



计算机科学

COMPUTER SCIENCE

结合注意力机制的多尺度特征融合图像去雾算法

范新南, 赵忠鑫, 严炜, 严锡君, 史朋飞

引用本文

范新南, 赵忠鑫, 严炜, 严锡君, 史朋飞. [结合注意力机制的多尺度特征融合图像去雾算法](#)[J]. 计算机科学, 2022, 49(5): 50-57.

FAN Xin-nan, ZHAO Zhong-xin, YAN Wei, YAN Xi-jun, SHI Peng-fei. [Multi-scale Feature Fusion Image Dehazing Algorithm Combined with Attention Mechanism](#)[J]. Computer Science, 2022, 49(5): 50-57.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于改进 CenterNet 的航拍绝缘子缺陷实时检测模型](#)

Real-time Detection Model of Insulator Defect Based on Improved CenterNet

计算机科学, 2022, 49(5): 84-91. <https://doi.org/10.11896/jsjcx.210400142>

[结合注意力机制与几何信息的特征融合框架](#)

Feature Fusion Framework Combining Attention Mechanism and Geometric Information

计算机科学, 2022, 49(5): 129-134. <https://doi.org/10.11896/jsjcx.210300180>

[面向对话的融入知识的实体关系抽取](#)

Dialogue-based Entity Relation Extraction with Knowledge

计算机科学, 2022, 49(5): 200-205. <https://doi.org/10.11896/jsjcx.210300198>

[基于用户关联的立场检测](#)

Stance Detection Based on User Connection

计算机科学, 2022, 49(5): 221-226. <https://doi.org/10.11896/jsjcx.210400135>

[基于时空自适应图卷积神经网络的脑电信号情绪识别](#)

EEG Emotion Recognition Based on Spatiotemporal Self-Adaptive Graph Convolutional Neural Network

计算机科学, 2022, 49(4): 30-36. <https://doi.org/10.11896/jsjcx.210900200>

结合注意力机制的多尺度特征融合图像去雾算法

范新南¹ 赵忠鑫² 严 炜¹ 严锡君² 史鹏飞¹

1 淮海大学物联网工程学院 江苏 常州 213022

2 淮海大学计算机与信息学院 南京 211100

(fanxn@hhu.edu.cn)

摘 要 针对传统图像去雾算法容易受到先验知识制约以及颜色失真等问题,提出了一种结合注意力机制的多尺度特征融合图像去雾算法。该算法首先通过下采样操作得到多个尺度的特征图,然后在不同尺度的特征图之间采用跳跃连接的方式将编码器部分的特征图与解码器部分的特征图连接起来以进行特征融合。同时,在网络中加入一个由通道注意力子模块和像素注意力子模块组成的特征注意力模块来控制不同通道与像素的重要性,这种特征注意力模块让网络更加关注细节信息和重要特征,因此能取得更好的去雾效果。为了验证所提算法的有效性,首先在 RESIDE 数据集上将所提算法与 5 种流行的去雾算法进行定性与定量对比实验。实验结果表明,所提算法能比较完全地除雾,而且对图像色彩的保持度较好;同时,在两个评价指标峰值信噪比(Peak Signal to Noise Ratio, PSNR)和结构相似性(Structure Similarity, SSIM)上的平均值分别为 28.83 dB 和 0.9575,相较于对比算法中性能位居第二的模型分别提高了 2.23 dB 和 0.0172。然后在 MSD 数据集以及真实图像上将所提算法与 5 种对比算法进行了定性对比实验。实验结果进一步证明了所提算法的去雾性能以及色彩保持度良好。

关键词: 注意力机制;多尺度特征融合;去雾算法;去雾性能;色彩保持度

中图法分类号 TP391

Multi-scale Feature Fusion Image Dehazing Algorithm Combined with Attention Mechanism

FAN Xin-nan¹, ZHAO Zhong-xin², YAN Wei¹, YAN Xi-jun² and SHI Peng-fei¹

1 College of Internet of Things Engineering, Hohai University, Changzhou, Jiangsu 213022, China

2 School of Computer and Information, Hohai University, Nanjing 211100, China

Abstract Aiming at the problems that traditional image dehazing algorithms are easily restricted by prior knowledge and color distortion, a multi-scale feature fusion image dehazing algorithm combined with attention mechanism is proposed. The algorithm first obtains feature maps of multiple scales through down-sampling operations, and then uses skip connections between feature maps of different scales to connect the feature maps of the encoder part and the decoder part for feature fusion. At the same time, a feature attention module composed of channel attention submodule and pixel attention submodule is added to the network to control the importance of different channels and pixels. This feature attention module allows the network to pay more attention to detailed information and important features, therefore, a better dehazing effect can be achieved. In order to verify the effectiveness of the proposed algorithm, qualitative and quantitative comparative experiments are carried out on the RESIDE dataset with five popular dehazing algorithms. The experimental results show that the proposed algorithm can more completely dehazing, and the preservation of image color is better. At the same time, the average values on the two evaluation indicators PSNR (Peak Signal to Noise Ratio) and SSIM (Structure Similarity) are 28.83 dB and 0.9575, respectively, which are 2.23 dB and 0.0172 higher than the second-performance model in the comparison algorithm. Then, qualitative comparison experiments are performed between the proposed algorithm and five comparison algorithms on the MSD data set and real images. The experimental results further prove that the proposed algorithm has good dehazing performance and color retention.

Keywords Attention mechanism, Multi-scale feature fusion, Dehazing algorithm, Dehazing performance, Color retention

到稿日期:2021-04-09 返修日期:2021-09-09

基金项目:常州市应用基础研究计划项目(CJ20200061)

This work was supported by the Applied Basic Research Programs of Changzhou (CJ20200061).

通信作者:史鹏飞(shipf@hhu.edu.cn)

1 引言

大气中存在雾、烟以及粉尘等颗粒,光经过这些颗粒时会发生折射和散射等现象^[1-3],这使得在大气中拍摄的图像经常出现色彩失真、对比度低和模糊等情况^[4-5]。这些低质量的图像将进一步影响其他高级视觉任务,如目标检测和分类^[6],而去雾任务就是将雾图像恢复成干净清晰的图像。近几十年来,去雾任务作为高级视觉任务的预处理步骤越来越受到研究人员的关注^[7-8]。

2 相关工作

传统去雾方法主要利用先验知识进行去雾。如 He 等^[9]提出暗通道先验方法(Dark Channel Prior, DCP),该方法的先验知识为雾图中总有一个灰度值很低的通道,He 等基于这个先验知识求解传输图,再利用大气散射模型进行图像去雾,该方法在当时取得了较好的去雾效果,但是在某些场景下会引起颜色失真。Meng 等^[10]提出一种有效的正则化方法去雾。该方法首先对传输函数的固有边界进行约束,再将该约束与基于 L1 范数的加权上下文正则化相结合。这种基于边界约束的去雾方法能解决去雾图像亮度偏低的问题。Zhu 等^[11]提出颜色衰减先验方法,该方法首先建立雾图深度的线性模型,并采用监督学习的方法学习线性模型中的参数,接着利用该模型估计传输图,并利用大气散射模型得到去雾图像。上述早期的图像去雾方法在某些特定场景下能够取得较好的效果,对去雾技术的发展做出了巨大贡献。但这些方法大多数依赖先验信息,并且采用人工方式提取雾图中的相关特征,导致传输图和大气光值的估计出现较大误差。因此,传统的去雾方法在很多情况下具有局限性,容易出现去雾不彻底和颜色失真等现象。

随着深度学习的迅速发展,研究人员提出了许多基于深度学习的去雾方法,该类方法大多利用卷积神经网络来构建可训练的去雾网络。如 Cai 等^[12]提出了一种 DehazeNet 去雾网络,该网络通过学习雾图中的特征来估计有雾图像与传输图之间的映射关系,然后根据输入的有雾图像的特征得出传输图,最后通过大气散射模型恢复出无雾图像。该算法的主要贡献在于首次提出了一个端到端的去雾网络。Ren 等^[13]提出了一种基于多尺度卷积神经网络的去雾方法(Multi-Scale Convolutional Neural Networks, MSCNN),该方法首先用大尺度网络估计整体传输图,然后用小尺度网络进行细化。通过该方法得到的传输图更加真实,一定程度上避免了细节信息丢失。Li 等^[14]通过重新推导大气散射模型,用参数 K 表示传输图和大气光值,然后设计 K 估计模块来估计参数 K ,最后通过 K 参数便可得到去雾图像。上述基于深度学习的方法首先采用卷积神经网络估计传输图、大气光值和其他中间变量,再利用大气散射模型去雾。这类方法虽然取得了不错的去雾效果,但实际上把去雾任务分成了两步,不是真正意义上的端到端去雾算法。同时,要从单幅雾图中估计传输图和大气光值非常困难。为解决该问题,利用深度卷积神经网络直接或迭代地估计出无雾图像的方法被提出。这些方法

主要采用通用的网络架构来直接估计传输图、大气光值和无雾图像,在保证鲁棒性的前提下提升了去雾性能。如 Ren 等^[15]提出了一种门控融合网络(Gated Fusion Network, GFN),该网络首先估计输入图像对应的权重图,然后以权重图为引导将输入图像进行加权融合,得到无雾图像。Chen 等^[16]提出利用一种门控上下文聚合网络来直接恢复最终的无雾图像。该网络采用最新的平滑膨胀技术来去除由膨胀卷积引起的网格伪影,并利用门控子网络融合不同层次的特征。上述算法虽然在一定程度上提升了图像的去雾性能,但是由于这些算法平等地对待有雾图像中的通道特征和像素特征,导致对图像中的浓雾像素区域和重要通道信息关注不足,最终影响了去雾性能。

本文针对上述方法存在的不足,提出了一种结合注意力机制的多尺度特征融合图像去雾算法。该算法利用类似于 U-Net 网络的编码器-解码器结构以实现一种直接学习输入自适应的去雾模型。具体来说,首先在编码器部分采用卷积神经网络得到不同尺度的特征图,然后在解码器部分依次从编码的特征图中恢复出图像细节。同时,为了充分利用输入信息和准确估计结构细节,采用跳跃连接的方式将编码器部分的特征图与解码器部分的特征图连接起来,以实现不同尺度特征图之间的融合。最后在编码器-解码器结构中加入由通道注意力子模块和像素注意力子模块组成的特征注意力模块。该模块能让网络更加关注重要的通道与像素,将更多的注意力集中在浓雾像素区域和重要通道信息上。特征注意力模块提高了网络处理不同类型信息的灵活性,使得本文算法能更加有效地处理雾浓度高且细节丰富的雾图。

3 本文算法

3.1 大气散射模型

根据雾图的成像原理,研究人员通常使用大气散射模型来模拟雾图的成像过程^[17],因此,该模型也是图像去雾的重要依据^[18]。大气散射模型的表达式如式(1)所示:

$$I(x) = J(x)t(x) + A(1-t(x)) \quad (1)$$

其中, $I(x)$ 代表成像设备获取的有雾图像; $J(x)$ 代表无雾的清晰图像; A 代表大气光值; $t(x)$ 代表传输图。为了更加清晰地表示去雾过程,可将式(1)进一步表示为:

$$J(x) = \frac{I(x) - A}{t(x)} + A \quad (2)$$

从式(2)中可以看出,在得到大气光值 A 和传输图 $t(x)$ 两个先验信息的基础上便可进行图像去雾。因此,去雾算法的核心任务就是估计得到大气光值和传输图。在一些基于深度学习的去雾算法中,首先用卷积神经网络估计传输图,再用传统的方法估计大气光值。这类算法的去雾性能较好,但不是真正意义上的端到端图像去雾方法。本文则是利用 U-Net 这一通用网络架构来直接估计传输图、大气光值和无雾图像,目的在于实现一种直接学习输入自适应的去雾模型。

3.2 网络结构

本文提出一种结合注意力机制的多尺度特征融合图像去雾算法。该算法的网络结构图如图 1 所示。

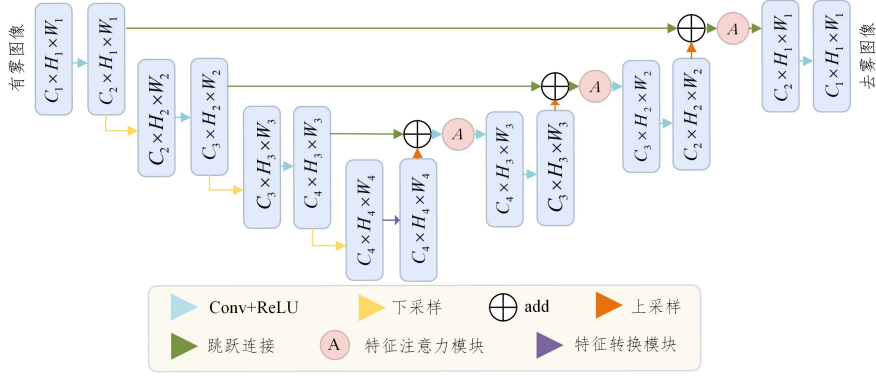


图1 本文方法的网络结构图

Fig. 1 Network structure diagram of the method in this paper

从图1可以看出,该算法的主体结构使用了类似于U-Net网络的编码器-解码器网络结构,网络主要由4个模块组成:编码器模块、特征转换模块、特征注意力模块和解码器模块。在本文算法中,首先使用卷积神经网络提取图像中多个尺度的特征图,然后采用跳跃连接的方式将不同尺度的特征图进行融合,使得最终得到的特征图中既包含深层特征,又包含浅层特征,从而提高了算法对特征的表达力。同时,在网络中加入由通道注意力子模块和像素注意力子模块组成的特征注意力模块,让网络更加关注浓雾像素区域和重要通道信息。下面将详细描述本文提出的结合注意力机制的多尺度特征融合图像去雾算法的各个模块。

3.2.1 编码器模块

编码器模块中包含3个卷积层和3个下采样层。编码器模块中的卷积操作如下式所示:

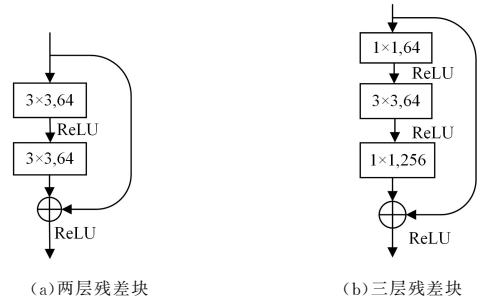
$$D_i = \text{Conv}_m(D_{i-1}), i \in \{1, 2, 3\} \quad (3)$$

其中, D_i 代表编码阶段的第 i 层特征图; D_0 为原有雾图像; Conv_m 代表编码阶段的卷积操作,在该卷积操作中,卷积核的大小为 3×3 ,原因在于,相比较大的卷积核, 3×3 的卷积核在有效提取特征的同时,还能减少网络的参数量以提高计算效率,卷积的步长为1,且通道数 C 将变为前一层特征图的2倍。同时,在每个卷积操作后都有一个非线性 ReLU 激活函数。在下采样操作中,采样的核大小为 2×2 ,即代表特征图经一次下采样后,特征图的宽 W 与高 H 均变为原来的 $1/2$,该下采样操作也可看作是一个核大小为 2×2 的平均池化操作。总的来说,编码器模块是以金字塔尺度的方式对图像信息进行编码,从而获得多个尺度的特征图。假设输入的有雾图像的尺寸为 $W \times H \times C$,而编码器模块中的3个卷积层使得特征图的通道 C 放大8倍,3个下采样层使得特征图的宽 W 与高 H 分别缩小8倍,那么经编码器模块之后产生的特征图的尺寸将变为 $1/8W \times 1/8H \times 8C$,即 $1/8W \times H \times C$ 。

3.2.2 特征转换模块

深度卷积神经网络的主要优势在于它能够利用复杂的运算来表示复杂的函数,学习到更深层且更为抽象的特征,从而提升网络的特征表达能力,并且在一般情况下,网络的性能将随着网络的加深而提升。因此,早期研究人员想通过不断加深网络的深度来提升网络的性能。但事实上,当网络达到一定深度之后,网络的性能不升反降,导致网络退化。后经研究

发现,该现象主要是因网络过深导致反向传播时梯度逐渐消失,无法调整前面网络层的权重所引起,而深度残差网络^[19]的出现解决了该问题。深度残差网络采用恒等映射的方式,在保证网络的性能不会随着网络的加深而退化的同时还能提取到更深层次的特征。深度残差网络主要由两种残差块组成,这两种残差块的结构示意图如图2所示。



(a) 两层残差块

(b) 三层残差块

图2 两种典型残差块结构

Fig. 2 Two typical residual block structures

由图2可以看出,两种残差块主要包含 1×1 和 3×3 的卷积核以及 ReLU 激活函数。同时,在残差块中,除了原始的卷积操作之外,还引入了跨层连接的短接结构,该结构能保证网络至少不会退化。本文利用残差块构成如图3中的特征转换模块,该模块可提取特征图中更为深层的特征。

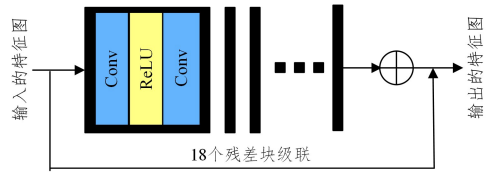


图3 特征转换模块

Fig. 3 Feature conversion module

由图3可以看出,为了平衡网络性能和计算效率,该特征转换模块由18个两层残差块组成,残差块中包含两个卷积运算和一个激活函数,卷积运算中的卷积核大小均为 3×3 ,卷积的步长始终保持为1,激活函数选用 ReLU 激活函数。同时,特征转换模块的输入为编码器模块所输出的特征图。

3.2.3 特征注意力模块

现有的大多数去雾网络都没有区别处理不同的通道和像素,但大量研究表明,不同的通道特征具有完全不同的加权

信息。此外,雾图中的雾分布也不均匀,薄雾区域的像素权重和厚雾区域的像素权重完全不同。因此,如果平等地对待雾图中的通道特征和像素特征将会花费大量资源来计算不必要的信息,同时会忽略重要的信息,导致网络缺乏覆盖所有通道和像素的能力,最终限制网络的表示。自从注意力机制在研究中取得良好效果之后^[20],许多研究都开始加入注意力机制。本文在编码器-解码器网络中加入由通道注意力子模块和像素注意力子模块组成的特征注意力模块。该模块让网络更加关注重要的通道与像素,将更多的注意力集中在浓雾像素和重要的通道信息上,从而更灵活地处理不同类型的信息,对一些雾浓度高且细节丰富的有雾图像的去雾性能提升更为显著。特征注意力模块的结构图如图4所示。

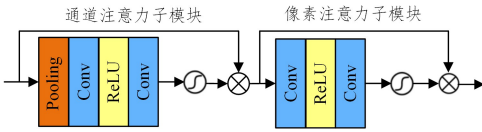


图4 特征注意力模块

Fig. 4 Feature attention module

从图4中可以更加清晰地看出,该特征注意力模块由通道注意力子模块与像素注意力子模块组成。通道注意力子模块包含一个池化层 Pooling、两个卷积层 Conv、一个非线性 ReLU 激活函数以及一个非线性 Sigmoid 激活函数。像素注意力子模块包含两个卷积层 Conv、一个非线性 ReLU 激活函数以及一个非线性 Sigmoid 激活函数。下面对通道注意力子模块和像素注意力子模块的具体实现细节进行详细的介绍。

(1) 通道注意力子模块

本文采用平均池化的方式来描述通道中的全局空间信息,平均池化操作如式(4)所示:

$$g_c = H_p(F_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X_c(i, j) \quad (4)$$

其中, F_c 代表输入的第 c 个通道的特征图, H_p 代表池化函数, g_c 代表输出的第 c 个通道的特征图, H 和 W 分别代表特征图的高和宽, $X_c(i, j)$ 代表第 c 个通道的特征图中像素点 (i, j) 处的值。假设输入的特征图的尺寸为 $C \times H \times W$, 经该平均池化操作后的特征图的尺寸将变为 $C \times 1 \times 1$, 这使得网络只需关注特征图的通道信息即可。为了进一步获得不同通道的权重信息, 特征图后续再经过两个卷积层、一个非线性 ReLU 激活函数以及一个非线性 Sigmoid 激活函数, 如式(5)所示:

$$CA_c = \sigma(\text{Conv}(\delta(\text{Conv}(g_c)))) \quad (5)$$

其中, g_c 代表经平均池化操作输出的第 c 个通道的特征图; Conv 代表卷积层, 卷积核的大小为 3×3 , 通道数不发生变化; δ 代表 ReLU 激活函数; σ 代表 Sigmoid 激活函数; CA_c 代表 c 通道的权重。经此操作后便可获得不同通道特征的不同加权信息。最后将通道权重 CA_c 与原输入特征图 F_c 进行逐像素相乘得到经通道注意力子模块的特征图 F_c^* , 如式(6)所示。该特征图 F_c^* 的尺寸恢复为输入时的尺寸 $C \times H \times W$, 同时该特征图实现了区别对待雾图中的通道特征, 而不再平等地处理雾图中的通道特征。

$$F_c^* = CA_c \otimes F_c \quad (6)$$

(2) 像素注意力子模块

为了进一步获得各个像素的权重信息, 把通道注意力子模块输出的特征图 F^* 经过两个卷积层、一个非线性 ReLU 激活函数以及一个非线性 Sigmoid 激活函数, 如式(7)所示:

$$PA_c = \sigma(\text{Conv}(\delta(\text{Conv}(F^*)))) \quad (7)$$

其中, F^* 代表通道注意力子模块输出的特征图; Conv 代表卷积层, 卷积核的大小为 3×3 ; δ 代表 ReLU 激活函数; σ 代表 Sigmoid 激活函数; PA 代表像素的权重。假设通道注意力子模块输出的特征图的尺寸为 $C \times H \times W$, 经像素注意力子模块后的特征图的尺寸变为 $1 \times H \times W$, 可以看出, 该模块中的卷积操作改变了特征图的通道数 C 。最后将像素权重 PA 与输入的特征图 F^* 逐像素相乘得到经像素注意力子模块的特征图 \bar{F} , 如式(8)所示。该特征图 \bar{F} 的尺寸恢复为输入时的尺寸 $C \times H \times W$, 同时该特征图实现了区别对待雾图中的像素特征, 而不再平等地处理雾图中的像素特征。

$$\bar{F} = F^* \otimes PA \quad (8)$$

3.2.4 解码器模块

解码器模块中包含 3 个卷积层和 3 个上采样层。解码器模块中的卷积操作如式(9)所示:

$$U_{i-1} = \text{Conv}_{de}(U_i), i \in \{3, 2, 1\} \quad (9)$$

其中, U_i 代表解码阶段第 i 层的特征图; U_0 为去雾图像; Conv_{de} 代表解码阶段的卷积操作, 在该卷积操作中, 卷积核的大小依然为 3×3 , 卷积的步长为 1, 且通道数 C 将变为前一层特征图的 $1/2$ 。同时, 在每个卷积操作后都有一个非线性 ReLU 激活函数。在上采样操作中, 特征图的宽 W 与高 H 均变为原来的 2 倍。总的来说, 解码器与编码器相反, 解码器的上采样层顺序恢复图像细节, 最终得到去雾图像。同时, 为了最大化多层次之间的信息流, 采用跳跃连接的方式将编码器和解码器之间的特征图连接起来, 并在跳跃连接的同时加入特征注意力模块, 使得网络更加关注图像中重要的通道信息和像素信息。假设通过特征转换模块后, 特征图的尺寸为 $1/8 W \times H \times C$, 而解码器模块中的 3 个卷积层使得特征图的通道 C 缩小为原来的 $1/8$, 3 个上采样层使得特征图的宽 W 与高 H 分别放大 8 倍, 那么经解码器模块之后所产生的特征图的尺寸将变为 $W \times H \times C$, 这与输入的有雾图像的尺寸一致。

3.2.5 损失函数

本文所提网络采用的损失函数为 smooth L_1 , 即用 smooth L_1 损失函数来定量描述有雾图像和真实图像之间的差异大小。该损失函数对异常值的敏感程度较低, 因而可以有效防止梯度爆炸的问题。smooth L_1 损失函数如式(10)所示:

$$L = \frac{1}{N} \sum_{p=1}^N \sum_{c=1}^3 (F(\bar{y}_c(p)) - y_c(p)) \quad (10)$$

其中, N 代表一幅图像中的总像素个数, p 代表第 p 个像素, c 代表第 c 个通道, 那么 $\bar{y}_c(p)$ 代表预测图像中第 c 个通道的第 p 个像素的像素值, $y_c(p)$ 代表真实图像中第 c 个通道的第 p 个像素的像素值, F 函数可表示为:

$$F(x) = \begin{cases} 0.5x^2, & \text{if } |x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases} \quad (11)$$

4 实验结果与分析

4.1 实验环境

本实验所用计算机处理器为 Intel Core i7-7700, 内存为 128 GB, 显存为 11 GB, GPU 为 NVIDIA TITAN V。基于 python3.6 编程实现, 深度学习框架为 pytorch1.0.0。

4.2 实验数据集

一般来说, 从真实世界中采集大量的无雾图像及相对应的有雾图像非常困难。因此, 通常用合成的有雾图像来评价算法的去雾性能。这些有雾图像是通过在真实无雾图像上加入适当的散射系数和大气光值合成得到的。本文首先使用评估去雾性能常用的 RESIDE 合成数据集来验证本文算法的性能。该数据集包含 NYU Depth V2 深度数据集^[21]和 Middlebury Stereo 立体数据集。RESIDE 室内训练集 ITS (Indoor Training Set) 由 1399 张清晰图像合成而来, 生成图像的大气光值范围为 [0.7, 1.0], 大气散射系数范围为 [0.6, 1.8], 每张清晰图像产生 10 张有雾图像。因此, 室内数据集总共包含 13990 张合成有雾图像, 其中 13000 张为训练集, 剩余 990 张为验证集。RESIDE 室外训练集 (Outdoor Training Set, OTS) 由 8477 张清晰图像合成而来, 生成图像的大气光值范围为 [0.8, 1.0], 散射系数范围为 [0.04, 0.2], 室外数据集总共包含 296695 张合成有雾图像。本文所用测试集为 RESIDE 数据集的综合目标测试集 (Synthetic Objective Testing Set, SOTS), 该测试集包含 500 对室内图像和 500 对室外图像。另外, 为了进一步证明本文方法的有效性以及鲁棒性, 本文还利用大气散射模型把 (Middlebury Stereo Datasets, MSD) 中两张色彩较为鲜艳的图片合成有雾图像作为测试集, 主要用于验证各算法在图像去雾之后是否会造成图像颜色失真。最后在真实雾图上进行本文算法与其他算法的对比实验, 以验证所提算法在真实场景中的去雾效果。

4.3 评价指标

由于主观判断不能完全说明本文方法的有效性, 为了与现有的去雾方法进行性能对比, 本文采用峰值信噪比 (PSNR) 和结构相似性 (SSIM) 来量化去雾图像的恢复质量。PSNR 是最大信号量与噪声强度的比值, PSNR 的值越大, 表明去雾图像越接近于无雾图像。SSIM 是衡量两幅图像相似度的指标, SSIM 的值越大, 表明去雾图像的失真程度越小。

4.4 参数设置

本文重要的实验参数如表 1 所列。

表 1 实验参数
Table 1 Experimental parameters

参数	值
Optimizer	Adam
Batch Size	1
Epoch	100
Learning Rate	0.0001

优化器 (Optimizer) 选用 Adam, 其中 β_1 和 β_2 分别采用默认值 0.9 和 0.999; 批处理大小 (Batch Size) 设为 1, 即每次只训练 1 张图片; 训练集的迭代轮次 (Epoch) 设为 100; 初始学习率 (Learning Rate) 设为 0.0001, 并采用余弦退火函数^[22]来

调整学习率。学习率的变化函数如式 (12) 所示:

$$\eta = \frac{1}{2} \left(1 + \cos \left(\frac{t\pi}{T} \right) \right) \eta_0 \quad (12)$$

其中, η 代表初始学习率, η_t 代表当前学习率, T 代表总批次, t 代表某一批次。

4.5 实验结果

为了验证本文算法的有效性, 本节将本文算法与目前流行的几种去雾算法进行定性与定量的比较, 对比算法包括 DCP^[9], DehazeNet^[12], MSCNN^[13], AOD-Net^[14], GFN^[15] 等。为了公平地进行比较, 所有方法均采用相同的环境进行训练, 具体的训练环境如 4.1 节所述。接着分别在 SOTS 数据集、MSD 数据集以及真实雾图上, 在相同的测试环境下测试各算法的去雾性能, 测试环境与训练环境一致。

4.5.1 在合成数据集上实验

首先将本文算法与其余对比算法在 SOTS 数据集上进行定量对比实验, 实验结果如表 2 所列。由表 2 可以看出, 本文方法得到的去雾图像的 PSNR 和 SSIM 值均大于其他对比算法, 证明了本文去雾算法的去雾性能优于对比算法。

表 2 各算法在评价指标上的对比结果

Table 2 Comparison results of each algorithm on evaluation indicators

方法	Indoor		Outdoor	
	PSNR/dB	SSIM	PSNR/dB	SSIM
DCP	16.62	0.8179	19.13	0.8148
DehazeNet	21.14	0.8472	22.46	0.8514
MSCNN	19.84	0.8327	22.06	0.9078
AOD-Net	19.06	0.8504	20.29	0.8765
GFN	24.91	0.9186	28.29	0.9621
本文方法	28.24	0.9440	29.42	0.9710

注: 加粗字体为每列最优值

然后, 将本文算法与其余对比算法在 SOTS 数据集上进行定性对比实验, 实验结果如图 5 所示。从图 5 可以看出, 由于 DCP 算法遵循暗通道先验原则, 因此该算法在处理雾浓度较高的有雾图像时, 无法准确估计雾的浓度, 导致去雾之后的图像通常比真实图像暗 (如图 5(b) 中第一张图像的桌面区域和墙壁区域), 并且会引起颜色失真 (如图 5(b) 中第三张图像的天空区域), 最终影响视觉效果。而由于 DehazeNet, MSCNN 和 AOD-Net 这 3 种基于深度学习的算法不受先验知识制约, 因此去雾效果比 DCP 算法更接近于实际情况。尽管如此, 经这 3 种算法处理后的图像仍然存在一些残留的雾气, 出现该现象的主要原因在于这 3 种算法在学习雾图特征的过程中忽略了有雾图像自身的结构信息, 估计得到的透射率误差较大, 最终导致图像去雾不彻底, 去雾之后的图像整体色调偏白, 如图 5(c)~图 5(e) 所示。GFN 算法解决了以上算法的问题, 因此其去雾效果较好, 雾图中的大多数雾都可去除。但是该算法去厚雾的能力有限, 因此雾浓度高的区域仍然存在未去除的雾气 (如图 5(f) 中第一张图像的椅子与椅子之间的区域)。相比之下, 由于本文算法加入了由通道注意力模块与像素注意力模块组成的特征注意力模块, 能区别处理不同的通道和像素信息, 对雾浓度高的区域赋予更多的关注, 因此该算法能够较好地处理雾浓度高的区域, 同时能清晰地保留图像中的细节, 减少颜色失真。

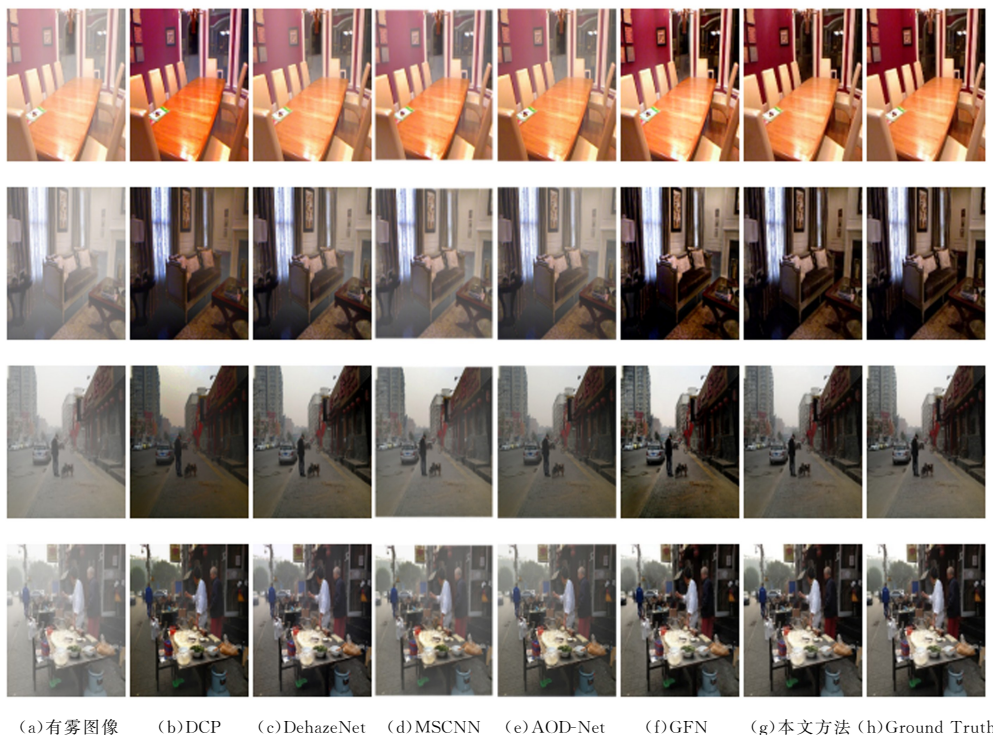


图 5 各算法在 SOTS 上的去雾效果

Fig. 5 Dehazing effect of each algorithm on SOTS

最后,将本文算法与对比算法在 MSD 数据集上进行定性对比实验,实验结果如图 6 所示。从图中可以看出,由于 DCP 算法受暗通道先验知识的制约,去雾之后的图像颜色失真较为严重,因此去雾后的图像整体色调偏暗,视觉效果很差。尽管 DehazeNet, MSCNN 和 AOD-Net 这 3 种基于深度学习的算法的去雾效果相比 DCP 算法较好,视觉效果更接近真实图像,但是由于这 3 种算法在去雾过程中忽略了图像自身

的结构信息,所以估计得到的透射率误差较大,最终导致图像去雾不彻底,去雾之后的图像整体色调偏白,图像颜色失真较为严重,整体色调严重偏离真实图像。GFN 算法去雾之后的图像色彩保持度较好,但由于原图像雾浓度过高,仍存在少量的雾未被去除。而本文算法由于加入了特征注意力模块,对浓雾区域给予了更多关注,因此去浓雾效果较 GFN 算法更好,同时本文算法的色彩保持度最好,去雾后的图像视觉效果最佳。

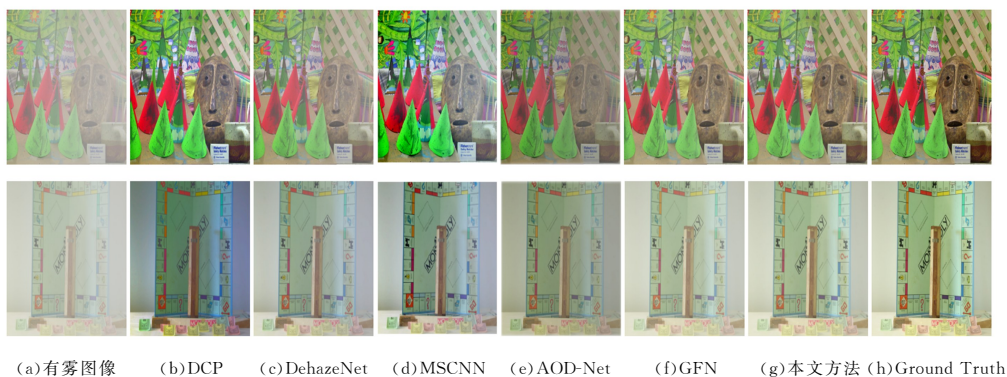


图 6 各算法在 MSD 上的去雾效果

Fig. 6 Dehazing effect of each algorithm on MSD

4.5.2 在真实雾图上实验

为了验证所提算法在真实场景中的去雾效果,本节将本文算法与对比算法在 3 张经典的真实雾图上进行定性对比实验,实验结果如图 7 所示。由图 7(b)可以看出,经 DCP 算法去雾之后的图像依然存在图像偏暗的问题(如图 7(b)的第一幅图中女孩的皮肤颜色偏暗,第三幅图中山峰区域的颜色偏暗),而且 DCP 算法会造成去雾图像颜色严重失真(如图 7(b)的第三幅图中天空颜色过蓝),这将影响视觉效果。其主要原

因是该算法十分依赖暗通道先验知识,无法准确地估计雾的浓度,导致透射率估计不准确。由图 7(c)一图 7(e)可以看出,对于 DehazeNet, MSCNN 和 AOD-Net 算法来说,它们的主要问题依然在于平等地对待雾图中的通道与像素信息,而忽略了图像自身结构信息,所以估计得到的透射率误差较大,从而导致图像去雾效果差,特别是图像中远景处的雾气有较多未被去除。由图 7(f)可以看出,GFN 算法在色彩保持和去雾效果上都较好,但依然无法去除浓雾区域

的雾气。由图 7(g)可以看出,本文算法由于加入了特征注意力模块,使得该算法能够有效去除各种浓度的雾气,

同时该算法能较好地保持图像的色彩不失真,视觉效果较好。



图 7 各算法在真实雾图上的去雾效果(电子版为彩图)

Fig. 7 Dehazing effect of each algorithm on real hazy images

4.6 运行时间对比

表 3 列出了各算法的平均去雾时间。所有算法使用的测试数据集均为 SOTS,即各算法均对 SOTS 数据集中的所有图像进行去雾操作,最后取所有图像去雾时间的平均值;而且所有算法的测试工作均是在同一台机器(CPU 为 Intel Core i7-7700,GPU 为 NVIDIA TITAN V, RAM 为 128GB)上进行;同时,所有算法均使用 python 编程语言实现。

表 3 各算法的平均去雾时间

Table 3 Average dehazing time of each algorithm

方法	时间/s
DCP	1.53
DehazeNet	0.30
MSCNN	0.26
AOD-Net	0.08
GFN	0.37
本文方法	0.20

从表 3 可以看出,传统的 DCP 去雾算法的平均去雾时间最长,原因主要在于该算法是根据大气散射模型来实现去雾的分阶段算法。因此该算法在去雾效率方面远低于后来提出的基于深度学习的端到端去雾算法。而与其余几种基于深度学习的去雾算法相比,AOD-Net 算法的去雾效率最优,主要原因在于 AOD-Net 算法具有轻质的网络结构,使得去雾效率大大提升,这一特性也使得 AOD-Net 算法经常与目标检测等高级视觉任务配合使用,从而提升雾图中的目标检测准确率。本文算法在去雾效率方面仅次于 AOD-Net 算法,平均去雾时间为 0.20s。本文算法的去雾效率之所以较高,原因主要在于该算法采用了编码器-解码器的简洁结构,而且所涉及的模块较少,是一种基于深度学习的端到端的去雾算法。

结束语 本文提出了一种有效的类似于 U-Net 网络的端到端去雾网络。该网络使用下采样层来提取图像中多个尺度的特征图,并且用跳跃连接的方式对不同尺度的特征图进

行融合。同时在跳跃连接的过程中加入由通道注意力子模块和像素注意力子模块组成的特征注意力模块来更加关注雾图中的重要通道与像素区域。文中采用公开 RESIDE 数据集对该模型进行训练,并且利用训练好的模型在测试数据集 SOTS,MSD 以及真实图像上进行了测试。实验结果表明,所提算法克服了传统单幅图像去雾算法容易受到先验知识制约以及颜色失真问题,可得到较为清晰的无雾图像,而且图像色彩失真小,去雾性能优于其他对比算法。

参考文献

- [1] GAO Z Z, WEI W B, PAN Z K, et al. Image dehazing combining dark channel prior and Hessian regular term [J]. Journal of Graphics, 2020, 41(1): 73-80.
- [2] CHEN Y, GUO H G, AI Y P. Single image dehazing method based on multi-scale convolution neural network [J]. Acta Optica Sinica, 2019, 39(10): 1-10.
- [3] LENG J W, LI P. Single image defogging algorithm based on high-order markov random fields [J]. Journal of Graphics, 2017, 38(5): 745-753.
- [4] LI B Y, REN W Q, FU D P, et al. Benchmarking single-image dehazing and beyond [J]. IEEE Transactions on Image Processing, 2018, 28(1): 492-505.
- [5] XIAO M X, LU C H, WEI H C, et al. Study on monitoring technology of straw burning based on UAV aerial image dehazing enhancement algorithm [J]. Journal of Electronic Measurement and Instrumentation, 2017, 31(4): 543-548.
- [6] ZHANG H, PATEL V M. Densely connected pyramid dehazing network [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE Press, 2018: 3194-3203.
- [7] CHU G S, SONG Y L, LI X C, et al. Near infrared dehazed imaging system based on SCMOS sensor [J]. Chinese Journal of

- Scientific Instrument, 2014, 35(6):138-141.
- [8] WU J W, YU Z C, LI Z Y, et al. A two-stage image dehazing network based on deep learning[J]. Computer Applications and Software, 2020, 37(4):197-202.
- [9] 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.
- [10] MENG G F, WANG Y, DUAN J Y, et al. Efficient image dehazing with boundary constraint and contextual regularization [C]//2013 IEEE International Conference on Computer Vision. Sydney: IEEE Press, 2013:617-624.
- [11] ZHU Q S, MAI J M, SHAO L. A fast single image haze removal algorithm using color attenuation prior[J]. IEEE Transactions on Image Processing, 2015, 24(11):3522-3533.
- [12] 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.
- [13] REN W Q, LIU S, ZHANG H, et al. Single image dehazing via multi-scale convolutional neural networks[C]//2016 European Conference on Computer Vision. Amsterdam: Springer Press, 2016:154-169.
- [14] 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. Venice: IEEE Press, 2017:4780-4788.
- [15] REN W Q, MA L, ZHANG J W, et al. Gated fusion network for single image dehazing[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE Press, 2018:3253-3261.
- [16] CHEN D D, HE M M, FAN Q N, et al. Gated context aggregation network for image dehazing and deraining[C]//2019 IEEE Winter Conference on Applications of Computer Vision. Waikoloa Village: IEEE Press, 2019:1375-1383.
- [17] CHOI H, LIM H, YU S, et al. Haze removal of multispectral remote sensing imagery using atmospheric scattering model-based haze thickness map[C]//2020 IEEE International Conference on Consumer Electronics-Asia. Seoul: IEEE Press, 2020:1-2.
- [18] 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. Columbus: IEEE Press, 2014:2995-3002.
- [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. Las Vegas: IEEE Press, 2016:770-778.
- [20] CHEN L, ZHANG H W, XIAO J, et al. Sca-cnn: Spatial and channel-wise attention in convolutional networks for image captioning[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE Press, 2017:6298-6306.
- [21] SILBERMAN N, HOIEM D, KOHLI P, et al. Indoor segmentation and support inference from rgbd images[C]//12th European Conference on Computer Vision. Florence: Springer Press, 2012:746-760.
- [22] HE T, ZHANG Z, ZHANG H, et al. Bag of tricks for image classification with convolutional neural networks [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE Press, 2019:558-567.



FAN Xin-nan, born in 1965, Ph.D, professor, Ph.D supervisor. His main research interests include information acquisition and processing, image processing and machine vision.



SHI Peng-fei, born in 1985, Ph.D, associate professor. His main research interests include machine vision, image processing and multi-source information fusion.

(责任编辑:柯颖)