



基于全局与局部感知网络的超高清图像去雾方法

郑卓然, 魏绎汶, 贾修一

引用本文:

郑卓然, 魏绎汶, 贾修一. 基于全局与局部感知网络的超高清图像去雾方法[J]. 智能系统学报, 2024, 19(1): 89–96.

ZHENG Zhuoran, WEI Yiwen, JIA Xiuyi. UHD image dehazing method based on global and local aware networks[J]. *CAAI Transactions on Intelligent Systems*, 2024, 19(1): 89–96.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202304013>

您可能感兴趣的其他文章

基于生成对抗网络的机载遥感图像超分辨率重建

Super-resolution reconstruction of airborne remote sensing images based on the generative adversarial networks

智能系统学报. 2020, 15(1): 74–83 <https://dx.doi.org/10.11992/tis.202002002>

基于注意力融合的图片描述生成方法

An image caption generation method based on attention fusion

智能系统学报. 2020, 15(4): 740–749 <https://dx.doi.org/10.11992/tis.201910039>

一种基于雨线主方向自适应的全局稀疏去雨模型

A global sparse rain removal model based on rain streaks main direction adaptation

智能系统学报. 2020, 15(2): 271–280 <https://dx.doi.org/10.11992/tis.201809042>

一种恢复图像自然色彩的重构方法

Reconstruction method for restoring the natural color of an image

智能系统学报. 2019, 14(5): 877–881 <https://dx.doi.org/10.11992/tis.201805011>

面向社群图像的显著区域检测方法

Salient region detection for social images

智能系统学报. 2018, 13(2): 174–181 <https://dx.doi.org/10.11992/tis.201706043>

一种特征字典映射的图像盲评价方法研究

Blind quality evaluation with image features codebook mapping

智能系统学报. 2018, 13(6): 989–993 <https://dx.doi.org/10.11992/tis.201805027>

DOI: 10.11992/tis.202304013

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20230801.0911.002>

基于全局与局部感知网络的超高清图像去雾方法

郑卓然, 魏绎汶, 贾修一

(南京理工大学 计算机科学与工程学院, 江苏 南京 210094)

摘要: 当前, 为实现图像全局建模的目的, 基于多层感知机 (multi-layer perceptron, MLP) 的模型通常需要将图像上的像素进行平铺, 之后实施一个自注意力机制或“混合”增强方案以获得图像的长范围依赖。然而, 这些方法通常消耗大量的计算资源来弥补图像重建丢失的空间拓扑信息。特别是对于超高清图像去雾任务, 大量堆积 MLP 的模型在资源受限的设备上执行一张超高清带雾图像时会出现内存溢出的问题。为了解决这个问题, 本文提出了一种可以在单个 GPU 上对分辨率为 4 k 的图像进行实时去雾 (110 f/s) 的模型, 该模型的建模过程中保持了图像空间结构信息, 同时具有低计算复杂度的优点。

关键词: 图像去雾; 超高清图像; 多层感知机; 空间拓扑信息; 局部特征提取; 全局特征提取; 深度学习; 实时去雾
中图分类号: TP181 **文献标志码:** A **文章编号:** 1673-4785(2024)01-0089-08

中文引用格式: 郑卓然, 魏绎汶, 贾修一. 基于全局与局部感知网络的超高清图像去雾方法 [J]. 智能系统学报, 2024, 19(1): 89-96.

英文引用格式: ZHENG Zhuoran, WEI Yiwen, JIA Xiuyi. UHD image dehazing method based on global and local aware networks[J]. CAAI transactions on intelligent systems, 2024, 19(1): 89-96.

UHD image dehazing method based on global and local aware networks

ZHENG Zhuoran, WEI Yiwen, JIA Xiuyi

(School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract: Current multilayer perceptron (MLP)-based models usually require flattening pixels on an image and subsequently enforce a self-attention mechanism or “Mix” enhancement scheme to achieve global modeling of images and obtain long-range dependence of the image. However, these approaches generally consume considerable computing resources to bridge the loss of spatial topological information in image reconstruction. Particularly for UHD image dehazing tasks, numerous stacked MLP models suffer from memory overflow when running a UHD-hazed image on a resource-constrained device. A novel model for real-time dehazing of 4 K images on a single GPU (110 fps) is proposed here to address this issue. This model is advantageous because it maintains spatial information of the raw image and has low computational complexity.

Keywords: image dehazing; UHD image; multilayer perceptron; spatial topology information; local feature extraction; global feature extraction; deep learning; real-time image dehazing

带有雾霾的图像具有低对比度和模糊的特性, 这会严重影响下游图像处理模型的表现, 例如行人检测、图像分割等。对此, 大量的单幅图像去雾方法被开发出来, 它们的目的在于把输入的带有雾霾的图像转换成一张清晰图像。然而, 伴随着移动设备和边缘设备对分辨率为 4 k 图像处理方法的不断增长, 现存的图像去雾的方法很少能高效地处理一张带雾的超高清

图像^[1]。

对于传统算法来说, 大量的研究人员专注于雾霾和环境的物理性质, 他们采用各种清晰的图像先验来规范解空间, 但这些方法通常需要复杂的迭代优化方案才能找到最佳解。而且, 这些手工制作的图像先验知识的复杂性远远不能满足实际应用的要求。例如 Tan^[2] 开创了在没有任何额外信息的情况下在单图像实现去雾的可能性。He 等^[3] 使用暗通道先验 (dark channel prior, DCP) 借助统计学来估计图像的雾霾以实现图像去雾。Zhu 等^[4] 提出了颜色衰减先验, 通过估计场景深

收稿日期: 2023-04-07. 网络出版日期: 2023-08-01.

基金项目: 国家自然科学基金项目 (62176123).

通信作者: 贾修一. E-mail: jiaxy@njut.edu.cn.

©《智能系统学报》编辑部版权所有

度来消除雾霾。Berman 等^[5]观察到,无雾图像的颜色可以很好地近似为 RGB 空间中形成紧密簇的数百种不同颜色,然后基于这一先验知识提出了一种去雾算法。Chen 等^[6]提出了一种改进的评价彩色图像去雾效果的方法。该方法考虑了对图像边缘信息的评估以及对颜色失真的评估。

最近,基于 CNN 的方法已被应用于图像去雾,并且与传统方法相比取得了显著的性能改进。早期的算法^[7-9]使用可学习的参数代替传统框架中的某些模块或步骤(例如估计透射图或大气光),并使用外部数据来学习参数。从那时起,更多的研究使用端到端的数据驱动的方法来消除图像雾化^[10-14]。例如,Cai 等^[7]提出了 DehazeNet 来生成端到端的传输图。Zhang 等^[15]将大气散射模型嵌入到网络中,允许 CNNs 同时输出传输图、大气光和去雾图像。GandelSman 等^[11]借助于图像先验知识提出了一种无监督的图像去雾方法。Chen 等^[16]在合成数据集中预先训练了去雾模型,之后使用无监督学习方法使用各种物理先验微调网络参数,以提高其在真实雾霾图像上的去雾性能。还有一系列研究放弃了传统的物理模型,并使用直接的端到端方法来生成去雾图像。Li 等^[8]设计了一个 AOD 网络,通过重新制定的大气散射模型直接生成去雾图像。Qu 等^[14]将去雾任务转换为图像到图像的转换任务,并增强了网络以进一步生成更逼真的无雾图像。尽管基于 CNN 的方法已经取得了最先进的结果,但它们通常需要堆叠更多的卷积层才能获得更好的性能,从而导致在资源受限的设备上计算成本过高。

除此之外,基于 MLP 的方法已被应用于图像增强任务,例如图像超分辨率^[17-18]、图像去噪^[19]和图像去雨^[20-21]。与 CNN 相比,这些方法在低运算量的基础上取得了更好的视觉效果。不幸的是,目前基于 MLP 的方法有 2 个主要限制。首先,上述方法将图像划分为多个块,以捕获图像上的全局感受野,导致图像像素之间的空间拓扑信息丢失;其次,图像去雾是一个高度不适定的问题,因此需要大量的 MLP 层或一些注意力机制来重建更好的高频细节。为此,这些结构和模块的大量堆叠会严重增加计算负担。例如,Uformer 结构^[22]只能使用 24 GB RAM 处理一张 360×360 分辨率的图像。

针对上述存在的问题,本文提出了一种不带有图像补丁的全局和局部感知网络。其中,全局感知网络基于 MLP-Mixer 的设计原则,在多尺度框架中捕获图像的全局特征。此外,局部信息的

抽取使用 U-Net 来捕捉图像的局部特征以弥补全局信息建模的不足。最后,通过融合全局和局部特征图生成一个高质量的系数张量,它用于输入图像的仿射变换。值得注意的是系数张量可以看作是一种注意力机制,它表示了带雾图像的局部区域应该有相似的变换。经过大量的实验分析表明,所提出的用于 UHD 图像去雾任务的全局感知网络具有两个优点:1)该模型能够有效地建模出图像的全局特性,同时保留了图像上的元素之间空间拓扑信息。2)全局特征和局部特征相辅相成,协同产生一张高质量的超高清去雾图像。本文算法有能力在单个 24 GB RAM 的 RTX 3090 上以 110 f/s 的速度处理一张 4 k 分辨率的图像,并实现最佳性能。值得注意的是该模型在 4KID 数据集中的峰值信噪比指标达到了 26.99 dB。

1 全局和局部感知网络的结构

图 1 给出了 4 k 分辨率图像去雾网络的架构,该网络主要由两个分支网络组成,一个全局信息提取网络和另一个是局部信息提取网络。

1.1 全局信息提取

传统的基于 MLP 的图像重构模型需要将图像分割成若干个块再进行特征抽取,这无疑会丢失图像的空间拓扑信息。灵感来自于 MLP-Mixer 的设计原则,本文设计了一个空间 MLP 混合器 (spatial-MLP-mixer, SMM)。具体来说, SMM 将完整的特征图 X 作为输入,其中特征图 X 的长度域,宽度域和通道域分别为 H 、 W 和 C , $X \in \mathbf{R}^{(C \times H \times W)}$ 。然后分别使用相同的投影矩阵和激活函数以“滚动的方式”对一张图像的宽度域、长度域和通道域进行非线性的投影。混合器块由尺寸相等的多层 MLP 组成,每层由 3 个 MLP 块组成。第 1 个块是图像的宽度混合 MLP,它作用于 X 的行,映射 $\mathbf{R}^W \mapsto \mathbf{R}^W$,并在所有行之间共享。第 2 个块是图像的长度混合 MLP,它作用于 X 的列(即它应用于转置的输入 X^T),映射 $\mathbf{R}^H \mapsto \mathbf{R}^H$,并在所有列之间共享。第 3 个块是图像的通道混合 MLP:它作用于 X 的通道维度,映射 $\mathbf{R}^C \mapsto \mathbf{R}^C$,并在所有通道之间共享。每个 MLP 块包含两个完全连接层和一个独立应用于输入数据张量每个维度的非线性层。具体如下:

$$Y_{*,*,i} = X_{*,*,i} + \omega_2 S(\omega_1 L(X)_{*,*,i}), i = 1, 2, \dots, W$$

$$Z_{*,j,*} = Y_{*,j,*} + \omega_4 S(\omega_3 L(Y)_{*,j,*}), j = 1, 2, \dots, H$$

$$U_{k,*,*} = Z_{k,*,*} + \omega_6 S(\omega_5 L(Z)_{k,*,*}), k = 1, 2, \dots, C$$

其中: L 表示层归一化, S 是 Sigmoid 函数, ω 表示全连接层参数。该结构的整体复杂性在图像中的像

素数上是线性的, 这与 ViT (vision transformer) 不同, ViT 的复杂性是二次的。SMM 可以通过“滚动”提取图像的空间域信息进行长范围依赖建模以更好地恢复图像的颜色与纹理信息。

除此之外, 多尺度特性也被考虑。多尺度特性是空间 MLP 学习高分辨率 (high resolution, HR) 图像的高质量特征的关键。为了实现更多的跨分辨率特征交互, 在 SMM 开始时以不同的尺度插入交叉分辨率特征信息。为了帮助低分辨率 (low resolution, LR) 特征保持更多图像细节和准确的位置信息, 该算法把低分辨率特征与高分

分辨率特征融合。HR 路径在 LR 路径中增加了更多的图像信息以减少信息损失, 并增强了反向传播过程中的梯度流, 以促进 LR 变换模块的训练。另一方面, 将 LR 特征合并到 HR 路径中, 以帮助模型获得具有更大感受野的抽象层次的特征。具体来说, 该网络有 3 种规模 (256、128 和 64) 的多尺度 SMM, 框架与 HRNet 相同。它始终保持高分辨率表示, 以获得空间准确的全局特征图。通过迭代融合由 HR 和 LR 子网络生成的特征来合成可靠的高分辨率特征。所有的图像下采样和上采样的方式都使用了双线性插值。

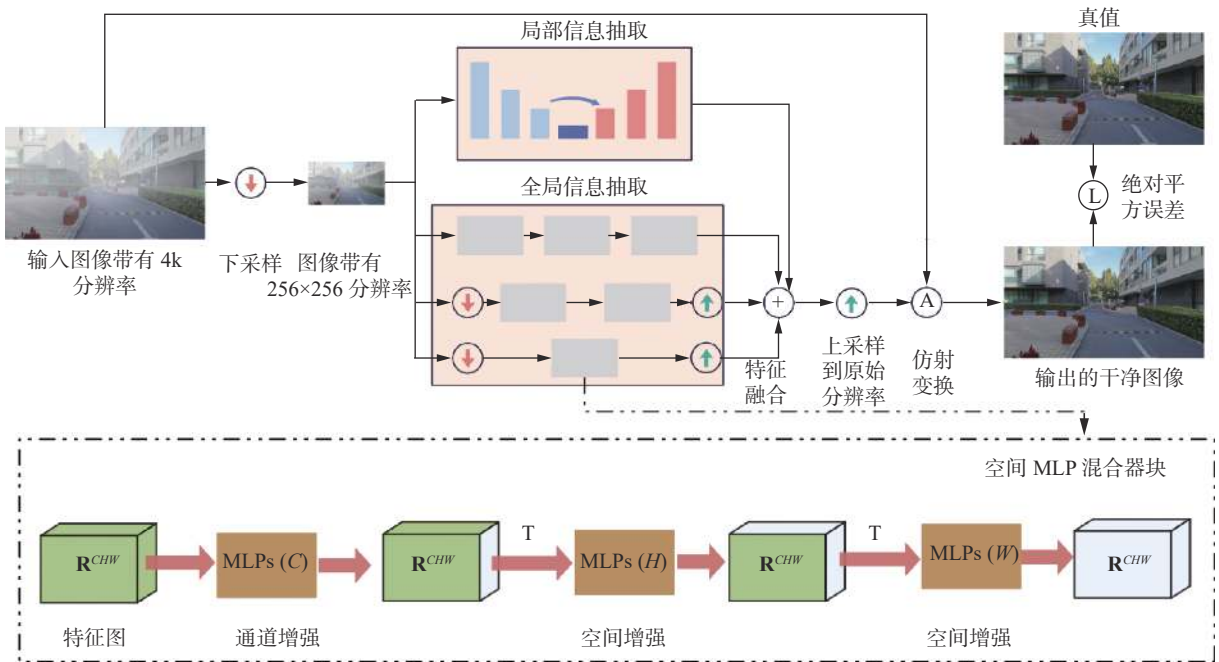


图 1 全局和局部感知网络框架

Fig. 1 Framework of global and local aware network

1.2 局部信息提取

为了进一步增强模型生成一张清晰的超高清去雾图像的能力, 该模型引入了图像的局部信息提取网络。该网络首先将 4 k 分辨率带雾输入降低到 256×256 的固定分辨率 (双线性插值的方法), 然后由 U-Net 获取其局部特征图。U-Net 添加了一个 3×3 卷积层, 将解码器最后一层的通道数从 64 映射到 3。局部提取模块通过堆叠卷积层和池化层, 可以更好地关注图像中的局部信息关系以消除冗余的特征信息。

此外, 图像局部信息的抽取可以用于恢复清晰的边缘特征, 这些特征可以通过依赖图像的短距离依赖进行恢复。如图 2(b) 所示, 本地信息提取模块的输出图像具有更清晰的边缘。相比之下, 图像的色彩信息不能仅根据该像素及其附近

像素的色彩信息进行恢复, 还需要考虑全局的长距离依赖才能正确恢复图像颜色。因此, 通过 SMM 来提取图像的长距离依赖色彩空间信息, 以更好地恢复图像颜色。如图 2(a) 所示, 全局信息提取模块更侧重于图像的颜色特征。

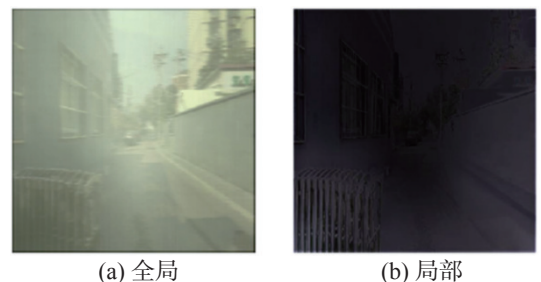


图 2 全局和局部分支归一化特征结果

Fig. 2 Results of normalized output feature maps of the global and the local branches

2 实验与结果分析

在本节中,通过对合成数据集和真实世界图像进行实验来评估所提出的方法。将所有结果与 9 种先进的去雾方法进行比较: AOD^[17]、PSD^[16]、DCP^[3]、CAP^[4]、NL^[5]、GCANet^[23]、MGBL^[1]、FDMHN^[24] 和 PFFNet^[25]。此外,还进行消融研究,以表明该网络在图像去雾任务上每个模块的有效性。

2.1 评价指标

为定量的评估去雾算法的表现,本文使用了峰值信噪比 P 和结构相似性 H 作为评估指标,其中 K 表示最大值, E 表示方差。

$$P = 10 \cdot \lg \frac{K^2}{E}$$

H 表示干净图像与噪声图像之间的均方差。

$$H(x, y) = \frac{(2\mu_x\mu_y + c)(2\sigma_{xy} + c)}{(\mu_x^2 + \mu_y^2 + c)(\sigma_x^2 + \sigma_y^2 + c)}$$

2.2 数据集

训练数据集总共包含 13 136 张雾化/真实图像。它包括来自 4KID 的 12 861 张包含建筑物、人物、车辆、道路等的图像和来自 I-HAZE 的 25 张室内场景图像和来自 O-HAZE 的 40 张室外场景图像进一步扩充了数据的多样性。相应地,实验对来自 4Kdehaze 的 200 张图像,来自 I-HAZE 的 5 张图像和来自 O-HAZE 的 5 张图像进行测试。

2.3 实验细节

该模型是使用 PyTorch 1.7 实现的,网络是使用 AdamW 优化器训练的。在这种情况下,一张分辨率为 512×512 的图像作为输入(输入到模型后会借助双线性插值被强行下采样到 256×256 的分辨率),并使用 8 的批量大小来训练网络。初始学习率设置为 0.001。整个模型的使用了 50 轮次的训练。

对于 DCP,将窗口大小设置为 60×60 用于测

试。对于去雾模型 PSD, GCANet 和 FDMHN,它们分别在 4KID、I-HAZE 和 O-HAZE 数据集上进行微调。网络使用 AdmaW 优化器进行训练,学习率为 0.000 1。特别是对于 PSD,本文使用作者提供的 PSD-MSDBN 模型系数进行微调。此外,对于 AOD、PFFNet 和 MGBL,应用 Adam 优化器并将学习率设置为 0.001 以训练网络。对于去雾算法 NL,灰度系数 γ 设置为 1 进行测试。

2.4 实验结果

所有方法都在 3 个数据集上进行评估,即 4KID、O-HAZE 和 I-HAZE 数据集。图 3 和图 4 中给出了在 4KID 数据集中的一张分辨率为 4 k 的图像和 I-HAZE 数据集中的一张图像的对比结果。可以观察到,传统的基于物理的方法(NL、DCP、CAP)倾向于过度增强结果,导致颜色失真。最近的深度模型(GCANet、FDMHN、AOD、PFFNet、MGBL)由于缺乏全局建模能力,结果中仍然存在一些模糊。虽然 PSD 的结构相似性优于本文算法,但局部与全局感知网络可以更快地处理分辨率为 4 k 的图像并获得更好的色彩结果。图 3(k)、4(k) 中局部与全局感知网络法生成的去雾结果接近图 3(l)、4(l) 中的真实无雾图像。表 1 表明了本文方法的有效性。同时,在同一台具有 NVIDIA 24GB RAM RTX 3 090 GPU 的机器上评估所有深度模型。运行时只是 GPU 的处理时间,不考虑 I/O 操作。4KID、I-HAZE 和 O-HAZE 数据集的平均运行时间如表 1 所示。传统方法(NL、DCP、CAP)需要解决复杂的函数,这不可避免地增加了计算成本。虽然一些轻量级网络(FDMHN、AOD、PFFNet、MGBL)可以实时消除分辨率为 4 k 的图像的雾霾,但它们的性能不如本文模型。此外,虽然一些大型网络(GCANet、PSD)实现了更好的性能,但它们无法实时去除单个分辨率为 4 k 的图像的雾霾。

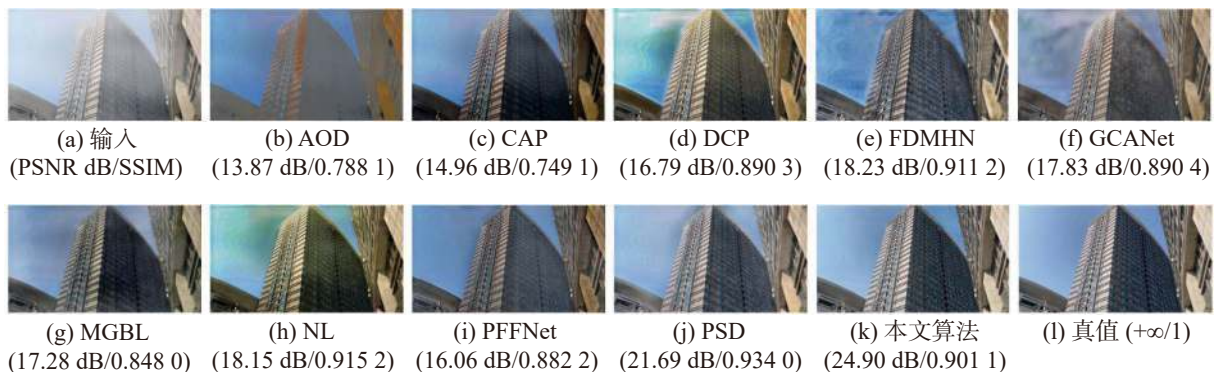


图 3 4KID 数据集上的测试结果

Fig. 3 Dehazed results on the 4KID test dataset

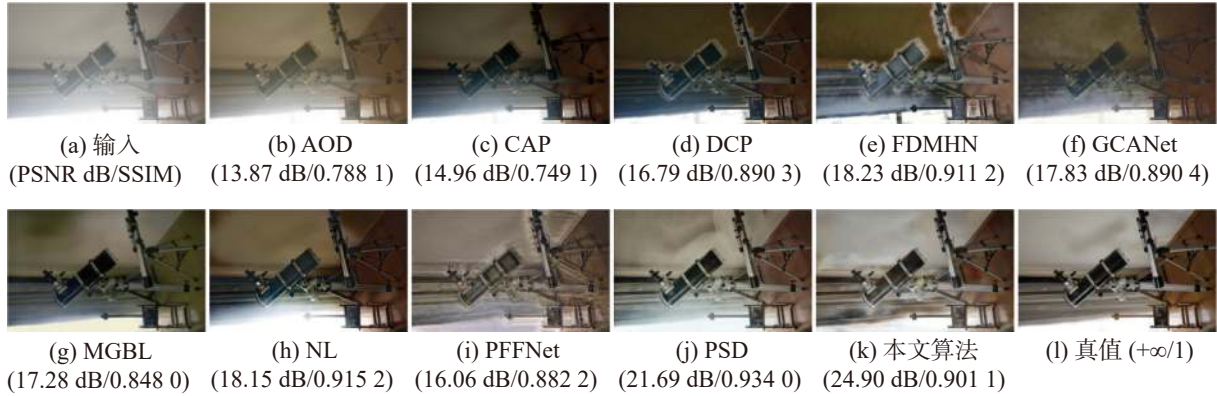


图 4 在 I-HAZE 数据集上的测试结果

Fig. 4 Dehazed results on the I-HAZE dataset

表 1 4KID、I-HAZE 和 O-HAZE 数据集上的定量评估

Table 1 Quantitative evaluation of the 4KID, I-HAZE and O-HAZE datasets

数据集	指标	输入	AOD	CAP	DCP	FDMHN	GCANet	MGBL	NL	PFFNet	PSD	本文算法
4KID	峰值信噪比/dB	10.87	14.34	12.38	15.21	17.83	15.38	22.97	13.83	15.69	23.19	26.99
	结构相似性	0.7506	0.8059	0.7269	0.8400	0.8878	0.8498	0.8955	0.8084	0.8580	0.9391	0.9000
	时间	—	1 ms	6 s	3 s	20 ms	1 s	9 ms	22 s	13 ms	382 ms	9 ms
O-HAZE	峰值信噪比/dB	13.36	18.09	15.20	16.57	18.55	16.35	19.26	14.65	17.66	19.77	20.54
	结构相似性	0.7109	0.7720	0.6241	0.7712	0.7730	0.7253	0.8179	0.6662	0.7453	0.8238	0.8108
	时间	—	5 ms	9 s	5 s	20 ms	2 s	16 ms	38 s	19 ms	2 s	13 ms
I-HAZE	峰值信噪比/dB	15.50	15.55	11.74	11.55	15.49	12.25	16.75	12.18	15.13	16.41	19.32
	结构相似性	0.8181	0.8271	0.5413	0.6710	0.7135	0.7447	0.8016	0.6788	0.7702	0.8352	0.8848
	时间	—	2 ms	9 s	4 s	19 ms	2 s	17 ms	55 s	16 ms	602 ms	14 ms

然后, 在真实世界的带雾图像上评估所提出的算法。首先, 在真实捕获的 4 k 分辨率带雾图像上与不同的先进方法进行比较。图 5 给出了两张具有挑战性的真实世界图像的结果的定性比较。如图所示, DCP 使去雾结果中的某些

区域变暗, CAP 和 PSD 遭受颜色失真, 而 AOD、GCANet、MGBL、FDMHN 和 PFFNet 生成的结果有一些残留的雾霾。相比之下, 本文算法能够生成逼真的颜色, 同时更好地消除雾霾, 如图 5(j) 所示。

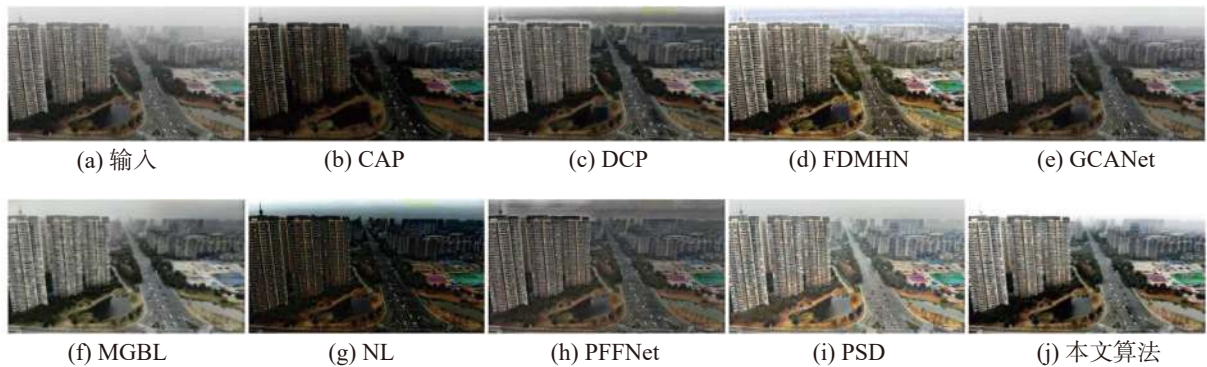


图 5 在真实 4 k 分辨率图像上的去雾结果

Fig. 5 Dehazed results on real-world 4 k resolution images

除了 4 k 分辨率图像, 在其他公共数据库下载的几个低分辨率带雾图像上评估了超高清去雾算法与其他的对比方法。去雾结果如图 6 所示。

可以看出, 除本文之外的所有型号都存在颜色失真。相反, 本文方法可以更好地消除雾霾并有效地产生逼真的色彩。

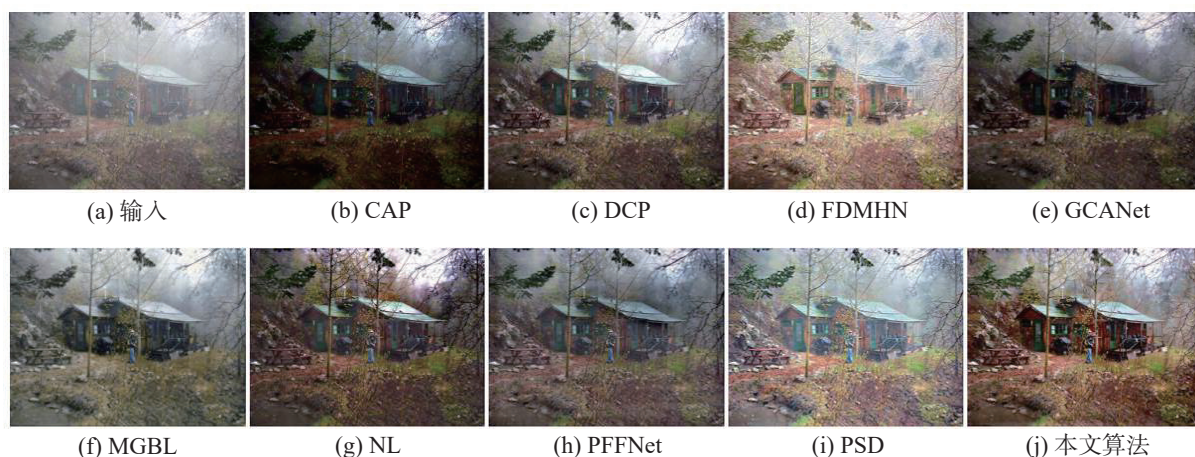


图 6 在低分辨率图像上的去雾结果

Fig. 6 Dehazed results on low-resolution hazy images of real-world

为了检验 SMM 的有效性,全局与局部网络还与 SwinIR^[19] 和 MLP-Mixer^[26] 进行了比较。所有 3 个模型应用大致相同数量的参数。SwinIR 和 MLP-Mixer 都需要将图像分割成块,导致空间拓扑信息的丢失和模型的图像增强能力的降低。此外, SwinIR 对 Transformer 的使用增加了其计算能力但减慢了模型的速度。如图 7 所示, MLP 混合器产生了模糊的结果,图像中存在可见的斑块,而 SwinIR 的输出不能完全消除雾霾,并且存在颜色失真。但是,本文提出的 SMM 能够更好恢复纹理和颜色。值得注意的是,对于大致相同数量的参数, SMM 是最快的,而 SwinIR 是最慢的。

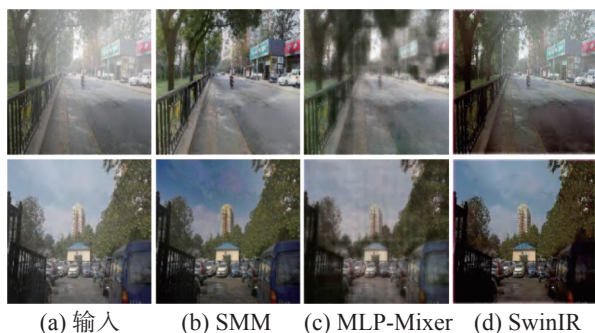


图 7 空间 MLP 混合器、MLP 混合器以及 SwinIR 效果对比
Fig. 7 The results of spatial MLP-mixer, MLP-mixer and SwinIR

2.5 消融实验

为了表明所提出的网络中引入的每个模块的

有效性,进行了一项消融研究,涉及以下 3 个实验:全局分支的有效性,该模型移除了全局特征提取分支并直接回归图像以获得最终输出;局部分支的有效性,该模型移除 UNet,直接将图像回归到多尺度空间 MLP 混合器中,以获得最终结果;多尺度的有效性,比较了分别使用单一尺度和两个尺度的效果,同时保持相同数量的参数。

如表 2 和图 8 所示,局部分支依靠图像的局部特征,在两个指标上取得了较好的结果,但其颜色恢复能力仍然不足。仅使用全局分支并不能给出令人满意的结果,但其更好地提取全局信息的能力可以增强局部分支对图像颜色的恢复。值得注意的是,对不同尺度的 SMM 进行消融实验时,该模型为单尺度和双尺度 SMM 堆叠了更多的 MLP 层,以达到与多尺度 SMM 相似数量的参数。显然,出色的多尺度性能是由于多分辨率图像提供的丰富细节。

表 2 全局分支、局部分支以及多尺度的消融实验
Table 2 Ablation studies of global branch, local branch and multi-scale

方法	峰值信噪比/dB	结构相似性
全局分支	18.61	0.8022
局部分支	22.82	0.8445
单尺度SMM	25.77	0.8498
双尺度SMM	26.03	0.8977
本文	26.99	0.9000

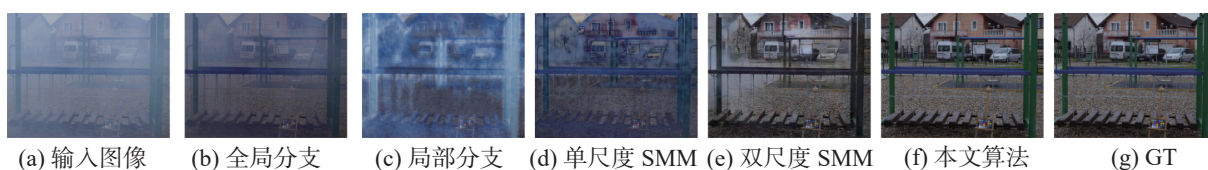


图 8 在低分辨率图像上的消融结果

Fig. 8 Dehazed results on low-resolution hazy images

3 结束语

本文提出了一种具有全局和局部感知的超高清分辨率图像去雾的新模型。该模型的关键是使用全局特征提取分支的空间 MLP 混合器。空间 MLP 混合器可以帮助模型从超高清分辨率的 (4k) 图像中恢复颜色特征。使用局部特征分支来恢复高质量的细节特征, 为图像去雾提供丰富的纹理信息。定量和定性结果表明, 该网络在准确性和推理速度方面与先进的去雾方法相比更好, 并在真实世界的 4k 雾霾图像上产生了视觉上令人满意的结果。

参考文献:

- [1] ZHENG Zhuoran, REN Wenqi, CAO Xiaochun, et al. Ultra-high-definition image dehazing via multi-guided bilateral learning[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 16180–16189.
- [2] TAN R T. Visibility in bad weather from a single image[C]//2008 IEEE Conference on Computer Vision and Pattern Recognition. Anchorage: IEEE, 2008: 1–8.
- [3] HE Kaiming, SUN Jian, TANG Xiaoou. Single image haze removal using dark channel prior[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2011, 33(12): 2341–2353.
- [4] ZHU Qingsong, MAI Jiaming, SHAO Ling. A fast single image haze removal algorithm using color attenuation prior[J]. *IEEE transactions on image processing: a publication of the IEEE signal processing society*, 2015, 24(11): 3522–3533.
- [5] BERMAN D, TREIBITZ T, AVIDAN S. Non-local image dehazing[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 1674–1682.
- [6] 陈珂, 柯文德, 许波, 等. 改进的彩色图像去雾效果评价方法 [J]. *智能系统学报*, 2015, 10(5): 803–809.
CHEN Ke, KE Wende, XU Bo, et al. An improved assessment method for the color image defogging effect[J]. *CAAI transactions on intelligent systems*, 2015, 10(5): 803–809.
- [7] CAI Bolun, XU Xiangmin, JIA Kui, et al. DehazeNet: an end-to-end system for single image haze removal[J]. *IEEE transactions on image processing: a publication of the IEEE signal processing society*, 2016, 25(11): 5187–5198.
- [8] LI Boyi, PENG Xiulian, WANG Zhangyang, et al. AOD-net: all-in-one dehazing network[C]//2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 4780–4788.
- [9] REN Wenqi, PAN Jinshan, ZHANG Hua, et al. Single image dehazing via multi-scale convolutional neural networks with holistic edges[J]. *International journal of computer vision*, 2020, 128(1): 240–259.
- [10] DENG Zijun, ZHU Lei, HU Xiaowei, et al. Deep multi-model fusion for single-image dehazing[C]//2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2020: 2453–2462.
- [11] GANDELSMAN Y, SHOCHER A, IRANI M. “double-DIP”: unsupervised image decomposition via coupled deep-image-priors[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2020: 11018–11027.
- [12] HONG Ming, XIE Yuan, LI Cuihua, et al. Distilling image dehazing with heterogeneous task imitation[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 3459–3468.
- [13] LI Runde, PAN Jinshan, LI Zechao, et al. Single image dehazing via conditional generative adversarial network[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 8202–8211.
- [14] QU Yanyun, CHEN Yizi, HUANG Jingying, et al. Enhanced Pix2pix dehazing network[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2020: 8152–8160.
- [15] ZHANG He, PATEL V M. Densely connected pyramid dehazing network[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 3194–3203.
- [16] CHEN Zeyuan, WANG Yangchao, YANG Yang, et al. PSD: principled synthetic-to-real dehazing guided by physical priors[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 7176–7185.
- [17] LIANG Jingyun, CAO Jiezhong, SUN Guolei, et al. SwinIR: image restoration using swin transformer[C]//2021 IEEE/CVF International Conference on Computer Vision Workshops. Montreal: IEEE, 2021: 1833–1844.
- [18] YANG Fuzhi, YANG Huan, FU Jianlong, et al. Learning texture transformer network for image super-resolution[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 5790–5799.

- [19] LUTHRA A, SULAKHE H, MITTAL T, et al. Eformer: edge enhancement based transformer for medical image denoising[EB/OL]. (2021-09-06)[2023-04-07]. <https://arxiv.org/abs/2109.08044>.
- [20] CHEN Hanting, WANG Yunhe, GUO Tianyu, et al. Pre-trained image processing transformer[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 12294-12305.
- [21] CHEN Xiang, LI Hao, LI Mingqiang, et al. Learning A sparse transformer network for effective image deraining[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023: 5896-5905.
- [22] WANG Zhendong, CUN Xiaodong, BAO Jianmin, et al. Uformer: a general U-shaped transformer for image restoration[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022: 17662-17672.
- [23] CHEN Dongdong, HE Mingming, FAN Qingnan, et al. Gated context aggregation network for image dehazing and deraining[C]//2019 IEEE Winter Conference on Applications of Computer Vision. Waikoloa: IEEE, 2019: 1375-1383.
- [24] 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. Seattle: IEEE, 2020: 1994-2001.
- [25] MEI Kangfu, JIANG Aiwen, LI Juncheng, et al. Progressive feature fusion network for realistic image dehazing[C]//Asian Conference on Computer Vision. Cham: Springer, 2019: 203-215.
- [26] TOLSTIKHIN I, HOULSBY N, KOLESNIKOV A, et al. MLP-mixer: an all-MLP architecture for vision[EB/OL]. (2021-05-30)[2023-04-07]. <https://arxiv.org/abs/2105.01601.pdf>.

作者简介:



郑卓然, 博士研究生, 主要研究方向为深度学习和图像增强。E-mail: zhengzr@njust.edu.cn。



魏绎汶, 硕士研究生, 主要研究方向为图像增强。E-mail: weiyw@njust.edu.cn。



贾修一, 教授, 博士生导师, CCF 高级会员, 主要研究方向为机器学习、粒计算和数据挖掘。主持国家自然科学基金项目 3 项, 参与重点研发计划项目、国家自然科学基金企业联合基金重点项目等多项。发表学术论文 100 余篇。E-mail: jiaxy@njust.edu.cn。