基于自注意力与多损优化的低光增强

张闵航[1] 宋小齐[2] 孙雨[2]
1. 电子科技大学 四川 成都 611731;
2. 中国铁路兰州局集团有限公司 甘肃 兰州 730000

摘 要 零样本低照度图像增强无须配对数据，但光照噪声干扰先验提取易产生伪影。为了解决这一问题，本文提出自注意力生成对抗网络（SA-GAN），集成亮度、空间、通道三重注意力强化暗部建模。设计梯度范数自适应损失平衡算法动态协调多目标约束。多基准集实验表明，该方法性能已媲美主流方法，提升了图像的自然度和可见性。研究结论表明，本文提出的自注意力生成对抗网络（SA-GAN）在低照度图像增强领域展现出显著优势。通过集成亮度、空间、通道三重注意力机制，SA-GAN有效强化了暗部区域的建模能力，提升了图像的亮度和细节重建质量。同时，引入的梯度范数自适应损失平衡算法，动态协调了多目标约束，增强了训练的稳定性和适应性。来研究可探索更轻量的注意力结构，以拓展至视频低光增强场景。

关键词 低照度图像；自注意力生成；多损优化

Low-Light Enhancement via Self-Attention and Multi-Loss Optimization

Zhang Min-hang[1], Song Xiao-qi[2], Sun Yu[2]
1. University of Electronic Science and Technology of China, Chengdu 611731, Sichuan Province, China;
2. China Railway Lanzhou Group Co., Ltd., Lanzhou 730000, Gansu Province, China

Abstract Zero-shot low-light image enhancement avoids paired data, yet prior extraction disturbed by illumination noise easily produces artifacts. To solve this problem, we propose a Self-Attention Generative Adversarial Network (SA-GAN) that integrates brightness, spatial, and channel attention to strengthen dim-region modeling. A gradient-norm adaptive loss-balancing algorithm is designed to dynamically coordinate multi-objective constraints. Experiments on multiple benchmarks show that the proposed method rivals state-of-the-art approaches and improves image naturalness and visibility. The study concludes that SA-GAN exhibits significant advantages in low-light enhancement. The triple-attention mechanism effectively reinforces modeling of dark regions, enhancing brightness and detail reconstruction, while the adaptive loss-balancing algorithm stabilizes training and improves adaptability. Future work can explore lighter attention structures for extension to video low-light enhancement.

Key words Low-Light Image; Self-Attention Generation; Multi-Loss Optimization

低照度图像普遍存在亮度不足、对比度偏低、细节信息丢失等问题，严重制约了视觉感知效果及后续高层任务的性能。近年来，零样本学习方法（如ExCNet[1]、Zero-DCE、RRDNet等）通过自监督学习摆脱了对成对训练数据的依赖，然而这类方法通常依赖简化的先验假设，易受图像噪声干扰，在恢复过程中常出现细节结构模糊或局部曝光失衡等现象；此外，其损失函数中各项权重多为固定设置，需针对不同场景反复调参，限制了模型的泛化能力与实用效率。

针对上述问题，本文在文献[2]研究基础上，提出一种基于自注意力生成对抗网络的低照度增强模型（SA-GAN）。该模型在生成器中集成三重自注意力机制，能够有效聚焦图像暗部区域并增强局部结构表征；判别器采用相对判别策略，提升生成结果在纹理与光照方面的真实感；同时，引入梯度范数驱动的自适应损失平衡算法，实现多个无参考损失项之间的动态权重协调，避免人工调参的不确定性。实验结果表明，该方法在主观视觉质量与多项客观评价指标上均优于现有主流方法，展现出良好的细节恢复与曝光控制能力。

1 相关工作

1.1 零样本低照度图像增强方法

零样本低照度增强主要是为了解决数据依赖问题，Zero-DCE[3]估计迭代曲线参数，结合无参考损失实现自监督增强；RRDNet分解光照、反射与噪声分量；RetinexDIP是深度图像先验和Retinex理论的结合。这些方法侧重结构设计，但多损失权重平衡问题仍待解决。

1.2 多目标损失平衡

EnlightenGAN[4]提出基于梯度归一化的动态权重调整策略，根据各损失梯度幅值自适应分配权重，避免单一损失主导训练。本文借鉴该思想，设计边界约束机制，稳定低照度增强训练过程。

2 实验方法

生成器使用亮度、空间、通道三类自注意力模块提高暗部、细节建模，判别器使用多层卷积网络考察真实性，多目标损失自适应平衡法则用于无参考训练的图像增强。

2.1 生成器网络

设计思路，生成器通过迭代曲线映射来增强图像。过程如下：

$$LE_n(x)=LE_{n-1}(x)+A_n(x)LE_{n-1}(x)[1-LE_{n-1}(x)] \quad (1)$$

式中，x—待处理像素位置；$LE_n(x)$—第n次增强对应图像在像素x处亮度值；$A_n(x)$—对应参数图取值；$LE_0(x)$—输入图像归一化亮度值；n—迭代次数。$A_n(x)>0$时增强暗部区域，$A_n(x)<0$时抑制过亮区域。

为精准预测参数图 $\{A_n\}_{n=1}^{N}$ ，生成器卷积骨干网络中加入三类自注意力模块，强化暗部区域与结构细节建模。

1.亮度注意力模块。由亮度图$L\in[0,1]^{H\times W}$构造亮度权重图，公式如下：

$$M_{LG}=1-\mathrm{Y} \quad (2)$$

该过程使暗部区域具有较大的权重。中间特征图为$F\in\mathbb{R}^{H\times W\times C}$的逐点加权公式：

$$Fb(x,y,c)=M_{LG}(x,y)F(x,y,c) \quad (3)$$

式中，(x,y)—空间位置；c—通道索引，引导网络聚焦暗部亮度提升与细节恢复。

2.空间注意力模块。该模块捕捉特征图中的空间关系，定位重要区域。对输入特征图$F\in\mathbb{R}^{H\times W\times C}$经过通道维度平均池化和最大池化，得到两个单通道特征图$M_1,M_2\in\mathbb{R}^{H\times W\times 1}$。

拼接后经7×7卷积和Sigmoid激活生成空间权重图$M_{SL}\in\mathbb{R}^{H\times W\times 1}$：

$$M_{SL}=F_{ex}(M_1,M_2)=\sigma[\mathrm{Conv}_{7\times 7}([M_1,M_2])] \quad (4)$$

空间权重图M_{SL}值越高，表明对应区域亮度越高、细节越丰富。

定义空间注意力权重图M_{SG}定义如下：

$$M_{SG}=F_M(M_{SL})=1-M_{SL} \quad (5)$$

式中，F_M—权重映射转换操作。$1-M_{SL}$可提升暗区域注意力权重，聚焦暗部亮度改善与细节恢复。

3.通道注意力模块。该模块强化通道间依赖，通过空间层面池化产生特征图$M_1,M_2\in\mathbb{R}^{1\times 1\times C}$；拼接后经7×7卷积和ReLU激活函数生成通道注意力图$M_{CL}\in\mathbb{R}^{1\times 1\times C}$：

$$M_{CL}=\mathrm{ReLU}[\mathrm{Conv}_{7\times 7}([M_1,M_2])] \quad (6)$$

定义通道注意力权重M_{CG}如下：

$$M_{CG}=1-M_{CL} \quad (7)$$

输入特征图X与M_{CG}加权后得到特征图$X_C\in\mathbb{R}^{H\times W\times C}$：

$$X_C(x,y,c)=M_{CG}(c)\cdot X(x,y,c) \quad (8)$$

实现对暗通道低光区域增强。

2.2 判别器网络

判别器采用多层卷积网络提取特征并评估真实性，判别形式：

$$D_R(x_r,x_f)=\sigma\left(C(x_r)-\mathbb{E}_{x_f\in\mathbb{P}_{\mathrm{fake}}}\left[C\left(x_f\right)\right]\right) \quad (9)$$

$$D_R(x_f,x_r)=\sigma\left(C(x_f)-\mathbb{E}_{x_r\in\mathbb{P}_{\mathrm{real}}}\left[C\left(x_r\right)\right]\right) \quad (10)$$

式中，x_r—真实图像；x_f—生成图像；$C(\cdot)$—判别器对输入图像的判分函数。

在训练中，生成器通过公式（10）最大化生成图像的真实性；判别器通过公式（9）和公式（10）增强区分能力，二者对抗博弈推动结果逼近正常照度图像分布。

2.3 多目标损失函数

零样本设定下，本文在对抗损失L_{adv}的基础上，引入了对抗损失L_{gen}、自保持特征损失L_{SPF}、曝光控制损失L_{exp}和色彩恒定性损失L_{col}。生成器总损失：

$$\mathscr{L}_G=\lambda_{gen}L_{gen}+\lambda_{SPF}L_{SPF}+\lambda_{exp}L_{exp}+\lambda_{col}L_{col} \quad (11)$$

式中，λ—是各损失项权重。

2.4 自适应多目标损失平衡机制

针对固定损失权重导致的训练失衡问题，依据损失梯度范数动态调整权重。首先，生成器总损失定义为：

$$\mathscr{L}_{total}(t)=\sum_i\lambda_i(t)\cdot L_i(t) \quad (12)$$

式中，$L_i(t)$—第i项损失在第t次迭代中的值；$\lambda_i(t)$—对应的权重。

计算损失项相对于网络参数W的梯度范数及最大值：

$$G_W^{(i)}=\left\|\nabla_W L_i(t)\right\|_2 \quad (13)$$

$$G_W^{\max}=\max\left\{GW_{it}\right\} \quad (14)$$

权重更新公式：

$$\lambda_i(t)=\begin{cases}0.9\cdot\lambda(t-1)+(1-a)\cdot\dfrac{G_W^{\max}(t)}{G_W^{(i)}(t)} & \lambda_i\leqslant\theta\\ \theta & \text{others wise}\end{cases} \quad (15)$$

式中，λ_i—初始权重；$\lambda_i(t-1)$—上一轮权重，实验取θ=10效果最佳。

算法流程：初始化损失权重 $\lambda_i^{(0)}=0\forall i$ 与网络权重W；迭代计算梯度范数并更新权重与网络参数至收敛。

3 实验结果与分析

3.1 数据集与训练设置

模型在EnlightenGAN的无配对数据集训练，在LIME、DICM、NPE、MEF数据集测试。基于PyTorch，RTX 3080 GPU，Adam优化器（初始学习率0.000 1，100 epoch后衰减至0），batch size大小为32，生成器/判别器参数尺寸为33.0/42.5 MB，网络结构较小，适用于对模型复杂度较低的低照度图像增强。

3.2 实验对比

在DICM、LIME、MEF和NPE数据集上，原始图像、EnlightenGAN、SCI、StableLLVE、URetinex-Net、Zero-DCE以及本文SA-GAN方法的增强结果。从可视化结果来看，在DICM、LIME和MEF上SA-GaN在增强亮度的同时，能够更好地恢复细节，抑制噪声和伪影，提高色彩的层次，视觉效果较对比方法更佳。在NPE上，虽然SA-GAN增强的图像亮度较低，但是可以有效抑制伪影和噪声。在纹理和非均匀照明情况下，SA-GAN可以防止过度增强和局部过曝，增强结果在真实感和可见性之间处于较好的平衡，与生成对抗学习和自注意力机制密切相关[5-6]。

4 结束语

本文提出的SA-GAN采用亮度、空间、通道注意力模块，对暗部与细节建模；利用自适应损失平衡机制动态调节权重，进行零样本低光增强的亮度、细节、自然度优化。实验结果表明，SA-GAN在主观和客观指标上均优于主流方法，后续可以将更轻量的注意力结构应用于视频低光增强场景。

参考文献

[1]BISCHOF R, KRAUS M A. Multi-objective loss balancing for physics-informed deep learning[J]. Computer Methods in Applied Mechanics and Engineering,2025,439:1179-1184.

[2]COLTUC D, BOLON P, CHASSERY J M.Exact histogram specification[J]. IEEE Transactions on Image Processing,2006,15(5):1143-1152.

[3]FAN M H, WANG W J, YANG W H, et al. Integrating semantic segmentation and Retinex model for low-light image enhancement[C]// Proceedings of the 28th ACM International Conference on Multimedia. 2020:2317-2325.

[4]GUO X J, LI Y, LING H B. LIME: low-light image enhancement via illumination map estimation[J]. IEEE Transactions on Image Processing, 2016,26(2):982-993.

[5]IBRAHIM H, KONG N S P. Brightness preserving dynamic histogram equalization for image contrast enhancement[J]. IEEE Transactions on Consumer Electronics,2007,53(4):1752-1758.

[6]JIANG Y F, GONG X Y, LIU D, et al. Enlightengan: deep light enhancement without paired supervision[J]. IEEE Transactions on Image Processing,2021,30:2340-2349.

（本文编辑　曹　阳）

