

# 并行多尺度注意力映射图像去雾算法

袁姮, 颜廷昊\*

辽宁工程技术大学软件学院, 辽宁 葫芦岛 125105

**摘要** 针对当前去雾算法中容易产生的图像伪影、图像颜色失真、图像细节模糊不清等问题,提出一种并行多尺度注意力映射图像去雾算法,通过端到端方式以编码器解码器结构实现图像去雾。在编码阶段,采用连续下采样层降低特征维度,避免过拟合。在特征转换阶段,采用并行分支结构设计并行多尺度注意力映射模块,使模型能够在关注图像重要特征的同时充分利用多尺度特征,并通过并行连接选择性特征融合模块有效收集图像空间结构信息。解码阶段,采用上采样层重构图像,并通过上下采样层融合更好地保留图像边缘信息。实验结果表明,该算法在合成雾天数据集以及真实雾天图像上均具有较好的去雾效果,相较于传统去雾算法,可更好地保留图像细节,具备较好的色彩保持度。

**关键词** 图像去雾; 卷积神经网络; 并行分支结构; 多尺度映射; 注意力机制

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP230921

## Parallel Multi Scale Attention Mapping Image Dehazing Algorithm

Yuan Heng, Yan Tinghao\*

College of Software, Liaoning Technical University, Huludao 125105, Liaoning, China

**Abstract** Problems such as image color distortion, blurred image details, and image artifacts are prone to occur in the current dehazing algorithm. In order to solve the above problems, an image dehazing algorithm with parallel multi scale attention mapping is proposed. The algorithm achieves image defogging through an end-to-end encoder decoder structure. In the encoder stage, the continuous downsampling layer is used to reduce feature dimension and avoid over-fitting. In the feature transformation stage, a parallel multi scale attention mapping block with a parallel branch structure is designed, so that the model can make full use of multi scale features while focusing on important features of the image, and effective collection of image spatial structure information by connecting selective feature fusion block in parallel. In the decoding stage, the upsampling layer is used to reconstruct the image, and through skip connections of up and down sampling to better preserve image edge information. Experimental results show that the algorithm has better dehazing effects on both synthetic hazy datasets and real hazy images. Compared with traditional dehazing methods, this algorithm better preserves image details and has better color retention.

**Key words** image dehazing; convolutional neural network; parallel branch structure; multi scale mapping; attention mechanism

## 1 引言

在雾、雾霾、烟尘等天气条件下拍摄的图像往往会存在图像细节模糊不清、图像颜色失真、对比度低等问题,而低质量图像会对后续高级视觉处理任务产生严重影响<sup>[1]</sup>。因此,通过图像去雾算法消除图像中雾气影响,恢复清晰高质量图像非常重要。

传统去雾算法主要基于图像增强技术实现图像去雾,常采用同态滤波<sup>[2]</sup>、直方图均衡化<sup>[3]</sup>以及 Retinex 理

论<sup>[4-5]</sup>等增强图片对比度与图像特征,达到去除雾气、恢复清晰图像的目的。韦春苗等<sup>[6]</sup>提出一种结合多尺度 Retinex 理论与小波变换的图像融合算法,采用多尺度 Retinex 理论增强雾天图像,并利用 db5 小波基对雾天图像与增强图像的亮度与饱和分量进行处理,输出去雾图像。方丹阳等<sup>[7]</sup>提出一种自适应动态范围 CLAHE 的雾天图像增强算法,通过改进自适应直方图均衡化算法,引入自适应参数,获取自适应阈值,增强图像对比度。同时由同态滤波改善图像过亮或过暗

收稿日期: 2023-03-22; 修回日期: 2023-05-03; 录用日期: 2023-06-01; 网络首发日期: 2023-06-11

基金项目: 国家自然科学基金(61172144)、辽宁省自然科学基金(20170540426)、辽宁省教育厅重点基金(LJYL049)。

通信作者: isaebellayth@163.com

区域,并结合二者获得清晰图像。基于图像增强去雾算法操作简单,但在增强过程中容易产生其他噪声,导致处理后图像视觉效果不理想。

随着人们对光学成像本质的不断研究,基于物理模型的图像去雾算法逐渐被广泛应用。基于物理模型图像去雾算法大多利用大气散射模型<sup>[8-10]</sup>,通过反演大气光值和透射图估计获得清晰图像。He 等<sup>[11]</sup>提出的基于暗通道先验(DCP)的图像去雾算法通过先验知识求解透射矩阵,由软抠图算法优化透射率,最后通过大气散射模型恢复清晰图像。曾接贤等<sup>[12]</sup>提出双边滤波与暗通道结合的图像保边去雾算法,对 DCP 算法加以改进,利用双边滤波方式求解介质透射率,降低了 DCP 算法的时间复杂度。基于 DCP 的算法在大多数场景能够可靠地计算传输矩阵,但当图像中出现白墙、天空等与大气光类似场景时,DCP 容易失效,造成图像对比度过强、颜色失真等问题。

目前,随着深度学习在不同领域的快速发展,许多基于深度学习的图像去雾算法相继被提出。Cai 等<sup>[13]</sup>提出用于介质传输估计的去雾网络模型 DehazeNet,通过卷积神经网络深层架构估计透射图,再由大气散射模型恢复无雾图像。Ren 等<sup>[14]</sup>提出的基于多尺度卷积神经网络去雾(MSCNN)算法,通过多尺度卷积神经网络更好地学习特征,由粗比例尺网络估计传输图,再由细比例尺网络局部细化去雾效果。Li 等<sup>[15]</sup>提出的一体化轻量级去雾网络模型(AOD-NET)通过重新配制大气散射模型,将传输矩阵与大气光值统一为一个参数,并由卷积神经网络进行估算,减小传输矩阵与大气光值单独估算过程中产生的误差。但通过卷积神经网络估计传输图与大气光值的方式容易忽略图像自身结构信息,处理后图像容易产生伪影。

为更好提取图像自身结构信息,由卷积神经网络直接输出清晰图像的端到端去雾算法受到越来越多关注。其中,编解码架构作为深度学习中常见框架,被广泛应用于端到端去雾算法中。Zhao 等<sup>[16]</sup>提出一种基于生成对抗网络的编解码器架构图像去雾算法(AEDGAN),通过注意力模块、蒸馏模块、增强模块构建生成网络,有效提取局部和全局特征信息,并通过多尺度判别网络进一步提升去雾性能。Feng 等<sup>[17]</sup>提出的基于 U-Net<sup>[18]</sup>的残差网络(URNET)将 ResNet<sup>[19]</sup>构建块嵌入编码器模块和解码器模块之间的连接处来提高网络性能。Bianco 等<sup>[20]</sup>提出一种使用深度编解码结构的高分辨率图像去雾算法(HR-Dehazer),并通过自定义损失函数,迫使网络考虑输入图像的语义,促进局部结构之间一致性。上述算法虽能够在一定程度上提升去雾处理后的图像质量,但在特征提取时缺乏全局特征与局部特征之间的联系,容易导致去雾不彻底。为解决以上问题,许多去雾算法将注意力机制引入模型来提升去雾效果。Zhang 等<sup>[21]</sup>提出的基于金字塔通道的图像去雾特征注意网络(PCFAN),利用不同层次特征之间

的互补性,通过金字塔结构结合通道注意力机制,捕获重要的通道映射进行特征重建,最终实现图像去雾。曹锐虎等<sup>[22]</sup>提出一种新型去雾网络模型(DRS-Dehaze Net),结合 Resnet 优势,设计了密集残差架构并引入注意力机制完成特征信息的提取与再分配,最后通过特征融合输出无雾图像。Jia 等<sup>[23]</sup>提出的元注意去雾网络(MADN),由特征提取和权重生成两部分构成元注意力模块,通过元注意力模块,由当前输入的有雾图像自动选择最合适的去雾网络结构来恢复清晰图像。

上述去雾算法虽取得一定成效,但在图像特征提取和细节的保留上仍存在缺陷,部分算法参数量大、模型复杂度高且处理后图像存在偏色失真现象。针对以上问题,本文提出一种并行多尺度注意力映射图像去雾算法,以编码器解码器结构实现图像去雾,该网络较为轻量化且不依赖于物理模型,通过卷积神经网络由有雾图像恢复清晰无雾图像。实验结果表明,所提算法具有较好的去雾效果,且处理速度快、图像恢复质量高,处理后图像更加符合人类视觉感知。

## 2 并行多尺度注意力映射图像去雾算法

在如 HR-Dehazer 算法等采用深度编解码结构图像去雾算法中,常利用大量下采样层操作去除冗余特征,捕获图像关键信息,并通过对应数量上采样层重构图像。但在采用下采样层缩减特征图尺寸的过程中,不可避免会造成图像信息损失,且图像包含信息量越大,下采样过程中对图像细节信息的损失就越大,导致解码器无法较好地从剩余的不完全信息中重构图像。因此,提出并行多尺度注意力映射图像去雾算法,从如何降低编解码阶段图像信息损失、提升已保留图像特征信息的利用率这两个角度出发,从多方面对深度编解码结构加以改进,使模型能够在降低特征维度的同时保留有效信息,具体结构如图 1 所示。

为降低编解码阶段图像信息损失,所提算法在深度编解码结构基础上减少上下采样层数,保留 3 层上下采样层对图像进行上下采样操作,去除冗余特征,保留图像关键信息,使模型在降低计算量与减少图像信息损失之间达到相对平衡的状态。同时将相同维度上下采样层进行跳跃连接。相较于单一上下采样层拼接,跳跃连接能够更好地融合深浅层特征,保留更多语义信息与图像细节。其次,设计选择性特征融合模块(SFFB),将选择性特征融合模块并行连接于下采样操作前,通过对原始尺寸特征图进行选择与聚合操作来更好地提取图像空间结构信息,并将提取到的特征信息与上采样输出特征融合,补偿编解码阶段造成的图像信息损失。

为提升已保留图像特征信息的利用率,在编码器与解码器阶段之间设计并行多尺度注意力映射模块

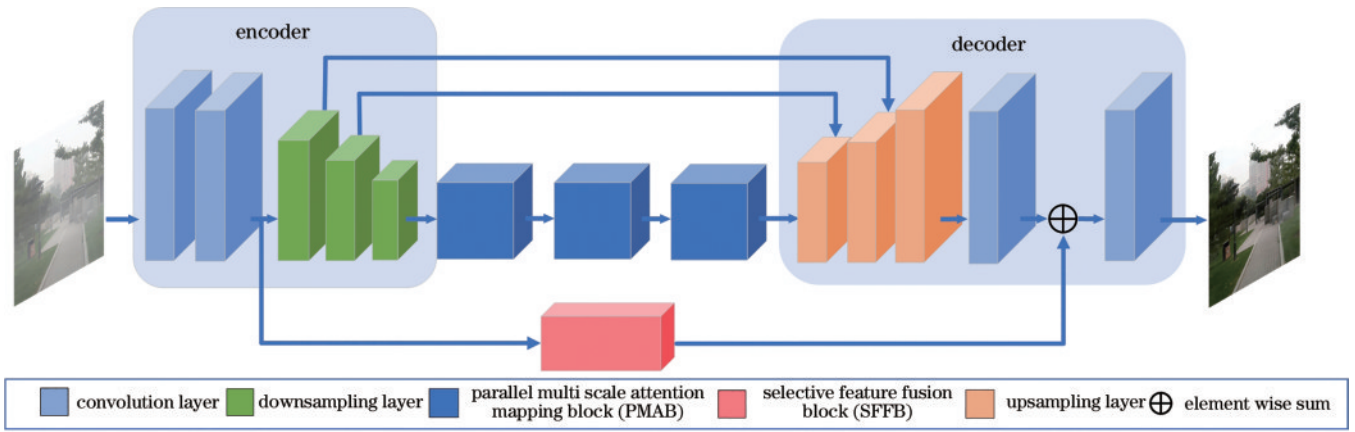


图 1 所提网络结构

Fig. 1 Proposed network structure

(PMAB),该模块由并行多尺度映射与注意力残差块结合而来,通过并行多尺度注意力映射模块对编码器阶段输入进行特征转换,使输入特征能够得到充分利用。并行多尺度注意力映射模块采用并行分支结构提取并融合多尺度特征信息,通过引入双注意力融合机制,使模型在充分利用多尺度特征的同时能够重点关注图像的有用特征。并行多尺度注意力映射模块能够为图像全局特征与局部特征建立联系,使模块在产生每个输出时,都能够充分利用输入图像包含的特征信息。

### 2.1 并行多尺度注意力映射模块

PMAB结构如图2所示。首先,在残差块中采用混合空洞卷积<sup>[24]</sup>来增大模型感受野。其次,通过并行连接混合空洞卷积残差块与多尺度映射,获取不同尺寸特征映射,并将不同尺寸特征通过聚合操作进行特征融合。同时,在并行多尺度映射中引入双注意力融合机制来自适应校正图像通道特征与空间特征,增加模型对图像有用特征关注度<sup>[25]</sup>。模型在获取图像关键信息的同时,又充分利用图像的局部特征与全局特征,增加了网络模型宽度,丰富了网络模型的表达能力。

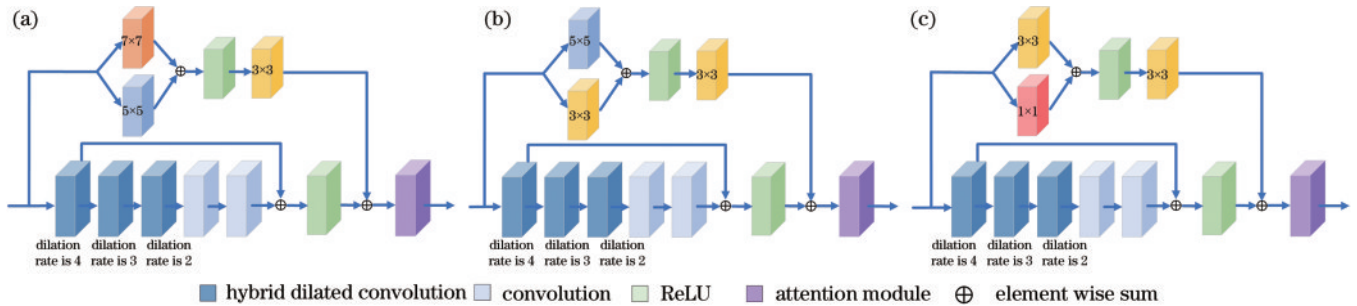


图 2 PMAB 结构。(a)模块 1;(b)模块 2;(c)模块 3

Fig. 2 Structure of PMAB. (a) Module 1; (b) module 2; (c) module 3

为进入残差块进行特征学习的同时能够保留更多图像特征信息,采用不同大小卷积核提取并融合不同尺度特征后与残差块并行连接。并行分支结构能够在同一层级获取不同感受野特征,融合后传递到下一层,可以更加灵活地降低计算量,提升模型性能。因此 PMAB 采用并行分支结构,通过并行连接 3 组不同大小卷积核,收集上一层输入的不同特征,再将所获特征融合,丰富输出特征图中包含的语义信息。

相较于深度学习模型中常用的多个小卷积核堆叠,大卷积核能够更直接高效地拓展有效感受野,同时规避模型深度增加时的优化难题<sup>[26]</sup>,但尺寸过大的卷积核容易造成模型计算量的较大增长,为平衡模型性能与计算量,PMAB 将卷积核按大小分为  $1 \times 1$  和  $3 \times 3$ 、 $3 \times 3$  和  $5 \times 5$ 、 $5 \times 5$  和  $7 \times 7$  等 3 组,由于浅层图像的

分辨率较高,图像像素值的方差较小,因此将  $5 \times 5$  和  $7 \times 7$  设置为 1 组,  $3 \times 3$  和  $5 \times 5$  为第 2 组,  $1 \times 1$  和  $3 \times 3$  为第 3 组,以此更完整地获取浅层特征信息,同时增加网络模型结构的多样性。

为确保网络能够捕获更多图像特征,在 PMAB 模块中引入注意机制,在训练过程中发现,注意力模块单独与多尺度映射模块或混合空洞卷积残差块分支连接时,所获权重概括性较大,权重分布较为集中,导致模型对局部信息过于关注,容易造成去雾不均匀现象。因此 PMAB 模块将注意力模块设置于混合空洞卷积残差块与多尺度映射的并行连接之后,利用注意力机制重构混合空洞卷积残差块与多尺度映射融合后的输出特征,以不同维度为并行多尺度映射输出特征分配权重,提升模型对于多尺度特征的关注度,提高注意力

的准确性。

### 2.1.1 混合空洞卷积残差块

为更好提取图像特征,增大模型感受野,采用混合空洞卷积残差块,当叠加多个相同空洞率的空洞卷积时,会造成空间层次和局部信息连续性丢失的网格效应,如图 3 所示。为解决网格效应问题,所提算法设置了无公因数关系的空洞率,使模型在获得更多密集数据的同时,减少图像信息丢失,保证了图像信息的连续性。混合空洞卷积相较标准卷积,能够为模型提供更大感受野,卷积层感受野计算公式为

$$r_n = r_{n-1} + (k-1) \prod_{i=1}^{n-1} s_i, \quad (1)$$

式中: $r_n$ 表示当前层感受野大小; $r_{n-1}$ 表示上一层感受野大小; $k$ 表示卷积核大小; $s_i$ 表示卷积层步长。若残差块采用步长为 1 的 5 层  $3 \times 3$  大小卷积核常规卷积顺序连接,则通过式(1)可求得该模块感受野大小为  $11 \times 11$ 。若残差块采用 5 层步长为 1 的  $3 \times 3$  大小卷积核且空洞率分别为 4、3、2、1、1 的混合空洞卷积连接,则通过式(1)可求得采用混合空洞卷积的模块感受野大小为  $23 \times 23$ 。由此可见,在相同深度网络模型中采用混合空洞卷积能够获得更大感受野,提升模型特征提取能力。

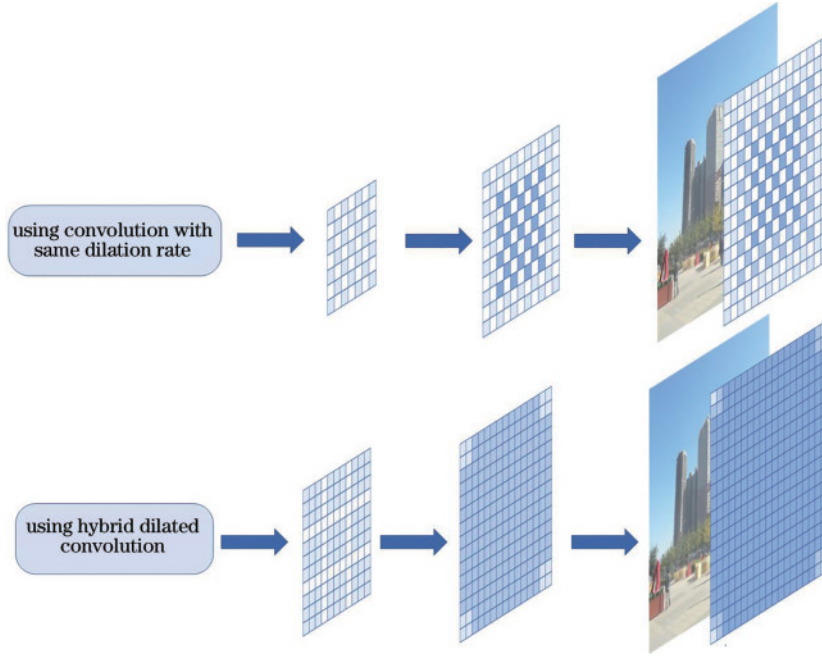


图 3 混合空洞卷积感受野对比

Fig. 3 Receptive field of hybrid dilated convolution contrast

### 2.1.2 双注意力融合机制

与常见采用单注意力机制图像去雾算法不同,受 convolutional block attention module (CBAM)<sup>[27]</sup> 注意力机制启发并加以改进,所提算法采用通道注意力与空间注意力相结合的双注意力融合机制,相较于单注意力机制,双注意力融合机制可以兼顾通道维度与空间维度的重要特征。通道注意力模块能够为每个通道上的特征图都施加一个权重,用于表示该通道与关键信息的相关程度,提高模型对通道关键信息的敏感度。空间注意力模块主要用于获取图像信息量最大区域,是对通道注意力模块的补充,将通道注意力模块与空间注意力模块顺序组合,能够使网络更加关注图像高频区域和边缘细节特征,能够更好消除雾气影响。注意力模块结构如图 4 所示。

输入特征首先经过最大池化层处理,相较于通道注意力中原本采用的平均池化层,最大池化层能够减小由卷积层参数误差造成估计均值的偏移,能够更好保留纹理信息,最大池化层计算过程如下:

$$M_c = \max \left[ \sum_{i=1}^H \sum_{j=1}^W F_{\text{input}}(i, j) \right], \quad (2)$$

式中: $M_c$ 表示第  $c$  通道最大池化层输出特征; $\max$  表示取最大值; $F_{\text{input}}(i, j)$  表示  $(i, j)$  位置上的输入特征。最大池化层输出特征经过一层卷积与非线性激活函数 ReLU 层处理后,由上采样层恢复图像大小,再由 Sigmoid 激活函数自适应获得权重:

$$F_{CA} = \gamma \left\{ \text{upsampling} \left\{ \text{Conv} \left\{ \sigma \left[ \text{Conv} \left( M_c \right) \right] \right\} \right\} \right\}, \quad (3)$$

式中: $F_{CA}$  表示通道注意力输出特征; $\gamma$  表示 Sigmoid 激活函数;upsampling 表示上采样操作;Conv 表示深度可分离卷积<sup>[28]</sup>操作; $\sigma$  表示 ReLU 激活函数。将通道注意力输出特征作为空间注意力输入特征,首先经过并行连接的平均池化层与最大池化层,获得两个  $H \times W \times 1$  大小特征图,再将特征图进行通道拼接,送入 Sigmoid 激活函数自适应分配权重,经过一层卷积操作后获得空间注意力输出特征:

$$F_{SA} = \text{Conv} \left\{ \gamma \left[ \text{MAX} \left( F_{CA} \right); \text{AVG} \left( F_{CA} \right) \right] \right\}, \quad (4)$$

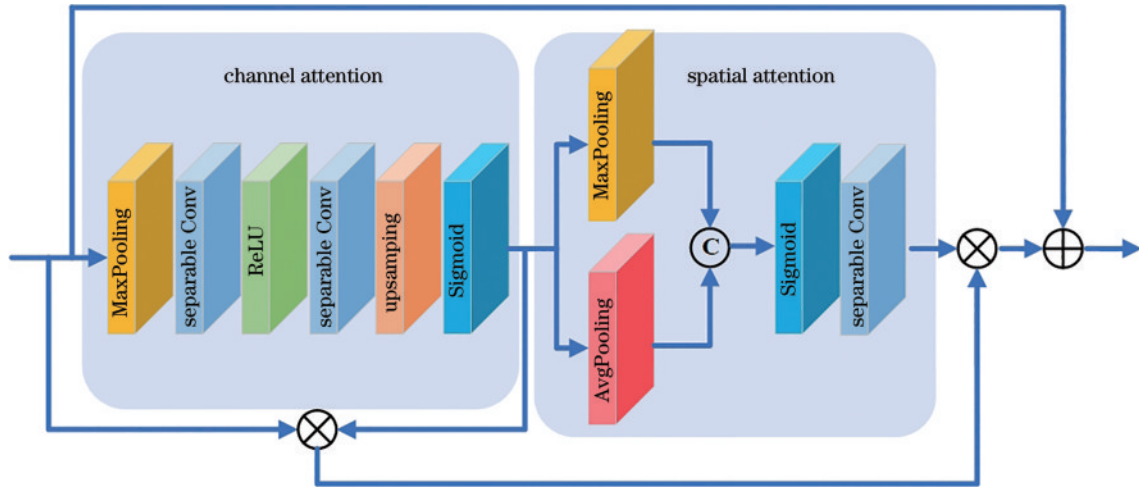


图 4 注意力模块结构

Fig. 4 Attention module structure

式中： $F_{SA}$  表示空间注意力输出特征；MAX 表示最大池化处理，AVG 表示平均池化处理。双注意力融合机制通过融合通道注意力与空间注意力输出以及注意力模块输入特征，获得注意力模块的最终输出特征：

$$F_{AM} = [(F_{CA} \otimes F_{input}) \otimes F_{SA}] \oplus F_{input}, \quad (5)$$

式中： $F_{AM}$  表示注意力模块输出特征。注意力模块输出特征图可视化效果如图 5 所示，通过图 5 能够发现，双注意力融合机制能够为有雾图像中物体边缘、纹理细节赋予较大权重，使模型更加关注图像高频区域，更好保留图像细节信息。

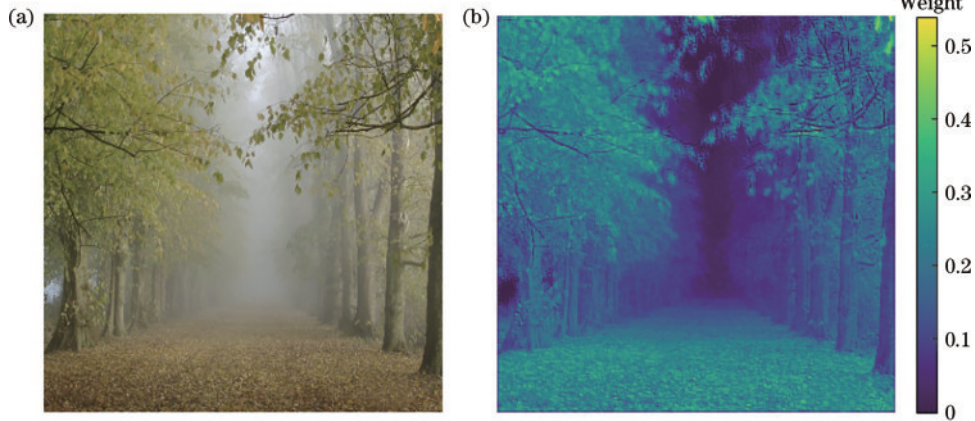


图 5 注意力模块特征图可视化。(a) 有雾图像；(b) 注意力模块输出特征图

Fig. 5 Attention module feature map visualization. (a) Hazy image; (b) attention module output feature map

为减少模型计算量，在注意力模块中采用  $3 \times 3$  大小的深度可分离卷积代替常规卷积，深度可分离卷积将卷积操作分解为逐通道卷积和逐点卷积，通过逐通道卷积对输入层每个通道独立卷积，再由逐点卷积在深度方向上进行加权组合，生成新的特征图。深度可分离卷积相对于常规卷积，能够减少模型参数量，在注意力模块中采用深度可分离卷积与常规卷积的参数数量对比如表 1 所示。通过对比可知，相较于常规卷积，深度可分离卷积参数量减少约 57%，同时降低了模型计算量，使模型更加轻量化，加快了网络模型的去雾处理速度。

## 2.2 选择性特征融合模块

为降低模型在编码阶段造成的图像信息损失，设计了选择性特征融合模块(SFFB)，通过选择性特征融合模块对压缩前原尺寸图像特征进行选择与聚合操

表 1 不同类型卷积参数量对比

Table 1 Comparison of parameter quantities for different types of convolutions

Convolution type	Parameter quantity
Convolution	113792
Depthwise separable convolution	48402

作，并将输出特征与模型上采样层输出融合，减少处理后图像纹理、细节的丢失，选择性特征融合模块结构如图 6 所示。

选择性特征融合模块采用双分支结构，分别通过  $3 \times 3$  和  $5 \times 5$  大小卷积核来获取不同尺度特征，然后将输出特征分别送入 ReLU 层处理，再通过最大池化层对特征进行压缩，去除冗余特征的同时更好地保留图

像边缘及纹理信息。然后分别对不同尺度特征图进行通道降维处理,将获得的紧凑特征送入 Softmax 激活函数中自适应获得权重,由上采样层恢复原图像大小,再将处理后的不同尺度特征进行聚合:

$$F_{3 \times 3} = \text{upsampling}$$

$$\left\{ \delta \left\{ \text{Conv}_{3 \times 3} \left\{ \text{MAX} \left\{ \sigma \left[ \text{Conv}_{3 \times 3} (F_{\text{input}}) \right] \right\} \right\} \right\} \right\}, \quad (6)$$

$$F_{5 \times 5} = \text{upsampling}$$

$$\left\{ \delta \left\{ \text{Conv}_{5 \times 5} \left\{ \text{MAX} \left\{ \sigma \left[ \text{Conv}_{5 \times 5} (F_{\text{input}}) \right] \right\} \right\} \right\} \right\}, \quad (7)$$

$$F_{\text{MIX}} = \delta \left[ \text{Conv}_{3 \times 3} (F_{3 \times 3} \oplus F_{5 \times 5}) \right], \quad (8)$$

式中: $F_{\text{MIX}}$ 表示多尺度特征融合后输出特征。对聚合后特征图进行通道升维,将升维后特征图送入

Softmax 激活函数再次获得权重后,分别与各分支输入特征进行元素相乘,最后将输出的多尺度特征元素相加,输出融合后特征图:

$$F_{\text{SFFB}} = \left\{ F_{\text{MIX}} \otimes \left[ \text{Conv}_{3 \times 3} (F_{\text{input}}) \right] \right\} \oplus \left\{ F_{\text{MIX}} \otimes \left[ \text{Conv}_{5 \times 5} (F_{\text{input}}) \right] \right\}, \quad (9)$$

式中: $F_{\text{SFFB}}$ 表示选择性特征融合模块输出特征。选择性特征融合模块通过输入的不同尺度特征信息来指导 Softmax 函数自适应产生不同权重,各分支所获权重在特征融合层中产生不同大小的有效感受野,因此选择性特征融合模块设置的动态选择机制相对于线性聚合,能更灵活地调整感受野大小,减少冗余信息,更有效地收集多尺度空间结构信息。

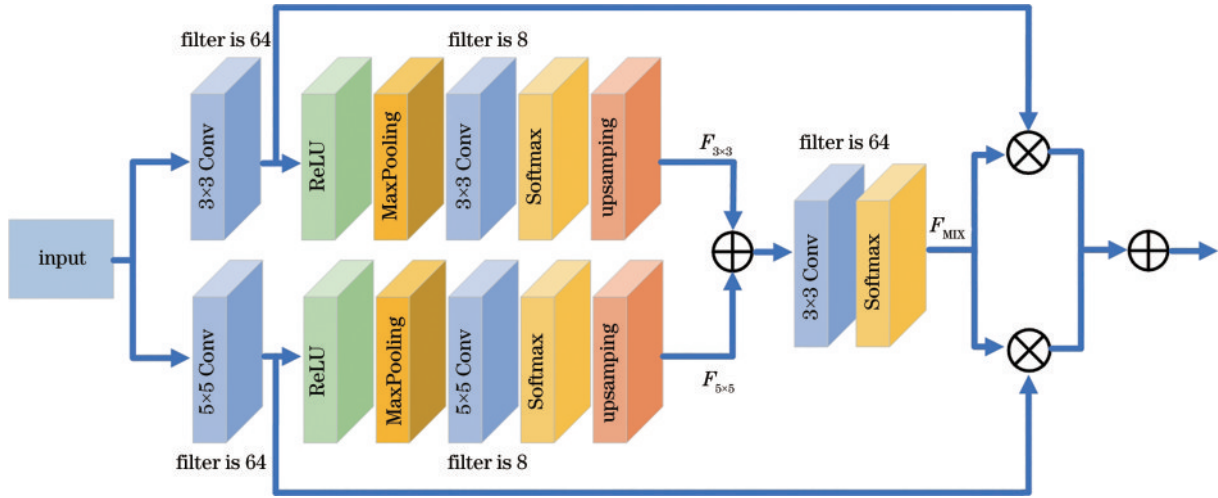


图 6 选择性特征融合模块结构

Fig. 6 Selective feature fusion block structure

### 2.3 损失函数

所提网络训练时采用 L2 损失, L2 损失即均方差 (MSE) 损失, 用于计算清晰无雾图像与经网络模型处理后的去雾图像之间的差值平方和, L2 损失常用于图像重构损失, 相对于 L1 损失来说 L2 损失对于异常值更加敏感, 收敛速度更快, 能够求出更接近、更稳定的解。L2 损失公式定义如下:

$$E_{\text{MSE}} = \frac{1}{n} \sum_{i=1}^n [f(x) - y_i]^2, \quad (10)$$

式中: $n$ 为训练集的样本总数; $f(x)$ 为清晰无雾图像; $y_i$ 为处理后去雾图像。L2 损失函数通过监测  $f(x)$  和  $y_i$  的差值平方和来优化模型。

为使模型能够准确达到最优解,将初始学习率设置为 0.0001,通过监测验证损失的变化情况来动态调整学习率,动态学习率能够有效防止网络在最优值附近震荡,使网络向正确方向加速收敛。

## 3 实验结果与分析

为了验证网络模型性能,选用合成雾天图像数据集和真实雾天图像进行实验,同时与经典的去雾算法

和性能较好的深度学习去雾算法进行比较,通过主观视觉评价与客观量化指标两方面对所提网络模型去雾效果进行分析。

### 3.1 数据集与实验设置

本实验采用 Python 语言编译,基于 TensorFlow 框架搭建网络模型,实验环境采用 Intel(R) Core (TM) i5-7300HQ CPU, GPU 采用 NVIDIA Tesla P100,显存容量为 16 GB,模型采用 Adam 优化器进行优化,初始学习率为 0.0001,通过监测验证损失动态调整学习率,批处理图像数量为 16,损失函数采用 L2 损失函数。

训练模型使用公开数据集 RESIDE 数据集<sup>[29]</sup>, RESIDE 数据集由 NYU V2 深度数据集<sup>[30]</sup>和米德尔伯里立体数据库中的清晰图像生成,其中,室内训练集 (ITS)由 1399 张清晰图像生成,依据大气散射模型调整大气光值范围为  $[0.7, 1.0]$ ,大气散射系数范围为  $[0.6, 1.8]$ ,每张清晰图像对应合成 10 张有雾图像,室内训练集共包含 13990 张合成有雾图像。RESIDE 室外训练集 (OTS)由 8970 张清晰图像生成,调整大气光值范围为  $[0.8, 1.0]$ ,散射系数范围为  $[0.04, 0.2]$ ,每

张清晰图像对应合成 35 张有雾图像,室外训练集共包含 313950 张合成有雾图像。模型性能测试采用 RESIDE 数据集中合成雾天数据集 SOTS、真实雾天数据集 HSTS 以及随机选取的 5 张真实场景雾天图像。

### 3.2 主观视觉评价

将所提算法分别在合成雾天图像以及真实雾天图像中进行测试,将测试结果与其他算法进行对比,对比算法包括:DCP 算法<sup>[11]</sup>、MSCNN 算法<sup>[14]</sup>、AOD-NET 算法<sup>[15]</sup>、AEDGAN 算法<sup>[16]</sup>、URNET 算法<sup>[17]</sup>、HR-Dehazer 算法<sup>[20]</sup>、PCFAN 算法<sup>[21]</sup>、MADN 算法<sup>[23]</sup>。

图 7 为不同算法在 SOTS 数据集上的去雾效果。通过观察发现:DCP 算法处理后图像色彩对比度普遍较高,容易造成色彩失真,如图 7(b)中天空部分。

MSCNN 算法对室内有雾图像去雾不彻底,处理后图像局部边缘易出现伪影,如图 7(c)中自行车轮毂位置。AOD-NET 算法在阴影部分处理效果偏暗,在局部细节上可见性较低,如图 7(d)中椅子部分。AEDGAN 算法与 PCFAN 算法不能较为均匀去除雾气影响,对局部雾气去除不彻底。URNET 算法处理后图像整体色温偏高,且对室内图像去雾效果一般,如图 7(f)中第 4 张图像。HR-Dehazer 算法处理后图像存在伪影且细节模糊不清,如图 7(g)中第 2 张图像中人群部分。MADN 算法处理后图像存在伪影,如图 7(i)中第 5 张图中地板部分。通过观察对比可以发现,所提算法在室内合成有雾图像和室外合成有雾图像上均取得较好的去雾效果,细节保留度较高且色彩对比度适中,比较符合人类视觉感知。

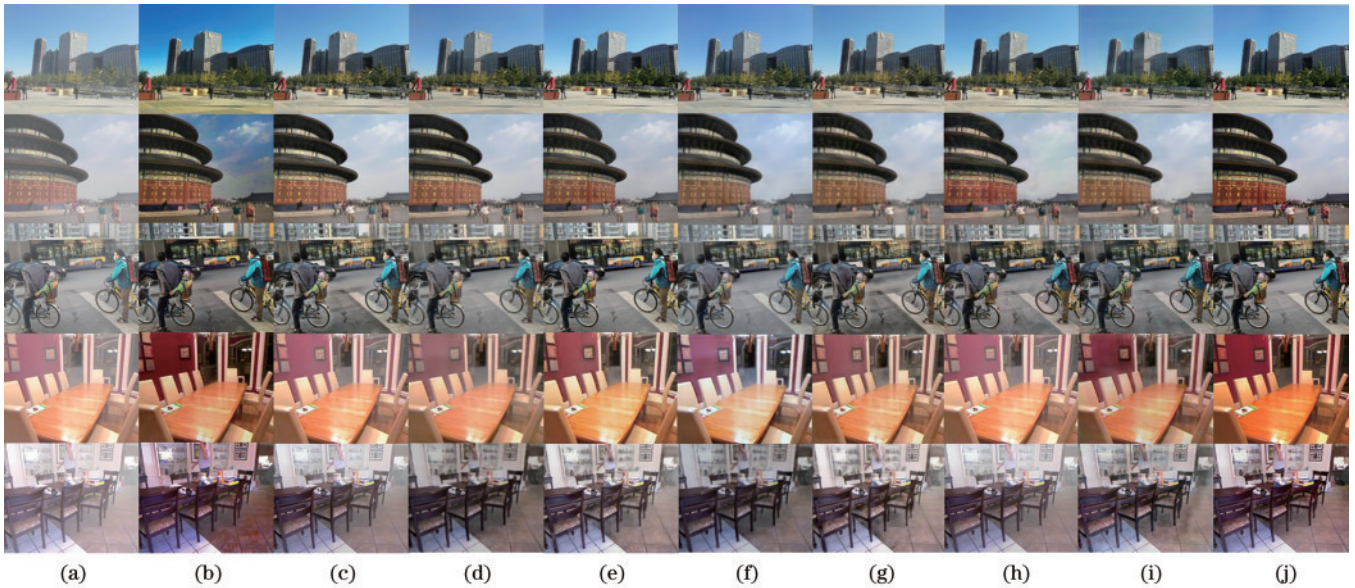


图 7 不同算法对合成有雾图像实验结果对比图。(a)有雾图像;(b)DCP;(c)MSCNN;(d)AOD-NET;(e)AEDGAN;(f)URNET;(g)HR-Dehazer;(h)PCFAN;(i)MADN;(j)所提算法

Fig. 7 Comparison of experimental results of synthetic hazy images by different algorithms. (a) Hazy images; (b) DCP; (c) MSCNN; (d) AOD-NET; (e) AEDGAN; (f) URNET; (g) HR-Dehazer; (h) PCFAN; (i) MADN; (j) proposed algorithm

为验证所提算法在真实雾天图像上的处理效果,通过不同算法对真实场景雾天图像进行处理,并对比去雾效果。图 8 展示了不同算法在真实场景雾天图像中去雾效果,可以看出:DCP 算法在图像天空区域误差较大,色彩失真较严重,如图 8(b)中天空部分。MSCNN 算法去雾不彻底,如图 8(c)中第 2、4、5 张图。AOD-NET 算法处理后图像偏暗,如图 8(d)中第 4、5 张图中山体区域,细节处理较差。AEDGAN 算法处理后图片对比度较高,存在颜色失真现象,如图 8(e)中第 2、4、5 张图。URNET 算法处理后图像色调偏冷,且对远处雾气影响处理效果不明显。HR-Dehazer 算法处理后图像存在伪影,且边缘细节比较模糊,如图 8(g)中第 3、4 张图。PCFAN 算法去雾不均匀,如图 8(h)中第 3、4 张图,局部仍残留雾气影响。MADN 算法对浓雾

区域去雾效果不明显,如图 8(i)中第 1、4 张图中深处浓雾部分,且存在一定程度颜色失真,如图 8(i)中湖面部分。通过对比能够发现,所提算法在真实雾天图像中具有一定去雾效果,处理后图像色彩合理,符合人类视觉感知,且不存在图像伪影,对图像细节具有较好的保留。

### 3.3 客观量化指标

为验证所提算法去雾效果,以峰值信噪比 (PSNR)、结构相似性 (SSIM)、信息熵 (Entropy)、雾密度估计 (FADE)<sup>[31]</sup>、图像视觉信息保真程度评价 (VIF)<sup>[32]</sup>、平均梯度 (AG),以及单幅图像处理时间为量化指标对不同算法去雾效果进行客观评估。

合成有雾图像去雾效果评估在 SOTS 数据集中进行,在相同实验设置中分别对 500 对室内图像和 500 对室外图像进行去雾效果评估,采用去雾处理后图像与

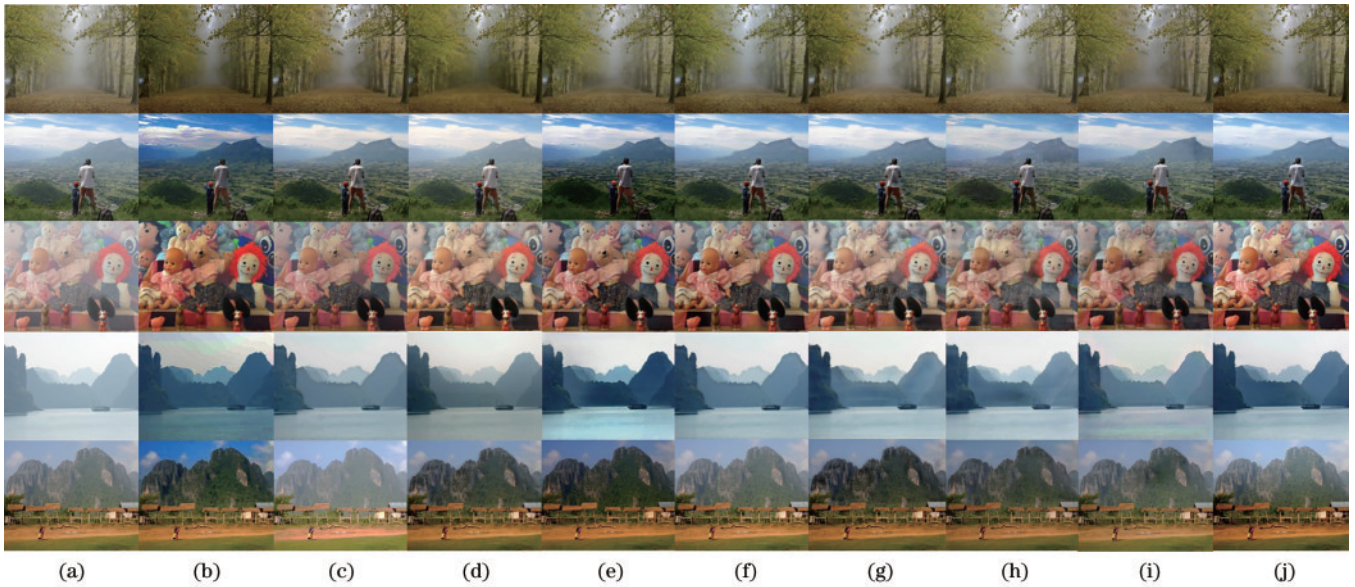


图 8 不同算法对真实有雾图像实验结果对比图。(a)有雾图像;(b)DCP;(c)MSCNN;(d)AOD-NET;(e)AEDGAN;(f)URNET;(g)HR-Dehazer;(h)PCFAN;(i)MADN;(j)所提算法

Fig. 8 Comparison of experimental results of different algorithms on real hazy images. (a) Hazy images; (b) DCP; (c) MSCNN; (d) AOD-NET; (e) AEDGAN; (f) URNET; (g) HR-Dehazer; (h) PCFAN; (i) MADN; (j) proposed algorithm

合成雾数据集中合成前清晰图像进行 PSNR、SSIM 计算。PSNR 是评价图像质量的一种客观标准,常通过均方差进行定义,PSNR 值越高,代表图像质量越好、失真越少、处理后图像与清晰图像之间误差越小。

SSIM 常用于衡量两幅图像相似程度,SSIM 值越接近于 1 代表两幅图结构越相似。不同算法在合成雾天数据集中,PSNR 与 SSIM 结果比较如图 9 和表 2 所示。

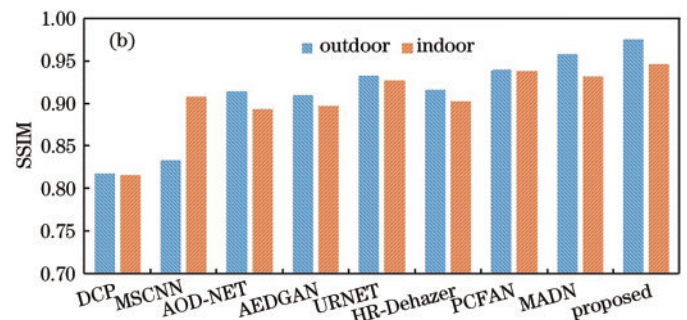
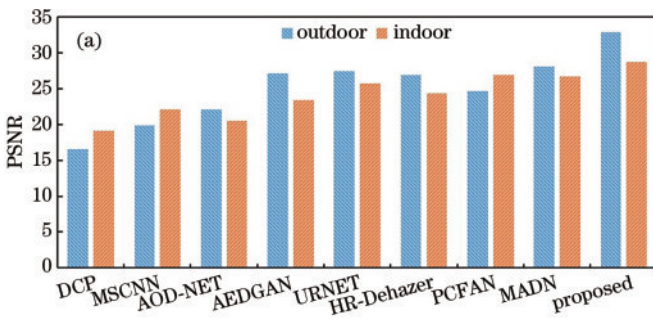


图 9 不同算法对合成雾天数据集 PSNR 及 SSIM 结果对比图。(a)峰值信噪比;(b)结构相似性

Fig. 9 Comparison chart of PSNR and SSIM results of different algorithms on SOTS dataset. (a) PSNR; (b) SSIM

从图 9 和表 2 可以看出,在相同实验环境下,所提算法相较于对比算法 PSNR、SSIM 结果在室外合成雾天图像数据集以及室内合成雾天图像数据集中位居前列,并在室外合成雾天图像数据集中,PSNR、SSIM 值相较第 2 名算法分别提升 4.7433 dB 与 0.0172。

密度,雾密度估计值越低表明图像中残留雾气影响越小;图像视觉信息保真程度评价常用于评估处理后图像失真程度,图像视觉信息保真程度评价值越大表明图像质量越高;平均梯度常用于评估图像的相对清晰程度,反映了图像对细节对比的表达能力,平均梯度值越大,表示图像包含越多的边缘信息。测试采用 HSTS 数据集,采用不同算法分别对真实雾天图像进行去雾处理,并对不同评估指标结果取平均值,不同算法在真实雾天图像 HSTS 数据集中测试结果如图 10 和表 3 所示。

由于难以获取真实雾天图像的清晰参考图像,因此,为评估算法在真实雾天图像中的去雾效果,采用以下 4 种不同的客观评价指标:信息熵、雾密度估计、图像视觉信息保真程度评价、平均梯度,对经不同算法处理后的真实有雾图像去雾效果进行评估。信息熵常用于评估去雾后图像清晰度与细节强度,信息熵值越高表明处理后图像能够保留越多细节特征;雾密度估计常用于评估图像中雾气浓度,预测整个图像的感知雾

从图 10 和表 3 可以看出,所提算法在信息熵、图像视觉信息保真程度以及平均梯度结果中均优于其他对比算法,在雾密度估计结果中位列第 3,因此综合多个



表 2 不同算法对合成雾天数据集峰值信噪比、结构相似性测试结果

Table 2 Test results of PSNR and SSIM of different algorithms on SOTS dataset

Algorithm	Outdoor		Indoor	
	PSNR	SSIM	PSNR	SSIM
DCP	16.6142	0.8176	19.1214	0.8159
MSCNN	19.8431	0.8327	22.0685	0.9078
AOD-NET	22.1389	0.9142	20.5477	0.8933
AEDGAN	27.1182	0.9095	23.3657	0.8968
URNET	27.4725	0.9326	25.7226	0.9269
HR-Dehazer	26.8715	0.9163	24.3136	0.9028
PCFAN	24.6931	0.9395	26.9713	0.9382
MADN	28.1353	0.9579	26.6972	0.9316
Proposed algorithm	<b>32.8786</b>	<b>0.9751</b>	<b>28.7852</b>	<b>0.9458</b>

表 3 不同算法对真实雾天数据集信息熵、雾密度估计、图像视觉信息保真程度、平均梯度测试结果

Table 3 Test results of Entropy, FADE, VIF, AG of different algorithms on HSTS dataset

Algorithm	Entropy	FADE	VIF	AG
Hazy image	5.8062	0.8116	0.6529	4.9983
DCP	6.7643	0.5726	0.7916	6.8955
MSCNN	6.3147	0.6891	0.816	7.0683
AOD-NET	6.6159	0.6377	0.8944	7.1924
AEDGAN	7.1625	<b>0.4914</b>	0.9125	7.5382
URNET	6.8693	0.5973	0.8362	7.6531
HR-Dehazer	7.1319	0.5834	0.8967	7.4938
PCFAN	7.0726	0.5046	0.9035	7.5976
MADN	6.8931	0.6193	0.9386	6.9043
Proposed algorithm	<b>7.2968</b>	0.5112	<b>0.9769</b>	<b>7.9306</b>

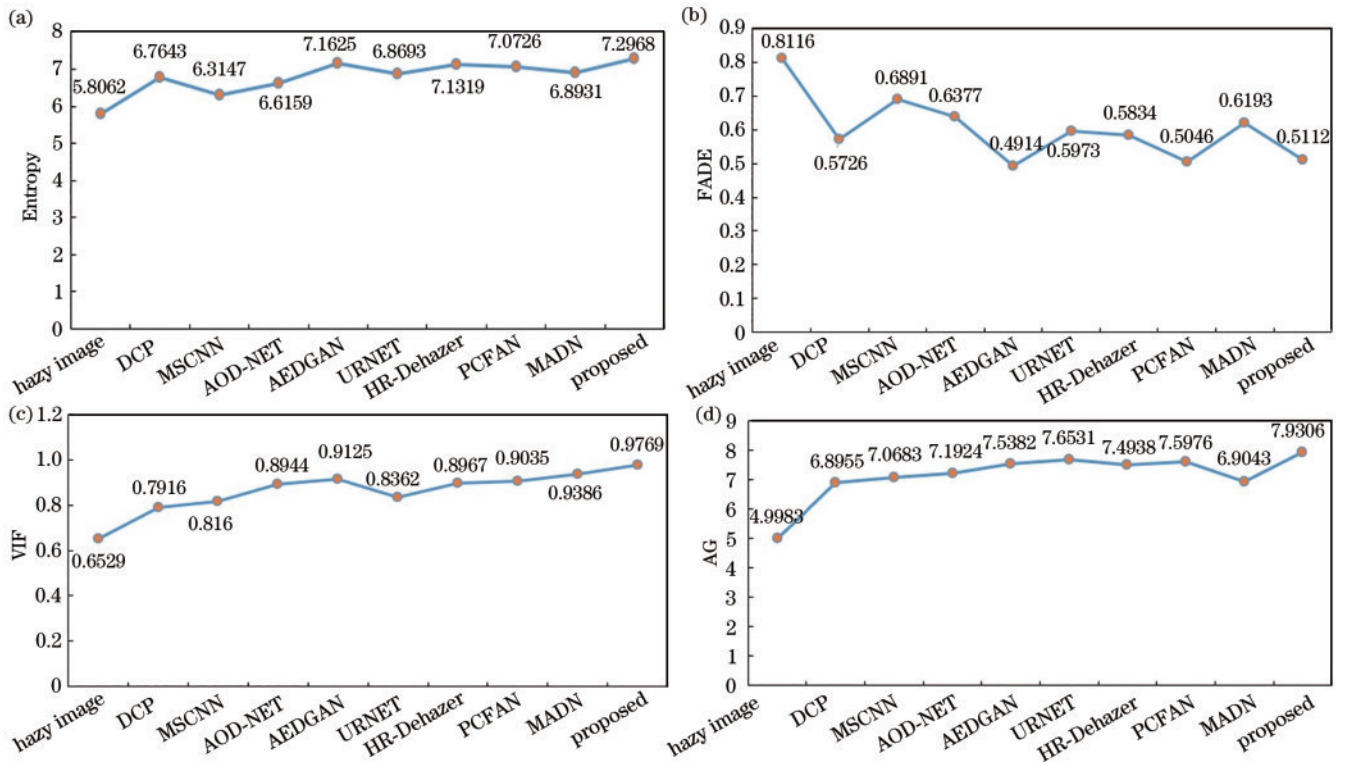


图 10 不同算法在真实雾天数据集中信息熵、雾密度估计、图像视觉信息保真程度、平均梯度结果对比图。(a)信息熵；(b)雾密度估计；(c)图像视觉信息保真程度；(d)平均梯度

Fig. 10 Comparison chart of Entropy, FADE, VIF, AG results of different algorithms on HSTS dataset. (a) Entropy; (b) FADE; (c) VIF; (d) AG

评价指标能够发现,所提算法对真实雾天图像具有一定去雾效果。相较于对比算法,处理后图像能够保留较多边缘细节,去雾图像质量较高。在实际应用中,为满足后续高级视觉处理任务的实时性需求,除了去雾效果外,去雾处理速度也是评价去雾算法优劣的另一重要指标。因此,通过不同算法分别对 SOTS 数据集中室外雾天图像进行去雾处理,记录处理图像的平均时间,单幅图像平均处理时间测试结果如表 4 所示。

相较于对比模型,所提模型参数量以及计算量较

少,模型较为轻量化。从表 4 可以看出,所提算法耗时最短,在对比实验中去雾处理速度较快、效率较高,能够更好应用于实际情况。

### 3.4 消融实验结果及分析

在并行多尺度注意力映射图像去雾算法中设计了并行多尺度注意力映射模块以及选择性特征融合模块,并在残差块中采用混合空洞卷积来提升模型去雾效果,为验证所提网络模型构造的合理性以及上述设计与改进的有效性,以单一残差块作为对比基线,在

表 4 不同算法图像处理时间结果对比

Table 4 Comparison of image processing time results of different algorithms

Algorithm	Processing time /s
DCP	1.0257
MSCNN	1.2331
AOD-NET	0.5543
AEDGAN	0.6378
URNET	0.4923
HR-Dehazer	0.6079
PCFAN	0.4063
MADN	0.6579
Proposed algorithm	<b>0.2016</b>

RESIDE 室外雾天图像数据集中对各个模块进行训练,采用 SOTS 数据集中室外雾天图像对消融实验进行测试,消融实验评价指标结果如表 5 所示。

从表 5 可以看出,单一注意力机制以及并行多尺度映射均能够在残差块基础上提升模型去雾效果,模型 PSNR 分别增长 1.2615 dB 和 2.0344 dB。但相较于单一的注意力残差块与单一的并行多尺度映射模块,将注意力残差块与并行多尺度映射模块并行连接的 PMAB 模块使 PSNR 相较基线模型增长了 4.5441 dB,去雾效果有更加明显的提升。同时,将网络模型上下采样层进行跳跃连接使模型 PSNR 增长 1.4339 dB,在网络模型中加入并行的选择性特征融合模块使模型 PSNR 增长 0.8801 dB。相较于常规卷积,混合空洞卷积的使用使模型 PSNR 增长 1.1054 dB、SSIM 增长 0.0042,提升了模型的去雾效果。不同模

表 5 消融实验结果对比

Table 5 Comparison of ablation experiment results

Experimental module	PSNR / dB	SSIM
Residual block	24.9151	0.9159
Attention residual block	26.1766	0.9375
Parallel multi scale mapping block	26.9495	0.9391
PMAB	29.4592	0.9573
PMAB+up-down sampling fusion	30.8931	0.9542
PMAB+up-down sampling fusion+SFFB	31.7732	0.9709
Hybrid dilated convolution	32.8786	0.9751

块去雾效果对比如图 11 所示。

从图 11 可以看出,单一残差块具备一定的去雾效果,但处理后图像存在严重图像伪影。单一注意力残差模块以及单一并行多尺度映射模块能够去除部分图像伪影,但处理后图像部分区域存在严重颜色失真现象,如图 11(c)、(d)中建筑主体。经并行多尺度注意力映射模块处理后,图像去雾效果明显提升,且图像伪影现象大幅度减少,但部分区域仍存在略微颜色失真,如图 11(e)中左侧建筑。在模型加入上下采样层后,图像伪影现象能够基本去除,但少部分区域仍残留雾气,如图 11(f)天空部分。加入选择性特征融合模块后图像中天空部分雾气影响减少,去雾更彻底,同时图像细节更清晰。混合空洞卷积的使用使雾气消除得更均匀,且色彩适中,不存在颜色失真现象。综上所述,所设计的并行多尺度注意力映射模块和选择性特征融合模块以及混合空洞卷积的使用,能够大幅度提升网络模型去雾能力,恢复出较高质量、满足人类视觉感知的清晰图像。

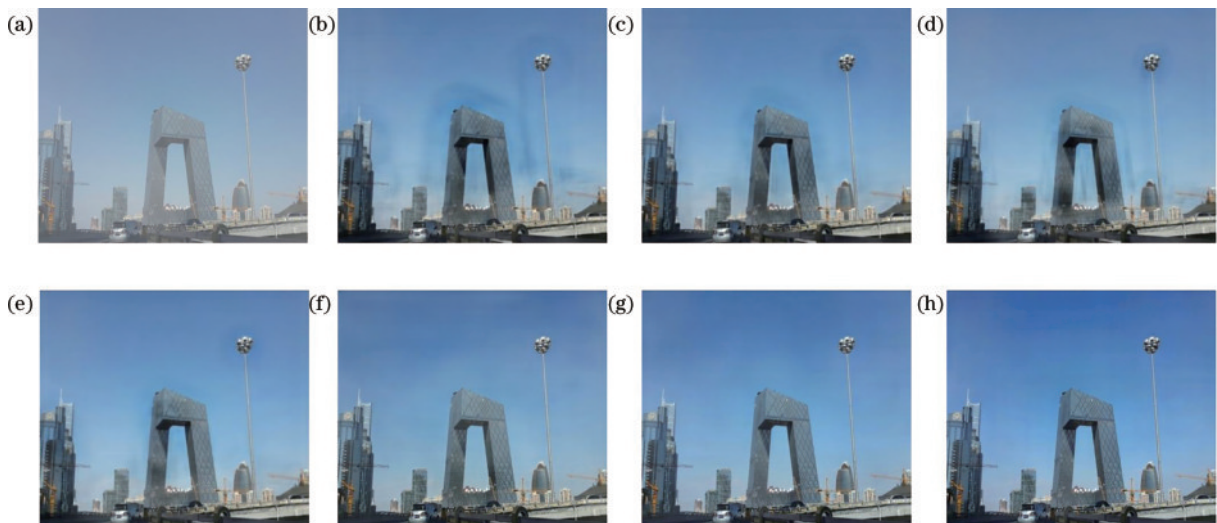


图 11 不同模块消融实验去雾效果对比图。(a)有雾图像;(b)残差块;(c)注意力残差块;(d)并行多尺度映射模块;(e)并行多尺度注意力映射模块;(f)并行多尺度注意力映射模块+上下采样层融合;(g)并行多尺度注意力映射模块+上下采样层融合+选择性特征融合模块;(h)混合空洞卷积

Fig. 11 Comparison diagrams of dehazing effects of different module ablation experiments. (a) Hazy image; (b) residual block; (c) attention residual block; (d) parallel multi scale mapping block; (e) PMAB; (f) PMAB+up-down sampling fusion; (g) PMAB+up-down sampling fusion+SFFB; (h) hybrid dilated convolution

## 4 结 论

针对去雾算法中容易产生的图像细节模糊不清、颜色失真、图像伪影等问题,提出一种并行多尺度注意力映射图像去雾算法,以端到端方式由有雾图像直接输出无雾图像。网络模型通过并行多尺度注意力映射模块与选择性特征融合模块,使模型在更加关注于图像特征信息的同时,扩大模型感受野,更好地获取全局特征,减少冗余特征信息,最后通过深浅层特征融合更好保留图像细节。所提算法在合成有雾图像数据集 RESIDE 以及真实有雾图像中均取得较好去雾效果,在室外有雾图像数据集中 PSNR 和 SSIM 分别达到 32.8786 dB 和 0.9751。同时,所提模型处理有雾图像速度较快,能够在高效率的同时保证图片恢复的高质量。由于所提算法针对浓雾区域去雾效果仍有进步空间,因此下一步工作重点在于继续优化网络模型,提升模型对浓雾图像去雾效果。

### 参 考 文 献

- [1] 张世辉, 路佳琪, 宋丹丹, 等. 基于多尺度特征结合细节恢复的单幅图像去雾方法[J]. 电子与信息学报, 2022, 44(11): 3967-3976.  
Zhang S H, Lu J Q, Song D D, et al. A single image defogging method based on multi-scale features combined with detail restoration[J]. Journal of Electronics & Information Technology, 2022, 44(11): 3967-3976.
- [2] Seow M J, Asari V K. Ratio rule and homomorphic filter for enhancement of digital colour image[J]. Neurocomputing, 2006, 69(7/8/9): 954-958.
- [3] Stark J A. Adaptive image contrast enhancement using generalizations of histogram equalization[J]. IEEE Transactions on Image Processing, 2000, 9(5): 889-896.
- [4] Land E H. The retinex theory of color vision[J]. Scientific American, 1977, 237(6): 108-128.
- [5] Land E H. The retinex[J]. American Scientist, 1964, 52(2): 247-264.
- [6] 韦春苗, 徐岩, 李媛. 基于小波变换的迭代融合去雾算法[J]. 激光与光电子学进展, 2021, 58(20): 2010018.  
Wei C M, Xu Y, Li Y. Iterative fusion defogging algorithm based on wavelet transform[J]. Laser & Optoelectronics Progress, 2021, 58(20): 2010018.
- [7] 方丹阳, 付青青, 吴爱平. 基于自适应动态范围 CLAHE 的雾天图像增强[J]. 激光与光电子学进展, 2023, 60(4): 0410008.  
Fang D Y, Fu Q Q, Wu A P. Foggy image enhancement based on adaptive dynamic range CLAHE [J]. Laser & Optoelectronics Progress, 2023, 60(4): 0410008.
- [8] Narasimhan S G, Nayar S K. Vision and the atmosphere [J]. International Journal of Computer Vision, 2002, 48(3): 233-254.
- [9] Narasimhan S G, Nayar S K. Contrast restoration of weather degraded images[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(6): 713-724.
- [10] Tang K T, Yang J C, Wang J. Investigating haze-relevant features in a learning framework for image dehazing[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 2995-3002.
- [11] 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.
- [12] 曾接贤, 余永龙. 双边滤波与暗通道结合的图像保边去雾算法[J]. 中国图象图形学报, 2017, 22(2): 147-153.  
Zeng J X, Yu Y L. Image defogging and edge preserving algorithm based on dark channel prior and bilateral filtering[J]. Journal of Image and Graphics, 2017, 22(2): 147-153.
- [13] 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.
- [14] Ren W Q, Liu S, Zhang H, et al. Single image dehazing via multi-scale convolutional neural networks[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9906: 154-169.
- [15] Li B Y, Peng X L, Wang Z Y, et al. AOD-net: all-in-one dehazing network[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 4780-4788.
- [16] Zhao L Q, Zhang Y P, Cui Y. An attention encoder-decoder network based on generative adversarial network for remote sensing image dehazing[J]. IEEE Sensors Journal, 2022, 22(11): 10890-10900.
- [17] Feng T, Wang C S, Chen X W, et al. URNet: a U-Net based residual network for image dehazing[J]. Applied Soft Computing, 2021, 102: 106884.
- [18] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [19] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [20] Bianco S, Celona L, Piccoli F, et al. High-resolution single image dehazing using encoder-decoder architecture [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 16-17, 2019, Long Beach, CA, USA. New York: IEEE Press, 2020: 1927-1935.
- [21] Zhang X Q, Wang T, Wang J X, et al. Pyramid channel-based feature attention network for image dehazing[J].

- Computer Vision and Image Understanding, 2020, 197/198: 103003.
- [22] 曹锐虎, 张鹏超, 王磊, 等. 基于注意力机制的单图像去雾算法[J]. 激光与光电子学进展, 2023, 60(2): 0210011.  
Cao R H, Zhang P C, Wang L, et al. Single image defogging algorithm based on attention mechanism[J]. Laser & Optoelectronics Progress, 2023, 60(2): 0210011.
- [23] Jia T Y, Li J F, Zhuo L, et al. Effective meta-attention dehazing networks for vision-based outdoor industrial systems[J]. IEEE Transactions on Industrial Informatics, 2022, 18(3): 1511-1520.
- [24] Wang P Q, Chen P F, Yuan Y, et al. Understanding convolution for semantic segmentation[C]//2018 IEEE Winter Conference on Applications of Computer Vision (WACV), March 12-15, 2018, Lake Tahoe, NV, USA. New York: IEEE Press, 2018: 1451-1460.
- [25] 王科平, 段雨朦, 杨艺, 等. 基于传递注意力机制的非均匀雾图去雾算法[J]. 模式识别与人工智能, 2022, 35(7): 575-588.  
Wang K P, Duan Y M, Yang Y, et al. Uneven hazy image dehazing based on transmitted attention mechanism [J]. Pattern Recognition and Artificial Intelligence, 2022, 35(7): 575-588.
- [26] Ding X H, Zhang X Y, Han J G, et al. Scaling up your kernels to  $31 \times 31$ : revisiting large kernel design in CNNs [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 18-24, 2022, New Orleans, LA, USA. New York: IEEE Press, 2022: 11953-11965.
- [27] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 3-19.
- [28] Chollet F. Xception: deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1800-1807.
- [29] 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.
- [30] Silberman N, Hoiem D, Kohli P, et al. Indoor segmentation and support inference from RGBD images [M]//Fitzgibbon A, Lazebnik S, Perona P, et al. Computer vision-ECCV 2012. Lecture notes in computer science. Heidelberg: Springer, 2012, 7576: 746-760.
- [31] Choi L K, You J, Bovik A C. Referenceless prediction of perceptual fog density and perceptual image defogging [J]. IEEE Transactions on Image Processing, 2015, 24(11): 3888-3901.
- [32] Sheikh H R, Bovik A C. Image information and visual quality[J]. IEEE Transactions on Image Processing, 2006, 15(2): 430-444.