

# 多深度自适应融合去雾生成网络

文 斌<sup>1,2</sup>, 彭 顺<sup>1,2</sup>, 杨 超<sup>1</sup>, 沈艳军<sup>1</sup>, 李 辉<sup>3</sup>

(1. 三峡大学电气与新能源学院, 湖北宜昌 443002; 2. 湖北省输电线路工程技术研究中心, 湖北宜昌 443002; 3. 电子科技大学航空航天学院, 成都 611731)

**摘要:** 针对图像去雾不彻底和失真等问题, 采用生成对抗的训练方式, 提出了一种多深度自适应融合的去雾生成网络。首先, 采用 U-Net++ 结构和雾霾感知单元来学习雾霾特征。其次, 提出一种向量混合注意力模块扩张底层信息, 补充去雾图像的细节; 再次, 从部分和整体维度构建自适应权重来选取不同深度的特征, 提升网络对有效信息的利用率; 最后, 采用混合损失来确保生成图像的质量, 并在对抗损失中引入 Wasserstein 距离。为了验证本文算法的有效性, 在 RESIDE 和 Haze4k 数据集上将本文算法和 10 种主流去雾算法进行客观定量对比, 然后在真实图像上进行主观评价。实验结果表明: 在 SOTS Outdoor 验证集上 PSNR 达到 36.20 dB, SSIM 达到 0.988 4, 具有更好的去雾效果。

**关键词:** 信息处理技术; 图像去雾; 生成对抗网络; 多深度特征; 自适应特征融合

**中图分类号:** TP391.4 **文献标志码:** A **文章编号:** 1671-5497(2025)06-2103-11

**DOI:** 10.13229/j.cnki.jdxbgxb.20230813

## Multi-depth adaptive fusion dehazing generation network

WEN Bin<sup>1,2</sup>, PENG Shun<sup>1,2</sup>, YANG Chao<sup>1</sup>, SHEN Yan-jun<sup>1</sup>, LI Hui<sup>3</sup>

(1. School of Electrical and New Energy, China Three Gorges University, Yichang 443002, China; 2. Hubei Provincial Engineering Technology Research Center for Power Transmission Line, Yichang 443002, China; 3. School of Aeronautics and Astronautics, University of Electronic Science and Technology of China, Chengdu 611731, China)

**Abstract:** To address issues such as incomplete dehazing and distortion in image dehazing, a multi-depth adaptive fusion dehazing generation network is proposed using generative adversarial training. Firstly, the network utilizes U-Net++ architecture and haze perception units to learn the haze features. Secondly, a vector mixed attention module is proposed to expand the bottom-layer information and supplement details in the dehazed image. Thirdly, adaptive weights are constructed from partial and global dimensions to select features at different depths and improve the utilization of effective information in the network. Finally, a mixed loss is employed to ensure the quality of the generated image, and Wasserstein distance is introduced into the adversarial loss. To validate the effectiveness of the proposed algorithm, objective quantitative comparisons are conducted with 10 popular dehazing algorithms on the RESIDE and Haze4k

收稿日期: 2023-08-03.

基金项目: 国家自然科学基金项目(62273200, 61876097); 湖北省输电线路工程技术研究中心研究基金项目(2022KXL03).

作者简介: 文斌(1985-)男, 讲师, 博士. 研究方向: 数字视频信号处理. E-mail: wenbin\_08@126.com

通信作者: 杨超(1989-)男, 讲师, 硕士. 研究方向: 工业智能控制技术. E-mail: yangchao\_0305@126.com

datasets, followed by subjective evaluations on real images. The experimental results demonstrate that the proposed algorithm achieves a PSNR of 36.20 dB and SSIM of 0.988 4 on the SOTS Outdoor validation set, showing superior dehazing performance.

**Key words:** information processing technology; image dehazing; generate adversarial network; multiple depth feature; adaptive feature fusion

## 0 引言

雾是一种大气散射现象,在雾霾天气中,由于光的反射会受到大气中颗粒的散射和衰减的影响,造成图像严重退化,使图像的信息丢失。严重的雾霾还会影响人们的出行,给交通系统和户外工作带来巨大挑战。因此,图像的去雾任务在近年来吸引了越来越多的研究者。

由于雾霾图像中雾气随场景深度增加,浓度也不断提升,导致获取理想的去雾图片十分困难。当前的去雾算法主要可以分为两类:基于先验信息<sup>[1]</sup>的去雾算法和基于深度学习的去雾算法。基于先验信息的去雾方法通常以大气散射模型为基础,以先验信息为媒介对透射率和大气光值进行估计,然后反演去雾图像。其中代表的算法有:He等<sup>[2]</sup>利用发现的暗通道先验进行去雾,但该方法在处理和大气光类似的白色物体时,会出现局部色斑和色偏效应;Raanan<sup>[3]</sup>通过构建雾霾场景中的颜色线模型来进行去雾,但该模型在户外场景中表现不佳;Tan<sup>[4]</sup>发现清晰图像比有雾图像有更高的对比度,基于该发现构建一个代价函数进行去雾,该方法不能很好地平衡饱和度和色彩的关系,容易造成图片失真的现象;Zhu等<sup>[5]</sup>在颜色衰弱的先验条件下,提出一种线性模型对模糊图像的景深图进行估计,由景深图来实现去雾。但是,上述基于先验信息的去雾方法不具有通用性,在实际中面对复杂场景时,先验信息往往会失效,去雾的效果就会变得很差。

随着深度学习的发展,神经网络在去雾领域也大放异彩。最初,神经网络的设计并未摆脱大气散射模型的束缚,主要思想是用神经网络代替先验信息去估算透射率和大气光值,代表方法有:Cai等<sup>[6]</sup>提出一个可训练的去雾网络(DehazeNet),利用输出的介质传输图和大气散射模型来进行去雾;Li等<sup>[7]</sup>为了简化计算,将模型参数进行统一,提出一种轻量级一体化的去雾网络(AOD-Net)。由于神经网络具有强大的学习能力,后续的网络设计逐渐摆脱了大气散射模型,采

用端对端的生成方式,直接将雾图片恢复成无雾图片。代表的方法有:Liu等<sup>[8]</sup>提出一种基于注意力的多尺度网络GridDehazeNet,该网络由3个模块组成:预处理、主干和后处理。Qu等<sup>[9]</sup>提出一种名为EPDN(Enhanced pix2pix dehazing network)的网络,利用生成对抗网络(Generative adversarial network, GAN)<sup>[10,11]</sup>进行去雾,在该网络中采用分级恢复的思想,通过多分辨率生成器和增强器将图像由粗糙变精细。Qin等<sup>[12]</sup>提出一种特征融合注意力网络FFA-Net(Feature fusion attention network),该网络利用残差嵌套注意力的方式实现去雾。上述基于深度学习的算法虽然在一定程度上优化了去雾图像的效果,但是网络结构并未充分考虑雾霾的特点,在场景深度大的区域,去雾表现较差,同时在色彩的保真度和细节的还原度上还有待提升。

针对上述问题,本文基于生成对抗网络,提出一种端到端的多深度自适应融合去雾生成网络(Multi-depth adaptive fusion dehaze generation network,MDAFG-Net),主要创新如下:

(1)引入多深度特征提取模块(Multi-depth feature extraction module,MDFM)获取不同语义信息的特征图,并采用雾霾感知单元(Haze perception block,HPB)来学习雾霾特征,可有效解决远景处雾霾难以消除的问题。

(2)提出了一种向量混合注意力模块(Vector mixed attention module,VMAM),弥补卷积过程中底层丢失的信息。

(3)对多特征的选取,采用自适应权重模块(Adaptive weight module,AWM)从部分和整体两个维度进行有侧重筛选。

(4)在原有的对抗损失中引入Wasserstein<sup>[13]</sup>距离,并采取混合损失去优化网络,保证图像的纹理细节和色彩还原度。

## 1 本文方法

Pix2pixgan<sup>[14]</sup>是一种典型的生成对抗网络,

具有强大的成对图片转化能力,可以将白天转化为黑夜、将图片转化为语义标签等。由于本文任务中有雾和无雾图片也是一组成对图片,因此,本文将去雾过程视为图片风格的转化过程,并借鉴 Pix2pixgan 的训练思路,开展训练工作。具体训练方法如图 1 所示:为了让网络具有分辨雾霾的能力,首先训练判别器(Discriminator, D)学习雾

霾的相关特征;然后训练生成器(Generator, G)将有雾图像转递到该模块直接生成去雾图像。在这个过程中,判别器需要不断纠正生成器,两者不断进行对抗进化,直至达到理想的去雾效果。然而, Pix2pixgan 中生成器 G 结构过于简单,难以应对复杂的去雾任务,因此在本文中重新优化设计了生成器 G,提高网络的去雾能力。

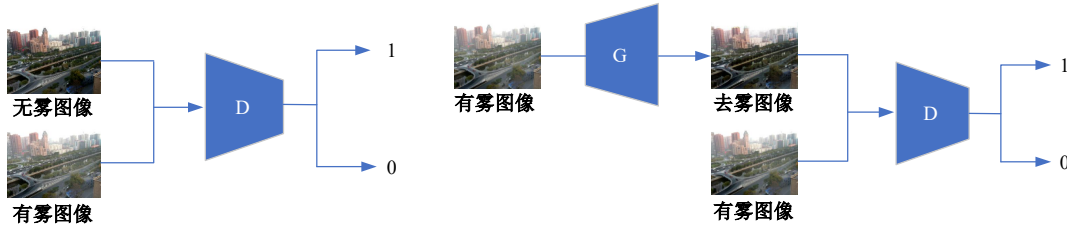


图 1 训练框架

Fig. 1 Training framework

### 1.1 多深度自适应融合去雾生成网络

生成器的设计在生成对抗网络中十分重要,其性能优劣直接关系到图像去雾的效果,本文所设计的生成器网络具体结构如图 2 所示,共分为 3

个部分:多深度特征提取模块 MDFM 和雾霾感知单元 HPB、向量混合注意力模块 VMAM 和自适应权重模块 AWM。在下节中,本文将针对这些模块的结构和功能做出详细介绍。

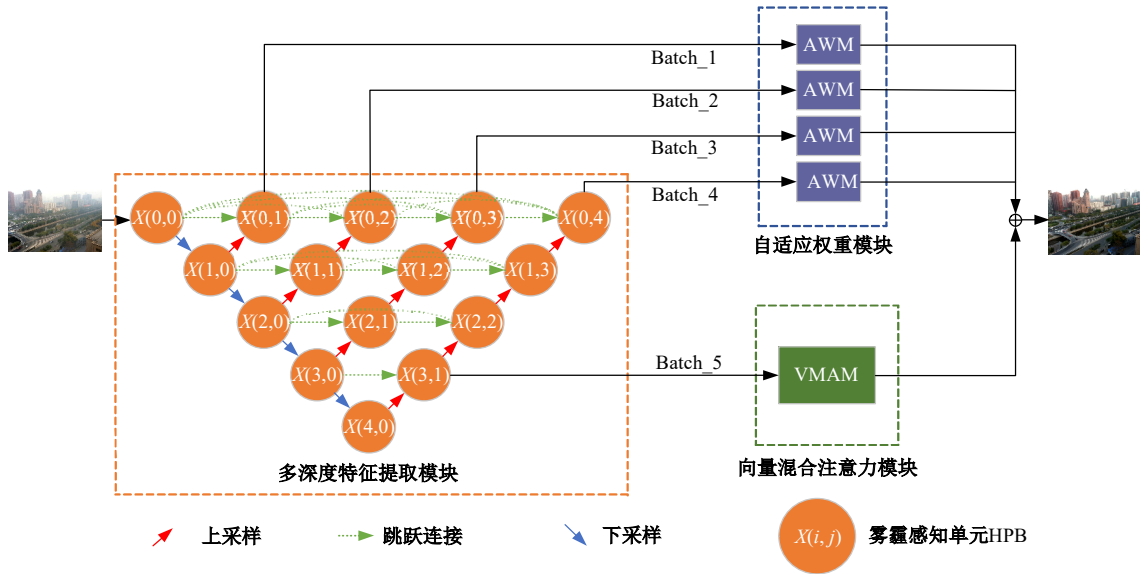


图 2 多深度自适应融合网络结构

Fig. 2 Multi-depth adaptive fusion network structure

#### 1.1.1 多深度特征提取模块

特征提取是整个生成器网络中的关键部分,在先前的研究中,一些工作使用 U-Net++<sup>[15]</sup>作为特征提取网络,取得了不错的成果,其具体结构如图 3 所示。从图中可以看到,该结构由 4 个深度依次递增的 U 型网络组成,并采用稠密连接方式提取不同深度的特征。通常来说,不同深度的网络具有不同的表达能力,浅层网络特征感受野较小,局部信息丰富,而深层网络特征感受野大,全

局信息丰富。在图像去雾任务中,需要利用充足的全局信息恢复图像主体,并结合细腻的局部信息对图像细节进行还原,以此达到理想的去雾效果。因此,本文采用 U-Net++ 的网络结构作为特征提取部分的整体架构。

虽然利用多深度的特征提取模块能够有效提取高级语义信息和浅层细节信息,但雾霾图像中由于存在遮掩导致部分信息丢失,难以还原。为了学习该部分特征,提高对雾霾的去除能力,增强

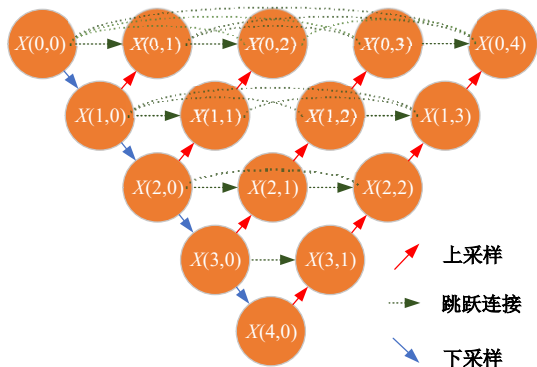


图 3 多深度特征提取模块结构

Fig. 3 Multi-depth feature extraction module structure

图像还原,本文提出了雾霾感知单元HPB,具体结构如图4所示。

首先,采用深度可分离卷积和不同卷积混合的方式,减少网络参数量,扩展网络的深度。其次,为了更加聚焦于有雾区域,提出全局感知模块,在水平和垂直方向上进行全局池化,通过拆分学习再合并的方式获取全局感知特征图。然后,用全连接的方式代替卷积,实现通道数的改变,相较于卷积只关注局部信息,全连接的方式能获取图像的整体信息,并采用反瓶颈结构,在中间层4倍扩张向量,引入一个比输入和输出更高的维度,从而增加网络的深度和复杂度,增强模型的表达能力。考虑到网络在训练的过程中,易发生过拟合和网络退化的问题,提出层缩放和随机子路径删除的策略来解决。最后,采用跳跃连接的方式,增加输入特征和输出特征的交互,从而使网络更容易学习到雾霾的特征,提高图像的保真度。为了保证梯度传输过程的稳定性,在非线性激活函数的选取上,采用 GELU (Gaussian error linear

unit)激活函数代替常用的 Relu(Recitified linear unit)激活函数。相较于 Relu 激活函数, GELU 激活函数在零处的梯度更加平滑。在归一化函数的选取上,采用 LN (Layer normalization) 层代替 BN (Batch normalization) 层,由于 BN 层是从 Batch 维度进行标准化,所以当 Batch 很小时,效果并不理想,而 LN 层是从通道维度进行标准化,并不受 Batch size 的影响。

1.1.2 向量混合注意力模块

为了让底层特征与浅层的特征相融合,在多深度特征提取模块中,采用了大量的上采样操作用于恢复特征分辨率。然而,网络在上采样过程中会损失大量的语义信息和通道信息,导致融合后的特征不完全、不充分,使底层特征的提取闲置。鉴于此,本文在雾霾感知单元 \$X(3,1)\$ 出一条分支 Batch\_5 (见图 2),并提出一种混合向量注意力模块对特征进行补充。

与其他注意力机制所不同的是,本文采用一种基于向量相似性度量的计算方法去计算信息权重,其计算公式如下所示:

$$\cos \theta = \frac{\vec{A} \cdot \vec{B}}{|\vec{A}| \cdot |\vec{B}|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \quad (1)$$

式中: \$A\$、\$B\$ 为不同向量; \$|\*|\$ 为模值计算,当向量 \$A\$、\$B\$ 的相似度越高时,向量的乘积越大,则其夹角越小。

向量混合注意力模块结构如图 5 所示,首先进行卷积操作改变通道数,然后采用像素注意力和通道注意力混合的方式去关注空间像素的相关

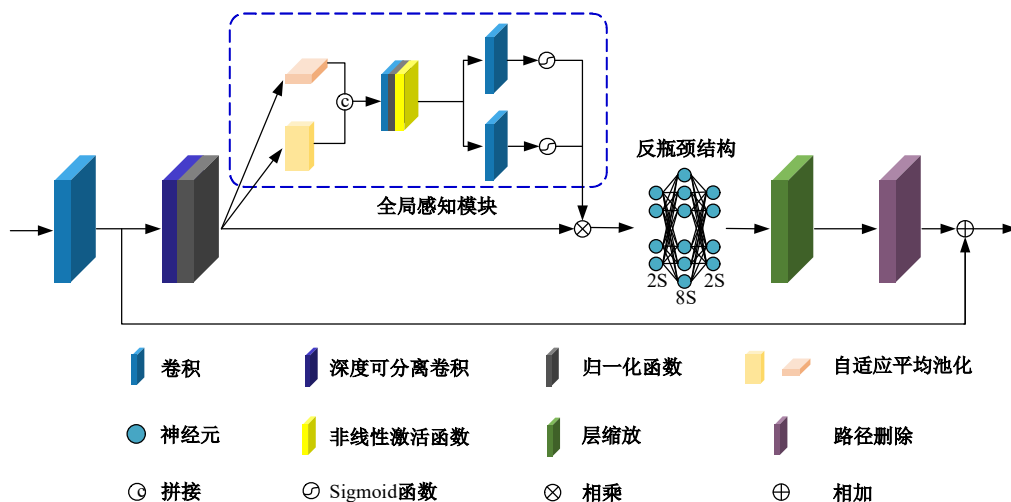


图 4 雾霾感知单元结构

Fig. 4 Haze perception block structure

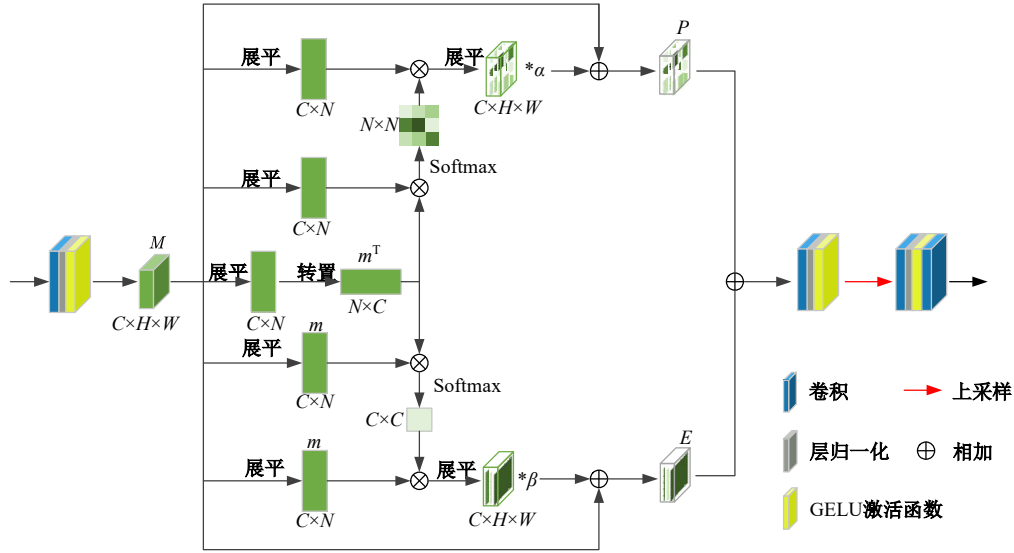


图 5 向量混合注意力模块结构

Fig. 5 Vector mixed attention module structure

性和通道间联系,弥补卷积操作局部性过强,全局性不足的问题,最后直接将图像恢复到原图大小进行特征补充。像素注意力  $P$  的计算公式如下所示:

$$p_j = \alpha \sum_{i=1}^N \frac{\exp(m_i^T \cdot m_i)}{\sum_{i=1}^N \exp(m_i^T \cdot m_i)} \cdot m_i + M_j \quad (2)$$

式中:  $M$  为三维特征图;  $m$  为二维矩阵;  $\exp$  为进行 softmax 归一化计算;  $p_j$  为经过像素注意力机制增强后,在空间位置  $j$  上的新特征图;  $\alpha$  为可学习的参数。

通过展平的方法将  $M$  转化为  $m$ ,并将矩阵  $m$  与自身转置矩阵  $m^T$  相乘,得到任意两个像素之间的空间相关性,通过 Softmax 函数得到不同像素的权重值。同时,引入参数  $\alpha$ ,让权重分配更加合理。

通道注意力  $E$  的计算公式如下所示:

$$E_j = \beta \sum_{i=1}^C \frac{\exp(m_i \cdot m_i^T)}{\sum_{i=1}^C \exp(m_i \cdot m_i^T)} \cdot m_i + M_j \quad (3)$$

式中:  $E_j$  为经过通道注意力机制增强后,在空间位置  $j$  上的新特征图;  $\beta$  为可学习的参数。

通道注意力与像素注意力相同,唯一的区别在于通道注意力通过  $m \cdot m^T$  获取大小  $C \times C$  的矩阵,在  $C \times C$  矩阵中包含了任意两个通道的联系。

### 1.1.3 自适应权重模块

来自多深度特征提取模块的 5 条分支特征被用于恢复无雾图像,考虑到不同路径的特征所携带的信息不同,对去雾效果也具有不同的影响,如

果平等对待每条路径,通过硬融合的方式对特征进行串联,会限制网络表达,造成有效信息利用率低下的问题。对此,本文提出自适应权重模块 AWM,采用 CBAM<sup>[16]</sup> 和自适应权重机制去筛选所需特征,其结构如图 6 所示。

其中  $F$  代表特征提取后的特征图,  $f$  代表筛选后的特征图,  $M_c$  代表从图像通道维度的选取,首先通过平均池化和最大池化的方式,从通道维度对特征图进行筛选,然后采用多层感知机 MLP 对通道信息进一步整合,最后经过 Sigmoid 激活得到通道维度的权重筛选图  $M_c(F)$ ,其计算过程如下所示。

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) = \sigma(m_1(m_0(F_{\text{avg}}^c)) + m_1(m_0(F_{\text{max}}^c))) \quad (4)$$

式中:  $\sigma(\cdot)$  为 Sigmoid 函数; MLP 为多层感知机;  $m_0$  和  $m_1$  分别为 MLP 中进行层缩放权重和层还原权重,  $m_0 \in \mathbb{R}^{C \times C}$ ,  $m_1 \in \mathbb{R}^{C \times C}$ ,  $r$  的大小取 8;  $F_{\text{avg}}^c$  和  $F_{\text{max}}^c$  分别为进行平均池化和最大池化后的结果图。

$M_p$  表示从图像像素维度的选取,将平均池化和最大池化后的特征图进行拼接,然后通过卷积操作来提高对空间中有效像素的利用率,计算公式如下所示:

$$M_p(F) = \sigma(g^{3 \times 3}([\text{AvgPool}(F); \text{MaxPool}(F)])) = \sigma(g^{3 \times 3}([F_{\text{avg}}^s; F_{\text{max}}^s])) \quad (5)$$

式中:  $g^{3 \times 3}$  为  $3 \times 3$  的卷积运算;  $F_{\text{avg}}^s$  和  $F_{\text{max}}^s$  分别为进行平均池化和最大池化后的结果图。

$W$  代表从路径这个整体维度进行选取,故整

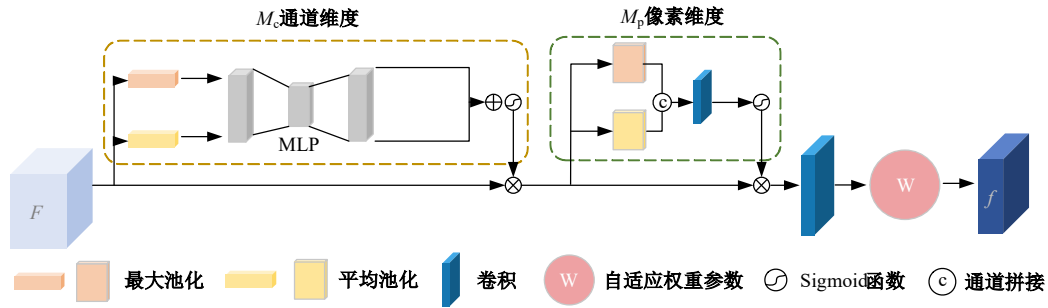


图 6 自适应权重模块结构

Fig. 6 Vector mixed attention module structure

个自适应权重模块 AWM 的流程,如下所示:

$$AWM = W \cdot \text{Conv}(M_s(M_c(F))) \quad (6)$$

式中:Conv 为卷积操作,进行通道数的改变;W 为可改变的自适应权重参数,随着网络的训练,根据不同的数据特征不断优化,直至达到最优的路径权重分配。

最终的输出 output 由 5 条路径构成,4 个 AWM 模块和一个向量混合注意力模块 VMAM,为了让特征的融合更加合理,在向量混合注意力模块 VMAM 对应的 Batch\_5 路径中也加入了自适应权重,如下所示:

$$\begin{cases} \text{output} = \sum_{i=1}^4 AWM_i + W_5 \cdot \text{VMAM} \\ w_i = \frac{e^{\alpha_i}}{\sum_j e^{\alpha_j}}, i = j = 1, 2, \dots, 5 \end{cases} \quad (7)$$

式中:W<sub>i</sub>为自适应权重,α<sub>i</sub>为初始化权重系数,起始值均为 0.2。

在图 7 中记录了权重值和峰值信噪比随训练轮次变化的过程。由图 7 可知,在网络训练的过程中,w<sub>1</sub>、w<sub>2</sub> 权重呈上升趋势,w<sub>3</sub>、w<sub>4</sub>、w<sub>5</sub> 权重呈

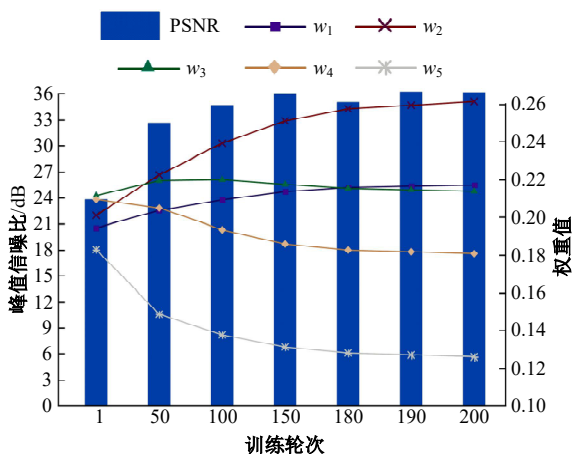


图 7 权重变化过程

Fig. 7 Weight change process

下降趋势,在 epoch=190 轮达到最优效果,其中 w<sub>2</sub> 占比相对较大,w<sub>5</sub> 占比相对较小。这表明在 w<sub>2</sub> 对应的 Bath\_2 路径中,该深度的网络相较其他深度的网络具有更强的雾霾去除能力。同时 w<sub>5</sub> 对应的 VMAM 模块也并非占比越小越好,在后面的消融实验中也验证了 VMAM 模块的有效性。

### 1.2 损失函数

为了优化去雾生成模型,本文使用了混合损失函数,该混合损失由 3 个部分组成,具体公式如下:

$$L_{\text{gen}} = \lambda_1 L_{\text{PM}} + \lambda_2 L_P + L_A \quad (8)$$

式中:L<sub>PM</sub>为像素匹配损失;L<sub>P</sub>为视觉感知损失;L<sub>A</sub>为对抗损失。对抗损失 L<sub>A</sub>和视觉感知损失 L<sub>P</sub>用于恢复图像的全局信息,像素匹配损失 L<sub>PM</sub>用于增强图片的细节并确保颜色的一致性。

像素匹配损失(Pixel matching loss):为了获得真实的图像,本文通过像素之间的逐一比较来完善对抗损失,如下所示:

$$L_{\text{PM}} = E_{x,k}(\|y - k\|_1) \quad (9)$$

式中:‖\*‖<sub>1</sub>表示 L<sub>1</sub> 范式。

通过像素匹配损失可以有效提升生成器的学习效果,恢复图像中被雾霾所遮挡的高频信息。

视觉感知损失(Perceptual loss):传统的 MSE 损失忽略了图片的高级语义信息,导致生成的图片不符合视觉系统的直观感受。为了解决该问题,提高生成图片视觉感知的相似性,本文提取 resnet-18 网络的前 n 层(表示为 φ<sub>n</sub>)获取无雾图像 y 和去雾图像 k 的特征图;将感知损失定义为两个特征图之间的 MSE 距离,如下所示:

$$L_P = \frac{1}{CWH} \sum_{m=1}^C \sum_{i=1}^W \sum_{j=1}^H [\phi_n(y_{m,i,j}) - \phi_n(k_{m,i,j})]^2 \quad (10)$$

式中: $C$ 为图像通道数; $W$ 为图像宽度; $H$ 为图像高度。

特征图所带来的抽象信息可以对图像的高级语义进行还原。

对抗损失(Adversarial loss):在 GAN 的训练中,生成器的损失由判别器传递,判别器采用 JS 散度去衡量两个分布之间的距离,如下所示:

$$\begin{cases} D_{\text{KL}}(p // q) = \int_x p(x) \log \frac{p(x)}{q(x)} dx \\ D_{\text{JS}}(p // q) = \frac{1}{2} D_{\text{KL}}(p // \frac{p+q}{2}) + \frac{1}{2} D_{\text{KL}}(q // \frac{p+q}{2}) \end{cases} \quad (11)$$

式中: $p(x)$ 为真实数据分布; $q(x)$ 为生成数据分布; $D_{\text{KL}}$ 为 KL 散度,用于测量一个分布相对于另一个分布的信息损失或不确定性; $D_{\text{JS}}$ 为 JS 散度,通过计算 KL 散度的平均值衡量两个分布之间的相似性。

JS 散度具有局限性,如图 8 所示,只有当两个分布有不可忽略的重叠时,才能通过优化 JS 散度去拉近生成分布和真实分布的距离,如果两者并未重叠或者重叠部分可忽略,其 JS 散度则为  $\log 2$ 。由于当两个分布的支撑集是高维空间的低维流形时,两分布重叠部分测度为 0 的概率为 100%,所以生成分布和原始分布几乎不能有不可忽略重叠,使用 JS 散度测量的结果始终为  $\log 2$ ,这就导致生成器会产生梯度消失和生成样本单一的问题。为了解决上述由 JS 散度带来的局限,本文决定引入 Wasserstein 距离用于衡量两个分布。相较于 JS 散度,Wasserstein 距离的优点在于,即使两个分布没有重叠的情况下,也可以反映它们之间的距离。引入 Wasserstein 距离后对抗损失如下所示:

$$L_A = \max_{D \in 1\text{-lipschitz}} \{E_{x,y}[D(x,y)] - E_{x,k}[D(x,k)]\} \quad (12)$$

式中: $E(\cdot)$ 为分布函数的期望值; $D$ 为判别器, $x$ 为有雾图像; $y$ 为无雾图像; $k = G(x, z)$ 为生成的样本, $G$ 为生成器; $D \in 1\text{-lipschitz}$ 为对判别器  $D$  参数的限定条件,为了让判别器的变化更平缓,本文中判别器参数固定在  $(-0.01, 0.01)$  之间。

本文记录了原始对抗损失的损失曲线和引入 Wasserstein 距离后的损失曲线,如图 9 所示。从损失曲线的比较上看,采用原始损失训练时,如图

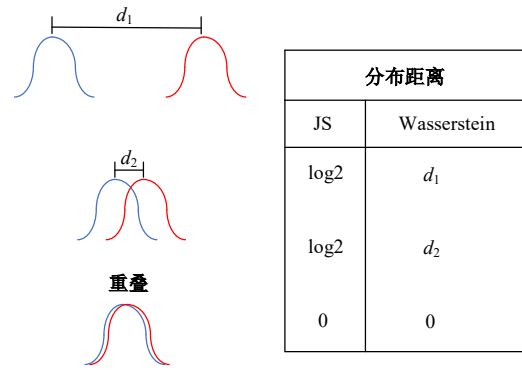
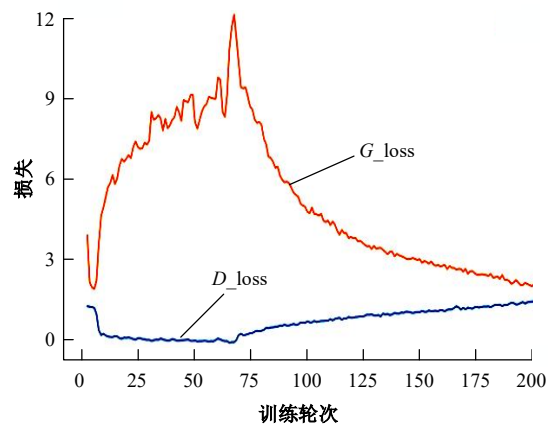


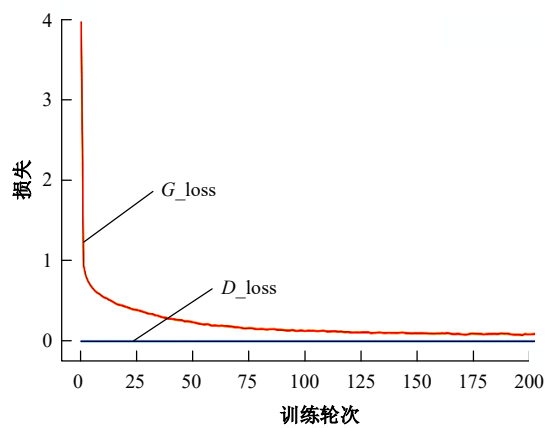
图 8 不同距离的衡量结果

Fig. 8 Measurement of different distances

9(a)所示,前 100 轮生成器损失较大,难以学习有效特征。引入 Wasserstein 距离后,如图 9(b)所示,生成器迅速收敛,在第 75 轮就趋于平缓,大大提高了有效特征的学习能力。



(a)原始损失



(b)引入Wasserstein距离的损失

图 9 训练损失曲线

Fig. 9 Training loss curve

## 2 实验及结果分析

### 2.1 实验设置及数据集

为了验证本文算法的有效性,在合成图片和

真实图片上进行实验,采用 RESIDE<sup>[17]</sup> 和 Haze4k<sup>[18]</sup> 两个公共数据集。在 RESIDE 数据集上进行实验时,分室内和室外两组进行。室内训练采用室内合成数据集(Indoor training set, ITS)中的 13 990 幅图片,室外训练采用室外合成数据集(Outdoor training set, OTS)中不同场景的 8 970 幅图像,并在合成客观数据集(Synthetic objective testing set, SOTS)上进行测试。在 Haze4k 数据集上进行实验时,采用训练集和测试集中的全部图片。

在进行网络训练的过程中,采用 RMSProp 优化器,学习率设置为  $2 \times 10^{-4}$ 。损失函数中超参数设置为  $\lambda_1, \lambda_2=5, \lambda_3=10$ ,使用 PyTorch 框架搭建模型, GPU 为 NVIDIA GeForce GTX 2080 Ti,进行 200 轮训练。

在图像质量评价中,分为合成数据和真实数据集两个部分。在合成数据中,将本文的算法和其他 10 种主流算法(DCP<sup>[2]</sup>、CAP<sup>[3]</sup>、AOD-Net<sup>[7]</sup>、Dehazed<sup>[6]</sup>、FD-GAN<sup>[10]</sup>、HIDE-GAN<sup>[11]</sup>、GridDehazed<sup>[8]</sup>、FFA-Net<sup>[12]</sup>、SADNet<sup>[19]</sup>、GANID<sup>[20]</sup>)从客观层面进行对比,并挑选其中 6 种算法进行可视化的主观分析。在真实数据集中由于缺少无雾的对比图,同样采用主观分析的方式。在客观对比时,采用峰值信噪比(Peak signal to noise ratio, PSNR)、结构相似性(Structural similarity, SSIM)。峰值信噪比 PSNR 通过计算两幅图像各个像素的差异来实现对图像的评价,其值越大代表图像质量越高。结构相似性 SSIM 通过分析图片亮度、对比度和结构 3 个因素来判断图片质量的好坏,其设计更符合人眼视觉感知系统的特点,取值范围为 0~1,相似度越大,图片质量越高。

## 2.2 合成数据集上实验

在客观数据层面上分析,结果如表 1 所示。在 Haze4k 和 SOTS Outdoor 数据集的定量比较中,本文方法在 PSNR 和 SSIM 均取得了最佳的表现,在 PSNR 指标的比较上,本文方法相较于排名第二的 FFA-Net 算法,在 SOTS Outdoor 数据集上提升了 7%,在 Haze4k 数据集上提升了 2%。并且,本文方法在两个数据集上的 SSIM 均超过 0.96。在 SOTS Indoor 数据集上本文方法在 PSNR 指标中排名第二,SSIM 指标排名第二,在室内场景的表现稍逊一筹。综合多个数据集来看,本文算法去雾和泛化能力更强。

表 1 合成数据集上不同方法定量比较  
Table 1 Comparison of different methods on synthetic data sets

Method	SOTS(indoor)		SOTS(outdoor)		Haze4k	
	PSNR /dB	SSIM /%	PSNR /dB	SSIM /%	PSNR /dB	SSIM /%
DCP	16.62	0.817 9	19.13	0.814 8	14.01	0.767 9
CAP	19.05	0.843 1	15.93	0.815 0	14.46	0.672 4
AOD-Net	19.06	0.850 4	20.29	0.876 5	17.15	0.833 9
Dehazed	21.14	0.847 2	22.46	0.851 4	19.12	0.846 7
FD-GAN	23.45	0.925 8	23.68	0.901 7	20.39	0.894 3
HIDE-GAN	25.01	0.871 2	26.64	0.892 3	21.72	0.915 2
GridDehazed	32.16	0.983 6	30.86	0.981 9	23.29	0.931 2
FFA-Net	36.39	0.988 6	33.57	0.984 0	26.96	0.957 3
SADNet	23.85	0.902 3	26.97	0.927 4	21.40	0.907 8
GANID	32.52	0.936 7	33.96	0.949 1	—	—
本文	34.20	0.985 1	36.20	0.988 4	27.67	0.968 9

在主观视觉层面上,挑选了其中具有代表性的 6 种算法进行比较,如图 10 所示。DCP 由于使用了暗通道先验的理论,对不满足先验信息的区域,会造成画面严重的失真,比如图片中的天空区域,经 DCP 处理后变成深蓝色;CAP 通过判断有雾区域与无雾区域之间的亮度和饱和度差异去雾,常造成过曝的现象;HIDE 虽然能去除一定的雾霾,但是受雾霾的影响,去雾后的图片整体色彩的还原度不高,会导致整体亮度过暗;GridDehazeNet 采用基于学习的方法,在图像对比度和饱和度方面有明显提升,但是依然存在去雾不彻底的问题,比如在图 10 中第二行图片的街道处,第五行图片中房子区域依然有大量白色雾霾;FFA-Net 从整体效果和色彩的还原度上,达到了令人满意的视觉效果。但是在细节上,对场景深处区域的雾霾,并未做到有效的处理,比如在图 10 第三行图中远处人行道、第四行图中远处石桥和第五行图中绿色公交车区域,均未能做到有效去雾。SADNet 是一种去噪网络,在对雾霾的处理上效果并不理想,依然可以看见大量的雾霾残留。而本文算法在对远景处雾霾的处理和色彩的还原度上均超过其他算法,从主观角度来看和标签无雾图片并无差别。

## 2.3 真实数据集上实验

为了进一步验证本文算法的有效性,在真实户外图片上进行实验。从主观视觉层面上分析,如图 11 所示,DCP 方法在处理天空区域时造成严重失真,图 11 中的 pic 1 和 pic 3 天空发生了严重

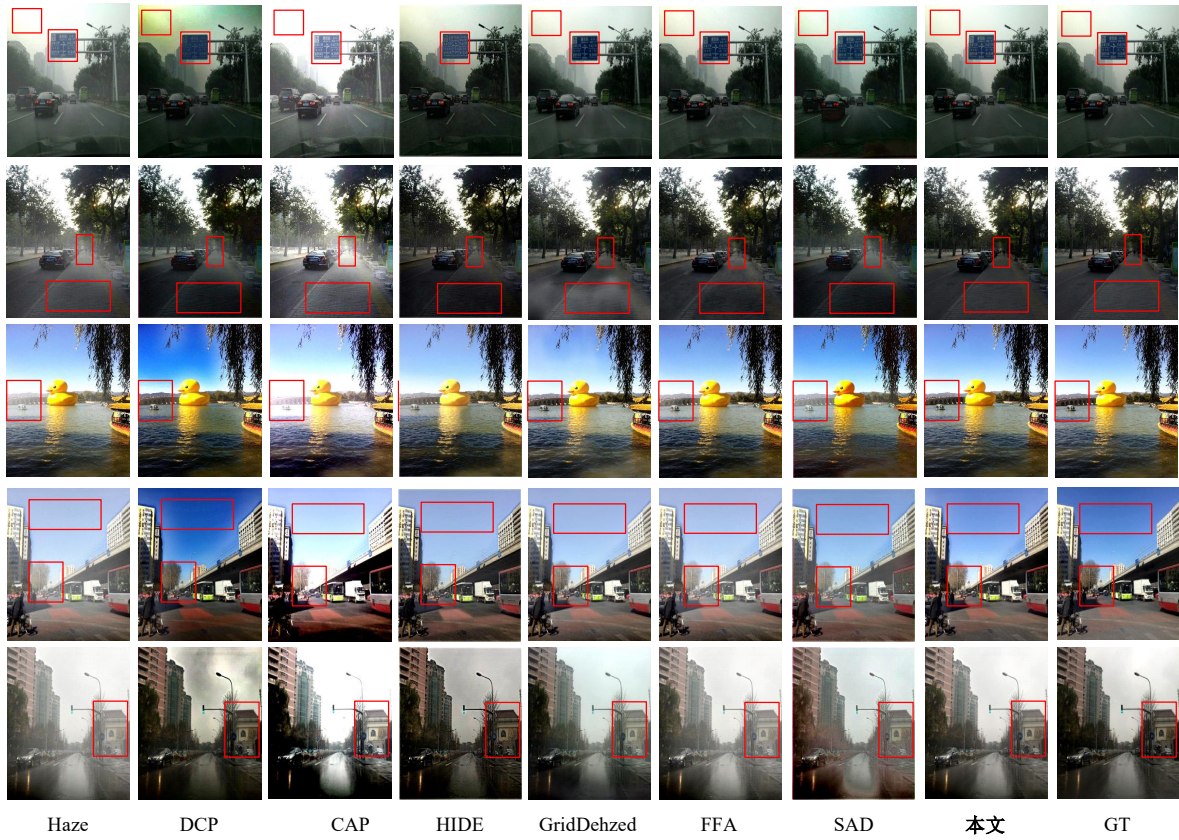


图 10 合成数据集上不同算法对比图

Fig. 10 Comparison of different algorithms on the synthetic data set

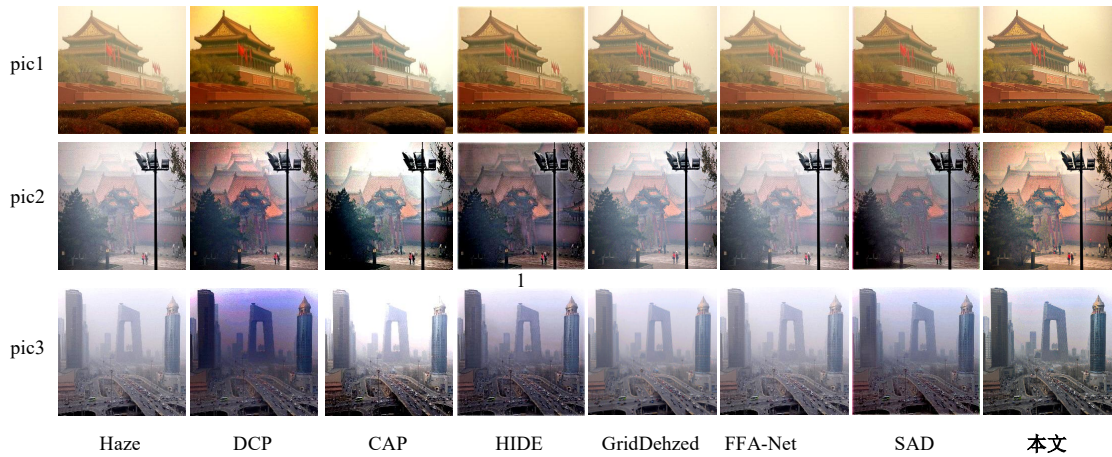


图 11 真实数据集上不同算法对比图

Fig. 11 Comparison of different algorithms on the real dataset

的畸变;CAP在处理的过程中天空会出现过曝的现象;HIDE对远景处的雾霾难以消除,同时会造成伪影和饱和度过低的现象;SAD虽然在色彩的还原度和饱和度上有较好还原,但是对雾霾的消除并不彻底,依然存在明显的雾霾残留;Grid-Dehazed和FFA-Net方法从人眼的角度,并未发现和原图的区别,去雾效果很不理想。相较上述算法,本文算法在去除雾霾的同时,也能很好保留

图片的颜色信息,符合人眼的视觉感知。

#### 2.4 消融实验

为了验证本文方法在去雾过程中的效果,在SOTS Outdoor上对生成器中对不同模块和不同损失函数进行消融实验。在对模块的消融实验中,主要关注以下3个要素:①向量混合注意力模块VMAM;②自适应权重模块AWM;③雾霾感知单元HPB。首先,在模型a中采用全部要素,为

后续实验对比带来参考。考虑向量混合注意力模块 VMAM 对模型结果的影响,模型 b 去除了向量混合注意力模块 VMAM。考虑自适应权重模块 AWM 对模型结果的影响,模型 c 去除了自适应权重模块 AWM。考虑雾霾感知单元 HPB 对模型结果的影响,模型 d 采取原始 U-Net++ 中的特征提取单元进行替换。

不同模块消融实验结果见表 2,不同模块对 PSNR 均有提升,但是 SSIM 的变化不大,均维持在 0.98 附近,其中 VMAM 模块提升 1.16 dB,AWM 模块提升 1.82 dB,HPB 模块提升 0.81 dB,AWM 模块对 PSNR 的提升最大。

表 2 不同模块消融实验

Table 2 Different module ablation experiments

模型	VMAM	AWM	HPB	PSNR/dB	SSIM/%
模型 a	✓	✓	✓	36.20	0.988 4
模型 b		✓	✓	35.04	0.985 3
模型 c			✓	33.22	0.984 3
模型 d				32.41	0.983 9

在对损失函数的消融实验中,均采用相同的网络模块,主要关注以下 3 个要素:①像素匹配损失  $L_{PM}$ ;②视觉感知损失  $L_p$ ;③引入 Wasserstein 的对抗损失  $W_{loss}$ 。

不同损失函数消融实验结果见表 3,所提的每个损失都能增强网络的去雾效果,其中  $W_{loss}$  带来的提升最大,在使用  $W_{loss}$  后,PSNR 提升了 2.06 dB,SSIM 提升了 0.02。由对比训练的曲线(见图 9)可知,Wasserstein 距离有效解决了判别器过强生成器较弱带来的难以训练的问题,可以让生成器在每一轮的训练中都学到有用的特征,从而快速收敛。综合表 2 和表 3 结果可知,不同模块和损失均对去雾效果带来一定的提升,证明了每个策略的有效性。

表 3 不同损失消融实验

Table 3 Different loss ablation experiments

Method	PSNR/dB	SSIM/%
Baseline	27.98	0.938 2
$W_{loss}$	30.04	0.960 1
$L_{PM} + W_{loss}$	30.89	0.969 8
$L_p + L_{PM} + W_{loss}$	32.41	0.983 9

### 3 结束语

本文采用生成对抗网络的形式,提出了一种多深度自适应地去雾生成网络,该网络摆脱了大

气散射模型的束缚,实现了端到端的去雾。在网络的结构设计中,首先利用多个不同深度的 U 型网络提取不同感受野的特征,为了解决场景深度雾霾难以消除的问题,采用雾霾感知单元 HPB 学习雾霾特征。此外,基于向量相似性原理,提出向量混合注意力模块 VMAM,提升网络的视觉感知效果。最后,通过自适应权重模块 AWM 来关注有效特征,提高了信息的利用率。在损失函数中,由于 JS 散度的局限,在原始的对抗损失中引入 Wasserstein 距离,有效避免了生成器模式崩塌的问题,大大提升了网络的学习能力。在合成数据集和真实数据集上的实验表明:本文算法在图像色彩还原和雾霾去除中均取得优秀表现。但是,本文算法依然存在不足之处:在图像纹理信息的恢复上仍有上升空间,未来对算法的改进可以通过与传统的图像处理方法结合,利用传统图像的方法去提取图像的高频特征来修复纹理信息。

### 参考文献:

- [1] Yao D N L, Bade A, Zolkifly I A, et al. A naive but effective post-processing approach for dark channel prior (DCP)[C]//Data Science and Emerging Technologies: Proceedings of DaSET 2022, Petaling Jaya, Malaysia, 2022: 67-76.
- [2] 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.
- [3] Raanan F. Dehazing using color-lines[J]. ACM Transactions on Graphics, 2014, 34(1): 1-14.
- [4] Tan R T. Visibility in bad weather from a single image[C]//Proceedings of the 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008), Anchorage, USA, 2008: 1-8.
- [5] Zhu Q S, Mai J M, Shao L. A fast single image haze removal algorithm using color attenuation prior[J]. 2015, 24(11): 3522-3533.
- [6] Cai B, Xu X, 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.
- [7] Li B, Peng X, Wang Z, et al. Aod-net: all-in-One dehazing network[C]//proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017: 4780-4788.

- [8] Liu X, Ma Y, Shi Z, et al. Grid dehaze net: attention-based multi-scale network for image dehazing[C]//Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea, 2019: 7313-7322.
- [9] Qu Y, Chen Y, Huang J, et al. Enhanced pix2pix dehazing network[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, 2019: 8152-8160.
- [10] Dong Y, Liu Y, Zhang H, et al. Fd-gan: generative adversarial networks with fusion-discriminator for single image dehazing[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 10729-10736.
- [11] Mehta A, Sinha H, Narang P, et al. HIDEgan: a hyperspectral-guided image dehazing GAN[J]. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, USA, 2020: 846-856.
- [12] Qin X, Wang Z L, Bai Y C, et al. FFA-Net: feature fusion attention network for single image dehazing[J/OL]. [2023-07-22]. <https://doi.org/10.48550/arXiv.1911.07559>
- [13] Yang X, Yan J C, Ming Q, et al. Rethinking rotated object detection with gaussian wasserstein distance loss[J/OL]. [2023-07-23]. <https://doi.org/10.48550/arXiv.2101.11952>
- [14] Isola P, Zhu J Y, Zhou T, et al. Image-to-image translation with conditional adversarial networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 5967-5976.
- [15] Zhao C, Shuai R, Ma L, et al. Segmentation of skin lesions image based on U-Net++ [J]. Multimedia Tools and Applications, 2022, 81(6): 8691-8717.
- [16] Fu H X, Song G Q, Wang Y C. Improved YOLOv4 marine target detection combined with CBAM[J]. Symmetry, 2021, 13(4): No. 623.
- [17] Li B, Ren W, Fu D, et al. Benchmarking single-image dehazing and beyond[J]. IEEE Transactions on Image Processing, 2018, 28(1): 492-505.
- [18] Liu Y, Zhu L, Pei S D, et al. From synthetic to real: image dehazing collaborating with unlabeled real data [J/OL]. [2023-07-23]. <https://doi.org/10.48550/arXiv.2108.02934>
- [19] Sun Z, Zhang Y, Bao F, et al. SADnet: semi-supervised single image dehazing method based on an attention mechanism[J]. ACM Transactions on Multimedia Computing, Communications, and Applications, 2022, 18(2): 1-23.
- [20] Manu C M, Sreeni K G. GANID: a novel generative adversarial network for image dehazing[J]. The Visual Computer, 2022, 39: 2923-3936.