

# 基于条件生成对抗网络与知识蒸馏的单幅图像去雾方法

何涛<sup>1</sup>, 俞舒曼<sup>1</sup>, 徐鹤<sup>2,3</sup>

(1.南京邮电大学 电子与光学工程学院、微电子学院,南京 210023; 2.南京邮电大学 计算机学院、软件学院、网络空间安全学院,南京 210023; 3.江苏省无线传感网高技术研究重点实验室,南京 210023)

**摘要:** 生成对抗网络广泛应用于图像去雾领域,但通常需要较大的计算量和存储空间,从而限制了其在移动设备上的应用。针对该问题,提出一种基于条件生成对抗网络与知识蒸馏的去雾方法 KD-GAN。将频率信息作为去雾的附加约束条件,通过傅里叶变换、拉普拉斯算子、高斯滤波器分别滤除原始图像的高频或低频信息,生成对应的高频和低频图像,并将融合得到的图像作为判别器的输入,以改进雾天图像的去雾效果。在此基础上,将原重型教师网络的知识迁移到具有较少权值参数的轻量型学生网络生成器中,并对轻量型学生网络进行训练,使其以更快的收敛速度达到与教师网络相近的去雾性能。在 OTS 和 HSTS 数据集上的实验结果验证了该方法的有效性,在学生网络的参数规模仅为教师网络 1/2 的条件下,学生网络在迭代第  $3 \times 10^4$  次时,生成器输出图像的峰值信噪比和结构相似性已接近于教师网络迭代第  $5 \times 10^4$  次时的数值,训练速度加快了约 1.67 倍。

**关键词:** 图像去雾;生成对抗网络;知识蒸馏;教师网络;学生网络

开放科学(资源服务)标志码(OSID):



**中文引用格式:** 何涛,俞舒曼,徐鹤. 基于条件生成对抗网络与知识蒸馏的单幅图像去雾方法[J]. 计算机工程, 2022, 48(4): 165-172.

**英文引用格式:** HE T, YU S M, XU H. Single image dehazing method based on conditional generative adversarial network and knowledge distillation[J]. Computer Engineering, 2022, 48(4): 165-172.

## Single Image Dehazing Method Based on Conditional Generative Adversarial Network and Knowledge Distillation

HE Tao<sup>1</sup>, YU Shuman<sup>1</sup>, XU He<sup>2,3</sup>

(1.College of Electronic and Optical Engineering, College of Microelectronics, Nanjing University of Posts and Telecommunications, Nanjing 210023, China; 2.College of Computer, College of Software, College of Cyberspace Security, Nanjing University of Posts and Telecommunications, Nanjing 210023, China; 3.Jiangsu Key Laboratory of Wireless Sensor Network High Technology Research, Nanjing 210023, China)

**[Abstract]** The Generative Adversarial Network (GAN) is widely used in the field of image dehazing; however it usually requires extensive computations and a large amount storage space, limiting its application in mobile devices. To solve this problem, this study proposes a dehazing method based on knowledge distillation and conditional generative adversarial network (KD-GAN). Taking the frequency information as the additional constraint condition for dehazing, the high-frequency and low-frequency information of the original image are filtered using the Fourier transform, Laplace operator, and Gaussian filter to generate the corresponding high-frequency and low-frequency images, whereby the fused image is used as the input of the discriminator to further improve the dehazing of the fog image. In the process, the knowledge of the original heavy teacher network is transferred to the lightweight student network generator with fewer weight parameters, and the lightweight student network is trained to achieve a dehazing performance similar to that of the teacher network with faster convergence speed. The experimental results on the OTS and HSTS datasets verify the effectiveness of this method. Under the condition that the parameter scale of the student network is only half that of the teacher network, the student network is iterated  $3 \times 10^4$  times. The peak signal-to-noise ratio and structural similarity of the output image of the generator are close to those of the teacher network iterated  $5 \times 10^4$  times, while the training speed is about 1.67 times faster than that of the teacher network.

**[Key words]** image dehazing; Generative Adversarial Network (GAN); knowledge distillation; teacher network; student network

DOI: 10. 19678/j. issn. 1000-3428. 0060818

**基金项目:** 国家重点研发计划(2019YFB2103003); 江苏省科技支撑计划项目(BE2019740); 江苏省高等学校自然科学研究重大项目(18KJA520008); 江苏省“六大人才高峰”高层次人才项目(RJFW-111)。

**作者简介:** 何涛(1972—),男,讲师、博士,主研方向为无线传感器网络、移动认知无线网络;俞舒曼,硕士研究生;徐鹤,副教授、博士。

**收稿日期:** 2021-02-05 **修回日期:** 2021-05-06 **E-mail:** het@njupt.edu.cn

## 0 概述

雾是由细尘或烟雾颗粒在大气中聚集得到的产物,在雾霾天气下采集的图片通常会失去对比度、颜色保真度和边缘信息,从而降低场景的可视性。例如,若环境中存在大量雾霾,在执行视频监控任务时将难以从视频中定位到有效的目标信息,进而阻碍任务的顺利执行。因此,去除图像中的雾霾并提高图像清晰度成为计算机视觉领域的研究热点。

近年来,研究人员提出基于先验知识和基于学习的去雾方法。基于先验知识的去雾方法是根据手工设计的特征,并结合文献[1]提出的大气散射模型以恢复清晰的无雾图像。文献[2]提出一种经典的统计观测方法,即在大多数非天空区域中,某些像素中至少有一个颜色通道的强度较低,从而根据暗通道先验知识可以粗略估计出透射率和大气光值。文献[3]是针对模糊图像的场景区构建一种线性模型,根据颜色衰减来恢复深度信息,利用模糊图像的深度图估计透射率。因透射率估计不准确,导致在输出去雾图像中存在明显的伪影问题。

基于学习的去雾方法是利用卷积神经网络(Convolutional Neural Network, CNN)从大量训练数据中提取图像特征,同时学习图像对之间的映射关系。早期的大多数方法是基于大气退化模型恢复无雾图像,例如 MSCNN<sup>[4]</sup>和 DehazeNet<sup>[5]</sup>方法是最早采用 CNN 进行图像去雾,再基于大气散射模型得到无雾图像。由于这两种方法缺乏有效的先验知识,因此难以准确估计大气光值和传输率,去雾效果并不理想。端到端的学习方法是将有雾图像输入到网络模型中直接得到无雾图像,无需估计大气散射模型的中间参数。例如,文献[6]将去雾任务看成是图像到图像的转换问题,构建一种基于增强的 pix2pix 去雾网络 EPDN,获得具有较优视觉效果的无雾图像。文献[7]提出一种基于残差模块和编码器-解码器结构的去雾模型 GMAN,该模型在一些标准的去雾数据集上具有较优的去雾效果。

本文提出一种基于条件生成对抗网络(Conditional Generative Adversarial Network, CGAN)与知识蒸馏的单幅图像去雾方法 KD-GAN。该方法以改进的 U-Net 网络<sup>[8]</sup>作为生成器的基本架构,并在判别器的输入端融合图像高、低频分量,将频率信息作为去雾的附加约束条件与原始图像一起输入到判别器中。在此基础上,通过知识蒸馏的方法减小条件生成对抗网络的计算量和存储空间,从而加快模型的训练速度。

## 1 生成对抗网络

近年来,生成对抗网络<sup>[9]</sup>(Generative Adversarial Network, GAN)在图像生成、图像插值和目标分类等领域中取得较大进展。GAN 由生成器和判别器两个部

分组成。生成器  $G: Z \rightarrow Y$ , 将输入的随机噪声  $z \in Z$  映射到输出图像  $y \in Y$  中; 判别器  $D: y \rightarrow [0, 1]$ , 将其输入的图像  $y$  映射成概率值。针对去雾网络模型, 生成器的作用是利用给定的一组有雾图像  $\{y_1, y_2, \dots, y_n\}$  生成伪无雾图像  $\{y_1^*, y_2^*, \dots, y_n^*\}$ 。判别器是通过对比真实图像  $x$  与生成器生成的图像  $y^*$  的特征分布, 从而预测输入的图像为真实图像  $x$  的概率。生成器的优化目标是在保留输入图像的结构和细节信息的同时, 尽可能地生成接近真实无雾图像特征分布的去雾图像, 从而使得判别器无法辨别图像来源。判别器的优化目标则是尽可能地辨别出输入图像是生成的去雾图像还是真实无雾图像。生成器和判别器相互博弈, 以达到纳什均衡<sup>[10]</sup>。

GAN 的损失函数如式(1)所示:

$$\min_G \max_D V(D, G) = E_{x \sim P_{\text{data}}(x)} [\ln D(x)] + E_{z \sim P_z(z)} [\ln(1 - D(G(z)))] \quad (1)$$

其中:  $P_{\text{data}}(x)$  为真实图像数据的分布;  $P_z(z)$  为输入生成器噪声  $z$  的分布;  $D(x)$  为判别器判断真实图像为真的概率;  $D(G(z))$  为判别器判断生成图像为真的概率。

由于生成器需要大量的参数来生成具有高维语义信息的图像, 因此大多数 GAN 的执行会消耗较多的计算资源, 从而限制了 GAN 在 CPU 性能有限的移动设备以及边缘设备上的应用。近年来, 研究人员对重型深度模型的压缩和加速方法进行研究。如文献[11]提出一种压缩技术, 通过将教师网络产生的分类概率分布作为软目标来训练小模型, 从而将集成模型中的知识迁移到单个模型中, 这种压缩技术又称知识蒸馏。在知识蒸馏模型中将具有复杂网络结构、学习能力更强的重型网络视为教师网络, 将网络结构简单、学习能力较弱的轻量型网络视为学生网络。文献[12]指出, 浅层网络可以学习到深层网络的复杂函数, 并实现以前只有深度模型才能达到的精度。受上述方法的启发, 本文提出一种基于条件生成对抗网络和知识蒸馏的单幅图像去雾方法(KD-GAN), 本文将蒸馏模型中的教师网络设计成具有较多网络层数及权值参数的重型网络, 并设计3种不同参数规模且精简于教师网络的学生网络模型, 以对比在不同简化程度下学生网络的去雾性能。在室外模糊图像公开数据集 OTS 和室内模糊图像公开数据集 ITS 上进行网络的训练和测试。本文通过对比不同参数规模学生网络的去雾性能变化, 同时对比训练过程中相同迭代次数下教师网络和学生网络的去雾效果, 并将本文方法与其他去雾方法进行对比, 验证知识蒸馏在模型训练速率和图像去雾中的有效性。

## 2 整体架构

本文提出的基于 CGAN 和知识蒸馏的单幅图像去雾方法, 网络的整体架构包括教师网络、学生网络和知识蒸馏结构3个部分。知识蒸馏<sup>[11]</sup>将提炼的教师网络

的知识传递给结构简单、学习能力较弱的学生网络,在提高学生模型训练速度的同时,尽可能使其性能接近于具有复杂结构、学习能力强的教师网络。

2.1 教师网络

文献[13]提出 CGAN,其在生成器和判别器的输入端增加了额外的限制条件,通过源域图像代替

噪声信息,以指导生成的目标图像。本文将原始有雾图像作为条件信息添加到判别器的输入端,判别器既要判断输入图像是否真实,同时判断输入图像与作为条件信息的原始有雾图像是否匹配。基于引入频率信息作为附加条件网络(HLF-CGAN)的去雾方法中教师网络结构如图1所示。

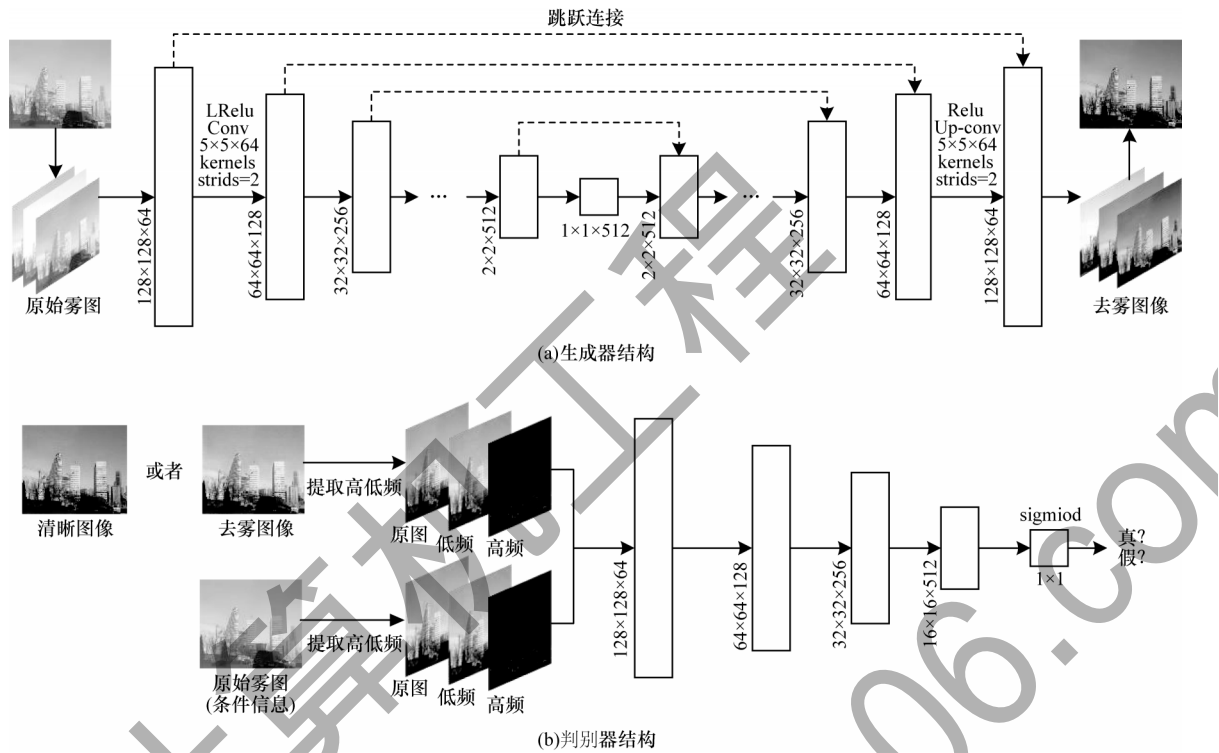


图1 教师网络结构

Fig.1 Structure of teacher network

图中生成器是基于 U-Net 网络<sup>[8]</sup>进行改进, U-Net网络是由编码器-解码器结构组成的U型对称网络,最初用于图像分割领域。编码器由4个编码块组成,每个编码块包含2个3×3大小的卷积核和1个2×2大小的最大池化层,用于下采样,解码器包含4次反卷积和4次上采样过程。虽然 U-Net 网络采用最大池化的方式进行下采样,不需要额外增加参数,但容易丢失高频特征。因此本文采用卷积步长为2的下采样方式,以减少特征的丢失。从图1可以看出,生成器网络是由编码器、解码器和跳跃连接层组成。编码器部分经过8次卷积操作将大小为256×256×3的输入图像转换成大小为1×1×512的特征图,其主要作用是提取图像特征信息。每层的卷积步长为2,卷积核大小为5×5,卷积核个数前3层分别为64、128、256个,后5层均为512个。

解码器部分进行8次反卷积操作,在与编码器相对应的位置上采用同样大小和数量的卷积核,直到最终输出大小为256×256×3的特征图。跳跃连接结构将编码器和解码器中相同大小的特征图按通道拼在一起,以更好地利用在下采样过程中保留的局部细节信息,从而加快网络训练过程。

高频和低频分量是图像的基本特征,去雾图像和对应的真实无雾图像的高、低频分量也会存在一定差异。高频分量包括图像中强度变化较快的部分,如锐利的边缘、纹理和精细的细节。相反,低频分量一般是图像强度值变化缓慢的部分,更加强调图像的亮度、颜色和对比度信息<sup>[14]</sup>。将高频和低频信息同时作为判别网络的约束条件,有利于判别器对图像边缘、纹理和颜色上的差异性进行判断。不同网络的去雾结果对比如图2所示。

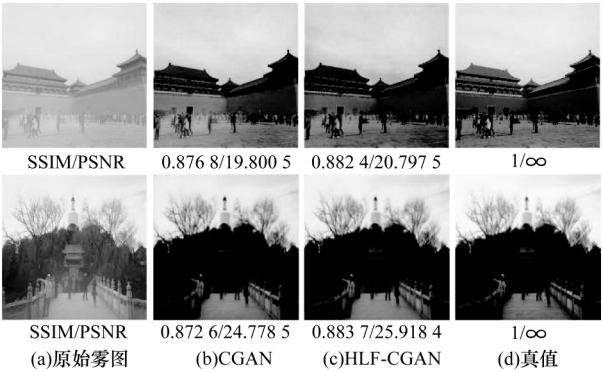


图2 不同网络的去雾结果对比

Fig.2 Dehazing results comparison among different networks

图片下方为去雾量化指标(结构相似性 SSIM/峰值信噪比 PSNR)。从图 2 可以看出,HLF-CGAN 相比直接将原图输入到判别器的网络(CGAN)具有更优的去雾效果。

图 1 中判别器采用 PatchGAN 结构, PatchGAN 与普通 GAN 的主要区别在于:普通 GAN 的判别器是将输入数据直接映射成一个概率值,而 PatchGAN 是将输入数据映射成一个  $N \times N$  大小的矩阵,矩阵中每个元素代表输入图像中一个较大的感受野。该矩阵的每个元素只有 0 或 1,即该元素的概率为真,最后再将求得的平均值作为判别器的最终输出。PatchGAN 的优点是可以从原输入图像较大的感受野中提取图像更多的细节信息。判别器对输入图像进行 4 次卷积操作,并转换成大小为  $16 \times 16 \times 512$  的特征图,再将该特征图拉伸成一维矩阵,通过 sigmoid 函数得到每个元素的概率值后再求取平均值,并将其作为最终判别器的输出。

本文通过 L1 正则项(最小绝对值偏差)约束生成图像与真实图像之间的距离,使得生成图像的特征分布更加接近真实图像的特征分布。与 L2 正则项(最小均方误差)相比, L1 正则项具有更优的鲁棒性。教师网络的总损失函数如式(2)所示:

$$L_{GAN}(G, D) = \arg \min_G \max_D \{ E_{x \sim P_{data}(x)} [\ln(1 - D(x \circ x_{LF} \circ x_{HF}, G(x) \circ G(x)_{LF} \circ G(x)_{HF}))] + E_{x, y \sim P_{data}(x)} [\ln(D(x \circ x_{LF} \circ x_{HF}, y \circ y_{LF} \circ y_{HF}))] + \lambda E[\|y - G(x)\|_1] \} \quad (2)$$

其中:  $P_{data}(x)$  为输入生成器图像的数据分布;  $E(\cdot)$  为期望值;  $x, y$  分别为原始有雾图像和真实无雾图像;

$D(x, y), D(x, G(x))$  分别为判别器判断真实图像为真的概率及判断生成图像为真的概率;符号  $\circ$  为拼接的含义;  $x \circ x_{LF} \circ x_{HF}$  为将原始雾图与其滤波后的低频、高频图像拼接;  $\lambda$  为可调节的超参数,代表损失项的权重。

## 2.2 学生网络

虽然教师网络具有更多权值参数和网络层数,能够优化去雾效果,但是其模型规模庞大,导致预训练耗时久,从而难以应用到实际场景中。例如,在自动驾驶领域中,当图像识别系统遇到雾天路况时,为达到安全驾驶的目的,系统需要及时去雾处理,此时自动驾驶的安全性将与系统的响应速度、去雾性能有着密切的关系。因此,如何压缩模型规模并提高模型训练速度成为研究热点。

本文设计了 3 个与教师网络结构相似,但具有不同参数规模的学生网络,分别记为学生网络-1、学生网络-2、学生网络-3。经计算,教师网络共包含  $3\,052.30 \times 10^4$  个权值参数,模型大小为 509 852 KB。学生网络-1 的生成器由 6 层下采样和 6 层上采样的编码器-解码器结构组成,包含  $1\,741.47 \times 10^4$  个权值参数,模型大小为 197 367 KB,约为教师网络的 1/2。其他 2 个学生网络的生成器分别由 5 层和 4 层的上下采样编码器-解码器结构组成,模型大小分别约为教师网络的 1/5 和 1/10,与教师网络的 8 层卷积与反卷积结构相比,学生网络具有更少的权值参数,并占用更小的内存空间,从而节省大量的计算资源。

## 2.3 知识蒸馏

本文提出的知识蒸馏网络结构如图 3 所示,其中学生网络结构以学生网络-1 为例。

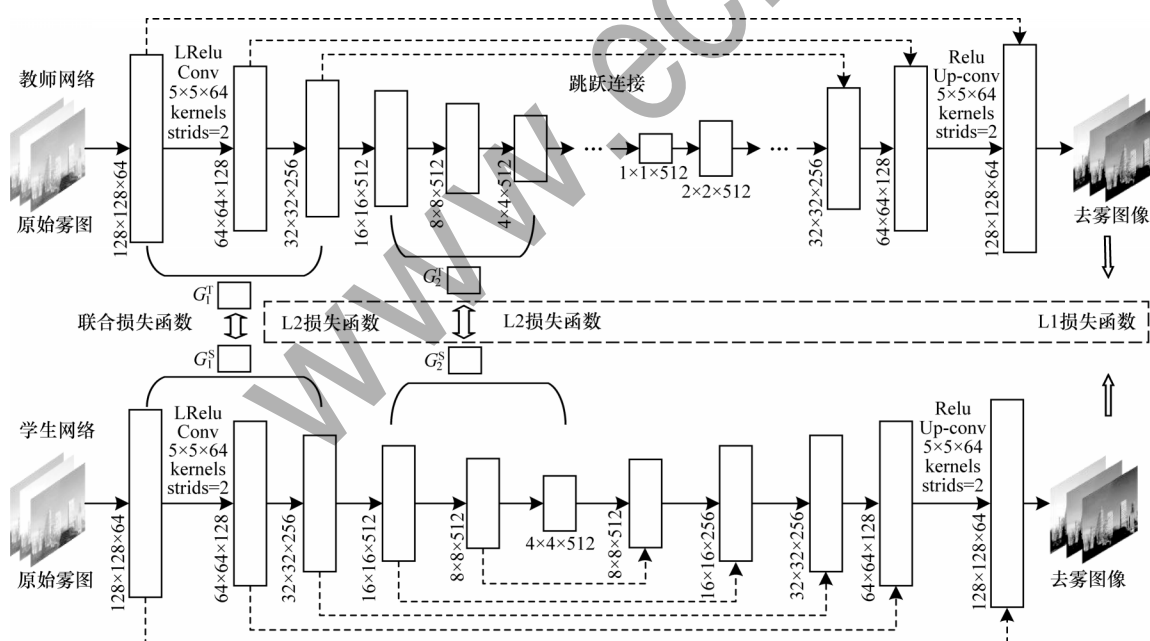


图 3 知识蒸馏结构

Fig.3 Structure of knowledge distillation

文献[15]提出的知识蒸馏方法是学生网络简单模仿教师网络的中间结果,考虑到教师网络和学生网络两层特征图之间的关系,本文首先通过最大池化将两层卷积层转化成相同的尺寸,再将其对应的通道两两作内积得到FSP矩阵,即图3中的 $G^T$ 和 $G^S$ 。FSP矩阵是由Yim等<sup>[16]</sup>提出的一种用于解决问题的流,若将流定义为2个不同层的特征图关系,知识蒸馏则是解决问题的流的方法。从图3可以看出,本文教师网络的2个FSP矩阵是由教师网络第1层、第3层、第4层和第6层的特征图分别两两作内积得到。同理,学生网络的2个FSP矩阵也是由网络相同层的特征图作内积得到。本文在得到教师网络和学生网络的FSP矩阵后,再利用L2损失函数缩短这2个网络FSP矩阵之间的距离,以达到知识蒸馏的目的。FSP矩阵的定义如式(3)所示:

$$G_{i,j}(x; W) = \sum_{s=1}^h \sum_{t=1}^w \frac{F_{s,t,i}^1(x; W) \times F_{s,t,j}^2(x; W)}{h \times w} \quad (3)$$

其中: $x$ 、 $W$ 分别为输入图像和网络的权值参数; $F_{s,t,i}^1 \in \mathbb{R}^{h \times w \times m}$ 为网络中第 $i$ 层特征图; $h$ 、 $w$ 、 $m$ 分别为特征图的高度、宽度和通道数; $F_{s,t,j}^2$ 为网络中第 $j$ 层特征图。考虑到L2正则项对异常点具有放大效果,本文选择L2范数的平方作为约束2个网络FSP矩阵的代价函数,因此中间层知识蒸馏的损失函数如式(4)所示:

$$L_{\text{FSP}}(w_t, w_s) = \frac{1}{N} \sum_x \sum_{i=1}^n \lambda_i \|G^T(x; w_t) - G^S(x; w_s)\|_2^2 \quad (4)$$

其中: $N$ 为损失函数的个数; $\lambda_i$ 为损失项的权值,取值为10~100<sup>[17]</sup>。通常 $\lambda$ 值越大,代价函数在参数为0时取到最小值,因为正则化系数越大,正则化的函数图形会加速向坐标轴原点收缩。本文在基于 $\lambda$ 为固定值100时进行实验。本文提出联合损失函数来衡量教师网络和学生网络之间的特征相似性,即同时使用L1正则项对2个网络生成器的输出端特征图进行约束。因此,学生网络生成器的总损失函数如式(5)所示:

$$G^* = \arg \min \{E_{x \sim p_{\text{data}(x)}} [\ln(1 - D(x, G(x)))] + \lambda E[\|y - G(x)\|_1] + L_{\text{FSP}}(w_t, w_s) + \lambda E[\|G^T(x) - G^S(x)\|_1]\} \quad (5)$$

其中: $G^T(x)$ 、 $G^S(x)$ 分别为教师网络和学生网络生成器输出的特征图。

本文实验主要分为以下5个步骤:

1)首先冻结学生网络权值参数,开始训练教师模型,学生网络不参与训练,直到教师网络的生成器和判别器的损失函数收敛。损失函数如式(2)所示。

2)当迭代第 $1 \times 10^4$ 、 $2 \times 10^4$ 、 $3 \times 10^4$ 、 $5 \times 10^4$ 次时,记录教师网络训练过程中测试集图像的去雾结果。

3)冻结并保存预训练好的教师网络的权值参数,在相同的数据集上开始训练学生网络,此时教师网络不参与训练,直到学生网络的生成器和判别器的损失函数收敛。损失函数如式(5)所示。

4)记录学生网络在训练过程迭代相同次数时测试集图像的去雾结果。

5)对比教师网络与学生网络在迭代相同次数时测试集中图像的去雾效果,得到相应的实验结论。

### 3 实验结果与分析

#### 3.1 实验数据与设置

本文从室外合成雾天图像数据集(OTS)和室内合成图像数据集(ITS)中分别选取5 000张组成模型的训练集。在OTS数据集上选取不重复的1 000张雾图和室外真实雾天图像(HSTS)数据集中选取1 000张雾图组成模型的测试集。

本文方法在NVIDIA RTX 1080Ti GPU 计算机的Ubuntu16.0系统上运行,并使用Tensorflow1.4.0深度学习框架进行实验。生成器输入图像大小设置为 $256 \times 256 \times 3$ ,学习率设置为固定值0.000 2,动量参数设置为0.5,采用Adam优化器对网络进行优化。

#### 3.2 实验结果分析

为客观地评估去雾效果,本文使用峰值信噪比PSNR<sup>[18]</sup>和结构相似性SSIM<sup>[19]</sup>作为去雾评价指标。PSNR值取决于原始雾图与生成图像对应像素间的误差,其值越大,恢复得到的无雾图像清晰度越高<sup>[20]</sup>。SSIM从亮度、对比度、结构3个方面对图像相似性进行评估,其值越大,代表图像失真越小。

##### 3.2.1 不同参数规模的学生网络去雾性能对比

图4为教师网络与3种不同参数规模学生网络的去雾结果对比,记为A组和B组。表1为教师网络与不同参数规模学生网络的评价指标与网络规模对比。学生网络-1在具有更少模型参数的情况下,可以达到与教师网络相近的去雾性能;学生网络-2和学生网络-3的去雾性能呈现较明显的下降趋势。从图4中方框选出的部分可以看出,学生网络-3对远景及细节部分恢复得比较模糊,失真较严重。实验结果表明,本文基于蒸馏架构,当学生网络参数规模简化为教师网络的1/2时,能够达到最佳的蒸馏效果,当继续简化学生网络参数规模时,网络的去雾评价指标与视觉感受都会逐渐下降。

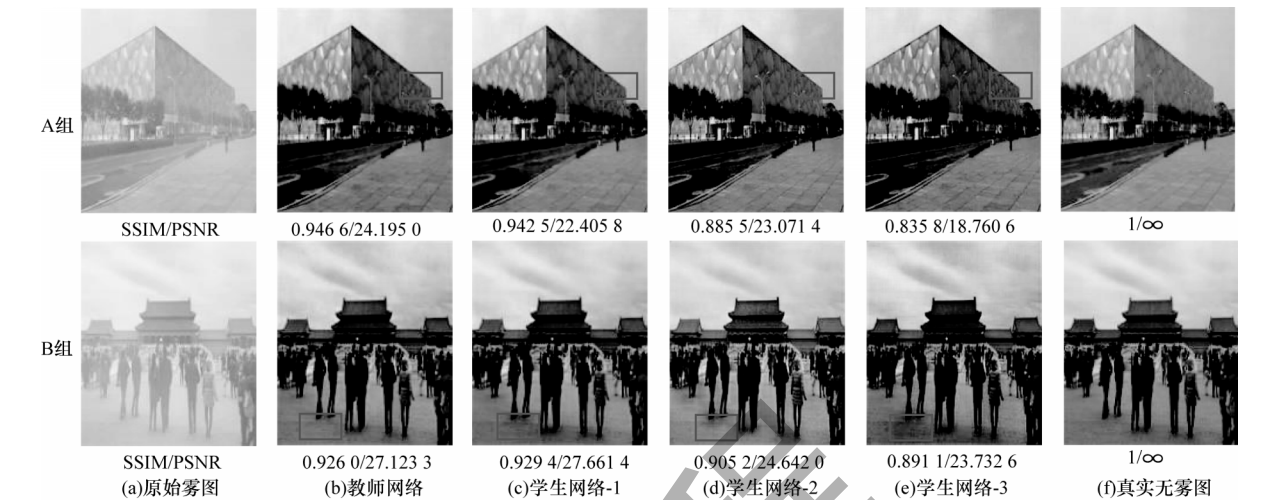


图4 教师网络与学生网络的去雾结果对比

Fig.4 Dehazing results comparison of teacher network and student network

表1 教师网络与学生网络的评价指标与网络规模对比

方法	SSIM/PSNR		权值参数量/ $10^4$	内存/KB
	A 组	B 组		
教师网络	0.946 6/24.195 0	0.926 0/27.123 3	305 2.30	509 852
学生网络-1	0.942 5/22.405 8	0.929 4/27.661 4	174 1.47	197 367
学生网络-2	0.885 5/23.071 4	0.905 2/24.642 0	108 6.06	120 563
学生网络-3	0.835 8/18.760 6	0.891 1/23.732 6	430.65	56 559

3. 2. 2 知识蒸馏在模型收敛速率中的有效性

本文对不同迭代次数下教师网络和学生网络(学生网络-1)的测试结果进行分析,分别记录了迭代第 $1\times 10^4$ 、 $2\times 10^4$ 、 $3\times 10^4$ 和 $5\times 10^4$ 次时2个网络的去雾结果。图5是教师网络和学生网络在迭代第 $1\times 10^4$ 次和第 $2\times 10^4$ 次时生成器输出的去雾图像与原始雾图的对比。

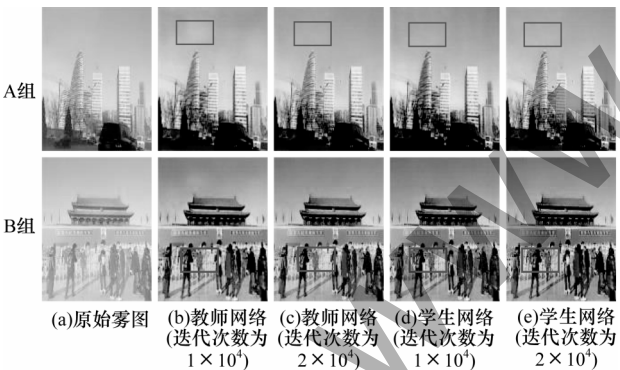


图5 在不同迭代次数下教师网络和学生网络的去雾结果对比

Fig.5 Dehazing results comparison of teacher networks and student networks under different iteration times

从迭代了相同次数的教师网络和学生网络的去雾结果可以看出,学生网络具有更优的去雾性能。在不同迭代次数下教师网络和学生网络的去雾评价指标如表2所示。

表2 在不同迭代次数下教师网络和学生网络去雾评价指标对比

网络	SSIM/PSNR		迭代次数
	A 组	B 组	
教师网络	0.864 9/19.781 1	0.832 6/17.037 1	$1\times 10^4$
	0.869 7/19.772 8	0.867 8/17.242 3	$2\times 10^4$
	0.878 1/20.218 8	0.889 3/19.940 8	$3\times 10^4$
	0.891 2/24.406 7	0.902 8/20.829 6	$5\times 10^4$
学生网络	0.858 0/20.742 4	0.868 4/18.671 3	$1\times 10^4$
	0.886 0/21.312 2	0.889 7/19.665 5	$2\times 10^4$
	0.887 7/20.718 5	0.896 0/19.612 7	$3\times 10^4$
	0.899 2/21.844 0	0.910 0/22.545 0	$5\times 10^4$

从表2可以看出,在迭代第 $1\times 10^4$ 和 $2\times 10^4$ 次时,学生网络的PSNR、SSIM值略优于迭代相同次数时的教师网络,并且学生网络在迭代到第 $3\times 10^4$ 次时PSNR、SSIM数值已接近于教师网络迭代第 $5\times 10^4$ 次时的值。这是由于训练好的重型网络通常具有大量信息,轻量型网络通过重型网络的知识迁移优化权值参数时,只需要对参数进行微调,因

此,轻量型网络的训练速度远优于从头开始训练的网络。

由此可以得出:迭代次数的增加对于网络性能的提升起着至关重要的作用,由于学生网络得到了从教师网络迁移的知识,它可以使用比教师网络更少的迭代次数,从而达到与其相同的去雾性能。



图6 不同方法的去雾结果对比

Fig.6 Dehazing results comparison among different methods

从图6可以看出,DCP方法对天空区域和白色区域的去雾失真严重,会将天空区域处理成深蓝色,白色物体会产生过高的对比度。文献[21]提出基于暗原色先验去雾方法,也存在天空区域去雾效果失真的问题。虽然NLP方法去雾性能较优,但是与真实图像对比存在颜色失真的问题。基于DNN的EPDN方法可以生成比较自然的去雾图像,然而在雾浓度大的区域中去雾结果会留有一些薄雾。基于编码-解码结构的端到端去雾网络GMAN通过残差模块提取雾天图像的特征信息,具有良好的去雾效果,但恢复的图像中存在较明显的网格状条纹,影响主观视觉感受。由于本文设计的基于CGAN网络不需要生成器中特定的函数形式,也无需利用马尔科夫链反复采样、推断,因此不会存在概率问题,生成的图像也更加清晰。图6中每幅图的下方标出了该图的去雾量化指标SSIM/PSNR,本文提出的KD-GAN去雾方法得到的量化指标与其他方法相比取得了较高的值。表3为在两种数据集上不同去雾方法的去雾评价指标PSNR、SSIM的平均值。从表3可以看出,本文提出的KD-GAN方法的PSNR和

### 3.2.3 知识蒸馏在图像去雾中的有效性

为验证知识蒸馏在图像去雾中的有效性,将本文提出的方法与DCP方法<sup>[2]</sup>、文献[21]方法、NLP方法<sup>[22]</sup>、GMAN方法<sup>[7]</sup>和EPDN方法<sup>[6]</sup>进行对比,对比结果如图6所示(彩色效果见《计算机工程》官网HTML版)。其中A组和B组图像选自OTS数据集,C组和D组图像选自HSTS数据集。

SSIM在OTS和HSTS数据集上均取得了较高的值,表明本文提出方法的去雾效果更优。

表3 不同方法的评价指标平均值对比

Table 3 The average value of evaluation indexes comparison among different methods

方法	SSIM/PSNR	
	OTS数据集	HSTS数据集
DCP方法	0.855 3/17.709 8	0.874 5/20.671 8
文献[21]方法	0.886 0/20.529 7	0.859 7/20.136 2
NLP方法	0.913 2/21.016 2	0.843 3/18.623 5
GMAN方法	0.892 4/21.294 7	0.883 8/22.265 6
EPDN方法	0.900 1/19.303 4	0.861 3/17.503 2
KD-GAN方法	0.902 0/22.418 8	0.895 8/22.888 5

## 4 结束语

本文提出一种基于条件生成对抗网络与知识蒸馏的单幅图像去雾方法。通过设计基于编码器-解码器结构的条件生成对抗网络,将原始有雾图像作为条件信息添加到判别器的输入端,以改进雾天图像的去雾效果。为降低生成对抗网络计算量和存储成本,通过选择L2范数的平方来约束重型网络和轻量化网络的

FSP矩阵,使用L1正则项对两个网络生成器输出端的特征图进行约束。在OTS和HSTS数据集上的实验结果表明,本文方法的峰值信噪比和结构相似性分别为0.902 0/22.418 8和0.895 8/22.888 5,能够有效改善去雾效果,同时加快训练速度。后续将优化以过程为导向的损失函数,进一步提升学生网络的学习效率和去雾性能。

### 参考文献

- [1] NARASIMHAN S G, NAYAR S K. Vision and the atmosphere[J]. *International Journal of Computer Vision*, 2002, 48(3): 233-254.
- [2] HE K, 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] 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: a Publication of the IEEE Signal Processing Society*, 2015, 24(11): 3522-3533.
- [4] REN W Q, PAN J S, ZHANG H. Single image dehazing via multi-scale convolutional neural networks with holistic edges[J]. *International Journal of Computer Vision*, 2020, 128(1): 240-259.
- [5] CAI B L, XU X M, JIA K. DehazeNet: an end-to-end system for single image haze removal [EB/OL]. [2021-01-03]. <https://arxiv.org/pdf/1601.07661.pdf>.
- [6] QU Y Y, CHEN Y Z, HUANG J Y. Enhanced Pix2pix dehazing network [C]//*Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Washington D. C., USA: IEEE Press, 2019: 8152-8160.
- [7] LIU Z, XIAO B T, ALRABEIAH M. Single image dehazing with a generic model-agnostic convolutional neural network[J]. *IEEE Signal Processing Letters*, 2019, 26(6): 833-837.
- [8] DUAN Y X, HE S Y, GUO D. General image segmentation by deeper residual U-Net [C]//*Proceedings of International Conference on Mathematics and Artificial Intelligence*. Chendu, China: [s. n.], 2019: 123-127.
- [9] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M. Generative adversarial nets[C]//*Proceedings of the 27th Advances in Neural Information Processing Systems*. Berlin, Germany: Springer, 2014: 2672-2680.
- [10] SETHI R, WEIBULL J. Nash equilibrium?[J]. *Notices of the AMS*, 2016, 63(5): 1-10.
- [11] HINTON G, VINYALS O, DEAN J. Distilling the knowledge in a neural network[EB/OL]. [2021-01-03]. <https://arxiv.org/abs/1503.02531>.
- [12] BA L J, CARUANA R. Do deep nets really need to be deep? [EB/OL]. [2021-01-03]. <https://arxiv.org/pdf/1312.6184v1.pdf>.
- [13] MIRZA M, OSINDERO S. Conditional generative adversarial nets[EB/OL]. [2021-01-03]. <http://de.arxiv.org/pdf/1411.1784>.
- [14] IGNATOV A, KOBYSHEV N, TIMOFTE R, et al. DSLR-quality photos on mobile devices with deep convolutional networks[C]//*Proceedings of IEEE International Conference on Computer Vision*. Washington D. C., USA: IEEE Press, 2017: 3277-3285.
- [15] ROMERO A, BALLAS N, KAHOU S E, et al. FitNets: hints for thin deep nets[EB/OL]. [2021-01-03]. <https://arxiv.org/pdf/1412.6550.pdf>.
- [16] YIM J, JOO D, BAE J. A gift from knowledge distillation: fast optimization, network minimization and transfer learning[C]//*Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Washington D. C., USA: IEEE Press, 2017: 7130-7138.
- [17] ISOL P, ZHU J Y, ZHOU T H. Image-to-image translation with conditional adversarial networks[EB/OL]. [2021-01-03]. <https://arxiv.org/pdf/1611.07004.pdf>.
- [18] ZHOU W, BOVIK A C, SHEIKH H R. Image quality assessment: from error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612.
- [19] MANNOS J, SAKRISON D. The effects of a visual fidelity criterion of the encoding of images[J]. *IEEE Transactions on Information Theory*, 1974, 20(4): 525-536.
- [20] ZHANG J W, WANG X Y, YANG C H, et al. Image dehazing based on dark channel prior and brightness enhancement for agricultural remote sensing images from consumer-grade cameras[J]. *Computers and Electronics in Agriculture*, 2018, 151: 196-206.
- [21] 蒋建国, 侯天峰, 齐美彬. 改进的基于暗原色先验的图像去雾算法[J]. *电路与系统学报*, 2011, 16(2): 7-12.
- [22] JIANG J G, HOU T F, QI M B. Image dehazing algorithm based on dark primary color prior[J]. *Journal of Circuits and Systems*, 2011, 16(2): 7-12. (in Chinese)
- [23] BERMAN D, TREIBITZ T, AVIDAN S. Non-local image dehazing [C]//*Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Washington D. C., USA: IEEE Press, 2016: 1674-1682.

编辑 薛晋栋