

# 结合密集注意力的自适应特征融合图像去雾网络<sup>①</sup>



王 燕, 他 雪, 卢鹏屹

(兰州理工大学 计算机与通信学院, 兰州 730050)

通信作者: 王 燕, E-mail: [wangyan@lut.edu.cn](mailto:wangyan@lut.edu.cn)

**摘 要:** 目前, 大多数图像去雾算法忽视图像的局部细节信息, 无法充分利用不同层次的特征, 导致恢复的无雾图像仍存在颜色失真、对比度下降和雾霾残留现象, 针对这一问题, 提出结合密集注意力的自适应特征融合图像去雾网络. 该网络以编码器-解码器结构为基本框架, 中间嵌入特征增强部分与特征融合部分, 通过在特征增强部分叠加由密集残差网络与 CS 联合注意模块构成的密集特征注意块, 使网络可以关注图像的局部细节信息, 同时增强特征的重复利用, 有效防止梯度消失; 在特征融合部分构建自适应特征融合模块融合低级与高级特征, 防止因网络加深而造成浅层特征退化. 实验结果表明, 所提算法在合成有雾图像数据集和真实有雾图像数据集上均表现优异, 在 SOTS 室内合成数据集上的峰值信噪比和结构相似性分别达到了 35.81 dB 和 0.9889, 在真实图像数据集 O-HAZE 上的峰值信噪比和结构相似性分别达到了 22.75 dB 和 0.7788, 有效解决了颜色失真、对比度下降和雾霾残留等问题.  
**关键词:** 图像去雾; 深度学习; 编码器-解码器; 密集连接; 注意力机制; 特征融合

引用格式: 王燕, 他雪, 卢鹏屹. 结合密集注意力的自适应特征融合图像去雾网络. 计算机系统应用, 2024, 33(2): 72-82. <http://www.c-s-a.org.cn/1003-3254/9407.html>

## Adaptive Feature Fusion Image Dehazing Network Combined with Dense Attention

WANG Yan, TA Xue, LU Peng-Yi

(School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China)

**Abstract:** At present, most image dehazing algorithms ignore the local details of the image and fail to make full use of features at different levels, resulting in color distortion, contrast reduction, and haze residual phenomena in the restored image without fog. To solve this problem, this study proposes an adaptive feature fusion image dehazing network combined with dense attention. The network takes the encoder-decoder structure as the basic framework, and the feature enhancement part and the feature fusion part are embedded in the middle. The dense feature attention block composed of the dense residual network and the Channel-Spatial attention combination module is superimposed on the feature enhancement part. In this way, the network can pay attention to the local details of the image, enhance the reuse of features, and effectively prevent the disappearance of gradients. In the feature fusion part, an adaptive feature fusion module is constructed to fuse low-level and high-level features to prevent shallow feature degradation caused by the deepening of the network. The experimental results show that the proposed algorithm performs well on both synthetic and real fog image datasets. The peak signal-to-noise ratio and structural similarity on SOTS indoor synthetic datasets reach 35.81 dB and 0.9889, respectively, and those on the real image datasets O-HAZE reach 22.75 dB and 0.7788 respectively. The proposed algorithm effectively solves the problems of color distortion, contrast reduction, and haze residue.

**Key words:** image dehazing; deep learning; encoder-decoder; dense connection; attention mechanism; feature fusion

① 基金项目: 国家自然科学基金 (61863025)

收稿时间: 2023-08-21; 修改时间: 2023-09-26; 采用时间: 2023-10-09; csa 在线出版时间: 2023-12-25

CNKI 网络首发时间: 2023-12-27

## 1 引言

雾、霾等复杂天气下捕获到的图像受到大气中混浊介质的影响,具有明显的退化现象,如颜色失真、饱和度低、纹理细节模糊等,而低质量的图像输入将难以解决其他高级计算机视觉任务,导致图像分类、语义分割、目标检测及跟踪等任务的准确性降低<sup>[1-4]</sup>.因此,图像去雾<sup>[5]</sup>作为高级视觉任务的预处理步骤,在计算机视觉任务中起着至关重要的作用.图像去雾主要包括基于图像增强的去雾方法,基于物理模型与先验知识的去雾方法和基于深度学习的去雾算法.

基于图像增强的去雾方法不依赖大气散射模型<sup>[6]</sup>,通过去除图像噪声、提高图像对比度来恢复清晰无雾图像,如 Retinex 方法<sup>[7]</sup>、直方图均衡化<sup>[8]</sup>和小波变换<sup>[9]</sup>等,此类方法并没有从根本上分析影响图像退化的因素,导致去雾后的图像不够自然,因此其去雾效果以及算法鲁棒性有限.

基于物理模型与先验知识的去雾算法通过先验知识或约束方程对物理模型参数进行估计,以实现图像去雾.最具代表性的是 He 等人<sup>[10]</sup>提出的基于暗通道先验 (dark channel prior, DCP) 的去雾算法,该算法基于暗通道先验理论估算透射率与大气光值,利用大气散射模型恢复无雾图像. Zhu 等<sup>[11]</sup>提出颜色衰减先验 (color attenuation prior, CAP) 算法,通过统计大量有雾图像,发现雾霾浓度和图像的亮度与饱和度之差成正比,由此建立了一个线性去雾模型. Meng 等<sup>[12]</sup>提出一种边界约束和正则化去雾方法 (BCCR),使用边界约束估计透射率,并增加上下文正则化来优化透射率,可以较好地恢复图像中的细节信息.虽然基于物理模型与先验信息的方法取得了显著的进展,但是,此类方法在无约束条件下,通常会发​​生不准确的估计,从而导致恢复出无雾图像的质量不够理想,且这些手工制作的先验并不总是适用于各种现实世界的模糊场景.

与传统方法不同,基于深度学习的图像去雾方法通过直接或间接地学习基于大规模数据集的映射函数来得到无雾图像. Cai 等<sup>[13]</sup>提出基于卷积神经网络的介质传输估计网络 DehazeNet,通过估计模糊图像与其介质传输之间的映射关系实现端到端去雾; Li 等<sup>[14]</sup>改进大气散射模型,提出基于深度学习的图像去雾模型 AOD-Net (all-in-one network),将透射率图和大气光放到一个新的变量中,通过轻量级卷积网络直接生成去雾图像; Chen 等<sup>[15]</sup>提出一种基于平滑扩张卷积的门控

上下文聚合网络 GCANet,解决了原始扩张卷积造成的网格化伪影问题; Qin 等<sup>[16]</sup>提出一种特征融合注意力网络 FFANet,将像素注意机制与通道注意机制结合在一起构建双注意力模块,应用注意力机制自适应学习不同特征信息的权重以实现图像去雾; Dong 等<sup>[17]</sup>基于生成对抗网络 (generative adversarial network, GAN) 提出了 FD-GAN,该网络将频率信息作为融合鉴别器的先验,指导网络生成更加真实自然的去雾图像; Wu 等<sup>[18]</sup>提出了一种基于对比学习的对比正则化算法 AEER-Net,利用清晰图像和有雾图像的信息分别作为正负样本,并借助提出的 CR (contrastive regularization) 恢复无雾图像; Song 等<sup>[19]</sup>提出一种 U-Net 的变体网络 gUNet,使用具有门控机制的残差块代替卷积块,通过选择性内核融合主路径和跳过连接的特征图,显著降低网络开销; Zheng 等<sup>[20]</sup>提出的 C<sup>2</sup>PNet,结合物理感知双分支单元和课程学习对比正则化,重新调整对比正则化中不同负样本的重要性,提高去雾特征空间的可解释性.除此之外,随着 Transformer 逐渐应用于图像领域, Guo 等<sup>[21]</sup>提出的 DeHamer 通过学习基于 Transformer 特征的调制矩阵来调制 CNN 特征,解决 Transformer 与 CNN 之间的特征不一致问题,同时提出一种新的传输感知三维位置嵌入,有效地将先验信息引入 Transformer,以此提高去雾性能. Song 等<sup>[22]</sup>提出了一种 Swin Transformer 和 U-Net 结合的图像去雾网络 DehazeFormer,通过对 Swin Transformer 的归一化层、激活函数和空间信息聚合方案进行改进,使 Transformer 更适合于图像去雾.

然而,随着数据驱动方法的发展和去雾性能的提高,去雾网络的复杂性也在增加.现有的基于深度学习的算法仍然存在以下挑战:极易出现去雾不彻底、对比度下降和颜色失真现象;模型参数量大;训练策略复杂;在真实有雾图像数据集上的去雾效果不理想.

对于以上问题,本文提出了一种结合密集注意力的自适应特征融合图像去雾网络 (adaptive feature fusion image dehazing network combined with dense attention, DAAF-Net),该网络不依赖于大气散射模型,而是通过学习有雾图像到无雾图像之间的映射关系来恢复无雾图像. DAAF-Net 可以很好地解决局部雾霾残留及去雾不彻底现象,在合成有雾图像以及真实有雾图像上具有强大的还原图像细节和色彩保真度的能力.本文的具体贡献如下.

(1) 提出结合密集注意力的自适应特征融合图像去雾网络. 通过一个包含特征增强部分 (feature enhancement part) 和特征融合部分 (feature fusion part) 的基于编解码器结构的去雾网络来有效地提取图像特征, 融合不同尺度特征信息, 恢复清晰无雾图像.

(2) 提出一种嵌入 CS 联合注意模块 (channel-spatial attention combination module) 的密集特征注意块 (dense feature attention block, DFAB). DFAB 结合密集残差网络和注意力机制, 增强特征的重复利用, 有效防止梯度消失; CS 联合注意模块依次从通道维度和空间维度聚焦上下文特征信息, 解决不同级别的不均匀雾度分布.

(3) 提出融合低级与高级特征的自适应特征融合模块 (adaptive feature fusion module, AFM), 动态融合下采样层和上采样层之间的信息以防止浅层特征逐渐退化.

## 2 DAAF-Net 模型设计

### 2.1 网络总体架构

本文提出了结合密集注意力的自适应特征融合图像去雾网络 DAAF-Net. 该网络由编码器部分 (encoder part)、特征增强部分 (feature enhancement part)、解码器部分 (decoder part) 以及特征融合部分 (feature fusion part)

组成. DAAF-Net 首先将输入的雾霾图像送入编码器中进行下采样, 利用多层卷积和特征增强残差块 (feature-enhanced residual block, RB) 提取浅层特征, 如图像轮廓、纹理等, 之后将编码器提取的图像特征送入特征增强部分, 由密集特征注意块学习低分辨率空间中的特征表示, 从而达到特征增强的目的, 最后将增强后的特征送入解码器部分来恢复清晰无雾图像, 与编码器部分不同的是, 解码器采用步长为 2 的反卷积将特征图像恢复为原始输入图像尺寸. 同时, 为了防止浅层特征丢失, 本文设计了自适应特征融合模块来充分融合上下采样层之间特征. 网络整体框架如图 1 所示.

### 2.2 密集特征注意块 (DFAB)

为获取图像的深层特征, 本文在特征增强部分叠加 6 个密集特征注意块 DFAB, 进一步提高了 DAAF-Net 的性能. 密集特征注意块如图 2 所示. DFAB 采用密集残差连接和注意力机制的思想, 一方面, 密集连接通过特征在通道上的连接来实现特征重用, 缓解梯度消失问题, 而局部残差学习绕过薄雾区域和低频信息, 使得主网络更多关注浓雾, 学习到更多有用的信息; 另一方面, CS 联合注意模块依次从通道维度和空间维度聚焦上下文特征信息, 在处理不同类型的信息时提供额外的灵活性, 并且可以扩展 CNN 的表示能力.

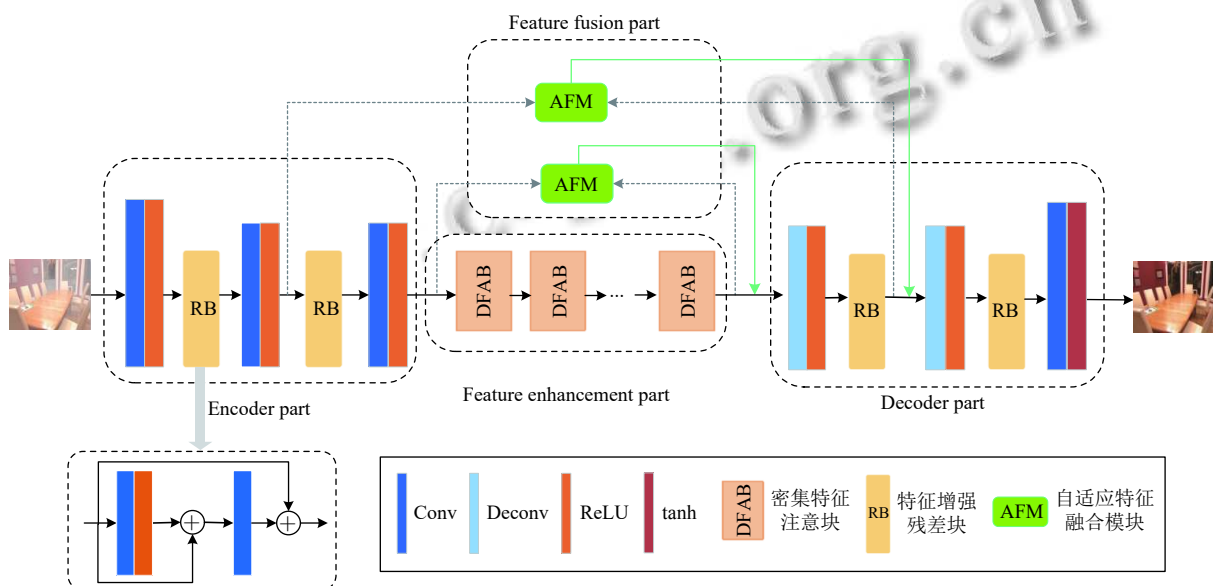


图 1 DAAF-Net 网络整体框架

CS 联合注意模块如图 3 所示. 该模块将通道注意力和空间注意力顺序连接以计算通道和空间维度上的

注意权重, 使得输出的深层特征具有全局视野, 从而使网络能够更好地辨识重要特征依赖关系, 过滤非重要

信息, 进而提升网络的性能.

通道注意力关注通道差异, 计算通道注意力特征图, 而不考虑上下文信息, 其权重系数  $W_C \in R^{C \times 1 \times 1}$ ; 空间注意力则从空间维度对重要信息进行加权, 使网络更加关注于图像的空间信息, 其权重系数  $W_S \in R^{1 \times H \times W}$ .

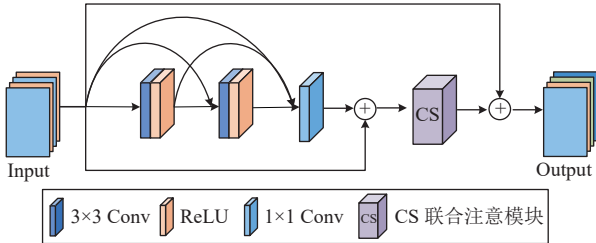


图2 密集特征注意块

设  $F_{in} \in R^{C \times H \times W}$  表示 CS 联合注意模块的输入特征, 通道注意力通过平均池化得到  $C \times 1 \times 1$  维的特征, 将基于通道的全局空间信息转化为通道描述  $G_C$ , 经过两个卷积层及 Sigmoid 函数, 得到通道权重  $W_C$ . 计算过程如式 (1) 和式 (2) 所示:

$$G_C = \text{AvgPool}(F_{in}) \quad (1)$$

$$W_C = \sigma(\text{Conv}_{1 \times 1}(\delta(\text{Conv}_{1 \times 1}(G_C)))) \quad (2)$$

其中, AvgPool表示自适应平均池化操作,  $\sigma$ 是 Sigmoid 函数,  $\delta$ 是 ReLU 函数.

将通道权重  $W_C$  与输入特征  $F_{in}$  中各通道的对应元素相乘, 输出大小为  $C \times H \times W$  的特征图  $F_C$ . 如式 (3) 所示:

$$F_C = W_C \otimes F_{in} \quad (3)$$

在通道注意力之后, 引入空间注意力,  $F_C$  即为空间注意力的输入特征, 首先对输入特征  $F_C$  分别进行一个通道维度的最大池化和平均池化, 使网络分别获取纹理特征信息和全局上下文信息, 然后通过通道相加将池化操作后得到的包含不同空间信息的重要特征进行融合, 得到  $2 \times H \times W$  维的特征. 此后经过卷积和 Sigmoid 函数, 得到权重系数  $W_S$ . 计算过程如式 (4) 和式 (5) 所示:

$$G_S = \text{Concat}(\text{AvgPool}(F_C), \text{MaxPool}(F_C)) \quad (4)$$

$$W_S = \sigma(\text{Conv}_{7 \times 7}(G_S)) \quad (5)$$

其中, AvgPool表示自适应平均池化操作, MaxPool 表示自适应最大池化操作,  $\sigma$ 是 Sigmoid 函数.

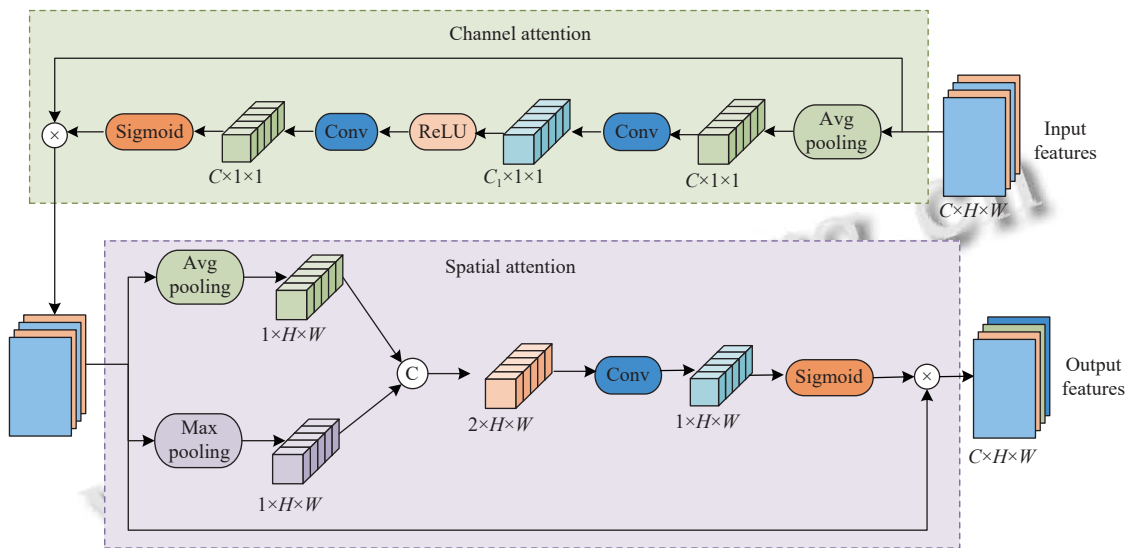


图3 CS联合注意模块

将空间权重  $W_S$  与空间注意力的输入特征  $F_C$  中各通道的对应元素相乘, 最终得到空间注意力加权后的特征图  $F_S$  即 CS 联合注意模块的输出特征图. 如式 (6) 所示:

$$F_S = W_S \otimes F_C \quad (6)$$

### 2.3 自适应特征融合模块 (AFM)

大量实验表明, 将编码器部分与解码器部分的特

征融合在图像去雾以及其他低级视觉任务中是一种有效的技巧. 通常情况下, 去雾网络的低层次特征 (如边缘和轮廓) 可以在神经网络的浅层特征提取层捕获, 然而, 随着网络深度的增加, 网络的浅层特征逐渐退化, 使得图像恢复阶段所得到的特征图存在边缘模糊, 细节丢失等现象. 为了解决这个问题, 一部分工作通过引入包含逐元素相加 (element-wise addition) 或者通道相

加 (Concat) 操作的跳过连接来生成新的特征. 然而, 对于本文的去雾网络, 简单的相加或串联操作并不能充分融合上采样层和下采样层之间特征. 鉴于此, 本文提出了自适应特征融合模块 AFM, 该方法通过 Parameter 函数引入可训练的权重  $w$  来调整融合比例, 采用自适应

混合操作<sup>[23]</sup>将这两层的信息进行融合, Parameter 函数初始化的权重会被添加到优化器更新的参数中, 使其在学习过程中会自适应调整以达到最优. 该模块在引入极少的参数量的情况下有效提高了模型的性能. 自适应特征融合模块如图 4 所示.

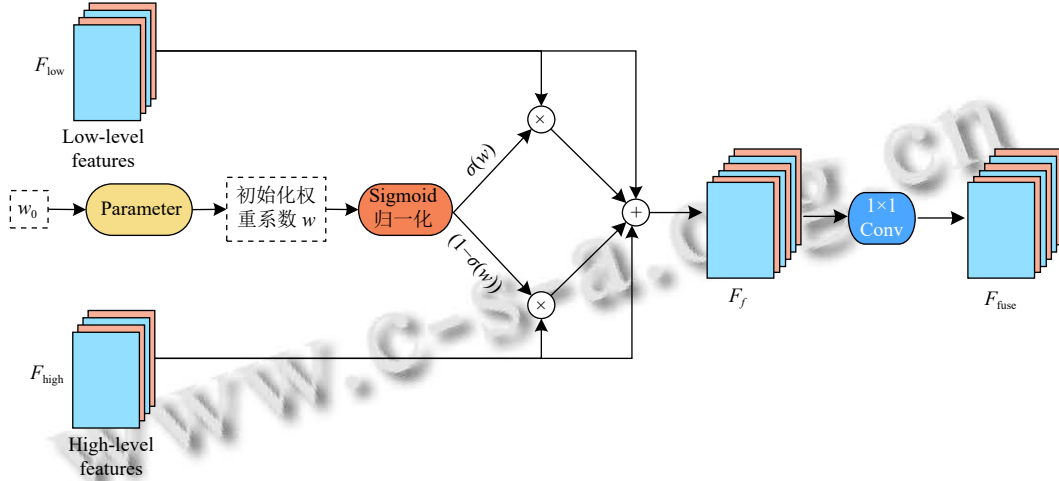


图 4 自适应特征融合模块

编码器部分中的低级特征和相应的高级特征被传送到 AFM 模块, AFM 通过自适应学习权重调整融合比例, 使下采样部分的浅层特征信息自适应地流向上采样部分, 这对于低级特征保留是十分有效的, 比加法操作更灵活. 该模块还通过跳过连接来添加输入特征, 以减轻梯度消失问题并简化学习过程. 最后, 通过  $1 \times 1$  卷积层将融合后的特征进行投影, 得到最终的输出特征 (即  $\sigma$ ). 融合操作的最终输出公式如式 (7) 和式 (8) 所示:

$$F_f = F_{low} + F_{high} + F_{low} \times \sigma(w) + F_{high} \times (1 - \sigma(w)) \quad (7)$$

$$F_{fuse} = \text{Conv}_{1 \times 1}(F_f) \quad (8)$$

其中,  $F_{fuse}$  是 AFM 处理后的最终输出,  $F_f$  是融合低级与高级特征后的特征图,  $F_{low}$  和  $F_{high}$  分别是来自下采样层和上采样层的低级特征和高级特征,  $\sigma$  是 Sigmoid 函数.  $w$  为通过 Parameter 函数引入的初始化融合权重, 采用 Sigmoid 函数对其进行归一化得到  $\sigma(\theta)$ . 在训练期间, AFM 可以有效地学习  $\sigma(\theta)$ , 使其达到一个最优值, 这比常数因子实现了更好的性能.

## 2.4 损失函数

本文通过 Smooth  $L_1$  损失和感知损失加权相加来训练所提出的网络.

Smooth  $L_1$  损失: Smooth  $L_1$  损失结合了  $L_1$  与

$L_2$  范数的部分优点, 提供了去雾图像和真实图像之间差异的定量测量,  $L_1$  范数可以防止潜在的梯度爆炸, 它对异常值的敏感性低于 MSE 损失. Smooth  $L_1$  损失的表达式如式 (9) 和式 (10) 所示:

$$L_S = \frac{1}{N} \sum_{x=1}^N \sum_{i=1}^3 F_S(I_i(x) - J_i(x)) \quad (9)$$

$$F_S(e) = \begin{cases} 0.5e^2, & \text{if } |e| < 1 \\ |e| - 0.5, & \text{otherwise} \end{cases} \quad (10)$$

其中,  $N$  表示像素总数;  $I_i(x)$  和  $J_i(x)$  分别为雾霾图像和去雾图像中像素  $x$  的第  $i$  个颜色通道的像素强度;  $e$  表示  $I_i(x)$  和  $J_i(x)$  的误差.

感知损失: 不同于逐像素损失, 感知损失 (perceptual loss) 通过从预训练的神经网络中提取特征来量化估计去雾图像和真实图像之间的差异, 相比普通的  $L_2$  损失而言, 可以增强输出特征的细节信息. 本文使用预训练的 VGG16 作为损失网络, 并从前 3 级 (即 Conv1-2、Conv2-2 和 Conv3-3) 的最后一层提取特征. 感知损失的表达式如式 (11) 所示:

$$L_P = \sum_{j=1}^3 \frac{1}{C_j H_j W_j} \|\phi_j(\hat{J}) - \phi_j(J)\|_2^2 \quad (11)$$

其中,  $C_j$ 、 $H_j$ 、 $W_j$ 分别代表特征图 $\phi_j(\hat{J})$ 和 $\phi_j(J)$ 的通道数、高度和宽度,  $\hat{J}$ 表示去雾图像,  $J$ 表示真实图像,  $\phi_j(\hat{J})$ 和 $\phi_j(J)$ 表示从VGG16模型提取到的去雾图像和真实图像对应的特征图。

总损失: 总损失函数表达式如式(12)所示:

$$L = L_S + \lambda L_P \quad (12)$$

其中,  $\lambda$ 为感知损失对应的权重系数, 本文将 $\lambda$ 设为0.04。

### 3 实验结果与分析

#### 3.1 实验数据集与评价指标

##### 3.1.1 实验数据集

本文评估模型的数据集主要采用 RESIDE 数据集<sup>[24]</sup>和 O-HAZE 数据集<sup>[25]</sup>。RESIDE 是一个大规模的合成有雾图像数据集, 共分为5个子集, 本文选取室内数据集 ITS (indoor training set) 和室外数据集 OTS (outdoor training set) 作为训练集, 综合目标测试集 SOTS (synthetic objective testing set) 为测试集。ITS 包含1399张清晰图像, 每个清晰图像根据不同的大气光值和 beta 值合成10张模糊图像, 共包含13990张合成模糊图像。与 ITS 相似, OTS 选取2061张来自北京实时天气的真实室外图, 每张清晰图像合成35张有雾图像, 共包含72135张户外模糊图像。为了加快模型的收敛速度, 本文从 OTS 中选取18200对合成室外图像来训练网络。SOTS 测试集包括500对室内图像和500对室外图像, 由于 SOTS 室外测试集所包含的图像尺寸大小不同, 通过缩放操作将测试集图片调整为统一大小以便于与其他模型作对比。除此之外, 为了验证模型的泛化性, 本文还采用真实世界数据集 O-HAZE 评估模型性能, O-HAZE 数据集包含45张有雾图像和相应的清晰图像, 随机选取39张有雾图像作为训练集, 6张有雾图像作为测试集, 采用双线性插值的方式将这些图像 Resize 成380×360大小以便于训练。

##### 3.1.2 评价指标

为了客观评价本文算法的优劣, 采用峰值信噪比 (peak signal-to-noise ratio, PSNR) 和结构相似性 (structural similarity, SSIM) 作为去雾图像质量评价的客观指标。PSNR 是一种评价图像质量的度量标准, PSNR 越大, 图像失真越小。其计算公式如式(13)所示:

$$PSNR = 10 \log_{10} \left( \frac{MAX_I^2}{MSE} \right) \quad (13)$$

其中,  $MAX_I$ 表示图像的最大像素值,  $MSE$ 表示图像处理前后的均方误差, 其表达式如式(14)所示:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (14)$$

其中,  $I$ 表示处理前的图像,  $K$ 表示处理后的图像, 图像大小为  $m \times n$ 。

SSIM 是衡量两幅图像相似度的指标, 主要考量图像亮度、对比度和结构, 能够评价处理前后两张图片的接近程度, SSIM 越大, 处理后的图片越接近原始图片。其计算公式如式(15)所示:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (15)$$

其中,  $\mu_x$ 和 $\mu_y$ 分别表示图像 $x$ 和 $y$ 的像素均值,  $\sigma_x^2$ 和 $\sigma_y^2$ 分别表示图像 $x$ 和 $y$ 的方差,  $\sigma_{xy}$ 表示图像 $x$ 和 $y$ 的协方差,  $c_1$ 、 $c_2$ 为常数。

#### 3.2 实验环境与参数

本文使用 Python 3.7 和 PyTorch 1.10 搭建深度学习环境, 在 NVIDIA GeForce RTX 3060 GPU 上开展实验。本文在 RGB 通道中训练网络, 训练图像被随机裁剪为240×240大小作为网络的输入, 网络模型训练使用 Adam 优化器, 批量大小为8, 其中 $\beta_1$ 和 $\beta_2$ 分别采用默认值0.9和0.999, 初始学习率设置为0.0001。对于 ITS, 每20轮迭代后学习率衰减为原来的一半, 总迭代次数为150次; 对于 OTS, 网络迭代100次, 学习率每10轮迭代后降低一半。对于真实世界数据集 O-HAZE, 设置总迭代次数为500次, 学习率每50轮迭代后降低为原来的0.8。

#### 3.3 实验结果及对比分析

将本文算法与图像去雾领域的多个经典网络进行比较, 比较方法包括 DCP<sup>[10]</sup>、AOD-Net<sup>[14]</sup>、GCANet<sup>[15]</sup>、GridDehazeNet<sup>[26]</sup>、MSBDN<sup>[27]</sup>、MSTN<sup>[28]</sup>和 DehazeFormer-T<sup>[22]</sup>。其中, DCP 为经典的基于先验知识的方法, AOD-Net 为改进大气散射模型的端到端去雾算法, 其他5种对比方法都是最近几年基于深度学习的图像去雾方法, GCANet 为基于编解码器结构的门控上下文聚合网络, GridDehazeNet 为基于注意力机制的一个3行6列的多尺度网格网络, MSBDN 为基于 U-Net 的多尺度增强去雾网络, MSTN 为基于多尺度特征融合模块和自适应特征选择模块的多尺度拓扑网络, Dehaze-

Former-T 为 Swin Transformer 和 U-Net 结合的图像去雾网络 DehazeFormer 的一个变体网络. 为了定性地比较算法优劣, 本文采用统一的数据集及训练方案对各网络进行了重新训练, 将最终测试图像裁剪为统一大小用于结果展示.

### 3.3.1 合成有雾图像数据集上的结果

不同去雾方法在 SOTS 合成测试集上对比实验结果如表 1 所示, 表 1 中加粗数据表示该组数据的最优结果. 与现有的方法相比, 本文提出的 DAAF-Net 在 SOTS 室内测试集上 PSNR 为 35.81 dB 和 SSIM 为 0.9889 的最佳性能, 在 SOTS 室外测试集上 PSNR 为 27.96 dB 和 SSIM 为 0.9197 的性能, 在 SOTS 室外测试集上 SSIM 略低于 MSTN, 差距仅为 0.000 8. 除此之外, 本文算法的参数量仅为 3.19M, 明显低于 MSTN 算法.

图 5 和图 6 分别展示了本文的 DAAF-Net 与现有的经典方法在 SOTS 室内测试集和 SOTS 室外测试集上的视觉效果. 在 SOTS 室内测试集上, AOD-Net 整体去雾功能几乎失效, 存在大量的雾霾残留; DCP 算法存

在对比度偏低和颜色失真现象, 且在图像光线较暗的区域有明显的雾霾残留现象; 与 DCP 算法相反, GCANet 对比度偏低、整体色彩较暗; 其他 4 种对比算法整体去雾效果较好, 但仍然有轻微的雾霾残留, 如 Grid-DehazeNet 算法的第 1 幅图片的桌角上方亮暗交替区域存在未恢复的薄雾; MSBDN 算法的第 1 幅图片左面的一排椅子旁边存在雾霾残留; DehazeFormer-T 算法第 2 幅图片左下角地面有雾霾残留; MSTN 去雾效果较好, 但其定量评价结果仍低于本文算法.

表 1 不同方法在 SOTS 测试集上的定量评价结果

方法	SOTS indoor		SOTS outdoor		#Param
	PSNR (dB)	SSIM	PSNR (dB)	SSIM	
DCP <sup>[10]</sup>	18.63	0.8616	15.39	0.7963	—
AOD-Net <sup>[14]</sup>	20.55	0.8560	23.40	0.8897	1.8k
GCANet <sup>[15]</sup>	26.68	0.9286	25.53	0.9048	0.70M
GridDehazeNet <sup>[26]</sup>	33.37	0.9854	25.86	0.9110	0.96M
MSBDN <sup>[27]</sup>	31.23	0.9736	26.63	0.9085	31.35M
MSTN <sup>[28]</sup>	32.64	0.9810	27.94	<b>0.9205</b>	19.41M
DehazeFormer-T <sup>[22]</sup>	33.21	0.9841	27.50	0.9172	0.69M
DAAF-Net (本文)	<b>35.81</b>	<b>0.9889</b>	<b>27.96</b>	0.9197	3.19M



图 5 不同方法在 SOTS 室内图像上的测试结果

在 SOTS 室外测试集上, DCP 去雾效果极差, 颜色偏移严重, 尤其是图像存在大面积白色区域和天空部分时颜色失真严重; AOD-Net 整体图片较为模糊, 色彩较暗, 且第 3 幅图片远处的楼层模糊, 去雾不彻底; GCANet 去雾效果较 AOD-Net 略好, 但第 3 幅图片远处的楼层仍有雾霾残留, 第 2 幅和第 4 幅图片天空区域色彩较暗; GridDehazeNet 算法有明显的颜色伪影, 如第 2 幅图片的天空区域与第 4 幅图片的楼层旁边; MSBDN 较清晰图片色彩略微偏暗; DehazeFormer-T 算法第 3 幅图片远处的楼层颜色略暗; 本文模型与 MSTN 去雾效果较好, 颜色和细节体现上最接近标准

无雾图.

### 3.3.2 真实有雾图像数据集上的结果

合成有雾图像雾霾分布均匀, 因此各算法去雾效果较好, 而现实中真实世界拍摄的有雾图片通常伴随浓雾, 雾霾分布不均匀, 为进一步验证本文方法对真实有雾图像的有效性, 本文采用真实世界数据集 O-HAZE 对网络进行训练, 同时与上述其他去雾方法作对比, 不同去雾方法在 O-HAZE 数据集上对比实验如表 2 所示. 由表 2 可知, 本文提出的 DAAF-Net 在 O-HAZE 数据集上 PSNR 为 22.75 dB, SSIM 为 0.778 8, 获得了最优性能.

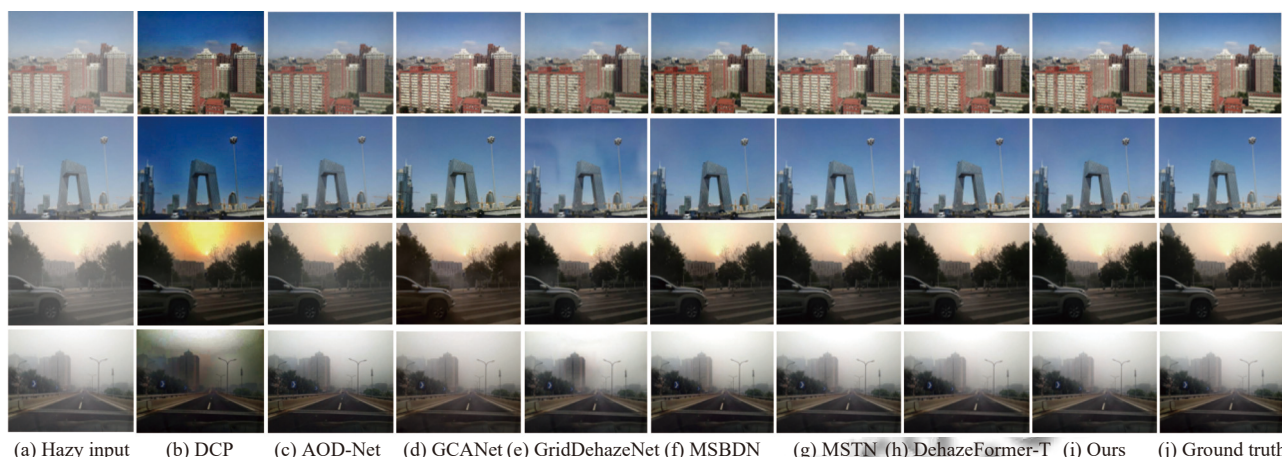


图6 不同方法在 SOTS 室外图像上的测试结果

相应的视觉效果对比图如图7所示。从图7可以观察到,当使用真实有雾图像训练时,模型测试结果普遍偏低。其中,DCP算法图像整体色彩偏蓝,存在严重的颜色偏差;而AOD-Net色彩较暗,图片对比度偏低,视觉效果较差;GCANet去雾效果较好,但图像色彩整体偏亮,且有明显的雾霾残留;GridDehazeNet、MSBDN、MSTN和DehazeFormer-T都存在不同程度的雾霾残留现象,如第1幅图片的木头在浓雾处明显去雾不彻底,第2幅图片远处的树木、第3幅图片的草地和远处的树木以及第4幅图片的地面及栏杆处都有雾霾残留。整体来说,本文算法无论从去雾效果还是图像的色彩对比度恢复方面均优于其他算法。

表2 不同方法在 O-HAZE 数据集上的定量评价结果

方法	O-HAZE		#Param
	PSNR (dB)	SSIM	
DCP <sup>[10]</sup>	15.38	0.6474	—
AOD-Net <sup>[14]</sup>	19.07	0.6757	1.8k
GCANet <sup>[15]</sup>	21.58	0.7663	0.70M
GridDehazeNet <sup>[26]</sup>	22.21	0.7734	0.96M
MSBDN <sup>[27]</sup>	21.23	0.7142	31.35M
MSTN <sup>[28]</sup>	21.58	0.7758	19.41M
DehazeFormer-T <sup>[22]</sup>	21.66	0.7457	0.69M
DAAF-Net (本文)	<b>22.75</b>	<b>0.7788</b>	3.19M

### 3.4 消融实验

为进一步验证网络中各模块的有效性,在 SOTS 室内测试集上设计如下消融实验。

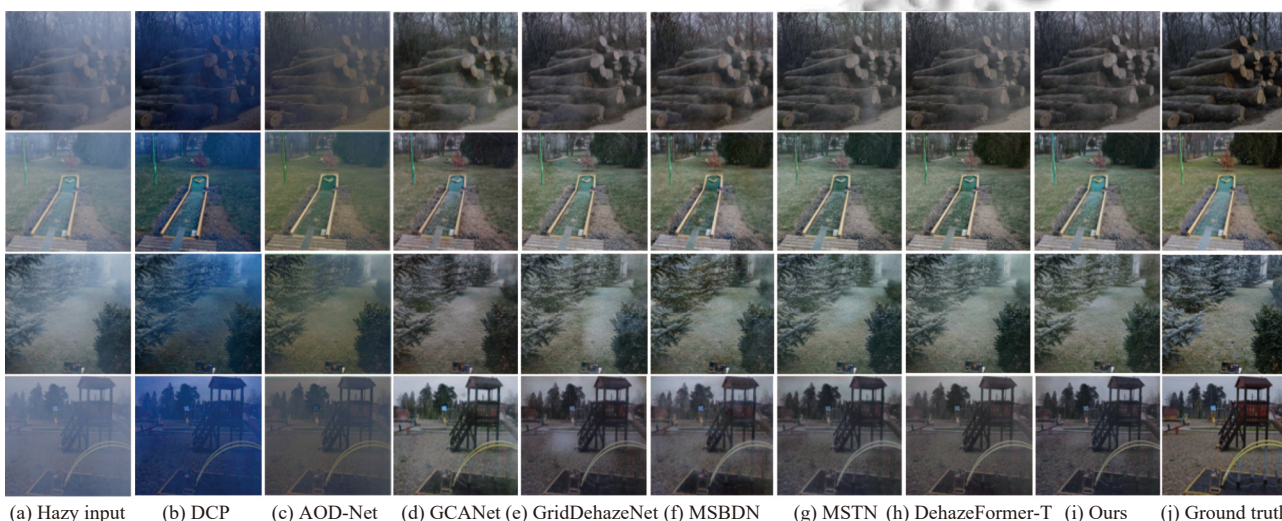


图7 不同方法在 O-HAZE 数据集上的测试结果

#### 3.4.1 不同模块组合消融实验

通过对网络中各模块进行不同的组合来构建消融

实验。RB、DFAB、AFM 模块有效性实验如表3所示,可视化结果如图8所示。其中包括以下组合:① Base+



DFAB; ② Base+RB+AFM; ③ Base+DFAB+AFM; ④ Base+RB+DFAB; ⑤ Base+RB+DFAB+AFM, 即本文网络. 其中, Base 表示两个下采样层和两个上采样层构成的基础网络.

表3 不同模块组合消融实验结果

网络结构组成	评价指标	
	PSNR (dB)	SSIM
Base+DFAB	30.91	0.9760
Base+RB+AFM	25.70	0.9531
Base+DFAB+AFM	32.03	0.9822
Base+RB+DFAB	34.62	0.9875
Base+RB+DFAB+AFM (本文)	<b>35.81</b>	<b>0.9889</b>



图8 不同模块组合消融实验可视化结果

### 3.4.2 DFAB 模块消融实验

本节对 DFAB 模块中不同注意力机制进行消融实验来评估所提模块的影响, 将不包含 DFAB 模块的网络结构记为 Base1, DFAB 模块中去除 CS 联合注意模块的基础模块记为 DB 模块. 分别对以下组合进行实验: ① Base1+DB; ② Base1+DB+空间注意机制; ③ Base1+DB+通道注意机制; ④ Base1+DB+CS 联合注意模块, 即本文网络. 定量评估结果如表4所示, 可视化结果如图9所示.

由表4可以看出, 无论缺少通道注意力还是空间注意力, 去雾性能都有一定程度的下降, 且 DFAB 引入 CS 联合注意模块比不使用该模块性能提高明显, PSNR 提高 4.57 dB, SSIM 提高 0.008 8. 另外, 从图9

由表3可以观察到, 无论是去除 RB、DFAB、AFM 模块中的哪一个模块, 网络性能都有一定程度的下降. 其中, 网络去除 DFAB 模块比原去雾网络性能下降明显, PSNR 下降 10.11 dB, SSIM 下降 0.035 8. 另外, 网络去除 RB 残差块, PSNR 下降 3.78 dB; 网络去除 AFM 模块, PSNR 下降 1.19 dB. 从图8也可以看出, 组合②在红色墙面、桌面以及光线昏暗处都存在明显雾霾残留现象, 且整体对比度较低; 组合①和组合④在远处光线暗的区域细节模糊, 组合③和组合④在红色墙面处均有不同程度雾霾残留. 经对比可验证本文所提方法中各个模块的有效性.

展示的可视化效果来看, 缺少 CS 联合注意模块的网络图像恢复效果较差, 存在边缘细节模糊现象; 单独使用通道注意力或空间注意力后得到的无雾图效果较好, 但整体不如引入 CS 联合注意模块取得的效果好. 由此可以证明本文提出的 CS 联合注意模块以及特征增强注意块 DFAB 的有效性.

表4 DFAB 模块消融实验结果

网络结构组成	PSNR (dB)	SSIM
Base1+DB	31.24	0.9801
Base1+DB+空间注意机制	32.70	0.9833
Base1+DB+通道注意机制	33.07	0.9832
Base1+DB+CS联合注意模块 (本文)	<b>35.81</b>	<b>0.9889</b>

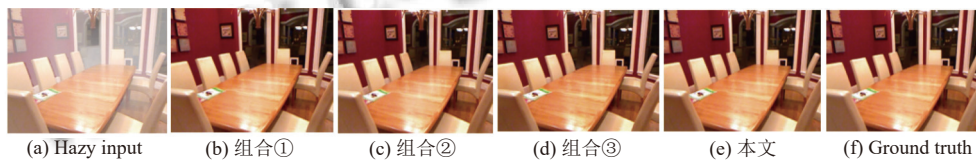


图9 DFAB 模块消融实验可视化结果

## 4 结束语

针对目前图像去雾算法存在颜色失真、对比度下降和雾霾残留等现象, 本文提出结合密集注意力的自适应特征融合图像去雾网络, 在编解码器结构中引入密集残差网络和 CS 联合注意模块构建的密集特征注意块, 使网络依次从通道维度和空间维度聚焦上下文特征信息, 更多地关注图像的局部细节信息, 防止雾霾

残留, 同时, 密集连接可以使网络提高特征重用, 具有很好的抗过拟合性能; 为充分融合不同级别的图像特征, 提出自适应特征融合模块, 使网络可以利用非相邻层级的特征信息, 避免浅层特征丢失. 本文网络不依赖于大气散射模型, 是一种端到端的图像去雾网络, 且参数量较小. 实验结果表明, 本文网络在合成有雾图像数据集和真实世界有雾图像上均可取得理想的去雾效果,

其恢复图像能够有效保留纹理细节,同时在颜色、亮度与对比度方面也最接近真实图像.尽管如此,本文网络还存在一些不足:模型的 *SSIM* 指标结果相对较低,在真实世界数据集上的性能仍有待提高.因此,后续工作中,将研究如何去除真实世界非均匀雾霾,构建更适用于真实场景的小数据集雾图的去雾网络,可考虑迁移学习等方法使模型更容易学习,实现更贴合实际的应用研究.

### 参考文献

- 1 Li JF, Zhuo L, Zhang H, *et al.* Effective data-driven technology for efficient vision-based outdoor industrial systems. *IEEE Transactions on Industrial Informatics*, 2020, 16(7): 4344–4354. [doi: [10.1109/TII.2019.2936467](https://doi.org/10.1109/TII.2019.2936467)]
- 2 赵茂军,郭凰,白俊峰,等.基于通道注意力与条件生成对抗网络的图像去雾. *计算机系统应用*, 2022, 31(11): 167–174. [doi: [10.15888/j.cnki.csa.008804](https://doi.org/10.15888/j.cnki.csa.008804)]
- 3 Xiong YY, Liu HX, Gupta S, *et al.* MobileDets: Searching for object detection architectures for mobile accelerators. *Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE, 2021. 3824–3833. [doi: [10.1109/CVPR46437.2021.00382](https://doi.org/10.1109/CVPR46437.2021.00382)]
- 4 朱钦权,王同罕,贾惠珍.基于雾密度感知增强的去雾图像质量评价方法. *计算机工程*, 2023, 49(6): 208–216, 226. [doi: [10.19678/j.issn.1000-3428.0066460](https://doi.org/10.19678/j.issn.1000-3428.0066460)]
- 5 王科平,杨艺,费树岷.雾霾图像清晰化算法综述. *智能系统学报*, 2023, 18(2): 217–230.
- 6 Cantor A. Optics of the atmosphere—Scattering by molecules and particles. *IEEE Journal of Quantum Electronics*, 1978, 14(9): 698–699. [doi: [10.1109/JQE.1978.1069864](https://doi.org/10.1109/JQE.1978.1069864)]
- 7 张晓东,秦娟娟,贾仲仲.多尺度 Retinex 在图像去雾算法中的应用研究. *西昌学院学报(自然科学版)*, 2021, 35(3): 60–65. [doi: [10.16104/j.issn.1673-1891.2021.03.013](https://doi.org/10.16104/j.issn.1673-1891.2021.03.013)]
- 8 Kim TK, Paik JK, Kang BS. Contrast enhancement system using spatially adaptive histogram equalization with temporal filtering. *IEEE Transactions on Consumer Electronics*, 1998, 44(1): 82–87. [doi: [10.1109/30.663733](https://doi.org/10.1109/30.663733)]
- 9 Hsu WY, Chen YS. Single image dehazing using wavelet-based haze-lines and denoising. *IEEE Access*, 2021, 9: 104547–104559. [doi: [10.1109/ACCESS.2021.3099224](https://doi.org/10.1109/ACCESS.2021.3099224)]
- 10 He KM, Sun J, Tang XO. Single image haze removal using dark channel prior. *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*. Miami: IEEE Press, 2009. 1956–1963. [doi: [10.1109/CVPR.2009.5206515](https://doi.org/10.1109/CVPR.2009.5206515)]
- 11 Zhu QS, Mai JM, Shao L. A fast single image haze removal algorithm using color attenuation prior. *IEEE Transactions on Image Processing*, 2015, 24(11): 3522–3533. [doi: [10.1109/TIP.2015.2446191](https://doi.org/10.1109/TIP.2015.2446191)]
- 12 Meng GF, Wang Y, Duan JY, *et al.* Efficient image dehazing with boundary constraint and contextual regularization. *Proceedings of the 2013 IEEE International Conference on Computer Vision*. Sydney: IEEE Press, 2013. 617–624. [doi: [10.1109/ICCV.2013.82](https://doi.org/10.1109/ICCV.2013.82)]
- 13 Cai BL, Xu XM, Jia K, *et al.* DehazeNet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing*, 2016, 25(11): 5187–5198. [doi: [10.1109/TIP.2016.2598681](https://doi.org/10.1109/TIP.2016.2598681)]
- 14 Li BY, Peng XL, Wang ZY, *et al.* AOD-Net: All-in-one dehazing network. *Proceedings of the 2017 IEEE International Conference on Computer Vision*. Venice: IEEE Press, 2017. 4780–4788. [doi: [10.1109/ICCV.2017.511](https://doi.org/10.1109/ICCV.2017.511)]
- 15 Chen DD, He MM, Fan QN, *et al.* Gated context aggregation network for image dehazing and deraining. *Proceedings of the 2019 IEEE Winter Conference on Applications of Computer Vision*. Waikoloa: IEEE Press, 2019. 1375–1383. [doi: [10.1109/WACV.2019.00151](https://doi.org/10.1109/WACV.2019.00151)]
- 16 Qin X, Wang ZL, Bai YC, *et al.* FFA-Net: Feature fusion attention network for single image dehazing. *Proceedings of the 34th AAAI Conference on Artificial Intelligence*. New York: AAAI Press, 2020. 11908–11915.
- 17 Dong Y, Liu YH, Zhang H, *et al.* FD-GAN: Generative adversarial networks with fusion-discriminator for single image dehazing. *Proceedings of the 34th AAAI Conference on Artificial Intelligence*. New York: AAAI Press, 2020. 10729–10736. [doi: [10.1609/aaai.v34i07.6701](https://doi.org/10.1609/aaai.v34i07.6701)]
- 18 Wu HY, Qu YY, Lin SH, *et al.* Contrastive learning for compact single image dehazing. *Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE Press, 2021. 10546–10555. [doi: [10.1109/CVPR46437.2021.01041](https://doi.org/10.1109/CVPR46437.2021.01041)]
- 19 Song YD, Zhou Y, Qian H, *et al.* Rethinking performance gains in image dehazing networks. *arXiv:2209.11448*, 2022.
- 20 Zheng Y, Zhan JH, He SF, *et al.* Curricular contrastive regularization for physics-aware single image dehazing. *Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver: IEEE Press, 2023. 5785–5794. [doi: [10.1109/CVPR52729.2023.00560](https://doi.org/10.1109/CVPR52729.2023.00560)]
- 21 Guo CL, Yan QX, Anwar S, *et al.* Image dehazing transformer with transmission-aware 3D position embedding. *Proceedings of the 2022 IEEE/CVF Conference on Computer*

- Vision and Pattern Recognition. New Orleans: IEEE Press, 2022. 5802–5810. [doi: [10.1109/CVPR52688.2022.00572](https://doi.org/10.1109/CVPR52688.2022.00572)]
- 22 Song YD, He ZQ, Qian H, *et al.* Vision transformers for single image dehazing. *IEEE Transactions on Image Processing*, 2023, 32: 1927–1941. [doi: [10.1109/TIP.2023.3256763](https://doi.org/10.1109/TIP.2023.3256763)]
- 23 Zhang HY, Cissé M, Dauphin YN, *et al.* mixup: Beyond empirical risk minimization. *Proceedings of the 6th International Conference on Learning Representations*. Vancouver: OpenReview.net, 2018. 1–13.
- 24 Li BY, Ren WQ, Fu DP, *et al.* Benchmarking single-image dehazing and beyond. *IEEE Transactions on Image Processing*, 2019, 28(1): 492–505. [doi: [10.1109/TIP.2018.2867951](https://doi.org/10.1109/TIP.2018.2867951)]
- 25 Ancuti CO, Ancuti C, Timofte R, *et al.* O-HAZE: A dehazing benchmark with real hazy and haze-free outdoor images. *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. Salt Lake City: IEEE Press, 2018. 867–8678. [doi: [10.1109/CVPRW.2018.00119](https://doi.org/10.1109/CVPRW.2018.00119)]
- 26 Liu XH, Ma YG, Shi ZH, *et al.* GridDehazeNet: Attention-based multi-scale network for image dehazing. *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision*. Seoul: IEEE Press, 2019. 7314–7323. [doi: [10.1109/ICCV.2019.00741](https://doi.org/10.1109/ICCV.2019.00741)]
- 27 Dong H, Pan JS, Xiang L, *et al.* Multi-scale boosted dehazing network with dense feature fusion. *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE Press, 2020. 2154–2164. [doi: [10.1109/CVPR42600.2020.00223](https://doi.org/10.1109/CVPR42600.2020.00223)]
- 28 Yi QS, Li JC, Fang FM, *et al.* Efficient and accurate multi-scale topological network for single image dehazing. *IEEE Transactions on Multimedia*, 2022, 24: 3114–3128. [doi: [10.1109/TMM.2021.3093724](https://doi.org/10.1109/TMM.2021.3093724)]

(校对责编: 牛欣悦)