA-ESRGAN借助注意力机制增强了关键区域的细节保留与全局信息融合能力;GTISR依托Transformer结构在全局特征建模与长距离依赖信息捕获上具有明显优势。

为确保实验公平性,ESRGAN、Real-ESRGAN和GTISR算法采用官方预训练模型,CA-ESRGAN和本文算法则从头训练。所有算法均设置相同的实验环境、数据集和参数,以确保一致性和可比性。

2.2 评价指标

图像增强算法的评价指标主要分为两大类,即

主观评价和客观评价。主观评价是通过人的主观感受来评估增强图像的质量和特性,判断增强效果的好坏。为更客观地分析算法性能,使用3种客观评价指标:PSNR、SSIM和LPIPS,对图像增强的效果进行量化,能够较为全面地反映不同算法的优劣。

2.3 对比实验及结果分析

FLIR官方红外图像数据集 D_1 的效果图如图4所示,变电站采集数据集 D_2 的效果图如图5所示。由图4(b)ESRGAN结果看,其一定程度上改善了对比度,但细节表现不足,边缘模糊;图4(c)Real-ESRGAN显示,图像细节得以增强,但部分区域过度锐化,并存在伪影;图4(d)CA-ESRGAN突显了目标区域,但边缘轮廓不够清晰;图4(e)GTISR提升了对比度和清晰度,但背景噪声有所增加。

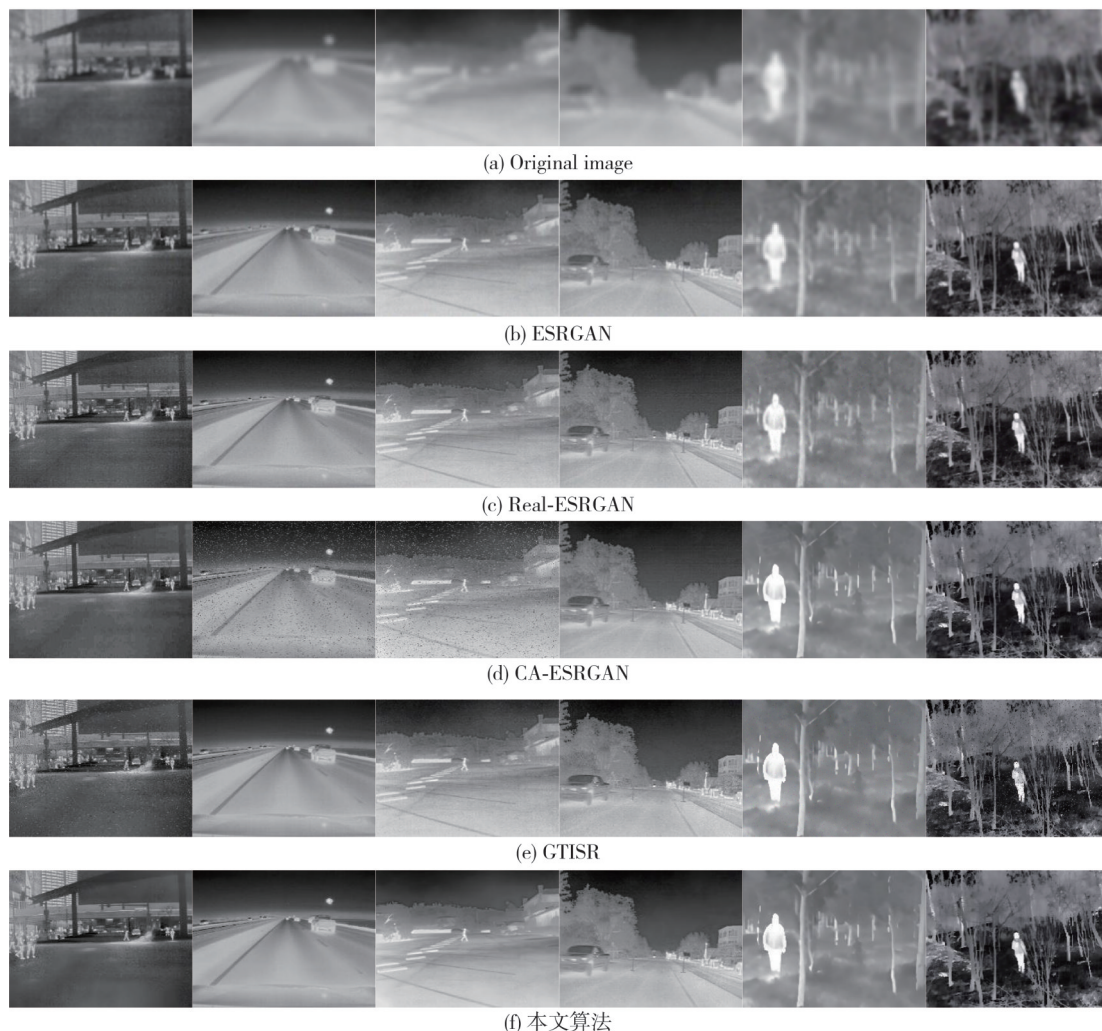


图4 红外图像数据集 D_1 增强结果

Fig. 4 Enhancement results on infrared image dataset D_1

图5(b)ESRGAN局部细节增强效果显著,如接线端子的温度轮廓更加清晰,但部分区域过度增

强,高温点周围出现伪影;图5(c)Real-ESRGAN在整体清晰度和纹理还原方面有较大提升,但设备

周围有较多高亮斑块,且设备冷区与背景之间对比度不足,视觉效果混乱;图 5(d) CA-ESRGAN 增强关键区域表现突出,接线端子高温区域自然,变压

器轮廓也较清晰,但对比度和细节信息不足;图 5(e) GTISR 整体提升明显,但热分布的边界表现较模糊,电力设备与周围区域的温度梯度失真。



图 5 红外图像数据集 D_2 增强结果

Fig. 5 Enhancement results on infrared image dataset D_2

相比之下,本文算法(图 4(f)、图 5(f))在多个场景下表现突出。数据集 D_1 中,其能够显著提升目标清晰度,物体边缘轮廓更加明显,图像纹理细节得到有效恢复。数据集 D_2 中,其表现更为突出,不仅细节表现完整,改善了复杂热分布区域的可读性,而且显著减少了背景区域的噪声干

扰,使冷区与背景的边界更加清晰分明。

从表 1、图 4~图 8 可知,本文算法相较对比算法在弱化背景的同时突出了主体目标,视觉观感有了较大的提升;且 PSNR、SSIM 和 LPIPS 等客观评价指标均优于对比算法,显示出较好的重建效果。

表 1 对比实验客观指标对比

Tab. 1 Comparison of objective metrics in contrast experiments

算法	D_1			D_2		
	PSNR/dB	SSIM	LPIPS	PSNR/dB	SSIM	LPIPS
ESRGAN	26.01	0.703 1	0.156 7	27.52	0.725 3	0.150 1
Real-ESRGAN	28.49	0.724 7	0.139 8	29.57	0.745 8	0.135 3
CA-ESRGAN	28.44	0.725 5	0.140 0	29.43	0.746 6	0.138 5
GTISR	30.45	0.743 1	0.129 5	31.05	0.751 6	0.130 7
本文算法	30.58	0.752 2	0.125 3	32.89	0.789 7	0.117 6

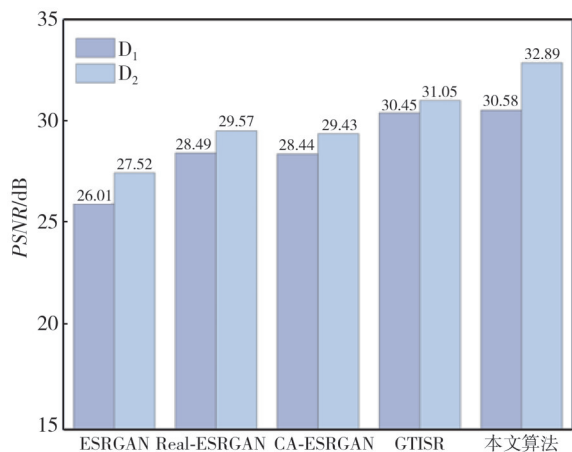


图6 对比实验PSNR对比

Fig. 6 PSNR comparison in contrast experiments

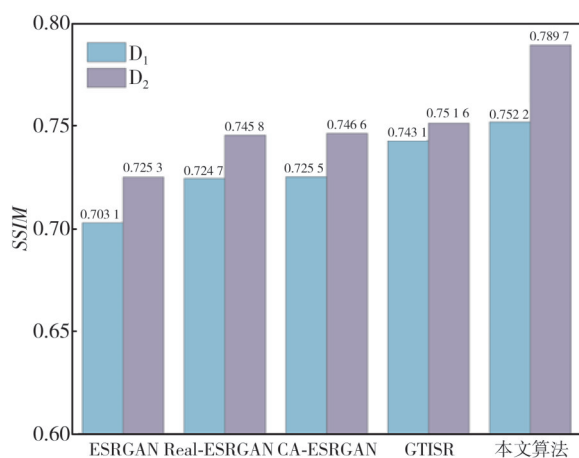


图7 对比实验SSIM对比

Fig. 7 SSIM comparison in contrast experiments

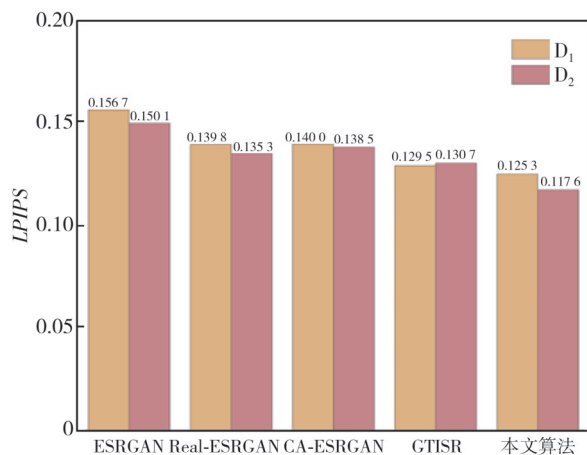


图8 对比实验LPIPS对比

Fig. 8 LPIPS comparison in contrast experiments

2.4 消融实验及结果分析

为验证本文算法引入LDConv、SCSA和R-RaGAN对性能的提升,使用FLIR官方数据集D₁和变电站采集数据集D₂进行消融实验。

使用完整算法分别与未引入LDConv、SCSA或R-RaGAN函数的不完整算法对比,结果如图9所示。

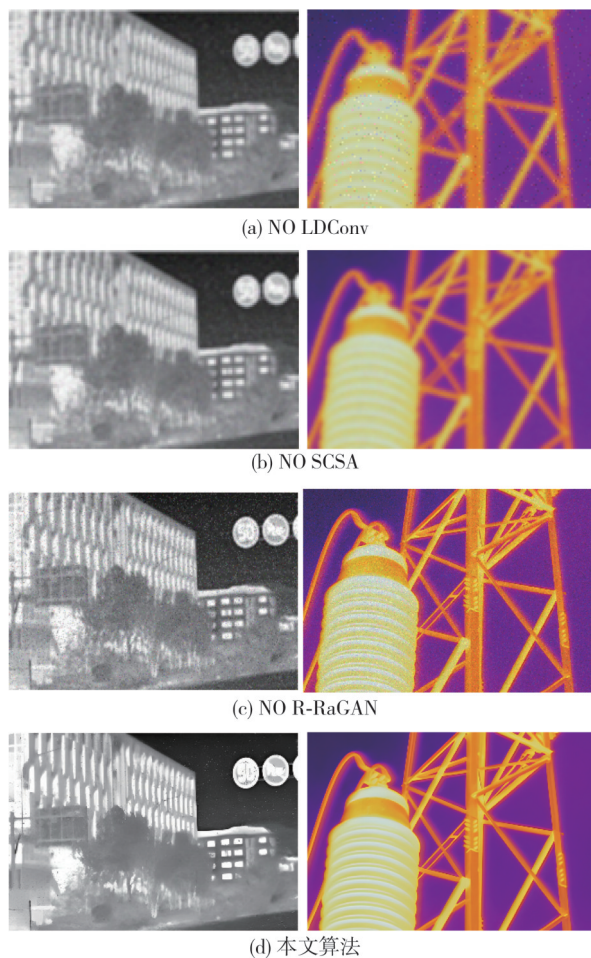


图9 消融实验效果图

Fig. 9 Melting experiment effect diagram

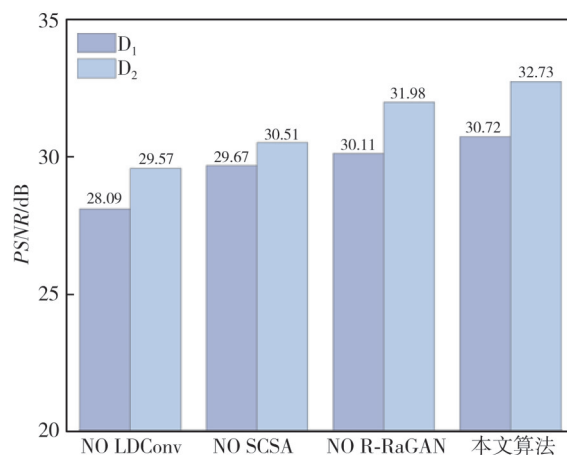


图10 消融实验PSNR对比

Fig. 10 PSNR comparison in ablation experiments

与完整算法相比,图9(a)中未引入LDConv的图像在细节还原上存在明显不足,边缘模糊且纹理不清晰;图9(b)中未引入SCSA的图像局部细节缺

失,目标区域对比度不足;图 9(c)中未引入 R-RaGAN 的图像出现伪影和纹理不清晰现象,且模型收敛速度较慢。完整算法在图像增强时表现出了显

著的性能提升,不仅提高了生成图像的质量,还增强了稳定性。由表 2、图 10~图 12 定量分析可得,本文算法在客观指标上均有一定提升。

表 2 消融实验客观指标对比

Tab. 2 Objective comparison of ablation experiment metrics

算法	D ₁			D ₂		
	PSNR/dB	SSIM	LPIPS	PSNR/dB	SSIM	LPIPS
NO-LDConv	28.09	0.721 1	0.148 7	29.57	0.702 3	0.135 5
NO-SCSA	29.67	0.698 7	0.139 9	30.51	0.754 4	0.150 5
NO-R-RaGAN	30.11	0.695 4	0.135 6	31.98	0.725 3	0.128 8
本文算法	30.72	0.781 2	0.120 7	32.73	0.793 4	0.119 8

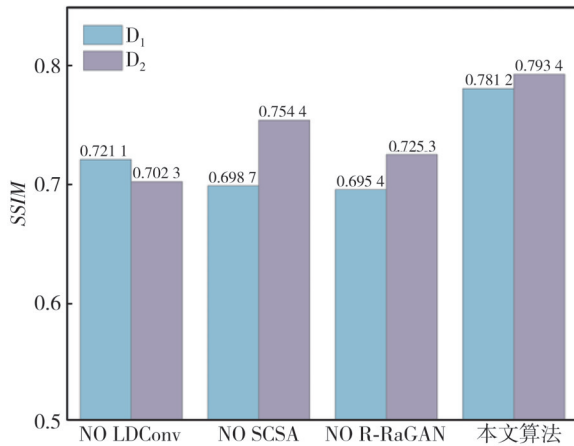


图 11 消融实验 SSIM 对比

Fig. 11 SSIM comparison in ablation experiments

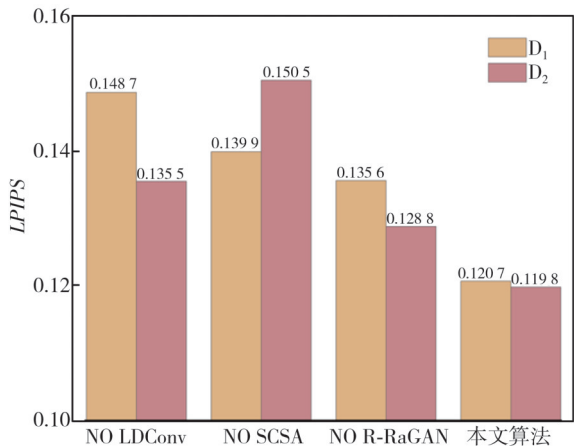


图 12 消融实验 LPIPS 对比

Fig. 12 LPIPS comparison in ablation experiments

3 结论

本文提出了一种基于改进 CA-ESRGAN 的红外图像增强算法,分别通过在特征提取中引入 LDConv,生成器中引入 SCSA 和训练优化中引入 R-RaGAN,显著提升了算法性能。LDConv 通过动态调整卷积核的采样形状,增强了模型对复杂纹理与局部细节的捕获能力;SCSA 通过同时建模空间与通道维度的特征关联性,强化了对图像

细节与目标区域的关注;R-RaGAN 在对抗训练中引入相对性判别与正则化约束,缓解了模式崩溃与收敛不稳定问题。消融实验结果表明,所引入模块均有效增强了模型性能。在对比实验中,本文算法在 FLIR 公开数据集(D₁)和变电所采集的红外图像数据集(D₂)上均取得增强效果,全面优于 ESRGAN、Real-ESRGAN、CA-ESRGAN 和 GTISR。尤其在存在低照度、高噪声、小目标密集等复杂巡检环境的 D₂ 数据集中,PSNR、SSIM 和 LPIPS 指标分别较 GTISR 提升 1.84 dB、0.038 1 和 0.013 1。实验结果充分验证了算法的有效性与鲁棒性,尤其在复杂电力巡检环境中,为电力设备设备巡检提供更可靠的红外图像支持。

利益冲突声明/Conflict of Interests

所有作者声明不存在利益冲突。

All authors declare no relevant conflict of interests.

参考文献

- [1] 王贺,张震. 基于 ResNeSt 和改进 Transformer 的多标签图像分类算法[J]. 测试技术学报, 2024, 38(1): 48-53.
Wang He, Zhang Zhen. Multi-label image classification algorithm based on ResNeSt and improved transformer[J]. Journal of Test and Measurement Technology, 2024, 38(1): 48-53. (in Chinese)
- [2] Guo Y, Du L. Infrared and visible image fusion algorithm based on Retinex-enhanced multiscale decomposition[J]. Journal of Measurement Science and Instrumentation, 2024, 15(2): 176-184.
- [3] 高金金,李潞洋. 一种改进的点云 Transformer 深度学习模型[J]. 中北大学学报(自然科学版), 2021, 42(6): 515-523.
Gao Jinjin, Li Luyang. Deep learning model based on improved point cloud transformer[J]. Journal of North

- University of China (Natural Science Edition), 2021, 42(6): 515-523. (in Chinese)
- [4] Shen Z, Qin F, Ge R, et al. IDTransformer: infrared image denoising method based on convolutional transposed self-attention[J]. Alexandria Engineering Journal, 2025, 110: 310-321.
- [5] Schmidt E L, Ou Z, Ximendes E, et al. Near-infrared II fluorescence imaging[J]. Nature Reviews Methods Primers, 2024, 4: 23.
- [6] Wang W, Li Z, Siddique A. Infrared maritime small-target detection based on fusion gray gradient clutter suppression[J]. Remote Sensing, 2024, 16(7): 1255.
- [7] Yang W, Wang S, Fang Y, et al. Band representation-based semi-supervised low-light image enhancement: bridging the gap between signal fidelity and perceptual quality [J]. IEEE Transactions on Image Processing, 2021, 30: 3461-3473.
- [8] Wang Q, Ma Q. Super-resolution reconstruction algorithm for medical images by fusion of wavelet transform and multi-scale adaptive feature selection [J]. IET Image Processing, 2024, 18(13): 4297-4309.
- [9] Zhu J, Meng L, Wu W, et al. Generative adversarial network-based atmospheric scattering model for image dehazing[J]. Digital Communications and Networks, 2021, 7(2): 178-186.
- [10] Sun C, Wang C, He C. Image super-resolution reconstruction algorithm based on SRGAN and swin transformer[J]. Symmetry, 2025, 17(3): 337.
- [11] Zhao D. SRGAN in underwater vision[J]. Cognitive Robotics, 2024, 4: 1-7.
- [12] Rifai R, Putra Utama H, Astina Tasmara F, et al. Enhancing resolution of raster-scan photoacoustic imaging using Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN) [J]. Journal of Physics: Conference Series, 2025, 2945(1): 012040.
- [13] 李刚, 张亚兵, 杨庆贺, 等. 基于 Real-ESRGAN 的岩石 CT 图像超分辨率重建[J]. 工矿自动化, 2023, 49(11): 84-91.
- Li Gang, Zhang Yabing, Yang Qinghe, et al. Super-resolution reconstruction of rock CT images based on Real-ESRGAN [J]. Journal of Mine Automation, 2023, 49(11): 84-91. (in Chinese)
- [14] Mukadam S B, Patil H Y. Skin cancer classification framework using enhanced super resolution generative adversarial network and custom convolutional neural network[J]. Applied Sciences, 2023, 13(2): 1210.
- [15] Suárez P L, Carpio D, Sappa A D. Enhancement of guided thermal image super-resolution approaches[J]. Neurocomputing, 2024, 573: 127197.