

文章编号:1007-2780(2025)10-1557-11

基于多级特征编码与双分支引导重构的 图像去雾网络

张鑫琪, 田莹*, 李义荣

(辽宁科技大学 计算机与软件工程学院, 辽宁 鞍山 114051)

摘要:针对现有去雾技术在处理有雾图像时因细节特征丢失和全局信息不足而导致图像质量下降的问题,提出了一种基于多级特征编码与双分支引导重构的图像去雾网络 HEDGR-Net。首先,在编码阶段引入结构增强卷积与非局部注意力机制,有效提取图像中不同尺度的关键结构与上下文信息;其次,在解码阶段设计线性解码与上采样解码并行的双分支结构,并设计全局感知引导模块将结果进行融合,提取图像整体风格向量以增强重建一致性;最后,在输出阶段引入注意力增强残差连接模块,将解码阶段的特征与原始输入特征加权融合,进一步提升去雾图像的视觉质量。此外,采用 L2 损失、SSIM 损失和 TV 损失相结合的复合损失函数,有效提升了图像细节的恢复能力和结构的一致性。在合成客观数据集 SOTS-outdoor 中,PSNR 和 SSIM 分别达到了 27.54 dB 和 0.957 0;在真实有雾数据集 RTTS 中,BRISQUE 和 Entropy 分别为 26.99 和 8.77。与传统算法 AOD-Net 相比,HEDGR-Net 在 PSNR 上提升了约 21.7%,SSIM 提升了约 5.0%,BRISQUE 下降了约 19.2%,Entropy 提升了约 3.9%,在多个指标上均表现出明显优势。所提方法能够提高图像的细节恢复能力与整体亮度的一致性,有效改善传统去雾方法去雾不彻底、色彩偏移等问题。

关键词:计算机视觉;图像去雾;多级特征融合;双分支结构;注意力机制

中图分类号:TP391.4 文献标识码:A doi:10.37188/CJLCD.2025-0163 CSTR:32172.14.CJLCD.2025-0163

Hierarchical encoding and dual-branch guided reconstruction network for image dehazing

ZHANG Xinqi, TIAN Ying*, LI Yirong

(School of Computer Science and Software Engineering, University of Science and Technology Liaoning,
Anshan 114051, China)

Abstract: To address the limitations of existing dehazing methods in detail restoration and global modeling, this paper proposes HEDGR-Net, a deep dehazing network based on hierarchical encoding and dual-branch guided reconstruction. HEDGR-Net adopts structure-enhanced convolution and non-local attention in the encoder to capture multi-scale structural and contextual features. In the decoder, a dual-branch design with a global context-guided module integrates local detail recovery and global semantic consistency. At the output stage, an attention-enhanced residual fusion module combines reconstructed and input features to refine image quality. A composite loss combining L2, SSIM, and TV is further

收稿日期:2025-08-13;修订日期:2025-08-29.

基金项目:辽宁省教育厅资助项目(No.LJ212410146058)

Supported by Project of Education Department of Liaoning Province (No.LJ212410146058)

*通信联系人, E-mail:astianying@126.com

employed to balance pixel fidelity, structural similarity, and smoothness. On the synthetic SOTS-outdoor dataset, HEDGR-Net achieves 27.54 dB PSNR and 0.957 0 SSIM. On the real RTTS dataset, it obtains 26.99 BRISQUE and 8.77 Entropy. Compared with AOD-Net, HEDGR-Net improves PSNR by 21.7%, SSIM by 5.0%, reduces BRISQUE by 19.2%, and increases Entropy by 3.9%, showing clear advantages across multiple metrics. The proposed method enhances detail restoration and global brightness consistency, effectively overcoming the limitations of traditional dehazing methods such as incomplete haze removal and color distortion.

Key words: computer vision; image dehazing; multi-level feature fusion; dual-branch structure; attention mechanisms

1 引言

在计算机视觉技术快速发展的背景下,雾和霾等大气条件在图像采集与传输过程中常导致视觉信息退化,给视觉系统带来挑战^[1],尤其是安全监控、交通管理以及自动驾驶等领域。近年来,受气候变化影响,雾霾发生的频率和范围持续增加,不仅降低图像质量,还会削弱传感器感知能力,对后续图像处理的决策和行动造成干扰^[2]。因此,研究和新的图像去雾方法对提高图像质量和增强视觉系统的可靠性具有重要意义^[3]。

传统图像的去雾方法大多基于先验信息,通过假设统计特性来估计退化程度并恢复清晰图像^[4]。He等^[5]提出暗通道先验方法,通过估计透射矩阵并结合大气光估计值去雾。Berman等^[6]提出了一种基于Non-Local颜色簇的图像去雾方法,利用颜色空间分布恢复图像特征。Zhu等^[7]提出基于颜色衰减先验的景深线性模型有效恢复单幅雾图像,但对先验假设依赖较强,在复杂天气条件下效果往往不理想^[8]。

基于卷积神经网络的去雾方法取得了显著进展,其通过学习大量有雾与无雾图像对,能够自动提取特征并实现去雾^[9]。Cai等^[10]通过卷积神经网络估计透射图以恢复清晰图像。Dong等^[11]利用多尺度卷积结构在不同尺度捕获图像细节。Li等^[12]通过同时估计透射矩阵值和大气光值,简化了去雾流程并提高了处理速度。Qin等^[13]引入注意力机制增强关键特征提取能力。Luo等^[14]通过大感受野卷积结构提升全局信息建模能力。Qui等^[15]结合多分支结构与Taylor展开理论实现复杂场景下的自适应处理。然而,现有

方法在复杂雾霾环境下仍存在一定局限。一方面,多数模型结构较为单一,难以充分刻画多尺度深层特征,导致细节恢复不足^[16];另一方面,不同尺度特征之间缺乏有效的信息交互与融合机制,在光照剧烈变化或雾霾分布不均时易出现色彩偏移和结构模糊^[17]。

为了解决这些问题,本文提出了一个全新的深度学习去雾网络HEDGR-Net(Hierarchical Encoding and Dual-branch Guided Reconstruction Network),设计了多级深度特征编码器、双分支引导聚合解码器,并通过跳跃连接和注意力机制对结果进行融合,在不丢失重要特征的同时有效整合信息,更精准地处理图像中的雾霾分布信息。本文的主要贡献包括:

(1)设计了多级深度特征编码器(Hierarchical Deep Feature Encoder, HDF-Encoder),高效捕捉多尺度语义信息,显著增强了模型对雾霾分布的感知与建模能力。

(2)构建了双分支引导聚合解码器(Dual-branch Guided Aggregation Decoder, DGA-Decoder),通过结构恢复与全局引导的深度融合,实现了更加自然且信息丰富的图像重建。

(3)提出了注意力增强残差融合模块(Attention Enhanced Residual Fusion Module, AERF Module),通过结合跳跃连接与注意力机制,进一步提升了图像的结构一致性和视觉质量。

2 HEDGR-Net网络结构

本文提出的图像去雾网络HEDGR-Net由多级深度特征编码器、双分支引导聚合解码器以及注意力增强残差融合模块构成,其整体结构如

图 1 所示。HDF-Encoder 用于提取多尺度特征并兼顾局部结构与全局上下文建模;DGA-Decoder 采用双分支设计,实现更自然的信息整合;AERF Module 进一步结合注意力机制和残差连接,增强特征融合效果与最终图像的结构保真度。

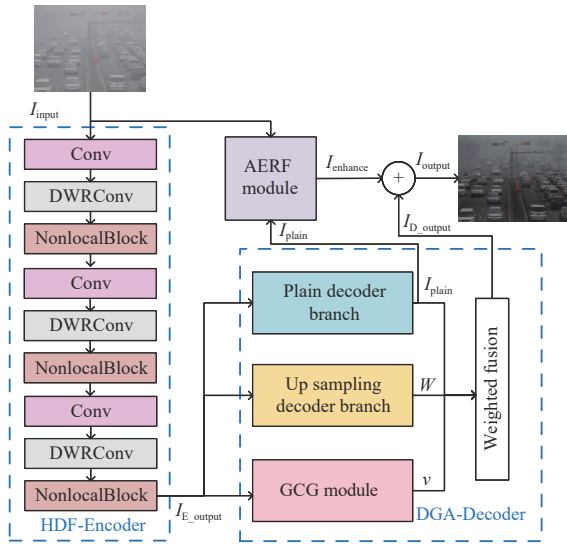


图 1 HEDGR-Net 结构图

Fig. 1 HEDGR-Net structure diagram

2.1 多级深度特征编码器

编码器作为整个模型的关键,负责提取和编码输入图像的特征。为提升模型特征表示与层次感知,提出了一种多级深度特征编码器(HDF-Encoder),其整体结构如图 2 所示。

在特征提取阶段,输入图像首先经过带 ReLU 激活的标准卷积实现通道扩展并提取初级语义特征。随后,设计结构增强模块(DWRConv)以提升对结构细节的感知能力。DWRConv 首先利用 1×1 的卷积进行特征降维以压缩冗余信息,然后通过动态加权残差(DWR)结构^[18]在多个扩张率($d=1,3,5$)下执行卷积操作,以增强模型对多尺度特征的响应能力。最后,结合批归一化与 GELU 激活函数^[19]进一步提升其非线性表达能力,GELU 的计算方式如式(1)所示:

$$\text{GELU}(x) = 0.5 \times$$

$$x \left[1 + \tanh \left(\sqrt{\frac{2}{\pi}} (x + 0.047715 \times x^3) \right) \right], \quad (1)$$

其中, x 表示编码特征在 DWRConv 中经过批归一化(BN)处理过后得到的线性结果。

为增强网络对关键区域的捕捉能力,多级深度特征编码器在各个阶段均引入非局部注意力模块(NonLocalBlock),通过计算特征各位置间的关系建立长距离依赖,从而获得全局特征表达。为了降低计算复杂度,首先在通道维度上将输入特征由 $[c, w, h]$ 降维至 $[c/2, w, h]$,再通过线性变换生成特征映射矩阵 t 及其转置 t' ,其形状为 $[c/2, w \times h]$ 和 $[w \times h, c/2]$,通过矩阵乘法获得 t 与 t' 的空间位置之间的相似度矩阵 f 。采用 Softmax 操作对相似度矩阵进行归一化,得到注意力权重矩阵 f_c ,并与特征映射矩阵 t 相乘获得包

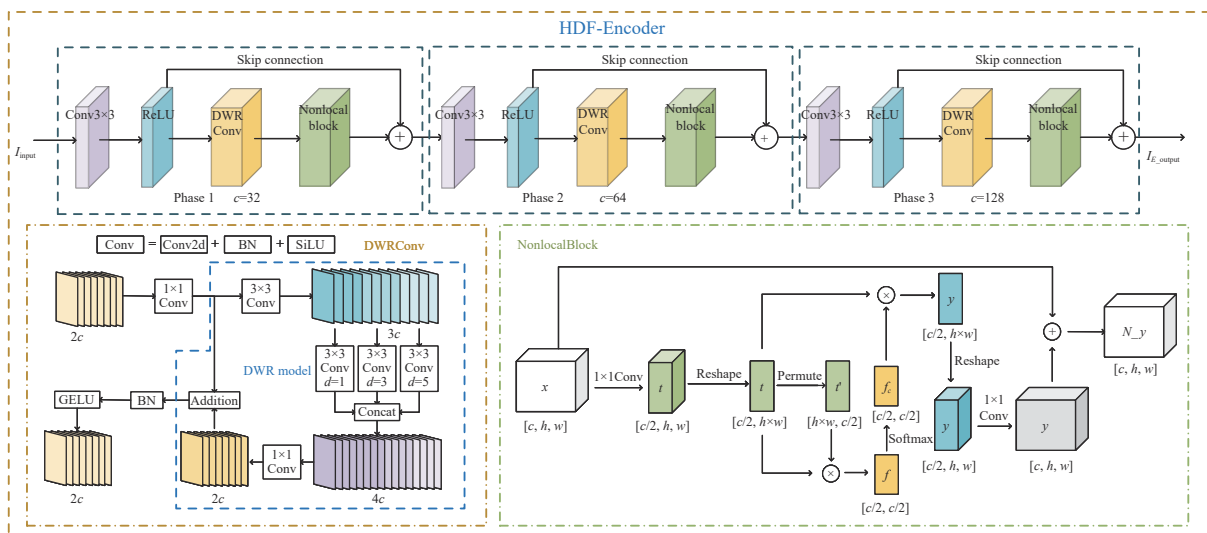


图 2 多级深度特征编码器结构图

Fig. 2 HDF-Encoder structure diagram

含融合全局上下文信息的特征图 y 。最终将 y 通过一次升维操作恢复至原始通道维度后与输入特征 x 相加,以残差方式实现全局信息建模与局部特征保持的协同融合。

每个非局部注意力模块的输出均与对应阶段 DWRCConv 之前的特征进行跳跃连接,有效保持深层次信息提取过程中的连续性与稳定性。此外,多级深度特征编码器在多级编码中采用逐层扩展通道数的方式,在第二阶段和第三阶段分别拓展至 64 和 128 通道,使网络能够在更宽的特征空间中逐步建立起从底层纹理到高层语义的多尺度特征层次结构。

2.2 双分支引导聚合解码器

为提升模型在图像重建任务中的结构建模能力与特征表达能力,本文设计了一种双分支引导聚合解码器(DGA-Decoder)。该解码模块旨在充分整合深层编码特征所蕴含的多尺度信息,精准还原图像中的结构边界与细节特征,其结构如图 3 所示。

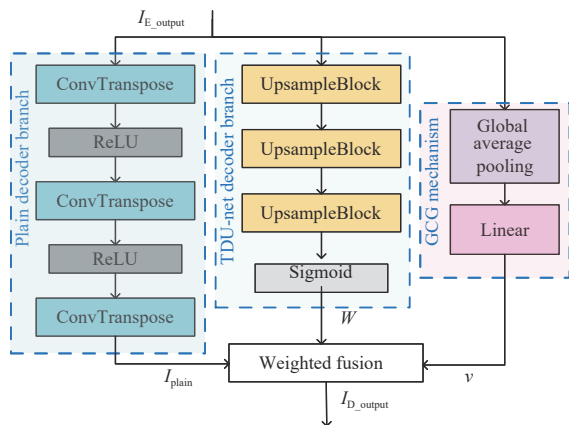


图 3 双分支引导聚合解码器结构图

Fig. 3 DGA-Decoder structure diagram

双分支引导聚合解码器由 3 类功能性子模块构成：

(1) 线性解码分支(Plain Decoder Branch), 通过逐级卷积映射实现特征的线性还原, 强调结构连续性与低频恢复;

(2) 上采样解码分支(Upsampling Decoder Branch), 采用上采样与卷积操作融合, 关注细节补偿与空间分辨率的恢复;

(3) 全局感知引导(Global Context-Guided, GCG)模块, 通过捕捉跨区域依赖关系, 引导解码器在融合过程中充分考虑全局语义信息。

线性解码分支由转置卷积(ConvTranspose)操作与非线性激活单元(ReLU)组成, 主要用于恢复编码器提取的中间特征中的空间结构信息, 并逐步还原图像的基础细节层次。ConvTranspose 输出图像的尺寸计算形式如式(2)和式(3)所示:

$$H_{out} = s \times (H_{in} - 1) + k - 2p, \quad (2)$$

$$W_{out} = s \times (W_{in} - 1) + k - 2p, \quad (3)$$

其中: H_{in} 和 W_{in} 分别是转置卷积层输入特征的高度和宽度, H_{out} 和 W_{out} 表示转置卷积层输出特征图的高度和宽度, s 表示步长, p 表示填充大小, k 代表卷积核大小。

上采样解码分支由多个上采样模块(UpsampleBlock)组成, 能够逐级提升特征图的空间分辨率, 并学习用于图像融合的像素级引导图。上采样模块的具体结构如图 4 所示。该引导图通过末端的 Sigmoid 激活函数进行归一化处理, 数值范围限定在 0~1 之间, 用于在重建过程中对局部特征与全局信息的权重进行动态调节。

在双分支基础上, 引入全局感知引导模块以增强语义建模。通过全局平均池化(Global

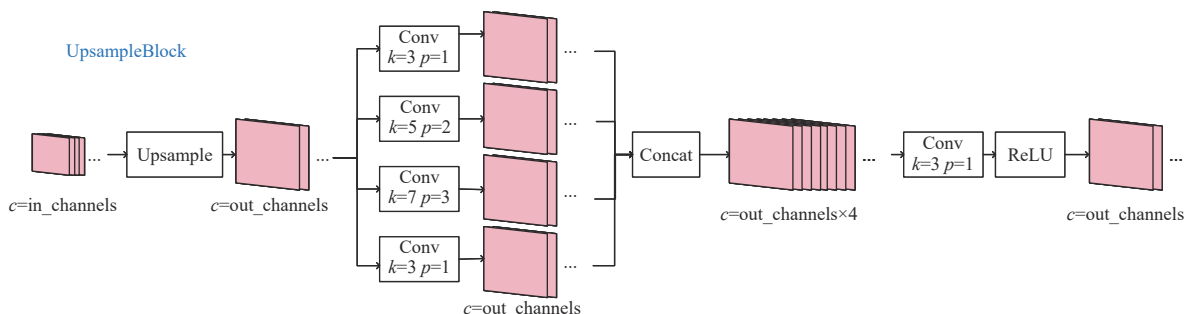


图 4 上采样模块结构图

Fig. 4 UpsampleBlock structure diagram

Average Pooling)和全连接层(Linear)生成低维RGB引导向量,作为全局补偿项与解码器输出加权融合,提升重建图像的结构一致性与语义完整性。最终,将双分支解码器的输出与全局感知引导模块的引导向量进行加权融合(Weighted Fusion),形成双分支引导聚合解码器的复原图像,其计算如式(4)所示:

$$I_{D_output} = I_{plain} \times W + v \times (1 - W), \quad (4)$$

其中: I_{plain} 表示线性解码分支生成的初步重建图像; $W \in [0, 1]$ 为像素级的引导图,用以刻画不同空间位置对局部解码结果的置信度; v 为全局感知引导模块输出的RGB引导向量,用于提供全图范围内的语义补偿。

2.3 注意力增强残差融合模块

为进一步提升模型对局部特征的精细建模能力,并增强解码阶段的残差融合效果,本文设计了一种注意力增强残差融合(AERF)模块,其结构图如图5所示。

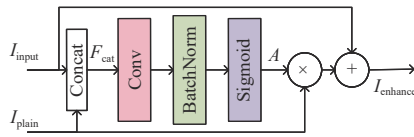


图5 注意力增强残差融合模块结构图

Fig. 5 AERF module structure diagram

该模块基于特征融合中的注意力机制思想,通过对线性解码分支的输出 I_{plain} 与原始图像 I_{input} 之间的关系建模,实现连接特征的自适应加权调节,从而有效提升融合特征的判别性与重建精度。首先对解码器输出特征与原始输入特征在通道维度进行拼接,得到联合特征 F_{cat} 。随后,该联合特征通过一个由 1×1 卷积、批归一化(BN)以及Sigmoid激活函数组成的注意力计算单元,生成一个通道注意力图 A ,其计算过程如式(5)所示:

$$A = \sigma(\text{BN}(\text{Conv}(F_{cat}))), \quad (5)$$

其中, σ 代表Sigmoid激活函数。

A 用于对由 F_{cat} 拼接而成的联合特征进行加权调节,进一步增强其中与当前解码特征 I_{plain} 高度相关的局部信息。将加权后的特征与输入图像 I_{input} 进行残差相加,最终输出融合后的增强图像 $I_{enhance}$,计算方式如式(6)所示:

$$I_{enhance} = A \times I_{plain} + I_{input}. \quad (6)$$

该过程不仅提升了信息在图像重建中的有效性,也强化了模型对关键区域的响应能力,使得最终重建图像更具结构连贯性和细节完整性。

2.4 损失函数

在图像去雾任务中,损失函数用于衡量去雾后图像与真实无雾图像之间的差异,指导模型在训练过程中进行优化。为进一步提高去雾模型的性能,本文提出了一种结合L2损失、SSIM损失和TV损失的复合损失函数,其结构如式(7)所示:

$$L_{mix} = L_{L2} + \lambda_1 L_{SSIM} + \lambda_2 L_{TV}, \quad (7)$$

其中: L_{L2} 为L2损失函数, L_{SSIM} 为SSIM损失函数, L_{TV} 为TV损失函数。

在复合损失设计中,权重系数的设定对模型性能起决定性作用。L2损失强调像素一致性,但忽略感知质量;SSIM损失突出结构与纹理约束,但过大时会抑制像素重建;TV损失能够降低噪声与伪影,但若权重过高则会造成纹理丢失。综合梯度强度与训练稳定性,SSIM项的约束应保持与L2项相近,以避免结构约束过强,而TV项因未归一化需设定在极小范围,仅作为弱正则。基于上述分析以及Optuna调参框架得到的实验结果,最终采用 $\lambda_1 = 0.85, \lambda_2 = 1.2 \times 10^{-7}$ 。

L2损失函数定义为:

$$L_{L2} = \frac{1}{N} \sum_{i,j} [x_{i,j} - y_{i,j}]^2, \quad (8)$$

其中: $x_{i,j}$ 是图像中位置 (i,j) 的像素值, $y_{i,j}$ 为真实无雾图像在位置 (i,j) 的像素值, N 为图像像素总数。

L2损失函数能够抑制噪声并加速收敛,但忽视了结构与纹理信息,可能导致图像模糊。为此,本文引入基于亮度、对比度和结构一致性的SSIM损失函数,定义为:

$$L_{SSIM}(x, y) = 1 - \text{SSIM}(x, y), \quad (9)$$

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (10)$$

其中: μ_x 和 μ_y 分别表示图像 x 与 y 的平均亮度, σ_x 和 σ_y 分别为其方差, σ_{xy} 为协方差, C_1 与 C_2 为常数。

SSIM损失函数关注去雾图像与真实无雾图像之间在亮度、对比度以及纹理结构上的相似

性,能更全面地反映图像的质量。此外,本文还引入了TV损失,通过约束相邻像素的差异来增强平滑性并减少噪声,其计算方法如式(11)所示:

$$L_{TV} = \sum_{i,j} (|x_{i,j} - x_{i+1,j}| + |x_{i,j} - x_{i,j+1}|), \quad (11)$$

其中, $x_{i,j}$ 是图像中位置 (i,j) 的像素值。

最终,结合L2损失、SSIM损失和TV损失的复合损失函数如式(12)所示:

$$L_{\text{mix}} = \frac{1}{N} \sum_{i,j} [x_{i,j} - y_{i,j}]^2 + \lambda_1 \times \left(1 - \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \right) + \lambda_2 \times \sum_{i,j} (|x_{i,j} - x_{i+1,j}| + |x_{i,j} - x_{i,j+1}|). \quad (12)$$

该复合损失能够在训练过程中同时平衡像素精度、结构保真与图像平滑性,使模型在细节恢复与整体质量优化之间取得更优的折中,从而显著提升去雾效果。

3 实验结果与分析

3.1 数据集与实验环境

本实验从公开数据集 RESIDE^[20] 的子数据集 OTS (Outdoor Training Set) 中随机抽取了 10 000 张合成的室外有雾图像以及相应的 2 049 张真实室外图像作为训练集。OTS 是一个大规模的合成数据集,包含多种天气条件下模拟生成的有雾图像。测试阶段使用合成有雾数据集 SOTS-outdoor (Synthetic Objective Testing Set-outdoor) 和真实场景数据集 RTTS (Real Task-driven Testing Set)。SOTS-outdoor 包含 500 对合成有雾图像与对应的无雾参考图像;RTTS 包含 4 322 张来源于真实自然场景的室外有雾图像,图像类型丰富,内容复杂,常用于衡量模型在实际应用中的泛化能力与稳健性。

实验平台为 Ubuntu 22.04.3 LTS,采用 NVIDIA GeForce GTX 1080Ti GPU 进行加速训练,开发框架为 Python 3.8.16 与 PyTorch 1.8.1。优化器选择 Adam,用于更新网络参数,初始学习率设置为 0.000 1,权重衰减系数为 0.000 1。模型训练共进行 10 个 epoch,批量大小为 8。

3.2 评价指标

在模型性能评估中,本文根据测试集的性质选择不同的指标。对于 SOTS-outdoor 测试集,由于其包含成对的有雾与清晰图像,采用全参考指标峰值信噪比 (PSNR) 与结构相似性 (SSIM)。PSNR 通过计算重建图像与参考图像之间的均方误差 (MSE) 衡量图像的失真程度,单位为分贝 (dB),数值越高表示图像质量越接近参考图像,如式 (13) 所示:

$$\text{PSNR} = 10 \times \lg \left(\frac{\text{MAX}^2}{\text{MSE}} \right), \quad (13)$$

其中,MAX 为像素取值的最大值。

SSIM 是一种基于人类视觉感知特性的图像结构相似性度量方法,综合考虑亮度、对比度和结构信息,取值范围为 $[0, 1]$,值越大表示结构一致性越强,SSIM 的公式表达如式 (10) 所示。

对于 RTTS 测试集,由于缺乏参考图像,采用无参考指标信息熵 Entropy 与 BRISQUE^[22]。Entropy 通过分析图像像素值分布所包含的信息量来衡量图像的复杂度与细节丰富度,熵值越高通常表示图像中包含更多纹理信息。BRISQUE 基于自然图像统计特性,通过分析图像局部区域的亮度归一化残差,衡量图像与自然图像在统计分布上的偏离程度,其得分越低表示图像越接近真实自然分布。

3.3 λ_1 和 λ_2 的选择与调优

为了筛选式 (7) 中 SSIM 损失和 TV 损失在总损失函数中的最优权重比例 λ_1 和 λ_2 ,使用 Optuna 调参框架进行了 100 次实验,以验证集的 PSNR 作为选择指标,部分实验在测试集上的结果如表 1 所示。

结果表明,当 λ_1 在较低区间时结构约束不足,去雾效果和细节还原较弱;当 λ_1 提升至 0.80 和 0.90 时,结构先验增强,有助于提升整体一致性和细节质量;继续增大 λ_1 则会抑制像素级重建,导致性能下降。从 λ_2 的影响来看,较大的 λ_2 会显著削弱模型的性能。当 λ_2 取 6.76×10^{-4} 或 6.71×10^{-4} 时,PSNR 仅为 21~22 dB。将 λ_2 控制在 $1.00 \times 10^{-6} \sim 5.00 \times 10^{-5}$ 范围内可获得更稳健的结果。当 λ_1 取 0.85 且 λ_2 取 1.2×10^{-7} 时,模型在两项评价指标上达到最优。

表 1 λ_1 和 λ_2 损失权重对去雾性能的影响Tab. 1 Effect of the loss weights λ_1 and λ_2 on dehazing performance

λ_1	$\lambda_2(\times 10^{-4})$	PSNR/dB	SSIM
0.03	6.710 0	22.38	0.800 8
0.09	0.248 0	22.35	0.820 4
0.24	0.035 1	23.99	0.893 3
0.32	0.083 3	23.36	0.880 2
0.38	0.026 8	22.34	0.819 7
0.45	0.055 9	20.74	0.751 4
0.50	6.760 0	21.18	0.753 3
0.56	0.232 0	22.90	0.861 3
0.64	0.017 3	22.98	0.857 7
0.70	0.023 9	23.11	0.859 1
0.76	0.171 0	24.08	0.899 5
0.80	0.070 1	25.67	0.944 5
0.83	0.082 8	26.53	0.952 0
0.84	0.175 0	26.42	0.952 8
0.85	0.001 2	27.54	0.957 0
0.89	0.005 5	27.09	0.951 3
0.94	0.266 0	25.83	0.946 6
0.97	0.017 4	25.13	0.919 6
0.99	0.003 2	24.43	0.903 9

注:粗体数字代表最优结果

3.4 对比实验

为全面验证所提去雾网络 HEDGR-Net 的性能,本文选取 DCP、AOD、PFFN^[23]、FFA、MSBDN、LKD、MB_TaylorFormer、DEA^[25] 及 MixDehazeNet^[26] 等具有代表性的经典与深度学习方法,在合成数据集 SOTS-outdoor 与真实数据集 RTTS 上进行了对比实验,结果如表 2 与表 3 所示。

在 SOTS-outdoor 合成数据集上, HEDGR-Net 显著优于对比算法,在 PSNR 与 SSIM 指标上分别达到 27.54 dB 和 0.957 0,较 DCP 提升 10.49 dB 和 0.088 4,较 AOD 提升 4.92 dB 和 0.045 7,较 DEA 提升 3.22 dB 和 0.008 0,较 MixDehazeNet 提升 2.89 dB 和 0.027 1,较当前表现优的 MB_TaylorFormer 分别提升 2.25 dB 和 0.013 0,表明本文模型在合成雾图像条件下能够更好地恢复被遮挡的细节特征,在整体结构建模上实现更精确的还原。

在 RTTS 真实数据集上, HEDGR-Net 同样展现出优越性能, BRISQUE 值为 26.99, Entropy

表 2 合成图像数据集上的客观评价结果

Tab. 2 Objective evaluation results on synthetic image dataset

Models	PSNR/dB	SSIM
DCP ^[6]	17.05	0.868 6
AOD ^[13]	22.62	0.911 3
PFFN ^[23]	24.17	0.927 2
FFA ^[14]	24.16	0.935 9
MSBDN ^[12]	23.17	0.929 4
LKD ^[15]	24.24	0.928 4
MB_TaylorFormer ^[16]	25.29	0.944 0
DEA ^[24]	24.32	0.949 0
MixDehazeNet ^[25]	24.65	0.929 9
HEDGR-Net	27.54	0.957 0

注:粗体数字代表最优结果

表 3 真实图像数据集上的客观评价结果

Tab. 3 Objective evaluation results on real image dataset

Models	BRISQUE	Entropy
DCP ^[6]	35.69	8.53
AOD ^[13]	33.41	8.44
PFFN ^[23]	34.61	8.68
FFA ^[14]	35.15	8.75
MSBDN ^[12]	32.38	8.65
LKD ^[15]	34.17	8.74
MB_TaylorFormer ^[16]	33.41	8.73
DEA ^[24]	33.22	8.64
MixDehazeNet ^[25]	31.37	8.71
HEDGR-Net	26.99	8.77

注:粗体数字代表最优结果

为 8.77。其中, HEDGR-Net 在 BRISQUE 指标上相较于 AOD、MSBDN、MB_TaylorFormer 和 DEA 分别下降 6.42、5.39、6.42 和 6.23,显著降低了感知层级的结构失真;在 Entropy 指标上, HEDGR-Net 相较 DCP 提升 0.24,相较于 LKD、FFA、DEA 和 MixDehazeNet 分别提升 0.03、0.02、0.13 和 0.06,说明本文方法在真实复杂雾霾分布下依然能够保持较高的信息量与纹理细节,避免了过度平滑或细节缺失的问题。

从表 2 和表 3 可以看出,部分算法在合成与真实场景中的表现差异较大,泛化能力有限。例如, FFA 在合成数据集上的 SSIM 达到 0.935 9,说明其在合成条件下对图像结构有一定保持能

力,但在真实数据集中 BRISQUE 却高达 35.15,说明其对真实复杂雾状分布的适应能力有限。相比之下,本文方法在两个数据集上均表现稳定,既能在合成场景中准确重建纹理与边缘细节,也能在真实场景中保持光照与色彩的一致性。

3.5 消融实验

为了验证所提去雾网络 HEDGR-Net 中各个模块的有效性,对每一个部分进行消融实验,并将结果汇总在表 4 中。以图 6 中基本的编码器(Encoder)和解码器(Decoder)模块作为实验的基线模型,其 PSNR 为 19.62 dB,SSIM 为 0.791 4。

表 4 消融实验

Tab. 4 Ablation experiment

Models	PSNR/dB	SSIM
Baseline	19.62	0.791 4
Baseline+HDF-Encoder	24.60	0.922 2
Baseline+DGA-Decoder	22.43	0.861 1
Baseline+HDF-Encoder+ DGA-Decoder	26.82	0.953 6
HEDGR-Net	27.54	0.957 0

注:粗体数字代表最优结果

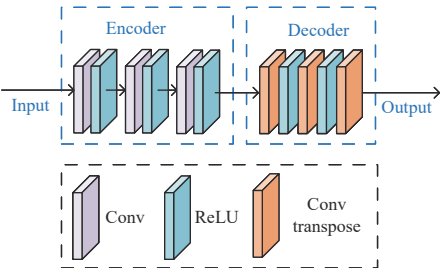


图 6 基线模型

Fig. 6 Baseline model

首先引入 HDF-Encoder, PSNR 较基线模型提高 4.98 dB, SSIM 提升 0.130 8, 说明其在细节刻画和结构保持方面作用显著。单独引入 DGA-Decoder 时, PSNR 较基线模型提高 2.81 dB, SSIM 增加 0.069 7, 说明双分支解码结构在局部与全局信息融合方面具有优势。当同时引入 HDF-Encoder 与 DGA-Decoder 时, 模型性能进一步提升, PSNR 达到 26.82 dB, SSIM 达到 0.953 6, 显示出两者在特征提取与图像重建上的协同效应。最后, 加入注意力增强残差融合模块, 构建完整的 HEDGR-Net, PSNR 达到 27.54 dB, SSIM

达到 0.957 0, 进一步增强了模型对关键区域的响应能力, 使最终结果在结构一致性和细节完整性方面更优。综上, 各模块既具独立有效性, 又具协同作用, 共同构成了高效、鲁棒的 HEDGR-Net。

3.6 可视化分析

为了更直观地验证所提算法的有效性, 从合成客观数据集 SOTS-outdoor 和真实有雾数据集 RTTS 选取图片对不同算法进行对比。

在合成数据集 SOTS-outdoor 上的实验结果如图 7 所示, 各方法在典型场景下的表现差异明显。DCP 去雾强度较大, 但普遍伴随色彩偏暗与过度处理, 在城市高楼场景中出现明显色偏; AOD 色彩较自然, 但建筑远景和居民楼场景中雾气残留明显, 远景细节模糊; PFFN 与 FFA 在边缘与高饱和区域易出现模糊、色块和色偏等不稳定现象; MBSDN 整体较均衡, 但在高楼远景和水面场景中整体对比度偏低; LKD 在天空与建筑边缘的恢复效果更接近清晰图像, 但高亮区容易过度曝光; MB_TaylorFormer 的全局色彩一致性较好, 但容易在强光区域产生失真; DEA 保持了边缘平滑, 但在高楼与水面场景中远景偏暗; MixDehazeNet 在近景增强明显, 但在天空和水面场景中存在亮度不均与轻微色彩偏差。相比之下, 本文算法在自然风景、城市高楼、道路居民楼以及水面场景中均能实现色彩自然、细节清晰与亮度平衡, 整体视觉效果最接近清晰图像。

在真实数据集 RTTS 上的实验结果如图 8 所示。DCP 虽能增强对比, 但图像颜色过深; AOD 色彩自然却去雾不足; PFFN 在行人与车辆以及林荫道路的边缘区域存在明显的模糊和色彩不连续现象, 在人物场景中出现色彩失真; FFA 和 MBSDN 在人物衣物、车辆和树木的高饱和区域常产生色块分离和异常色偏; LKD 和 MB_TaylorFormer 在高亮区域容易过度曝光, 在车辆、树木和人物衣服的细节恢复上仍有不足; DEA 对比度偏低, 远景部分模糊发灰; MixDehazeNet 在近景主体增强明显, 但在远处区域往往亮度不足或存在色彩偏差, 人物衣服出现色块偏差。相较而言, HEDGR-Net 在人物、车辆、林荫道路等多类真实场景中均能保持色彩自然、细节清晰和亮度均衡。

整体而言, HEDGR-Net 在细节恢复和全局

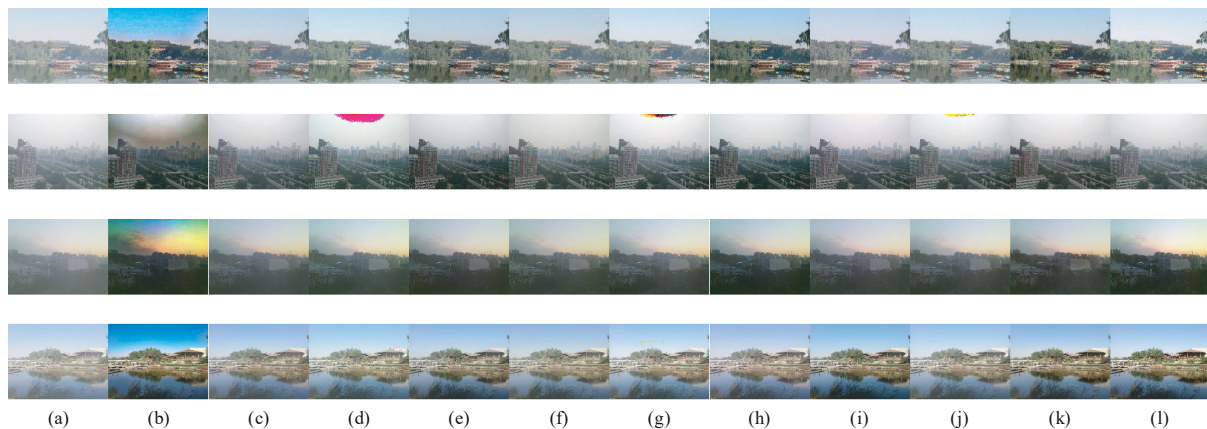


图 7 不同算法在合成客观数据集 SOTS-outdoor 中的去雾效果 (a) 有雾图像; (b) DCP 算法; (c) AOD 算法; (d) PFFN 算法; (e) FFA 算法; (f) MBSN 算法; (g) LKD 算法; (h) MB_TaylorFormer 算法; (i) DEA 算法; (j) MixDehazeNet 算法; (k) 本文算法; (l) 清晰图像。

Fig. 7 Dehazing effect of different algorithms on synthetic objective SOTS-outdoor dataset. (a) Hazy images; (b) DCP; (c) AOD; (d) PFFN; (e) FFA; (f) MBSN; (g) LKD; (h) MB_TaylorFormer; (i) DEA; (j) MixDehazeNet; (k) Ours; (l) Clear images.

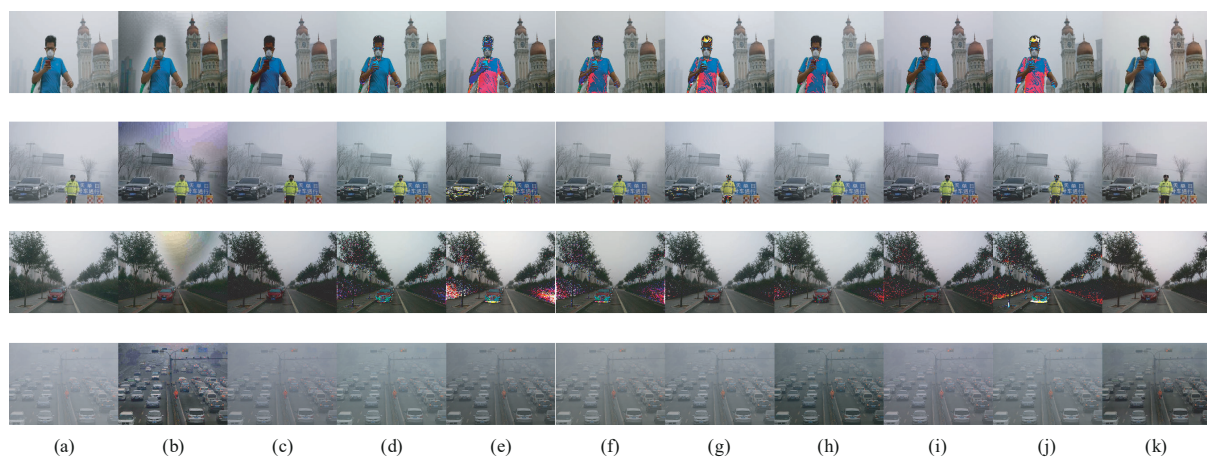


图 8 不同算法在真实有雾数据集 RTTS 中的细节对比。(a) 有雾图像; (b) DCP 算法; (c) AOD 算法; (d) PFFN 算法; (e) FFA 算法; (f) MBSN 算法; (g) LKD 算法; (h) MB_TaylorFormer 算法; (i) DEA 算法; (j) MixDehazeNet 算法; (k) 本文算法。

Fig. 8 Comparison of the details obtained by different algorithms on real fog dataset RTTS. (a) Hazy images; (b) DCP; (c) AOD; (d) PFFN; (e) FFA; (f) MBSN; (g) LKD; (h) MB_TaylorFormer; (i) DEA; (j) MixDehazeNet; (k) Ours.

一致性方面均优于现有方法,既能保证纹理与边缘的清晰度,又能维持色彩和亮度的整体协调,展现出更强的去雾能力。

4 结 论

本文针对现有图像去雾方法在复杂场景中存在的细节信息恢复不足与全局建模能力薄弱

等问题,提出了一种基于多级特征编码与双分支引导重构的图像去雾算法。所提算法构建了一个端到端的深度卷积网络框架,包含多级深度特征编码器、双分支引导解码器与注意力增强残差融合模块。在编码阶段引入结构增强卷积与非局部注意力机制,以建模多尺度语义结构和全局上下文关系。在解码阶段采用线性分支与上采样分支结合全局感知向量,实现局部细节恢复与

全局信息补偿的动态融合。在特征融合阶段设计注意力增强的残差连接模块,提升图像的结构一致性与视觉自然性。为进一步提升去雾效果,本文提出了结合 L2 损失、SSIM 损失和 TV 损失的复合损失函数,有效平衡像素误差与结构约束。实验结果表明,在合成有雾数据集 SOTS-

outdoor 上,所提方法的 PSNR 和 SSIM 分别达到 27.54 dB 与 0.957 0;在真实雾图像数据集 RTTS 上,BRISQUE 和 Entropy 分别为 26.99 与 8.77,表现出更高的图像质量与更低的感知失真,验证了所提算法在图像结构建模、细节还原与复杂场景适应性方面的有效性与鲁棒性。

参 考 文 献:

- [1] 谢欣丹,李晓艳,王鹏,等. 基于多尺度残差特征融合的图像去雾算法[J]. 液晶与显示,2024,39(6):822-832.
XIE X D, LI X Y, WANG P, *et al.* Image defogging algorithm based on multi-scale residual feature fusion [J]. *Chinese Journal of Liquid Crystals and Displays*, 2024, 39(6): 822-832. (in Chinese)
- [2] 傅妍芳,尹诗白,邓箴,等. 多级特征逐步细化及边缘增强的图像去雾[J]. 光学精密工程,2022,30(9):1091-1100.
FU Y F, YIN S B, DENG Z, *et al.* Multi-level features progressive refinement and edge enhancement network for image dehazing [J]. *Optics and Precision Engineering*, 2022, 30(9): 1091-1100. (in Chinese)
- [3] 陈永,郭红光,艾亚鹏. 基于多尺度卷积神经网络的单幅图像去雾方法[J]. 光学学报,2019,39(10):1010001.
CHEN Y, GUO H G, AI Y P. Single image dehazing method based on multi-scale convolution neural network [J]. *Acta Optica Sinica*, 2019, 39(10): 1010001. (in Chinese)
- [4] DENG J W, CHAN P P K, YEUNG D S. Real-world nighttime image dehazing using contrastive and adversarial learning [J]. *Pattern Recognition*, 2025, 165: 111596.
- [5] HE K M, SUN J, TANG X O. Single image haze removal using dark channel prior [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(12): 2341-2353.
- [6] BERMAN D, TREIBITZ T, AVIDAN S. Non-local image dehazing [C]//*Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas: IEEE, 2016: 1674-1682.
- [7] 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.
- [8] 汤永华,孟妍君,林森,等. 基于层级权值交互和拉普拉斯先验的非均匀去雾[J]. 激光与光电子学进展,2024,61(14):1437003.
TANG Y H, MENG Y J, LIN S, *et al.* Nonuniform defogging based on hierarchical weight interaction and laplacian prior [J]. *Laser & Optoelectronics Progress*, 2024, 61(14): 1437003. (in Chinese)
- [9] 王楚宁,赵开春. 基于暗通道先验原理的偏振图像去雾[J]. 光学精密工程,2023,31(23):3474-3481.
WANG C N, ZHAO K C. Polarization image defogging based on dark channel prior principle [J]. *Optics and Precision Engineering*, 2023, 31(23): 3474-3481. (in Chinese)
- [10] 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.
- [11] DONG H, PAN J S, XIANG L, *et al.* Multi-scale boosted dehazing network with dense feature fusion [C]//*Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle: IEEE, 2020: 2154-2164.
- [12] LI B Y, PENG X L, WANG Z Y, *et al.* AOD-Net: all-in-one dehazing network [C]//*Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV)*. Venice: IEEE, 2017: 4780-4788.
- [13] QIN X, WANG Z L, BAI Y C, *et al.* FFA-Net: feature fusion attention network for single image dehazing [C]//*Proceedings of the AAAI Conference on Artificial Intelligence*. New York: AAAI Press, 2020: 11908-11915.
- [14] 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 (ICME). Brisbane: IEEE, 2023: 1601-1606.
- [15] QIU Y W, ZHANG K H, WANG C X, *et al.* MB-TaylorFormer: multi-branc efficient transformer expanded by

- taylor formula for image dehazing [C]//*Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. Paris: IEEE, 2023: 12756-12767.
- [16] 杨燕,张全君,梁皓博. 混合注意力与多特征交互的去雾算法[J]. 哈尔滨工业大学学报,2024,56(9):56-64.
YANG Y, ZHANG Q J, LIANG H B. Dehazing algorithm based on mixed attention and multi-feature interaction [J]. *Journal of Harbin Institute of Technology*, 2024, 56(9): 56-64. (in Chinese)
- [17] SUI T Y, XIANG G F, CHEN F N, *et al.* U-shaped dual attention vision mamba network for satellite remote sensing single-image dehazing [J]. *Remote Sensing*, 2025, 17(6): 1055.
- [18] WEI H R, LIU X, XU S C, *et al.* DWRSeg: rethinking efficient acquisition of multi-scale contextual information for real-time semantic segmentation [J/OL]. *arXiv*, 2022: 2212.01173.
- [19] HENDRYCKS D, GIMPEL K. Gaussian error linear units (GELUs) [J/OL]. *arXiv*, 2016: 1606.08415.
- [20] LI B Y, REN W Q, FU D P, *et al.* Benchmarking single-image dehazing and beyond [J]. *IEEE Transactions on Image Processing*, 2019, 28(1): 492-505.
- [21] WANG Z, BOVIK A C, SHEIKH H R, *et al.* Image quality assessment: from error visibility to structural similarity [J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612.
- [22] MITTAL A, MOORTHY A K, BOVIK A C. No-reference image quality assessment in the spatial domain [J]. *IEEE Transactions on Image Processing*, 2012, 21(12): 4695-4708.
- [23] MEI K F, JIANG A W, LI J C, *et al.* Progressive feature fusion network for realistic image dehazing [C]. 14th Asian Conference on Computer Vision. Perth: Springer, 2018: 203-215.
- [24] 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.
- [25] LU L P, XIONG Q, XU B R, *et al.* MixDehazeNet: mix structure block for image dehazing network [C]//*Proceedings of the 2024 International Joint Conference on Neural Networks (IJCNN)*. Yokohama: IEEE, 2024: 1-10.

作者简介:



张鑫琪,女,硕士研究生,2023年于辽宁科技大学获得学士学位,主要从事计算机视觉及深度学习的研究。E-mail: 2663458995@qq.com



田莹,女,博士,教授,2008年于沈阳工业大学获得博士学位,主要从事模式识别、数字图像处理等方面的研究。E-mail:astianying@126.com