

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

基于双频域特征聚合的低照度图像增强

徐胜军, 杨华, 李明海, 刘光辉, 孟月波, 韩九强

引用本文:

徐胜军, 杨华, 李明海, 等. 基于双频域特征聚合的低照度图像增强[J]. *光电工程*, 2023, 50(12): 230225.

Xu S J, Yang H, Li M H, et al. Low-light image enhancement based on dual-frequency domain feature aggregation[J]. *Opto-Electron Eng*, 2023, 50(12): 230225.

<https://doi.org/10.12086/oe.2023.230225>

收稿日期: 2023-09-08; 修改日期: 2023-12-02; 录用日期: 2023-12-03

相关论文

一种并行混合注意力的渐进融合图像增强方法

刘光辉, 杨琦, 孟月波, 赵敏华, 杨华

光电工程 2023, 50(4): 220231 doi: 10.12086/oe.2023.220231

多尺度注意力与领域自适应的小样本图像识别

陈龙, 张建林, 彭昊, 李美惠, 徐智勇, 魏宇星

光电工程 2023, 50(4): 220232 doi: 10.12086/oe.2023.220232

融合暗通道先验损失的生成对抗网络用于单幅图像去雾

程德强, 尤杨杨, 寇旗旗, 徐进洋

光电工程 2022, 49(7): 210448 doi: 10.12086/oe.2022.210448

渐进式多粒度ResNet车型识别网络

徐胜军, 荆扬, 李海涛, 段中兴, 刘福友, 李明海

光电工程 2023, 50(7): 230052 doi: 10.12086/oe.2023.230052

更多相关论文见光电期刊集群网站 



<http://cn.ojournal.org/oe>



 OE_Journal



Website

DOI: 10.12086/oe.2023.230225

基于双频域特征聚合的低照度图像增强

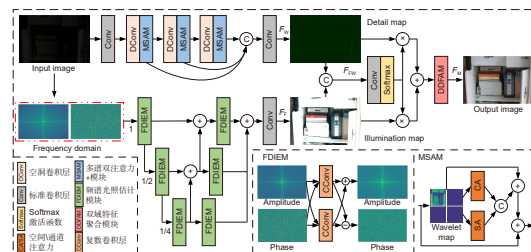
徐胜军^{1,2}, 杨华^{1,2*}, 李明海¹, 刘光辉^{1,2},
孟月波^{1,2}, 韩九强^{1,2}¹西安建筑科技大学 信息与控制工程学院, 陕西 西安 710055;²西安市建筑制造智能化技术重点实验室, 陕西 西安 710055

摘要: 针对低照度图像质量较差、噪声多、纹理模糊等问题, 提出一种基于双频域特征聚合的低照度增强网络 (dual frequency-domain feature aggregation network, DF-DFANet)。首先, 构建频谱光照估计模块 (frequency domain illumination estimation module, FDIEM) 实现跨域特征提取, 通过共轭对称约束调整频域特征图抑制噪声信号, 并采用逐层融合方式提高多尺度融合效率以扩大特征图感受野范围。其次, 设计多谱双注意力模块 (multiple spectral attention module, MSAM) 聚焦图像局部频率特征, 通过小波域空间、通道注意力机制关注图像细节信息。最后, 提出双域特征聚合模块 (dual domain feature aggregation module, DDFAM) 融合傅里叶域和小波域特征信息, 利用激活函数计算自适应调整权重实现像素级图像增强, 并结合傅里叶域全局信息提高融合效果。实验结果表明, 在 LOL 数据集上所提网络的 PSNR 达到 24.3714, SSIM 达到 0.8937。与对比网络相比, 所提网络增强效果更具自然性。

关键词: 深度学习; 图像增强; 傅里叶变换; 小波变换; 双域融合; 注意力机制

中图分类号: TP391

文献标志码: A



徐胜军, 杨华, 李明海, 等. 基于双频域特征聚合的低照度图像增强 [J]. 光电工程, 2023, 50(12): 230225

Xu S J, Yang H, Li M H, et al. Low-light image enhancement based on dual-frequency domain feature aggregation[J]. *Opto-Electron Eng*, 2023, 50(12): 230225

Low-light image enhancement based on dual-frequency domain feature aggregation

Xu Shengjun^{1,2}, Yang Hua^{1,2*}, Li Minghai¹, Liu Guanghui^{1,2}, Meng Yuebo^{1,2}, Han Jiuqiang^{1,2}¹College of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an, Shaanxi 710055, China;²Xi'an Key Laboratory of Building Manufacturing Intelligent & Automation Technology, Xi'an, Shaanxi 710055, China

Abstract: Aiming at the problems of poor low-light image quality, noise, and blurred texture, a low-light enhancement network (DF-DFANet) based on dual-frequency domain feature aggregation is proposed. Firstly, a spectral illumination estimation module (FDIEM) is constructed to realize cross-domain feature extraction, which can adjust the frequency domain feature map to suppress noise signals through conjugate symmetric constraints and improve the multi-scale fusion efficiency by layer-by-layer fusion to expand the range of the feature map.

收稿日期: 2023-09-08; 修回日期: 2023-12-02; 录用日期: 2023-12-03

基金项目: 国家自然科学基金面上项目 (52278125); 陕西省重点研发计划 (2020GY-186, 2021SF-429); 陕西省自然科学基金基础研究计划 (2023-JC-YB-532, 2022JQ-681)

*通信作者: 杨华, yhxuat@163.com.

版权所有©2023 中国科学院光电技术研究所

Secondly, the multispectral dual attention module (MSAM) is designed to focus on the local frequency characteristics of the image, and pay attention to the detailed information of the image through the wavelet domain space and channel attention mechanism. Finally, the dual-domain feature aggregation module (DDFAM) is proposed to fuse the feature information of the Fourier domain and the wavelet domain, and use the activation function to calculate the adaptive adjustment weight to achieve pixel-level image enhancement and combine the Fourier domain global information to improve the fusion effect. The experimental results show that the PSNR of the proposed network on the LOL dataset reaches 24.3714 and the SSIM reaches 0.8937. Compared with the comparison network, the proposed network enhancement effect is more natural.

Keywords: deep learning; image enhancement; fourier transform; wavelet transform; dual-domain convergence; attention mechanism

1 引言

道路监控是智慧交通领域的重要组成部分,然而在低照度条件下的夜间场景中,摄像头采集到的图像亮度、对比度较低,且存在较多噪声颗粒,这为交通监管领域重要目标检测、识别等视觉任务带来了难度。低照度图像增强可以有效改善图像视觉效果,然而原始图像画质较低,不合理的降噪策略容易造成纹理模糊等图像退化问题,使得低照度图像增强任务具有较大的挑战性^[1]。

已有低照度图像增强算法分为传统方法和基于深度学习的方法两类。传统低照度图像增强方法主要以直方图均衡化 (histogram equalization, HE) 和基于 Retinex 理论的方法为主。Li 等^[2]针对低画质图像难以保留细节信息的问题,提出了一种基于融合的 QCFs 低照度图像增强方法,将全局 QCFs 图像和局部对比度受限 HE 图像相融合来提高图像画面质量。Pan 等^[3]基于 Retinex 理论利用亮度增强函数 (BEF) 和自适应对比度增强 (IACE) 进行多尺度融合,使用高斯拉普拉斯金字塔对增强估计的光照图进行改善。李平等^[4]设计了单尺度 Retinex 自适应增强方法,通过引导滤波对图像照度分量进行评估,再经过 Gamma 矫正光照图以实现自适应图像增强。

由于深度学习发展迅速,在低照度图像增强领域的应用也越来越广泛。Zhao 等^[5]受到 Retinex 理论的启发,提出了一种基于视网膜的光照注意力网络 (retinex-based illumination attention network, RIANet),采用多分支结构来提取不同感受野下的光照特征,构建照明注意力块来缓解图像光照不均、对比度低等问题。Wu 等^[6]提出了一种基于视网膜的深度展开网络 (retinex-based deep unfolding network, URetinexNet),通过隐式先验正则化模型设计迭代学习模块,循环分

解图像实现噪声抑制和照度调整。Jiang 等^[7]针对分解得到反射图存在较多噪声的问题,提出了一种无监督分解与校正网络 (unsupervised decomposition and correction network, UDCN),通过无监督噪声降噪网络 (noise remove network, NRN) 以 HE 增强图像为参照对象,在内部一致性约束、混合无参考质量感知损失函数的指导下去除阴影噪声。虽然基于 Retinex 理论和卷积网络通过调整分解光照分量的方式实现了低照度增强,但是基于 Retinex 理论和卷积神经网络的结构中,空间域的卷积堆叠操作容易损失图像结构信息,不合理的降噪策略难以区分高频信息与噪声信号,导致光照图对图像空间信息表征能力减弱,反射图在降噪过程中容易丢失纹理结构。

为解决上述问题,研究者们通过二维傅里叶变换 (fast Fourier transform, FFT) 将空间图像分解为不同频率、方向复平面波之和的形式,并用复数矩阵存储平面波的波动大小和平移量,利用卷积神经网络学习复平面信息分布以抑制异常峰值点,从而提高网络区分噪声信号和纹理信息的能力。Li 等^[8]提出了一种超高清傅里叶低照度增强网络 (embedding Fourier for ultra-high-definition network, UHDFour),构建傅里叶卷积模块对振幅图和相位图进行分离处理,在傅里叶频域通过振幅图调整低照度图像亮度值,融合后的增强图像避免在增强亮度的同时放大噪声。Zhang 等^[9]为避免图像数据非线性映射到 RGB 色彩空间引入额外的噪声和伪影,提出了一种基于空间到光谱的跨尺度低照度增强网络,通过在三分支上嵌入快速傅里叶卷积块 (FFC) 和 Transformer 块,以扩大感受野范围和聚合多尺度特征。Hai 等^[10]提出了 R2RNet 网络实现低照度图像增强,利用空间信息提高图像亮度、对比度,同时通过 FFT 提取频域特征来保留图像细节信息。Lin 等^[11]为解决弱光环境下的增强过程中,容

易造成过度曝光和局部模糊的问题, 提出了一种基于退化特征向量引导的低照度网络, 在提取到暗区和亮区上退化特征向量的基础上, 设计快速傅里叶残差注意力块 (FFR-DG), 指导网络恢复图像细节特征。然而, 频域卷积是通过引入快速傅里叶变换将 RGB 图像信号映射到频域向量空间并进行处理的方法, 频域特征图上很小的扰动都会影响空间域全局图像的变化, 容易导致低照度增强图像纹理细节模糊和空间结构信息损失。

小波变换 (wavelet transform, WT) 能够实现了对图像全局空间信息的局部分析, 通过伸缩平移运算对特征图逐步进行多尺度细化, 区分高、低频率特征以避免相互之间的干扰影响, 从而聚焦到任意频率特征细节处进行分析, 提高网络恢复低照度图像细节方面的能力。Xu 等^[12] 为充分利用光照信息改善图像质量, 提出了一种基于光照引导的小波注意力网络 (illumination guided attentive wavelet network, IGAWN), 通过注意力小波变换层 (attention wavelet transform layer, ADWT) 对图像高、低频率进行分离, 增强图像小波域的重要特征的表达并抑制冗余的噪声信号。Fan 等^[13] 提出了一种 M-Net 的半小波注意力网络 (half wavelet attention M-Net, HWMNet), 在不同尺度的特征层上使用半小波注意力块以丰富小波域的特征, 通过融合不同分辨率图像来丰富特征图的上下文信息, 改善下采样造成的空间信息损失。胡聪等^[14] 为改善空间域卷积运算造成结构信息损失, 提出了一种融合半波注意力模块的低光照图像增强算法, 利用半波注意力模块获得采集小波域特征的能力, 提高网络恢复图像细节特征的能力。Chen 等^[15] 提出了一种基于注意力的广义引导网络 (attention-based broad self-guided network, ABSGN), 在小波域的低分辨率特征图上利用全局空间注意力块 (GSA) 关注上下文信息, 采用多级引导密集块 (MGDB) 融合多尺度特征图, 提高特征映射重用性, 但是直接在小波域进行图像降噪任务需要网络具有较好的特征表征能力, 以充分提取全局和局部的特征, 导致网络运算成本的增加。

综上所述, 虽然深度学习实现低照度图像增强已取得一定的效果, 但在增强亮度、对比度的同时, 容易放大阴影噪声, 不合理的降噪策略常导致图像信息损失严重, 出现不同程度的细节模糊, 尤其对于画面质量较差的低照度图像, 往往难以恢复丢失的纹理结

构。为解决上述问题, 提出一种基于双频域的特征聚合网络 (dual frequency domain feature aggregation network, DF-DFANet)。首先, 设计频谱光照估计模块 (FDIEM), 通过傅里叶域频谱特征图提取低照度图像全局特征, 在频域拉升图像亮度的同时减少对噪声信号的响应。其次, 提出多谱双注意力模块 (MSAM), 利用空间、通道注意力机制使得网络关注小波特征子图上重要特征, 提高网络恢复图像细节信息的能力。最后, 构建双域特征聚合模块 (DDFAM) 自适应学习不同像素级特征调整权重参数, 并通过复数域卷积促进融合特征信息, 增强图像色彩表现的自然性和纹理细节的丰富性。

2 基于双频域的特征聚合网络

为了解决低照度图像增强过程中容易出现噪声放大、纹理模糊等退化问题, 提出了一种基于双频域的特征聚合网络, 该网络主要由傅里叶域分支、小波域分支和双域特征聚合模块构成。其中傅里叶域分支, 通过快速傅里叶变化使得浅层特征图具有全局感受野信息, 由频谱光照估计模块 (FDIEM) 构建共轭约束条件提取丰富语义信息, 逐层融合不同尺寸特征图并得到精细化的光照图; 小波域分支主要由空洞卷积和多谱双注意力模块 (MSAM) 组成, 结合通道和空间维度强化子空间特征纹理细节信息; 双域特征聚合模块 (DDFAM) 自适应融合分支特征图, 引导不同维度特征图之间相互信息互补, 进一步提升融合图像颜色纹理特征的表达能力。所提 DF-DFANet 网络结构如图 1 所示。

2.1 频谱光照估计模块

针对空间域卷积神经网络在低照度图像增强过程中, 提升图像亮度、对比度的同时容易放大阴影噪声的问题, 提出一种频谱光照估计模块 (frequency domain illumination estimation module, FDIEM), 首先, 在傅里叶分支上低照度输入图像通过快速傅里叶变换从空域转换到频域表示, 频域特征矩阵中实部表示为振幅, 虚部表示为相位, 则分解得到振幅图和相位图分别包含图像的亮度和噪声信息^[8], 对其各自进行复数卷积运算^[16] 可有效避免图像亮度提升对噪声放大的影响。此外, 傅里叶域特征图具有覆盖全局图像的感受野^[17], 使得浅层网络获得包含全局上下文信息的特征图, 提升网络对低照度图像的提取和表征能力。二维图像连续傅里叶变换、离散傅里叶变换和离

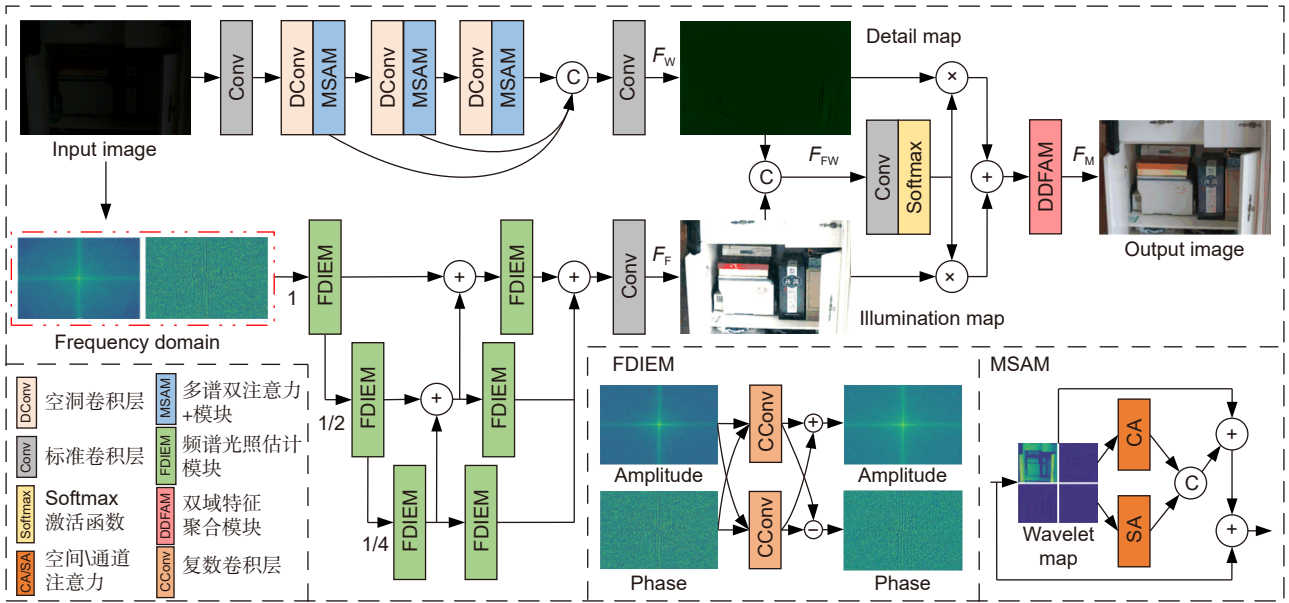


图 1 DF-DFANet 网络结构
Fig. 1 DF-DFANet network structure

散傅里叶逆变换计算表达式如下:

$$F(u, v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) e^{-j2\pi(ux+vy)} dx dy, \quad (1)$$

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)}, \quad (2)$$

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)}, \quad (3)$$

式中: x, y 表示图像空间域像素点坐标, u, v 表示傅里叶频域特征图点坐标, M, N 表示分别图像长、宽, $f(x, y)$ 表示输入的低照度图像, $F(u, v)$ 表示傅里叶频域特征图。

其次, 通过频谱光照估计模块分离频谱特征图复数矩阵的实部与虚部, 获得包含颜色纹理信息的振幅图和包含语义信息的相位图。根据傅里叶变换的共轭对称性, 任意一个信号序列都可由其共轭对称序列与共轭反对称序列和的形式表示, 利用复数卷积层构建二维傅里叶共轭对称约束条件对, 从而实现频谱光照估计模块。在经过卷积网络反向传播求导更新权重参数后, 实现利用相位图的语义信息引导振幅图关注重要频率特征, 并抑制无关噪声信号响应以达到图像降噪效果。频谱光照估计模块构建的共轭对称约束条件对计算表达式如下:

$$\begin{cases} \hat{F}_P = \phi(F_P) + \varphi(F_A) \\ \hat{F}_A = \phi(F_A) - \varphi(F_P) \end{cases}, \quad (4)$$

式中: F_P 为原始相位特征, F_A 原始振幅特征, $\varphi(\cdot)$ 、

$\phi(\cdot)$ 为互不相关的卷积运算, \hat{F}_P 为相位特征增强结果, \hat{F}_A 为振幅特征增强结果。

最后, 高分辨率特征图和低分辨率特征图具有相似的振幅图, 从而实现通过获取低分辨率的振幅图来调整高分辨率的亮度值, 降低网络计算成本, 所以将不同尺度的特征图采用逐步上采样组合的方式^[18] 获得亮度增强后的照度图, 提高网络特征图的融合效率^[19]。频谱光照估计模块如图 2 所示。

在频谱光照估计模块中, 令输入振幅矩阵和相位矩阵分别为 F_A, F_P , 并送入各自分支的复数卷积层中, 再经过层归一化和激活函数得到共轭约束关系输出结果 \hat{F}_A, \hat{F}_P , 并且循环其输出结果作为下一层复数卷积输入, 从而获得增强频域特征的振幅和相位图。频谱光照估计模块计算表达式如下:

$$\begin{cases} F_{A_{i+1}} = R(N(C(F_{A_i}))) - R(N(C(F_{P_i}))), \{i = 0, 1\} \\ F_{P_{i+1}} = R(N(C(F_{P_i}))) + R(N(C(F_{A_i}))), \{i = 0, 1\} \end{cases}, \quad (5)$$

式中: R 表示 ReLU 激活函数, N 表示层归一化操作, C 表示复数卷积层, F_{A_i} 表示第 i 个层的振幅特征图, F_{P_i} 表示第 i 个层的相位特征图。

2.2 多谱双注意力模块

针对复杂弱光环境下难以提取低照度图像深层特征的问题, 常用网络通过空间域卷积层堆叠的方式从而获得更大的感受野范围, 但容易导致图像结构信息的损失。为解决此问题, 提出一种多谱双注意力模块 (multiple spectral attention module, MSAM), 在小

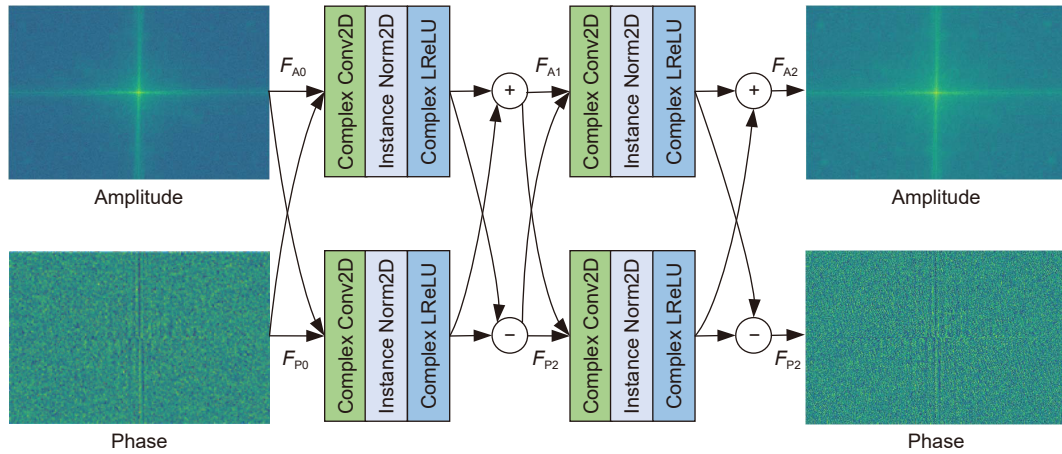


图 2 频谱光照估计模块结构
Fig. 2 Structure of spectral illumination estimation module

波域分支上, 通过连续空洞卷积扩大特征图感受野范围, 并引入多谱双注意力模块。该分支利用小波变换将图像投影到由尺度函数和小波函数簇构成的子空间, 实现对特定信息分量的分离与观测, 提高网络对细节特征的代表能力。然而对所得特征图进行简单的相加或拼接操作未曾考虑小波低频和高频特征系数分布的不同, 容易造成细节信息丢失, 并且在提取高频细节特征操作中容易保留阴影噪声, 因此引入注意力机制可以进一步促进有用特征的频率响应, 通道和空间注意力机制并联的结构设计有效缓解了不同注意力先后排列顺序造成的特征差异^[20]。二维小波变换、离散小波变换和离散小波反变换计算表达式如下:

$$\begin{cases} \varphi(x, y) = \varphi(x)\varphi(y) \\ \psi^H(x, y) = \psi(x)\varphi(y) \\ \psi^V(x, y) = \varphi(x)\psi(y) \\ \psi^D(x, y) = \psi(x)\psi(y) \end{cases}, \quad (6)$$

$$\begin{cases} \varphi_{j,m,n} = 2^{\frac{j}{2}}\varphi(2^jx - m, 2^jy - n) \\ \psi_{j,m,n}^i(x, y) = 2^{\frac{j}{2}}\psi^i(2^jx - m, 2^jy - n), i = \{H, V, D\} \\ W_\varphi(j_0, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)\varphi_{j_0,m,n}(x, y) \\ W_\psi^i(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)\psi_{j,m,n}^i(x, y), i = \{H, V, D\} \end{cases}, \quad (7)$$

$$\begin{aligned} f(x, y) &= \frac{1}{\sqrt{MN}} \sum_m \sum_n W_\varphi(j_0, m, n)\varphi_{j_0,m,n}(x, y) \\ &+ \frac{1}{\sqrt{MN}} \sum_{i=H,V,D} \sum_{j=j_0}^{\infty} \sum_m \sum_n W_\psi^i(j, m, n)\psi_{j,m,n}^i(x, y), \end{aligned} \quad (8)$$

式中: $\varphi(x, y)$ 表示二维尺度函数, $\varphi(x)$ 、 $\varphi(y)$ 表示一维尺度函数, $\psi^H(x, y)$ 、 $\psi^V(x, y)$ 、 $\psi^D(x, y)$ 分别表示沿水

平、垂直、对角线方向变化的二维小波函数, $\psi(x)$ 、 $\psi(y)$ 表示一维小波函数。 i 表示方向小波, M 、 N 表示图像长、宽, $f(x, y)$ 表示输入图像, j_0 表示任意的开始尺度, $W_\varphi(j_0, m, n)$ 表示图像在尺度 j_0 处的近似系数, $W_\psi^i(j, m, n)$ 表示对图像附加水平、垂直和对角方向的细节系数, H、V、D分别表示水平、垂直、对角线方向。

首先, 堆叠多个空洞卷积提取图像结构特征并设置不同大小空洞率的卷积, 在扩大特征提取范围的同时避免出现网格伪影。其次, 在级联卷积层之间嵌入多谱双注意力模块, 对输入图像进行分解得到低频特征和高频特征, 并沿其水平、垂直和对角线方向进行重构, 所得子特征图是空间域特征在小波组合频段的表示, 从而提升网络对图像边缘和纹理结构的特征响应。然后, 通过并行注意力机制对特征图在空间和通道维度上进行加权, 强化重要目标特征的显著性区域, 自适应调整不同通道特征图的权重分布。最后, 利用 1×1 卷积将高频特征系数重构到低频特征图, 对其进行逆小波变换形成包含丰富上下文信息的图像轮廓特征图。浅层特征图经过级联空洞卷积和多谱双注意力模块提取小波域细节特征, 输入特征先经过单个空洞卷积层获得输出特征图 F_{Di} , $i \in \{1, 2, 3, \dots, n\}$, 令其空洞率为 d , $d \in \{1, 3, \dots, 2 \times (n-1) + 1\}$ 。然后, 多谱双注意力模块将空间域特征图转换到小波域进行处理, 令模块输出特征图为 F_{WAMi} , $i \in \{1, 2, 3, \dots, n\}$ 。最后, 将级联空洞卷积和多谱双注意力模块各阶段输出特征图进行通道堆叠操作, 令卷积融合后的输出特征图为 F_w 。小波域分支计算表达式如下:

$$F_W = \text{Conv}\{\text{Cat}\{D\text{Conv}(F_{WAM1}), \dots, D\text{Conv}(F_{WAMi})\}, \quad (9)$$

式中： Conv 为卷积融合操作， Cat 为通道堆叠， $D\text{Conv}$ 为空洞卷积， i 为第*i*个小波注意力块， F_{WAMi} 为第*i*个小波注意力块输出特征， F_W 为小波域分支部分输出特征图。

由于网络空间域和傅里叶域的卷积操作，容易造成图像空间信息损失，加入多谱双注意力模块实现恢复图像细节信息。首先，可逆的小波变换过程对图像特征具有分离特性，将包含低频信息、水平高频信息、垂直高频信息、对角高频信息的子空间特征图进行通道堆叠。其次，利用空间和通道注意力机制给予不同频率的特征图以相应的权重参数，增强特征图中细节信息的表达，同时抑制不同频率中的噪声信号。最后，将各个维度特征图进行通道堆叠操作，经过卷积融合操作后的特征图与输入特征图逐像素相加，增强特征重用和信息流动。多谱双注意力模块结构如图3所示。

首先，在多谱双注意力模块中，令模块输入特征为 F_D ，经过小波变换操作后获得小波域的特征图为 F_{fw} ， $F_{fw} \in R^{C \times H \times W}$ ，将其沿水平和垂直方向对其低频、水平、垂直、对角信息子空间特征图做通道堆叠处理，得到堆叠后特征图为 F_{fwh} ， $F_{fwh} \in R^{4C \times \frac{H}{2} \times \frac{W}{2}}$ 。其次，令经过空间和通道注意力机制的输出堆叠特征图为 F_{wsca} ， $F_{wsca} \in R^{8C \times \frac{H}{2} \times \frac{W}{2}}$ 。最后，利用 3×3 卷积对特征图进行降维，并通过残差连接对输入和输出特征图进行逐像素相加操作。多谱双注意力模块计算表达式如下：

$$\begin{cases} F_{fwh} = \text{Cat}_{\text{DHVA}}[DWT[Conv_3(Conv_3(F_D))]] \\ F_{wsca} = \text{Cat}[SA(F_{fwh}) + CA(F_{fwh})] \\ F_{WAM} = Conv_1(F_D) + Conv_1(F_{wsca}) + F_{fwh} \end{cases}, \quad (10)$$

式中： F_D 为模块输入特征， $Conv_3(\cdot)$ 、 $Conv_1(\cdot)$ 分别为卷积核大小为 3×3 、 1×1 的卷积运算， $DWT[\cdot]$ 为小波变换操作， $\text{Cat}_{\text{DHVA}}[\cdot]$ 为按水平、垂直方向堆叠操作， F_{fwh} 为子空间特征图通道堆叠输出， $SA(\cdot)$ 、 $CA(\cdot)$ 分别为空间、通道注意力运算， F_{wsca} 为空间通道注意力输出特征， F_{WAM} 为多谱双注意力模块特征输出。

2.3 双域特征聚合模块

虽然融合不同维度的特征图可以互补相关信息，有效增强网络对特征的表达能力，但是特征图逐像素相加或相乘的融合策略难以有效利用不同特征图相关信息的空间映射关系，常导致特征融合操作图像细节增强效果不佳。为解决此问题，提出一种双域特征聚合模块(dual domain feature aggregation module, DDFAM)。由于傅里叶域和小波域的处理过程相互独立，作为该模块输入的光照图和细节图具有不同的特征分布，通过自适应计算两个特征图的权重比例关系以实现在融合过程中的软注意力机制^[21]，提高网络在恢复低照度图像光照和细节方面的能力。因此，首先将傅里叶域分支输出光照图和小波域分支输出细节图进行通道堆叠，并利用softmax激活函数赋予每个像素权重系数，实现特征图之间相关信息的映射。然后通过得到的权重系数对两个特征图进行逐像素相乘与相加操作，对融合特征图进行像素级细节调整。最后结合傅里叶域

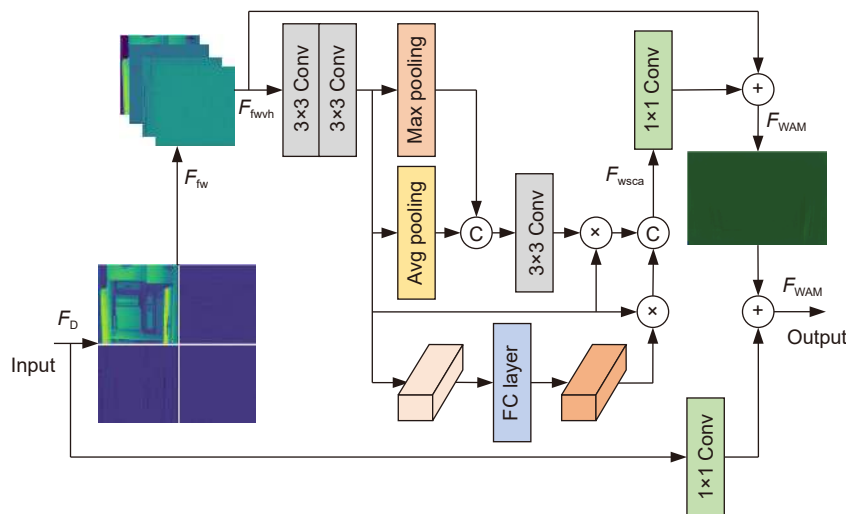


图3 多谱双注意力模块结构
Fig. 3 Structure of multiple spectral attention module

特征图具有全局感受野的特性^[22], 利用复数卷积将实数域特征图映射到复数域空间, 提升网络在复数特征空间的表征能力, 融合不同分支上特征图可有效增强网络对图像恢复纹理结构的能力, 提高增强图像的视觉感知质量。双域特征聚合模块结构如图 4 所示。

在双域特征聚合模块中, 令傅里叶域分支和小波域分支输出特征作为模块输入特征 F_F 和 F_W , 按通道维度堆叠后的特征图为 F_{FW} 。然后, 经过 softmax 激活函数计算得到参数权重矩阵为 w_f 、 w_s , 其与输入特征图逐像素相乘得到像素级增强后的特征图 F_{fs} 和 F_{ws} 。最后, 增强后特征图 F_{fs} 和 F_{ws} 进行相加融合并作为傅里叶域卷积块输入特征 F_{fcin} , 分别对其实部和虚部进行卷积运算, 在全局范围感受野下增强融合效果, 获得双域特征聚合模块输出特征图 F_M 。其计算表达式如下:

$$\begin{cases} w_{fs}, w_{ws} = \text{Softmax}(\text{Conv}(\text{Cat}(F_F, F_W))) \\ F_{fcin} = (F_F \otimes w_f) \oplus (F_W \otimes w_s) \\ F_M = \text{Conv}(f_{\text{FFT}}(f_{\text{IFFT}}(F_{fcin})) + F_{fcin}) \end{cases}, \quad (11)$$

式中: $\text{Softmax}(\cdot)$ 为 softmax 激活函数运算, w_{fs} 、 w_{ws} 为激活函数计算结果权重矩阵, \otimes 为逐像素相乘运算, \oplus 为逐像素相加运算, F_{fcin} 为像素级增强后的特征图, f_{FFT} 、 f_{IFFT} 为傅里叶变换和傅里叶反变换, F_M 为模块输出特征。

2.4 损失函数

损失函数使用常用的结构损失函数、感知损失函数和傅里叶损失函数, 则网络总损失函数定义为三部分损失函数加权之和的形式, 如式 (12) 所示。

$$L_{\text{total}} = L_{\text{ssim}} + \alpha L_{\text{perc}} + \beta L_{\text{fourier}}, \quad (12)$$

式中: 根据经验值设置权重 α 、 β 分别为 0.2、0.1,

L_{ssim} 表示结构损失函数, L_{perc} 表示感知损失函数, L_{fourier} 表示傅里叶损失函数。

结构损失函数为避免逐像素计算差异的方法造成图像模糊, 通过将图像属性中的亮度、对比度和结构信息进行分离, 用均值作为亮度的估计, 标准差作为对比度的估计, 协方差作为结构相似程度的度量, 排除亮度和对比度对结构相似性的影响, 提高网络对纹理结构的重建恢复能力。令 L_{ssim} 表示结构损失函数, 如式 (13) 所示:

$$L_{\text{ssim}} = 1 - \frac{(2\mu_S\mu_{\hat{S}} + c_1)(2\sigma_{S\hat{S}} + c_2)}{(\mu_S^2 + \mu_{\hat{S}}^2 + c_1)(\sigma_S^2 + \sigma_{\hat{S}}^2 + c_2)}, \quad (13)$$

式中: μ_S 、 $\mu_{\hat{S}}$ 分别表示预测图像 \hat{S} 和真实图像 S 的均值, σ_S 、 $\sigma_{\hat{S}}$ 分别表示预测图像 \hat{S} 和真实图像 S 的方差, $\sigma_{S\hat{S}}$ 表示预测图像 \hat{S} 和真实图像 S 的协方差, c_1 、 c_2 表示维持数值稳定的常数, 沿用文献 [23] 的方法设定 $c_1 = (k_1L)^2$ 、 $c_2 = (k_2L)^2$, L 表示图像像素值的动态范围, $k_1 = 0.01$ 、 $k_2 = 0.03$ 。

感知损失函数^[24] 利用欧式距离计算特征图之间的差值, 通过固定 VGG19 网络提高对浅层颜色纹理特征和深层高级语义特征的感知能力, 使得增强图像具有更好的视觉性能。令 L_{perc} 表示感知损失函数, 如式 (14) 所示:

$$L_{\text{perc}} = \frac{1}{C_t H_t W_t} \|P_t(\hat{S}) - P_t(S)\|_2^2, \quad (14)$$

式中: t 表示 VGG19 预训练权重第 t 层, C_t 、 H_t 、 W_t 分别表示 VGG19 预训练权重第 t 层特征图的通道数、宽、高, 第 t 层的特征图的尺寸大小, P_t 表示 VGG19 预训练权重第 t 层特征, \hat{S} 表示预测图像, S 表示真实图像, $\|\cdot\|_2$ 表示 L2 范数。

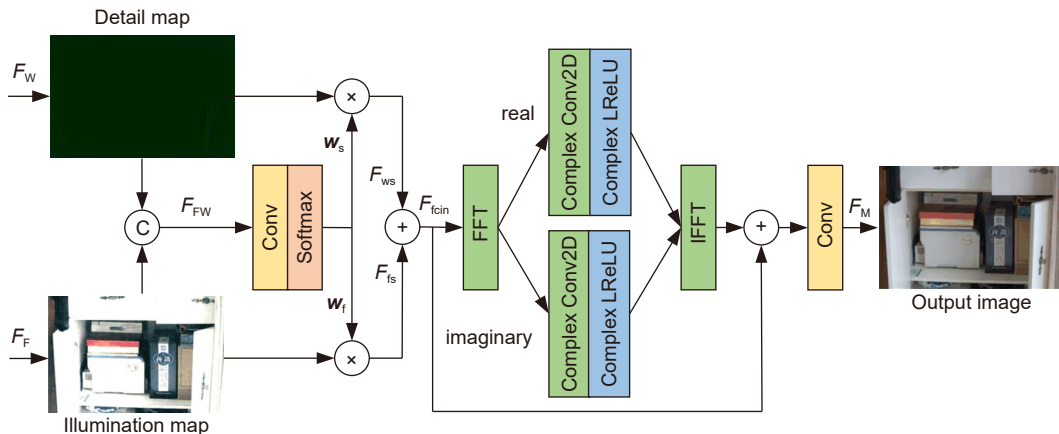


图 4 频域特征聚合模块结构
Fig. 4 Structure of frequency domain feature aggregation module

傅里叶损失函数通过计算真实图像和预测图像的振幅图和相位图之间平均差^[25], 引导网络从频域角度恢复纹理特征实现对空间域特征图的补充, 最小化真实图像与预测图像之间的差距, 提高网络对频域特征的敏感性。令 L_{fourier} 表示傅里叶损失函数, 如式(15)所示:

$$L_{\text{fourier}} = \frac{1}{N} \sum_{i=0}^n \| \text{Cat}[S^A, S^P] - \text{Cat}[\hat{S}^A, \hat{S}^P] \|^2, \quad (15)$$

式中: N 表示图像像素点, $\text{Cat}[\cdot]$ 表示特征图拼接操作, \hat{S} 表示预测图像, S 表示真实图像, A 、 P 分别表示图像的振幅图和相位图。

3 实验和结果分析

3.1 实验数据集及平台参数设置

实验平台采用 Intel(R) Xeon(R) Gold 6130 处理器, 128 GB 内存, NVIDIA RTX 2080Ti 11 GB 显卡, Ubuntu 18LTS 操作系统, 环境配置为 CUDA10.2+ Python3.6.9+Pytorch1.6。低照度图像增强对比实验数据集包括: LOL、MIT-Adobe FiveK。

LOL 数据集是来自于真实场景下的 500 对低照度图像, 通过改变采集设备的感光度和曝光值, 并进行拍摄以获得低照度图像和相应的正常曝光图像。由于 LOL 训练数据集数量较少, 为避免出现过拟合现象, 对正常曝光图像进行亮度调整, 得到不同亮度等级的低照度图像 10000 张, 并对其进行反转、旋转等数据增强处理以达到扩充数据集的目的。MIT-Adobe FiveK 数据集是通过调整相机传感器参数模拟模拟相机曝光过程制作, 对 5000 张原始 RGB 图像调整曝光值, 模拟非线性渲染过程实现不同曝光水平的图像, 最终得到 24330 张具有不同曝光设置的 8 位 sRGB 图像。实验所用数据集组成图像拍摄场景多样化且内容丰富, 图像光照条件多变, 使得这些数据集在低照度实验中具有较强的泛化能力和鲁棒性。

网络训练优化器使用 Adam 优化器, 初始学习率设置为 0.0001, 在训练任务完成 20 轮之后学习率衰减 10 倍, 网络迭代次数 epoch 设置为 200, batch size 为 8, patch size 为 256。训练过程中, 按照 8:2 划分为训练和测试数据集, 对 400×600 原始图像进行随机裁剪选取 256×256 图像块作为网络输入。

3.2 评价指标

客观评价评价指标采用峰值信噪比 (PSNR)、结

构相似性 (SSIM)、学习感知图像块相似度 (LPIPS)。各类评价指标定义如下:

峰值信噪比 (Peak signal to noise ratio, PSNR) 是衡量图像含噪声程度的指标, 单位为 dB, 如式 (16) 所示:

$$PSNR = 10 \times \log_{10} \left(\frac{(2^n - 1)^2}{MSE} \right). \quad (16)$$

结构相似性 (structural similarity, SSIM) 是一种基于人眼视觉感知的图像质量评估方法, 从亮度、对比度和结构三个维度来评估图像之间的相似度, 如式 (17) 所示:

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}. \quad (17)$$

学习感知图像块相似度 (Learned perceptual image patch similarity, LPIPS) 是用于衡量增强图像和真实图像之间差异量, 如式 (18) 所示:

$$d(x, x_0) = \sum_l \frac{1}{H_l W_l} \sum_{h,w} \|w_l \odot (\hat{y}^l - \hat{y}_{hw}^l)\|_2^2, \quad (18)$$

式中: x 、 y 分别为真实图像和增强图像, MSE 为均方误差计算结果, n 为图像存储比特位数, μ_x 和 μ_y 分别为 x 与 y 的均值, σ_x 、 σ_y 、 σ_{xy} 分别为各自方差和协方差, c_1 和 c_2 分别为常数, W_l 、 H_l 为第 l 提取到特征图的宽高, w_l 为放缩激活通道数。

3.3 LOL 数据集实验与分析

为验证所提网络对低照度图像的增强效果, 与其他对比网络进行对比实验, 实验结果如图 5 所示, 图 5(a) 为低照度输入图像, 从图 5(b-j) 分别为对比网络 RetinexNet^[26]、Zero-DCE^[27]、DSLRL^[28]、KinD^[29]、EnGAN^[30]、RUAS^[31]、URetinexNet^[6]、R2RNet^[10] 和 UHDFour^[8] 的增强结果, 图 5(k) 为所提网络 DF-DFANet 的增强结果, 图 5(l) 为真实图像。

3.3.1 定性分析

由图 5 可知, RetinexNet 增强图像的颜色饱和度和明亮度较高, 如图 5(b) 所示, 第 (3) 行图像墙体颜色艳丽、明亮度较高, 与真实图像视觉差距较大, 因其光照分量和反射分量融合过程简单, 导致颜色特征出现偏差。Zero-DCE 的增强图像曝光值偏低, 如图 5(c) 所示, 第 (1) 行橱柜的亮度水平较低且存在大面积阴影区域, 因为利用简单的亮度调节参数曲线使得网络泛化能力降低。DSLRL 利用拉普拉斯金字塔结构提升对局部细节的表征能力, 但不同尺寸特征图的简单融合导致图像伪影现象, 如图 5(d) 所示, 第 (1) 行橱柜挡板区域出现图像叠影而显得模糊、清晰度不高。由于 KinD

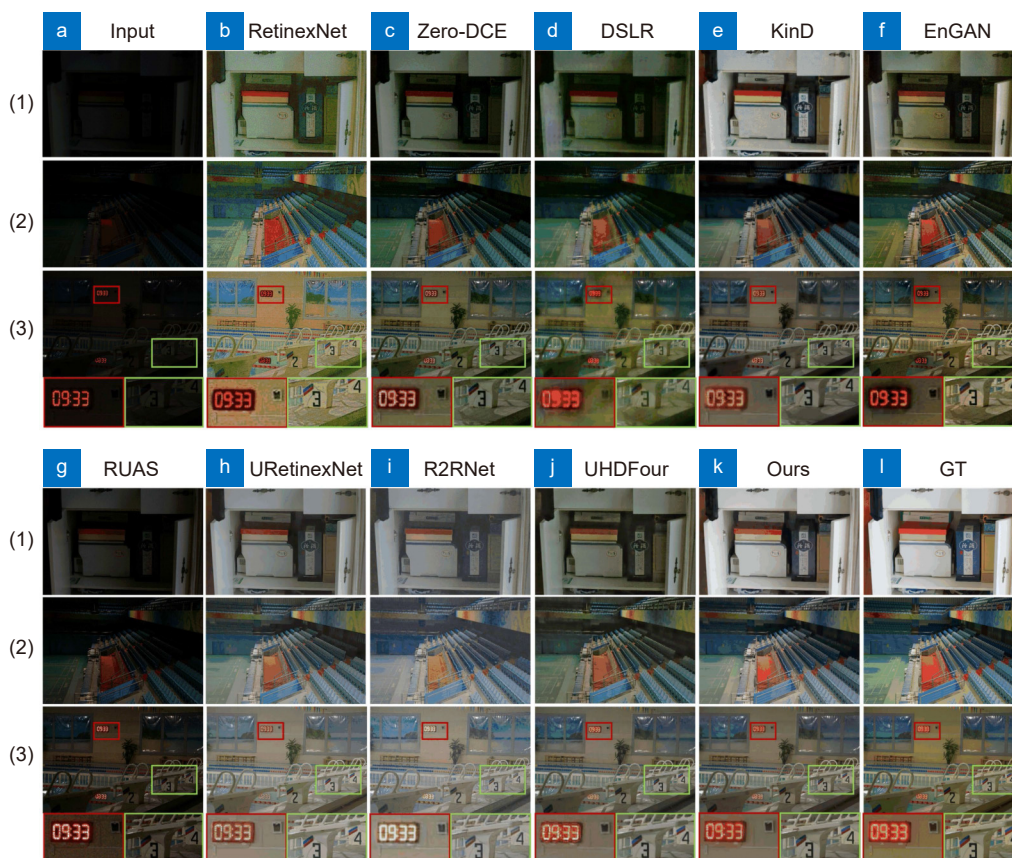


图 5 LOL 数据集增强结果对比

Fig. 5 LOL dataset enhancement results comparison

的降噪子网络对反射图进行下采样操作导致纹理信息损失, 易出现边缘模糊, 如图 5(e) 所示, 第 (3) 行数字时钟区域边缘失去纹理细节而变得光滑。基于生成对抗网络的 EnGAN 训练难度较大, 增强图像的色彩表现和真实图像存在一定差距, 如图 5(f) 所示, 第 (3) 行数字时钟颜色饱和度过高, 边缘出现明显图像阴影。RUAS 是基于 Retinex 理论的无参考优化模型, 通过搜索空间架构优化网络光照估计和降噪结构, 但是其适应不同数据集场景的能力较弱, 如图 5(g) 所示, 整体图像亮度值较低。URetinexNet 通过设计优化模块迭代细化反射图和光照图的方式, 自适应恢复光照和反射率, 提高增强图像恢复细节, 但其存在一定程度的褪色现象, 如图 5(h) 所示, 第 (2) 行观众席彩色墙体相较于真实图像颜色饱和度较低。R2RNet 通过傅里叶变换提高增强图像细节信息, 但基于 Retinex 理论模型的子空间特征图融合过程, 导致图像色彩表现上饱和度、对比度较低, 如图 5(i) 所示, 第 (3) 行数字时钟区域颜色明显鲜艳度较低。UHDFour 通过空间域和频域分离的特征提取方式, 降低特征空间的相互作用影响, 但空域卷积仍对增强图像纹理结构造

成影响, 如图 5(j) 所示, 第 (3) 行数字时钟区域纹理粗糙, 且边缘细节结构扭曲。所提网络 DF-DFANet 利用频谱光照估计模块提升在复杂弱光环境中对深层特征的提取能力, 并通过频域特征抑制噪声干扰, 多谱双注意力模块有效增强网络图像细节表征能力, 双域特征聚合模块自适应融合不同分支特征图, 对增强图像进行像素级细节调整, 更好保留了增强图像的颜色纹理特征。

3.3.2 定量分析

为进一步验证所提网络的有效性, 在 LOL 测试数据集上进行对比实验, 对比指标结果如表 1 所示。由表 1 可知, 所提网络的增强图像相较于其他对比网络评价指标均有一定的提升。所提网络与基于 Retinex 理论的 RUAS、RetinexNet、URetinexNet 相比, PSNR 指标分别提升了 48.56%、45.29%、14.26%, SSIM 指标分别提升了 77.53%、110.28%、7.05%; 所提网络与具有较好降噪结构的对比网络 DSLR、KinD 相比, PSNR 指标分别提升了 62.66%、34.70%, SSIM 指标分别提升了 49.84%、15.83%; 所提网络与无监督对比网络 Zero-DCE、RUAS 相比, PSNR 指标

分别提升了 63.99%、48.56%，SSIM 指标分别提升了 58.90%、77.53%；所提网络相较于频域特征增强的网络 R2RNet、UHDFour 相比，PSNR 指标分别提升了 20.60%、5.53%，SSIM 指标分别提升了 9.52%、2.48%。实验结果验证了所提网络 DF-DFANet 的有效性，其对低照度图像增强的效果最佳，PSNR 为 24.3714，SSIM 为 0.8937，LPIPS 为 0.1525，表明所提网络在恢复低照度图像上，增强图像画面质量更高，保留图像细节信息更丰富，与真实图像结构相似性更高。

表 1 LOL 真实低照度数据集测试结果

Table 1 LOL real-world dataset results

Method	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow
RetinexNet ^[26]	16.7740	0.4250	0.4739
Zero-DCE ^[27]	14.8607	0.5624	0.3352
DSLR ^[28]	14.9822	0.5964	0.3757
KinD ^[29]	17.6476	0.7715	0.1750
EnGAN ^[30]	17.4829	0.6515	0.3223
GLAD ^[32]	19.7182	0.6820	0.3994
RUAS ^[31]	16.4047	0.5034	0.2078
R2RNet ^[10]	20.2070	0.8160	-
UHDFour ^[8]	23.0926	0.8720	-
URetinexNet ^[6]	21.3282	0.8348	-
Ours	24.3714	0.8937	0.1525

3.4 MIT-Adobe FiveK 数据集实验与分析

为验证所提网络对不同场景的泛化性能，在 MIT-

Adobe FiveK 数据集上进行测试实验，对比网络分别为 RetinexNet^[26]、Zero-DCE^[27]、DSLR^[28]、KinD^[29]、EnGAN^[30]、RUAS^[31]、URetinexNet^[6]、R2RNet^[10] 和 UHDFour^[8]。实验结果如图 6 所示，从图中可发现，在低照度增强方面，Zero-DCE、RUAS 的增强图像曝光值仍较低，如图 6(c)、6(g) 所示，第 (2) 行树丛阴影面积较大，难以识别具体内容，由于无监督网络训练较难，且其特征提取部分网络结构设计过于简单，导致特征图难以精确映射图像不同区域亮度水平，而所提网络在极低照度区域仍能有效提取丰富底层特征，获得精细化的特征图以反应图像光照条件变化。在图像降噪方面，DSLR、KinD、URetinexNet 的增强图像对比度较低，颜色饱和度较低，如图 6(d)、6(e)、6(h) 所示，第 (1) 行停车场区域降噪效果优于其他对比网络，但因地面区域纹理结构复杂，降噪后仍存在较多噪声颗粒，而所提网络通过频域自适应学习噪声类型，区分噪声信号和高频纹理细节，使得增强图像纹理细节清晰度更高。在图像颜色恢复方面，RetinexNet、EnGAN 的增强结果存在较为明显的颜色失衡，如图 6(b)、6(f) 所示，第 (2) 行花丛区域调整曝光水平效果较好，但颜色饱和度过高，与真实图像存在较大差距，而所提网络通过双域特征聚合模块进行像素级恢复低照度图像颜色、纹理信息，提高增强图像视觉感知质量。在图像细节恢复方面，R2RNet、UHDFour 利用频域特征增强图像细节信息，如图 6(i)、6(j) 所示，与其他对比网络相比，增强图像对比度略有提高，

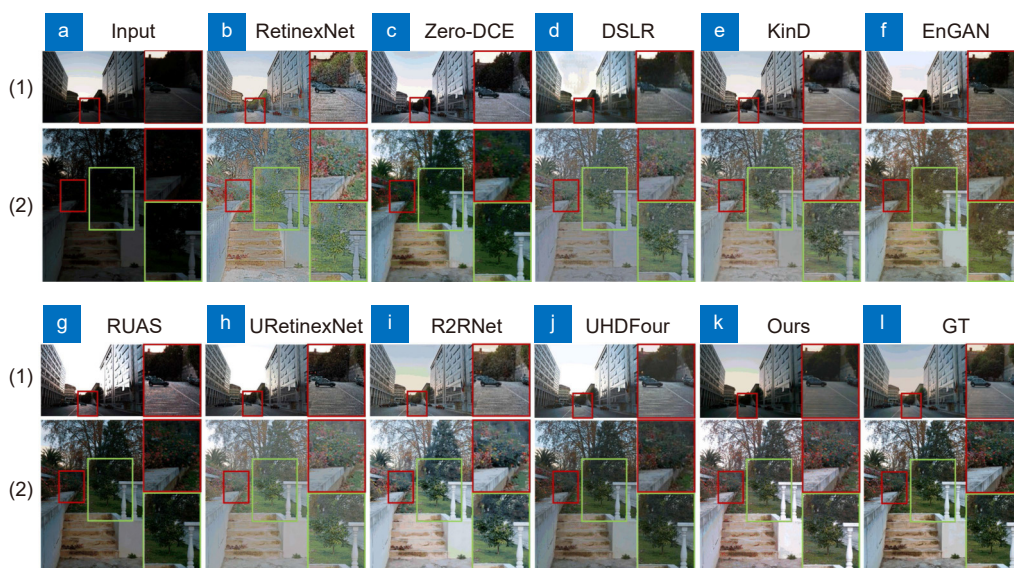


图 6 MIT-Adobe FiveK 数据集增强结果对比

Fig. 6 Comparison of enhancement results of mit-adobe fivek dataset

但所提网络通过多谱双注意力模块提取低照度图像细节特征图, 获得融合增强图像清晰度更好, 视觉效果更具自然性。综上所述, 所提网络在 MIT-Adobe FiveK 数据集上的效果图主观评价对比, 验证了网络具有较好的鲁棒性和泛化能力。

为进一步定量评价所提网络的泛化性能, 采用 PSNR、SSIM、LPIPS 作为衡量增强效果的评价指标, 对比结果如表 2 所示。由表中可知, 在 MIT-Adobe FiveK 数据集上, 由于数据集取景场地室外风景、日常生活较多, 采集图像内容丰富且光照变化剧烈, 导致对比网络难以适应不同场景的低照度图像, 所提网络相比于 DSLR、KinD 测试结果的 PSNR 指标分别提高了 12.24%、40.22%, SSIM 提升了 5.27%、11.28%, LPIPS 降低了 24.44%、23.03%。数据集中部分低照度低质图像大部分来自室外建筑, 其内容细节丰富、复杂弱光条件下的特征密集区域提升特征提取的难度较大, 所提网络相比于 Zero-DCE、RUAS, PSNR 指标分别提高了 42.62%、42.05%, SSIM 提升了 13.79%、10.97%, LPIPS 降低了 29.99%、17.46%, 这是因为对比网络的降噪部分将建筑物边缘细节视为噪声, 影响最终还原图像质量, 而所提网络利用注意力机制实现在低照度增强过程中抑制了不同类型图像噪声, 较好保留物体边缘细节。

表 2 MIT-Adobe FiveK 数据集测试结果

Table 2 MIT-Adobe FiveK dataset results

Method	PSNR↑	SSIM↑	LPIPS↓
Exposure ^[33]	18.7412	0.8159	0.1674
CycleGAN ^[34]	19.3823	0.7852	0.1636
RetinexNet ^[26]	12.5146	0.6708	0.2535
DSLR ^[28]	20.2435	0.8289	0.1526
KinD ^[29]	16.2032	0.7841	0.1498
EnGAN ^[30]	17.9050	0.8361	0.1425
Zero-DCE ^[27]	15.9312	0.7668	0.1647
Zero-DCE++ ^[35]	14.6111	0.4055	0.2309
RUAS ^[31]	15.9953	0.7863	0.1397
Ours	22.7214	0.8726	0.1153

3.5 消融实验与性能分析

3.5.1 多谱双注意力模块的消融实验

为进一步验证多谱双注意力模块中的通道和空间注意力之间连接方式的有效性, 在 LOL 数据集上对所提网络的 DDFAM 模块进行消融实验。Baseline 组设置普通卷积层替换通道和空间注意力块, Serial of CA & SA 组设置注意力块组合方式为串行结构, Parallel of CA & SA 组设置注意力块组合方式为并行结构, 实验效果如图 7 所示, 训练过程 PSNR 变化曲线对比如图 8 所示, 实验数据结果如表 3 所示。从中可看出, DDFAM 模块中的通道和空间注意力采用并行

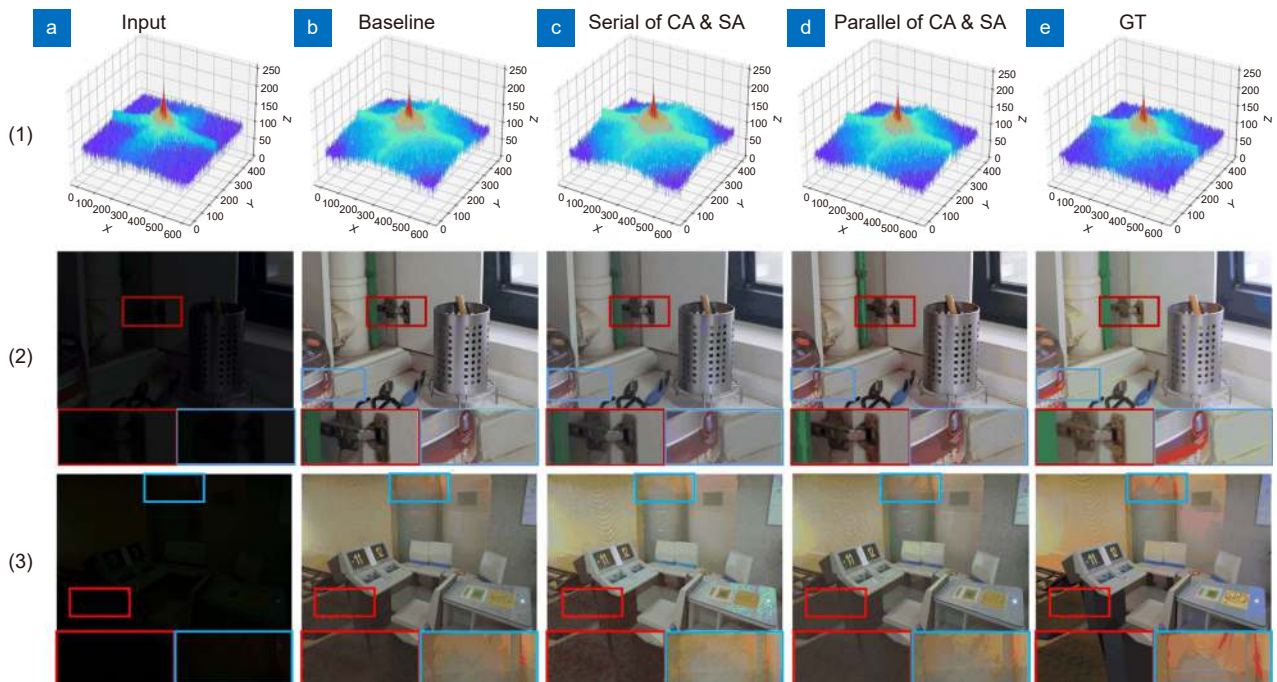


图 7 模块注意力结构实验效果对比

Fig. 7 Comparison of experimental effects of modular attention structure

的结构设计时, 如图 7 (b)、7(d)、7(e) 第 (2) 行所示, 带注意力机制的模块在降噪方面效果更好, 铁锈区域纹理颜色特征恢复更接近真实图像的表达效果, 且如图 7 第一行的三维频谱特征图所示, 所提网络采用并行结构的特征图在保留低频和高频部分更具优势; 而当模块中注意力组合采用串行结构时, 如图 7 (c)、7(d)、7(e) 第 (3) 行所示, 由于模块中不同的注意力机制排列方式, 使用注意力实现降噪任务的过程中仍造成较多图像的浅层纹理和颜色信息的损失, 灯光区域颜色饱和略低于真实图像, 且区域纹理过于平滑。如图 8 所示, 在网络拟合过程中的 PSNR 指标略高于模块注意力采用串行结构的设计方式, 将两路注意力分支并联的结构设计, 实现在通道堆叠后特征图仍能保留较多各自通道、空间维度信息, 缓解特征图混合后出现信息丢失的现象, 再通过卷积层进行融合的方式能够较多保留原有的图像特征, 从而提高增强图像的恢复质量。从表 3 可知, 和串行注意力结构对比, DDFAM 模块采用并行注意力结构, PSNR 提高了 3.25%, SSIM 提高了 8.09%, LPIPS 降低了 16.98%。

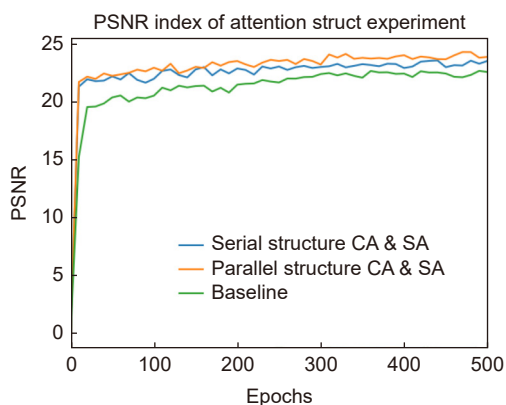


图 8 模块注意力结构的 PSNR 结果对比

Fig. 8 Comparison of PSNR results for module attention structure

表 3 模块注意力结构测试对比结果

Table 3 Comparison results of module attention structure testing

Method	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow
Baseline	22.7052	0.8147	0.2078
With serial of CA & SA	23.6042	0.8283	0.1837
With parallel of CA & SA	24.3714	0.8937	0.1525

3.5.2 网络模块消融实验

为验证所提网络各模块对低照度图像增强的有效性, 在 LOL 低照度数据集上对各模块进行消融实验。Baseline 设置一路分支为 U-Net 网络, 另一路为标准卷积层堆叠结构, 特征融合部分为逐像素相加的操作。FDIEM 表示基于频谱光照估计模块的傅里叶域分支,

MSAM 表示嵌入多谱双注意力模块和空洞卷积的小波域分支, DDFAM 表示双域特征聚合模块。实验结果如表 4 所示。由表中可知, 所提网络选用的 Baseline 网络能有效实现低照度图像增强, 但是其 PSNR 和 SSIM 仅为 20.8620、0.8515; 对比 Baseline、Model-1 和 Ours 的实验数据, FDIEM 模块有效提升网络对特征的提取能力, 逐层融合所提取到的特征信息, 扩大多尺度感受野范围的同时有利于丰富增强图像的语义信息; 对比 Baseline、Model-2 和 Ours 的实验数据可知, MSAM 模块对提升图像质量具有较好的效果; 对比 Baseline、Model-3 和 Ours 实验组, 表明在图像融合阶段引入激活函数对像素进行权重分配, 有利于提升图像色彩的恢复效果, 通过频域融合特征有利于提升图像质量; 最终, 所提网络 DF-DFANet 的 PSNR 和 SSIM 分别达到了 24.3714、0.8937, 各个模块相互作用使得增强图像各项评价指标表现更好, 视觉效果更具自然性。

表 4 网络模块消融实验结果

Table 4 Experimental results of network module ablation

Model	FDIEM	MSAM	DDFAM	PSNR	SSIM
Baseline	×	×	×	20.8620	0.8515
Model-1	×	√	√	21.3582	0.8653
Model-2	√	×	√	22.0401	0.8878
Model-3	√	√	×	21.9068	0.8919
Ours	√	√	√	24.3714	0.8937

对所提网络各个模块进行消融实验, 其效果图对比如图 9 所示。对比如 9 (b)、9 (e)、9 (f) 第 (2) 行可知, 9 (b) 列图中木质细节纹理保持较好但亮度矫正效果差且存在较多噪声颗粒, 验证 FDIEM 模块能够有效提取低照度图像轮廓信息, 通过频域特征图具有丰富上下文语义信息引导网络提取不同频域信息分量, 并且具有较好的图像降噪效果; 对比如 9 (c)、9 (e)、9 (f) 第 (3) 行可知, 9 (c) 列图中场地白线边缘出现局部模糊现象, 验证 MSAM 模块具有较好的细节特征提取能力, 提高网络恢复图像细节信息的能力; 对比如 9 (d)、9 (e)、9 (f) 第 (1) 行可知, 9 (d) 列中图像细节信息和纹理结构都保持较好, 但是三维频谱特征图频率分量与所提网络和真实图像相比仍存在不足, 影响增强图像视觉感知效果, 验证所提网络加入 DDFAM 模块后融合特征图能够缓解包含不同信息的特征之间相互作用影响, 提高图像融合效果。对比如 9 (e)、

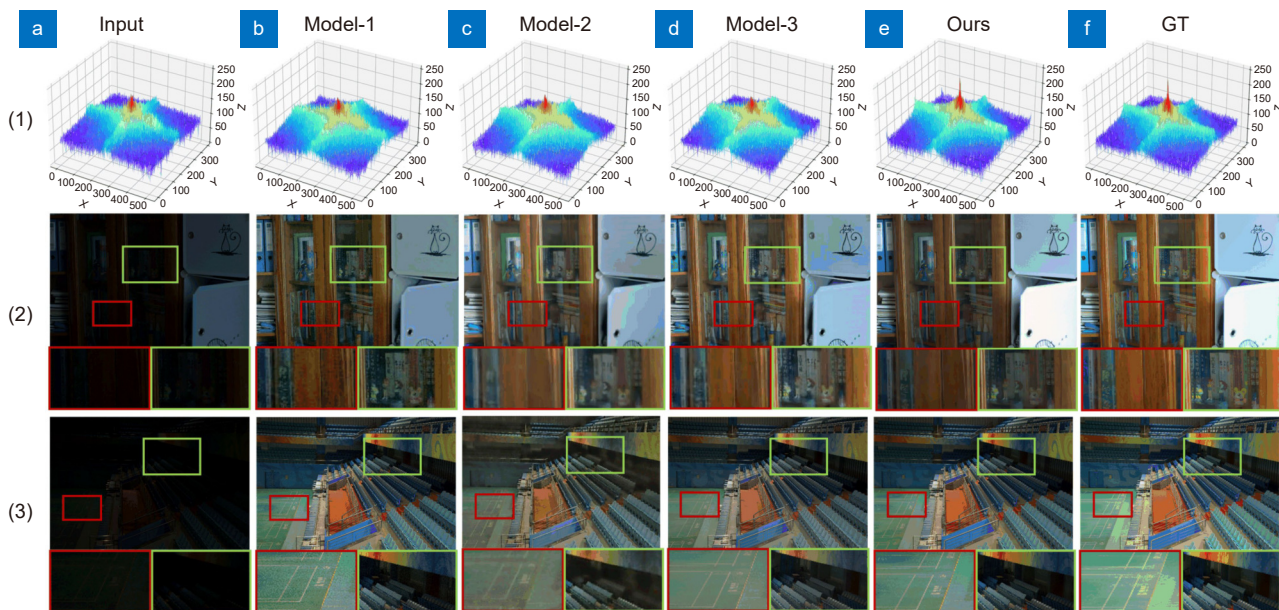


图 9 模块消融实验效果图对比

Fig. 9 Comparison of effect diagrams of modular ablation experiments

9(f) 第 (1) 行三维频谱图可知, 增强图像和真实图像在频域特征上相似性更高, 保持低、高频特征更具优势。

3.5.3 网络性能对比实验

为验证所提网络的性能, 与对比网络在 LOL 低照度测试数据集上进行性能测试。低照度图像增强处理单张图像的平均时间、模型大小、浮点运算量与 PSNR 对比实验数据如表 5 所示。由表 5 可知, ZeroDCE 虽然处理速度最快, 用时仅 2 ms, 利用亮度估计曲线实现快速增强处理, 且模型大小和浮点运算量最少, 仅为 0.97 M、5.2112 G, 但是其 PSNR、SSIM 指标不高, 仅为 14.8671、0.5624。MBLLEN 的模型参数量较小, 大小为 1.95 M, 但通过堆叠式去噪自编码器实现低照度图像增强, 导致其单图处理速度最慢, 用时为 80 ms。GLAD、RetinexNet、RDGAN 等网络处理速度保持在 25 ms 左右, 且模型大小均在 10 M 左右, 但是 PSNR 和 SSIM 指标均较低, 同时 FLOPs 浮点运算量略有增加。KinD 处理单张处理速度较快保持在 10 ms 左右, 并且 FLOPs 运算量较少, 仅为 29.1303 G, 但其 PSNR 指标不高且模型较大为 35 M。URetinexNet 增强效果较好, PSNR 和 SSIM 分别为 21.3282、0.8348, 且代码参数量 Para ms 仅为 0.3401, 但是其较高的浮点运算量 FLOPs 占用过多的计算资源, 达到 1801.4110 G。最后, 虽然所提网络浮点运算量较大, 达到 288.3776 G, 影响网络训练和推理速度, 但是增强图像具有较好的

图像质量, PSNR 为 24.3714, SSIM 为 0.8937。

表 5 不同网络的 PSNR 和平均处理时间、模型大小和浮点运算量对比

Table 5 Comparison of different network average processing time, model size and floating-point operations

Model	Time/ms	Params/M	FLOPs/G	PSNR	SSIM
RetinexNet ^[26]	20	9.2	136.0151	16.7740	0.4250
Zero-DCE ^[27]	2	0.97	5.2112	14.8671	0.5624
KinD ^[29]	10	35	29.1303	20.3792	0.7715
EnGAN ^[30]	20	33	61.0102	17.4828	0.6515
GLAD ^[32]	25	11	252.1410	19.7182	0.6820
MBLLEN ^[36]	80	1.95	19.9560	17.8583	0.7247
LPNet ^[37]	18	0.15	0.7700	21.4612	0.8020
URetinexNet ^[6]	2.93	0.34	1801.4110	21.3282	0.8348
Ours	48	1.61	288.3776	24.3714	0.8937

3.6 真实场景增强实验

低照度图像增强在视频监控领域有着广泛的应用, 对车辆管理、交通安全和车流控制都有着极大的应用价值, 为了验证所提网络的实际应用效果, 在现实车辆监控场景下进行实验与分析, 实验结果如图 10 所示。从图中可知, RetinexNet 的增强图像中车辆颜色存在色彩失衡问题, 如图 10(b) 所示, 整体图像的车辆颜色饱和度太高, 图像色彩扭曲影响检测任务的有效性。ZeroDCE 增强图像的车牌字符较为清晰, 边缘纹理保持较好, 但增强图像亮度是仍偏低, 如

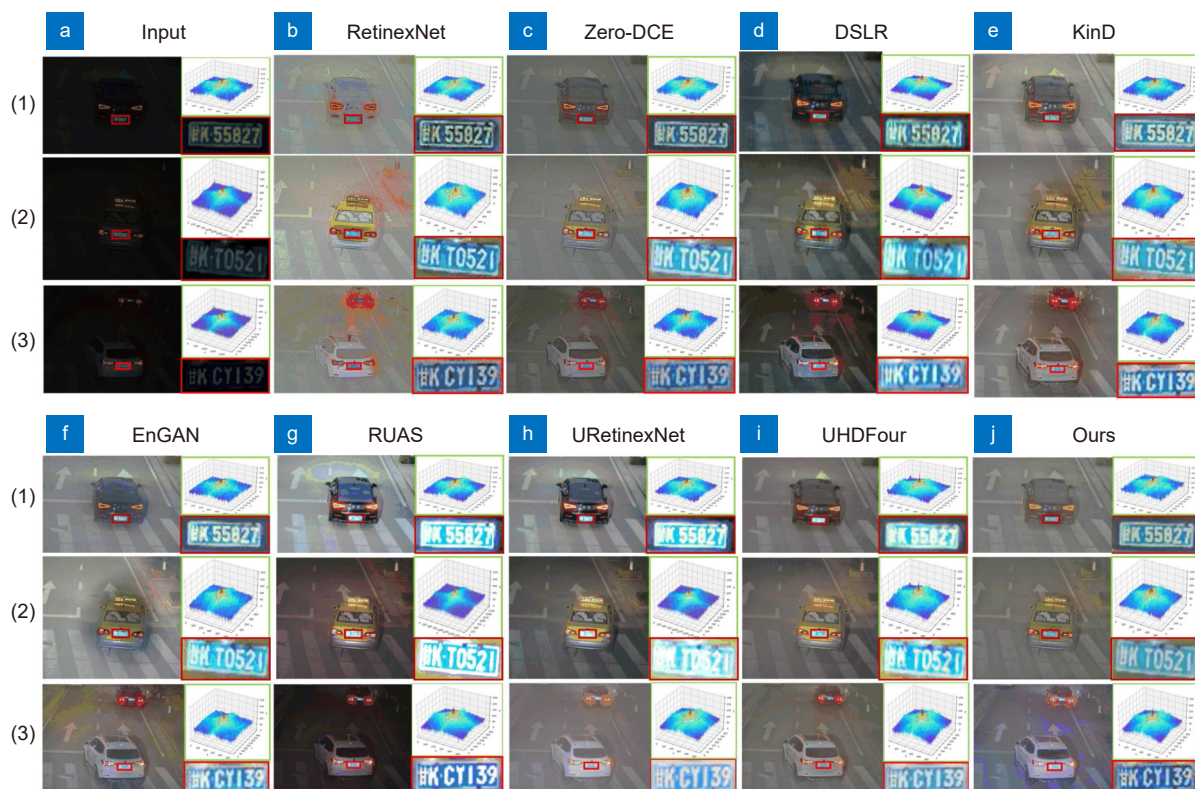


图 10 夜间低照度车辆监控图像测试结果

Fig. 10 Test results of monitoring images of low-light vehicles at night

图 10(c) 所示, 整体图像与所提网络相比曝光值较低。DSLR 增强图像如图 10(d) 所示, 第 (2) 行出租车区域出现伪影和模糊现象, 图像结构信息损失严重, 且从三维频谱图看出, 其高频分量与所提网络相比大大降低。kinD、EnGAN 恢复低照度车辆监控图像效果较好, 但车辆周围区域存在大面积阴影, 影响视觉感官体验, 如图 10(e)、图 10(f) 所示, 第 (2) 行出租车周围存在阴影。RUAS、URetinexNet 对车辆监控弱光图像进行增强处理, 但图像亮度值难以保持在正常范围内, 影响检测任务的准确率, 如图 10(g) 所示, 第 (2) 行增强图像亮度值偏低, 但车牌区域出现过曝现象, 第 (1) 行增强图像亮度值过高, 车牌区域出现过曝, 影响车牌识别精确度, 且从三维频谱图可知, 图像高频分量损失严重。所提网络调节低照度图像曝光值方面表现优异, 图像画面清晰明亮, 对车牌字符边缘细节保留较好, 提升车牌定位和识别任务准确性, 具有较强稳定性和泛化能力。

4 结论

针对低照度图像增强中出现的边缘模糊、图像噪

声等退化问题, 提出一种基于双频域特征聚合的低照度增强网络。在傅里叶域分支上通过逐层融合频谱光照估计模块提取的频域特征图, 扩大特征图感受野范围, 结合丰富的上下文语义信息获得精细化光照图; 在小波域分支上嵌入多谱双注意力模块, 对小波特征图利用空间和通道注意力, 提升网络关注图像高频细节特征的能力; 双域特征聚合模块利用激活函数获得图像像素分配权重, 实现对增强图像更精细化的调整, 提高网络恢复图像颜色、纹理的能力。LOL 数据集对比实验表明, 所提网络 PSNR 达到 24.3714, SSIM 达到 0.8937。在 MIT-Adobe FiveK 数据集上, PSNR、SSIM 分别达到了 22.7214、0.8726。此外, 本文在实际应用场景进行了实验测试, 增强效果稳定性良好且具有较好鲁棒性和泛化能力, 但所提网络浮点运算量 FLOPs 较大, 故降低网络计算量是下一步改进方向。

参考文献

- [1] Zhu M F, Pan P B, Chen W, et al. EEMEFN: low-light image enhancement via edge-enhanced multi-exposure fusion network[C]//Proceedings of the 34th AAAI Conference on Artificial Intelligence, 2020: 13106-13113. <https://doi.org/10.1609/aaai.v34i07.7013>.

- [2] Li C L, Tang S Q, Yan J W, et al. Low-light image enhancement based on quasi-symmetric correction functions by fusion[J]. *Symmetry*, 2020, **12**(9): 1561.
- [3] Pan X X, Li C L, Pan Z G, et al. Low-light image enhancement method based on retinex theory by improving illumination map[J]. *Appl Sci*, 2022, **12**(10): 5257.
- [4] Li P, Liang D, Liang D T, et al. Research on defect inspection method of pipeline robot based on adaptive image enhancement[J]. *Opto-Electron Eng*, 2020, **47**(1): 190304.
李平, 梁丹, 梁冬泰, 等. 自适应图像增强的管道机器人缺陷检测方法[J]. *光电工程*, 2020, **47**(1): 190304.
- [5] Zhao R N, Han Y, Zhao J. End-to-end retinex-based illumination attention low-light enhancement network for autonomous driving at night[J]. *Comput Intell Neurosci*, 2022, **2022**: 4942420.
- [6] Wu W H, Weng J, Zhang P P, et al. Uretinex-Net: retinex-based deep unfolding network for low-light image enhancement[C]//*Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022: 5891–5900. <https://doi.org/10.1109/CVPR52688.2022.00581>.
- [7] Jiang Q P, Mao Y D, Cong R M, et al. Unsupervised decomposition and correction network for low-light image enhancement[J]. *IEEE Trans Intell Transp Syst*, 2022, **23**(10): 19440–19455.
- [8] Li C Y, Guo C L, Zhou M, et al. Embedding fourier for ultra-high-definition low-light image enhancement[C]//*The Eleventh International Conference on Learning Representations*, 2023.
- [9] Zhang Y C, Liu H Y, Ding D D. A cross-scale framework for low-light image enhancement using spatial-spectral information[J]. *Comput Electr Eng*, 2023, **106**: 108608.
- [10] Hai J, Xuan Z, Yang R, et al. R2RNet: low-light image enhancement via real-low to real-normal network[J]. *J Vis Commun Image Represent*, 2023, **90**: 103712.
- [11] Lin X, Yue J T, Ren C, et al. Unlocking low-light-rainy image restoration by pairwise degradation feature vector guidance[Z]. arXiv: 2305.03997, 2023. <https://doi.org/10.48550/arXiv.2305.03997>.
- [12] Xu J Z, Yuan M K, Yan D M, et al. Illumination guided attentive wavelet network for low-light image enhancement[J]. *IEEE Trans Multimedia*, 2023, **25**: 6258–6271.
- [13] Fan C M, Liu T J, Liu K H. Half wavelet attention on M-Net+ for low-light image enhancement[C]//*2022 IEEE International Conference on Image Processing (ICIP)*, 2022: 3878–3882. <https://doi.org/10.1109/ICIP46576.2022.9897503>.
- [14] Hu C, Chen X J, Wu Y K. Research on image enhancement algorithm of low illumination image based on half wave attention mechanism[J]. *Laser J*, 2023.
胡聪, 陈绪君, 吴雨锴. 融合半波注意力机制的低光照图像增强算法研究[J]. *激光杂志*, 2023.
- [15] Chen Z L, Liang Y L, Du M H. Attention-based broad self-guided network for low-light image enhancement[C]//*2022 26th International Conference on Pattern Recognition (ICPR)*, 2022: 31–38. <https://doi.org/10.1109/ICPR56361.2022.9956143>.
- [16] Chi L, Jiang B R, Mu Y D. Fast Fourier convolution[C]//*Proceedings of the 34th International Conference on Neural Information Processing Systems*, 2020: 376.
- [17] Suvorov R, Logacheva E, Mashikhin A, et al. Resolution-robust large mask inpainting with Fourier convolutions[C]//*Proceedings of 2022 IEEE/CVF Winter Conference on Applications of Computer Vision*, 2022: 3172–3182. <https://doi.org/10.1109/WACV51458.2022.00323>.
- [18] Zamir S W, Arora A, Khan S, et al. Learning enriched features for fast image restoration and enhancement[J]. *IEEE Trans Pattern Anal Mach Intell*, 2022, **45**(2): 1934–1948.
- [19] Zhang G, Li Z Y, Li J M, et al. CFNet: cascade fusion network for dense prediction[Z]. arXiv: 2302.06052, 2023. <https://doi.org/10.48550/arXiv.2302.06052>.
- [20] Liu G H, Yang Q, Meng Y B, et al. A progressive fusion image enhancement method with parallel hybrid attention[J]. *Opto-Electron Eng*, 2023, **50**(4): 220231.
刘光辉, 杨琦, 孟月波, 等. 一种并行混合注意力的渐进融合图像增强方法[J]. *光电工程*, 2023, **50**(4): 220231.
- [21] Li X, Wang W H, Hu X L, et al. Selective kernel networks[C]//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019: 510–519. <https://doi.org/10.1109/CVPR.2019.00060>.
- [22] Jiang J X, Ye T, Bai J B, et al. Five A* network: you only need 9k parameters for underwater image enhancement[Z]. arXiv: 2305.08824, 2023. <https://doi.org/10.48550/arXiv.2305.08824>.
- [23] Starovoitov V V, Eldarova E E, Iskakov K T. Comparative analysis of the SSIM index and the Pearson coefficient as a criterion for image similarity[J]. *Eurasian J Math Comput Appl*, 2020, **8**(1): 76–90.
- [24] Tao X C, Zhu T, Huang Y L, et al. Low-quality image enhancement algorithm based on DDR GAN[J]. *Laser Technol*, 2023, **47**(3): 322–328.
陶昕辰, 朱涛, 黄玉玲, 等. 基于DDR GAN的低质量图像增强算法[J]. *激光技术*, 2023, **47**(3): 322–328.
- [25] Fuoli D, Van Gool L, Timofte R. Fourier space losses for efficient perceptual image super-resolution[C]//*Proceedings of 2021 IEEE/CVF International Conference on Computer Vision*, 2021: 2340–2349. <https://doi.org/10.1109/ICCV48922.2021.00236>.
- [26] Wei C, Wang W J, Yang W H, et al. Deep retinex decomposition for low-light enhancement[C]//*British Machine Vision Conference 2018*, 2018.
- [27] Guo C L, Li C Y, Guo J C, et al. Zero-reference deep curve estimation for low-light image enhancement[C]//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020: 1777–1786. <https://doi.org/10.1109/CVPR42600.2020.00185>.
- [28] Lim S, Kim W. DSLR: deep stacked Laplacian restorer for low-light image enhancement[J]. *IEEE Trans Multimedia*, 2021, **23**: 4272–4284.
- [29] Zhang Y H, Zhang J W, Guo X J. Kindling the darkness: a practical low-light image enhancer[C]//*Proceedings of the 27th ACM International Conference on Multimedia*, 2019: 1632–1640. <https://doi.org/10.1145/3343031.3350926>.
- [30] Jiang Y F, Gong X Y, Liu D, et al. EnlightenGAN: deep light enhancement without paired supervision[J]. *IEEE Trans Image Process*, 2021, **30**: 2340–2349.
- [31] Liu R S, Ma L, Zhang J A, et al. Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement[C]//*Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021: 10556–10565. <https://doi.org/10.1109/CVPR46437.2021.01042>.
- [32] Jiao Y, Zheng X T, Lu X Q. Attention-based multi-branch network for low-light image enhancement[C]//*2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*, 2021: 401–407. <https://doi.org/10.1109/ICBAIE52039.2021.9389960>.
- [33] Hu Y M, He H, Xu C X, et al. Exposure: a white-box photo post-

- processing framework[J]. *ACM Trans Graph*, 2018, **37**(2): 26.
- [34] Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]// *Proceedings of 2017 IEEE International Conference on Computer Vision*, 2017: 2242–2251. <https://doi.org/10.1109/ICCV.2017244>.
- [35] Li C Y, Guo C L, Loy C C. Learning to enhance low-light image via zero-reference deep curve estimation[J]. *IEEE Trans Pattern Anal Mach Intell*, 2022, **44**(8): 4225–4238.
- [36] Lv F F, Lu F, Wu J H, et al. MBLLN: low-light image/video enhancement using CNNs[C]// *British Machine Vision Conference 2018*, 2018.
- [37] Li J, Li J, Fang F, et al. Luminance-aware pyramid network for low-light image enhancement[J]. *IEEE Trans Multimedia*, 2020, **23**: 3153–3165.

作者简介



徐胜军(1976-), 男, 工学博士, 副教授, 硕士生导师, 主要研究方向为图像处理、模式识别。
E-mail: duplin@sina.com



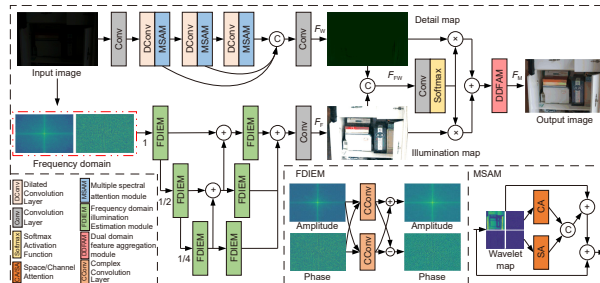
【通信作者】杨华(1997-), 男, 硕士研究生, 主要研究方向为图像处理、深度学习。
E-mail: yhxuat@163.com



扫描二维码, 获取PDF全文

Low-light image enhancement based on dual-frequency domain feature aggregation

Xu Shengjun^{1,2}, Yang Hua^{1,2*}, Li Minghai¹, Liu Guanghui^{1,2}, Meng Yuebo^{1,2}, Han Jiuqiang^{1,2}



Overview: Road monitoring is an important part of the field of intelligent transportation. However, in the night scene under the condition of low illumination, the brightness and contrast of the images collected by the camera are low, and there are more noise particles, which brings difficulty to the visual tasks such as detection and recognition of important targets in the field of traffic supervision. Although deep learning has achieved certain results in the enhancement of low-light images, it is easy to amplify shadow noise while enhancing brightness and contrast. Unreasonable noise reduction strategies often lead to different degrees of detail blur in the image, especially for low-light images with poor picture quality, it is often difficult to restore the lost texture structure. To solve these problems, a dual-frequency domain based feature aggregation network (DF-DFANet) is proposed. Firstly, the spectral illumination estimation module (FDIEM) is designed to extract the global features of the image through the Fourier domain spectral feature map and reduce the response to the noise signal while pulling up the brightness of the image in the frequency domain. Secondly, a multispectral dual attention module (MSAM) is proposed, which uses the spatial and channel attention mechanism to make the network focus on the important features of the Baud sign subgraph and improves the ability of the network to recover image details. Finally, a dual-domain feature aggregation module (DDFAM) was constructed to learn the adaptive weight parameters of different pixel level features, and the complex domain convolution was used to promote the fusion of feature information, which enhanced the naturalness of image color performance and the richness of texture details. In the Fourier domain branch, the frequency domain feature map extracted by the spectral illumination estimation module is fused layer by layer, the range of the sensitivity field of the feature map is expanded, and the refined illumination map is obtained by combining rich contextual semantic information. The multi-spectral dual attention module is embedded in the branch of the wavelet domain, and the space and the channel attention are used to improve the ability of the network to pay attention to the high-frequency detail features of the image. Dual-domain feature aggregation module uses an activation function to obtain image pixel allocation weight, realizes more refined adjustment of the enhanced image, and improves the ability of the network to restore image color and texture. Comparative experiments on the LOL dataset show that the PSNR and SSIM of the proposed network reach 24.3714 and 0.8937. On the MIT-Adobe FiveK dataset, PSNR and SSIM reach 22.7214 and 0.8726, respectively. In addition, the proposed method has been tested in practical application scenarios, and the enhancement effect has good stability, robustness, and generalization ability.

Xu S J, Yang H, Li M H, et al. Low-light image enhancement based on dual-frequency domain feature aggregation[J]. *Opto-Electron Eng*, 2023, 50(12): 230225; DOI: [10.12086/oe.2023.230225](https://doi.org/10.12086/oe.2023.230225)

Foundation item: Project supported by General Program of the National Natural Science Foundation of China (52278125), Shaanxi Provincial Key R&D Plan (2020GY-186, 2021SF-429), and Shaanxi Provincial Natural Science Basic Research Plan (2023-JC-YB-532, 2022JQ-681)

¹College of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an, Shaanxi 710055, China; ²Xi'an Key Laboratory of Building Manufacturing Intelligent & Automation Technology, Xi'an, Shaanxi 710055, China

* E-mail: yhxauat@163.com