

基于全局残差注意力和门控特征融合的 CNN-Transformer去雾算法

李海燕¹, 乔仁超¹, 李海江², 陈泉³

(1. 云南大学 信息学院, 云南 昆明 650050; 2. 云南省交通投资建设集团有限公司, 云南 昆明 650000;

3. 阿尔多瓦大学 人文学院, 法国 阿拉斯 62000)

摘要:为解决现有图像去雾算法因缺乏全局上下文信息、处理分布不均匀的雾时效果差且复用细节信息时引入噪声的缺陷,提出了基于全局残差注意力和门控特征融合的CNN-Transformer去雾算法.首先,引入全局残差注意力机制编码模块自适应地提取非均匀雾区的细节特征,设计跨维度通道空间注意力优化信息权重.然后,提出全局建模Transformer模块加深编码器的特征提取过程,设计带有并行卷积的Swin Transformer捕捉特征之间的依赖关系.最后,设计门控特征融合解码模块复用图像重建所需的纹理信息,滤除不相关的雾噪声,提高去雾性能.在4个公开数据集上进行定性和定量实验,实验结果表明:所提算法能够有效地处理非均匀雾区域,重建纹理细腻且语义丰富的高保真无雾图像,其峰值信噪比和结构相似性指数都优于经典对比算法.

关键词: 图像去雾;全局残差注意力机制;CNN-Transformer架构;门控特征融合;图像重建

中图分类号: TP 391 **文献标志码:** A **文章编号:** 1005-3026(2025)01-0026-09

CNN-Transformer Dehazing Algorithm Based on Global Residual Attention and Gated Feature Fusion

LI Hai-yan¹, QIAO Ren-chao¹, LI Hai-jiang², CHEN Quan³

(1. School of Information Science and Engineering, Yunnan University, Kunming 650050, China; 2. Yunnan Communications Investment and Construction Group Co., Ltd., Kunming 650000, China; 3. Faculty of Humanities, University of Artois, Arras 62000, France. Corresponding author: LI Hai-jiang, E-mail: li_cannie@163.com)

Abstract: To address the shortcomings of existing image dehazing algorithms, such as the lack of global contextual information, inadequate performance in dealing with non-uniform haze, and the introduction of noise during the reuse of detailed information, a CNN-Transformer dehazing algorithm based on global residual attention and gated feature fusion is proposed. Firstly, a global residual attention mechanism is introduced to adaptively extract the detailed features from non-uniform haze regions, and cross-dimensional channel-spatial attention is designed to optimize information weights. Thereafter, a global modelling Transformer module is proposed to deepen the feature extraction process of the encoder, and a Swin Transformer with parallel convolutions is constructed to capture the inter-feature dependencies. Finally, a gated feature fusion decoder module is designed to reuse the texture information required for image reconstruction, to filter out irrelevant haze noise, and thereby to improve dehazing performance. Qualitative and quantitative experiments conducted on four publicly available datasets indicate that the proposed algorithm can effectively handle non-uniform haze regions, reconstruct high-fidelity haze-free images with fine textures and rich semantics, and achieve higher peak signal-to-noise ratio and structural similarity index compared to the classic algorithm.

收稿日期: 2023-08-07

基金项目: 国家自然科学基金资助项目(62266049); 云南省万人计划“云岭教学名师”项目(20190101055); 2023年度云南省档案局科技项目(2023016607).

作者简介: 李海燕(1976—),女,云南红河人,云南大学教授,博士生导师.

Key words: image dehazing; global residual attention mechanism; CNN-Transformer architecture; gated feature fusion; image reconstruction

图像中的雾会降低图像的可见度,造成色彩失真、模糊和降低对比度.图像去雾的目标是从模糊的输入中恢复清晰的图像.现有图像去雾算法主要分为基于大气散射模型算法和基于深度学习的端对端算法两类.基于大气散射模型的算法^[1-4]利用物理先验估计透射图,然后使用大气散射模型反求解清晰图像,但是这些物理先验在复杂场景下的可靠性差,可能导致传输估计不准确以及雾残留.Ren等^[1]提出基于多尺度卷积神经网络的去雾算法,学习模糊图像与其对应透射图之间的映射关系,估计大气模型的参数.但是,在雾不均匀时,透射图估计的准确性低.为解决此缺陷,Zhang等^[5]提出用卷积神经网络(CNNs)估计透射率和全局大气光值,并引入生成对抗网络(generative adversarial network, GAN)^[6]区分真实和生成的无雾图像和透射率图.但是,微小的估计误差也会显著影响去雾后图像的重建效果.针对此问题,研究者提出采用端到端的深度学习网络直接将雾图像映射为清晰图像.Mei等^[7]提出了类似于U-net的渐进特征融合去雾网络,从观察到的雾图中恢复出无雾图像,克服了大气模型的局限性.Wu等^[8]迁移向量离散生成对抗网络(VQGAN)的高质量先验知识码本进入去雾网络.尽管这些算法可以避免由于复杂场景造成的参数估计误差,但它仍不能有效去除非均匀的雾.为了去除非均匀分布的雾,Qin等^[9]设计了一种特征融合注意力网络聚焦于通道和空间中相对重要的信息,显著提高了图像去雾效果.但是该方法忽略了通道-空间之间的相互作用,导致了跨维度信息丢失.Das等^[10]提出了一种快速深层多补丁分层网络,通过合并不同空间区域的多个图像补丁的特征来处理不均匀的雾图像.然而,由于卷积神经网络的感受野有限,它无法有效地捕捉特征之间的长距离依赖关系,同时缺乏全局上下文信息,因此重建效果差.为了学习模糊图像的全局背景特征,Wang等^[11]采用Transformer提取长距离的特征.Song等^[12]提出了一种轻量级的除雾网络,通过用映射填充取代变压器的多头自关注来减少参数量.但是该方法保留局部细节特征的能力较弱,导致重建后的图像细节模糊和粗糙.为解决此不足,Guo等^[13]研究了两个特征提取器的特性,并设计了结合CNN编码器和Transformer

编码器的方法,用于提取全局和细节信息.但是,该方法中采用的跳跃连接不能正确地融合编码器和解码器的特征.为了有效地复用编码器浅层的浅层特征,Wu等^[14]引入了自适应混合连接,动态地融合下采样和上采样层之间的特征.但是,该方法不能有效挖掘图像的高级特征,而仅能重复使用相似的低级特征.为了有效融合编码器浅层的特征,Romano等^[15]提出了一种增强算法,迭代应用增加一操作一减法修改解码器的特征图.Dong等^[16]证明了增加一操作一减法策略在图像去雾中的有效性.然而,这种网络在复用编码器特征时仍然引入雾噪声,导致图像去雾不完全.

综上所述,当前的图像去雾算法存在不能有效处理非均匀分布的雾及无法获取全局上下文信息的缺陷,在复用编码信息时,出现纹理细节丢失、与真实图像相似性不够、引入噪声等问题.针对这些问题,本文提出基于全局残差注意力和门控特征融合的CNN-Transformer去雾算法.

1 基于全局残差注意力和门控特征融合的CNN-Transformer去雾算法

1.1 整体去雾模型

如图1所示,本文提出的算法主要由卷积头、金字塔池化模块、全局残差注意力模块、Swin Transformer增强特征模块、门控增强融合模块以及卷积尾组成.首先,设计4层卷积编码器提取特征.在该过程中,采用金字塔池化^[17]模块提升维度,捕捉图像中不同尺度下的细节和语义信息.此外,设计全局残差注意力模块,有效处理不同像素信息,强调高度显著的信息,从浓雾中提取背景信息.然后,采用包含并行卷积的Swin Transformer模块增强CNN编码器输出的特征,更好地捕获全局上下文信息.最后,通过门控增强融合模块复用丰富的语义信息和高分辨率信息,重建清晰的背景图像.图1中 C, H, W 分别表示特征图的通道数、高和宽.

1.2 全局残差注意力模型

在特征编码过程中,若网络平等地处理所有通道信息和空间信息,则无法专注于重要的背景特征,例如被浓雾遮挡的物体.为了高效地提取特征信息,本文设计了全局残差注意力模块,有

效地结合了全局通道空间注意力和残差精细化单元,重点关注被遮挡物体的形状和纹理,如图2

所示.这样,去雾网络能够更好地捕捉到被浓雾遮挡物体的细节信息,实现更准确的去雾效果.

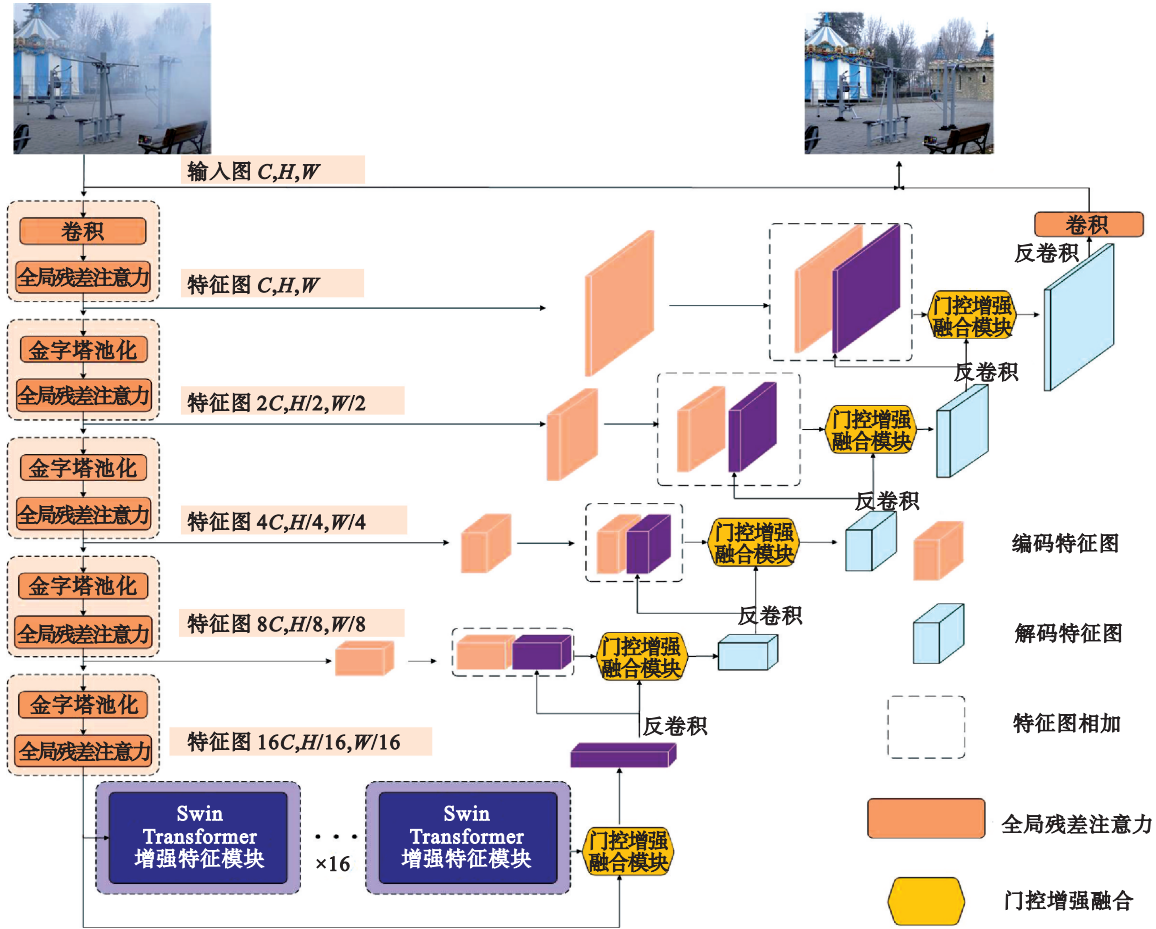


图1 本文提出算法的整体网络结构

Fig. 1 The overall network architecture of the proposed algorithm

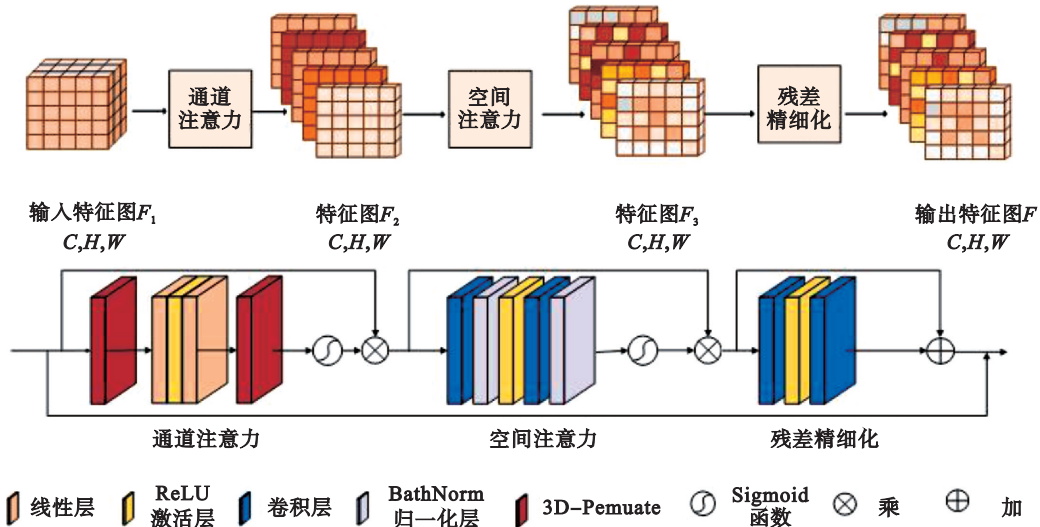


图2 全局残差注意力模型

Fig. 2 Global residual attention model

1. 2. 1 跨维度通道注意力

CBAM (convolutional block attention module)^[18]结合了通道注意力模块和空间注意力

模块,能动态地调整输入特征图的通道权重和空间权重,从而使模型更加关注图像中的重要特征.因此,本算法提出通道、高度和宽度3维排列

(3D-Permute)操作来保留 3 个维度的信息,以增强跨维度的信息交互.如图 2 所示,首先对输入的特征图进行 3D-Permute 操作,将 3 个维度进行排列,然后将其送入多层感知器(MLP)进行计算.在这个过程中,MLP 将特征图的通道维度压缩到原来的 $1/r$ 倍, r 为变量,然后再将其恢复到原始的通道数,并重新排列为原始的 3 个维度.通道注意力模块权重的计算可表示为

$$w_c = \sigma(\text{Permute}(\text{MLP}(\text{Permute}(F_1))))). \quad (1)$$

式中, σ 是 Sigmoid 函数.

输入特征图 F_1 与每个通道的权重 w_c 逐元素相乘,产生最终输出特征图 F_2 .与 CBAM 相比,本文算法保持了通道和空间的跨维度信息.通道注意力过程的数学表达式为

$$F_2 = F_1 \otimes w_c. \quad (2)$$

1.2.2 空间注意力

空间注意力模块专注于目标所在的空间位置上信息.首先,使用一个卷积层将通道注意力计算得到的特征图的通道数缩减至原通道数的 $1/r$,并进行归一化和激活函数处理.将处理后的特征图经过另一个卷积操作,使其恢复到初始特征的大小.最后,通过 Sigmoid 函数对特征进行处理.空间注意力模块权重的计算式表示为

$$w_s = \sigma(\delta(N_B(\text{Conv}_2(\delta(N_B(\text{Conv}_1(F_2))))))). \quad (3)$$

其中: δ 是 ReLU 激活函数; N_B 是 BatchNorm 归一化函数.

最后,输入特征图 F_2 和空间权重 w_s 相乘得到输出特征图 F_3 ,空间注意力过程的数学表达式为

$$F_3 = F_2 \otimes w_s. \quad (4)$$

1.2.3 残差精细化单元

残差精细化单元不太关注低频率或不太显著的信息,例如雾层,使网络架构更注重关注重要信息.残差精细化单元的计算,如式(5)所示,在空间注意力机制之后对特征图进行卷积激活再卷积,残差细化过程可用式(6)表示.

$$A_R = \text{Conv}_2(\delta(\text{Conv}_1(F_3))), \quad (5)$$

$$F = F_3 \oplus A_R + F_1. \quad (6)$$

式中, A_R 为 F_3 的精细化特征.

1.3 PC-SwinTransformer 层

Transformer 提取的特征通常较粗糙,而使用具有局部注意力和清晰细节的 CNN 特征可以弥补这一不足.此外,针对图像去雾任务,类似于 Trans-Unet 的结构非常适用.因此,提出了一种包含并行卷积的 Swin Transformer 作为编码器模块,捕捉全局特征的长期注意力,并与 CNN 获取的

特征细节相结合.如图 3a 所示,Swin Transformer 通过将图像分割为 A, B, C, D 子图像,并计算子图像的自注意力,通过滑动窗口促进子图像之间的交互,实现在整个特征图上进行自注意力计算,从而捕捉不同特征之间的远距离关系.

对于给定的输入特征, Swin Transformer 利用线性层和群组标记将其投影为 Q (query), K (key) 和 V (value).如图 3b 所示, CNN 的特征图 $F_c \in \mathbf{R}^{b \times H \times W \times C}$, 其中 b 表示 batch size. 首先通过线性层和滑动窗口划分的群组标记进行投影,得到 $Q \in \mathbf{R}^{b \times l \times d}$, $K \in \mathbf{R}^{b \times l \times d}$, $V \in \mathbf{R}^{b \times l \times d}$. 其中 l 表示窗口内的令牌数量, d 表示维度, 自注意力过程计算如下:

$$\text{Attention}(Q, K, V) = \text{Softmax}(QK^T / \sqrt{d} + B)V. \quad (7)$$

其中: B 表示相对位置偏置项; Softmax 为激活函数; K^T 是 K 的转置.

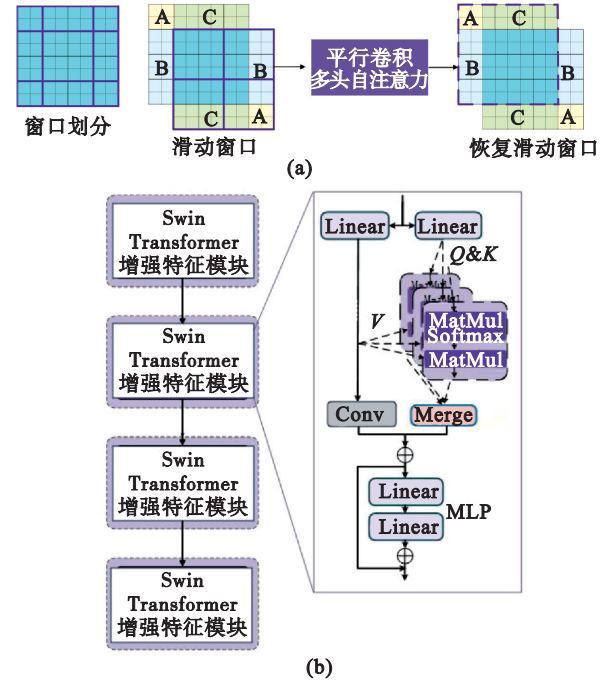


图 3 PC-Swin Transformer 结构

Fig. 3 The structure of PC-Swin Transformer block (a)—窗口滑动示意图; (b)—PC-Swin Transformer 结构.

乘法多头自注意力(MHSA)类似于低通滤波器^[19],其信息聚合权重会导致平滑效果,并且可能引入输入数据的偏差.每个头在计算注意力权重时会根据不同的位置、语义等因素赋予不同权重.当某些输入模式被多个注意力头共同关注并赋予权重时,这些模式的表示在聚合过程中会被放大,从而引入偏差.因此,本文算法引入并行卷积改进MHSA的空间信息聚合方式,提出空间信息聚合计算如下:

$$\text{Aggregation} = \text{Softmax}(\mathbf{Q}\mathbf{K}^T / \sqrt{d} + B)\mathbf{V} + \text{Conv}(\tilde{\mathbf{V}}). \quad (8)$$

其中: $\text{Conv}(\tilde{\mathbf{V}})$ 表示并行卷积; $\tilde{\mathbf{V}} \in \mathbf{R}^{b \times h \times w \times c}$ 表示窗口划分前的 \mathbf{V} .

最后,聚合的信息被输入到 MLP 进行进一步处理.为了确保网络的稳定性和深度,使用了局部残差连接.特征增强模块的计算如下:

$$\hat{F}_T = W - \text{MSA} - \text{PC}(N_L(F_C)) + F_C, \quad (9)$$

$$F_T = \text{MLP}(N_L(\hat{F}_T)) + \hat{F}_T. \quad (10)$$

式中: $W - \text{MSA} - \text{PC}$ 表示带有并行卷积的多头自注意力; N_L 表示 LayerNorm 归一化.

1.4 门控特征融合模型

根据 Dong 等^[16] 在“Strengthen-Operate-Subtract”增强策略的研究中所述,该增强算法可精细化输入图像特征,提高输出图像的峰值信噪比,其算法可以表示如下:

$$j^{n+1} = G_n(i^n + (j^n)_{\uparrow_2}) - (j^n)_{\uparrow_2}. \quad (11)$$

其中: j^n 表示第 n 次迭代的预测结果; $G_n(\cdot)$ 表示去雾操作; \uparrow_2 表示上采样; i^n 表示对雾图第 n 次编码.为了滤除多余的雾噪声并为图像重建提供多尺度和多层级的信息,本文算法将门控机制引入“Strengthen-Operate-Subtract”增强策略以合理融合编码特征和解码特征.如图 4 所示,通过 BatchNorm 归一化对 $i^n + (j^n)_{\uparrow_2}$ 进行处理,减少不同批次数据之间的内部协变量偏移,门控算法计算如下:

$$\left. \begin{aligned} x1 &= \text{Sigmoid}\left(\text{PWConv}_1(\text{BN}(i^n + (j^n)_{\uparrow_2}))\right), \\ x2 &= \text{DWConv}\left(\text{PWConv}_2(\text{BN}(i^n + (j^n)_{\uparrow_2}))\right); \end{aligned} \right\} \quad (12)$$

$$\text{PWConv}_3((x1 \odot x2) \oplus (i^n + (j^n)_{\uparrow_2})). \quad (13)$$

其中: PWConv 代表逐点卷积层; $\text{BN}(\cdot)$ 表示 BatchNorm 归一化; DWConv 代表深度卷积层.将 $x1$ 作为 $x2$ 门控信号,使用另一个逐点卷积层对结果进行投影,并将输出与恒等式 $i^n + (j^n)_{\uparrow_2}$ 相加.

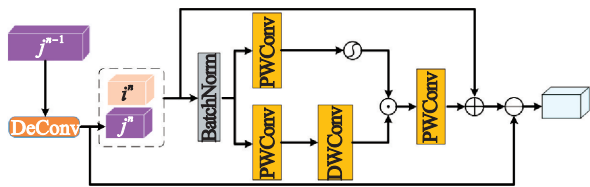


图 4 门控特征融合模块

Fig. 4 Gated feature fusion module

2 实验结果与分析

2.1 数据集及实验环境

实验使用 NH-Haze 2020^[20], Smoke-Haze

2022^[21], O-Haze 2018^[22] 和 Dense-Haze^[23] 4 个数据集验证算法的有效性. NH-Haze 数据集是第一个包含非均匀特征和逼真有雾场景的数据集,其中包括 55 个不同的户外场景.通过引入专业级的雾生成器模拟真实的雾气场景,使每个场景中的雾气分布不均匀.每张图像的尺寸为 1 200 像素 \times 1 600 像素. Smoke-Haze 数据集是 1 个真实的烟雾数据集,包括有雾图像和对应的真实图像,这些图像由固定位置的摄像机捕捉到.该数据集使用的烟雾生成设备与 O-Haze 数据集相同,但烟雾密度更高,雾气分布更不均匀.每张图像的尺寸为 640 像素 \times 512 像素. O-Haze 数据集包含了 45 个不同的户外场景,每张图像尺寸为 2 833 像素 \times 4 657 像素. Dense-Haze 数据集包含了 55 个真实的有雾图像,每幅图像都包含密集而均匀的雾气.该数据集中的图像是在多云天气的静态条件下拍摄的,拍摄时间为早晨或者日落时分.每张图像的尺寸为 1 200 像素 \times 1 600 像素.实验将数据集划分为训练集和测试集,其比例大致为 8:1.

实验运行于 Windows 10 系统上的单个 NVIDIA GeForce RTX 3090(24 GB)显卡上,使用 pycharm 平台,代码框架为 pytorch 1.8.1.采用 Adam 优化器,初始学习率设置为 1×10^{-4} ,并通过余弦退火策略进行调整,直至训练结束.采用结构相似性指数(SSIM)作为损失函数,约束网络输出与真实值之间的差异.在训练过程中,随机裁剪尺寸为 512 像素 \times 512 像素的雾图像块进行训练.为了验证本文算法的有效性,将其与 5 种基于深度神经网络的去雾方法 MSBDN^[16], FFANet^[9], AECR^[14], Dehazer^[13] 和 2023_ITBdehaze^[24] 算法进行对比.

2.2 实验一结果

在非均匀雾数据集(NH-Haze 和 Smoke-Haze)上的对比实验如图 5a 和图 6a 所示,可以清楚地观察到雾的分布呈现不均匀性,图中存在着不同浓度的雾气区域,一些区域具有密集的浓雾,而其他区域的雾相对较薄.图 5b~5f 展示了不同对比算法的去雾结果,在某些区域(箭头所指之处),部分雾气没有完全去除.图 5e 是 Dehazer 算法去雾的结果,雾气残留较少,但是绿色树木边缘的纹理细节不够清晰.图 5f 是 2023_ITBdehaze 算法去雾的结果,树木表面产生伪影,路灯周围有雾的残留.与此相比,图 5g 是本文算法的去雾效果,能够正确重建背景图像的纹理细

节,完全去除了噪声,去雾效果要优于对比算法.图 6b 展示了 MSBDN 算法去雾的结果,背景细节恢复较差,没有完全去除部分浓雾.虽然 Dehamer 算法去雾效果较好,但是红框处存在模糊.与其

他方法相比,本文算法能更干净地去除浓雾区域并恢复更多的细节,能有效处理非均匀雾情况,准确保留物体的颜色和丰富的细节.

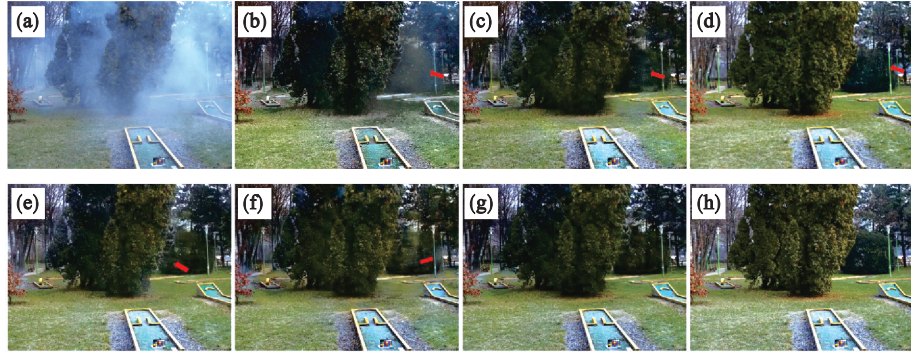


图 5 非均匀数据集 NH-Haze 不同算法去雾效果对比

Fig. 5 Comparison of dehazing effect of different algorithms on non-uniform data set NH-Haze

(a)—雾图; (b)—FFANet 算法; (c)—MSBDN 算法; (d)—AECR 算法; (e)—Dehamer 算法; (f)—2023_ITBdehaze 算法; (g)—本文算法; (h)—原图.

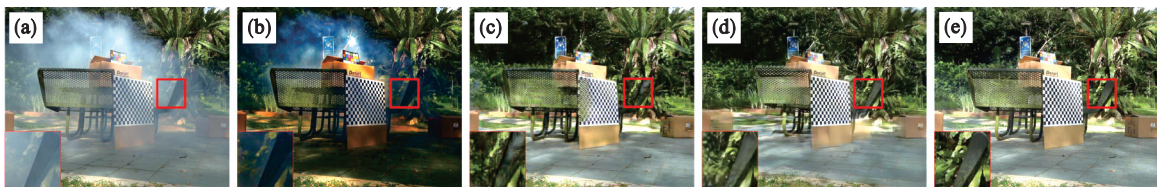


图 6 非均匀数据集 Smoke-Haze 不同算法去雾效果对比

Fig. 6 Comparison of dehazing effect of different algorithms on non-uniform data set Smoke-Haze

(a)—雾图; (b)—MSBDN 算法; (c)—Dehamer 算法; (d)—本文算法; (e)—原图.

2.3 实验二结果

为了进一步验证本文算法的有效性,在均匀薄雾图像数据集(O-Haze)和均匀厚雾图像数据集(Dense-Haze)上进行对比实验.如图 7a 和图 8a 所示,雾的分布相对均匀,MSBDN 算法的去雾结果在框内包含黑雾,且细节模糊不清. Dehamer 算法在天空区域呈现出灰暗效果.与其他比较算

法相比,本文算法在处理高清图片时去雾更彻底,细节更清晰.在图 8 中,只有 Dehamer 算法、2023_ITBdehaze 算法以及本文算法能恢复出密集雾气环境下的图像底层结构,但是对比算法在框内水果的色彩有失真.相比之下,本文算法能有效消除均匀雾气,并恢复接近真实图像的颜色.



图 7 均匀数据集 O-Haze 不同算法去雾效果对比

Fig. 7 Comparison of dehazing effect of different algorithms on uniform data set O-Haze

(a)—雾图; (b)—MSBDN 算法; (c)—Dehamer 算法; (d)—本文算法; (e)—原图.

2.4 定量评价

为了客观评价本文算法的图像去雾效果,使用峰值信噪比(PSNR)度量去雾图像与真实图像之间的差别,其值越高表示去雾效果越好,使用

SSIM 对比图像在结构、亮度以及对比度上的差别.表 1 和表 2 分别表示对比算法和本文算法在非均匀数据集和均匀数据集的去雾性能指标.从表中可看出,本文算法的 PSNR 和 SSIM 指标均优

于对比算法.其中,非均匀数据集的NH-Haze和Smoke-Haze去雾数据表明,本文算法较MSBDN算法的PSNR提高了2.38,4.04 dB,SSIM提高了5.77%,18.87%.均匀数据集O-Haze和Dense-Haze去雾数据表明,本文算法较MSBDN算法的

PSNR分别提高了1.56,1.59 dB;SSIM分别提高了6.49%,10.90%.评价结果表明本文算法在去除雾噪声、还原背景全局结构和局部细节方面均优于对比算法.

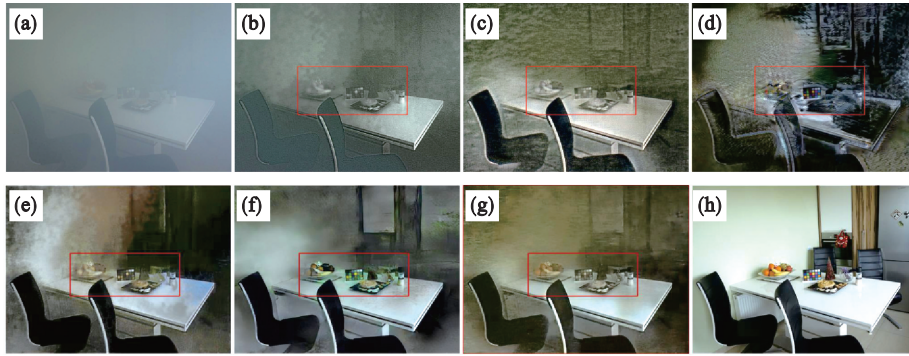


图8 均匀数据集Dense-Haze不同算法去雾效果对比

Fig. 8 Comparison of dehazing effect of different algorithms on uniform data set Dense-Haze

(a)—雾图; (b)—FFANet算法; (c)—MSBDN算法; (d)—AECR算法; (e)—Dehamer算法; (f)—2023_ITBdehaze算法; (g)—本文算法; (h)—原图.

表1 非均匀去雾数据集的定量结果对比
Table 1 Quantitative comparison of non-uniform dehazing datasets

算法	NH-Haze		Smoke-Haze	
	PSNR/dB	SSIM	PSNR/dB	SSIM
FFANet	18.13	0.6473	15.20	0.53
MSBDN	17.97	0.6591	15.19	0.53
AECR	19.24	0.5962	16.57	0.58
Dehamer	20.23	0.6844	18.83	0.62
2023_ITBdehaze	20.31	0.6260	19.01	0.63
本文算法	20.35	0.6971	19.23	0.63

表2 均匀去雾数据集的定量结果对比
Table 2 Quantitative comparison of uniform dehazing datasets

算法	O-Haze		Dense-Haze	
	PSNR/dB	SSIM	PSNR/dB	SSIM
FFANet	22.12	0.77	12.22	0.44
MSBDN	24.36	0.77	15.13	0.55
AECR	22.90	0.72	15.35	0.52
Dehamer	24.61	0.75	16.62	0.56
2023_ITBdehaze	25.84	0.78	16.31	0.56
本文算法	25.92	0.82	16.72	0.61

2.5 特征可视化

本文算法是以MSBDN去雾算法为基础模型,添加全局残差注意力模块、PC-Swin Transformer增强模块及门控特征融合模块得到.

为了深入研究本文算法每个模块的具体作用,对特征图进行可视化操作.图9b,9c是经过全局残差注意力模块处理前和处理后编码器特征的可视化结果,经注意力处理后,能更有效地识别和定位浓厚的雾霾区域.图9d,9e是经过PC-Swin Transformer特征增强模块处理前和处理后的特征可视化结果,特征增强后网络模型编码的特征包含全局上下文信息,因此本文算法能够更好地理解图像内容.图9f,9g是有和无门控特征融合解码器的输出特征可视化结果,带有该模块的解码器能捕捉到更多的重构细节,因此本文算法能够很好地处理分布不均匀的雾,获取足够的全局上下文信息重建美观的背景图像.

2.6 消融实验

为了验证本文算法每个模块的有效性,在NH-Haze数据集上对本文算法移除不同模块后进行评估,为保证实验公平,所有模型均采用相同设置和训练策略.研究结果如图10所示,定量评价结果如表3所示.

由图10可看出,与本文算法相比,去除注意力模块后红框内存在大面积雾层残留,去除增强模块后红框内的浓雾未能完全去除,去除融合模块后红框内的细节显得不太自然,无法生成精细的结构.相对而言,本文算法去雾性能好,保持了颜色和纹理的一致性,并且还还原了更多的细节信息.表3是消融实验的定量对比结果,将全局残差

注意力模型、PC-Swin Transformer 特征增强模块以及门控特征融合模块叠加到基础模型后,图像去雾的各项指标均得到改善. PSNR 提升了 13.24%, SSIM 提升了 5.77%, 表明全局残差注意

力模块、PC-Swin Transformer 增强模块以及门控特征融合模块优良的去噪性能,并显著改善了纹理和结构的重建.

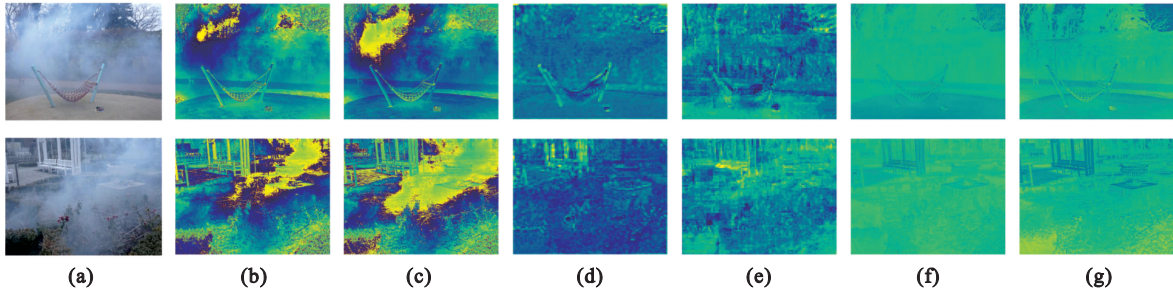


图 9 特征可视化结果

Fig. 9 Visual result of features

(a)一雾图; (b)一无全局残差注意力; (c)一有全局残差注意力; (d)一无 PC-Swin Transformer; (e)一有 PC-Swin Transformer; (f)一无门控特征融合; (g)一有门控特征融合.

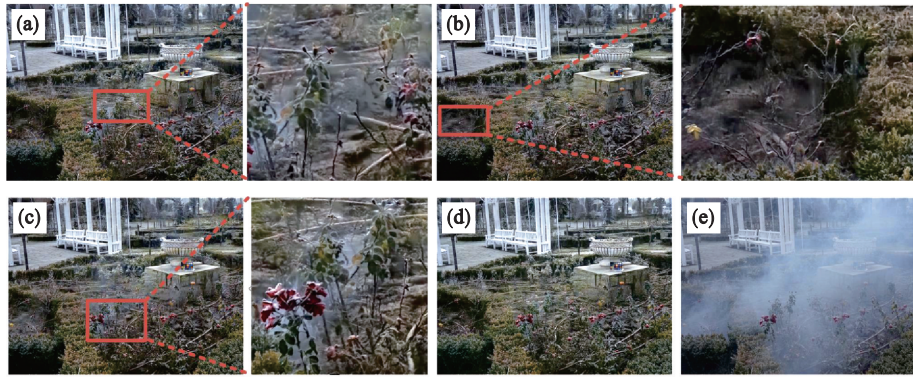


图 10 消融实验定性研究

Fig. 10 Qualitative study of ablation experiment

(a)一去除全局残差注意力模块; (b)一去除门控特征融合模块; (c)一去除 PC-Swin Transformer 增强模块; (d)一本文算法; (e)一雾图.

表 3 消融实验定量评估

Table 3 Quantitative evaluation of ablation experiments

去雾算法	PSNR/dB	SSIM
MSBDN 基础模型	17.97	0.659 1
去除全局残差注意力模块	18.84	0.667 3
去除 PC-Swin Transformer 增强模块	20.01	0.686 1
去除门控特征融合模块	19.85	0.679 5
本文算法	20.35	0.697 1

3 结 语

本文提出了一种基于 CNN-Transformer 架构结合全局残差注意力和门控特征融合的图像去雾算法. 设计了一种新颖的全局残差注意力模块, 能自适应地编码不同雾霾浓度区域的特征,

提高了网络的特征表达能力. 提出了 PC-Swin Transformer 特征增强模块增强图像编码能力, 减少全局上下文信息丢失. 采用门控特征融合模块, 合理地重复使用浅层卷积层的特征, 提高去雾率和重建效果的纹理清晰度. 在 2 类公共去雾数据集进行对比实验. 实验结果显示, 本文提出的算法在 4 个公开数据集上相较于基准方法 MSBDN, 峰值信噪比 (PSNR) 提升了最低 1.56 dB, 最高 4.04 dB; 结构相似性指数 (SSIM) 提升了最低 5.77%, 最高 10.90%. 这表明本文算法能够有效处理不同场景的雾噪声, 并成功重建细节丰富的背景图像.

参考文献:

[1] Ren W Q, Liu S, Zhang H, et al. Single image dehazing via multi-scale convolutional neural networks [M]//Lecture Notes in Computer Science. Cham: Springer International

- Publishing, 2016: 154–169.
- [2] McCartney E J, Hall F F. Optics of the atmosphere: scattering by molecules and particles [J]. *IEEE Journal of Quantum Electronics*, 1977, 30(5): 76–77.
- [3] Narasimhan S G, Nayar S K. Chromatic framework for vision in bad weather [C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2000). Hilton Head Island, 2000: 598–605.
- [4] Narasimhan S G, Nayar S K. Vision and the atmosphere [J]. *International Journal of Computer Vision*, 2002, 48: 233–254.
- [5] Zhang H, Patel V M. Densely connected pyramid dehazing network [C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, 2018: 3194–3203.
- [6] Creswell A, White T, Dumoulin V, et al. Generative adversarial networks: an overview [J]. *IEEE Signal Processing Magazine*, 2018, 35(1): 53–65.
- [7] Mei K F, Jiang A W, Li J C, et al. Progressive feature fusion network for realistic image dehazing [M]// Jawahar C V, Li H D, Mori G, et al. *Lecture Notes in Computer Science*. Cham: Springer International Publishing, 2019: 203–215.
- [8] Wu R X, Duan Z P, Guo C L, et al. RIDCP: revitalizing real image dehazing via high-quality codebook priors [C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Paris: IEEE, 2023: 22282–22291.
- [9] 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.
- [10] Das S D, Dutta S. Fast deep multi-patch hierarchical network for nonhomogeneous image dehazing [C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Seattle, 2020: 482–483.
- [11] Wang Z D, Cun X D, Bao J M, et al. Uformer: a general U-shaped transformer for image restoration [C]// 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, 2022: 17683–17693.
- [12] 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.
- [13] Guo C L, Yan Q X, Anwar S, et al. Image dehazing transformer with transmission-aware 3D position embedding [C]// 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, 2022: 5812–5820.
- [14] Wu H Y, Qu Y Y, Lin S H, et al. Contrastive learning for compact single image dehazing [C]// 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, 2021: 10551–10560.
- [15] Romano Y, Elad M. Boosting of image denoising algorithms [J]. *SIAM Journal on Imaging Sciences*, 2015, 8(2): 1187–1219.
- [16] 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 (CVPR), Seattle, 2020: 2157–2167.
- [17] Zhao H S, Shi J P, Qi X J, et al. Pyramid scene parsing network [C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017). Honolulu, 2017: 2881–2890.
- [18] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, 2018: 7132–7141.
- [19] Chen J N, Lu Y Y, Yu Q H, et al. TransUNet: transformers make strong encoders for medical image segmentation [EB/OL]. (2021–02–08) [2023–05–15]. <https://arxiv.org/abs/2102.04306v1> 2023, 12, 01.
- [20] Ancuti C O, Ancuti C, Timofte R. NH-HAZE: an image dehazing benchmark with non-homogeneous hazy and haze-free images [C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Seattle, 2020: 444–445.
- [21] Jin Y Y, Yan W D, Yang W H, et al. Structure representation network and uncertainty feedback learning for dense non-uniform fog removal [C]// ACCV 2022: 16th Asian Conference on Computer Vision. Macao, 2022: 155–172.
- [22] Ancuti C O, Ancuti C, Timofte R, et al. O-Haze: a dehazing benchmark with real hazy and haze-free outdoor images [C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Salt Lake City, 2018: 754–762.
- [23] Ancuti C O, Ancuti C, Sbert M, et al. Dense-Haze: a benchmark for image dehazing with dense-haze and haze-free images [C]// 2019 IEEE International Conference on Image Processing (ICIP). Taipei, 2019: 1014–1018.
- [24] Liu Y Y, Liu H, Li L Y, et al. A data-centric solution to nonhomogeneous dehazing via vision transformer [C]// 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Vancouver, 2023: 1406–1415.