

DOI: 10.13382/j.jemi.2508211

基于误差反馈和雾霾感知的图像去雾*

高雯婷 廖苗 荣佳丽 郭娟秀

(湖南科技大学计算机科学与工程学院 湘潭 411100)

摘要: 雾霾环境下拍摄的图像通常受到对比度降低、细节退化或色彩失真等多种因素的干扰,严重影响视觉体验和后续高级视觉任务的准确性。为了有效去除图像中的雾霾,提出了一种基于误差反馈的多尺度密集残差去雾网络(multi-scale dense residual dehazing network, MDRD-Net)。该网络在编、解码路径中对称引入误差反馈模块(error feedback module, EFM),以补偿下采样导致的信息丢失,并在误差反馈模块之间引入密集连接,加强非邻层级间的信息交互。为了让网络重点关注图像中的浓雾和细节区域,网络在特征提取阶段级联了多个雾霾感知模块(haze aware module, HAM)。此外,网络在跳跃连接中引入注意力机制,实现对编、解码端特征的自适应融合,克服深、浅层特征之间存在的语义鸿沟。在RESIDE公开数据集上进行了广泛实验,结果表明,提出方法可有效去除图像中的雾霾干扰,获得色彩真实、对比度高、细节丰富的高质量无雾图像,在定量和定性分析中均显著优于多种现有先进方法。

关键词: 图像去雾;误差反馈;雾霾感知;注意力机制

中图分类号: TP39;TN911.73 **文献标识码:** A **国家标准学科分类代码:** 510.40

Image dehazing based on error feedback and haze aware

Gao Wenting Liao Miao Rong Jiali Guo Juanxiu

(School of Computer Science and Engineering, Hunan University of Science and Technology, Xiangtan 411100, China)

Abstract: Images captured in haze are often affected by contrast reduction, detail degradation, or color distortion, which significantly impair visual quality and affect the performance of high-level vision tasks. To effectively remove the haze from images, a multi-scale dense residual dehazing network (MDRD-Net) based on error feedback is proposed. In this network, error feedback modules (EFM) are symmetrically introduced in the encoding and decoding paths to compensate for the information loss caused by downsampling. Dense connections are introduced between error feedback modules to enhance information interaction between non-adjacent layers. To make the network focus on regions with thick haze and rich details, multiple haze aware modules (HAM) are cascaded in the feature extraction stage. Additionally, an attention mechanism is introduced in the skip connections to adaptively fuse the features from the encoder and decoder to overcome the semantic gap between deep and shallow features. Extensive experiments on the RESIDE public dataset demonstrate that the proposed method can effectively remove the haze interference and obtain clear images with true colors, high contrast, and rich details. The results, both quantitatively and qualitatively, show a significant improvement over those of many existing state-of-the-art methods.

Keywords: image dehazing; error feedback; haze aware; attention mechanism

0 引言

雾霾环境下,大气中的雾、霾等悬浮颗粒会对可见光

产生散射和吸收作用。该气候条件下采集拍摄的图像普遍存在色彩失真、对比度降低和细节退化等问题,不仅影响人眼的视觉体验,也会降低目标检测^[1]、图像识别^[2]、语义分割^[3]等高级视觉任务的性能。

收稿日期: 2025-03-04 Received Date: 2025-03-04

* 基金项目: 湖南省自然科学基金(2025JJ50363)、湖南省教育厅科研项目(24A0356)、湖南省科技创新计划(2024RC3216)项目资助

为了去除雾霾、提高图像质量,国内外学者们提出了多种去雾算法。根据去雾原理,现有去雾方法主要分为以下 3 类:基于图像增强的去雾方法、基于物理模型的去雾方法和基于深度学习的去雾方法。基于图像增强的去雾方法主要通过直接修改图像对比度或饱和度增强图像中的有用信息,达到提升图像质量、去除雾霾的目的。常用的增强方法有直方图均衡化^[4]、同态滤波^[5]、小波变换^[6]、Retinex 算法^[7]等。由于上述方法未考虑图像降质的原因,仅依靠增强操作改善图像视觉效果,往往会造成去雾后图像失真、去雾效果欠佳等问题。

基于物理模型的去雾方法通常以大气散射模型为基础,通过估计雾图成像过程中的传输率和大气光,实现图像去雾。Ling 等^[8]提出了一种基于饱和度线性先验的去雾方法,通过挖掘局部区域内共享相同表面反射率像素的饱和度与亮度倒数的线性关系,精确估计透射率并实现有效的图像去雾。Zhu 等^[9]提出了一种雾图成像的颜色衰减先验,并构建基于该理论的线性模型,通过监督学习的方式获取线性模型中的参数。Berman 等^[10]利用雾线先验估计大气光和每个像素的传输率,同时通过正则化对传输率进行优化。尽管基于物理模型的去雾方法不需要复杂的训练过程,具有较高的实时性,但该方法依赖于先验知识,且需要人工提取雾图中的特征信息,鲁棒性较差。若对雾图成像过程中的传输率和大气光估计出现偏差,该方法将造成去雾不彻底、细节丢失、图像偏色等问题。

近年来,随着 GPU 算力的提高及大型去雾数据集^[11-12]的建立,基于深度学习的图像去雾已成为当前的研究热点。Cai 等^[13]首次采用卷积神经网络预测雾图的传输率。该方法将大气光视为常量,并利用预测的传输率,根据大气散射模型恢复清晰无雾图。Li 等^[14]重构了大气散射模型,将透射率和大气光合并为一个参数,并设计轻量级去雾网络(all-in-one dehazing network, AOD-Net)预测该参数。与文献^[13-14]方法不同,Chen 等^[15]提出了一种门控融合网络(gated context aggregation, GCA)。该网络不依赖先验知识,通过运用平滑膨胀卷积和门控子系统实现图像端到端去雾,可有效解决去雾后图像的网络伪影问题。Qin 等^[16]通过连续堆叠多个不同尺度的注意力块,构建了一种特征融合注意力去雾网络(feature fusion attention network, FFA-Net)。该网络需要对输入图像进行大量级联卷积,这将导致模型计算复杂度、推理速度慢。Cui 等^[17]提出了一种多分支动态选择频率模块,将特征图动态解耦为不同的频率分量,并选择信息最丰富的分量来恢复图像。目前大部分基于深度学习的去雾方法仅通过加深或加宽网络提升去雾性能,该做法将极大提升网络的参数量和计算成本。此外,为了获取多尺度特征,网络通常在编码阶段引入下采样

操作,然而该步骤将不可避免造成细节信息丢失,导致去雾后的图像纹理细节模糊。

针对上述问题,本文提出了一种基于误差反馈的多尺度密集残差去雾网络(multi-scale dense residual dehazing network, MDRD-Net)。该网络在编-解码器两端对称引入误差反馈模块(error feedback module, EFM),并在特征提取阶段设计了一种雾霾感知模块(haze aware module, HAM),使网络能够重点关注浓雾及细节区域。此外,在跳跃连接层引入注意力机制,实现编码-解码特征信息的自适应融合。

本文提出了一种误差反馈模块,并将其对称部署于编-解码端,可补偿下采样丢失的空间和细节信息,同时促进网络中非邻层级之间的特征交互。设计了一种基于密集残差和通道注意力的雾霾感知模块,利用密集残差结构提取多尺度细节特征,并通过通道注意力机制突出浓雾及细节区域,抑制无关信息,从而增强网络的特征表达能力。提出了带注意力机制的跳跃连接(skip connection with attention mechanism, SCAM),自适应地融合编-解码路径中的深、浅层特征。

1 本文方法

基于误差反馈的多尺度密集残差去雾网络整体结构如图 1 所示。该网络采用 U 形编码-解码框架,从左到右包含编码器(encoder)、特征提取模块(feature extraction)和解码器(decoder)3 部分。编码器由卷积层、下采样和误差反馈模块组成。首先,采用 3×3 卷积将有雾图像映射至高维空间。为了增大网络感受野、降低网络计算量,对特征图进行两次下采样操作,并在每次下采样后部署一个误差反馈模块。误差反馈模块可有效捕获下采样过程中丢失的信息,并将这些信息融合至更深层网络中,弥补由下采样引起的空间信息缺失。该误差反馈不仅可捕获相邻层级之间的下采样误差,还通过密集连接的方式,增加非相邻层级之间信息的交互和融合。随后,在特征提取阶段级联多个雾霾感知模块,使得网络能在下采样后的小尺度特征空间中进一步深入、全面提取图像特征。解码端与编码端具有对称结构,由卷积层、误差反馈模块、上采样组成。上采样操作用于将特征图恢复至原始图像分辨率,最后,采用 3×3 卷积层用于重建清晰无雾图像。此外,为了自适应融合编解码端特征,在跳跃连接中引入注意力机制。该操作不仅可避免直接拼接造成信息冗余,同时可缓解因网络加深导致的浅层信息衰退问题。

1.1 误差反馈模块

U-net 架构中,为了扩大网络感受野、获取不同尺度特征信息,在编码端引入了下采样操作。然而,下采样过

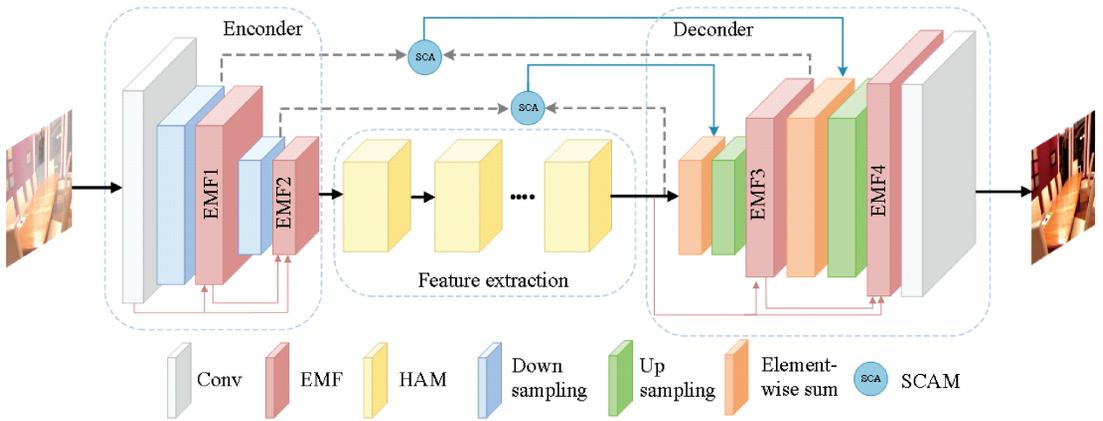


图 1 基于误差反馈的多尺度密集残差去雾网络

Fig. 1 Multi-scale dense residual dehazing network based on error feedback

程将不可避免地造成特征图结构细节信息和上下文信息的丢失。此外,U-net 网络编码和解码端的各网络层级均以级联方式相连,非相邻层级间缺乏足够的连接。针对上述问题,本文提出了基于残差连接的 EFM 模块,结构如图 2 所示。在 MDRD-Net 网络中,误差反馈模块对称部署于编解码路径,对下采样导致的信息丢失进行补偿,同时对不同层级特征进行融合,加强非相邻层级之间的特征联系。由图 1 可见,在编码和解码端均引入两个 EFM,分别位于下采样和上采样之后。编码端中的第 1 个 EFM 涉及两个输入和一个输出,结构如图 2(a)所示,具体操作如下:

$$E_{out}^1 = x^1 + Down(Up(x^1) - x^2) \quad (1)$$

式中: x^1 表示第 1 次下采样后的特征图; x^2 表示第 1 个卷积层的输出; Up 和 $Down$ 分别表示上、下采样,分别采用步长为 2 的反卷积和卷积实现。由于密集连接的引

入,编码端的第 2 个 EFM 涉及 3 个输入和 1 个输出,结构如图 2(b)所示,具体操作如下:

$$e_x^1 = Down(Down(Up(Up(x^3)) - x^2)) + x^3 \quad (2)$$

$$E_{out}^2 = (Down(Up(e_x^1)) - E_{out}^1) + e_x^1 \quad (3)$$

式中: x^3 表示编码端第 2 次下采样后的特征图。在误差反馈模块采用密集连接,可促进网络中非邻层级间的交互,同时减少特征传递导致的信息损失。

解码端与编码端的误差反馈互为逆操作,解码端误差反馈模型结构如图 2(c)和(d)所示,其中 y^1 为解码端第 1 次上采样后的特征图, y^2 为特征提取阶段的输出特征图, y^3 为解码端第 2 次上采样后的特征图。在解码端中,EFM 通过融合不同尺度之间的特征信息,改善网络深浅层之间信息流的传递方式,使网络在解码过程中充分利用不同尺度下的图像全局和细节特征,提升网络的去雾能力。

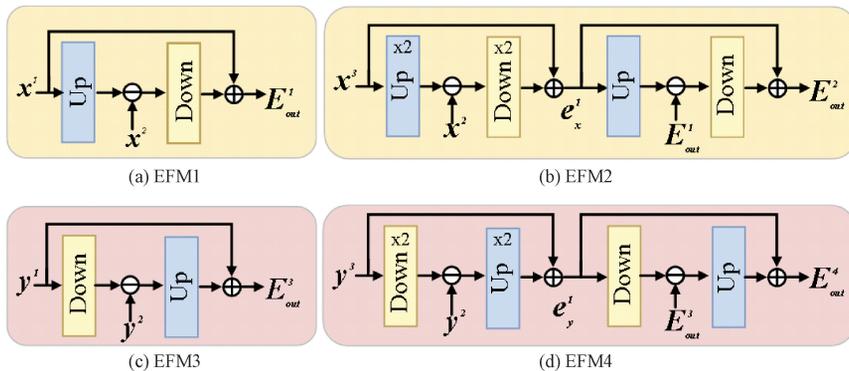


图 2 编解码端误差反馈模块具体结构

Fig. 2 Structure of error feedback modules in encoder and decoder

1.2 雾霾感知模块

现实世界中,雾霾的分布通常是复杂且不均匀的。若对整幅图像进行无差别去雾,则可能导致浓雾区域去

雾不彻底,薄雾区域结构信息丢失等问题。为了聚焦浓雾及目标区域,本文设计了一种 HAM 模块,结构如图 3 所示。HAM 由密集残差块(residual dense block,RDB)和

通道注意力模块 (channel attention, CA) 组成。相较于简单堆叠的卷积层, RDB 能更好地提取图像的局部和细节特征。CA 则能对 RDB 所提取到的特征图进行通道加权, 自适应增强重要通道特征, 同时对无关信息进行抑制。此外, 将 CA 的输出与原始输入以通道相加的方式相融合, 以缓解随着网络加深导致的梯度消失问题。

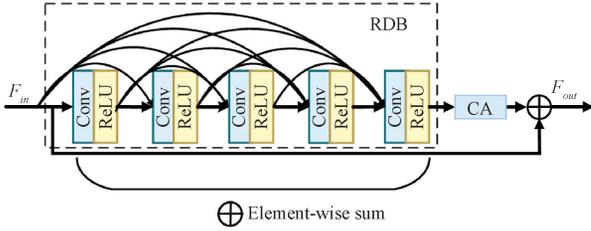


图 3 雾霾感知模块

Fig. 3 Haze aware module

由图 3 可见, RDB 共包含 5 个卷积层, 前 4 个级联卷积层的核大小为 3×3 , 通过引入密集残差连接, 充分提取图像的上下文特征, 最后 1×1 卷积层用于融合前四层所提取的特征图并调整输出通道数。

CA 通过赋予特征图不同通道不同权值, 促使网络重点关注雾图中的浓雾和细节丰富区域, 其结构如图 4 所示。首先对输入特征 $F \in R^{H \times W \times C}$ 进行全局平均池化, 采用特征降维的方式, 将空间信息转化为通道描述符 $G \in R^{1 \times 1 \times C}$:

$$G_i = D_p(F_i) = \frac{1}{H \times W} \sum_{x=1}^H \sum_{y=1}^W F_i(x, y) \quad i = 1, \dots, C \quad (4)$$

式中: $F_i(x, y)$ 表示位于输入特征 F 第 i 个通道上坐标为 (x, y) 的像素值; C 为特征图的总通道数; D_p 表示全局平均池化操作; H 和 W 分别表示特征图的高度和宽度。接着, 对 G_i 进行如下操作, 获取各通道的注意力权重:

$$\partial_i = \sigma(\text{Conv}(\delta(\text{Conv}(G_i)))) \quad i = 1, \dots, C \quad (5)$$

式中: Conv 表示大小为 3×3 的卷积层; δ 表 ReLU 示激活函数; σ 表示 sigmoid 激活函数。最后, 利用广播机制, 将注意力权重与输入特征进行点乘, 获取通道注意力结果:

$$F_i^* = \partial_i \otimes F_i \quad i = 1, \dots, C \quad (6)$$

式中: F_i 表示输入特征图的第 i 个通道; ∂_i 表示第 i 个通道的注意力权重; \otimes 表示元素相乘。CA 模块可以有效增强重要通道的特征信息, 同时对无关信息进行抑制。

1.3 带注意力机制的跳跃连接

网络浅层可提取图像中的底层信息, 如纹理、边缘等, 深层结构则可捕获图像的高级语义信息, 如形状、轮廓等。为了融合网络各层级所提取到的特征, 最直接的方式是在编-解码端引入跳跃连接, 对不同层级特征采取

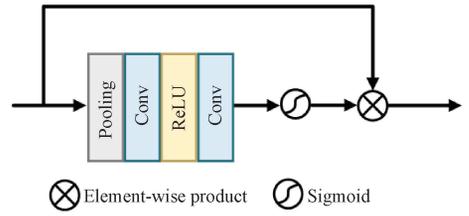


图 4 通道注意力机制

Fig. 4 Channel attention mechanism

相加或拼接的方式进行融合。该做法虽然能起到一定的特征融合作用, 但同时也带来了部分特征表达不充分、信息冗余等问题。针对上述问题, 本文设计了 SCAM 模块, 自适应融合不同尺度特征, 并缓解因网络加深导致的信息衰退问题。该模块结构如图 5 所示, 其输出可由式 (7) 获取。

$$F^{out} = w \times F_{en} + (1 - w) \times F_{de} \quad (7)$$

式中: F_{en} 和 F_{de} 分别表示编码和解码端特征; w 为可学习权重。为了自适应获取权重 w , 首先将特征 F_{en} 和 F_{de} 进行拼接, 随后运用 3×3 卷积层对其进行特征提取, 并采用图 3 中 CA 模块获取注意力矩阵, 最后经过 sigmoid 函数得到取值为 $0 \sim 1$ 的权重矩阵 w 。

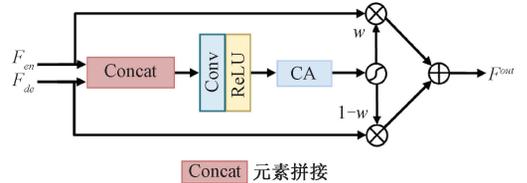


图 5 带注意力机制的跳跃连接

Fig. 5 Skip connection with attention mechanism

1.4 损失函数

本文采用基于平滑损失 L_s 和感知损失 L_p 的混合损失函数进行网络训练。平滑损失 L_s 用于衡量网络预测的去雾图像与真实清晰图像之间的像素级差异, 计算公式如下:

$$L_s = \frac{1}{N} \sum_{i=1}^N \sum_{c \in (R, G, B)} H_s(G_c(i) - J_c(i)) \quad (8)$$

$$H_s(e) = \begin{cases} 0.5e^2, & |e| < 1 \\ |e - 0.5|, & \text{其他} \end{cases} \quad (9)$$

式中: G 为网络预测的无雾图像; J 表示真实的清晰无雾图像; $G_c(i)$ 和 $J_c(i)$ 分别表示图像 G 和 J 中第 i 个像素点在通道 c 上的值; N 表示图像的像素数目。

与平滑损失 L_s 不同, 感知损失^[18]用于衡量图像的高级语义特征差异。该损失指标通常需要利用预训练好的深度神经网络提取预测图像和真实清晰图像不同尺度的特征图, 并运用 L2 范式量化两者之间的高级语义差异。

鉴于 ImageNet 上预训练好的 VGG-16^[19]能较好地提取图像高级语义特征,本文将网络预测的无雾图像和其对应的真实清晰图像同时输入至 VGG-16 网络中,提取其位于 VGG-16 网络中第 i 层的特征图,并通过计算其对应卷积层特征图的差值,获取去雾图像的感知损失。在模型训练阶段加入感知损失可以促使网络预测的无雾图像具有与清晰图像一致的高级语义信息,使得恢复的无雾图像更加符合人眼感知系统。感知损失具体计算公式如下:

$$L_p = \sum_i \|\phi_i(\hat{J}) - \phi_i(J)\|_2^2 \quad i \in \{3, 8, 15\} \quad (10)$$

式中: \hat{J} 表示网络预测的无雾图像; J 表示真实的清晰无雾图像; $\phi_i(\hat{J})$ 和 $\phi_i(J)$ 分别表示将 \hat{J} 和 J 输入 VGG-16 得到的第 i 层特征图; C_i 、 H_i 和 W_i 分别表示第 i 层特征图的通道数、高度和宽度。

综合考虑图像底层像素级和高层语义级损失,本文采用平滑和感知损失的加权和作为网络的总损失:

$$L_p = L_s + \lambda L_p \quad (11)$$

式中: λ 为权重系数。

2 实验结果分析

2.1 实验数据集

本文采用 RESIDE^[20]公开数据集中的室内合成训练集 (Indoor training set, ITS)、户外合成训练集 (Outdoor training set, OTS)、SOTS (Synthetic objective testing set) 进行网络训练与测试。ITS 为室内合成训练集包含了 1 399 幅室内清晰图像和其对应的 13 990 幅合成有雾图像。OTS 包含了 8 477 幅户外清晰图像和其对应的 296 695 幅合成雾图。本文采用 ITS 和 OTS 作为训练集。SOTS 为包含 500 幅室内和 500 幅户外雾图的综合数据集,作为本文测试集。

2.2 实验设置

本文方法基于 Pytorch 框架,在配备了 NVIDIA GeForce RTX 3060 GPU,内存 32 GB 和 Intel (R) Core (TM) i5-11400F 处理器的硬件平台上进行实验。在训练期间,使用默认动量 $\beta_1 = 0.9$ 和 $\beta_2 = 0.999$ 的 ADAM 优化器,所有网络层的初始学习率设置为 1×10^{-3} ,批次大小为 28。为了平衡感知损失和平滑损失,将 λ 设置为 0.4。

2.3 评价指标

本文采用峰值信噪比^[21] (peak signal to noise ratio, PSNR) 和结构相似度^[22] (structural similarity Index, SSIM) 对去雾图像进行定量评价。PSNR 是一种广泛使用的全参考图像质量评估标准,通过比较去雾图像与原始无雾

图像之间的像素级差异,量化图像失真程度。PSNR 值越小,表示图像失真越严重,即去雾图像的质量越差。SSIM 则是一种基于人眼视觉特征的图像质量评价指标,其综合考虑了图像在亮度、对比度和结构 3 个方面的相似性。SSIM 的取值范围在 0~1 之间,值越大,表明去雾后图像的特征信息保留程度越好,与原始无雾图越相似。

2.4 实验结果比较与分析

表 1 为不同模型在包含室内和室外场景的 SOTS 测试集上的平均 PSNR 和 SSIM 值。Dehaze-Net^[13] 和 AOD-Net^[14] 通过 CNN 预测中间参数,不依赖于人工先验知识。然而,该类方法的去雾性能主要取决于模型参数估计的准确性。从表 1 可以看出,Dehaze-Net^[13] 和 AOD-Net^[14] 在室外场景下的性能优于室内。这主要是因为 OTS 训练集远大于 ITS 训练集,Dehaze-Net^[13] 和 AOD-Net^[14] 在较大规模数据集上进行训练,可以获取更准确的模型参数。而在训练数据规模不足时,物理模型驱动的网络易出现特征学习不足的问题。Y-Net^[23]、AECR-Net^[24]、MSFDN^[25]、SGID-PFF^[26] 和本文方法均不依赖于手工先验知识和大气散射模型,可对图像进行直接端到端去雾。其中,Y-Net^[23] 采用一种类似 Y 型的网络结构,通过多次上、下采样充分提取图像的多尺度信息。MSFDN^[25] 在每次上、下采样操作后引入注意力模块,使网络在获取多尺度特征的同时对其进行自适应融合。MSFDN^[25] 在室内和室外场景获得的 PSNR 均超过了 30 dB,同时 SSIM 也分别达到了 0.967 0 和 0.975 6。AECR-Net^[24] 将无雾和有雾图像分别作为正、负样本,并利用对比正则化损失对网络进行约束。由于对比样本之间的特征不一致,正、负样本在解空间中距离较远,容易导致网络训练不稳定。AECR-Net^[24] 获得的 PSNR 和 SSIM 均明显低于 MSFDN^[25]。SGID-PFF^[26] 首先采用深度预去雾器生成包含清晰纹理结构的中间结果,并将其作为参考图像。为了充分利用雾图中的有用信息引导去雾,SGID-PFF^[26] 设计了一个渐进式特征融合模块,通过融合有雾图像和参考图像特征,引导网络恢复无雾图。提出的 MDRD-Net 在每次下、上采样后引入误差反馈模块,减少细节和纹理信息损失。考虑到雾图中雾霾分布不均匀,MDRD-Net 在特征提取阶段级联了多个雾霾感知模块,促使网络重点关注雾图中的浓雾与细节区域。此外,MDRD-Net 在跳跃连接中引入注意力机制,对深层和浅层特征进行自适应融合,减少信息冗余。相较于其他方法,本文提出的 MDRD-Net 在室内和室外图像上均可获得较高的 PSNR 和 SSIM。

从 SOTS 数据集中随机挑选的几幅室内和室外图像的去雾结果如图 6 所示。图 6(a) 为原始雾图,图 6(b)~(h) 分别为 Dehaze-Net^[13]、AOD-Net^[14]、Y-Net^[23]、AECR-Net^[24]、MSFDN^[25]、SGID-PFF^[26] 和本文方法结果,

表 1 不同方法在 SOTS 数据集上的去雾性能定量比较

Table 1 Quantitative comparison of dehazing performance achieved by different methods on SOTS dataset

方法	室内		室外	
	PSNR \uparrow	SSIM \uparrow	PSNR \uparrow	SSIM \uparrow
DehazeNet ^[13]	21.14	0.847 2	22.46	0.851 4
AOD-Net ^[14]	19.06	0.832 7	20.29	0.876 5
Y-Net ^[23]	22.52	0.918 3	26.39	0.949 7
AECR-Net ^[24]	23.21	0.872 4	29.53	0.957 4
MSFDN ^[25]	30.36	0.967 0	32.40	0.975 6
SGID-PFF ^[26]	30.05	0.977 3	30.20	0.975 4
本文	33.69	0.986 4	32.08	0.984 1

图 6(i) 为真实无雾图, 作为参考。图 6B、D、F 分别为图 6A、C、E 中绿色矩形框所示区域的放大图。Dehaze-Net^[13] 和 AOD-Net^[14] 通过搭建 CNN 网络预测大气光和透射率, 并根据物理模型恢复清晰无雾图像。模型结构相对简单, 难以充分捕获和挖掘图像特征, 其去雾性能有限。从图 6(b) 和 (c) 可以看到, 去雾后的图像存在对比度低、细节模糊等问题。Y-Net^[23]、AECR-Net^[24]、MSFDN^[25]、SGID-PFF^[26] 和本文方法属于端到端去雾算法, 不依赖先验知识和物理模型, 不涉及中间变量或参数

的预测, 可避免参数估计不准确带来误差累积。具体地, Y-Net^[23] 和 MSFDN^[25] 通过连续多次下采样获取不同尺度的特征信息, 然而, 该操作将不可避免造成图像细节信息丢失。从图 6(d) 和 (f) 可以看出, 上述方法去雾后的图像存在细节模糊和伪影问题。AECR-Net^[24] 利用对比学习, 将原始清晰图像和有雾图像作为正、负样本, 使输入图像在特征空间中靠近正样本, 远离负样本。然而, 不同雾霾程度的负样本在特征空间中分布不平衡, 这将严重干扰网络的特征学习与表达。从图 6(e) 可以看出, 浓雾区域的去雾效果不佳。SGID-PFF^[26] 利用雾图本身所蕴含的纹理细节等有用信息指导图像去雾。该方法的去雾性能较好, 但在图像高频区域仍会出现边缘或纹理信息模糊。本文方法采用编解码结构, 在每次下、上采样后引入误差反馈模块, 在减少信息丢失、加强非邻层级间的特征交互。同时, 特征提取阶段级联多个雾霾感知模块, 使网络重点关注浓雾和细节区域, 并在跳跃连接中引入注意力机制, 以克服编、解码端之间的语义鸿沟, 增强深、浅层特征之间的深度融合。从图 6(h) 可以看出, 本文提出的 MDRD-Net 能更好地保持图像结构和细节信息, 在去除图像雾霾的前提下, 恢复出色彩真实、对比度高的清晰无雾图。



图 6 不同方法在 SOTS 测试集上的去雾结果视觉比较

Fig. 6 Visual comparison of dehazing results by different methods on SOTS dataset

SOTS 为根据大气散射模型建立的人工合成数据集,其雾霾模式单一。然而,在自然环境下,雾霾分布更加复杂且不均。不同方法在真实雾图上的去雾效果如图 7 所示。图 7(a)为原始雾图,图 7(b)~(h)分别为 Dehaze-Net^[13]、AOD-Net^[14]、Y-Net^[23]、AECR-Net^[24]、MSFDN^[25]、SGID-PFF^[26]和本文方法结果。Dehaze-Net^[13]受物理模型约束,去雾后图像色彩饱和度过高。AOD-Net^[14]和 AECR-Net^[24]在真实场景中对细节恢复能力有限,导致去雾图像存在整体对比度低,细节模糊等问题。Y-Net^[23]、MSFDN^[25]和 SGID-PFF^[26]在真实雾图中虽能一

定程度保留结构信息,但在平坦区域产生伪影,如图 7(b)、(c)所示。本文提出的 MDRD-Net 利用误差反馈模块可有效补偿采样过程中的信息丢失,减少细节缺失。提出的雾霾感知模块使网络在真实场景中能准确关注浓雾与细节区域,而带注意力机制的跳跃连接则在融合深浅层特征的同时抑制冗余信息,避免伪影生成。从图 7 可以看到,相较于其他方法,MDRD-Net 在复杂雾霾分布下依然保持稳定性能,不仅能有效去除图像中的雾霾,同时可确保去雾后的图像色彩真实、细节丰富,避免产生伪影。



图 7 不同方法在真实雾图上的实验结果比较

Fig. 7 Experimental results of different methods on real hazy images

2.5 消融实验

1) SCAM 模块讨论

本文提出的 SCAM 可以自适应融合编码端浅层信息和解码端高级语义信息。为了深入研究跳跃连接中不同融合机制对图像去雾的影响,表 2 为采用不同跳跃连接结构进行去雾的性能比较。其中,Add 表示在跳跃连接中将编码、解码端特征进行通道相加操作,Cat 表示通道拼接操作,Mix 表示对编解码端特征进行 Mix 操作^[24],即通过引入一个可学习的参数,自动调节编码和解码端特征权重。

SCAM 表示采用本文提出的带注意力机制的跳跃连接。SCAM 通过引入注意力机制,自适应融合编、解码端不同层级的特征信息,减少信息冗余,在不引入大量参数的前提下,引导网络重点关注浓雾和细节区域。从表 2

可以看出,提出方法可获得更高的 PSNR 和 SSIM。

表 2 不同跳跃连接结构的去雾性能比较

Table 2 Dehazing performance comparison of different skip connection structures

跳跃连接结构	PSNR ↑	SSIM ↑
Add	28.88	0.972 1
Cat	29.02	0.971 1
Mix	30.83	0.978 1
SCAM	32.08	0.984 1

采用不同跳跃连接在户外雾图上的去雾结果比较如图 8 所示,其中图 8(a)为原始雾图,图 8(b)~(e)为分别采用 Add、Cat、Mix、SCAM 作为跳跃连接的去雾结果。可以看出,采用 SCAM 能够有效去除雾霾的同时,获得对比度高、细节清晰的无雾图像。

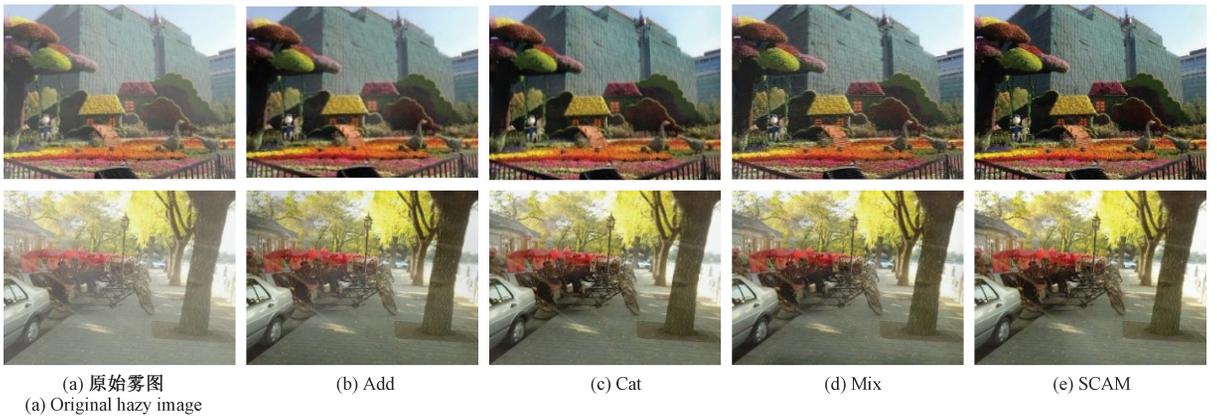


图 8 不同跳跃连接结构的去雾结果比较

Fig. 8 Comparison of dehazed results with different skip connection structures

2) EFM 和 HAM 模块讨论

MDRD-Net 采用编码-解码结构作为网络框架,在编、解码阶段对称引入 EFM,补偿下采样丢失的信息,并在不同误差反馈模块之间引入密集连接机制,促进不同层级之间特征的传递与交互。此外,为了高效、深入提取图像特征,在网络深层级联了多个 HAM。

为了验证 EFM 和 HAM 模块的有效性,本文设计了 4 种不同网络配置。(1) 基准网络 Base,将 HAM 替换成常规 3×3 大小的卷积层,且不包含 EFM 的编码-解码结构;(2) Base+HAM,在 Base 的基础上引入 HAM 模块;(3) Base+EFM,在 Base 的基础上引入 EFM 模块;(4) Base+HAM+EFM,在 Base 的基础上引入 HAM 和 EFM 模块。

表 3 为不同网络配置在 SOTS 室外测试集上的性能比较。Base 网络仅获得 26.9 dB 的 PSNR 和 0.968 9 的 SSIM,整体去雾性能最差。随着 HAM 或 EFM 模块的引入,网络在 PSNR 和 SSIM 上均有一定程度提升。在网络中同时部署 HAM 和 EFM 模块时,网络性能最佳,证实了网络架构设计和各模块的有效性。

表 3 不同网络配置在 SOTS 数据集上的性能比较

Table 3 Performance comparison of different network configurations on SOTS datasets

不同网络配置	PSNR \uparrow	SSIM \uparrow
Base	26.90	0.968 9
Base+HAM	28.29	0.979 8
Base+EFM	31.17	0.983 2
Base+HAM+EFM	32.08	0.984 1

2.6 运行时间

不同方法的参数数量、浮点计算数、运行平台和平均运行时间对比如表 4 所示。Dehaze-Net^[13] 和 AOD-Net^[14] 均依赖于大气散射模型,方法相对简单,涉及的参数数量与

浮点计算次数较少,可直接在 CPU 上运行。但该类方法鲁棒性不高,去雾性能有限。其他深度学习方法,如 Y-Net^[23]、AECR-Net^[24]、MSFDN^[25]、SGID-PFF^[26] 和本文提出的 MDRD-Net,均在 Pytorch 环境下编程,依赖 GPU 加速运行。MDRD-Net,参数数量仅为 1.195×10^6 ,且其浮点计算次数和测试一幅图像的平均运行时间分别为 0.89 GFLOPs 和 0.18 s,明显低于 AECR-Net^[24]、MSFDN^[25] 和 SGID-PFF^[26],表明本文方法在空间和时间效率上均具有明显优势。

表 4 不同方法的运行时间比较

Table 4 Running time comparison of different methods

方法	参数数量/ ($\times 10^6$)	浮点计算次数/ GFLOPs	运行平台	运行 时间/s
Dehaze-Net ^[13]	0.010	0.28	Matlab(CPU)	0.98
AOD-Net ^[14]	0.002	0.01	Pytorch(CPU)	0.18
Y-Net ^[23]	0.2547	0.02	Pytorch(GPU)	0.19
AECR-Net ^[24]	2.003	3.70	Pytorch(GPU)	0.19
MSFDN ^[25]	7.670	0.82	Pytorch(GPU)	0.22
SGID-PFF ^[26]	13.87	15.48	Pytorch(GPU)	0.78
本文方法	1.195	0.89	Pytorch(GPU)	0.18

3 结论

针对现有深度学习去雾方法计算开销过大、下采样造成细节丢失的问题,本文提出了一种基于误差反馈的多尺度密集残差去雾网络。该网络在上、下采样之后引入误差反馈模块,补偿采样过程造成的信息损失及加强非邻层级间的信息交互;并在网络瓶颈处级联多个雾霾感知模块,使网络重点关注雾图中细节丰富的高频区域;此外,为了自适应融合编、解码器之间不同语义特征,提出了带注意力机制的跳跃连接。提出方法在 RESIDE 公开数据集上进行了大量消融和对比实验,结果表明,相较

于其他方法,本文方法可在不引入大量参数和计算成本的前提下,有效去除图像中雾霾干扰,同时避免产生偏色、伪影、和细节丢失等问题。未来工作将探索任务驱动的协同训练机制,通过将图像去雾与目标检测、图像识别等高级视觉任务联合训练与优化,进一步提升模型的跨任务泛化能力和应用性能。

参考文献

- [1] 吴斌,王世杰,卢轶,等. 基于激光雷达与相机融合的 AGV 动态环境目标检测算法[J]. 仪器仪表学报, 2025, 46(4): 163-172.
WU B, WANG SH J, LU Y, et al. Dynamic environment target detection algorithm for AGV based on lidar and camera fusion [J]. Chinese Journal of Scientific Instruments, 2025, 46(4): 163-172.
- [2] 封皓元,段勇. 基于深度迁移学习的天气图像识别[J]. 电子测量与仪器学报, 2023, 37(4): 223-230.
FENG H Y, DUAN Y. Weather image recognition based on fusing deep transfer learning [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(4): 223-230.
- [3] WANG X, LI S, KALLIDROMITIS K, et al. Hierarchical open-vocabulary universal image segmentation [C]. Advances in Neural Information Processing Systems, 2023, 36: 21429-21453.
- [4] MONDAL K, RABIDAS R, DASGUPTA R. Single image haze removal using contrast limited adaptive histogram equalization based multiscale fusion technique [J]. Multimedia Tools and Applications, 2024, 83 (5): 15413-15438.
- [5] AL SOBBAHI R, TEKLI J. Low-light homomorphic filtering network for integrating image enhancement and classification [J]. Signal Processing: Image Communication, 2022, 100: 116527.
- [6] OTHMAN G, ZEEBAREE D Q. The applications of discrete wavelet transform in image processing: A review [J]. Journal of Soft Computing and Data Mining, 2020, 1(2): 31-43.
- [7] WU, W, WENG, J, ZHANG, P, et al. URetinex-Net: Retinex-based deep unfolding network for low-light image enhancement [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 5901-5910.
- [8] LING P, CHEN H, TAN X, et al. Single image dehazing using saturation line prior [J]. IEEE Transactions on Image Processing, 2023, 32: 3238-3253.
- [9] ZHU Q, MAI J, SHAO L. A fast single image haze removal algorithm using color attenuation prior [J]. IEEE Transactions on Image Processing, 2015, 24(11): 3522-3533.
- [10] BERMAN D, TREIBITZ T, AVIDAN S. Non-local image dehazing [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 1674-1682.
- [11] ANCUTI C O, ANCUTI C, TIMOFTE R. NH-HAZE: An image dehazing benchmark with non-homogeneous hazy and haze-free images [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020: 444-445.
- [12] ZHANG X, DONG H, PAN J, et al. Learning to restore hazy video: A new real-world dataset and a new method [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 9239-9248.
- [13] CAI B, XU X, 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.
- [14] LI B, PENG X, WANG Z, et al. AOD-Net: All-in-one dehazing network [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017.
- [15] CHEN D, HE M, FAN Q, et al. Gated context aggregation network for image dehazing and deraining [C]. 2019 IEEE Winter Conference on Applications of Computer Vision, 2019: 1375-1383.
- [16] QIN X, WANG Z, BAI Y, et al. FFA-Net: Feature fusion attention network for single image dehazing [C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 11908-11915.
- [17] CUI Y, TAO Y, BING Z, et al. Selective frequency network for image restoration [C]. The Eleventh International Conference on Learning Representations, 2023.
- [18] LI Q, WANG X, MA B, et al. Concealed attack for robust watermarking based on generative model and perceptual loss [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32 (8): 5695-5706.
- [19] ALAM M K, HAQUE T, AKHTER F, et al. Teeth segmentation by optical radiographic images using VGG-16 deep learning convolution architecture with R-CNN network approach for biomedical sensing applications [J]. Opt Quant Electron, 2023, 55: 808.
- [20] LI B, REN W, FU D, et al. Benchmarking single-image dehazing and beyond [J]. IEEE Transactions on Image Processing, 2019, 28(1): 492-505.
- [21] VENU D N. PSNR based evaluation of spatial Gaussian

- kernels for FCM algorithm with mean and median filtering based denoising for MRI segmentation [J]. IJFANS International Journal of Food and Nutritional Sciences, 2023, 12(1): 928-939.
- [22] SAURAV J R, NASR M S, SHANG H H, et al. A SSIM guided cGAN architecture for clinically driven generative image synthesis of multiplexed spatial proteomics channels [C]. 2023 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology, 2023: 1-8.
- [23] YANG H H, YANG C H H, TSAI Y C J. Y-Net: Multi-scale feature aggregation network with wavelet structure similarity loss function for single image dehazing [C]. ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing, 2020: 2628-2632.
- [24] WU H, QU Y, LIN S, et al. Contrastive learning for compact single image dehazing [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 10551-10560.
- [25] WANG G, YU X. MSF²DN: Multi scale feature fusion dehazing network with dense connection [C]. 16th Asian Conference on Computer Vision, 2022: 444-459.
- [26] BAI H, PAN J, XIANG X, et al. Self-guided image dehazing using progressive feature fusion [J]. IEEE Transactions on Image Processing, 2022, 31: 1217-1229.

作者简介



高雯婷, 湖南科技大学本科生, 主要研究方向为图像处理、深度学习。

E-mail: 1353646103@qq.com

Gao Wenting is now a B. Sc. candidate at Hunan University of Science and Technology. Her main research interests

include image processing and deep learning.



廖苗(通信作者), 分别在 2010 年、2013 年和 2016 年于中南大学获得学士学位、硕士学位和博士学位, 现为湖南科技大学副教授, 主要研究方向为图像处理与模式识别。

E-mail: liaomiaohi@163.com

Liao Miao (Corresponding author) received her B. Sc. degree, M. Sc. degree, and Ph. D. degree all from Central South University in 2010, 2013, and 2016, respectively. She is now an associate professor at Hunan University of Science and Technology. Her main research interests include image processing and pattern recognition.