

基于双域特征交互和局部相关性上采样的 单幅图像去雾方法

刘明杰¹, 吕梦琳¹, 刘平¹, 陈俊生¹, 朴昌浩^{1*}, 康宗绪²

(1. 重庆邮电大学自动化学院, 重庆 400065; 2. 重庆金美通信有限责任公司, 重庆 400060)

摘要: 雾霾天大气中的悬浮颗粒会导致可见光系统的成像质量显著降低, 产生图像对比度下降、颜色失真、细节信息丢失等图像退化现象. 这种退化会严重影响计算机视觉下游任务的性能. 因此, 通常会把图像去雾作为诸如目标检测、目标分割等高级视觉任务的预处理过程, 旨在为后续任务提供高质量的图像内容信息. 基于 U 型网络的图像去雾架构因其具有高效、注重细节特征、轻量化等特点而得到了广泛关注. 然而, 当前该类架构主要聚焦于通过提取图像的空间域特征达成去雾效果, 忽略了频域特征对图像去雾的影响. U 型结构中解码层采用的最近邻插值法上采样方法还会导致图像空间信息的丢失, 进而影响深层语义信息向浅层的有效传递, 使得重构的清晰图像不够理想. 针对上述问题, 本文以 U-Net 为基础结构提出一种双域特征交互和局部相关性上采样的单幅图像去雾方法. 首先在特征提取部分设计了双域特征交互模块, 分别通过双路特征融合子模块和频域特征增强子模块针对图像的空间域与频域特征分别提取, 并对这些特征进行交互通过引入频域信息提升模型对图像结构特征的捕捉能力. 接着在解码层上采样部分设计局部相关性上采样模块, 利用注意力机制捕捉每个特征图中局部信息的内在相关性, 并将含有特征补偿的深层语义特征传递至浅层, 有效地实现了不同尺度特征中语义信息的提取和融合. 此外, 为了直观表征算法的去雾性能, 本文提出一种基于热力图的对比分析方法, 即通过颜色梯度直观量化去雾效果差异, 其可以有效反映各去雾方法在图像细节还原方面的性能差异. 实验结果表明, 本文提出的图像去雾方法在确保较低参数量和模型计算复杂度的同时, 在客观指标和主观视觉效果上均取得了较好的效果, 在 SOTS-Indoor, SOTS-Outdoor 和 Haze4K3 个数据上的峰值信噪比 (Peak Signal Noise Ratio, PSNR) 和结构相似性指数 (Structural Similarity Index Measure, SSIM) 分别达到了 41.46 dB 和 0.994 3、37.73 dB 和 0.993 6、34.72 dB 和 0.993.

关键词: 图像去雾; U 型网络结构; 空间域与频域特征交互; 局部相关性上采样; 信息融合

基金项目: 重庆市技术创新与应用发展专项重大项目 (No.CSTB2023TIAD-STX0035)

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 0372-2112(2025)12-4349-15

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20250429

Dual-Domain Feature Interaction and Local Correlation Upsampling Network for Single Image Dehazing

LIU Ming-jie¹, LYU Meng-lin¹, LIU Ping¹, CHEN Jun-sheng¹, PIAO Chang-hao^{1*}, KANG Zong-xu²

(1. School of Automation, Chongqing University of Posts and Telecommunications, Chongqing 400065, China;

2. Aerospace Jinmei (Chongqing) Communication Co., Ltd., Chongqing 400060, China)

Abstract: Suspended particulates of the atmosphere in hazy weather markedly degrade the imaging quality of visible light systems, which manifests as reduced image contrast, color distortion, and loss of fine-grained details. Such image deterioration substantially impairs the performance of computer vision tasks. Consequently, image dehazing is commonly employed as a preprocessing step for high-level visual tasks to furnish processes with high-quality visual data. U-Net-based image dehazing architecture has garnered widespread attention due to its efficiency, detail-oriented feature extraction, and lightweight characteristics. However, current U-Net-based networks realize image dehazing based on features extracted from space domain, ignoring the impact of features in frequency domain. In addition, the decoder of U-Net-based networks always realizes feature upsampling by nearest neighbor interpolation. It may cause spatial information loss and impact se-

mantic information transmission from high-level to low-level, which adversely affects clear image restriction. To address the above issues, this paper proposes a novel image dehazing algorithm with dual-domain feature interaction and local correlation upsampling. Specifically, the dual-domain feature interaction module, including dual-path feature fusion submodule and frequency domain feature enhancement sub-module, is designed to extract and fuse the spatial domain and frequency domain features of the image. It can enhance the ability to capture the structural features of the image by introducing frequency domain information. Local correlation upsampling module embedded in decoder of U-Net is designed to capture the intrinsic correlation of local information of each feature map by attention mechanism, and transmit the high-level features with the compensatory information the low-level features simultaneously. In addition, we propose a contrast analysis method based on heat maps to visually the dehazing performance of different methods, which uses color gradients to quantitatively measure the differences in the dehazing effect. It can effectively reflect the performance differences of various dehazing methods in terms of image detail restoration. The experimental results demonstrate that the dehazing effect of our proposed method is superior to that of the compared method in both quantitative and qualitative evaluations. The peak signal noise ratio (PSNR) and structural similarity index measure (SSIM) values on the SOTS-Indoor, SOTS-Outdoor and Haze4K datasets achieve 41.46 dB and 0.994 3, 37.73 dB and 0.993 6, 34.72 dB and 0.993, respectively.

Key words: image dehaze; U-Net-based architecture; spatial and frequency domain feature interaction; local correlation upsampling; information fusion

Foundation Item(s): Chongqing Municipal Major Project for Technological Innovation and Application Development (No.CSTB2023TIAD-STX0035)

1 引言

视觉感知是人类获取外界信息的重要途径。图像作为视觉信息的载体,在智能交通、智能监控等实际应用中发挥着关键作用^[1]。然而,雾霾天大气中的悬浮颗粒会导致可见光系统的成像质量显著降低,产生图像对比度下降、颜色失真、细节信息丢失等图像退化现象^[2]。这种退化会严重影响诸如目标检测、目标跟踪、语义分割等计算机视觉任务的性能^[3,4]。因此,通常会把图像去雾作为高阶视觉任务的预处理过程,旨在为后续任务提供高质量的图像内容信息^[5]。

早期的图像去雾算法多是利用先验信息估计传输图和全局大气光值等中间参数,然后通过大气散射模型求出清晰图像,进而达到图像去雾的效果。常见的利用先验信息的去雾方法如暗通道先验^[6]、非局部先验^[7]、颜色衰减先验^[8]等^[9-12]。虽然基于先验信息的去雾方法取得了显著进展,但是,这类方法通常带有统计特性,与真实场景存在误差,且这类算法的泛化能力也相对有限,因此很难达到较好的去雾效果。

随着深度学习技术的发展,研究者们开始利用深度学习来解决图像去雾问题,基于深度学习的图像去雾方法凭借不依赖物理先验建模、去雾精度较高等优点,成为当前去雾图像的主流技术。其主要分为基于中间参数估计的去雾方法和基于端到端模型的去雾方法。其中,基于中间参数估计的去雾方法利用卷积神经网络(Convolutional Neural Networks, CNNs)估计传输图和大气光值等大气散射模型的中间参数,再根据大气散射模型得到清晰的图像。该类方法取得了比传统的参数估计更好的去雾效果。但是在中间参数估计不准

的情况下,该类方法仍然会导致图像的颜色失真和伪影现象。

相较于基于中间参数估计的去雾方法,基于端到端的去雾方法通过数据驱动的方式直接建立有雾图像与清晰图像之间的映射关系。基于端到端的去雾方法又可细分为3类:基于生成对抗网络^[13]的去雾方法、基于Transformer^[14]的去雾方法和基于U-Net^[15]的去雾方法。基于生成对抗网络的去雾方法通过生成器和判别器的对抗训练,生成器被迫学习到有雾图像与清晰图像之间更本质的映射关系,从而生成高质量的去雾结果。尽管改良的生成对抗网络能处理非配对数据,有效解决真实场景中难以获取大量“有雾-无雾”图像对的问题,但需要额外设计判别器和对抗损失,训练过程复杂且不稳定,容易发生模式崩溃或训练振荡。并且,为了通过判别器的验证,生成器有时会生成不真实的纹理或伪影,影响图像的真实性。为了克服生成对抗网络在生成结果真实性方面的局限,并更好地建模图像全局的依赖关系,研究者们引入了基于Transformer的去雾方法。该方法通过自注意力机制建立图像所有区域的全局依赖关系,最后通过一个解码器将处理后的特征序列重新组合成去雾图像。虽然基于Transformer的去雾方法具有强大的全局信息建模能力,但是自注意力机制的复杂度与图像序列长度的平方成正比,导致对计算资源和内存的需求非常高。并且,Transformer模型参数量大,需要大量数据才能充分训练,但有雾数据集往往数据量有限。相较于Transformer对计算资源和数据集大小的高要求,基于U-Net的去雾方法则提供了一种更为高效且注重细节特征的轻量化架构。基于U-Net

的去雾方法利用编码器提取图像的多尺度上下文特征,利用解码器恢复图像分辨率,利用跳跃连接将编码器每层特征图与解码器对应层特征图进行通道拼接,使得图像重建过程能直接利用编码器保留的边缘信息和纹理信息,避免了细节信息的丢失。但是原始的U-Net网络全局感知能力弱,需要进一步改进网络来增强全局信息的提取能力。

考虑到U-Net网络训练的稳定性和轻巧的模型架构,本文选择基于U-Net网络构建单幅图像去雾网络。为了提高U-Net网络的全局感知能力,改进的去雾方法通常聚焦于利用卷积采样或者注意力机制提取图像空间域下的全局特征,从而重建去雾图像,在研究过程中往往忽视图像频域信息。Yu等人^[5]的研究揭示了雾退化与图像幅值谱和相位谱之间的关系,雾霾引起的退化特性主要体现在幅值谱中,有雾图像和清晰图像的相位谱差异较小。即雾度主要改变图像的照明对比使整个图像的可见性降低,图像的结构受到了温和的影响。因此,提取并分析图像的频域特征能够弥补U-Net网络全局感知能力较弱的不足。在此基础上,进一步将空间域的细节信息与频域的结构特征相融合,可以有效提升图像去雾的整体性能。并且,当前基于U-Net改进的端到端去雾网络多是关注编码过程中多尺度特征的融合,以及编码层和解码层之间设计特征交互融合,鲜少关注解码层上采样过程中方法的局限性。解码层多采用最近邻插值法进行上采样,该方法未考虑相邻像素之间的相关性,导致了插值后的图像出现像素值不连续、锯齿状图案或马赛克外观等现象,进而造成空间信息丢失,影响深层语义信息向浅层的有效传递。

为实现去雾图像空间域和频域特征的交互,且实现去雾网络解码层语义信息自下而上的有效传递,本文以U-Net为基础结构提出一种双域交互和局部相关性上采样的图像去雾网络(Dual-domain Feature interaction and Local attention Upsampling image dehazing Network, DFLU-Net)。具体而言,为了有效地利用不同域的信息,设计空间域-频域特征交互模块(Spatial Frequency Feature Interaction Module, SFFIM)在空间域和频域分别提取图像局部细节特征和全局结构(幅值谱和相位谱)特征,然后将两者进行深度融合,从而实现双域信息的有效交互。这些编码层提取的双域特征也将被传递到对应的解码层,减少编码层提取特征过程中的特征稀释。此外,为了提升解码层上采样过程中特征的语义信息自下而上的传递效果,本文设计了一个局部相关性上采样模块(Local Correlation Upsampling Module, LCUM),利用注意力机制捕捉每个特征图中局部信息的内在相关性,再利用双线性插值将信息向浅层传递。该模块实现了从不同尺度特征中有效提取图

像的语义信息,并实现该信息的层级传递,有效地丰富了解码层的语义特征表达,提高了网络的去雾效果。综上所述,本文的主要贡献如下。

(1)构建空间域-频域特征交互模块。分别通过双路特征融合子模块(Dual-Path Feature Fusion Module, DPFFM)和频域特征增强子模块(Frequency domain Feature Enhancement Module, FFEM)提取去雾图像空间域的局部细节特征和频域的全局结构特征,并对这些特征进行交互,有效地利用双域信息进行图像去雾处理。

(2)设计局部相关性上采样模块。利用注意力机制捕捉特征图中局部信息的内在相关性,并将其在不同层级进行传递,有效地实现了不同尺度特征中语义信息的提取和融合,丰富了解码层的语义特征表达,提高了网络的去雾效果。

(3)在数据集 Haze4K、Reside-SOTS-Indoor、Reside-SOTS-Outdoor上的实验结果表明,本文提出的图像去雾方法确保较低参数量和模型计算复杂度的同时,在客观指标和主观视觉效果上均取得了良好的去雾效果。

2 相关工作

传统的图像去雾方法大多是利用先验信息估计传输图和全局大气光值等中间参数,然后通过大气散射模型求出清晰图像,进而达到图像去雾的效果。He等人^[6]依据清晰图像至少在一个颜色通道上的像素值接近于零这一现象,提出了基于暗通道先验的图像去雾算法,首先根据图像的暗通道先验知识估计传输图,再通过大气散射模型得到去雾图像。Berman等人^[7]依据有雾图像的RGB空间中颜色簇分布规律,提出基于非局部先验的图像去雾算法。近年来,基于深度学习的去雾方法得到了广泛关注,主要包括基于参数估计的去雾方法和基于端到端的去雾方法。

2.1 基于参数估计的去雾模型

基于参数估计的去雾模型利用CNNs估计传输图和大气光值等大气散射模型的中间参数,再根据大气散射模型得到清晰的图像。Cai等人^[16]提出了一种DehazeNet去雾网络,首次利用CNN对传输图进行估计,建立了有雾图像和传输图之间的映射关系,然后通过大气散射模型实现图像去雾。Li等人^[17]提出一体化图像去雾网络AOD-Net,该网络基于转换变形后的大气散射模型设计,无需对传输图和全局大气光值进行估计,可以通过CNN直接实现图像去雾。Zhang等人^[18]设计了一种基于稠密连接的金塔网络用于估计传输图和全局大气光值等参数,接着利用大气散射模型完成图像去雾。Li等人^[19]提出了一种针对单图像去雾的感知渐进网络LAP-Net,该网络以不同雾霾级别为监督逐级

学习传输图,同时利用稠密池化块学习全局大气光值,促使网络能够处理雾霾浓度较厚的图像. Ren 等人^[20]设计了一种基于整体边缘多尺度 CNN 的去雾模型,该模型利用粗细尺度网络学习有雾图像和传输图直接的映射关系,实现单图去雾. 为进一步提升性能,研究者们采用深度学习对多类参数进行先验估计,如暗通道先验、亮通道先验^[21]和颜色衰减先验等. 然而,这些方法所得到的估计值与真实值之间依旧存在着一定的差异,导致最终图像去雾的效果不够理想.

2.2 基于端到端的去雾模型

基于端到端的去雾方法通过数据驱动的方式直接建立有雾图像与清晰图像之间的映射关系,无需估计任何中间参数. 基于端到端的去雾方法又可细分为3类:基于生成对抗网络的去雾方法、基于 Transformer 的去雾方法和基于 U-Net 的去雾方法.

随着生成对抗网络在图像生成领域的广泛应用,其强大的生成能力也引起了图像去雾领域研究者的关注. 基于生成对抗网络的去雾方法通过生成器和判别器的对抗训练,生成器被迫学习到有雾图像与清晰图像之间更本质的映射关系,从而生成高质量的去雾结果. Zhu 等人^[22]提出了 CycleGAN 网络,该模型能够利用非配对数据集实现图像转换. Engin 等人^[23]提出了 Cycle-Dehaze 单幅图像去雾网络,并增加循环感知损失提升去雾图像质量. 但志平等^[24]提出了基于双判别器异构 CycleGAN 框架下多阶通道注意力校准的图像去雾方法,通过异构归一化的生成器和约束生成器局部视野的方式,提升算法的收敛效果和增加局部区域关注. 然而,生成对抗网络的训练过程往往面临不稳定和不收敛的问题,这些问题可能导致生成的去雾图像质量不稳定或难以达到预期效果.

基于 Transformer 的去雾方法通过自注意力机制建立图像所有区域的全局依赖关系,最后通过一个解码器将处理后的特征序列重新组合成去雾图像. Liu 等人^[25]提出了基于注意力机制的多尺度去雾神经网络 GridDehazeNet,旨在通过结合注意力机制和多尺度特征融合提升去雾效果. Qin 等人^[26]设计了一种端到端的特征融合注意力网络 (Feature Fusion Attention Network, FFA-Net),该网络将通道注意力和像素注意力机制相融合,将其引入去雾网络以处理不同类型的信息. 高继蕊等人^[27]设计了一种双注意力引导的细节和结构信息融合去雾网络,该网络通过空间-通道双注意力联合模块联合空间和通道两个维度的注意力进行特征提取,实现雾图像中细节和结构信息的增强和融合. Lu 等人^[28]提出了 MixDehazeNet 网络,该网络将多尺度大卷积核与并行注意力机制相结合并应用到神经网络中,该方法提升了神经网络对全局上下文信息的感知能

力. 然而,自注意力机制的复杂度与图像序列长度的平方成正比,导致对计算资源和内存的需求非常高. 并且,Transformer 模型参数量大,需要大量数据才能充分训练,但有雾数据集往往数据量有限.

U-Net 网络最初常用于医学图像分割领域,后来凭借其轻巧的模型架构和出色的图像细节提取能力被引入图像去雾领域. 基于 U-Net 的去雾方法利用编码器提取图像的多尺度上下文特征,利用解码器恢复图像分辨率,利用跳跃连接将编码器每层特征图与解码器对应层特征图进行通道拼接,使得图像重建过程能直接利用编码器保留的边缘信息和纹理信息,避免了细节信息的丢失. 为了更好地达到去雾效果,研究人员在 U-Net 网络中嵌入不同的特征提取方式以提升图像特征的表达力. Dong 等人^[29]提出了一种稠密特征融合的多尺度增强图像去雾网络 MSBDN,模型利用增强策略和反射技术分别在 U-Net 的解码和编码层进行多尺度特征交互融合,MSBDN 包含密集特征融合模块,该模块能够矫正高分辨率特征中的空间信息,从而恢复图像的空间细节. Luo 等人^[30]提出了 LKD-Net 模型,通过引入大核卷积去雾模块 (Large Kernel convolution Dehaze Block, LKDB) 实现去雾模型的显著优化. 这一模块巧妙地将大核卷积分解为小核卷积与扩展卷积,既避免了额外的计算负担又提升了模型性能. Chen 等人^[2]提出了基于细节增强卷积和内容引导注意力的单图像去雾网络 DEA-Net (Detail-enhanced attention network). DEA-net 在 U-Net 网络的编码层和解码层中引入了差分卷积,进一步提取细节特征,还设计了一种新的注意力机制内容引导注意力 (Content Guided Attention, CGA) 来提取全局特征. 孙航等人^[31]基于 U-Net 结构提出了一种跨层注意力特征交互和多通道注意力的图像去雾网络,该网络一方面利用跨层特征聚合减少去雾网络重构清晰图像过程中的特征稀释,另一方面通过多尺度通道注意力机制实现对图像上下文信息的利用. 这些方法多关注编码层的图像特征表征,以及编码层与解码层的特征融合,忽略了解码层不同尺度特征间语义信息的传递效率. 此外,多数端到端的去雾方法从空间域的角度提取全局信息来弥补 U-Net 网络提取全局信息的局限性,未考虑频域特征对去雾效果的影响. 基于此,本文提出了基于双域特征交互和全局注意力上采样的图像去雾网络.

3 本文方法

3.1 基于双域特征交互和局部相关性上采样的图像去雾网络结构

本文提出的去雾网络 DFLU-Net 的整体框架如图 1 所示,图像采用编码-解码的 U-Net 结构,由编码层、中

间特征转换层和解码层 3 部分组成. 其中, 编码层包含两层由 1 个 3×3 卷积和 N 个 SFFIM 串联构成的特征提取模块, 中间特征转换层由 1 个 3×3 卷积和 Q 个细节特

征增强模块 (Detail Enhanced Attention Block, DEAB)^[2] 组成, 解码层则包含两层由 1 个 LCUM、 N 个 SFFIM 和 1 个 3×3 卷积串联构成的特征上采样模块.

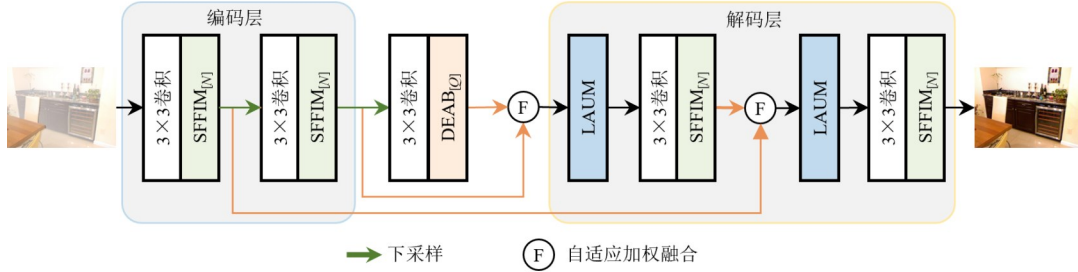


图1 网络架构图

网络以单幅有雾图像 $I \in \mathbb{R}^{3 \times H \times W}$ 作为输入, 首先在编码层通过 3×3 卷积得到尺寸为 $C \times H \times W$ 的特征图, 其中 C, H 和 W 分别为特征图的深度、高度和宽度. 然后, 利用 SFFIM 的两个分支 DPFFM 和 FFEM 分别提取图像的空间域特征和频域特征, 并将两者进行交互融合得到 $F_1 \in \mathbb{R}^{2C \times H/2 \times W/2}$. 将 F_1 重复上述过程, 得到编码层的输出特征 $F_2 \in \mathbb{R}^{4C \times H/4 \times W/4}$; 接着利用中间特征转换层的 DEAB 针对 F_2 进行细节特征增强, 得到中间特征转换层的输出特征 $F_m \in \mathbb{R}^{4C \times H/4 \times W/4}$. DEAB 由细节增强卷积 (Detail-Enhanced Convolution, DEConv) 和 CGA 两部分组成. 其中, DEConv 采用普通卷积并行的方式将图像特征的传统局部描述编码至卷积层, 增强其表征能力和泛化性能. CGA 则是以由粗到精 (coarse-to-fine) 的方式为输入特征的每个通道生成特定的空间重要性图, 并依次将通道注意力和空间注意力进行权重自适应融合, 实现有效的信息交互和梯度流动; 最后, 在解码层利用 SFFIM 和 3×3 卷积的串联进行相对于编码层的反向操作, 以恢复图像尺寸至 $3 \times H \times W$. 在此过程中, 为实现解码层上采样过程中语义信息自下而上的传递, 将 LCUM 嵌入至 SFFIM 之前, 确保输入 SFFIM 的特征图包含丰富的语义信息. 至此, 模型将输出清晰的无雾图像 $J_{\text{out}} \in \mathbb{R}^{3 \times H \times W}$. 值得注意的是, 为减少编码层提取特征过程中的特征稀释, 本文利用 CGA 模块将编码层提取的特征传递至相应的解码层.

3.2 空间域-频域特征交互模块

多数基于 U-Net 结构的去雾网络主要聚焦于从空间域提取图像的全局特征, 鲜少考虑频域中图像特征的特点. 通过对比有雾图像和无雾图像的频域特征可以发现, 两者在幅值谱上存在巨大差异, 而相位谱只受到轻微的影响. 基于此, 本文提出了空间域-频域交互模块, 其核心目标就是实现双域互补信息的高效融合, 结构如图 2 所示.

该模块针对双域特征的交互融合步骤如下: (1) 对输入 SFFIM 的特征进行归一化处理, 得到归一化后的

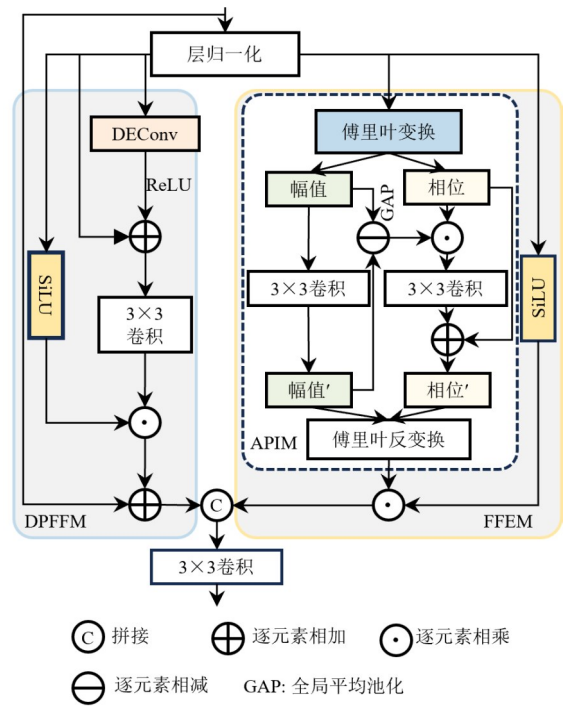


图2 空间域-频域特征交互模块

特征表达 F_{LN} ; (2) 分别利用 DPFFM 分支和 FFEM 分支提取图像的空间域特征和频域特征. 具体来说, 在空间域特征提取部分通过 DPFFM 提取局部细节特征, 用于恢复图像复杂纹理和边缘结构, 提升模型对物体轮廓和方向性特征的敏感度, 频域特征提取部分的核心是首先使用傅里叶变换提取图像的幅值谱和相位谱, 幅值谱包含图像对比度、纹理等细节信息, 相位谱包含图像结构信息, 再利用幅值谱补充相位谱得到更准确的全局结构特征, 接着利用傅里叶反变换将补偿后的频域特征变换回空间域; (3) 利用特征拼接和 3×3 卷积对 DPFFM 和 FFEM 输出的特征进行交互融合, 得到 SFFIM 的输出. SFFIM 通过双域特征交互, 实现对空间域边缘细节与方向敏感度, 以及对频域物体位置与排列方式的充分利用, 并提升特征表达的泛化性和鲁棒性.

3.2.1 双路特征融合模块

DPFFM 主要用于图像空间域特征的提取. 为实现对图像空间域全局特征和局部细节特征的同步提取, DPFFM 采用两条路径的结构设计. 其中, 第一条通路直接利用 SFFIM 的输入 F_m ; 第二条通路由两部分逐像素相加得到输出, 第一部分利用激活函数 SiLU 对归一化的特征 F_{LN} 进行非线性处理, 得到空间域的全局特征表达 F_{SiLU} , 第二部分则是利用以 DEConv^[2] 为核心的残差结构提取图像的空间细节特征 F_{DEC} , 通过 3×3 卷积后, 与第一部分输出 F_{SiLU} 逐像素相乘, 得到融合特征 F_{fused} . 最后, 将 F_{fused} 与原始输入特征 F_m 逐像素相加, 最终输出空间域分支特征 F_s . 该过程的数学表达如下:

$$F_{fused} = \text{Conv}(\text{DEConv}(F_{LN}) + \text{ReLU}(F_{LN}))_{3 \times 3} \odot F_{SiLU} \quad (1)$$

$$F_s = F_m \oplus F_{SiLU} \quad (2)$$

其中, m 表示编码层的第 m 次下采样编码, 其取值为 1 或 2; \odot 表示逐元素相乘; \oplus 表示逐元素相加. DPFFM 模块中的第一条通路旨在直接传递 SFFIM 模块的输入特征, 缓解深度网络中的梯度消失问题, 确保原始特征信息能够向前传播; 第二条通路则专注学习细节特征, 一部分通过以 DEConv 为核心的残差结构提取精细的空间细节和纹理信息, 再与 F_{SiLU} 逐像素相乘, 利用 SiLU 的门控特性自适应地筛选特征; 最后逐像素相加融合两条通路, 实现空间域中全局信息与细节信息的协同优化.

在 DPFFM 中, 提取图像空间域细节特征的残差网络的核心是 DEConv. 传统稠密连接块内部由于连续使用相同尺度的卷积核导致图像细节的特征多样性不足, 进而影响特征提取的全面性. DEConv 则是将差分卷积组和 3×3 卷积进行稠密连接, 这样在不显著增加模型参数数量和计算复杂度的情况下, 有效提升了针对细节特征的提取能力, 详细结构如图 3 所示. 其中, 差分卷积组包括了中心差分卷积 (Central Difference Convolution, CDC)、基于角度差异的像素差分卷积 (Angular Difference Convolution, ADC)、垂直差分卷积 (Vertical Difference Convolution, VDC) 以及水平差分卷积 (Horizontal Difference Convolution, HDC).

CDC 通过计算中心像素与邻域像素的差值提取特

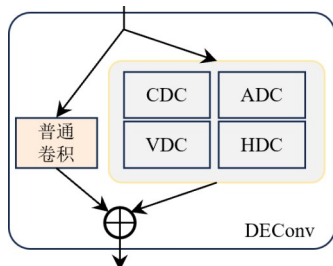


图3 DEConv 结构图

征, 重点关注中心像素与周围像素的变化, 从而有效捕捉图像的局部变化、纹理及边缘信息. 图像纹理通常具有方向性规律, 普通卷积难以捕捉此类特征, 而 ADC 能够根据不同方向的变化对特征进行处理以提取更具方向性的特征. VDC 和 HDC 则分别专注于提取图像垂直和水平方向的边缘及纹理变化. 对于图像去雾任务, 强度级语义信息对提取图像的纹理特征和轮廓至关重要. 因此, 将普通卷积与上述 4 种差分卷积结合, 能够显著提升模型的鲁棒性和特征提取能力. 然而, 多个卷积的并行部署不可避免地增加了模型的参数数量和推理时间. 为了使 DEConv 达到与普通卷积相近的计算成本和推理时间, 本文利用同一卷积层的可加性 (未经过激活函数的非线性处理前), 将 5 个卷积核进行逐元素相加后, 再与输入特征进行卷积, 数学表达式如式 (3) 所示:

$$F_{\text{DEConv}} = \sum_{i=1}^5 F_{LN} * K_i = F_{LN} * \left(\sum_{i=1}^5 K_i \right) = F_{LN} * K_{\text{DEConv}} \quad (3)$$

其中, F_{DEConv} 表示细节增强卷积; $K_{i=1,5}$ 分别表示 VC、CDC、ADC、VDC 和 HDC 的卷积核; $*$ 表示卷积操作; K_{DEConv} 表示叠加转换后的卷积核.

3.2.2 频域特征增强模块

FFEM 主要用于图像频域特征的提取, 同样采用两条路径的结构设计. 其中, 第一条路径是将特征 F_{LN} 通过激活函数 SiLU 得到特征 F_{SiLU} ; 第二条通路则是通过频谱幅相整合模块 (Amplitude-Phase Integration Module, APIM) 得到输出特征 F_f . 然后, 将两条路径的输出特征 F_{SiLU} 和 F_f 逐像素相乘, 得到频域的输出特征 F_j , 具体公式如下:

$$F_j = \text{SiLU}(F_{LN}) \odot \text{APIM}(F_{LN}) \quad (4)$$

不同空间位置对频域特征的依赖程度不同, 两条通路逐像素相乘, F_{SiLU} 通过学习到的空间权重对频域特征进行局部调制.

由前文可知, 雾霾引起的图像退化主要体现在幅值谱中, 有雾图像与无雾图像的相位谱差异较小. 因此, APIM 主要针对幅值谱进行补偿, 并利用补偿后的幅值谱求解对应的相位谱, 以便得到更准确的全局结构特征, 具体结构如图 2 所示. APIM 首先将归一化的空间域特征 F_{LN} 通过傅里叶变换转换为频域特征 Γ_{LN}^{FT} :

$$\Gamma_{LN}^{\text{FT}}(u, v) = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} F_{LN}(h, w) e^{-j2\pi \left(\frac{h}{H}u + \frac{w}{W}v \right)} \quad (5)$$

$$\Gamma_{LN}^{\text{FT}} = \text{Re}(\Gamma_{LN}^{\text{FT}}) + j\text{Im}(\Gamma_{LN}^{\text{FT}}) \quad (6)$$

其中, $\text{Re}(\Gamma_{LN}^{\text{FT}})$ 表示实部; $\text{Im}(\Gamma_{LN}^{\text{FT}})$ 表示虚部. 然后将实部和虚部转换为幅值和相位:

$$A(\Gamma_{LN}^{\text{FT}})(u, v) = \left[\text{Re}^2(\Gamma_{LN}^{\text{FT}})(u, v) + \text{Im}^2(\Gamma_{LN}^{\text{FT}})(u, v) \right]^{1/2} \quad (7)$$

$$P(\Gamma_{LN}^{FT})(u, v) = \arctan \left[\frac{\text{Im}(\Gamma_{LN}^{FT})(u, v)}{\text{Re}(\Gamma_{LN}^{FT})(u, v)} \right] \quad (8)$$

其中, $A(\Gamma_{LN}^{FT})$ 表示幅值谱; $P(\Gamma_{LN}^{FT})$ 表示相位谱. 由于有雾图像频域信息中的幅值谱相较于无雾图像存在严重退化, 因此需要对其进行补偿. 然而, 为了得到图像更准确的频域全局结构特征, 仍需要准确求解补偿后的幅值谱对应的相位谱. 具体地, 首先采用 1×1 卷积恢复对 $A(\Gamma_{LN}^{FT})$ 进行处理, 得到恢复后的幅值谱 $A'(\Gamma_{LN}^{FT})$:

$$A'(\Gamma_{LN}^{FT})(u, v) = \text{Conv}(A(\Gamma_{LN}^{FT})(u, v))_{1 \times 1} \quad (9)$$

接着, 对恢复后的幅值谱 $A'(\Gamma_{LN}^{FT})$ 和原始幅值谱 $A(\Gamma_{LN}^{FT})$ 进行逐元素求差, 得到幅值谱残差 $A_{\text{res}}(\Gamma_{LN}^{FT})$:

$$A_{\text{res}}(\Gamma_{LN}^{FT})(u, v) = A'(\Gamma_{LN}^{FT})(u, v) - A(\Gamma_{LN}^{FT})(u, v) \quad (10)$$

然后, 为了提取 $A'(\Gamma_{LN}^{FT})$ 和 $A(\Gamma_{LN}^{FT})$ 显著差异, 利用 $A_{\text{res}}(x)$ 注意力机制对幅值残差进行处理, 得到其注意力图 $A_{\text{atten}}(x)$, 并依据此求解 $A'(\Gamma_{LN}^{FT})$ 对应的相位谱 $P'(\Gamma_{LN}^{FT})$:

$$A_{\text{atten}}(\Gamma_{LN}^{FT})(u, v) = \text{softmax}[\text{GAP}(A_{\text{res}}(\Gamma_{LN}^{FT})(u, v))] \quad (11)$$

$$P(\Gamma_{LN}^{FT})(u, v) = P'(\Gamma_{LN}^{FT})(u, v) + \text{Conv} \left(\left[A_{\text{atten}}(\Gamma_{LN}^{FT})(u, v) \odot P'(\Gamma_{LN}^{FT})(u, v) \right] \right)_{1 \times 1} \quad (12)$$

其中, GAP 表示全局平均池化. 当幅值谱和相位谱恢复后, 再将其转为对应的实部和虚部表达:

$$\text{Re}(\Gamma_{LN}^{FT})(u, v) = A'(\Gamma_{LN}^{FT})(u, v) \cdot \cos[P'(\Gamma_{LN}^{FT})(u, v)] \quad (13)$$

$$\text{Im}(\Gamma_{LN}^{FT})(u, v) = A'(\Gamma_{LN}^{FT})(u, v) \cdot \sin[P'(\Gamma_{LN}^{FT})(u, v)] \quad (14)$$

最后, 通过傅里叶反变换将其频域特征转换回空间域, 得到其特征表示 F_{LN}^{-1} :

$$F_{LN}^{-1}(\Gamma_{LN}^{FT}) = \text{Re}(\Gamma_{LN}^{FT}) + j\text{Im}(\Gamma_{LN}^{FT}) \quad (15)$$

3.3 局部相关性上采样模块

编码层中上采样的主要目的是在还原特征尺寸的同时, 将深层特征的语义信息自下而上地传递至浅层特征中. 在传统的 U-Net 结构中, 通常采用最近邻插值法对特征图进行上采样, 然后通过一个普通卷积层对特征进行细化, 并调整其通道数. 然而, 该方法未考虑相邻像素之间的内部相关性, 进而可能导致插值后的图像出现像素值不连续、锯齿状图案或马赛克状外观, 最终造成特征的空间信息丢失, 影响深层语义信息向浅层的有效传递. 为解决上述上采样方法的局限性, 本文设计了局部相关性上采样模块. 由于插值是还原特征图尺度的必要步骤, 因此需要在上采样前对特征表示进行增强. 注意力机制可以有效捕捉单一特征图中局部信息的内在关联性, 从而增强其空间信息并提取

丰富的语义信息. 基于此, LCUM 首先通过注意力机制增强特征表示, 随后采用双线性插值法将信息传递到浅层特征中, 具体结构如图 4 所示.

首先, 利用 3 个 1×1 卷积将 LCUM 的输入特征 $F_{\text{in}}^{\text{GAU}} \in \mathbb{R}^{C_{\text{in}} \times H_{\text{in}} \times W_{\text{in}}}$ 映射为 3 个特征, 分别表述注意力机制中的查询 Q 、键 K 和值 V , 其中 C_{in} 、 H_{in} 和 W_{in} 分别为输入 LCUM 特征的深度、高度和宽度, 即

$$Q = \text{Conv}(W^Q F_{\text{in}}^{\text{GAU}})_{1 \times 1} \quad (16)$$

$$K = \text{Conv}(W^K F_{\text{in}}^{\text{GAU}})_{1 \times 1} \quad (17)$$

$$V = \text{Conv}(W^V F_{\text{in}}^{\text{GAU}})_{1 \times 1} \quad (18)$$

其中, $Q, K, V \in \mathbb{R}^{C_{\text{in}}/2 \times H_{\text{in}} \times W_{\text{in}}}$; W^Q, W^K, W^V 均是可学习的映射权重. 然后, 对 Q, K, V 进行矩阵变换, 分别得到 $Q_{\text{reshape}} \in \mathbb{R}^{(H_{\text{in}} \times W_{\text{in}}) \times (C_{\text{in}}/2)}$, $K_{\text{reshape}} \in \mathbb{R}^{(C_{\text{in}}/2) \times (H_{\text{in}} \times W_{\text{in}})}$, $V_{\text{reshape}} \in \mathbb{R}^{(C_{\text{in}}/2) \times (H_{\text{in}} \times W_{\text{in}})}$, 对 Q_{reshape} 和 K_{reshape} 执行点乘运算, 得到聚合权重 W_{agg} , 其代表了特征中每个位置的重要程度.

$$W_{\text{agg}} = Q_{\text{reshape}} \otimes K_{\text{reshape}} \quad (19)$$

然后, 利用 Softmax 函数对聚合权重 W_{agg} 进行归一化, 并将其与特征 V_{reshape} 进行点乘运算, 再进行矩阵变换得到特征 Y . 由于注意力机制的作用, 特征 Y 包含了丰富的语义信息.

最后, 对特征 Y 进行 1×1 卷积运算扩展其通道数, 与输入特征 $F_{\text{in}}^{\text{GAU}}$ 逐像素相加, 并采用双线性插值法得到最终的输出特征 $F_{\text{out}}^{\text{GAU}}$, 计算公式如下:

$$Y = V_{\text{reshape}} \otimes \text{Softmax}(W_{\text{agg}}) \quad (20)$$

$$F_{\text{out}}^{\text{GAU}} = f_{\text{bilinear}}((k_1 * Y) \oplus F_i) \quad (21)$$

其中, $f_{\text{bilinear}}(\cdot)$ 表示双线性插值操作.

由于注意力机制可以捕获更多深层特征的全局信息, 因此 LCUM 可以在补偿空间信息丢失的同时, 有效地将深层特征抽象信息和补偿的信息传递给浅层特征. 对于通道维数高、分辨率大的特征而言, 矩阵乘法运算会增加计算成本, LCUM 则在基于注意力机制的特征增强过程开始和结束时利用 1×1 卷积分别进行特征通道压缩和扩展.

3.4 损失函数

在网络模型学习过程中, 损失函数用以表示系统真值与模型预测值之间的差异. 本文采用 L_1 损失和对比损失联合监督网络模型训练过程.

其中, L_1 损失用于最大程度地减小模型预测输出的无雾图 J 与相应真实图像 GT 之间的像素差异, 表达式如下:

$$L_1 = \|J - \text{GT}\|_1 \quad (22)$$

其中, J, GT 分别表示本文算法预测输出的无雾图像和相应的真实的无雾图像; $\|\cdot\|_1$ 表示 L_1 范数.

此外, 对比损失用于在特征提取过程中对预测值

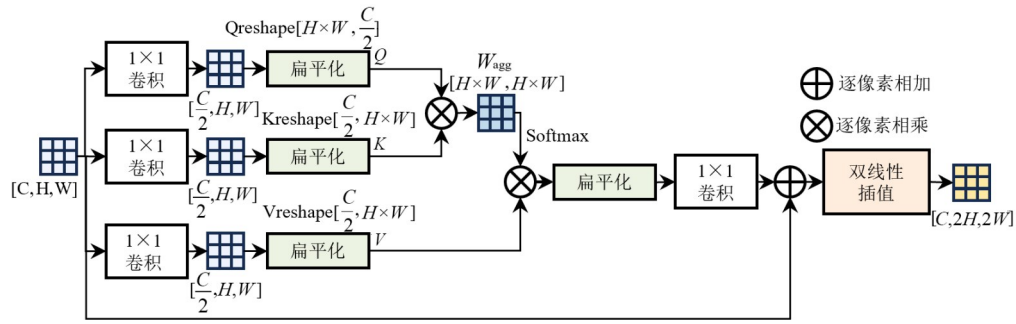


图4 局部相关性上采样模块

和真实值的匹配程度进行约束. 具体而言, 在各个特征层级中, 分别计算模型输出的预测图像与真实值之间的欧式距离 L_{J-GT} , 以及有雾图像与真实值之间的欧式距离 L_{J-X} , L_{J-GT} 和 L_{J-X} 之间的比值即为对比损失. 对比损失实现了对低级纹理到高级语义各个层级的有力约束, 使得生成的图像更好地匹配真实无雾图像的纹理、细节和颜色. 表达式如下:

$$L_{J-GT(i)} = \|\varphi_i(J) - \varphi_i(GT)\|_1 \quad (23)$$

$$L_{J-X(i)} = \|\varphi_i(J) - \varphi_i(X)\|_1 \quad (24)$$

$$L_C = \sum_{i=1}^N \left(w_i \times \frac{L_{J-GT(i)}}{L_{J-X(i)} + \varepsilon} \right) \quad (25)$$

其中, X 表示本文模型真实的有雾图像; $\varphi_i(\cdot)$ 表示获取 VGG16 预训练模型第 i 层的特征图; N 表示在对比损失中使用 VGG16 预训练模型特征层的数量; w_i 表示每层特征对应的对比损失权重, ε 设置为 10^{-7} , 用以防止分母为零.

为了更好地训练本文提出的模型, 总损失函数如式(26)所示:

$$L_{\text{total}} = \lambda_1 L_1 + \lambda_2 L_C \quad (26)$$

其中, λ_1, λ_2 为每个损失函数的超参数.

4 实验分析

4.1 实验配置与数据集

本文实验的训练和测试过程均基于 PyTorch 深度学习框架完成. 实验硬件环境为: 英特尔 i7-13700K@3.4 GHz CPU 和 NVIDIA RTX 4090 GPU, 显存为 24 GB. 每张输入图像尺寸随机裁剪为 256×256 , 采用图像尺寸随机缩放和图像方向随机旋转等方式进行数据增强. 本文使用 Adam 优化器^[32]进行模型训练优化, 并将其中参数 $\beta_1, \beta_2, \varepsilon$ 分别设置为 0.9、0.999、 1×10^{-8} . 批处理大小为 8, 模型训练初始学习率和权重衰减分别为 1×10^{-4} 和 1×10^{-7} . 模型模块数量 N 和 Q 分别设置为 4 和 8.

模型的评估在 RESIDE 合成有雾数据集^[33]和 Haze4K 室外合成有雾数据集^[34]上进行. 其中, 在

RESIDE 数据集中, 选择室内训练集 (Indoor Training Set, ITS) 和室外训练集 (Outdoor Training Set, OTS) 作为训练集, SOTS-Indoor 和 SOTS-Outdoor 作为测试集, 分别用于评估在 ITS 和 OTS 上训练的模型. ITS 包含 1 399 张室内无雾图像, 对于每一张无雾图像, 根据物理散射模型生成 10 张模拟有雾图像. 由于 OTS 数据集和 SOTS-Outdoor 数据集有重合, 在剔除 OTS 里的重合图像后, 剩余大约 296 000 张的图像用于训练过程. SOTS-Indoor 和 SOTS-Outdoor 分别包含 500 张室内和 500 张室外测试图像. 此外, 本文也使用 Haze4K 数据集进行实验, 其中包含 3 000 张综合训练图像和 1 000 张合成测试图像.

4.2 评价指标

为评估模型性能, 本文使用峰值信噪比 (Peak Signal Noise Ratio, PSNR) 和结构相似性指数 (Structural Similarity Index Measure, SSIM) 为定量评估评价指标. 为了公平比较, 本文基于 RGB 彩色图像计算指标, 且不裁剪像素.

PSNR 通常用来评价经过处理之后的图像与参考图像相比质量的好坏, 其值越高, 表明处理后的图像失真越小. 其计算原理是基于对应像素点间的误差, 计算方式如式(27)、式(28)所示:

$$\text{MSE} = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H (X(i,j) - Y(i,j))^2 \quad (27)$$

$$\text{PSNR} = 10 \times \lg \left(\frac{(2^n - 1)^2}{\text{MSE}} \right) \quad (28)$$

其中, MSE 表示处理后的图像 X 与参考图像 Y 之间的均方误差; W 和 H 分别代表图像的宽度和高度; n 表示像素的比特数, 一般取 8, 即代表像素灰级为 256.

SSIM 通常用于测量计算机视觉的图像质量, 用于去雾性能评估. 自然图像具有极高的结构性, 表现在图像的像素间存在很强的相关性, 尤其是在空间相似的情况下. 这些相关性在视觉场景中携带与物体结构有关的重要信息. 相比于未涉及到像素相关性的基于误差敏感度的质量评估方法使用线性变换来分解图像信号, SSIM 利用相关性特征能更加直接地比较失真图像和参考图像的结构. SSIM 主要从亮度、对比度和结构 3 个

方面比较预测值和真实值,最终两者的相似度计算如式(29)所示:

$$\text{SSIM}(x, y) = [l(x, y)]^\alpha [c(x, y)]^\beta [s(x, y)]^\gamma \quad (29)$$

其中, $l(x, y)$ 表示亮度关系; $c(x, y)$ 表示对比度关系; $s(x, y)$ 表示结构关系; α, β, γ 代表调整各部分数值的参数.

4.3 定量和定性结果评析

为验证本文方法的有效性,将本文方法 DDGA-Net 与 13 种先进的去雾方法 DCP^[6]、DehazeNet^[16]、AOD-Net^[17]、SGID-PFF^[35]、FFA-Net^[26]、MSBDN^[29]、FSDGN^[5]、DehazeFormer^[36]、LKD-Net^[30]、CycleGAN^[22]、DFR-Net^[37]、DEA-Net^[2]以及 DBDE-Net^[38]进行性能对比实验. 分别对 PSNR、SSIM、模型参数量和运算量等方面进行评估,表 1 展示了本文方法和其他先进去雾方法在 SOTS-Indoor、SOTS-Outdoor 和 Haze4K 数据集上的定量评估结果.

实验结果表明,DFLU-Net 在实现网络架构轻量化设计的同时,也达到了较好的去雾性能. 在模型去雾性

能方面,本文方法在 SOTS-Indoor、SOTS-Outdoor 和 Haze4K 三个数据上的 PSNR 和 SSIM 都达到了最优,分别为 41.46 dB 和 0.994 3、37.73 dB 和 0.993 6、34.72 dB 和 0.993 0. DFLU-Net 与基于参数估计的去雾模型 SGID-PFF 对比,在 SOTS-Indoor 和 SOTS-Outdoor 数据集上的去雾性能大幅领先;与基于生成对抗网络的去雾模型 CycleGAN 对比,在 SOTS-Outdoor 数据集上 PSNR 和 SSIM 指标上分别提升了 10.56 dB 和 0.031 6;与基于 Transformer 的去雾模型 FFA-Net 对比,在 SOTS-Indoor 数据集上 PSNR 和 SSIM 指标上分别提升了 5.07 dB 和 0.005 7,在 SOTS-Outdoor 数据集上 PSNR 和 SSIM 指标上分别提升了 4.16 dB 和 0.009 6,在 Haze4K 数据集上 PSNR 和 SSIM 指标上分别提升了 7.75 dB 和 0.043;与基于 U-Net 的最新方法 DBDE-Net 对比,在 SOTS-Indoor 数据集上 PSNR 和 SSIM 指标上分别提升了 1.77 dB 和 0.000 3,在 Haze4K 数据集上 PSNR 和 SSIM 指标上分别提升了 2.68 dB 和 0.009. 与不同去雾架构的横向对比表明,本文方法能够在轻巧的模型架构的基础上保证去雾性能,在性能与模型复杂度之间实现了良好的平衡.

表 1 三种数据集上定量比较 13 种模型的去雾性能

方法	SOTS-Indoor		SOTS-Outdoor		Haze4K		参数量($\times 10^6$)	运算量($\times 10^9$)
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM		
(TPAMI 10)DCP ^[6]	16.61	0.854 6	19.14	0.860 5	14.01	0.760 0	—	—
(TIP 16)DehazeNet ^[16]	19.82	0.820 9	27.75	0.926 9	19.12	0.840 0	0.008	0.540 9
(ICCV 17)AOD-Net ^[17]	20.51	0.816 2	24.14	0.919 8	17.15	0.830 0	0.001 8	0.114 6
(TIP 22)SGID-PFF ^[35]	38.52	0.991 3	30.20	0.975 4	—	—	13.870 0	152.8
(AAAI 20)FFA-Net ^[26]	36.39	0.988 6	33.57	0.984 0	26.97	0.950 0	4.456 0	287.5
(CVPR 20)MSBDN ^[29]	32.77	0.981 2	34.81	0.985 7	22.99	0.850 0	31.350 0	24.44
(ECCV 22)FSDGN ^[5]	38.63	0.990 3	—	—	—	—	2.730 0	19.6
(TIP 23)DehazeFormer ^[36]	—	—	34.95	0.984 0	30.29	0.985 0	2.514 0	—
(ICME '23)LKD-Net ^[30]	39.44	0.994 0	34.82	0.983 0	—	—	2.38	23.93
(电子学报 '23)CycleGAN ^[22]	—	—	27.17	0.962 0	—	—	—	—
(TMM '24)DFR-Net ^[37]	—	—	35.34	0.993 0	34.63	0.993 0	42.110 0	—
(TIP '24)DEA-Net ^[2]	40.20	0.993 4	36.03	0.989 1	33.19	0.987 0	3.653 0	32.23
(西北工业大学学报 '25) DBDE-Net ^[38]	39.69	0.994 0	—	—	32.04	0.984 0	—	—
(Ours) DFLU-Net	41.46	0.994 3	37.73	0.993 6	34.72	0.993 0	3.160 0	27.47

为了直观表征各算法的去雾性能,本文提出一种基于热力图的对比分析方法,即通过颜色梯度直观量化去雾效果差异,其可以有效反映各去雾方法在图像细节还原方面的性能差异. 具体而言,将有雾图像、模型预测的去雾图像,以及真实无雾图的 RGB 通道值分别叠加并归一化后,并计算真实无雾图与有雾图像、真实无雾图与去雾图像的像素绝对值差异,进而生成热力图. 热力图中,深色表示模型预测的去雾图像与真实无雾图之间的像素差异较大、还原度较低,浅色表示模型预测的去雾

图像与真实无雾图之间的像素差异较小、还原度较高.

图 5 展示了本文方法与现有方法在室内有雾测试集 SOTS-Indoor 上的视觉对比结果,色柱统一采用灰度色阶. 其中,有雾图像与真实无雾图的热力图差值区间 $[0, 0.7)$,表明有雾图像清晰度不足. 传统 DCP 模型输出的去雾结果与真实无雾图的热力图差值范围缩减,但去雾性能提升有限. SGID-PFF、FFA、DEA-Net 及本文方法的热力图色阶范围均缩减至 $[0, 0.14)$. 在相同度量标准下,首先比较本文方法、SGID-PFF 和 DEA-Net 的热

力图,在纹理复杂的区域,本文方法热力图颜色最浅,表明能够更好地保留和还原细节信息;然后比较本文

方法和FFA的热力图,在低对比度区域,本文方法能有效恢复图像真实的色彩和亮度.

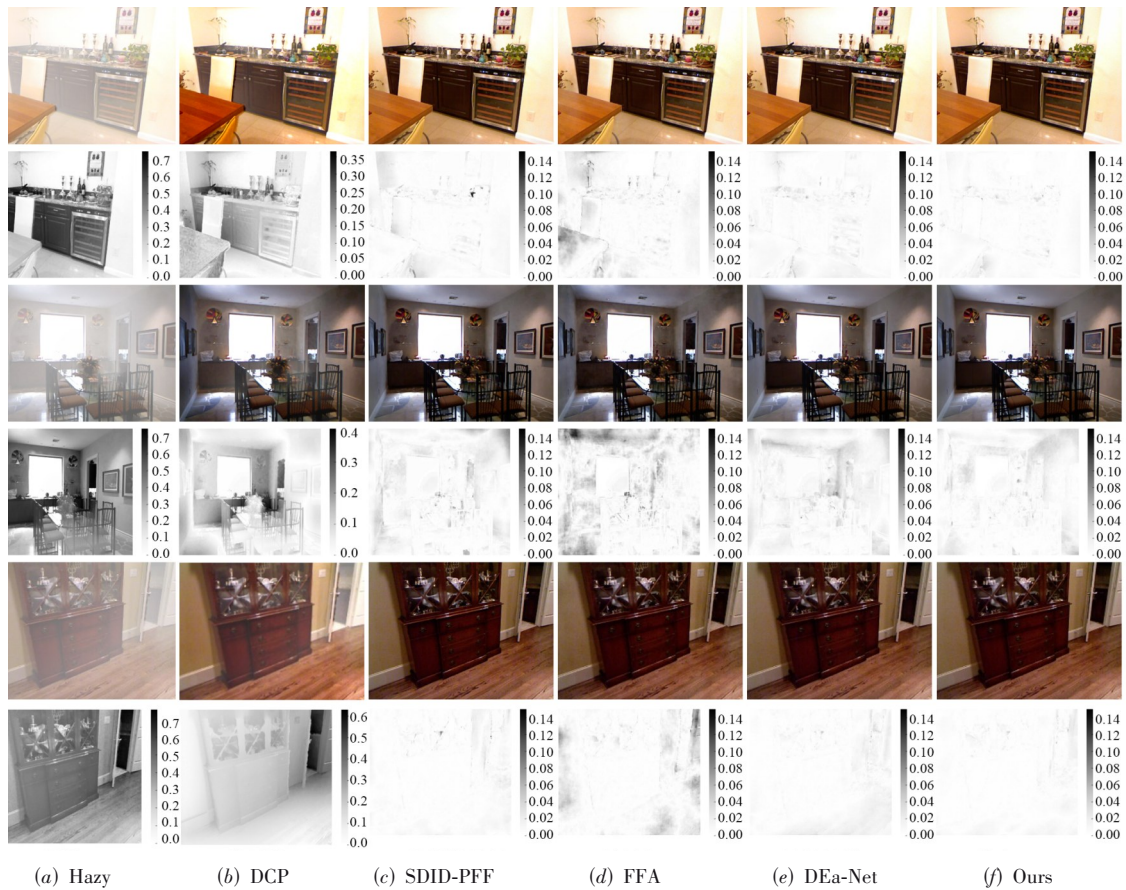


图5 SOTS-Indoor数据集可视化结果

图6展示了本文方法与现有方法在室外有雾测试集SOTS-Outdoor上的视觉对比结果.其中,有雾图像与真实无雾图的热力图差值分布区间较大,表明有雾图像清晰度不足.比较本文方法和传统DCP模型的去雾图像,传统DCP模型输出的去雾结果与真实无雾图的热力图差值区间有一定缩小,去雾能力有限,而本文方法热力图差值分布区间小,在天空区域避免了过曝和颜色偏差问题.SGID-PFF、FFA、DEa-Net及本文方法的热力图色阶范围均缩减至 $[0, 0.2)$.在相同度量标准下,首先比较本文方法和SGID-PFF的热力图,本文方法能够在低光照区域恢复更清晰的图像结构和整体外观;然后比较本文方法、FFA和DEa-Net的热力图,本文方法能够在处理远景区域时捕获更多细节信息,清晰地恢复了远处建筑物的窗户、砖块等.本文还在不同雾浓度的真实雾天图像上测试本文方法,并将结果与其他方法进行比较,视觉对比结果如图7所示.在浓雾的景观场景下,DCP存在天空色彩失真的问题,而SGID-PFF在整体去雾效果上优于DCP. FFA和DEa-Net在塔楼细节恢复上有明显不足,仍残存明显的雾气.相较于对比

方法,本文方法色彩更接近真实图像,还显著提升了塔楼下半部分轮廓与结构的恢复质量,展现出更优的细节恢复能力.在浓雾的交通场景下,DCP、FFA和DEa-Net均难以有效还原远景车辆,近景区域也仍存在雾气遮蔽.本文方法相较于SGID-PFF,在近景车辆的颜色还原和图像清晰度方面表现更为出色,同时对远景车辆的轮廓重建也更为准确.在中等雾浓度的交通场景下,SGID-PFF、FFA、DEa-Net的去雾效果相近,道路与远景区域均存在一定雾气残留.DCP虽然在增强道路标线色彩方面表现较好,但远景细节的可见度有所损失.本文方法则实现了更均衡的去雾效果,不仅还原了近景的交通道路线,远景的细节还原也更清晰.综上所述,本文方法在物体轮廓重建、边缘细节增强及色彩还原方面均优于对比方法,并且在去雾结果中残留雾气区域最少,展现出更强的综合去雾能力与视觉恢复效果.

4.4 消融实验

为了评估本文模型各模块的有效性,本节根据框架和提出的模块设计消融实验,其中包括3组消融实

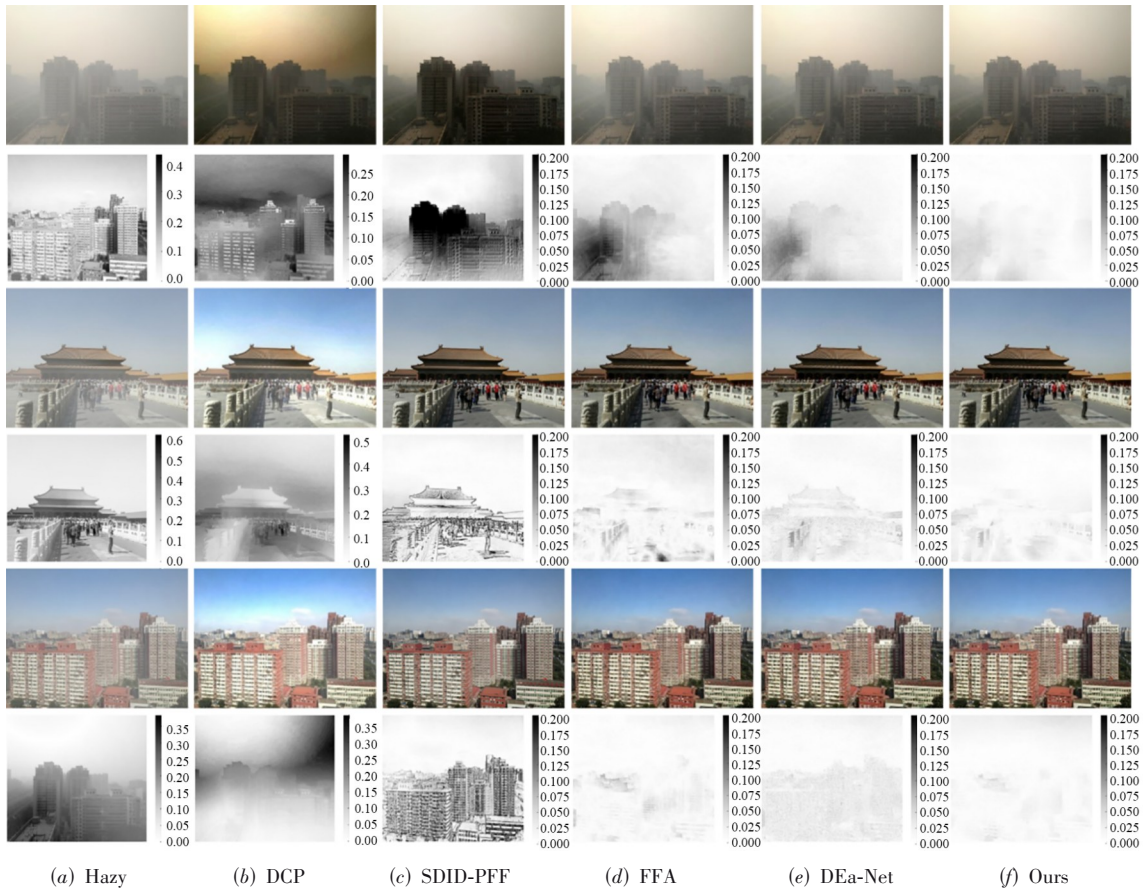


图6 SOTS-Outdoor数据集可视化结果

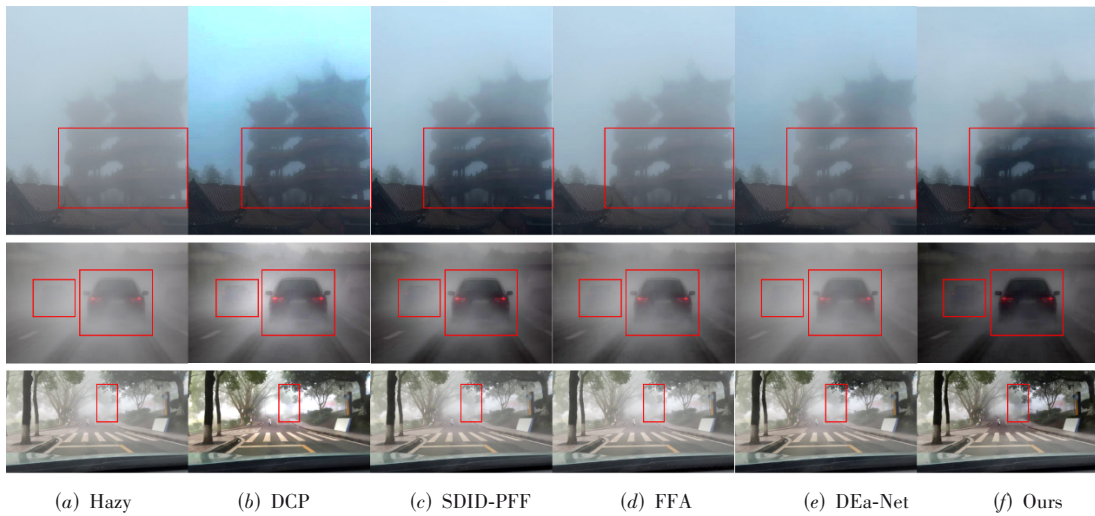


图7 不同雾浓度图像去雾可视化结果

验,分别验证各子模块(SFFIM、DEAB、LCUM)的有效性,空间域与频域特征融合的有效性和超参数设置的合理性.

4.4.1 验证各子模块有效性

为了验证本文方法中各子模块 SFFIM、DEAB、LCUM 的有效性,消融实验中保持 DFLU-Net 编码层、中

间特征转换层、解码层中的模块数量不变. 本文选择 ITS 数据集作为训练集,SOTS-Indoor 作为测试集. 其中包括 8 个实验.

(1) Baseline: 将 DFLU-Net 编码层、中间特征转换层、解码层的相应模块替换为残差块,一共五层. 使用最近邻插值法进行上采样,编码层和解码层通过自适

应加权融合跳跃连接。

(2)Baseline+SFFIM:在基线上替换编码层和解码层部分的残差块为SFFIM。

(3)Baseline+DEAB:在基线上替换中间层的8个残差块为8个DEAB。

(4)Baseline+LCUM:在基线上将最近邻插值上采样替换为LCUM。

(5)Baseline+SFFIM+DEAB:在(2)的基础上替换中间层的8个残差块为8个DEAB。

(6)Baseline+SFFIM+LCUM:在(2)的基础上将最近邻插值上采样替换为LCUM。

(7)Baseline+DEAB+LCUM:在(3)的基础上将最近邻插值上采样替换为LCUM。

(8)本文方法:在(7)的基础上替换编码器和解码器部分的残差块为SFFIM。

如表2所示,Baseline在PSNR和SSIM上为31.03 dB、0.961 2,通过替换编码器和解码器部分的残差块为SFFIM,PSNR达到35.85 dB,SSIM达到0.985 6。相较于基准模型PSNR提升了4.82 dB,SSIM提升了0.024 4,运算量下降了 4.09×10^9 ,参数量下降了 0.41×10^6 。验证了跨域特征交互有效融合了全局上下文信息和局部细节信息,提升了去雾精度。基于Baseline替换中间层的8个残差块为8个DEAB,PSNR达到36.67 dB,SSIM达到0.987 2。相较于基准模型PSNR提升了5.64 dB,SSIM提升了0.026 0,运算量下降了 4.45×10^9 ,参数量下降了 1.04×10^6 。验证了DEAB能够高效地建模雾度分布差异并优化特征传递,有效提取图像级别和通道级别的雾气特征。将Baseline中的最近邻插值上采样替换为LCUM,PSNR达到32.82 dB,SSIM达到0.963 1。相较于基准模型PSNR提升了1.79 dB,SSIM提升了0.001 9,运算量下降了 2.07×10^9 ,参数量下降了 0.04×10^6 。验证了LCUM在特征上采样过程中通过注意力机制补偿特征信息的有效性。此外,与基准模型相比,集成3个模块中的任何两个都会带来更好的性能体现。在Baseline基础上同时集成SFFIM、DEAB、LCUM,模型能在保持较低运算量和参数量的同时性能达到最优,PSNR达到41.46 dB,SSIM达到0.994 3,表明各模块在增强特征表达与优化计算效率之间达到了平衡。

4.4.2 SFFIM的消融实验

本文为了验证融合空间域特征和频域特征对图像去雾的有效性,对空间域DPFFM、频域FFEM和双域融合SFFIM进行了评估。本文选择ITS作为训练集,SOTS-Indoor作为测试集进行消融实验。其中包括4个实验。

(1)残差块:保持DFLU-Net编码器、特征转换层、解码器中的模块不变,跳跃连接方式不变,替换SFFIM

表2 在SOTS-Indoor数据集上验证各子模块的有效性结果

基线	SF-FIM	CKC-CGM	LCUM	PSNR /dB	SSIM	运算量 ($\times 10^9$)	参数量 ($\times 10^6$)
√	—	—	—	31.03	0.961 2	33.86	3.52
√	√	—	—	35.85	0.985 6	29.77	3.11
√	—	√	—	36.67	0.987 2	29.41	2.48
√	—	—	√	32.82	0.963 1	31.79	3.48
√	√	√	—	40.11	0.991 8	25.32	2.06
√	√	—	√	36.78	0.987 5	27.70	3.07
√	—	√	√	38.24	0.990 1	27.34	2.44
√	√	√	√	41.46	0.994 3	27.47	3.16

为残差块。

(2)DPFFM:在(1)的基础上,将残差块替换为DPFFM。

(3)FFEM:在(1)的基础上,将残差块替换为FFEM。

(4)SFFIM:在(1)的基础上,将残差块替换为SFFIM。

如表3所示,将DFLU-Net中编码器和解码器部分的特征提取模块全部替换为残差块,PSNR和SSIM为38.24 dB、0.990 1。将特征提取模块替换为空间域的DPFFM,PSNR达到40.15 dB,SSIM达到0.992 1。相较于残差块,PSNR提升了1.91 dB,SSIM提升了0.002 0,这是因为DPFFM中将普通卷积与上述4种差分卷积结合,而差分卷积组提取的语义信息有助于提取图像的纹理特征和轮廓,从而提升模型的鲁棒性和特征提取能力。将特征提取模块替换为频域的FFEM,PSNR达到38.92 dB,SSIM达到0.990 7。相较于残差块,PSNR提升了0.68 dB,SSIM提升了0.000 6,验证了频域中捕获的全局特征能有效弥补模型中对结构信息提取的不足。将残差块替换为SFFIM,虽然适当地提高了运算量,但是SFFIM结合了DPFFM和FFEM的优点,既在空间域中提取局部细节特征与纹理特征,又能在频域中提取图像结构信息,模型性能达到了最优,PSNR达到41.46 dB,SSIM达到了0.994 3。

4.5 超参数对比实验

为了验证本文模型中设置的SFFIM和DEAB数量 N 和 Q 的合理性,本文设置了对比实验。在模型架构不变的情况下,分别将超参数 N 和 Q 设置为3,6,4,8和5,10。根据模型性能和复杂度综合比较不同参数下模型在SOTS-Indoor、SOTS-Outdoor和Haze4K数据集上的PSNR、SSIM、模型运算量和参数量,确定最优的 N 和 Q 的值,具体对比结果如表4所示。

当 N 和 Q 的值分别设置为4和8时,模型在SOTS-Indoor、SOTS-Outdoor和Haze4K数据集上的PSNR、SSIM指标分别为41.46 dB、0.994 3,37.73 dB、0.993 6,以及34.72 dB、0.992 0。相较于 N 和 Q 的值分别设置为3和6时,在三个数据集上的PSNR、SSIM指标分别提升了1.29 dB、0.001 8,1.91 dB、0.001 8,以及2.42 dB、

表 3 在 SOTO-Indoor 数据集上验证 DPFFM 和 FFEM 的有效性结果

残差块	DPFFM	FFEM	SFFIM	PSNR/dB	SSIM	运算量($\times 10^9$)	参数量($\times 10^6$)
√	—	—	—	38.24	0.990 1	22.34	2.44
—	√	—	—	40.15	0.992 1	20.16	1.89
—	—	√	—	38.92	0.990 7	10.90	1.82
—	—	—	√	41.46	0.994 3	27.47	3.16

表 4 在三个数据集上的超参数对比实验结果

SFFIM	DEAB	SOTS-Indoor		SOTS-Outdoor		Haze4K		运算量($\times 10^9$)	参数量($\times 10^6$)
		PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM		
3	6	40.17	0.992 5	35.82	0.991 8	32.30	0.987 0	22.54	2.54
4	8	41.46	0.994 3	37.73	0.993 6	34.72	0.993 0	27.47	3.16
5	10	41.52	0.994 9	37.88	0.994 2	34.83	0.993 2	32.41	3.79

0.005 0,且模型的运算量和参数量分别仅增加了 4.93×10^9 和 0.62×10^6 . 相较于 N 和 Q 的值分别设置为 5 和 10 时,虽然,在三个数据集上的 PSNR、SSIM 指标略有下降,但模型的运算量和参数量分别降低了 4.94×10^9 和 0.63×10^6 . 综合评估模型的性能和复杂度,文中 N 和 Q 的值分别设置为 4 和 8 时,模型综合性能达到最优.

5 结论

本文提出了一种双域特征交互和局部相关性上采样的图像去雾网络. 该网络充分考虑了有雾图像和无雾图像在频域信息的差异,为此设计了空间域-频域特征交互模块,该模块能够有效融合了全局上下文信息和局部细节信息,增强网络对纹理信息和结构信息的提取能力. 此外,为实现 U-Net 结构上采样过程中语义信息自下而上地传递,本文设计了局部相关性上采样模块,利用注意力机制和双线性插值法捕捉每个特征图中局部信息的内在相关性,并将信息向浅层传递,有效克服了最近邻插值法的局限性. 最后,本文在定性比较不同模型的去雾结果时创新性地引入了差值热力图,通过热力图颜色的深浅更直观地展示了不同模型的去雾性能差异和图像复原能力. 在 RESIDE 数据集和 Haze4K 数据集上的实验表明,本文方法能够在提高去雾精度的同时,有效降低模型参数量和计算复杂度.

参考文献

- [1] 贾童瑶, 卓力, 李嘉锋, 等. 基于深度学习的单幅图像去雾研究进展[J]. 电子学报, 2023, 51(1): 231-245.
JIA T Y, ZHUO L, LI J F, et al. Research advances on deep learning based single image dehazing[J]. Acta Electronica Sinica, 2023, 51(1): 231-245. (in Chinese)
- [2] CHEN Z X, HE Z W, LU Z M. DEA-net: Single image dehazing based on detail-enhanced convolution and content-guided attention[J]. IEEE Transactions on Image Processing, 2024, 33: 1002-1015.
- [3] 付文晓, 张锐, 林丽丽, 等. 雾天室外场景光照参数估计[J]. 软件学报, 2014, 25(增刊2): 268-277.
FU W X, ZHANG R, LIN L L, et al. Illumination estimation of hazing outdoor scene[J]. Journal of Software, 2014, 25(S2): 268-277. (in Chinese)
- [4] 任东东, 李金宝. 低光照环境下的图像恢复与目标识别方法[J]. 软件学报, 2019, 30(增刊1): 94-104.
REN D D, LI J B. Methods of image restoration and object detection in low-light environment[J]. Journal of Software, 2019, 30(S1): 94-104. (in Chinese)
- [5] YU H, ZHENG N S, ZHOU M, et al. Frequency and spatial dual guidance for image dehazing[C]//Computer Vision-ECCV 2022. Cham: Springer, 2022: 181-198.
- [6] HE K M, SUN J, TANG X O. Single image haze removal using dark channel prior[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2009: 1956-1963.
- [7] BERMAN D, TREIBITZ T, AVIDAN S. Non-local image dehazing[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 1674-1682.
- [8] ZHU Q S, MAI J M, SHAO L. A fast single image haze removal algorithm using color attenuation prior[J]. IEEE Transactions on Image Processing, 2015, 24(11): 3522-3533.
- [9] LIU R S, FAN X, HOU M J, et al. Learning aggregated transmission propagation networks for haze removal and beyond[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(10): 2973-2986.
- [10] LIU Y, PAN J S, REN J, et al. Learning deep priors for image dehazing[C]//2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2020: 2492-2500.
- [11] ZHAO S Y, ZHANG L, SHEN Y, et al. RefineDNet: A

- weakly supervised refinement framework for single image dehazing[J]. *IEEE Transactions on Image Processing*, 2021, 30: 3391-3404.
- [12] ZHOU H, CHEN Z K, LIU Y, et al. Physical-priors-guided dehazeformer[J]. *Knowledge-Based Systems*, 2023, 266: 110410.
- [13] KRICHEN M. Generative adversarial networks[C]//2023 14th International Conference on Computing Communication and Networking Technologies. Piscataway: IEEE, 2023: 1-7.
- [14] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[EB/OL]. (2023-08-02)[2025-06-10]. <https://arXiv.org/abs/1706.03762>.
- [15] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]//*Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer, 2015: 234-241.
- [16] CAI B L, XU X M, JIA K, et al. DehazeNet: An end-to-end system for single image haze removal[J]. *IEEE Transactions on Image Processing*, 2016, 25(11): 5187-5198.
- [17] LI B Y, PENG X L, WANG Z Y, et al. AOD-net: All-in-one dehazing network[C]//2017 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2017: 4780-4788.
- [18] ZHANG H, PATEL V M. Densely connected pyramid dehazing network[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 3194-3203.
- [19] LI Y N, MIAO Q G, OUYANG W L, et al. LAP-net: Level-aware progressive network for image dehazing[C]//2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2020: 3275-3284.
- [20] REN W Q, PAN J S, ZHANG H, et al. Single image dehazing via multi-scale convolutional neural networks with holistic edges[J]. *International Journal of Computer Vision*, 2020, 128(1): 240-259.
- [21] ZHANG L B, WANG S, WANG X H. Single image dehazing based on bright channel prior model and saliency analysis strategy[J]. *IET Image Processing*, 2021, 15(5): 1023-1031.
- [22] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//2017 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2017: 2242-2251.
- [23] ENGIN D, GENÇ A, EKENEL H K. Cycle-dehaze: Enhanced CycleGAN for single image dehazing[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE, 2018: 938-9388.
- [24] 但志平, 方帅领, 孙航, 等. 基于双判别器异构 CycleGAN 框架下多阶通道注意力校准的室外图像去雾[J]. *电子学报*, 2023, 51(9): 2558-2571.
- DAN Z P, FANG S L, SUN H, et al. Outdoor image dehazing based on multi-order channel attention calibration using a dual-discriminator heterogeneous CycleGAN framework[J]. *Acta Electronica Sinica*, 2023, 51(9): 2558-2571. (in Chinese)
- [25] LIU X H, MA Y R, SHI Z H, et al. GridDehazeNet: Attention-based multi-scale network for image dehazing[C]//2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2020: 7313-7322.
- [26] QIN X, WANG Z L, BAI Y C, et al. FFA-net: Feature fusion attention network for single image dehazing[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34(7): 11908-11915.
- [27] 高继蕊, 李华锋, 张亚飞, 等. 双注意力引导的细节和结构信息融合图像去雾网络[J]. *电子学报*, 2023, 51(1): 160-171.
- GAO J R, LI H F, ZHANG Y F, et al. Dual attention-guided detail and structure information fusion network for image dehazing[J]. *Acta Electronica Sinica*, 2023, 51(1): 160-171. (in Chinese)
- [28] LU L P, XIONG Q, XU B R, et al. MixDehazeNet: Mix structure block for image dehazing network[C]//2024 International Joint Conference on Neural Networks. Piscataway: IEEE, 2024: 1-10.
- [29] DONG H, PAN J S, XIANG L, et al. Multi-scale boosted dehazing network with dense feature fusion[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2020: 2154-2164.
- [30] LUO P J, XIAO G Q, GAO X B, et al. LKD-net: Large kernel convolution network for single image dehazing[C]//2023 IEEE International Conference on Multimedia and Expo. Piscataway: IEEE, 2023: 1601-1606.
- [31] 孙航, 付秋月, 李勃辉, 等. 基于跨层注意力特征交互和多尺度通道注意力的单幅图像去雾网络[J]. *电子学报*, 2024, 52(11): 3711-3726.
- SUN H, FU Q Y, LI B H, et al. Cross-layer attention feature interaction and multi-scale channel attention network for single image dehazing[J]. *Acta Electronica Sinica*, 2024, 52(11): 3711-3726. (in Chinese)
- [32] KINGMA D P, BA J L. Adam: A method for stochastic optimization[EB/OL]. (2017-01-30)[2025-06-10]. <https://arXiv.org/abs/1712.04752>.

arXiv.org/abs/1412.6980.

- [33] LI B Y, REN W Q, FU D P, et al. Benchmarking single image dehazing and beyond[J]. IEEE Transactions on Image Processing, 2018, 28(1): 492-505.
- [34] LIU Y, ZHU L, PEI S D, et al. From synthetic to real: Image dehazing collaborating with unlabeled real data[C]// Proceedings of the 29th ACM International Conference on Multimedia. New York: ACM, 2021: 50-58.
- [35] BAI H R, PAN J S, XIANG X G, et al. Self-guided image dehazing using progressive feature fusion[J]. IEEE Transactions on Image Processing, 2022, 31: 1217-1229.
- [36] SONG Y D, HE Z Q, QIAN H, et al. Vision transformers

for single image dehazing[J]. IEEE Transactions on Image Processing, 2023, 32: 1927-1941.

- [37] WANG Z Z, ZHAO H T, YAO L J, et al. DFR-net: Density feature refinement network for image dehazing utilizing haze density difference[J]. IEEE Transactions on Multimedia, 2024, 26: 7673-7686.
- [38] 翟凤文, 朱玉彤, 金静, 等. 双分支卷积结合细节增强的图像去雾[J]. 西北工业大学学报, 2025, 43(1): 109-118.
- ZHAI F W, ZHU Y T, JIN J, et al. Image dehazing based on double branch convolution and detail enhancement[J]. Journal of Northwestern Polytechnical University, 2025, 43(1): 109-118. (in Chinese)

作者简介



刘明杰 男, 1987年出生于河南省辉县市。现为重庆邮电大学副教授、博士生导师。主要研究方向为智能网联汽车环境智能感知及控制、机器学习。
E-mail: liumj@cqupt.edu.cn



陈俊生 男, 1989年1月出生于安徽省六安市。现为重庆邮电大学讲师、硕士生导师。主要研究方向为多源信息融合、机器学习和人工智能。
E-mail: chenjunsheng@cqupt.edu.cn



吕梦琳 女, 2001年8月出生于重庆市南岸区。现为重庆邮电大学硕士研究生。主要研究方向为图像去雾算法、知识蒸馏方法。
E-mail: S230301033@stu.cqupt.edu.cn



朴昌浩 男, 1978年11月出生于吉林省延边朝鲜族自治州。现为重庆邮电大学教授、博士生导师。主要研究方向为智能网联汽车。
E-mail: piaoch@cqupt.edu.cn