

引用格式: 杜帆, 蒋品群, 宋树祥, 等. 基于注意力机制的单幅图像去雾算法[J]. 微电子学与计算机, 2021, 38(4): 52-56. [DUF, JIANG P Q, SONG S X, et al. Single image defogging algorithm based on attention mechanism[J]. Microelectronics & Computer, 2021, 38(4): 52-56.]

# 基于注意力机制的单幅图像去雾算法

杜 帆, 蒋品群, 宋树祥, 夏海英  
(广西师范大学 电子工程学院, 广西 桂林 541004)

**摘 要:** 针对现有去雾算法去雾效果不理想的问题, 提出了一种基于注意力机制的单幅图像去雾算法. 通过引入注意力机制构建通道注意和像素注意, 并将两者结合实现特征注意模块; 再通过多尺度卷积、局部残差学习和特征注意搭建基本模块; 最后结合全局残差学习实现端到端的去雾处理. 实验结果表明, 该算法在峰值信噪比 (PSNR)、结构相似度 (SSIM)、特征相似度 (FSIM) 三种图像评价指标上均高于对比算法, 取得了良好的去雾效果, 有效地解决了去雾效果不理想的问题.

**关键词:** 注意力; 图像去雾; 局部残差学习; 多尺度卷积

**中图分类号:** TP391.4

**文献标识码:** A

**文章编号:** 1000-7180(2021)04-0052-05

## Single image defogging algorithm based on attention mechanism

DU Fan, JIANG Pin-qun, SONG Shu-xiang, XIA Hai-ying

(College of Electronic Engineering, Guangxi Normal University, Guilin 541004, Guangxi China)

**Abstract:** Aiming at the problem that the existing defogging algorithm is not ideal for defogging, a single image defogging algorithm based on attention mechanism is proposed. Channel attention and pixel attention are constructed by introducing attention mechanism, and the feature attention module is realized by combining the two. Then the basic module is built through multi-scale convolution, local residual learning, and feature attention. At last, the end-to-end defogging is realized by global residual learning. The experimental results show that the algorithm is superior to the comparison algorithm in three image evaluation indexes of peak signal-to-noise ratio (PSNR), structural similarity (SSIM) and feature similarity (FSIM), and it achieves a good defogging effect and effectively solves the problems of undesirable defogging effect

**Key words:** attention; image defogging; local residual learning; multi-scale convolution

### 1 引言

雾的存在使得许多计算机视觉任务的有效性大打折扣, 现有的去雾算法主要有两类: 一类是基于物理模型的去雾算法, 一类是基于端到端的去雾算法. 基于物理模型的去雾算法分为基于手工先验的方法和基于学习先验的方法, 其区别在于基于手工

先验的方法需要手工制作图像先验, 如 He 等人<sup>[1]</sup>提出了基于暗通道先验的方法来估计透射率; D. Berman 等人<sup>[2]</sup>使用图像像素点先验来估计透射率, 然而此类方法会造成图像色彩保真度差, 导致不理想的去雾效果; 而基于学习先验的方法能够主动学习图像先验, 如 Cai 等人<sup>[3]</sup>提出了基于卷积神经网络的模型 DehazeNet 用于估计透射率; Li 等人<sup>[4]</sup>通过将介质透射率和大气光值合并为一个未知数并提

出了卷积神经网络模型 AOD-Net 来学习这个未知数. 此类方法对透射率的估计偏差会导致不理想的去雾效果.

基于端到端的去雾算法通过卷积神经网络直接恢复出无雾图像,例如 Chen 等人<sup>[5]</sup>提出了一种基于端到端的去雾算法,通过门控上下文聚合网络 GCANet 进行无雾图像的恢复;Liu 等人<sup>[6]</sup>在图像分割网络 GridNet 的基础上提出了一个端到端的去雾网络 GridDehazeNet;刘广洲等人<sup>[7]</sup>提出了一种端到端的密集连接扩张卷积神经网络 DCDN. 尽管现有的端到端的去雾算法比基于物理模型的去雾算法具有更好的去雾效果,但是由于容易忽略对于多尺度空间信息的聚合,导致图像细节信息的丢失,因此仍会存在去雾效果不理想的问题.

基于对上述问题的改进,本文在端到端去雾的基础上提出了基于注意力机制的单幅图像去雾算法.

## 2 本文方法

### 2.1 大气散射模型

大气散射模型通常用来模拟有雾图像的退化过程,其被广泛应用于图像去雾技术. 大气散射模型可表示为:

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

式中,  $I(x)$  为有雾图像;  $t(x)$  为介质透射率;  $J(x)$  为无雾图像;  $A$  为大气光值,表示光的强度. 从公式(1)中可以知道,要想恢复一个清晰的无雾图像,需要正确地估计出介质透射率和大气光值,然而,估计得到的透射率和大气光值的精度具有不确定性,这将对去雾效果造成很大的影响. 本文提出的算法采用端到端的网络结构摆脱了大气散射模型的束缚,从而能够达到更好的去雾效果.

### 2.2 特征注意 (Feature Attention, FA)

传统的卷积神经网络去雾模型会平等地处理通道和像素方向的特征,然而大部分有雾图像的雾度分布并不均匀,不同雾度的区域对应的权重不相同,此外不同通道特征也具有完全不同的加权信息. 如果网络平等地处理这些特征,将会消耗大量资源在不必要的信息上,降低去雾的有效性,因此本文在网络中引入注意力机制. 注意力机制类似于人的视觉注意力,它的目标是从众多信息中筛选出对当前任务最重要的信息. 随着注意力机制在神经网络中被越来越多地使用,它对于网络性能的提升起到了重要作用. 因此,本文基于注意力机制设计了一个特征

注意模块,其结构图如图 1 所示.

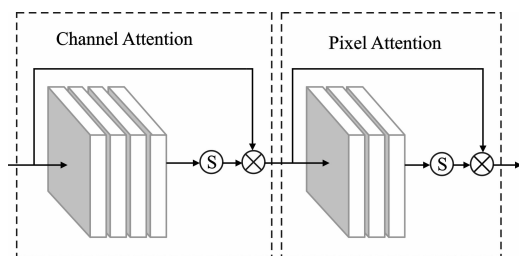


图 1 特征注意模块结构图

此模块由通道注意(Channel Attention, CA)和像素注意(Pixel Attention, PA)两部分组成. 特征注意模块不平等地对待不同的通道和像素,为每个通道和像素特征生成不同的权重,提高了网络的特征表达能力.

#### 2.2.1 通道注意

通道注意主要关注不同的通道特征具有完全不同的加权信息. 首先,通过全局平均池化将通道全局空间信息转化为通道描述符,即:

$$g_c = H_p(F_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X_c(i, j) \quad (2)$$

式中,  $X_c(i, j)$  表示第  $c$  个通道  $X_c$  在  $(i, j)$  位置的值;  $H_p$  为全局平均池化函数. 全局平均池化使得特征图的大小从  $C \times H \times W$  变为  $C \times 1 \times 1$ ,为了得到不同通道的不同权值,特征图随后通过两个卷积层以及 Sigmoid, ReLU 激活函数,即:

$$A_c = \sigma(\text{Conv}(\delta(\text{Conv}(g_c)))) \quad (3)$$

式中,  $\sigma$  为 Sigmoid 函数;  $\delta$  为 ReLU 函数;  $A_c$  为通道级别的权重. 最后,按元素顺序将输入  $F_c$  与权重  $A_c$  进行对应元素相乘得到通道注意的输出,即:

$$F_c^* = A_c \otimes F_c \quad (4)$$

通道注意模块使用的卷积均为  $1 \times 1$  卷积,这种逐点卷积计算得到的特征能够融合通道中的信息,但不会将跨空间的信息混合在一起,因此这种卷积有利于区分开通道特征学习和空间特征学习.

#### 2.2.2 像素注意

与通道注意类似,不同图像像素上的雾度分布不均匀,因此使用像素注意模块,使网络更加关注雾度较高的像素区域和高频区域的特征. 首先,通过通道数为 1 的卷积层使得特征图大小从  $C \times H \times W$  变为  $1 \times H \times W$ ,将通道注意的输出  $F_c^*$  直接输入到卷积层中,即:

$$A_p = \sigma(\text{Conv}(\delta(\text{Conv}(F_c^*)))) \quad (5)$$

式中,  $\sigma$  为 Sigmoid 函数;  $\delta$  为 ReLU 函数,得到的  $A_p$  为像素级别的权重. 然后将  $F_c^*$  与  $A_p$  进行对应元素

相乘,即:

$$F_p^* = A_p F_c^* \tag{6}$$

式中, $F_p^*$  为整个特征注意模块的输出.

2.3 网络结构

2.3.1 基本模块

不同尺度的卷积核可以提高多尺度下特征的鲁棒性,本文在基本模块中引入了并行多尺度卷积层,采用  $1\times 1,3\times 3,5\times 5,7\times 7$  这 4 组不同尺度的卷积核,以并行的方式进行特征提取,使得输入图像的更多细节得以保留.通过 concat 操作连接 4 组不同尺度的卷积层,有效保证了卷积输出特征图的数量,并且使得网络训练速度加快,防止网络出现局部最优的问题,同时保证不同尺度之间细节特征的有效传递,提高恢复图像的质量.多尺度卷积层参数如表 1 所示.

表 1 多尺度卷积层参数

| 卷积核尺寸  | 步长 | 填充 |
|--------|----|----|
| 16,1,1 | 1  | 0  |
| 16,3,3 | 1  | 1  |
| 16,5,5 | 1  | 2  |
| 16,7,7 | 1  | 3  |

基本模块结构图如图 2 所示,由多尺度卷积层、局部残差学习和特征注意组成.先通过并行的多尺度卷积层提取不同尺度信息,将这些特征融合后送入一个  $3\times 3$  卷积层,再通过特征注意模块实现对不同通道和像素的处理.局部残差学习通过残差连接使得模块在提取特征时绕过了薄雾或低频区域,使主网络学习更多有用的信息.

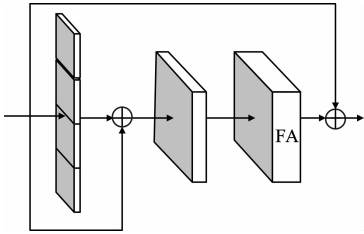


图 2 基本模块结构图

2.3.2 整体结构

为了增加网络的深度和表现力,本文设计了群组模块,其结构图如图 3 所示.将连续 5 个基本模块和一个用于特征提取的卷积核进行叠加,并在此基础上加入跃层连接以避免网络的训练困难.对于跃层连接本文使用 eltwise 融合操作,这样可以直接关联上下文的语义信息,尽管因此要花费更多的时间,但是对于精度的提升明显,效果更好.

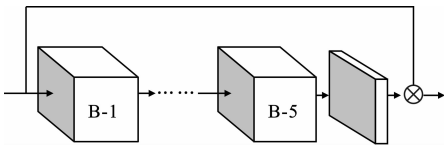


图 3 群组模块结构图

本文网络整体结构如图 4 所示,网络卷积层除了通道注意模块和多尺度卷积层以外均采用  $3\times 3$  卷积核.当有雾图像输入到网络中,它首先被传递到一个浅层特征提取部分,然后被输入到连续三个群组模块 G-1、G-2、G-3 中,并将输出的所有特征在通道方向上连接起来,随后通过乘以特征注意模块获得的自适应学习权重,将群组模块得到的输出特征融合到一起,最后这些特征被传递到全局残差学习结构和重构部分中,进而得到无雾图像.整个网络保留了低层的信息并将其传递到了深层,由于注意力机制的存在,使得网络加重对高雾度区域和低频纹理区域的处理.

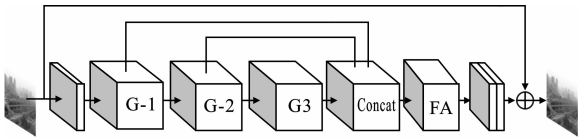


图 4 网络整体结构图

本文损失函数使用 L1 损失:

$$L(\theta) = \frac{1}{N} \sum_{i=1}^N \|I_{g^i} - FAA(I_h^i)\| \tag{7}$$

式中, $I_{g^i}$  表示无雾图像; $I_h$  表示有雾图像; $FAA$  表示本文网络.

3 实验

3.1 去雾图像数据集

本文数据集选用合成有雾图像数据集 RESIDE<sup>[8]</sup>的 OTS(Outdoor Training Set)数据集作为训练集,它是根据收集到的室外场景通过大气散射模型得到的,整个数据集包含 313,950 个合成有雾图像,随机选取其中的 20000 张图像作为训练集.随机选取 RESIDE 的 SOTS(Synthetic Objective Testing Set)数据集中的 500 张图像作为验证集.

在训练之前,对图像进行尺度归一化操作,将选用的所有图像尺寸统一为  $240\times 240$  大小.

3.2 实验环境配置

实验所用计算机网络训练配置信息如表 2 所示.

表 2 计算机网络训练配置信息

|      |                   |
|------|-------------------|
| CPU  | i7-9700K          |
| 实验框架 | Pytorch           |
| 显卡   | Nvidia GTX-1080Ti |
| 操作系统 | Ubuntu18. 04      |

3.3 网络参数设置

训练时,对输入图像进行随机旋转 90°、180°、270°和水平翻转以达到训练数据增强的效果. 设置 batchsize 为 2,训练迭代次数为  $1\times10^6$  次. 梯度下降优化算法采用 Adam 算法. 设置初始学习率为  $1\times10^{-4}$ ,采用余弦退火策略来调整学习率.

3.4 基于合成有雾图像的实验结果与分析



图 5 不同去雾算法的效果

为验证本文去雾算法对合成有雾图像的去雾效果,将本文算法与现有去雾算法进行对比,不同去雾算法的效果如图 5 所示.

从图中可以看出,DCP 算法在处理近景区域时去雾效果不理想,仍存在大量的雾,在处理高频区域

如图中的天空部分时,会导致图像颜色的失真;DehazeNet 算法存在的问题与 DCP 算法类似,图像去雾并不彻底,且天空部分颜色出现轻微失真;AOD-Net 算法处理后的图像颜色整体偏暗,导致掩盖住原本图像应有的细节;GCANet 算法同样使得图像色彩变暗;GridDehazeNet 算法与本文算法处理的图像并未出现以上现象,且本文算法处理后的图像在主观评价上比 GridDehazeNet 算法的图像更加自然. 综上,本文算法去雾效果显著,去雾后图像保存完整,相比较其他算法更加清晰,纹理和细节明显,由于引入了注意力机制,使得去雾后的天空区域色彩保真度更高,效果更加自然,与真实清晰图像更接近.

仅通过主观评价不足以验证本文算法去雾的有效性,因此本文选取三种常用图像评价指标峰值信噪比(PSNR)、结构相似度(SSIM)和特征相似度(FSIM)对不同算法的去雾效果进行对比. 由图 5 可知,无论是本文算法还是对比算法,对于近景图像和远景图像的处理效果并不相同,因此随机等量选取 SOTS 数据集中的近景有雾图像和远景有雾图像对其使用不同算法进行去雾,对去雾后的图像计算其 PSNR、SSIM 和 FSIM,结果如表 3 所示.

从表 3 可以看到,对于近景图像和远景图像,GridDehazeNet 算法的评价指标均为对比算法中最高. 对于远景图像,本文算法在 PSNR、SSIM 和 FSIM 上比 GridDehazeNet 分别高出 6. 79 db、0. 012 5 和 0. 004 6. 对于近景图像,本文算法在 SSIM 上比 GridDehazeNet 低了 0. 003 2,但在 PSNR 和 FSIM 上分别比 GridDehazeNet 高出 1. 2 db 和 0. 004 1. 综上,本文算法在远景图像上的指标均高于对比算法,在近景图像上的指标高于大部分算法,说明本文算法的有效性,经本文算法处理的去雾图像更接近真实无雾图像.

4 消融实验

为了进一步验证网络的有效性,对所提出的不同模块进行了消融实验. 主要关注基本模块中的三个结构:多尺度卷积、局部残差学习和特征注意对网络性能的影响. 将图像统一为  $48\times48$ ,训练次数为  $3\times10^5$ ,其它实验细节不变. 最终结果见表 4. 结果表明,基本模块中的三个结构均对网络性能有着重要的影响,它们的结合对于提升网络性能起到了明显的作用.

表 3 不同算法评价指标结果

| 方法                           | 近景图像    |       |         | 远景图像    |       |         |
|------------------------------|---------|-------|---------|---------|-------|---------|
|                              | SSIM    | PSNR  | FSIM    | SSIM    | PSNR  | FSIM    |
| DCP <sup>[1]</sup>           | 0.765 7 | 12.76 | 0.910 7 | 0.867 5 | 16.89 | 0.944 3 |
| DehazeNet <sup>[3]</sup>     | 0.778 3 | 13.88 | 0.905 9 | 0.962 7 | 24.48 | 0.981 3 |
| AOD-Net <sup>[4]</sup>       | 0.866 2 | 21.21 | 0.939 7 | 0.792 1 | 16.55 | 0.937 4 |
| GCANet <sup>[5]</sup>        | 0.893 4 | 22.61 | 0.967 5 | 0.8673  | 20.31 | 0.965 7 |
| GridDehazeNet <sup>[6]</sup> | 0.980 4 | 27.52 | 0.987 8 | 0.981 2 | 27.77 | 0.993 1 |
| This work                    | 0.977 2 | 28.72 | 0.991 9 | 0.993 7 | 34.56 | 0.997 7 |

表 4 不同模块对网络性能的影响

|        |       |       |       |
|--------|-------|-------|-------|
| 特征注意   | ✓     | ✓     | ✓     |
| 局部残差学习 |       | ✓     | ✓     |
| 多尺度卷积  |       |       | ✓     |
| PSNR   | 29.75 | 30.26 | 31.40 |

5 结束语

本文提出了一种基于注意力机制的端到端的单幅图像去雾算法,通过引入注意力机制构建特征注意模块,以及多尺度卷积对细节特征的提取,结合残差学习的使用,使得本文算法在图像去雾的有效性方面较之其他算法更加优异. 通过对比 DCP 算法、DehazeNet 算法、AOD-Net 算法、GCANet 算法以及 GridDehazeNet 算法与本文算法得到的去雾图像,表明使用本文提出的去雾算法无论是在主观视觉上还是在客观评价上均优于其他算法,又通过消融实验证明本文网络结构的有效性.

参考文献:

[1] 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. DOI: 10.1109/TPAMI.2010.168.

[2] BERMAN D, TREIBITZ T, AVIDAN S. Non-local imagedehazing[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas; IEEE, 2016: 1674-1682. DOI: 10.1109/CVPR.2016.185.

[3] 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. DOI: 10.1109/TIP.2016.2598681.

[4] LI B Y, PENG X L, WANG Z Y, et al. AOD-Net: all-in-one dehazing network[C]//Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV). Venice; IEEE, 2017: 4770-4778. DOI:

10.1109/ICCV.2017.511.

[5] CHEN D D, HE M M, FAN Q N, et al. Gated context aggregation network for image dehazing and deraining[C]//Proceedings of the 2019 IEEE Winter Conference on Applications of Computer Vision (WACV). Waikoloa Village; IEEE, 2019: 1375-1383. DOI: 10.1109/WACV.2019.00151.

[6] LIU X H, MA Y R, SHI Z H, et al. GridDehazeNet: attention-based multi-scale network for image dehazing[C]//Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul; IEEE, 2019: 7314-7323. DOI: 10.1109/ICCV.2019.00741.

[7] 刘广洲, 李金宝, 任东东, 等. 密集连接扩张卷积神经网络的单幅图像去雾[J/OL]. 计算机科学与探索: 1-13 [2020-06-10]. <https://kns.cnki.net/kcms/detail/11.5602.TP.20200609.1749.006.html>.

LIU G Z, LI J B, REN D D, et al. Single Image Dehazing Method Based on Densely Connected Dilated Convolution Neural Network[J/OL]. Journal of Frontiers of Computer Science and Technology: 1-13 [2020-06-10]. <https://kns.cnki.net/kcms/detail/11.5602.TP.20200609.1749.006.html>.

[8] 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. DOI: 10.1109/TIP.2018.2867951.

作者简介:

杜帆 男, (1996-), 硕士研究生. 研究方向为图像处理.

蒋品群(通讯作者) 男, (1970-), 博士, 副教授. 研究方向为集成电路设计和机器视觉.

E-mail: pqjiang@mailbox.gxnu.edu.cn.

宋树祥 男, (1970-), 博士, 教授. 研究方向为集成电路设计和机器视觉.

夏海英 女, (1983-), 博士, 教授. 研究方向为深度学习和数字图像处理.